# Apuntes de Álgebra Lineal

Francisco Martínez Pería
Noelia Belén Rios

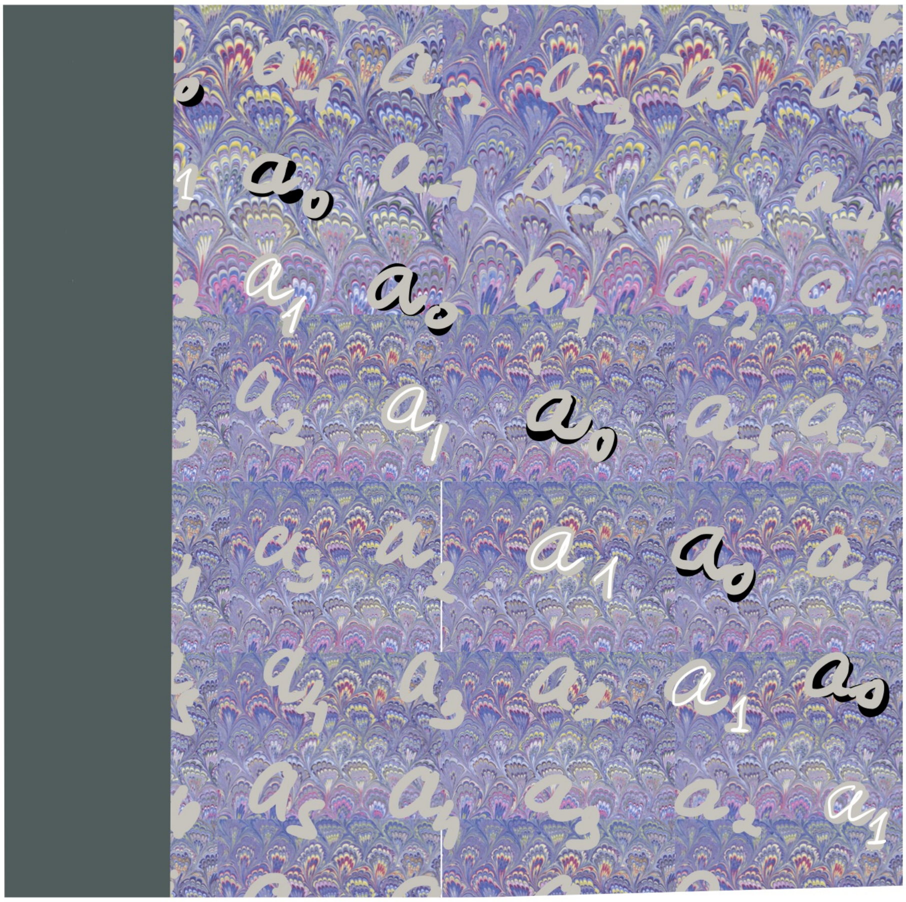


Figura 1: El arte de tapa es una gentileza de la Dra. Alejandra Maestripieri.

# Índice general

## Prólogo

Este libro de cátedra es una consecuencia (a nuestro entender, positiva) de la pandemia de COVID-19. A partir del Aislamiento Social Preventivo Obligatorio (ASPO) decretado por el Gobierno Nacional a mediados de marzo de 2020, nos encontramos ante la necesidad de compilar y sistematizar los apuntes teóricos y los trabajos prácticos que utilizamos para dictar un curso de álgebra lineal.

Nuestra intención no es presentar una obra original sobre álgebra lineal, sino realizar una recopilación del material que suele utilizarse para el curso. La mayor parte del contenido de estos apuntes (así como el orden en que se presentan la mayoría de los temas) está basado en el clásico libro de K. Hoffman y R. Kunze [4], el cual utilizamos desde hace muchos años. Sin embargo, en varios puntos decidimos tomar enfoques diferentes. Por ejemplo, el Capítulo 5 está influenciado por el libro de K. Nomizu [7] y el Capítulo 10 por el de E. Vinberg.

Otra fuente de información a la que hemos recurrido son notas de clase producidas por los profesores Pedro Massey, Marila Etchechoury y Gastón A. García en años anteriores. A todos ellos va nuestro agradecimiento.

*Francisco Martínez Pería y Noelia Belén Rios*

# Capítulo 1

# Espacios Vectoriales

Comenzamos recordando la definición de cuerpo, ya que utilizaremos uno de éstos como el conjunto de escalares necesarios para combinar con los "vectores" de un espacio vectorial abstracto. Después presentamos algunos ejemplos, los cuales utilizaremos frecuentemente.

**Definición 1.1.** Un **cuerpo** $\mathbb{K}$ es un conjunto dotado con dos operaciones binarias

$$+ : \mathbb{K} \times \mathbb{K} \to \mathbb{K}, \qquad \cdot : \mathbb{K} \times \mathbb{K} \to \mathbb{K},$$

que satisfacen las siguientes propiedades:

i) $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$ para todo $\alpha, \beta, \gamma \in \mathbb{K}$;

ii) $\alpha + \beta = \beta + \alpha$ para todo $\alpha, \beta \in \mathbb{K}$;

iii) existe $0 \in \mathbb{K}$ tal que $\alpha + 0 = 0 + \alpha = \alpha$ para todo $\alpha \in \mathbb{K}$;

iv) para cada $\alpha \in \mathbb{K}$ existe un único $-\alpha \in \mathbb{K}$ tal que $\alpha + (-\alpha) = (-\alpha) + \alpha = 0$;

v) $(\alpha \cdot \beta) \cdot \gamma = \alpha \cdot (\beta \cdot \gamma)$ para todo $\alpha, \beta, \gamma \in \mathbb{K}$;

vi) $\alpha \cdot \beta = \beta \cdot \alpha$ para todo $\alpha, \beta \in \mathbb{K}$;

vii) existe $1 \in \mathbb{K}$ tal que $\alpha \cdot 1 = 1 \cdot \alpha = \alpha$ para todo $\alpha \in \mathbb{K}$;

viii) para cada $\alpha \in \mathbb{K}$ existe un único $\alpha^{-1} \in \mathbb{K}$ tal que $\alpha \cdot \alpha^{-1} = \alpha^{-1} \cdot \alpha = 1$;

ix) $(\alpha + \beta) \cdot \gamma = \alpha \cdot \gamma + \beta \cdot \gamma$ para todo $\alpha, \beta, \gamma \in \mathbb{K}$.

A los elementos distinguidos 0 y 1 de $\mathbb{K}$ los llamaremos el **elemento neutro** y la **elemento unidad** del cuerpo $\mathbb{K}$, respectivamente.

**Ejemplos 1.2.**

1. $\mathbb{Q}$, $\mathbb{R}$ y $\mathbb{C}$ son cuerpos con las operaciones usuales.
2. Si $p \in \mathbb{N}$ es un número primo, el conjunto $\mathbb{Z}_p$ formado por las clases de equivalencia módulo $p$ es un cuerpo con las operaciones

$$\overline{n} + \overline{m} := \overline{n + m}, \qquad \overline{n} \cdot \overline{m} := \overline{n \cdot m}, \qquad n, m \in \mathbb{Z}.$$

3. $\mathbb{N}$ y $\mathbb{Z}$ no son cuerpos.

4. Si $i$ denota a la unidad imaginaria del cuerpo $\mathbb{C}$, el subcunjunto de $\mathbb{C}$ dado por
$$\mathbb{Q}[\sqrt{2}] = \{p + i\sqrt{2}q : \ p, q \in \mathbb{Q}\},$$
dotado con las operaciones que hereda de $\mathbb{C}$, también es un cuerpo.

A lo largo de estas notas, $\mathbb{K}$ denotará al cuerpo $\mathbb{C}$ ó a algún subcuerpo de $\mathbb{C}$, salvo que se indique lo contrario. Sin embargo, la próxima definición tiene sentido para un cuerpo $\mathbb{K}$ arbitrario.

**Definición 1.3.** Dado un cuerpo $\mathbb{K}$, un **espacio vectorial sobre** $\mathbb{K}$ es un conjunto $V$ dotado de una operación binaria, $+ : V \times V \to V$, y de una acción del cuerpo $\mathbb{K}$ sobre $V$,
$$\cdot : \mathbb{K} \times V \to V,$$
que satisfacen:

i) $(u + v) + w = u + (v + w)$ para todo $u, v, w \in V$;
ii) $u + v = v + u$ para todo $u, v \in V$;
iii) existe $\vec{0} \in V$ tal que $v + \vec{0} = \vec{0} + v = v$ para todo $v \in V$;
iv) para cada $v \in V$ existe un único $-v \in V$ tal que $v + (-v) = (-v) + v = \vec{0}$;
v) $(\alpha \cdot \beta) \cdot v = \alpha \cdot (\beta \cdot v)$ para todo $\alpha, \beta \in \mathbb{K}$ y todo $v \in V$;
vi) la unidad $1 \in \mathbb{K}$ verifica que $1 \cdot v = v$ para todo $v \in V$;
vii) $(\alpha + \beta) \cdot v = \alpha \cdot v + \beta \cdot v$ para todo $\alpha, \beta \in \mathbb{K}$ y todo $v \in V$;
viii) $\alpha \cdot (u + v) = \alpha \cdot u + \alpha \cdot v$ para todo $\alpha \in \mathbb{K}$ y todo $u, v \in V$.

Al elemento distinguido $\vec{0} \in V$ lo llamaremos el **vector nulo** del $\mathbb{K}$-espacio vectorial $V$.

A los elementos de un $\mathbb{K}$-espacio vectorial los llamaremos **vectores**, y a la operación binaria $+$ y a la acción $\cdot$ les diremos coloquialmente la **suma** en $V$ y el **producto por escalares** de $\mathbb{K}$, respectivamente.

**Ejemplos 1.4.** Dado un cuerpo $\mathbb{K}$, los siguientes son algunos ejemplos de $\mathbb{K}$-espacios vectoriales:

1. Dado $n \in \mathbb{N}$, $\mathbb{K}^n := \{(x_1, \ldots, x_n) : \ x_k \in \mathbb{K}, \ k = 1, \ldots, n\}$ con las operaciones:
$$\begin{aligned} (x_1, \ldots, x_n) + (y_1, \ldots, y_n) &:= (x_1 + y_1, \ldots, x_n + y_n), \\ \alpha \cdot (x_1, \ldots, x_n) &:= (\alpha \cdot x_1, \ldots, \alpha \cdot x_n), \qquad \alpha \in \mathbb{K}, \end{aligned}$$
es un $\mathbb{K}$-espacio vectorial al que llamaremos el espacio de las $n$-uplas en $\mathbb{K}$.

2. Dados $n, m \in \mathbb{N}$, el conjunto $\mathbb{K}^{n\times m}$ de matrices de $n \times m$ con entradas en $\mathbb{K}$, dotado con las operaciones:
$$\begin{aligned} (A + B)_{ij} &:= A_{ij} + B_{ij}, \qquad A, B \in \mathbb{K}^{n\times m} \\ (\alpha \cdot A)_{ij} &:= \alpha \cdot A_{ij}, \qquad A \in \mathbb{K}^{n\times m}, \ \alpha \in \mathbb{K}, \end{aligned}$$
es un $\mathbb{K}$-espacio vectorial.

3. Dado un conjunto $D$, el conjunto de funciones con dominio $D$ a valores en $\mathbb{K}$:

$$\mathcal{F}(D,\mathbb{K}) := \{f : D \to \mathbb{K} \ : \ f \text{ es una función}\},$$

dotado con las operaciones:

$$\begin{aligned}(f+g)(x) &:= f(x)+g(x), \qquad f,g \in \mathcal{F}(D,\mathbb{K}),\\ (\alpha \cdot f) &:= \alpha \cdot f(x), \qquad f \in \mathcal{F}(D,\mathbb{K}),\ \alpha \in \mathbb{K},\end{aligned}$$

también es un $\mathbb{K}$-espacio vectorial.

4. Si $V$ es un $\mathbb{K}$-espacio vectorial entonces $\mathcal{F}(D,V)$, con las operaciones definidas como en el ejemplo anterior, también es un $\mathbb{K}$-espacio vectorial.

5. El conjunto de polinomios $\mathbb{K}[x]$ con coeficientes en el cuerpo $\mathbb{K}$ en la indeterminada $x$ es un $\mathbb{K}$-espacio vectorial con las operaciones usuales de suma y producto por escalares. La definición formal de $\mathbb{K}[x]$ la dejamos para más adelante.

6. El subconjunto de $\mathcal{F}(\mathbb{K},\mathbb{K})$ formado por las **funciones polinomiales**, es decir, aquellas cuya ley de asignación es de la forma

$$f(x) = a_0 + a_1 x + \ldots + a_n x^n, \quad \text{para algún } n \in \mathbb{N}, \text{ y } a_0, a_1 \ldots, a_n \in \mathbb{K},$$

también es un $\mathbb{K}$-espacio vectorial con las operaciones que hereda de $\mathcal{F}(\mathbb{K},\mathbb{K})$.

Si $\mathbb{K}$ tiene una cantidad infinita de elementos, el $\mathbb{K}$-espacio vectorial de las funciones polinomiales puede identificarse con $\mathbb{K}[x]$, pero esta identificación falla si $\mathbb{K}$ es un cuerpo finito.

7. El espacio de sucesiones sobre $\mathbb{K}$:

$$\mathbb{K}^\infty := \{(a_j)_{j\in\mathbb{N}} : \ a_j \in \mathbb{K},\ j \in \mathbb{N}\},$$

dotado con las operaciones:

$$\begin{aligned}(a+b)_j &:= a_j + b_j, \qquad a,b \in \mathbb{K}^\infty\\ (\alpha \cdot a)_j &:= \alpha \cdot a_j, \qquad a \in \mathbb{K}^\infty, \alpha \in \mathbb{K},\end{aligned}$$

es un $\mathbb{K}$-espacio vectorial. Notemos que $\mathbb{K}^\infty$ puede identificarse con $\mathcal{F}(\mathbb{N},\mathbb{K})$.

A continuación presentamos algunas consecuencias inmediatas de la definición de $\mathbb{K}$-espacio vectorial. Para ello, supongamos que $V$ es un $\mathbb{K}$-espacio vectorial.

- Si $\vec{0}$ es el vector nulo de $V$ y $\alpha \in \mathbb{K}$ entonces $\alpha \cdot \vec{0} = \vec{0}$.

  De hecho, como $\alpha \cdot \vec{0} = \alpha \cdot (\vec{0}+\vec{0}) = \alpha \cdot \vec{0} + \alpha \cdot \vec{0}$, resulta que

$$\vec{0} = \alpha \cdot \vec{0} + \big(-(\alpha \cdot \vec{0})\big) = \alpha \cdot \vec{0} + \alpha \cdot \vec{0} + \big(-(\alpha \cdot \vec{0})\big) = \alpha \cdot \vec{0}.$$

- Si $0$ es el neutro de $\mathbb{K}$ y $v \in V$ entonces $0 \cdot v = \vec{0}$.

  En efecto, $0 \cdot v = (0+0) \cdot v = 0 \cdot v + 0 \cdot v$. Luego,

$$\vec{0} = 0 \cdot v + \big(-(0 \cdot v)\big) = (0 \cdot v + 0 \cdot v) + \big(-(0 \cdot v)\big) = 0 \cdot v.$$

  Luego, dados $v \in V$ y $\alpha \in \mathbb{K}$,

$$\text{si } \alpha \cdot v = \vec{0} \text{ entonces} \quad \alpha = 0 \quad \text{ó} \quad v = \vec{0},$$

  porque si $\alpha \neq 0$ tenemos que $v = 1 \cdot v = (\alpha^{-1} \cdot \alpha) \cdot v = \alpha^{-1} \cdot (\alpha \cdot v) = \alpha^{-1} \cdot \vec{0} = \vec{0}$.

- Además, si $v \in V$ entonces $(-1) \cdot v = -v$.

  Esto último sale de que $\vec{0} = 0 \cdot v = (1 + (-1)) \cdot v = 1 \cdot v + (-1) \cdot v = v + (-1) \cdot v$, y en consecuencia,

$$-v = -v + (v + (-1) \cdot v) = (-v + v) + (-1) \cdot v = \vec{0} + (-1) \cdot v = (-1) \cdot v.$$

A partir de aquí obviaremos el uso de $\cdot$ para anotar al producto por escalares. Es decir, dados $\alpha \in \mathbb{K}$ y $v \in V$, en lugar de anotar $\alpha \cdot v$ escribiremos simplemente $\alpha v$.

**Observación 1.5.** Las propiedades asociativa y conmutativa de la adición de vectores implican que la suma de una cantidad finita de vectores cualesquiera $v_1 \ldots, v_k \in V$ es independiente del orden de los sumandos y de cómo se asocien entre ellos, es decir,

$$v_1 + \ldots + v_k \text{ está bien definida.}$$

Antes de pasar a la próxima sección, formalicemos la idea de que ciertos vectores de un $\mathbb{K}$-espacio vectorial $V$ pueden descomponerse en términos de una familia distinguida $\{v_1, \ldots, v_k\}$ contenida en $V$.

**Definición 1.6.** Dado un $\mathbb{K}$-espacio vectorial $V$, sean $v_1, \ldots, v_k \in V$. Un vector $v \in V$ es una **combinación lineal** de los vectores $v_1, \ldots, v_k$ si existen escalares $\alpha_1, \ldots, \alpha_k \in \mathbb{K}$ tales que

$$v = \sum_{i=1}^{k} \alpha_i v_i = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_k v_k.$$

## 1.1. Subespacios

Entre los subconjuntos de un $\mathbb{K}$-espacio vectorial nos interesa destacar a aquellos que son también un $\mathbb{K}$-espacio vectorial.

**Definición 1.7.** Dado un $\mathbb{K}$-espacio vectorial $V$, un **subespacio** de $V$ es un subconjunto $W \subseteq V$ tal que, con las operaciones que hereda de $V$, resulta también un $\mathbb{K}$-espacio vectorial.

Para probar que un subconjunto $W \subseteq V$ es un subespacio de $V$, alcanza con probar que $W$ es **no vacío** y que las operaciones son **cerradas** en $W$, es decir,

- si $v, w \in W$ entonces $v + w \in W$;
- si $\alpha \in \mathbb{K}$ y $w \in W$ entonces $\alpha w \in W$.

**Observación 1.8.** Si $W$ es un subespacio de un $\mathbb{K}$-espacio vectorial $V$, entonces $\vec{0} \in W$. Esto es fácil de verificar, porque si $w \in W$ (alguno hay, porque $W \neq \varnothing$) entonces $\vec{0} = 0 \cdot w \in W$.

Por este motivo, muchas veces la condición " $W \neq \varnothing$ " se reemplaza por la condición equivalente " $\vec{0} \in W$ ".

Si $V$ es un $\mathbb{K}$-espacio vectorial, entonces $\{\vec{0}\}$ es un subespacio de $V$, lo llamaremos el **subespacio nulo**. Además, $V$ también es un subespacio de sí mismo. Entonces, a estos dos los llamaremos los **subespacios triviales** de $V$.

**Definición 1.9.** Un subespacio $W$ de $V$ es un **subespacio propio** si $W \neq \{\vec{0}\}$ y $W \neq V$.

**Ejemplos 1.10.**

1. Las funciones polinomiales sobre un cuerpo $\mathbb{K}$ son un subespacio propio de $\mathcal{F}(\mathbb{K}, \mathbb{K})$.

2. En particular, si $\mathbb{K} = \mathbb{R}$, las funciones polinomiales también son un subespacio propio de $C(\mathbb{R})$, el $\mathbb{R}$-espacio vectorial de funciones continuas (definidas en $\mathbb{R}$). A su vez, $C(\mathbb{R})$ es un subespacio de $\mathcal{F}(\mathbb{R}, \mathbb{R})$.

3. El conjunto $C^1(\mathbb{R})$ formado por las funciones con derivada continua es un subespacio propio de $C(\mathbb{R})$. Además, el conjunto de soluciones de la ecuación diferencial
$$y''(t) + y(t) = 0,$$
es un subespacio de $C^1(\mathbb{R})$. En efecto, este conjunto puede describirse como
$$\{\alpha \cos(t) + \beta \operatorname{sen}(t) : \ \alpha, \beta \in \mathbb{R}\}.$$

4. Un subconjunto $S \subseteq \mathbb{R}^2$ es un subespacio propio de $\mathbb{R}^2$ si y sólo si $S$ es una recta que pasa por el origen.

5. Un subconjunto $S \subseteq \mathbb{R}^3$ es un subespacio propio de $\mathbb{R}^3$ si y sólo si $S$ es una recta ó un plano que pasa por el origen.

6. Si $m < n$ el conjunto de soluciones del sistema de ecuaciones lineales homogéneo determinado por una matriz $A \in \mathbb{K}^{m\times n}$,
$$A \cdot X = \vec{0},$$
es un subespacio propio de $\mathbb{K}^m$.

El último ejemplo de la lista anterior puede pensarse como un caso particular del siguiente resultado, ya que un vector $X_0 \in \mathbb{K}^m$ es solución del sistema de ecuaciones si es solución de cada una de las ecuaciones individuales que lo conforman.

**Teorema 1.11.** *Sea $V$ un $\mathbb{K}$-espacio vectorial. Si $\{W_i\}_{i\in I}$ es una familia arbitraria de subespacios de $V$, entonces $\bigcap_{i\in I} W_i$ también es un subespacio de $V$.*

*Demostración.* Dada una familia arbitraria $\{W_i\}_{i\in I}$ de subespacios de $V$, sea $W := \bigcap_{i\in I} W_i$. Sean $v, w \in W$ y $\alpha \in \mathbb{K}$. Como $v, w \in W_i$ para cada $i \in I$, resulta que $\alpha v + w \in W_i$ para todo $i \in I$. Es decir, $\alpha v + w \in W$.

Entonces, como las operaciones son cerradas en $W$ y además $W \neq \varnothing$ (porque contiene a $\vec{0}$), $W$ es un subespacio de $V$. □

Es muy importante aclarar que, en general, la unión de subespacios de un $\mathbb{K}$-espacio vectorial $V$ puede no ser un subespacio de $V$.

**Ejemplo 1.12.** En $\mathbb{R}^2$, fijados dos escalares distintos $r_1, r_2 \in \mathbb{R}$, $r_1 \neq r_2$, consideremos los conjuntos:
$$R_1 := \big\{(x, r_1 x) : \ x \in \mathbb{R}\big\},$$
$$R_2 := \big\{(x, r_2 x) : \ x \in \mathbb{R}\big\}.$$
Es fácil ver que $R_1$ y $R_2$ son dos subespacios de $\mathbb{R}^2$, pero $R_1 \cup R_2$ no es un subespacio de $\mathbb{R}^2$ ya que
$$(1, r_1) \in R_1, \quad (1, r_2) \in R_2 \quad \text{pero} \quad (0, r_2 - r_1) = (1, r_2) - (1, r_1) \notin R_1 \cup R_2.$$
Gráficamente, si $r_1 = \frac{1}{2}$ y $r_2 = \frac{3}{2}$ entonces el punto $(0, 1)$ no pertenece a la unión de las rectas $R_1$ y $R_2$.

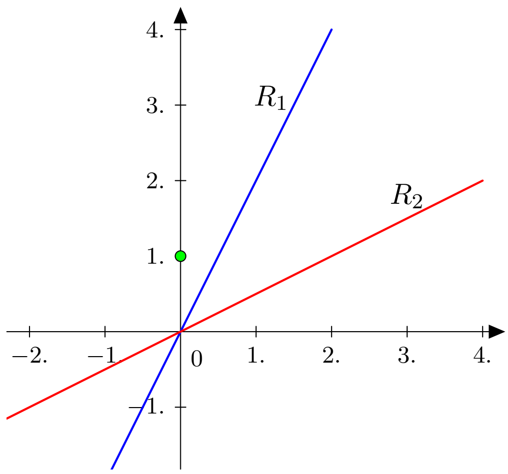



Figura 1.1: $R_1 \cup R_2$ no es un subespacio de $\mathbb{R}^2$.

Como la unión de dos subespacios $W_1$ y $W_2$ puede no ser un subespacio, debemos sustituir la idea de la unión por otra que represente al menor subespacio de $V$ que contiene tanto a $W_1$ como a $W_2$. Comencemos definiendo cual es el menor subespacio que contiene a un subconjunto arbitrario.

**Definición 1.13.** Dado un $\mathbb{K}$-espacio vectorial $V$, sea $S \subseteq V$ un subconjunto no vacío. El **subespacio generado por** $S$, al que anotaremos $\overline{S}$, es el conjunto de todas las combinaciones lineales (finitas) que pueden construirse con los vectores de $S$, es decir, $v \in \overline{S}$ si y sólo si existen $k \in \mathbb{N}$, vectores $v_1, \ldots, v_k \in S$ y escalares $\alpha_1, \ldots, \alpha_k \in \mathbb{K}$ tales que

$$v = \sum_{i=1}^{k} \alpha_i v_i = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_k v_k.$$

**Ejercicio 1.14.** Mostrar que $\overline{S}$ es el menor subespacio que contiene a $S$, es decir, probar que $\overline{S}$ es la intersección de todos los subespacios que contienen a $S$.

**Definición 1.15.** Dado un $\mathbb{K}$-espacio vectorial $V$, sean $W_1$ y $W_2$ dos subespacios de $V$. Definimos la **suma de los subespacios** $W_1$ y $W_2$ como el conjunto

$$W_1 + W_2 := \{u + v : \ u \in W_1, \ v \in W_2\}.$$

**Ejercicio 1.16.** Probar que $W_1 + W_2 = \overline{W_1 \cup W_2}$ y, por lo tanto, $W_1 + W_2$ es un subespacio de $V$.

**Ejemplo 1.17.** Si $V = \mathbb{C}^{2\times 2}$ es el $\mathbb{C}$-espacio vectorial de matrices de $2 \times 2$ con entradas complejas, consideremos los subespacios

$$W_1 = \left\{ \begin{pmatrix} x & y \\ z & 0 \end{pmatrix} : \ x, y, z \in \mathbb{C} \right\},$$
$$W_2 = \left\{ \begin{pmatrix} u & 0 \\ 0 & v \end{pmatrix} : \ u, v \in \mathbb{C} \right\}.$$

En este caso es fácil ver que $W_1 + W_2$ coincide con $V = \mathbb{C}^{2\times 2}$, mientras que

$$W_1 \cap W_2 = \left\{ \begin{pmatrix} x & 0 \\ 0 & 0 \end{pmatrix} : \ x \in \mathbb{C} \right\}.$$

La primera afirmación es consecuencia de que, si $a, b, c, d \in \mathbb{C}$,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \overbrace{\begin{pmatrix} a & b \\ c & 0 \end{pmatrix}}^{\in W_1} + \overbrace{\begin{pmatrix} 0 & 0 \\ 0 & d \end{pmatrix}}^{\in W_2}.$$

## 1.2. Bases y dimensión

Dado un $\mathbb{K}$-espacio vectorial $V$ arbitrario, necesitamos comprobar que existe (al menos un) subconjunto distinguido de $V$ que nos permita describir a los vectores de $V$, de manera única, como combinación lineal de los elementos de éste. A un subconjunto tal lo llamaremos **base de** $V$. Para presentar la definición formal necesitamos introducir primero cierta idea de independencia.

**Definición 1.18.** Sea $V$ un $\mathbb{K}$-espacio vectorial. Un conjunto $S \subseteq V$ se dice **linealmente dependiente** si existen $k \in \mathbb{N}$, vectores $v_1, \ldots, v_k \in S$ y escalares $\alpha_1, \ldots, \alpha_k \in \mathbb{K}$ no todos nulos, tales que

$$\alpha_1 v_1 + \ldots + \alpha_k v_k = \vec{0}.$$

Un conjunto que no es linealmente dependiente se dice **linealmente independiente**.

**Observación 1.19.** Si $S$ es un conjunto finito, digamos $S = \{v_1, \ldots, v_n\}$, entonces $S$ es linealmente independiente si y sólo si

$$\alpha_1 v_1 + \ldots + \alpha_n v_n = \vec{0} \qquad \Longrightarrow \qquad \alpha_1 = \ldots = \alpha_n = 0.$$

Los siguientes son algunos criterios que pueden utilizarse para decidir si un conjunto es linealmente dependiente o linealmente independiente.

i) Si $\vec{0} \in S$ entonces $S$ es linealmente dependiente, ya que $\vec{0} = \alpha \cdot \vec{0}$ con $\alpha \neq 0$.

ii) Si $S$ es linealmente dependiente y $S \subset S'$, entonces $S'$ es linealmente dependiente.

iii) Si $S$ es linealmente independiente y $T \subset S$, entonces $T$ es linealmente independiente.

iv) $S$ es linealmente independiente si y sólo si todo subconjunto finito de $S$ es linealmente independiente.

**Ejemplos 1.20.**

1. Tres vectores no nulos en $\mathbb{R}^2$ forman siempre un conjunto linealmente dependiente.

   En efecto, dados $\vec{u} = (u_1, u_2), \vec{v} = (v_1, v_2), \vec{w} = (w_1, w_2) \in \mathbb{R}^2$ y escalares $\alpha, \beta, \gamma \in \mathbb{R}$, notemos que $\alpha\vec{u} + \beta\vec{v} + \gamma\vec{w} = \vec{0}$ si y sólo si

   $$\begin{cases} \alpha u_1 + \beta v_1 + \gamma w_1 &= 0 \\ \alpha u_2 + \beta v_2 + \gamma w_2 &= 0 \end{cases}.$$

   Este es un sistema (homogéneo) de dos ecuaciones lineales con 3 incógnitas, con lo cual admite infinitas soluciones. En particular, existen $\alpha_0, \beta_0, \gamma_0 \in \mathbb{R}$ no todos nulos tales que

   $$\alpha_0\vec{u} + \beta_0\vec{v} + \gamma_0\vec{w} = \vec{0}.$$

   Por lo tanto, $\{\vec{u}, \vec{v}, \vec{w}\}$ es linealmente dependiente.

2. En $\mathbb{K}[x]$, la familia de polinomios

$$\{p_0(x) = 1,\ p_1(x) = x,\ p_2(x) = x^2,\ \dots,\ p_m(x) = x^m,\ p_{m+1}(x) = x^{m+1},\ \dots\}, \quad (1.1)$$

   es un conjunto linealmente independiente.

   Para probarlo, usaremos el ítem *iv*) de la lista de criterios anteriores, es decir, veremos que todo subconjunto finito de esta familia es linealmente independiente.

   Supongamos que $F$ es un subconjunto finito, y para fijar ideas, supongamos que tiene cardinal $k$. Luego, existen $k$ números naturales distintos $n_1, \dots, n_k \in \mathbb{N}$ (a los que suponemos ordenados de menor a mayor), tales que $F = \{p_{n_1}, p_{n_2}, \dots, p_{n_k}\}$.

   Dados $k$ escalares $\alpha_1, \dots, \alpha_k \in \mathbb{K}$, notemos que

$$\alpha_1 p_{n_1} + \alpha_2 p_{n_2} + \ldots + \alpha_k p_{n_k} = \vec{0},$$

   equivale a que el polinomio $\alpha_1 x^{n_1} + \alpha_2 x^{n_2} + \ldots + \alpha_k x^{n_k}$ coincida con el polinomio nulo. Para lo cual es necesario que todos los coeficientes $\alpha_1, \dots, \alpha_k$ sean nulos. Por lo tanto, $F$ es linealmente independiente.

   Como $F$ era un subconjunto finito arbitrario de la familia, tenemos que la familia (1.1) es linealmente independiente.

3. En $C(\mathbb{R})$, las funciones

$$f(x) = \cos^2(x), \quad g(x) = \operatorname{sen}^2(x), \quad h(x) = 1,$$

   forman un conjunto linealmente dependiente.

   En efecto, la identidad pitagórica asegura que

$$\cos^2(x) + \operatorname{sen}^2(x) = 1, \qquad \text{para todo } x \in \mathbb{R}.$$

   Es decir, $1 \cdot f + 1 \cdot g + (-1) \cdot h = \vec{0}$ y, por lo tanto, $\{f, g, h\}$ es linealmente dependiente.

**Definición 1.21.** Sea $V$ un $\mathbb{K}$-espacio vectorial. Un subconjunto $\mathcal{B} \subset V$ es una **base** de $V$ si verifica las siguientes condiciones:

- $\mathcal{B}$ es linealmente independiente;
- $\mathcal{B}$ es un **sistema de generadores** de $V$, es decir, $V = \overline{\mathcal{B}}$.

Diremos que $\mathcal{B}$ es una **base ordenada** si existe un conjunto ordenado de índices $I$ que nos permita describir a $\mathcal{B}$ como $\mathcal{B} = \{b_i\}_{i \in I}$.

A lo largo de estas notas consideraremos siempre bases ordenadas, sin embargo omitiremos el término "ordenada" cada vez que nos refiramos a una base de un $\mathbb{K}$-espacio vectorial.

**Ejemplos 1.22.**

1. En $\mathbb{K}^n$, consideremos los vectores $e_1, \dots, e_n$ dados por

$$\begin{aligned} e_1 &= (1, 0, 0, \dots, 0, 0), \\ e_2 &= (0, 1, 0, \dots, 0, 0), \\ \vdots\ & \quad \vdots \\ e_n &= (0, 0, 0, \dots, 0, 1). \end{aligned}$$

   Entonces, $\{e_1, \dots, e_n\}$ es una base de $\mathbb{K}^n$, a la que llamaremos la **base canónica** de $\mathbb{K}^n$.

2. En $\mathbb{K}^{n\times m}$, dados $i = 1, \ldots, n$ y $j = 1, \ldots, m$, consideremos la matriz $E^{ij}$ dada por
$$(E^{ij})_{kl} = \begin{cases} 1 & \text{si } k = i \text{ y } l = j, \\ 0 & \text{si } k \neq i \text{ o } l \neq j. \end{cases} \tag{1.2}$$
Entonces, $\left\{E^{ij} : \ i = 1, \ldots, n, \ j = 1, \ldots, m\right\}$ es una base de $\mathbb{K}^{n\times m}$.

3. El conjunto $\{p_0(x) = 1, \ p_1(x) = x, \ p_2(x) = x^2, \ \ldots, \ p_m(x) = x^m\}$ es una base de $\mathbb{K}^{(m)}[x]$, el subespacio de $\mathbb{K}[x]$ formado por los polinomios de grado menor o igual a $m$ (y por el polinomio nulo).

4. La familia dada en (1.1) es una base de $\mathbb{K}[x]$. Recordemos que en un ejemplo anterior vimos que es linealmente independiente. Por otra parte, es fácil ver que esta familia es un sistema de generadores de $\mathbb{K}[x]$, ya que todo polinomio puede escribirse (de manera única) como una combinación lineal de elementos del conjunto (1.1).

Acabamos de ver que $\mathbb{K}[x]$ es un $\mathbb{K}$-espacio vectorial con una "base infinita", es decir, una base con una cantidad infinita de elementos. ¿Esto significa que no existe una base finita de $\mathbb{K}[x]$? El siguiente resultado prueba que este es el caso.

**Teorema 1.23.** *Sea $V$ un $\mathbb{K}$-espacio vectorial generado por un conjunto finito de vectores $\{b_1, \ldots, b_m\}$. Entonces, todo conjunto linealmente independiente de vectores de $V$ contiene a lo sumo $m$ vectores.*

*Demostración.* Para probar este resultado, alcanza con probar que todo subconjunto $S \subset V$ que contiene más de $m$ vectores es linealmente dependiente.

Dados $S \subset V$ y $n > m$, supongamos que $v_1, \ldots, v_n \in S$ son distintos entre sí. Como $\{b_1, \ldots, b_m\}$ genera a $V$, para cada $j = 1, \ldots, n$ existen $\alpha_{ij} \in \mathbb{K}$ tales que
$$v_j = \sum_{i=1}^{m} \alpha_{ij} b_i.$$
Dados $c_1, \ldots, c_n \in \mathbb{K}$ cualesquiera, tenemos que
$$\begin{aligned} \sum_{j=1}^{n} c_j v_j = \sum_{j=1}^{n} c_j \left( \sum_{i=1}^{m} \alpha_{ij} b_i \right) &= \sum_{j=1}^{n} \sum_{i=1}^{m} (\alpha_{ij} c_j)\, b_i \\ &= \sum_{i=1}^{m} \left( \sum_{j=1}^{n} \alpha_{ij} c_j \right) b_i. \end{aligned}$$
Si llamamos $A = (\alpha_{ij}) \in \mathbb{K}^{m\times n}$ y $\vec{c} = (c_1, \ldots, c_n)^\top \in \mathbb{K}^{n\times 1}$, notemos que el escalar que multiplica al vector $b_i$ en la ecuación anterior es justamente la $i$-ésima coordenada del vector $A\vec{c} \in \mathbb{K}^m$.

Como $n > m$ el sistema de ecuaciones lineales homogéneo $A\vec{c} = \vec{0}$ tiene infinitas soluciones. En particular, existen $c_1, \ldots, c_n \in \mathbb{K}$ no todos nulos, tales que
$$\sum_{j=1}^{n} \alpha_{ij} c_j = 0, \quad \text{para } i = 1, \ldots, m.$$
Entonces, tenemos que existen $c_1, \ldots, c_n \in \mathbb{K}$, no todos nulos, tales que $\sum_{j=1}^{n} c_j v_j = \vec{0}$. Por lo tanto, $S$ es linealmente dependiente. □

__Nota__: Usualmente utilizaremos el símbolo $\top$ para denotar a la traspuesta de una matriz o de un vector. Para ahorrar espacio, muchas veces escribiremos $(x_1, \ldots, x_n)^\top$ en lugar de escribir el vector columna correspondiente.

El resultado anterior nos permite separar a los $\mathbb{K}$-espacios vectoriales en dos clases, los de dimensión finita y los de dimensión infinita.

**Definición 1.24.** Dado un $\mathbb{K}$-espacio vectorial $V$, diremos que $V$ es de **dimensión finita** si $V$ admite una base finita. En caso contrario, diremos que $V$ es de **dimensión infinita**.

A continuación, probaremos un corolario que será fundamental para cuantificar la dimensión de un $\mathbb{K}$-espacio vectorial de dimensión finita.

**Corolario 1.25.** *Sea $V$ un $\mathbb{K}$-espacio vectorial de dimensión finita. Entonces, dos bases cualesquiera de $V$ tienen el mismo cardinal.*

*Demostración.* Como $V$ es de dimensión finita, existe una base $\mathcal{B} = \{b_1, \ldots, b_m\}$ de $V$. Supongamos que $\mathcal{E} = \{e_1, \ldots, e_n\}$ es otra base de $V$. El teorema anterior asegura que $n \leq m$, pues $\mathcal{E}$ es un conjunto linealmente independiente.

Usando el mismo razonamiento, ahora usando a $\mathcal{E}$ como sistema de generadores de $V$, resulta que $m \leq n$ ya que $\mathcal{B}$ es linealmente independiente. Por lo tanto, $n = m$. Es decir, las dos bases tienen el mismo cardinal. $\square$

**Definición 1.26.** Dado un $\mathbb{K}$-espacio vectorial $V$ de dimensión finita, diremos que la **dimensión de** $V$ es el cardinal de cualquiera de sus bases. La anotaremos $\dim V$, ó $\dim_{\mathbb{K}} V$ si fuera necesario remarcar el cuerpo en cuestión.

**Ejemplos 1.27.** Calculemos las dimensiones de algunos $\mathbb{K}$-espacios vectoriales de dimensión finita.

1. En primer lugar, notemos que $\dim\{\vec{0}\} = 0$ ya que no contiene subconjuntos linealmente independientes.

2. Teniendo en cuenta las bases calculadas en los Ejemplos 1.22, tenemos que
$$\dim \mathbb{K}^n = n, \quad \dim \mathbb{K}^{n\times m} = n.m, \quad \text{y} \quad \dim \mathbb{K}^{(m)}[x] = m + 1.$$

3. Si consideramos su representación cartesiana, cada número complejo $z$ puede escribirse como
$$z = x + iy, \qquad x, y \in \mathbb{R}.$$
Además, la adición y el producto por escalares (reales) se realizan coordenada a coordenada, es decir,
$$\begin{aligned} z_1 + z_2 &= (x_1 + iy_1) + (x_2 + iy_2) = (x_1 + x_2) + i(y_1 + y_2), \\ \alpha z &= \alpha(x + iy) = \alpha x + i\alpha y, \qquad \alpha \in \mathbb{R}. \end{aligned}$$
Entonces, podemos pensar a $\mathbb{C} = \{z = x + iy : \ x, y \in \mathbb{R}\}$ como un $\mathbb{R}$-espacio vectorial, y además $\{1, i\}$ es una base de $\mathbb{C}$ en este sentido.

   Por lo tanto, $\dim_{\mathbb{R}} \mathbb{C} = 2$.

**Lema 1.28.** *Dado un $\mathbb{K}$-espacio vectorial $V$, sea $\{v_1, \ldots, v_m\} \subset V$ un conjunto linealmente independiente. Luego, si $w \in V \setminus \overline{\{v_1, \ldots, v_m\}}$ entonces*
$$\{v_1, \ldots, v_m, w\} \quad \textit{es linealmente independiente.}$$

*Demostración.* Supongamos que $\alpha_1, \ldots, \alpha_m, \alpha_{m+1} \in \mathbb{K}$ son tales que

$$\alpha_1 v_1 + \ldots + \alpha_m v_m + \alpha_{m+1} w = \vec{0}.$$

Entonces, $-\alpha_{m+1} w = \alpha_1 v_1 + \ldots + \alpha_m v_m$, y como $w \notin \overline{\{v_1, \ldots, v_m\}}$, resulta que $\alpha_{m+1} = 0$. Luego, la igualdad $\alpha_1 v_1 + \ldots + \alpha_m v_m = \vec{0}$ y la independencia lineal de $\{v_1, \ldots, v_m\}$ implican que $\alpha_1 = \ldots = \alpha_m = 0$.

Por lo tanto, $\{v_1, \ldots, v_m, w\}$ es linealmente independiente. □

**Teorema 1.29.** *Si $V$ es un $\mathbb{K}$-espacio vectorial de dimensión finita, entonces todo conjunto de vectores linealmente independiente en $V$ puede extenderse a una base de $V$.*

*Demostración.* Supongamos que $V$ es un $\mathbb{K}$-espacio vectorial con $\dim V = n$. Dado un conjunto linealmente independiente $S$, $S \subset V$, el Teorema 1.23 garantiza que $S$ tiene a lo sumo $n$ elementos. Supongamos entonces que $S = \{e_1, \ldots, e_k\}$ con $k \leq n$.

Si $\overline{S} = V$ entonces $S$ es una base de $V$ y no es necesario extenderlo (notemos que en este caso $k = n$). Por otra parte, si $\overline{S} \neq V$ (o lo que es lo mismo, $k < n$), existe un vector

$$e_{k+1} \in V \setminus \overline{S}.$$

Entonces, aplicando el Lema 1.28, tenemos que $\{e_1, \ldots, e_k, e_{k+1}\}$ es linealmente independiente.

Repitiendo el razonamiento anterior otras $n - (k+1)$ veces, completaremos el conjunto $S$ a una base de $V$. □

**Corolario 1.30.** *Sea $V$ un $\mathbb{K}$-espacio vectorial de dimensión finita. Si $W$ es un subespacio propio de $V$, entonces*

$$0 < \dim W < \dim V.$$

Dados dos subespacios $W_1$ y $W_2$ de un $\mathbb{K}$-espacio vectorial $V$, vimos que el menor subespacio que contiene tanto a $W_1$ como a $W_2$ es la suma $W_1 + W_2$. Veamos qué relación hay entre sus dimensiones, cuando éstas son finitas.

**Teorema 1.31.** *Dado un $\mathbb{K}$-espacio vectorial $V$, sean $W_1$ y $W_2$ dos subespacios de $V$, ambos de dimensión finita. Entonces, $W_1 + W_2$ también es un subespacio de dimensión finita, y además,*

$$\dim(W_1 + W_2) = \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2). \tag{1.3}$$

*Demostración.* La idea de la prueba es comenzar con una base de $W_1 \cap W_2$ y extenderla, por un lado a una base de $W_1$, y por el otro a una base de $W_2$. Finalmente hay que probar que la unión de los conjuntos resultantes es una base del subespacio suma, con el cardinal indicado en (1.3).

Como $W_1 \cap W_2$ es un subespacio de $W_1$, éste también es de dimensión finita. Supongamos que $\{u_1, \ldots, u_k\}$ es una base de $W_1 \cap W_2$. Luego, por el teorema anterior, podemos extenderla a una base de $W_1$,

$$\{u_1, \ldots, u_k, v_1, \ldots, v_m\},$$

y también a una base de $W_2$,

$$\{u_1, \ldots, u_k, w_1, \ldots, w_n\}.$$

De esta manera, es fácil convencerse de que

$$\{u_1, \ldots, u_k, v_1, \ldots, v_m, w_1, \ldots, w_n\}, \tag{1.4}$$

es un conjunto de generadores para $W_1 + W_2$. Pero además el conjunto (1.4) es linealmente independiente: supongamos que existen $\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_m, \gamma_1, \ldots, \gamma_n \in \mathbb{K}$ tales que

$$\sum_{i=1}^{k} \alpha_i u_i + \sum_{j=1}^{m} \beta_j v_j + \sum_{l=1}^{n} \gamma_l w_l = \vec{0}.$$

Luego, tenemos que

$$\overbrace{\sum_{i=1}^{k} \alpha_i u_i + \sum_{j=1}^{m} \beta_j v_j}^{\in W_1} = \overbrace{-\sum_{l=1}^{n} \gamma_l w_l}^{\in W_2} \in W_1 \cap W_2, \tag{1.5}$$

y existen escalares $\delta_1, \ldots, \delta_k$ tales que $-\sum_{l=1}^{n} \gamma_l w_l = \sum_{i=1}^{k} \delta_i u_i$. Entonces,

$$\sum_{i=1}^{k} \delta_i u_i + \sum_{l=1}^{n} (-\gamma_l) w_l = \vec{0},$$

y por la independencia lineal de $\{u_1, \ldots, u_k, w_1, \ldots, w_n\}$ resulta que $\delta_1 = \ldots = \delta_k = \gamma_1 = \ldots = \gamma_n = 0$. Además, volviendo a (1.5), vemos que

$$\sum_{i=1}^{k} \alpha_i u_i + \sum_{j=1}^{m} \beta_j v_j = \vec{0},$$

y la independencia lineal de $\{u_1, \ldots, u_k, v_1, \ldots, v_m\}$ garantiza que $\alpha_1 = \ldots = \alpha_k = \beta_1 = \ldots = \beta_m = 0$. Por lo tanto, el conjunto (1.4) es linealmente independiente.

Tenemos entonces que (1.4) es una base de $W_1 + W_2$ y, en consecuencia,

$$\dim(W_1 + W_2) = k + m + n = (k + m) + (k + n) - k = \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2).$$

□

## 1.3. Coordenadas con respecto a una base – Cambio de base

Supongamos que $V$ es un $\mathbb{K}$-espacio vectorial con $\dim V = n$ y que $\mathcal{B} = \{b_1, \ldots, b_n\}$ es una base de $V$. Para lo que sigue, será esencial recordar que una base es un conjunto ordenado, y tendremos que tener en cuenta este orden específico.

Dado un vector $v \in V$, sabemos que existen $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$ tales que

$$v = \sum_{i=1}^{n} \alpha_i b_i. \tag{1.6}$$

Notemos que estos escalares son **únicos**, ya que si $v$ admite otra escritura alternativa

$$v = \sum_{i=1}^{n} \beta_i b_i,$$

con $\beta_1, \ldots, \beta_n \in \mathbb{K}$, entonces

$$\sum_{i=1}^{n} (\alpha_i - \beta_i) b_i = \sum_{i=1}^{n} \alpha_i b_i - \sum_{i=1}^{n} \beta_i b_i = v - v = \vec{0},$$

y la independencia lineal de la base $\mathcal{B}$ asegura que $\alpha_i - \beta_i = 0$ para cada $i = 1, \ldots, n$.

**Definición 1.32.** Dado un vector $v \in V$, a los (únicos) escalares $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$ que satisfacen (1.6) los llamaremos las **coordenadas de** $v$ **con respecto a la base** $\mathcal{B}$, y las ordenaremos en un vector columna $[v]_{\mathcal{B}} \in \mathbb{K}^{n\times 1}$:

$$[v]_{\mathcal{B}} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix},$$

respetando el orden de los vectores de la base $\mathcal{B}$, es decir, si $b_i$ es el $i$-ésimo vector de la base $\mathcal{B}$ entonces la $i$**-ésima coordenada de** $v$ será el escalar $\alpha_i \in \mathbb{K}$ que multiplica a $b_i$ en (1.6).

Dados $v, w \in V$ y $\gamma \in \mathbb{K}$, supongamos que los vectores se escriben como

$$v = \sum_{i=1}^{n} \alpha_i b_i \qquad \text{y} \qquad w = \sum_{i=1}^{n} \beta_i b_i,$$

con respecto a la base $\mathcal{B}$. Luego,

$$\gamma v + w = \gamma \left( \sum_{i=1}^{n} \alpha_i b_i \right) + \sum_{i=1}^{n} \beta_i b_i = \sum_{i=1}^{n} (\gamma\alpha_i + \beta_i) b_i.$$

En consecuencia las coordenadas del vector $\gamma v + w$ están determinadas por las de $v$ y $w$, más precisamente,

$$[\gamma v + w]_{\mathcal{B}} = \gamma [v]_{\mathcal{B}} + [w]_{\mathcal{B}}.$$

Por lo tanto, la aplicación de $V$ en $\mathbb{K}^{n\times 1}$ dada por

$$v \mapsto [v]_{\mathcal{B}} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix},$$

preserva las operaciones de estos dos $\mathbb{K}$-espacios vectoriales.

**Ejemplos 1.33.** Calculemos las coordenadas de ciertos vectores en algunos casos concretos.

1. Consideremos la base canónica $\mathcal{E} = \{e_1 = (1,0,0), e_2 = (0,1,0), e_3 = (0,0,1)\}$ en $\mathbb{R}^3$. Luego, las coordenadas de un vector genérico $(x,y,z) \in \mathbb{R}^3$ son

$$[(x,y,z)]_{\mathcal{E}} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

En cambio, si consideramos la base $\mathcal{B} = \{b_1 = (0,0,1), b_2 = (1,0,0), b_3 = (0,1,0)\}$, entonces

$$[(x,y,z)]_{\mathcal{B}} = \begin{pmatrix} z \\ x \\ y \end{pmatrix}.$$

2. En $\mathbb{C}^{2\times 2}$ consideremos la base $\mathcal{E} = \{E^{11}, E^{12}, E^{21}, E^{22}\}$, donde $E^{ij}$ es la dada en (1.2). Luego, si

$$A = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix},$$

entonces las coordenadas de $A$ con respecto a la base $\mathcal{E}$ son:

$$[A]_{\mathcal{E}} = \begin{pmatrix} 0 \\ i \\ -i \\ 0 \end{pmatrix}.$$

3. En $\mathbb{Q}^{(3)}[x]$, el $\mathbb{Q}$-espacio vectorial de polinomios sobre $\mathbb{Q}$ de grado menor o igual que 3, consideremos la base $\mathcal{B} = \{q_0(x) = \frac{1}{2}, q_1(x) = \frac{1}{3}x, q_2(x) = \frac{1}{4}x^2, q_3(x) = x^3\}$. Entonces, las coordenadas de un polinomio genérico

$$p(x) = a_0 + a_1x + a_2x^2 + a_3x^3,$$

con respecto a la base $\mathcal{B}$ son

$$[p]_{\mathcal{B}} = \begin{pmatrix} 2a_0 \\ 3a_1 \\ 4a_2 \\ a_3 \end{pmatrix}.$$

Supongamos ahora que $\mathcal{B} = \{b_1, \ldots, b_n\}$ y $\mathcal{B}' = \{b'_1, \ldots, b'_n\}$ son dos bases distintas de un $\mathbb{K}$-espacio vectorial $V$ con $\dim V = n$. Fijada la base $\mathcal{B}$ de $V$, cada elemento $b'_j$ de $\mathcal{B}'$ se escribe de manera única como combinación lineal de los vectores de la base $\mathcal{B}$, es decir, existen únicos escalares $P_{ij} \in \mathbb{K}$, $i = 1, \ldots, n$, tales que

$$b'_j = \sum_{i=1}^{n} P_{ij} b_i, \qquad j = 1, \ldots, n.$$

Dado un vector $v \in V$, cuyas coordenadas con respecto a la base $\mathcal{B}'$ están dadas por $[v]_{\mathcal{B}'} = (\alpha'_1, \ldots, \alpha'_n)^\top$, notemos que

$$\begin{aligned} v &= \sum_{j=1}^{n} \alpha'_j b'_j = \sum_{j=1}^{n} \alpha'_j \left( \sum_{i=1}^{n} P_{ij} b_i \right) = \sum_{j=1}^{n} \sum_{i=1}^{n} \left( P_{ij} \alpha'_j \right) b_i \\ &= \sum_{i=1}^{n} \left( \sum_{j=1}^{n} P_{ij} \alpha'_j \right) b_i. \end{aligned}$$

Como las coordenadas del vector $v$ con respecto a la base $\mathcal{B}$ están unívocamente determinadas por, digamos $[v]_{\mathcal{B}} = (\alpha_1, \ldots, \alpha_n)^\top$, resulta que

$$\alpha_i = \sum_{j=1}^{n} P_{ij} \alpha'_j, \qquad i = 1, \ldots, n.$$

Luego, si $P := (P_{ij}) \in \mathbb{K}^{n \times n}$ tenemos que

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

**Observación 1.34.** Como $\mathcal{B}$ y $\mathcal{B}'$ son conjuntos linealmente independientes, sabemos además que

$$[v]_{\mathcal{B}} = \vec{0} \qquad \text{si y sólo si} \qquad [v]_{\mathcal{B}'} = \vec{0}.$$

Por lo tanto, $P$ es una matriz invertible. Además,

$$[v]_{\mathcal{B}'} = (P^{-1}P)[v]_{\mathcal{B}'} = P^{-1}(P[v]_{\mathcal{B}'}) = P^{-1}[v]_{\mathcal{B}},$$

en consecuencia, $P^{-1}$ también es una matriz que envía a las coordenadas de un vector con respecto a una base en las coordenadas de ese vector con respecto a otra base.

En conclusión, hemos probado buena parte del siguiente resultado.

**Teorema 1.35.** *Dado un $\mathbb{K}$-espacio vectorial $V$ con $\dim V = n$, sean $\mathcal{B}$ y $\mathcal{B}'$ dos bases (ordenadas) de $V$. Entonces, existe una única matriz invertible $P \in \mathbb{K}^{n\times n}$ tal que*

$$[v]_{\mathcal{B}} = P[v]_{\mathcal{B}'} \qquad \textit{para todo } v \in V. \tag{1.7}$$

*Demostración.* Como ya vimos antes del enunciado, dadas dos bases $\mathcal{B}$ y $\mathcal{B}'$ de $V$, existe una matriz invertible $P \in \mathbb{K}^{n\times n}$ que verifica (1.7). Sólo nos falta probar que esta matriz es única. Supongamos entonces que existen $P_1, P_2 \in \mathbb{K}^{n\times n}$ que cumplen (1.7). Entonces, para cada $v \in V$,

$$(P_1 - P_2)[v]_{\mathcal{B}'} = P_1[v]_{\mathcal{B}'} - P_2[v]_{\mathcal{B}'} = [v]_{\mathcal{B}} - [v]_{\mathcal{B}} = \vec{0}.$$

Esto implica que $P_1 = P_2$, como queríamos probar. $\square$

**Definición 1.36.** A la única matriz $P \in \mathbb{K}^{n\times n}$ que satisface (1.7) la llamaremos la **matriz cambio de base** de $\mathcal{B}'$ en $\mathcal{B}$, y la anotaremos $P_{\mathcal{B},\mathcal{B}'}$.

Si revisamos como fue construída la matriz cambio de base, veremos que

$$\mathrm{Col}_j(P_{\mathcal{B},\mathcal{B}'}) = [b_j']_{\mathcal{B}}, \qquad j = 1, \ldots, n,$$

es decir, las columnas de la matriz $P_{\mathcal{B},\mathcal{B}'}$ son los vectores de coordenadas con respecto a la base $\mathcal{B}$, de los elementos de la base $\mathcal{B}'$.

**Ejemplo 1.37.** En $\mathbb{R}^3$ sean $\mathcal{E}$ y $\mathcal{B}$ las bases dadas por

$$\mathcal{E} = \{e_1 = (1,0,0), e_2 = (0,1,0), e_3 = (0,0,1)\},$$
$$\mathcal{B} = \{b_1 = (1,0,0), e_2 = (1,1,0), e_3 = (1,1,1)\}.$$

Dado $(x,y,z) \in \mathbb{R}^3$, recordemos que

$$[(x,y,z)]_{\mathcal{E}} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Calculemos ahora las coordenadas con respecto a la base $\mathcal{B}$. Supongamos que $[(x,y,z)]_{\mathcal{B}} = (\alpha, \beta, \gamma)^\top$, es decir,

$$(x,y,z) = \alpha(1,0,0) + \beta(1,1,0) + \gamma(1,1,1) = (\alpha+\beta+\gamma, \beta+\gamma, \gamma).$$

Entonces,

$$[(x,y,z)]_{\mathcal{B}} = \begin{pmatrix} x-y \\ y-z \\ z \end{pmatrix}.$$

En particular,

$$[e_1]_{\mathcal{B}} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad [e_2]_{\mathcal{B}} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \quad [e_3]_{\mathcal{B}} = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}.$$

Por lo tanto, la matriz cambio de base de $\mathcal{E}$ en $\mathcal{B}$ es

$$P_{\mathcal{B},\mathcal{E}} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Además, es fácil ver que la matriz cambio de base de $\mathcal{B}$ en $\mathcal{E}$ es $P_{\mathcal{E},\mathcal{B}} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$.

El Teorema 1.35 también admite cierta clase de recíproca.

**Teorema 1.38.** *Dada una matriz invertible $P \in \mathbb{K}^{n\times n}$, si $V$ en un $\mathbb{K}$-espacio vectorial con $\dim V = n$ y $\mathcal{B}$ es una base de $V$, entonces existe una única base $\mathcal{B}'$ de $V$ tal que*

$$[v]_{\mathcal{B}} = P[v]_{\mathcal{B}'} \qquad \textit{para todo } v \in V, \tag{1.8}$$

$$[v]_{\mathcal{B}'} = P^{-1}[v]_{\mathcal{B}} \qquad \textit{para todo } v \in V. \tag{1.9}$$

*Demostración.* Supongamos que $P = (P_{ij}) \in \mathbb{K}^{n\times n}$ es invertible. Dada una base $\mathcal{B} = \{b_1, \ldots, b_n\}$ de $V$, definamos los vectores

$$b'_j = \sum_{i=1}^{n} P_{ij} b_i, \qquad j = 1, \ldots, n. \tag{1.10}$$

En primer lugar, veamos que $\mathcal{B}' = \{b'_1, \ldots, b'_n\}$ es una base de $V$. Para ello, aprovechando que $P$ es inversible, notemos que

$$\begin{aligned}\sum_{j=1}^{n} (P^{-1})_{jk} b'_j &= \sum_{j=1}^{n} (P^{-1})_{jk} \left(\sum_{i=1}^{n} P_{ij} b_i\right) = \sum_{j=1}^{n}\sum_{i=1}^{n} \left((P^{-1})_{jk} P_{ij}\right) b_i \\ &= \sum_{i=1}^{n} \left(\sum_{j=1}^{n} P_{ij}(P^{-1})_{jk}\right) b_i = b_k,\end{aligned}$$

ya que $\sum_{j=1}^{n} P_{ij}(P^{-1})_{jk}$ es el lugar $(i,k)$ de la matriz $P^{-1}P = I$.

El cálculo anterior muestra que la base $\mathcal{B}$ está contenida en el subespacio $\overline{\{b'_1, \ldots, b'_n\}}$. Entonces, $V \subseteq \overline{\{b'_1, \ldots, b'_n\}}$. Como además el cardinal de $\mathcal{B}'$ coincide con la dimensión de $V$, el conjunto $\mathcal{B}'$ resulta linealmente independiente, y en consecuencia, es una base de $V$.

Veamos ahora que se cumplen (1.8) y (1.9): dado $j = 1, \ldots, n$,

$$[b'_j]_{\mathcal{B}} = \left[\sum_{i=1}^{n} P_{ij} b_i\right]_{\mathcal{B}} = \sum_{i=1}^{n} P_{ij}[b_i]_{\mathcal{B}} = \sum_{i=1}^{n} P_{ij} e_i^{\top} = \mathrm{Col}_j(P) = P \cdot e_j^{\top} = P[b'_j]_{\mathcal{B}'},$$

donde $\{e_1, \ldots, e_n\}$ denota a la base canónica de $\mathbb{K}^n$. Luego, si $v \in V$ es un vector arbitrario con coordenadas $[v]_{\mathcal{B}'} = (\alpha_1, \ldots, \alpha_n)^{\top}$, entonces

$$\begin{aligned}[v]_{\mathcal{B}} &= \left[\sum_{j=1}^{n} \alpha_j b'_j\right]_{\mathcal{B}} = \sum_{j=1}^{n} \alpha_j [b'_j]_{\mathcal{B}} = \sum_{j=1}^{n} \alpha_j P[b'_j]_{\mathcal{B}'} = P\left(\sum_{j=1}^{n} \alpha_j [b'_j]_{\mathcal{B}'}\right) \\ &= P\left[\sum_{j=1}^{n} \alpha_j b'_j\right]_{\mathcal{B}'} = P[v]_{\mathcal{B}'},\end{aligned}$$

lo que prueba (1.8). Multiplicando a izquierda por $P^{-1}$ en la ecuación anterior, obtenemos (1.9). □

**Ejemplo 1.39.** Dado $\theta \in \mathbb{R}$, consideremos la matriz $P \in \mathbb{R}^{2\times 2}$ dada por

$$P = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}.$$

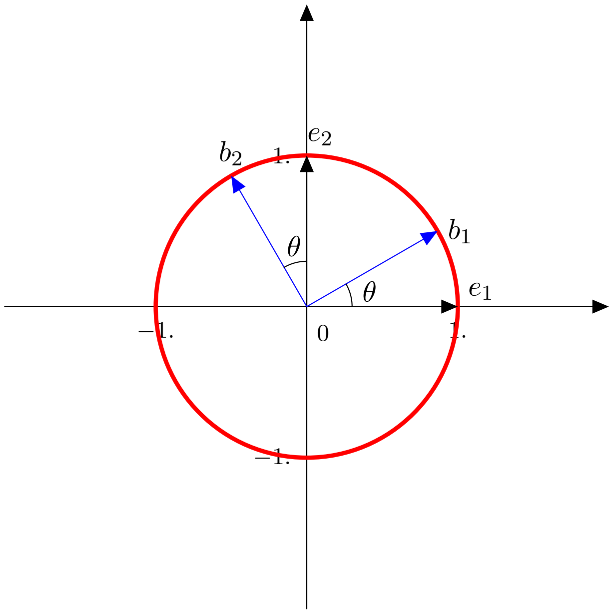



Figura 1.2: La base $\mathcal{B}_\theta = \{b_1, b_2\}$ se obtiene al rotar a la base canónica $\mathcal{E} = \{e_1, e_2\}$ un ángulo de $\theta$ radianes en sentido antihorario.

Como $\det P = \cos^2\theta + \operatorname{sen}^2\theta = 1$, $P$ es invertible. Además, es fácil ver que

$$P^{-1} = \begin{pmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{pmatrix}.$$

De acuerdo al teorema anterior,

$$\mathcal{B}_\theta = \{(\cos\theta, -\operatorname{sen}\theta), (\operatorname{sen}\theta, \cos\theta)\},$$

es una base de $\mathbb{R}^2$, ya que los vectores de $\mathcal{B}_\theta$ son los dados por (1.10) para esta matriz $P$ particular.

Dadas dos bases $\mathcal{B}$ y $\mathcal{B}'$ de un $\mathbb{K}$-espacio vectorial $V$, para calcular la matriz cambio de base de $\mathcal{B}'$ en $\mathcal{B}$, algunas veces conviene considerar una base adicional $\mathcal{E}$ (que usaremos como base intermedia) ya que

$$P_{\mathcal{B},\mathcal{B}'} = P_{\mathcal{B},\mathcal{E}} P_{\mathcal{E},\mathcal{B}'} = (P_{\mathcal{E},\mathcal{B}})^{-1} P_{\mathcal{E},\mathcal{B}'}.$$

**Ejemplo 1.40.** Sea $V$ el subespacio de $\mathbb{R}^{2\times 2}$ formado por las matrices triangulares superiores:

$$V = \left\{ \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} : \ a, b, c \in \mathbb{R} \right\},$$

y consideremos las bases de $V$ dadas por:

$$\mathcal{B} = \left\{ B_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \ B_2 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \ B_3 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \right\},$$
$$\mathcal{B}' = \left\{ C_1 = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \ C_2 = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}, \ C_3 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \right\}.$$

Si queremos calcular $P_{\mathcal{B}',\mathcal{B}}$, como tercera base utilicemos a

$$\mathcal{E} = \left\{ E_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},\ E_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},\ E_3 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\},$$

ya que es inmediato que

$$P_{\mathcal{E},\mathcal{B}} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix} \qquad \text{y} \qquad P_{\mathcal{E},\mathcal{B}'} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

Además, como $\det(P_{\mathcal{E},\mathcal{B}'}) = 1$, resulta que

$$(P_{\mathcal{E},\mathcal{B}'})^{-1} = \frac{1}{\det(P_{\mathcal{E},\mathcal{B}'})} \operatorname{Adj}(P_{\mathcal{E},\mathcal{B}'})^{\top} = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 1 & -1 \\ -1 & 0 & 1 \end{pmatrix}^{\top} = \begin{pmatrix} 0 & 1 & -1 \\ -1 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix}.$$

Por lo tanto,

$$P_{\mathcal{B}',\mathcal{B}} = (P_{\mathcal{E},\mathcal{B}'})^{-1} P_{\mathcal{E},\mathcal{B}} = \begin{pmatrix} 0 & 1 & -1 \\ -1 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 1 & 1 \\ -1 & -1 & 1 \\ 2 & 0 & -1 \end{pmatrix},$$

con lo cual,

$$\begin{aligned} B_1 &= (-1)C_1 + (-1)C_2 + 2C_3, \\ B_2 &= C_1 + (-1)C_2, \\ B_3 &= C_1 + C_2 + (-1)C_3. \end{aligned}$$

## 1.4. Ejercicios

1. Analizar si los siguientes conjuntos son $\mathbb{R}$-espacios vectoriales con las operaciones $+$ y $\cdot$ usuales.

    *a*) Los puntos de un recta de $\mathbb{R}^2$ que pasa por el origen de coordenadas.

    *b*) Las funciones lineales cuya gráfica pertenece a $\mathbb{R}^2$ y contiene al origen de coordenadas.

    *c*) Los polinomios de grado menor o igual a 3, que tienen el mismo término independiente.

    *d*) Las matrices de $\mathbb{R}^{2\times 2}$ cuya diagonal principal es nula.

    *e*) Las matrices de $\mathbb{R}^{3\times 3}$ inversibles.

2. Sea $\mathrm{GL}(3,\mathbb{C}) := \left\{ A \in \mathbb{C}^{3\times 3} : A \quad \text{es inversible} \right\}$. ¿Es $\mathrm{GL}(3,\mathbb{C})$ un $\mathbb{C}$-espacio vectorial con la suma dada por $A + B := AB$ y el producto por escalares usual?

3. Determinar si existe $t \in \mathbb{R}$ tal que los vectores $(t-1,\ 0,\ 1)$, $(t,\ 1,\ 2)$ y $(-1,\ 1,\ -1)$ sean linealmente independientes.

4. Demostrar que en un espacio vectorial de dimensión $n$, todo conjunto de $n+1$ vectores es linealmente dependiente.

5. Mostrar que $\overline{S}$ es el menor subespacio que contiene a $S$, es decir, probar que $\overline{S}$ es la intersección de todos los subespacios que contienen a $S$.

6. Sea $S$ el subespacio de $\mathbb{R}^4$ generado por $\mathcal{B} = \{(-1,0,-1,1),\ (0,1,0,-1),\ (-1,1,-1,0)\}$. ¿Es $\mathcal{B}$ una base de $S$? En caso contrario, hallar una base de $S$ y extenderla a una de $\mathbb{R}^4$.

7. Probar que el conjunto $\mathcal{B} = \{(1,-1,-1,-1),\ (0,1,1,0),\ (1,2,0,0),\ (0,1,2,-1)\}$ es una base de $\mathbb{C}^4$.

8. Hallar una base para $\mathbb{C}^{2\times 2}$, primero pensándolo como un $\mathbb{R}$-espacio vectorial y luego como un $\mathbb{C}$-espacio vecorial. ¿Qué dimensión tiene $\mathbb{C}^{2\times 2}$ como $\mathbb{R}$-espacio vectorial? ¿y como $\mathbb{C}$-espacio vectorial?

9. Analizar en cada caso si el conjunto correspondiente es linealmente independiente, y hallar el subespacio generado por cada uno ellos. Calcular además la dimensión que tiene cada subespacio.

   *a)* $\mathcal{B} = \{(2,0,1),\ (3,1,2),\ (1,1,1),\ (7,3,5)\} \subset \mathbb{R}^3$.
   
   *b)* $\mathcal{B} = \left\{ \begin{pmatrix} 2 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 2 & 1 \end{pmatrix}, \begin{pmatrix} 3 & -1 \\ 0 & 0 \end{pmatrix} \right\} \subset \mathbb{R}^{2\times 2}$.
   
   *c)* $\mathcal{B} = \left\{1-x,\ 2-x^2,\ x+x^2\right\} \subset \mathbb{R}^{(2)}[x]$.

10. Sea $V$ un $\mathbb{K}$-espacio vectorial de dimensión finita y $U$ un subespacio de $V$. Probar que, si la dimensión de $U$ coincide con la de $V$ entonces $U = V$.

11. Sea $V$ un $\mathbb{K}$-espacio vectorial, supongamos que $\{v_1, v_2, \dots, v_m\}$ es un conjunto linealmente independiente en $V$. Dado $w \in V$, mostrar que

$$\dim \overline{\left\{v_1 + w, v_2 + w, \dots, v_m + w\right\}} \geq m - 1.$$

12. Probar que $W_1 + W_2 = \overline{W_1 \cup W_2}$ y, por lo tanto, $W_1 + W_2$ es un subespacio de $V$.

13. Sean $V_1 = \left\{A \in \mathbb{C}^{2\times 2} : a_{11} + a_{12} = 0\right\}$ y $V_2 = \left\{A \in \mathbb{C}^{2\times 2} : a_{11} + a_{21} = 0\right\}$.

   *a)* Probar que $V_1$ y $V_2$ son subespacios de $\mathbb{C}^{2\times 2}$.
   
   *b)* Hallar $V_1 \cap V_2$ y $V_1 + V_2$.
   
   *c)* Hallar las dimensiones de $V_1$, $V_2$, $V_1 \cap V_2$ y $V_1 + V_2$.

14. Sea $S = \{(x,y,z) : x - y + z = 0\}$.

   *a)* Probar que $S$ es un subespacio de $\mathbb{R}^3$ y hallar una base para $S$.
   
   *b)* Hallar un subespacio $T$ de $\mathbb{R}^3$ tal que $S + T = \mathbb{R}^3$ ¿Es único?

15. Hallar una base de $V_1 + V_2 + V_3$ para los siguientes subespacios de $\mathbb{R}^5$.

   *a)* $V_1 = \overline{\{(1,1,2,0,1),\ (2,0,3,0,1)\}}$, $V_2 = \overline{\{(-1,1,-2,1,1)\}}$ y $V_3 = \overline{\{(0,1,0,1,1)\}}$.
   
   *b)* $V_1 = \overline{\{(1,1,2,0,1),\ (2,0,3,0,1)\}}$, $V_2 = \overline{\{(1,0,-2,1,1)\}}$ y $V_3 = \overline{\{(1,1,1,2,2)\}}$.

16. Sea $C(\mathbb{R})$ el $\mathbb{R}$-espacio vectorial de las funciones continuas de $\mathbb{R}$ en $\mathbb{R}$ dotado con las operaciones usuales. Si $V \subset C(\mathbb{R})$ es el conjunto de funciones pares y $W \subset C(\mathbb{R})$ el de las impares, probar que:

*a*) $V$ y $W$ son subespacios de $C(\mathbb{R})$.

*b*) $V \cap W = \{0\}$.

*c*) $V + W = C(\mathbb{R})$.

17. Sean $W_1$ y $W_2$ subespacios de un espacio vectorial $V$ tales que, $W_1 \cap W_2 = \{0\}$ y $V = W_1 + W_2$. Probar que, dado $v \in V$, existen únicos $w_1 \in W_1$ y $w_2 \in W_2$ tales que $v = w_1 + w_2$.

18. Sea $V = \{p \in \mathbb{Z}_2[x] : \operatorname{gr} p \leq 3\}$.

    *a*) Probar que $V$ es un $\mathbb{Z}_2$-subespacio vectorial de $\mathbb{Z}_2[x]$.

    *b*) Probar que $\mathcal{B}_1 = \left\{1, 1+x, x^2, x^3+1\right\}$ y $\mathcal{B}_2 = \left\{1,\, 1+x\,, 1+x+x^2\,, x^3\right\}$ son bases de $V$.

    *c*) Hallar las coordenadas de $p = x + x^2 + x^3$ en cada una de las bases del ítem anterior.

19. Sea $\mathcal{B} = \left\{1+x,\ 1+x^2,\ x+x^2\right\}$.

    *a*) Probar que $\mathcal{B}$ es una base para $\mathbb{R}^{(2)}[x]$.

    *b*) Encontrar las coordenadas de $p(x) = 3x^2 + 2x - 1$ en la base $\mathcal{B}$.

    *c*) Hallar las coordenadas de los elementos de $\mathcal{B}$ con respecto a la base canónica de $\mathbb{R}^{(2)}[x]$.

    *d*) Escribir a los vectores de la base canónica de $\mathbb{R}^{(2)}[x]$ como combinación lineal de los elementos de $\mathcal{B}$.

20. Sean $\mathcal{E}$ la base canónica de $\mathbb{R}^3$ y $\mathcal{B} = \{(1,1,0),\ (1,1,1),\ (0,1,1)\}$.

    *a*) Hallar la matriz $P_{\mathcal{E},\mathcal{B}}$ de cambio de base de $\mathcal{B}$ en $\mathcal{E}$.

    *b*) Hallar la matriz $P_{\mathcal{B},\mathcal{E}}$ de cambio de base de $\mathcal{E}$ en $\mathcal{B}$.

    *c*) Comprobar que la matriz hallada en el ítem *a*) es la inversa de la del ítem *b*).

    *d*) Comprobar que $P_{\mathcal{B},\mathcal{E}}\,[(1,2,0)]_{\mathcal{E}} = [(1,2,0)]_{\mathcal{B}}$.

21. Sea $\mathcal{E} = \{e_1, \ldots, e_n\}$ la base canónica de $\mathbb{R}^n$ y sean

    $$u_1 = e_2 - e_1, \quad u_2 = e_3 - e_2, \ \ldots\ , \quad u_{n-1} = e_n - e_{n-1}, \quad u_n = e_n\,.$$

    *a*) Probar que $\mathcal{B} = \{u_1, \ldots, u_n\}$ es una base de $\mathbb{R}^n$.

    *b*) Hallar las matrices de cambio de base $P_{\mathcal{E},\mathcal{B}}$ y $P_{\mathcal{B},\mathcal{E}}$ para $n = 3$.

    *c*) Probar que para todo $v \in \mathbb{R}^3$, se tiene que $[v]_{\mathcal{B}} = P_{\mathcal{B},\mathcal{E}}[v]_{\mathcal{E}}$.

22. Hallar las matrices de cambio de base $P_{\mathcal{B}_1,\mathcal{B}_2}$ y $P_{\mathcal{B}_2,\mathcal{B}_1}$ para cada uno de los siguientes casos.

    *a*) En $\mathbb{R}^3$, consideremos los conjuntos $\mathcal{B}_1 = \{(5,3,1), (1,-3,-2), (1,2,1)\}$ y $\mathcal{B}_2 = \{(-2,1,0), (-1,-3,0), (-2,-3,1)\}$.

    *b*) En $\mathbb{C}^{(3)}[x]$, sean $\mathcal{B}_1 = \left\{1, x-1, x^2-x, x^3\right\}$ y $\mathcal{B}_2 = \left\{x, x-g1, x^2, x^3+1\right\}$.

23. Sea $\mathcal{B}_1 = \{(1,0,1), (1,1,0), (0,0,1)\}$ una base de $\mathbb{R}^3$. Si

    $$P_{\mathcal{B}_2,\mathcal{B}_1} = \begin{pmatrix} 1 & 1 & 2 \\ 2 & 1 & 1 \\ -1 & -1 & 1 \end{pmatrix}.$$

    Hallar los vectores de la base $\mathcal{B}_2$.

# Capítulo 2

# Transformaciones Lineales

En este capítulo comenzaremos a estudiar a las funciones entre espacios vectoriales que "preservan" o "respetan" a las operaciones.

**Definición 2.1.** Dados dos $\mathbb{K}$-espacios vectoriales $V$ y $W$, una **transformación lineal** de $V$ en $W$ es una función $T : V \to W$ tal que

i) $T(u+v) = Tu + Tv$, para cualesquiera $u, v \in V$,

ii) $T(\alpha v) = \alpha\, Tv$, para cualquier $v \in V$ y cualquier $\alpha \in \mathbb{K}$.

**Ejemplos 2.2.**

1. Dado un $\mathbb{K}$-espacio vectorial $V$, la **transformación identidad** $I : V \to V$ está dada por $I(v) = v$ para todo $v \in V$. Otra transformación lineal siempre presente es la **transformación nula** $0 : V \to V$ dada por $0(v) = \vec{0}$ para todo $v \in V$.

2. Consideremos la función $D : \mathbb{K}[x] \to \mathbb{K}[x]$ dada por

$$(Dp)(x) = a_1 + 2a_2 x + \ldots + k a_k x^{k-1}, \qquad \text{si } p(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_k x^k.$$

   Esta es una transformación lineal de $\mathbb{K}[x]$ en sí mismo.

3. Cada matriz $A \in \mathbb{K}^{n\times m}$ determina una transformación lineal $T : \mathbb{K}^{m\times 1} \to \mathbb{K}^{n\times 1}$ definida por

$$T(X) = A \cdot X, \qquad X \in \mathbb{K}^{m\times 1}.$$

4. El ejemplo anterior puede generalizarse de varias formas. Por ejemplo, dadas matrices $A \in \mathbb{K}^{n\times m}$ y $B \in \mathbb{K}^{p\times q}$, consideremos las funciones $L_A : \mathbb{K}^{m\times k} \to \mathbb{K}^{n\times k}$ y $R_B : \mathbb{K}^{r\times p} \to \mathbb{K}^{r\times q}$ dadas por:

$$\begin{aligned} L_A(X) &= A \cdot X, \qquad \text{si } X \in \mathbb{K}^{m\times k}, \\ R_B(X) &= X \cdot B, \qquad \text{si } X \in \mathbb{K}^{r\times p}. \end{aligned}$$

   Es fácil ver que tanto $L_A$ como $R_B$ son transformaciones lineales.

5. La función $T : \mathcal{F}(\mathbb{K},\mathbb{K}) \to \mathcal{F}(\mathbb{K},\mathbb{K})$ definida por $(Tf)(x) = f(x+1)$, es una transformación lineal. En cambio, la función $S : \mathcal{F}(\mathbb{K},\mathbb{K}) \to \mathcal{F}(\mathbb{K},\mathbb{K})$ dada por $(Sf)(x) = f(x)+1$ **no** es una transformación lineal.

El siguiente resultado muestra que toda transformación lineal está determinada por su acción sobre los elementos de una base del dominio.

**Teorema 2.3.** *Dado un $\mathbb{K}$-espacio vectorial $V$ con $\dim V = n$, sea $\mathcal{B} = \{b_1, \ldots, b_n\}$ una base de $V$. Dado otro $\mathbb{K}$-espacio vectorial $W$, sean $w_1, \ldots, w_n \in W$ cualesquiera. Entonces, existe una* **única** *transformación lineal $T : V \to W$ tal que*

$$Tb_j = w_j, \qquad j = 1, \ldots, n. \tag{2.1}$$

*Demostración.* Primero, veamos que existe una transformación lineal $T : V \to W$ que satisface (2.1). Dado un vector $v \in V$ arbitrario, sea $[v]_{\mathcal{B}} = (\alpha_1, \ldots, \alpha_n)^\top$ su vector de coordenadas con respecto a la base $\mathcal{B}$, y consideremos la función $T : V \to W$ dada por

$$Tv = \sum_{i=1}^{n} \alpha_i w_i.$$

Como las coordenadas con respecto a una base son únicas, la función $T$ está bien definida, y además cumple (2.1). Sólo falta probar que $T$ es lineal. Para ello, consideremos otro vector $u = \sum_{i=1}^{n} \beta_i b_i$ y un escalar $\gamma \in \mathbb{K}$. Luego, $\gamma u + v = \sum_{i=1}^{n} (\gamma\beta_i + \alpha_i) b_i$ y resulta que

$$T(\gamma u + v) = \sum_{i=1}^{n} (\gamma\beta_i + \alpha_i) w_i = \gamma \sum_{i=1}^{n} \beta_i w_i + \sum_{i=1}^{n} \alpha_i w_i = \gamma Tu + Tv.$$

Por lo tanto, $T$ es una transformación lineal que satisface (2.1).

Veamos ahora que ésta es la única transformación lineal que satisface (2.1). Supongamos que $U : V \to W$ es una transformación lineal que verifica

$$U(b_j) = w_j, \qquad j = 1, \ldots, n.$$

Luego, si $v = \sum_{i=1}^{n} \alpha_i b_i$ es un vector arbitrario de $V$, tenemos que

$$U(v) = U\left(\sum_{i=1}^{n} \alpha_i b_i\right) = \sum_{i=1}^{n} \alpha_i U(b_i) = \sum_{i=1}^{n} \alpha_i w_i = T(v).$$

Por lo tanto, $U = T$. $\square$

**Ejemplo 2.4.** Consideremos la base $\mathcal{B} = \{b_1 = (1,2), b_2 = (0,2)\}$ de $\mathbb{R}^2$. Dados $w_1 = (1,2,3), w_2 = (4,5,6) \in \mathbb{R}^3$, calculemos la única transformación lineal $T : \mathbb{R}^2 \to \mathbb{R}^3$ tal que

$$T(1,2) = (1,2,3) \qquad \text{y} \qquad T(0,2) = (4,5,6).$$

Dado un vector arbitrario $(x,y) \in \mathbb{R}^2$, notemos que

$$(x,y) = x(1,2) + \left(\frac{y}{2} - x\right)(0,2),$$

con lo cual, $[(x,y)]_{\mathcal{B}} = \left(x, \frac{y}{2} - x\right)^\top$. Entonces,

$$\begin{aligned}
T(x,y) &= T\left(x(1,2) + \left(\frac{y}{2} - x\right)(0,2)\right) = x\,T(1,2) + \left(\tfrac{y}{2} - x\right)\,T(0,2) \\
&= x(1,2,3) + \left(\tfrac{y}{2} - x\right)(4,5,6) = \left(x + 4\left(\tfrac{y}{2} - x\right), 2x + 5\left(\tfrac{y}{2} - x\right), 3x + 6\left(\tfrac{y}{2} - x\right)\right) \\
&= \left(2y - 3x, \tfrac{5}{2}y - 3x, 3y - 3x\right).
\end{aligned}$$

Cada transformación lineal determina autómaticamente dos subespacios, uno del dominio y el otro del codominio.

**Definición 2.5.** Dados dos $\mathbb{K}$-espacios vectoriales $V$ y $W$, sea $T : V \to W$ una transformación lineal. El **nucleo (ó espacio nulo)** de $T$ es el conjunto

$$N(T) = \{v \in V : \ Tv = \vec{0}\},$$

y la **imagen** de $T$ es su imagen como función, es decir,

$$\mathrm{Im}(T) = \{w \in W : \ w = Tv \text{ para algún } v \in V\} = \{Tv : \ v \in V\}.$$

**Ejercicio 2.6.**

i) Dada una transformación lineal $T : V \to W$, probar que $N(T)$ es un subespacio de $V$ e $\mathrm{Im}(T)$ es un subespacio de $W$.

ii) Además, si $\dim V < \infty$, mostrar que $\dim(N(T)) \leq \dim V$ y $\dim(\mathrm{Im}(T)) \leq \dim V$.

**Definición 2.7.** Si $V$ es un $\mathbb{K}$-espacio vectorial de dimensión finita, llamaremos **nulidad** de $T$ a la dimensión del núcleo de $T$:

$$\mathrm{nul}(T) := \dim(N(T)),$$

y llamaremos **rango** de $T$ a la dimensión de la imagen de $T$:

$$\mathrm{rg}(T) := \dim(\mathrm{Im}(T)).$$

El próximo resultado muestra que, si $V$ es un $\mathbb{K}$-espacio vectorial de dimensión finita, la suma del rango y la nulidad de cualquier transformación lineal con dominio en $V$ coincide con la dimensión de $V$.

**Teorema 2.8.** *Dados dos $\mathbb{K}$-espacios vectoriales $V$ y $W$, sea $T : V \to W$ una transformación lineal. Si $\dim V < \infty$ entonces*

$$\dim V = \dim \mathrm{Im}(T) + \dim N(T) = \mathrm{rg}(T) + \mathrm{nul}(T). \tag{2.2}$$

*Demostración.* Comencemos considerando una base $\{b_1, \ldots, b_k\}$ de $N(T)$. Por el Teorema 1.29, existen $b_{k+1}, \ldots, b_n \in V$ que extienden la base anterior a una base $\{b_1, \ldots, b_n\}$ de $V$. Veamos ahora que $\{Tb_{k+1}, \ldots, Tb_n\}$ es una base de $\mathrm{Im}(T)$. Sabemos que

$$\mathrm{Im}(T) = T(V) = T\big(\overline{\{b_1, \ldots, b_n\}}\big).$$

Además, como $T$ es lineal, es fácil convencerse de que $T\big(\overline{\{b_1, \ldots, b_n\}}\big) = \overline{\{Tb_1, \ldots, Tb_n\}}$. Pero como $b_1, \ldots, b_k \in N(T)$, resulta que $Tb_1 = \ldots = Tb_k = \vec{0}$. Entonces,

$$\mathrm{Im}(T) = T\big(\overline{\{b_1, \ldots, b_n\}}\big) = \overline{\{Tb_1, \ldots, Tb_n\}} = \overline{\{Tb_{k+1}, \ldots, Tb_n\}}.$$

Esto dice que $\{Tb_{k+1}, \ldots, Tb_n\}$ es un sistema de generadores de $\mathrm{Im}(T)$, así que sólo falta probar que es un conjunto linealmente independiente. Para esto, supongamos que $\alpha_{k+1}, \ldots, \alpha_n \in \mathbb{K}$ son tales que

$$\sum_{i=k+1}^{n} \alpha_i T b_i = \vec{0}.$$

Utilizando una vez más la linealidad de $T$, tenemos que $\vec{0} = \sum_{i=k+1}^{n} \alpha_i T b_i = T\left(\sum_{i=k+1}^{n} \alpha_i b_i\right)$, o equivalentemente, $v := \sum_{i=k+1}^{n} \alpha_i b_i \in N(T)$.

Como $\{b_1, \ldots, b_k\}$ es una base de $N(T)$, existen $\beta_1, \ldots, \beta_k \in \mathbb{K}$ tales que $v = \sum_{j=1}^{k} \beta_j b_j$. Entonces,

$$\sum_{j=1}^{k} \beta_j b_j - \sum_{i=k+1}^{n} \alpha_i b_i = v - v = \vec{0}.$$

Luego, la independencia lineal de $\{b_1, \ldots, b_n\}$ garantiza que

$$\beta_1 = \ldots = \beta_k = \alpha_{k+1} = \ldots = \alpha_n = 0.$$

Por lo tanto, $\{Tb_{k+1}, \ldots, Tb_n\}$ es una base de $\mathrm{Im}(T)$ y además,

$$\mathrm{rg}(T) = \dim(\mathrm{Im}(T)) = n - k = \dim V - \dim N(T) = \dim V - \mathrm{nul}(T),$$

como queríamos probar. $\square$

## 2.1. El espacio de las transformaciones lineales

Dados dos $\mathbb{K}$-espacios vectoriales $V$ y $W$, anotaremos $L(V, W)$ al conjunto de transformaciones lineales de $V$ en $W$:

$$L(V, W) = \{T : V \to W \ : \ T \text{ es una transformación lineal}\}.$$

**Teorema 2.9.** *Dados dos $\mathbb{K}$-espacios vectoriales $V$ y $W$, el conjunto $L(V, W)$ dotado con las operaciones*

$$\begin{aligned} (S + T)(v) &= Sv + Tv, \qquad \text{si } S, T \in L(V, W), \\ (\alpha T)(v) &= \alpha \cdot Tv, \qquad \text{si } \alpha \in \mathbb{K},\ T \in L(V, W), \end{aligned}$$

*es un $\mathbb{K}$-espacio vectorial.*

*Demostración.* La dejamos como ejercicio. Una forma de hacerlo es comprobar que $L(V, W)$ es un subespacio del $\mathbb{K}$-espacio vectorial de funciones $\mathcal{F}(V, W)$. $\square$

Antes de enunciar el próximo resultado, recordemos la definición de la función indicatriz o **delta de Kronecker**:

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

Generalmente, las variables $i$ y $j$ tomarán valores en un subconjunto de índices, el cual puede ser finito ó infinito.

**Teorema 2.10.** *Sean $V$ y $W$ dos $\mathbb{K}$-espacios vectoriales de dimensión finita. Si* $\dim V = n$ *y* $\dim W = m$*, entonces* $\dim L(V, W) = nm$*.*

*Demostración.* Sean $\mathcal{B} = \{b_1, \ldots, b_n\}$ y $\mathcal{B}' = \{w_1, \ldots, w_m\}$ bases de $V$ y $W$, respectivamente. Para cada par $(p, q)$, $p = 1, \ldots, m$, $q = 1, \ldots, n$, definimos $T^{p,q} : V \to W$ como

$$T^{p,q} b_i = \delta_{iq} w_p = \begin{cases} w_p & \text{si } i = q, \\ \vec{0} & \text{si } i \neq q. \end{cases}$$

Por el Teorema 2.3, hay una única transformación lineal que cumple estas condiciones.

A continuación veremos que las $mn$ transformaciones lineales $T^{p,q}$ forman una base de $L(V, W)$. Dada $T \in L(V, W)$, para cada $j = 1, \ldots, n$ supongamos que $Tb_j = \sum_{p=1}^{m} A_{pj} w_p$, es decir,

$$[Tb_j]_{\mathcal{B}'} = (A_{1j}, A_{2j}, \ldots, A_{mj})^\top,$$

y comprobemos que

$$T = \sum_{p=1}^{m} \sum_{q=1}^{n} A_{pq} T^{p,q}.$$

Para ello, alcanza con verificar que coinciden en los vectores de la base $\mathcal{B}$: fijado $j = 1, \ldots, n$,

$$\begin{aligned} \left( \sum_{p=1}^{m} \sum_{q=1}^{n} A_{pq} T^{p,q} \right) b_j &= \sum_{p=1}^{m} \sum_{q=1}^{n} A_{pq} T^{p,q}(b_j) = \sum_{p=1}^{m} \sum_{q=1}^{n} A_{pq} \delta_{jq} w_p \\ &= \sum_{p=1}^{m} A_{pj} w_p = T b_j. \end{aligned}$$

Entonces, $T \in \overline{\{T^{p,q} : \ p = 1, \ldots, m, \ q = 1, \ldots, n\}}$. Por lo tanto,

$$\{T^{p,q} : \ p = 1, \ldots, m, \ q = 1, \ldots, n\}, \tag{2.3}$$

es un sistema de generadores de $L(V, W)$.

Además, este conjunto es linealmente independiente, ya que si suponemos que

$$\sum_{p=1}^{m} \sum_{q=1}^{n} \alpha_{pq} T^{p,q} = 0,$$

para ciertos escalares $\alpha_{pq} \in \mathbb{K}$, entonces, al evaluarla en los vectores de $\mathcal{B}$ resulta:

$$\vec{0} = \left( \sum_{p=1}^{m} \sum_{q=1}^{n} \alpha_{pq} T^{p,q} \right) b_j = \sum_{p=1}^{m} \sum_{q=1}^{n} \alpha_{pq} T^{p,q}(b_j) = \sum_{p=1}^{m} \alpha_{pj} w_p,$$

para cada $j = 1, \ldots, n$. Luego, la independencia lineal de $\mathcal{B}'$ implica que

$$\alpha_{pj} = 0 \quad \text{para todo } p = 1, \ldots, m \text{ y todo } j = 1, \ldots, n.$$

Por lo tanto, el conjunto (2.3) es una base de $L(V, W)$ y, en consecuencia, $\dim L(V, W) = nm$. $\square$

El lema siguiente muestra que la composición de transformaciones lineales, cuando tiene sentido, también es una transformación lineal.

**Lema 2.11.** *Sean $V$, $W$ y $V'$ tres $\mathbb{K}$-espacios vectoriales. Si $S : V \to W$ y $T : W \to V'$ son transformaciones lineales, entonces $T \circ S : V \to V'$ también lo es.*

*Demostración.* Dados $u, v \in V$ y $\alpha \in \mathbb{K}$, tenemos que

$$\begin{aligned} (T \circ S)(\alpha u + v) &= T\big(S(\alpha u + v)\big) = T(\alpha S u + S v) = \alpha T(S u) + T(S v) \\ &= \alpha (T \circ S) u + (T \circ S) v. \end{aligned}$$

Por lo tanto, $T \circ S$ es una transformación lineal. $\square$

**Definición 2.12.** Dado un $\mathbb{K}$-espacio vectorial $V$, si $T : V \to V$ es una transformación lineal, diremos que $T$ es un **operador lineal sobre** $V$.

Para alivianar la notación, al espacio de operadores lineales sobre $V$ lo anotaremos $L(V)$ en lugar de $L(V,V)$. Notemos que el lema anterior, considerando el caso particular $V = W = V'$, dice que la composición funciona como una segunda operación binaria en el espacio $L(V)$, una suerte de "producto" de operadores:

$$\text{si } S, T \in L(V), \text{ entonces } ST := S \circ T.$$

Hay que tener muy en cuenta que esta operación no es conmutativa, en general $ST \neq TS$.

**Ejemplo 2.13.** En $\mathbb{K}[x]$ consideremos los siguientes operadores lineales. En primer lugar, sea $T \in L(\mathbb{K}[x])$ dado por

$$(Tp)(x) = x \cdot p(x), \quad p \in \mathbb{K}[x].$$

En segundo término, consideremos el operador $D \in L(\mathbb{K}[x])$ que en la base usual de $\mathbb{K}[x]$ actúa como:

$$D(1) = \vec{0}, \quad D(x^k) = k\ x^{k-1}, \quad \text{si } k \in \mathbb{N}.$$

Luego, es fácil ver que $DT \neq TD$, Más aún, veamos que $DT - TD = I$. Para cada $k \in \mathbb{N}$,

$$\begin{aligned}(DT - TD)(x^k) &= D\big(T(x^k)\big) - T\big(D(x^k)\big) = D(x \cdot x^k) - T(k\ x^{k-1}) = D(x^{k+1}) - k\ T(x^{k-1}) \\ &= (k+1)x^k - k\ x \cdot x^{k-1} = x^k,\end{aligned}$$

y además $(DT - TD)(1) = D(x) - T(\vec{0}) = 1 - \vec{0} = 1$. Por lo tanto, $DT - DT = I$.

**Notación 2.14.** Dado un $\mathbb{K}$-espacio vectorial $V$, si $T \in L(V)$ anotaremos $T^2$ en lugar de $T \cdot T = T \circ T$. Análogamente, para cada $k \in \mathbb{N}$,

$$T^k := T \cdot T^{k-1}.$$

Además, adoptaremos la convención de que $T^0 = I$, en el caso de que $T \neq 0$.

**Lema 2.15.** *Dado un $\mathbb{K}$-espacio vectorial $V$, sean $S, T, U \in L(V)$ y sea $\alpha \in \mathbb{K}$. Entonces,*

*i)* $S \cdot I = I \cdot S = S$;

*ii)* $S \cdot (T + U) = S \cdot T + S \cdot U$ *y también* $(T + U) \cdot S = T \cdot S + U \cdot S$;

*iii)* $\alpha(S \cdot T) = (\alpha S) \cdot T = S \cdot (\alpha T)$.

*Demostración.* Ejercicio. □

El lema anterior completa la información necesaria para asegurar que $L(V)$ no sólo es un $\mathbb{K}$-espacio vectorial, sino que es una $\mathbb{K}$-álgebra lineal con identidad.

**Definición 2.16.** Dado un cuerpo $\mathbb{K}$, un **algebra lineal sobre** $\mathbb{K}$ es un $\mathbb{K}$-espacio vectorial $\mathfrak{A}$ con una operación binaria adicional (el producto de vectores) $\cdot : \mathfrak{A} \times \mathfrak{A} \to \mathfrak{A}$ tal que,

i) es asociativa,

$$u \cdot (v \cdot w) = (u \cdot v) \cdot w, \quad \text{para cualesquiera } u, v, w \in \mathfrak{A};$$

ii) es distributiva con respecto a la adición,

$$u \cdot (v + w) = u \cdot v + u \cdot w, \quad (v + w) \cdot u = v \cdot u + w \cdot u, \quad \text{para cualesquiera } u, v, w \in \mathfrak{A};$$

iii) para cada $\alpha \in \mathbb{K}$,

$$\alpha(u \cdot v) = (\alpha u) \cdot v = u \cdot (\alpha v), \quad \text{para cualesquiera } u, v \in \mathfrak{A}.$$

Si además existe un elemento distinguido $1 \in \mathfrak{A}$ tal que $1 \cdot v = v \cdot 1 = v$ para todo $v \in \mathfrak{A}$, diremos que $\mathfrak{A}$ es un **álgebra lineal con identidad sobre** $\mathbb{K}$, y al 1 lo llamaremos la identidad de $\mathfrak{A}$. El álgebra $\mathfrak{A}$ se dice **conmutativa** si $v \cdot w = w \cdot v$ para todo $v, w \in \mathfrak{A}$.

El espacio de polinomios $\mathbb{K}[x]$ es un ejemplo de álgebra conmutativa con unidad. En cambio, el conjunto $\mathbb{K}^{n\times n}$ de matrices de $n \times n$ sobre el cuerpo $\mathbb{K}$ es otro ejemplo de un álgebra lineal con identidad, pero no es conmutativa. El siguiente comentario muestra una íntima relación entre los operadores lineales sobre un $\mathbb{K}$-espacio vectorial de dimensión $n$ y $\mathbb{K}^{n\times n}$, la cual estudiaremos en detalle más adelante.

**Observación 2.17.** Si $\mathcal{B} = \{b_1, \ldots, b_n\}$ es una base de $V$, y consideramos los operadores $T^{p,q}$ dados por:

$$T^{p,q} b_i = \delta_{iq} b_p,$$

recordemos que estos $n^2$ operadores forman una base de $L(V)$.

Además, es fácil ver que

$$T^{p,q}T^{r,s} = \begin{cases} 0 & \text{si } q \neq r, \\ T^{p,s} & \text{si } q = r. \end{cases}$$

De hecho, si $i = 1, \ldots, n$ entonces

$$\big(T^{p,q}T^{r,s}\big)b_i = T^{p,q}\big(T^{r,s}b_i\big) = T^{p,q}\big(\delta_{is}b_r\big) = \delta_{is}\big(T^{p,q}b_r\big) = \delta_{is}\big(\delta_{rq}b_p\big) = \delta_{rq}\big(\delta_{is}b_p\big) = \delta_{rq}\big(T^{p,s}b_i\big).$$

La ecuación anterior dice que $T^{p,q}T^{r,s} = 0$ si $q \neq r$, mientras que en el caso $q = r$ obtenemos

$$T^{p,r}T^{r,s}b_i = T^{p,s}b_i \qquad \text{para todo } i = 1, \ldots, n.$$

Esto último implica que $T^{p,r}T^{r,s}v = T^{p,s}v$ para todo $v \in V$, es decir, $T^{p,r}T^{r,s} = T^{p,s}$.

Luego, si $S, T \in L(V)$ se escriben con respecto a la base $\mathcal{B}$ como

$$T = \sum_{p=1}^{n}\sum_{q=1}^{n} A_{pq}T^{p,q} \quad \text{y} \quad S = \sum_{r=1}^{n}\sum_{s=1}^{n} B_{rs}T^{r,s},$$

tenemos que

$$\begin{aligned} TS &= \left(\sum_{p=1}^{n}\sum_{q=1}^{n} A_{pq}T^{p,q}\right)\left(\sum_{r=1}^{n}\sum_{s=1}^{n} B_{rs}T^{r,s}\right) \\ &= \sum_{p=1}^{n}\sum_{q=1}^{n}\sum_{r=1}^{n}\sum_{s=1}^{n} A_{pq}B_{rs}T^{p,q}T^{r,s} = \sum_{p=1}^{n}\sum_{q=1}^{n}\sum_{r=1}^{n}\sum_{s=1}^{n} A_{pq}B_{rs}\delta_{qr}T^{p,s} \\ &= \sum_{p=1}^{n}\sum_{q=1}^{n}\sum_{s=1}^{n} A_{pq}B_{qs}T^{p,s} = \sum_{p=1}^{n}\sum_{s=1}^{n}\left(\sum_{q=1}^{n} A_{pq}B_{qs}\right)T^{p,s} \\ &= \sum_{p=1}^{n}\sum_{s=1}^{n} (AB)_{ps}\, T^{p,s}. \end{aligned}$$

Por lo tanto, la matriz de coordenadas de $TS$ es el producto de las matrices de coordenadas de $T$ y de $S$ (todas con respecto a la base $\mathcal{B}$).

## 2.2. Isomorfismos

Dados dos conjuntos $D$ y $D'$, sabemos que una función $f : D \to D'$ es *inversible* si existe otra función $g : D' \to D$ tal que

$$g \circ f = id_D \quad \text{y} \quad f \circ g = id_{D'}, \tag{2.4}$$

siendo $id_D : D \to D$ la función identidad, es decir $id_D(x) = x$ para todo $x \in D$.

En tal caso, la función $g$ que satisface (2.4) es única, la llamaremos la *función inversa de* $f$ y anotaremos $g = f^{-1}$. Además, recordemos que $f$ es inversible si y sólo si $f$ es biyectiva, es decir,

i) $f$ es inyectiva, i.e. $f(x) = f(y) \Rightarrow x = y$;

ii) $g$ es suryectiva, i.e. $\mathrm{Im}(f) = D'$.

**Proposición 2.18.** *Dados dos $\mathbb{K}$-espacios vectoriales $V$ y $W$, sea $T : V \to W$ una transformación lineal. Si $T$ es inversible, entonces la función inversa $T^{-1} : W \to V$ también es una transformación lineal.*

*Demostración.* Dados dos vectores cualesquiera $w_1, w_2 \in W$ y un escalar $\alpha \in \mathbb{K}$, queremos ver que
$$T^{-1}(w_1 + \alpha\, w_2) = T^{-1}w_1 + \alpha\, T^{-1}w_2.$$
Para $i = 1, 2$, sea $v_i = T^{-1}w_i \in V$. Como $T$ es lineal, sabemos que $T(v_1 + \alpha\, v_2) = Tv_1 + \alpha\, Tv_2 = w_1 + \alpha\, w_2$. Luego,

$$T^{-1}(w_1 + \alpha\, w_2) = v_1 + \alpha\, v_2 = T^{-1}w_1 + \alpha\, T^{-1}w_2,$$

como queríamos probar. □

**Corolario 2.19.** *Sean $V$, $W$ y $V'$ tres $\mathbb{K}$-espacios vectoriales. Si $T : V \to W$ y $S : W \to V'$ son transformaciones lineales inversibles, entonces $ST$ también es inversible. Además*

$$(ST)^{-1} = T^{-1}S^{-1}.$$

*Demostración.* Ejercicio. □

**Definición 2.20.** Dados dos $\mathbb{K}$-espacios vectoriales $V$ y $W$, sea $T : V \to W$ una transformación lineal. Diremos que:

- $T$ es un **monomorfismo** si $T$ es inyectiva;
- $T$ es un **epimorfismo** si $T$ es suryectiva, y;
- $T$ es un **isomorfismo** si $T$ es inversible.

**Teorema 2.21.** *Dados dos $\mathbb{K}$-espacios vectoriales $V$ y $W$, sea $T : V \to W$ una transformación lineal. Las siguientes condiciones son equivalentes:*

*i)* *$T$ es un monomorfismo;*

*ii)* *$N(T) = \{\vec{0}\}$;*

*iii)* *$T$ aplica a cada subconjunto linealmente independiente de $V$ en un subconjunto linealmente independiente de $W$.*

*Demostración.*
*i)* $\leftrightarrow$ *ii)* : Dados $u, v \in V$, notemos que

$$Tu = Tv \quad \Leftrightarrow \quad \vec{0} = Tu - Tv = T(u - v) \quad \Leftrightarrow \quad u - v \in N(T).$$

Es evidente entonces que $T$ es un monomorfismo si y sólo si $N(T)$ es trivial.
*ii)* $\rightarrow$ *iii)* : Supongamos que $N(T) = \{\vec{0}\}$.

Dado un subconjunto linealmente independiente $\{v_1, \ldots, v_m\}$ en $V$, veamos que el conjunto $\{Tv_1, \ldots, Tv_m\}$ es linealmente independiente en $W$. Supongamos que existen escalares $\alpha_1, \ldots, \alpha_m \in \mathbb{K}$ tales que

$$\sum_{i=1}^{m} \alpha_i \, Tv_i = \vec{0}.$$

Luego, $\sum_{i=1}^{m} \alpha_i \, v_i \in N(T)$, y por la hipótesis tenemos que $\sum_{i=1}^{m} \alpha_i \, v_i = \vec{0}$. Como $\{v_1, \ldots, v_m\}$ es linealmente independiente en $V$, resulta que $\alpha_1 = \ldots = \alpha_m = 0$.

Por lo tanto, $\{Tv_1, \ldots, Tv_m\}$ es linealmente independiente en $W$.
*iii)* $\rightarrow$ *ii)* : Por la contrarecíproca, supongamos que $N(T) \neq \{\vec{0}\}$. Es decir, supongamos que existe $v \in V$, $v \neq \vec{0}$ tal que $Tv = \vec{0}$. Entonces, $T$ aplica al conjunto linealmente independiente $\{v\}$ de $V$ en el conjunto linealmente dependiente $\{\vec{0}\}$ de $W$. □

En el Ejemplo 2.13 consideramos el operador lineal $T : \mathbb{K}[x] \rightarrow \mathbb{K}[x]$ definido por

$$(Tp)(x) = x \cdot p(x), \quad p \in \mathbb{K}[x].$$

Este es un ejemplo de un monomorfismo que no es suryectivo (Probarlo!).

A continuación mostraremos que esta situación es imposible para transformaciones lineales entre $\mathbb{K}$-espacios vectoriales de la misma dimensión (finita).

**Teorema 2.22.** *Sean $V$ y $W$ dos $\mathbb{K}$-espacios vectoriales de dimensión finita tales que* $\dim V = \dim W = n$. *Si $T \in L(V, W)$, las siguientes condiciones son equivalentes:*

*i)* *$T$ es un monomorfismo;*

*ii)* *$T$ es un epimorfismo;*

*iii)* *$T$ aplica cada base $\{b_1, \ldots, b_n\}$ de $V$ en una base de $W$;*

*iv)* *existe una base $\{b_1, \ldots, b_n\}$ de $V$ tal que $\{Tb_1, \ldots, Tb_n\}$ es una base de $W$;*

*v)* *$T$ es un isomorfismo.*

*Demostración.*
*i)* $\rightarrow$ *ii)* : Supongamos que $N(T) = \{\vec{0}\}$. Como $\dim V < \infty$, tenemos que

$$\dim V = \dim N(T) + \dim \operatorname{Im}(T) = \dim \operatorname{Im}(T) \leq \dim W,$$

ya que $\operatorname{Im}(T)$ es un subespacio de $W$. Pero como supusimos que $\dim V = \dim W$, resulta que $\operatorname{Im}(T) = W$. Por lo tanto, $T$ es un epimorfismo.
*ii)* $\rightarrow$ *iii)* : Dada una base $\mathcal{B} = \{b_1, \ldots, b_n\}$ de $V$, consideremos el conjunto $\mathcal{B}' = \{Tb_1, \ldots, Tb_n\}$ en $W$. Como suponemos que $T$ es un epimorfismo, veamos que $\mathcal{B}'$ es un conjunto de generadores de $W$. Para cada $w \in W$ existe un vector $v \in V$ tal que $w = Tv$. Luego, existen $\alpha_1, \ldots \alpha_n \in \mathbb{K}$ tales que $v = \sum_{k=1}^{n} \alpha_k b_k$ y

$$w = Tv = T\left(\sum_{k=1}^{n} \alpha_k b_k\right) = \sum_{k=1}^{n} \alpha_k \; Tb_k.$$

Además, como el cardinal de $\mathcal{B}'$ coincide con $\dim W$, es necesario que $\mathcal{B}'$ sea un conjunto linealmente independiente. Por lo tanto, $\mathcal{B}'$ es una base de $W$.

$iii) \rightarrow iv)$ : Es trivial.

$iv) \rightarrow v)$ : Supongamos que $\mathcal{B} = \{b_1, \ldots, b_n\}$ es una base de $V$ tal que $\mathcal{B}' = \{Tb_1, \ldots, Tb_n\}$ es una base de $W$.

Veamos primero que $T$ es inyectiva. Dado un vector $v \in N(T)$, escribámoslo con respecto a la base $\mathcal{B}$, es decir, sean $\alpha_1, \ldots \alpha_n \in \mathbb{K}$ tales que $v = \sum_{k=1}^n \alpha_k b_k$. Luego,

$$\vec{0} = Tv = T\left(\sum_{k=1}^n \alpha_k b_k\right) = \sum_{k=1}^n \alpha_k \; Tb_k,$$

y, como $\mathcal{B}'$ es linealmente independiente en $W$, resulta que $\alpha_1 = \ldots = \alpha_k = 0$. Entonces $v = \vec{0}$ y $N(T) = \{\vec{0}\}$.

Por otra parte, la suryectividad de $T$ está garantizada por el hecho de que $\mathcal{B}'$ es un conjunto de generadores de $W$. De hecho, dado $w \in W$ existen $\beta_1, \ldots, \beta_n \in \mathbb{K}$ tales que

$$w = \sum_{k=1}^n \beta_k \; Tb_k = T\left(\sum_{k=1}^n \beta_k b_k\right).$$

Luego, $v := \sum_{k=1}^n \beta_k b_k \in V$ es tal que $Tv = w$. Como $w$ era arbitrario, resulta que $T$ es suryectiva.

En conclusión, $T$ es una transformación lineal biyectiva, es decir, un isomorfismo.

$v) \rightarrow i)$ : También es trivial. □

**Definición 2.23.** Dados dos $\mathbb{K}$-espacios vectoriales $V$ y $W$, diremos que $V$ es **isomorfo** a $W$ si existe un isomorfismo $T : V \rightarrow W$.

Es fácil ver que la relación "$V$ es isomorfo a $W$" es una relación de equivalencia en el conjunto de $\mathbb{K}$-espacios vectoriales de dimensión finita. Mas aún, el siguiente resultado muestra que la dimensión de $V$ caracteriza a la clase de equivalencia del $\mathbb{K}$-espacio vectorial $V$.

**Teorema 2.24.** *Si $V$ es un $\mathbb{K}$-espacio vectorial con $\dim V = n$, entonces $V$ es isomorfo a $\mathbb{K}^n$.*

*Demostración.* Fijemos una base arbitraria $\mathcal{B} = \{b_1, \ldots, b_n\}$ de $V$ y consideremos la base canónica $\mathcal{E} = \{e_1, \ldots, e_n\}$ de $\mathbb{K}^n$. Luego, sea $T : V \rightarrow \mathbb{K}^n$ la (única) transformación lineal que satisface

$$Tb_j = e_j, \quad j = 1, \ldots, n,$$

es decir, si $v = \sum_{j=1}^n \alpha_j b_j \in V$ entonces

$$Tv = (\alpha_1, \ldots, \alpha_n) = [v]_{\mathcal{B}}^{\top}.$$

Como ya mostramos que esta aplicación es biyectiva, tenemos que $T$ es un isomorfismo. □

## 2.3. Representación matricial de transformaciones lineales

Dados $V$ y $W$ dos $\mathbb{K}$-espacios vectoriales con $\dim V = n$ y $\dim W = m$, sean $\mathcal{B} = \{b_1, \ldots, b_n\}$ y $\mathcal{B}' = \{b'_1, \ldots, b'_m\}$ bases de $V$ y $W$, respectivamente.

Vimos que toda transformación lineal $T : V \to W$ está determinada por su efecto sobre una base de $V$. Luego, si para $j = 1, \dots, n$

$$Tb_j = \sum_{i=1}^{m} A_{ij} b'_j,$$

entonces $T$ está unívocamente determinada por los $mn$ escalares de la matriz $A = (A_{ij})$. A esta matriz la llamaremos la **matriz de $T$ con respecto a las bases $\mathcal{B}$ y $\mathcal{B}'$**, y la anotaremos $[T]_{\mathcal{B}',\mathcal{B}}$:

$$[T]_{\mathcal{B}',\mathcal{B}} = \begin{bmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \dots & A_{mn} \end{bmatrix}.$$

Notemos que, para cada $j = 1, \dots, n$, la columna $j$-ésima de $[T]_{\mathcal{B}',\mathcal{B}}$ está compuesta por las coordenadas del vector $Tb_j$ con respecto a la base $\mathcal{B}'$:

$$\mathrm{Col}_j([T]_{\mathcal{B}',\mathcal{B}}) = [Tb_j]_{\mathcal{B}'}.$$

Si tenemos en cuenta además que el vector de coordenadas de $b_j$ con respecto a la base $\mathcal{B}$ coincide con el $j$-ésimo vector de la base canónica de $\mathbb{K}^{n\times 1}$, $\{e_1^\top, \dots, e_n^\top\}$, resulta que

$$[T]_{\mathcal{B}',\mathcal{B}} \cdot [b_j]_{\mathcal{B}} = [T]_{\mathcal{B}',\mathcal{B}} \cdot e_j^\top = \mathrm{Col}_j([T]_{\mathcal{B}',\mathcal{B}}) = [Tb_j]_{\mathcal{B}'}.$$

**Ejercicio 2.25.** Dado un vector arbitario $v \in V$, probar que

$$[T]_{\mathcal{B}',\mathcal{B}} \cdot [v]_{\mathcal{B}} = [Tv]_{\mathcal{B}'}.$$

(Ayuda: Si $[v]_{\mathcal{B}} = (\alpha_1, \dots, \alpha_n)^\top$ son las coordenadas de $v$ con respecto a la base $\mathcal{B}$, i.e. $v = \sum_{j=1}^{n} \alpha_j b_j$, calcular las coordenadas de $Tv$ con respecto a la base $\mathcal{B}'$.)

**Teorema 2.26.** *Sean $V$ y $W$ dos $\mathbb{K}$-espacios vectoriales con $\dim V = n$ y $\dim W = m$. Dadas $\mathcal{B}$ y $\mathcal{B}'$ bases ordenadas de $V$ y $W$, respectivamente, para cada $T \in L(V, W)$ existe una única matriz $[T]_{\mathcal{B}',\mathcal{B}} \in \mathbb{K}^{m\times n}$ tal que*

$$[T]_{\mathcal{B}',\mathcal{B}} \cdot [v]_{\mathcal{B}} = [Tv]_{\mathcal{B}'} \quad \textit{para todo } v \in V. \tag{2.5}$$

*Más aún, la aplicación $T \mapsto [T]_{\mathcal{B}',\mathcal{B}}$ es un isomorfismo de $L(V, W)$ en $\mathbb{K}^{m\times n}$.*

*Demostración.* Los detalles quedan a cargo del lector, pero el argumento es el desarrollado al inicio de esta sección. □

En el caso en que $T$ es un operador lineal sobre un $\mathbb{K}$-espacio vectorial $V$, normalmente consideraremos la misma base $\mathcal{B}$ tanto para el dominio como para el codominio. En tal caso, diremos simplemente que la matriz que lo representa es la **matriz de $T$ con respecto a la base $\mathcal{B}$** y la anotaremos $[T]_{\mathcal{B}}$. En este caso, la condición que la caracteriza es

$$[T]_{\mathcal{B}} \cdot [v]_{\mathcal{B}} = [Tv]_{\mathcal{B}} \quad \text{para todo } v \in V. \tag{2.6}$$

**Ejemplos 2.27.**

1. Sea $T : \mathbb{R}^2 \to \mathbb{R}^2$ dada por $T(x, y) = (x + y, x - y)$. Si $\mathcal{E} = \{e_1 = (1, 0), e_2 = (0, 1)\}$ es la base canónica de $\mathbb{R}^2$, entonces
$$[T]_{\mathcal{E}} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix},$$
ya que $[Te_1]_{\mathcal{E}} = [(1, 1)]_{\mathcal{E}} = (1, 1)^\top$ y $[Te_2]_{\mathcal{E}} = [(1, -1)]_{\mathcal{E}} = (1, -1)^\top$.

   En cambio, si $\mathcal{B} = \{b_1 = (1, 0), b_2 = (1, -1)\}$, resulta que
$$[T]_{\mathcal{B}} = \begin{pmatrix} 2 & 2 \\ -1 & -2 \end{pmatrix},$$
ya que $[Tb_1]_{\mathcal{B}} = [(1, 1)]_{\mathcal{B}} = [2.(1, 0)+(-1).(1, -1)]_{\mathcal{B}} = (2, -1)^\top$, mientras que $[Tb_2]_{\mathcal{B}} = [(0, 2)]_{\mathcal{B}} = [2.(1, 0) + (-2).(1, -1)]_{\mathcal{B}} = (2, -2)^\top$.

2. Sea $V = \mathbb{C}^{(3)}[x]$, el $\mathbb{C}$-espacio vectorial de polinomios con coeficientes complejos de grado menor o igual que 3 (además del polinomio nulo). Si $D : V \to V$ está definido por
$$D(a_0 + a_1x + a_2x^2 + a_3x^3) = a_1 + 2a_2x + 3a_3x^2,$$
y $\mathcal{E} = \{p_0(x) = 1, p_1(x) = x, p_2(x) = x^2, p_3(x) = x^3\}$ entonces
$$[D]_{\mathcal{E}} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

3. Dados dos $\mathbb{K}$-espacios vectoriales $V$ y $W$ con $\dim V = n$ y $\dim W = m$, sean $\mathcal{B} = \{b_1, \ldots, b_n\}$ y $\mathcal{B}' = \{b'_1, \ldots, b'_m\}$ bases de $V$ y $W$, respectivamente.

   Fijados $p$ y $q$, $1 \leq p \leq m$, $1 \leq q \leq n$, consideremos la transformación lineal $T^{p,q} : V \to W$ dada por
$$T^{p,q}b_j = \delta_{q,j} \cdot b'_p,$$
(una de las que utilizamos para mostrar que $L(V, W)$ es un $\mathbb{K}$-espacio vectorial con dimensión $nm$). Luego,
$$[T^{p,q}]_{\mathcal{B}',\mathcal{B}} = E^{p,q}$$
donde $E^{p,q} \in \mathbb{K}^{m\times n}$ es la matriz que tiene un 1 en la entrada $(p, q)$ y ceros en todas las restantes.

**Proposición 2.28.** *Sean $V_1$, $V_2$ y $V_3$ tres $\mathbb{K}$-espacios vectoriales de dimensión finita, dotados con bases $\mathcal{B}_1$, $\mathcal{B}_2$ y $\mathcal{B}_3$, respectivamente. Luego,*

*i) Si $S, T \in L(V_1, V_2)$ y $\alpha, \beta \in \mathbb{K}$, entonces*
$$[\alpha S + \beta T]_{\mathcal{B}_2,\mathcal{B}_1} = \alpha[S]_{\mathcal{B}_2,\mathcal{B}_1} + \beta[T]_{\mathcal{B}_2,\mathcal{B}_1}.$$

*ii) Si $T \in L(V_1, V_2)$ y $U \in L(V_2, V_3)$, entonces*
$$[UT]_{\mathcal{B}_3,\mathcal{B}_1} = [U]_{\mathcal{B}_3,\mathcal{B}_2}[T]_{\mathcal{B}_2,\mathcal{B}_1}.$$

*Demostración.* La prueba de $i$) queda como ejercicio para el lector. Para probar $ii$), recordemos que $[UT]_{\mathcal{B}_3,\mathcal{B}_1}$ está caracterizada por ser la única matriz tal que

$$[UT]_{\mathcal{B}_3,\mathcal{B}_1}[v]_{\mathcal{B}_1} = [(UT)v]_{\mathcal{B}_3} \quad \text{para todo } v \in V_1.$$

Fijado $v \in V_1$, notemos que $(UT)v = U(Tv)$ y las coordenadas de este último también las podemos calcular usando las representaciones matriciales de $U$ y $T$:

$$[(UT)v]_{\mathcal{B}_3} = [U(Tv)]_{\mathcal{B}_3} = [U]_{\mathcal{B}_3,\mathcal{B}_2}[Tv]_{\mathcal{B}_2} = [U]_{\mathcal{B}_3,\mathcal{B}_2}([T]_{\mathcal{B}_2,\mathcal{B}_1}[v]_{\mathcal{B}_1}) = ([U]_{\mathcal{B}_3,\mathcal{B}_2}[T]_{\mathcal{B}_2,\mathcal{B}_1})[v]_{\mathcal{B}_1}\ .$$

Por lo tanto, $[UT]_{\mathcal{B}_3,\mathcal{B}_1} = [U]_{\mathcal{B}_3,\mathcal{B}_2}[T]_{\mathcal{B}_2,\mathcal{B}_1}$. □

**Corolario 2.29.** *Dados $V$ y $W$ dos $\mathbb{K}$-espacios vectoriales con* $\dim V = \dim W < \infty$*, sean $\mathcal{B}$ y $\mathcal{B}'$ bases de $V$ y $W$, respectivamente. Si $T \in L(V,W)$ es un isomorfismo, entonces*

$$[T^{-1}]_{\mathcal{B},\mathcal{B}'} = ([T]_{\mathcal{B}',\mathcal{B}})^{-1}.$$

*Demostración.* Supongamos que $\dim V = \dim W = n$. En primer lugar, si $I_V$ es el operador identidad sobre $V$, notemos que $[I_V]_{\mathcal{B}}$ es la matriz identidad $I \in \mathbb{K}^{n\times n}$ independientemente de cual sea la base $\mathcal{B}$. Luego, como $I_V = T^{-1}\cdot T$, la proposición anterior garantiza que

$$I = [I_V]_{\mathcal{B}} = [T^{-1}\cdot T]_{\mathcal{B}} = [T^{-1}]_{\mathcal{B},\mathcal{B}'}[T]_{\mathcal{B}',\mathcal{B}}.$$

De la misma manera, si consideramos al operador identidad $I_W$ sobre $W$, tenemos que $[I_W]_{\mathcal{B}'}$ es la matriz identidad $I \in \mathbb{K}^{n\times n}$ independientemente de cual sea la base $\mathcal{B}'$. Entonces, de la identidad $I_W = T\cdot T^{-1}$ obtenemos que

$$I = [I_W]_{\mathcal{B}'} = [T\cdot T^{-1}]_{\mathcal{B}'} = [T]_{\mathcal{B}',\mathcal{B}}[T^{-1}]_{\mathcal{B},\mathcal{B}'},$$

como queríamos probar. □

La Proposición 2.28 y el Corolario 2.29 aplicados en el caso particular en que $W = V$ es un $\mathbb{K}$-espacio vectorial con $\dim V = n$, muestran que la aplicación $T \mapsto [T]_{\mathcal{B}}$ no sólo es un isomorfismo de $\mathbb{K}$-espacios vectoriales entre $L(V)$ y $\mathbb{K}^{n\times n}$, sino que también es un isomorfismo entre $\mathbb{K}$-álgebras lineales.

Dado un $\mathbb{K}$-espacio vectorial $V$ con $\dim V = n$, sea $I_V$ el operador identidad sobre $V$. En la prueba del corolario anterior mencionamos que $[I_V]_{\mathcal{B}}$ es la matriz identidad $I$ independientemente de cual sea la base $\mathcal{B}$. Veamos ahora cual es la matriz asociada a $I_V$ cuando consideramos dos bases distintas de $V$, una en el dominio del operador y la otra en el codominio.

Supongamos que $\mathcal{B}$ y $\mathcal{B}'$ son dos bases de $V$, y calculemos $[I_V]_{\mathcal{B}',\mathcal{B}}$. Si $\mathcal{B} = \{b_1,\ldots,b_n\}$, sabemos que esta matriz está caracterizada por

$$\mathrm{Col}([I_V]_{\mathcal{B}',\mathcal{B}}) = [I_V b_j]_{\mathcal{B}'}, \quad j = 1,\ldots,n.$$

Pero $[I_V b_j]_{\mathcal{B}'} = [b_j]_{\mathcal{B}'} = \mathrm{Col}(P_{\mathcal{B}',\mathcal{B}})$. Por lo tanto, cada columna de $[I_V]_{\mathcal{B}',\mathcal{B}}$ coincide con la correspondiente columna de la matriz cambio de base $P_{\mathcal{B}',\mathcal{B}}$, es decir,

$$[I_V]_{\mathcal{B}',\mathcal{B}} = P_{\mathcal{B}',\mathcal{B}}\ .$$

Fijada una transformación lineal entre $\mathbb{K}$-espacios vectoriales de dimensión finita, este "descubrimiento" nos permitirá pasar fácilmente de una representación matricial a otra. Dados dos $\mathbb{K}$-espacios vectoriales $V$ y $W$ de dimensión finita, supongamos que $T : V \to W$ es una transformación lineal de la cual conocemos $[T]_{\mathcal{B}_1',\mathcal{B}_1}$ para ciertas bases $\mathcal{B}_1$ y $\mathcal{B}_1'$ de $V$ y $W$, respectivamente.

Consideremos ahora otro par de bases $\mathcal{B}_2$ y $\mathcal{B}_2'$ (de $V$ y $W$, respectivamente) y veamos como calcular $[T]_{\mathcal{B}_2',\mathcal{B}_2}$ a partir de $[T]_{\mathcal{B}_1',\mathcal{B}_1}$. Dado un vector arbitrario $v \in V$,

$$\begin{aligned}
[Tv]_{\mathcal{B}_2'} &= P_{\mathcal{B}_2',\mathcal{B}_1'}[Tv]_{\mathcal{B}_1'} \\
&= P_{\mathcal{B}_2',\mathcal{B}_1'}\left([T]_{\mathcal{B}_1',\mathcal{B}_1}[v]_{\mathcal{B}_1}\right) \\
&= \left(P_{\mathcal{B}_2',\mathcal{B}_1'}[T]_{\mathcal{B}_1',\mathcal{B}_1}\right) P_{\mathcal{B}_1,\mathcal{B}_2}[v]_{\mathcal{B}_2} \\
&= \left(P_{\mathcal{B}_2',\mathcal{B}_1'}[T]_{\mathcal{B}_1',\mathcal{B}_1}P_{\mathcal{B}_1,\mathcal{B}_2}\right) [v]_{\mathcal{B}_2}.
\end{aligned}$$

Por lo tanto,

$$[T]_{\mathcal{B}_2',\mathcal{B}_2} = P_{\mathcal{B}_2',\mathcal{B}_1'}\,[T]_{\mathcal{B}_1',\mathcal{B}_1}\,P_{\mathcal{B}_1,\mathcal{B}_2} \tag{2.7}$$

Comprobemos esta igualdad retomando un ejemplo anterior.

**Ejemplo 2.30.** Sea $T:\mathbb{R}^2 \to \mathbb{R}^2$ tal que $[T]_{\mathcal{E}} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$, si $\mathcal{E}$ es la base canónica de $\mathbb{R}^2$. Dada la base $\mathcal{B} = \{b_1 = (1,0), b_2 = (1,-1)\}$ de $\mathbb{R}^2$, calculemos $[T]_{\mathcal{B}}$ utilizando la fórmula (2.7):

$$\begin{aligned}
[T]_{\mathcal{B}} = P_{\mathcal{B},\mathcal{E}}\,[T]_{\mathcal{E}}\,P_{\mathcal{E},\mathcal{B}} = P_{\mathcal{E},\mathcal{B}}^{-1}\,[T]_{\mathcal{E}}\,P_{\mathcal{E},\mathcal{B}} &= \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix} \\
&= \begin{pmatrix} 2 & 2 \\ -1 & -2 \end{pmatrix},
\end{aligned}$$

como ya habíamos visto en el item 1. de los Ejemplos 2.27.

**Ejercicio 2.31.** Al hacer las cuentas en el ejemplo anterior usamos que

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

¿Qué curioso, no? ¿Se animan a describir qué condiciones deben cumplir $a, b, c \in \mathbb{R}$ para que

$$\begin{pmatrix} a & b \\ 0 & c \end{pmatrix}^{-1} = \begin{pmatrix} a & b \\ 0 & c \end{pmatrix}?$$

Noten que este problema podemos plantearlo para un cuerpo arbitrario $\mathbb{K}$, otra curiosidad.

Dado un $\mathbb{K}$-espacio vectorial $V$ de dimensión finita, sea $T$ un operador lineal sobre $V$. Si $\mathcal{B}_1$ y $\mathcal{B}_2$ son dos bases arbitrarias de $V$, entonces

$$[T]_{\mathcal{B}_2} = P_{\mathcal{B}_2,\mathcal{B}_1}\,[T]_{\mathcal{B}_1}\,P_{\mathcal{B}_1,\mathcal{B}_2} = (P_{\mathcal{B}_1,\mathcal{B}_2})^{-1}\,[T]_{\mathcal{B}_1}\,P_{\mathcal{B}_1,\mathcal{B}_2},$$

es decir, $[T]_{\mathcal{B}_2}$ es semejante a $[T]_{\mathcal{B}_1}$ en el siguiente sentido:

**Definición 2.32.** Dadas $A, B \in \mathbb{K}^{n\times n}$, diremos que $B$ **es semejante a** $A$ si existe una matriz inversible $P \in \mathbb{K}^{n\times n}$ tal que

$$B = P^{-1}\,A\,P.$$

Recíprocamente, supongamos que $B$ es una matriz semejante a $[T]_{\mathcal{B}_1}$, es decir, supongamos que existe una $P$ inversible tal que

$$B = P^{-1}\,[T]_{\mathcal{B}_1}\,P.$$

Entonces, podemos interpretar a $P$ como una matriz cambio de base. Si $\mathcal{B}_1 = \{b_1, \ldots, b_n\}$, consideremos la base $\mathcal{B}_2 = \{b'_1, \ldots, b'_n\}$ cuyos elementos están definidos de la siguiente manera: dado $j = 1, \ldots, n$,

$$b'_j = \sum_{i=1}^{n} P_{ij} b_i.$$

Luego, es fácil ver que $[b'_j]_{\mathcal{B}_1} = \mathrm{Col}_j(P)$, y en consecuencia,

$$P = P_{\mathcal{B}_1,\mathcal{B}_2}.$$

**Ejercicio 2.33.** Mostrar que la semejanza es una relación de equivalencia en el álgebra de matrices cuadradas $\mathbb{K}^{n\times n}$.

Hay algunas cantidades asociadas a una matriz cuadrada que se preservan por la semejanza, por ejemplo, el determinante. Otra de estas cantidades invariantes es la traza de la matriz: si $A \in \mathbb{K}^{n\times n}$, la **traza de** $A$ se define como $\mathrm{tr}(A) = \sum_{i=1}^{n} A_{ii}$. La propiedad que caracteriza a la traza es la siguiente: si $A \in \mathbb{K}^{n\times m}$ y $B \in \mathbb{K}^{m\times n}$ entonces

$$\mathrm{tr}(AB) = \mathrm{tr}(BA), \tag{2.8}$$

la prueba de esta propiedad queda como ejercicio.

**Proposición 2.34.** *Si $A, B \in \mathbb{K}^{n\times n}$ son semejantes, entonces*

$$\mathrm{tr}(B) = \mathrm{tr}(A) \quad y \quad \det(B) = \det(A).$$

*Demostración.* Sea $P \in \mathbb{K}^{n\times n}$ tal que $B = P^{-1}AP$. Luego, $\det(P) \neq 0$ y

$$\det(B) = \det(P^{-1}AP) = \det(P^{-1})\det(A)\det(P) = \det(P)^{-1}\det(A)\det(P) = \det(A).$$

Por otra parte,

$$\mathrm{tr}(B) = \mathrm{tr}(P^{-1}AP) = \mathrm{tr}((P^{-1}A)P) = \mathrm{tr}(P(P^{-1}A)) = \mathrm{tr}((PP^{-1})A) = \mathrm{tr}(A),$$

donde fue necesario utilizar la propiedad de la traza (2.8). □

La invarianza del determinante y la traza bajo semejanzas nos permite extender estos conceptos a operadores lineales actuando en un $\mathbb{K}$-espacio vectorial de dimensión finita arbitrario.

**Definición 2.35.** Dado un $\mathbb{K}$-espacio vectorial $V$ de dimensión finita, sea $\mathcal{B}$ una base arbitraria de $V$. Si $T \in L(V)$, definimos la **traza** y el **determinante** de $T$ como:

$$\mathrm{tr}(T) = \mathrm{tr}([T]_{\mathcal{B}}) \quad \text{y} \quad \det(T) = \det([T]_{\mathcal{B}}),$$

respectivamente.

## 2.4. Ejercicios

1. Determinar si las siguientes aplicaciones son transformaciones lineales.

   *a*) $T : \mathbb{R}^3 \to \mathbb{R}^3$ dada por $T(x, y, z) = (x - y, x^2, 2z)$.

   *b*) $T : \mathbb{R}^{3\times1} \to \mathbb{R}^{3\times1}$ dada por $T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x - 3y \\ 3y - 2z \\ 2z \end{pmatrix}$.

   *c*) $T : \mathbb{R}^4 \to \mathbb{R}^2$ dada por $T(x_1, x_2, x_3, x_4) = (0, 0)$.

   *d*) $T : \mathbb{C} \to \mathbb{C}$ dada por $T(z) = \bar{z}$.

   *e*) La función traza $\mathrm{tr} : \mathbb{C}^{n\times n} \to \mathbb{C}$ dada por $\mathrm{tr}(A) = \sum_{i=1}^{n} a_{ii}$.

   *f*) La función determinante $\det : \mathbb{C}^{n\times n} \to \mathbb{C}$.

   *g*) Considerando al $\mathbb{Z}_2$-espacio vectorial $\mathbb{Z}_2 \times \mathbb{Z}_2$, $T : \mathbb{Z}_2 \times \mathbb{Z}_2 \to \mathbb{Z}_2 \times \mathbb{Z}_2$ dada por $T(a, b) = (b, a + b)$.

2. *a*) Dar un ejemplo de una función $f : \mathbb{R}^2 \to \mathbb{R}$ tal que

   $$f(\alpha v) = \alpha f(v) \quad \text{para todo } v \in \mathbb{R}^2 \text{ y todo } \alpha \in \mathbb{R}\,,$$

   pero que $f$ no sea una transformación lineal.

   *b*) Dar un ejemplo de una función $g : \mathbb{C} \to \mathbb{C}$ tal que

   $$g(u + v) = g(u) + g(v) \quad \text{para todo } u, v \in \mathbb{C}\,,$$

   pero que $g$ no sea una transformación lineal.

3. Dado $\alpha \in [0, 2\pi)$ fijo, sea $R_\alpha : \mathbb{R}^2 \to \mathbb{R}^2$ la función que define la rotación en un ángulo $\alpha$ en sentido antihorario, es decir,

   $$R_\alpha(x, y) = (x \cos(\alpha) - y \operatorname{sen}(\alpha), x \operatorname{sen}(\alpha) + y \cos(\alpha)), \quad (x, y) \in \mathbb{R}^2\,.$$

   *a*) Probar que $R_\alpha$ es una transformación lineal para cualquier valor del ángulo $\alpha$ fijo.

   *b*) Determinar la imagen por $R_{\frac{\pi}{2}}$ de los puntos $P = (0, 3)$, $Q = (3, 1)$ y $S = (1, -1)$. Graficar el triángulo $PQS$ y su transformado en el mismo sistema de coordenadas.

4. Sea $S_Y : \mathbb{R}^2 \to \mathbb{R}^2$ la función simetría con respecto del eje $y$, es decir,

   $$S_Y(x, y) = (-x, y), \quad (x, y) \in \mathbb{R}^2\,.$$

   *a*) Probar que $S_Y$ es una transformación lineal.

   *b*) Determinar la imagen por $S_Y$ de $A = (0, 1)$, $B = (2, 4)$, $C = (4, 3)$ y $D = (2, 0)$. Graficar el paralelogramo $ABCD$ y su transformado en el mismo sistema de coordenadas.

5. Dado $r \in \mathbb{R}$ fijo, sea $H_r : \mathbb{R}^2 \to \mathbb{R}^2$ la función homotecia de razón $r$, es decir,

   $$H_r(x, y) = (rx, ry), \quad (x, y) \in \mathbb{R}^2\,.$$

   *a*) Probar que $H_r$ es una transformación lineal.

   b) Fijando $r = 2$, determinar la imagen por $H_2$ de $A = (0, -1)$, $B = (1, 2)$ y $C = (3, 1)$. Graficar el triángulo $ABC$ y su transformado en el mismo sistema de coordenadas.

6. Sea $P_X : \mathbb{R}^2 \to \mathbb{R}^2$ la función proyección sobre el eje $x$.

   a) Hallar una expresión analítica para $P_X$ y probar que es una transformación lineal.
   b) Determinar la imagen por $P_X$ de $A = (0, 1)$, $B = (2, 4)$, $C = (4, 3)$ y $D = (2, 0)$. Graficar el rectángulo $ABCD$ y su transformado, en el mismo sistema de coordenadas.

7. Se dice que una transformación $T : \mathbb{R}^n \to \mathbb{R}^n$ es una isometría si preserva distancias, es decir, si
$$\operatorname{dist}(P, Q) = \operatorname{dist}(T(P), T(Q)) \quad \text{para todo } P, Q \in \mathbb{R}^n .$$
   Determinar cuáles de las transformaciones consideradas en los ejercicio 3, 4, 5 y 6 son isometrías.

8. a) Probar que existe una única transformación lineal $T : \mathbb{R}^2 \to \mathbb{R}^2$ tal que $T(1, 1) = (-5, 3)$ y $T(-1, 1) = (5, 2)$. Determinar $T(5, 3)$ y $T(-1, 2)$.
   b) Determinar si existe una transformación lineal $T : \mathbb{R}^2 \to \mathbb{R}^2$ tal que $T(1, 1) = (2, 6)$, $T(-1, 1) = (2, 5)$ y $T(2, 7) = (5, 3)$.

9. Sea $T : V \to W$ una transformación lineal.

   i) Probar que $N(T)$ es un subespacio de $V$ e $\operatorname{Im}(T)$ es un subespacio de $W$.
   ii) Además, si $\dim V < \infty$, mostrar que $\dim(N(T)) \leq \dim V$ y $\dim(\operatorname{Im}(T)) \leq \dim V$.

10. En cada caso, determinar una transformación lineal $T : \mathbb{R}^3 \to \mathbb{R}^3$ que verifique lo pedido.

    a) $(1, 1, 0) \in N(T)$ y $\dim(\operatorname{Im}(T)) = 1$.
    b) $N(T) \cap \operatorname{Im}(T) = \overline{\{(1, 1, 2)\}}$.
    c) $N(T) \neq \{(0, 0, 0)\}$, $\operatorname{Im}(T) \neq \{(0, 0, 0)\}$ y $N(T) \cap \operatorname{Im}(T) = \{(0, 0, 0)\}$.

11. *Extensión de transformaciones lineales.* Dado un $\mathbb{K}$-espacio vectorial $V$ de dimensión finita, sea $U$ un subespacio propio de $V$. Probar que si $S : U \to W$ es una transformación lineal, entonces existe $T : V \to W$ tal que
$$Tu = Su \quad \text{para todo } u \in U .$$

12. Probar que si $V$ es un $\mathbb{K}$-espacio vectorial con $\dim(V) = 1$ y $T \in L(V)$, entonces $T$ es un múltiplo (escalar) de la identidad.

13. Dado un $\mathbb{K}$-espacio vectorial $V$ de dimensión finita, sea $T : V \to V$ un operador lineal tal que $\operatorname{Im}(T^2) = \operatorname{Im}(T)$. Probar que $\operatorname{Im}(T) \cap N(T) = \{0\}$.

14. Dado un $\mathbb{K}$-espacio vectorial $V$, sea $T : V \to V$ un operador lineal. Probar que son equivalentes:

    a) $N(T) \cap \operatorname{Im}(T) = \{0\}$.
    b) Si $T(T(v)) = 0$ entonces $T(v) = 0$ para $v \in V$.

15. Sean $V$ y $W$ dos $\mathbb{K}$-espacios vectoriales. Supongamos que $V$ es de dimensión finita y sea $T : V \to W$ una transformación lineal. Probar que:

    *a*) $T$ es inyectiva si y sólo si $\dim(\mathrm{Im}(T)) = \dim(V)$.

    *b*) $T$ es suryectiva si y sólo si $\dim(\mathrm{Im}(T)) = \dim(W)$.

16. Dado un $\mathbb{K}$-espacio vectorial $V$ de dimensión finita, sea $S = \{v_1, \cdots, v_m\}$ un conjunto finito de vectores de $V$. Definimos $E : \mathbb{K}^m \to V$ como la función dada por

$$E(k_1, \cdots, k_m) = k_1\, v_1 + \cdots + k_m\, v_m\,.$$

    *a*) Probar que $E$ es una transformación lineal.

    *b*) Probar que $E$ es inyectiva si y sólo si $S$ es un conjunto de vectores linealmente independientes.

    *c*) Probar que $E$ es suryectiva si y sólo si $S$ es un conjunto de vectores que generan a $V$.

17. Dada la matriz

$$A = \begin{pmatrix} -1 & 0 & 2 \\ 5 & 4 & 2 \\ -4 & -3 & 2 \end{pmatrix} \in \mathbb{R}^{3\times 3},$$

consideremos las funciones $L_A : \mathbb{R}^{3\times 1} \to \mathbb{R}^{3\times 1}$ y $R_A : \mathbb{R}^3 \to \mathbb{R}^3$ dadas por

$$L_A \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = A \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \qquad \text{y} \qquad R_A(x_1, x_2, x_3) = (x_1, x_2, x_3)A,$$

respectivamente.

    *a*) Probar que $L_A \in L(\mathbb{R}^{3\times 1})$ y $R_A \in L(\mathbb{R}^3)$.

    *b*) Hallar el núcleo y la imagen de ambas transformaciones. Analizar su inyectividad, suryectividad y biyectividad.

18. Para cada una de las transformaciones lineales del ejercicio 1, determinar su núcleo e imagen. Además analizar inyectividad, suryectividad y biyectividad.

19. Dados $V$ y $W$ dos $\mathbb{K}$-espacios vectoriales, sea $T : V \to W$ una transformación lineal. Supongamos que $S_V$ y $S_W$ son subespacios de $V$ y $W$, respectivamente. Probar que

$$\begin{aligned} T(S_V) &:= \{w \in W :\ w = T(v) \text{ para algún } v \in S_V\}\,, \quad \text{y} \\ T^{-1}(S_W) &:= \{v \in V :\ T(v) \in S_W\}\,, \end{aligned}$$

son subespacios vectoriales de $W$ y $V$, respectivamente. Al primero lo llamamos la **imagen de $S_V$ por $T$** y al segundo la **preimagen de $S_W$ por $T$**.

20. Sea $T : \mathbb{R}^3 \to \mathbb{R}^2$ dada por $T(x, y, z) = (x + y, x + z)$.

    *a*) Probar que $T$ es una transformación lineal.

    *b*) Hallar el núcleo de $T$ ¿Qué dimensión tiene?

    *c*) Hallar la imagen de $T$ ¿Qué dimensión tiene?

    *d*) Sea $C = \{(x, y) :\ x = 1\}$. Hallar $T^{-1}(C)$ ¿Es este un subespacio de $\mathbb{R}^3$? ¿Contradice esto el ejercicio anterior? ¿Por qué?

21. *a*) Dado un $\mathbb{K}$-espacio vectorial $V$ de dimensión finita, sea $T : V \to V$ un operador lineal. Probar que $T$ es inyectivo si y sólo si $T^{-1}(0) = \{0\}$.

    *b*) Dar ejemplos de funciones $f : \mathbb{R} \to \mathbb{R}$ que no sean inyectivas pero que verifican $f^{-1}(0) = \{0\}$.

22. Sean $U$ y $V$ dos $\mathbb{K}$-espacios vectoriales, y sea $L(U, V)$ el conjunto de todas las transformaciones lineales de $U$ en $V$. Probar que $L(U, V)$ es un $\mathbb{K}$-espacio vectorial con las operaciones dadas por:

    - $(S + T)(u) = S(u) + T(u)$ para $S\,, T \in L(U, V)$ y $u \in U$.
    - $(\alpha\, T)(u) = \alpha \cdot T(u)$ para $T \in L(U, V)$, $\alpha \in \mathbb{K}$ y $u \in U$.

23. Sean $U$ y $V$ dos $\mathbb{K}$-espacios vectoriales de dimensión finita tales que $2 \leq \dim(U) \leq \dim(V)$. Probar que

$$S = \{T \in L(U, V) : T \text{ no es inyectiva}\}\ ,$$

    no es un subespacio de $L(U, V)$.

24. Dado un $\mathbb{K}$-espacio vectorial $V$, sean $S, T, U \in L(V)$ y sea $\alpha \in \mathbb{K}$. Probar que:

    *a*) $S \cdot I = I \cdot S = S$;

    *b*) $S(T + U) = ST + SU$ y también $(T + U)S = TS + US$;

    *c*) $\alpha(ST) = (\alpha S)T = S(\alpha T)$.

25. Sean $V_1$, $V_2$ y $V_3$ tres $\mathbb{K}$-espacios vectoriales. Probar que, si $T : V_1 \to V_2$ y $S : V_2 \to V_3$ son transformaciones lineales inversibles, entonces $ST$ también es inversible. Además

$$(ST)^{-1} = T^{-1}S^{-1}.$$

26. Considerar las transformaciones lineales $R_\alpha$, $S_Y$, $H_r$ y $P_X$, de los ejercicio 3, 4, 5 y 6, respectivamente. En cada caso:

    *a*) Hallar su representación matricial con respecto a la base canónica de $\mathbb{R}^2$.

    *b*) Determinar su núcleo e imagen, y calcular la dimensión de ambos subespacios.

    *c*) Analizar inyectividad, suryectividad y biyectividad.

    *d*) Hallar su representación matricial con respecto a la base $\mathcal{B} = \{(1, 1), (1, -1)\}$.

27. Sea $T : \mathbb{R}^3 \to \mathbb{R}^3$ la transformación lineal dada por $T(x, y, z) = (z, y + z, x + y + z)$.

    *a*) Hallar la representación matricial de $T$ con respecto a la base canónica de $\mathbb{R}^3$.

    *b*) Sea $\mathcal{B} = \{(1,\ 1,\ 0), (1,\ 2,\ 3), (1,\ 3,\ 5)\}$. Hallar las matrices de cambio de base $P_{\mathcal{E},\mathcal{B}}$ y $P_{\mathcal{B},\mathcal{E}}$.

    *c*) Hallar la representación matricial de $T$ con respecto a la base $\mathcal{B}$.

28. Dado un vector arbitario $v \in V$, probar que

$$[T]_{\mathcal{B}',\mathcal{B}} \cdot [v]_{\mathcal{B}} = [Tv]_{\mathcal{B}'}.$$

    (Ayuda: Si $[v]_{\mathcal{B}} = (\alpha_1, \ldots, \alpha_n)^\top$ son las coordenadas de $v$ con respecto a la base $\mathcal{B}$, i.e. $v = \sum_{j=1}^{n} \alpha_j b_j$, calcular las coordenadas de $Tv$ con respecto a la base $\mathcal{B}'$.)

29. Sea $T : \mathbb{R}^2 \to \mathbb{R}^3$ la transformación lineal dada por $T(x, y) = (3x + y, -y, x + y)$. Consideremos las bases de $\mathbb{R}^2$ y $\mathbb{R}^3$ dadas por $\mathcal{B} = \{(1, 1), (-1, 0)\}$ y $\mathcal{B}' = \{(1, 0, 1), (1, -1, 0), (0, 2, 1)\}$. Hallar la matriz de la transformación $T$ con respecto a las bases $\mathcal{B}$ y $\mathcal{B}'$, es decir, la matriz $[T]_{\mathcal{B}',\mathcal{B}}$.

30. Sea $T : \mathbb{R}^3 \to \mathbb{R}^3$ la transformación lineal dada por $T(x, y, z) = (2y + z, x - 4y, 3x)$.

    *a*) Hallar $[T]_{\mathcal{B}}$ para la base $\mathcal{B} = \{(1, 1, 1), (1, 1, 0), (1, 0, 0)\}$ de $\mathbb{R}^3$.

    *b*) Probar que $[T(v)]_{\mathcal{B}} = [T]_{\mathcal{B}}\, [v]_{\mathcal{B}}$ para cada $v \in \mathbb{R}^3$.

31. Sea $T : \mathbb{C}^2 \to \mathbb{C}^2$ la transformación lineal dada por $T(x, y) = (y, 0)$. Consideremos la base $\mathcal{B} = \{(1, i),\ (-i, 2)\}$ de $\mathbb{C}^2$ como $\mathbb{C}$-espacio vectorial.

    *a*) Hallar la representación matricial de $T$ con respecto a las bases $\mathcal{E}$ y $\mathcal{B}$.

    *b*) Hallar la representación matricial de $T$ con respecto a las bases $\mathcal{B}$ y $\mathcal{E}$.

    *c*) Hallar la representación matricial de $T$ con respecto a la base $\mathcal{B}$. Calcular el determinante y la traza.

    *d*) Elegir otra base de $\mathbb{C}^2$ y hallar la representación matricial de $T$ con respecto a esa base. ¿Puede decir cuánto valen el determinante y la traza sin hacer las cuentas? Justificar.

32. Consideremos a $\mathbb{C}$ como un $\mathbb{R}$-espacio vectorial y sea $T : \mathbb{C} \to \mathbb{C}$ dada por $T(z) = \overline{z}$.

    *a*) Probar que $T$ es una transformación lineal.

    *b*) Hallar la representación matricial $[T]_{\mathcal{B}}$ para la base $\mathcal{B} = \{1, i\}$ y $[T]_{\mathcal{B}'}$ para la base $\mathcal{B}' = \{1 + i, 1 - i\}$.

    *c*) Hallar $P_{\mathcal{B},\mathcal{B}'}$ y $P_{\mathcal{B}',\mathcal{B}}$.

    *d*) Verificar que $[T]_{\tilde{B}} = P_{\mathcal{B}',\mathcal{B}}\, [T]_{\mathcal{B}}\, P_{\mathcal{B},\mathcal{B}'}$.

33. Sea $T : \mathbb{R}^{(2)}[x] \to \mathbb{R}^3$ la transformación lineal dada por

$$T(ax^2 + bx + c) = (a + 2b + c, a + b, 2a + b - c)\,.$$

    *a*) Hallar una representación matricial de $T$.

    *b*) Hallar la dimensión del núcleo y de la imagen de $T$.

    *c*) Dado un subespacio $W$ de $\mathbb{R}^3$ tal que $W \dot{+} \mathrm{Im}(T) = \mathbb{R}^3$, hallar $T^{-1}(W)$.

34. Sean $\mathcal{B} = \{b_1, b_2, b_3\}$ y $\mathcal{B}' = \{c_1, c_2, c_3, c_4\}$ bases de $\mathbb{R}^3$ y de $\mathbb{R}^4$, respectivamente. Sea $T \in L(\mathbb{R}^3, \mathbb{R}^4)$ tal que

$$[T]_{\mathcal{B}',\mathcal{B}} = \begin{pmatrix} 1 & -2 & 1 \\ -1 & 1 & 1 \\ 2 & 1 & 4 \\ 3 & -2 & 5 \end{pmatrix}.$$

    *a*) Hallar $T(3b_1 + 2b_2 - b_3)$ y decir cuáles son sus coordenadas en la base $\mathcal{B}'$.

    *b*) Hallar una base para el núcleo y otra para la imagen de $T$.

    *c*) Hallar $T^{-1}(c_1 - 3c_3 - c_4)$.

35. *a*) Sean $V$ un $\mathbb{K}$-espacio vectorial con $\dim(V) = 5$ y $\mathcal{B} = \{b_1, \ldots, b_5\}$ una base (ordenada) de $V$. Supongamos también que $T \in L(V)$ está dada por

$$T(b_j) = \begin{cases} b_{j+1} & \text{si } j = 1, 2, 3, 4, \\ 0 & \text{si } j = 5. \end{cases}$$

   1) Hallar la representación matricial de $T$ con respecto a la base $\mathcal{B}$.
   2) Probar que $T^5 = 0$ pero $T^4 \neq 0$.

   *b*) Supongamos ahora que $\dim(V) = n$ y que $\mathcal{B} = \{b_1, \ldots, b_n\}$ es una base (ordenada) de $V$. Si $T \in L(V)$ está dada por

$$T(b_j) = \begin{cases} b_{j+1} & \text{si } j = 1, \ldots, n-1, \\ 0 & \text{si } j = n. \end{cases}$$

   hallar la representación matricial de $T$ en la base $\mathcal{B}$. Además, probar que ocurre lo mismo que en el inciso anterior, es decir $T^n = 0$ pero $T^{n-1} \neq 0$.

   *Sugerencia:* Notar que $b_j = T^{j-1}(b_1)$ y probar que $T^n(b_j) = 0$ para $j = 1, \ldots, n$.

   *c*) Supongamos que $S \in L(V)$ es cualquier operador lineal tal que $S^n = 0$ y $S^{n-1} \neq 0$. Probar que existe una base (ordenada) de $V$ tal que la representación matricial de $S$ con respecto a esa base, coincide con la matriz hallada en el ítem anterior.

36. Mostrar que la semejanza es una relación de equivalencia en el álgebra de matrices cuadradas $\mathbb{K}^{n \times n}$.

37. Probar que si $M, N \in \mathbb{K}^{n \times n}$ son tales que $M^n = N^n = 0$, $M^{n-1} \neq 0$ y $N^{n-1} \neq 0$, entonces $M$ y $N$ son semejantes.

# Capítulo 3

# El álgebra lineal de polinomios

Dado un cuerpo $\mathbb{K}$ comencemos recordando que, si $p(x) = a_0 + a_1x + a_2x^2 + \ldots + a_nx^n$ y $q(x) = b_0 + b_1x + b_2x^2 + \ldots + b_mx^m$ son dos polinomios en $\mathbb{K}[x]$, el producto de $p$ y $q$ se define como

$$(p \cdot q)(x) = c_0 + c_1x + c_2x^2 + \ldots + c_{n+m}x^{n+m},$$

donde los coeficientes $c_k$ para $k = 1, \ldots, n+m$ están dados por:

$$c_k = \sum_{j=0}^{k} a_j b_{k-j},$$

con la condición adicional de que $a_{n+1} = \ldots = a_{n+m} = 0$ y también $b_{m+1} = \ldots = b_{m+n} = 0$.

Luego, el espacio $\mathbb{K}[x]$ de polinomios con coeficientes en el cuerpo $\mathbb{K}$ es un álgebra lineal sobre $\mathbb{K}$, ya que es un $\mathbb{K}$-espacio vectorial con una operación binaria adicional (en este caso, el producto de polinomios) $\cdot : \mathbb{K}[x] \times \mathbb{K}[x] \to \mathbb{K}[x]$ tal que, para cualesquiera $p, q, r \in \mathbb{K}[x]$:

- es asociativa:
$$(p \cdot q) \cdot r = p \cdot (q \cdot r);$$
- es distributiva con respecto a la adición:
$$p \cdot (q + r) = p \cdot q + p \cdot r \quad , \quad (p + q) \cdot r = p \cdot r + q \cdot r;$$
- para cada $\alpha \in \mathbb{K}$, vale que
$$\alpha \cdot (p \cdot q) = (\alpha \cdot p) \cdot q = p \cdot (\alpha \cdot q).$$

Además, es un álgebra lineal conmutativa con unidad, porque el producto de polinomios es conmutativo y el polinomio $p_0(x) = 1$ es el neutro de esta operación.

Recordemos que el conjunto $\{1, x, x^2, x^3, \ldots\}$ es una "base ordenada de $\mathbb{K}[x]$" en el siguiente sentido: para cada polinomio no nulo $p \in \mathbb{K}[x]$, existe un único $n \in \mathbb{Z}$, $n \geq 0$, tal que la $n$-ésima coordenada de $p$ con respecto a esta base es **no nula** y las siguientes son todas cero

$$p(x) = a_0 + a_1x + a_2x^2 + \ldots + a_nx^n, \quad a_n \neq 0.$$

En otras palabras, $n$ es el menor entero tal que $p$ puede escribirse como combinación lineal de los polinomios en el subconjunto $\{1, x, x^2, \ldots, x^n\}$. Dicho entero es el **grado de** $p$ y lo anotaremos $\mathrm{gr}(p)$.

**Observación 3.1.** Al polinomio nulo no le asignaremos grado alguno.

**Teorema 3.2.** *Si $p, q \in \mathbb{K}[x]$ son polinomios no nulos, entonces:*

*i)* *$p \cdot q$ también es no nulo;*

*ii)* $\operatorname{gr}(p \cdot q) = \operatorname{gr}(p) + \operatorname{gr}(q)$*;*

*iii)* *si $p$ y $q$ son mónicos, $p \cdot q$ también lo es;*

*iv)* *si $p + q$ es no nulo, entonces* $\operatorname{gr}(p+q) \leq \text{máx}\{\operatorname{gr}(p), \operatorname{gr}(q)\}$.

*Demostración.* Supongamos que $\operatorname{gr}(p) = n$ y que $\operatorname{gr}(q) = m$. Entonces, si $k \geq 0$ tenemos que

$$(p \cdot q)_{n+m+k} = \sum_{i=0}^{m+n+k} p_i q_{m+n+k-i},$$

donde $p(x) = p_0 + p_1 x + \ldots + p_n x^n$ y $q(x) = q_0 + q_1 x + \ldots q_m x^m$.

Para que $p_i q_{m+n+k-i}$ sea distinto de cero es necesario que $i \leq n$ y que $m+n+k-i \leq m$, o equivalentemente,

$$n + k \leq i \leq n.$$

Por lo tanto, $(p \cdot q)_{n+m} = p_n q_m \neq 0$ y $(p \cdot q)_{n+m+k} = 0$ para todo $k \geq 1$.

De lo anterior se desprenden *i*), *ii*) y *iii*). El item *iv*) queda como ejercicio para el lector. □

Como consecuencia de lo anterior, tenemos que la propiedad de cancelación también vale en ecuaciones polinomiales.

**Corolario 3.3.** *Si $p, q, r \in \mathbb{K}[x]$ son tales que $p \neq 0$ y $p \cdot q = p \cdot r$, entonces $q = r$.*

*Demostración.* Si $p \cdot q = p \cdot r$ entonces $p \cdot (q - r) = 0$. Si además $p \neq 0$, el item *i*) del teorema anterior implica que $q - r = 0$, es decir, $q = r$. □

## 3.1. Interpolación de Lagrange

Dado un cuerpo $\mathbb{K}$, consideremos $n + 1$ pares ordenados en $\mathbb{K} \times \mathbb{K}$

$$(t_0, \alpha_0), (t_1, \alpha_1), \ldots (t_n, \alpha_n),$$

con la condición de que los escalares $t_0, t_1, \ldots, t_n$ sean distintos entre sí. La interpolación polinómica es una técnica en la cual, a partir de un conjunto de puntos como los anteriores, se pretende encontrar una función polinomial $p$ cuyo gráfico pase por todos los puntos en cuestión, es decir, que satisfaga las condiciones

$$p(t_j) = \alpha_j, \quad j = 0, 1, \ldots, n. \tag{3.1}$$

Existen infinitas soluciones al problema de interpolación polinómica, con lo cual es necesario imponer condiciones adicionales para establecer un problema con una única solución.

En nuestro caso estamos interesados en encontrar el polinomio interpolante cuyo grado sea el menor posible. La solución a este problema fue publicada por primera vez por Joseph-Louis Lagrange (1736–1813) en 1795, y habitualmente se la conoce como **interpolación de Lagrange**.

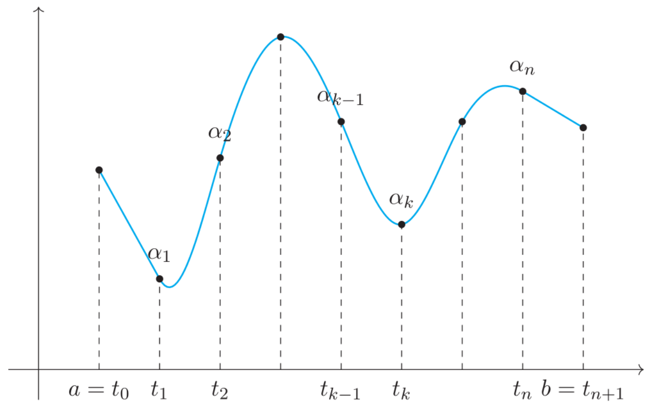



Figura 3.1: Un ejemplo de polinomio interpolante (curva celeste).

Dados $n+1$ escalares $t_0, t_1, \ldots, t_n$ distintos entre sí, se definen los siguientes polinomios $p_0, p_1, \ldots, p_n \in \mathbb{K}[x]$:

$$p_i(x) = \tfrac{(x-t_0)(x-t_1)\cdots(x-t_{i-1})(x-t_{i+1})\cdots(x-t_n)}{(t_i-t_0)(t_i-t_1)\cdots(t_i-t_{i-1})(t_i-t_{i+1})\cdots(t_i-t_n)} = \prod_{j\neq i} \tfrac{x-t_j}{t_i-t_j}, \qquad j = 0, \ldots, n.$$

Notemos que éstos son $n+1$ polinomios de grado $n$. Además,

$$p_i(t_j) = \begin{cases} 1, & \text{si } j = i, \\ 0, & \text{si } j \neq i, \end{cases}$$

es decir, $p_i(t_j) = \delta_{ij}$.

Consideremos el subespacio $\mathbb{K}^{(n)}[x]$ de $\mathbb{K}[x]$ formado por los polinomios de grado menor o igual que $n$, además del polinomio nulo. Sabemos que éste es un $\mathbb{K}$-espacio vectorial con $\dim \mathbb{K}^{(n)}[x] = n+1$, y una de sus bases ordenadas es $\mathcal{E} = \{1, x, x^2, \ldots, x^n\}$. Queremos mostrar que $\mathcal{L} = \{p_0, p_1, \ldots, p_n\}$ también es una base de $\mathbb{K}^{(n)}$, para lo cual alcanza con probar que $\mathcal{L}$ es linealmente independiente.

Supongamos que existen $\beta_0, \beta_1, \ldots, \beta_n \in \mathbb{K}$ tales que

$$\beta_0 p_0 + \beta_1 p_1 + \ldots + \beta_n p_n = \vec{0}_{\mathbb{K}^{(n)}[x]}.$$

Luego, evaluando la identidad anterior en $t_i$, obtenemos

$$\begin{aligned} 0 &= (\beta_0 p_0 + \beta_1 p_1 + \ldots + \beta_n p_n)(t_i) \\ &= \beta_0\, p_0(t_i) + \beta_1\, p_1(t_i) + \ldots + \beta_n\, p_n(t_i) \\ &= \beta_0\, \delta_{0i} + \beta_1\, \delta_{1i} + \ldots + \beta_n\, \delta_{ni} = \beta_i, \qquad \text{para } i = 0, \ldots, n. \end{aligned}$$

Por lo tanto, $\mathcal{L}$ es linealmente independiente y, en consecuencia, una base de $\mathbb{K}^{(n)}[x]$.

Si $p$ es un polinomio cualquiera en $\mathbb{K}^{(n)}[x]$, entonces existen $c_0, c_1, \ldots, c_n \in \mathbb{K}$ tales que $p = c_0 p_0 + c_1 p_1 + \ldots + c_n p_n$, y estos coeficientes podemos despejarlos evaluando a $p$ en los puntos de interpolación $t_j$:

$$\begin{aligned} p(t_j) &= (c_0 p_0 + c_1 p_1 + \ldots + c_n p_n)(t_j) \\ &= c_0\, p_0(t_j) + c_1\, p_1(t_j) + \ldots + c_n\, p_n(t_j) = \\ &= c_0\, \delta_{0j} + c_1\, \delta_{1j} + \ldots + c_n\, \delta_{nj} \\ &= c_j. \end{aligned}$$

Por lo tanto,

$$p = p(t_0)p_0 + p(t_1)p_1 + \ldots + p(t_n)p_n. \tag{3.2}$$

Esta fórmula se conoce como la **fórmula de interpolación de Lagrange** del polinomio $p$.

Volviendo al problema con el que iniciamos la sección, dados $n+1$ pares ordenados

$$(t_0, \alpha_0), (t_1, \alpha_1), \ldots, (t_n, \alpha_n),$$

tales que $t_i \neq t_j$ si $i \neq j$, la fórmula de interpolación de Lagrange nos permite obtener al **único** polinomio en $p \in \mathbb{K}^{(n)}[x]$ (y en consecuencia, el de menor grado) que satisface las $n+1$ condiciones de interpolación:

$$p(t_j) = \alpha_j \qquad j = 0, 1, \ldots, n.$$

**Ejemplo 3.4.** Supongamos que queremos hallar el polinomio en $\mathbb{R}[x]$ de menor grado que satisface las condiciones de interpolación:

$$p(-1) = -6, \quad p(0) = 2, \quad p(1) = -2, \quad y \quad p(2) = 6. \tag{3.3}$$

En este caso, $t_0 = -1$, $t_1 = 0$, $t_2 = 1$ y $t_3 = 2$. Luego,

$$\begin{aligned}
p_0(x) &= \tfrac{(x-t_1)(x-t_2)(x-t_3)}{(t_0-t_1)(t_0-t_2)(t_0-t_3)} = \tfrac{x(x-1)(x-2)}{(-1)(-2)(-3)} = -\tfrac{1}{6}(x^3 - 3x^2 + 2x),\\
p_1(x) &= \tfrac{(x-t_0)(x-t_2)(x-t_3)}{(t_1-t_0)(t_1-t_2)(t_1-t_3)} = \tfrac{(x+1)(x-1)(x-2)}{1.\ (-1)(-2)} = \tfrac{1}{2}(x^3 - 2x^2 - x + 2),\\
p_2(x) &= \tfrac{(x-t_0)(x-t_1)(x-t_3)}{(t_2-t_0)(t_2-t_1)(t_2-t_3)} = \tfrac{(x+1)x(x-2)}{2.\ 1.\ (-1)} = -\tfrac{1}{2}(x^3 - x^2 - 2x),\\
p_3(x) &= \tfrac{(x-t_0)(x-t_1)(x-t_2)}{(t_3-t_0)(t_3-t_1)(t_3-t_2)} = \tfrac{(x+1)x(x-1)}{3.\ 2.\ 1} = \tfrac{1}{6}(x^3 - x).
\end{aligned}$$

Aplicando la fórmula de interpolación de Lagrange con las condiciones (3.3), obtenemos el polinomio:

$$\begin{aligned}
p(x) &= \sum_{i=0}^{3} p(t_i)p_i(x) = p(t_0)p_0(x) + p(t_1)p_1(x) + p(t_2)p_2(x) + p(t_3)p_3(x)\\
&= (x^3 - 3x^2 + 2x) + (x^3 - 2x^2 - x + 2) + (x^3 - x^2 - 2x) + (x^3 - x)\\
&= 4x^3 - 6x^2 - 2x + 2,
\end{aligned}$$

el cual verifica las condiciones requeridas. □

Como ya mencionamos antes, dados $n+1$ escalares $t_0, t_1, \ldots, t_n \in \mathbb{K}$ tales que $t_i \neq t_j$ si $i \neq j$, los conjuntos $\mathcal{L} = \{p_0, p_1, \ldots, p_n\}$ y $\mathcal{E} = \{1, x, x^2, \ldots, x^n\}$ son dos bases de $\mathbb{K}^{(n)}[x]$. En particular, para cada $k = 0, 1, \ldots, n$ tenemos que

$$x^k = \sum_{i=0}^{n} t_i^k . p_i(x).$$

Por lo tanto,

$$[x^k]_{\mathcal{L}} = \begin{pmatrix} t_0^k \\ t_1^k \\ \vdots \\ t_n^k \end{pmatrix}, \qquad k = 0, 1, \ldots, n.$$

Luego, la matriz cambio de base $P_{\mathcal{L},\mathcal{E}}$ tiene el siguiente aspecto:

$$P_{\mathcal{L},\mathcal{E}} = \begin{pmatrix} 1 & t_0 & t_0^2 & \dots & t_0^n \\ 1 & t_1 & t_1^2 & \dots & t_1^n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & t_n & t_n^2 & \dots & t_n^n \end{pmatrix}.$$

Esta matriz es naturalmente inversible, y recibe el nombre de **matriz de Vandermonde** de orden $n+1$. Probar que ésta es inversible por otros medios no resulta para nada sencillo.

## 3.2. El algoritmo de la división - Raíces de polinomios

Comencemos recordando el algoritmo de la división para polinomios.

**Teorema 3.5.** *Si $p, d \in \mathbb{K}[x]$ y $d \neq 0$, entonces existen* **únicos** *$q, r \in \mathbb{K}[x]$ tales que*

*i)* $p = d \cdot q + r$;

*ii)* $r = 0$ *ó* $\mathrm{gr}(r) < \mathrm{gr}(d)$.

Sea $d \in \mathbb{K}[x]$, $d \neq 0$. Dado $p \in \mathbb{K}[x]$, el algoritmo de la división dice que existe, a lo sumo, un $q \in \mathbb{K}[x]$ tal que $p = d \cdot q$.

**Definición 3.6.** Dado $p \in \mathbb{K}[x]$, si existe un $q \in \mathbb{K}[x]$ tal que $p = d \cdot q$, se dice que $d$ **divide a** $p$ y se anota $d \,|\, p$.

**Corolario 3.7.** *Sean $p \in \mathbb{K}[x]$ y $\alpha \in \mathbb{K}$, entonces*

$$x - \alpha \,|\, p(x) \quad \text{si y sólo si} \quad p(\alpha) = 0.$$

*Demostración.* Aplicando el algoritmo de la división, existen $q, r \in \mathbb{K}[x]$ tales que $p(x) = (x-\alpha)q(x) + r(x)$ con $r = 0$ ó $\mathrm{gr}(r) < \mathrm{gr}(x-\alpha) = 1$. Esto asegura que el resto es un escalar. Pero como

$$p(\alpha) = (\alpha - \alpha)q(\alpha) + r(\alpha) = r(\alpha),$$

resulta que $x - \alpha \,|\, p(x)$ si y sólo si $p(\alpha) = 0$. □

**Definición 3.8.** $\alpha \in \mathbb{K}$ es una **raíz (o cero)** de un polinomio $p \in \mathbb{K}[x]$ si $p(\alpha) = 0$.

**Corolario 3.9.** *Sea $p \in \mathbb{K}[x]$. Si $\mathrm{gr}(p) = n$ entonces $p$ tiene a lo sumo $n$ raíces en $\mathbb{K}$.*

*Demostración.* Utilicemos el principio de inducción completa sobre el grado del polinomio. El caso $n = 1$ es trivial. Supongamos, como hipótesis inductiva, que el resultado vale para polinomios de grado $n - 1$.

Luego, dado $p \in \mathbb{K}[x]$ con $\mathrm{gr}(p) = n$, supongamos que $\alpha \in \mathbb{K}$ es una raíz de $p$. Entonces, $p(x) = (x - \alpha)q(x)$ para cierto polinomio $q \in \mathbb{K}[x]$ con $\mathrm{gr}(q) = n - 1$.

Entonces, dado $\beta \in \mathbb{K}$, $p(\beta) = 0$ si y sólo si $\beta = \alpha$ ó $q(\beta) = 0$. Por la hipótesis inductiva, $q$ tiene a lo sumo $n - 1$ raíces. Por lo tanto, $p$ tiene a lo sumo $n$ raíces en $\mathbb{K}$. □

**Definición 3.10.** Dado $p \in \mathbb{K}[x]$, supongamos que $\alpha \in \mathbb{K}$ es una raíz de $p$. La **multiplicidad de la raíz** $\alpha$ de $p$ es el mayor $r \in \mathbb{N}$ tal que $(x - \alpha)^r \,|\, p$.

El operador derivada $D : \mathbb{K}[x] \to \mathbb{K}[x]$ definido por

$$(Dp)(x) = a_1 + 2a_2x + \ldots + na_nx^{n-1}, \qquad \text{si } p(x) = a_0 + a_1x + a_2x^2 + \ldots + a_nx^n,$$

es de suma utilidad para calcular la multiplicidad de las raíces de un polinomio dado, si el cuerpo $\mathbb{K}$ tiene **característica cero**. Recordemos que la **característica** de un cuerpo $\mathbb{K}$ se define como el menor entero positivo $k$ tal que

$$\overbrace{1 + 1 + \ldots + 1}^{k \text{ sumandos}} = 0.$$

Si no existe tal $k \in \mathbb{N}$, se dice que la característica de $\mathbb{K}$ es cero.

**Teorema 3.11.** *Sea $\mathbb{K}$ un cuerpo con característica cero. Dado $p \in \mathbb{K}[x]$ con $\mathrm{gr}(p) = n$, sea $\alpha \in \mathbb{K}$ una raíz de $p$. Entonces, $\alpha$ es una raíz de $p$ con multiplicidad $r$ si y sólo si*

$$\begin{aligned} &(D^kp)(\alpha) = 0 \qquad \textit{para } k = 0, 1, \ldots, r-1, \\ &(D^rp)(\alpha) \neq 0. \end{aligned}$$

*Demostración.* Este resultado es una consecuencia de la fórmula de Taylor, la cual probaremos a continuación. Comencemos notando que, si $\mathbb{K}$ es un cuerpo con característica cero entonces

$$(\alpha + \beta)^m = \sum_{k=0}^{m} \binom{m}{k} \alpha^{m-k}\beta^k, \qquad \alpha, \beta \in \mathbb{K}. \tag{3.4}$$

Esta fórmula para una potencia natural del binomio se verifica fácilmente usando el principio de inducción completa. Ahora, dado $p \in \mathbb{K}[x]$ con $\mathrm{gr}(p) = n$, veamos que si $\alpha \in \mathbb{K}$

$$p(x) = \sum_{k=0}^{n} \frac{(D^kp)(\alpha)}{k!}(x - \alpha)^k. \tag{3.5}$$

Si $p(x) = x^m$ con $1 \leq m \leq n$, aplicando la fórmula del binomio obtenemos que

$$x^m = \big(\alpha + (x - \alpha)\big)^m = \sum_{k=0}^{m} \binom{m}{k} \alpha^{m-k}(x - \alpha)^k = \sum_{k=0}^{m} \frac{1}{k!}\frac{m!}{(m-k)!}\alpha^{m-k}(x - \alpha)^k,$$

mientras que $(D^kp)(\alpha) = \frac{m!}{(m-k)!}\alpha^{m-k}$ para cada $k = 1, \ldots, n$, lo que prueba la fórmula de Taylor en este caso particular. La prueba para el caso general es consecuencia de la linealidad de los operadores $D, D^2, \ldots, D^n$, ya que si $p(x) = \sum_{m=0}^{n} a_mx^m$ entonces

$$(D^kp)(\alpha) = \sum_{m=0}^{n} a_m(D^kx^m)(\alpha) = \sum_{m=0}^{n} a_m \frac{m!}{(m-k)!}\alpha^{m-k},$$

y en consecuencia

$$\begin{aligned} \sum_{k=0}^{n} \frac{(D^kp)(\alpha)}{k!}(x - \alpha)^k &= \sum_{k=0}^{n} \frac{1}{k!} \sum_{m=0}^{n} a_m \frac{m!}{(m-k)!}\alpha^{m-k}(x - \alpha)^k \\ &= \sum_{m=0}^{n} a_m \left( \sum_{k=0}^{n} \frac{m!}{k!(m-k)!}\alpha^{m-k}(x - \alpha)^k \right) = \sum_{m=0}^{n} a_mx^m = p(x), \end{aligned}$$

lo que completa la demostración de (3.5).

Ahora, supongamos que $\alpha$ es una raíz de $p$ con multiplicidad $r$. Entonces, existe $q \in \mathbb{K}[x]$ tal que

$$p(x) = (x-\alpha)^r q(x),$$

y además $q(\alpha) \neq 0$, porque en caso contrario $\alpha$ sería una raíz de $q$ (y, por lo tanto, una raíz de $p$ de orden mayor que $r$). Aplicando la fórmula de Taylor al polinomio $q$, tenemos que

$$p(x) = (x-\alpha)^r q(x) = (x-\alpha)^r \left[\sum_{m=0}^{n-r} \frac{(D^k q)(\alpha)}{m!}(x-\alpha)^m\right] = \sum_{m=0}^{n-r} \frac{(D^k q)(\alpha)}{m!}(x-\alpha)^{m+r}.$$

Como el conjunto $\{1, x-\alpha, (x-\alpha)^2, \ldots, (x-\alpha)^n\}$ es linealmente independiente en $\mathbb{K}[x]$, hay una única forma de escribir a $p$ como combinación lineal de estos polinomios, y entonces

$$\frac{(D^k p)(\alpha)}{k!} = \begin{cases} 0 & \text{si } 0 \leq k \leq r-1 \\ \frac{(D^{k-r} q)(\alpha)}{(k-r)!} & \text{si } r \leq k \leq n. \end{cases}$$

Por lo tanto, $(D^k p)(\alpha) = 0$ para $k = 0, \ldots, r-1$ y $(D^r p)(\alpha) = q(\alpha) \neq 0$.

Recíprocamente, supongamos que $(D^k p)(\alpha) = 0$ para $k = 0, \ldots, r-1$ y $(D^r p)(\alpha) \neq 0$. Luego, por la fórmula de Taylor, vemos que

$$\begin{aligned} p(x) &= \sum_{k=0}^{n} \frac{(D^k p)(\alpha)}{k!}(x-\alpha)^k = \sum_{k=r}^{n} \frac{(D^k p)(\alpha)}{k!}(x-\alpha)^k = \sum_{j=0}^{n-r} \frac{(D^{r+j} p)(\alpha)}{(r+j)!}(x-\alpha)^{r+j} \\ &= (x-\alpha)^r \left[\sum_{j=0}^{n-r} \frac{(D^{r+j} p)(\alpha)}{(r+j)!}(x-\alpha)^j\right], \end{aligned}$$

y además el polinomio que multiplica a $(x-\alpha)^r$ no se anula al evaluarlo en $\alpha$, ya que $(D^r p)(\alpha) \neq 0$. Pero esto implica que $\alpha$ es una raíz de $p$ con multiplicidad igual a $r$. □

## 3.3. Ideales de polinomios

**Definición 3.12.** Sea $\mathbb{K}$ un cuerpo. Un **ideal** del álgebra de polinomios $\mathbb{K}[x]$ es un subespacio $M$ de $\mathbb{K}[x]$ tal que

$$\text{si } p \in \mathbb{K}[x] \text{ y } q \in M \text{ entonces } p \cdot q \in M.$$

A esta última propiedad de los ideales la llamaremos informalmente **la propiedad de absorción**, ya que el producto de un polinomio arbitrario por uno que pertenezca al ideal termina siendo también un polinomio del ideal.

**Ejemplo 3.13.** Dado un cuerpo $\mathbb{K}$, sea $d \in \mathbb{K}[x]$. El conjunto de todos los polinomios divisibles por $d$,

$$d \cdot \mathbb{K}[x] := \{d \cdot p : \ p \in \mathbb{K}[x]\},$$

es un ideal de $\mathbb{K}[x]$.

En primer lugar, notemos que $d \cdot \mathbb{K}[x] \neq \varnothing$ porque $d \in d \cdot \mathbb{K}[x]$. Además es fácil ver que $d \cdot \mathbb{K}[x]$ es un subespacio de $\mathbb{K}[x]$: si $p, q \in d \cdot \mathbb{K}[x]$ y $\alpha \in \mathbb{K}$, entonces

$$\alpha(d \cdot p) + (d \cdot q) = d \cdot (\alpha p + q) \in d \cdot \mathbb{K}[x].$$

Por último verifiquemos la propiedad de absorción. Dados $p, q \in \mathbb{K}[x]$,

$$\overbrace{p}^{\in \mathbb{K}[x]} \cdot \overbrace{d \cdot q}^{\in d \cdot \mathbb{K}[x]} = d \cdot (p \cdot q) \in d \cdot \mathbb{K}[x].$$

Por lo tanto, $d \cdot \mathbb{K}[x]$ es un ideal de $\mathbb{K}[x]$. $\square$

El ideal del ejemplo anterior se denomina **ideal principal generado por** $d$.

**Ejemplo 3.14.** Dado un cuerpo $\mathbb{K}$, sea $\alpha \in \mathbb{K}$. El conjunto de polinomios que tienen a $\alpha$ como raíz

$$M = \{p \in \mathbb{K}[x] : \ p(\alpha) = 0\},$$

también es un ideal de $\mathbb{K}[x]$. Más aún, $M = (x - \alpha) \cdot \mathbb{K}[x]$. (Verificarlo!)

En general, dados $d_1, d_2, \ldots, d_n \in \mathbb{K}[x]$, la suma de los ideales principales $d_i \cdot \mathbb{K}[x]$,

$$M = d_1 \cdot \mathbb{K}[x] + d_2 \cdot \mathbb{K}[x] + \ldots + d_n \cdot \mathbb{K}[x],$$

también es un ideal de $\mathbb{K}[x]$ al que llamaremos el **ideal generado por** $d_1, d_2, \ldots, d_n$.

**Ejemplo 3.15.** En $\mathbb{K}[x]$ consideremos los polinomios $x - 1$ y $x^2 - 1$, y el conjunto

$$M = \{(x-1)p(x) + (x^2-1)q(x) : \ p, q \in \mathbb{K}[x]\}.$$

Éste también es un ideal de $\mathbb{K}[x]$.

Una vez más, es claro que $0 \in M$. Además, si $m_1, m_2 \in M$ se descomponen como $m_i(x) = (x-1)p_i(x) + (x^2-1)q_i(x)$ para $i = 1, 2$, y $\alpha \in \mathbb{K}$, entonces

$$\begin{aligned}(\alpha m_1 + m_2)(x) &= \alpha m_1(x) + m_2(x) \\ &= \alpha\big((x-1)p_1(x) + (x^2-1)q_1(x)\big) + (x-1)p_2(x) + (x^2-1)q_2(x) \\ &= (x-1)\big(\alpha p_1(x) + p_2(x)\big) + (x^2-1)\big(\alpha q_1(x) + q_2(x)\big) \in M.\end{aligned}$$

Por último, si $m(x) = (x-1)p(x) + (x^2-1)q(x) \in M$ y $r \in \mathbb{K}[x]$ entonces

$$(m \cdot r)(x) = m(x)r(x) = (x-1)p(x)r(x) + (x^2-1)q(x)r(x) \in M.$$

$\square$

**Ejemplo 3.16.** El ideal de $\mathbb{C}[x]$ generado por los polinomios $x + 2$ y $x^2 + 8x + 16$ coincide con $\mathbb{C}[x]$, es decir,

$$(x+2) \cdot \mathbb{C}[x] + (x^2 + 8x + 16) \cdot \mathbb{C}[x] = \mathbb{C}[x].$$

Para verificarlo, veamos que los escalares pertenecen a $M$: dividiendo a $x^2 + 8x + 16$ por $x + 2$ tenemos que

$$x^2 + 8x + 16 = (x+2)(x+6) + 4.$$

Luego, $4 = (x^2 + 8x + 16) - (x+2)(x+6) \in M$. En consecuencia, $1 \in M$ y la propiedad de absorción implica que $\mathbb{C}[x] \subseteq M$. $\square$

A pesar de que nos tomamos el trabajo de desarrollar varios ejemplos, el siguiente teorema muestra que todo ideal no trivial de $\mathbb{K}[x]$ es un ideal principal.

**Teorema 3.17.** *Si $\mathbb{K}$ es un cuerpo y $M$ es un ideal no trivial de $\mathbb{K}[x]$, existe un único polinomio mónico $d \in \mathbb{K}[x]$ tal que $M = d \cdot \mathbb{K}[x]$.*

*Demostración.* Supongamos que $M$ es un ideal no trivial de $\mathbb{K}[x]$. Sea $d \in M$ un polinomio de grado mínimo entre los polinomios de $M$. Además, podemos suponer que $d$ es mónico (si no lo fuera podemos hacerlo mónico multiplicando por una constante de $\mathbb{K}$, y este producto permanece en $M$).

Dado $p \in M$ cualquiera, por el algoritmo de la división sabemos que existen $q, r \in \mathbb{K}[x]$ tales que $p = d \cdot q + r$ con $r = 0$ ó $\mathrm{gr}(r) < \mathrm{gr}(d)$. Como $r = p - d \cdot q \in M$, el caso en que $\mathrm{gr}(r) < \mathrm{gr}(d)$ deriva en un absurdo porque estamos suponiendo que $d$ es de grado mínimo entre los polinomios de $M$. Por lo tanto, $r = 0$ y $p = d \cdot q \in d \cdot \mathbb{K}[x]$.

Lo anterior muestra que $M \subseteq d \cdot \mathbb{K}[x]$, mientras que la otra inclusión es trivial porque $d \in M$. En consecuencia, $M = d \cdot \mathbb{K}[x]$.

Supongamos ahora que $d' \in M$ es otro polinomio mónico tal que $M = d' \cdot \mathbb{K}[x]$. Entonces, existen polinomios no nulos $p, q \in \mathbb{K}[x]$ tales que $d = d' \cdot p$ y $d' = d \cdot q$. Luego, $d = d' \cdot p = (d \cdot q) \cdot p = d \cdot (q \cdot p)$ y además

$$\mathrm{gr}(d) = \mathrm{gr}(d) + \mathrm{gr}(q) + \mathrm{gr}(p).$$

Esto implica que $\mathrm{gr}(p) = \mathrm{gr}(q) = 0$, y como tanto $d$ como $d'$ son mónicos resulta que $p = q = 1$. Por lo tanto, $d' = d$. □

**Ejercicio 3.18.** En el álgebra de polinomios con coeficientes reales $\mathbb{R}[x]$, consideremos el ideal $M = (x-1)\mathbb{R}[x] + (x^2-1)\mathbb{R}[x]$. Mostrar que $M = (x-1)\mathbb{R}[x]$.

¿Cambia algo si reemplazamos al cuerpo $\mathbb{R}$ por otro cuerpo arbitrario $\mathbb{K}$?

**Corolario 3.19.** *Sean $p_1, \ldots, p_n \in \mathbb{K}[x]$ polinomios no todos nulos. Entonces, existe un único polinomio mónico $d \in \mathbb{K}[x]$ tal que*

*i) $d$ pertenece al ideal generado por $p_1, \ldots, p_n$;*

*ii) $d$ divide a cada $p_i$, para $i = 1, \ldots, n$.*

*Además, todo polinomio que satisface i) y ii) necesariamente satisface también*

*iii) $d$ es divisible por todos los polinomios que dividen a cada uno de los $p_i$.*

*Demostración.* Llamemos $d \in \mathbb{K}[x]$ al generador mónico del ideal $M = p_1 \cdot \mathbb{K}[x] + \ldots + p_n \cdot \mathbb{K}[x]$, y veamos primero que $d$ cumple con las condiciones *i*), *ii*) y *iii*).

Las dos primeras las cumple por definición de generador del ideal. Naturalmente $d \in M$ y, si $q \in M$ entonces $d \,|\, q$. En particular, $d \,|\, p_i$ para $i = 1, \ldots, n$.

Queremos ver ahora que $p \,|\, d$ para todo $p \in \mathbb{K}[x]$ tal que $p \,|\, p_i$ para $i = 1, \ldots, n$. Supongamos que $p \in \mathbb{K}[x]$ divide a todos los $p_i$. Luego, existen $r_1, \ldots, r_n \in \mathbb{K}[x]$ tales que

$$p_i = p \cdot r_i, \qquad i = 1, \ldots, n.$$

Por otro lado, como $d \in M$, existen $q_1, \ldots, q_n \in \mathbb{K}[x]$ tales que $d = p_1 \cdot q_1 + \ldots + p_n \cdot q_n$. Luego,

$$d = p_1 \cdot q_1 + \ldots + p_n \cdot q_n = p \cdot (r_1 \cdot q_1 + \ldots + r_n \cdot q_n),$$

es decir, $p \,|\, d$. Hemos probado entonces que $d$ verifica las condiciones *i*), *ii*) y *iii*).

Supongamos ahora que existe otro $d' \in \mathbb{K}[x]$ que verifica *i*) y *ii*). Repitiendo el mismo argumento anterior, vemos que $d'$ también satisface *iii*).

Como $d' \in M = d \cdot \mathbb{K}[x]$, tenemos que $d \,|\, d'$. Por otro lado, como $d' \,|\, p_i$ para todo $i = 1, \ldots, n$, y $d$ satisface *iii*), vale también que $d' \,|\, d$. Finalmente, si $d'$ es mónico, resulta que $d' = d$. □

El polinomio $d$ del resultado anterior no es ni más ni menos que el máximo común divisor entre los polinomios $p_1, \dots, p_n$.

**Definición 3.20.** Si $p_1, \dots, p_n \in \mathbb{K}[x]$ no son todos nulos, el generador mónico $d \in \mathbb{K}[x]$ del ideal $p_1 \cdot \mathbb{K}[x] + \dots + p_n \cdot \mathbb{K}[x]$ se llama el **maximo común divisor** (m. c. d.) de $p_1, \dots, p_n$. Lo anotaremos $d = (p_1, \dots, p_n)$.

Además, diremos que $p_1, \dots, p_n \in \mathbb{K}[x]$ son **primos entre sí** si su m. c. d. es 1, es decir, si el ideal generado por ellos es todo $\mathbb{K}[x]$.

**Ejemplo 3.21.** En $\mathbb{C}[x]$, el m. c. d. entre $(x-2)^2(x-i)$ y $(x-2)(x^2+1)$ es $(x-2)(x-i)$.

En efecto, si $M = (x-2)^2(x-i) \cdot K[x] + (x-2)(x^2+1) \cdot \mathbb{K}[x]$, eligiendo los polinomios $p(x) = 1$ y $q(x) = -1$ resulta que

$$\begin{aligned}(x-2)^2(x-i) \cdot p(x) + (x-2)(x^2+1) \cdot q(x) &= (x-2)^2(x-i) - (x-2)(x^2+1) \\ &= (i-2)(x-2)(x-i) \in M.\end{aligned}$$

Luego, $d := (x-2)(x-i) \in M$ es mónico y divide a los dos polinomios iniciales. Por lo tanto, es el m. c. d. entre ellos.

**Ejemplo 3.22.** En $\mathbb{Q}[x]$, los polinomios $(x-1)(x+2)^2$, $(x+2)^2(x-3)$ y $x-3$ son primos entre sí.

De hecho,

$$(x+2)^2 = \tfrac{1}{2}(x-1)(x+2)^2 + \big(-\tfrac{1}{2}\big)(x+2)^2(x-3) + 0.(x-3) \in M,$$

pero, dividiendo a $(x+2)^2$ por $x-3$, obtenemos $(x+2)^2 = (x-3)(x+7) - 17$. Esto implica entonces que

$$1 = \big(\tfrac{1}{17}(x+7)\big)(x-3) + \big(-\tfrac{1}{17}\big)(x+2)^2 \in M.$$

Por lo tanto, $M = \mathbb{Q}[x]$.

## 3.4. Factorización prima de un polinomio

Para finalizar este capítulo presentaremos al teorema de la factorización prima, el cual garantiza que todo polinomio puede factorizarse de manera única en términos de polinomios más sencillos.

**Definición 3.23.** Dado un cuerpo $\mathbb{K}$, decimos que $p \in \mathbb{K}[x]$ es **reducible sobre** $\mathbb{K}$ si existen $q_1, q_2 \in \mathbb{K}[x]$ con $\mathrm{gr}(q_1) \geq 1$, $\mathrm{gr}(q_2) \geq 1$, tales que $p = q_1 \cdot q_2$. En caso contrario, diremos que $p$ es **irreducible sobre** $\mathbb{K}$.

**Ejemplo 3.24.** El polinomio $p(x) = x^2 + 1$ es irreducible sobre $\mathbb{R}$, pero es reducible sobre $\mathbb{C}$ ya que $x^2 + 1 = (x-i)(x+i)$.

Naturalmente los polinomios escalares son irreducibles, pero no tienen gracia. Los polinomios que nos interesarán en esta parte son los polinomios primos.

**Definición 3.25.** Si $p \in \mathbb{K}[x]$ tiene $\mathrm{gr}(p) \geq 1$ y es irreducible sobre $\mathbb{K}$, diremos que $p$ es **primo en** $\mathbb{K}[x]$.

A continuación enunciaremos dos resultados importantes sobre factorización de polinomios. En primer lugar, veamos que si un polinomio primo divide a un producto de polinomios entonces forzosamente divide a alguno de los factores.

**Teorema 3.26.** *Sean $q_1, \ldots, q_n \in \mathbb{K}[x]$. Si $p \in \mathbb{K}[x]$ es primo en $\mathbb{K}[x]$ y además*

$$p \,|\, q_1 \cdot q_2 \cdot \ldots \cdot q_n,$$

*entonces $p \,|\, q_i$ para algún $i = 1, \ldots, n$.*

*Demostración.* La haremos por induccción sobre la cantidad $n$ de polinomios $q_i$. Sin pérdida de generalidad, podemos suponer que $p$ es un polinomio mónico.

En primer lugar, supongamos que $n = 2$, es decir $p \,|\, q_1 \cdot q_2$. Como $p$ es primo, sus únicos divisores son $p$ y 1. Luego, los únicos candidatos posibles para el m.c.d. entre $p$ y $q_1$ son $p$ y 1. Por un lado, si $(p, q_1) = p$, listo. Por el otro, si $p$ y $q_1$ son primos entre sí, existen polinomios $f, g \in \mathbb{K}[x]$ tales que $1 = p \cdot f + q_1 \cdot g$. Además, como $p \,|\, q_1 \cdot q_2$, existe $d \in \mathbb{K}[x]$ tal que $q_1 \cdot q_2 = p \cdot d$. Luego,

$$q_2 = \big(p \cdot f + q_1 \cdot g\big) \cdot q_2 = p \cdot f \cdot q_2 + (q_1 \cdot q_2) \cdot g = p \cdot f \cdot q_2 + (p \cdot d) \cdot g = p \cdot \big(f \cdot q_2 + d \cdot g),$$

lo que muestra que $p \,|\, q_2$.

Como hipótesis inductiva, supongamos que el resultado es cierto para una cantidad $k$ de polinomios, es decir, si $p \,|\, q_1 \cdot q_2 \cdot \ldots \cdot q_k$ entonces $p \,|\, q_i$ para algún $i = 1, \ldots, k$.

Ahora, supongamos que $p \,|\, q_1 \cdot q_2 \cdot \ldots \cdot q_{k+1}$. Asociando a los primeros $k$ polinomios, tenemos que $p \,|\, (q_1 \cdot q_2 \cdot \ldots \cdot q_k) \cdot q_{k+1}$. Entonces, por el caso $n = 2$ que probamos al comienzo, $p \,|\, q_1 \cdot q_2 \cdot \ldots \cdot q_k$ ó $p \,|\, q_{k+1}$. Por último, aplicando la hipótesis inductiva, tenemos que $p \,|\, q_i$ para algún $i = 1, \ldots, k$ ó $p \,|\, q_{k+1}$. Por lo tanto, $p \,|\, q_i$ para algún $i = 1, \ldots, k + 1$. □

**Teorema 3.27 (Teorema de la factorización prima).** *Si $\mathbb{K}$ es un cuerpo, cada polinomio no escalar $q \in \mathbb{K}[x]$ puede factorizarse como producto de un escalar $\alpha \in \mathbb{K}$ y de factores primos mónicos $p_1, \ldots, p_m \in \mathbb{K}[x]$ de manera* **única***, salvo permutación de los factores,*

$$q = \alpha \cdot p_1 \cdot \ldots \cdot p_m = \alpha \prod_{i=1}^{m} p_i.$$

*Además, si definimos $q_j = \prod_{i \neq j} p_i$ para $j = 1, \ldots, m$, entonces $q_1, \ldots, q_m$ son polinomios primos entre sí.*

*Demostración.* Lo probaremos por inducción en el grado del polinomio $q$. Si $\mathrm{gr}(q) = 1$ no hay nada que probar, ya que si $q(x) = a_0 + a_1 x$ la única factorización posible de $q$ es $q = \alpha \cdot p_1$ con $\alpha = a_1 \in \mathbb{K}$ y $p_1 = \frac{a_0}{a_1} + x \in \mathbb{K}[x]$.

Ahora, sea $q \in \mathbb{K}[x]$ un polinomio con $\mathrm{gr}(q) = k$. Como hipótesis inductiva, supongamos que el resultado es cierto para cualquier polinomio de grado menor estricto que $k$. Si $q$ es irreducible, listo. Si no lo es, supongamos que $q = \alpha \cdot f \cdot g$, donde $\alpha$ es el coeficiente principal de $q$ y $f, g \in \mathbb{K}[x]$ son polinomios mónicos con $\mathrm{gr}(f) > 1$ y $\mathrm{gr}(g) > 1$. Como $k = \mathrm{gr}(q) = \mathrm{gr}(f) + \mathrm{gr}(g)$, tenemos que $1 < \mathrm{gr}(f) < k$ y también $1 < \mathrm{gr}(g) < k$. Aplicando la hipótesis inductiva a estos dos polinomios, tenemos que existen $q_1, \ldots, q_h \in \mathbb{K}[x]$ y $q_{h+1}, \ldots, q_{h+l} \in \mathbb{K}[x]$, todos mónicos, tales que

$$f = \prod_{i=1}^{h} q_i \qquad \text{y} \qquad g = \prod_{j=1}^{l} q_{h+j},$$

con lo cual

$$q = \alpha \cdot f \cdot g = \alpha \cdot \left( \prod_{i=1}^{h} q_i \right) \cdot \left( \prod_{j=1}^{l} q_{h+j} \right) = \alpha \cdot \prod_{i=1}^{h+l} q_i.$$

□

Para finalizar, vamos a diferenciar entre los cuerpos en los cuales todos los polinomios primos son de grado 1 y aquellos que admiten polinomios primos de grado mayor a 1.

**Definición 3.28.** Un cuerpo se dice **algebraicamente cerrado** si todo polinomio primo en $\mathbb{K}[x]$ tiene grado 1.

Si $\mathbb{K}$ es un cuerpo algebraicamente cerrado y $p \in \mathbb{K}[x]$, el grado de $p$ determina unívocamente la cantidad de factores primos en la factorización de $p$.

**Ejemplo 3.29.** El Teorema Fundamental del Álgebra garantiza que $\mathbb{C}$ es un cuerpo algebraicamente cerrado, mientras que $\mathbb{R}$ y $\mathbb{Q}$ no lo son.

Cada polinomio $p \in \mathbb{R}[x]$ es también un polinomio en $\mathbb{C}[x]$. Aplicando el Teorema de la Factorización Prima a un polinomio $p \in \mathbb{R}[x]$ con $\mathrm{gr}(p) = m$ (pensado como un polinomio complejo), sabemos que existen únicos $\alpha \in \mathbb{R}$ y $z_1, \ldots, z_m \in \mathbb{C}$ tales que

$$p(x) = \alpha(x - z_1)(x - z_2) \ldots (x - z_m).$$

Además, sabemos que si $z \in \mathbb{C} \setminus \mathbb{R}$ es una raíz de $p$, entonces el conjugado $\bar{z}$ también es una raíz de $p$. Es decir, las raíces complejas aparecen de a **pares conjugados**.

Luego, si $(z_1, \bar{z_1}), \ldots, (z_r, \bar{z_r})$ son los $r$ pares conjugados de raíces complejas de $p$, y $z_{2r+1}, \ldots, z_m$ son las raíces reales restantes, podemos factorizar al polinomio $p$ como

$$p(x) = \alpha \cdot p_1(x) \ldots p_r(x)(x - z_{2r+1}) \ldots (x - z_m),$$

donde $p_i(x) = (x - z_i)(x - \bar{z_i}) \in \mathbb{R}[x]$ para cada $i = 1, \ldots, r$.

## 3.5. Apéndice: Construcción formal de $\mathbb{K}[x]$

Dado un cuerpo $\mathbb{K}$, aprovechando la idea intuitiva que cada lector tiene sobre el tema, hasta el momento hemos trabajado con los polinomios con coeficientes en $\mathbb{K}$ sin definir formalmente de qué clase de objetos estamos hablando. A continuación presentaremos una construcción formal del álgebra de polinomios $\mathbb{K}[x]$.

Dado un cuerpo $\mathbb{K}$, entre los primeros ejemplos de $\mathbb{K}$-espacios vectoriales presentamos al conjunto $\mathbb{K}^\infty$ formado por las sucesiones con coeficientes en $\mathbb{K}$ indexadas sobre el conjunto de los números naturales $\mathbb{N}$. Ahora haremos un pequeño ajuste y supondremos que $\mathbb{K}^\infty$ es el conjunto de sucesiones con coeficientes en $\mathbb{K}$ indexadas sobre el conjunto $\mathbb{N}_0$ formado por los números enteros no negativos.

Dado un par de sucesiones $a = (a_n)_{n\in\mathbb{N}_0}, b = (b_n)_{n\in\mathbb{N}_0} \in \mathbb{K}^\infty$, definimos el **producto** de $a$ y $b$ como la sucesión $a \cdot b = (c_n)_{n\in\mathbb{N}_0} \in \mathbb{K}^\infty$ cuyos coeficientes están dados por

$$c_n = \sum_{i=0} a_i b_{n-i}, \qquad n \in \mathbb{N}_0,$$

es decir,

$$a \cdot b = \big(a_0b_0, a_0b_1 + a_1b_0, a_0b_2 + a_1b_1 + a_2b_0, \ldots\big).$$

Es medianamente fácil convencerse de que $\mathbb{K}^\infty$ dotado con esta operación binaria adicional es un álgebra lineal (sobre $\mathbb{K}$), la cual es conmutativa y tiene unidad (la sucesión $1 = (1, 0, \ldots, 0, \ldots)$).

Si consideramos el vector

$$x = (0, 1, 0, \ldots, 0, \ldots) \in \mathbb{K}^\infty,$$

observemos que éste tiene un papel destacado, ya que

$$x \cdot x = (0, 0, 1, 0, \ldots, 0, \ldots), \qquad x \cdot x \cdot x = (0, 0, 0, 1, 0, \ldots, 0, \ldots),$$

y, si anotamos con $x^k$ al producto de $x$ por sí mismo $k$ veces, tenemos que $(x^k)_n = 0$ si $n \neq k$, y $(x^k)_k = 1$. Además, haremos la convención de que $x^0 = 1 = (1, 0, \ldots, 0, \ldots)$.

Notemos que el conjunto dado por $\{1, x, x^2, \ldots\}$ es un conjunto linealmente independiente e infinito, con lo cual $\mathbb{K}^\infty$ es un $\mathbb{K}$-espacio vectorial de dimensión infinita. Sin embargo, este conjunto no es una base de $\mathbb{K}^\infty$ ya que no es suficiente para generar a todas las sucesiones de $\mathbb{K}^\infty$. Por ejemplo la sucesión

$$(1, 1, \ldots, 1, \ldots),$$

no puede escribirse como una combinación lineal finita de los vectores en $\{1, x, x^2, \ldots\}$.

Ahora sí, estamos en condiciones de presentar la definición formal del álgebra de polinomios sobre $\mathbb{K}$:

**Definición 3.30.** Anotaremos con $\mathbb{K}[x]$ al subespacio de $\mathbb{K}^\infty$ generado por el conjunto $\{1, x, x^2, \ldots\}$, y si $p \in \mathbb{K}[x]$ diremos que $p$ es un **polinomio sobre el cuerpo** $\mathbb{K}$.

Como $\mathbb{K}[x]$ consta de todas los combinaciones lineales (finitas) de $x$ y sus potencias no negativas, tenemos que $p = (p_n)_{n \in \mathbb{N}_0} \in \mathbb{K}[x]$ si y sólo si existe $m \in \mathbb{N}_0$ tal que $p_m \neq 0$ y $p_n = 0$ para todo $n > m$, lo cual se puede escribir como

$$p = p_0 x^0 + p_1 x^1 + p_2 x^2 + \ldots + p_m x^m.$$

## 3.6. Ejercicios

1. Sea $\mathbb{K}$ un cuerpo. Dados $p, q, r \in \mathbb{K}[x]$, probar que:

   *a*) Si $p \,|\, q$ y $p \,|\, r$ entonces $p \,|\, (mq + nr)$ para todo $m, n \in \mathbb{K}[x]$.

   *b*) Si $p \,|\, q$ y $p \,|\, q + r$ entonces $p \,|\, r$.

   *c*) Si $p \,|\, q$ y $\mathrm{gr}(p) = \mathrm{gr}(q)$ entonces existe $\alpha \in \mathbb{K}$ tal que $p = \alpha\, q$.

2. Sean $p(x) = 2x^3 - 3x^2 + 1$ y $q(x) = 2x^2 + 4x - 1$, mostrar que no existen $c$ y $r$ en $\mathbb{Z}[x]$ tales que $p = cq + r$ y $\mathrm{gr}(r) < \mathrm{gr}(q)$. ¿Qué puede decirse sobre la existencia del algoritmo de división en $\mathbb{Z}[x]$?

3. Probar que en $\mathbb{R}[x]$ no existen polinomios no nulos tales que $p^2 + q^2 = 0$. ¿Ocurre lo mismo en $\mathbb{C}[x]$?

4. Determinar cuáles de los siguientes subconjuntos son ideales:

   *a*) $\{p \in \mathbb{K}[x] : \mathrm{gr}\, p \geq 2\} \subset \mathbb{K}[x]$.

   *b*) $\big\{p\,(x^2 + 4) \in \mathbb{R}[x] : p \in \mathbb{R}[x]\big\} \subset \mathbb{R}[x]$.

   *c*) $p\,\mathbb{K}[x] \subset \mathbb{K}[x]$, para $p \in \mathbb{K}[x]$.

   *d*) Dado $\alpha \in \mathbb{K}$, $\{p \in \mathbb{K}[x] : p(\alpha) = 0\} \subset \mathbb{K}[x]$.

5. Dados $d_1, d_2, \ldots, d_n \in \mathbb{K}[x]$, probar que la suma de los ideales principales $d_i \cdot \mathbb{K}[x]$,

   $$M = d_1 \cdot \mathbb{K}[x] + d_2 \cdot \mathbb{K}[x] + \ldots + d_n \cdot \mathbb{K}[x],$$

   también es un ideal de $\mathbb{K}[x]$.

6. En $\mathbb{R}[x]$, consideremos el ideal $M = (x-1)\mathbb{R}[x] + (x^2-1)\mathbb{R}[x]$. Mostrar que
$$M = (x-1)\mathbb{R}[x].$$
¿Cambia algo si reemplazamos al cuerpo $\mathbb{R}$ por otro cuerpo arbitrario $\mathbb{K}$?

7. Hallar el m. c. d. entre los siguientes polinomios.
   *a*) $p(x) = x^5 - 4x^4 - 3x + 1$ y $q(x) = 3x^2 + 2x + 1$.
   *b*) $p(x) = x^4 - 2x^3 + 1$ y $q(x) = x^2 - x + 2$.
   *c*) $p(x) = 2x^3 - 4x^2 + x - 1$ y $q(x) = x^3 - x^2 + 2x$.

8. Sea $p \in \mathbb{K}[x]$ con $\operatorname{gr} p = n$. Probar que $p$ tiene exactamente $n$ raíces distintas si y sólo si $(p, p') = 1$.

9. Sea $p \in \mathbb{C}[x]$ con $\operatorname{gr} p = n$. Probar que si existen $x_1, \ldots, x_n \in \mathbb{R}$, distintos entre sí, tales que $p(x_i) \in \mathbb{R}$ para todo $i = 1, \cdots, n$ entonces $p$ tiene coeficientes reales.

10. Hallar el generador mónico de los siguientes ideales de $\mathbb{Q}[x]$.
    *a*) $M = \{p \in \mathbb{Q}[x] : p(1) = p(2) = 0\}$.
    *b*) $M = \{p \in \mathbb{Q}[x] : (x - \pi) \mid p\}$.
    *c*) $M = \big\{p \in \mathbb{Q}[x] : p \text{ es divisible por } x^2 + 4 \text{ y } x^4 - 16\big\}$.

11. Sea $T : \mathbb{R}[x] \to \mathcal{F}(\mathbb{R}, \mathbb{R})$ dada por
$$(Tp)(x) = \begin{cases} \dfrac{p(x) - p(0)}{x} & \text{si } x \neq 0 \\ p'(0) & \text{si } x = 0\,. \end{cases}$$
Probar que $T$ es una transformación lineal y que $\operatorname{Im}(T)$ es un subconjunto de las funciones polinomiales.

# Capítulo 4

# Formas canónicas elementales – Diagonalización de operadores

En este capítulo (y los próximos) nos limitaremos a estudiar operadores lineales sobre $\mathbb{K}$-espacios vectoriales de dimensión finita. Dado un operador lineal $T$ actuando en un $\mathbb{K}$-espacio vectorial $V$ de dimensión finita, el objetivo es encontrar una base $\mathcal{B}$ de $V$ tal que la representación matricial $[T]_{\mathcal{B}}$ de $T$ sea lo más "sencilla" posible.

¿Qué significará para nosotros una representación matricial "sencilla"? En este capítulo buscamos caracterizar aquellos operadores lineales que admiten una representación matricial **diagonal**.

Comencemos buscando condiciones necesarias para garantizar una representación diagonal. Dado un $\mathbb{K}$-espacio vectorial $V$ con $\dim V = n$, sea $T \in L(V)$ y supongamos que $\mathcal{B} = \{b_1, \ldots, b_n\}$ es una base ordenada de $V$ tal que la representación matricial $[T]_{\mathcal{B}}$ de $T$ es diagonal:

$$[T]_{\mathcal{B}} = \begin{bmatrix} \lambda_1 & 0 & 0 & \ldots & 0 \\ 0 & \lambda_2 & 0 & \ldots & 0 \\ 0 & 0 & \lambda_3 & \ldots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \ldots & \lambda_n \end{bmatrix} = D.$$

¿Por qué consideramos que una representación matricial diagonal es simple? Con una representación así es fácil calcular, por ejemplo, el determinante de $T$:

$$\det(T) = \prod_{i=1}^{n} \lambda_i.$$

Además los vectores de la base $\mathcal{B}$ satisfacen:

$$Tb_i = \lambda_i \, b_i, \qquad i = 1, \ldots, n. \tag{4.1}$$

A partir de estas condiciones también es inmediato calcular el núcleo y la imagen de $T$:

$$N(T) = \overline{\{b_i : \ \lambda_i = 0\}} \quad \text{y} \quad \operatorname{Im}(T) = \overline{\{b_j : \ \lambda_j \neq \vec{0}\}}.$$

Ahora la cuestión es ¿para todo $T \in L(V)$ existe una base $\mathcal{B}$ tal que $[T]_{\mathcal{B}}$ es diagonal?

Si esto no es así ¿cuáles son los operadores que se pueden diagonalizar? Además, si un operador no se puede diagonalizar, ¿qué otra expresión "sencilla" podemos encontrar para representarlo?

Notemos que en (4.1) encontramos una serie de condiciones necesarias para los elementos de la base $\mathcal{B}$ que permite diagonalizar a $T$. En lo que sigue estudiaremos este fenómeno con más detalle.

## 4.1. Autovalores y autovectores

Comencemos definiendo que entendemos por un autovalor (y un autovector). La nomenclatura puede variar de acuerdo a la fuente bibliográfica consultada, lo que aquí llamamos autovalor otros autores lo llaman eigenvalor, valor propio ó valor característico.

**Definición 4.1.** Sea $T$ un operador lineal actuando en un $\mathbb{K}$-espacio vectorial $V$. Un **autovalor** de $T$ es un escalar $\lambda \in \mathbb{K}$ tal que existe algún vector $v \in V$, $v \neq \vec{0}$ que satisface la ecuación:

$$Tv = \lambda\, v.$$

En tal caso, cualquier $v \in V$, $v \neq \vec{0}$ tal que $Tv = \lambda v$ es un **autovector de** $T$ **asociado a** $\lambda$.

En primer lugar notemos que, si $v \in V$ es un autovector de $T$ asociado a $\lambda$, entonces para cada $\alpha \in \mathbb{K}$, $\alpha \neq 0$, $\alpha v$ también es un autovector de $T$ asociado a $\lambda$. Es decir, todos los vectores del subespacio generado por $v$ (salvo el vector nulo) son también autovectores de $T$ asociados a $\lambda$.

Luego, si $\lambda$ es un autovalor de $T$, definimos el **autoespacio de** $T$ **asociado a** $\lambda$ como

$$E(\lambda) := \{v \in V : \ Tv = \lambda v\}.$$

A pesar de que el vector nulo no es un autovector, también lo aceptamos dentro de $E(\lambda)$.

¿Por qué el vector nulo no es un autovector? Porque el vector nulo satisface $T\vec{0} = \lambda\,\vec{0}$ para todo $\lambda \in \mathbb{K}$.

**Ejercicio 4.2.** Dados $T \in L(V)$ y $\lambda \in \mathbb{K}$, mostrar que $E(\lambda)$ es un subespacio de $V$. Más aún, probar que

$$E(\lambda) = N(T - \lambda I).$$

¿En qué casos $E(\lambda)$ es un subespacio propio de $V$?

El resultado que sigue caracteriza a los autovalores de un operador lineal dado. La prueba es muy simple y la dejamos como ejercicio para el lector.

**Teorema 4.3.** *Dado un $\mathbb{K}$-espacio vectorial $V$ de dimensión finita, sea $T \in L(V)$. Fijado $\lambda \in \mathbb{K}$, las siguientes condiciones son equivalentes:*

*i)* *$\lambda$ es un autovalor de $T$;*

*ii)* *$T - \lambda I$* **no** *es inversible;*

*iii)* $\det(T - \lambda I) = 0$.

Lo importante de este resultado es que nos provee de un método práctico (y relativamente sencillo) para encontrar los autovalores de un operador dado. Si $V$ es un $\mathbb{K}$-espacio vectorial con $\dim V = n$ entonces

$$p(x) = \det(T - xI),$$

es un polinomio de grado $n$ con respecto a la variable $x$, i.e. $p \in \mathbb{K}^{(n)}[x]$. Luego, de acuerdo al teorema anterior,

$$\lambda \text{ es un autovalor de } T \quad \Leftrightarrow \quad \lambda \text{ es una raíz de } p.$$

Esto refuerza aún más la idea de que los autovalores dependen del cuerpo $\mathbb{K}$ para el cual $V$ es un $\mathbb{K}$-espacio vectorial. Veamos un ejemplo de esto.

**Ejemplo 4.4.** Sea $T : \mathbb{K}^2 \to \mathbb{K}^2$ dado por $T(w, z) = (-z, w)$.

Calculemos primero los autovalores de $T$ suponiendo que $\mathbb{K} = \mathbb{R}$. En este caso, observemos que $T$ representa una rotación de 90° (en sentido antihorario) alrededor del origen en $\mathbb{R}^2$. Dado un vector no nulo arbitrario $v = (x, y) \in \mathbb{R}^2$, es obvio que al rotarlo 90° no obtendremos un múltiplo escalar de $v$ (verificarlo analíticamente). En conclusión, si $\mathbb{K} = \mathbb{R}$ entonces $T$ no tiene autovectores (y en consecuencia no tiene autovalores).

Ahora, supongamos que $\mathbb{K} = \mathbb{C}$. Buscamos aquellos $\lambda \in \mathbb{C}$ tales que $T(w, z) = \lambda(w, z)$ para algún vector no nulo $v = (w, z) \in \mathbb{C}^2$. Esto equivale a encontrar soluciones para las ecuaciones

$$-z = \lambda w, \quad w = \lambda z,$$

de las cuales se deduce que $-z = \lambda^2 z$. Como $z \neq 0$ (de lo contrario, las ecuaciones anteriores implican que $w = 0$ y el vector $v$ sería el nulo), buscamos las soluciones de la ecuación

$$\lambda^2 = -1,$$

las cuales son $\lambda_1 = i$ y $\lambda_2 = -i$. Por lo tanto, los autovalores de $T$ son exactamente éstos.

**Definición 4.5.** Sea $V$ un $\mathbb{K}$-espacio vectorial con $\dim V = n$. Dado $T \in L(V)$, **el polinomio característico de** $T$ se define como

$$p_T(x) = \det(xI - T), \tag{4.2}$$

el cual es un polinomio mónico de grado $n$.

En general, la forma más simple de calcular el polinomio característico (y los autovalores) es mediante una representación matricial del operador $T$. Como la noción de determinante no depende de la representación matricial elegida, y como

$$[xI - T]_{\mathcal{B}} = x[I]_{\mathcal{B}} - [T]_{\mathcal{B}} = xI_n - [T]_{\mathcal{B}},$$

para cualquier base $\mathcal{B}$ de $V$, el polinomio característico tampoco depende de la representación matricial.

**Ejemplo 4.6.** Sea $T \in L(\mathbb{R}^3)$ tal que $[T]_{\mathcal{E}} = \begin{bmatrix} 3 & 1 & -1 \\ 2 & 2 & -1 \\ 2 & 2 & 0 \end{bmatrix} = A.$

Entonces,

$$p_T(x) = \det(xI - T) = \det(xI - A) = \begin{vmatrix} x-3 & -1 & 1 \\ -2 & x-2 & 1 \\ -2 & -2 & x \end{vmatrix} =$$
$$= x^3 - 5x^2 + 8x - 4 = (x-1)(x-2)^2.$$

Por lo tanto, los autovalores de $T$ son $\lambda_1 = 1$ y $\lambda_2 = 2$.

Calculemos ahora los autoespacios asociados. Un vector $v = (x, y, z) \in E(1) = N(T - I)$ si y sólo si

$$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \vec{0} = (A - I)[v]_{\mathcal{E}} = \begin{bmatrix} 2 & 1 & -1 \\ 2 & 1 & -1 \\ 2 & 2 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix},$$

de lo cual se deduce que $y = 0$ y $z = 2x$, es decir, $v = (x, 0, 2x)$. Por lo tanto,

$$E(1) = \overline{\{(1, 0, 2)\}}.$$

En cambio, $v = (x, y, z) \in E(2) = N(T - 2I)$ si y sólo si

$$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \vec{0} = (A - 2I)[v]_{\mathcal{E}} = \begin{bmatrix} 1 & 1 & -1 \\ 2 & 0 & -1 \\ 2 & 2 & -2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix},$$

de donde obtenemos que $z = 2x$ y $z = x + y$. Luego,

$$E(2) = \overline{\{(1, 1, 2)\}}.$$

**Definición 4.7.** Dado un $\mathbb{K}$-espacio vectorial $V$ de dimensión finita, sea $T \in L(V)$. Diremos que $T$ es **diagonalizable** si existe una base $\mathcal{B}$ de $V$ conformada por autovectores del operador $T$.

El operador $T$ del Ejemplo 4.6 no es diagonalizable ya que tenemos a lo sumo dos autovectores de $T$ linealmente independientes y $\dim \mathbb{R}^3 = 3$. El operador $T : \mathbb{R}^2 \to \mathbb{R}^2$ del Ejemplo 4.4 tampoco es diagonalizable porque no tiene autovalores (y por consiguiente, no existen autovectores de $T$). En cambio, si pensamos a $T$ como un operador lineal sobre $\mathbb{C}^2$, éste sí es diagonalizable (comprobarlo!).

**Ejemplo 4.8.** Consideremos ahora el operador $S : \mathbb{R}^3 \to \mathbb{R}^3$ que representa a la simetría con respecto al plano $x = y$ en $\mathbb{R}^3$. Es decir, $S(x, y, z)$ es el vector que resulta de reflejar a $(x, y, z)$ con respecto al plano $x = y$. O sea,

$$S(x, y, z) = (y, x, z).$$

Veamos si $S$ es o no un operador diagonalizable. Su representación matricial con respecto a la base canónica de $\mathbb{R}^3$ es $[S]_{\mathcal{E}} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = A$, con lo cual

$$p_S(x) = \det(xI - S) = \begin{vmatrix} x & -1 & 0 \\ -1 & x & 0 \\ 0 & 0 & x - 1 \end{vmatrix} = (x - 1)(x^2 - 1) = (x - 1)^2(x + 1).$$

Por lo tanto, los autovalores de $S$ son $\lambda_1 = 1$ y $\lambda_2 = -1$.

Calculemos ahora los autoespacios asociados. Un vector $v = (x, y, z) \in E(1) = N(S - I)$ si y sólo si

$$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \vec{0} = (A - I)[v]_{\mathcal{E}} = \begin{bmatrix} -1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix},$$

de lo cual se deduce que $y = x$, es decir, $v = (x, x, z)$. Por lo tanto,

$$E(1) = \overline{\{(1, 1, 0), (0, 0, 1)\}}.$$

En cambio, $v = (x, y, z) \in E(2) = N(S + I)$ si y sólo si

$$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \vec{0} = (A - 2I)[v]_{\mathcal{E}} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix},$$

de donde obtenemos que $y = -x$ y $z = 0$. Luego,

$$E(-1) = \overline{\{(1, -1, 0)\}}.$$

Notemos que $\mathcal{B} = \{(1,1,0), (0,0,1), (1,-1,0)\}$ es una base de $\mathbb{R}^3$ conformada por autovectores de $S$. Por lo tanto, $S$ es diagonalizable.

A continuación veremos que el fenómeno del ejemplo anterior no es una casualidad, autovectores correspondientes a autovalores diferentes resultan ser siempre linealmente independientes.

**Lema 4.9.** *Dado un $\mathbb{K}$-espacio vectorial $V$ de dimensión finita, sea $T \in L(V)$. Si $\lambda_1, \ldots, \lambda_k$ son autovalores de $T$ distintos entre sí y $v_1, \ldots, v_k$ son autovectores de $T$ asociados a $\lambda_1, \ldots, \lambda_k$, respectivamente, entonces*

$$\{v_1, \ldots, v_k\} \quad \textit{es linealmente independiente.}$$

*Demostración.* La realizaremos utilizando el principio de inducción completa sobre la cantidad de autovalores $k$. Si $k = 2$, sean $v_1, v_2 \in V$ vectores no nulos tales que

$$Tv_1 = \lambda_1 v_1 \quad \text{y} \quad Tv_2 = \lambda_2 v_2.$$

Supongamos que existen $\alpha_1, \alpha_2 \in \mathbb{K}$ tales que $\alpha_1 v_1 + \alpha_2 v_2 = \vec{0}$. Aplicando el operador $T$ a cada lado de esta ecuación obtenemos:

$$\vec{0} = T(\vec{0}) = T(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 T v_1 + \alpha_2 T v_2 = \alpha_1 \lambda_1 v_1 + \alpha_2 \lambda_2 v_2.$$

Como sabemos que $\alpha_1 v_1 = -\alpha_2 v_2$, resulta que

$$\vec{0} = \lambda_1 \alpha_1 v_1 + \lambda_2 \alpha_2 v_2 = \lambda_1(-\alpha_2 v_2) + \lambda_2 \alpha_2 v_2 = (\lambda_2 - \lambda_1)\alpha_2 v_2.$$

De esto se deduce que $\alpha_2 = 0$, y reemplazándolo en la ecuación original resulta también que $\alpha_1 = 0$. Por lo tanto, $\{v_1, v_2\}$ es linealmente independiente.

Como hipótesis inductiva, supongamos que si $v_1, \ldots, v_{k-1}$ son autovectores asociados a autovalores $\lambda_1, \ldots, \lambda_{k-1}$ de $T$ distintos entre sí, entonces $\{v_1, \ldots, v_{k-1}\}$ es linealmente independiente.

Ahora, sean $v_1, \ldots, v_k$ autovectores asociados a autovalores $\lambda_1, \ldots, \lambda_k$ de $T$, distintos entre sí. Supongamos que existen $\alpha_1, \ldots, \alpha_k \in \mathbb{K}$ tales que

$$\alpha_1 v_1 + \ldots + \alpha_{k-1} v_{k-1} + \alpha_k v_k = \vec{0}. \tag{4.3}$$

Una vez más, aplicando el operador $T$ a cada lado de esta ecuación obtenemos:

$$\begin{aligned}\vec{0} = T(\vec{0}) = T(\alpha_1 v_1 + \ldots + \alpha_{k-1} v_{k-1} + \alpha_k v_k) &= \alpha_1 T v_k + \ldots + \alpha_{k-1} T v_{k-1} + \alpha_k T v_k \\ &= \alpha_1 \lambda_1 v_1 + \ldots + \alpha_{k-1} \lambda_{k-1} v_{k-1} + \alpha_k \lambda_k v_k.\end{aligned}$$

De (4.3) sabemos que $\alpha_k v_k = -\alpha_1 v_1 - \ldots - \alpha_{k-1} v_{k-1}$, y reemplazando en la ecuación anterior tenemos

$$\begin{aligned}\vec{0} &= \lambda_1 \alpha_1 v_1 + \ldots + \lambda_{k-1} \alpha_{k-1} v_{k-1} + \lambda_k(-\alpha_1 v_1 - \ldots - \alpha_{k-1} v_{k-1}) \\ &= (\lambda_1 - \lambda_k)\alpha_1 v_1 + \ldots + (\lambda_{k-1} - \lambda_k)\alpha_{k-1} v_{k-1}.\end{aligned}$$

Por la hipótesis inductiva, los escalares que acompañan a los vectores $v_1, \ldots, v_{k-1}$ deben ser todos nulos y, como los autovalores son distintos entre sí, resulta que $\alpha_1 = \ldots = \alpha_{k-1} = 0$. Finalmente, por (4.3), tenemos también que $\alpha_k = 0$.

Por lo tanto, $\{v_1, \ldots, v_k\}$ es linealmente independiente. $\square$

**Corolario 4.10.** *Si $V$ es un $\mathbb{K}$-espacio vectorial con $\dim V = n$, entonces cada operador $T \in L(V)$ tiene a lo sumo $n$ autovalores distintos.*

*Demostración.* Dado $T \in L(V)$, supongamos que $\lambda_1, \ldots, \lambda_k$ son autovalores de $T$ distintos entre sí. Sean $v_1, \ldots, v_k$ autovectores asociados a $\lambda_1, \ldots, \lambda_k$, respectivamente. Por la lema anterior, éstos son linealmente independiente. Por lo tanto, $k \leq n$, como queríamos probar. □

**Corolario 4.11.** *Dado un $\mathbb{K}$-espacio vectorial $V$ con $\dim V = n$, si $T \in L(V)$ tiene $n$ autovalores distintos, entonces $T$ es diagonalizable.*

## 4.2. Subespacios independientes

El concepto que sigue será esencial para poder descomponer a un espacio vectorial en suma de subespacios.

**Definición 4.12.** Dado un $\mathbb{K}$-espacio vectorial $V$, sean $W_1, \ldots, W_k$ subespacios de $V$. Diremos que los subespacios $W_1, \ldots, W_k$ son **subespacios independientes** si cada vector $v \in W_1 + \ldots + W_k$ admite una única representación de la forma

$$v = w_1 + \ldots + w_k,$$

donde cada $w_i \in W_i$, $i = 1, \ldots, k$.

**Observación 4.13.** Los subespacios $W_1, \ldots, W_k$ son independientes si y sólo si

$$\vec{0} = w_1 + \ldots + w_k \;\text{ con } w_i \in W_i \quad \Longrightarrow \quad w_i = \vec{0} \;\text{ para todo } i = 1, \ldots, k, \tag{4.4}$$

es decir, la única escritura posible para el vector nulo como combinación de vectores $w_i \in W_i$ es la obvia.

Si $W_1, \ldots, W_k$ son subespacios independientes es inmediato que la implicación (4.4) es verdadera. Recíprocamente, supongamos que (4.4) es cierta. Supongamos además que cierto vector $v \in W_1 + \ldots + W_k$ admite dos representaciones, digamos

$$v = v_1 + v_2 + \ldots + v_k \quad \text{y} \quad v = w_1 + w_2 + \ldots + w_k,$$

donde $v_i, w_i \in W_i$ para cada $i = 1, \ldots, k$. Luego, la diferencia entre ambas expresiones da como resultado:

$$\vec{0} = (v_1 - w_1) + (v_2 - w_2) + \ldots + (v_k - w_k).$$

Aplicando la implicación (4.4) obtenemos que $v_i = w_i$ para cada $i = 1, \ldots, k$, es decir, el vector $v$ admite una única representación. Por lo tanto, $W_1, \ldots, W_k$ son subespacios independientes.

En el ejercicio 17 del Capítulo 1 caracterizamos cuando un par de subespacios (cuya suma es todo el $\mathbb{K}$-espacio vectorial $V$) son independientes. La siguiente proposición generaliza este resultado para una familia finita arbitraria de subespacios.

**Proposición 4.14.** *Dado un $\mathbb{K}$-espacio vectorial $V$, sean $W_1, \ldots, W_k$ subespacios de $V$. Las siguientes condiciones son equivalentes:*

*i)* $W_1, \ldots, W_k$ *son subespacios independientes.*

*ii)* $W_l \cap (W_1 + \ldots + W_{l-1}) = \{\vec{0}\}$ *para cada* $l = 2, \ldots, k$.

*iii)* *Si $\mathcal{B}_i$ es una base ordenada de $W_i$ para cada $i = 1, \dots, k$, entonces la yuxtaposición de las mismas $\mathcal{B} = \{\mathcal{B}_1, \dots, \mathcal{B}_k\}$ es una base oredenada de $W_1 + \dots + W_k$.*

*Demostración.*
*i)* $\Rightarrow$ *ii)*: Supongamos que $W_1, \dots, W_k$ son subespacios independientes. Ahora, si $v \in W_l \cap (W_1 + \dots + W_{l-1})$, existen vectores $w_1 \in W_1, \dots, w_{l-1} \in W_{l-1}$ tales que $v = w_1 + \dots + w_{l-1}$. Entonces,

$$\vec{0} = \overbrace{w_1}^{\in W_1} + \dots + \overbrace{w_{l-1}}^{\in W_{l-1}} + \overbrace{(-v)}^{\in W_l} + \overbrace{\vec{0} + \dots + \vec{0}}^{k-l \text{ veces}}.$$

Luego, la implicación (4.4) garantiza que $w_1 = \dots = w_{l-1} = \vec{0}$ y $v = \vec{0}$. Por lo tanto, $W_l \cap (W_1 + \dots + W_{l-1}) = \{\vec{0}\}$.
*ii)* $\Rightarrow$ *iii)*: Si suponemos que vale *ii)*, en particular tenemos que $W_k \cap (W_1 + \dots + W_{k-1}) = \{\vec{0}\}$. Si $\mathcal{B}_1, \dots, \mathcal{B}_k$ son bases de $W_1, \dots, W_k$, respectivamente, es fácil convencerse de que $\mathcal{B}$ es un sistema de generadores de $W_1 + \dots + W_k$.

Sólo resta probar que $\mathcal{B}$ es linealmente independiente. Para esto, supongamos que cada base $\mathcal{B}_i$ es de la forma $\mathcal{B}_i = \{b_1^{(i)}, \dots, b_{n_i}^{(i)}\}$, donde $n_i = \dim W_i$. Luego, supongamos que existen escalares $\alpha_j^{(i)} \in \mathbb{K}$ tales que

$$\sum_{i=1}^{k} \sum_{j=1}^{n_i} \alpha_j^{(i)} b_j^{(i)} = \sum_{i=1}^{k} \left( \sum_{j=1}^{n_i} \alpha_j^{(i)} b_j^{(i)} \right) = \vec{0}. \tag{4.5}$$

Entonces,

$$\sum_{i=1}^{k-1} \left( \sum_{j=1}^{n_i} \alpha_j^{(i)} b_j^{(i)} \right) = \sum_{j=1}^{n_k} \alpha_j^{(k)} b_j^{(k)}.$$

Como cada suma $\sum_{j=1}^{n_i} \alpha_j^{(i)} b_j^{(i)}$ es un vector en $W_i$, la ecuación anterior dice que $\sum_{j=1}^{n_k} \alpha_j^{(k)} b_j^{(k)} \in W_k \cap (W_1 + \dots + W_{k-1})$. Entonces, $\sum_{j=1}^{n_k} \alpha_j^{(k)} b_j^{(k)} = \vec{0}$ y la independencia lineal de $\mathcal{B}_k$ garantiza que $\alpha_j^{(k)} = 0$ para $j = 1, \dots, n_k$.

Volviendo a (4.5), tenemos que $\sum_{i=1}^{k-1} \left( \sum_{j=1}^{n_i} \alpha_j^{(i)} b_j^{(i)} \right) = \vec{0}$, o equivalentemente,

$$\sum_{i=1}^{k-2} \left( \sum_{j=1}^{n_i} \alpha_j^{(i)} b_j^{(i)} \right) = \sum_{j=1}^{n_{k-1}} \alpha_j^{(k-1)} b_j^{(k-1)} \in W_{k-1} \cap (W_1 + \dots + W_{k-2}).$$

Como por hipótesis este subespacio también es trivial, repitiendo el argumento anterior resulta que $\alpha_j^{(k-1)} = 0$ para $j = 1, \dots, n_{k-1}$.

Repitiendo este proceso las $k-2$ veces que faltan, obtendremos que $\alpha_j^{(i)} = 0$ para todo $i$ y para todo $j$. Por lo tanto, $\mathcal{B}$ es linealmente independiente.
*iii)* $\Rightarrow$ *i)*: Dadas bases $\mathcal{B}_i = \{b_1^{(i)}, \dots, b_{n_i}^{(i)}\}$ de $W_i$, para $i = 1, \dots, k$, supongamos que $\mathcal{B} = \{\mathcal{B}_1, \dots, \mathcal{B}_k\}$ es una base ordenada de $W_1 + \dots + W_k$. Para ver que $W_1, \dots, W_k$ son subespacios independientes utilizaremos (4.4). Es decir, supongamos que existen vectores $w_i \in W_i$ tales que $w_1 + \dots + w_k = \vec{0}$ y veamos que entonces cada $w_i$ debe ser el vector nulo.

Escribamos a cada $w_i$ como combinación lineal de los elementos de la base $\mathcal{B}_i$, es decir, sean $\beta_1^{(i)}, \dots, \beta_{n_i}^{(i)} \in \mathbb{K}$ tales que $w_i = \beta_1^{(i)} b_1^{(i)} + \dots + b_{n_i}^{(i)} \beta_{n_i}^{(i)}$. Entonces,

$$\vec{0} = w_1 + \dots + w_k = \sum_{i=1}^{k} \left( \sum_{j=1}^{n_i} \beta_j^{(i)} b_j^{(i)} \right).$$

Ahora, como $\mathcal{B}$ es una base de $W_1+\ldots+W_k$, resulta que $\beta_j^{(i)}=0$ para todo $i$ y para todo $j$. En particular, tenemos que $w_i=\vec{0}$ para $i=1,\ldots,k$ y, por (4.4), los subespacios $W_1,\ldots,W_k$ son independientes. □

**Definición 4.15.** Si $W_1,\ldots,W_k$ son subespacios independientes, diremos que la suma de éstos es una **suma directa** y la anotaremos

$$W_1 \dotplus \ldots \dotplus W_k. \tag{4.6}$$

Por ejemplo, en el ejercicio 16 del Capítulo 1 mostramos que el espacio de funciones continuas sobre la recta real se descompone como la suma directa del subespacio de funciones pares y el de funciones impares.

**Ejercicio 4.16.** Dado un $\mathbb{K}$-espacio vectorial $V$, sean $W_1,\ldots,W_k$ subespacios independientes tales que $V=W_1\dotplus\ldots\dotplus W_k$. Supongamos que $S_1,\ldots,S_k$ es otra familia de subespacios independientes que satisfacen $V=S_1\dotplus\ldots\dotplus S_k$ y además $S_i\subseteq W_i$ para todo $i=1,\ldots,k$. Probar que entonces forzosamente $S_i=W_i$ para todo $i=1,\ldots,k$.

A continuación probaremos que los autoespacios asociados a autovalores diferentes son subespacios independientes.

**Proposición 4.17.** *Dado un $\mathbb{K}$-espacio vectorial $V$ de dimensión finita, sea $T\in L(V)$. Si $\lambda_1,\ldots,\lambda_k$ son autovalores de $T$ distintos entre sí, entonces $E(\lambda_1),\ldots,E(\lambda_k)$ son subespacios independientes.*

*Demostración.* Este resultado es consecuencia del Lema 4.9. Sean $v_i\in E(\lambda_i)$, $i=1,\ldots,k$ tales que $v_1+\ldots+v_k=\vec{0}$. Queremos ver que $v_1=\ldots=v_k=\vec{0}$. Por el absurdo, supongamos que alguno de los $v_i$ es no nulo. Si $v_{j_1},\ldots,v_{j_l}$ son los vectores no nulos entre $v_1,\ldots,v_k$, entonces

$$v_{j_1}+v_{j_2}+\ldots+v_{j_l}=\vec{0}.$$

Pero, por el Lema 4.9, $v_{j_1},\ldots,v_{j_l}$ son linealmente independiente, lo que resulta una contradicción. □

**Ejercicio 4.18.** Utilizar la Proposición 4.14 para probar (usando el principio de inducción completa) que, si $W_1,\ldots,W_k$ son subespacios independientes, entonces

$$\dim(W_1+\ldots+W_k)=\sum_{i=1}^{k}\dim(W_i).$$

En particular, si $\lambda_1,\ldots,\lambda_k$ son los distintos autovalores de un operador $T$ actuando en un $\mathbb{K}$-espacio vectorial $V$ con $\dim V=n$, entonces

$$\sum_{i=1}^{k}\dim E(\lambda_i)\leq n.$$

**Definición 4.19.** Si $\lambda$ es un autovalor del operador lineal $T$, consideraremos dos nociones distintas de multiplicidad:

- la **multiplicidad geométrica** de $\lambda$ es la dimensión del autoespacio asociado,

$$d:=\dim E(\lambda);$$

- la **multiplicidad algebraica** de $\lambda$ es el mayor $k \in \mathbb{N}$ tal que $(x-\lambda)^k$ divide al polinomio característico de $T$, es decir,
$$(x-\lambda)^k \mid p_T(x) \quad \text{y} \quad (x-\lambda)^{k+1} \nmid p_T(x).$$

**Observación 4.20.** Notemos que, en general, la multiplicidad geométrica de un autovalor es menor o igual a la multiplicidad algebraica correspondiente.

De hecho, dado un autovalor $\lambda$ de $T$ supongamos que su multiplicidad geométrica es $d$. Luego, tenemos $d$ vectores linealmente independientes $v_1, \ldots, v_d \in V$ tales que $Tv_i = \lambda v_i$ para todo $i = 1, \ldots, d$. Si extendemos este conjunto a una base $\mathcal{B} = \{v_1, \ldots, v_d, v_{d+1}, \ldots, v_n\}$ de $V$ resulta que
$$A = [T]_{\mathcal{B}} = \begin{pmatrix} \lambda \cdot I_d & B \\ 0 & C \end{pmatrix},$$
donde $I_d \in \mathbb{K}^{d\times d}$ es la matriz identidad, $B \in \mathbb{K}^{d\times(n-d)}$, $C \in \mathbb{K}^{(n-d)\times(n-d)}$ y $0 \in \mathbb{K}^{(n-d)\times d}$ es la matriz nula.

Como el polinomio característico de $T$ coincide con el de $A$, tenemos que
$$\begin{aligned} p_T(x) &= \det(xI - T) = \det(xI_n - A) = \det \begin{pmatrix} (x-\lambda)\cdot I_d & -B \\ 0 & xI_{n-d} - C \end{pmatrix} \\ &= (x-\lambda)^d \det(xI_{n-d} - C), \end{aligned}$$
es decir, $(x-\lambda)^d \mid p_T(x)$. Por lo tanto, la multiplicidad algebraica de $\lambda$ es mayor o igual a $d$.

El próximo resultado dice que $T$ es diagonalizable si y sólo si las multiplicidades geométrica y algebraica de $\lambda$ coinciden, para cada autovalor $\lambda$ de $T$.

**Teorema 4.21.** *Dado un $\mathbb{K}$-espacio vectorial $V$ con dimensión finita, sea $T \in L(V)$. Sean $\lambda_1, \ldots, \lambda_k$ los distintos autovalores de $T$ y $d_1, \ldots, d_k$ sus respectivas multiplicidades geométricas. Entonces, las siguientes condiciones son equivalentes:*

*i) $T$ es diagonalizable;*

*ii) el polinomio característico de $T$ puede factorizarse como*
$$p_T(x) = (x-\lambda_1)^{d_1}(x-\lambda_2)^{d_2}\ldots(x-\lambda_k)^{d_k}; \tag{4.7}$$

*iii) $d_1 + d_2 + \ldots + d_k = \dim V$.*

*Demostración.*
*i)* $\Rightarrow$ *ii)*: Si $T$ es diagonalizable y $d_i = \dim E(\lambda_i)$ para $i = 1, \ldots, k$, entonces existe una base $\mathcal{B}$ de $V$ tal que
$$[T]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 \cdot I_{d_1} & 0 & \ldots & 0 \\ 0 & \lambda_2 \cdot I_{d_2} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_k \cdot I_{d_k} \end{pmatrix},$$
es decir, $[T]_{\mathcal{B}}$ es una matriz diagonal con el autovalor $\lambda_1$ repetido $d_1$ veces, el autovalor $\lambda_2$ repetido $d_2$ veces, etc. En efecto, si para cada $i = 1, \ldots, k$ elegimos una base $\mathcal{B}_i = \{b_1^{(i)}, \ldots, b_{d_i}^{(i)}\}$ de $E(\lambda_i)$, la yuxtaposición $\mathcal{B} = \{\mathcal{B}_1, \ldots, \mathcal{B}_k\}$ es una base de $V$ (porque $T$ es diagonalizable). Además, para cada $i = 1, \ldots, k$ y para cada $j = 1, \ldots, d_i$ tenemos que
$$Tb_j^{(i)} = \lambda_i b_j^{(i)},$$

es decir, la representación matricial de $T$ con respecto a esta base coincide con la expuesta más arriba. A partir de esta expresión para $[T]_{\mathcal{B}}$ es inmediato que vale la factorización (4.7).

*ii)* $\Rightarrow$ *iii)*: Supongamos que vale (4.7). Luego,

$$\dim V = \mathrm{gr}(p_T) = \mathrm{gr}\left(\prod_{i=1}^{k}(x-\lambda_i)^{d_i}\right) = \sum_{i=1}^{k} d_i.$$

*iii)* $\Rightarrow$ *i)*: Por lo Proposición 4.17, tenemos que $E(\lambda_1), \ldots, E(\lambda_k)$ son subespacios independientes. En particular, si $\mathcal{B}_1, \ldots, \mathcal{B}_k$ son bases de $E(\lambda_1), \ldots, E(\lambda_k)$, respectivamente, entonces $\mathcal{B} = \{\mathcal{B}_1, \ldots, \mathcal{B}_k\}$ es una base de $E(\lambda_1) \dot{+} \ldots \dot{+} E(\lambda_k)$. Como

$$\dim(E(\lambda_1) \dot{+} \ldots \dot{+} E(\lambda_k)) = \sum_{i=1}^{k} \dim E(\lambda_i) = \sum_{i=1}^{k} d_i = \dim V,$$

resulta que $V = E(\lambda_1) \dot{+} \ldots \dot{+} E(\lambda_k)$ y $\mathcal{B}$ es una base de $V$ compuesta por autovectores de $T$. Por lo tanto, $T$ es diagonalizable. □

## 4.3. Descomposiciones en sumas directas – Proyecciones

En esta sección discutiremos sobre las proyecciones, las cuales son los operadores lineales más simples que podemos encontrar. Las utilizaremos como los "ladrillitos" con los que construiremos a los demás operadores.

**Definición 4.22.** Dado un $\mathbb{K}$-espacio vectorial $V$, un operador lineal $E \in L(V)$ es una **proyección** si es un idempotente, es decir, si

$$E^2 = E.$$

En primer lugar, notemos que si $E$ es una proyección entonces $E^k = E$ para todo $k \in \mathbb{N}$.

**Observación 4.23.** Dado un $\mathbb{K}$-espacio vectorial $V$, sea $E \in L(V)$ una proyección. Si llamamos $\mathcal{R}$ a la imagen de $E$ y $\mathcal{N}$ al núcleo de $E$, entonces:

i) $w \in \mathcal{R}$ si y sólo si $Ew = w$.

ii) Para cualquier $v \in V$ se tiene que $Ev \in \mathcal{R}$ y $(I-E)v \in \mathcal{N}$. Además,

$$v = Ev + (I-E)v. \tag{4.8}$$

iii) Lo anterior garantiza que $V = \mathcal{R} \dot{+} \mathcal{N}$.

La prueba de los ítems *i)* y *ii)* es inmediata, así que queda como ejercicio para el lector. Además, (4.8) muestra que la suma de los subespacios $\mathcal{R}$ y $\mathcal{N}$ coincide con $V$. Sólo nos falta probar que la suma es directa. Pero si $v \in \mathcal{R} \cap \mathcal{N}$ entonces $Ev = v$ (porque $v \in \mathcal{R}$) y $Ev = \vec{0}$ (porque $v \in \mathcal{N}$), en consecuencia $v = Ev = \vec{0}$. Por lo tanto, $V = \mathcal{R} \dot{+} \mathcal{N}$.

Recíprocamente, veamos que si $W_1$ y $W_2$ son dos subespacios de $V$ tales que

$$V = W_1 \dot{+} W_2,$$

existe una **única proyección** que tiene como imagen a $W_1$ y como núcleo a $W_2$: si $V = W_1 \dot{+} W_2$, para cada $v \in V$ existen únicos vectores $w_1 \in W_1$ y $w_2 \in W_2$ tales que $v = w_1 + w_2$. Luego, definamos el operador $P_{W_1//W_2} : V \to V$ de la siguiente manera:

$$P_{W_1//W_2}\, v = w_1.$$

Es fácil ver que $P_{W_1//W_2}$ es lineal y que es una proyección. Además, notemos que

$$\begin{aligned} v \in \operatorname{Im}(P_{W_1//W_2}) \quad &\Leftrightarrow \quad v = P_{W_1//W_2}\, v = w_1 \quad \Leftrightarrow \quad v \in W_1, \\ v \in N(P_{W_1//W_2}) \quad &\Leftrightarrow \quad P_{W_1//W_2}\, v = \vec{0} \quad \Leftrightarrow \quad v \in W_2, \end{aligned}$$

es decir, $\operatorname{Im}(P_{W_1//W_2}) = W_1$ y $N(P_{W_1//W_2}) = W_2$. La unicidad de tal proyección es consecuencia del siguiente ejercicio.

**Ejercicio 4.24.** Sean $E_1, E_2 \in L(V)$ dos proyecciones. Probar que, si $\operatorname{Im}(E_1) = \operatorname{Im}(E_2)$ y $N(E_1) = N(E_2)$ entonces $E_1 = E_2$.

Si $V = W_1 \dot{+} W_2$, a la proyección $P_{W_1//W_2}$ la llamaremos **la proyección sobre $W_1$ a lo largo de $W_2$**.

Si $V$ es un $\mathbb{K}$-espacio vectorial con $\dim V = n$, entonces cualquier proyección $E \in L(V)$ es diagonalizable, ya que si $\{b_1, \ldots, b_r\}$ es una base de $\operatorname{Im}(E)$ y $\{b_{r+1}, \ldots, b_n\}$ es una base de $N(E)$ entonces $\mathcal{B} = \{b_1, \ldots, b_n\}$ es base de $V$, y además

$$[E]_{\mathcal{B}} = \begin{pmatrix} I_r & 0_{r,n-r} \\ 0_{n-r,r} & 0_{n-r,n-r} \end{pmatrix},$$

donde $I_r$ es la matriz identidad de $\mathbb{K}^{r\times r}$ y $0_{p,q}$ indica a la matriz nula de $\mathbb{K}^{p\times q}$.

**Observación 4.25.** Si $V = W_1 \dot{+} W_2$ entonces

$$P_{W_2//W_1} = I - P_{W_1//W_2}.$$

En general, si $W_1, \ldots, W_k$ son subespacios de $V$ tales que $V$ se descompone como

$$V = W_1 \dot{+} \ldots \dot{+} W_k,$$

entonces existen $k$ proyecciones que descomponen al operador identidad.

**Teorema 4.26.** *Si $W_1, \ldots, W_k$ son subespacios de $V$ tales que $V = W_1 \dot{+} \ldots \dot{+} W_k$, entonces existen $E_1, \ldots, E_k \in L(V)$ tales que:*

*i) cada $E_i$ es una proyección, i.e. $E_i^2 = E_i$;*

*ii) $E_j E_i = 0$ si $i \neq j$;*

*iii) $I = E_1 + \ldots + E_k$;*

*iv) $\operatorname{Im}(E_i) = W_i$.*

*Recíprocamente, si existen operadores $E_1, \ldots, E_k \in L(V)$ que satisfacen i), ii) y iii) y llamamos $W_i = \operatorname{Im}(E_i)$, entonces*

$$V = W_1 \dot{+} \ldots \dot{+} W_k.$$

*Demostración.* Supngamos que $V = W_1 \dot{+} \ldots \dot{+} W_k$. Luego, dado $v \in V$ existen únicos $w_i \in W_i$ tales que $v = w_1 + \ldots + w_k$. Para cada $j = 1, \ldots, k$ definamos

$$E_j(v) = w_j.$$

Es fácil ver que $E_j \in L(V)$, que $\operatorname{Im}(E_j) = W_j$ y que $E_j^2 = E_j$. Entonces, para cada $v \in V$ tenemos que

$$v = w_1 + \ldots + w_k = E_1(v) + \ldots + E_k(v) = (E_1 + \ldots + E_k)(v),$$

de donde se desprende $iii$).

Además, es inmediato que el núcleo de $E_j$ es

$$N(E_j) = W_1 \dot{+} \ldots \dot{+} W_{j-1} \dot{+} W_{j+1} \dot{+} \ldots \dot{+} W_k.$$

Entonces, si $i \neq j$ tenemos que $\mathrm{Im}(E_i) = W_i \subseteq N(E_j)$. En consecuencia, $E_j E_i = 0$.

Recípocramente, supongamos que $E_1, \ldots, E_k \in L(V)$ son operadores que satisfacen $i$), $ii$) y $iii$) y llamemos $W_i = \mathrm{Im}(E_i)$. Dado un vector $v \in V$ cualquiera, tenemos que

$$v = I(v) = (E_1 + \ldots + E_k)(v) = E_1(v) + \ldots + E_k(v).$$

Como $E_i v \in W_i$ para cada $i = 1, \ldots, k$, tenemos que $V \subseteq W_1 + \ldots + W_k \subseteq V$.

Veamos que la expresión anterior de $v$ es la única posible como suma de vectores de los subespacios $W_1, \ldots, W_k$. Supongamos que $v = w_1 + \ldots + w_k$ con $w_i \in W_i = \mathrm{Im}(E_i)$ y veamos que $w_i = E_i v$ para cada $i = 1, \ldots, k$. Para cada $j$, como $w_j \in \mathrm{Im}(E_j)$ y $E_j$ es una proyección, tenemos que $E_j w_j = w_j$. Entonces, fijado un $i = 1, \ldots, k$,

$$\begin{aligned} E_i v &= E_i(w_1 + \ldots + w_k) = E_i w_1 + \ldots + E_i w_{i-1} + E_i w_i + E_i w_{i+1} + \ldots + E_i w_k \\ &= E_i E_1 w_1 + \ldots + E_i E_{i-1} w_{i-1} + E_i E_i w_i + E_i E_{i+1} w_{i+1} + \ldots + E_i E_k w_k \\ &= E_i^2 w_i = E_i w_i = w_i. \end{aligned}$$

Por lo tanto, para cada $v \in V$ existe una única representación posible como suma de vectores de los subespacios $W_1, \ldots, W_k$, es decir, $V = W_1 \dot{+} \ldots \dot{+} W_k$. □

## 4.4. Sumas directas de subespacios invariantes

En esta sección desarrollamos otra herramienta que nos ayudará a comprender la estructura de los operadores lineales. Supongamos que $T$ es un operador lineal sobre un $\mathbb{K}$-espacio vectorial $V$. Si tenemos una descomposición en suma directa

$$V = W_1 \dot{+} \ldots \dot{+} W_k,$$

donde cada $W_j$ es un subespacio propio de $V$, entonces para entender el comportamiento de $T$ nos alcanza con entender el comportamiento de cada una de las restricciones $T|_{W_j}$. Recordemos que $T|_{W_j}$, la **restricción de $T$ al subespacio $W_j$**, es la transformación lineal $T|_{W_j} : W_j \to V$ definida por $T|_{W_j} w = Tw$ para $w \in W_j$.

Tratar con las restricciones $T|_{W_j}$ debería ser más fácil que con el operador $T$, porque los subespacios $W_j$ son de menor dimensión. Sin embargo, tenemos un problema: las restricciones son transformaciones y no operadores lineales (su codominio es siempre el espacio $V$). Esto nos induce a considerar sólo descomposiciones en suma directa como la de arriba, pero para las cuales $T$ mapee cada subespacio $W_j$ sobre sí mismo.

La noción de un subespacio al cual $T$ mapea sobre sí mismo es lo suficientemente importante para merecerse un nombre.

**Definición 4.27.** Dado un $\mathbb{K}$-espacio vectorial $V$, sean $W$ un subespacio de $V$ y $T \in L(V)$. Diremos que $W$ es **invariante por $T$ (ó $T$-invariante)** si

$$T(W) \subseteq W,$$

es decir, si la imagen de $W$ por $T$ es un subespacio de $W$.

En otras palabras, $W$ es invariante por $T$ si $T|_W$ es un operador lineal sobre $W$.

**Ejemplo 4.28.** Dado un $\mathbb{K}$-espacio vectorial $V$ de dimensión finita, sea $T \in L(V)$. Entonces, los siguientes subespacios de $V$ son invariantes por $T$:

- $\{\vec{0}\}$;
- $V$;
- $N(T)$;
- $\mathrm{Im}(T)$.

Los dos primeros ejemplos no tienen gracia, la prueba de que los dos últimos son $T$-invariantes queda como ejercicio.

**Ejemplo 4.29.** Sea $D \in L(\mathbb{K}[x])$ el operador derivada, definido por $Dp = p'$.

Notemos que $\mathbb{K}^{(4)}[x]$ es un subespacio $D$-invariante de $\mathbb{K}[x]$, ya que la derivada de un polinomio de grado menor o igual que 4 es otro polinomio de grado menor o igual que 4.

¿Será cierto que $\mathbb{K}^{(n)}[x]$ es un subespacio $D$-invariante de $\mathbb{K}[x]$ para cualquier $n \in \mathbb{N}$?

**Ejercicio 4.30.** Dado un $\mathbb{K}$-espacio vectorial $V$, sea $T \in L(V)$. Probar que, si $\lambda \in \mathbb{K}$ es un autovalor de $T$, entonces el autoespacio $E(\lambda)$ asociado a $\lambda$ es un subespacio $T$-invariante.

El problema más famoso aún no resuelto del Análisis Funcional (primo hermano del Álgebra lineal que se dedica a los espacios vectoriales de dimensión infinita) se conoce como el **problema del subespacio invariante**: si $V$ es un espacio de Hilbert ¿es cierto que para todo operador lineal (y continuo) $T$ sobre $V$, siempre existe algún subespacio (cerrado) $W$ de $V$ que es $T$-invariante y no trivial? (es decir, $T(W) \subseteq W$, $W \neq V$ y $W \neq \{\vec{0}\}$). Su enunciado es bastante simple, pero involucra varias nociones topológicas (espacio de Hilbert, operador continuo, subespacio cerrado) que exceden el nivel de este curso.

¿Qué ocurre con el problema del subespacio invariante en $\mathbb{K}$-espacios vectoriales de dimensión finita? ¿todo operador $T \in L(V)$ tiene un subespacio invariante no trivial?

Más adelante veremos que la respuesta es afirmativa si $\mathbb{K} = \mathbb{C}$ y $\dim V > 1$, y también si $\mathbb{K} = \mathbb{R}$ y $\dim V > 2$.

Volvamos ahora a las descomposiciones en suma directa de subespacios invariantes, y tratemos de justificar porqué estás descomposiciones nos permiten escribir al operador $T$ de una forma más sencilla. Dado un $\mathbb{K}$-espacio vectorial $V$ de dimensión finita, sea $T \in L(V)$. Supongamos que $W_1$ y $W_2$ son dos subespacios propios de $V$, que resultan $T$-invariantes, y tales que

$$V = W_1 \dot{+} W_2.$$

Consideremos los operadores $T_1 := T|_{W_1} \in L(W_1)$ y $T_2 := T|_{W_2} \in L(W_2)$ y veamos que el efecto de $T$ está determinado por el de $T_1$ y el de $T_2$. Para cada $v \in V$ existen únicos $w_1 \in W_1$ y $w_2 \in W_2$ tales que $v = w_1 + w_2$. Luego,

$$Tv = Tw_1 + Tw_2 = T|_{W_1}\, w_1 + T|_{W_2}\, w_2 = T_1 w_1 + T_2 w_2.$$

Ahora, si $\mathcal{B}_1 = \{b_1, \ldots, b_m\}$ es una base de $W_1$ y $\mathcal{B}_2 = \{b_{m+1}, \ldots, b_n\}$ es una base $W_2$, tenemos que $\mathcal{B} = \{\mathcal{B}_1, \mathcal{B}_2\} = \{b_1, \ldots, b_n\}$ es una base de $V$ (ver Proposición 4.14). Calculemos la matriz de $T$ con respecto a la base $\mathcal{B}$:

- si $i = 1, \ldots, m$, entonces $Tb_i \in W_1$ y

  $$[Tb_i]_{\mathcal{B}} = (a_{1i}, a_{2i}, \ldots, a_{mi}, \overbrace{0, \ldots, 0}^{n-m \text{ veces}})^\top,$$

  donde $(a_{1i}, a_{2i}, \ldots, a_{mi})^\top = [Tb_i]_{\mathcal{B}_1}$;

- en cambio, si $j = m+1, \ldots, n$ entonces $T_1 b_j \in W_2$ y

$$[Tb_j]_{\mathcal{B}} = (\overbrace{0, \ldots, 0}^{m \text{ veces}}, a'_{1j}, a'_{2j}, \ldots, a'_{n-m,j})^\top,$$

  donde $(a'_{1j}, a'_{2j}, \ldots, a'_{n-m,j})^\top = [T_2 b_j]_{\mathcal{B}_2}$.

Recopilando toda esta información, es fácil convencerse de que

$$[T]_{\mathcal{B}} = \begin{pmatrix} [T_1]_{\mathcal{B}_1} & 0 \\ 0 & [T_2]_{\mathcal{B}_2} \end{pmatrix}.$$

Un razonamiento análogo es válido para una descomposición en suma directa con $k$ sumandos invariantes por $T$.

Si $V = W_1 \dotplus \ldots \dotplus W_k$, ¿cómo podemos asegurarnos de que cada $W_i$ es invariante por $T$?

**Teorema 4.31.** *Dado un $\mathbb{K}$-espacio vectorial $V$, sean $W_1, \ldots, W_k$ subespacios tales que*

$$V = W_1 \dotplus \ldots \dotplus W_k.$$

*Para cada $i = 1, \ldots, k$, sea $E_i$ la proyección sobre $W_i$ construída en el Teorema 4.26.*

*Entonces, cada $W_i$ es $T$-invariante si y sólo si*

$$TE_i = E_i T, \quad i = 1, \ldots, k.$$

*Demostración.* Supongamos primero que cada subespacio $W_i$ es $T$-invariante. Como $E_1 + \ldots + E_k = I$, a cada $v \in V$ podemos escribirlo como $v = E_1 v + \ldots + E_k v$. Luego,

$$Tv = TE_1 v + \ldots + TE_k v$$

Además, como $TE_i v \in W_i$, sabemos que $E_i T E_i v = TE_i v$. Entonces,

$$E_j T E_i v = E_j E_i T E_i v = \vec{0} \quad \text{si } j \neq i$$

y resulta que

$$\begin{aligned} E_j\, Tv &= E_j TE_1 v + \ldots + E_j TE_{j-1} v + E_j TE_j v + E_j TE_{j+1} v + \ldots + E_j TE_k v \\ &= \vec{0} + \ldots + \vec{0} + E_j TE_j v + \vec{0} + \ldots + \vec{0} \\ &= E_j TE_j v = TE_j v. \end{aligned}$$

Como $v \in V$ era un vector arbitrario, hemos probado que $T$ conmuta con $E_j$.

Recíprocamente, supongamos que $TE_i = E_i T$ para todo $i = 1, \ldots, k$. Fijado un $i = 1, \ldots, k$, consideremos un vector $w \in W_i$. Entonces,

$$Tw = TE_i w = E_i Tw \in W_i$$

Por lo tanto, cada $W_i$ es $T$-invariante. □

Veamos ahora como descomponer a un operador lineal diagonalizable como una combinación lineal de proyecciones que conmutan entre sí. Esta será nuestra primera versión del **Teorema espectral**.

**Teorema 4.32.** *Dado un $\mathbb{K}$-espacio vectorial $V$ de dimensión finita, sea $T \in L(V)$. Si $T$ es diagonalizable y $\lambda_1, \ldots, \lambda_k \in \mathbb{K}$ son los distintos autovalores de $T$, entonces existen proyecciones $E_1, \ldots, E_k \in L(V)$ tales que*

$$T = \sum_{i=1}^{k} \lambda_i \ E_i, \qquad \sum_{i=1}^{k} E_i = I, \qquad E_iE_j = 0 \ \text{ si } i \neq j. \tag{4.9}$$

*Además,* $\mathrm{Im}(E_i) = E(\lambda_i)$ *es el autoespacio asociado a $\lambda_i$.*

*Recíprocamente, si existen $\lambda_1, \ldots, \lambda_k \in \mathbb{K}$ y $E_1, \ldots, E_k \in L(V)$ (no nulos) que satisfacen* (4.9)*, entonces $T$ es diagonalizable, los $\lambda_i$ son los autovalores de $T$ y las $E_i$ son proyecciones sobre los autoespacios $E(\lambda_i)$.*

*Demostración.* Supongamos que $T$ es diagonalizable y definamos $W_i := E(\lambda_i)$ para cada autovalor $\lambda_i$, $i = 1, \ldots, k$. Como la yuxtaposición de bases de los subespacios $W_1, \ldots, W_k$ forma una base de $V$, la Proposición 4.14 garantiza que

$$V = W_1 \dot{+} \ldots \dot{+} W_k.$$

Para cada $j = 1, \ldots, k$ sea $\mathcal{N}_j := W_1 \dot{+} \ldots \dot{+} W_{j-1} \dot{+} W_{j+1} \dot{+} \ldots \dot{+} W_k$ y consideremos la proyección

$$E_j = P_{W_j//\mathcal{N}_j} \in L(V).$$

Es inmediato que $\sum_{i=1}^{k} E_i = I$ y que $E_iE_j = 0$ si $i \neq j$. Además, para cada $v \in V$,

$$Tv = T\left(\sum_{i=1}^{k} E_i\right) v = \sum_{i=1}^{k} TE_iv = \sum_{i=1}^{k} \lambda_i E_i v,$$

ya que $E_iv \in W_i$ es un autovector de $T$ asociado al autovalor $\lambda_i$ (o el vector nulo). Por lo tanto, $T = \sum_{i=1}^{k} \lambda_i E_i$.

Recíprocamente, supongamos que $\lambda_1, \ldots, \lambda_k \in \mathbb{K}$ y $E_1, \ldots, E_k \in L(V)$ son operadores no nulos que satisfacen (4.9). Veamos primero que los $E_i$ son proyecciones:

$$E_i = E_iI = E_i\left(\sum_{j=1}^{k} E_j\right) = \sum_{j=1}^{k} E_iE_j = E_i^2.$$

Por otra parte,

$$TE_i = \left(\sum_{j=1}^{k} \lambda_j E_j\right) E_i = \sum_{j=1}^{k} \lambda_j E_j E_i = \lambda_i E_i,$$

es decir, $(T - \lambda_i I)E_i = 0$. Esto implica que $\mathrm{Im}(E_i) \subseteq N(T - \lambda_i I)$. Entonces, $\lambda_i$ es un autovalor de $T$ y además $\mathrm{Im}(E_i) \subseteq E(\lambda_i)$.

Veamos que éstos son los únicos autovalores de $T$. Si $\lambda \in \mathbb{K}$,

$$T - \lambda I = \sum_{j=1}^{k} \lambda_j E_j - \lambda\left(\sum_{j=1}^{k} E_j\right) = \sum_{j=1}^{k} (\lambda_j - \lambda) E_j.$$

Luego, dado $v \in V$, $v \neq \vec{0}$, $(T - \lambda I)v = \vec{0}$ si y sólo si $(\lambda_j - \lambda)E_jv = \vec{0}$ para todo $j = 1, \ldots, k$. Pero si $v \neq \vec{0}$, existe un $i = 1, \ldots, k$ tal que $E_iv \neq \vec{0}$, lo que implica que $\lambda_i - \lambda = 0$, es decir, $\lambda = \lambda_i$.

Finalmente, como $\sum_{i=1}^{k} E_i = I$ y $\text{Im}(E_i) \subseteq E(\lambda_i)$, resulta que existe una base de $V$ formada por autovectores de $T$, i.e. $T$ es diagonalizable.

Sólo nos falta probar que $E(\lambda_i) \subseteq \text{Im}(E_i)$. Pero si $v$ es un autovector asociado a $\lambda_i$, tenemos

$$\vec{0} = (T - \lambda_i I)v = \sum_{j=1}^{k} (\lambda_j - \lambda_i) E_j v.$$

Entonces, como las imágenes de las proyecciones son subespacios independientes, tenemos que $(\lambda_j - \lambda_i)E_j v = \vec{0}$ para todo $j = 1, \ldots, k$. Esto implica que $E_j v = \vec{0}$ si $j \neq i$, y en consecuencia $E_i v = v$, o sea, $v \in \text{Im}(E_i)$. □

Una de las razones por la cual la teoría de operadores es más rica que la teoría de transformaciones lineales en general es que se pueden calcular las potencias de un operador, y en consecuencia también se puede evaluar a un polinomio en un operador lineal.

**Definición 4.33.** Dado un operador no nulo $T \in L(V)$, usaremos la convención de que

$$T^0 := I.$$

Luego, para cualquier $m \in \mathbb{N}$ la $m$-ésima potencia de $T$ se define recursivamente como

$$T^m := T \cdot T^{m-1}.$$

Además, si $T$ es inversible, definimos

$$T^{-m} := (T^{-1})^m.$$

**Ejercicio 4.34.** Si $T \in L(V)$, mostrar que

$$T^m \cdot T^n = T^{n+m} \quad \text{y} \quad (T^m)^n = T^{mn}, \tag{4.10}$$

donde $m, n \in \mathbb{N}$. Además, suponiendo que $T$ es inversible, verificar que $T^{-m}$ es el inverso de $T^m$ para todo $m \in \mathbb{N}$ y que (4.10) vale para todo $m, n \in \mathbb{Z}$.

**Definición 4.35.** Dado un $\mathbb{K}$-espacio vectorial $V$, sea $T \in L(V)$, $T \neq 0$. Si $p \in \mathbb{K}[x]$ es el polinomio dado por $p(x) = a_0 + a_1 x + \ldots + a_m x^m$, definimos el operador $p(T)$ como

$$p(T) := a_0 I + a_1 T + \ldots + a_m T^m.$$

Si $T \in L(V)$ es diagonalizable y $T = \sum_{i=1}^{k} \lambda_i E_i$ es la descomposición dada por el Teorema 4.32, para cada polinomio $p \in \mathbb{K}[x]$ se tiene que

$$p(T) = \sum_{i=1}^{k} p(\lambda_i) E_i \qquad \text{(verificarlo!)}$$

Luego, el operador $p(T)$ es el operador nulo si y sólo si $p(\lambda_i) = 0$ para todo $i = 1, \ldots, k$, o equivalentemente,

$$x - \lambda_i \,|\, p(x) \qquad \text{para todo } i = 1, \ldots, k.$$

Por lo tanto, $p(T) = 0$ si y sólo si $(x - \lambda_1) \ldots (x - \lambda_k) \,|\, p(x)$.

En la próxima sección estudiaremos al conjunto de polinomios que anula a un operador dado, y veremos que lo anterior no es una casualidad ni una particularidad de los operadores diagonalizables.

## 4.5. El ideal de polinomios anuladores

Sea $V$ un $\mathbb{K}$-espacio vectorial con dimensión finita. Dado $T \in L(V)$ no nulo, consideremos

$$\mathcal{A}(T) = \{p \in \mathbb{K}[x] : \; p(T) = 0\}, \tag{4.11}$$

el conjunto de los polinomios que anulan a $T$.

En primer lugar, veamos que $\mathcal{A}(T) \neq \varnothing$. Para esto, supongamos que $\dim V = n$ y notemos que las primeras $n^2 + 1$ potencias de $T$

$$I, T, T^2, \ldots, T^{n^2-1}, T^{n^2},$$

son $n^2 + 1$ vectores en el $\mathbb{K}$-espacio vectorial $L(V)$, mientras que $\dim L(V) = n^2$. Entonces, el conjunto $\{I, T, T^2, \ldots, T^{n^2-1}, T^{n^2}\}$ es linealmente dependiente, es decir, existen $c_0, c_1, \ldots, c_{n^2} \in \mathbb{K}$ tales que

$$c_0 I + c_1 T + \ldots + c_{n^2} T^{n^2} = 0.$$

Por lo tanto, el polinomio $q(x) = c_0 + c_1 x + \ldots + c_{n^2} x^{n^2}$ pertenece a $\mathcal{A}(T)$.

También es fácil ver que $\mathcal{A}(T)$ es un $\mathbb{K}$-espacio vectorial, porque si $p, q \in \mathcal{A}(T)$ y $\alpha \in \mathbb{K}$ entonces,

$$(\alpha p + q)(T) = \alpha\, p(T) + q(T) = \alpha\, 0 + 0 = 0,$$

es decir, $\alpha p + q \in \mathcal{A}(T)$.

Maás aún, $\mathcal{A}(T)$ es un ideal de $\mathbb{K}[x]$: si $p \in \mathcal{A}(T)$ y $q \in \mathbb{K}[x]$ entonces

$$(p \cdot q)(T) = p(T) \cdot q(T) = 0 \cdot q(T) = 0.$$

Como vimos en el Teorema 3.17, todo ideal de $\mathbb{K}[x]$ es un ideal principal, es decir existe un único polinomio mónico $m \in \mathbb{K}[x]$ tal que $\mathcal{A}(T) = m \cdot \mathbb{K}[x]$.

**Definición 4.36.** Dado un $\mathbb{K}$-espacio vectorial $V$ de dimensión finita, sea $T \in L(V)$. El **polinomio minimal de** $T$ es el (único) polinomio mónico que genera al ideal $\mathcal{A}(T)$. Lo anotaremos $m_T$.

El polinomio minimal de $T$ es el único polinomio en $\mathbb{K}[x]$ que satisface las siguientes condiciones:

i) $m_T$ es mónico;

ii) $m_T(T) = 0$;

iii) si $p \in \mathbb{K}[x]$ tiene $\mathrm{gr}(p) < \mathrm{gr}(m_T)$, entonces $p(T) \neq 0$.

**Ejemplo 4.37.** Sea $S : \mathbb{R}^3 \to \mathbb{R}^3$ dado por $S(x, y, z) = (y, x, z)$, la simetría con respecto al plano $x = y$. En el Ejemplo 4.8 vimos que los autovalores de $S$ son $\lambda_1 = 1$ y $\lambda_2 = -1$. Tratemos de calcular el polinomio minimal de $S$ utilizando la descripción de más arriba:

- Si $\alpha \in \mathbb{K}$ y $\alpha$ no es uno de los autovalores, entonces $x - \alpha$ no anula a $S$ porque $S - \alpha I$ es inversible.

- Veamos que $S - I$ y $S + I$ tampoco son iguales al operador nulo (i.e. $x - 1$ y $x + 1$ no anulan a $T$).

$$(S - I)(x, y, z) = (y, x, z) - (x, y, z) = (y - x, x - y, 0) \quad \Longrightarrow \quad S - I \neq 0,$$
$$(S + I)(x, y, z) = (y, x, z) + (x, y, z) = (y + x, x + y, 2x) \quad \Longrightarrow \quad S + I \neq 0.$$

- Finalmente, veamos que $x^2 - 1 = (x-1)(x+1)$ anula al operador $S$:
$$S^2(x,y,z) = S(y,x,z) = (x,y,z) = I(x,y,z),$$
con lo cual $(S^2 - I)(x,y,z) = (0,0,0)$ para todo $(x,y,z) \in \mathbb{R}^3$, es decir, $S^2 - I = 0$.

Por lo tanto, $m_S(x) = x^2 - 1 = (x-1)(x+1)$.

Si $A \in \mathbb{K}^{n\times n}$ también podemos considerar el ideal de polinomios que anulan a $A$:
$$\mathcal{A}(A) = \{p \in \mathbb{K}[x] : \ p(A) = 0\},$$
el cual admite un único generador mónico al que llamaremos el **polinomio minimal de** $A$, y lo anotaremos $m_A$.

**Ejercicio 4.38.** Si $A, B \in \mathbb{K}^{n\times n}$ son matrices semejantes, probar que $m_A = m_B$.

Si $A$ es la representación matricial de un operador $T \in L(V)$ sobre cierto $\mathbb{K}$-espacio vectorial $V$, veamos que relación hay entre los polinomios minimales de $A$ y $T$. Supongamos que $\mathcal{B}$ es una base de $V$ tal que $A = [T]_{\mathcal{B}}$. Luego, dado un polinomio arbitrario $p \in \mathbb{K}[x]$, es fácil ver que
$$[p(T)]_{\mathcal{B}} = p([T]_{\mathcal{B}}) = p(A).$$
De esto se desprende que $\mathcal{A}(A) = \mathcal{A}(T)$ y, por lo tanto, $m_A = m_T$.

Por lo tanto, para calcular el polinomio minimal de un operador lineal, lo que haremos generalmente es calcular el polinomio minimal de cualquiera de sus representaciones matriciales.

A continuación queremos calcular cuales son las raíces del polinomio minimal. Para ello necesitaremos del siguiente lema.

**Lema 4.39.** *Sea $\lambda \in \mathbb{K}$ un autovalor de un operador lineal $T \in L(V)$. Supongamos que $v \in V$ cumple $Tv = \lambda v$. Si $p \in \mathbb{K}[x]$ es un polinomio arbitrario, entonces*
$$p(T)v = p(\lambda)v.$$

*Demostración.* Ejercicio. □

**Teorema 4.40.** *Dado un $\mathbb{K}$-espacio vectorial $V$ de dimensión finita, sea $T \in L(V)$. Entonces, el polinomio característico y el polinomio minimal de $T$ tienen las mismas raíces (salvo multiplicidades).*

*Demostración.* Si $m_T$ es el polinomio minimal de $T$ y $\lambda \in \mathbb{K}$, queremos ver que $m_T(\lambda) = 0$ si y sólo si $\lambda$ es un autovalor de $T$.

Supongamos primero que $\lambda$ es una raíz de $m_T$. Luego, existe $q \in \mathbb{K}[x]$ tal que
$$m_T(x) = (x - \lambda)q(x).$$
Como $\mathrm{gr}(q) < \mathrm{gr}(m_T)$ sabemos que $q(T) \neq 0$, es decir, existe algún $v \in V$ tal que $q(T)v \neq \vec{0}$. Si definimos $w := q(T)v$, resulta que $w \neq \vec{0}$ y
$$(T - \lambda I)w = (T - \lambda I)q(T)v = m_T(T)v = 0\,v = \vec{0},$$
es decir, $\lambda$ es un autovalor de $T$.

Recíprocamente, si $\lambda$ es un autovalor de $T$ consideremos un autovector $v \in V$ asociado a $\lambda$: $v \neq \vec{0}$ y $Tv = \lambda v$. Aplicando el lema anterior, tenemos que
$$m_T(\lambda)v = m_T(T)v = 0\,v = \vec{0},$$
i.e. $\lambda$ es una raíz del polinomio minimal. □

Finalmente podemos caracterizar a los operadores diagonalizables en función de su polinomio minimal. Fíjense que en la demostración utilizaremos los polinomios interpoladores de Lagrange que definimos en la Sección 3.1.

**Teorema 4.41.** *Dado un $\mathbb{K}$-espacio vectorial $V$ de dimensión finita, sea $T \in L(V)$. Si $\lambda_1, \ldots, \lambda_k \in \mathbb{K}$ son los distintos autovalores de $T$, entonces*

$$T \text{ es diagonalizable} \qquad \Longleftrightarrow \qquad m_T(x) = (x-\lambda_1)(x-\lambda_2)\ldots(x-\lambda_k).$$

*Demostración.* Supongamos que $T$ es diagonalizable. Por el Teorema 4.32, existen proyecciones $E_1, \ldots, E_k \in L(V)$ tales que $T = \sum_{i=1}^k \lambda_i E_i$. Como $p(T) = \sum_{i=1}^k p(\lambda_i)E_i$ para cualquier $p \in \mathbb{K}[x]$, resulta que

$$p(T) = 0 \qquad \Longleftrightarrow \qquad (x-\lambda_1)(x-\lambda_2)\ldots(x-\lambda_k) \mid p(x),$$

de lo cual se deduce que $m_T(x) = (x-\lambda_1)(x-\lambda_2)\ldots(x-\lambda_k)$.

Recíprocamente, supongamos que $m_T(x) = (x-\lambda_1)(x-\lambda_2)\ldots(x-\lambda_k)$. Consideremos los polinomios de Lagrange

$$p_j(x) = \prod_{i \neq j} \frac{x-\lambda_i}{\lambda_j - \lambda_i}, \quad j = 1, \ldots, k.$$

Recordemos que $p_j(\lambda_i) = \delta_{ij}$, y si $q \in \mathbb{K}^{(k-1)}[x]$ entonces

$$q = q(\lambda_1)p_1 + q(\lambda_2)p_2 + \ldots + q(\lambda_k)p_k.$$

En particular,

$$1 = p_1 + p_2 + \ldots + p_k, \quad \text{y} \quad x = \lambda_1 p_1 + \lambda_2 p_2 + \ldots + \lambda_k p_k.$$

Ahora, si definimos los operadores $E_j := p_j(T)$ para $j = 1, \ldots, k$, resulta que

$$I = E_1 + E_2 + \ldots + E_k \quad \text{y} \quad T = \lambda_1 E_1 + \lambda_2 E_2 + \ldots + \lambda_k E_k.$$

Además, si $i \neq j$ es fácil ver que $m_T \,|\, p_i p_j$, con lo cual

$$0 = (p_i \cdot p_j)(T) = p_i(T)p_j(T) = E_i E_j.$$

Finalmente, notemos que cada $E_i \neq 0$ porque $\mathrm{gr}(p_i) < \mathrm{gr}(m_T)$. Por lo tanto, aplicando el Teorema 4.32, resulta que $T$ es diagonalizable. □

Para terminar con esta sección, enunciaremos el Teorema de Cayley–Hamilton. La demostración del mismo pueden encontrarla, por ejemplo, en [4].

**Teorema 4.42** (Cayley-Hamilton)**.** *Dado un $\mathbb{K}$-espacio vectorial $V$ de dimensión finita, sea $T \in L(V)$. Si $p_T$ es el polinomio característico de $T$ entonces $p_T(T) = 0$, es decir, el polinomio minimal $m_T$ divide al polinomio característico $p_T$.*

Este teorema nos permite reducir el número de candidatos a polinomio minimal de $T$, ya que si $p_T(x) = (x-\lambda_1)^{d_1}(x-\lambda_2)^{d_2}\ldots(x-\lambda_k)^{d_k}$ entonces

$$m_T(x) = (x-\lambda_1)^{r_1}(x-\lambda_2)^{r_2}\ldots(x-\lambda_k)^{r_k}, \qquad \text{con } 1 \leq r_i \leq d_i.$$

Por otra parte, en algunos casos también puede servir como herramienta para calcular el polinomio característico de un operador. Veamos cómo en un ejemplo.

**Ejemplo 4.43.** Sea $A \in \mathbb{Q}^{4\times 4}$ dada por

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

Es fácil ver que

$$A^2 = \begin{pmatrix} 2 & 0 & 2 & 0 \\ 0 & 2 & 0 & 2 \\ 2 & 0 & 2 & 0 \\ 0 & 2 & 0 & 2 \end{pmatrix} \quad \text{y} \quad A^3 = \begin{pmatrix} 0 & 4 & 0 & 4 \\ 4 & 0 & 4 & 0 \\ 0 & 4 & 0 & 4 \\ 4 & 0 & 4 & 0 \end{pmatrix}.$$

En consecuencia, $A^3 = 4A$. Luego, si $p(x) = x^3 - 4x = x(x-2)(x+2)$, vemos que $p \in \mathcal{A}(A)$. Entonces, $\mathrm{gr}(m_A) \leq \mathrm{gr}(p) = 3$.

Como $A$ no es un múltiplo de la matriz identidad, $\mathrm{gr}(m_A) > 1$. Con lo cual, los posibles candidatos a polinomio minimal son:

$$\begin{aligned} m_1(x) &= x^2 - 2x, \\ m_2(x) &= x^2 + 2x, \\ m_3(x) &= x^2 - 4, \quad \text{y} \\ m_4(x) &= p(x). \end{aligned}$$

Es fácil ver que $A^2 \neq 2A$, $A^2 \neq -2A$ y $A^2 \neq 4I$, de donde deducimos que

$$m_T(x) = p(x) = x^3 - 4x.$$

Por lo tanto, los autovalores de $A$ son $\lambda_1 = 0$, $\lambda_2 = 2$ y $\lambda_3 = -2$, y $A$ es diagonalizable.

Para calcular el polinomio característico de $A$, notemos que $A$ es una matriz de rango 2. Entonces,

$$\dim N(A) = \dim \mathbb{Q}^4 - \dim \mathrm{Im}(A) = 4 - 2 = 2.$$

Por lo tanto, $p_A(x) = x^2(x-2)(x+2)$.

## 4.6. Teorema de la descomposición prima

Dado un operador lineal $T$ actuando en un $\mathbb{K}$-espacio vectorial $V$ de dimensión finita, hasta el momento hemos intentado diagonalizarlo. Como ya vimos anteriormente, esto no siempre es posible, y hay dos motivos por los cuales $T$ puede no ser diagonalizable:

1. El operador no tiene suficientes autovalores, es decir, el polinomio minimal de $T$ no tiene todas sus raíces en el cuerpo $\mathbb{K}$. Esto ocurre si $\mathbb{K}$ no es algebraicamente cerrado.

2. La suma directa de los autoespacios no es lo suficientemente grande, es decir, no genera todo el $\mathbb{K}$-espacio vectorial $V$.

El segundo motivo sí es una situación particular que depende del operador, y tenemos que buscar una forma alternativa de representarlo como una suma de operadores "elementales". Para esto debemos reemplazar a los autoespacios por otros subespacios $T$-invariantes que nos permitan generar a todo el $\mathbb{K}$-espacio vectorial $V$. Veamos cómo hacerlo en un ejemplo.

**Ejemplo 4.44.** Dado un cuerpo arbitrario $\mathbb{K}$, sea $T \in L(\mathbb{K}^3)$ el operador representado en la base canónica por

$$A = [T]_{\mathcal{E}} = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix} = \left(\begin{array}{cc|c} 2 & 0 & 0 \\ 1 & 2 & 0 \\ \hline 0 & 0 & -1 \end{array}\right).$$

Es fácil convencerse de que $T$ no es diagonalizable, ya que el polinomio minimal de $T$ es

$$m_T(x) = p_T(x) = (x+1)(x-2)^2$$

y en este caso coincide con el característico. Sin embargo, es evidente que la matriz $A$ es una matriz diagonal por bloques, de acuerdo a la descomposición

$$\mathbb{K}^3 = \overline{\{(1,0,0),(0,1,0)\}} \dot{+} \overline{\{(0,0,1)\}}. \tag{4.12}$$

Veamos como podemos interpretar a estos subespacios en términos del operador $T$ y sus autovalores. Por un lado,

$$E(-1) = \overline{\{(0,0,1)\}} \quad \text{y} \quad E(2) = \overline{\{(0,1,0)\}}.$$

Por otra parte,

$$(T-2I)^2 = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -3 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 9 \end{pmatrix}$$

Entonces, si en lugar de calcular $E(2) = N(T-2I)$ calculamos $N((T-2I)^2)$ tenemos:

$$N((T-2I)^2) = \overline{\{(1,0,0),(0,1,0)\}}.$$

Por lo tanto, la descomposición (4.12) puede describirse como

$$\mathbb{K}^3 = N((T-2I)^2) \dot{+} N(T-(-1)I).$$

En el ejemplo anterior vimos que reemplazando a $N(T-2I)$ por $N((T-2I)^2)$ conseguíamos descomponer al espacio como suma de subespacios $T$-invariantes. Notemos que si $T \in L(V)$ y $\lambda \in \mathbb{K}$ entonces

$$N\big((T-\lambda I)^k\big) \subseteq N\big((T-\lambda I)^{k+1}\big) \quad \text{para todo } k \geq 1.$$

Esto dice que, para cada autovalor $\lambda$ de $T$, tenemos una familia creciente de subespacios $T$-invariantes. A continuación mostraremos que, si el polinomio minimal $m_T$ de $T$ se descompone como

$$m_T(x) = (x-\lambda_1)^{r_1}(x-\lambda_2)^{r_2}\ldots(x-\lambda_k)^{r_k}, \tag{4.13}$$

con $\lambda_1,\ldots,\lambda_k \in \mathbb{K}$ distintos entre sí, entonces

$$V = N\big((T-\lambda_1 I)^{r_1}\big) \dot{+} N\big((T-\lambda_2 I)^{r_2}\big) \dot{+} \ldots \dot{+} N\big((T-\lambda_k I)^{r_k}\big).$$

En realidad, probaremos algo más fuerte, porque en la factorización prima de $m_T$ admitiremos factores primos no lineales.

**Teorema 4.45 (de la descomposición prima).** *Dado un $\mathbb{K}$-espacio vectorial $V$ de dimensión finita, sea $T \in L(V)$. Si la factorización prima del polinomio minimal de $T$ es*

$$m_T = p_1^{r_1} p_2^{r_2} \dots p_k^{r_k},$$

*donde los polinomios $p_1, p_2, \dots, p_k$ son mónicos, irreducibles sobre $\mathbb{K}$ y distintos entre sí, consideremos los subespacios*

$$W_i = N\big(p_i(T)^{r_i}\big).$$

*Entonces,*

*i)* $V = W_1 \dot{+} W_2 \dot{+} \dots \dot{+} W_k$;

*ii)* *cada* $W_i$ *es* $T$*-invariante;*

*iii)* *si* $T_i = T|_{W_i}$ *entonces* $m_{T_i} = p_i^{r_i}$.

*Demostración.* La idea de la demostración consiste en encontrar $k$ proyecciones $E_1, \dots, E_k$ tales que $E_i w = w$ si $w \in W_i$, y $E_i w = \vec{0}$ si $w \in W_j$ con $j \neq i$. Además necesitamos que estas proyecciones conmuten con $T$, para que los subespacios $W_i$ sean $T$-invariantes (ver Teoremas 4.26 y 4.31). Por lo tanto, buscamos una familia de polinomios $q_1, q_2, \dots, q_k \in \mathbb{K}[x]$ tales que $E_i = q_i(T)$ para $i = 1, \dots, k$.

Para cada $i = 1, \dots, k$, sea $f_i := \frac{m_T}{p_i^{r_i}} = \prod_{j \neq i} p_j^{r_j}$. Por el Teorema de la factorización prima (Teorema 3.27), los polinomios $f_1, f_2, \dots, f_k$ son primos entre sí. Luego, existen $g_1, g_2, \dots, g_k \in \mathbb{K}[x]$ tales que

$$f_1 \cdot g_1 + f_2 \cdot g_2 + \dots + f_k \cdot g_k = 1.$$

Notemos además que, si $i \neq j$ entonces $m_T \mid f_i \cdot f_j$.

Veamos ahora que, si definimos $q_i := f_i \cdot g_i$, los polinomios $q_1, \dots, q_k$ cumplen las condiciones requeridas. Para cada $i = 1, \dots, k$, sea

$$E_i := q_i(T) = f_i(T) g_i(T).$$

Como $q_1 + q_2 + \dots + q_k = 1$ y $m_T \mid q_i \cdot q_j$ si $i \neq j$, es inmediato que

$$E_1 + E_2 + \dots + E_k = I \quad \text{y} \quad E_i E_j = 0 \ \text{ si } i \neq j.$$

Entonces, las $E_i$ son proyecciones y además

$$V = \mathrm{Im}(E_1) \dot{+} \mathrm{Im}(E_2) \dot{+} \dots \dot{+} \mathrm{Im}(E_k).$$

Veamos ahora que $\mathrm{Im}(E_i) = W_i$. Si $w \in \mathrm{Im}(E_i)$, sabemos que $E_i w = w$. Luego,

$$p_i(T)^{r_i} w = p_i(T)^{r_i} E_i w = p_i(T)^{r_i} f_i(T) g_i(T) w = m_T(T) g_i(T) w = \vec{0},$$

ya que $m_T$ anula a $T$. Por lo tanto, $w \in W_i$. Recíprocamente, si suponemos que $w \in W_i$, tenemos que $p_i(T)^{r_i} w = \vec{0}$. Esto implica que, si $j \neq i$, entonces $f_j(T) w = \vec{0}$ porque $p_i^{r_i} \mid f_j$. Luego, $E_j w = g_j(T) f_j(T) w = \vec{0}$ si $i \neq j$ y

$$w = E_1 w + E_2 w + \dots + E_k w = E_i w.$$

Por lo tanto, $w \in \mathrm{Im}(E_i)$. Esto completa la demostración del ítem *i)*.

El ítem *ii*) se verifica trivialmente, ya que las $E_i$ así construídas conmutan con $T$, ver el Teorema 4.31. Sólo falta probar el ítem *iii*). En primer lugar veamos que $p_i(T_i)^{r_i} = 0$: dado un vector arbitrario $v \in W_i$,

$$p_i(T_i)^{r_i} v = p_i(T)^{r_i} v = p_i(T)^{r_i} E_i v = p_i(T)^{r_i} f_i(T) g_i(T) v = m_T(T) g_i(T) v = \vec{0},$$

ya que $m_T$ anula a $T$. Entonces, el polinomio minimal $m_{T_i}$ de $T_i$ divide a $p_i^{r_i}$.

Veamos ahora que $p_i^{r_i} \mid m_{T_i}$. Para esto, veremos que

$$\text{si } g \in \mathbb{K}[x] \text{ es tal que } g(T_i) = 0, \text{ entonces } (g \cdot f_i)(T) = 0. \tag{4.14}$$

Esto dice que $m_T \mid g \cdot f_i$, de donde se deduce que $p_i^{r_i} \mid g$ (porque $m_T = p_i^{r_i} \cdot f_i$ y además $p_i^{r_i}$ y $f_i$ son coprimos). Como $p_i^{r_i}$ divide a todos los polinomios que anulan a $T_i$, en particular $p_i^{r_i} \mid m_{T_i}$. Finalmente, tenemos que $p_i^{r_i} \mid m_{T_i}$, $m_{T_i} \mid p_i^{r_i}$ y ambos son mónicos. Por lo tanto, $m_{T_i} = p_i^{r_i}$, completando la demostración.

Probemos entonces la afirmación (4.14). Si $g \in \mathbb{K}[x]$ es tal que $g(T_i) = 0$, dado un vector arbitrario $v \in V$,

$$(g \cdot f_i)(T)v = (g \cdot f_i)(T)(E_1 v + E_2 v + \ldots + E_k v) = (g \cdot f_i)(T) E_i v,$$

porque $f_i(T)E_j = f_i(T) f_j(T) g_j(T) = 0$ si $i \neq j$ (recordemos que $m_T \mid f_i \cdot f_j$). Luego,

$$(g \cdot f_i)(T)v = (g \cdot f_i)(T) E_i v = f_i(T) g(T) E_i v = f_i(T) g(T_i) E_i v = \vec{0},$$

ya que $E_i v \in W_i$ y $g$ anula a $T_i$. En conclusión, si $g \in \mathbb{K}[x]$ es tal que $g(T_i) = 0$ entonces $(g \cdot f_i)(T) = 0$, como queríamos probar. □

**Observación 4.46.** En la demostración anterior utilizamos un argumento bastante útil que vale la pena destacar:

$$\text{Si } W \text{ es } T\text{-invariante y } q \in \mathbb{K}[x], \text{ entonces } q(T|_W) = q(T)|_W.$$

Aunque resulta evidente, tómense unos minutos para probarlo.

**Corolario 4.47.** *Si $E_1, \ldots, E_k$ son las proyecciones determinadas por el Teorema de la descomposición prima aplicado a $T \in L(V)$, y $U \in L(V)$ es tal que $TU = UT$, entonces*

$$E_i U = U E_i, \quad i = 1, \ldots, k$$

*es decir, cada $W_i := \operatorname{Im}(E_i)$ es invariante por $U$.*

*Demostración.* Recordemos que, para cada $i = 1, \ldots, k$, $E_i = q_i(T)$ para cierto polinomio $q \in \mathbb{K}[x]$. Luego, si $TU = UT$ entonces

$$p(T)U = Up(T) \quad \text{para todo } p \in \mathbb{K}[x].$$

En particular, $E_i U = q_i(T) U = U q_i(T) = U E_i$ para cada $i = 1, \ldots, k$. □

Dado un operador $T \in L(V)$, supongamos que $V = W_1 \dotplus \ldots \dotplus W_k$ es la descomposición prima de $V$ inducida por $T$. Como cada subespacio $W_i$ es invariante por $T$, si elegimos una base $\mathcal{B}_i$ de cada $W_i$ y luego consideramos la base de $V$ dada por

$$\mathcal{B} = \{\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_k\},$$

la representación matricial de $T$ con respecto a esta base será una matriz diagonal por bloques de la forma:

$$[T]_{\mathcal{B}} = \begin{pmatrix} [T_1]_{\mathcal{B}_1} & 0 & \dots & 0 \\ 0 & [T_2]_{\mathcal{B}_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & [T_k]_{\mathcal{B}_k} \end{pmatrix},$$

donde cada $T_i = T|_{W_i}$ es la restricción de $T$ al subespacio $W_i$. Además, cada uno de los operadores $T_i$ tiene a lo sumo un autovalor.

De aquí en más, estudiaremos como descomponer a los subespacios $W_i$ en subespacios $T$-invariantes más pequeños, de forma tal que nos permitan obtener una representación matricial triangular inferior.

Supongamos ahora que el polinomio minimal de $T \in L(V)$ tiene todas sus raíces en $\mathbb{K}$:

$$m_T(x) = (x - \lambda_1)^{r_1}(x - \lambda_2)^{r_2} \dots (x - \lambda_k)^{r_k}.$$

Siguiendo con la notación del Teorema de la descomposición prima, las imágenes de las proyecciones $E_1, \dots, E_k$ están dados por

$$W_i = N\big((T - \lambda_i I)^{r_i}\big), \quad i = 1, \dots, k.$$

Consideremos ahora el operador

$$D := \lambda_1 E_1 + \dots + \lambda_k E_k. \tag{4.15}$$

Por el Teorema 4.32, $D$ es diagonalizable. Además, como $T = T(E_1 + \dots + E_k) = TE_1 + \dots + TE_k$, resulta que

$$N := T - D = (T - \lambda_1 I)E_1 + (T - \lambda_2 I)E_2 + \dots + (T - \lambda_k I)E_k. \tag{4.16}$$

Como las proyecciones $E_i$ se anulan mutuamentente, i.e. $E_iE_j = 0$ si $i \neq j$, el operador $N$ satisface

$$\begin{aligned} N^2 &= (T - \lambda_1 I)^2 E_1 + (T - \lambda_2 I)^2 E_2 + \dots + (T - \lambda_k I)^2 E_k, \\ N^3 &= (T - \lambda_1 I)^3 E_1 + (T - \lambda_2 I)^3 E_2 + \dots + (T - \lambda_k I)^3 E_k, \\ \vdots &= \qquad\qquad \vdots \\ N^r &= (T - \lambda_1 I)^r E_1 + (T - \lambda_2 I)^r E_2 + \dots + (T - \lambda_k I)^r E_k, \end{aligned}$$

para cualquier $r \in \mathbb{N}$.

En particular, si elegimos un $r \geq \text{máx}\{r_1, \dots, r_k\}$ tenemos que $N^r = 0$ porque $(T - \lambda_i I)^r E_i = 0$ para cada $i = 1, \dots, k$.

**Definición 4.48.** Un operador lineal $N$ se dice **nilpotente** si existe un $r \in \mathbb{N}$ tal que $N^r = 0$.

**Ejercicio 4.49.** Probar que, si $N \in L(V)$ es nilpotente, entonces su único autovalor es $\lambda = 0$.

**Teorema 4.50.** *Dado un $\mathbb{K}$-espacio vectorial $V$ de dimensión finita, sea $T \in L(V)$. Si el polinomio minimal de $T$ se factoriza como $m_T(x) = (x - \lambda_1)^{r_1}(x - \lambda_2)^{r_2} \dots (x - \lambda_k)^{r_k}$ con $\lambda_1, \dots, \lambda_k \in \mathbb{K}$, entonces existen (únicos) operadores $D, N \in L(V)$ tales que $D$ es diagonalizable, $N$ es nilpotente y satisfacen*

$$T = D + N \quad y \quad DN = ND.$$

*Además, $D$ y $N$ son polinomios en $T$.*

La existencia de $D$ y $N$ quedó probada más arriba (además del hecho que $D$ y $N$ resultan ser polinomios evaluados en $T$). Para probar la unicidad, necesitaremos del siguiente resultado auxiliar.

**Teorema 4.51.** *Dado un $\mathbb{K}$-espacio vectorial $V$ de dimensión finita, sean $U, T \in L(V)$ diagonalizables. Si $U$ y $T$ conmutan (i.e. $UT = TU$) entonces existe una base $\mathcal{B}$ de $V$ tal que $[U]_{\mathcal{B}}$ y $[T]_{\mathcal{B}}$ son matrices diagonales.*

*Demostración.* La demostración la haremos por inducción, sobre la dimensión de $V$.

Si $\dim V = 1$, ho hay nada que hacer. Como hipótesis inductiva, supongamos que el resultado vale para todo $\mathbb{K}$-espacio vectorial con dimensión menor a $n$.

Dado un $\mathbb{K}$-espacio vectorial $V$ con $\dim V = n$, sean $U, T \in L(V)$ diagonalizables. Si $T = \alpha I$ para algún $\alpha \in \mathbb{K}$ no hay nada que probar, porque $[T]_{\mathcal{B}} = \alpha I_n$ en cualquier base $\mathcal{B}$, en particular en aquellas que diagonalizan a $U$.

Si $T$ no es un múltiplo de la identidad, sean $\lambda_1, \ldots, \lambda_k$ los distintos autovalores de $T$ y consideremos los autoespacios $W_i = N(T - \lambda_i I)$, $i = 1, \ldots, k$. Fijado un índice $j = 1, \ldots, k$, el autoespacio $W_j$ es invariante por $U$ ya que

$$U(T - \lambda_j I) = (T - \lambda_j I)U.$$

Consideremos las restricciones de $U$ y $T$ a $W_j$: $U_j = U|_{W_j}$ y $T_j = T|_{W_j}$. Como $m_{T_i} \mid m_T$ y $m_{U_i} \mid m_U$, los operadores $T_j$ y $U_j$ también son diagonalizables.

Luego, aplicando la hipótesis inductiva al subespacio $W_j$ (el cual tiene $\dim W_j < n$), obtenemos una base $\mathcal{B}_j$ de $W_j$ tal que $[T_j]_{\mathcal{B}_j}$ y $[U_j]_{\mathcal{B}_j}$ son diagonales.

Entonces, $\mathcal{B} = \{\mathcal{B}_1, \ldots, \mathcal{B}_k\}$ es una base de $V$ tal que $[T]_{\mathcal{B}}$ y $[U]_{\mathcal{B}}$ son diagonales. □

Con ésto podemos probar la unicidad de los operadores del Teorema 4.50.

*Demostración del Teorema 4.50.* Supongamos que además de los operadores $D$ y $N$ construidos a partir del Teorema de la descomposición prima en (4.15) y (4.16), respectivamente, existen operadores $D', N' \in L(V)$ tales que $D'$ es diagonalizable, $N'$ es nilpotente, y satisfacen $T = D' + N'$ y $D'N' = N'D'$. Queremos ver que $D' = D$ y que $N' = N$.

Como $D'$ y $N'$ conmutan entre sí y $T = D' + N'$, tenemos que $D'$ y $N'$ conmutan también con $T$. Luego,

$$D'p(T) = p(T)D' \quad \text{y} \quad N'p(T) = p(T)N' \quad \text{para todo } p \in \mathbb{K}[x].$$

En particular, $D'D = DD'$ y $N'N = NN'$. Además, $D + N = T = D' + N'$ implica que $D - D' = N' - N$.

Como $D$ y $D'$ son diagonalizables y $D'D = DD'$, el teorema anterior garantiza que son simultáneamente diagonalizables. Luego, $D - D'$ también es diagonalizable (con respecto a la misma base que digonaliza simultáneamente a $D$ y $D'$).

Por otra parte, como $N'N = NN'$, para cada $r \in \mathbb{N}$ podemos utilizar el binomio de Newton para calcular

$$(N' - N)^r = \sum_{i=0}^{r} \begin{pmatrix} r \\ i \end{pmatrix} (N')^{r-i}(-N)^i.$$

Luego, como $N$ y $N'$ son nilpotentes, si $r$ es suficientemente grande resulta que $(N' - N)^r = 0$, es decir, $N' - N$ también es nilpotente.

Entonces, $S := D - D' = N' - N$ es nilpotente y diagonalizable. Por el Ejercicio 4.49, el único autovalor de este operador es $\lambda = 0$. Pero como $S$ es diagonalizable, el minimal debe tener raíces simples. Entonces, $m_S(x) = x$, o equivalentemente, $S = 0$.

Por lo tanto, $D' = D$ y $N' = N$. □

Como $\mathbb{C}$ es un cuerpo algebraicamente cerrado, tenemos el siguiente corolario.

**Corolario 4.52.** *Si $V$ es un $\mathbb{C}$-espacio vectorial de dimensión finita, entonces todo $T \in L(V)$ puede descomponerse (de manera única) como*

$$T = D + N,$$

*donde $D \in L(V)$ es diagonalizable, $N \in L(V)$ es nilpotente, y además $D$ y $N$ conmutan entre sí. Más aún, $D$ y $N$ pueden escribirse como polinomios en $T$.*

El Teorema 4.50 muestra también que el estudio de los operadores lineales en un $\mathbb{K}$-espacio vectorial (de dimensión finita) sobre un cuerpo algebraicamente cerrado $\mathbb{K}$, se reduce al estudio de los operadores nilpotentes. Si $\mathbb{K}$ no es algebraicamente cerrado habrá que encontrar además un sustituto para los autovalores y autovectores. En el próximo capítulo veremos como tratar estos dos problemas a la vez.

## 4.7. Operadores triangulables

Volvamos ahora sobre la condición que impusimos en la ecuación (4.13). A continuación veremos que el hecho de que polinomio minimal tenga todas sus raíces en el cuerpo $\mathbb{K}$ implica que $T$ admite una representación matricial **triangular inferior**, es decir, que existe una base $\mathcal{B} = \{b_1, \ldots, b_n\}$ de $V$ tal que

$$[T]_{\mathcal{B}} = \begin{pmatrix} a_{11} & 0 & 0 & \ldots & 0 \\ a_{21} & a_{22} & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ a_{n-1\,1} & a_{n-1\,2} & a_{n-1\,3} & \ddots & 0 \\ a_{n1} & a_{n2} & a_{n3} & \ldots & a_{nn} \end{pmatrix}.$$

Comencemos caracterizando cuando un operador admite una representación triangular inferior. En tal caso, diremos que $T$ es **triangulable**.

**Proposición 4.53.** *Dado un $\mathbb{K}$-espacio vectorial $V$ con $\dim V = n$, sea $\mathcal{B} = \{b_1, \ldots, b_n\}$ una base de $V$. Si $T \in L(V)$, son equivalentes:*

*i) la matriz de $T$ con respecto a la base $\mathcal{B}$ es triangular inferior;*

*ii) $Tb_j \in \overline{\{b_j, \ldots, b_n\}}$ para cada $j = 1, \ldots, n$;*

*iii) $\overline{\{b_j, \ldots, b_n\}}$ es un subespacio $T$-invariante para cada $j = 1, \ldots, n$.*

*Demostración.* La equivalencia entre *i*) y *ii*) es inmediata. Además, es obvio que *iii*) implica *ii*). Luego, para completar la demostración sólo falta probar que *ii*) implica *iii*).

Supongamos que $Tb_j \in \overline{\{b_j, \ldots, b_n\}}$ para cada $j = 1, \ldots, n$. Fijado $i = 1, \ldots, n$, notemos que no sólo $Tb_i \in \overline{\{b_i, \ldots, b_n\}}$ sino que $Tb_j \in \overline{\{b_i, \ldots, b_n\}}$ para todo $j = i, \ldots, n$. Entonces, si $v \in \overline{\{b_i, \ldots, b_n\}}$ tenemos que $Tv \in \overline{\{b_i, \ldots, b_n\}}$. Por lo tanto, $\overline{\{b_i, \ldots, b_n\}}$ es un subespacio $T$-invariante. $\square$

Ahora sí podemos mostrar que si el polinomio minimal de un operador $T$ tiene todas sus raíces en el cuerpo $\mathbb{K}$ entonces $T$ es triangulable.

**Teorema 4.54.** *Dado un $\mathbb{K}$-espacio vectorial $V$ de dimensión finita, sea $T \in L(V)$. Si el polinomio minimal de $T$ tiene todas sus raíces en el cuerpo $\mathbb{K}$ entonces $T$ es triangulable.*

*Demostración.* Lo probaremos por inducción sobre la dimensión de $V$. Es claro que el resultado vale si $\dim V = 1$ (de hecho, en este caso todo operador sobre $V$ es diagonalizable).

Ahora, supongamos que $\dim V = n > 1$ y que el resultado vale para todo $\mathbb{K}$-espacio vectorial con dimensión menor que $n$. Sea $T \in L(V)$ tal que todas las raíces de su polinomio minimal pertenezcan a $\mathbb{K}$. Si $T$ es el operador nulo, es obviamente triangulable. Si $T \neq 0$ su polinomio minimal tiene grado mayor o igual a 1. Sea $\lambda \in \mathbb{K}$ una de sus raíces. Por el Teorema 4.40, $\lambda$ es un autovalor de $T$. Entonces,

$$W := \operatorname{Im}(T - \lambda I)$$

es un subespacio de $V$ con $\dim W = m < n$, porque $\dim N(T - \lambda I) > 0$. Veamos que $W$ es un subespacio $T$-invariante: dado $w \in W$, tenemos que

$$Tw = (T - \lambda I)w + \lambda w \in W,$$

ya que tanto $\lambda w$ como $(T - \lambda I)w$ son vectores en $\operatorname{Im}(T - \lambda I) = W$. Luego, $T_0 := T|_W$ es un operador lineal sobre $W$ y además su polinomio minimal tiene todas sus raíces en $\mathbb{K}$ (esto último se debe a que $m_{T_0} \,|\, m_T$ pues $m_T(T_0) = 0$). Entonces, por hipótesis inductiva, existe una base $\mathcal{B}_0 = \{w_1, \ldots, w_m\}$ de $W$ tal que $[T]_{\mathcal{B}_0}$ es una matriz triangular inferior. Luego, si $\mathcal{B}_1 = \{v_{m+1}, \ldots, v_n\}$ es una base de $N(T - \lambda I)$, resulta que $\mathcal{B} := \{w_1, \ldots, w_m, v_{m+1}, \ldots, v_n\}$ es una base de $V$ y además

$$[T]_{\mathcal{B}} = \begin{pmatrix} [T_0]_{\mathcal{B}_0} & 0 \\ 0 & \lambda \cdot I_{n-m} \end{pmatrix}.$$

En particular, $[T]_{\mathcal{B}}$ es triangular inferior. □

Una consecuencia inmediata del teorema anterior es que todo operador lineal sobre un $\mathbb{C}$-espacio vectorial de dimensión finita es triangulable. Esto se debe a que $\mathbb{C}$ es un cuerpo algebraicamente cerrado.

**Corolario 4.55.** *Sea $\mathbb{K}$ un cuerpo algebraicamente cerrado. Si $V$ es un $\mathbb{K}$-espacio vectorial de dimensión finita entonces todo $T \in L(V)$ es triangulable.*

*Demostración.* Supongamos que $V$ es un $\mathbb{K}$-espacio vectorial de dimensión finita y que $T \in L(V)$. Como $\mathbb{K}$ es algebraicamente cerrado, tenemos que el polinomio minimal de $T$ tiene todas sus raíces en $\mathbb{K}$. Entonces, aplicando el teorema anterior, resulta que $T$ es triangulable. □

Para finalizar, notemos que si $T$ es triangulable (es decir, admite una representación matricial triangular inferior) entonces $T$ también admite una representación matricial triangular superior. En efecto, si $\mathcal{B} = \{b_1, \ldots, b_n\}$ es una base de $V$ tal que $[T]_{\mathcal{B}}$ es triangular inferior, entonces $\mathcal{B}' = \{b_n, b_{n-1}, \ldots, b_2, b_1\}$ también es una base de $V$ y $[T]_{\mathcal{B}'}$ es una matriz triangular superior.

## 4.8. Ejercicios

1. Sea $A = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$. Probar que 2, $2 + \sqrt{2}$ y $2 - \sqrt{2}$ son autovalores de $A$ y hallar los autovectores correspondientes.

2. Sea $A \in \mathbb{K}^{n\times n}$ una matriz inversible ¿Puede ser $\lambda = 0$ un autovalor de $A$? Probar que si $\lambda$ es un autovalor de $A$, entonces $\lambda^{-1}$ es autovalor de $A^{-1}$ y además los autoespacios asociados a $\lambda$ y $\lambda^{-1}$ pertenecientes a $A$ y $A^{-1}$ respectivamente, coinciden.

3. Probar que si $A \in \mathbb{K}^{n\times n}$ es una matriz triangular, entonces los autovalores de $A$ son los elementos de la diagonal.

4. Sea $A \in \mathbb{K}^{n\times n}$ ¿Puede tener $A$ más de $n$ autovectores linealmente independientes?

5. 
   *a*) Construir una matriz $A \in \mathbb{R}^{2\times 2}$ que tenga un sólo autovalor.
   *b*) Construir una matriz $A \in \mathbb{R}^{2\times 2}$ que tenga un sólo autovalor con un autoespacio asociado de dimensión 1.
   *c*) Construir una matriz $A \in \mathbb{R}^{2\times 2}$ que no tenga autovalores. ¿Puede hacer lo mismo para una matriz de $\mathbb{C}^{2\times 2}$?

6. Consideremos las transformaciones lineales $R_{\frac{\pi}{2}}$, $S_Y$, $H_2$ y $P_X$ presentadas en los ejercicios 3, 4, 5 y 6 del Capítulo 2. Hallar los autovalores, autovectores y autoespacios asociados. ¿Es alguna de ellas diagonalizable?

7. Dado un $\mathbb{K}$-espacio vectorial $V$ de dimensión finita, sea $T \in L(V)$. Supongamos que $\lambda$ es un autovalor de $T$ y que $v \in V$ es un autovector asociado a $\lambda$. Probar que si $p \in \mathbb{K}[x]$, entonces $p(T)\, v = p(\lambda)\, v$.

8. Dada una matriz $A \in \mathbb{K}^{n\times n}$, probar que es $A$ es inversible si y sólo si $m_A(0) \neq 0$.

9. Sea $A = \begin{pmatrix} 0 & -2 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \in \mathbb{K}^{3\times 3}$.
   *a*) Hallar el polinomio minimal de $A$ suponiendo que $\mathbb{K} = \mathbb{C}$, $\mathbb{K} = \mathbb{R}$ y $\mathbb{K} = \mathbb{Z}_3$.
   *b*) Decir en cada caso si $A$ es diagonalizable.

10. Para cada una de las siguientes matrices hallar sus autovalores y autoespacios asociados. Decir si son diagonalizables y, en caso de serlo, hallar la matriz diagonal y el cambio de base correspondiente.
$$A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{pmatrix}, \quad C = \begin{pmatrix} -3 & 0 & 0 \\ 0 & 5 & -1 \\ 0 & 6 & -2 \end{pmatrix}.$$

11. Consideremos las matrices
$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 1 \end{pmatrix}.$$
Probar que $A$ y $B$ tienen polinomios característicos diferentes, pero sus polinomios minimales coinciden.

12. Sea $T \in L(\mathbb{R}^{(2)}[x])$ dado por $T(a_0 + a_1x + a_2x^2) = (a_0 + a_1) - (2a_1 + 3a_2)x$.
   *a*) Hallar la representación matricial de $T$ con respecto a la base usual de $\mathbb{R}^{(2)}[x]$.
   *b*) Hallar el polinomio característico y los autovalores de $T$.

*c*) ¿Es $T$ diagonalizable?

13. Sea $T \in L(\mathbb{R}^3)$ dado por $T(x, y, z) = (x, x + y, z)$.

    *a*) Hallar el polinomio característico y el polinomio minimal de $T$.

    *b*) Calcular sus autovalores y una base para cada uno de los autoespacios.

    *c*) Decidir si $T$ es o no es diagonalizable, justificando de dos maneras diferentes.

14. *a*) Sea $T \in L(\mathbb{R}^3)$ tal que:

    - Sus autovalores son 1 y $-1$.
    - $\{(0, 1, -1)\}$ es una base de $N(T + I)$ y $\{(0, 1, 1), (1, 0, 0)\}$ es una base de $N(T - I)$.

    ¿Se puede determinar si $T$ es diagonalizable? Hallar el polinomio característico de $T$.

    *b*) Sea $T \in L(\mathbb{R}^4)$ tal que:

    - Sus autovalores son 1 y $-1$.
    - $\{(0, -1, 0, 0)\}$ es una base de $N(T + I)$ y $\{(0, 0, 1, 1), (1, 0, 0, 0)\}$ es una base de $N(T - I)$.

    ¿Se puede determinar si $T$ es diagonalizable?

15. ¿Cuáles son los posibles autovalores de una matriz $A$ si se sabe que $A = A^2$?

16. Determinar para qué valores de $a$ y $b$ la siguiente matriz es diagonalizable:

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

17. Sea $A$ una matriz cuadrada tal que $A \neq I$ y $A^3 - A^2 + A = I$. ¿Es $A$ diagonalizable sobre $\mathbb{C}$? ¿Y sobre $\mathbb{R}$?

18. Probar que si $A \in \mathbb{R}^{2\times 2}$ es simétrica, entonces es semejante a una matriz diagonal.

19. Sea $A \in \mathbb{K}^{n\times n}$ una matriz tal que $A^2 = A$ pero $A \neq 0$ y $A \neq I$.

    *a*) Hallar el polinomio minimal de $A$.

    *b*) Probar que $A$ es semejante a la matriz diagonal $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$, donde $r = \mathrm{rg}(A)$.

20. En cada caso, hallar una matriz $A \in \mathbb{C}^{3\times 3}$ que verifique:

    *a*) $A^2 \neq 0$ y $A^3 = 0$.

    *b*) $A \neq 0$, $A \neq I$ y $A^2 = A$.

    *c*) $A \neq I$ y $A^3 = A$.

21. Sea $A = \begin{pmatrix} 2 & -1 \\ 1 & 3 \end{pmatrix}$. Usando el Teorema de Cayley-Hamilton probar que $A^3 = 5A^2 - 7A$.

22. Sea $T \in L(\mathbb{R}^4)$ tal que

$$[T]_{\mathcal{E}} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

*a*) Probar que los únicos autovalores de $T$ son 0 y 1, pero $T$ no es una proyección.

*b*) ¿$T$ es diagonalizable?

*c*) Sea $S$ un operador diagonalizable que tiene como únicos autovalores al 0 y al 1. Mostrar que $S$ es una proyección.

23. Probar que si $T$ es una proyección entonces $I - T$ también es una proyección.

24. Sea $T : \mathbb{C}^2 \to \mathbb{C}^2$ dado por $T(x, y) = (\frac{1}{2}x + \frac{1}{2}y, \frac{1}{2}x + \frac{1}{2}y)$.

*a*) Probar que $T$ es un operador lineal, hallar su núcleo y su imagen.

*b*) Probar que $T$ es una proyección.

25. Probar que los siguientes operadores son proyecciones:

*a*) $T : \mathbb{R}_3[x] \to \mathbb{R}_3[x]$ dado por $T(ax^2 + bx + c) = c$.

*b*) $\text{diag} : \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$ dado por

$$(\text{diag}(A))_{ij} = \begin{cases} A_{ii} & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases} \quad \text{para } i, j = 1, \ldots, n.$$

26. Dado un $\mathbb{K}$-espacio vectorial $V$, sea $T \in L(V)$. Probar que $N(T)$ e $\text{Im}(T)$ son subespacios $T$-invariantes.

27. Sea

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

y supongamos que $A = [T]_{\mathcal{E}}$ para $T \in L(\mathbb{C}^3)$. Hallar tres subespacios no nulos de $\mathbb{C}^3$ que sean invariantes por $T$ y tales que $\mathbb{C}^3$ se pueda escribir como suma directa de ellos.

28. *a*) Descomponer a $\mathbb{R}^3$ como suma de directa de tres subespacios $W_1, W_2, W_3$.

*b*) Hallar las proyecciones $P_1, P_2, P_3$ correspondientes a cada uno de los subespacios del inciso *a*, respectivamente.

*c*) Hallar el polinomio minimal y el polinomio característico de $\sqrt{2}P_1 + \pi P_2 + 3P_3$.

*d*) ¿Es $\sqrt{2}P_1 + \pi P_2 + 3P_3$ un operador diagonalizable? Hallar sus autovalores y autoespacios correspondientes.

29. Sea

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

una representación matricial de $T \in L(\mathbb{C}^4)$.

*a*) Hallar el polinomio característico y el polinomio minimal de $T$. ¿Es $T$ diagonalizable?

*b*) Hallar dos subespacios de $\mathbb{C}^4$ que sean $T$-invariantes y tales que su suma directa sea $\mathbb{C}^4$.

30. Sea $T \in L(\mathbb{R}^2)$ tal que
$$[T]_{\mathcal{E}} = A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}.$$

   a) Probar que $W_1 = \overline{\{(1,0)\}}$ es $T$-invariante.
   b) Probar que no existe un subespacio $W_2$ de $\mathbb{R}^2$ que sea $T$-invariante y que además verifique $\mathbb{R}^2 = W_1 \dot{+} W_2$.

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

   a) Si $T \in L(\mathbb{R}^2)$ es tal que $[T]_{\mathcal{E}} = A$ ¿existe algún subespacio propio de $\mathbb{R}^2$ que sea $T$-invariante?
   b) Si $S \in L(\mathbb{C}^2)$ es tal que $[S]_{\mathcal{E}} = A$ ¿existe algún subespacio propio de $\mathbb{C}^2$ que sea $S$-invariante?

32. Dado un $\mathbb{K}$-espacio vectorial $V$, sean $T, S \in L(V)$.

   a) Sea $W$ un subespacio de $V$ que es invariante por $T$ y $S$. Probar que $W$ también es invariante por los operadores $T + S$ y $T \circ S$.
   b) Supongamos que $S \circ T = T \circ S$. Probar que si $\lambda$ es un autovalor de $T$, entonces el autoespacio asociado a $\lambda$ es $S$-invariante.

33. Dado un $\mathbb{K}$-espacio vectorial $V$ con $\dim(V) = n$, probar que si $N \in L(V)$ es nilpotente entonces su polinomio característico es $p_N(x) = x^n$.

34. ¿Si un operador lineal $N$ tiene como único autovalor al cero, se puede afirmar entonces que $N$ es nilpotente?

35. Sea $N \in \mathbb{C}^{2\times 2}$ tal que $N^2 = 0$. Probar que, o bien $N = 0$, o $N$ es semejante a la matriz
$$A = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

36. Sea $T \in L(\mathbb{R}^5)$ dado por
$$T(x_1, x_2, x_3, x_4, x_5) = (x_2, x_3, x_4, x_5, 0).$$

   a) Hallar la representación matricial de $T$ en la base canónica.
   b) Probar que $T$ es nilpotente.
   c) ¿Existe algún vector $v \in \mathbb{R}^5$ que sea $T$-cíclico? Hallar el subespacio cíclico generado por $(1,1,1,1,1)$.
   d) Para cada $i = 1, \ldots, 5$ hallar un subespacio $S_i$ de $\mathbb{R}^5$ que sea $T$-invariante y tal que $\dim(S_i) = i$.
   e) Probar que no existen dos subespacios propios $S_1$ y $S_2$ de $\mathbb{R}^5$ que sean $T$-invariantes y que verifiquen $\mathbb{R}^5 = S_1 \dot{+} S_2$.

37. Dado un $\mathbb{K}$-espacio vectorial $V$ con $\dim(V) = n$, sea $B = \{b_1, \cdots, b_n\}$ una base (ordenada) de $V$. Sea $T \in L(V)$ dado por
$$T(b_j) = \begin{cases} b_{j+1} & \text{si } j = 1, \ldots, n-1, \\ 0 & \text{si } j = n. \end{cases}$$

*a*) Probar que el único autovalor de $T$ es cero.

*b*) Probar que $T$ es nilpotente.

*c*) ¿Cuánto vale la traza de $T$? ¿Y el determinante de $T$?

# Capítulo 5

# Teorema de la Descomposición Cíclica y Forma de Jordan

El objetivo de este capítulo es encontrar una forma canónica elemental para aquellos operadores lineales que no sean diagonalizables. Para esto será necesario desarrollar previamente algunas herramientas.

## 5.1. Espacio cociente

En esta primera sección, dado un subespacio $W$ de un $\mathbb{K}$-espacio vectorial $V$ contruiremos un nuevo $\mathbb{K}$-espacio vectorial, el cual es isomorfo a un complemento algebraico de $W$ en $V$.

**Definición 5.1.** Dado un $\mathbb{K}$-espacio vectorial $V$, fijemos un subespacio $W$ de $V$. Dados $u, v \in V$, diremos que $u$ **es congruente con** $v$ **módulo** $W$ si $v - u \in W$. En tal caso, anotaremos $u \equiv_W v$.

**Proposición 5.2.** *La relación de congruencia módulo $W$ es una relación de equivalencia:*

*i)* $v \equiv_W v$ *para cada* $v \in V$*;*

*ii)* *si* $u \equiv_W v$ *entonces* $v \equiv_W u$*, para cualesquiera* $u, v \in V$*,*

*iii)* *si* $u \equiv_W v$ *y* $v \equiv_W w$ *entonces* $u \equiv_W w$*, para cualesquiera* $u, v, w \in V$*.*

*Además, esta relación es compatible con las operaciones del* $\mathbb{K}$*-espacio vectorial* $V$*:*

*iv)* *si* $u_1 \equiv_W v_1$ *y* $u_2 \equiv_W v_2$ *entonces* $u_1 + u_2 \equiv_W v_1 + v_2$*;*

*v)* *si* $u \equiv_W v$ *y* $\alpha \in \mathbb{K}$ *entonces* $\alpha u \equiv_W \alpha v$*.*

*Demostración.* Ejercicio. □

Dado un vector $v \in V$, sea $[v]$ la clase de equivalencia módulo $W$. Notemos que

$$[v] = \{u \in V : \ v \equiv_W u\} = \{u \in V : \ u - v \in W\},$$

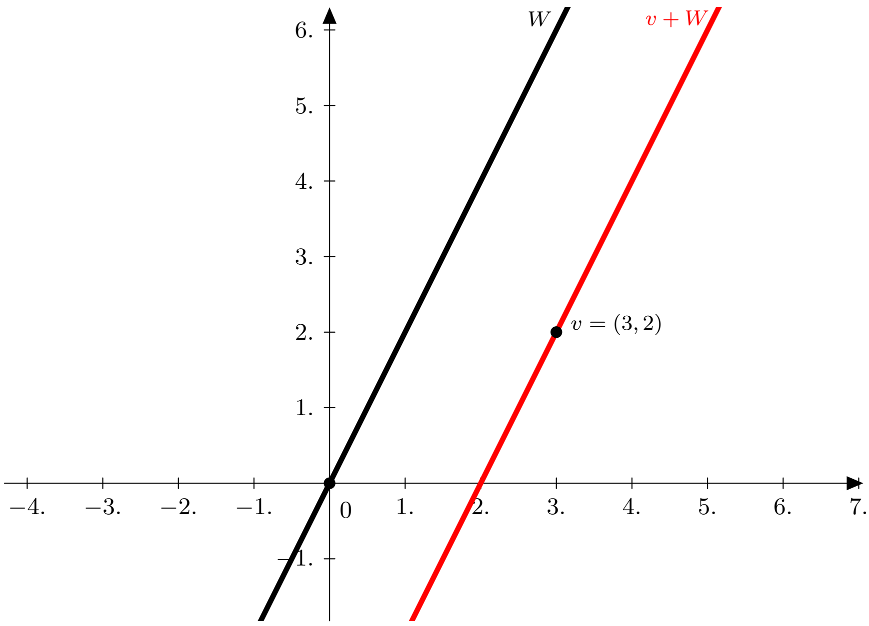
y como $u = v + (u - v)$, usualmente la clase $[v]$ se anota $v + W$.

**Ejemplo 5.3.** En $\mathbb{R}^2$, consideremos el subespacio

$$W = \{(x, 2x) : \ x \in \mathbb{R}\}.$$

La representación gráfica de $W$ es la recta con pendiente $m = 2$ que pasa por el origen.

Luego, si $v = (3, 2)$, la clase de equivalencia de $v$ es la recta con pendiente $m = 2$ que pasa por $v = (3, 2)$.




**Definición 5.4.** Si $W$ es un subespacio de $V$, el **espacio cociente** $V/W$ es el conjunto de clases de equivalencias módulo $W$:

$$V/W = \{v + W : \ v \in V\}.$$

Los ítems *iv*) y *v*) de la proposición anterior dicen que, si definimos

$$(u + W) + (v + W) := (u + v) + W,$$
$$\alpha \cdot (v + W) := \alpha u + W,$$

estas operaciones están bien definidas y, además, $V/W$ dotado con ellas resulta un $\mathbb{K}$-espacio vectorial. A la función $\pi : V \to V/W$ definida por

$$\pi(v) = v + W,$$

la llamaremos **la proyección al cociente**.

**Ejercicio 5.5.** Probar que la proyección al cociente $\pi$ es una transformación lineal suryectiva (i.e. un epimosfismo).

Veamos ahora que $V/W$ es isomorfo a un complemento algebraico de $W$ en $V$.

**Teorema 5.6.** *Sea $V$ un $\mathbb{K}$-espacio vectorial de dimensión finita. Si $W$ es un subespacio de $V$ entonces*

$$\dim(V/W) = \dim V - \dim W.$$

*Además, $V/W$ es isomorfo a un subespacio $S$ de $V$ tal que*

$$V = W \dotplus S.$$

*Demostración.* Si $W$ es un subespacio de $V$, consideremos la proyección al cociente $\pi : V \to V/W$. Dada una base $\mathcal{B} = \{b_1, \ldots, b_n\}$ de $V$, resulta que

$$\{b_1 + W, b_2 + W, \ldots, b_n + W\}$$

es un conjunto de generadores de $V/W$, ya que $\pi$ es un epimorfismo. En particular, vemos que $\dim(V/W) < \infty$.

Describamos ahora cual es el núcleo de $\pi$: dado $v \in V$, $\pi(v) = \vec{0}_{V/W} = \vec{0} + W = W$ si y sólo si $\vec{0} \equiv_W v$, o equivalentemente, si y sólo si $v \in W$. Por lo tanto, $N(\pi) = W$.

Luego, aplicando el teorema de la dimensión, obtenemos que

$$\dim V = \mathrm{rg}(\pi) + \mathrm{nul}(\pi) = \dim(V/W) + \dim W,$$

de donde se desprende que $\dim(V/W) = \dim V - \dim W$.

Por último, veamos que existe un subespacio $S$ de $V$ que es isomorfo a $V/W$ y además $V = W \dotplus S$. Supongamos que

$$\{v_1 + W, v_2 + W, \ldots, v_k + W\} \tag{5.1}$$

es una base de $V/W$, y consideremos el subespacio $S$ de $V$ generado por $\{v_1, v_2, \ldots, v_k\}$, i.e. $S := \overline{\{v_1, v_2, \ldots, v_k\}}$.

Veamos primero que $W \cap S = \{\vec{0}\}$. Si $v \in W \cap S$, existen $\alpha_1, \ldots, \alpha_k \in \mathbb{K}$ tales que $v = \sum_{i=1}^k \alpha_i v_i \in W$. Luego,

$$\vec{0} + W = \pi(v) = \pi\left(\sum_{i=1}^k \alpha_i v_i\right) = \sum_{i=1}^k \alpha_i \pi(v_i) = \sum_{i=1}^k \alpha_i (v_i + W).$$

Como el conjunto (5.1) es una base de $V/W$, en particular es linealmente independiente, y resulta que $\alpha_1 = \ldots = \alpha_k = 0$. Entonces, $v = \vec{0}$.

Por otra parte, dado un $u \in V$ arbitrario, $u + W \in V/W$. Entonces, existen $\beta_1, \ldots, \beta_k \in \mathbb{K}$ tales que

$$u + W = \sum_{i=1}^k \beta_i (v_i + W) = \left(\sum_{i=1}^k \beta_i v_i\right) + W,$$

es decir, $u \equiv_W \sum_{i=1}^k \beta_i v_i$. Luego,

$$u = \overbrace{\left(u - \sum_{i=1}^k \beta_i v_i\right)}^{\in W} + \overbrace{\sum_{i=1}^k \beta_i v_i}^{\in S}.$$

Por lo tanto, $V = W \dotplus S$.

Sólo falta probar que $\pi|_S : S \to V/W$ es un isomorfismo. En primer lugar notemos que los dos espacios tienen la misma dimensión. Además, $\pi|_S$ también es un epimorfismo, porque

$$\pi(\{v_1, \ldots, v_k\}) = \{v_1 + W, v_2 + W, \ldots, v_k + W\}.$$

En consecuencia, $\pi|_S$ es un isomorfismo entre $S$ y $V/W$. □

## 5.2. Operadores nilpotentes y subespacios cíclicos

Retomemos ahora el estudio de los operadores nilpotentes, con el fin de encontrar bases adecuadas para representarlos.

**Definición 5.7.** Sea $V$ un $\mathbb{K}$-espacio vectorial de dimensión finita. Dado un operador nilpotente $T \in L(V)$, llamaremos **indice de nilpotencia de** $T$ al grado del polinomio minimal de $T$, es decir, $T$ es nilpotente de orden $k$ si $m_T(x) = x^k$. Equivalentemente, $T$ es nilpotente de orden $k$ si

$$T^k = 0 \quad \text{pero} \quad T^{k-1} \neq 0.$$

Si $T \in L(V)$ es nilpotente y $\dim V = n$, entonces su polinomio característico es

$$p_T(x) = x^n.$$

En el siguiente resultado probaremos que si el polinomio minimal y el característico de un operador nilpotente coinciden, entonces el operador tiene una representación matricial elemental. Ésta es una matriz triangular inferior, cuyas únicas entradas no nulas son todas iguales a 1 y se encuentran en la diagonal por debajo de la diagonal principal.

**Proposición 5.8.** *Dado un* $\mathbb{K}$*-espacio vectorial con* $\dim V = n$*, sea* $T \in L(V)$ *un operador nilpotente de orden* $n$*. Entonces, existe una base ordenada* $\mathcal{B}$ *de* $V$ *tal que*

$$[T]_{\mathcal{B}} = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & 0 & \dots & 1 & 0 \end{pmatrix}. \tag{5.2}$$

*Demostración.* Supongamos que $T \in L(V)$ es un operador nilpotente de orden $n$, es decir $T^n = 0$ pero $T^{n-1} \neq 0$. Como $T^{n-1} \neq 0$, existe $b_1 \in V$ tal que $T^{n-1}b_1 \neq \vec{0}$. A partir de este vector construiremos una base de $V$ con las características deseadas. Para cada $k = 2, \dots n$, sea

$$b_k = T^{k-1}b_1.$$

Veamos que $\mathcal{B} := \{b_1, \dots, b_n\}$ es linealmente independiente. Supongamos que existen escalares $\alpha_1, \dots, \alpha_n \in \mathbb{K}$ tales que $\alpha_1 b_1 + \alpha_2 b_2 + \dots + \alpha_n b_n = \vec{0}$. Equivalentemente,

$$\alpha_1 b_1 + \alpha_2 T b_1 + \alpha_3 T^2 b_1 \dots + \alpha_n T^{n-1} b_1 = \vec{0}.$$

Aplicando $T^{n-1}$ a ambos lados de la igualdad anterior, obtenemos

$$\begin{aligned} \vec{0} = T^{n-1}\vec{0} &= T^{n-1}\big(\alpha_1 b_1 + \alpha_2 T b_1 + \alpha_3 T^2 b_1 \dots + \alpha_n T^{n-1} b_1\big) \\ &= \alpha_1 T^{n-1} b_1 + \alpha_2 T^n b_1 + \alpha_3 T^{n+1} b_1 + \dots + \alpha_n T^{2n-2} b_1 \\ &= \alpha_1 T^{n-1} b_1, \end{aligned}$$

ya que $T^n$ (y cualquier potencia de $T$ mayor que $n$) es igual al operador nulo. Además, como $T^{n-1}b_1 \neq \vec{0}$, resulta que

$$\alpha_1 = 0 \quad \text{y} \quad \alpha_2 b_2 + \dots + \alpha_n b_n = \vec{0}.$$

Ahora, aplicando $T^{n-2}$ a ambos lados de la igualdad anterior, obtenemos

$$\begin{aligned}\vec{0} = T^{n-2}\vec{0} &= T^{n-2}\big(\alpha_2 Tb_1 + \alpha_3 T^2 b_1 \ldots + \alpha_n T^{n-1} b_1\big)\\ &= \alpha_2 T^{n-1} b_1 + \alpha_3 T^n b_1 + \ldots + \alpha_n T^{2n-3} b_1\\ &= \alpha_2 T^{n-1} b_1.\end{aligned}$$

De donde resulta que

$$\alpha_2 = 0 \quad \text{y} \quad \alpha_3 b_3 + \ldots + \alpha_n b_n = \vec{0}.$$

Siguiendo con este proceso, obtendremos que $\alpha_1 = \alpha_2 = \ldots = \alpha_n = 0$. Por lo tanto, $\mathcal{B}$ es linealmente independiente.

Además, como $\dim V = n$, tenemos que $\mathcal{B}$ es una base de $V$ tal que

$$\begin{aligned}Tb_k &= T(T^{k-1}b_1) = T^k b_1 = b_{k+1} \quad \text{si } k = 1, \ldots, n-1,\\ Tb_n &= T(T^{n-1}b_1) = T^n b_1 = \vec{0},\end{aligned}$$

de donde se desprende la representación matricial anticipada. □

Dado un $\mathbb{K}$-espacio vectorial $V$ de dimensión finita, consideremos un operador $T \in L(V)$ cualquiera. Fijado un vector $v \in V$, ¿cuál es el menor subespacio $Z(v,T)$ que contiene a $v$ y es invariante por $T$?

Este subespacio puede definirse como la intersección de todos los subespacios $T$-invariantes que contienen a $v$, pero esta descripción es poco práctica. Por otra parte, además de a $v$, este subespacio debe contener a $Tv$, $T^2v$, etc., es decir, debe contener a todos los vectores de la forma $p(T)v$ con $p \in \mathbb{K}[x]$.

**Definición 5.9.** Si $v \in V$ es un vector arbitrario, el **subespacio $T$-ciclico generado por** $v$ es el subespacio

$$Z(v,T) := \big\{p(T)v : \ p \in \mathbb{K}[x] \ \big\}. \tag{5.3}$$

Si $Z(v,T) = V$, diremos que $v$ es un **vector ciclico para** $T$.

Si $\dim V = n$, es fácil convencerse de que

$$Z(v,T) = \overline{\{v, Tv, \ldots, T^{n-1}v\}},$$

pero $\{v, Tv, \ldots, T^{n-1}v\}$ podría no ser linealmente independiente, con lo cual esta descripción tampoco es del todo eficiente.

**Observación 5.10.** El vector $b_1$ que consideramos en la demostración de la Proposición 5.8 es un vector cíclico para $T$, porque el subespacio cíclico determinado por $b_1$,

$$Z(b_1,T) = \overline{\{b_1, Tb_1, \ldots, T^{n-1}b_1\}},$$

coincide con $V$. Notemos que (cuando existe) el vector cíclico no es único, ya que si lo multiplicamos por cualquier escalar $\alpha \in \mathbb{K}$, $\alpha \neq 0$, tendremos otro vector cíclico para $T$.

En general, si $T$ no es nilpotente ó si el orden de nilpotencia es menor que $\dim V$, no existe un vector cíclico para $T$. Sin embargo, más adelante veremos que el espacio puede descomponerse como una suma directa de subespacios cíclicos.

Consideremos ahora el conjunto de todos los polinomios tales que, evaluados en $T$, anulan a un vector fijo $v$:

$$\mathcal{M}(v,T) := \big\{q \in \mathbb{K}[x] : \ q(T)v = \vec{0} \ \big\}. \tag{5.4}$$

Es fácil ver que es un ideal de $\mathbb{K}[x]$, que contiene al ideal $\mathcal{A}(T)$ de polinomios anuladores de $T$: si $p \in \mathcal{A}(T)$ tenemos que $p(T) = 0$, en particular $p(T)v = 0v = \vec{0}$, con lo cual $p \in \mathcal{M}(v,T)$.

**Definición 5.11.** Al único polinomio mónico $p_v$ que genera al ideal $\mathcal{M}(v,T)$ lo llamaremos el **$T$-anulador de** $v$.

Como $\mathcal{M}(v,T)$ contiene a $\mathcal{A}(T)$, resulta que el $T$-anulador de $v$ divide al polinomio minimal $m_T$ de $T$:

$$p_v \mid m_T.$$

**Teorema 5.12.** *Dado un vector $v \in V$, $v \neq \vec{0}$, sea $p_v$ el $T$-anulador de $v$.*

*i) El grado de $p_v$ es igual a la dimensión del subespacio cíclico $Z(v,T)$.*

*ii) Si $\mathrm{gr}(p_v) = k$, entonces $\{v, Tv, \ldots, T^{k-1}v\}$ es una base de $Z(v,T)$.*

*iii) Si $U = T|_{Z(v,T)}$ entonces $m_U = p_v$.*

*Demostración.* Supongamos que $\mathrm{gr}(p_v) = k$. Cada $p \in \mathbb{K}[x]$ se escribe como

$$p = p_v q + r,$$

donde $r = 0$ ó $\mathrm{gr}(r) < \mathrm{gr}(p_v) = k$. Como $p_v q \in \mathcal{M}(v,T)$, tenemos que

$$p(T)v = r(T)v.$$

Pero las condiciones $r = 0$ ó $\mathrm{gr}(r) < \mathrm{gr}(p_v) = k$ implican que $r(T)v$ es combinación lineal de $v, Tv, \ldots, T^{k-1}v$. Luego, la ecuación anterior y (5.3) garantizan que $v, Tv, \ldots, T^{k-1}v$ generan a $Z(v,T)$. Veamos que además estos vectores son linealmente independientes. Supongamos que existen $\alpha_0, \ldots, \alpha_{k-1}$ tales que

$$\alpha_0 v + \alpha_1 Tv + \ldots + \alpha_{k-1}T^{k-1}v = \vec{0}.$$

Luego, el polinomio $g(x) := \alpha_0 + \alpha_1 x + \ldots + \alpha_{k-1}x^{k-1}$ tiene grado menor que $k$ y cumple $g(T)v = \vec{0}$, lo que contradice la definición de $p_v$. Esto completa la demostración de los ítems *i)* y *ii)*.

Por último, consideremos $U = T|_{Z(v,T)}$. Queremos ver que $m_U = p_v$, para lo cual necesitamos que $p_v(U)w = \vec{0}$ para cualquier $w \in Z(v,T)$. De (5.3) sabemos que cada $w \in Z(v,T)$ se escribe como $w = p(T)v$ para algún $p \in \mathbb{K}[x]$. Luego,

$$p_v(U)w = p_v(T)w = p_v(T)p(T)v = p(T)p_v(T)v = p(T)\vec{0} = \vec{0},$$

es decir, $p_v(U) = 0$. Ahora, supongamos que $q \in \mathbb{K}[x]$ tiene $\mathrm{gr}(q) < \mathrm{gr}(p_v) = k$ y que $q$ anula a $U$. Entonces, $q(U)v = \vec{0}$ y $\mathrm{gr}(q) < \mathrm{gr}(p_v) = k$, lo que contradice la definición de $p_v$. Por lo tanto, $m_U = p_v$. □

Consideremos ahora un operador $T \in L(V)$ nilpotente de orden $k$. En este caso, $m_T(x) = x^k$. Luego, para cada $v \in V$ el $T$-anulador de $v$ será de la forma

$$p_v(x) = x^r, \quad \text{con } r \leq k.$$

Como consecuencia del teorema anterior sabemos además que $\{v, Tv, \ldots, T^{r-1}v\}$ es linealmente independiente y

$$Z(v,T) = \overline{\{v, Tv, \ldots, T^{r-1}v\}}.$$

A continuación enunciaremos una primera versión del Teorema de la Descomposición Cíclica (para un operador nilpotente).

**Teorema 5.13.** *Dado un $\mathbb{K}$-espacio vectorial $V$ con $\dim V = n$, sea $T \in L(V)$ un operador nilpotente. Entonces existen vectores $v_1, \ldots, v_l \in V$ con polinomios $T$-anuladores*

$$p_{v_i}(x) = x^{k_i}, \quad i = 1, \ldots, l,$$

*que verifican:*

*i)* $V = Z(v_1, T) \dot{+} Z(v_2, T) \dot{+} \ldots \dot{+} Z(v_l, T)$*;*

*ii)* $p_{v_1} = m_T$*,* $p_{v_2} \mid p_{v_1}$*, ...,* $p_{v_l} \mid p_{v_{l-1}}$*;*

*es decir, $k_1$ es el orden de nilpotencia de $T$, $k_1 \geq k_2 \geq \ldots \geq k_l$, y $\sum_{i=1}^{l} k_i = n$.*

En lo que sigue utilizaremos la idea de espacio cociente para ver como encontrar a los vectores $v_1, \ldots, v_l$ del teorema anterior.

## 5.3. Teorema de la Descomposición Cíclica

El contenido de esta sección está basado en el libro "Fundamentals of Linear Algebra" de K. Nomizu [7].

Dado un $\mathbb{K}$-espacio vectorial $V$ de dimensión finita, supongamos que $T \in L(V)$ es nilpotente de orden $k$. Entonces,

$$N(T) \subseteq N(T^2) \subseteq \ldots \subseteq N(T^k) = V. \tag{5.5}$$

Intentaremos motivar el Teorema de la Descomposición Cíclica a partir de esta cadena creciente de subespacios, estudiando los espacios cociente

$$\frac{N(T^{j+1})}{N(T^j)} \qquad j = 1, \ldots, k-1.$$

En primer lugar, para garantizar que los cocientes anteriores no son triviales, veamos que las inclusiones en (5.5) son estrictas.

**Proposición 5.14.** *Dado un $\mathbb{K}$-espacio vectorial $V$ de dimensión finita, supongamos que $T \in L(V)$ es nilpotente de orden $k$. Entonces,*

$$N(T) \subset N(T^2) \subset \ldots \subset N(T^{k-1}) \subset N(T^k) = V, \tag{5.6}$$

*y todas las inclusiones son estrictas.*

*Demostración.* En primer lugar, como el orden de nilpotencia de $T$ es $k$, tenemos que $N(T^{k-1}) \neq V$ y $N(T^k) = V$. Para probar que el resto de las inclusiones son estrictas, veamos que si $N(T^m) = N(T^{m+1})$ para algún $m \in \mathbb{N}$ entonces $N(T^m) = N(T^{m+q})$ para todo $q \in \mathbb{N}$.

Supongamos que $N(T^m) = N(T^{m+1})$ para algún $m \in \mathbb{N}$ y veamos que $N(T^{m+1}) = N(T^{m+2})$. La inclusión $N(T^{m+1}) \subseteq N(T^{m+2})$ vale para cualquier operador $T$. Para probar la otra, supongamos que $v \in N(T^{m+2})$. Entonces, $Tv \in N(T^{m+1}) = N(T^m)$, o equivalentemente, $v \in N(T^{m+1})$. El resto de la prueba de la afirmación sale aplicando el principio de inducción completa.

Ahora bien, si suponemos que $N(T^m) = N(T^{m+1})$ para algún $m < k$ tendríamos en particular que $N(T^{k-1}) = N(T^k)$, lo cual resulta absurdo. Por lo tanto, $m = k$ y las inclusiones en (5.6) son todas estrictas. $\square$

Suponiendo que $T$ es nilpotente de orden $k$, aplicando el Teorema 5.6 tenemos que

$$\dim V = \dim N(T^k) = \dim \frac{N(T^k)}{N(T^{k-1})} + \dim N(T^{k-1}).$$

De la misma manera, tenemos que $\dim N(T^{k-1}) = \dim \frac{N(T^{k-1})}{N(T^{k-2})} + \dim N(T^{k-2})$ y así sucesivamente. Por lo tanto,

$$\dim V = \dim N(T) + \sum_{j=1}^{k-1} \dim \frac{N(T^{j+1})}{N(T^j)} = \sum_{h=1}^{k} \dim \frac{N(T^h)}{N(T^{h-1})}, \tag{5.7}$$

es decir, las dimensiones de $N(T)$ y los distintos espacios cocientes $\frac{N(T^{j+1})}{N(T^j)}$ forman una **partición** de la dimensión de $V$ de acuerdo a la siguiente definición.

**Definición 5.15.** Dado un número natural $n \in \mathbb{N}$, una sucesión finita $\{m_1, m_2, \ldots, m_k\}$ en $\mathbb{N}$ es una **partición de** $n$ si

$$m_1 \geq m_2 \geq \ldots \geq m_k \qquad \text{y} \qquad \sum_{j=1}^{k} m_j = n.$$

**Ejercicio 5.16.** Probar que efectivamente $\dim N(T) \geq \dim \frac{N(T^2)}{N(T)} \geq \ldots \geq \dim \frac{N(T^k)}{N(T^{k-1})}$.
(<u>Idea</u>: Para probar que $\dim N(T) \geq \dim \frac{N(T^2)}{N(T)}$ considerar una base $\{b_1, \ldots, b_p\}$ de $N(T)$ y extenderla a una base de $\{b_1, \ldots, b_p, b_{p+1}, \ldots, b_{p+q}\}$ de $N(T^2)$. Llamando $w_j = Tb_{p+j}$, para $j = 1, \ldots, q$, comprobar que $\{w_1, \ldots, w_q\}$ es un conjunto linealmente independiente en $N(T)$. Esto prueba que $q \leq p$, es decir, $\dim N(T) \geq \dim \frac{N(T^2)}{N(T)}$. La idea para probar que $\dim \frac{N(T^{h-1})}{N(T^{h-2})} \geq \dim \frac{N(T^h)}{N(T^{h-1})}$ para $h = 2, \ldots, k$ es similar.)

Por otra parte, si $T$ es nilpotente de orden $k$, cada clase de equivalencia (no nula) en $\frac{N(T^k)}{N(T^{k-1})}$ determina un subespacio cíclico $Z(v, T)$ de dimensión $k$, ya que si $[v] = v + N(T^{k-1})$ entonces $\{v, Tv, \ldots, T^{k-1}v\}$ es linealmente independiente (probarlo!).

**Definición 5.17.** Dado un $\mathbb{K}$-espacio vectorial $V$ de dimensión finita, sea $W$ un subespacio de $V$. Diremos que un subconjunto $\{v_1, \ldots, v_k\}$ de $V$ es **linealmente independiente módulo** $W$ si $\{v_1 + W, v_2 + W, \ldots, v_k + W\}$ es linealmente independiente en $V/W$, es decir, si $\sum_{i=1}^k \alpha_i v_i \in W$ implica que $\alpha_1 = \ldots = \alpha_k = 0$.

El siguiente lema muestra que cada conjunto linealmente independiente de clases de equivalencia en $\frac{N(T^k)}{N(T^{k-1})}$ determina una familia independiente de subespacios cíclicos.

**Lema 5.18.** *Dado un $\mathbb{K}$-espacio vectorial $V$ de dimensión finita, sea $T \in L(V)$ un operador nilpotente de orden $k$. Si $\{v_1, \ldots, v_m\}$ es linealmente independiente módulo $N(T^{k-1})$, entonces los vectores*

$$\begin{array}{llll} v_1, & v_2, & \ldots\,, & v_m, \\ Tv_1, & Tv_2, & \ldots\,, & Tv_m, \\ \vdots & \vdots & & \vdots \\ T^{k-1}v_1, & T^{k-1}v_2, & \ldots\,, & T^{k-1}v_m, \end{array}$$

*son linealmente independientes en $V$.*

*Demostración.* Supongamos que existen $\alpha_{ij} \in \mathbb{K}$ para $i = 1, \ldots, m$, $j = 1, \ldots, k$ tales que

$$\sum_{i=1}^{m} \sum_{j=1}^{k} \alpha_{ij} T^{j-1} v_i = \vec{0}. \tag{5.8}$$

Aplicando $T^{k-1}$ a ambos lados de la ecuación (recordando que $T^k = 0$) obtenemos

$$T^{k-1} \left( \alpha_{11} v_1 + \ldots + \alpha_{m1} v_m \right) = \vec{0},$$

es decir, $\alpha_{11} v_1 + \ldots + \alpha_{m1} v_m \in N(T^{k-1})$. Como por hipótesis $\{v_1, \ldots, v_m\}$ es linealmente independiente módulo $N(T^{k-1})$, resulta que $\alpha_{11} = \ldots = \alpha_{m1} = 0$.

Luego, (5.8) equivale a

$$\sum_{i=1}^{m} \sum_{j=2}^{k} \alpha_{ij} T^{j-1} v_i = \vec{0},$$

y aplicando $T^{k-2}$ a ambos lados de la ecuación obtenemos

$$T^{k-1} \left( \alpha_{12} v_1 + \ldots + \alpha_{m2} v_m \right) = T^{k-2} \left( \alpha_{12} T v_1 + \ldots + \alpha_{m2} T v_m \right) = \vec{0},$$

es decir, $\alpha_{12} v_1 + \ldots + \alpha_{m2} v_m \in N(T^{k-1})$. En consecuencia, $\alpha_{12} = \ldots = \alpha_{m2} = 0$.

Repitiendo este proceso las $k-2$ veces restantes, obtendremos que $\alpha_{ij} = 0$ para $i = 1, \ldots, m$, $j = 1, \ldots, k$, como queríamos probar. □

Como consecuencia del lema anterior, si $T$ es nilpotente de orden $k$ y $\{[v_1], \ldots, [v_{m_k}]\}$ es una base de $N(T^k)/N(T^{k-1})$, los subespacios

$$Z(v_1, T), Z(v_2, T), \ldots, Z(v_{m_k}, T)$$

son independientes. Por lo tanto, la suma de estos subespacios es una suma directa:

$$W_k := Z(v_1, T) \dot{+} Z(v_2, T) \dot{+} \ldots \dot{+} Z(v_{m_k}, T)$$

y es un subespacio $T$-invariante. El objetivo del siguiente resultado es probar que $W_k$ admite un complemento algebraico en $V$ que también es $T$-invariante.

**Lema 5.19.** *Bajo las hipótesis anteriores, existe un subespacio $S$ de $V$ tal que*

$$V = W_k \dot{+} S \quad y \quad T(S) \subseteq S.$$

*Demostración.* Haremos la prueba por inducción en la dimensión del espacio $V$. Si $\dim V = 1$, no hay nada que probar. Como hipótesis inductiva, supongamos que el resultado vale para cualquier $\mathbb{K}$-espacio vectorial con dimensión menor que $n$.

Supongamos ahora que $\dim V = n$ y sea $\{[v_1], [v_2], \ldots, [v_{m_k}]\}$ una base de $N(T^k)/N(T^{k-1})$. Por el Teorema 5.6, tenemos que

$$V = N(T^{k-1}) \dot{+} \overline{\{v_1, v_2, \ldots, v_{m_k}\}}.$$

Consideremos ahora la restricción de $T$ a $N(T^{k-1})$:

$$T_1 := T|_{N(T^{k-1})} : N(T^{k-1}) \to N(T^{k-1}),$$

la cual es obviamente un operador nilpotente de orden $k-1$. De acuerdo al Lema 5.18, los vectores $Tv_1, Tv_2, \ldots, Tv_{m_k}$ son linealmente independientes. En realidad, de dicho lema se desprende una conclusión aún más fuerte, estos vectores son linealmente independientes módulo $N(T^{k-2})$. De hecho, supongamos que existen $\alpha_1, \alpha_2, \ldots, \alpha_{m_k} \in \mathbb{K}$ tales que $\sum_{i=1}^{m_k} \alpha_i T v_i \in N(T^{k-2})$. Esto dice que $\sum_{i=1}^{m_k} \alpha_i v_i \in N(T^{k-1})$, y como $\{v_1, v_2, \ldots, v_{m_k}\}$ es linealmente independiente módulo $N(T^{k-1})$, resulta que $\alpha_1 = \alpha_2 = \ldots = \alpha_{m_k} = 0$.

Luego, podemos completar el conjunto $\{[Tv_1], [Tv_2], \ldots, [Tv_{m_k}]\}$ a una base del espacio $N(T^{k-1})/N(T^{k-2})$. Supongamos que $w_1, \ldots, w_{m_{k-1}} \in N(T^{k-1})$ son tales que

$$\{[Tv_1], [Tv_2], \ldots, [Tv_{m_k}], [w_1], \ldots, [w_{m_{k-1}}]\} \text{ es una base de } N(T^{k-1})/N(T^{k-2}).$$

Como $\dim N(T^{k-1}) < \dim V = n$, podemos aplicar la hipótesis inductiva al operador nilpotente $T_1$ (de orden $k-1$) que actúa en este $\mathbb{K}$-espacio vectorial, es decir, si $W$ es el subespacio generado por los vectores:

$$\begin{array}{llllllll}
Tv_1, & Tv_2, & \ldots, & Tv_{m_k}, & w_1, & w_2, & \ldots, & w_{m_{k-1}}, \\
T^2v_1, & T^2v_2, & \ldots, & T^2v_{m_k}, & Tw_1, & Tw_2, & \ldots, & Tw_{m_{k-1}}, \\
\vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\
T^{k-1}v_1, & T^{k-1}v_2, & \ldots, & T^{k-1}v_{m_k}, & T^{k-2}w_1, & T^{k-w}w_2, & \ldots, & T^{k-2}w_{m_{k-1}},
\end{array}$$

existe un subespacio $S_1$ de $N(T^{k-1})$ tal que

$$N(T^{k-1}) = W \dotplus S_1 \quad \text{y} \quad T(S_1) = T_1(S_1) \subseteq S_1.$$

Luego, tenemos que

$$V = N(T^{k-1}) \dotplus \overline{\{v_1, v_2, \ldots, v_{m_k}\}} = W \dotplus S_1 \dotplus \overline{\{v_1, v_2, \ldots, v_{m_k}\}}$$

Como $W = Z(Tv_1, T) \dotplus \ldots \dotplus Z(Tv_{m_k}, T) \dotplus Z(w_1, T) \dotplus \ldots \dotplus Z(w_{m_{k-1}}, T)$, y además

$$Z(v_i, T) = \overline{\{v_i\}} \dotplus Z(Tv_i, T), \quad i = 1, \ldots, m_k,$$

resulta que

$$\begin{aligned}
V &= S_1 \dotplus Z(w_1, T) \dotplus \ldots \dotplus Z(w_{m_{k-1}}, T) \dotplus Z(v_1, T) \dotplus \ldots \dotplus Z(v_{m_k}, T) \\
&= \big(S_1 \dotplus Z(w_1, T) \dotplus \ldots \dotplus Z(w_{m_{k-1}}, T)\big) \dotplus \big(Z(v_1, T) \dotplus \ldots \dotplus Z(v_{m_k}, T)\big) \\
&= \big(S_1 \dotplus Z(w_1, T) \dotplus \ldots \dotplus Z(w_{m_{k-1}}, T)\big) \dotplus W_k.
\end{aligned}$$

Llamando $S := S_1 \dotplus Z(w_1, T) \dotplus \ldots \dotplus Z(w_{m_{k-1}}, T)$ vemos que $S$ es $T$-invariante y además

$$V = W_k \dotplus S. \qquad \square$$

Notemos que la demostración anterior también nos da una idea de como encontrar los subespacios cíclicos necesarios para probar el Teorema de la Descomposición Cíclica. Comenzando con una base $\{[v_1], [v_2], \ldots, [v_{m_k}]\}$ de $\frac{V}{N(T^{k-1})} = \frac{N(T^k)}{N(T^{k-1})}$, desarrollamos los subespacios cíclicos correspondientes (los cuales son de dimensión $k$). Luego, si $\{[Tv_1], [Tv_2], \ldots, [Tv_{m_k}]\}$ no es una base de $\frac{N(T^{k-1})}{N(T^{k-2})}$, la completamos con clases $[w_1], \ldots, [w_{m_{k-1}}]$ y agregamos los subespacios cíclicos (de dimensión $k-1$) que éstas generan. A continuación tendremos que analizar si $\{[T^2v_1], [T^2v_2], \ldots, [T^2v_{m_k}], [Tw_1], \ldots, [Tw_{m_{k-1}}]\}$ es una base de $\frac{N(T^{k-2})}{N(T^{k-3})}$, y en caso contrario completarlo a una base. Y así sucesivamente ...

$$
\begin{array}{lccccccccc}
N(T^k) & \overset{v_1}{\bullet} & \overset{v_2}{\bullet} & \ldots & \overset{v_{m_k}}{\bullet} & & & & & \\
N(T^{k-1}) & \overset{Tv_1}{\bullet} & \overset{Tv_2}{\bullet} & \ldots & \overset{Tv_{m_k}}{\bullet} & \overset{w_1}{\bullet} & \overset{w_2}{\bullet} & \ldots & \overset{w_{m_{k-1}}}{\bullet} & \\
N(T^{k-2}) & \overset{T^2v_1}{\bullet} & \overset{T^2v_2}{\bullet} & \ldots & \overset{T^2v_{m_k}}{\bullet} & \overset{Tw_1}{\bullet} & \overset{Tw_2}{\bullet} & \ldots & \overset{Tw_{m_{k-1}}}{\bullet} & \ldots \\
\vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \\
N(T) & \overset{T^{k-1}v_1}{\bullet} & \overset{T^{k-1}v_2}{\bullet} & \ldots & \overset{T^{k-1}v_{m_k}}{\bullet} & \overset{T^{k-2}w_1}{\bullet} & \overset{T^{k-w}w_2}{\bullet} & \ldots & \overset{T^{k-2}w_{m_{k-1}}}{\bullet} & \ldots
\end{array}
$$

El gráfico anterior intenta representar la construcción de los subespacios cíclicos antes descripta. Aquí cada columna representa a la base de un subespacio cíclico. Además, debe interpretarse que si un vector está en una fila dada entonces no puede estar en una fila inferior, es decir, si $v$ aparece en la fila correspondiente a $N(T^j)$ entonces $v \in N(T^j) \setminus N(T^{j-1})$.

**Teorema** (de la Descomposición Cíclica para operadores nilpotentes)**.** *Dado un $\mathbb{K}$-espacio vectorial $V$ con $\dim V = n$, sea $T \in L(V)$ un operador nilpotente de orden $k$. Entonces, existen $k_1, \ldots, k_p \in \mathbb{N}$ tales que*

$$k = k_1 \geq k_2 \geq \ldots \geq k_{p-1} \geq k_p \quad y \quad \sum_{i=1}^{p} k_i = n, \tag{5.9}$$

*y vectores $v_1, \ldots, v_p \in V$ tales que $T^{k_i} v_i = \vec{0}$ pero $T^{k_i - 1} v_i \neq \vec{0}$, de forma tal que*

$$V = Z(v_1, T) \dotplus Z(v_2, T) \dotplus \ldots \dotplus Z(v_p, T).$$

*Además, la partición $\{k_1, \ldots, k_p\}$ está unívocamente determinada por $T$.*

*Demostración.* Una vez más, haremos la demostración por inducción sobre la dimensión del espacio $V$. Si $\dim V = 1$, no hay nada que probar.

Como hipótesis inductiva, supongamos que el resultado vale para todo $\mathbb{K}$-espacio vectorial con dimensión menor a $n$.

Ahora, dado un $\mathbb{K}$-espacio vectorial con $\dim V = n$ y un operador nilpotente $T \in L(V)$ de orden $k$, consideremos el subespacio $N(T^{k-1})$ y sean $v_1, \ldots, v_m \in V$ tales que $\{[v_1], \ldots, [v_m]\}$ es una base de $\frac{V}{N(T^{k-1})} = \frac{N(T^k)}{N(T^{k-1})}$.

Si llamamos $W_k := Z(v_1, T) \dotplus \ldots \dotplus Z(v_m, T)$, el lema anterior garantiza la existencia de un subespacio $S$ de $V$ tal que

$$V = W_k \dotplus S \quad \text{y} \quad T(S) \subseteq S.$$

Como $\dim S < n$, podemos aplicar la hipótesis inductiva al operador $T|_S \in L(S)$ que es nilpotente de orden $l < k$. Luego, existen $l = l_1 \geq l_2 \geq \ldots \geq l_q$ tales que $\sum_{i=1}^{q} l_i = \dim S$, y vectores $w_1, \ldots, w_q \in S$ tales que

$$T^{l_i} w_i = \vec{0} \quad \text{pero} \quad T^{l_i - 1} w_i \neq \vec{0},$$

y $S = Z(w_1, T) \dotplus Z(w_2, T) \dotplus \ldots \dotplus Z(w_q, T)$. Entonces, si definimos

- $p := m + q$;
- $k_j := k$ si $j = 1, \ldots, m$, y $k_{m+i} := l_i$ para $i = 1, \ldots, q$;
- $v_{m+i} := w_i$ para $i = 1, \ldots, q$;

tenemos que $\sum_{i=1}^{p} k_i = mk + \sum_{j=1}^{q} k_{m+j} = mk + \sum_{j=1}^{q} l_j = \dim W_k + \dim S = n$ y

$$\begin{aligned} V &= W_k \dotplus S \\ &= \big(Z(v_1,T) \dotplus \ldots \dotplus Z(v_m,T)\big) \dotplus \big(Z(w_1,T) \dotplus Z(w_2,T) \dotplus \ldots \dotplus Z(w_q,T)\big) \\ &= Z(v_1,T) \dotplus \ldots \dotplus Z(v_m,T) \dotplus Z(v_{m+1},T) \dotplus \ldots \dotplus Z(v_{m+q},T) \\ &= Z(v_1,T) \dotplus \ldots \dotplus Z(v_p,T), \end{aligned}$$

probando la existencia de la descomposición cíclica. □

Dado un subespacio cíclico $Z(v,T)$ de dimensión $k$, a la base correspondiente

$$\{v, Tv, \ldots, T^{k-1}v\},$$

se la suele llamar **cadena de Jordan de longitud** $k$ (asociada al autovalor $\lambda = 0$ de $T$). En algunos textos, en lugar de hablar sobre subespacios cíclicos se refieren a las respectivas cadenas de Jordan.

Notemos que en (5.7) y (5.9) obtuvimos dos particiones distintas de $\dim V$. Estas dos particiones se denominan **particiones conjugadas** porque puede obtenerse una a partir de la otra de la siguiente forma: el $h$-ésimo término de la partición (5.7) $\big($es decir, $\dim \frac{N(T^h)}{N(T^{h-1})}\big)$ se obtiene contando la cantidad de términos de la partición (5.9) mayores o iguales a $j$ $\big($es decir, $\#\{i : \ k_i \geq j\}\big)$, y viceversa. La partición (5.7) suele llamarse **característica de Weyr** de $T$, mientras que la partición (5.9) se denomina **característica de Segre** del operador $T$.

Los diagramas con "puntitos" como el que encontramos antes del enunciado del Teorema de la Descomposición Cíclica (aunque está incompleto) se denominan **diagramas de Ferrers**, a partir de ellos es fácil convencerse de que la característica de Weyr y la característica de Segre son particiones conjugadas.

## 5.4. Forma canónica de Jordan

Ahora combinaremos el Teorema de la Descomposición Prima con el Teorema de la Descomposición Cíclica, para obtener una forma canónica para aquellos operadores lineales cuyo polinomio minimal tiene todas las raíces en el cuerpo $\mathbb{K}$.

Dado un $\mathbb{K}$-espacio vectorial $V$ de dimensión finita, supongamos que $T \in L(V)$ es tal que su polinomio característico puede factorizarse como

$$p_T(x) = (x-\lambda_1)^{d_1}(x-\lambda_2)^{d_2}\ldots(x-\lambda_k)^{d_k},$$

con $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{K}$ distintos entre sí y $d_1, \ldots, d_k \in \mathbb{N}$, mientras que el polinomio minimal se escribe como

$$m_T(x) = (x-\lambda_1)^{r_1}(x-\lambda_2)^{r_2}\ldots(x-\lambda_k)^{r_k},$$

con $r_1, \ldots, r_k \in \mathbb{N}$ tales que $1 \leq r_i \leq d_i$ para cada $i = 1, \ldots, k$.

Si consideramos los subespacios $W_i := N\big((T-\lambda_i I)^{r_i}\big)$, el Teorema 4.45 dice que

$$V = W_1 \dotplus W_2 \dotplus \ldots \dotplus W_k,$$

y además $T_i := T|_{W_i}$ tiene como polinomio minimal a $(x-\lambda_i)^{r_i}$. Luego, si definimos

$$N_i := T - \lambda_i I|_{W_i},$$

el operador $N_i$ es nilpotente de orden $r_i$ para cada $i = 1, \ldots, k$. Fijado un índice $i = 1, \ldots, k$, aplicando el Teorema de la Descomposición Cíclica a $N_i$, tenemos que existen naturales $r_i = r_1^{(i)} \geq r_2^{(i)} \geq \ldots \geq r_{p_i}^{(i)}$ y vectores $v_{1,i}, v_{2,i}, \ldots, v_{p_i,i} \in W_i$ tales que

$$W_i = Z(v_{1,i}, N_i) \dotplus Z(v_{2,i}, N_i) \dotplus \ldots \dotplus Z(v_{p_i,i}, N_i),$$

con $\dim Z(v_{j,i}, N_i) = r_j^{(i)}$. En particular, $N_i|_{Z(v_{j,i},N_i)}$ es nilpotente de orden $r_j^{(i)}$. Luego, si consideramos la base

$$\mathcal{B}_{j,i} = \left\{ v_{j,i}\ , N_i v_{j,i}\ , \ldots,\ N_i^{r_j^{(i)}-1} v_{j,i} \right\}$$

de $Z(v_{j,i}, N_i)$, resulta que

$$\begin{aligned}
\big[T|_{Z(v_{j,i},N_i)}\big]_{\mathcal{B}_{j,i}} &= \big[\lambda_i I|_{Z(v_{j,i},N_i)}\big]_{\mathcal{B}_{j,i}} + \big[(T-\lambda_i I)|_{Z(v_{j,i},N_i)}\big]_{\mathcal{B}_{j,i}} \\
&= \lambda_i\ I_{k_j^{(i)}} \quad + \quad J\big(0, r_j^{(i)}\big) \\
&= \begin{pmatrix} \lambda_i & 0 & \ldots & 0 & 0 \\ 0 & \lambda_i & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & \lambda_i & 0 \\ 0 & 0 & \ldots & 0 & \lambda_i \end{pmatrix} + \begin{pmatrix} 0 & 0 & \ldots & 0 & 0 \\ 1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & 0 \end{pmatrix},
\end{aligned}$$

Entonces,

$$\big[T|_{Z(v_{j,i},N_i)}\big]_{\mathcal{B}_{j,i}} = \begin{pmatrix} \lambda_i & 0 & \ldots & 0 & 0 \\ 1 & \lambda_i & \ldots & 0 & 0 \\ 0 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & \lambda_i \end{pmatrix} =: J\big(\lambda_i, r_j^{(i)}\big) \in \mathbb{K}^{r_j^{(i)} \times r_j^{(i)}}.$$

Luego, $\mathcal{B}_i := \{\mathcal{B}_{1,i}, \mathcal{B}_{2,i}, \ldots, \mathcal{B}_{r_i,i}\}$ es una base del subespacio $W_i$ tal que

$$[T_i]_{\mathcal{B}_i} = \begin{pmatrix} J\big(\lambda_i, r_1^{(i)}\big) & 0 & \ldots & 0 \\ 0 & J\big(\lambda_i, r_2^{(i)}\big) & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & J\big(\lambda_i, r_{p_i}^{(i)}\big) \end{pmatrix} =: A_i \in \mathbb{K}^{d_i \times d_i} \tag{5.10}$$

Finalmente, $\mathcal{B} := \{\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_k\}$ es una base de $V$ tal que

$$[T]_{\mathcal{B}} = \begin{pmatrix} [T_1]_{\mathcal{B}_1} & 0 & \ldots & 0 \\ 0 & [T_2]_{\mathcal{B}_2} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & [T_k]_{\mathcal{B}_k} \end{pmatrix} = \begin{pmatrix} A_1 & 0 & \ldots & 0 \\ 0 & A_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & A_k \end{pmatrix}, \tag{5.11}$$

la que se denomina **forma canónica de Jordan de** $T$. A cada uno de los bloques $A_1, \ldots, A_k$ se lo denomina **bloque de Jordan asociado al autovalor** $\lambda_i$, mientras que cada uno de los $J\big(\lambda_i, r_j^{(i)}\big)$ es un **bloque elemental de Jordan de tamaño** $r_j^{(i)}$ **asociado a** $\lambda_i$.

Notemos que, de acuerdo al Teorema de la Descomposición Cíclica, para cada $i = 1, \ldots, k$ se tiene que

$$r_1^{(i)} \geq r_2^{(i)} \geq \ldots \geq r_{p_i}^{(i)}.$$

Luego, en cada bloque de Jordan $A_i$, los bloques elementales $J\big(\lambda_i, r_j^{(i)}\big)$ están ordenados de manera decreciente (o mejor dicho, no creciente) con respecto a sus tamaños.

Ahora queremos mostrar que la representación matricial (5.11) está unívocamente determinada por $T$, salvo el orden en el que aparecen los bloques de Jordan $A_1, \ldots, A_k$.

Supongamos que existe una base $\mathcal{B}$ de $V$ tal que la representación matricial de $T$ con respecto a esa base es (5.11). Si $A_i \in \mathbb{K}^{d_i \times d_i}$ entonces $d_i$ es la multiplicidad algebraica de $\lambda_i$ como autovalor de $T$ (la multiplicidad de $\lambda_i$ como raíz del polinomio característico $p_T$). Como además $d_1 + d_2 + \ldots + d_k = \dim V$, resulta que

$$p_T(x) = (x - \lambda_1)^{d_1}(x - \lambda_2)^{d_2} \ldots (x - \lambda_k)^{d_k}.$$

Además, como (5.11) es una matriz diagonal por bloques, es evidente que

$$V = W_1 \dotplus W_2 \dotplus \ldots \dotplus W_k, \tag{5.12}$$

y cada uno de los $W_i$ es un subespacio $T$-invariante, donde $W_1$ es el subespacio generado por los primeros $k_1$ vectores de la base $\mathcal{B}$, $W_2$ es el subespacio generado por los siguientes $k_2$ vectores de $\mathcal{B}$, y así sucesivamente.

Por otra parte, como en cada $A_i$ el bloque elemental de Jordan más grande tiene tamaño $r_1^{(i)} = r_i$, es fácil ver que $(A_i - \lambda_i I)^{r_i}$ es la matriz nula de $\mathbb{K}^{d_i \times d_i}$. A partir de ésto, es fácil convencerse de que el polinomio minimal de $T$ es

$$m_T(x) = (x - \lambda_1)^{r_1}(x - \lambda_2)^{r_2} \ldots (x - \lambda_k)^{r_k}.$$

Para comprobar que (5.12) es la descomposición de $V$ dada por el Teorema de la Descomposición Prima, tenemos que mostrar que $W_i = N\big((T - \lambda_i I)^{r_i}\big)$. Pero esto es consecuencia de que $A_i - \lambda_i I$ es nilpotente de orden $r_i$ (como mencionamos en el párrafo anterior) mientras que $A_j - \lambda_i$ es inversible si $j \neq i$ (ya que es una matriz triangular inferior con entradas no nulas a lo largo de la diagonal principal). Por lo tanto, los subespacio $W_1, \ldots, W_k$ son únicos.

Por último, si $T_i = T|_{W_i}$ es la restricción de $T$ a $W_i$, es evidente que su único autovalor es $\lambda_i$. Luego, $N_i := T_i - \lambda_i I$ es nilpotente y, por el Teorema de la Descomposición Cíclica, existen únicos $k_1, k_2, \ldots, k_{q_i} \in \mathbb{N}$ tales que $k_1 \geq k_2 \geq \ldots \geq k_{q_i}$ y $\sum_{j=1}^{q_i} k_j = \dim W_i = d_i$ y vectores $v_{i,1}, v_{i,2}, \ldots, v_{i,p_i} \in W_i$ tales que

$$W_i = Z(v_{1,i}, N_i) \dotplus Z(v_{2,i}, N_i) \dotplus \ldots \dotplus Z(v_{p_i,i}, N_i).$$

Luego, comparando con el tamaño de los bloques elementales de $A_i$, resulta que $q_i = p_i$ y

$$k_j = r_j^{(i)}, \quad j = 1, \ldots, p_i,$$

es decir, la matriz $A_i$ está unívocamente determinada por la descomposición cíclica de $W_i$ asociada al nilpotente $N_i$.

**Observación 5.20.** Si $T$ es un operador diagonalizable con polinomio característico $p_T(x) = (x - \lambda_1)^{d_1}(x - \lambda_2)^{d_2} \ldots (x - \lambda_k)^{d_k}$, entonces la forma canónica de Jordan de $T$ coincide con la matriz diagonal:

$$D = \begin{pmatrix} \lambda_1 I_{d_1} & 0 & \ldots & 0 \\ 0 & \lambda_2 I_{d_2} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_k I_{d_k} \end{pmatrix}.$$

Desarrollemos un ejemplo para practicar el cálculo de la forma de Jordan de un operador lineal.

**Ejemplo 5.21.** Dado $T \in L(\mathbb{R}^3)$ tal que $A := [T]_{\mathcal{E}} = \begin{pmatrix} 3 & 1 & -1 \\ 2 & 2 & -1 \\ 2 & 2 & 0 \end{pmatrix}$, calculemos la forma de Jordan de $T$.

Comencemos calculando el polinomio característico de $T$:

$$\begin{aligned} p_T(x) = \det(xI - A) &= \det \begin{pmatrix} x-3 & -1 & 1 \\ -2 & x-2 & 1 \\ -2 & -2 & x \end{pmatrix} \\ &= (x-3)(x-2)x + 2 + 4 + 2(x-2) + 2(x-3) - 2x \\ &= (x-3)(x-2)x + 2(x-2) = (x-2)(x^2-3x+2) = (x-2)^2(x-1). \end{aligned}$$

Luego, los posibles polinomios minimales de $T$ son $m_1(x) = (x-2)(x-1)$ y $m_2(x) = p_T(x)$. Pero como

$$(A-2I)(A-I) = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 0 & -1 \\ 2 & 2 & -2 \end{pmatrix} \begin{pmatrix} 2 & 1 & -1 \\ 2 & 1 & -1 \\ 2 & 2 & -1 \end{pmatrix} = \begin{pmatrix} 2 & 0 & -1 \\ 2 & 0 & -1 \\ 4 & 0 & -2 \end{pmatrix} \neq 0,$$

tenemos que $m_T(x) = p_T(x) = (x-2)^2(x-1)$. Entonces, $T$ no es diagonalizable.

Calculemos ahora los subespacios correspondientes a la descomposición prima, $W_1 = N\big((T-2I)^2\big)$ y $W_2 = N(T-I)$. Comencemos por $W_1 = N\big((T-2I)^2\big)$:

$$(A-2I)^2 = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 0 & -1 \\ 2 & 2 & -2 \end{pmatrix} \begin{pmatrix} 1 & 1 & -1 \\ 2 & 0 & -1 \\ 2 & 2 & -2 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \\ 2 & -2 & 0 \end{pmatrix}$$

Luego, $W_1 = N\big((T-2I)^2\big) = \overline{\{(1,1,0),(0,0,1)\}}$, pero necesitamos una base cíclica de $W_1$. Notemos que si $b_1 = (1,1,1)$ entonces

$$\begin{aligned} &(T-2I)b_1 = (1,1,2) =: b_2 \neq \vec{0}, \\ &(T-2I)b_2 = \vec{0}. \end{aligned}$$

Por otro lado, $W_2 = N(T-I) = \overline{\{(1,0,2)\}}$. Llamando $b_3 := (1,0,2)$, tenemos que $\mathcal{B} = \{b_1, b_2, b_3\}$ es una base de $\mathbb{R}^3$ tal que

$$\begin{aligned} &Tb_1 = (3,3,4) = 2\ (1,1,1) + (1,1,2) = 2\ b_1 + b_2, \\ &Tb_2 = 2\ b_2, \\ &Tb_3 = b_3, \end{aligned}$$

con lo cual $[T]_{\mathcal{B}} = \left(\begin{array}{cc|c} 2 & 0 & 0 \\ 1 & 2 & 0 \\ \hline 0 & 0 & 1 \end{array}\right)$, es la forma canónica de Jordan de $T$.

A continuación recopilaremos algunas características que tiene la forma canónica de Jordan, y que son datos útiles al momento de calcularla. Si $\dim V = n$ y la forma de Jordan de $T$ es

$$[T]_{\mathcal{B}} = \begin{pmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & A_k \end{pmatrix},$$

entonces:

- Cada bloque de Jordan $A_i$ es de tamaño $d_i \times d_i$, donde
$$d_i = \dim W_i = \dim N\big((T-\lambda_i I)^n\big),$$
es la multiplicidad algebraica del autovalor $\lambda_i$.
- Cada $A_i$ es una matriz diagonal por bloques, con $p_i$ matrices elementales de Jordan $J\big(\lambda_i, k_j^{(i)}\big)$ a lo largo de la diagonal principal, siendo
$$p_i = \dim N(T-\lambda_i I).$$
- Los bloques elementales de Jordan en $A_i$ están dispuestos en orden decreciente con respecto a su tamaño, y el primer bloque es de tamaño $r_i \times r_i$, siendo $r_i$ la multiplicidad del autovalor $\lambda_i$ como raíz del polinomio minimal.
- La suma de los tamaños de los bloques elementales $J\big(\lambda_i, k_j^{(i)}\big)$ coincide con $d_i$, el tamaño de $A_i$.

Estos datos no siempre son suficientes para calcular la forma de Jordan de un operador lineal, pero nos permiten generar una lista de posibles formas de Jordan.

**Ejemplo 5.22.** Sea $T \in L(\mathbb{C}^{10})$ tal que sus polinomios característico y minimal son, respectivamente,
$$p_T(x) = (x-2)^7(x-1)^2(x+1) \quad \text{y} \quad m_T(x) = (x-2)^4(x-1)(x+1).$$
Intentemos calcular la forma canónica de Jordan de $T$ con la información que tenemos disponible. Utilizando los datos que recopilamos anteriormente, vemos que la forma de Jordan de $T$ debe ser
$$A = \begin{pmatrix} 2 & 0 & 0 & 0 & & & & & & \\ 1 & 2 & 0 & 0 & & & & & & \\ 0 & 1 & 2 & 0 & & & & & & \\ 0 & 0 & 1 & 2 & & & & & & \\ & & & & 2 & 0 & 0 & & & \\ & & & & ? & 2 & 0 & & & \\ & & & & ? & ? & 2 & & & \\ & & & & & & & 1 & & \\ & & & & & & & & 1 & \\ & & & & & & & & & -1 \end{pmatrix},$$
pero no podemos determinar completamente cual es el aspecto del bloque de $3 \times 3$. Hay tres situaciones posibles, de acuerdo a cual sea la dimensión del autoespacio asociado al autovalor $\lambda = 2$:
$$\begin{pmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 1 & 2 \end{pmatrix} \qquad \text{si } \dim N(T-2I) = 2;$$
$$\begin{pmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \qquad \text{si } \dim N(T-2I) = 3;$$
$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \qquad \text{si } \dim N(T-2I) = 4.$$

**Ejemplo 5.23.** Sea $D$ el operador de derivación en el espacio $\mathbb{R}^{(3)}[x]$ de polinomios de grado menor o igual que 3. Si

$$\mathcal{E} = \{p_0(x) = 1, p_1(x) = x, p_2(x) = x^2, p_3(x) = x^3\},$$

en el Ejemplo 2.27 vimos que

$$[D]_{\mathcal{E}} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

¿Cuál es la forma canónica de Jordan de $D$?

Es fácil ver que $p_D(x) = m_D(x) = x^4$, con lo cual existe un vector en el espacio $\mathbb{R}^{(3)}[x]$ que es cíclico para $D$. Además, sabemos que

$$\begin{aligned} D(x^3) &= 3x^2, \\ D(3x^2) &= 6x, \\ D(6x) &= 6, \\ D(6) &= 0, \end{aligned}$$

Entonces, si consideramos la base

$$\mathcal{B} := \{q_0(x) = x^3, q_1(x) = 3x^2, q_2(x) = 6x, q_3(x) = 6\},$$

resulta que

$$[D]_{\mathcal{B}} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix},$$

es la forma de Jordan de $D$.

## 5.5. Ejercicios

1. Sea $W$ un subespacio de un $\mathbb{K}$-espacio vectorial $V$. Probar que:

    *a*) la relación de congruencia módulo $W$ es una relación de equivalencia;
    
    *b*) la proyección al cociente $\pi : V \to V/W$ es un epimosfismo.

2. Dado un $\mathbb{K}$-espacio vectorial $V$ de dimensión finita, sea $T \in L(V)$ un operador nilpotente de orden $k$. Probar que $\dim N(T) \geq \dim \frac{N(T^2)}{N(T)} \geq \ldots \geq \dim \frac{N(T^k)}{N(T^{k-1})}$.

3. Consideremos los operadores lineales $R_{\frac{\pi}{2}}$, $S_Y$, $H_2$ y $P_X$ introducidos en la sección de Ejercicios del Capítulo 2. Sea $\mathcal{E} = \{e_1, e_2\}$ la base canónica de $\mathbb{R}^2$.

    *a*) Probar que $Z(e_1, R_{\frac{\pi}{2}}) = \mathbb{R}^2 = Z(e_2, R_{\frac{\pi}{2}})$.
    
    *b*) Hallar $Z(e_1, S_Y)$ y $Z(e_2, S_Y)$.
    
    *c*) Probar que $H_2$ no tiene vectores cíclicos.
    
    *d*) Hallar $Z(e_1, P_X)$ y $Z(e_2, P_X)$. ¿Puede dar algún vector cíclico de $P_X$?

4. Sea $V$ un $\mathbb{K}$-espacio vectorial de dimensión finita. Dado $T \in L(V)$, probar que $T$ tiene un vector cíclico si y sólo si $m_T = p_T$.

5. Dada $A \in \mathbb{K}^{n\times n}$ fija, sea $T \in L(\mathbb{K}^{n\times n})$ dado por
$$T(B) = AB - BA\,.$$
   *a*) Probar que si $A$ y $B$ se diagonalizan simultáneamente, $T(B) = 0$.
   *b*) Probar que si $A$ y $B$ son nilpotentes y $T(B) = 0$ entonces $AB$ es nilpotente. Hallar un ejemplo en que $A$ y $B$ sean nilpotentes, pero $AB$ no lo sea.
   *c*) Probar que si $A$ es una matriz nilpotente de orden 2, entonces $T$ es un operador nilpotente de orden 3.
   *d*) Si $A = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$, hallar $Z(E_{12}, T)$ y el polinomio $T$-anulador de $E_{12}$. Verificar que el $T$-anulador de $E_{12}$ divide al polinomio minimal de $T$. ¿Puede hallar algún vector $T$-cíclico?

6. Sea $V$ un $\mathbb{K}$-espacio vectorial y $T \in L(V)$. ¿Es cierto que $v \in V$ es un autovector de $T$ asociado al autovalor $\lambda$ si, y sólo si, $p(x) = x - \lambda$ es el polinomio $T$-anulador de $v$?

7. Sea $T \in L(\mathbb{R}^5)$ cuya representación matricial en la base canónica de $\mathbb{R}^5$ es
$$A = \begin{pmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix}.$$
   *a*) Probar que $T$ no tiene ningún vector cíclico.
   *b*) Hallar el subespacio cíclico generado por $(1, 1, -1, -1, -1)$.

8. Sea $T \in L(\mathbb{C}^3)$ tal que
$$A = \begin{pmatrix} 3 & 1 & -2 \\ -1 & 0 & 5 \\ -1 & -1 & 4 \end{pmatrix},$$
es una representación matricial de $T$. Hallar la forma de Jordan de $T$ y una base $\mathcal{B}$ de $\mathbb{C}^3$ tal que $[T]_{\mathcal{B}}$ es una matriz de Jordan.

9. Sea $T \in L(\mathbb{R}^5)$ tal que
$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$
la representación matricial de $T$ en la base canónica de $\mathbb{R}^5$.
   *a*) ¿Es T diagonalizable?
   *b*) Hallar, en caso de que exista, una matriz de Jordan semejante a $A$ y la matriz de cambio de base correspondiente.

*c*) Probar que existe un operador diagonalizable $D \in L(\mathbb{R}^5)$ y un operador nilpotente $N \in L(\mathbb{R}^5)$, tales que $T = D + N$ y $DN = ND$. Escribir las representaciones matriciales de $D$ y $N$ en la base canónica de $\mathbb{R}^5$.

10. Sea $T \in L(\mathbb{R}^3)$ y sea $\mathcal{B}$ una base en la cual
$$[T]_{\mathcal{B}} = \begin{pmatrix} -2 & 0 & 0 \\ 1 & -2 & 0 \\ 0 & 1 & -2 \end{pmatrix}.$$
Determinar, observando la matriz, el polinomio minimal y el característico de $T$.

11. Hallar dos endomorfismos de $\mathbb{R}^3$ diferentes, que tengan como forma de Jordan a la matriz
$$A = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & -3 \end{pmatrix},$$

12. Encontrar la forma de Jordan de los operadores representados por las siguientes matrices
$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad \text{y} \quad B = \begin{pmatrix} 2 & -1 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & -1 & 2 \end{pmatrix}$$

13. Sean
$$A = \begin{pmatrix} 1 & -1 \\ 2 & 4 \end{pmatrix} \quad \text{y} \quad B = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$
Determinar si ambas matrices tienen la misma forma de Jordan, sin calcularla explícitamente.

14. ¿Cuáles son las posibles formas de Jordan de un operador lineal $T \in L(\mathbb{C}^5)$ si $p_T(x) = (x+4)^2(x-\pi)^2(x-i)$?

15. Sea $T \in L(\mathbb{C}^3)$ tal que
$$A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 2 \\ 1 & 2 & 0 \end{pmatrix},$$
es una representación matricial de $T$. Hallar la forma de Jordan de $T$ y una base $\mathcal{B}$ de $\mathbb{C}^3$ tal que $[T]_{\mathcal{B}}$ es una matriz de Jordan.

16. En cada caso, hallar todas las posibles formas de Jordan para una matriz $A$ con los siguientes polinomios minimales y característicos.

    *a*) $p_A(x) = (x-5)^4$ y $m_A(x) = (x-5)^2$.
    
    *b*) $p_A(x) = (x-2)^2(x-4)^2$ y $m_A(x) = p_A(x)$.
    
    *c*) $p_A(x) = (x+3)^4(x-2)^4$ y $m_A(x) = (x+3)^2(x-2)^2$.

17. Sea

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

   *a*) Hallar la forma de Jordan asociada a $A$.

   *b*) Hallar el polinomio característico, el minimal, los autovalores y la dimensión de los autoespacios asociados.

18. Hallar las posibles formas de Jordan de un operador $T \in L(\mathbb{C}^9)$ tal que

$$p_T(x) = (x+7)^4(x-3)^3(x+i)^2 \quad \text{y} \quad m_T(x) = (x+7)^2(x-3)^2(x+i)\,.$$

   Mencionar en cada caso la dimensión de los autoespacios asociados.

# Capítulo 6

# Funcionales lineales: el espacio dual

En este capítulo estudiaremos las transformaciones lineales que tienen como codominio al cuerpo $\mathbb{K}$.

**Definición 6.1.** Dado un $\mathbb{K}$-espacio vectorial $V$, un **funcional lineal** sobre $V$ es una transformación lineal de $V$ en $\mathbb{K}$, es decir, una función $f : V \to \mathbb{K}$ tal que

$$f(\alpha u + v) = \alpha\ f(u) + f(v) \qquad u, v \in V,\ \alpha \in \mathbb{K}.$$

Los siguientes son algunos ejemplos de funcionales lineales.

**Ejemplos 6.2.**

1. Dados $\alpha_1, \dots, \alpha_n \in \mathbb{K}$, sea $f : \mathbb{K}^n \to \mathbb{K}$ dada por

   $$f(x_1, \dots, x_n) = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n, \qquad (x_1, \dots, x_n) \in \mathbb{K}^n. \tag{6.1}$$

   Este $f$ es un funcional lineal sobre $\mathbb{K}^n$. Además, si $\mathcal{E} = \{e_1, e_2, \dots, e_n\}$ es la base canónica de $\mathbb{K}^n$, notemos que

   $$f(e_j) = f(0, \dots, 0, \overset{\overset{j}{\downarrow}}{1}, 0, \dots, 0) = \alpha_j, \quad j = 1, \dots, n.$$

   Recíprocamente, dado un funcional lineal $f$ sobre $\mathbb{K}^n$ cualquiera, tenemos que

   $$f(x_1, \dots, x_n) = f\left(\sum_{i=1}^{n} x_i \cdot e_i\right) = \sum_{i=1}^{n} x_i f(e_i).$$

   Luego, si para cada $j = 1, \dots, n$ definimos $\alpha_j := f(e_j)$, vemos que $f$ tiene el mismo aspecto que (6.1).

2. La traza $\operatorname{tr} : \mathbb{K}^{m\times m} \to \mathbb{K}$, definida por

   $$\operatorname{tr}(A) = \sum_{i=1}^{m} a_{ii}, \qquad \text{si } A = (a_{ij}) \in \mathbb{K}^{m\times m},$$

   es un funcional lineal sobre $\mathbb{K}^{m\times m}$.

3. Si $C([a,b])$ es el $\mathbb{R}$-espacio vectorial de funciones continuas sobre el intervalo cerrado $[a,b]$, sea $L : C([a,b]) \to \mathbb{R}$ dada por
$$L(g) = \int_a^b g(t)dt, \qquad g \in C([a,b]).$$
La integral definida $L$ es un funcional lineal sobre $C([a,b])$.

**Definición 6.3.** Si $V$ es un $\mathbb{K}$-espacio vectorial, llamaremos **espacio dual de** $V$ al $\mathbb{K}$-espacio vectorial de funcionales lineales sobre $V$, y lo anotaremos $V^*$:
$$V^* = L(V, \mathbb{K}).$$

Si $V$ es un $\mathbb{K}$-espacio vectorial de dimensión finita, entonces,
$$\dim V^* = \dim L(V, \mathbb{K}) = \dim V \cdot \dim \mathbb{K} = \dim V.$$
Por lo tanto, $V^*$ es isomorfo a $V$. Además, cada base de $V$ determina este isomorfismo de una forma muy particular.

**Teorema 6.4.** *Dado un $\mathbb{K}$-espacio vectorial $V$ con $\dim V = n$, sea $\mathcal{B} = \{b_1, \ldots, b_n\}$ una base de $V$. Entonces, existe una única base $\mathcal{B}^* = \{f_1, \ldots, f_n\}$ de $V^*$ tal que*
$$f_i(b_j) = \delta_{ij} = \begin{cases} 1 & si \;\; i = j, \\ 0 & si \;\; i \neq j. \end{cases} \tag{6.2}$$

*Demostración.* La existencia de los (únicos) funcionales $f_1, \ldots, f_n$ que satisfacen las condiciones (6.2) es consecuencia del Teorema 2.3, el cual asegura que una transformación lineal queda determinada por los valores que toma sobre los vectores de una base.

Veamos ahora que $\mathcal{B}^* = \{f_1, \ldots, f_n\}$ es linealmente independiente: supongamos que existen $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$ tales que $\sum_{i=1}^n \alpha_i f_i = \vec{0}_{V^*}$. Luego, para cada $j = 1, \ldots, n$,
$$0 = \vec{0}_{V^*}(b_j) = \left(\sum_{i=1}^n \alpha_i f_i\right) b_j = \sum_{i=1}^n \alpha_i f_i(b_j) = \sum_{i=1}^n \alpha_i \delta_{ij} = \alpha_j.$$
Como además $\dim V^* = \dim V = n$, tenemos que $\mathcal{B}^*$ es una base de $V^*$. □

La base $\mathcal{B}^* = \{f_1, \ldots, f_n\}$ que cumple $f_i(b_j) = \delta_{ij}$ se denomina **base dual** de la base $\mathcal{B} = \{b_1, \ldots, b_n\}$. Una explicación posible para tal nombre son las siguientes fórmulas:
$$f = \sum_{i=1}^n f(b_i) f_i \qquad \text{para cada } f \in V^*, \quad \text{y} \tag{6.3}$$
$$v = \sum_{j=1}^n f_j(v) b_j \qquad \text{para cada } v \in V, \tag{6.4}$$
las cuales muestran que $\mathcal{B}$ y $\mathcal{B}^*$ tienen roles duales, en el sentido de que usando la base $\mathcal{B}^*$ podemos calcular los coeficientes para representar a cada vector $v \in V$ con respecto a la base $\mathcal{B}$, y viceversa.

Para comprobar (6.4), fijado un vector $v \in V$, sabemos que existen únicos $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$ tales que $v = \sum_{i=1}^n \alpha_i b_i$. Además, para cada $j = 1, \ldots, n$, resulta que
$$f_j(v) = f_j\left(\sum_{i=1}^n \alpha_i b_i\right) = \sum_{i=1}^n \alpha_i f_j(b_i) = \sum_{i=1}^n \alpha_i \delta_{ij} = \alpha_j.$$

Por lo tanto, $v = \sum_{j=1}^{n} f_j(v)b_j$, como queríamos probar. Esto dice además que $f_1, \ldots, f_n$ son justamente los **funcionales coordenados** con respecto a la base $\mathcal{B}$.

Análogamente, fijado un funcional $f \in V^*$, existen únicos $\beta_1, \ldots, \beta_n \in \mathbb{K}$ tales que $f = \sum_{j=1}^{n} \beta_j f_j$. En este caso, para cada $i = 1, \ldots, n$, tenemos que

$$f(b_i) = \left( \sum_{j=1}^{n} \beta_j f_j \right) b_i = \sum_{j=1}^{n} \beta_j f_j(b_i) = \sum_{j=1}^{n} \beta_j \delta_{ji} = \beta_i,$$

lo que prueba (6.3), generalizando el primero de los Ejemplos 6.2.

**Ejemplo 6.5.** En el $\mathbb{K}$-espacio vectorial $\mathbb{K}^{(n)}[x]$ de polinomios de grado menor o igual a $n$, consideremos los funcionales $L_0, L_1, \ldots, L_n$ determinados por $n+1$ escalares $t_0, t_1, \ldots, t_n \in \mathbb{K}$, distintos entre sí, mediante

$$L_i(p) = p(t_i), \qquad i = 0, 1, \ldots, n.$$

Gracias a la matriz de Vandermonde es fácil convencerse de que $\{L_0, L_1, \ldots, L_n\}$ es una base de $(\mathbb{K}^{(n)}[x])^*$, con lo cual es natural preguntarse:

¿Cuál es la base de $\mathbb{K}^{(n)}[x]$ cuya base dual coincide con $\{L_0, L_1, \ldots, L_n\}$?

La respuesta es justamente la base formada por los polinomios interpoladores de Lagrange $\mathcal{L} = \{p_0, p_1, \ldots, p_n\}$, ya que éstos estaban definidos a partir de las condiciones $p_i(t_j) = \delta_{ij}$, es decir,

$$L_j(p_i) = p_i(t_j) = \delta_{ji}.$$

Estudiaremos ahora como describir a los subespacios de un $\mathbb{K}$-espacio vectorial dado a partir de los funcionales lineales definidos sobre el mismo. En primer lugar, notemos que si $V$ es un $\mathbb{K}$-espacio vectorial con $\dim V = n$ y $f \in V^*$ es un funcional no nulo, entonces existe un vector $v_0 \in V$ tal que

$$f(v_0) \neq 0.$$

Luego, multiplicando a $v_0$ por los escalares en $\mathbb{K}$ es fácil convencerse de que $\mathrm{Im}(f) = \mathbb{K}$, con lo cual $\dim(\mathrm{Im}\, f) = 1$. Entonces,

$$\dim N(f) = \dim V - \dim(\mathrm{Im}\, f) = n - 1.$$

**Definición 6.6.** Dado un $\mathbb{K}$-espacio vectorial $V$ con $\dim V = n$, diremos que un subespacio $W$ de $V$ es un **hiperespacio** si

$$\dim W = n - 1.$$

Lo que probamos antes de la definición, es que cada funcional lineal (no nulo) $f$ sobre $V$ determina un hiperespacio, el cual es $N(f)$. Recíprocamente, si $W$ es un hiperespacio de $V$ es fácil construir un funcional lineal no nulo $f$ tal que $N(f) = W$, esto último queda como ejercicio para el lector (sale usando una vez más el Teorema 2.3).

**Definición 6.7.** Si $S$ es un subconjunto de un $\mathbb{K}$-espacio vectorial $V$, el **anulador de** $S$ se define como

$$S^{\circ} = \{f \in V^* : \ f(s) = 0 \text{ para todo } s \in S\}.$$

**Ejemplo 6.8.** Consideremos en $\mathbb{R}^5$ el conjunto

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

y calculemos el anulador de $S$.

En general, un funcional sobre $\mathbb{R}^5$ es de la forma:

$$f(v,w,x,y,z) = \alpha v + \beta w + \gamma x + \delta y + \varepsilon z,$$

con $\alpha,\beta,\gamma,\delta,\varepsilon \in \mathbb{R}$. Ahora, si $f \in S^\circ$ entonces $f$ debe verificar:

$$\left\{\begin{array}{rcl} f(2,-2,3,4,-1) & = & 0 \\ f(-1,1,2,5,2) & = & 0 \\ f(0,0,-1,-2,3) & = & 0 \\ f(1,-1,2,3,0) & = & 0 \end{array}\right. \quad \Longrightarrow \quad \left\{\begin{array}{rcl} 2\alpha + (-2)\beta + 3\gamma + 4\delta + (-1)\varepsilon & = & 0 \\ (-1)\alpha + \beta + 2\gamma + 5\delta + 2\varepsilon & = & 0 \\ (-1)\gamma + (-2)\delta + 3\varepsilon & = & 0 \\ \alpha + (-1)\beta + 2\gamma + 3\delta & = & 0 \end{array}\right.$$

Sumando la segunda ecuación con la cuarta, resulta que $4\gamma + 8\delta + 2\varepsilon = 0$. Mientras que si multiplicamos a la tercera por $(-4)$, obtenemos $4\gamma + 8\delta - 12\varepsilon = 0$. Al restar estas dos últimas, concluímos que

$$\varepsilon = 0 \qquad \text{y} \qquad \gamma + 2\delta = 0.$$

Por otra parte, al reemplazar estas condiciones en la primer ecuación, resulta $2\alpha - 2\beta + \gamma = 0$. En conclusión, $\alpha,\beta,\gamma,\delta$ y $\varepsilon$ deben verificar:

$$\left\{\begin{array}{rcl} \varepsilon & = & 0 \\ \gamma + 2\delta & = & 0 \\ 2\alpha - 2\beta + \gamma & = & 0 \end{array}\right.$$

Entonces $\varepsilon = 0$ y, fijando los valores de dos de las variables, las restantes quedan unívocamente determinadas. Por ejemplo, si $\alpha = 1$ y $\beta = 0$ entonces $\gamma = -2$, $\delta = 1$ y el funcional

$$f_1(v,w,x,y,z) = v - 2x + y$$

está en el anulador de $S$.

En cambio, si $\alpha = 0$ y $\beta = 1$ entonces $\gamma = 2$, $\delta = -1$ y el funcional

$$f_2(v,w,x,y,z) = w + 2x - y$$

está en el anulador de $S$. Por lo tanto,

$$S^\circ = \overline{\{f_1, f_2\}} = \overline{\{f_1(v,w,x,y,z) = v - 2x + y, f_2(v,w,x,y,z) = w + 2x - y\}}.$$

**Ejercicio 6.9.** Probar que, para cualquier subconjunto $S$ de $V$:

1. el anulador $S^\circ$ es un subespacio de $V^*$;
2. si $\overline{S}$ es el subespacio generado por $S$, entonces $\left(\overline{S}\right)^\circ = S^\circ$.

Verificar además que $\{\vec{0}_V\}^\circ = V^*$ y $V^\circ = \{\vec{0}_{V^*}\}$.

**Ejercicio 6.10.** Probar que, si $S_1$ y $S_2$ son dos subconjuntos de un $\mathbb{K}$-espacio vectorial $V$, $S_1 \subseteq S_2$ implica que $S_2^\circ \subseteq S_1^\circ$.

**Teorema 6.11.** *Dado un $\mathbb{K}$-espacio vectorial $V$ de dimensión finita, sea $W$ un subespacio de $V$. Entonces,*

$$\dim V = \dim W + \dim W^\circ. \tag{6.5}$$

*Demostración.* Supongamos que $\dim V = n$ y que $\dim W = k$. Fijada una base $\{b_1, \ldots, b_k\}$ de $W$ extendámosla a una base de $V$, es decir, sean $b_{k+1}, \ldots, b_n \in V$ tales que $\mathcal{B} = \{b_1, \ldots, b_k, b_{k+1}, \ldots, b_n\}$. Si $\mathcal{B}^* = \{f_1, \ldots f_n\}$ es la base dual de $\mathcal{B}$, veamos que $\{f_{k+1}, \ldots, f_n\}$ es una base de $W^\circ$.

En primer lugar, como $f_i(b_j) = \delta_{ij}$, es fácil ver que los funcionales $f_{k+1}, \ldots, f_n$ pertenecen a $W^\circ$. Veamos ahora que $\{f_{k+1}, \ldots f_n\}$ es un conjunto de generadores de $W^\circ$. Como $\mathcal{B}^*$ es la base dual de la base $\mathcal{B}$, a cada $f \in V^*$ podemos descomponerlo como

$$f = \sum_{i=1}^{n} f(b_i) f_i.$$

Pero si $f \in W^\circ$ entonces $f(b_i) = 0$ para $i = 1, \ldots, k$. Luego,

$$f = \sum_{i=k+1}^{n} f(b_i) f_i,$$

es decir, $f \in \overline{\{f_{k+1}, \ldots, f_n\}}$. Por lo tanto, $W^\circ = \overline{\{f_{k+1}, \ldots f_n\}}$ y, en particular,

$$\dim W^\circ = n - k = \dim V - \dim W. \qquad \square$$

**Corolario 6.12.** *Dado un $\mathbb{K}$-espacio vectorial $V$ con $\dim V = n$, sea $W$ un subespacio de $V$ con $\dim W = k$. Entonces, existen $g_1, \ldots, g_{n-k} \in V^*$ tales que*

$$W = \bigcap_{i=1}^{n-k} N(g_i).$$

*Demostración.* Es una consecuencia de la prueba del teorema anterior, queda como ejercicio para el lector. $\square$

**Proposición 6.13.** *Dado un $\mathbb{K}$-espacio vectorial $V$ de dimensión finita, sean $W_1$ y $W_2$ dos subespacios de $V$. Entonces,*

$$W_1 = W_2 \quad \textit{si y sólo si} \quad W_1^\circ = W_2^\circ.$$

*Demostración.* Si $W_1 = W_2$, la igualdad entre los anuladores es una consecuencia del ejercicio anterior.

La otra implicación sale por la contrarrecíproca. Supongamos que $W_1 \neq W_2$, y para fijar ideas supongamos que existe un vector $w \in W_1 \setminus W_2$ (si esto no fuera cierto, intercambiamos los roles de $W_1$ y $W_2$).

Como $w \notin W_2$, existe un $f \in W_2^\circ$ tal que $f(w) \neq 0$. Entonces, $f \in W_2^\circ \setminus W_1^\circ$ y los anuladores también son diferentes. $\square$

**Ejemplo 6.14.** Consideremos los funcionales $f_1, f_2, f_3 \in (\mathbb{R}^4)^*$ dados por

$$\begin{aligned} f_1(w, x, y, z) &= w + 2x + 2y + z, \\ f_2(w, x, y, z) &= 2x + z, \\ f_3(w, x, y, z) &= -2w - 4y + 3z. \end{aligned}$$

El subespacio $W$ que es anulado por estos funcionales puede calcularse explícitamente, es el subespacio de soluciones del sistema de ecuaciones lineales homogéneo:

$$\begin{cases} w + 2x + 2y + z = 0 \\ 2x + z = 0 \\ -2w - 4y + 3z = 0 \end{cases} \Longrightarrow \begin{cases} w + 2y = 0 \\ 2x + z = 0 \\ -2(w + 2y) + 3z = 0 \end{cases},$$

es decir, $(w, x, y, z) \in W$ si y sólo si $w + 2y = 0$ y $x = z = 0$. Por lo tanto,

$$W = \overline{\{(-2, 0, 1, 0)\}}.$$

## 6.1. El doble dual

Dado un $\mathbb{K}$-espacio vectorial $V$ de dimensión finita, vimos que toda base $\mathcal{B}$ de $V$ tiene asociada su correspondiente base dual $\mathcal{B}^*$ en $V^*$. Ahora, podríamos hacernos otra pregunta: si $\mathcal{B}'$ es una base de $V^*$, ¿existe una base $\mathcal{B}$ de $V$ cuya base dual sea $\mathcal{B}'$?

Una forma de encarar esta pregunta es considerar el espacio dual de $V^*$, al que llamaremos el **doble dual de** $V$ y lo anotaremos $V^{**}$. Aplicando el Teorema 6.4, la base dual $(\mathcal{B}')^*$ es una base de $V^{**}$. Además, tenemos que $V^{**}$ es isomorfo a $V^*$, que a su vez es isomorfo a $V$. Sin embargo, hasta ahora la única forma de encontrar tales isomorfismos requiere fijar una base del espacio (y utilizar el Teorema 6.4). A continuación veremos que hay un isomorfismo entre $V$ y $V^{**}$ el cual no requiere prefijar una base, y por lo tanto se lo considera más natural.

Cada vector $v \in V$ induce un funcional lineal sobre $V^*$ de la siguiente manera: dado $v \in V$, sea $L_v : V^* \to \mathbb{K}$ definida por

$$L_v(f) := f(v), \quad f \in V^*.$$

Esta función está bien definida y además, si $f, g \in V^*$ y $\alpha \in \mathbb{K}$,

$$L_v(\alpha f + g) = (\alpha f + g)(v) = \alpha f(v) + g(v) = \alpha L_v(f) + L_v(g).$$

Por lo tanto, $L_v \in (V^*)^* = V^{**}$.

**Observación 6.15.** Si $v \in V$, $v \neq \vec{0}_V$ entonces $L_v \neq \vec{0}_{V^{**}}$. De hecho, si tomamos una base $\mathcal{B} = \{b_1, \ldots, b_n\}$ de $V$ tal que $b_1 = v$ y consideramos su base dual $\mathcal{B}^* = \{f_1, \ldots . f_n\}$, resulta que

$$L_v(f_1) = f_1(v) = f_1(b_1) = 1 \neq 0.$$

**Teorema 6.16.** *Sea $V$ un $\mathbb{K}$-espacio vectorial de dimensión finita. Luego, la aplicación*

$$\Phi : V \to V^{**} \quad \textit{dada por} \quad \Phi(v) = L_v,$$

*es un isomorfismo de $V$ en $V^{**}$.*

*Demostración.* La función $\Phi$ está bien definida, ya que para cada $v \in V$ tenemos que $\Phi(v) = L_v \in V^{**}$. Veamos ahora que $\Phi$ es lineal. Dados $v, w \in V$ y $\alpha \in \mathbb{K}$, para cada $f \in V^*$,

$$L_{\alpha v + w}(f) = f(\alpha v + w) = \alpha f(v) + f(w) = \alpha L_v(f) + L_w(f) = (\alpha L_v + L_w)(f),$$

es decir, $\Phi(\alpha v + w) = \alpha\Phi(v) + \Phi(w)$. Entonces, $\Phi$ es una transformación lineal de $V$ en $V^{**}$.

Además, por la observación anterior, si $\Phi(v) = \vec{0}_{V^{**}}$ entonces $v = \vec{0}_V$. Esto dice que $\Phi$ es un monomorfismo. Por último, como ya sabemos que $\dim V^{**} = \dim V$, resulta que $\Phi$ es un isomorfismo. □

**Corolario 6.17.** *Sea $V$ un $\mathbb{K}$-espacio vectorial de dimensión finita. Si $L \in V^{**}$ entonces existe un único $v \in V$ tal que $L = L_v$, es decir,*

$$L(f) = L_v(f) = f(v) \quad \textit{para todo } f \in V^*.$$

Ahora sí estamos en condiciones de mostrar que la respuesta a la pregunta con la que iniciamos la sección es afirmativa.

**Corolario 6.18.** *Si $V$ es un $\mathbb{K}$-espacio vectorial de dimensión finita entonces toda base de $V^*$ es la base dual de alguna base de $V$.*

*Demostración.* Supongamos que $\dim V = n$ y sea $\mathcal{B}' = \{f_1, \ldots, f_n\}$ una base de $V^*$. Su base dual $(\mathcal{B}')^* = \{L_1, \ldots, L_n\}$ es la única base de $V^{**}$ tal que $L_i(f_j) = \delta_{ij}$.

Por el corolario anterior, existen únicos vectores $v_1, \ldots, v_n \in V$ tales que

$$L_i(f) = f(v_i) \quad \text{para todo } f \in V^*.$$

Además, como $\Phi$ es un isomorfismo, tenemos que $\mathcal{B} := \{v_1, \ldots, v_n\}$ es una base de $V$ (recordar el Teorema 2.22), la cual verifica:

$$f_i(v_j) = L_j(f_i) = \delta_{ij},$$

i.e. $\mathcal{B}'$ es la base dual de la base $\mathcal{B}$. □

Si $\mathcal{B} = \{b_1, \ldots, b_n\}$ es una base de $V$ y $\mathcal{B}^* = \{f_1, \ldots, f_n\}$ es su base dual, recordemos que

$$v = \sum_{i=1}^{n} f_i(v) b_i \qquad \text{para todo } v \in V, \text{ y} \tag{6.6}$$

$$f = \sum_{j=1}^{n} f(b_j) f_j \qquad \text{para todo } f \in V^*. \tag{6.7}$$

Es fácil ver que la base dual de $\mathcal{B}^*$ es la base $\mathcal{B}^{**} = \{L_{b_1}, \ldots, L_{b_n}\}$, y en consecuencia

$$f = \sum_{j=1}^{n} L_{b_j}(f) f_j = \sum_{j=1}^{n} f(b_j) f_j \qquad \text{para todo } f \in V^*, \text{ y}$$

$$L = \sum_{i=1}^{n} L(f_i) L_{b_i} \qquad \text{para todo } L \in V^{**}. \tag{6.8}$$

Utilizando el isomorfismo $\Phi$ del Teorema 6.16, es fácil ver que (6.6) y (6.8) dicen lo mismo. Por este motivo, suele decirse que $V$ es el espacio dual de $V^*$, ó que los espacios $V$ y $V^*$ están en mutua dualidad.

Una ventaja de esta identificación es utilizarla al momento de calcular el **doble anulador** $(S^\circ)^\circ$ de un subconjunto $S$ de $V$. De hecho, para cada vector $v \in S$ tenemos que

$$L_v(f) = f(v) = 0 \qquad \text{para cada } f \in S^\circ,$$

es decir, el conjunto $\{L_v : \ v \in S\}$ está contenido en $(S^\circ)^\circ$. Además, como este conjunto es la imagen de $S$ por $\Phi$, el subespacio generado por $S$ es isomorfo al subespacio generado por $\{L_v : \ v \in S\}$,

$$\overline{S} \cong \overline{\{L_v : \ v \in S\}} \subseteq (S^\circ)^\circ.$$

Como además sabemos que

$$\dim \overline{S} + \dim S^\circ = \dim V = \dim V^* = \dim S^\circ + \dim (S^\circ)^\circ,$$

resulta que $\overline{S}$ y $(S^\circ)^\circ$ tienen la misma dimensión. Por lo tanto,

$$(S^\circ)^\circ \cong \overline{S}.$$

**Teorema 6.19.** *Si $V$ es un $\mathbb{K}$-espacio vectorial de dimensión finita y $S$ es un subconjunto de $V$, entonces $(S^\circ)^\circ$ es isomorfo a $\overline{S}$.*

**Definición 6.20.** $(S^\circ)^\circ$ es el **doble anulador de** $S$ y lo anotaremos $S^{\circ\circ}$.

## 6.2. Traspuesta de una transformación lineal

Supongamos que $V$ y $W$ son dos $\mathbb{K}$-espacios vectoriales. Cada transformación lineal $T : V \to W$ induce una transformación lineal de $W^*$ en $V^*$ de la siguiente manera: dado un funcional $g \in W^*$, notemos que si $f : V \to \mathbb{K}$ está definida por

$$f(v) = g(Tv), \qquad v \in V,$$

entonces $f \in V^*$ (ya que es composición de transformaciones lineales). Luego, definimos $T^t : W^* \to V^*$ como

$$T^t(g) := g \circ T, \qquad g \in W^*.$$

Veamos que $T^t$ es lineal. Dados $g_1, g_2 \in W^*$ y $\alpha \in \mathbb{K}$, para cada $v \in V$ tenemos que

$$\begin{aligned} T^t(\alpha g_1 + g_2)\ v &= (\alpha g_1 + g_2)\ Tv = \alpha\ g_1(Tv) + g_2(Tv) = \alpha\ (g_1 \circ T)v + (g_2 \circ T)v \\ &= (\alpha\ T^t(g_1) + T^t(g_2))\ v, \end{aligned}$$

es decir, $T^t(\alpha g_1 + g_2) = \alpha\ T^t(g_1) + T^t(g_2)$. En conclusión, hemos probado el siguiente teorema.

**Teorema 6.21.** *Sean $V$ y $W$ dos $\mathbb{K}$-espacios vectoriales. Para cada $T \in L(V, W)$ existe una única $T^t \in L(W^*, V^*)$ tal que*

$$(T^t g)(v) = g(Tv), \qquad \text{para todo } g \in W^* \text{ y todo } v \in V,$$

*A $T^t$ la llamaremos la* **transformación traspuesta** *de $T$.*

**Ejemplo 6.22.** Si $T \in L(\mathbb{R}^2, \mathbb{R}^3)$ está definida como

$$T(x, y) = (x + y, y - x, -2x + y),$$

calculemos $T^t : (\mathbb{R}^3)^* \to (\mathbb{R}^2)^*$. Dado $g \in (\mathbb{R}^3)^*$, supongamos que

$$g(x, y, z) = \alpha x + \beta y + \gamma z,$$

para ciertos escalares $\alpha, \beta, \gamma \in \mathbb{K}$. Luego, como $T^t g \in (\mathbb{R}^2)^*$, para determinar cual es su acción debemos evaluarlo en un vector arbitrario de $\mathbb{R}^2$:

$$\begin{aligned} (T^t g)(x, y) = (g \circ T)(x, y) &= g(x + y, y - x, -2x + y) \\ &= \alpha(x + y) + \beta(y - x) + \gamma(-2x + y) \\ &= (\alpha - \beta - 2\gamma)x + (\alpha + \beta + \gamma)y. \end{aligned}$$

**Ejercicio 6.23.** Sea $T : \mathbb{R}[x] \to \mathbb{R}[x]$ dado por $(Tp)(x) = x^2 p(x) + p''(x)$ si $p \in \mathbb{R}[x]$.

i) Si $f \in \mathbb{R}[x]^*$ está definido como $f(p) = p'(4)$, calcular el funcional $T^t f \in \mathbb{R}[x]^*$.

ii) Si $g \in \mathbb{R}[x]^*$ está definido por $g(p) = \int_0^1 p(x)dx$, evaluar $(T^t g)(x^3)$.

A continuación enumeraremos algunas propiedades algebraicas de la traspuesta. La demostración queda como ejercicio para el lector.

**Proposición 6.24.** *Dados $\mathbb{K}$-espacios vectoriales $V$, $W$ y $W'$,*

*i)* $(S + T)^t = S^t + T^t$ *para cualesquiera* $S, T \in L(V, W)$*;*

*ii)* $(\alpha T)^t = \alpha T^t$ *para todo* $T \in L(V, W)$ *y todo* $\alpha \in \mathbb{K}$*;*

*iii)* $(ST)^t = T^t S^t$ *para cualesquiera* $S \in L(W, W')$, $T \in L(V, W)$.

El núcleo y la imagen de la transformación traspuesta de $T : V \to W$ coinciden con los anuladores de ciertos subespacios de $W$ y $V$, respectivamente.

**Teorema 6.25.** *Sean* $V$ *y* $W$ *dos* $\mathbb{K}$*-espacios vectoriales. Si* $T \in L(V, W)$ *entonces*

$$N(T^t) = \operatorname{Im}(T)^\circ.$$

*Además, si los espacios son de dimensión finita entonces*

*i)* $\operatorname{rg}(T^t) = \operatorname{rg}(T)$, *y;*

*ii)* $\operatorname{Im}(T^t) = N(T)^\circ$.

*Demostración.* Dado $g \in W^*$, $g \in N(T^t)$ si y sólo si $g \circ T = \vec{0}_{V^*}$, o equivalentemente,

$$g(Tv) = 0 \qquad \text{para todo } v \in V.$$

Por lo tanto, $g \in N(T^t)$ si y sólo si $g \in \operatorname{Im}(T)^\circ$.

Ahora, supongamos que $\dim V = n$ y $\dim W = m$. Supongamos además que $r := \operatorname{rg}(T) = \dim \operatorname{Im}(T)$. Luego, usando la caracterización de $N(T^t)$ y (6.5) tenemos que

$$\dim N(T^t) = \dim \operatorname{Im}(T)^\circ = \dim W - \dim \operatorname{Im}(T) = m - r.$$

Además, como $T^t \in L(W^*, V^*)$ sabemos que

$$\dim \operatorname{Im}(T^t) = \dim W^* - \dim N(T^t) = m - \dim N(T^t) = m - (m - r) = r.$$

Por lo tanto, $\operatorname{rg}(T^t) = \operatorname{rg}(T)$.

Para verificar la igualdad entre $\operatorname{Im}(T^t)$ y el anulador de $N(T)$, mostraremos que $\operatorname{Im}(T^t) \subseteq N(T)^\circ$ y que ambos subespacios de $V^*$ tienen la misma dimensión.

En primer lugar, si $f \in \operatorname{Im}(T^t)$, existe un $g \in W^*$ tal que $f = T^t g = g \circ T$. Luego, para cada $v \in N(T)$, tenemos que

$$f(v) = (g \circ T)v = g(Tv) = g(\vec{0}) = 0,$$

es decir, $f \in N(T)^\circ$. Por lo tanto, $\operatorname{Im}(T^t)$ es un subespacio de $N(T)^\circ$. Además,

$$\dim \operatorname{Im}(T^t) = \operatorname{rg}(T^t) = \operatorname{rg}(T) = \dim \operatorname{Im}(T) = \dim V - \dim N(T) = \dim N(T)^\circ,$$

donde usamos (6.5) para la última igualdad. Por lo tanto, $\operatorname{Im}(T^t) = N(T)^\circ$. □

**Ejercicio 6.26.** Si $V$ y $W$ son dos $\mathbb{K}$-espacios vectoriales de dimensión finita y $T \in L(V, W)$, probar que

$$\dim N(T^t) = \dim N(T) + \dim W - \dim V.$$

**Teorema 6.27.** *Sean* $V$ *y* $W$ *dos* $\mathbb{K}$*-espacios vectoriales con* $\dim V = n$ *y* $\dim W = m$*. Sean* $\mathcal{B}_V$ *y* $\mathcal{B}_W$ *bases de* $V$ *y* $W$*, con bases duales* $\mathcal{B}_V^*$ *y* $\mathcal{B}_W^*$*, respectivamente. Si* $T \in L(V, W)$ *entonces*

$$\left[T^t\right]_{\mathcal{B}_V^*, \mathcal{B}_W^*} = \left([T]_{\mathcal{B}_W, \mathcal{B}_V}\right)^\top.$$

*Demostración.* Supongamos que las bases están dadas por

$$\begin{aligned} \mathcal{B}_V &= \{v_1, \ldots, v_n\}, \quad \mathcal{B}_W = \{w_1, \ldots, w_m\}, \\ \mathcal{B}_V^* &= \{f_1, \ldots, f_n\}, \quad \mathcal{B}_W^* = \{g_1, \ldots, g_m\}. \end{aligned}$$

Si $[T]_{\mathcal{B}_W,\mathcal{B}_V} = A = (A_{ij}) \in \mathbb{K}^{m\times n}$ y $\left[T^t\right]_{\mathcal{B}_V^*,\mathcal{B}_W^*} = B = (B_{kl}) \in \mathbb{K}^{n\times m}$, tenemos que

$$\begin{aligned} Tv_i &= \sum_{j=1}^{m} A_{ji} w_j \quad i = 1, \ldots, n, \\ T^t g_l &= \sum_{k=1}^{n} B_{kl} f_k \quad l = 1, \ldots, m. \end{aligned} \tag{6.9}$$

Por otra parte, para cada $f \in V^*$ sabemos que $f = \sum_{i=1}^n f(v_i) f_i$. En particular, si $f = T^t g_j$ para algún $j = 1, \ldots, m$ resulta

$$T^t g_j = \sum_{i=1}^{n} T^t g_j(v_i) f_i.$$

Entonces, necesitamos calcular los escalares $T^t g_j(v_i)$. Dados $i = 1, \ldots, n$ y $j = 1, \ldots, m$,

$$T^t g_j(v_i) = g_j(Tv_i) = g_j\left(\sum_{k=1}^{m} A_{ki} w_k\right) = \sum_{k=1}^{m} A_{ki} g_j(w_k) = \sum_{k=1}^{m} A_{ki}\delta_{jk} = A_{ji},$$

con lo cual $T^t g_j = \sum_{i=1}^n A_{ji} f_i$. Comparando esto último con (6.9), resulta que

$$B_{ij} = A_{ji} \qquad \text{para todo } i = 1, \ldots, n \text{ y } j = 1, \ldots, m,$$

es decir, $B = A^\top$. $\square$

El teorema anterior nos da una forma alternativa de calcular la traspuesta de una transformación lineal. Revisemos una vez más el Ejemplo 6.22.

**Ejemplo 6.28.** Si $T \in L(\mathbb{R}^2, \mathbb{R}^3)$ está definida como $T(x, y) = (x + y, y - x, -2x + y)$, calculemos $T^t : (\mathbb{R}^3)^* \to (\mathbb{R}^2)^*$ utilizando la matriz de $T$.

Si $\mathcal{E}_2$ y $\mathcal{E}_3$ denotan las bases canónicas de $\mathbb{R}^2$ y $\mathbb{R}^3$, respectivamente, comencemos calculando $[T]_{\mathcal{E}_3,\mathcal{E}_2}$. Como

$$T(1, 0) = (1, -1, -2) \quad \text{y} \quad T(0, 1) = (1, 1, 1),$$

tenemos que

$$[T]_{\mathcal{E}_3,\mathcal{E}_2} = \begin{pmatrix} 1 & 1 \\ -1 & 1 \\ -2 & 1 \end{pmatrix}.$$

Por el teorema anterior, resulta que

$$\left[T^t\right]_{\mathcal{E}_2^*,\mathcal{E}_3^*} = ([T]_{\mathcal{E}_3,\mathcal{E}_2})^\top = \begin{pmatrix} 1 & -1 & -2 \\ 1 & 1 & 1 \end{pmatrix}.$$

Veamos ahora como reconstruir a $T^t$ a partir de esta información. Recordemos que

$$\begin{aligned} \mathcal{E}_3^* &= \{f_1(x, y, z) = x,\ f_2(x, y, z) = y,\ f_3(x, y, z) = z\}, \quad \text{y} \\ \mathcal{E}_2^* &= \{g_1(x, y) = x, g_2(x, y) = y\}. \end{aligned}$$

Dado $f \in (\mathbb{R}^3)^*$, supongamos que $f(x, y, z) = \alpha x + \beta y + \gamma z$, es decir,

$$f = \alpha f_1 + \beta f_2 + \gamma f_3.$$

Luego,

$$\begin{aligned} T^t f &= T^t(\alpha f_1 + \beta f_2 + \gamma f_3) = \alpha T^t f_1 + \beta T^t f_2 + \gamma T^t f_3 \\ &= \alpha(1.g_1 + 1.g_2) + \beta((-1).g_1 + 1.g_2) + \gamma((-2).g_1 + 1.g_2) \\ &= (\alpha - \beta - 2\gamma)g_1 + (\alpha + \beta + \gamma)g_2. \end{aligned}$$

Por lo tanto, $(T^t f)(x, y) = (\alpha - \beta - 2\gamma)x + (\alpha + \beta + \gamma)y$, como ya habíamos visto antes.

## 6.3. Ejercicios

1. Hallar la base dual de cada una de las siguientes bases:

   *a*) $\mathcal{B} = \{(1, 0, -1), (1, 1, 0), (1, 0, 0)\}$ base de $\mathbb{R}^3$.

   *b*) $\mathcal{B} = \{(-i, 0), (1, 1 - i)\}$ base de $\mathbb{C}^2$.

   *c*) $\mathcal{B} = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right\}$ base de $\mathbb{R}^{2\times 2}$.

2. Consideremos los siguientes vectores de $\mathbb{R}^3$:

   $$v_1 = (-1, -1, 0); \;\; v_2 = (0, 1, -2); \;\; v_3 = (1, 0, 1)\,.$$

   *a*) Probar que forman una base de $\mathbb{R}^3$ y hallar su base dual.

   *b*) Supongamos que $f : \mathbb{R}^3 \to \mathbb{R}$ es un funcional lineal tal que

   $$f(v_1) = 1; \;\; f(v_2) = -f(v_1); \;\; f(v_3) = 3.$$

   Hallar explícitamente a $f(x, y, z)$ para todo $(x, y, z) \in \mathbb{R}^3$. ¿Cuánto vale $f(3, 4, 0)$?

   *c*) Hallar un funcional lineal $g : \mathbb{R}^3 \to \mathbb{R}$ tal que

   $$g(v_1) = g(v_2) = 0; \;\; g(v_3) \neq 0.$$

3. Sea $V$ un $\mathbb{K}$-espacio vectorial de dimensión finita y sea $f$ un funcional lineal sobre $V$, no nulo. ¿Qué dimensión tiene $N(f)$?

4. Hallar un hiperespacio de $\mathbb{R}^2$ y uno de $\mathbb{R}^3$. Interpretar gráficamente.

5. Sea $V$ un $\mathbb{K}$-espacio vectorial de dimensión finita y sean $f$ y $g$ funcionales lineales sobre $V$, tales que $h(x) = f(x)g(x)$ es un funcional lineal sobre $V$. Probar que, o bien $f = 0$, o bien $g = 0$.

   **Sugerencia:** Evaluar $h$ en $\alpha \cdot x$ para $\alpha \in \mathbb{K}$.

6. Consideremos $\mathbb{C}^{(2)}[x]$, el $\mathbb{C}$-espacio vectorial de los polinomios de grado menor o igual a 2 con coeficientes complejos.

   *a*) Si definimos las funciones $f_1$, $f_2$ y $f_3$ de $\mathbb{C}^{(2)}[x]$ en $\mathbb{C}$ mediante

   $$f_k(p) = p(k - 1) \quad \text{para } p \in \mathbb{C}_2[x],$$

   probar que $\mathcal{B}^* = \{f_1, f_2, f_3\}$ es una base del espacio dual de $\mathbb{C}^{(2)}[x]$.

   *b*) Hallar una base de $\mathbb{C}^{(2)}[x]$, tal que $\mathcal{B}^*$ sea su base dual.

7. Sea $V$ un $\mathbb{K}$-espacio vectorial. Probar que, para cualquier subconjunto $S$ de $V$:
   *a*) el anulador $S^\circ$ es un subespacio de $V^*$;
   *b*) si $\overline{S}$ es el subespacio generado por $S$, entonces $\left(\overline{S}\right)^\circ = S^\circ$.

   Verificar además que $\{\vec{0}_V\}^\circ = V^*$ y $V^\circ = \{\vec{0}_{V^*}\}$.

8. Hallar el subespacio $S^\circ$ en cada uno de los siguientes casos:
   *a*) $S = \overline{\{(1,1,0,0,0),(2,-1,1,2,0)\}} \subset \mathbb{R}^5$.
   *b*) $S = \overline{\{(x,y,z):\ x+y+2z=0 \wedge x-3z=0\}} \subset \mathbb{R}^3$.
   *c*) $S = \overline{\{1+x+x^2,\ 2+x^3+2x^4,\ x^3\}} \subset \mathbb{R}_4[x]$.

9. Sea $W$ el subespacio de $\mathbb{R}^4$ generado por $u = (2,3,1,1)$ y $v = (1,0,-1,2)$. ¿Qué funcionales lineales

$$f(x_1,x_2,x_3,x_4) = c_1x_1 + c_2x_2 + c_3x_3 + c_4x_4 \quad (c_1,c_2,c_3,c_4 \in \mathbb{R})$$

   pertenecen a $W^\circ$?

10. Sea $V$ un $\mathbb{K}$-espacio vectorial de dimensión finita. Probar que:
    *a*) Si $A$ y $B$ son subconjuntos de $V$ tales que $A \subseteq B$ entonces $B^\circ \subseteq A^\circ$.
    *b*) Si $W_1$ y $W_2$ son subespacios de $V$ entonces,
       *i*. $(W_1+W_2)^\circ = W_1^\circ \cap W_2^\circ$.
       *ii*. $(W_1 \cap W_2)^\circ = W_1^\circ + W_2^\circ$.
    *c*) Sean $W_1 = \{(x_1,x_2,x_3,x_4):\ x_1 - x_3 = 0 \wedge x_1+x_2+x_4 = 0\}$ y $W_2 = \overline{\{(2,1,3,1)\}}$ subespacios de $\mathbb{R}^4$. Hallar una base para $(W_1+W_2)^\circ$ y una base para $(W_1 \cap W_2)^\circ$.

11. Sean $U$, $V$ y $W$ tres $\mathbb{K}$-espacios vectoriales. Probar que:
    *a*) $(S+T)^t = S^t + T^t$ para $S,T \in L(U,V)$.
    *b*) $(\alpha T)^t = \alpha T^t$ para $\alpha \in \mathbb{K}$ y $T \in L(U,V)$.
    *c*) $(ST)^t = T^tS^t$ para $S \in L(V,U)$ y $T \in L(W,V)$.

12. Sea $f$ un funcional lineal sobre $\mathbb{C}^2$, dado por $f(u,v) = \alpha u + \beta v$ $(\alpha,\beta \in \mathbb{C})$. Hallar $g = T^tf = f \circ T$, siendo $T$ alguno de los operadores $P_X, S_Y, H_2$ introducidos en los Ejercicios del Capítulo 1.

13. Sea $T : \mathbb{R}^2 \to \mathbb{R}^4$ la transformación lineal dada por

$$T(x,y) = (x+y, x, x-2y, y)\,.$$

    Calcular $T^t(f)$, para el funcional lineal $f$ de $\mathbb{R}^4$ dado por

$$f(s,t,u,v) = s - 2t + u + v\,.$$

14. Sea $T : \mathbb{R}^4 \to \mathbb{R}^3$ la transformación lineal dada por

$$T(x,y,z,w) = (x-y, x-z, x-w)\,.$$

    Calcular $T^t(g)$, para el funcional lineal $g$ de $\mathbb{R}^3$ dado por

$$g(t,u,v) = 3t + 4u - v\,.$$

15. Sea $T : \mathbb{R}[x] \to \mathbb{R}[x]$ dado por $(Tp)(x) = x^2 p(x) + p''(x)$ si $p \in \mathbb{R}[x]$.

    *a)* Si $f \in \mathbb{R}[x]^*$ está definido como $f(p) = p'(4)$, calcular el funcional $T^t f \in \mathbb{R}[x]^*$.

    *b)* Si $g \in \mathbb{R}[x]^*$ está definido por $g(p) = \int_0^1 p(x)dx$, evaluar $(T^t g)(x^3)$.

16. Sea $T \in L(\mathbb{R}^2)$ dado por $T(x, y) = (-y, x)$.

    *a)* Hallar la representación matricial de $T^t \in L((\mathbb{R}^2)^*)$ en la base dual de la base canónica de $\mathbb{R}^2$.

    *b)* Hallar el núcleo y la imagen de $T^t$ y analizar inyectividad, suryectividad y biyectividad.

17. Dada una matriz $B \in \mathbb{K}^{n\times n}$ fija, sea $f_B : \mathbb{K}^{n\times n} \to \mathbb{K}$ la función definida por

    $$f_B(A) = \mathrm{tr}(B^\top A).$$

    Probar que:

    *a)* $f_B$ es un funcional lineal sobre $\mathbb{K}^{n\times n}$.

    *b)* La función $T : \mathbb{K}^{n\times n} \to (\mathbb{K}^{n\times n})^*$ dada por $T(B) = f_B$ es una transformación lineal de $\mathbb{K}^{n\times n}$ en $(\mathbb{K}^{n\times n})^*$.

    *c)* $T$ es un isomorfismo entre $\mathbb{K}^{n\times n}$ y $(\mathbb{K}^{n\times n})^*$.

18. Si $V$ y $W$ son dos $\mathbb{K}$-espacios vectoriales de dimensión finita y $T \in L(V, W)$, probar que

    $$\dim N(T^t) = \dim N(T) + \dim W - \dim V.$$

# Capítulo 7

# Espacios con producto interno

A lo largo de este capítulo $\mathbb{K} = \mathbb{R}$ ó $\mathbb{K} = \mathbb{C}$, es decir, nos restringiremos a considerar el cuerpo de los números reales $\mathbb{R}$ y el cuerpo de los números complejos $\mathbb{C}$, incluso en algunos casos tendremos que explicitar cual es el cuerpo sobre el que estamos trabajando.

Al presentar la definición de un $\mathbb{K}$-espacio vectorial hemos generalizado la estructura lineal (adición y multiplicación por un escalar) de los espacios $\mathbb{R}^2$ y $\mathbb{R}^3$, pero hemos ignorado otros aspectos importantes para poder desarrollar alguna clase de geometría sobre los $\mathbb{K}$-espacios vectoriales abstractos. En particular necesitamos introducir las nociones de "longitud de un vector" y "ángulo entre vectores", las cuales desarrollaremos a partir del concepto de producto interno.

Para esto usaremos como modelo al espacio $\mathbb{R}^3$. En dicho espacio la longitud y el ángulo se definen a partir del producto escalar:

$$\vec{x} \cdot \vec{y} = x_1y_1 + x_2y_2 + x_3y_3, \qquad \text{si } \vec{x} = (x_1, x_2, x_3),\ \vec{y} = (y_1, y_2, y_3).$$

Más precisamente, dado un vector $\vec{x} = (x_1, x_2, x_3) \in \mathbb{R}^3$ la longitud de $\vec{x}$ está dada por

$$\|\vec{x}\| = \sqrt{x_1^2 + x_2^2 + x_3^2} = \sqrt{\vec{x} \cdot \vec{y}},$$

y si $\vec{y} = (y_1, y_2, y_3)$ es otro vector de $\mathbb{R}^3$ el ángulo (mínimo) entre $\vec{x}$ e $\vec{y}$ es el único $\theta \in [0, \frac{\pi}{2}]$ que verifica

$$\cos\theta = \frac{\vec{x} \cdot \vec{y}}{\|\vec{x}\|\|\vec{y}\|}.$$

Intentaremos introducir una estructura similar al producto escalar de $\mathbb{R}^3$ en $\mathbb{K}$-espacios vectoriales abstractos, de la siguiente manera.

**Definición 7.1.** Dado un $\mathbb{K}$-espacio vectorial $V$, un **producto interno** sobre $V$ es una función $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{K}$ tal que, para $u, v, w \in V$ cualesquiera y para todo $\alpha \in \mathbb{K}$, verifica las siguientes condiciones:

i) $\langle \alpha u + v, w \rangle = \alpha \langle u, w \rangle + \langle v, w \rangle$;

ii) $\langle v, w \rangle = \overline{\langle w, v \rangle}$;

iii) $\langle v, v \rangle > 0$ si $v \neq \vec{0}$.

**Observación 7.2.**

- Del ítem $i$) se deduce que $\left\langle \vec{0}, v \right\rangle = 0$ para todo $v \in V$, ya que

$$\left\langle \vec{0}, v \right\rangle = \langle 0v, v\rangle = 0 \langle v, v\rangle = 0.$$

  En particular, $v = \vec{0}$ es el único vector en $V$ que satisface $\langle v, v\rangle = 0$.

- De los ítems $i$) y $ii$) se deduce que

$$\langle u, \alpha v + w\rangle = \overline{\alpha} \langle u, v\rangle + \langle u, w\rangle, \qquad \text{para cualesquiera } u, v, w \in V,\ \alpha \in \mathbb{K}.$$

- La barra sobre el miembro derecho en la igualdad del ítem $ii$) denota al complejo conjugado del escalar correspondiente. En el caso en que $\mathbb{K} = \mathbb{R}$, esta conjugación se omite (porque el conjugado de un número real es el mismo número). Pero en el caso en que $\mathbb{K} = \mathbb{C}$ esta conjugación es indispensable para la coexistencia de los ítem $i$) y $ii$), sin ella arribaríamos a la siguiente contradicción: dado $v \in V$ cualquiera,

$$\langle v, v\rangle > 0 \quad \text{y} \quad -1 \langle v, v\rangle = i^2 \langle v, v\rangle = \langle iv, iv\rangle > 0.$$

**Ejemplos 7.3.** A continuación enumeramos varios ejemplos de productos internos sobre diferentes $\mathbb{K}$-espacios vectoriales.

1. En $\mathbb{K}^n$, el **producto interno canónico** está definido por

$$\langle x, y\rangle = \sum_{i=1}^{n} x_i \overline{y_i}, \qquad \text{si } x = (x_1, \dots, x_n), y = (y_1, \dots, y_n) \in \mathbb{K}^n.$$

   Le adjudicamos el adjetivo "canónico" porque este producto interno se define a partir de las coordenadas de los vectores con respecto a la base canónica de $\mathbb{K}^n$.

2. En $\mathbb{R}^2$, consideremos el producto interno dado por

$$\langle x, y\rangle = x_1y_1 - x_2y_1 - x_1y_2 + 4x_2y_2, \qquad \text{si } x = (x_1, x_2), y = (y_1, y_2) \in \mathbb{R}^2. \tag{7.1}$$

   Es fácil comprobar que satisface $i$) y $ii$). Para verificar que vale $iii$), notemos que si $x = (x_1, x_2) \neq (0, 0)$ entonces

$$\langle x, x\rangle = x_1^2 - 2x_1x_2 + 4x_2^2 = (x_1 - x_2)^2 + 3x_2^3 > 0.$$

3. En $\mathbb{K}^{n\times m}$, podemos definir un producto interno mediante

$$\langle A, B\rangle = \operatorname{tr}(B^*A), \qquad \text{si } A, B \in \mathbb{K}^{n\times m},$$

   donde $B^* \in \mathbb{K}^{m\times n}$ denota a la **matriz adjunta** de $B$, es decir, $B^*_{ij} = \overline{B_{ji}}$.

   Aunque no lo parezca, este producto interno es muy similar a los anteriores, ya que si $A = (A_{ij}), B = (B_{ij}) \in \mathbb{K}^{n\times m}$ entonces

$$\begin{aligned}\langle A, B\rangle = \operatorname{tr}(B^*A) &= \sum_{j=1}^{m} (B^*A)_{jj} = \sum_{j=1}^{m} \left( \sum_{i=1}^{n} B^*_{ji} A_{ij} \right) = \sum_{i=1}^{n} \left( \sum_{j=1}^{m} A_{ij} B^*_{ji} \right) \\ &= \sum_{i=1}^{n} \sum_{j=1}^{m} A_{ij} \overline{B_{ij}}\end{aligned}$$

4. En $C([-1,1])$, consideremos

$$\langle f,g\rangle = \int_{-1}^{1} f(t)g(t)dt, \qquad f,g \in C([-1,1]).$$

Éste también es un producto interno sobre $C([-1,1])$.

**Ejercicio 7.4.** Comprobar que los anteriores son realmente productos internos sobre los diferentes $\mathbb{K}$-espacios vectoriales.

Los isomorfismos entre $\mathbb{K}$-espacios vectoriales nos permiten replicar productos internos de un espacio en otro. Dados dos $\mathbb{K}$-espacios vectoriales isomorfos $V$ y $W$, supongamos que $\langle\cdot,\cdot\rangle$ es un producto interno sobre $W$. Si $T \in L(V,W)$ es un isomorfismo, entonces

$$\langle u,v\rangle_T := \langle Tu, Tv\rangle, \qquad u,v \in V,$$

define un producto interno sobre $V$ (comprobarlo!).

En particular, si $\dim V = n$ y $\mathcal{B} = \{b_1,\ldots,b_n\}$ es una base de $V$, sea $T \in L(V,\mathbb{K}^n)$ el isomorfismo dado por

$$Tb_i = e_i, \qquad i = 1,\ldots,n,$$

siendo $\mathcal{E} = \{e_1,\ldots,e_n\}$ la base canónica de $\mathbb{K}^n$. Entonces, tenemos el producto interno sobre $V$ dado por

$$\left\langle \sum_{i=1}^n \alpha_i b_i, \sum_{j=1}^n \beta_j b_j \right\rangle_T = \left\langle T\left(\sum_{i=1}^n \alpha_i b_i\right), T\left(\sum_{j=1}^n \beta_j b_j\right)\right\rangle = \left\langle \sum_{i=1}^n \alpha_i e_i, \sum_{j=1}^n \beta_j e_j \right\rangle = \sum_{i=1}^n \alpha_i \overline{\beta_i}.$$

En particular, verifica que $\langle b_i, b_j\rangle_T = \delta_{ij}$.

**Definición 7.5.** A un $\mathbb{K}$-espacio vectorial $V$ dotado de un producto interno $\langle\cdot,\cdot\rangle$, lo llamaremos **espacio con producto interno** y lo anotaremos $(V,\langle\cdot,\cdot\rangle)$.

A un $\mathbb{R}$-espacio con producto interno de dimensión finita se lo suele llamar **espacio euclidiano**, y a un $\mathbb{C}$-espacio con producto interno de dimensión finita se lo suele llamar **espacio unitario**.

**Observación 7.6.** Si $(V,\langle\cdot,\cdot\rangle)$ es un espacio con producto interno con $\dim V = n$ y $\mathcal{B} = \{b_1,\ldots,b_n\}$ es una base de $V$, entonces $\langle\cdot,\cdot\rangle$ está completamente determinado por los valores

$$G_{ij} = \langle b_j, b_i\rangle, \qquad i,j = 1,\ldots,n.$$

En efecto, si $u,v \in V$ son vectores arbitrarios y sus coordenadas en la base $\mathcal{B}$ están dadas por

$$[u]_{\mathcal{B}} = (\alpha_1,\ldots,\alpha_n)^\top \quad \text{y} \quad [v]_{\mathcal{B}} = (\beta_1,\ldots,\beta_n)^\top,$$

entonces

$$\begin{aligned}
\langle u,v\rangle &= \left\langle \sum_{i=1}^n \alpha_i b_i, v\right\rangle = \sum_{i=1}^n \alpha_i \langle b_i, v\rangle = \sum_{i=1}^n \alpha_i \left\langle b_i, \sum_{j=1}^n \beta_j b_j\right\rangle = \sum_{i=1}^n \alpha_i \sum_{j=1}^n \overline{\beta_j}\langle b_i, b_j\rangle \\
&= \sum_{i=1}^n\sum_{j=1}^n \alpha_i\overline{\beta_j}G_{ji} = \sum_{j=1}^n \overline{\beta_j}\left(\sum_{i=1}^n G_{ji}\alpha_i\right) = \sum_{j=1}^n \overline{\beta_j}(G[u]_{\mathcal{B}})_j \\
&= [v]_{\mathcal{B}}^*\, G\, [u]_{\mathcal{B}},
\end{aligned}$$

donde $G \in \mathbb{K}^{n\times n}$ es la matriz dada por los coeficientes $G_{ij} = \langle b_j, b_i \rangle$ y $[v]_{\mathcal{B}}^{\,*} = (\overline{\beta_1}, \ldots, \overline{\beta_n})$.

Esta matriz hereda varias propiedades de los axiomas que definen al producto interno. El ítem *ii*) garantiza que $G$ es autoadjunta, es decir $G^* = G$, mientras que el ítem *iii*) dice que

$$X^*GX > 0 \qquad \text{para toda } X \in \mathbb{K}^{n\times 1}.$$

Por esto último diremos que $G$ es **definida positiva**. Lo que acabamos de mostrar es que todo producto interno de $V$ está determinado por una matriz definida positiva. Recíprocamente, fijada una base $\mathcal{B}$, toda matriz definida positiva $G \in \mathbb{K}^{n\times n}$ determina un producto interno en $V$ mediante $\langle u, v\rangle := [v]_{\mathcal{B}}^{\,*}\, G\, [u]_{\mathcal{B}}$ (comprobarlo!).

Dado un espacio con producto interno $(V, \langle\cdot,\cdot\rangle)$, sea $Q : V \to \mathbb{R}^+$ la función definida por

$$Q(v) = \langle v, v\rangle\,, \qquad v \in V.$$

Esta es la **forma cuadrática asociada** al producto interno, le decimos así porque si $v \in V$ y $\alpha \in \mathbb{K}$ entonces

$$Q(\alpha v) = \langle \alpha v, \alpha v\rangle = \alpha\overline{\alpha}\,\langle v, v\rangle = |\alpha|^2\,\langle v, v\rangle = |\alpha|^2 Q(v).$$

Veamos que el producto interno está determinado por $Q$, mediante la **identidad de polarización**. Esta identidad se presenta de dos formas diferentes, de acuerdo a si el cuerpo es el de los números reales ó el de los complejos.

Si $\mathbb{K} = \mathbb{R}$, dados $u, v \in V$ tenemos que

$$\begin{aligned} Q(u+v) &= \langle u+v, u+v\rangle = \langle u, u+v\rangle + \langle v, u+v\rangle = \langle u,u\rangle + \langle u,v\rangle + \langle v,u\rangle + \langle v,v\rangle \\ &= Q(u) + 2\,\langle u,v\rangle + Q(v), \end{aligned}$$

mientras que $Q(u-v) = Q(u) - 2\,\langle u,v\rangle + Q(v)$, de donde deducimos que

$$\langle u, v\rangle = \frac{1}{4}\big(Q(u+v) - Q(u-v)\big). \tag{7.2}$$

En cambio, si $\mathbb{K} = \mathbb{C}$, dados $u, v \in V$ repitiendo el argumento anterior vemos que

$$\begin{aligned} Q(u+v) &= \langle u+v, u+v\rangle = Q(u) + 2\,\mathrm{Re}(\langle u,v\rangle) + Q(v), \quad \text{y} \\ Q(u-v) &= \langle u+v, u+v\rangle = Q(u) - 2\,\mathrm{Re}(\langle u,v\rangle) + Q(v), \end{aligned}$$

con lo cual

$$\mathrm{Re}(\langle u,v\rangle) = \frac{1}{4}\big(Q(u+v) - Q(u-v)\big).$$

Ahora, recordando que $\mathrm{Im}(z) = \mathrm{Re}(-iz)$ si $z \in \mathbb{C}$, tenemos que

$$\mathrm{Im}(\langle u,v\rangle) = \mathrm{Re}(-i\,\langle u,v\rangle) = \mathrm{Re}(\overline{i}\,\langle u,v\rangle) = \mathrm{Re}(\langle u, iv\rangle) = \frac{1}{4}\big(Q(u+iv) - Q(u-iv)\big).$$

Por lo tanto,

$$\begin{aligned} \langle u,v\rangle &= \mathrm{Re}(\langle u,v\rangle) + i\,\mathrm{Im}(\langle u,v\rangle) \\ &= \frac{1}{4}\big(Q(u+v) - Q(u-v) + iQ(u+iv) - iQ(u-iv)\big) \\ &= \frac{1}{4}\sum_{k=0}^{3} i^k Q(u + i^k v). \end{aligned} \tag{7.3}$$

Como anticipamos al comienzo, las fórmulas (7.2) y (7.3) son las identidades de polarización para el caso real y para el caso complejo, respectivamente.

**Proposición 7.7.** *Si* $(V, \langle \cdot, \cdot \rangle)$ *es un espacio con producto interno y* $W$ *es un subespacio de* $V$*, entonces* $(W, \langle \cdot, \cdot \rangle)$ *también es un espacio con producto interno.*

*Demostración.* Sólo hace falta convencerse de que la restricción del producto interno $\langle \cdot, \cdot \rangle$ a $W$,

$$\langle \cdot, \cdot \rangle : W \times W \to \mathbb{K},$$

verifica las condiciones que definen a un producto interno, pero esto es inmediato ya que las verifica para un dominio mayor $(V \times V)$. □

## 7.1. Longitudes y ángulos entre vectores

Cada producto interno determina una norma, la que a su vez determina una distancia entre vectores.

**Definición 7.8.** Dado un espacio con producto interno $(V, \langle \cdot, \cdot \rangle)$, para cada $v \in V$ la **norma** de $v$ se define como

$$\|v\| := \sqrt{\langle v, v \rangle}.$$

La norma asociada a un producto interno tiene una serie de propiedades que reuniremos en el siguiente resultado.

**Proposición 7.9.** *Si* $(V, \langle \cdot, \cdot \rangle)$ *es un espacio con producto interno, para vectores* $u, v \in V$ *cualesquiera y para todo* $\alpha \in \mathbb{K}$ *tenemos*

*i)* $\|v\| > 0$ *si* $v \neq \vec{0}$*;*

*ii)* $\|\alpha v\| = |\alpha| \; \|v\|$*;*

*iii)* $|\langle u, v \rangle| \leq \|u\| \; \|v\|$*;*

*iv)* $\|u + v\| \leq \|u\| + \|v\|$*.*

La última desigualdad del enunciado se conoce como la **desigualdad triangular**, ya que puede interpretarse como que la longitud de cada lado de un triángulo es menor o igual a la suma de las longitudes de los lados restantes.

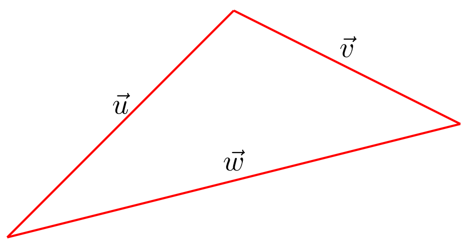



Figura 7.1: El vector $w$ es la suma de los vectores $u$ y $v$.

*Demostración.* Los ítems *i)* y *ii)* son consecuencias inmediatas de las propiedades del producto interno.

Probemos ahora el ítem *iii)*. Sean $u, v \in V$ vectores arbitrarios. Si $u = \vec{0}$, la conclusión es cierta. Supongamos en cambio que $u \neq \vec{0}$ y consideremos el vector

$$w := v - \frac{\langle v, u \rangle}{\|u\|^2} u \in V.$$

Entonces, $\langle w, u\rangle = \langle v, u\rangle - \frac{\langle v,u\rangle}{\|u\|^2}\langle u, u\rangle = 0$ y además,

$$\begin{aligned} 0 \leq \langle w, w\rangle &= \left\langle v - \frac{\langle v,u\rangle}{\|u\|^2}u, w\right\rangle = \langle v, w\rangle - \frac{\langle v,u\rangle}{\|u\|^2}\langle u, w\rangle = \langle v, w\rangle \\ &= \left\langle v, v - \frac{\langle v,u\rangle}{\|u\|^2}u\right\rangle = \langle v, v\rangle - \frac{\langle v,u\rangle}{\|u\|^2}\langle v, u\rangle \\ &= \|v\|^2 - \frac{|\langle v,u\rangle|^2}{\|u\|^2}, \end{aligned}$$

es decir, $|\langle v, u\rangle|^2 \leq \|v\|^2\|u\|^2$, o equivalentemente, $|\langle u, v\rangle| \leq \|u\|\,\|v\|$.
Finalmente, usaremos *iii*) para probar la desigualdad *iv*):

$$\begin{aligned} \|u+v\|^2 &= \langle u+v, u+v\rangle = \langle u, u\rangle + 2\operatorname{Re}(\langle u, v\rangle) + \langle v, v\rangle \\ &= \|u\|^2 + 2\operatorname{Re}(\langle u, v\rangle) + \|v\|^2 \\ &\leq \|u\|^2 + 2|\langle u, v\rangle| + \|v\|^2 \\ &\leq \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 = (\|u\| + \|v\|)^2, \end{aligned}$$

es decir, $\|u+v\| \leq \|u\| + \|v\|$. □

Por otra parte, la desigualdad del ítem *iii*) de la Proposición 7.9 es conocida como **desigualdad de Cauchy-Schwarz** y se manifiesta de distintas formas, de acuerdo al producto interno involucrado. Veamos su aspecto en los Ejemplos 7.3:

1. En $\mathbb{K}^n$, si $x = (x_1, \ldots, x_n)$ e $y = (y_1, \ldots, y_n)$,

$$\left|\sum_{i=1}^n x_i\overline{y_i}\right| \leq \left(\sum_{i=1}^n |x_i|^2\right)^{1/2}\left(\sum_{i=1}^n |y_i|^2\right)^{1/2}.$$

   Esta es la versión de la desigualdad que probó originalmente el matemático francés Augustin-Louis Cauchy (1789–1857) en 1821 (en el caso en que $\mathbb{K} = \mathbb{R}$).

2. En $\mathbb{R}^2$, si $x = (x_1, x_2)$ e $y = (y_1, y_2)$,

$$|x_1y_1 - x_2y_1 - x_1y_2 + 4x_2y_2| \leq \big((x_1 - x_2)^2 + 3x^2\big)^{1/2}\big((y_1 - y_2)^2 + 3y^2\big)^{1/2}.$$

3. En $\mathbb{K}^{n\times m}$,

$$|\operatorname{tr}(B^*A)| \leq \big(\operatorname{tr}(A^*A)\big)^{1/2}\big(\operatorname{tr}(B^*B)\big)^{1/2}.$$

4. En $C([-1, 1])$,

$$\left|\int_{-1}^1 f(t)g(t)dt\right| \leq \left(\int_{-1}^1 f(t)^2dt\right)^{1/2}\left(\int_{-1}^1 g(t)^2dt\right)^{1/2}.$$

   Esta es la versión de la desigualdad que probó el matemático alemán Hermann Schwarz (1843–1921) en 1886.

El siguiente resultado se conoce como la **identidad del paralelogramo** debido a su interpretación geométrica: en cada paralelogramo, la suma de los cuadrados de las longitudes de las diagonales es igual a la suma de los cuadrados de las longitudes de los cuatro lados. La demostración queda como ejercicio.

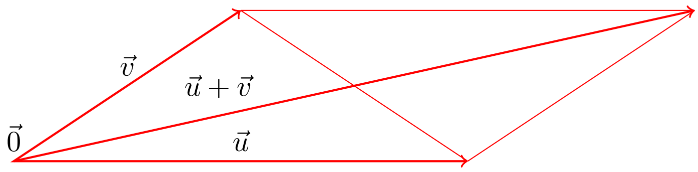



Figura 7.2: La suma de los cuadrados de las longitudes de $u+v$ y $u-v$ coincide con el doble de la suma de los cuadrados de las longitudes de $u$ y $v$.

**Proposición 7.10** (Identidad del paralelogramo)**.** *Sea* $(V, \langle\cdot,\cdot\rangle)$ *un espacio con producto interno. Si* $u, v \in V$ *entonces,*

$$\|u+v\|^2 + \|u-v\|^2 = 2(\|u\|^2 + \|v\|^2).$$

Asociada a cada norma tenemos una noción de **distancia**. Si $(V, \langle\cdot,\cdot\rangle)$ es un espacio con producto interno, consideremos la función $\mathrm{d} : V \times V \to [0, +\infty)$ definida por

$$\mathrm{d}(u, v) = \|u - v\|, \qquad u, v \in V. \tag{7.4}$$

Esta distancia nos permite interpretar a la norma de un vector $v \in V$ como la longitud de $v$, ya que $\|v\|$ coincide con la distancia entre el vector $v$ y el vector nulo.

**Ejercicio 7.11.** Si d es la distancia en $(V, \langle\cdot,\cdot\rangle)$ dada por (7.4), probar que verifica las siguientes propiedades: dados $u, v \in V$,

i) $\mathrm{d}(u, v) = 0$ si y sólo si $v = u$;

ii) $\mathrm{d}(u, v) = \mathrm{d}(v, u)$;

iii) $\mathrm{d}(u, v) \leq \mathrm{d}(u, w) + \mathrm{d}(w, v)$ para cualquier $w \in V$.

Ahora presentaremos una noción de perpendicularidad, la cual lleva implícita la idea de “ángulo” entre vectores.

**Definición 7.12.** Sea $(V, \langle\cdot,\cdot\rangle)$ un espacio con producto interno. Dados $u, v \in V$, diremos que son **vectores ortogonales** si $\langle u, v\rangle = 0$, y lo anotaremos $u \perp v$.

Notemos que $\vec{0} \perp v$ para todo $v \in V$. Además es el único vector con esta propiedad, porque si $u \perp v$ para todo $v \in V$ en particular vale que $\langle u, u\rangle = 0$, lo que implica que $u = \vec{0}$.

**Teorema 7.13** (Pitágoras)**.** *Si* $u$ *y* $v$ *son vectores ortogonales en* $V$*, entonces*

$$\|u+v\|^2 = \|u\|^2 + \|v\|^2.$$

*Demostración.* Si $u \perp v$ entonces

$$\begin{aligned} \|u+v\|^2 &= \langle u+v, u+v\rangle = \langle u, u\rangle + 2\,\mathrm{Re}(\langle u, v\rangle) + \langle v, v\rangle \\ &= \langle u, u\rangle + \langle v, v\rangle = \|u\|^2 + \|v\|^2. \end{aligned}$$

□

La prueba anterior muestra que $\|u+v\|^2 = \|u\|^2 + \|v\|^2$ si y sólo si $2\,\mathrm{Re}(\langle u, v\rangle) = 0$. Por lo tanto, la recíproca del Teorema de Pitágoras también es válida en $\mathbb{R}$-espacios vectoriales con producto interno.

**Definición 7.14.** Sea $(V, \langle\cdot,\cdot\rangle)$ un espacio con producto interno. Dados dos vectores no nulos $u, v \in V$, el **angulo (mínimo) entre** $u$ **y** $v$ es el único $\theta \in [0, \frac{\pi}{2}]$ tal que

$$\cos(\theta) = \frac{|\langle u, v\rangle|}{\|u\| \ \|v\|}.$$

La desigualdad de Cauchy-Schwarz dice que el cociente $\frac{|\langle u,v\rangle|}{\|u\| \ \|v\|}$ pertenece al intervalo $[0, 1]$, con lo cual garantiza que el ángulo entre $u$ y $v$ está bien definido.

**Definición 7.15.** Si $S$ es un subconjunto de $V$, diremos que $S$ es un **conjunto ortogonal** si cada par de vectores distintos de $S$ son ortogonales entre sí, es decir, si $u, v \in S$ y $u \neq v$ entonces $u \perp v$. Diremos que el conjunto $S$ es un **conjunto ortonormal** si es ortogonal y además $\|v\| = 1$ para todo $v \in S$.

Veamos algunos ejemplos de conjuntos ortonormales.

**Ejemplos 7.16.**

1. La base canónica $\mathcal{E} = \{e_1, \ldots, e_n\}$ es un conjunto ortonormal en $\mathbb{K}^n$ con respecto al producto interno canónico.

2. Si $V = \mathbb{K}^{n\times m}$, consideremos el conjunto $\{E^{p,q} : \ p = 1, \ldots, n, \ q = 1, \ldots, m\}$ donde $E^{p,q}$ es la matriz cuya única entrada no nula es un 1 en la entrada $p, q$ (fila $p$, columna $q$). Este conjunto es ortonormal con respecto al producto interno $\langle A, B\rangle = \operatorname{tr}(AB^*)$, ya que

$$\langle E^{p,q}, E^{r,s}\rangle = \operatorname{tr}(E^{p,q}(E^{r,s})^*) = \operatorname{tr}(E^{p,q}E^{s,r}) = \operatorname{tr}(\delta_{q,s}E^{p,r}) = \delta_{q,s}\operatorname{tr}(E^{p,r}) = \delta_{q,s}\delta_{p,r}.$$

   En el argumento de arriba hemos abusado sistemáticamente de la notación ya que utilizamos la misma notación para matrices de distintas dimensiones, por ejemplo, $(E^{r,s})^* = E^{s,r} \in \mathbb{K}^{m\times n}$ y $E^{p,q}E^{s,r} = \delta_{q,s}E^{p,r} \in \mathbb{K}^{n\times n}$.

3. Si $V = C([0, 1])$ dotado con el producto interno $\langle f, g\rangle = \int_0^1 f(t)g(t)dt$, consideremos las funciones

$$f_k(x) = \sqrt{2}\cos(2\pi kx) \quad \text{y} \quad g_k(x) = \sqrt{2}\operatorname{sen}(2\pi kx).$$

   Luego, el conjunto $\{1, f_1, g_1, f_2, g_2, \ldots\}$ es un conjunto ortonormal infinito. Para verificar esta última afirmación es necesario calcular las integrales

$$\int_0^1 \cos(2\pi kx)\operatorname{sen}(2\pi lx)dx,$$

   para $(k, l) \in \mathbb{N} \times \mathbb{N}$ cualquiera.

Que los conjuntos que propusimos en los ejemplos sean linealmente independientes no es un casualidad, ya que todo conjunto ortogonal es linealmente independiente. Antes de probarlo, veamos que los vectores del subespacio generado por un conjunto ortogonal (de vectores no nulos) admiten una única representación.

**Proposición 7.17.** *Dado un espacio con producto interno $(V, \langle\cdot,\cdot\rangle)$, sea $\{v_1, \ldots, v_k\}$ un conjunto ortogonal de vectores no nulos de $V$. Si $w \in \overline{\{v_1, \ldots, v_k\}}$, entonces $w$ se representa de manera única como combinación lineal de $v_1, \ldots, v_k$ como*

$$w = \sum_{i=1}^{k} \frac{\langle w, v_i\rangle}{\|v_i\|^2} \ v_i.$$

*Demostración.* Sea $w \in W := \overline{\{v_1, \ldots, v_k\}}$. Luego, existen $\alpha_1, \ldots, \alpha_k \in \mathbb{K}$ tales que $w = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_k v_k$. Veamos que $\alpha_1, \ldots, \alpha_k$ están unívocamente determinados: para cada $j = 1, \ldots, k$,

$$\langle w, v_j \rangle = \left\langle \sum_{i=1}^{k} \alpha_i v_i, v_j \right\rangle = \sum_{i=1}^{k} \alpha_i \langle v_i, v_j \rangle = \alpha_j \langle v_j, v_j \rangle = \alpha_j \|v_j\|^2,$$

es decir, $\alpha_j = \frac{\langle w, v_j \rangle}{\|v_j\|^2}$. □

**Teorema 7.18.** *Sea $(V, \langle \cdot, \cdot \rangle)$ un espacio con producto interno. Si $S \subset V$ es un conjunto ortogonal de vectores no nulos, entonces $S$ es linealmente independiente.*

*Demostración.* Supongamos que $S$ es un conjunto ortogonal de vectores no nulos. Recordemos que, para probar que $S$ es linealmente independiente, alcanza con mostrar que todo subconjunto finito de $S$ es linealmente independiente. Sea $S_F := \{s_1, \ldots, s_k\}$ un subconjunto finito de $S$. Por la Proposición 7.17, como $S_F$ es ortogonal, todo vector del subespacio generado por $S_F$ admite una única representación como combinación lineal de los elementos de $S_F$.

En particular, si $\vec{0}$ se escribe como combinación lineal de los vectores de $S_F$ como $\vec{0} = \beta_1 s_1 + \beta_2 s_2 + \ldots + \beta_k s_k$, resulta que cada $\beta_j$ es:

$$\beta_j = \frac{\left\langle \vec{0}, s_j \right\rangle}{\|s_j\|^2} = \frac{0}{\|s_j\|^2} = 0.$$

Por lo tanto, $S_F$ es linealmente independiente. Como $S_F$ era un subconjunto finito arbitrario de $S$, tenemos que $S$ es linealmente independiente. □

Como consecuencia del teorema anterior, si $(V, \langle \cdot, \cdot \rangle)$ es un espacio con producto interno de dimensión finita, el cardinal de cualquier subconjunto ortogonal (de vectores no nulos) está acotado superiormente por $\dim V$.

**Teorema 7.19.** *Sea $\{b_1, \ldots, b_n\}$ un conjunto linealmente independiente de un espacio con producto interno $(V, \langle \cdot, \cdot \rangle)$. Entonces, existe un subconjunto ortogonal $\{v_1, \ldots, v_n\}$ en $V$ tal que, para cada $k = 1, \ldots, n$,*

$$\overline{\{b_1, \ldots, b_k\}} = \overline{\{v_1, \ldots, v_k\}}.$$

*Demostración.* Construiremos los vectores $v_1, \ldots, v_n$ recursivamente. En primer lugar, sea $v_1 := b_1$. Luego, supongamos construidos $v_1, \ldots, v_m$ tales que, para cada $k = 1, \ldots, m$,

$$\{v_1, \ldots, v_k\} \quad \text{es una base ortogonal de} \quad \overline{\{b_1, \ldots, b_k\}}.$$

Entonces, definimos

$$v_{m+1} := b_{m+1} - \sum_{i=1}^{m} \frac{\langle b_{m+1}, v_i \rangle}{\|v_i\|^2} \, v_i.$$

Notemos que $v_{m+1} \neq \vec{0}$ porque, en caso contrario, $b_{m+1} \in \overline{\{v_1, \ldots, v_m\}} = \overline{\{b_1, \ldots, b_m\}}$ lo que resulta absurdo. Además, si $j = 1, \ldots, m$,

$$\begin{aligned} \langle v_{m+1}, v_j \rangle &= \left\langle b_{m+1} - \sum_{i=1}^{m} \frac{\langle b_{m+1}, v_i \rangle}{\|v_i\|^2} \, v_i, v_j \right\rangle = \langle b_{m+1}, v_j \rangle - \sum_{i=1}^{m} \frac{\langle b_{m+1}, v_i \rangle}{\|v_i\|^2} \langle v_i, v_j \rangle \\ &= \langle b_{m+1}, v_j \rangle - \sum_{i=1}^{m} \frac{\langle b_{m+1}, v_i \rangle}{\|v_i\|^2} \delta_{ij} \langle v_j, v_j \rangle = \langle b_{m+1}, v_j \rangle - \frac{\langle b_{m+1}, v_j \rangle}{\|v_j\|^2} \|v_j\|^2 \\ &= \langle b_{m+1}, v_j \rangle - \langle b_{m+1}, v_j \rangle = 0. \end{aligned}$$

Entonces, $\{v_1, \ldots, v_{m+1}\}$ es un sistema ortogonal de vectores no nulos en $\overline{\{b_1, \ldots, b_{m+1}\}}$. Por el Teorema 7.18, resulta que $\{v_1, \ldots, v_{m+1}\}$ es una base de $\overline{\{b_1, \ldots, b_{m+1}\}}$, como queríamos probar. □

El método utilizado en la demostración anterior se conoce como **proceso de ortogonalización de Gram-Schmidt**.

**Corolario 7.20.** *Todo espacio con producto interno de dimensión finita tiene una base ortonormal.*

*Demostración.* Supongamos que $(V, \langle \cdot, \cdot \rangle)$ es un espacio con producto interno con $\dim V = n$, y que $\{b_1, \ldots, b_n\}$ es una base de $V$. Por el Teorema 7.19, existe una base ortogonal $\{v_1, \ldots, v_n\}$ de $V$. Luego, si para cada $i = 1, \ldots, n$ definimos $e_i = \frac{v_i}{\|v_i\|}$, el conjunto $\mathcal{E} = \{e_1, \ldots, e_n\}$ es una base ortonormal de $V$. □

**Ejemplo 7.21.** En $\mathbb{R}^3$, consideremos la base $\mathcal{B} = \{b_1, b_2, b_3\}$ dada por

$$b_1 = (3, 0, 4), \quad b_2 = (-1, 0, 7), \quad \text{y} \quad b_3 = (2, 9, 11).$$

Aplicando el proceso de ortogonalización de Gram-Schmidt con respecto al producto interno canónico, obtenemos

$$\begin{aligned}
v_1 &:= b_1 = (3, 0, 4), \\
v_2 &:= b_2 - \frac{\langle b_2, v_1 \rangle}{\|v_1\|^2} v_1 = (-1, 0, 7) - \frac{25}{25}(3, 0, 4) = (-4, 0, 3), \\
v_3 &:= b_3 - \frac{\langle b_3, v_1 \rangle}{\|v_1\|^2} v_1 - \frac{\langle b_3, v_2 \rangle}{\|v_2\|^2} v_2 = (2, 9, 11) - \frac{50}{25}(3, 0, 4) - \frac{25}{25}(-4, 0, 3) = (0, 9, 0).
\end{aligned}$$

Entonces, $\{v_1, v_2, v_3\}$ es una base ortogonal de $\mathbb{R}^3$. Normalizando los vectores, vemos que

$$\mathcal{E} = \left\{e_1 := (\tfrac{3}{5}, 0, \tfrac{4}{5}), e_2 := (-\tfrac{4}{5}, 0, \tfrac{3}{5}), e_3 := (0, 1, 0)\right\} \quad \text{es una base ortonormal de } \mathbb{R}^3.$$

El proceso de ortogonalización de Gram-Schmidt es una aplicación repetida de una operación geométrica básica, la **proyección ortogonal sobre un subespacio**. Este método también aparece de forma natural en la solución de un importante problema de aproximación, **el problema de cuadrados mínimos**, el cual presentamos a continuación.

Dado un espacio con producto interno $(V, \langle \cdot, \cdot \rangle)$, sean $W$ un subespacio de $V$ y $v \in V$. El problema de cuadrados mínimos consiste en encontrar, si existe alguno, un $w_0 \in W$ tal que

$$\|v - w_0\| = \min_{w \in W} \|v - w\|, \tag{7.5}$$

es decir, un $w_0 \in W$ tal que $\|v - w_0\| \leq \|v - w\|$ para todo $w \in W$.

**Teorema 7.22.** *Dado un espacio con producto interno $(V, \langle \cdot, \cdot \rangle)$, sean $W$ un subespacio de $V$ y $v \in V$. Entonces,*

*i)* *$w_0 \in W$ satisface (7.5) si y sólo si $\langle v - w_0, w \rangle = 0$ para todo $w \in W$.*

*ii)* *Si existe un $w_0 \in W$ que satisface (7.5), éste es único.*

*iii)* *Si $\dim W = n$ y $\{e_1, \ldots, e_n\}$ es una base ortonormal de $W$, entonces*

$$w_0 := \sum_{i=1}^{n} \langle v, e_i \rangle \, e_i \quad \textit{es la única solución de } (7.5).$$

*Demostración.* Dados $w, w_0 \in W$, notemos que $v - w = (v - w_0) + (w_0 - w)$ y

$$\begin{aligned}\|v-w\|^2 &= \langle (v-w_0)+(w_0-w), (v-w_0)+(w_0-w)\rangle \\ &= \|v-w_0\|^2 + 2\operatorname{Re}(\langle v-w_0, w_0-w\rangle) + \|w_0-w\|^2. \end{aligned} \tag{7.6}$$

En primer lugar, probemos $i$). Supongamos que $w_0 \in W$ es tal que $\langle v-w_0, w\rangle = 0$ para todo $w \in W$. Entonces, como $w - w_0 \in W$, la igualdad anterior se transforma en

$$\|v-w\|^2 = \|v-w_0\|^2 + \|w_0-w\|^2,$$

y el valor mínimo del miembro derecho se alcanza cuando $w = w_0$, es decir $w_0$ satisface (7.5).

Recíprocamente, si $w_0 \in W$ es tal que $\|v-w_0\| \leq \|v-w\|$ para todo $w \in W$, de (7.6) se deduce que

$$2\operatorname{Re}(\langle v-w_0, w_0-w\rangle) + \|w_0-w\|^2 \geq 0 \qquad \text{para todo } w \in W.$$

Llamando $u = w_0 - w$, la ecuación anterior se lee como

$$2\operatorname{Re}(\langle v-w_0, u\rangle) + \|u\|^2 \geq 0 \qquad \text{para todo } u \in W. \tag{7.7}$$

Fijado un vector $u \in W$, $w \neq \vec{0}$ arbitrario, consideremos la función $f : \mathbb{R} \to \mathbb{R}$ dada por

$$f(t) = 2\operatorname{Re}(\langle v-w_0, tu\rangle) + \|tu\|^2 = 2t\operatorname{Re}(\langle v-w_0, u\rangle) + t^2\|u\|^2,$$

la cual toma valores no negativos, como consecuencia de (7.7). Pero llamando $A = \|u\|^2$ y $B = 2\operatorname{Re}(\langle v-w_0, u\rangle)$, vemos que $f$ es una función cuadrática, $f(t) = At^2 + Bt$ cuyo coeficiente principal $A$ es positivo y $f(0) = 0$. Luego, para que $f$ tome sólo valores no negativos es necesario que $B^2 - 4AC = 0$.

Por lo tanto, para que $f$ tome sólo valores no negativos es necesario que $B = 0$, es decir, $\operatorname{Re}(\langle v-w_0, u\rangle) = 0$.

De la misma manera, considerando la función $g : \mathbb{R} \to \mathbb{R}$ dada por

$$\begin{aligned} g(t) &= 2\operatorname{Re}(\langle v-w_0, itu\rangle) + \|itu\|^2 = 2t\operatorname{Re}(-i\,\langle v-w_0, u\rangle) + t^2\|u\|^2 \\ &= 2t\operatorname{Im}(\langle v-w_0, u\rangle) + t^2\|u\|^2, \end{aligned}$$

argumentando de manera análoga se deduce que $\operatorname{Im}(\langle v-w_0, u\rangle) = 0$. Por lo tanto,

$$\langle v-w_0, u\rangle = 0.$$

Esto completa la demostración del ítem $i$).

Supongamos que $w_0, w_1 \in W$ son dos vectores que verifican (7.5). Por el ítem $i$), resulta que

$$\langle v-w_0, w\rangle = 0 \quad \text{y} \quad \langle v-w_1, w\rangle = 0 \qquad \text{para todo } w \in W.$$

Entonces, para todo $w \in W$ tenemos que

$$\langle w_1-w_0, w\rangle = \langle (v-w_0)-(v-w_1), w\rangle = \langle v-w_0, w\rangle - \langle v-w_1, w\rangle = 0.$$

Luego, como $(W, \langle\cdot,\cdot\rangle)$ es un espacio con producto interno y $w_1 - w_0$ es ortogonal a todos los vectores de $W$, resulta que $w_1 - w_0 = \vec{0}$, es decir, $w_1 = w_0$.

Por último, supongamos que $\{e_1, \ldots, e_n\}$ es una base ortonormal de $W$. Definamos al vector $w_0 = \sum_{i=1}^{n} \langle v, e_i\rangle\, e_i \in W$ y veamos que $v - w_0$ es ortogonal a todos los vectores de

$W$. Para ello, alcanza probar que es ortogonal a los vectores de la base ortonormal, pero para cada $j = 1, \ldots, n$,

$$\begin{aligned} \langle v - w_0, e_j \rangle &= \left\langle v - \sum_{i=1}^{n} \langle v, e_i \rangle e_i, e_j \right\rangle = \langle v, e_j \rangle - \sum_{i=1}^{n} \langle v, e_i \rangle \langle e_i, e_j \rangle \\ &= \langle v, e_j \rangle - \langle v, e_j \rangle = 0. \end{aligned}$$

Luego, por el ítem $i)$ podemos asegurar que $w_0$ satisface (7.5). □

**Definición 7.23.** Dado un espacio con producto interno $(V, \langle \cdot, \cdot \rangle)$, sea $S$ un subconjunto cualquiera de $V$. El **complemento ortogonal de** $S$ es el conjunto definido por

$$S^{\perp} := \{v \in V : \langle v, s \rangle = 0 \text{ para todo } s \in S\}.$$

De la definición se sigue que $\{\vec{0}\}^{\perp} = V$ y que $V^{\perp} = \{\vec{0}\}$. Además,

**Proposición 7.24.** *Dado un espacio con producto interno $(V, \langle \cdot, \cdot \rangle)$, si $S$ es un subconjunto cualquiera de $V$ entonces $S^{\perp}$ es un subespacio de $V$.*

*Demostración.* Ejercicio. □

El siguiente resultado justifica el uso del término "complemento" para el ortogonal a un subespacio.

**Teorema 7.25.** *Dado un espacio con producto interno $(V, \langle \cdot, \cdot \rangle)$ de dimensión finita, sea $W$ un subespacio de $V$. Entonces,*

$$V = W \dotplus W^{\perp}. \tag{7.8}$$

*Demostración.* Supongamos que $\dim V = n$ y que $\dim W = m < n$. Sea $\mathcal{B} = \{b_1, \ldots, b_n\}$ una base de $V$ tal que $\{b_1, \ldots, b_m\}$ es una base de $W$ (recordemos que, para encontrar una base así, podemos comenzar con una base $\{b_1, \ldots, b_m\}$ de $W$ y luego extenderla a una base de $V$).

Luego, podemos ortogonalizarla usando el proceso de Gram-Schmidt (Teorema 7.19), obteniendo una base ortogonal $\mathcal{E} = \{e_1, \ldots, e_n\}$ de $V$ tal que $\{e_1, \ldots, e_m\}$ es una base ortogonal de $W$.

Veamos que $\{e_{m+1}, \ldots, e_n\}$ es una base ortogonal de $W^{\perp}$. Si $j = m+1, \ldots, n$, tenemos que

$$\langle e_j, e_i \rangle = 0 \quad \text{para } i = 1, \ldots, m.$$

Entonces $\langle e_j, w \rangle = 0$ para todo $w \in \overline{\{e_1, \ldots, e_n\}} = W$, es decir, $e_j \in W^{\perp}$. Por lo tanto, $\overline{\{e_{m+1}, \ldots, e_n\}} \subseteq W^{\perp}$.

Recíprocamente, supongamos que $v \in W^{\perp}$. Escribiéndolo con respecto a la base $\mathcal{E}$, vemos que

$$v = \sum_{i=1}^{n} \frac{\langle v, e_i \rangle}{\|e_i\|^2} e_i = \sum_{i=m+1}^{n} \frac{\langle v, e_i \rangle}{\|e_i\|^2} e_i,$$

porque $\langle v, e_i \rangle = 0$ para $i = 1, \ldots, m$. Entonces, $v \in \overline{\{e_{m+1}, \ldots, e_n\}}$.

Por lo tanto, $W^{\perp} = \overline{\{e_{m+1}, \ldots, e_n\}}$ y, en consecuencia, $V = W \dotplus W^{\perp}$. □

La suma directa de dos subespacios ortogonales se denomina **suma directa ortogonal**, y se anota con el símbolo $\oplus$ en lugar de $\dotplus$. En particular, el teorema anterior dice que

$$V = W \oplus W^{\perp},$$

para cualquier subespacio $W$ de un espacio con producto interno $(V, \langle \cdot, \cdot \rangle)$.

**Ejercicio 7.26.** Dado un subconjunto $S$ cualquiera de un espacio con producto interno $(V, \langle\cdot,\cdot\rangle)$, probar que el "doble ortogonal" de $S$ coincide con el subespacio generado por $S$:

$$(S^{\perp})^{\perp} = \overline{S}.$$

**Definición 7.27.** Dado un espacio con producto interno $(V, \langle\cdot,\cdot\rangle)$ de dimensión finita, sea $W$ un subespacio de $V$. La **proyección ortogonal sobre** $W$ es la proyección asociada a la descomposición $V = W \oplus W^{\perp}$, con imagen $W$ y núcleo $W^{\perp}$:

$$P_W := P_{W//W^{\perp}}.$$

Notemos que la proyección complementaria a una proyección ortogonal también es una proyección ortogonal:

$$I - P_W = P_{W^{\perp}}.$$

Dado un vector arbitrario $v \in V$, como $P_W v$ y $P_{W^{\perp}} v$ son ortogonales, el Teorema de Pitágoras dice que

$$\|v\|^2 = \|P_W v\|^2 + \|P_{W^{\perp}} v\|^2.$$

Pero como además $\langle v - P_W v, w\rangle = \langle P_{W^{\perp}} v, w\rangle = 0$ para todo $w \in W$, resulta que

$$\|v - P_W v\|^2 = \min_{w\in W} \|v - w\|^2,$$

es decir, $P_W v$ es la (única) solución al problema de cuadrados mínimos, y la distancia del vector $v$ al subespacio $W$ está dada por $\|P_{W^{\perp}} v\|$.

Otras consecuencias inmediatas de la definición de la proyección ortogonal sobre un subespacio son la **desigualdad de Bessel** y la **identidad de Parseval**:

**Proposición 7.28.** *Si $\{e_1, \ldots, e_m\}$ es un conjunto ortonormal de un espacio con producto interno $(V, \langle\cdot,\cdot\rangle)$, entonces*

$$\sum_{i=1}^{m} |\langle v, e_i\rangle|^2 \leq \|v\|^2 \qquad \textit{para todo } v \in V. \tag{7.9}$$

*Además, si $\mathcal{E} = \{e_1, \ldots, e_n\}$ es una base ortonormal de $V$, entonces*

$$\sum_{i=1}^{n} |\langle v, e_i\rangle|^2 = \|v\|^2 \qquad \textit{para todo } v \in V. \tag{7.10}$$

*Demostración.* Supongamos que $\{e_1, \ldots, e_m\}$ es un conjunto ortonormal de un espacio con producto interno $(V, \langle\cdot,\cdot\rangle)$, y sea $W$ el subespacio generado por estos vectores, $W = \overline{\{e_1, \ldots, e_m\}}$. Dado un vector arbitrario $v \in V$, tenemos que

$$\|v\|^2 = \|P_W v\|^2 + \|P_{W^{\perp}} v\|^2 \geq \|P_W v\|^2.$$

Además, como $P_W v$ es la única solución a (7.5), comparando con el ítem *iii*) del Teorema 7.22 tenemos que $P_W v = w_0 = \sum_{i=1}^{m} \langle v, e_i\rangle\, e_i$. Por lo tanto,

$$\begin{aligned}
\|v\|^2 \geq \|P_W v\|^2 &= \left\|\sum_{i=1}^{m} \langle v, e_i\rangle\, e_i\right\|^2 = \left\langle \sum_{i=1}^{m} \langle v, e_i\rangle\, e_i, \sum_{j=1}^{m} \langle v, e_j\rangle\, e_j \right\rangle \\
&= \sum_{i=1}^{m} \langle v, e_i\rangle \left\langle e_i, \sum_{j=1}^{m} \langle v, e_j\rangle\, e_j \right\rangle = \sum_{i=1}^{m} \langle v, e_i\rangle \sum_{j=1}^{m} \overline{\langle v, e_j\rangle}\, \langle e_i, e_j\rangle \\
&= \sum_{i=1}^{m} \langle v, e_i\rangle\, \overline{\langle v, e_i\rangle} = \sum_{i=1}^{m} |\langle v, e_i\rangle|^2.
\end{aligned}$$

Por último, si $\mathcal{E} = \{e_1, \dots, e_n\}$ es una base ortonormal de $V$, dado un vector arbitrario $v \in V$ tenemos que $v = \sum_{i=1}^n \langle v, e_i\rangle e_i$ y, repitiendo el argumento anterior,

$$\|v\|^2 = \left\| \sum_{i=1}^n \langle v, e_i\rangle e_i \right\|^2 = \sum_{i=1}^n |\langle v, e_i\rangle|^2.$$

□

Como anunciamos anteriormente, la desigualdad (7.9) se denomina desigualdad de Bessel y a la igualdad (7.10) se la conoce como la identidad de Parseval.

## 7.2. Funcionales lineales en un espacio con producto interno

En esta sección veremos que, para un espacio con producto interno $(V, \langle\cdot,\cdot\rangle)$ de dimensión finita, los funcionales lineales sobre $V$ pueden representarse de una forma muy particular. En primer lugar, notemos que cada vector $w \in V$ define un funcional lineal sobre $V$:

$$f_w : V \to \mathbb{K} \qquad \text{definido por} \qquad f_w(v) = \langle v, w\rangle .$$

**Ejemplo 7.29.** En el Ejemplo 7.21 vimos que la ortogonalización de la base $\mathcal{B} = \{b_1, b_2, b_3\}$ mediante el proceso de Gram-Schmidt daba como resultado la base ortogonal

$$\mathcal{B}_0 = \{v_1 = (3,0,4), v_2 = (-4,0,3), v_3 = (0,9,0)\}.$$

Luego, dado un vector arbitrario $(x_1, x_2, x_3) \in \mathbb{R}^3$, tenemos que

$$(x_1,x_2,x_3) = \sum_{i=1}^3 \frac{\langle (x_1,x_2,x_3), v_i\rangle}{\|v_i\|^2} v_i = \frac{3x_1+4x_3}{25} v_1 + \frac{-4x_1+3x_3}{25} v_2 + \frac{9x_2}{81} v_3.$$

En particular, la base dual de $\mathcal{B}_0$ está dada por $\mathcal{B}_0^* = \{f_1, f_2, f_3\}$, donde los funcionales están dados explícitamente por

$$\begin{aligned}
f_1(x_1,x_2,x_3) &= \frac{3x_1+4x_3}{25} = \left\langle (x_1,x_2,x_3), (\tfrac{3}{25}, 0, \tfrac{4}{25})\right\rangle ,\\
f_2(x_1,x_2,x_3) &= \frac{-4x_1+3x_3}{25} = \left\langle (x_1,x_2,x_3), (\tfrac{-4}{25}, 0, \tfrac{3}{25})\right\rangle ,\\
f_3(x_1,x_2,x_3) &= \frac{9x_2}{81} = \left\langle (x_1,x_2,x_3), (0, \tfrac{1}{9}, 0)\right\rangle .
\end{aligned}$$

Por lo tanto, los funcionales de la base dual son de la clase de funcionales que se realizan haciendo el producto interno contra un vector.

**Ejemplo 7.30.** La función $\varphi : \mathbb{R}^{(2)}[x] \to \mathbb{R}$ definida por

$$\varphi(p) = \int_{-1}^1 p(t)\cos(\pi t)dt,$$

es un funcional lineal sobre $\mathbb{R}^{(2)}[x]$ (dotado con el producto interno $\langle p, q\rangle = \int_0^1 p(t)q(t)dt$). En este caso no es obvio que exista un polinomio $p_0 \in \mathbb{R}^{(2)}[x]$ tal que

$$\varphi(p) = \langle p, p_0\rangle ,$$

para todo $p \in \mathbb{R}^{(2)}[x]$. El candidato natural sería la función $u(t) = \cos(\pi t)$, pero éste no es un elemento de $\mathbb{R}^{(2)}[x]$.

El resultado siguiente muestra que todo funcional sobre un espacio con producto interno de dimensión finita se realiza haciendo el producto interno contra un vector. El ejemplo anterior ilustra la profundidad de este teorema, porque para el funcional lineal presentado en dicho ejemplo no hay un candidato obvio para el vector que lo representa.

**Teorema 7.31 (Teorema de representación de Riesz).** *Sea $(V, \langle\cdot,\cdot\rangle)$ un espacio con producto interno de dimensión finita. Para cada $f \in V^*$ existe un único $w \in V$ tal que*

$$f(v) = \langle v, w\rangle \qquad \textit{para todo } v \in V. \tag{7.11}$$

Este resultado recibe el nombre del matemático húngaro Frigyes Riesz (1880–1956), quien probó varios resultados muy similares al teorema en cuestión, a comienzos del siglo XX. Una forma de probarlo sería convencerse de que la función $v \mapsto \langle\cdot, w\rangle$ es un isomorfismo de $V$ en $V^*$, pero la demostración de más abajo es aún más explícita.

*Demostración.* Veamos primero que, fijado $f \in V^*$, hay un vector $w \in V$ tal que $f(v) = \langle v, w\rangle$ para todo $v \in V$. Si $\mathcal{E} = \{e_1, \ldots, e_n\}$ es una base ortonormal de $V$, cada vector $v \in V$ se representa de manera única como $v = \sum_{i=1}^n \langle v, e_i\rangle\, e_i$. Entonces,

$$\begin{aligned} f(v) &= f\left(\sum_{i=1}^n \langle v, e_i\rangle\, e_i\right) = \sum_{i=1}^n \langle v, e_i\rangle\, f(e_i) = \sum_{i=1}^n \left\langle v, \overline{f(e_i)} e_i\right\rangle \\ &= \left\langle v, \sum_{i=1}^n \overline{f(e_i)} e_i\right\rangle, \end{aligned}$$

donde en un primer paso usamos la linealidad de $f$, y luego usamos que el producto interno es conjugado lineal en la segunda entrada.

Por lo tanto, dado $f \in V^*$, el vector $w := \sum_{i=1}^n \overline{f(e_i)} e_i$ es tal que $f(v) = \langle v, w\rangle$ para todo $v \in V$.

Veamos ahora que este $w$ es el único vector que verifica (7.11). Supongamos que existe $w' \in V$ tal que $f(v) = \langle v, w'\rangle$ para todo $v \in V$. Luego,

$$\langle v, w - w'\rangle = \langle v, w\rangle - \langle v, w'\rangle = f(v) - f(v) = 0 \qquad \text{para todo } v \in V,$$

es decir, $w - w' \perp v$ para todo $v \in V$. Por lo tanto, $w - w' = \vec{0}$, es decir, $w' = w$. □

**Observación 7.32.** Supongamos que $(V, \langle\cdot,\cdot\rangle)$ es un espacio con producto interno con $\dim V = n$. Dado un funcional no nulo $f \in V^*$, sabemos que su núcleo es un hiperespacio de $V$, es decir, $\dim N(f) = n - 1$. Veamos que el vector $w \in V$ que permite representar a $f$ de acuerdo a (7.11), genera el complemento ortogonal de $N(f)$.

Dado $v \in V$, tenemos que $v \in N(f)$ si y sólo si $\langle v, w\rangle = 0$. Luego, $w \perp v$ para todo $v \in N(f)$, o sea, $w \in N(f)^\perp$. Además, como $\dim N(f)^\perp = \dim V - \dim N(f) = n - (n-1) = 1$, resulta que

$$N(f)^\perp = \overline{\{w\}}.$$

**Ejercicio 7.33.** Calcular el polinomio $p_0 \in \mathbb{R}^{(2)}[x]$ que realiza al funcional lineal del Ejemplo 7.30.

Un último comentario antes de finalizar esta sección. Si el espacio con producto interno $(V, \langle\cdot,\cdot\rangle)$ es un $\mathbb{R}$-espacio vectorial, la función $\Phi : V^* \to V$ que asigna a cada $f \in V^*$ el único $w \in V$ que verifica (7.11) es una transformación lineal.

En cambio, si $V$ es un $\mathbb{C}$-espacio vectorial, la función $\Phi : V^* \to V$ **no** es una transformación lineal. Veamos que $\Phi$ es una transformación **conjugado lineal**, es decir.

$$\Phi(\alpha f + \beta g) = \overline{\alpha}\Phi(f) + \overline{\beta}\Phi(g),$$

para cualesquiera $f, g \in V^*$ y $\alpha, \beta \in \mathbb{C}$.

Dados $f, g \in V^*$, sean $w, u \in V$ (los únicos) tales que

$$f = \langle \cdot, w \rangle \quad \text{y} \quad g = \langle \cdot, u \rangle ,$$

o equivalentemente, $\Phi(f) = w$ y $\Phi(g) = u$. Luego, si $\alpha, \beta \in \mathbb{C}$ y $v \in V$,

$$(\alpha f + \beta g)(v) = \alpha f(v) + \beta g(v) = \alpha \langle v, w \rangle + \beta vu = \left\langle v, \overline{\alpha} w + \overline{\beta} u \right\rangle ,$$

es decir, $\Phi(\alpha f + \beta g) = \overline{\alpha} w + \overline{\beta} u = \overline{\alpha}\Phi(f) + \overline{\beta}\Phi(g)$, como queríamos probar.

## 7.3. Ejercicios

1. Determinar, en cada caso, si $(V, \langle \cdot, \cdot \rangle)$ es un espacio vectorial con producto interno.
   *a)* $V = \mathbb{R}^2$ dotado con $\langle (x_1, y_1), (x_2, y_2) \rangle = x_1 x_2 - x_1 y_2 - y_1 x_2 + 3 y_1 y_2$.
   *b)* $V = \mathbb{R}^{2\times 2}$ dotado con $\langle A, B \rangle = a_{11} b_{11} + a_{12} b_{12} + a_{21} b_{21} + a_{22} b_{22}$.
   *c)* $V = \mathbb{C}^2$ dotado con $\langle (x_1, y_1), (x_2, y_2) \rangle = T(x_1, y_1)\,\overline{T(x_2, y_2)}$, siendo $T \in L(\mathbb{C}^2, \mathbb{C})$.
   *d)* $V = \mathbb{C}[x]$ dotado con $\langle p, q \rangle = \int_0^1 p(t)\overline{q(t)}\, dt$.
2. Sean $V$ y $W$ dos $\mathbb{K}$-espacios vectoriales y sea $\langle \cdot, \cdot \rangle$ un producto interno sobre $W$. Probar que si $T \in L(V, W)$ es un isomorfismo de espacios vectoriales, entonces $\langle \cdot, \cdot \rangle_T$ dado por
   $$\langle v_1, v_2 \rangle_T = \langle T v_1, T v_2 \rangle \quad \text{para } v_1, v_2 \in V ,$$
   es un producto interno sobre $V$.
3. Determinar qué condiciones debe cumplir una matriz $B \in \mathbb{R}^{n\times n}$ para que
   $$\langle x, y \rangle = x^t B y ,$$
   sea un producto interno sobre $\mathbb{R}^n$.
4. Sea $V$ un $\mathbb{C}$-espacio vectorial dotado con un producto interno. Probar que la norma determinada por el producto interno cumple la "identidad del paralelogramo":
   $$\|v + w\|^2 + \|v - w\|^2 = 2(\|v\|^2 + \|w\|^2) \quad \text{para todo } v, w \in V .$$
5. Sea $(V, \langle \cdot, \cdot \rangle)$ un espacio con producto interno. Si d es la distancia en $(V, \langle \cdot, \cdot \rangle)$ dada por (7.4), probar que verifica las siguientes propiedades: dados $u, v \in V$,
   i) $\mathrm{d}(u, v) = 0$ si y sólo si $v = u$;
   ii) $\mathrm{d}(u, v) = \mathrm{d}(v, u)$;
   iii) $\mathrm{d}(u, v) \leq \mathrm{d}(u, w) + \mathrm{d}(w, v)$ para cualquier $w \in V$.
6. Consideremos a $\mathbb{R}^2$ dotado con el producto interno usual, y a los vectores $v_1 = (1, 2)$ y $v_2 = (-1, 1)$.

*a*) Probar que existe $w \in \mathbb{R}^2$ tal que $\langle v_1, w\rangle = -1$ y $\langle v_2, w\rangle = 3$. ¿Es único?

*b*) Si $\mathcal{E} = \{e_1, e_2\}$ es la base canónica de $\mathbb{R}^2$, probar que

$$v = \langle v, e_1\rangle e_1 + \langle v, e_2\rangle e_2 \quad \text{para todo } v \in \mathbb{R}^2\,.$$

7. *a*) Consideremos el $\mathbb{R}$-espacio vectorial $\mathbb{R}^{n\times m}$. Probar que la aplicación

$$\langle A, B\rangle = \operatorname{tr}(B^t A) \quad \text{para } A, B \in \mathbb{R}^{n\times m}\,,$$

es un producto interno para $\mathbb{R}^{n\times m}$.

*b*) Consideremos ahora al $\mathbb{C}$-espacio vectorial $\mathbb{C}^{n\times m}$ y la aplicación

$$\langle A, B\rangle = \operatorname{tr}(B^t A) \quad \text{para } A, B \in \mathbb{C}^{n\times m}\,,$$

¿es un producto interno para $\mathbb{C}^{n\times m}$? ¿De qué manera puede modificarse para que lo sea?

8. Consideremos a $\mathbb{C}^{2\times 2}$ dotado con el producto interno

$$\langle A, B\rangle = \operatorname{tr}(\overline{B}^t A) \quad \text{para } A, B \in \mathbb{C}^{2\times 2}\,,$$

*a*) Hallar la distancia entre

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

*b*) Hallar el complemento ortogonal del subespacio de $\mathbb{C}^{2\times 2}$ generado por $A$ y $B$.

9. ¿Qué dimensión tiene el subespacio de $\mathbb{R}^6$ que es ortogonal a los vectores $(1, 1, -2, 3, 4, 5)$ y $(0, 0, 1, 1, 0, 7)$?

10. Consideremos a $\mathbb{C}^8$ dotado con el producto interno usual y sea $\mathcal{E} = \{e_1, e_2, \ldots, e_8\}$ la base canónica. Para $j = 1, \ldots, 6$, sea

$$f_j = e_j - 2e_{j+2}.$$

Si $S = \{f_1, f_2, f_3, f_4, f_5, f_6\}$, hallar una base para $S^\perp$.

11. Consideremos a $\mathbb{R}^3$ dotado con el producto interno usual.

*a*) Hallar el subespacio ortogonal de $W = \big\{(x, y, z) \in \mathbb{R}^3 : 3x - y + 2z = 0\big\}$.

*b*) Hallar una base ortogonal de $\mathbb{R}^3$ que contenga al vector $(1, -1, 2)$.

*c*) Consideremos la base $\mathcal{B} = \{(1, 0, 1), (2, 0, 1), (1, 1, 0)\}$ de $\mathbb{R}^3$. A partir de $\mathcal{B}$, hallar una base $\mathcal{B}'$ que sea ortonormal y calcular las coordenadas de $(2, -1, 3)$ en esta nueva base.

12. Sea $V = \mathbb{C}^{(3)}[x]$ dotado con el producto interno

$$\langle p, q\rangle = \int_0^1 p(t)\overline{q(t)}\,dt \quad \text{para } p, q \in V\,.$$

*a*) Sean $p(x) = x + 2$ y $q(x) = x^2 - 2x - 3$. Calcular $\langle p, q\rangle$, $\|p\|$ y $\|q\|$.

*b*) Aplicar el proceso de ortogonalización de Gram-Schmidt a $\mathcal{B} = \big\{1, x, x^2, x^3\big\}$, para obtener una base ortonormal de $V$.

*c*) ¿Qué sucede si hacemos lo mismo para la base $\mathcal{B}$ reordenada de la siguiente manera, $\mathcal{B}' = \{x, 1, x^2, x^3\}$?

13. Hallar una base ortonormal de $\mathbb{C}^3$ formada por autovectores de

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

14. Sea $(V, \langle \cdot, \cdot \rangle)$ un espacio con producto interno (de dimensión finita) y sea $S$ un subconjunto de $V$, probar que $S^\perp$ es un subespacio de $V$ y que

$$(S^\perp)^\perp = \overline{S}.$$

15. Calcular el polinomio $p_0 \in \mathbb{R}^{(2)}[x]$ que realiza al funcional lineal del Ejemplo 7.30.

16. Dado un $\mathbb{K}$-espacio con producto interno $(V, \langle \cdot, \cdot \rangle)$ con $\dim V = n$, sean $R$ y $S$ dos subespacios de $V$.

   *a*) Probar que $V = (R + S^\perp) \oplus S \cap R^\perp$.

   *b*) Mostrar que las siguientes condiciones son equivalentes:

      (i) $\dim(R) = \dim(S)$;

      (ii) $\dim(R^\perp) = \dim(S^\perp)$;

      (iii) $\dim(R \cap S^\perp) = \dim(S \cap R^\perp)$.

17. Dado un espacio con producto interno $(V, \langle \cdot, \cdot \rangle)$, para cada $u \in V$ sea $\varphi_u : V \to \mathbb{K}$ el funcional lineal definido por

$$\varphi_u(v) = \langle v, u \rangle \quad \text{para todo } v \in V.$$

Luego, consideremos la función $\varphi : V \to V^*$ definida por $\varphi(u) = \varphi_u$ para $u \in V$.

   *a*) Probar que si $\mathbb{K} = \mathbb{R}$ entonces $\varphi : V \to V^*$ es una transformación lineal. ¿Ocurre lo mismo si $\mathbb{K} = \mathbb{C}$?

   *b*) Probar que $\varphi$ es inyectiva.

   *c*) Suponiendo que $\mathbb{K} = \mathbb{R}$ y que $V$ es de dimensión finita, deducir que $\varphi$ es un isomorfismo entre $V$ y $V^*$ (sin utilizar el Teorema de representación de Riesz).

# Capítulo 8

# Transformaciones lineales entre espacios con producto interno

Los resultados más profundos relacionados con espacios con producto interno tratan sobre transformaciones u operadores lineales sobre un espacio con producto interno. Utilizando las propiedades de la transformación adjunta, describiremos detalladamente varias clases de operadores lineales sobre un espacio con producto interno.

Comencemos mostrando que, en un espacio con producto interno de dimensión finita, las coordenadas de una transformación lineal con respecto a bases ortonormales pueden calcularse utilizando el producto interno.

**Teorema 8.1.** *Dados dos espacios con producto interno $(V, \langle \cdot, \cdot \rangle)$ y $(W, \langle \cdot, \cdot \rangle)$ con $\dim V = n$ y $\dim W = m$, sean $\mathcal{E} = \{e_1, \ldots, e_n\}$ y $\mathcal{F} = \{f_1, \ldots, f_m\}$ bases ortonormales de $V$ y $W$, respectivamente. Si $T \in L(V, W)$ entonces las entradas de la matriz $A := [T]_{\mathcal{F},\mathcal{E}} \in \mathbb{K}^{m \times n}$ se calculan como:*

$$A_{ij} = \langle Te_j, f_i \rangle, \qquad i = 1, \ldots, m, \; j = 1, \ldots, n.$$

*Demostración.* Dado $j = 1, \ldots, n$, recordemos que la $j$-ésima columna de la matriz $A$ está compuesta por las coordenadas del vector $Te_j$ con respecto a la base $\mathcal{F}$.

Como $\mathcal{F}$ es una base ortonormal, cada $w \in W$ se escribe (con respecto a esta base) como $w = \sum_{i=1}^{m} \langle w, f_i \rangle f_i$. En particular, si $j = 1, \ldots, n$,

$$Te_j = \sum_{i=1}^{m} \langle Te_j, f_i \rangle f_i.$$

Entonces,

$$[Te_j]_{\mathcal{F}} = (\langle Te_j, f_1 \rangle, \langle Te_j, f_2 \rangle, \ldots, \langle Te_j, f_m \rangle)^\top.$$

Por lo tanto, $A_{ij} = \langle Te_j, f_i \rangle$. □

El resultado anterior también vale para una base de $V$ que no sea ortonormal, pero lo enunciamos así porque será utilizado de esta forma más adelante.

Notemos que si un operador $T \in L(V)$ verifica que $\langle Tv, w \rangle = 0$ para todo $v, w \in V$ entonces $T$ es el operador nulo. Esto se debe a que, para cada $v \in V$, el vector $Tv$ resulta ortogonal a todo $w \in V$ y por lo tanto $Tv = 0$. El próximo resultado muestra que, si $V$ es un $\mathbb{C}$-espacio con producto interno, es suficiente que $\langle Tv, v \rangle$ se anule para todo $v \in V$ para garantizar que $T$ es el operador nulo.

**Proposición 8.2.** *Dado un $\mathbb{C}$-espacio con producto interno $(V, \langle \cdot, \cdot \rangle)$, sea $T \in L(V)$. Entonces,*

$$T = 0 \qquad \textit{si y sólo si} \qquad \langle Tv, v \rangle = 0 \textit{ para todo } v \in V.$$

*Demostración.* Si $T = 0$ es inmediato que $\langle Tv, v \rangle = \left\langle \vec{0}, v \right\rangle = 0$ para todo $v \in V$. Para probar la recíproca, notemos que si $T$ es lineal entonces la función $[\cdot, \cdot] : V \times V \to \mathbb{K}$ definida por

$$[v, w] := \langle Tv, w \rangle, \qquad v, w \in V,$$

es lineal en la primera entrada y conjugado lineal en la segunda, con lo cual verifica la identidad de polarización (revisar la demostración de (7.3)). Luego, si suponemos que $\langle Tv, v \rangle = 0$ para todo $v \in V$, dados $u, w \in V$ cualesquiera,

$$\langle Tu, w \rangle = [u, w] = \frac{1}{4} \sum_{k=0}^{3} i^k \left[ u + i^k w, u + i^k w \right] = \frac{1}{4} \sum_{k=0}^{3} i^k \left\langle T(u + i^k w), u + i^k w \right\rangle = 0.$$

Entonces $Tu \perp w$ para todo $w \in V$, es decir, $Tu = \vec{0}$. Como $u \in V$ también era arbitrario, resulta que $T = 0$. $\square$

En cambio, si $V$ es un $\mathbb{R}$-espacio con producto interno, es fácil construir ejemplos de operadores no nulos que verifican $\langle Tv, v \rangle = 0$ para todo $v \in V$.

**Ejemplo 8.3.** Sea $T \in L(\mathbb{R}^2)$ el operador que representa a una rotación de $90°$ en sentido antihorario alrededor del origen, el cual está dado por

$$T(x, y) = (-y, x), \qquad (x, y) \in \mathbb{R}^2.$$

Por un lado, para cada $(x, y) \in \mathbb{R}^2$ tenemos que

$$\langle T(x, y), (x, y) \rangle = \langle (-y, x), (x, y) \rangle = -yx + xy = 0.$$

Por otro lado, $T(x, y) = (0, 0)$ si y sólo si $x = y = 0$.

## 8.1. La transformación adjunta

**Definición 8.4.** Sean $(V, \langle \cdot, \cdot \rangle)$ y $(W, \langle \cdot, \cdot \rangle)$ dos espacios con producto interno. Dada una transformación $T \in L(V, W)$, a la (única) transformación lineal $T^* \in L(W, V)$ que satisface

$$\langle Tv, w \rangle = \langle v, T^* w \rangle, \quad \text{para todo } v \in V \text{ y todo } w \in W, \tag{8.1}$$

se la denomina **transformación adjunta de** $T$.

Es importante notar que en (8.1) aparecen involucrados dos productos internos distintos (el de $V$ y el de $W$). Aunque usemos la misma notación para ambos, esto no significa que sean el mismo producto interno, es simplemente un abuso de notación. Otra consecuencia inmediata de (8.1) es que $T^*$ no sólo depende de la transformación $T$ sino también de los productos internos de los espacios correspondientes.

Antes de continuar, verifiquemos que toda transformación lineal entre $\mathbb{K}$-espacios vectoriales de dimensión finita tiene una (única) transformación adjunta.

**Teorema 8.5.** *Sean $(V, \langle \cdot, \cdot \rangle)$ y $(W, \langle \cdot, \cdot \rangle)$ dos espacios con producto interno de dimensión finita. Si $T \in L(V, W)$ entonces existe una única $T^* \in L(W, V)$ tal que*

$$\langle Tv, w \rangle = \langle v, T^* w \rangle, \quad \textit{para todo } v \in V \textit{ y todo } w \in W.$$

*Demostración.* Dado $w \in W$ cualquiera, sea $g_w : V \to \mathbb{K}$ definida por

$$g_w(v) = \langle Tv, w \rangle , \quad v \in V.$$

Es fácil ver que $g_w \in V^*$. Luego, por el Teorema 7.31, existe un único $v' \in V$ tal que $g_w(v) = \langle v, v' \rangle$. Esto nos permite ver que $T^* : W \to V$ dada por $T^*w = v'$, es una función bien definida. Además,

$$\langle Tv, w \rangle = g_w(v) = \left\langle v, v' \right\rangle = \langle v, T^*w \rangle , \quad \text{para todo } v \in V.$$

Veamos ahora que $T^*$ es una transformación lineal: dados $w, u \in W$ y $\alpha \in \mathbb{K}$, para cualquier $v \in V$,

$$\begin{aligned} \langle v, T^*(\alpha w + u) \rangle &= \langle Tv, \alpha w + u \rangle = \overline{\alpha} \langle Tv, w \rangle + \langle Tv, u \rangle = \overline{\alpha} \langle v, T^*w \rangle + \langle v, T^*u \rangle \\ &= \langle v, \alpha T^*w + T^*u \rangle . \end{aligned}$$

Como $v \in V$ era arbitrario, resulta que $T^*(\alpha w + u) = \alpha T^*w + T^*u$. Por lo tanto, $T^* \in L(W, V)$.

Para la prueba de la unicidad, supongamos que existe otro $S \in L(W, V)$ tal que $\langle Tv, w \rangle = \langle v, Sw \rangle$ para todo $v \in V$ y todo $w \in W$. Luego, para cada $v \in V$ y cada $w \in W$,

$$\langle v, (T^* - S)w \rangle = \langle v, T^*w - Sw \rangle = \langle v, T^*w \rangle - \langle v, Sw \rangle = \langle Tv, w \rangle - \langle Tv, w \rangle = 0.$$

Fijado $w \in W$, la ecuación anterior dice que $(T^* - S)w \perp v$ para todo $v \in V$, entonces $(T^* - S)w = \vec{0}$. Como $w \in W$ era arbitrario, resulta que $Sw = T^*w$ para todo $w \in W$, es decir $S = T^*$. $\square$

**Ejemplo 8.6.** Consideremos a $\mathbb{R}^2$ y $\mathbb{R}^3$ dotados con el producto interno canónico. Dada $T \in L(\mathbb{R}^3, \mathbb{R}^2)$ definida por

$$T(x, y, z) = (y + 3z, 2x),$$

calculemos su transformación adjunta.

En primer lugar, notemos que en este caso $T^* : \mathbb{R}^2 \to \mathbb{R}^3$. Ahora, fijemos un punto $(u, v) \in \mathbb{R}^2$. Luego, para todo $(x, y, z) \in \mathbb{R}^3$ tenemos

$$\begin{aligned} \langle (x, y, z), T^*(u, v) \rangle &= \langle T(x, y, z), (u, v) \rangle = \langle (y + 3z, 2x), (u, v) \rangle = (y + 3z)u + 2xv \\ &= x(2v) + yu + z(3u) = \langle (x, y, z), (2v, u, 3u) \rangle . \end{aligned}$$

Entonces, $T^*(u, v) = (2v, u, 3u)$.

**Ejercicio 8.7.** Supongamos ahora que $\mathbb{R}^2$ está dotado con el producto interno dado en (7.1), mientras que $\mathbb{R}^3$ sigue dotado con el producto interno canónico. Calcular la $T^*$ correspondiente para la transformación $T$ del ejemplo anterior.

La aplicación $T \mapsto T^*$ se comporta de manera similar a la conjugación de números complejos, como lo muestra el siguiente resultado.

**Proposición 8.8** (**Propiedades de la adjunta**). *Dados dos espacios con producto interno $(V, \langle \cdot, \cdot \rangle)$ y $(W, \langle \cdot, \cdot \rangle)$, sean $S, T \in L(V, W)$ y $\alpha \in \mathbb{K}$. Entonces,*

*i)* $(S + T)^* = S^* + T^*$;

*ii)* $(\alpha T)^* = \overline{\alpha}\, T^*$;

*iii)* $(T^*)^* = T$;

*iv) si $I \in L(V)$ es el operador identidad entonces $I^* = I$.*

*Además, si $(V', \langle \cdot, \cdot \rangle)$ es un tercer espacio con producto interno, $T \in L(V, W)$ y $S \in L(W, V')$, entonces*

$$(ST)^* = T^* S^*.$$

*Demostración.* Ejercicio. □

**Ejemplos 8.9.** Los siguientes ejemplos son un poco más "teóricos", pero muestran el cálculo de los adjuntos de ciertos operadores.

1. Dado un espacio con producto interno $(V, \langle \cdot, \cdot \rangle)$ de dimensión finita, sea $W$ un subespacio de $V$. Entonces, la proyección ortogonal sobre $W$ es igual a su adjunta, es decir,

   $$P_W^* = P_W.$$

   De hecho, dados $u, v \in V$ cualesquiera,

   $$\langle P_W u, v \rangle = \langle P_W u, P_W v + (I - P_W) v \rangle = \langle P_W u, P_W v \rangle,$$

   ya que $P_W u \in W$ y $(I - P_W) v \in W^\perp$. De la misma forma, como $(I - P_W) u \in W^\perp$ y $P_W v \in W$,

   $$\langle P_W u, v \rangle = \langle P_W u, P_W v \rangle = \langle P_W u + (I - P_W) u, P_W v \rangle = \langle u, P_W v \rangle.$$

   Por lo tanto, $P_W^* = P_W$.

2. Sea $V = \mathbb{C}^{n \times n}$ dotado con el producto interno dado por

   $$\langle A, B \rangle = \operatorname{tr}(B^* A).$$

   Dada $M \in \mathbb{C}^{n \times n}$ fija, si $L_M : V \to V$ está definido por $L_M(A) = MA$ entonces,

   $${L_M}^* = L_{M^*},$$

   es decir, el adjunto también es un operador de multiplicación a izquierda, pero con $M^*$ en lugar de $M$. De hecho, si $A, B \in \mathbb{C}^{n \times n}$,

   $$\begin{aligned} \langle L_M(A), B \rangle &= \langle MA, B \rangle = \operatorname{tr}(B^*(MA)) = \operatorname{tr}((B^* M) A) = \operatorname{tr}((M^* B)^* A) \\ &= \langle A, M^* B \rangle = \langle A, L_{M^*}(B) \rangle. \end{aligned}$$

En el último ejemplo utilizamos el asterisco $*$ para denotar al adjunto del operador $L_M$ pero también para anotar las adjuntas de las matrices que fueron apareciendo. En lo que sigue intentaremos justificar este abuso de notación.

**Proposición 8.10.** *Dados dos espacios con producto interno $(V, \langle \cdot, \cdot \rangle)$ y $(W, \langle \cdot, \cdot \rangle)$ con $\dim V = n$ y $\dim W = m$, sean $\mathcal{E} = \{e_1, \ldots, e_n\}$ y $\mathcal{F} = \{f_1, \ldots, f_m\}$ bases ortonormales de $V$ y $W$, respectivamente. Si $T \in L(V, W)$ entonces*

$$[T^*]_{\mathcal{E},\mathcal{F}} = {[T]_{\mathcal{F},\mathcal{E}}}^*.$$

*Demostración.* Supongamos que $A := [T]_{\mathcal{F},\mathcal{E}} \in \mathbb{K}^{m \times n}$ y $B := [T^*]_{\mathcal{E},\mathcal{F}} \in \mathbb{K}^{n \times m}$. Por el Teorema 8.1, fijados $i = 1, \ldots, n$ y $j = 1, \ldots, m$, sabemos que

$$B_{ij} = \langle T^* f_j, e_i \rangle = \overline{\langle e_i, T^* f_j \rangle} = \overline{\langle T e_i, f_j \rangle} = \overline{A_{ji}}. \tag{8.2}$$

Por lo tanto, $B$ coincide con la matriz adjunta de $A$. □

A diferencia del Teorema 8.1, aquí sí es esencial que las dos bases involucradas sean bases ortonormales. La base $\mathcal{E}$ debe ser ortonormal para que sea cierta la primera igualdad en (8.2), y $\mathcal{F}$ debe ser ortonormal para que sea cierta la última igualdad en (8.2).

**Proposición 8.11.** *Sean* $(V, \langle\cdot,\cdot\rangle)$ *y* $(W, \langle\cdot,\cdot\rangle)$ *dos espacios con producto interno de dimensión finita. Si* $T \in L(V, W)$*, entonces*

$$N(T^*) = \operatorname{Im}(T)^\perp \quad y \quad \operatorname{Im}(T^*) = N(T)^\perp.$$

*Demostración.* Ejercicio. □

Teniendo en cuenta que $(T^*)^* = T$, también es cierto que

$$N(T) = \operatorname{Im}(T^*)^\perp \quad \text{y} \quad \operatorname{Im}(T) = N(T^*)^\perp.$$

Por último, veamos que la adjunta de una transformación lineal no es más que su traspuesta, identificando a los duales con los espacios originales vía el Teorema de representación de Riesz.

Sean $(V, \langle\cdot,\cdot\rangle)$ y $(W, \langle\cdot,\cdot\rangle)$ dos espacios con producto interno de dimensión finita. Sean $\Phi_V : V^* \to V$ y $\Phi_W : W^* \to W$ las funciones que asignan a cada funcional lineal el único vector que lo representa en $V$ y $W$, respectivamente (recordar el Teorema de representación de Riesz y (7.11)). Luego, el siguiente diagrama conmuta:

$$\begin{array}{ccc} W^* & \xrightarrow{T^t} & V^* \\ {\scriptstyle \Phi_W}\downarrow & & \downarrow{\scriptstyle \Phi_V} \\ W & \xrightarrow[T^*]{} & V \end{array}$$

es decir, $\Phi_V \circ T^t = T^* \circ \Phi_W$, y como $\Phi_W$ es inversible,

$$T^* = \Phi_V \circ T^t \circ \Phi_W^{-1}.$$

Otra propiedad de $\Phi_V$ es que mapea el anulador de un subespacio $W$ de $V$ en el complemento ortogonal a $W$,

$$\Phi_V(W^\circ) = W^\perp.$$

Teniendo esto en cuenta, comparar la Proposición 8.11 con el Teorema 6.25.

## 8.2. Isometrías e isomorfismos isométricos

**Definición 8.12.** Dados $(V, \langle\cdot,\cdot\rangle)$ y $(W, \langle\cdot,\cdot\rangle)$ dos espacios con producto interno, sea $T \in L(V, W)$. Diremos que $T$ es una **isometría** si

$$\|Tv\| = \|v\| \quad \text{para todo } v \in V. \tag{8.3}$$

Si además $T$ es un isomorfismo, diremos que es un **isomorfismo isométrico**.

**Ejemplos 8.13.** Consideremos a los espacios $\mathbb{R}^2$ y $\mathbb{R}^3$ dotados con los productos internos usuales, y veamos algunos ejemplos:

1. La transformación lineal $T : \mathbb{R}^2 \to \mathbb{R}^3$ definida por

   $$T(x, y) = (x, y, 0),$$

   es una isometría, pero obviamente $T$ no es un isomorfismo.

2. Por otro lado, el operador $R : \mathbb{R}^2 \to \mathbb{R}^2$ dado por
$$R(x, y) = (-y, x),$$
no sólo es una isometría sino que es un isomorfismo isométrico.

3. Para terminar, presentemos una transformación lineal que no es isométrica. Sea $S : \mathbb{R}^3 \to \mathbb{R}^2$ dada por
$$S(x, y, z) = (x, y).$$
Esta transformación no es isométrica ya que, por ejemplo, para cualquier $z \neq 0$ tenemos que $S(0, 0, z) = (0, 0)$, pero
$$\|(0, 0, z)\| = |z| \neq 0 = \|(0, 0)\|.$$

La ecuación (8.3) garantiza que $T$ preserva las normas, es por esto que utilizamos el apelativo "isometría" para estas transformaciones. Pero si $T$ satisface (8.3) entonces también preserva productos internos.

**Teorema 8.14.** *Dados $(V, \langle \cdot, \cdot \rangle)$ y $(W, \langle \cdot, \cdot \rangle)$ dos espacios con producto interno, sea $T \in L(V, W)$. Entonces, $T$ es una isometría si y sólo si*
$$\langle Tu, Tv \rangle = \langle u, v \rangle, \qquad \textit{para todo } u, v \in V. \tag{8.4}$$

*Demostración.* Supongamos que $V$ es un espacio vectorial complejo (si fuera un espacio vectorial real, hay que usar la fórmula de polarización correspondiente). Si $T$ es una isometría, por la identidad de polarización tenemos que, para $u, v \in V$ cualesquiera,
$$\langle Tu, Tv \rangle = \sum_{k=0}^{3} i^k \|Tu + i^k\ Tv\|^2 = \sum_{k=0}^{3} i^k \|T(u + i^k v)\|^2 = \sum_{k=0}^{3} i^k \|u + i^k v\|^2 = \langle u, v \rangle .$$
Por lo tanto, $T$ preserva productos internos.

Recíprocamente, supongamos que $T$ satisface (8.4). Luego, dado $v \in V$,
$$\|Tv\| = \langle Tv, Tv \rangle^{1/2} = \langle v, v \rangle^{1/2} = \|v\|. \tag{8.5}$$
Por lo tanto, $T$ es una isometría. □

Además, si $T : V \to W$ es un isomorfismo isométrico no sólo $T$ es una isometría sino que $T^{-1}$ también lo es. De hecho, si $u, v \in W$,
$$\left\langle T^{-1}u, T^{-1}v \right\rangle = \left\langle T(T^{-1}u), T(T^{-1}v) \right\rangle = \left\langle (TT^{-1})u, (TT^{-1})v \right\rangle = \langle u, v \rangle .$$

El siguiente resultado muestra que las únicas isometrías entre dos espacios con producto interno con la misma dimensión son los isomorfismos isométricos.

**Teorema 8.15.** *Sean $(V, \langle \cdot, \cdot \rangle)$ y $(W, \langle \cdot, \cdot \rangle)$ dos espacios con producto interno con la misma dimensión (finita). Si $T \in L(V, W)$, las siguientes condiciones son equivalentes:*

*i)* $T$ *es una isometría;*

*ii)* $T$ *es un isomorfismo isométrico;*

*iii)* $T$ *aplica cada base ortonormal de $V$ sobre una base ortonormal de $W$;*

*iv) $T$ aplica cierta base ortonormal de $V$ sobre una base ortonormal de $W$.*

*Demostración.*
*i)* $\rightarrow$ *ii)* : Supongamos que $T$ es una isometría. De (8.3) se desprende inmediatamente que $N(T) = \{\vec{0}\}$, y como $\dim W = \dim V$, resulta que $T$ es un isomorfismo isométrico.
*ii)* $\rightarrow$ *iii)* : Supongamos que $\dim W = \dim V = n$, y sea $\mathcal{E} = \{e_1, \ldots, e_n\}$ una base ortonormal de $V$ cualquiera. Entonces, $\{Te_1, \ldots, Te_n\}$ es una base de $W$, porque $T$ es un isomorfismo. Además, dados $i, j = 1, \ldots, n$,

$$\langle Te_i, Te_j \rangle = \langle e_i, e_j \rangle = \delta_{ij}.$$

Por lo tanto, $\{Te_1, \ldots, Te_n\}$ es una base ortonormal de $W$.
*iii)* $\rightarrow$ *iv)* : Es trivial.
*iv)* $\rightarrow$ *v)* : Supongamos que $\mathcal{E} = \{e_1, \ldots, e_n\}$ es una base ortonormal de $V$ tal que el conjunto $\{Te_1, \ldots, Te_n\}$ es una base ortonormal de $W$. Dados $u, v \in V$ cualesquiera, escribámoslos en términos de la base ortonormal $\mathcal{E}$: $u = \sum_{i=1}^{n} \langle u, e_i \rangle e_i$ y $v = \sum_{i=1}^{n} \langle v, e_i \rangle e_i$. Luego,

$$\langle u, v \rangle = \left\langle \sum_{i=1}^{n} \langle u, e_i \rangle e_i, \sum_{j=1}^{n} \langle v, e_j \rangle e_j \right\rangle = \sum_{j=1}^{n} \overline{\langle v, e_j \rangle} \left\langle \sum_{i=1}^{n} \langle u, e_i \rangle e_i, e_j \right\rangle = \sum_{j=1}^{n} \overline{\langle v, e_j \rangle} \langle u, e_j \rangle ,$$

mientras que

$$\begin{aligned} \langle Tu, Tv \rangle &= \left\langle \sum_{i=1}^{n} \langle u, e_i \rangle (Te_i), \sum_{j=1}^{n} \langle v, e_j \rangle (Te_j) \right\rangle = \sum_{j=1}^{n} \overline{\langle v, e_j \rangle} \left\langle \sum_{i=1}^{n} \langle u, e_i \rangle (Te_i), Te_j \right\rangle \\ &= \sum_{j=1}^{n} \overline{\langle v, e_j \rangle} \langle u, e_j \rangle , \end{aligned}$$

ya que $\{Te_1, \ldots, Te_n\}$ también es una base ortonormal. Comparando las dos ecuaciones anteriores, tenemos que

$$\langle Tu, Tv \rangle = \langle u, v \rangle \quad \text{para todo } u, v \in V,$$

es decir, $T$ es una isometría. $\square$

Del teorema anterior se desprende el siguiente corolario, el cual dice que si existe un isomorfismo entre dos espacios con producto interno de dimensión finita entonces también existe un isomorfismo isométrico entre ellos.

**Corolario 8.16.** *Sean $(V, \langle \cdot, \cdot \rangle)$ y $(W, \langle \cdot, \cdot \rangle)$ dos espacios con producto interno de dimensión finita. Entonces, $V$ y $W$ son isométricamente isomorfos si y sólo si* $\dim V = \dim W$.

**Ejemplo 8.17.** Sea $V$ el $\mathbb{R}$-espacio vectorial formado por las matrices antisimétricas reales de $3 \times 3$,

$$V = \{A \in \mathbb{R}^{3\times3} : \ A^\top = -A\},$$

dotado con el producto interno $\langle A, B \rangle = \operatorname{tr}(B^\top A)$. Notemos que $(V, \langle \cdot, \cdot \rangle)$ es un $\mathbb{R}$-espacio con producto interno con $\dim V = 3$, ya que

$$\mathcal{E} = \left\{ E_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, E_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, E_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \right\},$$

es una base ortonormal de $V$.

Luego, la transformación lineal $T : \mathbb{R}^3 \to V$ dada por

$$T(x, y, z) = \begin{pmatrix} 0 & -z & y \\ z & 0 & -x \\ -y & x & 0 \end{pmatrix},$$

es un isomorfismo isométrico, porque

$$T(1, 0, 0) = E_1, \quad T(0, 1, 0) = E_2 \quad \text{y} \quad T(0, 0, 1) = E_3,$$

es decir, $T$ aplica la base canónica de $\mathbb{R}^3$ en la base ortonormal $\mathcal{E}$ de $V$.

## 8.3. Operadores unitarios

**Definición 8.18.** Sea $(V, \langle \cdot, \cdot \rangle)$ un espacio con producto interno. Un operador $U$ sobre $V$ se dice **unitario** si es un isomorfismo isométrico de $V$ en sí mismo.

Si $\dim V < \infty$ y $U \in L(V)$, el Teorema 8.15 dice que las siguientes condiciones son equivalentes:

- $U$ es unitario;
- $U$ preserva el producto interno;
- $U$ preserva la norma;
- $U$ aplica bases ortonormales en bases ortonormales.

Un operador unitario también puede caracterizarse por medio de la relación con su adjunto.

**Teorema 8.19.** *Dado un espacio con producto interno $(V, \langle \cdot, \cdot \rangle)$ de dimensión finita, sea $U \in L(V)$. Entonces, $U$ es unitario si y sólo si*

$$U^*U = UU^* = I.$$

*Demostración.* Supongamos que $U$ es unitario. Luego, dados $v, w \in V$ cualesquiera,

$$\langle Uv, w \rangle = \left\langle Uv, (UU^{-1})w \right\rangle = \left\langle Uv, U(U^{-1}w) \right\rangle = \left\langle v, U^{-1}w \right\rangle,$$

es decir, $U^* = U^{-1}$. Por lo tanto, $U^*U = UU^* = I$.

Recíprocamente, supongamos que $U^*U = UU^* = I$. Esta igualdad dice que $U^* = U^{-1}$, en particular, $U$ es invertible. Además, si $v, w \in V$,

$$\langle Uv, Uw \rangle = \langle v, U^*(Uw) \rangle = \langle v, (U^*U)w \rangle = \langle v, w \rangle,$$

es decir, $U$ preserva el producto interno. Por lo tanto, $U$ es unitario. □

La siguiente proposición muestra que, los operadores unitarios sobre un espacio con producto interno, forman un grupo.

**Proposición 8.20.** *Sea $(V, \langle \cdot, \cdot \rangle)$ un espacio con producto interno. Luego,*

*i) si $U_1, U_2 \in L(V)$ son unitarios, entonces $U_1U_2$ también lo es;*

*ii) si $U \in L(V)$ es unitario, entonces $U^{-1}$ también lo es.*

*Demostración.* Supongamos primero que $U_1, U_2 \in L(V)$ son unitarios. Luego,

$$(U_1U_2)^* = U_2^*U_1^* = U_2^{-1}U_1^{-1} = (U_1U_2)^{-1}.$$

Por el teorema anterior, tenemos que $U_1U_2$ es unitario, lo que prueba $i$).

Por otra parte, si $U \in L(V)$ es unitario,

$$(U^{-1})^* = (U^*)^* = U = (U^{-1})^{-1},$$

de donde se deduce que $U^{-1}$ también es unitario. □

Para entender las representaciones matriciales de los operadores unitarios necesitamos introducir dos nuevas familias de matrices.

**Definición 8.21.** Dada una matriz $A \in \mathbb{C}^{n\times n}$, diremos que

- $A$ es **unitaria** si $A^*A = AA^* = I_n$;
- $A$ es **ortogonal** si $A^\top A = AA^\top = I_n$.

Si $A \in \mathbb{R}^{n\times n}$, estas dos nociones coinciden. Pero si alguna de las entradas de $A$ tiene parte imaginaria no nula, $A$ no puede ser unitaria y ortogonal simultáneamente.

**Ejemplo 8.22.** Si $\theta \in [0, 2\pi)$ entonces

$$A_\theta := \begin{pmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{pmatrix} \in \mathbb{R}^{2\times 2},$$

es ortogonal (y unitaria). En cambio, si $z, w \in \mathbb{C} \setminus \mathbb{R}$ son tales que $|w|^2 + |z|^2 = 1$, entonces

$$A_{z,w} := \begin{pmatrix} z & w \\ -\overline{w} & \overline{z} \end{pmatrix},$$

es unitaria pero no es ortogonal.

**Observación 8.23.** Notemos que si $A \in \mathbb{C}^{n\times n}$ es unitaria, entonces las columnas de $A$ forman una base ortonormal de $\mathbb{C}^n$. Lo mismo es cierto para las filas de $A$.

Si $A \in \mathbb{C}^{n\times n}$ es unitaria, entonces

$$\begin{aligned} 1 &= \det(I_n) = \det(A^*A) = \det(A^*)\det(A) = \det(\overline{A^\top})\det(A) = \overline{\det(A^\top)}\det(A) \\ &= \overline{\det(A)}\det(A) = |\det(A)|^2. \end{aligned}$$

Por lo tanto, $\det(A) = \cos\theta + i\operatorname{sen}\theta$ para algún $\theta \in [0, 2\pi)$. En cambio, si $A$ es ortogonal,

$$1 = \det(I_n) = \det(A^\top A) = \det(A^\top)\det(A) = \det(A)^2.$$

Por lo tanto, $\det(A) = 1$ ó $\det(A) = -1$.

Ahora sí estamos en condiciones de describir las representaciones matriciales de los operadores unitarios.

**Teorema 8.24.** *Dado un espacio con producto interno $(V, \langle\cdot,\cdot\rangle)$ de dimensión finita, sea $U \in L(V)$. Entonces, $U$ es unitario si y sólo si $[U]_{\mathcal{E}}$ es unitaria para alguna (para toda) base ortonormal $\mathcal{E}$ de $V$.*

*Demostración.* Sea $\mathcal{E}$ una base ortonormal de $V$ y llamemos $A = [U]_{\mathcal{E}}$. Por la Proposición 8.10, $[U^*]_{\mathcal{E}} = [U]^*_{\mathcal{E}} = A^*$. Luego,

$$A^*A = [U^*]_{\mathcal{E}}[U]_{\mathcal{E}} = [U^*U]_{\mathcal{E}} \quad \text{y} \quad AA^* = [U]_{\mathcal{E}}[U^*]_{\mathcal{E}} = [UU^*]_{\mathcal{E}}$$

Por lo tanto, $A^*A = AA^* = I_n$ si y sólo si $U^*U = UU^* = I$. □

Dado un espacio con producto interno $(V, \langle\cdot,\cdot\rangle)$ con $\dim V = n$, supongamos que $\mathcal{E} = \{e_1, \dots, e_n\}$ y $\mathcal{E}' = \{e'_1, \dots, e'_n\}$ son dos bases ortonormales de $V$. Si $U \in L(V)$ es el (único) operador lineal tal que

$$Ue_i = e'_i, \quad i = 1, \dots, n,$$

tenemos que la representación matricial $[U]_{\mathcal{E}}$ coincide con la matriz cambio de base $P := P_{\mathcal{E},\mathcal{E}'}$ (comprobarlo!). Como $U$ aplica una base ortonormal en otra, $U$ es unitario. Luego, $P$ es una matriz unitaria.

Recíprocamente, cada matriz unitaria $P$ puede interpretarse como la matriz cambio de base de la base canónica en la base ortonormal formada por las columnas de $P$ (recordar la Observación 8.23).

Luego, si $T \in L(V)$ es un operador cualquiera,

$$[T]_{\mathcal{E}'} = P^{-1}[T]_{\mathcal{E}}P = P^*[T]_{\mathcal{E}}P,$$

es decir, los cambios de bases (de una base ortonormal a otra) se realizan mediante la **conjugación con una matriz unitaria**.

**Definición 8.25.** Dadas dos matrices $A, B \in \mathbb{C}^{n\times n}$, diremos que $A$ y $B$ son:

- **unitariamente equivalentes** si existe una matriz unitaria $P \in \mathbb{C}^{n\times n}$ tal que
$$B = P^{-1}AP = P^*AP.$$
- **ortogonalmente equivalentes** si existe una matriz ortogonal $P \in \mathbb{C}^{n\times n}$ tal que
$$B = P^{-1}AP = P^\top AP.$$

## 8.4. Operadores autoadjuntos

**Definición 8.26.** Un operador $T \in L(V)$ se dice **autoadjunto** si $T^* = T$, es decir, $T$ es autoadjunto si y sólo si

$$\langle Tv, w\rangle = \langle v, Tw\rangle, \qquad \text{para todo } v, w \in V.$$

La primera consecuencia inmediata de la definición es la siguiente: si $T$ es autoadjunto entonces

$$\langle Tv, v\rangle \in \mathbb{R}, \qquad \text{para todo } v \in V. \tag{8.6}$$

Si $V$ es un $\mathbb{R}$-espacio vectorial el siguiente ejemplo muestra que la recíproca no es cierta.

**Ejemplo 8.27.** Sea $T \in L(\mathbb{R}^2)$ dado por

$$T(x, y) = (-y, x), \qquad (x, y) \in \mathbb{R}^2.$$

Como vimos en el Ejemplo 8.3, $\langle T(x, y), (x, y)\rangle = \langle(-y, x), (x, y)\rangle = -yx + xy = 0 \in \mathbb{R}$ para cada $(x, y) \in \mathbb{R}^2$. Pero $T$ no es autoadjunto, ya que

$$\begin{aligned}&\left\langle T(x,y),(x',y')\right\rangle = \left\langle(-y,x),(x',y')\right\rangle = -yx' + xy', \quad \text{mientras que}\\ &\left\langle(x,y),T(x',y')\right\rangle = \left\langle(x,y),(-y',x')\right\rangle = -xy' + yx' = -(-yx' + xy') = -\left\langle T(x,y),(x',y')\right\rangle.\end{aligned}$$

En cambio, si $V$ es un $\mathbb{C}$-espacio vectorial la recíproca sí vale. Es decir, si $V$ es un $\mathbb{C}$-espacio con producto interno y $T \in L(V)$ satisface $\langle Tv, v\rangle \in \mathbb{R}$ para todo $v \in V$, entonces $T$ es autoadjunto.

**Teorema 8.28.** *Supongamos que $(V, \langle\cdot,\cdot\rangle)$ es un $\mathbb{C}$-espacio con producto interno. Si $T \in L(V)$ satisface*

$$\langle Tv, v\rangle \in \mathbb{R} \qquad \text{para todo } v \in V,$$

*entonces el operador $T$ es autoadjunto.*

*Demostración.* Si $\langle Tv, v\rangle \in \mathbb{R}$ entonces $\langle Tv, v\rangle = \overline{\langle Tv, v\rangle} = \langle v, Tv\rangle = \langle T^*v, v\rangle$. Luego, si $\langle Tv, v\rangle \in \mathbb{R}$ para todo $v \in V$, tenemos que

$$\langle (T^* - T)v, v\rangle = 0 \quad \text{para todo } v \in V.$$

Ahora, aplicando la Proposición 8.2 resulta que $T^* - T = 0$, es decir, $T^* = T$. □

**Ejercicio 8.29.** Mostrar que, el conjunto $\mathcal{A}(V)$ de operadores autoadjuntos sobre un espacio con producto interno $(V, \langle\cdot,\cdot\rangle)$, es un $\mathbb{R}$-espacio vectorial, es decir, dados $S, T \in \mathcal{A}(V)$ y $\alpha, \beta \in \mathbb{R}$ entonces

$$\alpha S + \beta T \in \mathcal{A}(V).$$

Teniendo en cuenta la analogía que establecimos entre la adjunción de operadores y la conjugación compleja, y que un número complejo $z \in \mathbb{C}$ es real si y sólo si $\overline{z} = z$, podríamos argumentar que los operadores autoadjuntos cumplen el rol de los números reales. La siguiente proposición refuerza esta misma idea.

**Proposición 8.30.** *Dado un espacio con producto interno $(V, \langle\cdot,\cdot\rangle)$, sea $T \in L(V)$. Entonces, $T$ puede descomponerse como*

$$T = \mathrm{Re}(T) + i\,\mathrm{Im}(T), \tag{8.7}$$

*donde* $\mathrm{Re}(T)$ *y* $\mathrm{Im}(T)$ *son los operadores autoadjuntos dados por*

$$\mathrm{Re}(T) := \frac{T + T^*}{2} \quad y \quad \mathrm{Im}(T) := \frac{T - T^*}{2i}.$$

*Demostración.* De la misma definición de $\mathrm{Re}(T)$ e $\mathrm{Im}(T)$, se deduce que éstos son operadores autoadjuntos. Además es evidente que vale (8.7). □

**Teorema 8.31.** *Dado un espacio con producto interno $(V, \langle\cdot,\cdot\rangle)$, sea $T \in L(V)$ autoadjunto. Entonces, los autovalores de $T$ son* **reales**. *Además, si $v, w \in V$ son autovectores de $T$ asociados a distintos autovalores $\lambda, \mu \in \mathbb{R}$, respectivamente, i.e.*

$$Tv = \lambda v, \quad Tw = \mu w \quad y \quad \lambda \neq \mu,$$

*entonces $v \perp w$.*

*Demostración.* En primer lugar, supongamos que $\lambda \in \mathbb{C}$ es un autovalor de $T$, y que $v$ es un autovector asociado a $\lambda$, es decir, $v \neq \vec{0}$ y $Tv = \lambda v$. Luego,

$$\lambda\|v\|^2 = \lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle Tv, v\rangle = \langle v, Tv\rangle = \langle v, \lambda v\rangle = \overline{\lambda}\langle v, v\rangle = \overline{\lambda}\|v\|^2.$$

Como $v \neq \vec{0}$, la igualdad anterior implica que $\overline{\lambda} = \lambda$, es decir, $\lambda \in \mathbb{R}$.

Además, si $\mu$ es otro autovalor de $T$, $\mu \neq \lambda$, y $w$ es un autovector asociado a $\mu$, resulta

$$\lambda\langle v, w\rangle = \langle \lambda v, w\rangle = \langle Tv, w\rangle = \langle v, Tw\rangle = \langle v, \mu w\rangle = \overline{\mu}\langle v, w\rangle = \mu\langle v, w\rangle,$$

ya que $\mu \in \mathbb{R}$. De la igualdad anterior, deducimos que $(\lambda - \mu)\langle v, w\rangle = 0$. Por lo tanto, $\langle v, w\rangle = 0$. □

En capítulos anteriores vimos ejemplos de operadores lineales que no tienen autovectores. En particular, el operador del Ejemplo 8.3 es uno de ellos. El siguiente resultado muestra que esta situación no ocurre con operadores autoadjuntos.

**Teorema 8.32.** *Si $(V, \langle \cdot, \cdot \rangle)$ es un espacio con producto interno de dimensión finita, todo operador autoadjunto $T \in L(V)$ tiene un autovector.*

*Demostración.* En esta prueba tendremos que prestar especial atención al cuerpo $\mathbb{K}$ sobre el que $V$ es un $\mathbb{K}$-espacio vectorial (recordemos que puede ser $\mathbb{K} = \mathbb{R}$ ó $\mathbb{K} = \mathbb{C}$).

Supongamos que $\dim V = n$ y que $T \in L(V)$ es autoadjunto. Fijada una base ortonormal $\mathcal{B}$ de $V$, sea $A = [T]_{\mathcal{B}} \in \mathbb{K}^{n\times n}$. Como $T^* = T$, tenemos que $A^* = A$.

Ahora, sea $W = \mathbb{C}^{n\times 1}$ dotado con el producto interno dado por

$$\langle X, Y\rangle = Y^* X.$$

(Notemos que, como $Y^*X \in \mathbb{C}^{1\times 1} = \mathbb{C}$, no hace falta tomar la traza).

Luego, definamos $U : W \to W$ como $U(X) = AX$. Es fácil ver $U$ es un operador lineal sobre $W$. Además, $U$ es autoadjunto porque si $X, Y \in W$,

$$\begin{aligned}\langle U(X), Y\rangle &= \langle AX, Y\rangle = Y^*(AX) = (Y^*A)X = (Y^*A^*)X = (AY)^*X \\ &= \langle X, AY\rangle = \langle X, U(Y)\rangle .\end{aligned}$$

El polinomio característico de $A$, $p_A(x) := \det(xI_n - A)$, es también el polinomio característico de los operadores $T$ y $U$. Como es un polinomio (de grado $n$), tiene al menos una raíz en $\mathbb{C}$. Es decir, existe algún $\lambda \in \mathbb{C}$ tal que $\det(\lambda I_n - A) = 0$. Como $\lambda I_n - A \in \mathbb{C}^{n\times n}$ es una matriz singular, existe un $X \in \mathbb{C}^{n\times 1}$, $X \neq \vec{0}$ tal que $(\lambda I_n - A)X = 0$. Luego,

$$U(X) = AX = \lambda X,$$

es decir, $\lambda$ es un autovalor de $U$ y $X$ es un autovector asociado a $\lambda$. Por el teorema anterior, $\lambda \in \mathbb{R}$.

Para terminar la demostración sólo falta mostrar que podemos encontrar a un autovector de $T$ a partir del autovector $X \in \mathbb{C}^{n\times 1}$ de $U$. La idea es reconstruirlo combinando los vectores de la base $\mathcal{B}$ con las entradas de $X$. De hecho, el autovector será el vector de $V$ cuyas coordenadas con respecto a la base $\mathcal{B}$ sean las entradas del vector $X$ (respetando el orden).

Si $\mathbb{K} = \mathbb{C}$, no hay nada más que hacer. En cambio, si $\mathbb{K} = \mathbb{R}$ hay que convencerse de que las entradas de $X$ son reales. En este último caso, notemos que $A \in \mathbb{R}^{n\times n}$. Luego, como también $\lambda \in \mathbb{R}$ y $A - \lambda I_n \in \mathbb{R}^{n\times n}$ es singular, el sistema

$$(A - \lambda I_n)X = \vec{0},$$

tiene una solución no trivial $X \in \mathbb{R}^n$, como queríamos probar. □

Si $V$ es un $\mathbb{C}$-espacio vectorial de dimensión finita, la misma demostración anterior prueba que todo operador lineal $T \in L(V)$ tiene un autovector (y por consiguiente también un autovalor) aún cuando $T$ no es autoadjunto. El hecho de que $T^* = T$ sólo fue necesario para garantizar la existencia del autovector en el caso en que $\mathbb{K} = \mathbb{R}$.

**Definición 8.33.** Dada una matriz $A \in \mathbb{C}^{n\times n}$, diremos que

- $A$ es **autoadjunta** si $A^* = A$;
- $A$ es **simétrica** si $A^\top = A$.

Si $A \in \mathbb{R}^{n\times n}$, estas dos nociones coinciden. Pero si alguna de las entradas de $A$ tiene parte imaginaria no nula, $A$ no puede ser autoadjunta y simétrica simultáneamente.

**Ejercicio 8.34.** Si $A \in \mathbb{C}^{n\times n}$ es autoadjunta, mostrar que el polinomio característico de $A$ tiene coeficientes reales:

$$p_A(x) := \det(xI - A) \in \mathbb{R}[x].$$

**Proposición 8.35.** *Dado un espacio con producto interno $(V, \langle\cdot,\cdot\rangle)$ de dimensión finita, sea $T \in L(V)$. Entonces, $T$ es autoadjunto si y sólo si*

$$[T]_{\mathcal{E}} \text{ es autoadjunta,}$$

*para alguna (para toda) base ortonormal $\mathcal{E}$ de $V$.*

*Demostración.* Es una consecuencia inmediata de la Proposición 8.10. □

## 8.5. Operadores normales

En esta sección consideraremos sólo espacios vectoriales sobre $\mathbb{C}$.

**Definición 8.36.** Dado un espacio con producto interno $(V, \langle\cdot,\cdot\rangle)$, sea $T \in L(V)$. Diremos que $T$ es **normal** si $T$ conmuta con su adjunto, es decir,

$$T^*T = TT^*.$$

Los operadores unitarios y los operadores autoadjuntos son ejemplos de operadores normales, aunque no son los únicos.

**Ejemplo 8.37.** Sea $T$ el operador sobre $\mathbb{C}^3$ cuya matriz con respecto a la base canónica es

$$A = [T]_{\mathcal{E}} = \begin{pmatrix} 2 & 0 & 3 \\ 0 & i & 0 \\ -3 & 0 & 2 \end{pmatrix}.$$

Veamos que $T$ es normal, pero no es autoadjunto ni unitario.

Comencemos calculando la matriz adjunta de $A$:

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

Como $A^* \neq A$, el operador $T$ no es autoadjunto. Por otra parte,

$$A^*A = \begin{pmatrix} 2 & 0 & -3 \\ 0 & -i & 0 \\ 3 & 0 & 2 \end{pmatrix} \begin{pmatrix} 2 & 0 & 3 \\ 0 & i & 0 \\ -3 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 13 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 13 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 3 \\ 0 & i & 0 \\ -3 & 0 & 2 \end{pmatrix} \begin{pmatrix} 2 & 0 & -3 \\ 0 & -i & 0 \\ 3 & 0 & 2 \end{pmatrix} = AA^*.$$

Como $[T^*T]_{\mathcal{E}} = A^*A = AA^* = [TT^*]_{\mathcal{E}}$, tenemos que $T$ es normal. Por último, $T$ no es unitario porque $A^*A = AA^* \neq I_3$.

**Lema 8.38.** *Un operador $T \in L(V)$ es normal si y sólo si*

$$\|T^*v\| = \|Tv\| \qquad \text{para todo } v \in V \tag{8.8}$$

*Demostración.* Dado $T \in L(V)$, $T$ es normal si y sólo si $T^*T - TT^* = 0$. Como el operador $T^*T - TT^*$ es autoadjunto, en particular tenemos que $\langle (T^*T - TT^*)v, v\rangle \in \mathbb{R}$ para cada $v \in V$. Luego la Proposición 8.2 asegura que, $T^*T - TT^* = 0$ si y sólo si

$$\langle (T^*T - TT^*)v, v\rangle = 0 \quad \text{para todo } v \in V.$$

O equivalentemente,

$$\|Tv\|^2 = \langle T^*Tv, v\rangle = \langle TT^*v, v\rangle = \|T^*v\|^2 \quad \text{para todo } v \in V,$$

de lo cual se deduce el enunciado. $\square$

Entre otras cosas, el resultado anterior implica que $N(T^*) = N(T)$ para todo operador normal $T$.

**Proposición 8.39.** *Dado un espacio con producto interno $(V, \langle\cdot,\cdot\rangle)$, sea $T \in L(V)$ un operador normal. Si $v \in V$ es un autovector de $T$ asociado al autovalor $\lambda \in \mathbb{C}$, entonces $v$ también es un autovector de $T^*$ asociado al autovalor $\overline{\lambda}$.*

*Demostración.* Supongamos que $T$ es normal y que $\mu \in \mathbb{C}$. Entonces, el operador $T - \mu I$ también es normal, ya que $(T - \mu I)^* = T^* - \overline{\mu} I$ y

$$\begin{aligned}(T - \mu I)^*(T - \mu I) &= (T^* - \overline{\mu} I)(T - \mu I) = T^*T - \mu T^* - \overline{\mu} T + \overline{\mu}\mu I \\ &= TT^* - \overline{\mu} T - \mu T^* + \mu\overline{\mu} I = (T - \mu I)(T^* - \overline{\mu} I) \\ &= (T - \mu I)(T - \mu I)^*.\end{aligned}$$

Luego, supongamos que $v \in V$ es un autovector de $T$ asociado a $\lambda$, es decir,

$$Tv = \lambda v \quad \Leftrightarrow \quad (T - \lambda I)v = \vec{0} \quad \Leftrightarrow \quad \|(T - \lambda I)v\| = 0.$$

Como $T - \lambda I$ es normal, $\|(T - \lambda I)v\| = 0$ implica que $\|(T^* - \overline{\lambda} I)v\| = \|(T - \lambda I)^*v\| = 0$, o equivalentemente, $v$ es un autovector de $T^*$ asociado al autovalor $\overline{\lambda}$. $\square$

**Corolario 8.40.** *Sea $T \in L(V)$ un operador normal. Si $v$ y $w$ son autovectores de $T$ asociados a autovalores distintos, entonces $v \perp w$.*

*Demostración.* Supongamos que $\lambda, \mu \in \mathbb{C}$ son dos autovalores de $T$ y que $\lambda \neq \mu$. Sean $v$ y $w$ autovectores de $T$ asociados a $\lambda$ y $\mu$, respectivamente, es decir,

$$Tv = \lambda v \quad \text{y} \quad Tw = \mu w.$$

Por la proposición anterior, tenemos también que $T^*w = \overline{\mu} w$. Luego,

$$(\lambda - \mu)\langle v, w\rangle = \langle \lambda v, w\rangle - \langle v, \overline{\mu} w\rangle = \langle Tv, w\rangle - \langle v, T^*w\rangle = \langle Tv, w\rangle - \langle Tv, w\rangle = 0.$$

Como $\lambda \neq \mu$, la ecuación anterior implica que $\langle v, w\rangle = 0$, es decir, $v \perp w$. $\square$

**Definición 8.41.** Una matriz $A \in \mathbb{C}^{n\times n}$ es **normal** si $A^*A = AA^*$.

**Proposición 8.42.** *Dado un espacio con producto interno $(V, \langle\cdot,\cdot\rangle)$ de dimensión finita, sea $T \in L(V)$. Entonces, $T$ es normal si y sólo si*

$$[T]_{\mathcal{E}} \text{ es normal,}$$

*para alguna (para toda) base ortonormal $\mathcal{E}$ de $V$.*

*Demostración.* Es una consecuencia inmediata de la Proposición 8.10. □

**Proposición 8.43.** *Supongamos que $A \in \mathbb{C}^{n\times n}$ es triangular (inferior o superior). Entonces, $A$ es normal si y sólo si $A$ es diagonal.*

*Demostración.* En primer lugar, notemos que si $A$ es diagonal, entonces $A^*$ también es una matriz diagonal. Como dos matrices diagonales cualesquiera conmutan, en particular $A^*A = AA^*$, es decir, $A$ es normal.

Recíprocamente, supongamos que $A$ es normal y llamemos $\mathcal{E} = \{e_1, \dots, e_n\}$ a la base canónica de $\mathbb{C}^{n\times 1}$. Para fijar ideas, supongamos que $A = (a_{ij})$ es triangular superior, es decir, $a_{ij} = 0$ si $i > j$. Notemos que, en este caso,

$$Ae_1 = a_{11}e_1.$$

Por la Proposición 8.39, tenemos también que $A^*e_1 = \overline{a_{11}}\ e_1$. Por otro lado, sabemos que $A^*e_1$ es la primer columna de la matriz $A^*$, y como $A^* = (\overline{A})^\top$, resulta que las entradas de $A^*e_1$ son las conjugadas de las de la primer fila de $A$:

$$A^*e_1 = \sum_{j=1}^{n} \overline{a_{1j}}\ e_j.$$

Comparando las dos expresiones de $A^*e_1$, resulta que $a_{1j} = 0$ para $j = 2, \dots, n$.

En particular, $a_{12} = 0$, y como $A$ es triangular superior, resulta que

$$Ae_2 = a_{12}e_1 + a_{22}e_2 = a_{22}e_2.$$

Repitiendo el argumento anterior, vemos que $A^*e_2 = \overline{a_{22}}\ e_2$ y $a_{2j} = 0$ para $j = 3, \dots, n$. Continuando de esta manera se ve que

$$Ae_j = a_{jj}e_j, \quad \text{para todo } j = 1, \dots, n.$$

Por lo tanto, $A$ es diagonal.

Si $A$ fuera triangular inferior, entonces $A^*$ es triangular superior y normal. Resulta entonces que $A^*$ es diagonal, y evidentemente $A$ también lo es. □

## 8.6. El Teorema Espectral en espacios con producto interno

Al conjunto de autovalores de un operador lineal $T$ se lo suele llamar el **espectro (puntual) de** $T$, y se lo anota $\sigma(T)$. La teoría espectral en espacios con producto interno involucra al concepto de autovalor, la clasificación de operadores/matrices en distintas clases (simétricos, ortogonales, autoadjuntos, unitarios, etc.), los teoremas sobre la naturaleza de los autovalores asociados a las distintas clases y, sobre todo, aquellos que describen las formas canónicas de cada clase.

La teoría espectral para matrices fue creada en el período 1826–1876, comenzando por un tratado de Augustin Cauchy en el que demuestra que los autovalores de una matriz simétrica son reales, y que la forma cuadrática asociada puede transformarse en una suma de términos al cuadrado (es decir, “diagonalizada”) por medio de una sustitución ortogonal (un cambio de bases ortonormales).

Como ya mencionamos anteriormente, en Álgebra Lineal un **teorema espectral** es un resultado que explica cuando un operador lineal puede ser diagonalizado, es decir, representado como una matriz diagonal con respecto a alguna base. En esta sección en particular nos interesaremos por caracterizar, en un espacio con producto interno $(V, \langle\cdot,\cdot\rangle)$ de dimensión finita, aquellos operadores $T \in L(V)$ que determinan una base ortonormal formada por autovectores de $T$, es decir, aquellos $T$ para los cuales existe una base ortonormal $\mathcal{E}$ de $V$ tal que $[T]_{\mathcal{E}}$ es una matriz diagonal.

El cuerpo $\mathbb{K}$ sobre el cual está definido el espacio con producto interno $(V, \langle\cdot,\cdot\rangle)$ cumple un rol central en este tema, por lo cual trataremos de manera diferenciada al caso real del complejo.

### 8.6.1. Operadores ortogonalmente diagonalizables

Supongamos que $(V, \langle\cdot,\cdot\rangle)$ es un espacio euclideano (es decir, $V$ es un $\mathbb{R}$-espacio con producto interno de dimensión finita) y que $\dim V = n$.

Dado $T \in L(V)$, supongamos que $T$ es **ortogonalmente diagonalizable**, es decir, existe una base ortonormal $\mathcal{B}$ de $V$ tal que

$$Tb_j = \lambda_j b_j, \qquad j = 1, \dots, n, \tag{8.9}$$

para ciertos $\lambda_1, \dots, \lambda_n \in \mathbb{R}$, o equivalentemente,

$$[T]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} =: D.$$

**Observación 8.44.** La justificación para el apelativo de “ortogonalmente diagonalizable” se encuentra en que, si $(V, \langle\cdot,\cdot\rangle)$ es un espacio euclidiano, la matriz cambio de base entre dos bases ortonormales es una matriz ortogonal.

Como consecuencia de (8.9), tenemos que

$$\lambda_j = \lambda_j \langle b_j, b_j\rangle = \langle \lambda_j b_j, b_j\rangle = \langle Tb_j, b_j\rangle \in \mathbb{R} \quad \text{para todo } j = 1, \dots, n.$$

Luego,

$$[T^*]_{\mathcal{B}} = [T]_{\mathcal{B}}^* = [T]_{\mathcal{B}}^\top = D^\top = D = [T]_{\mathcal{B}}.$$

Por lo tanto, $T$ es autoadjunto.

A continuación probaremos que también vale la recíproca, todo operador autoadjunto en un espacio euclidiano es ortogonalmente diagonalizable.

**Lema 8.45.** *Dado un espacio con producto interno $(V, \langle\cdot,\cdot\rangle)$, sea $T \in L(V)$. Si $W$ es un subespacio de $V$ invariante por $T$, entonces $W^\perp$ es invariante por $T^*$.*

*Demostración.* Supongamos que $T(W) \subseteq W$. Dado $v \in W^\perp$, veamos que $T^*v \in W^\perp$. Para cada $w \in W$,

$$\langle T^*v, w\rangle = \langle v, Tw\rangle = 0,$$

ya que $Tw \in W$. Entonces, $T^*v \in W^\perp$. Como $v \in W^\perp$ era arbitrario, $W^\perp$ es invariante por $T^*$. □

**Teorema 8.46.** *Dado un espacio con producto interno $(V, \langle\cdot,\cdot\rangle)$ de dimensión finita, sea $T \in L(V)$. Si $T$ es autoadjunto, entonces existe una base ortonormal de $V$ compuesta por autovectores de $T$.*

*Demostración.* La prueba será por inducción sobre la dimensión de $V$. Si $\dim V = 1$, es una consecuencia del Teorema 8.32.

Como hipótesis inductiva, supongamos que el enunciado es cierto para todo espacio con producto interno con dimensión menor que $n$.

Supongamos también que $\dim V = n$. Por el Teorema 8.32, existen $\lambda_1 \in \mathbb{R}$ y $v \in V$, $v \neq \vec{0}$ tales que

$$Tv = \lambda_1 v.$$

Consideremos entonces $e_1 := \frac{v}{\|v\|}$ y $W := \overline{\{e_1\}}$. Como $W$ es $T$-invariante, el Lema 8.45 asegura que $W^\perp$ es invariante por $T^* = T$. Luego, $(W^\perp, \langle\cdot,\cdot\rangle)$ es un espacio con producto interno con $\dim W = n-1$ y además,

$$T|_{W^\perp} : W^\perp \to W^\perp \quad \text{es autoadjunto.}$$

Por la hipótesis inductiva, existe una base ortonormal $\{e_2, \ldots, e_n\}$ de $W^\perp$ tal que

$$Te_j = T|_{W^\perp} e_j = \lambda_j e_j, \qquad j = 2, \ldots, n,$$

para ciertos $\lambda_2, \ldots, \lambda_n \in \mathbb{K}$.

Entonces, $\mathcal{E} = \{e_1, e_2, \ldots, e_n\}$ es una base ortonormal de $V = W \oplus W^\perp$ compuesta por autovectores de $T$. □

**Corolario 8.47.** *Si $(V, \langle\cdot,\cdot\rangle)$ es un espacio euclidiano y $T \in L(V)$,*

$$T \text{ es ortogonalmente diagonalizable} \quad \Leftrightarrow \quad T^* = T.$$

### 8.6.2. Operadores unitariamente diagonalizables

Ahora, supongamos que $(V, \langle\cdot,\cdot\rangle)$ es un espacio unitario (es decir, $V$ es un $\mathbb{C}$-espacio con producto interno de dimensión finita) y que $\dim V = n$.

Dado $T \in L(V)$, supongamos que $T$ es **unitariamente diagonalizable**, es decir, existe una base ortonormal $\mathcal{B}$ de $V$ tal que

$$Tb_j = \lambda_j b_j, \qquad j = 1, \ldots, n, \tag{8.10}$$

para ciertos $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$, o equivalentemente,

$$[T]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix} =: D.$$

**Observación 8.48.** La justificación para el apelativo de “unitariamente diagonalizable” se encuentra en que, si $(V, \langle\cdot,\cdot\rangle)$ es un espacio unitario, la matriz cambio de base entre dos bases ortonormales es una matriz unitaria.

Si $[T]_{\mathcal{B}} = D$ entonces,

$$[T^*]_{\mathcal{B}} = [T]_{\mathcal{B}}^* = D^* = \overline{D}.$$

En general, $\overline{D} \neq D$ porque $D$ puede tener en la diagonal alguna entrada con componente imaginaria no nula. Pero como $\overline{D}D = D\overline{D}$, tenemos que

$$[T^*T]_{\mathcal{B}} = [T^*]_{\mathcal{B}}[T]_{\mathcal{B}} = \overline{D}D = D\overline{D} = [T]_{\mathcal{B}}[T^*]_{\mathcal{B}} = [TT^*]_{\mathcal{B}}.$$

Por lo tanto, $T$ es un operador normal.

A continuación probaremos la recíproca, todo operador normal en un espacio unitario es unitariamente diagonalizable. Comenzaremos mostrando que todo operador lineal en un espacio unitario es triangulable. Este resultado fue probado originalmente en 1909 por el matemático alemán Issai Schur (1875–1941).

**Teorema 8.49** (I. Schur)**.** *Dado un espacio unitario $(V, \langle\cdot,\cdot\rangle)$, sea $T \in L(V)$. Entonces existe una base ortonormal $\mathcal{B}$ de $V$ tal que $[T]_{\mathcal{B}}$ es una matriz triangular superior.*

*Demostración.* Lo probaremos por inducción sobre la dimensión del espacio $V$. Si $\dim V = 1$, no hay nada que demostrar.

Como hipótesis inductiva, supongamos que el resultado vale para espacios con producto interno con dimensión menor a $n$.

Supongamos ahora que $\dim V = n$, y sea $T \in L(V)$. Como $V$ es un $\mathbb{C}$-espacio vectorial, todo operador lineal tiene un autovector (ver el comentario que sigue al Teorema 8.32). En particular, $T^*$ tiene un autovector, es decir, existen $\lambda \in \mathbb{C}$ y $v \in V$ tales que

$$T^*v = \lambda v.$$

Luego, consideremos el subespacio $W := \overline{\{v\}}$ y el operador restringido

$$T|_{W^\perp} : W^\perp \to W^\perp.$$

Notemos que, como $W$ es $T^*$-invariante, el subespacio $W^\perp$ es $T$-invariante y el operador anterior está bien definido.

Por hipótesis inductiva, como $\dim W = n-1$, existe una base ortonormal $\mathcal{E}' = \{e_1, \ldots, e_{n-1}\}$ de $W^\perp$ tal que

$$[T|_{W^\perp}]_{\mathcal{E}'} \quad \text{es triangular superior.}$$

Luego, si llamamos $e_n = \frac{v}{\|v\|}$, $\mathcal{E} := \{e_1, \ldots, e_{n-1}, e_n\}$ es una base ortonormal de $V$, y además es fácil convencerse de que

$$[T]_{\mathcal{E}} = \begin{pmatrix} [T|_{W^\perp}]_{\mathcal{E}'} & * \\ 0 \cdots 0 & * \end{pmatrix},$$

es decir, $[T]_{\mathcal{E}}$ es triangular superior. □

**Corolario 8.50.** *Si $(V, \langle\cdot,\cdot\rangle)$ es un espacio unitario y $T \in L(V)$,*

$$T \text{ es unitariamente diagonalizable} \quad \Leftrightarrow \quad T^*T = TT^*.$$

*Demostración.* La primera implicación fue probada al comienzo de esta subsección.

Recíprocamente, supongamos que $T$ es normal. Por el teorema anterior, existe una base ortonormal $\mathcal{E}$ de $V$ tal que $A := [T]_{\mathcal{E}}$ es una matriz triangular superior (y normal). Luego, aplicando la Proposición 8.43, resulta que $A$ es diagonal. Por lo tanto, $T$ es unitariamente diagonalizable. □

## 8.7. Ejercicios

1. Consideremos $\mathbb{R}^2$ con el producto interno dado por

$$\langle (x_1, x_2), (y_1, y_2) \rangle = x_1y_1 - 2x_1y_2 - 2x_2y_1 + 8x_2y_2\,.$$

Sea $T \in L(\mathbb{R}^2)$ dado por $T(x_1, x_2) = (2x_1 - x_2, x_1 + x_2)$, calcular $T^*$.

2. Sea $V$ un $\mathbb{K}$-espacio vectorial de dimensión finita dotado con un producto interno. Dados $S, T \in L(V)$ y $\alpha \in \mathbb{K}$, probar que:

   *a*) $(S+T)^* = S^* + T^*$.

   *b*) $(\alpha T)^* = \overline{\alpha} T^*$.

   *c*) $(ST)^* = T^* S^*$.

   *d*) $(T^*)^* = T$.

   *e*) Si $T$ es inversible, entonces $(T^*)^{-1} = (T^{-1})^*$.

3. Sea $(V, \langle \cdot, \cdot \rangle)$ un espacio con producto interno de dimensión finita. Para $u \in V$ fijo, definimos $T_u : V \to V$ como

$$T_u(v) = \langle v, u \rangle \, u \quad \text{para } v \in V\,.$$

   *a*) Probar que $T_u \in L(V)$ y hallar $T_u^*$. ¿Bajo qué condiciones $T_u$ es una proyección ortogonal?

   *b*) Hallar $N(T_u)$. ¿Bajo qué condiciones $T_u$ es un isomorfismo?

4. Determinar si los siguientes operadores son autoadjuntos considerando en cada espacio vectorial el producto interno usual.

   *a*) La homotecia $H_2$ y la proyección $P_Y$ sobre el eje y en $L(\mathbb{R}^2)$.

   *b*) $T \in L(\mathbb{R}^3)$ dado por $T(x, y, z) = (x + y, x, -z)$.

   *c*) $T \in L(\mathbb{C}^{3\times 3})$ dado por $T(A) = A^\top$.

5. Sean $V$ y $W$ dos espacios con producto interno (de dimensión finita). Probar que si $T \in L(V, W)$ entonces

$$N(T^*) = \operatorname{Im}(T)^\perp \quad \text{y} \quad \operatorname{Im}(T^*) = N(T)^\perp.$$

6. Sea $T \in L(\mathbb{R}^4)$ dado por

$$T(x_1, x_2, x_3, x_4) = (x_2, x_3, x_4, 0)\,.$$

   *a*) Hallar $T^*$. ¿Es $T$ un operador normal?

   *b*) Hallar el núcleo y la imagen de $T^*$.

7. Dado $V$ un $\mathbb{C}$-espacio con producto interno, sea $T \in L(V)$. Probar que:

   *a*) Si $\lambda$ es un autovalor de $T$, entonces $\overline{\lambda}$ es un autovalor de $T^*$.

   *b*) Si $T$ es autoadjunto y $\lambda$ es un autovalor de $T$, entonces $\lambda \in \mathbb{R}$.

8. Dado un espacio con producto interno $(V, \langle \cdot, \cdot \rangle)$ de dimensión finita, sea $W$ un subespacio de $V$. Probar que si $U \in L(V)$ es unitario y $W$ es $U$-invariante, entonces $W^\perp$ es $U$-invariante.

9. Probar que si $\lambda$ es un autovalor de una isometría $S \in L(V)$ entonces $|\lambda| = 1$. ¿Cuáles son los valores posibles para el determinante de $S$?

10. Dado un $\mathbb{C}$-espacio con producto interno $V$ de dimensión finita, sea $U \in L(V)$ unitario.

    *a*) Probar que si $\lambda$ es un autovalor de $T$, entonces $|\lambda| = 1$.

*b*) Mostrar que si $x \in V$ es tal que $Ux = \lambda x$, entonces $U^*x = \lambda^{-1}x = \overline{\lambda}x$.

*c*) Dados dos autovalores de $U$ distintos, $\lambda$ y $\mu$, mostrar que los autovectores asociados a $\lambda$ son ortogonales a los autovectores asociados a $\mu$.

11. Dado $n \in \mathbb{N}$, hallar una matriz $A \in \mathbb{C}^{n\times n}$ que sea normal pero que no sea ni unitaria ni autoadjunta.

12. Sea $A = \begin{pmatrix} 5 & -2i & 4 \\ 2i & 8 & -2i \\ 4 & 2i & 5 \end{pmatrix}$. Probar que es autoadjunta y hallar una matriz unitaria $U$ tal que $U^{-1}AU$ sea diagonal.

13. Sea $T \in L(\mathbb{R}^3)$ tal que $[T]_{\mathcal{E}} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & -1 & 2 \\ 0 & 2 & -1 \end{pmatrix}$, donde $\mathcal{E}$ es la base canónica de $\mathbb{R}^3$.

*a*) Hallar una base ortonormal $\{v_1, v_2, v_3\}$ de $\mathbb{R}^3$, que conste de autovectores de $T$.

*b*) Para $i = 1, 2, 3$, sea $T_{v_i} \in L(\mathbb{R}^3)$ definido como en el ejercicio 3, y sea $\lambda_i \in \mathbb{R}$ tal que $Tv_i = \lambda_i v_i$. Probar que $T = \lambda_1 T_{v_1} + \lambda_2 T_{v_2} + \lambda_3 T_{v_3}$ y que $I = T_{v_1} + T_{v_2} + T_{v_3}$.

14. Dado un espacio con producto interno $(V, \langle\cdot,\cdot\rangle)$, sea $W$ un subespacio de $V$.

*a*) Si $v = w + w'$ con $w \in W$ y $w' \in W^\perp$, probar que $J : V \to V$ dado por

$$J(v) = w - w',$$

es un operador lineal autoadjunto y unitario.

*b*) Consideremos $V = \mathbb{R}^3$ dotado con el producto interno usual y $W = \overline{\{(1, 0, 1)\}}$.

*i*. Hallar la representación matricial de $J$ en la base canónica de $\mathbb{R}^3$.

*ii*. Hallar una base ortonormal de autovectores de $J$.

15. Dado un vector unitario $v \in \mathbb{R}^n$ (es decir, $v^\top v = 1$), consideremos la matriz simétrica

$$J = I - 2vv^\top.$$

*a*) Mostrar que $J^2 = I$. Esto dice que $J$ no sólo es simétrica, sino que además es una matriz ... (complete la oración y justifique adecuadamente).

*b*) Mostrar que $v$ es uno de los autovectores de $J$, y calcular el autovalor correspondiente.

*c*) Si $u$ es un vector perpendicular a $v$, mostrar que $u$ también es un autovector de $J$ y calcular el autovalor correspondiente.

*d*) Teniendo en cuenta todo lo anterior, ¿se anima a proponer una base de $\mathbb{R}^n$ que diagonalice a $J$? ¿cuál es la representación diagonal de $J$ con respecto a esta base?

16. Consideremos a $\mathbb{C}^2$ dotado con el producto interno usual $\langle\cdot,\cdot\rangle$ y sea $J \in L(\mathbb{C}^2)$ tal que $[J]_{\mathcal{E}} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Además, definamos $[x, y]_J = \langle Jx, y\rangle \quad$ para $x, y \in \mathbb{C}^2$.

*a*) Probar que $J = J^* = J^{-1}$ y $J^2 = I$.

*b*) Probar que:

*i.* $[\alpha\, x + z, y]_J = \alpha\, [x, y]_J + [z, y]_J$ para cualesquiera $x, y \in \mathbb{C}^2$ y $\alpha \in \mathbb{C}$.

*ii.* $\overline{[x, y]}_J = [y, x]_J$ para cualesquiera $x, y \in \mathbb{C}^2$.

*c*) ¿Podemos asegurar que $[\cdot, \cdot]_J$ es un producto interno sobre $\mathbb{C}^2$?
Sugerencia: considerar el vector $(1, 1)$.

*d*) Hallar $[\cdot, \cdot]_J$ para los vectores de la base canónica de $\mathbb{C}^2$. Comparar con las propiedades del producto interno usual.

*e*) Dado un subconjunto $S$ de $\mathbb{C}^2$, definamos

$$S^{[\perp]} := \left\{ y \in \mathbb{C}^2 : [x, y]_J = 0 \text{ para todo } x \in S \right\}.$$

Hallar $S^{[\perp]}$ para $S = \overline{\{(1, 1)\}}$. Comparar con $S^{\perp}$.

17. Sea $V$ un $\mathbb{C}$-espacio con producto interno de dimensión finita. Diremos que $T \in L(V)$ es **(semidefinido) positivo** si

$$\langle Tv, v\rangle \geq 0 \qquad \text{para todo } v \in V.$$

*a*) Mostrar que, si $T$ es positivo entonces $T$ es autoadjunto.

*b*) Dado $T \in L(V)$ autoadjunto, probar que las siguientes condiciones son equivalentes:

i) $T$ es positivo;

ii) todos los autovalores de $T$ son no negativos;

iii) $T$ tiene una **raíz cuadrada** positiva, es decir, existe $R \in L(V)$ positivo tal que $R^2 = T$;

iv) existe un operador $R \in L(V)$ tal que $T = R^*R$.

# Capítulo 9

# Formas bilineales

En este capítulo estudiaremos formas bilineales sobre espacios vectoriales de dimensión finita. En este capítulo supondremos que $\mathbb{K} = \mathbb{C}$ ó que $\mathbb{K}$ es un subcuerpo de $\mathbb{C}$. Esta hipótesis se justifica en el hecho de que, muchas veces, necesitaremos utilizar que $\mathbb{K}$ es un cuerpo de característica 0, es decir, si $n \in \mathbb{N}$ la suma $1 + \ldots + 1$ ($n$ veces) en $\mathbb{K}$ es distinta de 0.

**Definición 9.1.** Sea $V$ un $\mathbb{K}$-espacio vectorial. Una **forma bilineal** sobre $V$ es una función $\mathcal{A} : V \times V \to \mathbb{K}$ que satisface:

$$\begin{aligned} \mathcal{A}(\alpha u_1 + u_2, v) &= \alpha\, \mathcal{A}(u_1, v) + \mathcal{A}(u_2, v), \qquad && u_1, u_2, v \in V,\ \alpha \in \mathbb{K}; \\ \mathcal{A}(u, \beta v_1 + v_2) &= \beta\, \mathcal{A}(u, v_1) + \mathcal{A}(u, v_2), \qquad && u, v_1, v_2 \in V,\ \beta \in \mathbb{K}. \end{aligned}$$

Comencemos presentando algunos ejemplos.

**Ejemplo 9.2.** Si $\mathbb{K} = \mathbb{R}$ y $V = \mathbb{R}^2$, las funciones:

$$\begin{aligned} \mathcal{A}((x, y), (u, v)) &= (x + u)^2 - (x - u)^2 + yv, \\ \mathcal{A}'((x, y), (u, v)) &= xv + yu, \end{aligned}$$

son formas bilineales sobre $\mathbb{R}^2$.

**Ejemplo 9.3.** Dado un cuerpo $\mathbb{K}$, sea $V = \mathbb{K}^{m\times n}$. Fijada una matriz $A \in \mathbb{K}^{m\times m}$, sea $\mathcal{A} : V \times V \to \mathbb{K}$ dada por

$$\mathcal{A}(X, Y) = \operatorname{tr}(Y^\top A X), \qquad X, Y \in \mathbb{K}^{m\times n}.$$

Es fácil comprobar que $\mathcal{A}$ es una forma bilineal sobre $\mathbb{K}^{m\times n}$ (ejercicio).

En particular, si $n = 1$, entonces $V = \mathbb{K}^{m\times 1}$ y $\mathcal{A}$ tiene un aspecto similar a las formas del ejemplo anterior:

$$\begin{aligned} \mathcal{A}(X, Y) &= \operatorname{tr}(Y^\top A X) = Y^\top A X \\ &= (y_1, y_2, \ldots, y_m) \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1m} \\ a_{21} & a_{22} & \ldots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mm} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} \\ &= \sum_{i=1}^{m} \sum_{j=1}^{m} a_{ij} x_j y_i \ . \end{aligned}$$

**Ejemplo 9.4.** Dado un $\mathbb{K}$-espacio vectorial $V$, sean $f, g \in V^*$. Luego, sea $\mathcal{A} : V \times V \to \mathbb{K}$ definida por

$$\mathcal{A}(v, v) = f(u)g(v), \qquad u, v \in V.$$

A partir de la linealidad de los funcionales $f$ y $g$, se desprende que $\mathcal{A}$ es una forma bilineal sobre $V$.

Llamaremos $\mathrm{Bil}(V)$ al conjunto de formas bilineales sobre un $\mathbb{K}$-espacio vectorial $V$.

**Proposición 9.5.** *Si $V$ es un $\mathbb{K}$-espacio vectorial entonces $\mathrm{Bil}(V)$ también es un $\mathbb{K}$-espacio vectorial, es decir,*

$$\textit{si } \mathcal{A}, \mathcal{A}' \in \mathrm{Bil}(V) \textit{ y } \alpha \in \mathbb{K} \textit{ entonces } \alpha\mathcal{A} + \mathcal{A}' \in \mathrm{Bil}(V).$$

*Demostración.* Ejercicio. □

Dado un $\mathbb{K}$-espacio vectorial $V$ con $\dim V = n$, supongamos que $\mathcal{B} = \{b_1, \ldots, b_n\}$ es una base de $V$. Si $\mathcal{A} \in \mathrm{Bil}(V)$, veamos que $\mathcal{A}$ está determinada por los valores que toma sobre pares de elementos de la base $\mathcal{B}$. Dados $u, v \in V$, escribámoslos con respecto a la base $\mathcal{B}$:

$$u = \sum_{i=1}^{n} x_i b_i, \quad v = \sum_{j=1}^{n} y_j b_j,$$

para ciertos $x_1, \ldots, x_n, y_1, \ldots, y_n \in \mathbb{K}$. Luego,

$$\begin{aligned} \mathcal{A}(u, v) = \mathcal{A}\left(\sum_{i=1}^{n} x_i b_i, \sum_{j=1}^{n} y_j b_j\right) &= \sum_{i=1}^{n} x_i \ \mathcal{A}\left(b_i, \sum_{j=1}^{n} y_j b_j\right) \\ &= \sum_{i=1}^{n} x_i \sum_{j=1}^{n} y_j \ \mathcal{A}(b_i, b_j) \\ &= \sum_{j=1}^{n} y_j \left(\sum_{i=1}^{n} \mathcal{A}(b_i, b_j) x_i\right). \end{aligned}$$

Entonces, si consideramos la matriz $A = (A_{ij}) \in \mathbb{K}^{n\times n}$ dada por

$$A_{ij} = \mathcal{A}(b_j, b_i),$$

resulta que

$$\mathcal{A}(u, v) = [v]_{\mathcal{B}}^{\top} A [u]_{\mathcal{B}}. \tag{9.1}$$

Recíprocamente, dada una matriz $A \in \mathbb{K}^{n\times n}$ cualquiera, podemos definir una (única) forma bilineal $\mathcal{A}$ sobre $V$ mediante (9.1) y, en particular, tendremos que $\mathcal{A}(b_j, b_i) = A_{ij}$.

**Definición 9.6.** Dado un $\mathbb{K}$-espacio vectorial $V$ con $\dim V = n$, sea $\mathcal{B} = \{b_1, \ldots, b_n\}$ una base de $V$. Si $\mathcal{A} \in \mathrm{Bil}(V)$, definimos la **matriz de $\mathcal{A}$ con respecto a la base $\mathcal{B}$** como la matriz $A = (A_{ij}) \in \mathbb{K}^{n\times n}$ dada por

$$A_{ij} = \mathcal{A}(b_j, b_i).$$

Usualmente la anotaremos $[\mathcal{A}]_{\mathcal{B}}$.

**Teorema 9.7.** *Sea $V$ un $\mathbb{K}$-espacio vectorial con $\dim V = n$. Cada base $\mathcal{B}$ de $V$ determina un isomorfismo entre $\mathrm{Bil}(V)$ y $\mathbb{K}^{n\times n}$, mediante la aplicación:*

$$\mathcal{A} \mapsto [\mathcal{A}]_{\mathcal{B}}.$$

*Demostración.* Ya vimos que la aplicación $\mathcal{A} \mapsto [\mathcal{A}]_{\mathcal{B}}$ es una biyección entre $\operatorname{Bil}(V)$ y $\mathbb{K}^{n\times n}$, sólo resta probar que es una transformación lineal, lo que queda como ejercicio. $\square$

**Corolario 9.8.** *Si $\mathcal{B} = \{b_1, \ldots, b_n\}$ es una base de $V$ y $\mathcal{B}^* = \{f_1, \ldots, f_n\}$ es su base dual, entonces las $n^2$ formas bilineales definidas por*

$$\mathcal{A}_{ij}(u, v) := f_i(v)f_j(u), \qquad u, v \in V, \tag{9.2}$$

*para $i, j = 1, \ldots, n$, forman una base de* $\operatorname{Bil}(V)$.

*Demostración.* En primer lugar, notemos que las funciones $\mathcal{A}_{ij}$ definidas en (9.2) son formas bilineales del tipo considerado en el Ejemplo 9.4.

Además, dados $u, v \in V$ cualesquiera, escribámoslos con respecto a la base $\mathcal{B}$:

$$u = \sum_{i=1}^{n} u_i b_i, \qquad v = \sum_{i=1}^{n} v_i b_i,$$

para ciertos $u_1, \ldots, u_n, v_1, \ldots, v_n \in \mathbb{K}$. Luego,

$$\mathcal{A}_{ij}(u, v) = f_i(v)f_j(u) = v_i u_j.$$

Por otra parte, dada una forma $\mathcal{A} \in \operatorname{Bil}(V)$ cualquiera, si $A = [\mathcal{A}]_{\mathcal{B}} \in \mathbb{K}^{n\times n}$ entonces

$$\mathcal{A}(u, v) = \sum_{i=1}^{n}\sum_{j=1}^{n} A_{ij} u_j v_i = \sum_{i=1}^{n}\sum_{j=1}^{n} A_{ij}\mathcal{A}_{ij}(u, v),$$

es decir,

$$\mathcal{A} = \sum_{i=1}^{n}\sum_{j=1}^{n} A_{ij}\mathcal{A}_{ji}.$$

Por lo tanto, $\{\mathcal{A}_{ij} : \ i, j = 1, \ldots, n\}$ es un conjunto de generadores de $\operatorname{Bil}(V)$, y como tiene $n^2 = \dim \operatorname{Bil}(V)$ elementos, resulta que es una base de este $\mathbb{K}$-espacio vectorial. $\square$

De la prueba del corolario anterior se desprende que

$$[\mathcal{A}_{ij}]_{\mathcal{B}} = E^{ij},$$

la matriz cuya única entrada no nula es un 1 en la fila $i$ y columna $j$. Como estas matrices forman una base de $\mathbb{K}^{n\times n}$ es natural entonces que las formas $\mathcal{A}_{ij}$ conformen una base de $\operatorname{Bil}(V)$.

Veamos que ocurre con la matriz asociada a una forma bilineal cuando pasamos de una base de $V$ a otra. Supongamos que $\mathcal{B}' = \{b_1', \ldots, b_n'\}$ es otra base de $V$ y descubramos que relación hay entre $[\mathcal{A}]_{\mathcal{B}}$ y $[\mathcal{A}]_{\mathcal{B}'}$.

Si $A := [\mathcal{A}]_{\mathcal{B}}$ y $P := P_{\mathcal{B},\mathcal{B}'}$ es la matriz cambio de base de $\mathcal{B}'$ en $\mathcal{B}$, entonces

$$\begin{aligned}\mathcal{A}(u, v) &= [v]_{\mathcal{B}}^{\top} A [u]_{\mathcal{B}} = (P[v]_{\mathcal{B}'})^{\top} A (P[u]_{\mathcal{B}'}) = ([v]_{\mathcal{B}'}^{\top} P^{\top}) A (P[u]_{\mathcal{B}'}) \\ &= [v]_{\mathcal{B}'}^{\top} (P^{\top} A P)[u]_{\mathcal{B}'},\end{aligned}$$

con lo cual

$$[\mathcal{A}]_{\mathcal{B}'} = P^{\top}[\mathcal{A}]_{\mathcal{B}} P. \tag{9.3}$$

**Ejemplo 9.9.** Sea $\mathcal{A}$ la forma bilineal sobre $\mathbb{R}^2$ dada por

$$\mathcal{A}((x,y),(u,v)) = xu + xv + yu + yv.$$

Si $\mathcal{E} = \{e_1 = (1,0), e_2 = (0,1)\}$ es la base canónica de $\mathbb{R}^2$ entonces

$$[\mathcal{A}]_{\mathcal{E}} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

Por otra parte, si $\mathcal{B} = \{b_1 = (1,-1), b_2 = (1,1)\}$ entonces

$$P := P_{\mathcal{E},\mathcal{B}} = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}.$$

Por lo tanto,

$$\begin{aligned} [\mathcal{A}]_{\mathcal{B}} = P^\top [\mathcal{A}]_{\mathcal{E}} P = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} &= \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 2 \\ 0 & 2 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 \\ 0 & 4 \end{pmatrix}. \end{aligned}$$

Esto dice que, si escribimos

$$\begin{aligned} (x,y) &= x'(1,-1) + y'(1,1), \\ (u,v) &= u'(1,-1) + v'(1,1), \end{aligned}$$

entonces

$$\mathcal{A}((x,y),(u,v)) = 4y'v'.$$

De hecho, podemos ver que $y' = \frac{x+y}{2}$ y $v' = \frac{u+v}{2}$, y efectivamente,

$$4y'v' = 4\frac{(x+y)}{2}\frac{(u+v)}{2} = (x+y)(u+v) = xu + xv + yu + yv = \mathcal{A}((x,y),(u,v)).$$

Una consecuencia de la fórmula de cambio de bases (9.3) es que si $A$ y $A'$ son dos matrices de $n \times n$ que representan a la misma forma bilineal (con respecto a dos bases diferentes), entonces $A$ y $A'$ tienen el mismo rango. Por lo tanto, podemos definir el rango de una forma bilineal de la siguiente manera:

**Definición 9.10.** Si $V$ es un $\mathbb{K}$-espacio vectorial de dimensión finita y $\mathcal{A} \in \mathrm{Bil}(V)$, definimos el **rango de** $\mathcal{A}$ como el rango de la matriz $[\mathcal{A}]_{\mathcal{B}}$ para alguna (cualquier) base $\mathcal{B}$ de $V$.

**Proposición 9.11.** *Dado un $\mathbb{K}$-espacio vectorial $V$ con $\dim V = n$, sea $\mathcal{A} \in \mathrm{Bil}(V)$. Entonces, las siguientes condiciones son equivalentes:*

*i) $\mathcal{A}$ tiene rango completo, o sea, el rango de $\mathcal{A}$ es igual a $n$;*

*ii) para todo $u \in V$, $u \neq \vec{0}$, existe $v \in V$ tal que $\mathcal{A}(u,v) \neq 0$;*

*iii) para todo $v \in V$, $v \neq \vec{0}$, existe $u \in V$ tal que $\mathcal{A}(u,v) \neq 0$.*

*Si $\mathcal{A} \in \mathrm{Bil}(V)$ satisface alguna de las condiciones anteriores, diremos que $\mathcal{A}$ es una forma* **no degenerada**.

*Demostración.* Fijada una base $\mathcal{B} = \{b_1, \ldots, b_n\}$ de $V$, sea $A = [\mathcal{A}]_{\mathcal{B}}$ y recordemos que

$$\mathcal{A}(u,v) = [v]_{\mathcal{B}}^{\top} A [u]_{\mathcal{B}},$$

para cualesquiera $u, v \in V$. La condición $i)$ equivale a que $A \in \mathbb{K}^{n\times n}$ sea invertible.

Supongamos que $A \in \mathbb{K}^{n\times n}$ es invertible. Fijado $u \in V$, $u \neq \vec{0}$, notemos que $A[u]_{\mathcal{B}} \neq \vec{0}$. Luego, el vector $A[u]_{\mathcal{B}} \in \mathbb{K}^n$ tiene alguna coordenada no nula, supongamos que es la $i$-ésima. Si tomamos el vector $v := b_i$ entonces

$$[v]_{\mathcal{B}} = [b_i]_{\mathcal{B}} = e_i = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \begin{matrix} \\ \\ \\ \leftarrow\ i \\ \\ \\ \\ \end{matrix} ,$$

y por lo tanto, $\mathcal{A}(u,v) = [v]_{\mathcal{B}}^{\top} A[u]_{\mathcal{B}} \neq 0$ ya que coincide con la $i$-ésima coordenada de $A[u]_{\mathcal{B}}$. Esto prueba que $i)$ implica $ii)$.

Por otra parte, si $A$ no es invertible, existe un vector $X = (x_1, \ldots, x_n)^{\top} \in \mathbb{K}^{n\times 1}$ no nulo tal que $AX = \vec{0}$. Si definimos el vector $u := x_1 b_1 + \ldots + x_n b_n \in V$, tenemos que $u \neq \vec{0}$ y además, para cualquier $v \in V$,

$$\mathcal{A}(u,v) = [v]_{\mathcal{B}}^{\top} A[u]_{\mathcal{B}} = [v]_{\mathcal{B}}^{\top} AX = [v]_{\mathcal{B}}^{\top} \vec{0} = \vec{0}.$$

Acabamos de probar que $ii)$ implica $i)$ (por la contrarrecíproca).

Para probar la equivalencia entre $i)$ y $iii)$, consideremos la forma bilineal

$$\mathcal{A}^t(u,v) := \mathcal{A}(v,u).$$

Es fácil ver que $[\mathcal{A}^t]_{\mathcal{B}} = A^{\top}$. Luego, como $\operatorname{rg}(A) = \operatorname{rg}(A^{\top})$, la equivalencia se deduce de la prueba anterior. □

**Ejercicio 9.12.** Dado $V = \mathbb{R}^{2\times 3}$, sea $\mathcal{A}$ la forma bilineal definida por $\mathcal{A}(X,Y) = \operatorname{tr}(Y^{\top} AX)$, donde

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

Determinar si $\mathcal{A}$ es una forma degenerada o no degenerada.

## 9.1. Formas bilineales simétricas y antisimétricas

**Definición 9.13.** Dado un $\mathbb{K}$-espacio vectorial $V$, sea $\mathcal{A} \in \operatorname{Bil}(V)$. Diremos que $\mathcal{A}$ es una forma bilineal **simétrica** si

$$\mathcal{A}(u,v) = \mathcal{A}(v,u) \qquad \text{para todo } u, v \in V.$$

En la demostración de la Proposición 9.11, a partir de una forma $\mathcal{A} \in \operatorname{Bil}(V)$ definimos una segunda forma bilineal $\mathcal{A}^t \in \operatorname{Bil}(V)$, mediante

$$\mathcal{A}^t(u,v) := \mathcal{A}(v,u),$$

es decir, trasponiendo las entradas de $\mathcal{A}$. La definición anterior dice que $\mathcal{A}$ es simétrica si y sólo si $\mathcal{A}^t = \mathcal{A}$.

**Observación 9.14.** Si $V$ es un $\mathbb{K}$-espacio vectorial de dimensión finita, $\mathcal{A} \in \text{Bil}(V)$ es simétrica si y sólo si

$$[\mathcal{A}]_{\mathcal{B}} \text{ es una matriz simétrica,}$$

para cualquier base $\mathcal{B}$ de $V$.

**Definición 9.15.** Dado un $\mathbb{K}$-espacio vectorial $V$, sea $\mathcal{A} \in \text{Bil}(V)$. Diremos que $\mathcal{A}$ es una forma bilineal **antisimétrica** si

$$\mathcal{A}(u, v) = -\mathcal{A}(v, u) \qquad \text{para todo } u, v \in V.$$

Notemos que la definición anterior dice que $\mathcal{A}$ es antisimétrica si y sólo si $\mathcal{A}^t = -\mathcal{A}$.

**Proposición 9.16.** *Dado un $\mathbb{K}$-espacio vectorial $V$, sea $\mathcal{A} \in \text{Bil}(V)$. Entonces, $\mathcal{A}$ es antisimétrica si y sólo si*

$$\mathcal{A}(v, v) = 0 \quad \textit{para todo } v \in V. \tag{9.4}$$

*Demostración.* Supongamos que $\mathcal{A}$ es antisimétrica. Luego, dado $v \in V$,

$$\mathcal{A}(v, v) = -\mathcal{A}(v, v),$$

de donde se deduce que $\mathcal{A}(v, v) = 0$. Recíprocamente, supongamos que vale (9.4). Dados $u, v \in V$,

$$0 = \mathcal{A}(u + v, u + v) = \mathcal{A}(u, u) + \mathcal{A}(u, v) + \mathcal{A}(v, u) + \mathcal{A}(v, v) = \mathcal{A}(u, v) + \mathcal{A}(v, u).$$

Por lo tanto, $\mathcal{A}(u, v) = -\mathcal{A}(v, u)$ para todo $u, v \in V$. □

**Observación 9.17.** Si $V$ es un $\mathbb{K}$-espacio vectorial de dimensión finita, $\mathcal{A} \in \text{Bil}(V)$ es antisimétrica si y sólo si

$$[\mathcal{A}]_{\mathcal{B}} \text{ es una matriz antisimétrica,}$$

para cualquier base $\mathcal{B}$ de $V$.

En particular, los elementos de la diagonal de $[\mathcal{A}]_{\mathcal{B}} = (A_{ij})$ son todos nulos, ya que

$$A_{ii} = \mathcal{A}(b_i, b_i) = -\mathcal{A}(b_i, b_i) = -A_{ii}.$$

**Proposición 9.18.** *Dado un $\mathbb{K}$-espacio vectorial $V$, toda $\mathcal{A} \in \text{Bil}(V)$ puede descomponerse como la suma de una forma bilineal simétrica y una antisimétrica.*

*Demostración.* Dada una forma $\mathcal{A} \in \text{Bil}(V)$, consideremos las formas $\mathcal{A}_{sim}, \mathcal{A}_{ant} \in \text{Bil}(V)$ definidas por

$$\mathcal{A}_{sim} := \frac{1}{2}(\mathcal{A} + \mathcal{A}^t) \qquad \text{y} \qquad \mathcal{A}_{ant} := \frac{1}{2}(\mathcal{A} - \mathcal{A}^t).$$

Veamos que $\mathcal{A}_{sim}$ es simétrica: dados $u, v \in V$,

$$\begin{aligned} \mathcal{A}_{sim}(v, u) &= \frac{1}{2}(\mathcal{A}(v, u) + \mathcal{A}^t(v, u)) = \frac{1}{2}(\mathcal{A}(v, u) + \mathcal{A}(u, v)) \\ &= \frac{1}{2}(\mathcal{A}(u, v) + \mathcal{A}^t(u, v)) = \mathcal{A}_{sim}(u, v). \end{aligned}$$

Análogamente, $\mathcal{A}_{ant}$ es antisimétrica, ya que para $u, v \in V$,

$$\begin{aligned} \mathcal{A}_{ant}(v, u) &= \frac{1}{2}(\mathcal{A}(v, u) - \mathcal{A}^t(v, u)) = \frac{1}{2}(\mathcal{A}(v, u) - \mathcal{A}(u, v)) \\ &= -\frac{1}{2}(\mathcal{A}(u, v) - \mathcal{A}(v, u)) = -\frac{1}{2}(\mathcal{A}(u, v) - \mathcal{A}^t(u, v)) = -\mathcal{A}_{ant}(u, v). \end{aligned}$$

Además, es inmediato que $\mathcal{A}_{sim} + \mathcal{A}_{ant} = \mathcal{A}$. □

**Definición 9.19.** Sea $V$ un $\mathbb{K}$-espacio vectorial. Diremos que una función $Q : V \to \mathbb{K}$ es una **forma cuadrática** si

$$Q(v) = \mathcal{A}(v, v),$$

para alguna forma bilineal simétrica $\mathcal{A} \in \mathrm{Bil}(V)$.

Si $\dim V = n$ y $\mathcal{B} = \{b_1, \ldots, b_n\}$ es una base de $V$, supongamos que $A = (A_{ij}) \in \mathbb{K}^{n\times n}$ es la representación matricial de $\mathcal{A}$ con respecto a la base $\mathcal{B}$. Luego, dado $v \in V$,

$$\begin{aligned} Q(v) = \mathcal{A}(v,v) = [v]_{\mathcal{B}}^{\top} A [v]_{\mathcal{B}} &= (v_1, \ldots, v_n) \begin{pmatrix} A_{11} & A_{12} & \ldots & A_{1n} \\ A_{21} & A_{22} & \ldots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \ldots & A_{nn} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} A_{ij} v_i v_j = \sum_{i=1}^{n} A_{ii} v_i^2 + 2 \sum_{i=1}^{n} \sum_{j=1}^{i-1} A_{ij} v_i v_j, \end{aligned} \tag{9.5}$$

si $[v]_{\mathcal{B}} = (v_1, \ldots, v_n)^{\top}$.

Recíprocamente, cada expresión polinomial de la forma (9.5) en la cual las constantes $A_{ij} \in \mathbb{K}$ se suponen dadas y $v_1, \ldots, v_n$ representan a $n$ variables, define una forma cuadrática en cualquier $\mathbb{K}$-espacio vectorial con $\dim V = n$. De hecho, si $\mathcal{B} = \{b_1, \ldots, b_n\}$ es una base de $V$, dados

$$u = \sum_{i=1}^{n} u_i b_i, \qquad v = \sum_{i=1}^{n} v_i b_i,$$

consideremos la forma bilineal $\mathcal{A} : V \times V \to \mathbb{K}$ dada por

$$\mathcal{A}(u, v) = \sum_{i=1}^{n} A_{ii} u_i v_i + \sum_{i=1}^{n} \sum_{j=1}^{i-1} A_{ij} u_i v_j + \sum_{i=1}^{n} \sum_{j=1}^{i-1} A_{ij} v_i u_j, \tag{9.6}$$

entonces

$$Q(v) := \mathcal{A}(v, v) = \sum_{i=1}^{n} A_{ii} v_i^2 + 2 \sum_{i=1}^{n} \sum_{j=1}^{i-1} A_{ij} v_i v_j,$$

como queríamos probar.

**Observación 9.20.** Si $\mathcal{A} \in \mathrm{Bil}(V)$ es una forma bilineal arbitraria, la función $Q : V \to \mathbb{K}$ dada por $Q(v) = \mathcal{A}(v, v)$, también es una forma cuadrática. La hipótesis adicional de que la forma $\mathcal{A}$ sea simétrica (en la Definición 9.19) se justifica en el hecho de que una forma cuadrática puede ser determinada por más de una forma bilineal (no necesariamente simétrica). Por ejemplo, en $\mathbb{R}^2$, las formas bilineales

$$\begin{aligned} \mathcal{A}((x, y), (u, v)) &= xu + 3xv + 9yv, \\ \mathcal{A}'((x, y), (u, v)) &= xu + 3yu + 9yv, \end{aligned}$$

generan la misma forma cuadrática $Q(x, y) = x^2 + 3xy + 9y^2$.

Sin embargo, el próximo resultado muestra que cada forma cuadrática está determinada por una única forma bilineal simétrica. Notemos que las formas bilineales de la observación anterior tienen la misma parte simétrica

$$\mathcal{A}_{sim}((x, y), (u, v)) = xu + \frac{3}{2} xv + \frac{3}{2} yu + 9yv,$$

la cual genera también la forma cuadratica $Q(x, y) = x^2 + 3xy + 9y^2$.

**Proposición 9.21.** *Dado un $\mathbb{K}$-espacio vectorial $V$, sea $Q : V \to \mathbb{K}$ una forma cuadrática sobre $V$. Entonces, existe una única forma bilineal simétrica $\mathcal{A}_{sim} \in \mathrm{Bil}(V)$ tal que*

$$Q(v) = \mathcal{A}_{sim}(v, v) \quad \textit{para cada } v \in V.$$

*Demostración.* Por definición, si $Q : V \to \mathbb{K}$ es una forma cuadrática, existe una forma bilineal simétrica $\mathcal{A}_{sim} \in \mathrm{Bil}(V)$ tal que $Q(v) = \mathcal{A}_{sim}(v, v)$ para cada $v \in V$.

Además, $\mathcal{A}_{sim}$ está unívocamente determinada por $Q$, ya que

$$\begin{aligned} Q(u+v) &= \mathcal{A}_{sim}(u+v, u+v) = \mathcal{A}_{sim}(u,u) + \mathcal{A}_{sim}(u,v) + \mathcal{A}_{sim}(v,u) + \mathcal{A}_{sim}(v,v) \\ &= Q(u) + 2\,\mathcal{A}_{sim}(u,v) + Q(v), \\ Q(u-v) &= \mathcal{A}_{sim}(u-v, u-v) = \mathcal{A}_{sim}(u,u) - \mathcal{A}_{sim}(u,v) - \mathcal{A}_{sim}(v,u) + \mathcal{A}_{sim}(v,v) \\ &= Q(u) - 2\,\mathcal{A}_{sim}(u,v) + Q(v), \end{aligned}$$

de donde se desprende que

$$\begin{aligned} \mathcal{A}_{sim}(u,v) &= \frac{1}{2}\big(Q(u+v) - Q(u) - Q(v)\big) \qquad \text{(forma polar)}, \\ &= \frac{1}{4}\big(Q(u+v) - Q(u-v)\big) \qquad \text{(identidad de polarización)}. \end{aligned}$$

□

**Definición 9.22.** Dado un $\mathbb{K}$-espacio vectorial $V$ de dimensión finita, sea $\mathcal{B}$ una base de $V$. Si $Q : V \to \mathbb{K}$ es una forma cuadrática sobre $V$, la **matriz de $Q$ con respecto a la base $\mathcal{B}$** es la matriz asociada a la forma bilineal simétrica $\mathcal{A}_{sim}$ tal que $Q(v) = \mathcal{A}_{sim}(v, v)$, es decir,

$$[Q]_{\mathcal{B}} := [\mathcal{A}_{sim}]_{\mathcal{B}}.$$

Notemos que la forma bilineal dada en (9.6) es simétrica, y determina la forma cuadrática (9.5). Luego, la matriz de $Q$ en la base $\mathcal{B}$ es:

$$[Q]_{\mathcal{B}} = [\mathcal{A}]_{\mathcal{B}} = A = \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{21} & A_{22} & \dots & A_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \dots & A_{nn} \end{pmatrix}.$$

Una familia importante de formas bilineales simétricas es la compuesta por los productos internos sobre $\mathbb{R}$-espacios vectoriales, los cuales tratamos en el capítulo anterior. Si $V$ es un $\mathbb{R}$-espacio vectorial, un producto interno sobre $V$ no es más que una forma bilineal simétrica $\mathcal{A} \in \mathrm{Bil}(V)$ que verifica

$$\mathcal{A}(v, v) > 0 \quad \text{si } v \neq \vec{0}. \tag{9.7}$$

Una forma bilineal $\mathcal{A}$ que satisface (9.7) de denomina **definida positiva**. Por lo tanto, un producto interno sobre un $\mathbb{R}$-espacio vectorial es una forma bilineal no degenerada, simétrica y definida positiva sobre dicho espacio.

Teniendo esto en mente, podemos aprovechar parte de la terminología y la maquinaria desarrollada para los productos internos y extrapolarla a formas bilineales simétricas más generales. En particular,

**Definición 9.23.** Si $\mathcal{A}$ es una forma bilineal simétrica sobre un $\mathbb{K}$-espacio vectorial $V$, diremos que dos vectores $u, v \in V$ son **vectores ortogonales con respecto a $\mathcal{A}$** si $\mathcal{A}(u, v) = 0$.

Con los mismos argumentos utilizados para los espacios con producto interno, se puede mostrar que si $S$ es un subconjunto de $V$ entonces

$$S^{\perp} := \{v \in V : \mathcal{A}(v, s) = 0 \text{ para todo } s \in S\}, \tag{9.8}$$

es un subespacio de $V$.

**Ejercicio 9.24.** En $\mathbb{R}^4$, consideremos la forma cuadrática definida por

$$Q(t, x, y, z) = t^2 - x^2 - y^2 - z^2.$$

Si $\mathcal{A}$ es la forma bilineal simétrica asociada a $Q$ y $W = \{(t, t, 0, 0) : t \in \mathbb{R}\}$, mostrar que el subespacio ortogonal a $W$ con respecto a $\mathcal{A}$ es:

$$W^{\perp} = \{(x, x, y, z) : x, y, z \in \mathbb{R}\}.$$

En particular, $W \cap W^{\perp} \neq \{\vec{0}\}$ y $W^{\perp}$ no es un complemento algebraico de $W$, es decir,

$$W + W^{\perp} \neq \mathbb{R}^4.$$

El ejercicio anterior muestra que el subespacio ortogonal con respecto a una forma bilineal simétrica no es necesariamente un complemento. Por lo tanto, hay que ser extremadamente cuidadoso al manejarlo.

Ahora estamos en condiciones de probar el resultado principal de esta sección, el cual asegura que toda forma bilineal simétrica puede ser "diagonalizada", es decir, expresada de forma canónica.

**Teorema 9.25.** *Dado un $\mathbb{K}$-espacio vectorial $V$ de dimensión finita, sea $\mathcal{A} \in \mathrm{Bil}(V)$ una forma bilineal simétrica. Entonces, existe una base $\mathcal{B}$ de $V$ en la cual $[\mathcal{A}]_{\mathcal{B}}$ es una matriz diagonal.*

*Demostración.* Si $\mathcal{A} = 0$ ó si $\dim V = 1$, el resultado es obviamente cierto. Podemos suponer entonces que $\dim V = n > 1$ y que $\mathcal{A} \neq 0$.

Si $Q(v) := \mathcal{A}(v, v) = 0$ para todo $v \in V$, la forma polar de $\mathcal{A}$ (ver la prueba de la Proposición 9.21) muestra que $\mathcal{A} = 0$. Entonces, existe un vector $b_1 \in V$ tal que $\mathcal{A}(b_1, b_1) \neq 0$. Sea $W$ el subespacio generado por $b_1$, $W = \overline{\{b_1\}}$, y consideremos su complemento ortogonal $W^{\perp}$ (con respecto a la forma bilineal $\mathcal{A}$). Veamos que, en este caso particular,

$$V = W \dotplus W^{\perp}.$$

- $W \cap W^{\perp} = \{\vec{0}\}$: Supongamos que $w \in W \cap W^{\perp}$. Como $w \in W$, existe $\alpha \in \mathbb{K}$ tal que $w = \alpha b_1$. Pero además $w \in W^{\perp}$, así que

$$0 = \mathcal{A}(w, w) = \mathcal{A}(\alpha b_1, \alpha b_1) = \alpha^2 \mathcal{A}(b_1, b_1).$$

  Como $\mathcal{A}(b_1, b_1) \neq 0$, resulta que $\alpha = 0$ y, en consecuencia, $w = \vec{0}$.

- $W + W^{\perp} = V$: Dado $v \in V$ arbitrario, sea

$$w := v - \frac{\mathcal{A}(b_1, v)}{\mathcal{A}(b_1, b_1)} b_1.$$

  Luego,

$$\mathcal{A}(b_1, w) = \mathcal{A}\left(b_1, v - \frac{\mathcal{A}(b_1, v)}{\mathcal{A}(b_1, b_1)} b_1\right) = \mathcal{A}(b_1, v) - \frac{\mathcal{A}(b_1, v)}{\mathcal{A}(b_1, b_1)} \mathcal{A}(b_1, b_1) = 0.$$

  Entonces, $w \in W^{\perp}$. De la definición de $w$ se ve que $v$ puede escribirse como la suma de un vector en $W$ y otro en $W^{\perp}$. Por lo tanto, $V = W + W^{\perp}$.

Además, la restricción de $\mathcal{A}$ a $W^\perp$ es una forma bilineal simétrica sobre $W^\perp$. Como $\dim W^\perp = n-1$, podemos suponer por inducción que $W^\perp$ tiene una base $\{b_2, \ldots, b_n\}$ tal que

$$\mathcal{A}(b_i, b_j) = 0 \quad \text{si } i \neq j,\ i, j = 2, \ldots, n.$$

Entonces, $\mathcal{B} := \{b_1, \ldots, b_n\}$ es una base de $V$ tal que $\mathcal{A}(b_i, b_j) = 0$ si $i \neq j$, es decir, $[\mathcal{A}]_\mathcal{B}$ es una matriz diagonal. □

**Corolario 9.26.** *Si $A \in \mathbb{K}^{n\times n}$ es una matriz simétrica, entonces existe $P \in \mathbb{K}^{n\times n}$ invertible tal que $P^\top AP$ es una matriz diagonal.*

Recordemos que, si $\mathbb{K} = \mathbb{R}$, la matriz invertible $P$ del corolario anterior puede elegirse como una matriz ortogonal, i.e. $P^\top = P^{-1}$. Pero como vimos en el capítulo anterior, probar esto no es inmediato.

## 9.2. Formas bilineales simétricas reales

En esta sección trataremos formas bilineales simétricas y formas cuadráticas en espacios vectoriales sobre el cuerpo $\mathbb{R}$ de los números reales. Estas formas aparecen naturalmente en muchas ramas de la matemática y la física.

La siguiente es una clasificación de las formas bilineales simétricas reales utilizada, entre otras cosas, para clasificar los extremos locales de funciones de varias variables.

**Definición 9.27.** Dado un $\mathbb{R}$-espacio vectorial $V$, sea $\mathcal{A} \in \mathrm{Bil}(V)$ una forma simétrica. Diremos que:

i) $\mathcal{A}$ es **definida positiva** si $\mathcal{A}(v, v) > 0$ para todo $v \in V$, $v \neq \vec{0}$;

ii) $\mathcal{A}$ es **semidefinida positiva** si $\mathcal{A}(v, v) \geq 0$ para todo $v \in V$;

iii) $\mathcal{A}$ es **definida negativa** si $\mathcal{A}(v, v) < 0$ para todo $v \in V$, $v \neq \vec{0}$;

iv) $\mathcal{A}$ es **semidefinida negativa** si $\mathcal{A}(v, v) \leq 0$ para todo $v \in V$;

v) $\mathcal{A}$ es **indefinida** si existen $u, v \in V$ tales que $\mathcal{A}(u, u) > 0$ y $\mathcal{A}(v, v) < 0$.

**Ejemplos 9.28.** Consideremos las siguientes formas bilineales simétricas en $\mathbb{R}^n$:

1. el producto interno usual en $\mathbb{R}^n$ es una forma definida positiva;
2. $\mathcal{A}((x_1, \ldots, x_n), (y_1, \ldots, y_n)) := x_1y_1$ es una forma semidefinida positiva;
3. $\mathcal{A}((x_1, \ldots, x_n), (y_1, \ldots, y_n)) := -\sum_{k=1}^n 2^k x_k y_k$ es una forma definida negativa;
4. $\mathcal{A}((x_1, \ldots, x_n), (y_1, \ldots, y_n)) := -8x_ny_n$ es una forma semidefinida negativa;
5. $\mathcal{A}((x_1, \ldots, x_n), (y_1, \ldots, y_n)) := 2x_1y_1 - 5x_ny_n$ es una forma indefinida.

Las formas bilineales simétricas reales pueden diagonalizarse de una forma especial, la que describiremos en el próximo resultado.

**Teorema 9.29.** *Dado un $\mathbb{R}$-espacio vectorial $V$ con $\dim V = n$, sea $\mathcal{A} \in \mathrm{Bil}(V)$ una forma simétrica con rango $r$. Entonces, existen una base $\mathcal{E} = \{e_1, \ldots, e_n\}$ de $V$ y un $p \in \mathbb{N}$, $1 \leq p \leq r$ tales que $[\mathcal{A}]_\mathcal{E} =: D$ es diagonal, y además,*

$$D_{ii} = \begin{cases} 1 & si\ \ i = 1, \ldots, p, \\ -1 & si\ \ i = p+1, \ldots, r, \\ 0 & si\ \ i = r+1, \ldots, n. \end{cases}$$

*Demostración.* Por el Teorema 9.25, si $\mathcal{A} \in \mathrm{Bil}(V)$ es una forma simétrica con rango $r$, existe una base $\mathcal{B} = \{b_1, \dots, b_n\}$ de $V$ tal que

$$[\mathcal{A}]_{\mathcal{B}} = \begin{pmatrix} d_1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & d_p & & & & & & \\ & & & -d_{p+1} & & & & & \\ & & & & \ddots & & & & \\ & & & & & -d_r & & & \\ & & & & & & 0 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 0 \end{pmatrix} \in \mathbb{R}^{n\times n}, \tag{9.9}$$

para ciertos $d_1, \dots, d_r \in \mathbb{R}$ no negativos.

Sea $U \in \mathbb{R}^{n\times n}$ la matriz diagonal cuya diagonal principal está dada por

$$U_{ii} = \begin{cases} \frac{1}{\sqrt{d_i}} & \text{si } \ i = 1, \dots, r, \\ 1 & \text{si } \ i = r+1, \dots, n. \end{cases}$$

Es fácil ver que $D := U[\mathcal{A}]_{\mathcal{B}}U = U^{\top}[\mathcal{A}]_{\mathcal{B}}U$ cumple las condiciones del enunciado. Además, si definimos los vectores

$$e_i = \begin{cases} \frac{1}{\sqrt{d_i}} b_i & \text{si } i = 1, \dots, r, \\ b_i & \text{si } i = r+1, \dots, n, \end{cases}$$

entonces $\mathcal{E} = \{e_1, \dots, e_n\}$ es también una base de $V$ y $[\mathcal{A}]_{\mathcal{E}} = D$. □

El siguiente resultado muestra que la cantidad de 1, la de $-1$, y la de 0, a lo largo de la diagonal principal en la matriz $D$ del teorema anterior son cantidades invariantes, es decir, no dependen la base elegida para representar a $\mathcal{A}$.

**Teorema 9.30.** *Dado un $\mathbb{R}$-espacio vectorial $V$ de dimensión finita, sea $\mathcal{A} \in \mathrm{Bil}(V)$ una forma simétrica. Toda representación matricial diagonal de $\mathcal{A}$ tiene el mismo número $p$ de entradas positivas y el mismo número $q$ de entradas negativas.*

*Demostración.* Por el Teorema 9.25, existe una base $\mathcal{B} = \{b_1, \dots, b_n\}$ de $V$ tal que $[\mathcal{A}]_{\mathcal{B}}$ es una matriz diagonal con, digamos, $p$ entradas positivas y $q$ entradas negativas en la diagonal principal. Ahora, supongamos que $\mathcal{B}' = \{b_1', \dots, b_n'\}$ es otra base de $V$ tal que $[\mathcal{A}]_{\mathcal{B}'}$ es también una matriz diagonal con, en este caso, $p'$ entradas positivas y $q'$ entradas negativas en la diagonal principal. Como el rango de $\mathcal{A}$ no depende de la base elegida para representarla, tenemos que

$$p + q = r = p' + q'.$$

Por lo tanto, alcanza con probar que $p = p'$.

Sin pérdida de generalidad, podemos asumir que las entradas positivas son las que aparecen primero en cada una de las matrices. Luego, consideremos los subespacios

$$W := \overline{\{b_1, \dots, b_p\}} \quad \text{y} \quad W' := \overline{\{b'_{p'+1}, \dots, b'_n\}}.$$

Notemos que $\mathcal{A}(u,u) > 0$ para todo $u \in W$ y que $\mathcal{A}(v,v) \leq 0$ para todo $v \in W'$. Por lo tanto, $W \cap W' = \{\vec{0}\}$. Luego,

$$\dim(W + W') = \dim W + \dim W' - \dim(W \cap W') = p + (n - p') - 0 = n + (p - p'),$$

pero como $W + W'$ es un subespacio de $V$, su dimensión no puede exceder a $n$. Entonces, $p \leq p'$.

De manera análoga, considerando los subespacios $S := \overline{\{b_{p+1}, \ldots, b_n\}}$ y $S' := \overline{\{b'_1, \ldots, b'_{p'}\}}$ que complementan a $W$ y $W'$, respectivamente, se puede ver que $p' \leq p$. En consecuencia $p = p'$, como queríamos probar. □

**Observación 9.31.** Dado un $\mathbb{R}$-espacio vectorial $V$ con $\dim V = n$, sea $\mathcal{A} \in \mathrm{Bil}(V)$ una forma simétrica. Supongamos que $\mathcal{A}$ admite una representación matricial diagonal con $p$ entradas positivas y $q$ entradas negativas a lo largo de la diagonal principal. Entonces,

i) $\mathcal{A}$ es definida positiva $\Leftrightarrow$ $p = n$;

ii) $\mathcal{A}$ es semidefinida positiva $\Leftrightarrow$ $q = 0$;

iii) $\mathcal{A}$ es definida negativa $\Leftrightarrow$ $q = n$;

iv) $\mathcal{A}$ es semidefinida negativa $\Leftrightarrow$ $p = 0$;

v) $\mathcal{A}$ es indefinida $\Leftrightarrow$ $p \neq 0$ y $q \neq 0$.

La diferencia $s := p - q$ se denomina **la signatura de** $\mathcal{A}$.

**Ejercicio 9.32.** Dado un $\mathbb{R}$-espacio vectorial $V$ con $\dim V = n$, sea $\mathcal{A} \in \mathrm{Bil}(V)$ una forma simétrica con rango $r$ y signatura $s$. Probar que

i) $\mathcal{A}$ es semidefinida positiva $\Leftrightarrow$ $s = r$;

ii) $\mathcal{A}$ es semidefinida negativa $\Leftrightarrow$ $s = -r$.

**Corolario 9.33** (Ley de inercia de Sylvester)**.** *Toda forma cuadrática $Q : \mathbb{R}^n \to \mathbb{R}$ tiene una única representación de la forma*

$$Q(x_1, \ldots, x_n) = x_1^2 + \ldots + x_p^2 - x_{p+1}^2 - \ldots - x_{p+q}^2. \tag{9.10}$$

*En particular, $r = p + q$ es el rango de $Q$.*

La cantidad $p$ de términos positivos en (9.10) suele llamarse **indice de inercia positivo** de $Q$ y la cantidad $q$ de términos negativos en (9.10) suele llamarse **indice de inercia negativo** de $Q$.

## 9.3. Formas hermitianas

Las formas que consideraremos a continuación deben su nombre al matemático francés Charles Hermite (1822–1901), quien elaboró varios tratados sobre formas cuadráticas y polinomios ortogonales, entre otras cosas.

**Definición 9.34.** Sea $V$ un $\mathbb{C}$-espacio vectorial. Una función $\mathcal{A} : V \times V \to \mathbb{C}$ que satisface:

$$\begin{aligned} \mathcal{A}(\alpha u_1 + u_2, v) &= \alpha\, \mathcal{A}(u_1, v) + \mathcal{A}(u_2, v), \qquad && u_1, u_2, v \in V,\ \alpha \in \mathbb{K}; \\ \mathcal{A}(u, v) &= \overline{\mathcal{A}(v, u)}, && u, v \in V; \end{aligned}$$

se denomina **forma hermitiana** sobre $V$.

Dados $u, v_1, v_2 \in V$, y $\beta \in \mathbb{K}$, de las propiedades anteriores se deduce que

$$\begin{aligned}\mathcal{A}(u, \beta v_1 + v_2) &= \overline{\mathcal{A}(\beta v_1 + v_2, u)} = \overline{\beta \mathcal{A}(v_1, u) + \mathcal{A}(v_2, u)} = \overline{\beta}\ \overline{\mathcal{A}(v_1, u)} + \overline{\mathcal{A}(v_2, u)} \\ &= \overline{\beta}\ \mathcal{A}(u, v_1) + \mathcal{A}(u, v_2),\end{aligned}$$

es decir, $\mathcal{A}$ es una forma hermitiana si es lineal en la primera variable y conjugado lineal en la segunda variable. Además,

$$\mathcal{A}(v, v) \in \mathbb{R} \qquad \text{para todo } v \in V,$$

ya que $\mathcal{A}(v, v) = \overline{\mathcal{A}(v, v)}$.

Los resultados de la sección anterior para formas bilineales simétricas tienen sus análogos para formas hermitianas. Dada una forma hermitiana $\mathcal{A} : V \times V \to \mathbb{C}$, la función $Q : V \to \mathbb{R}$ definida por $Q(v) = \mathcal{A}(v, v)$ se denomina **forma cuadrática hermitiana** asociada a $\mathcal{A}$. Podemos reconstruir a $\mathcal{A}$ a partir de $Q$, mediante la identidad de polarización:

$$\mathcal{A}(u, v) = \frac{1}{4} \sum_{k=0}^{3} i^k Q(u + i^k v).$$

Si $\dim V = n$ y $\mathcal{B} = \{b_1, \ldots, b_n\}$ es una base de $V$, la matriz $H = (H_{ij}) \in \mathbb{C}^{n \times n}$ dada por

$$H_{ij} = \mathcal{A}(b_j, b_i),$$

es la **representación matricial de $\mathcal{A}$ con respecto a la base $\mathcal{B}$**, y verifica:

$$\mathcal{A}(u, v) = [v]_{\mathcal{B}}^* H [u]_{\mathcal{B}},$$

si $[u]_{\mathcal{B}}, [v]_{\mathcal{B}} \in \mathbb{C}^n$ son las coordenadas de los vectores $u, v \in V$.

Como $\mathcal{A}(b_i, b_j) = \overline{\mathcal{A}(b_j, b_i)}$, la matriz $H$ es autoadjunta y, en particular, las entradas de la diagonal de $H$ son reales. Por lo tanto, si $\mathcal{A}$ admite una representación matricial diagonal, ésta contiene sólo entradas reales.

**Teorema 9.35.** *Dado un $\mathbb{C}$-espacio vectorial $V$ con $\dim V = n$, sea $\mathcal{A}$ una forma hermitiana sobre $V$. Entonces, existe una base $\mathcal{B} = \{b_1, \ldots, b_n\}$ de $V$ tal que $[\mathcal{A}]_{\mathcal{B}}$ es una matriz diagonal, i.e. $\mathcal{A}(b_j, b_i) = 0$ si $i \neq j$. Más aún, toda representación matricial diagonal de $\mathcal{A}$ tiene la misma cantidad $p$ de entradas positivas y el mismo número $q$ de entradas negativas.*

*Demostración.* Ejercicio. $\square$

La suma $p + q$ coincide con el rango de la forma $\mathcal{A}$ y la diferencia $s := p - q$ se denomina la signatura de $\mathcal{A}$.

Las formas hermitianas también pueden clasificarse en formas (semi)definidas positivas, (semi)definidas negativas e indefinidas.

**Ejemplo 9.36.** El producto interno canónico de $\mathbb{C}^n$:

$$\langle (x_1, \ldots, x_n), (y_1, \ldots, y_n) \rangle = x_1 \overline{y_1} + \ldots + x_n \overline{y_n},$$

es una forma hermitiana sobre $\mathbb{C}^n$. Más aún, $\langle \cdot, \cdot \rangle$ es una forma definida positiva ya que

$$\langle (x_1, \ldots, x_n), (x_1, \ldots, x_n) \rangle = |x_1|^2 + \ldots + |x_n|^2 > 0 \qquad \text{si } (x_1, \ldots, x_n) \neq \vec{0}.$$

## 9.4. Ejercicios

1. Decir cuales de las siguientes aplicaciones $\mathcal{A} : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ son bilineales.

   *a*) $\mathcal{A}((x_1, y_1), (x_2, y_2)) = 1$.
   
   *b*) $\mathcal{A}((x_1, y_1), (x_2, y_2)) = (x_1 + y_1)^2 - (x_1 - y_1)^2$.
   
   *c*) $\mathcal{A}((x_1, y_1), (x_2, y_2)) = (x_1 - y_1)^2 + x_2 y_2$.
   
   *d*) $\mathcal{A}((x_1, y_1), (x_2, y_2)) = x_1 y_2 - x_2 y_1$.

2. Dada $A \in \mathbb{K}^{m\times m}$, sea $\mathcal{A} : \mathbb{K}^{m\times n} \times \mathbb{K}^{m\times n} \to \mathbb{K}$ definida por

$$\mathcal{A}(X, Y) = \mathrm{tr}(Y^t A X). \quad \text{para } X, Y \in \mathbb{K}^{m\times n}.$$

   Probar que $\mathcal{A}$ es una forma bilineal.

3. Dado un $\mathbb{K}$-espacio vectorial $V$, probar que $\mathrm{Bil}(V)$ es un $\mathbb{K}$-espacio vectorial.

4. Determinar si las siguientes formas bilineales son simétricas:

   *a*) $\mathcal{A} \in \mathrm{Bil}(\mathbb{R}^2)$ dada por $\mathcal{A}((x_1, y_1), (x_2, y_2)) = 3x_1x_2 + 2x_1y_2 + 2y_1x_2 + y_1y_2$.
   
   *b*) $\mathcal{A} \in \mathrm{Bil}(\mathbb{R}^2)$ dada por $\mathcal{A}((x_1, y_1), (x_2, y_2)) = x_1x_2 + 4x_1y_2 - 3y_1x_2$.
   
   *c*) $\mathcal{A} \in \mathrm{Bil}(\mathbb{R}^3)$ dada por $\mathcal{A}((x_1, y_1, z_1), (x_2, y_2, z_2)) = x_1x_2 + y_1y_2 + z_1z_2$.

5. Para las formas bilineales simétricas del punto anterior, escribir una representación matricial en una base del espacio vectorial involucrado diferente de la canónica. Hallar su rango.

6. Sea $\mathcal{A}$ la forma cuadrática sobre $\mathbb{R}^3$ dada por

$$\mathcal{A}(x, y, z) = x^2 - 4xz + 5y^2 + 4z^2.$$

   *a*) Determinar si $\mathcal{A}$ es una forma definida positiva, semidefinida positiva, definida negativa, semidefinida negativa o indefinida. Justificar.
   
   *b*) Encontrar una representación matricial de $\mathcal{A}$ que sea diagonal.

7. Dado un $\mathbb{R}$-espacio vectorial $V$ con $\dim V = n$, sea $\mathcal{A} \in \mathrm{Bil}(V)$ una forma simétrica con rango $r$ y signatura $s$. Probar que

   *a*) $\mathcal{A}$ es semidefinida positiva $\Leftrightarrow$ $s = r$;
   
   *b*) $\mathcal{A}$ es semidefinida negativa $\Leftrightarrow$ $s = -r$.

8. Dado un $\mathbb{K}$-espacio vectorial $V$ con $\dim V = n$, sea $\mathcal{B}$ una forma bilineal sobre $V$ con rango $r$. Consideremos los subespacios:

$$V^\perp = \{v \in V : \mathcal{B}(u, v) = 0 \text{ para todo } u \in V\}$$
$$\text{y} \quad V^\top = \{u \in V : \mathcal{B}(u, v) = 0 \text{ para todo } v \in V\}.$$

   Probar que $r = n - \dim V^\perp = n - \dim V^\top$, y en consecuencia $\dim V^\perp = \dim V^\top$.

   (Sugerencia: Recordar que el rango de $\mathcal{B}$ es el rango de su matriz asociada.)

9. Dado un $\mathbb{K}$-espacio vectorial $V$, sea $\mathcal{A} \in \mathrm{Bil}(V)$ simétrica. Probar que $\mathcal{A}$ es no degenerada si y sólo si $V^\perp = \{0\}$.

10. Sea $\mathcal{A} \in \mathrm{Bil}(\mathbb{R}^3)$ dada por

$$\mathcal{A}((x_1, y_1, z_1), (x_2, y_2, z_2)) = 3x_1x_2+4x_1y_2+x_1z_2+4y_1x_1+y_1y_2+6y_1z_2+z_1x_2+2z_1y_2+z_1z_2\,.$$

   a) ¿Es simétrica $\mathcal{A}$? ¿Es antisimétrica?
   b) Hallar $\mathcal{A}_{sim}$ y $\mathcal{A}_{ant}$ en $\mathrm{Bil}(\mathbb{R}^3)$ tales que $\mathcal{A} = \mathcal{A}_{sim} + \mathcal{A}_{ant}$.

11. Sea $B = \{u, v\}$ una base de $\mathbb{R}^2$ y sea $\mathcal{A} \in \mathrm{Bil}(\mathbb{R}^2)$ dada por

$$\mathcal{A}(u, u) = \mathcal{A}(v, v) = 0 \quad \mathcal{A}(u, v) = 1 = -\mathcal{A}(v, u)\,.$$

   Probar que $\mathcal{A}(x, x) = 0$ para todo $x \in V$ y que $\mathcal{A}$ es no degenerada.

12. Supongamos que $V = \mathbb{R}^{2\times 2}$. Dada $M = \begin{pmatrix} 1 & 2 \\ 3 & 5 \end{pmatrix}$, sea $\mathcal{B} : V \times V \to \mathbb{R}$ dada por

$$\mathcal{B}(A, B) = \mathrm{tr}\left(A^\top MB\right), \qquad A, B \in V.$$

   Mostrar que $\mathcal{B}$ es una forma bilineal sobre $V$ y encontrar la matriz de $\mathcal{B}$ con respecto a la base $\mathcal{E} = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$.

13. Sea $\mathcal{A} \in \mathrm{Bil}(\mathbb{R}^2)$ dada por

$$\mathcal{A}((x_1, y_1), (x_2, y_2)) = 2x_1x_2 + x_1y_2 + y_1x_2\,.$$

   Hallar su forma cuadrática asociada.

14. Para cada una de las siguientes formas cuadráticas sobre $\mathbb{R}^2$, hallar la forma bilineal simétrica asociada y representarla matricialmente en la base canónica.

   a) $Q(x, y) = \alpha x^2$ para $\alpha \in \mathbb{R}$.
   b) $Q(x, y) = \beta y^2$ para $\beta \in \mathbb{R}$.
   c) $Q(x, y) = \gamma xy$ para $\gamma \in \mathbb{R}$.
   d) $Q(x, y) = 2x^2 - \frac{1}{3}xy$.

15. Probar que:

   a) $Q(x, y) = x^2 - 4xy + 5y^2$ es una forma cuadrática definida positiva sobre $\mathbb{R}^2$.
   b) $Q(x, y, z) = 7x^2 + 4xy + y^2 - 8xz - 3z^2$ es una forma cuadrática indefinida sobre $\mathbb{R}^3$.
   c) $Q(x, y, z) = -x^2 - 2y^2 - 3z^2$ es una forma cuadrática definida negativa sobre $\mathbb{R}^3$.

16. Mostrar que la forma cuadrática en $\mathbb{R}^3$ asociada a la matriz

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

   es definida negativa.

17. Dado un $\mathbb{R}$-espacio con producto interno $(V, \langle\cdot,\cdot\rangle)$, sea $Q$ una forma cuadrática sobre $V$. Supongamos que $A$ es la matriz que la representa con respecto a una base $\mathcal{B}$ de $V$. Probar que si $\lambda$ es un autovalor de $A$ entonces existe un autovector $v$ asociado a $\lambda$ tal que
$$Q(v) = \lambda \|v\|^2 .$$
Deducir que:
    *a*) si $Q$ es definida positiva entonces todos los autovalores de $A$ son positivos;
    *b*) si $Q$ es definida negativa entonces todos los autovalores de $A$ son negativos;
    *c*) si $Q$ es indefinida entonces $A$ tiene autovalores positivos y autovalores negativos.

    Analizar si vale la recíproca en alguno de los casos anteriores.

18. Dado un $\mathbb{R}$-espacio con producto interno $(V, \langle\cdot,\cdot\rangle)$ con $\dim(V) = n$, sea $\mathcal{A} \in \mathrm{Bil}(V)$ una forma bilineal simétrica.
    *a*) Probar que existen escalares $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ y una base ortonormal $\mathcal{B} = \{v_1, \ldots, v_n\}$ de $V$ tales que, si $v = \sum_{i=1}^n \alpha_i\, v_i \in V$ entonces
    $$\mathcal{A}(v, v) = \sum_{i=1}^n \lambda_i\ \alpha_i^2 .$$
    *b*) Siguiendo con la notación del inciso anterior, mostrar que la forma cuadrática asociada a $\mathcal{A}$ es definida positiva si y sólo si $\lambda_i > 0$ para $1 \leq i \leq n$. Deducir que, en dicho caso, existe $\delta > 0$ tal que
    $$\mathcal{A}(v, v) \geq \delta \cdot \|v\|^2 \quad \text{para todo } v \in V.$$

# Capítulo 10

# Álgebra tensorial

El álgebra tensorial es más un lenguaje que una teoría en sí misma, pero nos permite presentar una descripción uniforme para todos los objetos tratados en estos apuntes, además de agruparlos en una estructura algebraica.

## 10.1. Producto tensorial de $\mathbb{K}$-espacios vectoriales

Comenzaremos generalizando la noción de forma bilineal tratada en el capítulo anterior. En particular, ahora las formas tomarán valores en un $\mathbb{K}$-espacio vectorial y no en el cuerpo $\mathbb{K}$.

**Definición 10.1.** Dados $\mathbb{K}$-espacios vectoriales $V$, $W$ y $U$, una función $\varphi : V \times W \to U$ es **bilineal** si es lineal en cada una de sus variables (cuando la otra permanece fija), i.e.

$$\begin{aligned} \varphi(\alpha v_1 + \beta v_2, w) &= \alpha\, \varphi(v_1, w) + \beta\, \varphi(v_2, w) \qquad \text{para todo } v_1, v_2 \in V,\ w \in W \text{ y } \alpha, \beta \in \mathbb{K}, \\ \varphi(v, \alpha w_1 + \beta w_2) &= \alpha\, \varphi(v, w_1) + \beta\, \varphi(v, w_2) \qquad \text{para todo } v \in V,\ w_1, w_2 \in W \text{ y } \alpha, \beta \in \mathbb{K}. \end{aligned}$$

Estas funciones forman un $\mathbb{K}$-espacio vectorial al que denotaremos $\operatorname{Bil}(V \times W, U)$. Si $V$, $W$ y $U$ son $\mathbb{K}$-espacios vectoriales de dimensión finita, entonces

$$\dim \operatorname{Bil}(V \times W, U) = \dim V \cdot \dim W \cdot \dim U.$$

El producto tensorial de dos $\mathbb{K}$-espacios vectoriales $V$ y $W$ aparece naturalmente al considerar funciones bilineales $\varphi : V \times W \to U$ en las cuales $U$ es un tercer $\mathbb{K}$-espacio vectorial. Veremos que uno de estos $U$ es "universal", en el sentido de que describe a todos los posibles $U$. Este $U$ "universal" será el producto tensorial entre $V$ y $W$.

**Proposición 10.2.** *Dados dos $\mathbb{K}$-espacios vectoriales $V$ y $W$ de dimensión finita, sean $\mathcal{B} = \{b_1, \ldots, b_n\}$ y $\mathcal{E} = \{e_1, \ldots, e_m\}$ bases de $V$ y $W$, respectivamente. Si $\varphi : V \times W \to U$ es bilineal, las siguientes condiciones son equivalentes:*

*i)* $\big\{\varphi(b_i, e_j) : \ i = 1, \ldots, n,\ j = 1, \ldots, m\big\}$ *es una base de $U$;*

*ii) para cada $z \in U$ existen únicos $y_1, \ldots, y_n \in W$ tales que* $z = \displaystyle\sum_{i=1}^{n} \varphi(b_i, y_i)$*;*

*iii) para cada $z \in U$ existen únicos $x_1, \ldots, x_m \in V$ tales que* $z = \displaystyle\sum_{j=1}^{m} \varphi(x_j, e_j)$.

*Demostración.* Supongamos que $\{\varphi(b_i, e_j) : \ i = 1, \dots, n, \ j = 1, \dots, m\}$ es una base de $U$. Dado $z \in U$, sean $\alpha_{ij} \in \mathbb{K}$ tales que

$$z = \sum_{i=1}^{n} \sum_{j=1}^{m} \alpha_{ij} \varphi(b_i, e_j).$$

Luego, si para $i = 1, \dots, n$ llamamos $y_i := \sum_{j=1}^{m} \alpha_{ij} e_j$, tenemos que

$$z = \sum_{i=1}^{n} \sum_{j=1}^{m} \alpha_{ij} \varphi(b_i, e_j) = \sum_{i=1}^{n} \varphi\left(b_i, \sum_{j=1}^{m} \alpha_{ij} e_j\right) = \sum_{i=1}^{n} \varphi(b_i, y_i),$$

y la unicidad de los $y_i$ es consecuencia de la unicidad de los escalares $\alpha_{ij}$.

Recíprocamente, supongamos que para cada $z \in U$ existen únicos $y_1, \dots, y_n \in W$ tales que $z = \sum_{i=1}^{n} \varphi(b_i, y_i)$. Como $\mathcal{E}$ es una base de $W$, para cada $i = 1, \dots, n$ existen únicos $\alpha_{i1}, \dots, \alpha_{im} \in \mathbb{K}$ tales que $y_i = \sum_{j=1}^{m} \alpha_{ij} e_j$. Entonces,

$$z = \sum_{i=1}^{n} \varphi(b_i, y_i) = \sum_{i=1}^{n} \varphi\left(b_i, \sum_{j=1}^{m} \alpha_{ij} e_j\right) = \sum_{i=1}^{n} \sum_{j=1}^{m} \alpha_{ij} \varphi\left(b_i, e_j\right).$$

Por lo tanto, $\{\varphi(b_i, e_j) : \ i = 1, \dots, n, \ j = 1, \dots, m\}$ es una base de $U$.

Esto completa la prueba de la equivalencia entre $i)$ y $ii)$. La equivalencia entre $i)$ y $iii)$ sale mutatis mutandis. □

**Corolario 10.3.** *Si la propiedad $i)$ vale para un par de bases $\mathcal{B}$ y $\mathcal{E}$ de $V$ y $W$, respectivamente, entonces vale para cualquier par de bases de $V$ y $W$.*

Dicho mal y pronto, un $\mathbb{K}$-espacio vectorial $T$ es un producto tensorial de $V$ y $W$ si es la imagen de una función bilineal definida sobre $V \times W$. Formalmente,

**Definición 10.4.** Dados dos $\mathbb{K}$-espacios vectoriales $V$ y $W$ de dimensión finita, un **producto tensorial** de $V$ y $W$ es un par $(T, \otimes)$ formado por un $\mathbb{K}$-espacio vectorial $T$ y una función bilineal

$$\otimes : V \times W \to T, \qquad (x, y) \mapsto x \otimes y,$$

que satisface la condición: si $\mathcal{B} = \{b_1, \dots, b_n\}$ y $\mathcal{E} = \{e_1, \dots, e_m\}$ son bases de $V$ y $W$, respectivamente, entonces

$$\{b_i \otimes e_j : \ i = 1, \dots, n, \ j = 1, \dots, m\} \text{ es una base de } T.$$

El corolario anterior garantiza que esta definición no depende de la elección de las bases de los $\mathbb{K}$-espacios vectoriales $V$ y $W$.

**Observación 10.5.** El producto tensorial de dos $\mathbb{K}$-espacios vectoriales $V$ y $W$ de dimensión finita es **único** en el siguiente sentido: si $(T_1, \otimes_1)$ y $(T_2, \otimes_2)$ son dos productos tensoriales de $V$ y $W$, existe un único isomorfismo $\psi : T_1 \to T_2$ tal que

$$\psi(x \otimes_1 y) = x \otimes_2 y,$$

para todo $x \in V$, $y \in W$. De hecho, si $\mathcal{B} = \{b_1, \dots, b_n\}$ y $\mathcal{E} = \{e_1, \dots, e_m\}$ son bases de $V$ y $W$, respectivamente, entonces $\psi$ puede definirse como la única transformación lineal tal que

$$\psi(b_i \otimes_1 e_j) = b_i \otimes_2 e_j, \qquad i = 1, \dots, n, \ j = 1, \dots, m.$$

Al producto tensorial de dos $\mathbb{K}$-espacios vectoriales $V$ y $W$ lo anotaremos $V \otimes W$. Si $\dim V = n$ y $\dim W = m$ entonces

$$\dim(V \otimes W) = nm = \dim V \cdot \dim W.$$

**Ejemplo 10.6.** Consideremos la función bilineal $\otimes : \mathbb{K}[x] \times \mathbb{K}[y] \to \mathbb{K}[x, y]$ dada por

$$(p \otimes q)(x, y) := p(x)q(y), \qquad p \in \mathbb{K}[x],\ q \in \mathbb{K}[y].$$

Los productos $x^k \otimes y^l = x^k y^l$ con $k, l \in \mathbb{N}$ forman una base de $\mathbb{K}[x, y]$, así que

$$\mathbb{K}[x, y] = \mathbb{K}[x] \otimes \mathbb{K}[y].$$

Análogamente,

$$\mathbb{K}[x_1, \ldots, x_n] \otimes \mathbb{K}[y_1, \ldots, y_m] = \mathbb{K}[x_1, \ldots, x_n, y_1, \ldots, y_m].$$

**Ejemplos 10.7.** Dados dos $\mathbb{K}$-espacios vectoriales $V$ y $W$, sean $\mathcal{B} = \{b_1, \ldots, b_n\}$ y $\mathcal{E} = \{e_1, \ldots, e_m\}$ bases de $V$ y $W$, respectivamente. Consideremos también como bases de $V^*$ y $W^*$, a las bases duales $\mathcal{B}^* = \{f_1, \ldots, f_n\}$ y $\mathcal{E}^* = \{g_1, \ldots, g_m\}$.

1. Para cada $f \in V^*$ y cada $y \in W$, definamos la transformación lineal

   $$f \otimes y : V \to W \quad \text{dada por} \quad (f \otimes y)(x) = f(x)y.$$

   Así, construímos una función bilineal

   $$\otimes : V^* \times W \to L(V, W).$$

   Es fácil ver que $[f_i \otimes e_j]_{\mathcal{E},\mathcal{B}} = E^{ji}$, la matriz en $\mathbb{K}^{m\times n}$ cuya única entrada no nula es un 1 en la fila $j$, columna $i$. Como $\{E^{ji} :\ j = 1, \ldots, m,\ i = 1, \ldots, n\}$ es una base de $\mathbb{K}^{m\times n}$, la familia de operadores $f_i \otimes e_j$ es una base de $L(V, W)$ y resulta que

   $$L(V, W) = V^* \otimes W.$$

2. Para cada par de funcionales $f, g \in V^*$, definimos la forma bilineal

   $$f \otimes g : V \times V \to \mathbb{K} \quad \text{dada por} \quad (f \otimes g)(u, v) = f(u)g(v).$$

   Tenemos entonces una función bilineal

   $$\otimes : V^* \times V^* \to \mathrm{Bil}(V).$$

   Como además $\{f_i \otimes f_j : i = 1, \ldots, n,\ j = 1, \ldots, n\}$ es una base de $\mathrm{Bil}(V)$, resulta que

   $$\mathrm{Bil}(V) = V^* \otimes V^*.$$

La propiedad de universalidad que caracteriza al producto tensorial de dos $\mathbb{K}$-espacios vectoriales puede enunciarse como:

**Proposición 10.8.** *Dados tres $\mathbb{K}$-espacios vectoriales $V$, $W$ y $U$, supongamos que $V$ y $W$ son de dimensión finita. Para cada función bilineal $\varphi : V \times W \to U$ existe una única transformación lineal $T \in L(V \otimes W, U)$ tal que*

$$\varphi(x, y) = T(x \otimes y),$$

*para cualesquiera $x \in V$, $y \in W$.*

Esta propiedad puede reinterpretarse diciendo que para cada bilineal $\varphi : V \times W \to U$ existe una única $T \in L(V \otimes W, U)$ tal que el siguiente diagrama conmuta:

$$\begin{array}{ccc} V \times W & \xrightarrow{\varphi} & U \\ \otimes \downarrow & \nearrow^{T} & \\ V \otimes W & & \end{array}$$

*Demostración.* Dadas $\mathcal{B} = \{b_1, \ldots, b_n\}$ y $\mathcal{E} = \{e_1, \ldots, e_m\}$ bases de $V$ y $W$, respectivamente, sabemos que $\mathcal{B} \otimes \mathcal{E} := \{b_i \otimes e_j : \ i = 1, \ldots, n, \ j = 1, \ldots, m\}$ es una base de $V \otimes W$. Luego, si para cada par $(i, j)$ definimos

$$T(b_i \otimes e_j) = \varphi(b_i, e_j), \qquad i = 1, \ldots, n, \ j = 1, \ldots, m, \tag{10.1}$$

tenemos una función bien definida de la base $\mathcal{B} \otimes \mathcal{E}$ en $U$. Entonces, sabemos que existe una única $T \in L(V \otimes W, U)$ que satisface (10.1), o equivalentemente,

$$\varphi(x, y) = T(x \otimes y), \qquad \text{para todo } x \in V, \ y \in W. \qquad \square$$

Dados dos $\mathbb{K}$-espacios vectoriales $V$ y $W$ de dimensión finita, supongamos que $\mathcal{B} = \{b_1, \ldots, b_n\}$ y $\mathcal{E} = \{e_1, \ldots, e_m\}$ son bases de $V$ y $W$, respectivamente. Todo vector $z \in V \otimes W$ se descompone de manera única como

$$z = \sum_{i=1}^{n} \sum_{j=1}^{m} \alpha_{ij} \ b_i \otimes e_j, \qquad \alpha_{ij} \in \mathbb{K}.$$

Los escalares $\alpha_{ij}$ son las **coordenadas de $z$ con respecto a la base**

$$\mathcal{B} \otimes \mathcal{E} := \big\{b_i \otimes e_j : \ i = 1, \ldots, n, \ j = 1, \ldots, m\big\}.$$

En particular, cada $z \in V \otimes W$ puede describirse por medio de su matriz de coordenadas

$$Z := (\alpha_{ij}) \in \mathbb{K}^{n \times m}.$$

**Definición 10.9.** Un elemento $z \in V \otimes W$ se dice **descomponible** si existe un par de vectores $x \in V$, $y \in W$ tales que $z = x \otimes y$.

Dado $z = x \otimes y$, es claro que si escribimos a $x$ e $y$ con respecto a las bases $\mathcal{B}$ y $\mathcal{E}$,

$$x = \sum_{i=1}^{n} \beta_i b_i, \quad y = \sum_{j=1}^{m} \gamma_j e_j,$$

entonces $z = \sum_{i=1}^{n} \sum_{j=1}^{m} \beta_i \gamma_j \ b_i \otimes e_j$, es decir,

$$(\alpha_{ij}) = Z = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} (\gamma_1, \gamma_2, \ldots, \gamma_m).$$

En particular, $Z$ es una matriz de rango menor o igual a 1.

Por lo tanto, los elementos descomponibles representan una porción muy pequeña de $V \otimes W$. Sin embargo, son suficientes para generar todo el espacio.

**Ejercicio 10.10.** Los vectores descomponibles no tienen una única representación posible. De hecho, si $z = x \otimes y$ y $\alpha \in \mathbb{K}$, $\alpha \neq 0$, entonces $z$ también puede representarse como

$$z = (\alpha x) \otimes (\alpha^{-1} y).$$

Probar que estas son todas las representaciones posibles de $z$, es decir, si $z = u \otimes v$ entonces existe $\lambda \in \mathbb{K}$, $\lambda \neq 0$, tal que

$$u = \lambda x, \quad v = \lambda^{-1} y.$$

La Proposición 10.2 sugiere otras descomposiciones posibles de un elemento $z \in V \otimes W$, las cuales pueden ser útiles en algunas situaciones. Si $\mathcal{B} = \{b_1, \ldots, b_n\}$ y $\mathcal{E} = \{e_1, \ldots, e_m\}$ son bases de $V$ y $W$, respectivamente, entonces existen únicos $x_1, \ldots, x_m \in V$ y también únicos $y_1, \ldots, y_n \in V$ tales que

$$z = \sum_{i=1}^{n} b_i \otimes y_i \quad \text{y} \quad z = \sum_{j=1}^{m} x_j \otimes e_j.$$

Veamos que el producto tensorial de $\mathbb{K}$-espacios vectoriales es, en cierto sentido, conmutativo y asociativo.

Recordemos que en la Observación 10.5 mencionamos que el producto tensorial es único, salvo isomorfismos. Notemos que si $V$ y $W$ son dos $\mathbb{K}$-espacios vectoriales, entonces hay un isomorfismo entre $V \otimes W$ y $W \otimes V$:

$$\psi : V \otimes W \to W \otimes V, \qquad x \otimes y \mapsto y \otimes x.$$

Como $\psi$ es lineal y mapea la base $\mathcal{B} \otimes \mathcal{E}$ de $V \otimes W$ en la base $\mathcal{E} \otimes \mathcal{B}$ de $W \otimes V$, el Teorema 2.22 asegura que $\psi$ es un isomorfismo.

Análogamente, si $U$ es un tercer $\mathbb{K}$-espacio vectorial, tenemos un isomorfismo $\eta : (U \otimes V) \otimes W \to U \otimes (V \otimes W)$ definido por

$$(u \otimes v) \otimes w \mapsto u \otimes (v \otimes w), \qquad u \in U, v \in V, w \in W.$$

Identificando los espacios $(U \otimes V) \otimes W$ y $U \otimes (V \otimes W)$, tenemos una buena definición para el producto tensorial $U \otimes V \otimes W$ de tres $\mathbb{K}$-espacios vectoriales. Análogamente, podemos escribir el producto tensorial de cualquier cantidad finita de $\mathbb{K}$-espacios vectoriales: si $V_1, V_2, \ldots, V_p$ son $\mathbb{K}$-espacios vectoriales, la asociatividad del producto tensorial nos permite escribir

$$V_1 \otimes V_2 \otimes \ldots \otimes V_p,$$

sin temor a malinterpretarlo. Los productos tensoriales de vectores básicos de $V_1, \ldots, V_p$ forman una base del espacio anterior.

### 10.1.1. Producto tensorial de operadores lineales

Dados dos $\mathbb{K}$-espacios vectoriales $V$ y $W$ de dimensión finita, sean $T \in L(V)$ y $S \in L(W)$. Podemos construir un operador sobre $V \otimes W$ definiendo:

$$(T \otimes S)(x \otimes y) := Tx \otimes Sy,$$

sobre los elementos descomponibles $x \otimes y \in V \otimes W$. Al operador $T \otimes S$ se lo conoce como el **producto tensorial de** $T$ **y** $S$.

Si $\mathcal{B} = \{b_1, \ldots, b_n\}$ y $\mathcal{E} = \{e_1, \ldots, e_m\}$ son bases de $V$ y $W$, respectivamente, sean $A = (a_{ij}) \in \mathbb{K}^{n\times n}$ y $B = (b_{kl}) \in \mathbb{K}^{m\times m}$ tales que $A = [T]_{\mathcal{B}}$ y $B = [S]_{\mathcal{E}}$. Luego, si

$$\begin{aligned}\mathcal{B}\otimes\mathcal{E} &= \{b_i \otimes e_j : \ i = 1, \ldots, n, \ j = 1, \ldots, m\}\\ &= \{b_1\otimes e_1, b_1\otimes e_2, \ldots, b_1\otimes e_m, b_2\otimes e_1, \ldots, b_2\otimes e_m, \ldots, b_n\otimes e_1, \ldots, b_n\otimes e_m\},\end{aligned}$$

tenemos que

$$A\otimes B := [T\otimes S]_{\mathcal{B}\otimes\mathcal{E}} = \begin{pmatrix} a_{11}B & a_{12}B & \ldots & a_{1n}B\\ a_{21}B & a_{22}B & \ldots & a_{2n}B\\ \vdots & \vdots & \ddots & \vdots\\ a_{n1}B & a_{n2}B & \ldots & a_{nn}B\end{pmatrix} \in \mathbb{K}^{nm\times nm},$$

es el **producto tensorial de las matrices** $A$ **y** $B$.

**Ejercicio 10.11.** Usando la matriz $A\otimes B$ definida más arriba, comprobar que:

$$\operatorname{tr}(T\otimes S) = (\operatorname{tr} T)(\operatorname{tr} S); \tag{10.2}$$

$$\det(T\otimes S) = (\det T)^m(\det S)^n. \tag{10.3}$$

**Ejercicio 10.12.** Supongamos que $\mathbb{K} = \mathbb{C}$. Si $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ son los $n$ autovalores de $T$ (contados con multiplicidades) y $\mu_1, \ldots, \mu_m \in \mathbb{C}$ son los $m$ autovalores de $S$, probar que los $nm$ autovalores de $T\otimes S$ están dados por

$$\eta_{ij} := \lambda_i\mu_j, \quad i = 1, \ldots, n, \ j = 1, \ldots, m.$$

Usando esta descripción de los autovalores de $T\otimes S$, deducir las fórmulas (10.2) y (10.3).

**Ejercicio 10.13.** Si $V$ y $W$ son dos $\mathbb{K}$-espacios vectoriales de dimensión finita, mostrar que $L(V\otimes W)$ es el producto tensorial de los espacios $L(V)$ y $L(W)$.

## 10.2. El álgebra tensorial de un $\mathbb{K}$-espacio vectorial

A lo largo de esta sección, supongamos que $V$ es un $\mathbb{K}$-espacio vectorial con $\dim V = n$.

**Definición 10.14.** Dados $p, q \in \mathbb{N}$, el $\mathbb{K}$-espacio vectorial

$$T^p_q(V) := \overbrace{V\otimes V\otimes\ldots\otimes V}^{p \text{ copias}} \otimes \overbrace{V^*\otimes V^*\otimes\ldots\otimes V^*}^{q \text{ copias}}, \tag{10.4}$$

es el espacio de tensores de tipo $(p, q)$ sobre $V$. Aceptaremos también la convención de que $T^0_0(V) = \mathbb{K}$.

Es fácil ver que $\dim T^p_q(V) = n^{p+q}$. Además,

$$\begin{array}{rcccl} T^1_0(V) &=& V & & \\ T^0_1(V) &=& V^* & & \\ T^1_1(V) &=& V\otimes V^* &=& L(V)\\ T^0_2(V) &=& V^*\otimes V^* &=& \operatorname{Bil}(V).\end{array}$$

Más generalmente,

$$\begin{array}{rcccl} T^0_q(V) &=& \overbrace{V^*\otimes V^*\otimes\ldots\otimes V^*}^{q \text{ copias}} &=& \operatorname{Mul}(\overbrace{V, \ldots, V}^{q \text{ copias}}, \mathbb{K}),\\ T^q_1(V) &=& V\otimes \overbrace{V^*\otimes V^*\otimes\ldots\otimes V^*}^{q \text{ copias}} &=& \operatorname{Mul}(\overbrace{V, \ldots, V}^{q \text{ copias}}, V),\end{array}$$

donde $\mathrm{Mul}(V_1, \ldots, V_q, U)$ denota al $\mathbb{K}$-espacio vectorial de funciones multilineales

$$\varphi : V_1 \times \ldots \times V_p \to U$$

es decir, lineales en cada una de las $q$ variables (cuando las restantes $q-1$ variables permanecen fijas).

### 10.2.1. Operaciones entre espacios de tensores

En este apartado presentaremos dos operaciones entre espacios de tensores, las cuales nos permitirán cierta interacción entre tensores de distinto tipo.

En primer lugar, podemos usar el producto tensorial para definir una operación binaria entre espacios de tensores:

$$\otimes : T_q^p(V) \times T_s^r(V) \to T_{q+s}^{p+r}(V),$$

la cual se define en los tensores descomponibles de cada espacio como

$$\begin{aligned}&(x_1 \otimes \ldots \otimes x_p \otimes f_1 \otimes \ldots \otimes f_q) \otimes (x_{p+1} \otimes \ldots \otimes x_{p+r} \otimes f_{q+1} \otimes \ldots \otimes f_{q+s})\\ &= x_1 \otimes \ldots \otimes x_p \otimes x_{p+1} \otimes \ldots \otimes x_{p+r} \otimes f_1 \otimes \ldots \otimes f_q \otimes f_{q+1} \otimes \ldots \otimes f_{q+s}.\end{aligned}$$

**Ejemplo 10.15.** Como vimos en el Ejercicio 10.13, el espacio

$$T_2^2(V) = V \otimes V \otimes V^* \otimes V^* = (V \otimes V) \otimes (V \otimes V)^*,$$

puede identificarse con el espacio $L(V \otimes V)$. Entonces, la "multiplicación tensorial"

$$\otimes : T_1^1(V) \times T_1^1(V) \to T_2^2(V),$$

coincide con el producto tensorial de operadores lineales descripto en la sección anterior.

De hecho, como ambas operaciones son bilineales, alcanza con probar que coinciden para vectores descomponibles (operadores lineales de rango 1). Dados $u, v \in V$ y $f, g \in V^*$, sean $T := u \otimes f \in L(V)$ y $S := v \otimes g \in L(V)$.

De acuerdo a lo visto en los Ejemplos 10.7, para cada par de vectores $x, y \in V$, la forma bilineal $f \otimes g$ actúa como

$$(f \otimes g)(x \otimes y) = f(x)g(y).$$

Entonces, dados $x, y \in V$,

$$\begin{aligned}(T \otimes S)(x \otimes y) &= Tx \otimes Sy = \big((u \otimes f)x\big) \otimes \big((v \otimes g)y\big)\\ &= \big(f(x)u\big) \otimes \big(g(y)v\big)\\ &= f(x)g(y)\ u \otimes v = \big((f \otimes g)(x \otimes y)\big)(u \otimes v)\\ &= \big((u \otimes v) \otimes (f \otimes g)\big)(x \otimes y).\end{aligned}$$

Por lo tanto, $T \otimes S = (u \otimes v) \otimes (f \otimes g)$.

Otra operación importante sobre los tensores es la **contracción**. Dados $p, q \in \mathbb{N}$, construiremos una transformación lineal

$$C_1^1 : T_q^p(V) \to T_{q-1}^{p-1}(V),$$

de la siguiente manera: consideremos la función multilineal

$$\varphi_1^1 : \overbrace{V \times V \times \ldots \times V}^{p \text{ copias}} \times \overbrace{V^* \times V^* \times \ldots \times V^*}^{q \text{ copias}} \to T_{q-1}^{p-1}(V),$$

dada por $\varphi_1^1(x_1,\ldots,x_p,f_1,\ldots,f_q) = f_1(x_1)(x_2\otimes\ldots\otimes x_p\otimes f_2\otimes\ldots\otimes f_q)$. Luego, por la Proposición 10.8, existe una única transformación lineal $C_1^1 : T_q^p(V)\to T_{q-1}^{p-1}(V)$ que verifica:

$$C_1^1(x_1\otimes x_2\otimes\ldots\otimes x_p\otimes f_1\otimes f_2\otimes\ldots\otimes f_q) = f_1(x_1)(x_2\otimes\ldots\otimes x_p\otimes f_2\otimes\ldots\otimes f_q).$$

Esta es justamente la contracción de las primeras componentes de los espacios

$$\overbrace{V\otimes V\otimes\ldots\otimes V}^{p\text{ copias}} \quad \text{y} \quad \overbrace{V^*\otimes V^*\otimes\ldots\otimes V^*}^{q\text{ copias}},$$

cuyo producto tensorial es $T_q^p(V)$.

De la misma manera, dados dos índices cualesquiera $i=1,\ldots,p$, $j=1,\ldots,q$, podemos definir $C_j^i : T_q^p(V)\to T_{q-1}^{p-1}(V)$ como la contracción de la $i$-ésima coordenada de $\overbrace{V\otimes V\otimes\ldots\otimes V}^{p\text{ copias}}$ con la $j$-ésima coordenada de $\overbrace{V^*\otimes V^*\otimes\ldots\otimes V^*}^{q\text{ copias}}$.

**Ejemplo 10.16.** *La contracción de un operador lineal es su traza.*

Recordemos que los operadores lineales son tensores de tipo $(1,1)$. Supongamos que $T\in L(V)$ es un operador de rango 1 de la forma

$$T = u\otimes f, \qquad \text{con } u\in V, f\in V^*,$$

es decir, un tensor descomponible. Supongamos además que ni $u$ ni $f$ son nulos. Como $\dim N(f) = n-1$, podemos considerar una base $\mathcal{B}=\{b_1,\ldots,b_n\}$ de $V$ tal que $\{b_1,\ldots,b_{n-1}\}$ sea una base de $N(f)$ y $f(b_n)\neq 0$. Luego, si $u$ se descompone en esta base como

$$u = \alpha_1 b_1 + \alpha_2 b_2 + \ldots + \alpha_n b_n,$$

tenemos que

$$\begin{aligned} T(b_i) &= \vec{0} && \text{si } i=1,\ldots,n-1,\\ T(b_n) &= (u\otimes f)(b_n) = f(b_n)u. \end{aligned}$$

Entonces,

$$[T]_{\mathcal{B}} = \begin{pmatrix} 0 & \ldots & 0 & f(b_n)\alpha_1\\ 0 & \ldots & 0 & f(b_n)\alpha_2\\ \vdots & \ddots & \vdots & \vdots\\ 0 & \ldots & 0 & f(b_n)\alpha_n \end{pmatrix},$$

y en particular,

$$\begin{aligned} \operatorname{tr}(T) &= f(b_n)\alpha_n = f(\alpha_n b_n) = f(\alpha_1 b_1 + \ldots + \alpha_{n-1}b_{n-1} + \alpha_n b_n) = f(u)\\ &= C_1^1(u\otimes f) = C_1^1(T). \end{aligned}$$

Si $T\in L(V)$ es un operador arbitrario, el resultado sale descomponiéndolo en suma de operadores de rango 1 y utilizando la linealidad (tanto de la traza como de la contracción).

**Ejemplo 10.17.** *La evaluación de un operador lineal en un vector, también es una contracción.*

Consideremos los tensores de tipo $(2,1)$, los cuales pueden identificarse con el producto tensorial de $V$ y $L(V)$:

$$T_1^2(V) = V\otimes V\otimes V^* = V\otimes(V\otimes V^*) = V\otimes L(V).$$

Luego, si $u, v \in V$, $f \in V^*$, notemos que

$$C_1^1\big(u \otimes (v \otimes f)\big) = f(u)v = (v \otimes f)(u).$$

Una vez más, usando la linealidad de $C_1^1$ y del producto tensorial,

$$C_1^1(u \otimes T) = Tu, \quad \text{si } u \in V \text{ y } T \in L(V) = V \otimes V^*.$$

**Ejemplo 10.18.** *El producto de operadores lineales, también es una contracción.*

Consideremos los tensores de tipo $(2, 2)$, los cuales pueden identificarse con el producto tensorial $L(V) \otimes L(V)$:

$$T_2^2(V) = V \otimes V \otimes V^* \otimes V^* = (V \otimes V^*) \otimes (V \otimes V^*) = L(V) \otimes L(V).$$

Luego, la contracción

$$C_1^2 : T_2^2(V) = L(V) \otimes L(V) \to T_1^1(V) = L(V),$$

representa al producto de operadores lineales. De hecho, si $T = u \otimes f$ y $S = v \otimes g$ son tensores descomponibles,

$$C_1^2(u \otimes v \otimes f \otimes g) = f(v) \cdot u \otimes g.$$

Por otra parte, dado $x \in V$,

$$\begin{aligned} TSx &= (u \otimes f)(v \otimes g)x = (u \otimes f)\big(g(x)v\big) = g(x)\ (u \otimes f)v = g(x)f(v)u \\ &= f(v)\big((u \otimes g)x\big) = \big(f(v)(u \otimes g)\big)x = C_1^2(u \otimes v \otimes f \otimes g)(x). \end{aligned}$$

Entonces, $TS = C_1^2(u \otimes v \otimes f \otimes g) = C_1^2(T \otimes S)$. Otra vez, el resto es consecuencia de la linealidad.

A la contracción del producto tensorial de los tensores $T$ y $S$ se la suele llamar la **contracción de $T$ con $S$**.

## 10.3. Coordenadas con respecto a una base

Fijada una base $\mathcal{B} = \{b_1, \ldots, b_n\}$ de $V$, sea $\mathcal{B}^* = \{f_1, \ldots, f_n\}$ la base dual correspondiente. Luego,

$$\big\{b_{i_1} \otimes \ldots \otimes b_{i_p} \otimes f_{j_1} \otimes \ldots \otimes f_{j_q} : \ 1 \leq i_1, \ldots, i_p, j_1, \ldots, j_q \leq n\big\},$$

es una base de $T_q^p(V)$, es decir, todo tensor $T \in T_q^p(V)$ puede escribirse de manera única como

$$T = \sum_{i_1,\ldots,i_p,j_1,\ldots,j_q} T_{i_1,\ldots,i_p,j_1,\ldots,j_q} b_{i_1} \otimes \ldots \otimes b_{i_p} \otimes f_{j_1} \otimes \ldots \otimes f_{j_q}, \tag{10.5}$$

para cierta familia de coeficientes $T_{i_1,\ldots,i_p,j_1,\ldots,j_q} \in \mathbb{K}$, a la que llamaremos las **coordenadas de** $T$ con respecto a la base $\mathcal{B}$ de $V$. Notemos que hemos escrito una única sumatoria con varios subíndices, lo que indica que debe sumarse sobre todos ellos (en este caso, de 1 a $n$).

**Ejemplo 10.19.** Las coordenadas de $T \in L(V)$, visto como un tensor de tipo $(1,1)$, son exactamente las entradas de $[T]_{\mathcal{B}}$. De hecho, si $T = \sum_{i,j} T_{ij}\ b_i \otimes f_j$ entonces

$$Tb_k = \left(\sum_{i,j} T_{ij}\ b_i \otimes f_j\right) b_k = \sum_{i,j} T_{ij}(b_i \otimes f_j)(b_k) = \sum_{i,j} T_{ij} f_j(b_k) b_i = \sum_{i} T_{ik} b_i,$$

con lo cual las coordenadas de la $k$-ésima columna de $[T]_{\mathcal{B}}$ son $(T_{1k}, T_{2k}, \ldots, T_{nk})^\top$.

De la misma manera, si $\mathcal{A} \in \mathrm{Bil}(V)$ las coordenadas de $\mathcal{A}$, vista como un tensor de tipo $(0,2)$, son las entradas de la matriz $[\mathcal{A}]_{\mathcal{B}}$.

El físico alemán Albert Einstein (1879–1955) propuso una notación alternativa para los tensores de tipo $(p,q)$ y sus coordenadas, en la cual se usan tanto subíndices como supraíndices. En ésta los vectores de la base de $V$ se indican con subíndices, y los de la base dual se indican con supraíndices. Por otra parte, los índices de las coordenadas de un tensor de tipo $(p,q)$ correspondientes a los vectores de la base de $V$ se indican como supraíndices, mientras que los índices de las coordenadas correspondientes a los vectores de la base dual se indican como subíndices. Además, si en una expresión aparece un índice repetido, una vez como subíndice y la otra como supraíndice (otras repeticiones no están permitidas), entonces supondremos que estamos sumando sobre ese índice sin escribir la sumatoria correspondiente. Por lo tanto, de acuerdo a la notación de Einstein, la fórmula (10.5) se escribe como:

$$T = T^{i_1,\ldots,i_p}_{j_1,\ldots,j_q}\ b_{i_1} \otimes \ldots \otimes b_{i_p} \otimes f^{j_1} \otimes \ldots \otimes f^{j_q}. \tag{10.6}$$

A continuación utilizaremos la notación de Einstein en algunos ejemplos, considerando siempre la base $\mathcal{B}$ de $V$, y su base dual como base de $V^*$.

**Ejemplo 10.20.** Dados $x \in V$ y $T \in L(V)$, supongamos que

$$x = \sum_{j=1}^{n} x_j b_j \qquad \text{y} \qquad T = \sum_{i=1}^{n}\sum_{j=1}^{n} A_{ij} b_i \otimes f_j,$$

para ciertos escalares $x_j \in \mathbb{K}$ y $A_{ij} \in \mathbb{K}$. Luego, en la notación de Einstein,

$$x = x^j b_j \qquad \text{y} \qquad T = A^i_j b_i \otimes f^j,$$

y sus coordenadas con respecto a la base $\mathcal{B}$ son

$$[x]_{\mathcal{B}} = (x^j) \qquad \text{y} \qquad [T]_{\mathcal{B}} = (A^i_j).$$

Luego, $[Tx]_{\mathcal{B}} = (A^i_j x^j)$ ya que

$$Tx = \left(A^i_j b_i \otimes f^j\right)\left(x^k b_k\right) = A^i_j x^k \left((b_i \otimes f^j) b_k\right) = A^i_j x^k f^j(b_k) b_i = (A^i_j x^j) b_i.$$

**Ejemplo 10.21.** Sean $S, T \in L(V)$ tales que $[T]_{\mathcal{B}} = (A^i_j)$ y $[S]_{\mathcal{B}} = (B^i_j)$. Entonces,

$$[TS]_{\mathcal{B}} = (A^i_k B^k_j).$$

De hecho, tenemos que

$$\begin{aligned} TS &= \left(A^i_j b_i \otimes f^j\right)\left(B^k_l b_k \otimes f^l\right) = A^i_j B^k_l (b_i \otimes f^j)(b_k \otimes f^l) = A^i_j B^k_l f^j(b_k)(b_i \otimes f^l) \\ &= A^i_k B^k_l (b_i \otimes f^l). \end{aligned}$$

Hemos utilizado el hecho de que $(b_i \otimes f^j)(b_k \otimes f^l) = f^j(b_k)(b_i \otimes f^l)$, el cual es fácilmente verificable (ya sea evaluando en un vector arbitrario $v \in V$, o notando que las matrices asociadas a $b_i \otimes f^j$, $b_k \otimes f^l$ y $b_i \otimes f^l$ son $E^{ji}$, $E^{lk}$ y $E^{li}$, respectivamente).

### 10.3.1. Coordenadas de operaciones entre tensores

Dados $T \in T^p_q(V)$ y $S \in T^r_s(V)$, supongamos que

$$T = T^{i_1,\ldots,i_p}_{j_1,\ldots,j_q}\; b_{i_1} \otimes \ldots \otimes b_{i_p} \otimes f^{j_1} \otimes \ldots \otimes f^{j_q}, \tag{10.7}$$

$$S = S^{i_{p+1},\ldots,i_{p+r}}_{j_{q+1},\ldots,j_{q+s}}\; b_{i_{p+1}} \otimes \ldots \otimes b_{i_{p+r}} \otimes f^{j_{q+1}} \otimes \ldots \otimes f^{j_{q+s}}. \tag{10.8}$$

Luego, las coordenadas de $T \otimes S$ son los productos

$$(T \otimes S)^{i_1,\ldots,i_{p+r}}_{j_1,\ldots,j_{q+s}} := T^{i_1,\ldots,i_p}_{j_1,\ldots,j_q}\; S^{i_{p+1},\ldots,i_{p+r}}_{j_{q+1},\ldots,j_{q+s}}$$

Por otra parte, si $T \in T^p_q(V)$ y $S = C^1_1(T) \in T^{p-1}_{q-1}(V)$ entonces,

$$S^{i_1,\ldots,i_{p-1}}_{j_1,\ldots,j_{q-1}} = T^{r,i_1,\ldots,i_{p-1}}_{r,j_1,\ldots,j_{q-1}},$$

donde $r$ es un índice adicional.

Esto es consecuencia de la fórmula (10.6), considerando los tensores de la base de $T^p_q(V)$ inducida por la base $\mathcal{B}$ de $V$, tenemos que

$$C^1_1\big(b_{i_1} \otimes \ldots \otimes b_{i_p} \otimes f^{j_1} \otimes \ldots \otimes f^{j_q}\big) = \delta_{i_1,j_1} \cdot b_{i_2} \otimes \ldots \otimes b_{i_p} \otimes f^{j_2} \otimes \ldots \otimes f^{j_q}.$$

De la misma manera, si $1 \leq k \leq p$ y $1 \leq l \leq q$ podemos calcular la contracción por los índices $k$ y $l$. Como

$$C^k_l\big(b_{i_1} \otimes \ldots \otimes b_{i_p} \otimes f^{j_1} \otimes \ldots \otimes f^{j_q}\big) = \delta_{i_k,j_l} \cdot b_{i_1} \otimes \ldots \otimes b_{i_{k-1}} \otimes b_{i_{k+1}} \ldots \otimes b_{i_p} \otimes f^{j_1} \otimes \ldots \otimes f^{j_{l-1}} \otimes f^{j_{l+1}} \otimes \ldots \otimes f^{j_q},$$

si $S = C^k_l(T)$ entonces sus coordenadas son:

$$S^{i_1,\ldots,i_{p-1}}_{j_1,\ldots,j_{q-1}} = T^{i_1,\ldots,i_{k-1},r,i_k,\ldots,i_{p-1}}_{j_1,\ldots,j_{l-1},r,j_l,\ldots,j_{q-1}}.$$

donde $r$ es un índice adicional.

## 10.4. Cambio de coordenadas en $T^p_q(V)$

Supongamos que $\mathcal{B} = \{b_1, \ldots, b_n\}$ y $\mathcal{E} = \{e_1, \ldots, e_n\}$ son dos bases de $V$. Dado $T \in T^p_q(V)$, supongamos que conocemos sus coordenadas con respecto a la base $\mathcal{B}$, $[T]_{\mathcal{B}} = (T^{i_1,\ldots,i_p}_{j_1,\ldots,j_q})$, y queremos calcular sus coordenadas con respecto a la base $\mathcal{E}$.

Si $P_{\mathcal{B},\mathcal{E}} = (\alpha^i_j)$ es la matriz cambio de base de $\mathcal{E}$ en $\mathcal{B}$, tenemos que

$$e_j = \alpha^i_j b_i, \qquad j = 1, \ldots, n.$$

De la misma manera, si $P_{\mathcal{E},\mathcal{B}} = (\beta^i_j)$ es la matriz cambio de base de $\mathcal{B}$ en $\mathcal{E}$, tenemos que

$$b_j = \beta^i_j e_i, \qquad j = 1, \ldots, n.$$

Llamemos $\mathcal{B}^* = \{f^1, \ldots, f^n\}$ y $\mathcal{E}^* = \{g^1, \ldots, g^n\}$ a las bases duales de $\mathcal{B}$ y $\mathcal{E}$, respectivamente. Si $P_{\mathcal{B}^*,\mathcal{E}^*} = (\gamma^i_j)$ entonces,

$$g^i = \gamma^i_k f^k, \qquad i = 1, \ldots, n.$$

Luego, como $g^i(e_j) = \delta_{ij}$ y $f^k(b_l) = \delta_{kl}$, tenemos que

$$\delta_{ij} = g^i(e_j) = (\gamma^i_k f^k)e_j = \gamma^i_k f^k(e_j) = \gamma^i_k f^k(\alpha^l_j b_l) = \gamma^i_k \alpha^l_j f^k(b_l) = \gamma^i_k \alpha^l_j \delta_{kl}.$$

Entonces, si $k = l$ resulta que $\delta_{ij} = \gamma^i_k \alpha^k_j$, es decir

$$P_{\mathcal{B}^*,\mathcal{E}^*} P_{\mathcal{B},\mathcal{E}} = I.$$

Con lo cual, $P_{\mathcal{E}^*,\mathcal{B}^*} = P_{\mathcal{E}^*,\mathcal{B}^*}(P_{\mathcal{B}^*,\mathcal{E}^*} P_{\mathcal{B},\mathcal{E}}) = (P_{\mathcal{E}^*,\mathcal{B}^*} P_{\mathcal{B}^*,\mathcal{E}^*}) P_{\mathcal{B},\mathcal{E}} = P_{\mathcal{B},\mathcal{E}}$. Por lo tanto,

$$f^j = \alpha^j_l g^l, \qquad j = 1, \ldots, n.$$

Utilizando toda la información recopilada, notemos que

$$\begin{aligned} T &= T^{i_1,\ldots,i_p}_{j_1,\ldots,j_q}\, b_{i_1} \otimes \ldots \otimes b_{i_p} \otimes f^{j_1} \otimes \ldots \otimes f^{j_q} \\ &= T^{i_1,\ldots,i_p}_{j_1,\ldots,j_q}\, \beta^{k_1}_{i_1} e_1 \otimes \ldots \otimes \beta^{k_p}_{i_p} e_p \otimes \alpha^{j_1}_{l_1} g^{l_1} \otimes \ldots \otimes \alpha^{j_q}_{l_q} g^{l_q} \\ &= \beta^{k_1}_{i_1} \ldots \beta^{k_p}_{i_p}\, T^{i_1,\ldots,i_p}_{j_1,\ldots,j_q}\, \alpha^{j_1}_{l_1} \ldots \alpha^{j_q}_{l_q}\, e_1 \otimes \ldots \otimes e_p \otimes g^{l_1} \otimes \ldots \otimes g^{l_q}. \end{aligned}$$

Por lo tanto, $[T]_{\mathcal{E}} = (\widetilde{T}\,^{k_1,\ldots,k_p}_{l_1,\ldots,l_q})$ están dadas por

$$\widetilde{T}\,^{k_1,\ldots,k_p}_{l_1,\ldots,l_q} = \beta^{k_1}_{i_1} \ldots \beta^{k_p}_{i_p}\, T^{i_1,\ldots,i_p}_{j_1,\ldots,j_q}\, \alpha^{j_1}_{l_1} \ldots \alpha^{j_q}_{l_q}.$$

## 10.5. Tensores sobre un espacio euclidiano

En cada espacio euclidiano hay un tensor especial que determina el producto interno.

Sea $(V, \langle \cdot, \cdot \rangle)$ un espacio euclidiano con $\dim V = n$. Como, en este caso, el producto interno es una forma bilineal simétrica sobre $V$, denotaremos $\mathfrak{g} \in T^0_2(V)$ al tensor que la representa. Lo llamaremos el **tensor métrico** de $V$.

Dada una base $\mathcal{B} = \{b_1, \ldots, b_n\}$ de $V$, supongamos que

$$[\mathfrak{g}]_{\mathcal{B}} = (g_{ij}).$$

Luego, dados dos vectores $x = x^i b_i$ e $y = y^i b_i$ en $V$, tenemos que

$$\langle x, y \rangle = x^i y^j g_{ij}.$$

Dado un vector $y \in V = T^1_0(V)$, la contracción del producto tensorial $\mathfrak{g} \otimes y$ (en cualquier índice de $\mathfrak{g}$ y el único de $y$) es un tensor de tipo $(0, 1)$, es decir, un funcional con coordenadas

$$y_i = g_{ij} y^j, \qquad i = 1, \ldots, n.$$

Recordemos que, por el Teorema de representación de Riesz, cada vector $y \in V$ determina un único funcional lineal $\varphi_y = \langle \cdot, y \rangle$ sobre $V$ (y viceversa). Entonces, a las coordenadas $(y^i)$ las llamaremos las **coordenadas contravariantes** de $y$, mientras que a las coordenadas $(y_i)$ las llamaremos las **coordenadas covariantes** de $y$.

Notemos que

$$y_i = g_{ij} y^j = \sum_{j=1}^n \langle b_i, b_j \rangle\, y^j = \left\langle b_i, \sum_{j=1}^n y^j b_j \right\rangle = \langle b_i, y \rangle\,.$$

Así que, en el caso particular de una base ortonormal $\mathcal{E} = \{e_1, \ldots, e_n\}$ de $V$, resulta que

$$y_i = y^i, \qquad i = 1, \ldots, n,$$

es decir, las coordenadas covariantes y contravariantes coinciden.

Más en general, la contracción del producto tensorial de $\mathfrak{g}$ con cualquier tensor $T \in T^p_q(V)$ (en el primer supraíndice de $T$ y cualquiera de los subíndices de $\mathfrak{g}$) es el tensor $\widetilde{T} \in T^{p-1}_{q+1}(V)$ con coordenadas

$$\widetilde{T}\,^{i_2,\ldots,i_p}_{j,j_1,\ldots,j_q} = g_{jk}\ T^{k,i_2,\ldots,i_p}_{j_1,\ldots,j_q}.$$

Esta operación se conoce como el **descenso** del primer supraíndice de $T$. De manera análoga podemos definir el descenso de cualquiera de los supraíndices de $T$.

Si $\mathcal{E} = \{e_1, \ldots, e_n\}$ es una base ortonormal de $V$, entonces

$$g_{jk} = \delta_{jk},$$

y en consecuencia tenemos que

$$\widetilde{T}\,^{i_2,\ldots,i_p}_{j,j_1,\ldots,j_q} = \delta_{jk}\ T^{k,i_2,\ldots,i_p}_{j_1,\ldots,j_q} = T^{j,i_2,\ldots,i_p}_{j_1,\ldots,j_q}.$$

Esto implica, en primer lugar, que el descenso de un supraíndice es una operación invertible. A su inversa, la llamaremos el **ascenso** de un subíndice de $T$. En segundo término, si nos restringimos a bases ortonormales, tenemos que no hay diferencias entre subíndices y supraíndices de tensores en un espacio euclidiano.

**Ejemplo 10.22.** Al descender el índice de un operador lineal $T \in L(V) = T^1_1(V)$, obtenemos la forma bilineal $\mathcal{A} \in \mathrm{Bil}(V)$ dada por

$$\mathcal{A}(x, y) = g_{jk} x^j T^k_l y^l = \langle x, Ty \rangle\,.$$

Esto establece un isomorfismo canónico entre el espacio de operadores lineales $L(V)$ y el espacio de formas bilineales $\mathrm{Bil}(V)$ sobre un espacio euclidiano.

## 10.6. Tensores covariantes y contravariantes

Los tensores de tipo $(p, 0)$ se denominan **tensores contravariantes de rango** $p$. A partir de aquí denotaremos

$$T^p(V) = T^p_0(V).$$

Los espacios $T^0(V) = \mathbb{K}$, $T^1(V) = V$, $T^2(V)$, ... pueden organizarse en un álgebra lineal, pero para ello necesitamos introducir la idea de "suma directa externa" de espacios vectoriales. El enfoque de este concepto no supone de antemano que los $\mathbb{K}$-espacios vectoriales $V_1, \ldots, V_k$ sean subespacios de un espacio en común.

**Definición 10.23.** Dada una familia finita $V_1, \ldots, V_k$ de $\mathbb{K}$-espacios vectoriales, la **suma directa externa** de ellos es el $\mathbb{K}$-espacio vectorial $V_1 \hat{\oplus} \ldots \hat{\oplus} V_k$ formado por las $k$-uplas

$$(x_1, \ldots, x_k), \quad \text{con } x_i \in V_i,\ i = 1, \ldots, k,$$

dotado con las operaciones de suma y producto por escalares definidas componente a componente, es decir,

$$(x_1, \ldots, x_k) + (y_1, \ldots, y_k) = (x_1 + y_1, \ldots, x_k + y_k),$$
$$\lambda(x_1, \ldots, x_k) = (\lambda x_1, \ldots, \lambda x_k).$$

Esta noción está íntimamente ligada al concepto de suma directa de subespacios que estudiamos previamente. De hecho, si consideramos los vectores de la forma

$$(0, \ldots, 0, \overbrace{\ x\ }^{\text{coord. i}}, 0, \ldots, 0), \quad x \in V_i,$$

éstos forman un subespacio de $V_1\hat{\oplus}\ldots\hat{\oplus}V_k$ que es isomorfo a $V_i$, llamémoslo $\hat{V}_i$. Más aún, cada elemento de $V_1\hat{\oplus}\ldots\hat{\oplus}V_k$ se descompone de manera única como una suma de vectores de estos subespacios, es decir,

$$V_1\hat{\oplus}\ldots\hat{\oplus}V_k = \hat{V}_1 \dot{+} \ldots \dot{+} \hat{V}_k.$$

Recíprocamente, si $V$ es un $\mathbb{K}$-espacio vectorial que se descompone como la suma directa de $k$ subespacios $W_1, \ldots, W_k$, entonces la aplicación

$$W_1\hat{\oplus}\ldots\hat{\oplus}W_k \to V, \qquad (x_1, \ldots, x_k) \mapsto x_1 + \ldots + x_k,$$

es un isomorfismo de $\mathbb{K}$-espacios vectoriales.

La discusión anterior se puede generalizar a una familia infinita de $\mathbb{K}$-espacios vectoriales $(V_j)_{j\in\mathbb{N}}$ siempre y cuando consideremos sólo sucesiones $(x_1, x_2, \ldots)$, $x_i \in V_i$, con una cantidad **finita** de términos no nulos.

Ahora podemos describir la construcción del álgebra tensorial de un $\mathbb{K}$-espacio vectorial $V$. Consideremos la suma directa externa infinita

$$T(V) := \widehat{\bigoplus}_{p\geq 0} T^p(V). \tag{10.9}$$

Como

$$T^p(V) \otimes T^q(V) \subset T^{p+q}(V),$$

$T(V)$ dotado con el producto tensorial resulta un álgebra lineal, llamada el **álgebra tensorial de** $V$. Ésta es un álgebra asociativa y tiene una unidad, que es el 1 del cuerpo $\mathbb{K} = T^0(V)$.

De manera análoga, los tensores de tipo $(0, q)$ se denominan **tensores covariantes de rango** $q$. Si denotamos $T_q(V) := T^0_q(V)$, el álgebra

$$T_*(V) = \widehat{\bigoplus}_{q\geq 0} T_q(V),$$

se llama el **álgebra de funciones multilineales sobre** $V$. Además, como

$$T_q(V) = T^q(V^*),$$

el álgebra de tensores covariantes puede interpretarse como el álgebra tensorial del espacio dual $V^*$.

## 10.7. Ejercicios

1. Sean $\mathbb{K}$ un cuerpo y $\varphi : \mathbb{K}^3 \times \mathbb{K}^4 \to \mathbb{K}^{3\times 4}$ la función bilineal dada por

$$(\varphi(x, y))_{ij} = x_i\, y_j \quad \text{para } x = (x_1, x_2, x_3) \in \mathbb{K}^3,\ y = (y_1, y_2, y_3, y_4) \in \mathbb{K}^4.$$

   Probar que $(\mathbb{K}^{3\times 4}, \varphi)$ es un producto tensorial.

2. Sean $V$ y $W$ dos $\mathbb{K}$-espacios vectoriales de dimensión finita. Probar que todo elemento $z \in V \otimes W$ se puede escribir como

$$z = \sum_{k=1}^{r} v_k \otimes w_k\,,$$

   para $v_1, \ldots, v_r \in V$ y $w_1, \ldots, w_r \in W$ vectores linealmente independientes de $V$ y $W$, respectivamente.

3. Dados dos $\mathbb{K}$-espacios vectoriales de dimensión finita $V$ y $W$, sea $z = x \otimes y \in V \otimes W$ un vector descomponible. Probar que para todo $\lambda \in \mathbb{K}$ (no nulo) $z = u \otimes v$, con
$$u = \lambda x \qquad \text{y} \qquad v = \lambda^{-1} y\,.$$
Mostrar además que éstas son todas las representaciones posibles de $z$, es decir, si $z = u \otimes v$ con $u \in V$ y $v \in W$ entonces existe $\lambda \in \mathbb{K}$, $\lambda \neq 0$, tal que
$$u = \lambda x, \quad v = \lambda^{-1} y.$$

4. Dar ejemplos de vectores no descomponibles en diferentes productos tensoriales de $\mathbb{K}$-espacios vectoriales.

5. Sean $V$ y $W$ dos $\mathbb{K}$-espacios vectoriales con $\dim V = n$ y $\dim W = m$, respectivamente. Consideremos bases $\mathcal{B}_V = \{v_1, \ldots, v_n\}$ y $\mathcal{B}_W = \{w_1, \ldots, w_m\}$ de $V$ y $W$, respectivamente. Dados $T \in L(V)$ y $S \in L(W)$, sean $A = [T]_{\mathcal{B}_V}$ y $C = [S]_{\mathcal{B}_W}$.

   *a*) Probar que
$$[T \otimes S]_{\mathcal{B}_V \otimes \mathcal{B}_W} = \begin{pmatrix} a_{11}C & a_{12}C & \cdots & a_{1n}C \\ a_{21}C & a_{22}C & \cdots & a_{2n}C \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}C & a_{n2}C & \cdots & a_{nn}C \end{pmatrix} \in \mathbb{K}^{nm \times nm}\,.$$

   *b*) Mostrar que $\operatorname{tr}(T \otimes S) = \operatorname{tr}(A)\operatorname{tr}(C)$ y $\det(T \otimes S) = \det(A)^m \det(C)^n$.

   *c*) Probar que si $\lambda_1, \ldots, \lambda_n$ son los autovalores de $T$ y $\mu_1, \ldots, \mu_m$ son los autovalores de $S$ (contando multiplicidades), entonces los $nm$ autovalores de $T \otimes S$ son exactamente
$$\eta_{ij} = \lambda_i\, \mu_j \quad \text{para } i = 1, \ldots, n\,,\ j = 1, \ldots, m\,.$$
   Deducir el ítem *b*) a partir de este resultado.

6. Probar que si $V$ y $W$ son dos $\mathbb{K}$-espacios vectoriales de dimensión finita, entonces $L(V \otimes W)$ es el producto tensorial de $L(V)$ y $L(W)$.

7. Dar ejemplos de tensores en $\mathbb{R}^4$ que sean:

   *a*) 2 veces covariantes.

   *b*) 4 veces covariantes.

   *c*) $n$ veces covariantes.

8. Dar ejemplos de tensores de tipo $(0,0)$, $(0,1)$, $(0,2)$, $(2,2)$. En cada caso aclarar el espacio vectorial considerado y su dimensión.

9. Probar que al contraer $p$ veces un tensor de tipo $(p,p)$, se obtiene como resultado un escalar.

10. Probar que la contracción de un operador lineal es su traza.

11. Probar que el producto escalar canónico es un tensor métrico cuyas componentes en la base usual están dadas por la delta de Kronecker.

12. Consideremos un tensor de tipo $(2,1)$, $a_k^{ij}$.

   *a*) ¿Qué tipo de tensor se obtiene al contraerlo con un tensor métrico $g_{pj}$?

   *b*) ¿Qué tipo de tensor se obtiene al contraerlo con un tensor métrico $g^{qk}$?

# Capítulo 11

# Grupos de matrices

Para comenzar, recordemos la definición de grupo.

**Definición 11.1.** Un **grupo** es un conjunto $G$ dotado con una operación binaria

$$\cdot : G \times G \to G, \qquad (x, y) \mapsto xy,$$

que verifica las siguientes condiciones:

i) $(xy)z = x(yz)$ para cualesquiera $x, y, z \in G$;

ii) existe un único $e \in G$ tal que $ex = xe = x$ para todo $x \in G$;

iii) para cada $x \in G$ existe un único $x^{-1} \in G$ tal que $xx^{-1} = x^{-1}x = e$.

Si además verifica

iv) $xy = yx$ para cualesquiera $x, y \in G$,

diremos que el grupo es un **grupo abeliano**.

**Ejemplos 11.2.** A continuación presentaremos algunos grupos abelianos con los cuales estuvimos trabajando durante todo el curso.

1. Si $(\mathbb{K}, +, \cdot)$ es un cuerpo, entonces
   i) $(\mathbb{K}, +)$ y $(\mathbb{K} \setminus \{0\}, \cdot)$ son grupos abelianos.
   ii) $(\mathbb{K}^n, +)$ y $(\mathbb{K}^{m\times n}, +)$ también son grupos abelianos.
2. Si $V$ es un $\mathbb{K}$-espacio vectorial, entonces $(V, +)$ es un grupo abeliano.
3. Si $S^1 := \{z \in \mathbb{C} : \ |z| = 1\}$, entonces $(S^1, \cdot)$ es un grupo abeliano.

El objetivo principal de este capítulo es presentar algunos grupos de matrices particulares, los cuales casualmente no son abelianos (en caso de que $n \geq 2$).

**Ejemplos 11.3.**

1. Si $\mathbb{K}$ es un cuerpo,

$$\mathrm{GL}(n, \mathbb{K}) = \{A \in \mathbb{K}^{n\times n} : \ \det(A) \neq 0 \ \}, \tag{11.1}$$

   es un grupo con la multiplicación usual de matrices. Se llama el **grupo lineal general** (de orden $n$) sobre $\mathbb{K}$.

2. Si $\mathbb{K} = \mathbb{C}$,
$$\mathrm{U}(n, \mathbb{C}) = \{U \in \mathrm{GL}(n, \mathbb{C}) : \ U^{-1} = U^* \ \}, \tag{11.2}$$
es el grupo de matrices (complejas) **unitarias**.

3. Si $\mathbb{K} = \mathbb{R}$ ó $\mathbb{K} = \mathbb{C}$,
$$\mathrm{O}(n, \mathbb{K}) = \{A \in \mathrm{GL}(n, \mathbb{K}) : \ A^{-1} = A^\top \ \}, \tag{11.3}$$
es el grupo de matrices **ortogonales**.

4. Si $\mathbb{K}$ es un cuerpo,
$$\mathrm{SL}(n, \mathbb{K}) = \{A \in \mathbb{K}^{n \times n} : \ \det(A) = 1 \ \}, \tag{11.4}$$
es el **grupo lineal especial** (de orden $n$) sobre $\mathbb{K}$.

**Definición 11.4.** Dado un grupo $G$, sea $H \subset G$. Se dice que $H$ es un subgrupo de $G$ si satisface:

i) $e \in H$;

ii) si $x, y \in H$ entonces $xy \in H$;

iii) si $x \in H$ entonces $x^{-1} \in H$.

Repasando los ejemplos anteriores, notemos que $\mathrm{SL}(n, \mathbb{K})$ y $\mathrm{O}(n, \mathbb{K})$ son subgrupos de $\mathrm{GL}(n, \mathbb{K})$. También es fácil ver que $\mathrm{U}(n, \mathbb{C})$ es un subgrupo de $\mathrm{GL}(n, \mathbb{C})$.

**Ejercicio 11.5.** Sea $G$ un grupo. Si $H_1$ y $H_2$ son dos subgrupos de $G$, probar que $H_1 \cap H_2$ también es un subgrupo de $G$.

Otro grupo destacado es
$$\mathrm{SO}(3) = \{A \in \mathrm{O}(3, \mathbb{R}) : \ \det(A) = 1 \ \}. \tag{11.5}$$
Este grupo es la intersección de $\mathrm{O}(3, \mathbb{R})$ y $\mathrm{SL}(3, \mathbb{R})$. Cada elemento de este grupo puede asociarse a una rotación en $\mathbb{R}^3$ (con respecto al origen). Recordemos que cada rotación (no trivial) queda determinada por su **eje de rotación**, el cual es una recta que pasa por el origen, y por su **ángulo de rotación**.

## 11.1. El grupo $\mathrm{SU}(2)$ y la superficie $S^3$

Consideremos el conjunto
$$\mathrm{SU}(2) = \{U \in \mathrm{U}(2, \mathbb{C}) : \ \det(U) = 1 \ \}. \tag{11.6}$$
En primer lugar, notemos que $\mathrm{SU}(2) = \mathrm{U}(2, \mathbb{C}) \cap \mathrm{SL}(2, \mathbb{C})$, con lo cual también es un (sub)grupo. Lo llamaremos el **grupo especial unitario** de orden 2.

Dada $U \in \mathrm{SU}(2)$, supongamos que
$$U = \begin{pmatrix} w & z \\ x & y \end{pmatrix}, \quad w, z, x, y \in \mathbb{C}.$$

Luego, $U^* = \begin{pmatrix} \overline{w} & \overline{x} \\ \overline{z} & \overline{y} \end{pmatrix}$ mientras que $U^{-1} = \dfrac{1}{\det(U)} \begin{pmatrix} y & -x \\ -z & w \end{pmatrix}^\top = \begin{pmatrix} y & -z \\ -x & w \end{pmatrix}$, ya que $\det(U) = 1$. Entonces, $U^* = U^{-1}$ implica que

$$U = \begin{pmatrix} w & z \\ -\overline{z} & \overline{w} \end{pmatrix}, \qquad w, z \in \mathbb{C},$$

mientras que $\det(U) = 1$ dice que $|w|^2 + |z|^2 = 1$. Esta es claramente una parametrización de $\mathrm{SU}(2)$ en función de dos variables complejas.

A continuación presentaremos una parametrización en términos de variables reales, utilizando las **matrices de Pauli**.

Supongamos que escribimos las descomposiciones cartesianas de $w$ y $z$ como:

$$\begin{aligned} w &= x_0 + ix_3, \quad x_0, x_3 \in \mathbb{R}, \\ z &= x_2 + ix_1, \quad x_1, x_2 \in \mathbb{R}. \end{aligned}$$

Luego,

$$\begin{aligned} U &= \begin{pmatrix} w & z \\ -\overline{z} & \overline{w} \end{pmatrix} = \begin{pmatrix} x_0 + ix_3 & x_2 + ix_1 \\ -x_2 + ix_1 & x_0 - ix_3 \end{pmatrix} \\ &= \begin{pmatrix} x_0 & 0 \\ 0 & x_0 \end{pmatrix} + i \begin{pmatrix} 0 & x_1 \\ x_1 & 0 \end{pmatrix} + \begin{pmatrix} 0 & x_2 \\ -x_2 & 0 \end{pmatrix} + i \begin{pmatrix} x_3 & 0 \\ 0 & -x_3 \end{pmatrix} \\ &= x_0 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + ix_1 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + ix_2 \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + ix_3 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \end{aligned}$$

Entonces, definiendo

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{11.7}$$

notemos que para $j = 1, 2, 3$, las matrices $\sigma_j$ cumplen $\sigma_j^* = \sigma_j$ y además $\mathrm{tr}(\sigma_j) = 0$.

Las tres matrices en (11.7) son llamadas **matrices de Pauli**, y si anotamos $\vec{x} = (x_0, x_1, x_2, x_3) \in \mathbb{R}^4$, resulta que

$$U = x_0 I + i(x_1\sigma_1 + x_2\sigma_2 + x_3\sigma_3), \qquad \text{con } x_0^2 + x_1^2 + x_2^2 + x_3^2 = 1.$$

Por lo tanto, si consideramos a $\mathbb{R}^4$ como un espacio euclidiano con el producto interno usual, podemos identificar a cada elemento de $\mathrm{SU}(2)$ con un único elemento vector en la superficie esférica de $\mathbb{R}^4$:

$$S^3 := \{\vec{x} \in \mathbb{R}^4 : \ \|\vec{x}\| = 1\}.$$

La correspondencia establecida entre $S^3$ y $\mathrm{SU}(2)$ se extiende a un isomorfismo entre $\mathbb{R}^4$ y el **espacio de cuaterniones**. Este último es un $\mathbb{R}$-espacio vectorial de matrices de $2 \times 2$:

$$\mathbb{H} := \left\{ \begin{pmatrix} x_0 + ix_3 & x_2 + ix_1 \\ -x_2 + ix_1 & x_0 - ix_3 \end{pmatrix} : \ x_0, x_1, x_2, x_3 \in \mathbb{R} \right\}. \tag{11.8}$$

La hache es un homenaje al físico-matemático irlandés William Rowan Hamilton (1805–1865), a quien se le atribuye la invención de los cuaterniones.

Es fácil verificar que $\mathbb{H}$ es un $\mathbb{R}$-espacio vectorial de dimensión 4, y que $\mathbb{H}$ es un conjunto cerrado con respecto a la multiplicación de matrices. Además, todas las matrices no nulas de $\mathbb{H}$ son inversibles, y su inversa también pertenece a $\mathbb{H}$. $\mathbb{H}$ es entonces una estructura algebraica con dos operaciones, suma y producto, que cumple con todas las propiedades para ser un cuerpo, con la excepción de que el producto no es conmutativo. Podemos decir entonces que $\mathbb{H}$ es un "cuerpo no conmutativo", o un **álgebra con división**, de acuerdo a la terminología actual.

Notemos que $\mathbb{H}$ contiene a una copia de $\mathbb{C}$ como subespacio (las matrices de la forma $x_0 I + i x_3 \sigma_3$).

## 11.2. Grupos que preservan formas bilineales

Si $V$ es un $\mathbb{K}$-espacio vectorial, entonces

$$\mathrm{GL}(V) := \{T \in L(V) : \ T \text{ es inversible } \}, \tag{11.9}$$

es un grupo, se llama el **grupo lineal general sobre** $V$.

**Definición 11.6.** Dado un $\mathbb{K}$-espacio vectorial $V$, supongamos que $\mathcal{A} \in \mathrm{Bil}(V)$ y $T \in L(V)$. Diremos que $T$ **preserva a** $\mathcal{A}$ si

$$\mathcal{A}(Tv, Tw) = \mathcal{A}(v, w), \qquad \text{para todo } v, w \in V.$$

El siguiente resultado dice que, si $\mathcal{A}$ es no degenerada, los operadores que preservan a $\mathcal{A}$ forman un subgrupo de $\mathrm{GL}(V)$.

**Teorema 11.7.** *Dado un $\mathbb{K}$-espacio vectorial $V$ de dimensión finita, sea $\mathcal{A} \in \mathrm{Bil}(V)$ una forma no degenerada. Entonces, el conjunto de los operadores lineales que preservan a $\mathcal{A}$ es un grupo (con respecto al producto de operadores).*

*Demostración.* Es inmediato que $I \in L(V)$ preserva a $\mathcal{A}$. Supongamos ahora que $S, T \in L(V)$ preservan a $\mathcal{A}$. Luego, dados $v, w \in V$,

$$\mathcal{A}(STv, STw) = \mathcal{A}(S(Tv), S(Tw)) = \mathcal{A}(Tv, Tw) = \mathcal{A}(v, w),$$

es decir, $ST$ preserva a $\mathcal{A}$.

Finalmente, veamos que si $T$ preserva a $\mathcal{A}$ entonces $T \in \mathrm{GL}(V)$ y además $T^{-1}$ también preserva a $\mathcal{A}$. Supongamos que $v \in N(T)$. Entonces, para todo $w \in V$,

$$\mathcal{A}(v, w) = \mathcal{A}(Tv, Tw) = \mathcal{A}(\vec{0}, Tw) = 0.$$

Luego, como $\mathcal{A}$ es no degenerada, esto implica que $v = \vec{0}$. Por lo tanto, $N(T) = \{\vec{0}\}$, o equivalentemente, $T \in \mathrm{GL}(V)$.

Además, dados $v, w \in V$,

$$\mathcal{A}\big(T^{-1}v, T^{-1}w\big) = \mathcal{A}\big(T(T^{-1}v), T(T^{-1}w)\big) = \mathcal{A}\big((TT^{-1})v, (TT^{-1})w\big) = \mathcal{A}(v, w),$$

es decir, $T^{-1}$ preserva a $\mathcal{A}$. $\square$

Si $V$ es un $\mathbb{K}$-espacio vectorial de dimensión finita, cada forma $\mathcal{A} \in \mathrm{Bil}(V)$ no degenerada determina, fijando una base $\mathcal{B}$ de $V$, un grupo de matrices:

$$G := \big\{ [T]_{\mathcal{B}} : \ T \text{ preserva a } \mathcal{A} \ \big\}.$$

También podemos describir a este grupo de la siguiente manera: supongamos que $A = [\mathcal{A}]_{\mathcal{B}}$, es decir,

$$\mathcal{A}(v, w) = [w]_{\mathcal{B}}^{\top} A [v]_{\mathcal{B}}, \qquad \text{para todo } v, w \in V.$$

Dado $T \in L(V)$, sea $M = [T]_{\mathcal{B}}$. Luego,

$$T \text{ preserva a } \mathcal{A} \quad \Leftrightarrow \quad M^\top A M = A.$$

De hecho, fijados $v, w \in V$,

$$\mathcal{A}(Tv, Tw) = [Tw]_{\mathcal{B}}^\top A[Tv]_{\mathcal{B}} = ([T]_{\mathcal{B}}[w]_{\mathcal{B}})^\top A([T]_{\mathcal{B}}[v]_{\mathcal{B}}) = [w]_{\mathcal{B}}^\top (M^\top AM)[v]_{\mathcal{B}},$$

de donde se deduce la equivalencia anterior.

Por lo tanto, si $\dim V = n$ y $A = [\mathcal{A}]_{\mathcal{B}}$ entonces

$$G = \big\{ M \in \mathrm{GL}(n, \mathbb{K}) : \ M^\top A M = A \ \big\}. \tag{11.10}$$

**Ejemplo 11.8.** Si $K = \mathbb{R}$ ó $\mathbb{K} = \mathbb{C}$, $\mathrm{O}(n, \mathbb{K})$ es el grupo de matrices que preservan la forma bilineal simétrica

$$\mathcal{A}(\vec{x}, \vec{y}) := \sum_{j=1}^{n} x_j y_j,$$

para $\vec{x} = (x_1, \ldots, x_n), \vec{y} = (y_1, \ldots, y_n) \in \mathbb{K}^n$.

En efecto, si $\mathcal{E}$ es la base canónica de $\mathbb{K}^n$, entonces la matriz asociada a $\mathcal{A}$ es la identidad.

Antes de pasar al siguiente ejemplo, prestemos atención al siguiente comentario. Supongamos que $\mathcal{A}$ es una forma bilineal simétrica sobre un $\mathbb{K}$-espacio vectorial $V$. Entonces, un operador $T \in L(V)$ preserva a $\mathcal{A}$ si y sólo si $T$ preserva a la forma cuadrática asociada a $\mathcal{A}$:

$$Q(v) = \mathcal{A}(v, v), \quad v \in V.$$

Si $T$ preserva a $\mathcal{A}$ es inmediato que $T$ también preserva a $Q$. Recíprocamente, si $T$ satisface $Q(Tv) = Q(v)$ para todo $v \in V$, utilizando la identidad de polarización podemos comprobar que $T$ preserva a $\mathcal{A}$.

**Ejemplo 11.9.** Sea $\mathcal{A}$ la forma bilineal simétrica sobre $\mathbb{R}^4$ asociada a la forma cuadrática:

$$Q(t, x, y, z) = t^2 - x^2 - y^2 - z^2.$$

Un operador $T \in L(\mathbb{R}^4)$ que preserva esta forma cuadrática se denomina **transformación de Lorentz**, y el grupo de operadores que preservan a $\mathcal{A}$ se llama el **grupo de Lorentz** y se anota $\mathrm{O}(1, 3)$. Este grupo debe su nombre al físico holandés Hendrik Antoon Lorentz (1853–1928), quien determinó las ecuaciones de transformación que sustentan la teoría especial de la relatividad de Albert Einstein.

A continuación presentaremos un método para describir al grupo de Lorentz. Sea $H$ el siguiente $\mathbb{R}$-espacio vectorial:

$$H = \{ A \in \mathbb{C}^{2\times 2} : \ A^* = A \ \}.$$

Como $\dim_{\mathbb{R}}(H) = 4$, sabemos que $H$ es isomorfo a $\mathbb{R}^4$. En particular, la transformación

$$\Phi : \mathbb{R}^4 \to H, \qquad \text{dada por} \quad \Phi(t, x, y, z) = \begin{pmatrix} t + x & y + iz \\ y - iz & t - x \end{pmatrix},$$

es un isomorfismo entre $\mathbb{R}^4$ y $H$.

Bajo este isomorfismo, la forma cuadrática $Q$ se transforma en la función determinante, es decir,

$$Q(t, x, y, z) = \det \begin{pmatrix} t + x & y + iz \\ y - iz & t - x \end{pmatrix},$$

o equivalentemente, $Q(\vec{v}) = \det(\Phi(\vec{v}))$ si $\vec{v} \in \mathbb{R}^4$.

Esto sugiere que podemos estudiar a las transformaciones de Lorentz en $\mathbb{R}^4$ estudiando

$$G = \{T \in L(H) : T \text{ preservan el determinante}\}.$$

Dada $M \in \mathbb{C}^{2\times 2}$, sea

$$\operatorname{Ad}_M : H \to H \quad \text{dado por} \quad \operatorname{Ad}_M(A) = MAM^*.$$

Veamos cómo debe ser $M$ para que $\operatorname{Ad}_M$ preserve el determinante:

$$\begin{aligned}\det(\operatorname{Ad}_M(A)) &= \det(MAM^*) = \det(M)\det(A)\det(M^*) = \det(M)\det(A)\overline{\det(M^\top)} \\ &= |\det(M)|^2\det(A),\end{aligned}$$

así que $\operatorname{Ad}_M \in G$ si y sólo si $|\det(M)| = 1$. Entonces, para cada $M \in \mathbb{C}^{2\times 2}$ con $|\det(M)| = 1$, sea

$$T_M : \mathbb{R}^4 \to \mathbb{R}^4 \quad \text{dado por} \quad T_M = \Phi^{-1} \circ \operatorname{Ad}_M \circ \Phi\ .$$

Como $\operatorname{Ad}_M \in G$, para cada $\vec{v} \in \mathbb{R}^4$,

$$\begin{aligned}Q(T_M\vec{v}) &= Q(\Phi^{-1}\operatorname{Ad}_M \Phi\vec{v}) = \det\left(\Phi(\Phi^{-1}\operatorname{Ad}_M \Phi\vec{v})\right) = \det(\operatorname{Ad}_M \Phi\vec{v}) = \det(\Phi\vec{v}) \\ &= Q(\vec{v}),\end{aligned}$$

es decir, $T_M$ es una transformación de Lorentz.

Sin embargo, no todas las trasformaciones de Lorentz pueden obtenerse con este método.

**Ejemplo 11.10.** Sea $\mathcal{A}$ la forma bilineal simétrica sobre $\mathbb{R}^n$ cuya forma cuadrática asociada es

$$Q(x_1, \dots, x_n) = \sum_{j=1}^{p} x_j^2 - \sum_{j=p+1}^{n} x_j^2.$$

Esta forma es no degenerada y tiene signatura $p - (n - p) = 2p - n$. El grupo de matrices (asociadas a los operadores lineales) que preservan una forma de este tipo se denomina **grupo pseudo-ortogonal**, y se anota $\mathrm{O}(p, n - p)$.

Cuando $p = n$, recuperamos el grupo ortogonal $\mathrm{O}(n, \mathbb{R})$ como un caso particular de esta familia de grupos. Además, para cada uno de los $n+1$ posibles valores de $p$, $p = 0, 1, \dots, n$, tenemos una forma bilineal $\mathcal{A}$ diferente; sin embargo, la forma que resulta para $p = n - k$ es la opuesta de la obtenida para $p = k$, con lo cual,

$$\mathrm{O}(p, n - p) = \mathrm{O}(n - p, p).$$

Entonces, cuando $n$ es impar tenemos $\frac{n+1}{2}$ grupos de matrices pseudo-ortogonales en $\mathbb{R}^{n\times n}$, y cuando $n$ es par tenemos $\frac{n}{2}$ de estos grupos.

# Glosario

**A**

**algebra lineal** sobre un cuerpo. 31

**algebraicamente cerrado** , cuerpo. 58

**angulo** mínimo entre dos vectores. 133

**antisimétrica** , forma bilineal. 170

**anulador** de un subespacio. 115

**autoadjunta** , matriz. 155

**autoadjunto** , operador lineal. 153

**autoespacio** asociado a un autovalor. 62

**autovalor** de un operador lineal. 62

**autovector** asociado a un autovalor. 62

**B**

**base** de un espacio vectorial. 13

**base dual** asociada a una base. 114

**base canónica** . 13

**base ordenada** . 13

**bloque de Jordan** asociado a un autovalor. 105

**bloque elemental de Jordan** asociado a un autovalor. 105

**C**

**cadena de Jordan** . 104

**característica** de un cuerpo. 52

**ciclico** , subespacio. 97

**combinación lineal** . 9

**complemento ortogonal** . 137

**M**



**N**



**O**



**P**

**R**



**S**



**T**

# Bibliografía


[1] Axler, S. (2015). Linear Algebra Done Right, Third Edition. Heidelberg: Springer.

[2] Gentile, E. (1968). Espacios vectoriales. Buenos Aires: Functor.

[3] Hawkings, T. Cauchy and the spectral theory of matrices, Historia Mathematica 2 (1975), 1–29.

[4] Hoffman, K. y Kunze, R. (1973). Álgebra Lineal. México D.F.: Prentice-Hall Interamericana.

[5] Lang, S. (1966). Linear Algebra. Reading: Addison-Wesley.

[6] Lipschutz, S. (1991). Schaum's Outline of Theory and Problems of LINEAR ALGEBRA, Second Edition. New York: McGraw-Hill.

[7] Nomizu, K. (1966). Fundamentals of Linear Algebra. New York: McGraw-Hill.

[8] Santaló, L. A. (1970). Vectores y tensores con sus aplicaciones, Octava Edición. Buenos Aires: Eudeba.

[9] Vinberg, E. B. (2003). A course in Algebra, Graduate Studies in Mathematics vol. 56. Providence: Amer. Math. Soc.