Publicaciones Electrónicas
Sociedad Matemática Mexicana

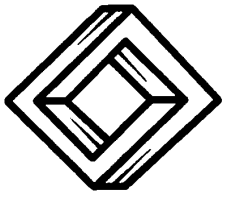
# Teoría de Espacios Vectoriales

Fernando Ignacio Becerra López
Alonso Castillo Ramírez
Alfonso Manuel Hernández Magdaleno
Osbaldo Mata Gutiérrez

www.smm.org.mx

Serie: Textos. Vol. 24 (2023)

ISBN: 978-607-8008-17-9



# Teoría
# de
# Espacios Vectoriales

**Fernando Ignacio Becerra López**
**Alonso Castillo Ramírez**
**Alfonso Manuel Hernández Magdaleno**
**Osbaldo Mata Gutiérrez**

*Centro Universitario de Ciencias Exactas e Ingenierías*
*Universidad de Guadalajara*

Publicaciones Electrónicas
Sociedad Matemática Mexicana

# Índice general

# Prefacio

La teoría de los espacios vectoriales engloba los fundamentos teóricos de la rama de las matemáticas conocida como "álgebra lineal", la cual tiene importantes aplicaciones en ingeniería, física, biología, ciencias computacionales, y economía, entre otras ciencias. Dentro de las matemáticas mismas, el álgebra lineal es una pieza fundamental en el desarrollo del álgebra multilineal, las ecuaciones diferenciales, la teoría de módulos, el análisis funcional, la teoría de representaciones y la geometría algebraica.

El álgebra lineal se originó como el estudio de los sistemas de ecuaciones lineales, el cual evolucionó naturalmente al estudio de matrices y vectores geométricos. La definición moderna de *espacio vectorial* fue presentada por Giuseppe Peano en 1888, y su desarrollo teórico se dio principalmente durante la primera mitad del siglo XX. Es posible que el lector conozca alguna definición del concepto de *vector*, como la siguiente:

**Definición [vector geométrico].** Un *vector geométrico* es un objeto que tiene magnitud y dirección.

Sin embargo, en este libro, nuestra definición de vector será la siguiente:

**Definición [vector].** Un *vector* es un elemento de un espacio vectorial.

De esta forma, los vectores son objetos abstractos que satisfacen los axiomas establecidos por la definición de espacio vectorial (ver Definición 1.1); no asumiremos que los vectores satisfacen ninguna otra propiedad (en particular, para nosotros no tienen magnitud ni dirección).

Este libro está dirigido a estudiantes de tercer semestre de la Licenciatura en Matemáticas del Centro Universitario de Ciencias Exactas e Ingenierías de la Universidad de Guadalajara. Sin embargo, estamos convencidos de que nuestro enfoque será benéfico a estudiantes avanzados de carreras afines.

Iniciamos el libro con el Capítulo 0 sobre temas preliminares, donde repasamos los conceptos de relación de equivalencia, clase de equivalencia, operación binaria, grupo y campo; además, abordamos algunos ejemplos básicos de campos que son indispensables posteriormente, como los números racionales, los números reales, los números complejos y los enteros módulo un primo.

En el Capítulo 1 estudiamos la definición de espacio vectorial, nuestro principal objeto de estudio en este libro. Demostramos algunas de las propiedades elementales de los espacios vectoriales y examinamos en detalle algunos ejemplos, incluyendo al espacio euclídeo $\mathbb{R}^n$, espacios de matrices, espacios de funciones y espacios de polinomios.

En el Capítulo 2 presentamos la definición de subespacio vectorial y demostramos su equivalencia con otras afirmaciones (test del subespacio 1 y 2). También estudiamos algunas formas de construir subespacios, entre las que están los subespacios generados por conjuntos, la intersección de subespacios y la suma de subespacios.

En el Capítulo 3 analizamos las funciones entre espacios vectoriales que preservan sus estructuras: las transformaciones lineales. Estudiamos sus propiedades básicas y definimos los conceptos de isomorfismo, imagen y kernel.

En el Capítulo 4, dado un subespacio $S$ de un espacio vectorial $V$, construimos un nuevo espacio vectorial $V/S$ llamado el espacio vectorial cociente. Los vectores de este nuevo espacio son clases laterales de la forma $v + S$. Es en este capítulo donde aparecen el Primer y el Segundo Teorema de Isomorfía.

En el Capítulo 5 estudiamos el concepto de independencia lineal para definir lo que es una *base* de un espacio vectorial. A grandes rasgos, una base es un conjunto de vectores que determinan la estructura de todo el espacio. La cardinalidad de una base se conoce como la *dimensión* de un espacio, y resulta ser una característica clave, ya que dos espacios vectoriales definidos sobre un mismo campo son isomorfos si y sólo si tienen la misma dimensión.

En el Capítulo 6 consideramos espacios vectoriales de dimensión finita. Analizamos cómo se comporta la dimensión en la suma de subespacios, el espacio cociente y el espacio de las transformaciones lineales, y, después de un breve repaso sobre matrices, mostramos cómo un vector puede representarse mediante sus coordenadas respecto a una base y cómo una transformación lineal puede representarse mediante una matriz. Mediante estas representaciones, descubrimos que aplicar una transformación lineal a un vector es lo mismo que multiplicar la matriz correspondiente por las coordenadas del vector, y que la composición de transformaciones lineales es igual a la multiplicación de sus matrices correspondientes.

En el Capítulo 7 estudiamos algunos temas relacionados con la teoría de matrices, como los determinantes y las matrices elementales, para terminar con el Teorema Fundamental de Matrices Invertibles.

En el Capítulo 8 definimos la relación de equivalencia de similitud de matrices: esencialmente, dos matrices son similares si y sólo si representan a la misma transformación lineal respecto a dos bases posiblemente distintas. Nuestro objetivo de estudiar propiedades comunes que poseen las matrices similares nos lleva a considerar los autovalores y autovectores de una matriz.

En el Capítulo 9 estudiamos principalmente dos teoremas muy importantes: el Teorema de Cayley-Hamilton, el cual establece que cualquier matriz es una raíz de su polinomio característico, y el Teorema de la Forma Canónica de Jordan, el cual establece que cualquier matriz compleja es similar a una matriz triangular superior en la forma de Jordan.

Finalmente, en el Capítulo 10 presentamos una aplicación del álgebra lineal a las ciencias computacionales y la teoría de la información. Mostramos cómo a través de los espacios vectoriales sobre campos finitos es posible diseñar un sistema de transmisión de datos en el que sea posible detectar, e incluso corregir, errores ocurridos en la transmisión. Creemos que esta aplicación ilustra el poder de la abstracción matemática, así como la belleza y creatividad detrás de muchos de sus conceptos.

# 0

# Conceptos Preliminares

Este capítulo es un repaso de algunos conceptos preliminares que serán necesarios en nuestro estudio de espacios vectoriales. Para un repaso más a fondo sobre estos temas sugerimos consultar el libro *Conjuntos y Números* [2].

## 0.1. Relaciones de equivalencia

Sean $A$ y $B$ conjuntos. Recordemos que una relación $R$ de $A$ en $B$ es un subconjunto del producto cartesiano $A \times B$. Escribimos $aRb$ si $(a, b) \in R$.

**Ejemplo 0.1.** Sea $f : A \to B$ cualquier función entre los conjuntos $A$ y $B$. La siguiente relación se llama *la gráfica de la función*:

$$R := \{(a, b) \in A \times B | f(a) = b\}.$$

**Ejemplo 0.2.** Sea $f : A \to C$ y $g : B \to C$ dos funciones. La siguiente relación se llama el *producto fibrado* de $f$ y $g$:

$$R := \{(a, b) \in A \times B | f(a) = g(b)\}.$$

Una relación *sobre* $A$ es simplemente una relación de $A$ en $A$.

**Definición 0.3 (relación de equivalencia).** Una relación $R$ sobre un conjunto $A$ es una *relación de equivalencia* si se cumplen las siguientes propiedades:

**(E1)** $R$ es *reflexiva*: $aRa$ para toda $a \in A$.

**(E2)** $R$ es *simétrica*: $aRb$ implica $bRa$.

**(E3)** $R$ es *transitiva*: $aRb$ y $bRc$ implican $aRc$.

**Ejemplo 0.4.** Sea $\mathbb{Z}$ el conjunto de los números enteros y sea $n \geq 1$ un número entero. La relación

$$R_n = \{(a, b) \in \mathbb{Z} \times \mathbb{Z} : n \mid (a - b)\}$$

es una relación de equivalencia llamada la *congruencia módulo* $n$. Si $(a, b) \in R_n$ escribimos $a \equiv b \mod (n)$. Es necesario demostrar que las propiedades **(E1)**, **(E2)** y **(E3)** se cumplen:

**(E1)** Para cualquier $a \in \mathbb{Z}$, $n \mid (a-a) = 0$, así que $a \equiv a \mod (n)$.

**(E2)** Si $n \mid (a-b)$, entonces $n \mid (b-a)$. Por lo tanto, $a \equiv b \mod (n)$ implica $b \equiv a \mod (n)$.

**(E3)** Ejercicio 0.36.

**Definición 0.5 (clase de equivalencia).** Sea $R$ una relación de equivalencia sobre $A$. La *clase de equivalencia* de un elemento $a \in A$, denotada como $[a]$, es el subconjunto de $A$ definido como

$$[a] = \{x \in A : xRa\}.$$

Al conjunto de todas las clases de equivalencia de los elementos de $A$ se le llama el *conjunto cociente* de $A$ por $R$, y se denota como $A/R$. En símbolos,

$$A/R = \{[a] : a \in A\}.$$

**Ejemplo 0.6.** Sea $R_n$ la relación de congruencia módulo $n$. Para cualquier $a \in \mathbb{Z}$, la clase de equivalencia módulo $n$ de $a$ es

$$[a] = \{x \in \mathbb{Z} : x \equiv a \mod (n)\}.$$

El conjunto cociente $\mathbb{Z}/R_n$, denotado en este caso simplemente como $\mathbb{Z}_n$, es

$$\mathbb{Z}_n := \mathbb{Z}/R_n = \{[a] : a \in \mathbb{Z}\} = \{[0], [1], [2], \ldots, [n-1]\}.$$

Es posible demostrar esta última igualdad usando el algoritmo de la división.

**Lema 0.7 (propiedades básicas de las clases de equivalencia).** Sea $R$ una relación de equivalencia sobre $A$.

**(1)** $a \in [a]$ para toda $a \in A$.

**(2)** $[a] = [b]$ si y sólo si $aRb$.

**(3)** Si $[a] \neq [b]$, entonces $[a] \cap [b] = \emptyset$.

**(4)** $A = \bigcup_{a \in A} [a]$.

**Demostración.** Ejercicio 0.37. □

## 0.2. Grupos

**Definición 0.8 (operación binaria).** Sea $G$ un conjunto no vacío. Una *operación binaria* de $G$ es una función de la forma $f : G \times G \to G$.

En general, para verificar que una operación binaria $f : G \times G \to G$ *está bien definida* hay que asegurarse que realmente $f(a,b) \in G$ para cualquier $(a,b) \in G \times G$. Comúnmente llamamos a esto la propiedad de *cerradura* de la operación.

**Ejemplo 0.9.** Consideremos algunos ejemplos y contraejemplos.

1. La función $+ : \mathbb{Z}\times\mathbb{Z} \to \mathbb{Z}$ definida como $+(n, m) := n+m$ es una operación binaria del conjunto $\mathbb{Z}$ llamada la *suma usual de números enteros.*

2. La función $+ : \mathbb{Q} \times \mathbb{Q} \to \mathbb{Q}$ definida como $+(\frac{a_1}{b_1}, \frac{a_2}{b_2}) := \frac{a_1b_2+a_2b_1}{b_1b_2}$ es una operación binaria del conjunto $\mathbb{Q}$ de números racionales llamada la *suma usual de números racionales.*

3. La resta **no** es una operación binaria del conjunto $\mathbb{N}$ de números naturales porque no cumple la propiedad de cerradura (por ejemplo, $3-4 = -1 \notin \mathbb{N}$).

4. La función $\text{mcd} : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ que asigna a cualquier par de números naturales su máximo común divisor es una operación binaria de $\mathbb{N}$.

5. La función $\cdot : 2\mathbb{Z}\times 2\mathbb{Z} \to 2\mathbb{Z}$ definida como $\cdot(n, m) := nm$ es una operación binaria del conjunto $2\mathbb{Z}$ de números enteros pares ya que el producto de dos enteros pares siempre es un entero par.

Una propiedad importante de las funciones es que cada elemento del dominio tiene una única imagen en el codominio. Por lo tanto, para demostrar que una operación binaria $f : G \times G \to G$ está bien definida, además de verificar la propiedad de cerradura, hay que verificar que si $(a, b) = (a', b')$, entonces $f(a, b) = f(a', b')$. Comprobar esto es trivial en los ejemplos anteriores, si embargo no es tan obvio cuando los elementos de $G$ son clases de equivalencia que dependen de un representante.

**Ejemplo 0.10 (suma módulo $n$).** Sea $\mathbb{Z}_n = \{[0], [1], \ldots, [n-1]\}$ el conjunto de clases de equivalencia módulo $n \in \mathbb{N}$, $n \neq 0$. Definimos una operación binaria $f : \mathbb{Z}_n \times \mathbb{Z}_n \to \mathbb{Z}_n$, llamada *suma módulo* $n$, como

$$f([a], [b]) := [a + b],$$

para cualquier $[a], [b] \in \mathbb{Z}_n$. Es obvio que $f$ cumple la propiedad de cerradura. Ahora demostraremos que si $[a] = [a']$ y $[b] = [b']$, entonces $f([a], [b]) = f([a'], [b'])$. Por el Lema 0.7,

$$a \equiv a' \mod (n) \text{ y } b \equiv b' \mod (n).$$

Luego $n \mid (a - a')$ y $n \mid (b - b')$, lo que implica que $a - a' = sn$ y $b - b' = tn$ para algunos $s, t \in \mathbb{Z}$. Sumando las ecuaciones previas, obtenemos

$$(a + b) - (a' + b') = (s + t)n,$$

y por lo tanto, $(a + b) \equiv (a' + b') \mod (n)$. Así, $f([a], [b]) = f([a'], [b'])$.

Normalmente, denotamos con un punto $\cdot$ a una operación binaria arbitraria de $G$, y denotamos como $a \cdot b$ a la imagen del par $(a, b) \in G \times G$.

Un *grupo* es una estructura algebraica que consiste en un conjunto y una operación binaria que cumple tres propiedades.

**Definición 0.11 (grupo).** Sea $G$ un conjunto no vacío y $\cdot$ una operación binaria de $G$. El par $(G, \cdot)$ es un *grupo* si se cumplen las siguientes propiedades:

**(G1)** *Asociatividad.* Para toda $a, b, c \in G$, se cumple que

$$a \cdot (b \cdot c) = (a \cdot b) \cdot c.$$

**(G2)** *Identidad.* Existe un elemento $e \in G$ tal que, para toda $a \in A$,

$$e \cdot a = a \cdot e = a.$$

**(G3)** *Inversos.* Para cualquier $a \in G$ existe un $b \in G$ tal que

$$a \cdot b = b \cdot a = e.$$

El elemento $e \in G$ de la propiedad **(G2)** es llamado la *identidad* de $G$. El elemento $b \in G$ de la propiedad **(G3)** es llamado el *inverso* de $a \in G$ y lo denotamos como $a^{-1}$.

**Ejemplo 0.12 (números enteros).** Sea $+$ la suma usual de $\mathbb{Z}$. El par $(\mathbb{Z}, +)$ es un grupo:

**(G1)** Para toda $n, m, k \in \mathbb{Z}$, se cumple que $(n + m) + k = n + (m + k)$.

**(G2)** La identidad es $0 \in \mathbb{Z}$ porque $0 + n = n + 0 = n$, para toda $n \in \mathbb{Z}$.

**(G3)** El inverso de cualquier $n \in \mathbb{Z}$ es $-n \in \mathbb{Z}$ porque $n + (-n) = (-n) + n = 0$.

**Ejemplo 0.13 (grupo trivial).** Consideremos un conjunto con un solo elemento $G = \{e\}$ y una operación binaria $\cdot$ definida como $e \cdot e = e$. El par $(\{e\}, \cdot)$ es un grupo: las propiedades **G1-G3** se cumplen obviamente. Llamamos a $(\{e\}, \cdot)$ el *grupo trivial*.

Enunciaremos algunos resultados básicos.

**Lema 0.14 (propiedades básicas de grupos).** Sea $(G, \cdot)$ un grupo.

**(1)** *Cancelación derecha.* Para toda $a, b, c \in G$, si $a \cdot c = b \cdot c$, entonces $a = b$.

**(2)** *Cancelación izquierda.* Para toda $a, b, c \in G$, si $c \cdot a = c \cdot b$, entonces $a = b$.

**(3)** *Unicidad de la identidad.* La identidad $e$ de $G$ es única.

**(4)** *Unicidad de los inversos.* Para toda $a \in G$, el inverso de $a$ es único.

**(5)** *Inverso del inverso.* Para toda $a \in G$, $(a^{-1})^{-1} = a$.

**Demostración.** Ejercicio 0.40. □

**Definición 0.15 (Grupo abeliano).** Decimos que un grupo $(G, \cdot)$ es *abeliano* si se cumple la siguiente propiedad:

**(G4)** *Conmutatividad.* Para toda $a, b \in G$, se cumple que

$$a \cdot b = b \cdot a.$$

**Ejemplo 0.16 (números racionales).** Sea $\cdot$ la multiplicación usual de números racionales: $\frac{a_1}{b_1} \cdot \frac{a_2}{b_2} := \frac{a_1 a_2}{b_1 b_2}$. Sea $\mathbb{Q}^* := \mathbb{Q} \setminus \{0\}$. Demostraremos que el par $(\mathbb{Q}^*, \cdot)$ es un grupo abeliano:

**(G1)** Para cualquier $\frac{a_i}{b_i} \in \mathbb{Q}^*$,

$$\begin{aligned} \frac{a_1}{b_1} \cdot \left( \frac{a_2}{b_2} \cdot \frac{a_3}{b_3} \right) &= \frac{a_1 \left(a_2 a_3\right)}{b_1 \left(b_2 b_3\right)} \\ &= \frac{\left(a_1 a_2\right) a_3}{\left(b_1 b_2\right) b_3} \\ &= \left( \frac{a_1}{b_1} \cdot \frac{a_2}{b_2} \right) \cdot \frac{a_3}{b_3}. \end{aligned}$$

**(G2)** La identidad es $\frac{1}{1} \in \mathbb{Q}^*$ porque $\frac{1}{1} \cdot \frac{a}{b} = \frac{a}{b} \cdot \frac{1}{1} = \frac{a}{b}$, para toda $\frac{a}{b} \in \mathbb{Q}^*$.

**(G3)** El inverso de cualquier $\frac{a}{b} \in \mathbb{Q}^*$ es $\frac{b}{a} \in \mathbb{Q}^*$ porque $\frac{a}{b} \cdot \frac{b}{a} = \frac{ab}{ba} = \frac{ab}{ab} = \frac{1}{1}$.

**(G4)** Para cualquier $\frac{a_i}{b_i} \in \mathbb{Q}^*$,

$$\frac{a_1}{b_1} \cdot \frac{a_2}{b_2} = \frac{a_1 a_2}{b_1 b_2} = \frac{a_2 a_1}{b_2 b_1} = \frac{a_2}{b_2} \cdot \frac{a_1}{b_1}.$$

Si $(G, \cdot)$ es un grupo y $H$ un subconjunto de $G$, denotamos por $\cdot_H$ a la restricción de $\cdot$ en $H$; en otras palabras, $\cdot_H$ es la función $\cdot_H : H \times H \to G$ definida como

$$a \cdot_H b = a \cdot b, \text{ donde } a, b \in H.$$

**Definición 0.17 (Subgrupo).** Sea $(G, \cdot)$ un grupo y $H \subseteq G$. Decimos que $(H, \cdot_H)$ es un *subgrupo* de $(G, \cdot)$ si $(H, \cdot_H)$ es en sí mismo un grupo.

Decimos que $(H, \cdot_H)$ es un subgrupo propio de $(G, \cdot)$ si $H \subsetneq G$.

**Teorema 0.18 (test del subgrupo).** Sea $(G, \cdot)$ un grupo y $H \subseteq G$. El par $(H, \cdot_H)$ es un subgrupo de $(G, \cdot)$ si y sólo si se cumplen las siguientes propiedades:

**(S1)** *Cerradura en $H$.* Para toda $a, b \in H$, se cumple que $a \cdot b \in H$.

**(S2)** *Identidad en $H$.* $e \in H$, donde $e$ es la identidad del grupo $(G, \cdot)$.

**(S3)** *Inversos en $H$.* Para cualquier $a \in H$, se cumple que $a^{-1} \in H$.

**Demostración.**

**($\Rightarrow$)** Si $(H, \cdot_H)$ es un subgrupo, claramente se cumplen las propiedades (**S1**)-(**S3**).

**(⇐)** Supongamos que el par $(H, \cdot_H)$ cumple las propiedades (**S1**)-(**S3**). La propiedad (**S1**) garantiza que $\cdot_H$ es una función de la forma $H \times H \to H$, así que es una operación binaria de $H$. Las propiedades (**S2**) y (**S3**) implican directamente que (**G2**) y (**G3**) se cumplen. Finalmente, $(H, \cdot_H)$ también cumple (**G1**) porque, para cualquier $a, b, c \in H$,

$$a \cdot_H (b \cdot_H c) = a \cdot (b \cdot c) = (a \cdot b) \cdot c = (a \cdot_H b) \cdot_H c.$$

Por lo tanto, $(H, \cdot_H)$ es un grupo en sí mismo.

□

Para simplificar notación, si $(H, \cdot_H)$ es un subgrupo de $(G, \cdot)$, denotamos la operación $\cdot_H$ con el mismo símbolo que la operación de $(G, \cdot)$.

**Ejemplo 0.19 ($n\mathbb{Z}$).** Sea $n \in \mathbb{N}$, $n \neq 0$. Consideremos al conjunto de los múltiplos enteros de $n$:

$$n\mathbb{Z} = \{nk : k \in \mathbb{Z}\}.$$

Claramente, $n\mathbb{Z}$ es un subconjunto de $\mathbb{Z}$ (un subconjunto propio si $n \neq 1$). Además, $(n\mathbb{Z}, +)$ es un subgrupo de $(\mathbb{Z}, +)$:

**(S1)** Sean $a, b \in n\mathbb{Z}$. Entonces $a = nk_1$ y $b = nk_2$, para algunos $k_1, k_2 \in \mathbb{Z}$. Por lo tanto,

$$a + b = nk_1 + nk_2 = n(k_1 + k_2) \in n\mathbb{Z}.$$

**(S2)** El conjunto $n\mathbb{Z}$ contiene a 0 porque $0 = n0$.

**(S3)** Si $a \in n\mathbb{Z}$, entonces $a = nk$, para algún $k \in \mathbb{Z}$, así que $-a = n(-k) \in n\mathbb{Z}$.

Un grupo $(G, \cdot)$ es *finito* si $G$ es un conjunto finito. Cuando $|G| = m$, podemos escribir una tabla, llamada la *tabla de Cayley* de $(G, \cdot)$, con $m$ filas y $m$ columnas, que determina completamente el comportamiento de la operación binaria del grupo. Para esto, ordenamos de manera arbitraria los elementos del grupo, $G = \{g_1, g_2, ..., g_m\}$, y escribimos $g_i \cdot g_j$ en la entrada $(i, j)$ de la tabla.

**Ejemplo 0.20 (enteros módulo $n$).** El par $(\mathbb{Z}_n, +)$ es un grupo abeliano finito (Ejercicio 0.41), donde $+$ es la suma módulo $n$. Si $n = 5$, el Cuadro 1 es la tabla de Cayley de $(\mathbb{Z}_5, +)$.

**Definición 0.21 (subgrupo cíclico generado por $g$).** Sea $(G, \cdot)$ un grupo. Definimos al *grupo cíclico generado por* $g \in G$ como el conjunto

$$\langle g \rangle = \{g^k : k \in \mathbb{Z}\},$$

donde $g^0 = e$ y

$$g^k = \begin{cases} \underbrace{g \cdot g \cdot g \ldots g}_{k \text{ veces}} & \text{para } k > 0, \\ \underbrace{g^{-1} \cdot g^{-1} \ldots g^{-1}}_{-k \text{ veces}} & \text{para } k < 0. \end{cases}$$

| $+$ | $[0]$ | $[1]$ | $[2]$ | $[3]$ | $[4]$ |
|---|---|---|---|---|---|
| $[0]$ | $[0]$ | $[1]$ | $[2]$ | $[3]$ | $[4]$ |
| $[1]$ | $[1]$ | $[2]$ | $[3]$ | $[4]$ | $[0]$ |
| $[2]$ | $[2]$ | $[3]$ | $[4]$ | $[0]$ | $[1]$ |
| $[3]$ | $[3]$ | $[4]$ | $[0]$ | $[1]$ | $[2]$ |
| $[4]$ | $[4]$ | $[0]$ | $[1]$ | $[2]$ | $[3]$ |

Cuadro 1: Tabla de Cayley de $(\mathbb{Z}_5, +)$

**Teorema 0.22 (subgrupo cíclico generado por $g$).** Sea $(G, \cdot)$ un grupo y $g \in G$. Entonces, $(\langle g \rangle, \cdot)$ es un subgrupo abeliano de $(G, \cdot)$.

**Demostración.** Usaremos el Teorema 0.18 del test del subgrupo.

**(S1)** Sean $g^k, g^s \in \langle g \rangle$, $k, s \in \mathbb{Z}$. Analizaremos varios casos. Si $k > 0$ y $s > 0$, entonces

$$g^k \cdot g^s = \underbrace{g \cdot g \dots g}_{k \text{ veces}} \cdot \underbrace{g \cdot g \dots g}_{s \text{ veces}} = \underbrace{g \cdot g \cdot g \dots g}_{k+s \text{ veces}} = g^{k+s} \in \langle g \rangle.$$

Si $k > 0$, $s < 0$, y $k > -s$, entonces

$$g^k \cdot g^s = \underbrace{g \cdot g \dots g}_{k \text{ veces}} \cdot \underbrace{g^{-1} \cdot g^{-1} \dots g^{-1}}_{-s \text{ veces}} = \underbrace{g \cdot g \dots g}_{k+s \text{ veces}} \cdot e \dots e = g^{k+s} \in \langle g \rangle.$$

Si $k > 0$, $s < 0$, y $k < -s$, entonces

$$g^k \cdot g^s = \underbrace{g \cdot g \dots g}_{k \text{ veces}} \cdot \underbrace{g^{-1} \cdot g^{-1} \dots g^{-1}}_{-s \text{ veces}} = e \dots e \cdot \underbrace{g^{-1} \cdot g^{-1} \dots g^{-1}}_{-s-k \text{ veces}} = g^{k+s} \in \langle g \rangle.$$

Si $k < 0$ y $s < 0$, entonces

$$g^k \cdot g^s = \underbrace{g^{-1} \cdot g^{-1} \dots g^{-1}}_{-k \text{ veces}} \cdot \underbrace{g^{-1} \cdot g^{-1} \dots g^{-1}}_{-s \text{ veces}} = \underbrace{g^{-1} \cdot g^{-1} \dots g^{-1}}_{-k-s \text{ veces}} = g^{k+s} \in \langle g \rangle.$$

Los otros casos (como $k < 0$ y $s > 0$, o $k = 0$) se analizan de manera similar. Por lo tanto, $g^k \cdot g^s = g^{k+s} \in \langle g \rangle$ para toda $k, s \in \mathbb{Z}$, lo que demuestra la cerradura en $\langle g \rangle$.

**(S2)** Por definición, $e = g^0 \in \langle g \rangle$.

**(S3)** El inverso de cualquier $g^k \in \langle g \rangle$ es $g^{-k}$, el cual es claramente un elemento de $\langle g \rangle$.

**(G4)** Sean $g^k, g^s \in \langle g \rangle$. Entonces $g^k \cdot g^s = g^{k+s} = g^{s+k} = g^s \cdot g^k$.

□

**Definición 0.23 (grupo cíclico).** Decimos que un grupo $(G, \cdot)$ es *cíclico* si existe $g \in G$ tal que $G = \langle g \rangle$.

**Ejemplo 0.24.** Para cualquier $n \in \mathbb{N}$, el grupo $(\mathbb{Z}_n, +)$ es cíclico. Demostraremos que $\mathbb{Z}_n = \langle [1] \rangle$. Sea $k \in \mathbb{Z}$. Siempre que la operación del grupo sea una suma, escribiremos $k \cdot [1]$ en lugar $[1]^k$, ya que $k \cdot [1]$ va más acorde con la notación aditiva. En este caso $k \cdot [1]$ significa:

$$k \cdot [1] = \begin{cases} \underbrace{[1] + [1] + \cdots + [1]}_{k \text{ veces}} & \text{para } k > 0, \\ \\ \underbrace{(-[1]) + (-[1]) + \cdots + (-[1])}_{-k \text{ veces}} & \text{para } k < 0. \end{cases}$$

Analizando los casos $k < 0$, $k = 0$ y $k > 0$ por separado, es sencillo comprobar que $k \cdot [1] = [k]$ para toda $k \in \mathbb{Z}$. Por lo tanto,

$$\langle [1] \rangle = \{k \cdot [1] : k \in \mathbb{Z}\} = \{[k] : k \in \mathbb{Z}\} = \mathbb{Z}_n.$$

Para el lector interesado en profundizar más en temas de teoría de grupos recomendamos el libro [3].

## 0.3. Campos

En esta sección definimos una nueva estructura algebraica que involucra dos operaciones binarias.

**Definición 0.25 (Campo).** Sea $F$ un conjunto no vacío. Sean $+$ y $\cdot$ dos operaciones binarias de $F$. La tríada $(F, +, \cdot)$ es un *campo* si se cumplen las siguientes propiedades:

**(C1)** $(F, +)$ es un grupo abeliano con identidad 0.

**(C2)** $(F \setminus \{0\}, \cdot)$ es un grupo abeliano con identidad 1.

**(C3)** *Distributividad.* Para toda $a, b, c \in F$,

$$a \cdot (b + c) = (a \cdot b) + (a \cdot c).$$

Las operaciones $+$ y $\cdot$ en un campo son llamadas "suma" y "multiplicación", respectivamente. Esto no significa que $+$ y $\cdot$ sean la suma y multiplicación usual de números; de hecho, el conjunto $F$ podría no contener números. Al elemento 0 se le llama *identidad aditiva* del campo, mientras que a 1 se le llama *identidad multiplicativa*.

Sea $(F, +, \cdot)$ un campo. Es costumbre denotar al inverso aditivo de $a \in F$ como $-a$ y, para simplificar notación, si $a, b \in F$, escribimos $a - b$ en lugar de $a + (-b)$. Sea $F^* = F \setminus \{0\}$ el conjunto de elementos del campo distintos de cero. Por la propiedad **(C2)**, cualquier $a \in F^*$ tiene inverso multiplicativo, al cual denotamos como $\frac{1}{a}$. Como ambas operaciones $+$ y $\cdot$ forman grupos abelianos, es claro que $a + b = b + a$ para toda $a, b \in F$, y que $a \cdot b = b \cdot a$ para toda $a, b \in F^*$.

**Lema 0.26 (propiedades básicas de campos).** Sea $(F, +, \cdot)$ un campo.

**(1)** *Multiplicación por* 0. Para cualquier $a \in F$, $0 \cdot a = a \cdot 0 = 0$.

**(2)** 0 *no tiene inverso multiplicativo.* No existe $a \in F$ tal que $a \cdot 0 = 1$.

**(3)** *No hay divisores de* 0. Si $a \cdot b = 0$, con $a, b \in F$, entonces $a = 0$ o $b = 0$.

**(4)** *Leyes de los signos.* Para toda $a, b \in F$, se cumple que

$$(-a) \cdot b = a \cdot (-b) = -(a \cdot b) \text{ y } (-a) \cdot (-b) = a \cdot b.$$

**Demostración.** Ejercicio 0.44. □

**Ejemplo 0.27 (campo con dos elementos).** Si $(F, +, \cdot)$ es un campo, sabemos que tiene una identidad aditiva 0 y una identidad multiplicativa 1. Obviamente, $0 \neq 1$ porque 1 pertenece al conjunto $F \setminus \{0\}$. Supongamos que $F = \{0, 1\}$. ¿Existe un campo $(\{0, 1\}, +, \cdot)$? Deduzcamos cómo deben ser sus operaciones. Por la definición de identidad aditiva, $0 + 1 = 1 + 0 = 1$ y $0 + 0 = 0$. Por la definición de la identidad multiplicativa, $1 \cdot 1 = 1$. Por el Lema 0.26 **(1)**, $0 \cdot 1 = 1 \cdot 0 = 0$. El único elemento que falta por determinar es $1+1$. Si $1+1 = 1$, por cancelación derecha tenemos que $1 = 0$, lo cual es una contradicción. Por lo tanto, $1 + 1 = 0$. Las tablas de sumar y multiplicar que hemos encontrado son:

| $+$ | 0 | 1 |
|---|---|---|
| 0 | 0 | 1 |
| 1 | 1 | 0 |

| $\times$ | 0 | 1 |
|---|---|---|
| 0 | 0 | 0 |
| 1 | 0 | 1 |

Ahora es fácil comprobar que $(\{0, 1\}, +, \cdot)$ satisface **(C1)**, **(C2)** y **(C3)**.

**Ejemplo 0.28 (números racionales).** Si $+$ es la suma usual y $\cdot$ es la multiplicación usual de números racionales, la tríada $(\mathbb{Q}, +, \cdot)$ es un campo. Veamos que se cumplen cada una de las propiedades.

**(C1)** Demostraremos que $(\mathbb{Q}, +)$ es un grupo abeliano:

**(G1)** Para toda $\frac{a_i}{b_i} \in \mathbb{Q}$,

$$\frac{a_1}{b_1} + \left(\frac{a_2}{b_2} + \frac{a_3}{b_3}\right) = \frac{a_1}{b_1} + \left(\frac{a_2 b_3 + a_3 b_2}{b_2 b_3}\right) = \frac{a_1 b_2 b_3 + a_2 b_1 b_3 + a_3 b_1 b_2}{b_1 b_2 b_3}.$$

Por otro lado,

$$\left(\frac{a_1}{b_1} + \frac{a_2}{b_2}\right) + \frac{a_3}{b_3} = \frac{a_1 b_2 + a_2 b_1}{b_1 b_2} + \frac{a_3}{b_3} = \frac{a_1 b_2 b_3 + a_2 b_1 b_3 + a_3 b_1 b_2}{b_1 b_2 b_3}.$$

Por lo tanto,

$$\frac{a_1}{b_1} + \left(\frac{a_2}{b_2} + \frac{a_3}{b_3}\right) = \left(\frac{a_1}{b_1} + \frac{a_2}{b_2}\right) + \frac{a_3}{b_3}.$$

**(G2)** $0 \in \mathbb{Q}$ es la identidad aditiva porque $0 + \frac{a}{b} = \frac{a}{b}$, para toda $\frac{a}{b} \in \mathbb{Q}$.

**(G3)** El inverso de cualquier $\frac{a}{b} \in \mathbb{Q}$ es $\frac{-a}{b} \in \mathbb{Q}$ ya que $\frac{a}{b} + \left(\frac{-a}{b}\right) = 0$.

**(G4)** Para toda $\frac{a_i}{b_i} \in \mathbb{Q}$,

$$\frac{a_1}{b_1} + \frac{a_2}{b_2} = \frac{a_1 b_2 + a_2 b_1}{b_1 b_2} = \frac{a_2}{b_2} + \frac{a_1}{b_1}.$$

**(C2)** Por el Ejemplo 0.16, $(\mathbb{Q}^\star, \cdot)$ es un grupo abeliano con identidad $\frac{1}{1}$.

**(C3)** Para toda $\frac{a_i}{b_i} \in \mathbb{Q}$,

$$\frac{a_1}{b_1} \cdot \left(\frac{a_2}{b_2} + \frac{a_3}{b_3}\right) = \frac{a_1}{b_1} \cdot \frac{a_2 b_3 + a_3 b_2}{b_2 b_3} = \frac{a_1(a_2 b_3 + a_3 b_2)}{b_1 b_2 b_3} = \frac{a_1 a_2 b_3 + a_1 a_3 b_2}{b_1 b_2 b_3}.$$

Por otro lado,

$$\left(\frac{a_1}{b_1} \cdot \frac{a_2}{b_2}\right) + \left(\frac{a_1}{b_1} \cdot \frac{a_3}{b_3}\right) = \frac{a_1 a_2}{b_1 b_2} + \frac{a_1 a_3}{b_1 b_3} = \frac{a_1 a_2 b_1 b_3 + a_1 a_3 b_1 b_2}{b_1 b_1 b_2 b_3} = \frac{a_1 a_2 b_3 + a_1 a_3 b_2}{b_1 b_2 b_3}.$$

Por lo tanto,

$$\frac{a_1}{b_1} \cdot \left(\frac{a_2}{b_2} + \frac{a_3}{b_3}\right) = \left(\frac{a_1}{b_1} \cdot \frac{a_2}{b_2}\right) + \left(\frac{a_1}{b_1} \cdot \frac{a_3}{b_3}\right).$$

**Ejemplo 0.29.** La tríada $(\mathbb{Z}, +, \cdot)$ no es un campo porque $(\mathbb{Z} \setminus \{0\}, \cdot)$ no es un grupo: ningún elemento $a \in \mathbb{Z}$, $a \neq \pm 1$, tiene inverso multiplicativo en $\mathbb{Z}$.

**Ejemplo 0.30 (campo de los números reales).** Sea $\mathbb{R}$ el conjunto de los números reales. La tríada $(\mathbb{R}, +, \cdot)$ es un campo, con identidad aditiva 0 e identidad multiplicativa 1, llamado el *campo de los números reales*.

**Ejemplo 0.31 (campo de los números complejos).** El conjunto de los *números complejos* es el producto cartesiano de los números reales:

$$\mathbb{C} = \{(x, y) : x, y \in \mathbb{R}\} = \mathbb{R} \times \mathbb{R}$$

La primera coordenada de $\mathbb{C}$ se llama *coordenada real*, mientras que la segunda se llama *coordenada imaginaria*. El número complejo $(x, y) \in \mathbb{C}$ es llamado *real puro* si $y = 0$, o *imaginario puro* si $x = 0$.

Para cualquier $(x_i, y_i) \in \mathbb{C}$, la *suma usual de números complejos* está definida como

$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2,\ y_1 + y_2) \in \mathbb{C},$$

mientras que la *multiplicación usual de números complejos* está definida como

$$(x_1, y_1) \cdot (x_2, y_2) = (x_1 x_2 - y_1 y_2,\ x_1 y_2 + x_2 y_1) \in \mathbb{C},$$

donde $x_i y_j$ representa la multiplicación usual de números reales.

Si $(x_1, 0)$ y $(x_2, 0)$ son reales puros, las operaciones definidas previamente coinciden con la suma y multiplicación usual de números reales:

$$(x_1, 0) + (x_2, 0) = (x_1 + x_2,\ 0)\,,$$
$$(x_1, 0) \cdot (x_2, 0) = (x_1 x_2,\ 0)\,.$$

Observemos que, para cualquier $(x, y) \in \mathbb{C}$,

$$(x, y) = (x, 0) \cdot (1, 0) + (y, 0) \cdot (0, 1)\,.$$

Para simplificar notación, identificamos a $(x, 0)$ y $(y, 0)$ con los números reales $x, y \in \mathbb{R}$, respectivamente. Si definimos

$$i = (0, 1) \in \mathbb{C},$$

podemos denotar al número complejo $(x, y)$ como

$$x + yi \in \mathbb{C}.$$

El imaginario puro $i = (0, 1)$ es llamado la *unidad imaginaria* y cumple que

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

al cual identificamos con $-1 \in \mathbb{R}$. Es por esta razón que comúnmente se dice que "$i$ es una raíz cuadrada de $-1$".

Con con esta nueva notación, la suma y multiplicación de números complejos puede escribirse como

$$(x_1 + y_1 i) + (x_2 + y_2 i) = x_1 + x_2 + (y_1 + y_2)\, i,$$
$$(x_1 + y_1 i) \cdot (x_2 + y_2 i) = x_1 x_2 - y_1 y_2 + (x_1 y_2 + x_2 y_1)\, i.$$

Usando la notación propuesta y las propiedades de los números reales, no es difícil demostrar que $(\mathbb{C}, +, \cdot)$ es un campo. Sin embargo, la demostración es algo laboriosa, así que se deja como ejercicio.

Un campo $(F, +, \cdot)$ es llamado un **campo finito** si $F$ es un conjunto finito.

**Ejemplo 0.32 ($\mathbb{Z}_p$).** Sea $p \in \mathbb{N}$ un número primo. Consideremos la tríada $(\mathbb{Z}_p, +, \cdot)$ donde $\mathbb{Z}_p = \{[0]\,, [1]\,, \ldots, [p-1]\}$, $+$ es la suma usual de clases módulo $p$, y $\cdot$ está definida como

$$[k] \cdot [s] = [ks] \in \mathbb{Z}_p,\ \text{donde } [k], [s] \in \mathbb{Z}_p.$$

Demostraremos que $(\mathbb{Z}_p, +, \cdot)$ es un campo finito.

**(C1)** Sabemos por el Ejercicio 0.41 que $(\mathbb{Z}_p, +)$ es un grupo abeliano con identidad $[0]$.

**(C2)** Por el Ejercicio 0.39, $\cdot$ es una operación binaria de $\mathbb{Z}_p$ bien definida. Demostraremos que $(\mathbb{Z}_p \setminus \{[0]\}\,, \cdot)$ cumple las propiedades **(G1)-(G4)**.

**(G1)** Para toda $[r], [k], [s] \in \mathbb{Z}_p \setminus \{[0]\}$,

$$[r] \cdot ([k] \cdot [s]) = [r\,(ks)] = [(rk)\,s] = ([r] \cdot [k]) \cdot [s]\,.$$

**(G2)** Para toda $[s] \in \mathbb{Z}_p \setminus \{[0]\}$, tenemos que $[1] \cdot [s] = [1s] = [s]$. Por lo tanto, $[1] \in \mathbb{Z}_p \setminus \{0\}$ es la identidad multiplicativa.

**(G3)** Sea $[s] \in \mathbb{Z}_p \setminus \{[0]\}$. Demostramos la existencia del inverso multiplicativo de $[s]$ usando el Lema de Bézout (Lema 4.23 en [2]). Como $[s] \neq [0]$, sabemos que $p$ no divide a $s$, y que $p$ sea primo implica que $\text{mcd}(s, p) = 1$. Por el Lema de Bézout, existen enteros $x, y \in \mathbb{Z}$ tales que

$$1 = sx + py.$$

Claramente, $p$ no divide a $x$, ya que de lo contrario $p \mid (sx + py) = 1$, lo cual es imposible. De esta forma, $[x] \in \mathbb{Z}_p \setminus \{[0]\}$ es el inverso multiplicativo de $[s]$ porque

$$[s] \cdot [x] = [sx] = [1 - py] = [1]\,.$$

**(G4)** Para toda $[k], [s] \in \mathbb{Z}_p \setminus \{[0]\}$,

$$[k] \cdot [s] = [ks] = [sk] = [s] \cdot [k]\,.$$

**(C3)** Para toda $[r], [k], [s] \in \mathbb{Z}_p$,

$$\begin{aligned}[r] \cdot ([k] + [s]) &= [r] \cdot [k + s] = [r(k + s)] \\ &= [rk + rs] = [rk] + [rs] \\ &= [r] \cdot [k] + [r] \cdot [s].\end{aligned}$$

El siguiente teorema establece bajo qué circunstancias es $\mathbb{Z}_m$ un campo.

**Teorema 0.33.** Sea $m \in \mathbb{N}$. La tríada $(\mathbb{Z}_m, +, \cdot)$ es un campo si y sólo si $m$ es un número primo.

**Demostración.** Demostraremos cada implicación.

**(⇒)** Supongamos $(\mathbb{Z}_m, +, \cdot)$ es un campo. Por reducción al absurdo, supongamos que $m$ no es un número primo. Entonces, $m = ks$, donde $k, s \in \mathbb{N}$, $1 < k, s < m$. Luego,

$$[0] = [m] = [ks] = [k] \cdot [s],$$

donde $[k] \neq [0]$ y $[s] \neq [0]$ porque $m \nmid k$ y $m \nmid s$. Sin embargo, esto contradice el Lema 0.26 **(3)**, ya que $[k] \cdot [s] = [0]$ implica $[k] = [0]$ o $[s] = [0]$. Por lo tanto, $m$ debe ser un número primo.

**(⇐)** Si $m = p$ es un número primo, sabemos que $(\mathbb{Z}_p, +, \cdot)$ es un campo por el Ejemplo 0.32.

□

Existen otros ejemplos de campos finitos además de $\mathbb{Z}_p$, pero, en general, su estructura es más difícil de describir. El siguiente teorema, del cual omitimos su demostración, establece todos los posibles tamaños de los campos finitos.

**Teorema 0.34 (campos finitos).** Para cualquier número primo $p$ y cualquier $k \in \mathbb{N}$, existe esencialmente un único campo finito con $p^k$ elementos, al cual denotamos[1] como $\mathrm{GF}(p^k)$. No existen campos finitos de otros tamaños.

Para cualquier número primo $p$ se cumple que

$$\mathrm{GF}(p) = \mathbb{Z}_p.$$

Sin embargo, $\mathrm{GF}(p^k) \neq \mathbb{Z}_{p^k}$ porque $\mathbb{Z}_{p^k}$ no es un campo. Por ejemplo, el Teorema 0.34 establece que hay un campo finito $\mathrm{GF}(4)$ con 4 elementos (el cual no puede ser igual a $\mathbb{Z}_4$) ¿Cómo son entonces las operaciones binarias $+$ y $\cdot$ en $\mathrm{GF}(4)$? En el Ejercicio 0.48 se pide encontrar tablas de sumar y multiplicar para estas operaciones.

**Observación 0.35.** A partir de ahora, haremos dos simplificaciones en nuestra notación:

**(1)** Denotaremos a los elementos de $\mathbb{Z}_m = \{[0], [1], \ldots, [m-1]\}$ simplemente como números; es decir, en lugar de escribir $[s]$ escribiremos $s$. Sin embargo, hay que tener siempre presente que estos "números" en $\mathbb{Z}_m$ son en realidad clases de equivalencia módulo $m$.

**(2)** Siempre que las operaciones binarias estén claramente establecidas por el contexto, denotaremos a los campos, y a los grupos, con el conjunto que contiene a sus elementos; por ejemplo, si $(F, +, \cdot)$ es un campo y $(G, \cdot)$ es un grupo, simplemente escribiremos $F$ y $G$, respectivamente.

***Palabras clave:*** *relación de equivalencia, clase de equivalencia, operacion binaria, grupo, grupo abeliano, campo, campo finito.*

[1] Las iniciales GF significan *Galois Field* y hacen referencia al matemático francés Évariste Galois (1811–1832).

## 0.4. Ejercicios de Conceptos Preliminares

**Ejercicio 0.36.** Demuestra que la relación de congruencia módulo $n$ cumple la propiedad transitiva.

**Ejercicio 0.37.** Demuestra el Lema 0.7 sobre las propiedades básicas de las clases de equivalencia.

**Ejercicio 0.38.** Demuestra que la función $\oplus$ definida como

$$\frac{a_1}{b_1} \oplus \frac{a_2}{b_2} = \frac{a_1 + a_2}{b_1 + b_2},$$

donde $\frac{a_i}{b_i} \in \mathbb{Q}$, no es una operación binaria de $\mathbb{Q}$.

**Ejercicio 0.39.** Sea $n \in \mathbb{N}$, $n \neq 0$. Si $[k], [s] \in \mathbb{Z}_n$, demuestra que la operación binaria $[k] \times [s] = [ks]$ está bien definida (es decir, que no depende de los representantes $k$ y $s$).

**Ejercicio 0.40.** Demuestra el Lema 0.14 sobre propiedades básicas de grupos.

**Ejercicio 0.41.** Sea $n \in \mathbb{N}$, $n \neq 0$. Demuestra que $(\mathbb{Z}_n, +)$ es un grupo abeliano finito.

**Ejercicio 0.42.** Sea $+$ la suma usual de números enteros.

**(1)** Demuestra que $(\mathbb{Z}, +)$ es un grupo cíclico.

**(2)** Si $n \in \mathbb{N}$ y $n\mathbb{Z} = \{nk : k \in \mathbb{Z}\}$, demuestra que $(n\mathbb{Z}, +)$ es un subgrupo cíclico de $\mathbb{Z}$.

**Ejercicio 0.43.** Encuentra todos los subgrupos cíclicos de $\mathbb{Z}_7$, $\mathbb{Z}_6$, $\mathbb{Z}_{10}$ y $\mathbb{Z}_{11}$.

**Ejercicio 0.44.** Demuestra el Lema 0.26 sobre propiedades básicas de campos.

**Ejercicio 0.45.** Demuestra que la tríada $(\mathbb{C}, +, \cdot)$ definida en el Ejemplo 0.31 es un campo.

**Ejercicio 0.46.** Sean $+$ and $\odot$ las siguientes operaciones binarias de $\mathbb{R} \times \mathbb{R}$:

$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2),$$
$$(x_1, y_1) \odot (x_2, y_2) = (x_1 x_2, y_1 y_2).$$

Demuestra que $(\mathbb{R} \times \mathbb{R}, +, \odot)$ no es un campo.

**Ejercicio 0.47.** Equipados con las operaciones binarias usuales de números (enteros, racionales, reales, complejos, clases módulo $m$) en cada caso, ¿cuáles de los siguientes conjuntos son campos? Justifica tu respuesta detalladamente.

**(1)** $\mathbb{N}$.

**(2)** $\mathbb{R} \setminus \mathbb{Q}$.

**(3)** $\mathbb{Z}_{14}$.

**(4)** $\mathbb{Z}_{17}$.

**(5)** $\mathbb{Q}(i) = \{x + yi : x, y \in \mathbb{Q}\}$.

**(6)** $\mathbb{Q}(\sqrt{6}) = \{x + y\sqrt{5} : x, y \in \mathbb{Q}\}$.

**(7)** $3\mathbb{Z} = \{3k : k \in \mathbb{Z}\}$.

**(8)** $\mathbb{Z}_1$.

**Ejercicio 0.48.** Sea $\mathrm{GF}(4) = \{0, 1, \alpha, \beta\}$ el campo finito con 4 elementos. Encuentra las tablas de sumar y multiplicar para las operaciones binarias de este campo.

# 1

# Espacios vectoriales y sus propiedades

Hasta ahora hemos estudiado grupos y campos. Ahora vamos a estudiar otra estructura conocida como *espacio vectorial*.

**Definición 1.1 (espacio vectorial).** Un *espacio vectorial sobre un campo* $F$ es una tríada $(V, +, \cdot)$, donde:

**(a)** $V$ es un conjunto no vacío;

**(b)** $+ : V \times V \to V$ es una operación binaria de $V$ llamada *suma*;

**(c)** $\cdot : F \times V \to V$ es una función llamada *multiplicación escalar*.

Además, deben satisfacerse las siguientes propiedades:

**(EV1)** El par $(V, +)$ es un grupo abeliano con identidad $\mathbf{0}$.

**(EV2)** La multiplicación escalar satisface: $\forall \alpha, \beta \in F$ y $\forall v, w \in V$,

- **(1)** $\alpha \cdot (v + w) = (\alpha \cdot v) + (\alpha \cdot w)$;
- **(2)** $(\alpha + \beta) \cdot v = (\alpha \cdot v) + (\beta \cdot v)$;
- **(3)** $(\alpha \cdot \beta) \cdot v = \alpha \cdot (\beta \cdot v)$;
- **(4)** $1 \cdot v = v$, donde 1 es la identidad multiplicativa de $F$.

Para simplificar, nos referimos a un espacio vectorial $(V, +, \cdot)$ simplemente como $V$, y abusando de la notación, escribimos $V = (V, +, \cdot)$. Los elementos de un espacio vectorial $V$ se llaman *vectores*, mientras que los del campo $F$ se llaman *escalares*. La suma de $V$ asigna a cada par $(u, v)$ de vectores en $V$ un vector $u + v \in V$. La multiplicación escalar asigna a cada par $(\alpha, v) \in F \times V$ un vector $\alpha \cdot u$ en $V$. A partir de ahora, escribiremos simplemente $\alpha v$ en lugar de $\alpha \cdot v$ para denotar la multiplicación escalar.

Si $F$ es el campo $\mathbb{R}$ de los números reales, $V$ es llamado *espacio vectorial real*; similarmente si $F$ es $\mathbb{Q}$, o $F$ es $\mathbb{C}$, hablaremos del *espacio vectorial racional*, o *espacio vectorial complejo*, respectivamente.

**Observación 1.2.** Usaremos letras Latinas $v, w, x$ etc. para denotar vectores, es decir elementos de $V$ y letras Griegas $\alpha, \beta, \gamma$ etc. para denotas escalares, es decir elementos del campo $F$.

**Observación 1.3.** No debería haber ninguna confusión sobre el uso de la palabra "vector". En otros textos, un "vector" es un objeto que posee magnitud, dirección y sentido, y es representado geométricamente por un segmento de línea; para nostros, un "vector" es simplemente un elemento abstracto de un espacio vectorial.

**Observación 1.4.** En la definición anterior, denotamos la suma de vectores con el símbolo $+$. Este símbolo también denota la suma en el campo $F$, es decir, la suma de escalares. No debe haber confusión aunque se ha utilizado el mismo símbolo para indicar esta operación en distintos conjuntos. De esta manera si $v, w \in V$, entonces $v + w$ representa suma en $V$, es decir la suma de vectores y $v + w \in V$. Si $\alpha, \beta \in F$, entonces $\alpha + \beta$ representa la suma de escalares, es decir suma en el campo $F$ y $\alpha + \beta \in F$. Tampoco debe haber confusión al multiplicar escalares. Si $\alpha, \beta \in F$, entonces $\alpha\beta$ representa la multiplicación en $F$ y $\alpha\beta \in F$. Si $\alpha \in F$ y $v \in V$, entonces $\alpha v$ representa la multiplicación escalar y $\alpha v \in V$.

**Observación 1.5.** Un espacio vectorial tiene dos tipos de elemento cero. Uno es el cero del grupo $(V, +)$, al que simbolizamos por $\mathbf{0}$ (o por $\vec{\mathbf{0}}$) y llamamos *vector cero*; el otro es el cero del campo $F$, al que simbolizamos simplemente por 0 y llamamos *cero escalar*.

Antes de examinar algunos ejemplos particulares de espacios vectoriales, demostraremos algunas propiedades generales que siempre deben cumplir. Hacemos énfasis que, en las siguientes proposiciones, no hablamos de un espacio vectorial en particular, así que no podemos asumir que los vectores tengan alguna forma particular. Por el contrario, las siguientes proposiciones enuncian propiedades de espacios vectoriales *generales* y *abstractos*. No sabemos nada sobre la naturaleza de los vectores en estos espacios vectoriales, pero sí sabemos una cosa: satisfacen las propiedades (a), (b), (c), (EV1) y (EV2) enunciadas en la definición de espacio vectorial. Son sólo estas propiedades (junto con el hecho de que $F$ es un campo) las que debemos usar para demostrar las propiedades básicas que se enuncian a continuación.

**Proposición 1.6 (propiedades de espacios vectoriales).** Sea $V$ un espacio vectorial sobre un campo $F$.

**(1)** $0v = \mathbf{0}$, $\forall v \in V$.

**(2)** $\alpha\mathbf{0} = \mathbf{0}$, $\forall \alpha \in F$.

**(3)** Si $\alpha v = \mathbf{0}$, entonces $\alpha = 0$ o $v = \mathbf{0}$.

**(4)** $(-\alpha)\, v = \alpha\,(-v) = -(\alpha v)$, $\forall \alpha \in F$, $v \in V$.

**Demostración.** Demostraremos **(1)**, **(2)** y **(3)**, mientras que **(4)** se deja como ejercicio.

**(1)** Tenemos que

$$\begin{aligned} 0v &= (0+0)\,v, && [\because\ 0 = 0+0] \\ &= 0v + 0v, && [\because\ (\mathbf{EV2.2})] \\ \mathbf{0} + 0v &= 0v + 0v. && [\because\ \mathbf{0} \text{ es la identidad aditiva en } V] \end{aligned}$$

Por la propiedad de cancelación derecha del grupo $(V, +)$, tenemos que $\mathbf{0} = 0v$.

**(2)** Observemos que

$$\begin{aligned} \alpha\mathbf{0} &= \alpha\,(\mathbf{0}+\mathbf{0})\,, && [\because\ \mathbf{0} = \mathbf{0}+\mathbf{0}] \\ &= \alpha\mathbf{0} + \alpha\mathbf{0}, && [\because\ (\mathbf{EV2.1})] \\ \mathbf{0} + \alpha\mathbf{0} &= \alpha\mathbf{0} + \alpha\mathbf{0}. && [\because\ \mathbf{0} \text{ es la identidad aditiva en } V] \end{aligned}$$

Por la propiedad de cancelación derecha del grupo $(V, +)$, tenemos que $\mathbf{0} = \alpha\mathbf{0}$.

**(3)** Si $\alpha = 0$, tenemos que $\alpha v = \mathbf{0}$ como se demostró en el punto **(1)**. Supongamos que $\alpha v = \mathbf{0}$ y $\alpha \neq 0$; demostraremos que $v = \mathbf{0}$. Debido a que $\alpha \in F \setminus \{0\}$ y $F$ es un campo, entonces el inverso multiplicativo $\alpha^{-1}$ existe. De esta manera,

$$\begin{aligned} \alpha^{-1}(\alpha v) &= \alpha^{-1}\mathbf{0}, \\ \left(\alpha^{-1}\alpha\right) v &= \mathbf{0}, \\ 1v &= \mathbf{0}, \\ v &= \mathbf{0}. \end{aligned}$$

**(4)** *Ejercicio* 1.20

□

**Proposición 1.7.** Sea $V$ un espacio vectorial sobre un campo $F$.

**(1)** Sean $\alpha, \beta \in F$ y $v \in V \setminus \{\mathbf{0}\}$. Si $\alpha v = \beta v$, entonces $\alpha = \beta$.

**(2)** Sean $v, w \in V$ y $\alpha \in F \setminus \{0\}$. Si $\alpha v = \alpha w$, entoces $v = w$.

**Demostración.**

**(1)** Si $\alpha v = \beta v$, entonces

$$\begin{aligned} \alpha v - \beta v &= \beta v - \beta v, \\ \alpha v - \beta v &= \mathbf{0}, \\ (\alpha - \beta)\, v &= \mathbf{0}, \end{aligned}$$

Como $v \neq \mathbf{0}$, la Proposición 1.6 **(3)** implica que

$$\begin{aligned} \alpha - \beta &= 0, \\ \alpha &= \beta. \end{aligned}$$

**(2)** *Ejercicio* 1.21.

□

## 1.1. Ejemplos de espacios vectoriales

**Ejemplo 1.8 (el campo $F$ sobre si mismo).** Sea $F$ un campo. Como consecuencia inmediata de la definición de espacio vectorial, mostraremos que $F$ es un espacio vectorial sobre $F$ (es decir, ambos conjuntos de vectores y escalares son iguales a $F$). La suma de vectores y multiplicación escalar en este caso coinciden con la suma y multiplicación de elementos del campo $F$. Por la definición de campo, $(F,+)$ es un grupo abeliano (**(C1)**). Además si $\alpha,\beta\in F$ son escalares cualesquiera, y $v,w\in F$ son vectores cualesquiera, entonces

$$\alpha\left(v+w\right)=\alpha v+\alpha w \quad \text{y} \quad \left(\alpha+\beta\right)v=\alpha v+\beta v,$$

debido a que $(F\backslash 0,\cdot)$ es un grupo abeliano (**C2**) y $F$ satisface la ley distributiva (**C3**). Además,

$$\left(\alpha\beta\right)v=\alpha\left(\beta v\right),$$

porque la multiplicación de $F$ es asociativa. Finalmente, si 1 es la identidad multiplicativa de $F$ y $v\in F$, entonces

$$1v=v.$$

Por lo tanto, $F$ es un espacio vectorial sobre sí mismo.

**Observación 1.9.** Si $F$ es un campo cualquiera, entonces $F$ es un espacio vectorial sobre cualquier subcampo $R$ de $F$.

**Observación 1.10.** Si $\mathbb{C}$ es el campo de los números complejos y $\mathbb{R}$ es el campo de los reales, entonces $\mathbb{C}$ es un espacio vectorial sobre $\mathbb{R}$ porque $\mathbb{R}$ es un subcampo de $\mathbb{C}$. Pero $\mathbb{R}$ no es un espacio vectorial sobre $\mathbb{C}$ porque $\mathbb{R}$ no es cerrado con respecto ala multiplicación escalar. Por ejemplo, $2\in\mathbb{R}$ y $\left(3+4i\right)2\notin\mathbb{R}$.

**Ejemplo 1.11 ($\mathbb{R}^n$).** El ejemplo básico más importante de espacio vectorial es, sin duda, el espacio vectorial real denotado por $\mathbb{R}^n=(\mathbb{R}^n,+,\cdot)$, donde $n\in\mathbb{N}$, y los elementos que conforman la tríada están definidos de la siguiente manera:

**(a)** $\mathbb{R}^n$ es la $n$ potencia cartesiana del campo real

$$\mathbb{R}^n=\underbrace{\mathbb{R}\times\mathbb{R}\times\cdots\times\mathbb{R}}_{n\text{ veces}}=\left\{(x_1,\ldots,x_n):x_i\in\mathbb{R},\forall i\right\}.$$

**(b)** Para $v=(v_1,\ldots,v_n),w=(w_1,\ldots,w_n)\in\mathbb{R}^n$, la suma de vectores está definida como

$$\begin{aligned} v+w &= (v_1,\ldots,v_n)+(w_1,\ldots,w_n)\\ &= (v_1+w_1,\ldots,v_n+w_n). \end{aligned}$$

**(c)** Para $\alpha \in \mathbb{R}$ y $v = (v_1, \ldots, v_n) \in \mathbb{R}^n$, la multiplicación escalar está definida como

$$\begin{aligned} \alpha v &= \alpha\,(v_1, \ldots, v_n) \\ &= (\alpha v_1, \ldots, \alpha v_n)\,. \end{aligned}$$

Demostraremos que $\mathbb{R}^n$ es un espacio vectorial sobre $\mathbb{R}$. Sean

$$u = (u_1, \ldots, u_n), v = (v_1, \ldots, v_n), w = (w_1, \ldots, w_n) \in \mathbb{R}^n,$$

vectores arbitrarios, y $\alpha, \beta \in \mathbb{R}$ escalares arbitrarios.

**(EV1)** Comprobaremos que $(\mathbb{R}^n, +)$ es un grupo abeliano.

**(G0)** *Cerradura*: claramente, por definición $v + w \in \mathbb{R}^n$.

**(G1)** *Asociatividad de la suma*:

$$\begin{aligned} u + (v + w) &= (u_1, \ldots, u_n) + [(v_1, \ldots, v_n) + (w_1, \ldots, w_n)] \\ &= (u_1, \ldots, u_n) + (v_1 + w_1, \ldots, v_n + w_n) \\ &= (u_1 + [v_1 + w_1], \ldots, u_n + [v_n + w_n]) \\ &= ([u_1 + v_1] + w_1, \ldots, [u_n + v_n] + w_n) \\ &= (u_1 + v_1, \ldots, u_n + v_n) + (w_1, \ldots, w_n) \\ &= [(u_1, \ldots, u_n) + (v_1, \ldots, v_n)] + (w_1, \ldots, w_n) \\ &= (u + v) + w. \end{aligned}$$

**(G2)** *Identidad aditiva en* $\mathbb{R}^n$: el vector $\mathbf{0} = (0, \ldots, 0) \in \mathbb{R}^n$ satisface que

$$\begin{aligned} v + \mathbf{0} &= (v_1, \ldots, v_n) + (0, \ldots, 0) \\ &= (v_1 + 0, \ldots, v_n + 0) \\ &= (v_1 \ldots, v_n) = v. \end{aligned}$$

Por lo tanto, $\mathbf{0} = (0, \ldots, 0)$ es la identidad aditiva en $\mathbb{R}^n$.

**(G3)** *Inversos aditivos*: el vector $-v = (-v_1, \ldots, -v_n) \in \mathbb{R}^n$ es el inverso aditivo de $v$:

$$\begin{aligned} (-v) + v &= (-v_1, \ldots, -v_n) + (v_1, \ldots, v_n) \\ &= (-v_1 + v_1, \ldots, -v_n + v_n) \\ &= (-v_1, \ldots, -v_n) \\ &= (0, \ldots, 0) = \mathbf{0}. \end{aligned}$$

**(G4)** *Conmutatividad de la suma*:

$$\begin{aligned} v + w &= (v_1, \ldots, v_n) + (w_1, \ldots, w_n) \\ &= (v_1 + w_1, \ldots, v_n + w_n) \\ &= (w_1 + v_1, \ldots, w_n + v_n) \\ &= (w_1, \ldots, w_n) + (v_1, \ldots, v_n) \\ &= w + v. \end{aligned}$$

**(EV2)** La multiplicación escalar satisface cada punto de la definición:

**(1)**

$$\begin{aligned}
\alpha(v+w) &= \alpha(v_1+w_1,\ldots,v_n+w_n)\\
&= (\alpha[v_1+w_1],\ldots,\alpha[v_n+w_n])\\
&= (\alpha v_1+\alpha w_1,\ldots,\alpha v_n+\alpha w_n)\\
&= (\alpha v_1,\ldots,\alpha v_n)+(\alpha w_1,\ldots,\alpha w_n)\\
&= \alpha(v_1,\ldots,v_n)+\alpha(w_1,\ldots,w_n)\\
&= (\alpha v)+(\alpha w).
\end{aligned}$$

**(2)**

$$\begin{aligned}
(\alpha+\beta)v &= ([\alpha+\beta]v_1,\ldots,[\alpha+\beta]v_n)\\
&= (\alpha v_1+\beta v_1,\ldots,\alpha v_n+\beta v_n)\\
&= (\alpha v_1,\ldots,\alpha v_n)+(\beta v_1,\ldots,\beta v_n)\\
&= \alpha(v_1,\ldots,v_n)+\beta(v_1,\ldots,v_n)\\
&= \alpha v+\beta v.
\end{aligned}$$

**(3)**

$$\begin{aligned}
(\alpha\beta)v &= ([\alpha\beta]v_1,\ldots,[\alpha\beta]v_n)\\
&= (\alpha[\beta v_1],\ldots,\alpha[\beta v_n])\\
&= \alpha(\beta v_1,\ldots,\beta v_n)\\
&= \alpha(\beta(v_1,\ldots,v_n))\\
&= \alpha(\beta v).
\end{aligned}$$

**(4)**

$$\begin{aligned}
1v &= 1(v_1,\ldots,v_n)\\
&= (1v_1,\ldots,1v_n)\\
&= (v_1,\ldots,v_n)\\
&= v.
\end{aligned}$$

**Observación 1.12.** El ejemplo anterior puede generalizarse a cualquier campo $F$ para formar el espacio vectorial $F^n$.

Una *matriz* con entradas en un campo $F$ es un arreglo rectangular de elementos de $F$. Típicamente, denotamos a cada elemento de una matriz como $a_{i,j} \in F$ (o cualquier otra letra con el subíndice $i,j$), donde $i$ representa el número de renglón donde se encuentra el elemento y $j$ representa el número de la columna. Decimos que una matriz es de $n \times m$ si tiene $n$ renglones y $m$

columnas. Una matriz genérica de $n \times m$ es la siguiente:

$$\begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \dots & a_{1,m} \\ a_{2,1} & a_{2,2} & a_{2,3} & \dots & a_{2,m} \\ a_{3,1} & a_{3,2} & a_{3,3} & \dots & a_{3,m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & a_{n,3} & \dots & a_{n,m} \end{pmatrix}.$$

Para simplificar, denotamos a la matriz de arriba simplemente como $(a_{i,j})$. Cuando $n = m$, decimos que la matriz es *cuadrada*. A los elementos $a_{i,i}$ de la matriz se les llama elementos *diagonales*.

**Ejemplo 1.13 (Espacio de las matrices).** Sea $F$ un campo. La tríada $M_{n\times m}(F) = (M_{n\times m}(F), +, \cdot)$ es un espacio vectorial sobre $F$ donde los elementos que conforman la tríada están definidos de la siguiente manera:

**(a)** $M_{n\times m}(F)$ es el conjunto de todas las matrices de $n \times m$ con entradas en $F$.

**(b)** La suma de dos matrices de $n \times m$ se realiza sumando respectivamente cada una de las entradas de ambas matrices:

$$(a_{i,j}) + (b_{i,j}) = (a_{i,j} + b_{i,j}).$$

**(c)** La multiplicación escalar de $\alpha \in \mathbb{F}$ por una matriz se realiza multiplicando todos los elementos de la matriz por $\alpha$:

$$\alpha(a_{i,j}) = (\alpha a_{i,j}).$$

Es rutinatio verificar que $M_{n\times m}(F)$ es un espacio vectorial sobre $F$.

**Ejemplo 1.14 (Espacio de las funciones).** Sea $S$ un conjunto no vacío y $F$ un campo. Consideremos la tríada $F^S = (F^S, +, \cdot)$ donde:

**(a)** $F^S$ es el conjunto de las funciones con dominio $S$ y codominio $F$

$$F^S = \{f : S \to F : f \text{ es una función}\}.$$

**(b)** Para $f, g \in F^S$, la suma $(f + g)$ es la función dada por

$$(f + g)(x) = f(x) + g(x), \quad \forall x \in S.$$

**(c)** Para $f$ y $\alpha \in F$, la multiplicación escalar $\alpha f$ es la función dada por

$$(\alpha f)(x) = \alpha f(x), \quad \forall x \in S.$$

Demostraremos que $F^S$ es un espacio vectorial sobre $F$. Sean $f, g, h \in F^S$ funciones arbitrarias y $\alpha, \beta \in F$ escalares arbitrarios.

**(EV1)** $\left(F^S, +\right)$ es un grupo abeliano.

**(G0)** *Cerradura*: por definición, $f + g \in F^S$.

**(G1)** *Asociatividad de la suma*: para toda $x \in S$,

$$\begin{aligned}
[(f+g)+h](x) &= (f+g)(x) + h(x) \\
&= [f(x) + g(x)] + h(x) \\
&= f(x) + [g(x) + h(x)] \\
&= f(x) + (g+h)(x) \\
&= [f + (g+h)](x).
\end{aligned}$$

Como dos funciones son iguales si y sólo si coinciden en todos sus valores, tenemos que

$$(f+g)+h = f+(g+h).$$

**(G2)** *Identidad aditiva*: definamos la *función cero* como $i_0 : S \to F$ como $i_0(x) = 0$, $\forall x \in S$. Para toda $x \in S$,

$$\begin{aligned}
(f+i_0)(x) &= f(x) + i_0(x) \\
&= f(x) + 0 \\
&= f(x).
\end{aligned}$$

Por lo tanto

$$f + i_0 = f.$$

Esto demuestra que $i_0$ es la identidad aditiva en $F^S$.

**(G3)** *Inversos aditivos*: Definamos la función $-f : S \to F$ por

$$(-f)(x) = -f(x), \quad \forall x \in S.$$

Entonces, para toda $x \in S$,

$$\begin{aligned}
[f + (-f)](x) &= f(x) + (-f)(x) \\
&= f(x) - f(x) \\
&= 0 \\
&= i_0(x).
\end{aligned}$$

Por lo tanto,

$$f + (-f) = i_0,$$

así que $-f$ es el inverso aditivo de $f$.

**(G4)** *Conmutatividad de la suma*: para toda $x \in S$,

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

Por lo tanto ,

$$f + g = g + f.$$

**(EV2)** La multiplicación escalar satisface cada punto de la definición:

**(1)** Para toda $x \in S$,

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

Por lo tanto,

$$\alpha(f+g)=\alpha f+\alpha g.$$

**(2)** Para toda $x \in S$,

$$\begin{aligned}[(\alpha+\beta)f](x) &= (\alpha+\beta)f(x) \\ &= \alpha f(x)+\beta f(x) \\ &= (\alpha f)(x)+(\beta f)(x) \\ &= (\alpha f+\beta f)(x).\end{aligned}$$

Por lo tanto,

$$(\alpha+\beta)f=\alpha f+\beta f.$$

**(3)** Para toda $x \in S$,

$$\begin{aligned}[(\alpha\beta)f](x) &= (\alpha\beta)f(x) \\ &= \alpha[\beta f(x)] \\ &= \alpha[(\beta f)(x)] \\ &= [\alpha(\beta f)](x).\end{aligned}$$

Por lo tanto,

$$(\alpha\beta)f=\alpha(\beta f).$$

**(4)** Para toda $x \in S$,

$$\begin{aligned}(1f)(x) &= 1f(x) \\ &= f(x).\end{aligned}$$

Por lo tanto,

$$1f=f.$$

Por lo tanto, $F^S$ es un espacio vectorial sobre $F$.

**Ejemplo 1.15 ($F[x]$).** Sea $F$ un campo y consideremos la tríada $F[x]=(F[x],+,\cdot)$ donde:

**(a)** $F[x]$ es el conjunto de todos los polinomios en la variable $x$ con coeficientes en $F$,

$$F[x] = \{a_0 + a_1x + \ldots + a_nx^n : a_i \in F,\ n \in \mathbb{N}\} = \left\{\sum_{i=0}^{n} a_ix^i : a_i \in F, n \in \mathbb{N}\right\}.$$

**(b)** Sean $p(x) = a_0 + a_1x + \ldots + a_nx^n = \sum_{i=0}^{n} a_ix^i$ y $q(x) = b_0 + b_1x + \ldots + b_mx^m = \sum_{i=0}^{m} b_ix^i$ elementos de $F[x]$. Sin perder generalidad supongamos que $n \geq m$. La suma de polinomios está definida como

$$\begin{aligned} p(x) + q(x) &= (a_0 + \ldots + a_nx^n) + (b_0 + \ldots + b_mx^m) \\ &= (a_0 + b_0) + \ldots + (a_m + b_m)x^m + a_{m+1}x^{m+1} + \ldots + a_nx^n \\ &= \sum_{i=0}^{m}(a_i + b_i)x^i + \sum_{i=m+1}^{n} a_ix^i. \end{aligned}$$

**(c)** Sean $p(x) = a_0 + \ldots + a_nx^n \in F[x]$ y $\alpha \in F$. La multiplicación por escalar está definida como

$$\alpha p(x) = \alpha(a_0 + a_1x + \ldots + a_nx^n) = \alpha\sum_{i=0}^{n} a_ix^i = \sum_{i=0}^{n} \alpha a_ix^i.$$

Demostremos que $F[x]$ es un espacio vectorial sobre $F$. Sean

$$p(x) = \sum_{i=0}^{n} a_ix^i,\ \text{y}\ q(x) = \sum_{i=0}^{m} b_ix^i,$$

elementos arbitrarios de $F[x]$, con $n \geq m$, y sean $\alpha, \beta \in F$ escalares arbitrarios.

**(EV1)** Demostrar que $(F[x], +)$ es un grupo abeliano se deja como ejercicio (*Ejercicio* 1.19).

**(EV2)** La multiplicación escalar satisface cada punto de la definición:

**(1)**

$$\begin{aligned}\alpha\left[p\left(x\right)+q\left(x\right)\right] &= \alpha\left[\sum_{i=0}^{n} a_i x^i + \sum_{i=0}^{m} b_i x^i\right] \\ &= \alpha\left[\sum_{i=0}^{m}(a_i+b_i)x^i + \sum_{i=m+1}^{n} a_i x^i\right] \\ &= \sum_{i=0}^{m}\alpha(a_i+b_i)x^i + \sum_{i=m+1}^{n}\alpha a_i x^i \\ &= \sum_{i=0}^{m}(\alpha a_i+\alpha b_i)x^i + \sum_{i=m+1}^{n}\alpha a_i x^i \\ &= \sum_{i=0}^{n}\alpha a_i x^i + \sum_{i=0}^{m}\alpha b_i x^i \\ &= \alpha p\left(x\right)+\alpha q\left(x\right).\end{aligned}$$

**(2)**

$$\begin{aligned}\left(\alpha+\beta\right)p\left(x\right) &= \left(\alpha+\beta\right)\sum_{i=0}^{n} a_i x^i \\ &= \sum_{i=0}^{n}(\alpha+\beta)a_i x^i \\ &= \sum_{i=0}^{n}(\alpha a_i+\beta a_i)x^i \\ &= \sum_{i=0}^{n}\alpha a_i x + \sum_{i=0}^{n}\beta a_i x^i \\ &= \alpha p\left(x\right)+\beta p\left(x\right).\end{aligned}$$

**(3)**

$$\begin{aligned}\left(\alpha\beta\right)p\left(x\right) &= \left(\alpha\beta\right)\sum_{i=0}^{n} a_i x^i = \sum_{i=0}^{n}(\alpha\beta)a_i x^i \\ &= \sum_{i=0}^{n}\alpha(\beta a_i)x^i = \alpha\sum_{i=0}^{n}\beta a_i x^i = \alpha\left[\beta p\left(x\right)\right].\end{aligned}$$

**(4)**

$$1p\left(x\right) = 1\sum_{i=0}^{n} a_i x^i = \sum_{i=0}^{n} 1a_i x^i = \sum_{i=0}^{n} a_i x^i = p\left(x\right).$$

Terminaremos este capítulo estudiando cómo construir un nuevo espacio vectorial a partir de dos espacios vectoriales dados.

**Definición 1.16 (suma directa externa).** Sean $V$ y $W$ espacios vectoriales sobre $F$. La *suma directa externa* de $V$ y $W$ es el espacio vectorial $V \boxplus W = (V \times W, +, \cdot)$ sobre $F$, donde cada elemento de la tríada está definido como sigue:

**(a)** $V \times W = \{(v, w) : v \in V, w \in W\}$.

**(b)** La suma de elementos de $V \times W$ es la *suma por coordenadas*: $(v_1, w_1) + (v_2, w_2) = (v_1 + v_2, w_1 + w_2)$, para toda $(v_1, w_1), (v_2, w_2) \in V \times W$.

**(c)** La multiplicación escalar es la *multiplicación escalar por coordenadas*: $\alpha\,(v, w) = (\alpha v, \alpha w)$, para toda $(v, w) \in V \times W$ y $\alpha \in F$.

La demostración de que $V \boxplus W$ es de hecho un espacio vectorial sobre $F$ se deja como ejercicio. Observemos que, tomando $V = W = \mathbb{R}$, los espacios vectoriales $\mathbb{R} \boxplus \mathbb{R}$ y $\mathbb{R}^2$ son iguales. En general, para cualquier campo $F$, tenemos que $F^n = \underbrace{F \boxplus F \boxplus \cdots \boxplus F}_{n \text{ veces}}$.

***Palabras clave:*** *espacio vectorial, espacio vectorial real, espacio vectorial de matrices, espacio vectorial de funciones, espacio vectorial de polinomios, suma directa externa.*

## 1.2. Ejercicios

**Ejercicio 1.17.** Sea $V$ un espacio vectorial sobre un campo $F$.

**(1)** Demuestra que si $\alpha, \beta \in F$ son escalares distintos, entonces $\alpha v \neq \beta v$, para todo $v \in V \setminus \{0\}$.

**(2)** Usa el punto anterior para demostrar que todo espacio vectorial no trivial sobre un campo infinito tiene un número infinito de vectores distintos.

**(3)** Si $F$ es un campo finito, da un ejemplo de espacio vectorial sobre $F$ que tenga un número finito de vectores, y otro que tenga un número infinito de vectores.

**Ejercicio 1.18.** Sean $V$ y $W$ espacios vectoriales sobre $F$. Demuestra que la suma directa externa $V \boxplus W$ es un espacio vectorial sobre $F$.

**Ejercicio 1.19.** Demuestra que $(F[x], +)$, definido en el Ejemplo 1.15, es un grupo abeliano.

**Ejercicio 1.20.** Demuestra la propiedad **(4)** de la Proposición 1.6.

**Ejercicio 1.21.** Demuestra la propiedad **(2)** de la Proposición 1.7.

**Ejercicio 1.22.** Demuestra que el conjunto de matrices

$$M_{2\times 2}(\mathbb{R}) = \left\{ \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} : x_{11}, x_{12}, x_{21}, x_{22} \in \mathbb{R} \right\},$$

es un espacio vectorial sobre $\mathbb{R}$, con la suma y multiplicación escalar usual de matrices.

**Ejercicio 1.23.** Sea $F$ un campo. Demuestra que el conjunto de series formales

$$F\left[[x]\right] = \left\{ \sum_{i=0}^{\infty} a_i x^i : a_i \in F \right\},$$

junto con las operaciones

$$\begin{aligned} \sum_{i=0}^{\infty} a_i x^i + \sum_{i=0}^{\infty} b_i x^i &= \sum_{i=0}^{\infty} (a_i + b_i)\, x^i, \\ \alpha \sum_{i=0}^{\infty} a_i x^i &= \sum_{i=0}^{\infty} \alpha a_i x^i, \text{ donde } \alpha \in F, \end{aligned}$$

es un espacio un espacio vectorial sobre $F$.

**Ejercicio 1.24.** Demuestra que $\mathbb{C} = \{x + yi : x, y \in \mathbb{R}\}$, junto con la suma usual de números complejos y la multiplicación escalar $\alpha\,(x + yi) = \alpha x + \alpha y i$, $\alpha \in \mathbb{R}$, es un espacio vectorial sobre $\mathbb{R}$.

**Ejercicio 1.25.** Considera el conjunto $\mathbb{R} \times \mathbb{R}$ y las operaciones

$$\begin{aligned}(x_1, x_2) \mathbin{\widehat{+}} (y_1, y_2) &= (x_1 y_1, x_2 y_2)\,, \\ \alpha\,(x_1, x_2) &= (\alpha x_1, \alpha x_2)\,,\end{aligned}$$

donde $x_i, y_i, \alpha \in \mathbb{R}$. Explica por qué $\mathbb{R} \times \mathbb{R}$, junto con estas operaciones, **no** es un espacio vectorial real.

**Ejercicio 1.26.** Sea $V = \{(x_1, x_2) : x_1, x_2 \in \mathbb{R}\}$. Para $(x_1, x_2), (y_1, y_2) \in V$ y $\alpha \in \mathbb{R}$, definamos

$$\begin{aligned}(x_1, x_2) \mathbin{\widehat{+}} (y_1, y_2) &= (x_1 + y_1, 0)\,, \\ \alpha \mathbin{\widehat{\cdot}} (x_1, x_2) &= (\alpha x_1, 0)\,.\end{aligned}$$

Demuestra que $V$ con estas operaciones **no** es un espacio vectorial sobre $\mathbb{R}$.

Existen ejemplos de espacios espacios raros que satisfacen los propiedades de espacio vectorial a pesar de contar con operaciones que no se parecen a la adición y la multiplicación escalar usuales. Estos espacios son, en esencia, espacios vectoriales ordinarios con sus elementos etiquetados de nuevo para disfrazar su naturaleza. Los siguientes ejercicios exploran esta posibilidad con frecuencia confusa.

**Ejercicio 1.27.** Sea $V = \{(x_1, x_2) : x_1, x_2 \in \mathbb{C}\}$. Para $(x_1, x_2), (y_1, y_2) \in V$ y $\alpha \in \mathbb{C}$, definamos

$$\begin{aligned}(x_1, x_2) \mathbin{\widehat{+}} (y_1, y_2) &= (x_1 + y_1 + 1, x_2 + y_2 + 1)\,, \\ \alpha \mathbin{\widehat{\cdot}} (x_1, x_2) &= (\alpha x_1 + \alpha - 1, \alpha x_2 + \alpha - 1)\,.\end{aligned}$$

Demuestra que $V$ es un espacio vectorial sobre $\mathbb{C}$. ¿Cuál es el vector cero?, ¿Cuál es el inverso aditivo?

**Ejercicio 1.28.** Sea $V$ un espacio vectorial sobre un campo $F$, $S$ un conjunto y $f : V \to S$ una función biyectiva (así que $f^{-1} : S \to V$ es una función bien definida). Para todo $v, w \in S$ y $\alpha \in F$, definamos

$$\begin{aligned}v \mathbin{\widehat{+}} w &= f\left(f^{-1}(v) + f^{-1}(w)\right), \\ \alpha \mathbin{\widehat{\cdot}} v &= f\left(\alpha f^{-1}(v)\right).\end{aligned}$$

Demuestra que $S$ con estas operaciones satisfacen las propiedades de espacio vectorial.

**Ejercicio 1.29.** Sea $S = \mathbb{R}^+$. Demuestra que $S$ es un espacio vectorial con las operaciones

$$\begin{aligned}v \mathbin{\widehat{+}} w &= vw, \\ \alpha \mathbin{\widehat{\cdot}} v &= v^\alpha.\end{aligned}$$

¿Qué elemento de $S$ es la identidad aditiva? ¿Qué significado tiene $-x$ en este contexto?

**Ejercicio 1.30.** Sea $S = \mathbb{R}$. Demuestra que $S$ es un espacio vectorial con las operaciones

$$v \widehat{+} w = v + w + 1,$$
$$\alpha \widehat{\cdot} v = \alpha v + \alpha - 1.$$

¿Qué elemento de $S$ es la identidad aditiva? ¿Qué significado tiene $-x$ en este contexto?

**Ejercicio 1.31.** Muestra que los espacios de los ejercicios 1.29 y 1.30 se pueden obtener de $V = \mathbb{R}$ por la construcción del ejercicio 1.28 y los mapeos $f(v) = e^v$ y $f(v) = v - 1$, respectivamente. De hecho, todo ejemplo de un espacio vectorial raro puede ser obtenido de un espacio vectorial simple por medio de esta construcción.

# 2

# Subespacios vectoriales

**Definición 2.1 (subespacio vectorial).** Sea $V$ un espacio vectorial sobre $F$. Un subconjunto $S$ de $V$ es un *subespacio vectorial* de $V$ si $S$ es en sí mismo un espacio vectorial sobre $F$ bajo la restricción de las operaciones de suma y multiplicación escalar definidas en $V$.

**Observación 2.2.** Si $+ : V \times V \to V$ y $\cdot : F \times V \to V$ son la operaciones de un espacio vectorial $V$, la restricción de estas operaciones a un subconjunto $S \subseteq V$ tienen la forma $+|_S : S \times S \to V$ y $\cdot|_S : F \times S \to V$. Para que $S$ sea un subespacio vectorial de $V$, es necesario que estas operaciones tengan la forma $+|_S : S \times S \to S$ y $\cdot|_S : F \times S \to S$; es decir, que cada una cumpla la propiedad de cerradura en $S$.

**Observación 2.3.** Para denotar que $S$ es un subespacio vectorial de $V$ escribiremos $S \leq V$. También, ocasionalmente, usaremos simplemente el término *subespacio* para referirnos a un subespacio vectorial.

**Teorema 2.4 (test del subespacio 1).** Sea $V$ un espacio vectorial sobre un campo $F$. Un subconjunto $S \subseteq V$ es un subespacio vectorial de $V$ si y sólo si se cumple lo siguiente:

**(SV1)** $\mathbf{0} \in S$.

**(SV2)** Si $u, v \in S$, entonces $u + v \in S$.

**(SV3)** Si $v \in S$, $\alpha \in F$, entonces $\alpha v \in S$.

**Demostración.**

($\Rightarrow$) Supongamos que $S \leq V$. Entonces $S$ cumple las propiedades **(SV2)** y **(SV3)** por la definición de espacio vectorial, pues las operaciones suma y multiplicación por escalar deben estar bien definidas sobre $S$. Además, $S$ debe tener un vector cero, al cual denotaremos por $\mathbf{0}_S$, el cual cumple que $\mathbf{0}_S + (-\mathbf{0}_S) = \mathbf{0}_S$. Sin embargo, $\mathbf{0}_S \in V$ debido a que $S \subseteq V$, y debemos tener que $\mathbf{0}_S + (-\mathbf{0}_S) = \mathbf{0}$, donde $\mathbf{0}$ es el vector cero de $V$. Esto demuestra la propiedad **(SV1)**, pues $\mathbf{0} = \mathbf{0}_S \in S$.

($\Leftarrow$) Supongamos que $S \subseteq V$ cumple las propiedades **(SV1)-(SV3)**. Demostraremos que $S$ es un espacio vectorial.

Para demostrar que $S$ es un espacio vectorial en sí mismo, mostremos primero que $(S, +)$ es un subgrupo abeliano de $(V, +)$:

**(EV1)** $(S, +)$ es un grupo abeliano.

- **(G0)** *Cerradura*: por la hipótesis **(SV2)**, $u + v \in S$, $\forall u, v \in S$.
- **(G1)** *Asociatividad*: Sabemos que $u+(v+w) = (u+v)+w$, $\forall u, v, w \in V$, porque $V$ es un espacio vectorial. Como $S \subseteq V$, en particular se cumple que $u + (v + w) = (u + v) + w$, $\forall u, v, w \in S$.
- **(G2)** *Identidad*: por la hipótesis **(SV1)**, la identidad aditiva $\mathbf{0}$ pertenece a $S$.
- **(G3)** *Inversos*: sea $v \in S$. Por la hipótesis **(SV3)**, tenemos que $(-1)\, v \in S$. Como $(-1)\, v = -(1v) = -v$ por la Proposición 1.6 **(4)**, entonces $-v \in S$.
- **(G4)** *Conmutatividad*: Sabemos que $v + w = w + v$, $\forall v, w \in V$, porque $V$ es un espacio vectorial. Como $S \subseteq V$, en particular se cumple que $v + w = w + v$, $\forall v, w \in S$.

**(EV2)** Por la hipótesis **(SV3)**, la multiplicación escalar cumple la propiedad de cerradura en $S$. El resto de las propiedades de la multiplicación escalar se cumplen automáticamente en $S$ porque son un caso particular de las propiedades que se satisfacen en $V$.

□

Si asumimos que $S$ es un subconjunto no vacío de $V$, entonces podemos eliminar la condición **(SV1)** del Teorema 2.4; en efecto, si $v \in S$, entonces $0v = \mathbf{0} \in S$ por la condición **(SV3)**. Esto indica que, para determinar si $S$ es un subespacio de $V$, no es necesario siempre verificar que $\mathbf{0} \in S$, basta con demostrar que $S$ es no vacío y que se cumplen las condiciones **(SV2)** y **(SV3)** del Teorema 2.4. Podemos simplificar aún más las cosas, según se indica en el siguiente resultado.

**Teorema 2.5 (test del subespacio 2).** Sea $V$ un espacio vectorial sobre un campo $F$ y $S$ un subconjunto no vacío de $V$. Entonces, $S$ es un subespacio vectorial de $V$ si y sólo si

$$\alpha v + \beta w \in S,$$

para toda $\alpha, \beta \in F$ y $v, w \in S$.

**Demostración.**

($\Rightarrow$) Supongamos que $S \leq V$. Como $S$ es cerrado bajo la la suma de vectores y la multiplicación escalar, obviamente se cumple que $\alpha v + \beta w \in S$, $\forall \alpha, \beta \in F$, $v, w \in V$.

($\Leftarrow$) Supongamos ahora que $S$ es un subconjunto no vacío de $V$ tal que para cualquier $\alpha$, $\beta \in F$ y $v, w \in S$ resulta que $\alpha v + \beta w \in S$. Usaremos la Proposición 2.4 para demostrar que $S$ es un subespacio vectorial.

**(SV1)** Como $S$ es no vacío, existe al menos un vector $v \in V$. Tomando $w = v$ y $\alpha = \beta = 0$, vemos que $0v + 0v = \mathbf{0} \in S$.

**(SV2)** Tomando $\alpha = \beta = 1$, vemos que $v + w \in S$, $\forall v, w \in S$.

**(SV3)** Tomando $\beta = 0$, vemos que $\alpha v + 0w = \alpha v \in S$, $\forall v \in S$, $\alpha \in F$.

□

**Ejemplo 2.6.** Sea $V$ un espacio vectorial sobre un campo $F$. Los conjuntos $V$ y $\{\mathbf{0}\}$ siempre son subespacios de $V$, llamados *subespacios triviales*.

**Ejemplo 2.7.** El conjunto

$$S = \left\{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_3 = 0\right\},$$

es un subespacio de $\mathbb{R}^3$. Verificaremos que se cumple el Test del Subespacio 1.

**(SV1)** $\mathbf{0} = (0, 0, 0) \in S$ porque la tercera coordenada es cero.

**(SV2)** Sean $v = (v_1, v_2, 0), w = (w_1, w_2, 0) \in S$. Entonces, $v + w = (v_1 + w_1, v_2 + w_2, 0) \in S$, puesto que la tercera coordenada es cero.

**(SV3)** Sean $v = (v_1, v_2, 0) \in S$ y $\alpha \in \mathbb{R}$. Entonces, $\alpha v = (\alpha v_1, \alpha v_2, 0) \in S$, puesto que la tercera coordenada es cero.

**Ejemplo 2.8.** El conjunto

$$S = \left\{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 + x_2 - 2x_3 = 0\right\},$$

es un subespacio de $\mathbb{R}^3$. Verificaremos que se cumple el Test del Subespacio 1.

**(SV1)** $\mathbf{0} = (0, 0, 0) \in S$ porque $0 + 0 - (2 \cdot 0) = 0$.

**(SV2)** Si $v = (v_1, v_2, v_3), w = (w_1, w_2, w_3) \in S$, por definición tenemos que

$$v_1 + v_2 - 2v_3 = 0,$$
$$w_1 + w_2 - 2w_3 = 0.$$

Ahora, $v + w = (v_1 + w_1, v_2 + w_2, v_3 + w_3) \in S$ porque

$$(v_1+w_1)+(v_2+w_2)-2(v_3+w_3) = (v_1+v_2-2v_3)+(w_1+w_2-2w_3) = 0+0 = 0.$$

**(SV3)** Para toda $v = (v_1, v_2, v_3) \in S$ y $\alpha \in \mathbb{R}$, tenemos que $\alpha v = (\alpha v_1, \alpha v_2, \alpha v_3) \in S$ porque

$$\alpha v_1 + \alpha v_2 - 2\alpha v_3 = \alpha(v_1 + v_2 - 2v_3) = \alpha \cdot 0 = 0.$$

**Ejemplo 2.9.** El conjunto

$$S := \{(0, 0, 0), (1, 1, 0), (1, 0, 1), (0, 1, 1)\}.$$

es un subespacio de $\mathbb{Z}_2^3$. Verificaremos que se cumple el Test del Subespacio 1.

**(SV1)** Claramente $(0,0,0) \in S$.

**(SV2)** Es fácil verificar que $S$ es cerrado bajo la suma usando el hecho que $v + v = \mathbf{0}$, $\forall v \in \mathbb{Z}_2^3$, y que

$$(1,1,0) + (1,0,1) = (0,1,1).$$

**(SV3)** En el campo $\mathbb{Z}_2$ sólo hay dos escalares: 0 y 1. Por lo tanto, para toda $v \in S$, tenemos $0v = \mathbf{0} \in S$ y $1v = v \in S$.

**Ejemplo 2.10.** Veremos algunos ejemplos de conjuntos que **no** son subespacios.

1. $A := \{(x_1, x_2) \in \mathbb{R}^2 : x_1 + x_2 = 1\}$ no es subespacio de $\mathbb{R}^2$ porque $(0,0) \notin A$ (ya que $0 + 0 \neq 1$).
2. $A := \{(x_1, x_2) \in \mathbb{R}^2 : x_1 x_2 = 0\}$ no es un subespacio de $\mathbb{R}^2$ porque no es cerrado bajo la suma: por ejemplo, $(1,0), (0,1) \in A$ pero $(1,0) + (0,1) = (1,1) \notin A$.
3. $A := \{(0,0), (1,0), (0,1), (1,1)\}$ no es subespacio de $\mathbb{R}^2$ porque no es cerrado bajo la suma: por ejemplo, $(1,0) + (1,0) = (2,0) \notin A$.
4. $A := \{(x_1, x_2) \in \mathbb{R}^2 : x_1, x_2 \in \mathbb{Z}\}$ no es subespacio de $\mathbb{R}^2$ porque no es cerrado bajo la multiplicación por escalar: si $(x_1, x_2) \in A$ existen $\alpha \in \mathbb{R}$ (por ejemplo, $\alpha = \sqrt{2}$ o $\alpha = \pi$) tales que $\alpha(x_1, x_2) \notin A$.

## 2.1. Subespacios generados

**Definición 2.11 (combinación lineal).** Sea $A$ un subconjunto no vacío de un espacio vectorial $V$ sobre $F$. Una *combinación lineal* de $A$ es una expresión de la forma

$$\alpha_1 v_1 + \alpha_2 v_2 + \ldots \alpha_n v_n,$$

donde $\alpha_i \in F$, $v_i \in A$, $i = 1, \ldots, n$. Los escalares $\alpha_1, \ldots, \alpha_n$ se llaman los *coeficientes* de la combinación lineal. Decimos que la combinación lineal es *trivial* si todos sus coeficientes son cero.

Obviamente, también podemos usar la notación de sumatoria para escribir una combinación lineal:

$$\alpha_1 v_1 + \alpha_2 v_2 + \ldots \alpha_n v_n = \sum_{i=1}^{n} \alpha_i v_i.$$

**Observación 2.12.** Por definición, el número de términos que aparece en una combinación lineal de $A$ es un número **finito**, aunque el subconjunto $A$ en sí mismo podría ser infinito.

La siguiente definición nos permite construir un subespacio a partir de un subconjunto.

**Definición 2.13 (subespacio generado).** Sea $A$ un subconjunto no vacío de un espacio vectorial $V$ sobre $F$. El *espacio generado por $A$ sobre $F$*, denotado por $\text{gen}_F(A)$ o por $\langle A\rangle$ (cuando $F$ está claro en el contexto), es el conjunto de todas las combinaciones lineales de $A$ con escalares en $F$. En otras palabras,

$$\text{gen}_F(A) = \langle A\rangle := \{\alpha_1 v_1 + \cdots + \alpha_n v_n : \alpha_i \in F, v_i \in A, n \in \mathbb{N}\}.$$

Si $A = \emptyset$, definimos $\langle\emptyset\rangle = \{\mathbf{0}\}$.

**Teorema 2.14 (subespacio generado).** Sea $A$ un subconjunto no vacío de un espacio vectorial $V$ sobre $F$. Entonces, $\langle A\rangle$ es un subespacio de $V$.

**Demostración.** Usaremos el Test del Subespacio 1.

**(SV1)** $\mathbf{0} \in \langle A\rangle$ porque $\mathbf{0} = 0v_1 + \cdots + 0v_n$ para cualquier $v_i \in A$.

**(SV2)** Sean $\sum_{i=1}^n \alpha_i v_i, \sum_{i=1}^m \beta_i w_i \in \langle A\rangle$ dos combinaciones lineales, donde $v_i, w_i \in A$, $\alpha_i, \beta_i \in F$. Entonces su suma es también una combinacion lineal de $A$, por lo que $\sum_{i=1}^n \alpha_i v_i + \sum_{i=1}^m \beta_i w_i \in \langle A\rangle$.

**(SV3)** Sea $\sum_{i=1}^n \alpha_i v_i \in \langle A\rangle$ y $\alpha \in F$. Entonces,

$$\alpha\left(\sum_{i=1}^n \alpha_i v_i\right) = \sum_{i=1}^n (\alpha\alpha_i) v_i \in \langle A\rangle.$$

□

**Ejemplo 2.15.** Sea $V$ un espacio vectorial sobre $F$. Si $v \in V$, entonces

$$\text{gen}_F(v) = \langle\{v\}\rangle = \langle v\rangle = \{\alpha v : \alpha \in F\}.$$

Por ejemplo, si consideramos $v = (1,0,0) \in \mathbb{R}^3$, entonces

$$\begin{aligned}\langle v\rangle &= \{\alpha(1,0,0) : \alpha \in \mathbb{R}\} \\ &= \{(\alpha,0,0) : \alpha \in \mathbb{R}\}.\end{aligned}$$

Similarmente, si $u = (1,0,0) \in \mathbb{Z}_3^3$, entonces

$$\begin{aligned}\langle u\rangle &= \{(\alpha,0,0) : \alpha \in \mathbb{Z}_3\} \\ &= \{(0,0,0),(1,0,0),(2,0,0)\}.\end{aligned}$$

**Ejemplo 2.16.** Sea $V$ un espacio vectorial sobre $F$. Si $v, u \in V$, entonces

$$\langle v, u\rangle = \{\alpha_1 v + \alpha_2 u : \alpha_i \in F\}.$$

Por ejemplo, si $v = (1,0), u = (0,1) \in \mathbb{R}^2$, entonces

$$\begin{aligned}\langle v, u\rangle &= \{\alpha_1(1,0) + \alpha_2(0,1) : \alpha_i \in \mathbb{R}\} \\ &= \{(\alpha_1,\alpha_2) : \alpha_i \in \mathbb{R}\} = \mathbb{R}^2.\end{aligned}$$

Cabe señalar que si $w = \alpha v + \beta u$ es cualquier combinación lineal de $v$ y $u$, entonces $\langle v, u, w\rangle = \langle v, u\rangle$, debido a que

$$\begin{aligned}\langle v, u, w\rangle &= \{\alpha_1 v + \alpha_2 u + \alpha_3 w : \alpha_i \in F\} \\ &= \{\alpha_1 v + \alpha_2 u + \alpha_3 (\alpha v + \beta u) : \alpha_i \in F\} \\ &= \{(\alpha_1 + \alpha_3\alpha) v + (\alpha_2 + \alpha_3\beta) u : \alpha_i \in F\} \\ &= \langle v, u\rangle .\end{aligned}$$

**Observación 2.17.** Sea $V$ un espacio vectorial sobre $F$. Sean $v_1, \ldots, v_n \in V$ vectores arbitrarios y sean $\alpha_1, \ldots, \alpha_n \in F$ escalares arbitrarios. Entonces,

$$\langle v_1, \ldots, v_n\rangle = \langle \alpha_1 v_1, \ldots, \alpha_n v_n\rangle.$$

**Lema 2.18 (propiedades básicas de subespacios generados).** Sea $A$ un subconjunto no vacío de un espacio vectorial $V$ sobre $F$. Entonces:

**(1)** $A \subseteq \langle A\rangle$

**(2)** Si $S \leq V$ satisface que $A \subseteq S$, entonces $\langle A\rangle \subseteq S$.

**(3)** Si $B$ es un subconjunto de $V$ tal que $A \subseteq B$, entonces $\langle A\rangle \subseteq \langle B\rangle$.

**(4)** $A$ es un subespacio de $V$ si y sólo si $\langle A\rangle = A$.

**(5)** $\langle A \cup \{v\}\rangle = \langle A\rangle$ si y sólo si $v \in \langle A\rangle$.

**Demostración.** Ejercicio 2.45. □

**Observación 2.19.** Los puntos **(1)** y **(2)** del lema anterior nos dice que $\langle A\rangle$ es el subespacio más pequeño que contiene a $A$.

**Ejemplo 2.20.** Consideremos el espacio vectorial de polinomios $\mathbb{R}[x]$. Si $A := \left\{1, x, x^2\right\}$, entonces

$$\langle A\rangle = \left\{\alpha_1 + \alpha_2 x + \alpha_3 x^2 : \alpha_i \in \mathbb{R}\right\}$$

es el subespacio de polinomios con grado menor o igual que dos.

**Ejemplo 2.21.** Consideremos el siguiente subconjunto infinito de $\mathbb{R}[x]$

$$A := \{1, x, x^2, \ldots\} = \{x^i : i \in \mathbb{N}\}.$$

Entonces, $\langle A\rangle = \mathbb{R}[x]$, porque cualquier polinomio en $\mathbb{R}[x]$ es una combinación lineal de elementos de $A$.

**Ejemplo 2.22.** Sea $A = \{(-3, 1, 1), (1, -3, 1), (1, 1, -3)\} \subseteq \mathbb{R}^3$. Queremos determinar si el vector $(1, 2, 4)$ pertenece al espacio generado por $A$. En caso de que así sea, deben existir escalares $\alpha_i \in \mathbb{R}$ tales que

$$\begin{aligned}(1, 2, 4) &= \alpha_1 (-3, 1, 1) + \alpha_2 (1, -3, 1) + \alpha_3 (1, 1, -3) \\ &= (-3\alpha_1 + \alpha_2 + \alpha_3, \alpha_1 - 3\alpha_2 + \alpha_3, \alpha_1 + \alpha_2 - 3\alpha_3) .\end{aligned}$$

Por lo tanto, tales escalares existen sí y sólo si el siguiente sistema de ecuaciones tiene solución:

$$\begin{aligned} 1 &= -3\alpha_1 + \alpha_2 + \alpha_3 \\ 2 &= \alpha_1 - 3\alpha_2 + \alpha_3 \\ 4 &= \alpha_1 + \alpha_2 - 3\alpha_3 \end{aligned}$$

Con cálculos directos obtenemos que $\left[\alpha_1 = -2, \alpha_2 = -\frac{9}{4}, \alpha_3 = -\frac{11}{4}\right]$ es una solución del sistema; por lo tanto, efectivamente $(1,2,4) \in \langle A \rangle$ y

$$(1,2,4) = -2\,(-3,1,1) - \frac{9}{4}\,(1,-3,1) - \frac{11}{4}\,(1,1,-3)\,.$$

**Ejemplo 2.23.** Consideremos el siguiente subespacio de $\mathbb{R}^3$

$$S := \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 + x_2 + 2x_3 = 0\}.$$

Podemos encontrar un conjunto generador de $S$ de la siguiente forma:

$$\begin{aligned} S &= \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 = -x_2 - 2x_3\} \\ &= \{(-x_2 - 2x_3, x_2, x_3) : x_2, x_3 \in \mathbb{R}\} \\ &= \{x_2(-1,1,0) + x_3(-2,0,1) : x_2, x_3 \in \mathbb{R}\} \\ &= \langle (-1,1,0), (-2,0,1) \rangle\,. \end{aligned}$$

## 2.2. Intersecciones y sumas de subespacios

Dados dos subespacios $S$ y $T$ de $V$, existen dos subespacios importantes relacionados: uno que está contenido en ambos, llamado el *subespacio intersección*, y otro que contiene a ambos, llamado el*subespacio suma*.

**Teorema 2.24 (intersección de subespacios).** Sean $S$ y $T$ subespacios de un espacio vectorial $V$ sobre $F$. Entonces, la intersección $S \cap T$ también es un subespacio de $V$

**Demostración.** Puesto que $\mathbf{0} \in S$ y $\mathbf{0} \in T$, tenemos que $\mathbf{0} \in S \cap T$, así que $S \cap T \neq \emptyset$. Usaremos el Teorema 2.5. Sean $v, w \in S \cap T$ y $\alpha, \beta \in F$ elementos arbitrarios. Ahora,

$$\begin{aligned} v \in S \cap T \quad &\Longrightarrow \quad v \in S \text{ y } v \in T \\ w \in S \cap T \quad &\Longrightarrow \quad w \in S \text{ y } w \in T. \end{aligned}$$

Puesto que $S$ y $T$ son subespacios, tenemos que $\alpha v + \beta w \in S$ y $\alpha v + \beta w \in T$. Por lo tanto, $\alpha v + \beta w \in S \cap T$. Esto demuestra que $S \cap T$ es un subespacio de $V$. □

**Corolario 2.25.** Sean $S$ y $T$ subespacios del espacio vectorial $V$. El subespacio $S \cap T$ es el subespacio de $V$ más grande contenido simultáneamente en $S$ y $T$.

**Demostración.** *Ejercicio* 2.46. □

**Observación 2.26.** La unión de dos subespacios de un espacio vectorial $V$, puede no ser un subespacio de $V$. Por ejemplo,

$$S = \{(0,0,z) : z \in \mathbb{R}\} \text{ y } T = \{(0,y,0) : y \in \mathbb{R}\}$$

son subespacios de $\mathbb{R}^3$, pero $S \cup T$ no es un subespacio de $\mathbb{R}^3$. Para comprobar esto, vemos que $(0,0,1) \in S \cup T$ y $(0,1,0) \in S \cup T$, pero

$$(0,0,1) + (0,1,0) = (0,1,1) \notin S \cup T$$

ya que $(0,1,1) \notin S$ y $(0,1,1) \notin T$. Esto demuestra que $S \cup T$ no es cerrado bajo la suma de vectores, por lo que no puede ser un subespacio.

**Teorema 2.27 (unión de subespacios).** Sean $S$ y $T$ subespacios de un espacio vectorial $V$ sobre $F$. La unión $S \cup T$ es un subespacio de $V$ si y sólo si $S \subseteq T$ o $T \subseteq S$.

**Demostración.**

**(⇒)** Asumimos que $S \cup T \leq V$. Supongamos que $S \nsubseteq T$, así que demostraremos que $T \subseteq S$. Como $S \nsubseteq T$, existe un vector $s \in S \setminus T$. Sea $t \in T$ un vector arbitrario. Debido a que $S \cup T$ es un subespacio, tenemos que $s+t \in S \cup T$; es decir, $s+t \in S$ o $s+t \in T$. Si $s+t \in T$, entonces, por cerradura de la suma, deducimos que $s = (s+t)+(-t) \in T$, lo cual contradice que $s \in S \setminus T$. Luego, $s+t \in S$. Nuevamente, por cerradura de la suma, $t = (s+t)+(-s) \in S$. Esto demuestra que si $t \in T$, entonces $t \in S$; en otras palabras, $T \subseteq S$.

**(⇐)** Si $S \subseteq T$, entonces $S \cup T = T$ es un subespacio. Similarmente, si $T \subseteq S$, entonces $S \cup T = S$ es un subespacio.

□

**Definición 2.28 (suma de subespacios).** Sean $S$ y $T$ subespacios de un espacio vectorial $V$ sobre $F$. La suma de $S$ y $T$, denotada por $S+T$, es el subconjunto de $V$ de todas las posibles sumas de vectores de $S$ y $T$, es decir

$$S + T = \{s + t \in V : s \in S, t \in T\}.$$

**Ejemplo 2.29.** Sean $S = \{(x_1,0,0) \in \mathbb{R}^3 : x_1 \in \mathbb{R}\}$ y $T = \{(0,x_2,0) \in \mathbb{R}^3 : x_2 \in \mathbb{R}\}$. Entonces

$$S + T = \{(x_1, x_2, 0) : x_1, x_2 \in \mathbb{R}\}.$$

**Proposición 2.30.** Sean $S$ y $T$ subespacios de un espacio vectorial $V$ sobre $F$.

**(1)** La suma $S+T$ es un subespacio de $V$.

**(2)** $S+T$ es el subespacio de $V$ más pequeño que contiene a $S$ y $T$.

**Demostración.**

**(1)** Como $\mathbf{0} \in S$ y $\mathbf{0} \in T$, tenemos que $\mathbf{0} = \mathbf{0}+\mathbf{0} \in S+T$. Por lo tanto, $S+T$ es un conjunto no vacío. Usaremos el Teorema 2.5. Sean $s_1+t_1, s_2+t_2 \in S+T$, donde $s_1, s_2 \in S$ y $t_1, t_2 \in T$. Sean $\alpha, \beta \in F$ escalares arbitrarios. Puesto que $S$ y $T$ son subespacios, $\alpha s_1 + \beta s_2 \in S$ y $\alpha t_1 + \beta t_2 \in T$. Por lo tanto,

$$\alpha\,(s_1 + t_1) + \beta\,(s_2 + t_2) = (\alpha s_1 + \beta s_2) + (\alpha t_1 + \beta t_2) \in S + T,$$

Esto demuestra que $S + T \leq V$.

**(2)** Para probar que $S + T$ es el subespacio de $V$ más pequeño que contiene a $S$ y $T$, necesitamos mostrar dos cosas: **(a)** $S \subseteq S + T$ y $T \subseteq S + T$, **(b)** si $W \leq V$ satisface que $S \subseteq W$ y $T \subseteq W$, entonces $S + T \subseteq W$.

- **(a)** Cualquier $s \in S$ puede escribirse como $s = s+\mathbf{0} \in S+T$, donde $\mathbf{0} \in T$, así que $S \subseteq S + T$. Análogamente demostramos que $T \subseteq S + T$.
- **(b)** Sea $W$ un subespacio de $V$ tal que $S \subseteq W$ y $T \subseteq W$. Sea $s+t \in S+T$ un elemento arbitrario, donde $s \in S$ y $t \in T$. Puesto que $S \subseteq W$ y $T \subseteq W$, tenemos que $s \in W$ y $t \in W$. Como $W \leq V$, deducimos que $s + t \in W$. Esto demuestra que $S + T \subseteq W$.

□

**Definición 2.31 (suma directa interna).** Sean $S$ y $T$ subespacios de un espacio vectorial $V$ sobre $F$. Decimos que la suma de $S+T$ es una *suma directa interna* de $S$ y $T$, y escribimos $S \oplus T$, si cada elemento $v \in S + T$ puede escribirse de forma única como $v = s + t$; donde $s \in S$, $t \in T$.

**Ejemplo 2.32.** Sean $S = \{(x_1, x_2, 0) : x_1, x_2 \in \mathbb{R}\}$ y $T = \{(0, x_2, x_3) : x_2, x_3 \in \mathbb{R}\}$ dos subespacios de $\mathbb{R}^3$. No es difícil demostrar que la suma $S+T$ es igual a todo el espacio $\mathbb{R}^3$. Sin embargo, la suma $S+T$ no es una suma directa interna porque no todos sus vectores tienen una representación única; por ejemplo, $(4, 6, 8) \in S+T$ puede escribirse de dos formas distintas como suma de elementos de $S$ y $T$:

$$(4, 6, 8) = (4, 5, 0) + (0, 1, 8)\,, \text{ y}$$
$$(4, 6, 8) = (4, -1, 0) + (0, 7, 8)\,.$$

**Ejemplo 2.33.** Sean $S = \{(x_1, x_2, 0) : x_1, x_2 \in \mathbb{R}\}$ y $T = \{(0, 0, x_3) : x_3 \in \mathbb{R}\}$ dos subespacios de $\mathbb{R}^3$. En este caso, cualquier vector $(x_1, x_2, x_3) \in S + T = \mathbb{R}^3$ se escribe de forma única como la suma de un vector en $S$ y otro en $T$:

$$(x_1, x_2, x_3) = (x_1, x_2, 0) + (0, 0, x_3)\,.$$

Por lo tanto, $\mathbb{R}^3 = S \oplus T$.

**Ejemplo 2.34.** Sea $V = M_n(\mathbb{R})$ el espacio vectorial de las matrices de $n \times n$ sobre el campo $\mathbb{R}$. Si $V_1$ y $V_2$ son los subespacios de las matrices simétricas y antisimétricas respectivamente, entonces $V = V_1 \oplus V_2$. De hecho, cualquier matriz $A \in V$ se puede escribir de forma única como el suma de una matriz simétrica y una matriz antisimétrica: la única manera de tener $A = B + C$ con $B$ simétrica y $C$ antisimétrica es a través de $B = \frac{1}{2}\left(A + A^\top\right)$ y $C = \frac{1}{2}\left(A - A^\top\right)$.

En capítulos posteriores demostraremos que si $S$ y $T$ son subespacios de un espacio vectorial $V$, entonces la suma directa interna $S \oplus T$ y la suma directa externa $S \boxplus T$ son espacios vectoriales "estructuralmente iguales" (es decir, *isomorfos*).

**Teorema 2.35.** Sea $S$ y $T$ subespacios del espacio vectorial $V$ sobre $F$. La suma $S + T$ es una suma directa interna si y sólo si $S \cap T = \{\mathbf{0}\}$.

**Demostración.**

($\Rightarrow$) Supongamos que la suma $S + T$ es una suma directa interna, es decir, $S + T = S \oplus T$. Esto significa que cada elemento $v \in S \oplus T$ se escribe de manera única como $v = s + t$, donde $s \in S$, $t \in T$. Sea $w \in S \cap T$. Claramente, $w \in S \oplus T$ y puede escribirse como

$$w = \mathbf{0} + w \text{ donde } \mathbf{0} \in S, w \in T,$$
$$w = w + \mathbf{0} \text{ donde } w \in S, \mathbf{0} \in T.$$

Por la unicidad de la representación, tenemos que $w = \mathbf{0}$. Esto demuestra que $S \cap T = \{\mathbf{0}\}$.

($\Leftarrow$) Supongamos que $S \cap T = \{\mathbf{0}\}$. Demostraremos que cualquier vector de $S + T$ tiene una representación única como la suma de un vector de $S$ y otro de $T$. Sea $v \in S + T$ y supongamos que $v$ puede escribirse como

$$v = s + t \text{ donde } s \in S, t \in T, \text{ y}$$
$$v = s' + t' \text{ donde } s' \in S, t' \in T.$$

Luego,

$$s + t = s' + t' \implies s - s' = t' - t.$$

Como $S$ es un subespacio y $s, s' \in S$, tenemos que $s - s' \in S$. Similarmente, $t' - t \in T$. Así, $s - s' = t' - t \in S \cap T = \{\mathbf{0}\}$, lo que implica que $s - s' = t' - t = \mathbf{0}$. Por lo tanto,

$$s - s' = \mathbf{0} \Rightarrow s = s',$$
$$t' - t = \mathbf{0} \Rightarrow t = t'.$$

Esto demuestra que cualquier vector $v \in V$ se expresa de manera única como la suma de un elemento de $S$ y otro de $T$.

□

**Observación 2.36.** Si $S_1, S_2, \ldots, S_n$ son subespacios de un espacio vectorial $V$, definimos de manera análoga la suma de estos subespacios

$$\sum_{i=1}^{n} S_i := S_1 + S_2 + \cdots + S_n := \{s_1 + s_2 + \cdots + s_n : s_i \in S_i\}.$$

También decimos que la suma $\sum_{i=1}^{n} S_i$ es una suma directa interna si para todo $v \in S_1 + S_2 + \cdots + S_n$ existen únicos $s_1 \in S_1, s_2 \in S_2, \ldots s_n \in S_n$ tales que

$$v = s_1 + s_2 + \cdots + s_n.$$

Sin embargo, en este caso, la versión análoga del Teorema 2.35 debe formularse de manera adecuada, como en el Ejercicio 2.53.

***Palabras clave:*** *subespacio vectorial, combinación lineal, subespacio generado, intersección y suma de subespacios, suma directa interna.*

## 2.3. Ejercicios

**Ejercicio 2.37.** Sea $\mathbb{R}$ el campo de los números reales. ¿Cuáles de los siguientes conjuntos son subespacios de $\mathbb{R}^3$?. Justifica tu respuesta.

**(a)** $W_1 = \{(x_1, 2x_2, 3x_3) : x_1, x_2, x_3 \in \mathbb{R}\}$.

**(b)** $W_2 = \{(x_1, x_2, x_3) : x_1, x_2, x_3 \in \mathbb{Q}\}$.

**(c)** $W_3 = \{(x_1, x_1, x_1) : x_1 \in \mathbb{R}\}$.

**(d)** $W_4 = \left\{(x_1, x_2, x_3) : x_1, x_2, x_3 \in \mathbb{R} \text{ y } x_1^2 + x_2^2 + x_3^2 \geq 1\right\}$.

**Ejercicio 2.38.** Determina si los conjuntos $S_i$ son subespacios del espacio vectorial $V_i$. Justifica detalladamente tu respuesta.

**(a)** $S_1 := \left\{(x_1, x_2) \in \mathbb{R}^2 : x_1 \leq x_2\right\}$, $V_1 := \mathbb{R}^2$.

**(b)** $S_2 := \left\{(x_1, \ldots, x_n) \in \mathbb{R}^n : x_1^2 + \cdots + x_n^2 = 1\right\}$, $V_2 := \mathbb{R}^n$.

**(c)** $S_3 := \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 : x_1 + x_2 = x_3 - x_4 \text{ y } x_1 + 2x_2 + 3x_3 + 4x_4 = 0\}$, $V_3 := \mathbb{R}^4$.

**(d)** $S_4 := \{(0,0,0), (1,1,1)\}$, $V_4 = \mathbb{Z}_2^3$.

**(e)** $S_4 := \{(0,0,0), (1,0,0), (0,1,0), (0,0,1)\}$, $V_4 = \mathbb{Z}_2^3$.

**(f)** $S_5 := \{p(x) \in \mathbb{R}[x] : p(x) \text{ tiene grado } 2\}$, $V_5 := \mathbb{R}[x]$.

**(g)** $S_6 := \{p(x) \in \mathbb{R}[x] : p(x) \text{ tiene grado menor o igual que } 2\}$, $V_6 := \mathbb{R}[x]$.

**(h)** $S_7 := \{p(x) \in \mathbb{R}[x] : p(3) = 0\}$, $V_7 := \mathbb{R}[x]$.

**(i)** $S_8 := \{p(x) \in \mathbb{R}[x] : p(3) = 1\}$, $V_8 := \mathbb{R}[x]$.

**Ejercicio 2.39.** Sea $V = M_{2\times 2}(\mathbb{R})$ el espacio vectorial de todas las matrices $2 \times 2$ sobre el campo $\mathbb{R}$. Muestra que el subconjunto $W$ de todas las matrices $2 \times 2$ con determinante cero no es un subespacio vectorial de $V$.

**Ejercicio 2.40.** Sea $M_{n\times n}(\mathbb{R})$ el espacio vectorial sobre $\mathbb{R}$ que consiste en todas las matrices de $n \times n$ con entradas en $\mathbb{R}$. Sea $T \in M_{n\times n}(\mathbb{R})$ una matriz dada, y consideremos el conjunto $S$ de las matrices que conmutan con $T$:

$$S = \{A \in M_{n\times n}(\mathbb{R}) : AT = TA\}.$$

Demuestra que $S$ es un subespacio de $M_{n\times n}(\mathbb{R})$.

**Ejercicio 2.41.** Considera el espacio vectorial de la funciones reales

$$\mathbb{R}^{\mathbb{R}} = \{f : \mathbb{R} \to \mathbb{R} : f \text{ es una función }\}.$$

¿Cuál de los siguientes conjuntos son subespacios de $\mathbb{R}^{\mathbb{R}}$? Justifica tu respuesta.

**(a)** $S = \{f \in \mathbb{R}^{\mathbb{R}} : f\left(x^2\right) = [f(x)]^2\}$.

**(b)** $T = \{f \in \mathbb{R}^{\mathbb{R}} : f(0) = f(1)\}$.

**(c)** $W = \{f \in \mathbb{R}^{\mathbb{R}} : f(3) = 1 + f(-5)\}$.

**Ejercicio 2.42.** Sea $\mathbb{R}^{\mathbb{R}}$ el espacio vectorial de las funciones reales. Decimos que $f \in \mathbb{R}^{\mathbb{R}}$ es par si $f(-x) = f(x)$ para todo $x \in \mathbb{R}$, y $g \in V$ es impar si $g(-x) = -g(x)$ para todo $x \in \mathbb{R}$. Sean

$$S = \left\{f \in \mathbb{R}^{\mathbb{R}} : f \text{ es par } \right\} \text{ y } T = \left\{g \in \mathbb{R}^{\mathbb{R}} : g \text{ es impar } \right\}.$$

Demuestra que $S$ y $T$ son subespacios de $\mathbb{R}^{\mathbb{R}}$ y que $\mathbb{R}^{\mathbb{R}} = S \oplus T$. (Sugerencia: si $f \in \mathbb{R}^{\mathbb{R}}$, define pares e impares como $a(x) = \frac{1}{2}[f(x) + f(-x)]$ y $b(x) = \frac{1}{2}[f(x) - f(-x)]$, respectivamente).

**Ejercicio 2.43.** Sea $\mathbb{R}[x]$ el conjunto de todos los polinomios en la variable $x$ con coeficientes en $\mathbb{R}$. Determina si los siguientes subconjuntos son subespacios de $\mathbb{R}[x]$:

**(a)** $S = \left\{\sum_{i=0}^{n} a_i x^i : a_i \in \mathbb{Z}, n \in \mathbb{N}\right\}$.

**(b)** $W = \left\{\sum_{i=0}^{n} a_i x^{2i} : a_i \in \mathbb{R}, n \in \mathbb{N}\right\}$.

**Ejercicio 2.44.** Investiga si

$$(3, -1, 0, -1) \in gen_{\mathbb{R}}\{(2, -1, 3, 2), (-1, 1, 1 - 3), (1, 1, 9, -5)\} \subset \mathbb{R}^4.$$

**Ejercicio 2.45.** Demuestra el Lemma 2.18 de propiedades básicas del subespacio generado.

**Ejercicio 2.46.** Demuestra el Corolario 2.25.

**Ejercicio 2.47.** Demuestra que la intersección de cualquier colección de subespacios de un espacio vectorial $V$ sobre $F$ es un subespacio de $V$.

**Ejercicio 2.48.** Sean $S$ y $T$ subconjuntos de $\mathbb{R}^3$ definidos como

$$S = \{(x_1, x_2, 0) : x_1, x_2 \in \mathbb{R}\}, \quad T = \{(0, x_2, 0) : x_2 \in \mathbb{R}\}.$$

Demuestra que $S$ y $T$ son subespacios de $\mathbb{R}^3$ y describe los subespacios $S \cap T$ y $S + T$.

**Ejercicio 2.49.** Demostrar que $\mathbb{R}^3$ es igual a la suma directa interna de los siguientes subespacios vectoriales:

$$U = \left\{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 + x_2 + x_3 = 0\right\},$$
$$W = \{(t, 2t, 3t) : t \in \mathbb{R}\}.$$

**Ejercicio 2.50.** Suponga que

$$S = \{(x_1, x_2, x_1 + x_2, x_1 - x_2, 2x_1) \in \mathbb{R}^5 : x_1, x_2 \in \mathbb{R}\}.$$

Encuentra un subespacio $T$ de $\mathbb{R}^5$ tal que $\mathbb{R}^5 = S \oplus T$.

**Ejercicio 2.51.** Sean $S$, $T$ y $U$ subespacios del espacio vectorial $V$ sobre $F$.

**(a)** Demuestra que si $S \subseteq U$, entonces

$$U \cap (S + T) = S + (U \cap T).$$

Esta es llamada la *ley modular de los subespacios vectoriales.*

**(b)** Muestra con un contra ejemplo que en general

$$U \cap (S + T) \neq (U \cap S) + (U \cap T).$$

Sugerencia: Sea $V = \mathbb{R}^2$ y sean $S, T$ y $U$ tres lineas rectas distintas que pasen por el origen.

**Ejercicio 2.52.** Sean $S$ y $T$ subespacios del espacio vectorial $V$ sobre $F$. Supongamos que $V = S \oplus T$, y sea $U$ un subespacio de $V$ tal que $S \subseteq U$. Demuestra que $U = S \oplus (U \cap T)$. Sugerencia: usa el ejercicio 2.51.

**Ejercicio 2.53.** Sean $S_1, S_2, \ldots, S_n$ subespacios de un espacio vectorial $V$. Demuestra que la suma $\sum_{i=1}^{n} S_i$ es una suma directa interna si y solo si para toda $i \in \{1, \ldots, n\}$ se tiene que

$$S_i \cap \left( \sum_{j \neq i} S_j \right) = \{\mathbf{0}\}.$$

**Ejercicio 2.54.** Sea $V$ un espacio vectorial y considera la familia de subespacios

$$\mathcal{A} = \{S \leq V : \forall T \leq V, \exists U \leq V \text{ tal que } S \leq U \text{ y } T \leq U\}.$$

(Esto es, cualesquier dos elementos de $\mathcal{A}$ están contenidos en un tercero). Demuestra que $\cup\mathcal{A}$ es un subespacio de $V$.

# 3
# Transformaciones lineales

Veamos como relacionar dos espacios vectoriales sobre un campo $F$ mediante una función que preserve la estructura de espacio vectorial.

**Definición 3.1 (transformación lineal).** Sean $V$ y $W$ dos espacios vectoriales sobre un mismo campo $F$. Una función $\phi : V \to W$ es una *transformación lineal*, o un *homomorfismo de espacios vectoriales*, si

$$\phi(\alpha v + \beta w) = \alpha\phi(v) + \beta\phi(w),$$

para todo $v, w \in V$ y $\alpha, \beta \in F$.

En la definición anterior hay que tener en cuenta que las operaciones realizadas del lado izquierdo de la igualdad son las operaciones definidas en $V$, mientras que las operaciones realizadas del lado derecho son las operaciones definidas en $W$.

**Definición 3.2 (Tipos de transformaciones lineales).** Sean $V$ y $W$ espacios vectoriales sobre $F$. Sea $\phi : V \to W$ una transformación lineal. Decimos que $\phi$ es un:

**(1)** **Monomorfismo** si $\phi$ es inyectivo.

**(2)** **Epimorfismo** si $\phi$ es sobreyectivo.

**(3)** **Isomorfismo** si $\phi$ es biyectivo.

**(4)** **Endomorfismo** si $V = W$.

**(5)** **Automorfismo** si $V = W$ y $\phi$ es biyectivo.

**(6)** **Funcional lineal** si $W = F$.

## 3.1. Ejemplos de transformaciones lineales

Veamos algunos ejemplos de transformaciones lineales importantes.

**Ejemplo 3.3 (endomorfismo identidad).** Sea $V$ un espacio vectorial. La función identidad $I_V : V \to V$ definida por

$$I_V(v) = v, \quad \forall v \in V$$

es un endomorfismo sobre $V$. Para comprobar esto, sean $v, w \in V$ y $\alpha, \beta \in F$ elementos arbitrarios. Entonces

$$\begin{aligned} I(\alpha v + \beta w) &= \alpha v + \beta w && \text{[por la definición de } I\text{]} \\ &= \alpha I(v) + \beta I(w). \end{aligned}$$

El endomorfismo $I$ es llamado el *endomorfismo identidad* sobre $V$.

**Ejemplo 3.4 (transformación cero).** Sean $V$ y $W$ espacios vectoriales sobre un mismo campo $F$. La función $\hat{\mathbf{0}} : V \to W$ definida por

$$\hat{\mathbf{0}}(v) = \mathbf{0}_W, \quad \forall v \in V,$$

donde $\mathbf{0}_W$ representa el vector cero de $W$, es una transformación lineal. Para comprobar esto, sean $v, w \in V$ y $\alpha, \beta \in F$ elementos arbitrarios. Entonces,

$$\begin{aligned} \hat{\mathbf{0}}(\alpha v + \beta w) &= \mathbf{0}_W && \text{[por la definición de } \phi\text{]} \\ &= \alpha \mathbf{0}_W + w\mathbf{0}_W \\ &= \alpha \hat{\mathbf{0}}(v) + \beta \hat{\mathbf{0}}(w). \end{aligned}$$

A esta transformación lineal se le llama la *transformación cero*.

**Ejemplo 3.5 (negativo de una transformación lineal).** Sean $V$ y $W$ espacios vectoriales sobre un mismo campo $F$. Sea $\phi : V \to W$ una transformación lineal. La función $-\phi$ definida por

$$(-\phi)(v) = -\phi(v), \quad \forall v \in V,$$

es también una transformación lineal. Para comprobar esto, sean $v, w \in V$ y $\alpha, \beta \in F$ elementos arbitrarios. Entonces,

$$\begin{aligned} (-\phi)(\alpha v + \beta w) &= -\phi(\alpha v + \beta w) \\ &= -(\alpha \phi(v) + \beta \phi(w)) \\ &= \alpha(-\phi(v)) + \beta(-\phi(w)) \\ &= \alpha(-\phi)(v) + \beta(-\phi)(w). \end{aligned}$$

**Ejemplo 3.6 (endomorfismo escalar).** Sea $V$ es un espacio vectorial sobre el campo $F$ y $\gamma$ un escalar fijo, la función $\phi : V \to V$ definida por

$$\phi(v) = \gamma v, \quad \forall v \in V,$$

es una endomorfismo. Para comprobar esto, sean $v, w \in V$ y $\alpha, \beta \in F$ elementos arbitrarios. Entonces,

$$\begin{aligned}(\phi)(\alpha v + \beta w) &= \gamma(\alpha v + \beta w) \\ &= (\gamma\alpha)v + (\gamma\beta)w \\ &= (\alpha\gamma)v + (\beta\gamma)w \\ &= \alpha(\gamma v) + \beta(\gamma w) \\ &= \alpha\phi(v) + \beta\phi(w).\end{aligned}$$

Observemos que la conmutatividad de la multiplicación de los elementos del campo de escalar juega aquí un papel crucial. Además, si $0 < \gamma < 1$, entonces $\phi$ es llamada *contracción* de $V$ con factor $\gamma$, y si $\gamma > 1$, es llamada *dilatación* de $V$ con factor $\gamma$.

**Ejemplo 3.7.** Definamos la función $\phi : \mathbb{R}^3 \to \mathbb{R}$ por

$$\phi(x_1, x_2, x_3) = 3x_1 - 2x_2 + 2x_3.$$

Veamos que $\phi$ es una transformación lineal. Sean $v = (v_1, v_2, v_3)$, $w = (w_1, w_2, w_3) \in \mathbb{R}^3$. Entonces para cualquier $\alpha, \beta \in \mathbb{R}$, tenemos que

$$\begin{aligned}\phi(\alpha v + \beta w) &= \phi(\alpha(v_1, v_2, v_3) + \beta(w_1, w_2, w_3)) \\ &= \phi((\alpha v_1 + \beta w_1, \alpha v_2 + \beta w_2, \alpha v_3 + \beta w_3)) \\ &= 3(\alpha v_1 + \beta w_1) - 2(\alpha v_2 + \beta w_2) + 2(\alpha v_3 + \beta w_3) \\ &= \alpha(3v_1 - 2v_2 + 2v_3) + \beta(3w_1 - 2w_2 + 2w_3) \\ &= \alpha\phi(v) + \beta\phi(w).\end{aligned}$$

Por lo tanto, $\phi$ es una trasformación lineal.

**Ejemplo 3.8.** Recordemos que los elementos del espacio $\mathbb{R}^n$ tienen la forma $v = (v_1, v_2, \ldots, v_n)$, donde $v_i \in \mathbb{R}$. Denotamos por $v^T$ al mismo vector $v$ pero escrito en forma de columna, es decir:

$$v^T = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}.$$

Recordemos también que el producto de una matriz $A = (a_{i,j}) \in M_{n\times n}(\mathbb{R})$ por

un vector columna $v^T$ es un vector columna definido como

$$Av^T = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,n} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{i=1}^{n} a_{1,i} v_i \\ \sum_{i=1}^{n} a_{2,i} v_i \\ \vdots \\ \sum_{i=1}^{n} a_{n,i} v_i \end{pmatrix}.$$

Debido a que la multiplicación de matrices por vectores columna satisface que

$$A(v^T + w^T) = Av^T + Aw^T,$$

no es difícil comprobar que la función $\phi_A : \mathbb{R}^n \to \mathbb{R}^n$ definida como

$$\phi_A(v) = Av^T$$

es un endomorfismo sobre $\mathbb{R}^n$.

**Ejemplo 3.9.** Sea $V = M_{m\times n}(F)$ el espacio vectorial de las matrices de $m \times n$ sobre el campo $F$. Sea $P \in M_{m\times m}(F)$ una matriz fija de $m \times m$ y $Q \in M_{n\times n}(F)$ una matriz fija de $n \times n$. La función $\phi : V \to V$ definida por

$$\phi(A) = PAQ, \ \forall A \in V,$$

es un endomorfismo sobre $V$. Primero observemos que el producto $PAQ$ es efectivamente una matriz de $m \times n$. Sean $A, B \in V$ y $\alpha, \beta \in F$. Entonces,

$$\begin{aligned} \phi(\alpha A + \beta B) &= P(\alpha A + \beta B)Q && \text{[por definición de } \phi\text{]} \\ &= (\alpha PA + \beta PB)Q \\ &= \alpha PAQ + \beta PBQ \\ &= \alpha\phi(A) + \beta\phi(B). \end{aligned}$$

Por lo tanto $\phi$ es un endomorfismo sobre $V$.

**Ejemplo 3.10 (derivada formal).** Sea $F[x]$ el espacio vectorial de todos los polinomios con coeficientes en el campo $F$. Definamos una función $D : F[x] \to F[x]$ como

$$D\left(\sum_{i=0}^{n} a_i x^i\right) = \begin{cases} \sum_{i=1}^{n} a_i i x^{i-1} & \text{si } n \geq 1, \\ 0 & \text{si } n = 0. \end{cases}$$

Demostraremos que $D$ es un endomorfismo sobre $V$. Sean $\sum_{i=0}^{n} a_i x^i, \sum_{i=0}^{m} b_i x^i \in F[x]$, con $n \geq m$, y $\alpha, \beta \in F$, elementos arbitrarios. Supongamos que $n, m \neq 0$. Entonces, tenemos que

$$\begin{aligned} D\left(\alpha \sum_{i=0}^{n} a_i x^i + \beta \sum_{i=0}^{m} b_i x^i\right) &= D\left(\sum_{i=0}^{m} (\alpha a_i + \beta b_i) x^i + \sum_{i=m+1}^{n} \alpha a_i x^i\right) \\ &= \sum_{i=1}^{m} (\alpha a_i + \beta b_i) i x^{i-1} + \sum_{i=m+1}^{n} \alpha a_i i x^{i-1} \\ &= \alpha \sum_{i=1}^{n} a_i i x^{i-1} + \beta \sum_{i=1}^{m} b_i i x^{i-1} \\ &= \alpha D\left(\sum_{i=0}^{n} a_i x^i\right) + \beta D\left(\sum_{i=0}^{m} b_i x^i\right). \end{aligned}$$

Los casos cuando $n = 0$ o $m = 0$ se demuestran similarmente. A este endomorfismo se le llama la *derivada formal* sobre el espacio de polinomios. Cabe señalar que esta definición de derivada para polinomios es puramente algebraica; en un curso de análisis real se estudia una generalización de esta definición para clases más generales de funciones usando el concepto de límite.

## 3.2. Propiedades de las transformaciones lineales

**Proposición 3.11 (propiedades básicas de las transformaciones lineales).** Sea $\phi : V \to W$ es una transformación lineal entre espacios vectoriales sobre $F$.

**(1)** $\phi(\mathbf{0}_V) = \mathbf{0}_W$, donde $\mathbf{0}_V$ y $\mathbf{0}_W$ son los vectores cero de $V$ y $W$, respectivamente.

**(2)** $\phi(-v) = -\phi(v)$, para todo $v \in V$.

**(3)** $\phi(v - w) = \phi(v) - \phi(w)$, para todo $v, w \in V$.

**(4)** $\phi(\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n) = \alpha_1 \phi(v_1) + \alpha_2 \phi(v_2) + \ldots + \alpha_n \phi(v_n)$, donde $v_i \in V$ y $\alpha_i \in F$.

**Demostración.**

**(1)** Sea $v \in V$. Entonces,

$$\begin{aligned} \phi(v) + \mathbf{0}_W &= \phi(v) \\ &= \phi(v + \mathbf{0}_V) \\ &= \phi(v) + \phi(\mathbf{0}_V). \end{aligned}$$

Por lo tanto $\mathbf{0}_W = \phi(\mathbf{0_V})$, por cancelación izquierda en $W$.

**(2)** Puesto que $\phi$ es una transformación lineal se cumple que $\phi(v+(-v)) = \phi(v)+\phi(-v)$. Sin embargo, por la propiedad **(1)**,

$$\phi(v+(-v)) = \phi(\mathbf{0}_V) = \mathbf{0}_W$$

Así, $\phi(v)+\phi(-v) = \mathbf{0}_W$ y, por lo tanto, $\phi(-v) = -\phi(v)$.

**(3)** Para todo $v, w \in V$, tenemos que

$$\begin{aligned}
\phi(v-w) &= \phi[v+(-w)] \\
&= \phi(v)+\phi(-w) && [\because \phi \text{ es lineal}] \\
&= \phi(v)+[-\phi(w)] && [\text{por la propiedad } \mathbf{(2)}] \\
&= \phi(v)-\phi(w).
\end{aligned}$$

**4)** Es una consecuencia directa de la definición de transformación lineal.

□

**Ejemplo 3.12.** Sea $\phi : \mathbb{R}^2 \to \mathbb{R}^2$ la función definida por $\phi(x,y) = (x+2, y+3)$, $\forall (x,y) \in \mathbb{R}^2$. Como $\phi(0,0) = (2,3) \neq (0,0)$, entonces $\phi$ no es una transformación lineal puesto que contradice la propiedad **(1)** de la Proposición 3.11.

**Proposición 3.13 (composición de transformaciones lineales).** Sean $\phi : V \to W$ y $\varphi : W \to U$ transformaciones lineales. La composición $\varphi \circ \phi : V \to U$ también es una transformación lineal.

**Demostración.** Sean $u, v \in V$ y $\alpha, \beta \in F$. Entonces $\alpha u + \beta v \in V$ , luego

$$\begin{aligned}
(\varphi \circ \phi)(\alpha u + \beta v) &= \varphi[\phi(\alpha u + \beta v)] \\
&= \varphi[\alpha\phi(u) + \beta\phi(v)] \\
&= \alpha\varphi[\phi(u)] + \beta\varphi[\phi(v)] \\
&= \alpha(\varphi \circ \phi)(u) + \beta(\varphi \circ \phi)(v).
\end{aligned}$$

Por lo tanto $\varphi \circ \phi$ es una transformación lineal. □

**Observación 3.14.** Si $\phi : V \to V$ es un endomorfismo, las potencia de $\phi$ se definen mediante la composición de funciones:

$$\phi^1 = \phi,\ \phi^2 = \phi \circ \phi,\ \phi^3 = \phi^2 \circ \phi,\ \ldots,\ \phi^k = \phi^{k-1} \circ \phi.$$

Acordamos que $\phi^0 = I$ es el endomorfismo identidad.

Recordemos que una función $f : X \to Y$ entre dos conjuntos $X$ y $Y$ es *invertible* si existe una función $g : Y \to X$ tal que $g \circ f = I_X$ y $f \circ g = I_Y$, donde $I_X$ e $I_Y$ son las funciones identidades en $X$ y $Y$, respectivamente. En este caso, la función $g$ se denota por $g = f^{-1}$ y se llama la *inversa* de $f$. Recordamos que una función es invertible si y sólo si es biyectiva (es decir, inyectiva y sobreyectiva).

El siguiente lema muestra que si una función dada es una transformación lineal invertible de un espacio vectorial en otro, entonces la linealidad se conserva también por la inversa.

**Lema 3.15 (inversa de una transformación lineal).** Sea $\phi : V \to W$ una transformación lineal entre espacios vectoriales. Si $\phi$ es invertible, entonces la inversa $\phi^{-1} : W \to V$ también una transformación lineal.

**Demostración.** Sean $w_1, w_2 \in W$ y $\alpha, \beta \in F$ elementos arbitrarios. Puesto que $\phi$ es invertible, es sobreyectiva, por lo que existen vectores $v_1, v_2 \in V$ tales que $\phi(v_1) = w_1$ y $\phi(v_2) = w_2$. Entonces

$$\begin{aligned}\phi^{-1}(\alpha w_1 + \beta w_2) &= \phi^{-1}(\alpha\phi(v_1) + \beta\phi(v_2)) \\ &= \phi^{-1}(\phi(\alpha v_1 + \beta v_2)) \\ &= \alpha v_1 + \beta v_2 \\ &= \alpha\phi^{-1}(w_1) + \beta\phi^{-1}(w_2).\end{aligned}$$

Esto demuestra que $\phi^{-1}$ es una transformación lineal. □

**Definición 3.16 (isomorfismo).** Sean $V$ y $W$ espacios vectoriales sobre un mismo campo $F$. Una transformación lineal $\phi : V \to W$ es un *isomorfismo* si es invertible (o, equivalentemente, biyectiva). Cuando existe un isomorfismo de $V$ a $W$, decimos que $V$ y $W$ son *isomorfos* y escribimos $V \cong W$.

Si $V$ y $W$ son isomorfos entre sí, esto significa que tienen el mismo aspecto como espacios vectoriales.

**Teorema 3.17 (relación de isomorfía).** La relación de isomorfía entre espacios vectoriales sobre un campo $F$ es una relación de equivalencia.

**Demostración.** Sean $V$, $W$ y $U$ espacios vectoriales sobre $F$. Demostraremos que se cumplen las propiedades de una relación de equivalencia.

- *Reflexividad.* Claramente, el endomorfismo identidad $I_V : V \to V$ es un isomorfismo, así que $V \cong V$.
- *Simetría.* Supongamos que $V \cong W$, y sea $\phi : V \to W$ un isomorfismo. Por el Lema 3.15, la inversa $\phi^{-1} : W \to V$ también es un isomorfismo. Por lo tanto, $W \cong V$.
- *Transitividad.* Supongamos que $V \cong W$ y $W \cong U$, y sean $\phi : V \to W$ y $\tau : W \to U$ isomorfismos. Por la Proposición 3.13, $\tau \circ \phi : V \to U$ también es una transformación lineal. Además, $\tau \circ \phi$ es invertible porque su inversa es $\phi^{-1} \circ \tau^{-1}$. Por lo tanto, $\tau \circ \phi : V \to U$ es un isomorfismo y $V \cong U$.

□

**Ejemplo 3.18.** El campo $\mathbb{C}$ de los números complejos puede ser visto como un espacio vectorial sobre $\mathbb{R}$. Sea $\varphi : \mathbb{C} \to \mathbb{R}^2$ la función definida por

$$\varphi(a + ib) = (a, b), \quad \forall (a + ib) \in \mathbb{C}$$

Demostraremos que $\varphi$ es un isomorfismo.

**1)** $\varphi$ *es una transformación lineal.* Sean $u = a + ib$, $v = c + id$ elementos arbitrarios de $\mathbb{C}$ y sean $k_1, k_2 \in \mathbb{R}$ dos escalares cualesquiera. Entonces,

$$\begin{aligned}
\varphi(k_1u + k_2v) &= \varphi(k_1(a+ib) + k_2(c+id)) \\
&= \varphi((k_1a + k_2c) + i(k_1b + k_2d)) \\
&= (k_1a + k_2c, k_1b + k_2d), && \text{[por la definición de } \varphi\text{]} \\
&= (k_1a, k_1b) + (k_2c, k_2d) \\
&= k_1(a,b) + k_2(c,d), \\
&= k_1\varphi(a+ib) + k_2\varphi(a+ib), && \text{[por la definición de } \varphi\text{]} \\
&= k_1\varphi(u) + k_2\varphi(v).
\end{aligned}$$

**2)** $\varphi$ *es inyectiva.* Sean $u = a + ib$, $v = c + id$ dos elementos cualesquera de $\mathbb{C}$ tales que $\varphi(u) = \varphi(u)$. Entonces,

$$\varphi(a+ib) = \varphi(c+id) \;\Rightarrow\; (a,b) = (c,d).$$

Por lo tanto, $a = c$, $b = d$, lo que implica que $u = v$. Por lo tanto, $\varphi$ es inyectiva.

**3)** $\varphi$ *es sobreyectiva.* Sea $(a,b) \in \mathbb{R}^2$ un elemento arbitrario. Observemos que $v = a + ib \in \mathbb{C}$ satisface que $\varphi(v) = (a,b)$. Por lo tanto, $\varphi$ es sobreyectiva.

Por lo tanto, $\mathbb{R}^2$ y $\mathbb{C}$ son isomorfos como espacios vectoriales.

## 3.3. Imagen y kernel de una transformación lineal

Para cualquier transformación lineal $\phi : V \to W$, existen dos subespacios importantes asociados con $\phi$. El primero es un subespacio de $V$ llamado el kernel (o núcleo) de $\phi$; el segundo es un subespacio de $W$ llamado la imagen de $\phi$. En esta sección definimos estos dos subespacios.

**Definición 3.19 (imagen).** Sea $\phi : V \to W$ una transformación lineal de espacios vectoriales sobre $F$. La *imagen* de $\phi$, denotada por $\mathrm{Im}(\phi)$, es el conjunto

$$\begin{aligned}
\mathrm{Im}(\phi) &= \{\phi(v) \in W : v \in V\} \\
&= \{w \in W : w = \phi(v) \text{ para algún } v \in V\}.
\end{aligned}$$

**Observación 3.20.** Puesto que $\mathbf{0}_W = \phi(\mathbf{0_V})$, el vector cero de $W$ está en $\mathrm{Im}(\phi)$, de modo que el conjunto $\mathrm{Im}(\phi)$ no es vacío.

**Teorema 3.21 (la imagen es subespacio).** Sea $\phi : V \to W$ una transformación lineal de espacios vectoriales sobre $F$. La imagen $\mathrm{Im}(\phi)$ es un subespacio de $W$.

**Demostración.** Supongamos que $w_1, w_2 \in \mathrm{Im}(\phi)$ y $\alpha, \beta \in F$. Debemos probar que $\alpha w_1 + \beta w_2 \in \mathrm{Im}(\phi)$.Por definición, existen vectores $v_1, v_2 \in V$ tales que $\phi(v_1) = w_1$ y $\phi(v_2) = w_2$. Ya que $V$ es un espacio vectorial, $\alpha v_1 + \beta v_2 \in V$. Ahora,

$$\begin{aligned} \phi(\alpha v_1 + \beta v_2) &= \alpha\phi(v_1) + \beta\phi(v_2) && [\because \phi \text{ es lineal}] \\ &= \alpha w_1 + \beta w_2. \end{aligned}$$

De esta manera, $\alpha w_1 + \beta w_2$ es la imagen del vector $\alpha v_1 + \beta v_2$ y por lo tanto, $\alpha w_1 + \beta w_2 \in \mathrm{Im}(\phi)$ □

**Definición 3.22 (kernel).** Sea $\phi : V \to W$ una transformación lineal de espacios vectoriales sobre $F$. El *kernel* (o *núcleo*, o *espacio nulo*) de $\phi$, denotado por $\ker(\phi)$, es el conjunto

$$\ker(\phi) = \{v \in V : \phi(v) = \mathbf{0}_W\}.$$

**Teorema 3.23 (el kernel es subespacio).** Sea $\phi : V \to W$ una transformación lineal de espacios vectoriales sobre $F$. El kernel $\ker(\phi)$ es un subespacio de $V$.

**Demostración.** Supongamos que $v_1, v_2 \in \ker(\phi)$ y $\alpha, \beta \in F$. Por definición de kernel, $\phi(v_1) = \phi(v_2) = \mathbf{0}_W$. Aplicando $\phi$ a $\alpha v_1 + \beta v_2 \in V$ obtenemos

$$\begin{aligned} \phi(\alpha v_1 + \beta v_2) &= \alpha\phi(v_1) + \beta\phi(v_2), && [\because \phi \text{ es lineal}] \\ &= \alpha\mathbf{0}_W + \beta\mathbf{0}_W, \\ &= \mathbf{0}_W + \mathbf{0}_W, \\ &= \mathbf{0}_W. \end{aligned}$$

Así, $\alpha v_1 + \beta v_2 \in \ker(\phi)$. Por lo tanto, $\ker(\phi)$ es un subespacio de $V$. □

**Teorema 3.24.** Sea $\phi : V \to W$ una transformación lineal de espacios vectoriales sobre $F$.

**(1)** $\phi$ es inyectiva si y sólo si $\ker(\phi) = \{\mathbf{0}_V\}$.

**(2)** $\phi$ es sobreyectiva si y sólo si $\mathrm{Im}(\phi) = W$.

**Demostración.**

**(1)** Demostraremos cada implicación.

**(⇒)** Supongamos que $\phi$ es inyectiva. Sea $v \in \ker(\phi)$. Entonces, $\phi(v) = \mathbf{0}_W = \phi(\mathbf{0_V})$. Puesto que $\phi$ es inyectiva, tenemos que $v = \mathbf{0}_V$. Por lo tanto $\ker(\phi) = \{\mathbf{0}_V\}$.

**(⇐)** Supongamos que $\ker(\phi) = \{\mathbf{0}_V\}$. Sean $v, w \in V$, tales que $\phi(v) = \phi(w)$. Puesto que $\phi$ es una transformación lineal, tenemos que

$$\phi(v - w) = \phi(v) - \phi(w) = \mathbf{0}_W.$$

Esto significa que $v - w \in \ker(\phi) = \{\mathbf{0}_V\}$. Luego, $v - w = \mathbf{0}_V$ y $v = w$. Por lo tanto, $\phi$ es inyectiva.

**(2)** Esto es cierto por la definición de sobreyectividad.

□

**Ejemplo 3.25.** Consideremos la trasformación lineal $\phi : \mathbb{R}^3 \to \mathbb{R}^3$ dada por

$$\phi(x_1, x_2, x_3) = (x_1 + x_2, x_2 + x_3, x_1 + 2x_2 + x_3).$$

Describamos el kernel y la imagen de $\phi$.

**1)** *Kernel.* El vector $(x_1, x_2, x_3) \in \ker(\phi)$ si y sólo si $\phi(x_1, x_2, x_3) = (0, 0, 0)$, es decir si

$$(x_1 + x_2, x_2 + x_3, x_1 + 2x_2 + x_3) = (0, 0, 0),$$

lo que significa

$$\begin{aligned} x_1 + x_2 + 0x_3 &= 0 \\ 0x_1 + x_2 + x_3 &= 0 \\ x_1 + 2x_2 + x_3 &= 0 \end{aligned} \tag{3.1}$$

Por lo tanto, el kernel de $\phi$ es el conjunto de soluciones del sistema de ecuaciones lineales homogéneas (3.1). Sea $A$ la matriz aumentada de coeficientes del sistema (3.1). Entonces, realizando eliminación Gaussiana

$$A = \left(\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 2 & 1 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

El sistema es consistente y las soluciones son: $x_1 = \alpha$, $x_2 = -\alpha$, $x_3 = \alpha$, para cualquier $\alpha \in \mathbb{R}$. El conjunto de soluciones representa una línea paralela a $(1, -1, 1)$ a través del origen. Por lo tanto,

$$\ker(\phi) = \left\{(\alpha, -\alpha, \alpha) \in \mathbb{R}^3 : \alpha \in \mathbb{R}\right\}.$$

**2)** *Imagen.* La imagen de $\phi$ consiste en aquellos $(y_1, y_2, y_3) \in \mathbb{R}^3$ tales que $\phi((x_1, x_2, x_3)) = (y_1, y_2, y_3)$, para algún $(x_1, x_2, x_3) \in \mathbb{R}^3$. Esto significa que

$$\begin{aligned} x_1 + x_2 + 0x_3 &= y_1 \\ 0x_1 + x_2 + x_3 &= y_2 \\ x_1 + 2x_2 + x_3 &= y_3 \end{aligned} \tag{3.2}$$

Por lo tanto, la imagen de $\phi$ es la solución del sistema de ecuaciones lineales no homogéneas (5.3). Sea $B$ la matriz aumentada de coeficientes del sistema (5.3). Entonces, realizando eliminación Gaussiana

$$B = \left(\begin{array}{ccc|c} 1 & 1 & 0 & y_1 \\ 0 & 1 & 1 & y_2 \\ 1 & 2 & 1 & y_3 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 1 & 0 & y_1 \\ 0 & 1 & 1 & y_2 \\ 0 & 0 & 0 & y_3 - y_1 - y_2 \end{array}\right).$$

Este sistema tiene soluciones si y sólo si $y_3 - y_1 - y_2 = 0$. Por lo tanto

$$\mathrm{Im}(\phi) = \left\{(y_1, y_2, y_3) \in \mathbb{R}^3 : y_3 - y_1 - y_2 = 0\right\}.$$

Así, la imagen de $\phi$ es el plano $y_3 = y_1 + y_2$ en $\mathbb{R}^3$.

**Teorema 3.26.** Sean $\phi : U \to V$ y $\tau : V \to W$ transformaciones lineales entre espacios vectoriales sobre un campo $F$ tales que $\tau \circ \phi$ es un isomorfismo. Entonces,

$$V = \mathrm{Im}\,(\phi) \oplus \ker(\tau).$$

**Demostración.** Haremos la demostración en dos pasos.

**1)** Demostraremos que $\mathrm{Im}\,(\phi) + \ker(\tau) = V$. Sea $v \in V$ un vector arbitrario. Como $\tau(v) \in W$ y $\tau \circ \phi : U \to W$ es un isomorfismo, existe $u \in U$ tal que $(\tau \circ \phi)(u) = \tau(v)$. Sea $v' = \phi(u) \in \mathrm{Im}\,(\phi)$ y $v'' = v - v'$. Claramente,

$$v = v' + v'',$$

donde $v' \in \mathrm{Im}\,(\phi)$. Demostraremos que $v'' \in \ker(\tau)$:

$$\begin{aligned}\tau(v'') &= \tau(v - v') \\ &= \tau(v) - \tau(v') \\ &= (\tau \circ \phi)(u) - \tau(\phi(u)) \\ &= \mathbf{0}_W.\end{aligned}$$

Luego, $\mathrm{Im}\,(\phi) + \ker(\tau) = V$, pues $v$ era arbitrario.

**2)** Demostraremos que la suma es directa comprobando que $\mathrm{Im}\,(\phi) \cap \ker(\tau) = \{\mathbf{0}_V\}$. Sea $v \in \mathrm{Im}\,(\phi) \cap \ker(\tau)$. Como $v \in \mathrm{Im}\,(\phi)$, existe $u \in U$ tal que $\phi(u) = v$. Como $v \in \ker(\tau), \tau(v) = \mathbf{0}_W$. Luego,

$$(\tau \circ \phi)(u) = \tau(v) = \mathbf{0}_W = (\tau \circ \phi)(\mathbf{0}_U).$$

Debido a que $\tau \circ \phi$ es un isomorfismo, $u = \mathbf{0}_U$. Luego, $v = \phi(u) = \phi(\mathbf{0}_U) = \mathbf{0}_V$. Por el Teorema 2.35, $V = \mathrm{Im}\,(\phi) \oplus \ker(\tau)$.

□

## 3.4. El espacio de las transformaciones lineales

**Definición 3.27 ($\mathbf{Hom}(V, W)$).** Sean $V$ y $W$ espacios vectoriales sobre el mismo campo $F$. Denotamos al conjunto de todas las transformaciones lineales de $V$ en $W$ por

$$\mathrm{Hom}(V, W) = \{\phi : V \to W : \phi \text{ es una transformación lineal}\}.$$

Cuando $V = W$, $\mathrm{Hom}(V, W)$ se abrevia $\mathrm{End}(V)$.

**Ejemplo 3.28 ($\hat{\mathbf{0}} \in \mathbf{Hom}(V, W)$).** La transformación cero es una transformación lineal (Ejemplo 3.4). En consecuencia $\hat{\mathbf{0}} \in \mathrm{Hom}(V, W)$, $\hat{\mathbf{0}}(v) = \mathbf{0}_W$ para todo $v \in V$.

**Ejemplo 3.29 ($I \in \mathbf{Hom}(V)$).** La función identidad Identidad es una transformación lineal (Ejemplo 3.3). En consecuencia $I \in \mathrm{Hom}(V)$, $I_V(v) = v$ para todo $v \in V$.

**Ejemplo 3.30 ($\varphi \otimes z$).** Si $\varphi : V \to F$ es un funcional lineal de $V$ (Definición 3.2) y $z$ un vector fijo en $W$, la función $\phi : V \to W$ definida por $\phi(v) = \varphi(v)z$ para todo $v \in V$ es una transformación lineal. Para comprobar esto, sean $v, w \in V$ y $\alpha, \beta \in F$ elementos arbitrarios. Entonces,

$$\begin{aligned}\phi(\alpha v + \beta w) &= \varphi(\alpha v + \beta w)z \\ &= [\alpha\varphi(v) + \beta\varphi(w)]\, z \\ &= \alpha\varphi(v)z + \beta\varphi(w)z \\ &= \alpha\phi(v) + \beta\phi(v).\end{aligned}$$

La dependencia de $\phi$ sobre $\varphi$ se expresa escribiendo $\phi = \varphi \otimes z$, el símbolo sugiere un tipo de producto de $\varphi$ con $z$. En consecuencia $\varphi \otimes z \in \mathrm{Hom}(V, W)$, $(\varphi \otimes z)(v) = \varphi(v)z$ para todo $v \in V$.

Ahora dotaremos de una estructura de espacio vectorial al conjunto $\mathrm{Hom}(V, W)$ sobre el campo $F$. Para ello se tendrá que definir adecuadamente la suma y la multiplicación escalar en $\mathrm{Hom}(V, W)$.

**Definición 3.31.** Sean $\phi, \varphi \in \mathrm{Hom}(V, W)$ y $\alpha \in F$. Definimos la suma es la función $\phi + \varphi : V \to W$ definida por

$$(\phi + \varphi)(v) = \phi(v) + \varphi(v), \quad \forall v \in V. \tag{3.3}$$

La multiplicación escalar es la función $\alpha\phi : V \to W$ definida por

$$(\alpha\phi)(v) = \alpha\phi(v), \quad \forall v \in V \text{ y } \alpha \in F. \tag{3.4}$$

**Lema 3.32.** Si $\phi, \varphi \in \mathrm{Hom}(V, W)$ y $\alpha \in F$, entonces $\phi + \varphi \in \mathrm{Hom}(V, W)$ y $\alpha\phi \in \mathrm{Hom}(V, W)$.

**Demostración.** Demostraremos primero que la función $\phi + \varphi : V \to W$ es una transformaciones lineal. Sean $v, w \in V$ y $\alpha, \beta \in F$, entonces

$$\begin{aligned}(\phi + \varphi)(\alpha v + \beta w) &= \phi(\alpha v + \beta w) + \varphi(\alpha v + \beta w) && [\text{por (3.3)}] \\ &= [\alpha\phi(v) + \beta\phi(w)] + [\alpha\varphi(v) + \beta\varphi(w)] && [\because \phi, \varphi \text{ son transformaciones lineales}] \\ &= \alpha\,[\phi(v) + \varphi(v)] + \beta\,[\phi(w) + \varphi(w)] \\ &= \alpha(\phi + \varphi)(v) + \beta(\phi + \varphi)(w) && [\text{por (3.3)}]\,.\end{aligned}$$

Por lo tanto $\phi + \varphi$ es una transformación lineal de $V$ en $W$ y $\phi + \varphi \in \mathrm{Hom}(V, W)$.

Demostremos ahora que la función $\alpha\phi : V \to W$ es una transformación lineal. Sean $v, w \in V$ y $\beta, \gamma \in F$, entonces

$$\begin{aligned}(\alpha\phi)(\beta v + \gamma w) &= \alpha\phi(\beta v + \gamma w) && \text{[por (3.4)]}\\ &= \alpha\left[\beta\phi(v) + \gamma\phi(w)\right] && [\because \phi, \text{ es una transformacion lineal}]\\ &= (\alpha\beta)\phi(v) + (\alpha\gamma)\phi(w)\\ &= \beta[\alpha\phi(v)] + \gamma[\alpha\phi(w)]\\ &= \beta[(\alpha\phi)(v)] + \gamma[(\alpha\phi)(w)] && \text{[por (3.4)]}\end{aligned}$$

Por lo tanto $\alpha\phi$ es una transformación lineal de $V$ en $W$ y $\alpha\phi \in \mathrm{Hom}(V, W)$. □

**Teorema 3.33.** Sean $V$ y $W$ espacios vectoriales sobre el mismo campo $F$, entonces $\mathrm{Hom}(V, W)$ es también un espacio vectorial sobre el campo $F$ con las operaciones de la Definición 3.31.

**Demostración.** Demostremos que $\mathrm{Hom}(V, W)$ cumple las propiedades de espacio vectorial

**(EV1)** Demostremos que $(\mathrm{Hom}(V, W), +)$ es un grupo abeliano. Si $\phi, \varphi, \psi \in \mathrm{Hom}(V, W)$, entonces

**(1)** *Cerradura*: por el Lema 3.32, $\mathrm{Hom}(V, W)$ es cerrado respecto a la suma.

**(2)** *Asociatividad de la suma*: tenemos que

$$\begin{aligned}\left[\phi + (\varphi + \psi)\right](v) &= \phi(v) + (\varphi + \psi)(v)\\ &= \phi(v) + [\varphi(v) + \psi(v)]\\ &= [\phi(v) + \varphi(v)] + \psi(v)\\ &= (\phi + \varphi)(v) + \psi(v)\\ &= [(\phi + \varphi) + \psi](v).\end{aligned}$$

**(3)** *Existencia de la identidad aditiva*: La transformación cero $\hat{\mathbf{0}}$ satisface

$$\begin{aligned}(\hat{\mathbf{0}} + \phi)(v) &= \hat{\mathbf{0}}(v) + \phi(v)\\ &= \mathbf{0}_W + \phi(v)\\ &= \phi(v).\end{aligned}$$

Por lo tanto $\hat{\mathbf{0}}$ es la identidad aditiva en $\mathrm{Hom}(V, W)$.

**(4)** *Existencia del inverso aditivo*: El negativo de una transformación lineal también es una transformación lineal (Ejemplo 3.5). De esta manera, si $\phi \in \mathrm{Hom}(V, W)$ tenemos que $-\phi \in \mathrm{Hom}(V, W)$. Entonces

$$\begin{aligned}(-\phi + \phi)(v) &= (-\phi)(v) + \phi(v)\\ &= -\phi(v) + \phi(v)\\ &= \mathbf{0}_W\\ &= \hat{\mathbf{0}}(v).\end{aligned}$$

Así cada elemento en $\mathrm{Hom}(V, W)$ posee inverso aditivo.

**(5)** *Conmutatividad de la suma*: Sean $\phi, \varphi \in \mathrm{Hom}(V, W)$. Si $v \in V$, entonces

$$\begin{aligned}(\phi + \varphi)(v) &= \phi(v) + \varphi(v) \\ &= \varphi(v) + \phi(v) \\ &= (\varphi + \phi)(v).\end{aligned}$$

Por lo tanto $\mathrm{Hom}(V, W)$ es un grupo abeliano con la operación suma definida.

**(EV2)** La verificación de las propiedades restantes es igualmente sencillo, se dejan al lector como ejercicio.

□

***Palabras clave:*** *transformación lineal, composición de transformaciones lineales, isomorfismo, relación de isomorfía, imagen y kernel de una transformación lineal, espacio de las transformaciones lineales.*

## 3.5. Ejercicios

**Ejercicio 3.34.** Determina si las siguientes funciones son transformaciones lineales. Justifica detalladamente tu respuesta.

1. $\tau : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ definida por $\tau(x_1, x_2, x_3) = (x_1, x_2)$.
2. $\sigma : \mathbb{R}^2 \longrightarrow \mathbb{R}$ definida por $\sigma(x_1, x_2) = x_1 x_2$.
3. $\varsigma : \mathbb{R}^2 \longrightarrow \mathbb{R}$ definida por $\varsigma(x_1, x_2) = x_1 + x_2$.
4. $\phi : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ definida por $\phi(x_1, x_2, x_3) = (x_1 + 1, x_2, x_3)$.
5. $\psi : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ definida por $\psi(x_1, x_2, x_3) = (x_1 + 5x_3, 9x_2)$.
6. $\exp : \mathbb{R} \longrightarrow \mathbb{R}$ definida por $\exp(x) = e^x$, $\forall x \in \mathbb{R}$.
7. $\phi : \mathbb{Z}_3 \longrightarrow \mathbb{Z}_3$ definida por $\varphi(x) = x^2$, $\forall x \in \mathbb{Z}_3$.
8. $\varphi : \mathbb{Z}_3 \longrightarrow \mathbb{Z}_3$ definida por $\varphi(x) = x^3$, $\forall x \in \mathbb{Z}_3$.
9. $I : \mathbb{R}[x] \longrightarrow \mathbb{R}[x]$ definida por

$$I(f(x)) = \int_0^x f(t)\, dt$$

donde $f(x) \in \mathbb{R}[x]$.

**Ejercicio 3.35.** Encuentra el rango y el kernel de las siguientes transformaciones lineales y determina si son funciones inyectivas o sobreyectivas.

1. $\phi_\gamma : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ definida por $\phi_\gamma(x_1, x_2) = (\gamma x_1, \gamma x_2)$, donde $\gamma \in \mathbb{R}$.
2. $\tau : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ definida por $\tau(x_1, x_2, x_3) = (x_3, x_2)$.
3. $\varphi : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ definida por
   $\varphi(x_1, x_2, x_3) = (2x_1 - x_2, 5x_2 + x_3, 2x_1 + 4x_2 + x_3)$.
4. $\sigma : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ definida por $\sigma(x_1, x_2, x_3) = (x_1 + x_3, x_1 + x_3, x_2 + x_3)$.
5. $\phi : \mathbb{R}^3 \longrightarrow \mathbb{R}$ definida por $\phi(x_1, x_2, x_3) = 2x_1 - x_2 + 3x_3$.
6. $\varsigma : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ definida por $\varsigma(x_1, x_2) = (x_1, x_2, x_1 + x_2)$.
7. $\tau : \mathbb{R}^4 \longrightarrow \mathbb{R}^2$ definida por $\tau(x_1, x_2, x_3, x_4) = (x_1 - x_4, x_2 - x_3)$.
8. $\varphi : \mathbb{Z}_3 \longrightarrow \mathbb{Z}_3$ definida por $\varphi(x) = x^3$, $\forall x \in \mathbb{Z}_3$.
9. $D : \mathbb{R}[x] \longrightarrow \mathbb{R}[x]$, donde $D$ es la derivada formal de polinomios.

**Ejercicio 3.36.** Determina si las siguientes funciones son automorfismos. Justifica detalladamente tu respuesta.

1. $\phi : \mathbb{R} \longrightarrow \mathbb{R}$ definida como $\phi(x) = 5x$.
2. $\xi : \mathbb{R} \longrightarrow \mathbb{R}$ definida como $\xi(x) = -2$.
3. $\varphi : \mathbb{R}^4 \longrightarrow \mathbb{R}^3$ definida como $\varphi(x_1, x_2, x_3, x_4) = (x_1, x_2, x_3)$.
4. $\tau : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ definida como $\tau(x_1, x_2) = (x_2, x_1)$.
5. $\omega : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ definida como $\omega(x_1, x_2) = (x_1, 0)$.
6. $\sigma : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ definida como
   $\sigma(x_1, x_2, x_3) = (-3x_1 + 2x_2 + x_3, 2x_1 - x_2, x_1 + x_3)$.
7. $\xi : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ definida como
   $\xi(x_1, x_2, x_3) = (2x_1 + 3x_2 - x_3, 3x_1 + 3x_3, x_2 + x_3)$.
8. $\sigma : \mathbb{C} \longrightarrow \mathbb{C}$ definida como $\sigma(x + iy) = x - iy$.
9. $\tau : \mathbb{C} \longrightarrow \mathbb{C}$ definida como $\tau(x + iy) = (x + iy)^2$.

**Ejercicio 3.37.** Demuestra los siguientes isomorfismos de espacios vectoriales.

1. $\mathbb{Z}_3^3 \cong (\mathbb{Z}_3)^{\mathbb{Z}_3}$.
2. $\mathbb{R}^4 \cong \mathbb{R}[x]_{\leq 3}$, donde $\mathbb{R}[x]_{\leq 3}$ es el espacio vectorial de polinomios de grado menor o igual que 3 con coeficientes en $\mathbb{R}$.

**Ejercicio 3.38.** Demuestra que la función $\phi : \mathbb{R}^3 \to \mathbb{R}^2$ definida como

$$\phi(x_1, x_2, x_3) = (3x_1 - 2x_2 + x_3, x_1 - 3x_2 - 2x_3),$$

es una transformación lineal de $\mathbb{R}^3$ sobre $\mathbb{R}^2$. Encuentra y describe el kernel y la imagen de $\phi$.

**Ejercicio 3.39.** Demuestra que la función $\phi : \mathbb{R}^2 \to \mathbb{R}^3$ definida como

$$\phi(x_1, x_2) = (x_1 + x_2,\ x_1 - x_2, x_2),$$

es una trasformación lineal de $\mathbb{R}^2$ sobre $\mathbb{R}^3$. Encuentra y describe el kernel y la imagen de $\phi$.

**Ejercicio 3.40.** Sean $\mathbb{C}$ el campo de los números complejos, muestra que la función $\phi : \mathbb{C}^3 \to \mathbb{C}^3$ definida como

$$\phi(x_1, x_2, x_3) = (x_1 - x_2 + 2x_3, 2x_1 + x_2 - x_3, -x_1 - 2x_2),$$

es una transformación lineal de $\mathbb{C}^3$ sobre $\mathbb{C}^3$. Encuentra y describe el kernel y la imagen de $\phi$.

**Ejercicio 3.41.** Sea $V = M_{n\times n}(\mathbb{R})$ el espacio vectorial de las matrices de $n \times n$ sobre el campo $\mathbb{R}$ y sea $B$ una una matriz fija de $n \times n$. Si

$$\phi(A) = AB - A, \ \forall A \in M_{n\times n}(\mathbb{R}).$$

Demuestra que $\phi$ es una transformación lineal de $V$ sobre $V$.

**Ejercicio 3.42.** Sean $U$ y $V$ espacios vectoriales sobre el campo $F$ y sean $\phi, \varphi$ dos transformaciones lineales de $U$ sobre $V$ sean $\alpha_1$ y $\alpha_2$ dos elementos de $F$. Entonces el mapeo $\psi$ definido como

$$\psi(x) = \alpha_1\phi(x) + \alpha_2\varphi(x), \ \forall x \in U,$$

es una transformación lineal.

**Ejercicio 3.43.** Sea $\phi : \mathbb{R}^3 \to \mathbb{R}^3$ una transformación lineal definida como

$$\phi(x_1, x_2, x_3) = (2x, 4x - y, 2x + 3y - z).$$

Demuestra que $\phi$ es invertible y encuentra las expresiones para $\phi^{-1}$ y $(\phi^{-1})^2$.

**Ejercicio 3.44.** Sean $S$ y $T$ subespacios de un espacio vectorial $V$ sobre $F$ tales que $S \cap T = \{\mathbf{0}\}$. Demuestra que la suma directa interna $S \oplus T$ y la suma directa externa $S \boxplus T$ son espacios vectoriales isomorfos.

**Ejercicio 3.45.** Si $T : U \to V$ y $S : V \to W$ son transformaciones lineales, demuestra que $\mathrm{Im}(S \circ T) \subseteq \mathrm{Im}(S)$ y $\ker(T) \subseteq \ker(S \circ T)$.

**Ejercicio 3.46.** Sea $\phi : V \to V$ una transformación lineal tal que $\mathrm{Im}(\phi) \subseteq \ker(\phi - I_V)$, donde $I_V$ es el endomorfismo identidad. Demuestra que $\phi^2 = \phi$. Recuerda que $\phi^2 = \phi \circ \phi$.

**Ejercicio 3.47.** Sea $V$ un espacio vectorial sobre $F$ y $\phi$ una transformación lineal de $V$ sobre $V$. Demuestra que las siguientes dos proposiciones acerca de $\phi$ son equivalentes

**(a)** La intersección de la imagen de $\phi$ y el kernel de $\phi$ es el subespacio cero de $V$, es decir $\mathrm{Im}(\phi) \cap \ker(\phi) = \{\mathbf{0}\}$.

**(b)** $\phi(\phi(x)) = \mathbf{0}$ entonces $\phi(x) = \mathbf{0}$.

**Ejercicio 3.48.** Completa la demostración del Teorema 3.33.

**Ejercicio 3.49.** Sea $V$ el espacio vectorial real o complejo y sea $\phi : V \to V$ un endomorfismo tal que $\phi^2 = I_V$. Definamos

$$S = \{x \in V : \phi(x) = x\}, \quad T = \{x \in V : \phi(x) = -x\}$$

Demuestra que $S$ y $T$ son subespacios de $V$ tales que $V = S \oplus T$. Sugerencia: para cada vector $x$, $x = \frac{1}{2}[x + \phi(x)] + \frac{1}{2}[x - \phi(x)]$.

# 4

# Espacio vectorial cociente

Sea $V$ un espacio vectorial sobre un campo $F$ y sea $S$ un subespacio de $V$. Formaremos un nuevo espacio vectorial a partir de $V$ y $S$ llamado el espacio cociente de $V$ por $S$, y lo escribiremos como $V/S$.

**Definición 4.1 (congruencia módulo $S$).** Sea $S$ un subespacio de un espacio vectorial $V$ sobre un campo $F$. Decimos que dos vectores $w, v \in V$ son *congruentes módulo* $S$, y escribimos $w \equiv v \bmod S$, si $w - v \in S$ .

**Lema 4.2 (congruencia módulo $S$).** Sea $S$ un subespacio de un espacio vectorial $V$ sobre un campo $F$. La relación de congruencia módulo $S$ es una relación de equivalencia sobre $V$.

**Demostración.** Demostraremos que se cumplen las propiedades de relación de equivalencia.

**1)** *Reflexividad.* Puesto que $\mathbf{0} \in S$, tenemos que $v - v = \mathbf{0} \in S$, para todo $v \in V$. Por lo tanto, $v \equiv v \bmod S$, para todo $v \in V$.

**2)** *Simetría.* Supongamos que $w \equiv v \bmod S$. Entonces $w - v \in S$, y, multiplicando por el escalar $-1 \in F$, obtenemos $(-1)(w - v) = v - w \in S$. Por lo tanto, $v = w \bmod S$.

**3)** *Transitividad.* Supongamos que $u \equiv v \bmod S$ y $v \equiv w \bmod S$. Entonces $u - v \in S$ y $v - w \in S$. Puesto que $S$ es un subespacio de $V$, $(u - v) + (v - w) = u - w \in S$. Por lo tanto, $u \equiv w \bmod S$.

□

Denotemos la clase de equivalencia de $v \in V$ para la relación congruencia módulo $S$ por $[v]_S$. Para describir cómo lucen estas clases de equivalencia, observemos que

$$\begin{aligned}
[v]_S &= \{w \in V : w \equiv v \bmod S\} \\
&= \{w \in V : w - v \in S\} \\
&= \{w \in V : w = v + s \text{ para algún } s \in S\} \\
&= \{v + s : s \in S\}\,.
\end{aligned}$$

La notación $v + S$ es una forma más intuitiva de representar la clase de equivalencia $[v]_S$:

$$[v]_S = v + S = \{v + s : s \in S\}.$$

**Definición 4.3 (clase lateral).** Sea $S$ un subespacio de un espacio vectorial $V$ sobre un campo $F$. Para cualquier $v \in V$, la clase de equivalencia $v + S$ se llama la *clase lateral de $S$ en $V$ con representante $v$*.

Observemos que la clase lateral $v + S$ puede ser representada por cualquiera de sus elementos en el sentido de que $(v + s) + S = v + S$, para todo $s \in S$. La clase lateral $v + S$ es un subconjunto de $V$, pero en general no es un subespacio de $V$.

**Proposición 4.4 (clases laterales).** Sea $S$ un subespacio de un espacio vectorial $V$ sobre un campo $F$. Sean $u, v \in V$ vectores arbitrarios.

**(1)** El conjunto de clases laterales de $S$ en $V$ forma una partición de $V$.

**(2)** $u + S = v + S$ si y sólo si $u - v \in S$.

**(3)** $v + S$ es un subespacio de $V$ si y sólo si $v \in S$.

**Demostración.**

**(1)** Las clases de equivalencia de una relación de equivalencia sobre un conjunto siempre forman una partición del conjunto (ver Lema 0.7 **(3)** y **(4)**).

**(2)** Sabemos que dos clases de equivalencia $u + S$ y $v + S$ son iguales si y sólo si $u \equiv v \bmod S$ (ver Lema 0.7 **(2)**). Esta última condición es equivalente a $u - v \in S$.

**(3)** Supongamos que $v + S \leq V$. Entonces, $\mathbf{0} \in v + S$, así que $v + s = \mathbf{0}$, para algún $s \in S$. Luego, $v = -s \in S$. Supongamos ahora que $v \in S$. Por la propiedad **(2)**, $v + S = \mathbf{0} + S = S$, ya que $v - \mathbf{0} \in S$. Como $S$ es un subespacio por hipótesis, entonces $v + S = S$ es un subespacio.

□

**Definición 4.5 (espacio cociente).** Sea $S$ un subespacio de un espacio vectorial $V$ sobre un campo $F$. El *espacio cociente*, denotado por $V/S$, es el conjunto de todas las clases laterales de $S$ en $V$. En otras palabras,

$$V/S = \{v + S : v \in V\}.$$

**Observación 4.6.** La notación $V/S$ se lee "espacio cociente de $V$ módulo $S$".

Definimos en $V/S$ las operaciones de suma y multiplicación escalar como sigue: para toda $(u + S), (v + S) \in V/S$ y $\alpha \in F$,

$$\begin{aligned} (u + S) + (v + S) &= (u + v) + S, \\ \alpha\,(u + S) &= \alpha u + S. \end{aligned} \tag{4.1}$$

Una dificultad aparece en la definición de estas operaciones. Debido a que los elementos de $V/S$ son clases de equivalencia (clases laterales), para demostrar que las operaciones dadas en (4.1) están bien definidas es necesario comprobar que el resultado en cada operación no depende de los representantes de las clases. Esta situación es similar a la ocurrida con el grupo cíclico $\mathbb{Z}_m$.

**Proposición 4.7.** Las operaciones $+ : V/S \times V/S \to V/S$ y $\cdot : F \times V/S \to V/S$ dadas en (4.1) están bien definidas.

**Demostración.**

**(1)** Para demostrar que la suma está bien definida, debemos comprobar que si $u + S = u' + S$ y $v + S = v' + S$, entonces $(u + S) + (v + S) = (u' + S) + (v' + S)$. Por la Proposición 4.4,

$$\begin{aligned} u + S = u' + S &\Rightarrow u - u' \in S \\ v + S = v' + S &\Rightarrow v - v' \in S. \end{aligned}$$

Como $S$ es un subespacio,

$$(u - u') + (v - v') \in S.$$

La expresión anterior puede escribirse como

$$(u + v) - (u' + v') \in S,$$

y esto implica que

$$(u + v) + S = (u' + v') + S,$$

Por lo tanto,

$$(u + S) + (v + S) = (u' + S) + (v' + S)\,.$$

**(2)** Para demostrar que la multiplicación escalar está bien definida, debemos comprobar que si $u + S = u' + S$, entonces $\alpha(u + S) = \alpha(u' + S)$, para todo $\alpha \in F$. Como $u + S = u' + S$ implica que $u - u' \in S$, entonces

$$\alpha\,(u - u') \in S.$$

La expresión anterior puede escribirse como

$$\alpha u - \alpha u' \in S,$$

y esto implica que

$$\alpha u + S = \alpha u' + S.$$

Por lo tanto, $\alpha(u + S) = \alpha(u' + S)$.

□

**Teorema 4.8 ($V/S$ es espacio vectorial).** Sea $S$ un subespacio de un espacio vectorial $V$ sobre un campo $F$. El espacio cociente $V/S$, junto con las operaciones definidas en (4.1), es un espacio vectorial sobre $F$.

**Demostración.**

**(EV1)** Mostraremos primero que $V/S$ es un grupo abeliano con respecto a la suma. Sean $u+S, v+S, w+S \in V/S$ elementos arbitrarios.

**(G0)** *Cerradura*: Por definición, $(u+S)+(v+S) = (u+v)+S \in V/S$.

**(G1)** *Asociatividad*: Usando la asociatividad de vectores en $V$, deducimos lo siguiente:

$$\begin{aligned}(u+S)+[(v+S)+(w+S)] &= (u+S)+[(v+w)+S] \\ &= [u+(v+w)]+S \\ &= [(u+v)+w]+S \\ &= [(u+v)+S]+(w+S) \\ &= [(u+S)+(v+S)]+(w+S)\,.\end{aligned}$$

**(G2)** *Identidad aditiva*: Si $\mathbf{0}$ es el vector cero de $V$, tenemos que

$$(\mathbf{0}+S)+(u+S) = (u+\mathbf{0})+S = u+S.$$

Por lo tanto, $\mathbf{0}+S = S$, es la identidad aditiva en $V/S$.

**(G3)** *Inversos aditivos*: Para cualquier $u+S \in V/S$, vemos que $(-u)+S$ satisface

$$(u+S)+((-u)+S) = (u-u)+S = \mathbf{0}+S.$$

Por lo tanto, $(-u)+S$ es el inverso aditivo de $u+S$.

**(G4)** *Conmutatividad*: Usando la conmutatividad de la suma en $V$, deducimos lo siguiente:

$$\begin{aligned}(u+S)+(v+S) &= (u+v)+S \\ &= (v+u)+S \\ &= (v+S)+(u+S)\,.\end{aligned}$$

**(EV2)** Sean $\alpha, \beta \in F$ y $u+S, v+S \in V/S$ elementos arbitrarios.

**(1)**

$$\begin{aligned}\alpha\,[(u+S)+(v+S)] &= \alpha\,[(u+v)+S] \\ &= \alpha\,(u+v)+S \\ &= (\alpha u+\alpha v)+S \\ &= (\alpha u+S)+(\alpha v+S) \\ &= \alpha\,(u+S)+\alpha\,(v+S)\,.\end{aligned}$$

**(2)**

$$\begin{aligned}(\alpha+\beta)(u+S) &= (\alpha+\beta)u+S\\ &= (\alpha u+\beta u)+S\\ &= (\alpha u+S)+(\beta u+S)\\ &= \alpha(S+u)+\beta(S+u).\end{aligned}$$

**(3)**

$$\begin{aligned}(\alpha\beta)(u+S) &= (\alpha\beta)u+S\\ &= \alpha(\beta u)+S\\ &= \alpha(\beta u+S)\\ &= \alpha[\beta(u+S)].\end{aligned}$$

**(4)**

$$1(u+S) = 1u+S = u+S.$$

□

**Ejemplo 4.9 (interpretación geométrica de $V/S$).** Veamos ahora la interpretación geométrica de algunos espacios cocientes.

**(1)** Sea $V=\mathbb{R}^2$ y $S=\{(x,0): x\in\mathbb{R}\}$. En la representación geométrica de $\mathbb{R}^2$ como plano cartesiano, $S$ representa el eje $x$. El espacio cociente

$$\mathbb{R}^2/S = \{(x,y)+S : (x,y)\in\mathbb{R}^2\}$$

es el conjunto de todas las líneas en $\mathbb{R}^2$ que son paralelas al eje $x$. Esto es porque para cualquier vector $v=(v_1,v_2)\in\mathbb{R}^2$, la clase lateral $v+S$ es igual a

$$v+S = (v_1,v_2)+S = \{(v_1+x,v_2): x\in\mathbb{R}\};$$

es decir, $v+S$ es la línea $y=v_2$ paralela al eje $x$. Esta línea está arriba o abajo del eje $x$ de acuerdo a $v_2>0$ o $v_2<0$. Si $v_2=0$, $v+S=S$ coincide con el eje $x$.

**(2)** Si $V=\mathbb{R}^3$ y $S=\{(x,y,0)\in\mathbb{R}^3\}$, entonces $S$ es el plano $xy$ y para cualquier $v=(v_1,v_2,v_3)\in\mathbb{R}^3$, la clase lateral $v+S$ representa geométricamente el plano paralelo al plano $xy$ a través del punto $v=(v_1,v_2,v_3)$ a una distancia $v_3$ del plano $xy$ (arriba o abajo del plano $xy$ de acuerdo a $v_3>0$ o $v_3<0$).

**(3)** Sea $V=\mathbb{R}^3$ y $S=\{(x,y,z)\in\mathbb{R}^3 : 5x-4y+3z=0 \text{ y } 2x-3y+4z=0\}$. Para cualquier $v=(v_1,v_2,v_3)\in\mathbb{R}^3$, la clase lateral $v+S\in V/S$ representa geométricamente la línea paralela a la línea determinada por la intersección de los dos planos:

$$\begin{aligned}&5(x-v_1)-4(y-v_2)+3(z-v_3)=0, \quad \text{y}\\ &2(x-v_1)-3(y-v_2)+4(z-v_3)=0.\end{aligned}$$

**Definición 4.10 (proyección canónica).** Sea $S$ un subespacio de un espacio vectorial $V$ sobre un campo $F$. La función $\mathrm{pr} : V \to V/S$ definida por

$$\mathrm{pr}(v) = v + S,$$

es llamada la *proyección canónica* de $V$ en $V/S$.

**Proposición 4.11 (proyección canónica).** La proyección canónica $\mathrm{pr} : V \to V/S$ es una transformación lineal sobreyectiva.

**Demostración.** Para cualquier $v, w \in V$ y $\alpha, \beta \in F$, se cumple que

$$\begin{aligned}\mathrm{pr}(\alpha v + \beta w) &= (\alpha v + \beta w) + S \\ &= (\alpha v + S) + (\beta w + S) \\ &= \alpha(v + S) + \beta(w + S) \\ &= \alpha\,\mathrm{pr}(v) + \beta\,\mathrm{pr}(w).\end{aligned}$$

Por lo tanto, pr es una transformación lineal.

Para demostrar que pr es sobreyectiva, sea $v + S$ cualquier elemento de $V/S$; entonces, $v \in V$ es la preimagen de $v + S$ bajo la proyección canónica: $\mathrm{pr}(v) = v + S$. □

**Proposición 4.12 (proyección canónica inyectiva).** La proyección canónica pr es inyectiva si y sólo si $S = \{\mathbf{0}\}$.

**Demostración.** Si $S \neq \{\mathbf{0}\}$ debe existir un $u \in S$ tal que $u \neq \mathbf{0}$, como $u - \mathbf{0} \in S$ entonces $u + S = \mathbf{0} + S$ (prop. 4.4), por lo que $\mathrm{pr}(u) = \mathrm{pr}(\mathbf{0})$ y pr no puede ser inyectiva. Para el converso, si $S = \{\mathbf{0}\}$ tenemos que para $v, w \in V$, el que $\mathrm{pr}(v) = \mathrm{pr}(w)$ implica que $v + S = w + S$ y de nuevo por la proposición 4.4 $v - w \in S$, por lo que $v - w = \mathbf{0}$ y $v = w$, así que pr es inyectiva. □

**Teorema 4.13 (primer teorema de isomorfía).** Sea $\phi : V \to W$ una transformación lineal entre espacios vectoriales sobre $F$. Entonces,

$$V/\ker(\phi) \cong \mathrm{Im}(\phi).$$

**Demostración.** Recordemos que los elementos del espacio cociente $V/\ker(\phi)$ son clases laterales:

$$V/\ker(\phi) = \{v + \ker(\phi) : v \in V\}.$$

Definamos una función $\xi : V/\ker(\phi) \to \mathrm{Im}(\phi)$ como

$$\xi\,(v + \ker(\phi)) = \phi(v).$$

Demostraremos que $\xi$ es un isomorfismo entre $V/\ker(\phi)$ y $\mathrm{Im}(\phi)$.

**(1)** *$\xi$ es una función bien definida.* Debido a que el dominio de $\xi$ es un conjunto de clases laterales, es necesario demostrar que la imagen de cada clase no depende del representante. Sean $v + \ker(\phi) = u + \ker(\phi)$ clases laterales iguales. Por la Proposición 4.4 **(2)**, $v - u = s \in \ker(\phi)$, para algún $s \in \ker(\phi)$. Aplicando $\phi$ a $v - u = s$ y usando que es una transformación lineal, obtenemos que

$$\phi(v - u) = \phi(v) - \phi(u) = \phi(s) = \mathbf{0}_W.$$

Por lo tanto, $\phi(v) = \phi(u)$. Esto demuestra que

$$\xi(v + \ker(\phi)) = \phi(v) = \phi(u) = \xi(u + \ker(\phi)).$$

**(2)** *$\xi$ es una transformación lineal.* Sean $v, u \in V$ y $\alpha, \beta \in F$ elementos arbitrarios. Usando el hecho que $\phi$ es una transformación lineal, deducimos lo siguiente:

$$\begin{aligned}
\xi\left(\alpha\left(u + \ker(\phi)\right) + \beta\left(v + \ker(\phi)\right)\right) &= \xi\left((\alpha u + \beta v) + \ker(\phi)\right) \\
&= \phi\left(\alpha u + \beta v\right) \\
&= \alpha\phi\left(u\right) + \beta\phi\left(v\right) \\
&= \alpha\xi\left(u + \ker(\phi)\right) + \beta\xi\left(v + \ker(\phi)\right).
\end{aligned}$$

**(3)** *$\xi$ es sobreyectiva.* Sea $w \in \mathrm{Im}(\phi)$ un elemento arbitrario. Por definición del conjunto $\mathrm{Im}(\phi)$, existe $v \in V$ tal que $\phi(v) = w$. Luego, $v + \ker(\phi)$ es la preimagen de $w$ bajo $\xi$:

$$\xi(v + \ker(\phi)) = \phi(v) = w.$$

**(4)** *$\xi$ es inyectiva.* Sean $v + \ker(\phi)$ y $u + \ker(\phi)$ elementos de $V/\ker(\phi)$ tales que $\xi(v + \ker(\phi)) = \xi(u + \ker(\phi))$. Por la definición de $\xi$, tenemos que $\phi(v) = \phi(u)$. Obtenemos las siguientes implicaciones:

$$\phi(v) = \phi(u) \;\Rightarrow\; \phi(v) - \phi(u) = \mathbf{0}_W \;\Rightarrow\; \phi(v - u) = \mathbf{0}_W \;\Rightarrow\; v - u \in \ker(\phi).$$

Por la Proposición 4.4 **(2)**, deducimos que $v + \ker(\phi) = u + \ker(\phi)$. Esto demuestra que $\xi$ es inyectiva.

□

***Palabras clave:*** *congruencia módulo un subespacio, clase lateral, espacio cociente, proyección canónica, primer teorema de isomorfía.*

## 4.1. Ejercicios

**Ejercicio 4.14.** Considera el espacio cociente $\mathbb{R}^3/\langle(1,0,1)\rangle$. Determina si las siguientes afirmaciones son verdaderas o falsas. Justifica detalladamente tu respuesta.

1. $(1,0,1) + \langle(1,0,1)\rangle = (0,0,0) + \langle(1,0,1)\rangle$.
2. $(1,0,0) + \langle(1,0,1)\rangle = (-1,0,1) + \langle(1,0,1)\rangle$.
3. $(1,0,0) + \langle(1,0,1)\rangle = (-1,0,-2) + \langle(1,0,1)\rangle$.
4. $((3,0,1) + \langle(1,0,1)\rangle) + ((-2,1,1) + \langle(1,0,1)\rangle) = (0,1,1) + \langle(1,0,1)\rangle$.
5. $((3,0,1) + \langle(1,0,1)\rangle) + ((-2,1,1) + \langle(1,0,1)\rangle) = (1,1,1) + \langle(1,0,1)\rangle$.

Ilustra la interpretación geométrica de cada inciso.

**Ejercicio 4.15.** Considera el espacio vectorial $V = \mathbb{Z}_2^3$ sobre $\mathbb{Z}_2$.

1. Si $S = \{(0,0,0),(1,1,0)\}$, comprueba que $S$ es un subespacio y escribe todos los elementos de $V/S$. ¿Cuántos elementos distintos hay en $V/S$?
2. Si $T = \{(0,0,0),(1,0,0),(0,1,1),(1,1,1)\}$, comprueba que $T$ es un subespacio y escribe todos los elementos de $V/T$. ¿Cuántos elementos distintos hay en $V/T$?

**Ejercicio 4.16.** Demuestra que si $V = S \oplus T$ entonces $V/S \cong T$. (*Sugerencia: Restringe la proyección canónica* $\mathrm{pr} : V \to V/S$ *a* $T$*, y calcula el kernel y la imagen de la restricción.*)

**Ejercicio 4.17.** Sea $V$ un espacio vectorial, sean $S \leq V$ y $T \leq V$, y sean $v, w \in V$ vectores fijos. Demuestra lo siguiente:

**a)** $v + S \subseteq w + T$ si y sólo si $S \subseteq T$ y $v - w \in T$.

**b)** $(v + S) \cap (w + T) \neq \emptyset$ si y sólo si $v - w \in S + T$.

**c)** si $z \in (v + S) \cap (w + T)$ entonces $(v + S) \cap (z + T) = z + S \cap T$.

(*Nota: Observa que este ejercicio compara clases laterales respecto a subespacios distintos*).

**Ejercicio 4.18.** Supongamos que $V = S \oplus T$ y sean $v, w \in V$. Demuestra que $(v+S) \cap (w+T)$ contiene exactamente un elemento. Sugerencia: use el ejercicio 4.17. Da una interpreta geométricamente si $V = \mathbb{R}^2$ y si $V = \mathbb{R}^3$.

**Ejercicio 4.19.** Sea $\phi : V \to W$ una transformación lineal, $S$ un subespacio de $V$ tal que $S \subseteq \ker\phi$, y $\mathrm{pr} : V \to V/S$ la proyección canónica. Demuestra lo siguiente:

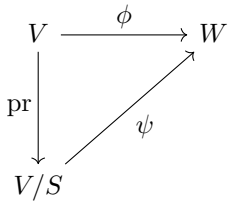



Figura 4.1:

**a)** Existe una (única) transformación lineal $\psi : V/S \to W$ tal que $\phi = \psi \circ \mathrm{pr}$ (vea Figura 4.1 ).

**b)** $\psi$ es inyectiva si y sólo si $S = \ker \phi$.

**Ejercicio 4.20.** Sea $S$ un subespacio de un espacio vectorial $V$ sobre un campo $F$. Sea $\mathrm{pr} : V \to V/S$ la proyección canónica, $\phi : V \to V$ una transformación lineal tal que $\phi(S) \subseteq S$. Demuestra que existe una (única) transformación lineal $\psi : V/S \to V/S$ tal que $\psi \circ \mathrm{pr} = \mathrm{pr} \circ \psi$ (vea Figura 4.2). Sugerencia: Aplica el Ejercicio 4.19 a la transformación lineal $\mathrm{pr} \circ \phi : V \to V/S$.

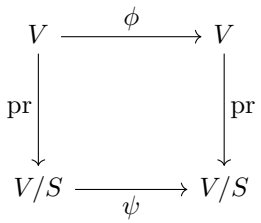



Figura 4.2:

**Ejercicio 4.21.** Donde existe un primero, existe un segundo. Demuestra el siguiente resultado llamado el *segundo teorema de isomorfía*. Sea $V$ un espacio vectorial, $S, T$ subespacios de $V$, $S \cap T$ su intersección y $S+T$ su suma. Entonces $S \cap T$ es un subespacio de $S$, $T$ es un subespacio de $S + T$, y

$$S/(S \cap T) \cong (S + T)/T.$$

Sugerencia: sea $\phi : V \to V/T$ el mapeo cociente y sea $\varphi = \phi|_S$ la restricción de $\phi$ a $S$, esto es, $\varphi : S \to V/T$ y $\varphi(x) = x + T$ para todo $x \in S$. Demuestra que $\varphi$ tiene imagen $(S + T)/T$ y kernel $S \cap T$.

# 5
# Bases y dimensión

## 5.1. Independencia Lineal

Recordemos que una combinación lineal de un subconjunto $A$ de un espacio vectorial $V$ sobre $F$ es una expresión de la forma

$$\alpha_1 v_1 + \alpha_2 v_2 + \dots \alpha_n v_n,$$

donde $\alpha_i \in F$, $v_i \in A$, $i = 1, \dots, n$. La combinación lineal es *trivial* si todos sus coeficientes son cero, y *no trivial* si al menos uno de los coeficientes es distinto de cero.

**Definición 5.1 (linealmente dependiente).** Sea $A$ un subconjunto de un espacio vectorial $V$ sobre $F$. Decimos que $A$ es *linealmente dependiente* si existe una combinación lineal de $A$ no trivial que sea igual al vector cero. En otras palabras, si existen vectores $v_1, v_2, \dots, v_n \in A$ y escalares $\alpha_1, \alpha_2, \dots, \alpha_n \in F$ no todos ceros, tales que

$$\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n = \sum_{i=1}^{n} \alpha_i v_i = \mathbf{0}. \tag{5.1}$$

**Definición 5.2 (linealmente independiente).** Decimos que un subconjunto $A$ de un espacio vectorial $V$ sobre $F$ es *linealmente independiente* si no existen combinaciones lineales de $A$ no triviales que sean iguales al vector cero. En otras palabras, si se cumple la siguiente implicación: si

$$\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n = \sum_{i=1}^{n} \alpha_i v_i = \mathbf{0}.$$

donde $\alpha_i \in F$, $v_i \in A$, $i = 1, \dots, n$, entonces $\alpha_i = 0$ para toda $i = 1, \dots, n$.

**Observación 5.3.** Por definición, el conjunto vacío $\varnothing$ es linealmente independiente, porque la frase "no existen combinaciones lineales de $\varnothing$ no triviales que sean iguales al vector cero" siempre es verdadera (no existen combinaciones lineales de $\varnothing$ en lo absuluto).

**Ejemplo 5.4.** Sea $V$ un espacio vectorial sobre $F$, y $A \subseteq V$ un subconjunto no vacío de vectores. Siempre que $\mathbf{0} \in A$, el conjunto $A$ será linealmente dependiente porque $\alpha\mathbf{0} = \mathbf{0}$ para cualquier $\alpha \in F$.

**Ejemplo 5.5.** Sea $F$ cualquier campo. En el espacio vectorial $F^n$, el conjunto de $n$ vectores

$$\{e_1 = (1, 0, \ldots, 0), e_2 = (0, 1, \ldots, 0), \ldots, e_n = (0, 0, \ldots, 1)\}, \tag{5.2}$$

es linealmente independiente. Para demostrar esto, supongamos que

$$\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n = \mathbf{0}$$

para algunos $\alpha_1, \alpha_2, \ldots, \alpha_n \in F$. Reescribimos la relación de arriba como

$$\alpha_1 (1, 0, \ldots, 0) + \alpha_2 (0, 1, \ldots, 0) + \ldots + \alpha_n (0, 0, \ldots, 1) = (0, 0, \ldots, 0).$$

Luego,

$$(\alpha_1, \alpha_2, \ldots, \alpha_n) = (0, 0, \ldots, 0),$$

y entonces $\alpha_1 = 0, \alpha_2 = 0, \ldots, \alpha_n = 0$. Por lo tanto, este conjunto de $n$ vectores de $F^n$ es linealmente independiente.

**Ejemplo 5.6.** El conjunto $A = \{(1, 0), (1, 1)\} \subseteq \mathbb{R}^2$ es linealmente independiente. Para demostrar esto, supongamos que

$$\alpha_1 (1, 0) + \alpha_2 (1, 1) = (0, 0),$$

para algunos $\alpha_i \in \mathbb{R}$. Entonces,

$$(\alpha_1 + \alpha_2, \alpha_2) = (0, 0)$$

lo que implica que

$$\alpha_1 + \alpha_2 = 0 \text{ y } \alpha_2 = 0.$$

Por lo tanto, $\alpha_1 = \alpha_2 = 0$.

**Ejemplo 5.7.** El subconjunto $A = \{(1, 0), (1, 1), (0, 1)\}$ de $\mathbb{R}^2$ es linealmente dependiente sobre $\mathbb{R}$ porque

$$(1, 0) - (1, 1) + (0, 1) = (0, 0) = \mathbf{0}.$$

**Ejemplo 5.8.** El subconjunto $\left\{1, x, 1 + x + x^2\right\}$ de $\mathbb{R}[x]$ es linealmente independiente. Para mostrar esto, sean $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$ escalares tales que

$$\alpha_1 (1) + \alpha_2 x + \alpha_3 \left(1 + x + x^2\right) = 0.$$

Entonces,

$$(\alpha_1 + \alpha_3) + (\alpha_2 + \alpha_3) x + \alpha_3 x^2 = 0,$$

de esta manera tenemos que $\alpha_1 + \alpha_3 = 0$, $\alpha_2 + \alpha_3 = 0$, $\alpha_3 = 0$ de donde $\alpha_1 = 0, \alpha_2 = 0, \alpha_3 = 0$. Por lo tanto, el subconjunto $\{1, x, 1 + x + x^2\} \subseteq \mathbb{R}[x]$ es linealmente independiente.

**Observación 5.9.** En ocasiones es conveniente hablar de *listas* de vectores linealmente independientes, en lugar de conjuntos, usando una definición análoga. Una *lista* de vectores de un espacio vectorial $V$ es simplemente una sucesión finita $v_1, v_2, \ldots, v_n$, donde $v_i \in V$. Decimos que la lista $v_1, v_2, \ldots, v_n$ es linealmente independiente si

$$\alpha_1 v_1 + \alpha_2 v_1 + \cdots + \alpha_n v_n = \mathbf{0},$$

implica que $\alpha_i = 0$ para toda $i = 0, \ldots, n$.

Por ejemplo, si $v \in \mathbb{R}^3$ es cualquier vector distinto de cero, el conjunto $\{v\}$ es linealmente **independiente**; sin embargo, la lista $v, v, v$ es linealmente **dependiente** debido a que existe una combinación lineal no trivial igual al vector cero:

$$v - \frac{1}{2}v - \frac{1}{2}v = \mathbf{0}.$$

Los resultados de esta sección serán enunciados para conjuntos linealmente independientes, pero hay que notar que también son válidos para listas linealmente independientes (ver Sec. 2.A en [1]).

**Definición 5.10 (Dimensión Infinita).** Decimos que un espacio vectorial $V$ es *de dimensión infinita* si existe un conjunto de cardinalidad infinita que sea linealmente independiente. En caso contrario, decimos que $V$ es *de dimensión finita*.

Aunque no hayamos definido la *dimensión* de un espacio vectorial, el párrafo anterior define las frases "de dimensión infinita" y "de dimensión finita".

**Ejemplo 5.11.** El espacio vectorial de polinomios $\mathbb{R}[x]$ es de dimensión infinita porque el conjunto $A = \{1, x, x^2, x^3, ...\}$ es un conjunto infinito linealmente independiente: claramente, para cualquier $n \in \mathbb{N}$,

$$\alpha_0 + \alpha_1 x + \alpha_2 x^2 + \alpha_3 x^3 + ... + \alpha_n x^n = \mathbf{0}$$

si y sólo si $\alpha_0 = \alpha_1 = ... = \alpha_n = 0$.

**Lema 5.12 (subconjuntos linealmente independientes).** Sea $V$ un espacio vectorial sobre $F$. Si $A \subseteq V$ es linealmente independiente, cualquier subconjunto de $A$ también es linealmente independiente.

**Demostración.** Ejercicio 5.51. □

**Teorema 5.13 (independencia lineal).** Sea $V$ un espacio vectorial sobre $F$. Un subconjunto $A \subseteq V$ es linealmente independiente si y sólo si ningún elemento de $A$ es igual a la combinación lineal de otros elementos de $A$.

**Demostración.** Demostraremos la siguiente afirmación equivalente: $A$ es linealmente dependiente si y sólo si algún elemento de $A$ es igual a la combinación lineal de otros elementos de $A$.

($\Rightarrow$) Supongamos que $A$ es linealmente dependiente. Por definición, existen $v_1, \ldots, v_n \in A$ y escalares $\alpha_1, \alpha_2, \ldots, \alpha_n \in F$ no todos cero, tales que

$$\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n = \mathbf{0}. \tag{5.3}$$

Sin perder generalidad (reetiquetando si es necesario), supongamos que $\alpha_1 \neq 0$. Luego,

$$v_1 = \left(\frac{-\alpha_2}{\alpha_1}\right) v_2 + \left(\frac{-\alpha_3}{\alpha_1}\right) v_3 + \ldots + \left(\frac{-\alpha_n}{\alpha_1}\right) v_n.$$

Por lo tanto $v_1$ es igual a una combinación lineal de otros vectores en $A$.

($\Leftarrow$) Supongamos que existe un vector $v_1 \in A$ que puede escribirse como combinación lineal de otros vectores $v_2, \ldots, v_n \in A$, donde $v \neq v_i$ para toda $i = 2, \ldots, n$. En otras palabras,

$$v_1 = \alpha_2 v_2 + \alpha_3 v_3 + \ldots + \alpha_n v_n,$$

Para algunos $\alpha_2, \ldots, \alpha_n \in F$. Entonces,

$$1v_1 - \alpha_2 v_2 - \alpha_2 v_2 - \ldots - \alpha_n v_n = \mathbf{0},$$

es una combinación lineal no trivial de elementos de $A$ (porque al menos el coeficiente de $v_1$ es distinto de cero). Por lo tanto, $A$ es linealmente dependiente.

□

**Ejemplo 5.14.** El subconjunto $A = \{(1,2), (1,1), (3,4)\}$ de $\mathbb{R}^2$ es linealmente dependiente. Observemos primero que

$$\alpha_1 (1,2) \neq (1,1),$$

para cualquier escalar $\alpha_1$. Ahora

$$\alpha_1 (1,2) + \alpha_2 (1,1) = (3,4),$$

entonces $\alpha_1 + \alpha_2 = 3$ y $2\alpha_1 + \alpha_2 = 4$. Resolviendo, tenemos que $\alpha_1 = 1$, $\alpha_2 = 2$. Por lo tanto, el conjunto $A$ es linealmente dependiente.

## 5.2. Conjuntos generadores

Sea $A$ un subconjunto de un espacio vectorial $V$ sobre $F$. Recordemos que $\text{gen}_F(A)$ (también denotado como $\langle A \rangle$ cuando $F$ está claro por el contexto), es el conjunto de todas las combinaciones lineales de $A$.

**Definición 5.15 (conjunto generador).** Sea $A$ un subconjunto de un espacio vectorial $V$ sobre $F$. Decimos que $A$ es un *conjunto generador* de $V$, o que *genera a* $V$, si $\text{gen}_F(A) = V$.

**Observación 5.16.** Si queremos demostrar que $\langle A\rangle = V$, entonces debemos probar solamente la contención $V \subseteq \langle A\rangle$, puesto que la otra contención $\langle A\rangle \subseteq V$ siempre se cumple ya que $\langle A\rangle$ es un subespacio de $V$.

**Ejemplo 5.17.** Consideremos el conjunto

$$W = \left\{(x_1, x_2, x_3) \in \mathbb{R}^3 : 2x_1 + x_2 - x_3 = 0\right\}$$

Encontraremos un conjunto generador para $W$. Observemos que

$$\begin{aligned} W &= \{(x_1, x_2, 2x_1 + x_2) : x_1, x_2 \in \mathbb{R}\} \\ &= \{x_1 (1, 0, 2) + x_2 (0, 1, 1) : x_1, x_2 \in \mathbb{R}\} \end{aligned}$$

Así que $A = \{(1, 0, 2), (0, 1, 1)\}$ es un conjunto de vectores de $\mathbb{R}^3$ tal que $\langle A\rangle = W$. Obviamente, este conjunto generador no es único; por ejemplo, $\{(2, 0, 4), (0, -1, -1)\}$ es otro conjunto generador de $W$.

**Teorema 5.18 (espacios generados e independencia lineal).** Sea $V$ un espacio vectorial sobre $F$ y sea $A \subseteq V$. Entonces:

**(1)** Si $A$ es linealmente dependiente, existe un $v \in A$ tal que $\langle A \setminus \{v\}\rangle = \langle A\rangle$.

**(2)** Si $A$ es linealmente independiente y $v \in V \setminus \langle A\rangle$, entonces $A \cup \{v\}$ es linealmente independiente.

**Demostración.**

**(1)** Por el Teorema 5.13, si $A$ es linealmente dependiente, existe $v \in A$ que puede escribirse como combinación lineal de otros vectores $v_1, \ldots, v_n \in A$, donde $v \neq v_j$, para toda $j$. Es decir,

$$v = \alpha_1 v_1 + \ldots \alpha_n v_n, \tag{5.4}$$

para algunos $\alpha_i \in F$, $i = 1, \ldots, n$. Obviamente, $\text{gen}_F(A \setminus \{v\}) \leq \text{gen}_F(A)$ porque $A \setminus \{v\} \subseteq A$. Para demostrar la otra inclusión, consideremos una combinación lineal $\sum_{i=1}^{k} \beta_i a_i \in \text{gen}_F(A)$. Si $v \neq a_i$, para toda $i$, entonces $\sum_{i=1}^{k} \beta_i a_i \in \text{gen}_F(A \setminus \{v\})$. Si $v = a_i$ para alguna $i$, digamos $v = a_1$, podemos substituir (5.4) en la combinación lineal:

$$\sum_{i=1}^{k} \beta_i a_i = \sum_{j=1}^{n} \alpha_i v_i + \sum_{i=2}^{k} \beta_i a_i \in \text{gen}_F(A \setminus \{v\}).$$

**(2)** Ejercicio.

□

**Teorema 5.19 (Lema del Intercambio).** Sea $V$ un espacio vectorial sobre $F$, y sean $A$ y $B$ subconjuntos finitos de $V$. Si $A$ genera a $V$ y $B$ es linealmente independiente, entonces $|A| \geq |B|$.

**Demostración.** Supongamos que $A = \{a_1, \ldots, a_n\}$ y $B = \{b_1, \ldots, b_m\}$. Por reducción al absurdo, supongamos que $m > n$.

Como $b_1 \in V = \langle A \rangle$ existen $\alpha_i \in F$ no todos cero (porque $b_1 \neq \vec{0}$ ya que $B$ es linealmente independiente), tales que

$$b_1 = \alpha_1 a_1 + \ldots \alpha_n a_n.$$

Sin perder generalidad, supongamos que $\alpha_1 \neq 0$. Luego,

$$a_1 = \frac{1}{\alpha_1} b_1 - \frac{\alpha_2}{\alpha_1} a_2 - \cdots - \frac{\alpha_n}{\alpha_1} a_n.$$

Esto demuestra que podemos reemplazar en $A$ a $a_1$ por $b_1$ para obtener el conjunto

$$A_1 := \{b_1, a_2, \ldots, a_n\}$$

que también genera a $V$ (por un argumento similar al de la demostración del Teorema 5.18 (1)).

Repetimos el procedimento anterior: como $b_2 \in V = \langle A_1 \rangle$, existen $\beta_i \in F$ no todos cero tales que

$$b_2 = \beta_1 b_1 + \beta_2 a_2 + \cdots + \beta_n a_n.$$

Si $\beta_i = 0$, para toda $i \geq 2$, obtenemos la relación $b_2 = \beta_1 b_1$, $\beta_1 \neq 0$, lo cual contradice que $\{b_1, b_2\}$ es linealmente independiente. Luego, podemos suponer que $\beta_2 \neq 0$, y entonces

$$a_2 = \frac{1}{\beta_2} b_2 - \frac{\beta_1}{\beta_2} b_1 - \frac{\beta_3}{\beta_2} a_3 - \cdots - \frac{\beta_n}{\beta_2} a_n.$$

Esto demuestra que podemos reemplazar en $A_1$ a $a_2$ por $b_2$ para obtener el conjunto

$$A_2 := \{b_1, b_2, a_3 \ldots, a_n\}$$

que también genera a $V$.

Después de $n$ repeticiones de este procedimento, podemos reemplazar en $A_{n-1} := \{b_1, b_2, \ldots, b_{n-1}, a_n\}$ a $a_n$ por $b_n$ (el cual existe porque $m > n$), y obtenemos que el conjunto

$$A_n := \{b_1, b_2, \ldots, b_n\}$$

también genera a $V$. Sin embargo $A_n \subset B$ y existe $b_{n+1} \in B \setminus A_n$, el cual puede ser escrito como combinación lineal de otros elementos de $B$. Esto contradice que $B$ sea linealmente independiente. Por lo tanto, $m \leq n$. □

## 5.3. Bases

**Definición 5.20 (Base).** Sea $V$ un espacio vectorial sobre $F$ y sea $B \subseteq V$. Decimos que $B$ es una base de $V$ sobre $F$ si $B$ es linealmente independiente y $\text{gen}_F(B) = V$.

**Ejemplo 5.21.** El conjunto $B$ de $n$ vectores

$$e_1 = (1, 0, \ldots, 0), e_2 = (0, 1, \ldots, 0), \ldots, e_n = (0, 0, \ldots, 1),$$

es una base para el espacio vectorial $\mathbb{R}^n$ sobre $\mathbb{R}$. En el ejemplo 5.5 ya se demostró que este conjunto es linealmente independiente sobre $\mathbb{R}$. Ahora debemos probar que $\text{gen}_{\mathbb{R}}(B) = \mathbb{R}^n$. Por cerradura, sabemos que $\text{gen}_{\mathbb{R}}(B) \subseteq \mathbb{R}^n$. Así que debemos demostrar que $\mathbb{R}^n \subseteq gen_{\mathbb{R}}(B)$. Sea $v = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{R}^n$ un elemento arbitrario. Entonces podemos escribir

$$(\alpha_1, \alpha_2, \ldots, \alpha_n) = \alpha_1 (1, 0, \ldots, 0) + \alpha_2 (0, 1, \ldots, 0) + \ldots + \alpha_n (0, 0, \ldots, 1),$$

es decir

$$v = \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n,$$

De esta forma $\mathbb{R}^n \subseteq gen_{\mathbb{R}}(B)$, y $\text{gen}_{\mathbb{R}}(B) = \mathbb{R}^n$. Por lo tanto $B$ es una base de $\mathbb{R}^n$ sobre $\mathbb{R}$. Llamaremos a esta base en particular la **base canónica** de $\mathbb{R}^n$.

**Ejemplo 5.22.** Sea $V = \{\mathbf{0}\}$ el espacio vectorial trivial sobre $F$. El conjunto $\{\mathbf{0}\}$ no es una base de $V$ porque no es linealmente independiente. Sin embargo, $\emptyset$ sí es una base para $V$ porque es linealmente independiente y $\langle \emptyset \rangle = \{\mathbf{0}\}$ (ver Definición 2.13).

**Teorema 5.23 (Tamaño de las Bases).** Sea $V$ un espacio vectorial sobre $\mathbf{F}$, y sean $B$ y $C$ bases de $V$. Entonces $|B| = |C|$.

**Demostración.** Si $V$ es de dimensión infinita, puede consultarse el Teorema 1.12 de [6]. El caso de dimensión finita puede demostrarse fácilmente usando el Lema del Intercambio, así que se deja como ejercicio. □

La motivación principal para la definición de base se origina de la posibilidad de obtener un conjunto **mínimo** de vectores que generen un espacio vectorial dado.

**Teorema 5.24 (definiciones equivalentes de base).** Sea $V$ un espacio vectorial sobre $F$, y $B \subseteq V$ un subconjunto no vacío. Las siguientes afirmaciones son equivalentes:

(1) $B$ es una base de $V$.

(2) Cualquier vector $v \in V$ puede ser escrito de forma *esencialmente única*[1] como una combinación lineal de los elementos de $B$.

(3) $B$ es un conjunto generador de $V$ minimal; es decir, $\text{gen}_F(B) = V$ y si $A \subsetneq B$, entonces $A$ no genera a $V$.

(4) $B$ es un conjunto linealmente independiente maximal; es decir, $B$ es linealmente independiente y si $A \supsetneq B$, entonces $A$ es linealmente dependiente.

[1] En este contexto, *esencialmente única* significa que las combinaciones lineales son iguales salvo por sumas de combinaciones lineales triviales.

**Demostración.** Demostraremos cada implicación.

(1) $\Rightarrow$ (2) Supongamos que $B$ es una base de $V$. Como $V = \langle B \rangle$, cualquier elemento de $V$ puede escribirse como una combinación lineal de vectores en $B$. Demostraremos la unicidad. Supongamos que algún $v \in V$ es igual a dos combinaciones lineales distintas de $B$. Permitiendo coeficientes iguales a 0 si es necesario, podemos suponer que

$$v = \alpha_1 v_1 + ... + \alpha_n v_n = \beta_1 v_1 + ... + \beta_n v_n$$

donde $\alpha_i, \beta_i \in \mathbf{F}$, y $v_i \in B$. Entonces

$$(\alpha_1 - \beta_1) v_1 + ... + (\alpha_n - \beta_n) v_n = \mathbf{0}.$$

Como $B$ es linealmente independiente, esto implica que $\alpha_i - \beta_i = 0$, para toda $i = 1, \ldots, n$. Por lo tanto, la representación de $v$ como combinación lineal de vectores de $B$ es única.

(2) $\Rightarrow$ (3) Sea $A \subsetneq B$. Por reducción al absurdo, supongamos que $\langle A \rangle = V$. Sea $w \in B \setminus A$, $w \neq \mathbf{0}$. Como $w$ también pertenece a $V$, tenemos que

$$w = \alpha_1 v_1 + ... + \alpha_n v_n$$

donde $\alpha_i \in \mathbf{F}$, $\alpha_i \neq 0$, y $v_i \in A \subsetneq B$. Luego,

$$\alpha_1 v_1 + ... + \alpha_n v_n - w = \mathbf{0}.$$

Como $\mathbf{0} = 0v_1 + ... + 0v_n$, la igualdad anterior implica que el vector $\mathbf{0}$ tiene dos representaciones distintas como combinación lineal de elementos de $B$. Esto contradice el punto (2).

(3) $\Rightarrow$ (4) Supongamos que $B$ es un conjunto generador de $V$ minimal. Primero demostraremos que $B$ es linealmente independiente. Por reducción al absurdo, supongamos que $B$ es linealmente dependiente. Por el Teorema 5.18 (1), existe $v \in B$ tal que $\langle B \setminus \{v\} \rangle = \langle B \rangle = V$, lo que contradice que $B$ sea generador minimal. Luego, $B$ es linealmente independiente.

Ahora demostraremos que $B$ es linealmente independiente maximal. Por reducción al absurdo, supongamos que existe $A \subseteq V$ tal que $A \supsetneq B$ y $A$ es linealmente independiente. Sea $u \in A \setminus B$. Como $\langle B \rangle = V$, podemos escribir a $u$ como combinación lineal de $A$: sin embargo, por el Teorema 5.13, esto implica que $A$ es linealmente dependiente, lo cual es una contradicción.

(4) $\Rightarrow$ (1) Supongamos que $B \subseteq V$ es un conjunto linealmente independiente maximal. Para demostrar que $B$ es una base de $V$, debemos mostrar que $\langle B \rangle = V$. Sea $v \in V$, y, por reducción al absurdo, supongamos que existe $v \in V \setminus \langle B \rangle$. Por el Teorema 5.18 (2), esto implica que el conjunto $A := B \cup \{v\}$ es linealmente independiente. Como $A \supsetneq B$, esto contradice que $B$ sea un conjunto linealmente independiente maximal. Por lo tanto, $\langle B \rangle = V$.

□

Para demostrar el Teorema de Existencia de Bases, necesitamos el Lema de Zorn, el cual es equivalente al Axioma de Elección. Antes de enunciar el Lema de Zorn, debemos recordar algunas definiciones.

**Definición 5.25 (Conjuntos parcialmente ordernados).** Sea $P$ un conjunto no vacío. Una *relación de orden* sobre $P$ es una relación $\preceq$ sobre $P$ que es reflexiva, *antisimétrica* (i.e. para todo $a, b \in P$, si $a \preceq b$ y $b \preceq a$, entonces $a = b$) y transitiva. Nos referimos a un conjunto equipado con una relación de orden como un *conjunto parcialmente ordenado*.

**Definición 5.26.** Sea $P$ un conjunto parcialmente ordenado.

1. Una *cadena* de $P$ es un subconjunto $C \subseteq P$ *totalmente ordenado*; es decir, para toda $a, b \in C$ se tiene que $a \preceq b$ o $b \preceq a$.
2. Una *cota superior* de un subconjunto $A \subseteq P$ es un elemento $c \in P$ tal que $a \preceq c$, para toda $a \in A$.
3. Un *elemento maximal* de $P$ es un elemento $m \in P$ tal que si $a \in P$ satisface $m \preceq a$, entonces $m = a$.

**Ejemplo 5.27.** Consideremos el conjunto $P = \mathcal{P}(\{1, 2, 3, 4\})$; es decir, $P$ es el conjunto potencia de $\{1, 2, 3, 4\}$. La inclusión de conjuntos $\subseteq$ es una relación de orden sobre $P$. Una cadena de $P$ es

$$C := \{\{1\}, \{1, 2\}, \{1, 2, 3\}\}.$$

Las cotas superiores de $C$ son $\{1, 2, 3\}$ y $\{1, 2, 3, 4\}$, mientras que $\{1, 2, 3\}$ es el único elemento maximal de $C$. Por otro lado, el conjunto

$$A := \{\{3\}, \{2, 4\}, \{2, 3\}\}$$

no es una cadena (por ejemplo, $\{3\}$ y $\{2, 4\}$ no se pueden comparar con la inclusión) y tiene dos elementos maximales: $\{2, 4\}$ y $\{2, 3\}$.

**Lema 5.28 (Zorn).** Sea $P$ un conjunto parcialmente ordenado con la propiedad de que cualquier cadena tiene una cota superior. Entonces $P$ contiene al menos un elemento maximal.

En el siguiente teorema demostramos que cualquier espacio vectorial distinto de cero tiene una base. Si el espacio es de dimensión finita, es posible demostrar esto sin usar el Lema de Zorn; sin embargo, este lema es indispensable cuando el espacio es de dimensión infinita.

**Teorema 5.29 (Existencia de bases).** Sea $V$ un espacio vectorial sobre $F$. Las siguientes afirmaciones se cumplen:

1. Cualquier conjunto linealmente independiente de $V$ está contenido en una base de $V$.

2. Cualquier conjunto generador de $V$ contiene a una base de $V$.

3. Existe una base de $V$.

**Demostración.** Demostraremos cada punto.

1. Sea $I \subseteq V$ un conjunto linealmente independiente. Demostraremos que $I$ está contenido en una base de $V$. Consideremos al conjunto

$$\mathcal{I} := \{A \subset V : I \subset A \text{ y } A \text{ es linealmente independiente}\},$$

Claramente, $\mathcal{I}$ es no vacío porque $I \in \mathcal{I}$. Además, $\mathcal{I}$ es un conjunto ordenado bajo la inclusión de conjuntos. Consideremos una cadena de $\mathcal{I}$:

$$\mathbf{C} = \{C_j : j \in J\}$$

donde $J$ es un conjunto de índices. Consideremos la unión de todos los elementos de $\mathbf{C}$:

$$U = \bigcup_{j \in J} C_j.$$

Demostraremos que $U$ es linealmente independiente. Por definición, si $v \in U$, entonces $v \in C_j$, para algún $j \in J$. Luego, si $v_1, ..., v_k \in U$, entonces, sin perder generalidad, podemos suponer que $v_1 \in C_1, v_2 \in C_2, ..., v_k \in C_k$. Como $\mathbf{C}$ es totalmente ordenado, siempre tenemos que $C_i \subseteq C_j$, o $C_j \subseteq C_i$, para cualquier $i, j \in J$. Por lo tanto, podemos suponer que

$$C_1 \subseteq C_2 \subseteq ... \subseteq C_k.$$

En particular, $\{v_1, ..., v_k\} \subseteq C_k$, lo implica que $\{v_1, ..., v_k\}$ es linealmente independiente (Ejercicio 5.51 (1.)). Esto demuestra que $U$ es linealmente independiente, así que $U \in \mathcal{I}$. Como $C_j \subseteq U$, para toda $j \in J$, entonces $U$ es una cota superior de $\mathbf{C}$. Debido a que $\mathbf{C}$ es una cadena arbitraria, podemos usar el Lema de Zorn; por lo tanto, $\mathcal{I}$ contiene al menos un elemento maximal $B \in \mathcal{I}$. Por el Teorema 5.24, $B$ es una base de $V$ que contiene a $I$.

2. Sea $D \subset V$ un subconjunto tal que $\langle D \rangle = V$. Consideremos el conjunto

$$\mathcal{S} := \{A \subseteq D : A \text{ es linealmente independiente}\}.$$

Ahora, imitando el argumento del punto anterior, podemos demostrar que $\mathcal{S}$ tiene un elemento maximal $B \in \mathcal{S}$ (Ejercicio 5.58). Por el Lema del Intercambio, $B$ es un subconjunto de $V$ linealmente independiente maximal. Por lo tanto, $B$ es una base de $V$ contenida en $D$.

3. Si $V = \{\mathbf{0}\}$, entonces $\emptyset$ es una base para $V$. Si $V \neq \{\mathbf{0}\}$, sea $v \in V$ tal que $v \neq \mathbf{0}$. El conjunto $A = \{v\}$ es linealmente independiente, así que por el punto (1), está contenido en una base de $V$. En cualquier caso, existe una base de $V$.

□

En el siguiente resultado, entendemos por *algoritmo* a un procedimiento que siempre termina después de un número finito de pasos.

**Teorema 5.30 (Existencia de Bases: caso de dimensión finita).** Sea $V$ un espacio vectorial de dimensión finita sobre $F$. Entonces, existe un algoritmo para encontrar una base de $V$.

**Demostración.** Si $V = \{\mathbf{0}\}$, entonces $\emptyset$ es una base para $V$, así que supongamos que $V \neq \{\mathbf{0}\}$. Sea $v_1 \in V$, $v_1 \neq \mathbf{0}$. Si $\langle v_1 \rangle = V$, entonces $\{v_1\}$ es una base para $V$. En caso contrario, sea $v_2 \in V \setminus \langle v_1 \rangle$. Por el Teorema 5.18, el conjunto $\{v_1, v_2\}$ es linealmente independiente. Si $\langle v_1, v_2 \rangle = V$, entonces $\{v_1, v_2\}$ es una base para $V$. En caso contrario, sea $v_3 \in \setminus \langle v_1, v_2 \rangle$, y repitamos el proceso anterior. Repetir este proceso debe terminar en algún momento porque $V$ es de dimensión finita, y no tiene subconjuntos infinitos linealmente independientes. Si el proceso termina después de $n$ pasos, obtenemos un conjunto $\{v_1, \ldots, v_n\}$ el cual genera a $V$ y es linealmente independiente; por lo tanto, éste es una base de $V$. □

## 5.4. Dimensión

Ahora definiremos uno de los conceptos más importantes relacionados con la teoría de espacios vectoriales.

**Definición 5.31 (dimensión).** La dimensión de un espacio vectorial $V$ sobre un campo $F$, denotada por $\dim_F(V)$, es la cardinalidad de cualquier base $B$ de $V$. Si el campo $F$ está claro por el contexto, escribimos simplemente $\dim(V)$.

**Observación 5.32.** La definición anterior tiene sentido gracias al Teorema 5.23.

**Observación 5.33.** Por el Teorema 5.29, $\dim(V) < \infty$ si y sólo si $V$ es de dimensión finita.

**Ejemplo 5.34.** La dimensión del espacio vectorial trivial es 0 porque $|\emptyset| = 0$.

**Ejemplo 5.35.** Como la base canónica $\{e_1, e_2, \ldots, e_n\}$ de $F^n$ tiene $n$ elementos, deducimos que

$$\dim_F(F^n) = n.$$

**Ejemplo 5.36.** Sea $V$ un espacio vectorial sobre $F$. Si $v \in V$, $v \neq \mathbf{0}$, entonces

$$\dim(\langle v \rangle) = 1,$$

ya que $\{v\}$ es una base de $\langle v \rangle$. Si $w \in V$, entonces $\{v, w\}$ es linealmente independiente si y sólo si

$$\dim(\langle v, w \rangle) = 2.$$

**Ejemplo 5.37.** La dimensión de $\mathbb{R}$, visto como espacio vectorial sobre $\mathbb{R}$, es 1 porque $\{1\}$ es una base. Sin embargo, si consideramos a $\mathbb{R}$ como espacio vectorial sobre $\mathbb{Q}$, entonces $\{1\}$ no es una base ya que el espacio generado $\langle 1\rangle = \{\alpha 1 : \alpha \in \mathbb{Q}\}$ no es igual a $\mathbb{R}$. De hecho, como $\mathbb{Q}$ es un conjunto numerable y $\mathbb{R}$ es no numerable, es posible demostrar que el espacio vectorial $\mathbb{R}$ sobre $\mathbb{Q}$ es de dimensión infinita. En conclusión,

$$\dim_{\mathbb{R}}(\mathbb{R}) = 1, \text{ pero } \dim_{\mathbb{Q}}(\mathbb{R}) = \infty.$$

**Observación 5.38.** Si $S \leq V$ y $A$ es una base de $S$, claramente, $A$ es un subconjunto de $V$ linealmente independiente. Por el Teorema 5.29 (1), podemos extender $A$ a una base de $V$; esto significa que la base de cualquier subespacio de $V$ puede extenderse a una base de $V$.

**Ejemplo 5.39.** Consideremos el subespacio

$$S = \left\{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 = x_2\right\} \subseteq \mathbb{R}^3.$$

Debido a que

$$A = \{(1, 1, 0), (0, 0, 1)\}$$

es un conjunto linealmente independiente que genera a $S$, tenemos que $\dim_{\mathbb{R}}(S) = 2$. El conjunto $A$ puede extenderse a una base de $V$ añadiendo algún vector que no sea una combinación lineal de los elementos de $A$; por ejemplo,

$$B = \{(1, 1, 0), (0, 0, 1), (1, 0, 0)\}$$

es una base de $V$ que contiene a $A$.

**Ejemplo 5.40.** Sabemos que tanto el espacio de polinomios $F[x]$ como el espacio de series formales $F[[x]]$ son de dimensión infinita. El conjunto

$$B = \{1, x, x^2, \dots\} = \{x^i : i \geq 0\},$$

es una base de $F[x]$ porque $B$ es linealmente independiente y $\text{gen}_F(B) = F[x]$, y por lo tanto

$$\dim_F(F[x]) = |B| = \aleph_0,$$

donde $\aleph_0$ es la cardinalidad de los números naturales. Sin embargo, $B$ no es una base de $F[[x]]$ porque $\text{gen}_F(B) \neq F[[X]]$. (Recordemos que las combinaciones lineales siempre son expresiones finitas; así por ejemplo, la serie formal $\sum_{i=0}^{\infty} x^i$ no puede ser expresada como una combinación lineal de $B$). Entonces, ¿podemos encontrar una base para $F[[x]]$? El Teorema de Existencia de Bases asegura que $F[[x]]$ debe tener una base, pero su demostración es no constructiva y no especifica cómo encontrarla. Sin embargo, aunque nadie sabe explícitamente cuál es la base de $F[[x]]$, es posible demostrar que debe ser un conjunto no numerable; por lo tanto, $\dim_F(F[x]) < \dim_F(F[[x]])$.

Nuestro objetivo ahora es demostrar el Teorema de Dimensión e Isomorfía, el cual establece que dos espacios vectoriales son isomorfos si y sólo si tienen la misma dimensión.

El siguiente teorema nos presenta un concepto que será de gran utilidad; su demostración involucra un procedimiento para 'extender' una función definida sobre una base a una transformación lineal.

**Teorema 5.41 (Extensión Lineal).** Sean $V$ y $W$ espacios vectoriales sobre $F$, y sea $B$ una base para $V$. Sea $f : B \longrightarrow W$ cualquier función que asigna vectores de $W$ a los elementos de $B$. Entonces existe una única transformación lineal $\phi : V \longrightarrow W$ tal que $\phi(b) = f(b)$, para toda $b \in B$.

**Demostración.** Para demostrar la existencia, definamos una función $\phi : V \longrightarrow W$ de la siguiente forma:

$$\phi(v) = \phi(\alpha_1 b_1 + ... + \alpha_n b_n) := \alpha_1 f(b_1) + ... + \alpha_n f(b_n),$$

donde $v = \alpha_1 b_1 + \cdots + \alpha_n b_n \in V$, $\alpha_i \in F$ y $b_i \in B$. Demostraremos que $\phi$ es una transformación lineal. Sea $u = \beta_1 b_1 + ... + \beta_m b_m \in V$, $\beta_i \in F$, y $\alpha, \beta \in F$. Sin perder generalidad, supongamos que $m \geq n$. Entonces:

$$\begin{aligned}
\phi(\alpha v + \beta u) &= \phi(\alpha(\alpha_1 b_1 + ... + \alpha_n b_n) + \beta(\beta_1 b_1 + ... + \beta_m b_m)) \\
&= \phi((\alpha\alpha_1 + \beta\beta_1) b_1 + ... + (\alpha\alpha_n + \beta\beta_n) b_n + \cdots + \beta\beta_m b_m) \\
&= (\alpha\alpha_1 + \beta\beta_1) f(b_1) + ... + (\alpha\alpha_n + \beta\beta_n) f(b_n) + \cdots + \beta\beta_m f(b_m) \\
&= \alpha[\alpha_1 f(b_1) + ... + \alpha_n f(b_n)] + \beta[\beta_1 f(b_1) + ... + \beta_m f(b_m)] \\
&= \alpha\phi(v) + \beta\phi(u).
\end{aligned}$$

Claramente, $\phi$ cumple que $\phi(b) = f(b)$, para toda $b \in B$. Para demostrar la unicidad, supongamos que $\varphi : V \longrightarrow W$, es una transformación lineal tal que $\varphi(b) = f(b)$, para toda $b \in B$. Entonces, para cualquier $v = \alpha_1 b_1 + ... + \alpha_i b_n \in V$,

$$\begin{aligned}
\varphi(v) &= \varphi(\alpha_1 b_1 + ... + \alpha_n b_n) \\
&= \alpha_1 \varphi(b_1) + ... + \alpha_n \varphi(b_n) \\
&= \alpha_1 f(b_1) + ... + \alpha_n f(b_n) \\
&= \phi(\alpha_1 b_1 + ... + \alpha_n b_n) = \phi(v).
\end{aligned}$$

Por lo tanto, $\varphi = \phi$. □

**Observación 5.42.** El teorema anterior implica que sólo es necesario definir las imágenes de los elementos de la base de un espacio vectorial para definir, de manera única, una transformación lineal.

**Definición 5.43 (Extensión Lineal).** La transformación lineal $\phi : V \longrightarrow W$ definida en el teorema anterior se llama la *extensión lineal* de $f : B \longrightarrow W$.

**Ejemplo 5.44.** Sea $B = \{(1,0),(0,1)\}$ la base canónica de $\mathbb{R}^2$. Consideremos la función $f : B \longrightarrow \mathbb{R}^3$ definida como

$$\begin{aligned} f(1,0) &= (1,0,1), \\ f(0,1) &= (0,2,0). \end{aligned}$$

Entonces, la extensión lineal de $f$ es la transformación lineal $\phi : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ definida como

$$\begin{aligned} \phi(x_1,x_2) &= x_1 f(1,0) + x_2 f(0,1) \\ &= x_1(1,0,1) + x_2(0,2,0) \\ &= (x_1, 2x_2, x_1), \end{aligned}$$

donde $x_i \in \mathbb{R}$.

**Ejemplo 5.45.** Consideremos ahora la base $B = \{(1,0),(1,1)\}$ de $\mathbb{R}^2$, y definamos $f : B \longrightarrow \mathbb{R}^2$ como

$$\begin{aligned} f(1,0) &= (1,-2), \\ f(1,1) &= (-2,1). \end{aligned}$$

La extensión lineal de $f$ es la transformación lineal $\phi : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ definida como

$$\begin{aligned} \phi(x_1,x_2) &= \phi((x_1-x_2)(1,0) + x_2(1,1)) \\ &= (x_1-x_2) f(1,0) + x_2 f(1,1) \\ &= (x_1-x_2)(1,-2) + x_2(-2,1) \\ &= (x_1 - 3x_2, -2x_1 + 3x_2). \end{aligned}$$

**Ejemplo 5.46.** Sea $B = \{(1,0,0),(0,1,0),(0,0,1)\}$ la base canónica de $\mathbb{R}^3$. Consideremos la función $f : B \longrightarrow \mathbb{R}$ definida como

$$\begin{aligned} f(1,0,0) &= 3, \\ f(0,1,0) &= -2, \\ f(0,0,1) &= 2. \end{aligned}$$

Entonces, la extensión lineal de $f$ es la transformación lineal $\phi : \mathbb{R}^3 \longrightarrow \mathbb{R}$ definida como

$$\begin{aligned} \phi(x_1,x_2,x_3) &= x_1 f(1,0,0) + x_2 f(0,1,0) + x_3 f(0,0,1) \\ &= 3x_1 - 2x_2 + 2x_3, \end{aligned}$$

donde $x_i \in \mathbb{R}$.

El siguiente teorema es uno de los más importantes de la teoría de espacios vectoriales: establece que las clases de isomorfía de espacios vectoriales están completamente determinadas por la dimensión de los espacios.

**Teorema 5.47 (Dimensión e Isomorfía).** Sean $V$ y $W$ espacios vectoriales sobre $F$. Entonces, $V \cong W$ si y sólo si $\dim(V) = \dim(W)$.

**Demostración.** Observemos que este teorema no asume que $V$ y $W$ sean de dimensión finita. Demostraremos cada implicación.

($\Longrightarrow$) Supongamos que $V \cong W$ y sea $\phi : V \longrightarrow W$ un isomorfismo. Sea $B$ una base de $V$. Demostraremos que $\phi(B)$ es una base de $W$. Sea $w \in W$. Como $\phi$ es sobreyectivo, $w = \phi(v)$ para algún $v \in V$. Ahora, $v = \alpha_1 b_1 + ... + \alpha_n b_n$, para algunos $\alpha_i \in F$, $b_i \in B$, así que

$$\begin{aligned} w &= \phi(\alpha_1 b_1 + ... + \alpha_n b_n) \\ &= \alpha_1 \phi(b_1) + ... + \alpha_n \phi(b_n). \end{aligned}$$

Esto demuestra que $\phi(B)$ genera a $W$. Para demostrar que $\phi(B)$ es linealmente independiente, supongamos que

$$\alpha_1 \phi(b_1) + ... + \alpha_n \phi(b_n) = \mathbf{0}_W.$$

Como $\phi$ es una transformación lineal,

$$\phi(\alpha_1 b_1 + ... + \alpha_n b_n) = \mathbf{0}_W = \phi(\mathbf{0}_V),$$

así que por inyectividad obtenemos que

$$\alpha_1 b_1 + ... + \alpha_n b_n = \mathbf{0}_V.$$

Como $B$ es linealmente independiente, esto implica que $\alpha_i = 0$, para toda $i$. Luego, $\phi(B)$ es linealmente independiente. Ahora,

$$|B| = |\phi(B)|$$

porque $\phi$ es una biyección, así que $\dim(V) = \dim(W)$.

($\Longleftarrow$) Supongamos que $\dim(V) = \dim(W)$. Sea $B = \{b_i : i \in I\}$ una base de $V$ y $C = \{c_i : i \in I\}$ una base de $W$. Definamos la función $f : B \longrightarrow C$ como $f(b_i) = c_i$. Observemos que $f$ es una biyección, lo que concuerda con el hecho de que $|B| = |C|$. Sea $\phi : V \longrightarrow W$ la extensión lineal de $f$. Demostraremos que $\phi$ es biyectiva, y por lo tanto un isomorfismo. Sea $w \in W$, y, reetiquetando si es necesario, supongamos que

$$w = \alpha_1 c_1 + ... + \alpha_n c_n$$

donde $\alpha_i \in F$. Sea $v = \alpha_1 b_1 + ... + \alpha_n b_n \in V$ . Observemos que

$$\begin{aligned} \phi(v) &= \phi(\alpha_1 b_1 + ... + \alpha_n b_n) \\ &= \alpha_1 f(b_1) + ... + \alpha_n f(b_n) \\ &= \alpha_1 c_1 + ... + \alpha_n c_n = w, \end{aligned}$$

lo que demuestra que $\phi$ es sobreyectiva. Para demostrar que $\phi$ es inyectiva, sea $v \in \ker(\phi)$. Luego, $v = \gamma_1 b_1 + ... + \gamma_n b_n$, para algunos $\gamma_i \in F$, y

$$\phi(\gamma_1 b_1 + ... + \gamma_n b_n) = \mathbf{0}_W$$

Esto implica que

$$\gamma_1 \phi(b_1) + ... + \gamma_n \phi(b_n) = \gamma_1 c_1 + ... + \gamma_n c_n = \mathbf{0}_W$$

lo que implica que $\gamma_i = 0$ para toda $i$, por la independencia lineal de $C$. Luego, $\ker(\phi) = \{\mathbf{0}_V\}$, y $\phi$ es inyectiva por el Teorema 3.24.

□

**Corolario 5.48.** Sea $V$ un espacio vectorial sobre $F$ con $\dim(V) = n \in \mathbb{N}$. Entonces $V \cong F^n$.

**Observación 5.49.** Si $B = \{b_1, \ldots, b_n\}$ es una base de $V$ y $\{e_1, \ldots, e_n\}$ es la base canónica de $F^n$, entonces podemos considerar la función $f : B \to F^n$ definida por $f(b_i) = e_i$. La extensión lineal de $f$ es un isomofismo muy importante entre $V$ y $F^n$ el cual estudiaremos más en la siguiente sección.

***Palabras clave:*** *conjunto linealmente independiente, conjunto generador, lema del intercambio, base, existencia de bases, dimensión, extensión lineal, teorema de dimensión e isomorfía.*

## 5.5. Ejercicios

**Ejercicio 5.50.** Determina si cada una de las siguientes sentencias son verdaderas o falsas. Justifica tu respuesta:

1. $\dim\left(\mathbb{Z}_3^5\right) = 5$.
2. $\dim\left(F^F\right) < \infty$, donde $F = \mathbb{R}$.
3. $\dim_{\mathbb{Q}}(\mathbb{C}) = 2$.
4. $\dim_{\mathbb{R}}(\mathbb{C}) = 2$.
5. Si $\dim(V) = n$, cualquier subconjunto de $V$ de cardinalidad mayor que $n$ es linealmente independiente.
6. Si $\dim(V) = n$, cualquier subconjunto de $V$ de cardinalidad mayor que $n$ genera a $V$.
7. $\dim\left(\mathbb{Z}_2^3\right) = 3$.
8. $\dim_{\mathbb{R}}\left(\mathbb{R}^2\right) = 2$.
9. Sea $A = \{(1,0,1),(0,1,1),(2,-1,1)\} \subset \mathbb{R}^3$, entonces $\dim(gen_{\mathbb{R}}A) = 3$.
10. Todo espacio vectorial sobre un campo finito tiene dimensión finita.

**Ejercicio 5.51.** Sea $V$ un espacio vectorial sobre $F$ y sea $A \subseteq V$ un subconjunto linealmente independiente.

1. Demuestra que cualquier subconjunto de $A$ es también linealmente independiente.
2. Sea $v \in V \setminus \langle A \rangle$. Demuestra que $A \cup \{v\}$ es linealmente independiente.

**Ejercicio 5.52.** Determina si los siguientes conjuntos $B$ son bases de los espacios vectoriales $V$ dados a continuación. Justifica detalladamente tu respuesta.

**(a)** $B := \left\{(1,0,-1),(-1,\tfrac{1}{3},0),(\tfrac{7}{2},-1,-\tfrac{1}{2})\right\}$; $V = \mathbb{R}^3$.

**(b)** $B := \{(1,1,2),(1,2,1),(2,1,1)\}$; $V = \mathbb{R}^3$.

**(c)** $B := \{(0,1,2),(2,1,1),(1,1,0)\}$; $V := \mathbb{Z}_3^3$.

**(d)** $B := \left\{x^2+1,\ x^2+x,\ x^2\right\}$; $V := \{p(x) \in \mathbb{R}[x] : \text{grad}(p(x)) \leq 2\}$.

**(e)** $B := \left\{x^2+x+1,\ x^2-x-1\right\}$; $V := \{p(x) \in \mathbb{R}[x] : \text{grad}(p(x)) \leq 2\}$.

**Ejercicio 5.53.** Encuentra una base y la dimensión de los siguientes subespacios de $\mathbb{R}^3$:

1. $S_1 := \langle (3,0,0) \rangle$.

2. $S_2 := \langle (2,0,0), (0,0,0), (1,0,0), (0,0,1) \rangle$.
3. $S_3 := \langle (1,0,0), (1,1,0), (1,1,1) \rangle$.
4. $S_4 := \langle (1,1,-1), (0,0,1), (1,1,0) \rangle$.
5. $S_5 := \langle (1,0,0), (0,0,0), (-1,0,0), (0,0,1) \rangle$.
6. $S_6 := \{(x,y,z) | x+y+z = 0\}$.
7. $S_7 := \{(x,y,z) | x+y = 0\}$.
8. $S_8 := S_6 \cap S_7$, donde $S_6$ y $S_7$ son los subespacios definidos en los puntos anteriores.

**Ejercicio 5.54.** Sea $V$ un espacio vectorial sobre $F$ y sea $S \leq V$ un subespacio.

1. Demuestra que $\dim(S) \leq \dim(V)$.
2. Demuestra que si $V$ es de dimensión finita y $S \neq V$, entonces $\dim(S) < \dim(V)$.
3. Da un ejemplo de un espacio $V$ y un subespacio $S$ tales que $S \neq V$ pero $\dim(S) = \dim(V)$.

**Ejercicio 5.55.** Sea $X := \{a, b, c\}$, y considera las funciones $f_a, f_b, f_c \in \mathbb{R}^X$ definidas de la siguiente manera:

$$f_a(x) := \begin{cases} 1 & \text{si } x = a, \\ 0 & \text{si } x \neq a, \end{cases} \quad f_b(x) := \begin{cases} 1 & \text{si } x = b, \\ 0 & \text{si } x \neq b, \end{cases} \quad f_c(x) := \begin{cases} 1 & \text{si } x = c, \\ 0 & \text{si } x \neq c. \end{cases}$$

Demuestra que $B := \{f_a, f_b, f_c\}$ es una base de $\mathbb{R}^X$.

**Ejercicio 5.56.** Usa el Lema del Intercambio para demostrar que todas las bases de un espacio vectorial tienen el mismo tamaño.

**Ejercicio 5.57.** Sin usar el Lema de Zorn, demuestra que cualquier espacio vectorial de dimensión finita tiene una base.

**Ejercicio 5.58.** Sea $V$ un espacio vectorial sobre $F$ y sea $D \subset V$ un subconjunto tal que $\langle D \rangle = V$. Consideremos el conjunto

$$\mathcal{S} := \{A \subseteq D : A \text{ es linealmente independiente}\}.$$

Usa el Lema de Zorn para demostrar que $\mathcal{S}$ tiene un elemento maximal.

# 6

# Dimensiones finitas y coordenadas

## 6.1. Dimensiones finitas

Recordemos del capítulo anterior que la dimensión de un espacio vectorial $V$ sobre $F$, denotada por $\dim_F(V)$ es igual a la cardinalidad de cualquier base de $V$. En esta sección, analizaremos algunas propiedades de la dimensión de subespacios y espacios vectoriales de dimensión finita.

**Teorema 6.1 (dimensión de la suma de subespacios).** Sean $S$ y $T$ subespacios de un espacio vectorial de dimensión finita $V$ sobre un campo $F$. Entonces

$$\dim(S+T) = \dim(S) + \dim(T) - \dim(S \cap T).$$

**Demostración.** Sea $B := \{v_1, v_2, \ldots, v_r\}$ una base de $S \cap T$. Usaremos el Teorema 5.29: como $B$ es linealmente independiente y $B \subseteq S$, existe una base $B'$ de $S$ que contiene a $B$. Supongamos que $B' = \{v_1, \ldots, v_r, s_1, \ldots, s_n\}$. Similarmente, $B \subseteq T$, así que existe una base $B''$ de $T$ que contiene a $B$; supongamos que $B'' = \{v_1, \ldots, v_r, t_1, \ldots, t_m\}$. Demostraremos que el conjunto

$$C := B' \cup B'' = \{v_1, \ldots, v_r, s_1, \ldots, s_n, t_1, \ldots, t_m\}$$

es una base de $S+T$. Sea $s+t \in S+T$, con $s \in S$, $t \in T$, un vector arbitrario de la suma. Sabemos que $s \in \langle B' \rangle$ y $t \in \langle B'' \rangle$, así que $s+t \in \langle B' \cup B'' \rangle = \langle C \rangle$. Por lo tanto, $S+T = \langle C \rangle$. Para demostrar que $C$ es linealmente independiente, supongamos que

$$\sum_{i=i}^{r} \alpha_i v_i + \sum_{i=1}^{n} \beta_i s_i + \sum_{i=1}^{m} \gamma_i t_i = \mathbf{0}, \tag{6.1}$$

para algunos $\alpha_i, \beta_i, \gamma_i \in F$. Consideremos

$$w := \sum_{i=i}^{r} \alpha_i v_i + \sum_{i=1}^{n} \beta_i s_i = -\sum_{i=1}^{m} \gamma_i t_i.$$

La igualdad anterior implica que $w \in S$ y $w \in T$, así que $w \in S \cap T$. Por lo tanto, existen escalares $\lambda_i \in F$ tales que

$$w = \sum_{i=1}^{r} \lambda_i v_i \;\Rightarrow\; -\sum_{i=1}^{m} \gamma_i t_i = \sum_{i=1}^{r} \lambda_i v_i \;\Rightarrow\; \sum_{i=1}^{r} \lambda_i v_i + \sum_{i=1}^{m} \gamma_i t_i = \mathbf{0}.$$

Como $B''$ es linealmente independiente, tenemos que $\lambda_i = 0$ y $\gamma_i = 0$ para toda $i$. Substituyendo en (6.1), obtenemos

$$\sum_{i=i}^{r} \alpha_i v_i + \sum_{i=1}^{n} \beta_i s_i = \mathbf{0}.$$

Como $B'$ es linealmente independiente, deducimos que $\alpha_i = 0$ y $\beta_i = 0$ para toda $i$. Por lo tanto, $C$ es linealmente independiente. Concluimos que

$$\dim(S+T) = r+n+m = (r+n)+(r+m)-r = \dim(S)+\dim(T)-\dim(S\cap T).$$

□

**Corolario 6.2 (dimensión de la suma directa).** Sean $U$ y $W$ subespacios de un espacio vectorial de dimensión finita $V$ sobre un campo $F$. Supongamos que $U \cap W = \{\mathbf{0}\}$, así que la suma de $U$ y $W$ es directa. Entonces,

$$\dim(U \oplus W) = \dim U + \dim W.$$

**Teorema 6.3 (dimensión del espacio cociente).** Sea $V$ un espacio vectorial sobre $F$ de dimensión finita, y sea $S \leq V$. Entonces,

$$\dim(V/S) = \dim(V) - \dim(S).$$

**Demostración.** Sea $B = \{s_1, \ldots, s_n\}$ una base de $S$ y sea $C = \{w_1 + S, \ldots, w_m + S\}$ una base de $V/S$. Demostraremos que el conjunto

$$D = \{s_1, \ldots, s_n, w_1, \ldots, w_m\}$$

es una base de $V$. Sea $v \in V$ un vector arbitrario y consideremos la clase lateral $v + S$. Como $C$ es una base de $V/S$, existen $\alpha_i \in F$ tales que

$$v + S = \alpha_1(w_1 + S) + \cdots + \alpha_m(w_m + S) = (\alpha_1 w_1 + \cdots + \alpha_m w_m) + S.$$

Por el Lema de Propiedades Básicas de Clases Laterales,

$$\alpha_1 w_1 + \cdots + \alpha_m w_m - v \in S.$$

Como $B$ es una base de $S$, existen escalares $\beta_i \in F$ tales que

$$\sum_{i=1}^{m} \alpha_i w_i - v = \sum_{i=1}^{n} \beta_i s_i \;\Rightarrow\; v = \sum_{i=1}^{n} \beta_i s_i + \sum_{i=1}^{m} \alpha_i w_i \in \langle D \rangle.$$

Esto demuestra que $V = \langle D \rangle$. Supongamos ahora que existen escalares $\gamma_i, \lambda_i \in F$ tales que

$$\sum_{i=1}^{n} \gamma_i s_i + \sum_{i=1}^{m} \lambda_i w_i = \mathbf{0} \;\Rightarrow\; \sum_{i=1}^{m} \lambda_i w_i = -\sum_{i=1}^{n} \gamma_i s_i \in S. \tag{6.2}$$

Por el lema de Propiedades Básicas de Clases Laterales,

$$\sum_{i=1}^{m} \lambda_i w_i + S = \mathbf{0} + S,$$

lo que implica que

$$\lambda_1(w_1 + S) + \cdots + \lambda_m(w_m + S) = \mathbf{0} + S.$$

Como $C$ es linealmente independiente, esto implica que $\lambda_i = 0$, para toda $i$. Substituyendo en (6.2), obtenemos que

$$\gamma_1 s_1 + \cdots + \gamma_n s_n = \mathbf{0}.$$

Como $B$ linealmente independiente, concluimos que $\gamma_i = 0$ para toda $i$, así que $D$ es linealmente independiente. Finalmente,

$$\dim(V) = n + m = \dim(S) + \dim(V/S).$$

□

**Definición 6.4 (rank y nulidad).** Sean $V$ y $W$ espacios vectoriales sobre $F$ y sea $\phi : V \to W$ una transformación lineal. Definimos el *rank* de $\phi$ como la dimensión de la imagen de $\phi$:

$$\mathrm{rk}(\phi) := \dim(\mathrm{im}(\phi)).$$

Definimos la *nulidad* de $\phi$ como la dimensión del kernel de $\phi$:

$$\mathrm{nul}(\phi) := \dim(\ker(\phi)).$$

El siguiente teorema tiene una gran variedad de aplicaciones para demostrar resultados generales sobre transformaciones lineales entre espacios vectoriales de dimensión finita (ver Ejercicio 6.62).

**Teorema 6.5 (rank + nulidad).** Sean $V$ y $W$ espacios vectoriales de dimensión finita sobre $F$ y sea $\phi : V \to W$ una transformación lineal. Entonces

$$\dim(V) = \mathrm{rk}(\phi) + \mathrm{nul}(\phi).$$

**Demostración.** Por el Primer Teorema de Isomorfía, sabemos que

$$V/\ker(\phi) \cong \mathrm{im}(\phi).$$

Como espacios isomorfos tienen la misma dimensión, el Teorema 6.3 implica que

$$\dim(\mathrm{im}(\phi)) = \dim(V/\ker(\phi)) = \dim(V) - \dim(\ker(\phi)).$$

□

**Teorema 6.6 (Dimensión del espacio de transformaciones lineales).** Si $V$ y $W$ son dos espacios vectoriales de dimensión finita sobre $F$. Entonces,

$$\dim\left(\mathrm{Hom}\left(V,W\right)\right)=\dim(V)\cdot\dim(W).$$

**Demostración.** Supongamos que $\dim(V)=m$ y $\dim(W)=n$, y sean $B=\{v_1,...,v_m\}$ y $B'=\{w_1,..,w_n\}$ bases de $V$ y $W$, respectivamente. Debemos encontrar una base para $\mathrm{Hom}\left(V,W\right)$ y demostrar que tiene $mn$ elementos. Para cada $i=1,\ldots,m$ y $j=1,\ldots,n$, definimos $f_{ij}:B\to W$ como

$$f_{ij}\left(v_k\right)=\begin{cases} w_j, & \text{si } k=i \\ 0, & \text{si } k\neq i \end{cases}$$

Sea $\tau_{ij}:V\to W$ la extensión lineal de $f_{ij}:V\to W$. Demostraremos que el conjunto $C:=\{\tau_{ij}:1\leq i\leq m,\ 1\leq j\leq n\}$ forma una base de $\mathrm{Hom}\left(V,W\right)$.

Para comprobar que $C$ es linealmente independiente, supongamos que

$$\sum_{i=1}^{m}\sum_{j=1}^{n}\alpha_{ij}\tau_{ij}=\mathbf{0},\ \alpha_{ij}\in F,$$

Evaluando de ambos lados en cada $v_i\in B$ tenemos

$$\sum_{i=1}^{m}\sum_{j=1}^{n}\alpha_{ij}\tau_{ij}\left(v_i\right)=\sum_{j=1}^{n}\alpha_{ij}w_j=0,$$

y como $B'$ es linealmente independiente, esto implica que $\alpha_{ij}=0$ para todo $i,j$.

Para demostrar que $C$ genera a $\mathrm{Hom}\left(V,W\right)$, sea $\tau\in\mathrm{Hom}\left(U,V\right)$. Para cada $v\in V$ existen escalares $\beta_i,\gamma_{ij}\in F$ tales que

$$\begin{aligned} v&=\beta_1v_1+\ldots\beta_mv_m \\ \tau(v_i)&=\gamma_{i1}w_1+\cdots+\gamma_{1n}w_n. \end{aligned}$$

Por lo tanto

$$\begin{aligned} \tau(v)&=\sum_{i=1}^{n}\beta_i\tau(v_i)=\sum_{i=1}^{n}\sum_{j=1}^{m}\beta_i\gamma_{ij}w_j \\ &=\sum_{i=1}^{n}\sum_{j=1}^{m}\beta_i\gamma_{ij}\tau_{ij}(v_i) \\ &=\sum_{i=1}^{n}\sum_{j=1}^{m}\gamma_{ij}\tau_{ij}(\beta_1v_1+\cdots+\beta_mv_m) \\ &=\sum_{i=1}^{n}\sum_{j=1}^{m}\gamma_{ij}\tau_{ij}(v) \end{aligned}$$

Luego $\langle C\rangle=\mathrm{Hom}\left(V,W\right)$.

Concluimos que $C$ es base para $\mathrm{Hom}\,(U,V)$, y por lo tanto $\dim\left(\mathrm{Hom}\,(V,W)\right) = |C| = mn$. □

Recordemos que $M_{n\times m}\,(F)$ es el espacio vectorial de matrices de $n \times m$ con entradas en $F$.

**Corolario 6.7.** Sean $V$ y $W$ espacios vectoriales de dimensión finita sobre $F$. Sean $n := \dim(V)$ y $m := \dim(W)$. Entonces,

$$\mathrm{Hom}(V,W) \cong M_{m\times n}(F).$$

**Demostración.** Observemos que $\dim(M_{m\times n}(F)) = mn$. El resultado queda establecido usando el Teorema 6.6 y el Teorema de Dimensión e Isomorfía. □

## 6.2. Repaso: Conceptos y operaciones básicas de matrices

Como vimos en el Corolario 6.7, existe una conexión muy estrecha entre transformaciones lineales y matrices. No solo la suma y multiplicación escalar de transformaciones lineales es análoga a la suma y multiplicación escalar de matrices; resulta que la composición de transformaciones lineales es análoga a la multiplicación de matrices. Por tal motivo, en esta sección daremos un breve repaso sobre algunos conceptos y operaciones básicas de matrices.

**Definición 6.8 (multiplicación de matrices).** Sea $A = (a_{i,j}) \in M_{n\times m}\,(F)$, $D = (d_{i,j}) \in M_{m\times k}\,(F)$. Definimos al producto $AD$ como la siguiente matriz de $n \times k$:

$$\begin{aligned} AD &= \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,m} \\ a_{2,1} & a_{2,2} & \dots & a_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,m} \end{pmatrix} \begin{pmatrix} d_{1,1} & d_{1,2} & \dots & d_{1,k} \\ d_{2,1} & d_{2,2} & \dots & d_{2,k} \\ \vdots & \vdots & \ddots & \vdots \\ d_{m,1} & d_{m,2} & \dots & d_{m,k} \end{pmatrix} \\ &= \begin{pmatrix} \sum_{j=1}^{m} a_{1,j}d_{j,1} & \sum_{j=1}^{m} a_{1,j}d_{j,2} & \dots & \sum_{j=1}^{m} a_{1,j}d_{j,k} \\ \sum_{j=1}^{m} a_{2,j}d_{j,1} & \sum_{j=1}^{m} a_{2,j}d_{j,2} & \dots & \sum_{j=1}^{m} a_{2,j}d_{j,k} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{j=1}^{m} a_{n,j}d_{j,1} & \sum_{j=1}^{m} a_{n,j}d_{j,2} & \dots & \sum_{j=1}^{m} a_{n,j}d_{j,k} \end{pmatrix}. \end{aligned}$$

En otras palabras, el elemento $(i,j)$ del producto $AD$ se obtiene multiplicando, respectivamente, los elementos del renglón $i$ de $A$ por los elementos de la columna $j$ de $D$, y sumándolos.

**Observación 6.9.** La multiplicación de una matriz $A \in M_{n\times m}\,(F)$ por $D \in M_{s\times k}\,(F)$ está definida si y sólo si $m = s$, y el producto $AD$ es una matriz de $m \times k$.

**Ejemplo 6.10.** Sean

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \text{ y } D = \begin{pmatrix} x_1 & x_4 & x_7 \\ x_2 & x_5 & x_8 \\ x_3 & x_6 & x_9 \end{pmatrix}.$$

Donde $x_i \in \mathbb{R}$. Entonces,

$$\begin{aligned} AD &= \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \begin{pmatrix} x_1 & x_4 & x_7 \\ x_2 & x_5 & x_8 \\ x_3 & x_6 & x_9 \end{pmatrix} \\ &= \begin{pmatrix} x_1 + 2x_2 + 3x_3 & x_4 + 2x_5 + 3x_6 & x_7 + 2x_8 + 3x_9 \\ 4x_1 + 5x_2 + 6x_3 & 4x_4 + 5x_5 + 6x_6 & 4x_7 + 5x_8 + 6x_9 \\ 7x_1 + 8x_2 + 9x_3 & 7x_4 + 8x_5 + 9x_6 & 7x_7 + 8x_8 + 9x_9 \end{pmatrix}. \end{aligned}$$

**Lema 6.11.** La multiplicación de matrices es una operación asociativa.

**Observación 6.12.** La multiplicación de matrices no es una operación conmutativa. Primero observemos que, para $A \in M_{n\times m}(F)$ y $D \in M_{s\times k}(F)$, los productos $AD$ y $DA$ están definidos si y sólo si $n = m = s = k$. Sin embargo, aún en este caso, es sencillo encontrar ejemplos de matrices cuadradas tales que

$$AD \neq DA.$$

**Ejemplo 6.13.** Sea $I_n$ la matriz de $n \times n$ con elementos diagonales igual a 1 y elementos no diagonales igual a 0:

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

La matriz $I_n$ se llama la *matriz identidad* de $n \times n$, y satisface que

$$I_n A = A I_n = A$$

para cualquier $A \in M_{n\times n}(F)$.

Las siguientes definiciones hacen referencia a tipos especiales de matrices cuadradas.

**Definición 6.14 (matriz triangular).** Sea $A = (a_{i,j}) \in M_{n\times n}(F)$. Decimos que $A$ es *triangular superior* si $a_{i,j} = 0$, para toda $i > j$ . Decimos que $A$ es *triangular inferior* si $a_{i,j} = 0$ para toda $i < j$.

**Definición 6.15 (matriz diagonal).** Sea $A = (a_{i,j}) \in M_{n\times n}(F)$. Decimos que $A$ es *diagonal* si $A$ es triangular superior e inferior al mismo tiempo.

**Ejemplo 6.16.** Cualquier matriz triangular superior de $M_{3\times 3}(\mathbb{R})$ tiene la forma

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

donde $a_{i,j} \in \mathbb{R}$. Cualquier matriz diagonal de $M_{3\times 3}(\mathbb{R})$ tiene la forma

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

donde $a_{i,i} \in \mathbb{R}$.

**Observación 6.17.** Los subconjuntos de $M_{n\times n}(F)$ de matrices triangulares superiores, triangulares inferiores y diagonales son subespacios de $M_{n\times m}(F)$.

**Definición 6.18 (transpuesta de una matriz).** Sea $A = (a_{i,j})$ una matriz de $n \times m$ con entradas en $F$. Definimos a la *transpuesta* de $A$ como $A^T = (a_{j,i})$; en otras palabras, las columnas de $A$ son las filas de $A^T$ y viceversa.

**Ejemplo 6.19.** Si

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

entonces

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

Observemos que $A$ es una matriz de $2 \times 3$ mientras que $A^T$ es una matriz de $3 \times 2$. Veamos también que

$$\left(A^T\right)^T = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{pmatrix} = A.$$

**Teorema 6.20 (propiedades de la transpuesta).** Sean $A, D \in M_{n\times m}(F)$. Entonces

1. $\left(A^T\right)^T = A$.
2. $(A + D)^T = A^T + D^T$.
3. $(\alpha A)^T = \alpha A^T$, $\forall \alpha \in F$.

**Demostración.** Ejericio 7.32. □

A partir de ahora nos enforaremos en estudiar matrices cuadradas.

**Definición 6.21 (inversa de una matriz).** Una matriz $A \in M_{n\times n}(F)$ es *invertible* si existe una matriz $A^{-1} \in M_{n\times n}(F)$ tal que

$$AA^{-1} = A^{-1}A = I_n.$$

En tal caso, decimos que $A^{-1}$ es la *matriz inversa* de $A$.

**Observación 6.22.** En la siguiente sección veremos que si $M \in M_{n\times n}(F)$ es una matriz tal que $AM = I_n$, entonces $MA = I_n$, y viceversa, si $MA = I_n$, entonces $AM = I_n$. Esta es una propiedad muy interesante de las matrices que puede enunciarse como sigue:

$$M \text{ es un } \textit{inverso izquierdo} \text{ de } A \Leftrightarrow M \text{ es un } \textit{inverso derecho} \text{ de } A.$$

Esta es una propiedad que siempre se cumple para los elementos de un grupo, pero no necesariamente para los elementos de cualquier *monoide*[1]. Así que, como veremos más adelante, la demostración de que esta propiedad se cumple para las matrices debe usar resultados propios de la teoría de espacios vectoriales.

**Ejemplo 6.23.** Consideremos las siguientes matrices en $M_{3\times 3}(\mathbb{R})$:

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

Observemos que

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

Por lo tanto, la matriz inversa de $A^{-1} = D$.

No es verdad que cualquier matriz tenga una matriz inversa.

**Ejemplo 6.24.** La matriz

$$A = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix} \in M_{2\times 2}(\mathbb{R})$$

no tiene inversa. Para demostrar esto por reducción al absurdo, supongamos que $AD = I_2$, donde

$$D = \begin{pmatrix} d_{1,1} & d_{1,2} \\ d_{2,1} & d_{2,2} \end{pmatrix}.$$

Entonces, la igualdad

$$AD = \begin{pmatrix} d_{1,1} - d_{2,1} & d_{1,2} - d_{2,2} \\ d_{1,1} - d_{2,1} & d_{1,2} - d_{2,2} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

[1] Un monoide es simplemente un conjunto equipado con una operación binaria asociativa y con identidad. Por ejemplo, $M_{n\times n}(F)$ es un monoide.

es equivalente al siguiente sistema de ecuaciones:

$$\begin{aligned} d_{1,1} - d_{2,1} &= 1; \ d_{1,2} - d_{2,2} = 0; \\ d_{1,1} - d_{2,1} &= 0; \ d_{1,2} - d_{2,2} = 1. \end{aligned}$$

Esto implica que $1 = 0$, lo cual es una contradicción. Por lo tanto, la matriz $A$ no tiene inversa en $M_{2\times 2}(\mathbb{R})$.

**Definición 6.25 (matriz invertible).** Decimos que una matriz cuadrada $A$ es invertible si su inversa $A^{-1}$ existe.

**Proposición 6.26 (matrices invertibles).** Sean $A, D \in M_{n\times n}(F)$ matrices invertibles. Entonces, el producto $AD$ es invertible y

$$(AD)^{-1} = D^{-1}A^{-1}.$$

**Demostración.** Observemos que

$$(AD)\left(D^{-1}A^{-1}\right) = A\left(DD^{-1}\right)A^{-1} = AI_nA^{-1} = I_n.$$

Por lo tanto, $(AD)^{-1} = D^{-1}A^{-1}$ por definición. □

**Observación 6.27.** Sea $n \geq 1$. El siguiente conjunto

$$GL_n(F) := \{A \in M_{n\times n}(F) : A \text{ es invertible}\},$$

junto con la multiplicación de matrices es un grupo no abeliano con identidad $I_n$. A este grupo se le conoce como el *grupo general lineal de grado n sobre F*.

Analicemos más a detalle el caso de las matrices cuadradas de $2 \times 2$.

**Teorema 6.28.** Sea $A \in M_{2\times 2}(F)$ definida como

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

Si $ad - bc \neq 0$, entonces la matriz inversa de $A$ está dada por

$$A^{-1} = \frac{1}{(ad-bc)}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

**Demostración.** Verificamos que

$$\begin{aligned} \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} \frac{d}{ad-bc} & -\frac{b}{ad-bc} \\ -\frac{c}{ad-bc} & \frac{a}{ad-bc} \end{pmatrix} &= \begin{pmatrix} \frac{ad}{ad-bc} - \frac{bc}{ad-bc} & -\frac{ab}{ad-bc} + \frac{ab}{ad-bc} \\ \frac{cd}{ad-bc} - \frac{cd}{ad-bc} & -\frac{bc}{ad-bc} + \frac{ad}{ad-bc} \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned}$$

□

**Definición 6.29 (determinante, caso $2 \times 2$).** Sea $A = (a_{i,j}) \in M_{2\times 2}(F)$. Definimos el *determinante* de $A$ como

$$\det(A) = a_{1,1}a_{2,2} - a_{1,2}a_{2,1}.$$

**Teorema 6.30 (determinante, caso $2 \times 2$).** Sea $A \in M_{2\times 2}(F)$. Entonces, $A$ es invertible si y sólo si $\det(A) \neq 0$.

El teorema anterior puede generalizarse para el caso de matrices de $n \times n$.

**Teorema 6.31 (Linealidad de Matrices).** Sea $F$ un campo. Sea $A \in M_{n\times m}(F)$ y sean $v, u \in M_{m\times 1}(F)$. Entonces,

$$A(\alpha v + \beta u) = \alpha(Av) + \beta(Au),$$

para cualquier $\alpha, \beta \in F$.

**Ejemplo 6.32.** Consideremos $v = (1,2,3)$, $u = (0,1,0) \in \mathbb{R}^3$, $\alpha = 2$, $\beta = 3$, y

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

Entonces,

$$\begin{aligned} A(\alpha v + \beta u) &= \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 2 \end{pmatrix} \left( 2 \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + 3 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right) \\ &= \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} 2 \\ 7 \\ 6 \end{pmatrix} \\ &= \begin{pmatrix} 16 \\ 19 \end{pmatrix}. \end{aligned}$$

Por otro lado,

$$\begin{aligned} \alpha(Av) + \beta(Au) &= 2 \left( \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \right) + 3 \left( \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right) \\ &= 2 \begin{pmatrix} 5 \\ 8 \end{pmatrix} + 3 \begin{pmatrix} 2 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} 16 \\ 19 \end{pmatrix}. \end{aligned}$$

## 6.3. Coordenadas

En esta sección sólo consideraremos espacios de dimensión finita sobre $F$. Como dichos espacios son isomorfos a $F^n$, donde $n$ es la dimensión del espacio (Teorema 5.47), nos enfocaremos sólo en estos casos.

**Definición 6.33 (Coordenadas).** Sea $B = \{b_1, b_2, \ldots, b_n\}$ una base de $F^n$. Sea $v \in F^n$ y sea

$$v = \alpha_1 b_1 + \alpha_2 b_2 + \cdots + \alpha_n b_n,\ \alpha_i \in F,$$

la representación de $v$ como combinación lineal de la base $B$. Las *coordenadas* de un vector $v \in F^n$ respecto a $B$ es la matriz de $n \times 1$ definida por

$$[v]_B := \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}.$$

Para ahorrar espacio, también escribimos $[v]_B = (\alpha_1, \alpha_2, \ldots, \alpha_n)^T$, donde $T$ denota la transpuesta.

**Ejemplo 6.34.** Consideremos la base $D = \{(0,1), (1,-1)\}$ de $\mathbb{R}^2$. Sea $u = (-4,1) \in \mathbb{R}^2$. Debido a que

$$(-4,1) = -3(0,1) - 4(1,-1),$$

tenemos que las coordenadas de $u$ respecto a $D$ son $[u]_D = (-3,-4)^T$.

**Ejemplo 6.35.** Consideremos la base canónica $B = \{(1,0), (0,1)\}$ de $\mathbb{R}^2$. Entonces, para todo $v = (x_1, x_2) \in \mathbb{R}^2$, sabemos que

$$v = x_1(1,0) + x_2(0,1).$$

Por lo tanto, las coordenadas de $(x_1, x_2)$ con respecto a $B$ son

$$[(x_1, x_2)]_B = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Sin embargo, esto es diferente si consideramos otra base de $\mathbb{R}^2$. Por ejemplo, si consideramos la base $C = \{(1,0), (1,1)\}$ de $\mathbb{R}^2$, tenemos que

$$(x_1, x_2) = (x_1 - x_2)(1,0) + x_2(1,1).$$

Por lo tanto, las coordenadas de $(x_1, x_2)$ con respecto a $C$ son

$$[(x_1, x_2)]_C = \begin{pmatrix} x_1 - x_2 \\ x_2 \end{pmatrix}.$$

**Ejemplo 6.36.** Consideremos la base $B = \{(1,0,1), (1,1,0), (0,1,1)\}$ de $\mathbb{R}^3$. Encontraremos las coordenadas de $v = (x_1, x_2, x_3) \in \mathbb{R}^3$ con respecto a $B$. Queremos encontrar escalares $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$, tales que

$$(x_1, x_2, x_3) = \alpha_1(1,0,1) + \alpha_2(1,1,0) + \alpha_3(0,1,1).$$

La igualdad anterior es equivalente al siguiente sistema de ecuaciones:

$$\begin{aligned} \alpha_1 + \alpha_2 &= x_1, \\ \alpha_2 + \alpha_3 &= x_2, \\ \alpha_1 + \alpha_3 &= x_3. \end{aligned}$$

La soluciones del sistema son

$$\begin{aligned} \alpha_1 &= \frac{1}{2}(x_1 - x_2 + x_3), \\ \alpha_2 &= \frac{1}{2}(x_1 + x_2 - x_3), \\ \alpha_3 &= \frac{1}{2}(-x_1 + x_2 + x_3). \end{aligned}$$

Por lo tanto,

$$[(x_1, x_2, x_3)]_B = \begin{pmatrix} \frac{1}{2}(x_1 - x_2 + x_3) \\ \frac{1}{2}(x_1 + x_2 - x_3) \\ \frac{1}{2}(-x_1 + x_2 + x_3) \end{pmatrix}.$$

**Observación 6.37.** Las coordenadas de un vector de $F^n$ con respecto a una base $B$ dependen del orden en el cual aparezcan los vectores base en $B$.

**Definición 6.38 (Matriz asociada a una transformación lineal).** Sea $\phi : F^m \to F^n$ una transformación lineal. Sea $B = \{b_1, \ldots, b_m\}$ una base de $F^m$ y $C = \{c_1, \ldots, c_n\}$ una base de $F^n$. La *matriz asociada a $\phi$ respecto a $B$ y $C$* es la matriz de $n \times m$ definida por

$$[\phi]_B^C := \begin{pmatrix} \uparrow & \uparrow & & \uparrow \\ [\phi(b_1)]_C & [\phi(b_2)]_C & \ldots & [\phi(b_m)]_C \\ \downarrow & \downarrow & & \downarrow \end{pmatrix}.$$

**Observación 6.39.** Si $B$ y $C$ son las bases canónicas de $F^m$ y $F^n$, respectivamente, denotamos a $[\phi]_B^C$ simplemente por $[\phi]$.

**Ejemplo 6.40.** Ahora tomemos $\phi : \mathbb{R}^3 \to \mathbb{R}^2$, dado por $\phi(x, y, z) = (x - 3y, 4y - z)$. Usando como bases $B = \{(1, 0, 0), (0, 1, 1), (0, 0, 2)\}$ y $B' = \{(1, 0), (1, 1)\}$ para $\mathbb{R}^3$ y $\mathbb{R}^2$ respectivamente, obtenemos la matriz asociada a $\phi$, encontrando primero las imágenes de la base $B$ mediante $\phi$,

$$\begin{aligned} \phi(1, 0, 0) &= (1, 0) \\ \phi(0, 1, 1) &= (-3, 3) \\ \phi(0, 0, 2) &= (0, -2) \end{aligned}$$

para luego obtener las coordenadas de dichas imágenes en base $B'$,

$$\begin{aligned}[(1,0)]_{B'} &= (1,0)^T \\ [(-3,3)]_{B'} &= (-6,3)^T \\ [(0,-2)]_{B'} &= (2,-2)^T\end{aligned}$$

para finalmente escribir la matriz asociada$[\phi]_B^{B'} = \begin{pmatrix} 1 & -6 & 2 \\ 0 & 3 & -2 \end{pmatrix}$.

**Ejemplo 6.41.** Sea $P_1$ el subespacio de $\mathbb{R}[x]$ de polinomios con grado menor o igual a 1. Sea $\phi : P_1 \to \mathbb{R}^2$, dado por $\phi(ax+b) = (a+2b, a)$. Tomemos bases $B = \{x-1, 1\}$ y $B' = \{(1,0),(1,1)\}$ para $P_1$ y $\mathbb{R}^2$ respectivamente. Para obtener la matriz asociada a $\phi$, encontramos las imágenes de la base $B$ mediante $\phi$,

$$\begin{aligned}\phi(x-1) &= (-1,1) \\ \phi(1) &= (2,0)\end{aligned}$$

y obtenemos las coordenadas de dichas imágenes en base $B'$, ya que $(-1,1) = -2(1,0) + (1,1)$ y $(2,0) = 2(1,0) + 0(1,1)$, entonces

$$\begin{aligned}(-1,1)_{B'} &= \begin{pmatrix} -2 \\ 1 \end{pmatrix} \\ (2,0)_{B'} &= \begin{pmatrix} 2 \\ 0 \end{pmatrix}\end{aligned}$$

por lo que la matriz asociada es $[\phi]_B^{B'} = \begin{pmatrix} -2 & 2 \\ 1 & 0 \end{pmatrix}$.

**Teorema 6.42 (Matriz de una transformación lineal).** Sea $F$ un campo, $B$ una base de $F^m$ y $C$ una base de $F^n$. Sea $\phi : F^m \to F^n$ una transformación lineal. Entonces, para todo $v \in F^m$,

$$[\phi]_B^C [v]_B = [\phi(v)]_C.$$

**Demostración.** Sea $B = \{b_1, ..., b_m\}$ una base de $F^m$ y $C = \{c_1, ..., c_n\}$ una base de $F^n$. Sea

$$v = \beta_1 b_1 + ... + \beta_m b_m \in F^m$$

un vector arbitrario. Debido a que

$$\phi(v) = \phi(\beta_1 b_1 + ... + \beta_m b_m) = \beta_1 \phi(b_1) + ... + \beta_m \phi(b_m), \qquad (*)$$

vemos que la transformación $\phi$ está completamente determinada por los vectores $\phi(b_j) \in F^n$, $j = 1, ..., m$. Ahora, como $C$ es una base de $F^n$, es posible escribir

cada $\phi(b_j)$ como una combinación lineal única de elementos de $C$:

$$\begin{aligned}
\phi(b_1) &= a_{1,1}c_1 + a_{2,1}c_2 + ... + a_{n,1}c_n, \\
\phi(b_2) &= a_{1,2}c_1 + a_{2,2}c_2 + ... + a_{n,2}c_n, \\
&\vdots \\
\phi(b_m) &= a_{1,m}c_1 + a_{2,m}c_2 + ... + a_{n,m}c_n.
\end{aligned}$$

Entonces, las coordenadas de $\phi(b_j)$ con respecto a $C$ son

$$[\phi(b_j)]_C = \begin{pmatrix} a_{1,j} \\ a_{2,j} \\ \vdots \\ a_{n,j} \end{pmatrix}.$$

Resulta que los escalares $a_{i,j} \in F$ determinan completamente la transformación lineal $\phi$. Efectivamente, sustituyendo en $(*)$

$$\begin{aligned}
\phi(v) &= \beta_1\phi(b_1) + ... + \beta_m\phi(b_m) \\
&= \beta_1 \sum_{i=1}^{n} a_{i,1}c_i + \beta_2 \sum_{i=1}^{n} a_{i,2}c_i + ... + \beta_m \sum_{i=1}^{n} a_{i,n}c_i \\
&= \left(\sum_{j=1}^{m} \beta_j a_{1,j}\right) c_1 + \left(\sum_{j=1}^{m} \beta_j a_{2,j}\right) c_2 + ... \left(\sum_{j=1}^{m} \beta_j a_{n,j}\right) c_n
\end{aligned}$$

por lo que las coordenadas de $\phi(v)$ con respecto a $C$ son

$$[\phi(v)]_C = \begin{pmatrix} \sum_{j=1}^{m} \beta_j a_{1,j} \\ \sum_{j=1}^{m} \beta_j a_{2,j} \\ \vdots \\ \sum_{j=1}^{m} \beta_j a_{n,j} \end{pmatrix}.$$

Por otro lado, la matriz asociada con $\phi$ respecto a las bases $B$ y $C$ es

$$[\phi]_B^C = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & ... & a_{1,m} \\ a_{2,1} & a_{2,2} & a_{2,3} & ... & a_{2,m} \\ a_{3,1} & a_{3,2} & a_{3,3} & ... & a_{3,m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & a_{n,3} & ... & a_{n,m} \end{pmatrix},$$

y las coordenadas de $v$ respecto a $B$ son $[v]_B = (\beta_1, \beta_2, \ldots, \beta_m)^T$. Por lo tanto,

el producto $[\phi]_B^C[v]_B$ es igual a $[\phi(v)]_C$:

$$\begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & ... & a_{1,m} \\ a_{2,1} & a_{2,2} & a_{2,3} & ... & a_{2,m} \\ a_{3,1} & a_{3,2} & a_{3,3} & ... & a_{3,m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & a_{n,3} & ... & a_{n,m} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{pmatrix} = \begin{pmatrix} a_{1,1}\beta_1 + a_{1,2}\beta_2 + ... + a_{1,m}\beta_m \\ a_{2,1}\beta_1 + a_{2,2}\beta_2 + ... + a_{2,m}\beta_m \\ \vdots \\ a_{n,1}\beta_1 + a_{n,2}\beta_2 + ... + a_{n,m}\beta_m \end{pmatrix}.$$

□

Si $B$ es la base canónica de $F^m$ y $C$ es la base canónica de $F^n$, las coordenadas $[v]_B$ y $[\phi(v)]_C$ coinciden con los vectores mismos (excepto que es $[v]_B$ una columna y $v$ una fila). En este caso, para simplificar notación, escribimos el resultado del teorema anterior como

$$[\phi]v = \phi(v),$$

donde los vectores $v$ y $\phi(v)$ deben ser escritos como columnas.

**Ejemplo 6.43.** Consideremos la transformación lineal $\phi : \mathbb{R}^3 \to \mathbb{R}^2$ definida como

$$\phi(x_1, x_2, x_3) = (2x_1 + x_3, 5x_2 - 2x_3).$$

Para obtener la matriz $[\phi]$ (con respecto a las bases canónicas), debemos encontrar las imágenes de los elementos de la base:

$$\begin{aligned} \phi(1,0,0) &= (2,0), \\ \phi(0,1,0) &= (0,5), \\ \phi(0,0,1) &= (1,-2). \end{aligned}$$

Entonces $[\phi]$ es una matriz de $2 \times 3$ y tiene como columnas a los vectores anteriores:

$$[\phi] = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 5 & -2 \end{pmatrix}$$

Para verificar que $M_\phi v = \phi(v)$, sea $v = (x_1, x_2, x_3) \in \mathbb{R}^3$. Entonces,

$$[\phi]v = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 5 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2x_1 + x_3 \\ 5x_2 - 2x_3 \end{pmatrix} = \phi(v).$$

**Ejemplo 6.44.** Sea $\phi : \mathbb{R}^3 \to \mathbb{R}^2$ la transformación lineal del ejemplo anterior. Ahora encontraremos la matriz $[\phi]_B^C$ con respecto a las bases $B = \{(1,0,1), (0,2,0), (0,0,1)\}$ y $C = \{(1,0), (1,1)\}$. Primero encontramos las imágenes de los elementos de $B$:

$$\begin{aligned} \phi(1,0,1) &= (3,-2), \\ \phi(0,2,0) &= (0,10), \\ \phi(0,0,1) &= (1,-2). \end{aligned}$$

Ahora encontramos las coordenadas de los vectores anteriores con respecto a $C$:

$$[\phi(1,0,1)]_C = \begin{pmatrix} 5 \\ -2 \end{pmatrix},\ [\phi(0,2,0)]_C = \begin{pmatrix} -10 \\ 10 \end{pmatrix},\ [\phi(0,0,1)]_C = \begin{pmatrix} 3 \\ -2 \end{pmatrix}.$$

Por lo tanto,

$$[\phi]_B^C = \begin{pmatrix} 5 & -10 & 3 \\ -2 & 10 & -2 \end{pmatrix}.$$

**Ejemplo 6.45.** Retomando el Ejemplo 6.40, sea $u = (3,-2,6) \in \mathbb{R}^3$, ya que $(3,-2,6) = 3(1,0,0) - 2(0,1,1) + 4(0,0,2)$, entonces $[u]_B = (3,-2,4)^T$. Por otro lado, $\phi(u) = (9,-14) = 23(1,0) - 14(1,1)$, por lo que $[\phi(u)]_{B'} = (23,-14)^T$. Es fácil comprobar que $[\phi(u)]_{B'} = [\phi]_B^{B'}[u]_B$.

**Observación 6.46.** Observemos que $[\mathrm{id}] = I_n$, donde $\mathrm{id} : F^n \to F^n$ es el endomorfismo identidad.

**Observación 6.47.** Sea $A \in M_{n\times m}(F)$ y sea $\{e_i : 1 \leq i \leq m\}$ la base canónica de $F^m$, donde

$$e_i = (0,...,0,1,0,...,0).$$

Entonces, por la definición de la multiplicación de una matriz por un vector, tenemos que $Ae_i$ es igual a la $i$-ésima columna de $A$.

**Observación 6.48.** El Teorema 6.31 establece que la multiplicación de una matriz $A \in M_{n\times m}(F)$ por un vector $v \in F^m$ define una tranformación lineal $\phi : F^m \to F^n$ como

$$\phi(v) = Av^T.$$

Como $\phi(e_i) = Ae_i$ coincide con la $i$-ésima columna de $A$, tenemos que $A = [\phi]$.

**Teorema 6.49 (Igualdad).** Sean $\phi, \tau \in \mathrm{Hom}(F^m, F^n)$ y sean $B$ y $C$ bases cualesquiera de $F^m$ y $F^n$, respectivamente. Entonces,

$$[\phi]_B^C = [\tau]_B^C \text{ si y sólo si } \phi = \tau.$$

**Demostración.** Si $\phi = \tau$, está claro que $[\phi]_B^C = [\tau]_B^C$. Para demostrar el converso, supongamos que $[\phi]_B^C = [\tau]_B^C$. Luego, las columnas de ambas matrices son iguales, así que $[\phi(b)]_C = [\tau(b)]_C$, para toda $b \in B$. Por propiedades de bases, las coordenadas de dos vectores respecto a $C$ son iguales si y sólo si los vectores son iguales; por lo tanto, $\phi(b) = \tau(b)$, para toda $b \in B$. Como una transformación lineal está completamente determinada por sus imágenes en cualquier base, deducimos que $\phi = \tau$. □

La importancia de la definición de multiplicación de matrices recae en la correspondencia entre multiplicar matrices y obtener la composición de transformaciones lineales. Sabemos que, si $\phi : F^m \to F^n$ y $\tau : F^k \to F^r$ son transformaciones lineales, la composición $\phi \circ \tau$ existe si y sólo si $m = r$.

**Teorema 6.50 (Multiplicación de matrices).** Sean $\tau : F^k \to F^m$ y $\phi : F^m \to F^n$ transformaciones lineales. Entonces,

$$[\phi \circ \tau]_R^C = [\phi]_B^C [\tau]_R^B,$$

donde $R$ una base de $F^k$, $B$ una base de $F^m$ y $C$ una base de $F^n$.

**Demostración.** Sea $v \in F^k$. Usaremos repetidas veces el Teorema 6.42:

$$\begin{aligned} [\phi \circ \tau]_R^C [v]_R &= [\phi \circ \tau(v)]_C \\ &= [\phi(\tau(v))]_C \\ &= [\phi]_B^C [\tau(v)]_B \\ &= [\phi]_B^C [\tau]_R^B [v]_R. \end{aligned}$$

Como el vector $v$ es arbitrario, en particular puede tomar los valores de la base $R$, lo que hace que $[v]_R$ sean vectores columna canónicos. Por la Observación 6.47, tenemos que las columnas de $[\phi \circ \tau]_R^C$ deben ser iguales a las columnas de $[\phi]_B^C [\tau]_R^B$, y por lo tanto deben ser matrices iguales. □

Si tomamos siempre las bases canonicas, el teorema anterior establece que la matriz que representa la transformación $\phi \circ \tau : F^k \to F^n$ coincide con el producto de $[\phi]$ por $[\tau]$.

**Observación 6.51.** Por el Teorema de Multiplicación de Matrices, la multiplicación de matrices es asociativa, pero no es conmutativa.

**Ejemplo 6.52.** Consideremos las transformaciones lineales $\phi : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ y $\tau : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ definidas como

$$\begin{aligned} \phi(x_1, x_2, x_3) &= (2x_1 + x_3, 5x_2 - 2x_3), \\ \tau(y_1, y_2, y_3) &= (y_1, y_2, y_1 + y_2 + y_3). \end{aligned}$$

En un ejemplo anterior encontramos que

$$[\phi] = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 5 & -2 \end{pmatrix}.$$

Para encontrar $[\tau]$ calculamos que

$$\begin{aligned} \tau(1,0,0) &= (1,0,1), \\ \tau(0,1,0) &= (0,1,1), \\ \tau(0,0,1) &= (0,0,1), \end{aligned}$$

así que

$$[\tau] = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}.$$

La composición $\phi \circ \tau : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ es

$$\begin{aligned} \phi \circ \tau (y_1, y_2, y_3) &= \phi (y_1, y_2, y_1 + y_2 + y_3) \\ &= (3y_1 + y_2 + y_3, -2y_1 + 3y_2 - 2y_3) . \end{aligned}$$

Para encontrar $[\phi \circ \tau]$ calculamos que

$$\begin{aligned} \phi \circ \tau (1, 0, 0) &= (3, -2) , \\ \phi \circ \tau (0, 1, 0) &= (1, 3) , \\ \phi \circ \tau (0, 0, 1) &= (1, -2) , \end{aligned}$$

así que

$$[\phi \circ \tau] = \begin{pmatrix} 3 & 1 & 1 \\ -2 & 3 & -2 \end{pmatrix} .$$

Finalmente, comprobamos que

$$\begin{aligned} [\phi][\tau] &= \begin{pmatrix} 2 & 0 & 1 \\ 0 & 5 & -2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 3 & 1 & 1 \\ -2 & 3 & -2 \end{pmatrix} = [\phi \circ \tau]. \end{aligned}$$

Recordemos que un automorfismo es un endomorfismo biyectivo.

**Teorema 6.53 (Endomorfismo).** Un endomorfismo $\phi : F^n \to F^n$ es un automorfismo si y sólo si $[\phi]$ es una matriz invertible.

**Demostración.** Demostraremos cada implicación.

($\Rightarrow$) Sabemos que $\phi : F^n \to F^n$ es biyectivo si y sólo si existe un endomorfismo $\phi^{-1} : F^n \to F^n$ tal que

$$\phi \circ \phi^{-1} = \phi^{-1} \circ \phi = \mathrm{id},$$

Por el Teorema 6.42,

$$[\phi \circ \phi^{-1}] = [\phi][\phi^{-1}] = I_n$$

Esto implica que $[\phi]^{-1} = [\phi^{-1}]$, así que $[\phi]$ es invertible.

($\Leftarrow$) Supongamos que existe una matriz $D \in M_{n \times n} (F)$ tal que $[\phi]D = D[\phi] = I_n$. Sabemos que $D$ define un endomorfismo $\tau : F^n \to F^n$ como $\tau (v) = Dv^T$, donde $v \in F^n$. Claramente, $D = [\tau]$, así que

$$[\phi][\tau] = [\phi \circ \tau] = I_n = [\mathrm{id}].$$

Por el Teorema 6.49, $\phi \circ \tau = \mathrm{id}$. Similarmente, demostramos que $\tau \circ \phi = \mathrm{id}$, así que $\phi$ es un automorfismo.

□

**Teorema 6.54 (Una matriz conmuta con su inversa).** Sean $M, A \in M_{n\times n}(F)$ matrices tales que $AM = I_n$. Entonces, $MA = I_n$.

**Demostración.** Sea $\phi : F^n \to F^n$ el endomorfismo asociado con $A$ (i.e. $\phi(v) = Av^T$, para todo $v \in F^n$). Definimos $\tau : F^n \to F^n$ como $\tau(v) = Mv^T$, para todo $v \in F^n$. Entonces,

$$AM = [\phi][\tau] = [\phi \circ \tau] = I_n = [\mathrm{id}].$$

Por el Teorema 6.49, $\phi \circ \tau = \mathrm{id}$. Luego, $\phi$ es sobreyectiva (ya que una preimagen de $v \in F^n$ es $w := \tau(v) \in F^n$ porque $\phi(w) = \phi(\tau(v)) = \mathrm{id}(v) = v$). Por el Teorema Rank + Nulidad,

$$n = \dim(F^n) = \mathrm{rk}(\phi) + \mathrm{nul}(\phi) = n + \mathrm{nul}(\phi) \;\Rightarrow\; \mathrm{nul}(\phi) = 0.$$

Esto demuestra que $\phi$ es inyectiva. Es un resultado elemental de funciones que una función es inyectiva si y sólo si tiene un inverso izquierdo; es decir, existe una transformación lineal $\sigma : F^n \to F^n$ tal que $\sigma \circ \phi = \mathrm{id}$. Luego,

$$\tau = \mathrm{id} \circ \tau = (\sigma \circ \phi) \circ \tau = \sigma \circ (\phi \circ \tau) = \sigma \circ \mathrm{id} = \sigma.$$

Esto demuestra que $[\tau][\phi] = I_n$, y $MA = I_n$. □

**Observación 6.55.** En general, si $X$ es un conjunto y $f : X \to X$ y $g : X \to X$ son funciones, no es verdad que $f \circ g = \mathrm{id}$ implique que $g \circ f = \mathrm{id}$. Sin embargo, como lo demuestra el teorema anterior, esto sí es verdad para endomorfismos de espacios vectoriales.

**Definición 6.56 (Matriz del Cambio de Base).** Sean $B$ y $C$ son dos bases de $F^n$. La *matriz del cambio de base* de $B$ a $C$ es

$$P := \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ [b_1]_C & [b_2]_C & \cdots & [b_n]_C \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}$$

donde $B = \{b_1, \ldots, b_n\}$.

**Observación 6.57.** Veamos que si $P$ es la matriz del cambio de base de $B$ a $C$, entonces $P = [\mathrm{id}]_B^C$, donde $\mathrm{id} : F^n \to F^n$ es la función identidad.

**Ejemplo 6.58.** Tomemos $V = \mathbb{R}^2$, y las bases $B = \{(0,1), (1,3)\}$, $B' = \{(1,1), (1,0)\}$. Primero calculamos las coordenadas de los elementos de base $B$ en la base $B'$. Como $(0,1) = (1,1) - (1,0)$ y $(1,3) = 3(1,1) - 2(1,0)$, entonces

$$[(0,1)]_{B'} = (1,-1)^T \text{ y } [(1,3)]_{B'} = (3,-2)^T,$$

por lo que la matriz del cambio de base de $B$ a $B'$ es

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

Por otro lado, $(1,1) = -2(0,1) + (1,3)$ y $(1,0) = -3(0,1) + (1,3)$, así que

$$[(1,1)]_B = (-2,1)^T \text{ y } [(1,0)]_B = (-3,1)^T,$$

por lo que la matriz del cambio de base de $B'$ a $B$ es

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

Es importante observar que $PQ = I_2$, por lo que $P^{-1} = Q$. Esto no es una coincidencia como se verá a continuación.

**Teorema 6.59 (Cambio de Base).** Sea $\phi : F^n \to F^n$ un endomorfismo, y sean $B$ y $C$ dos bases de $F^n$. Sea $P$ la matriz del cambio de base de $B$ a $C$. Entonces:

1. $P[v]_B = [v]_C$, para todo $v \in F^n$.
2. $P$ es invertible.
3. $[\phi]_C^C = P[\phi]_B^B P^{-1}$.

**Demostración.** Demostraremos cada punto:

1. Recordemos que $P = [\mathrm{id}]_B^C$. Por el Teorema 6.42,

$$P[v]_B = [\mathrm{id}]_B^C [v]_B = [\mathrm{id}(v)]_C = [v]_C,$$

para todo $v \in F^n$.

2. Demostraremos que la inversa de $P$ es $P^{-1} = [\mathrm{id}]_C^B$. Por el Teorema 6.50,

$$[\mathrm{id}]_B^C [\mathrm{id}]_C^B = [\mathrm{id} \circ \mathrm{id}]_C^C = [\mathrm{id}]_C^C = I_n.$$

3. De nuevo por el Teorema 6.50,

$$P[\phi]_B^B P^{-1} = [\mathrm{id}]_B^C [\phi]_B^B [\mathrm{id}]_C^B = [\mathrm{id}]_B^C [\phi \circ \mathrm{id}]_C^B = [\mathrm{id} \circ \phi]_C^C = [\phi]_C^C.$$

□

**Ejemplo 6.60.** Consideremos el endomorfimso $\phi : \mathbb{R}^2 \to \mathbb{R}^2$ definido como

$$\phi(x_1, x_2) = (x_1, 2x_1 + x_2).$$

Consideremos las bases $B = \{e_1 = (1,0), e_2 = (0,1)\}$ y $C = \{c_1 = (1,1), c_2 = (1,-1)\}$. Las matriz $[\phi]_B^B$ es

$$[\phi]_B^B = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}.$$

Para encontrar la matriz $[\phi]_C^C$ veamos que

$$\begin{aligned} [\phi(1,1)]_C &= [(1,3)]_C = [2c_1 - c_2]_C = \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \\ [\phi(1,-1)]_C &= [(1,1)]_C = [c_1]_C = \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \end{aligned}$$

Por lo tanto,

$$[\phi]_C^C = \begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix}.$$

Escribimos los vectores de $B$ como combinaciones lineales de los vectores de $C$:

$$\begin{aligned} (1,0) &= \frac{1}{2}(1,1) + \frac{1}{2}(1,-1), \\ (0,1) &= \frac{1}{2}(1,1) - \frac{1}{2}(1,-1). \end{aligned}$$

Por lo tanto, la matriz de cambio de base de $B$ a $C$ es

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

Por el Teorema 6.28,

$$P^{-1} = \frac{1}{\left(\frac{1}{2}\left(-\frac{1}{2}\right) - \frac{1}{2}\frac{1}{2}\right)} \begin{bmatrix} -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

Ahora podemos comprobar que

$$\begin{aligned} P[\phi]_B^B P^{-1} &= \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 3 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix} = [\phi]_C^C. \end{aligned}$$

**Observación 6.61.** Recordemos que

$$M_{n\times m}(F) \cong \mathrm{Hom}(F^m, F^n).$$

Para bases fijas $B$ y $C$ de $F^m$ y $F^n$, respectivamente, un isomorfismo entre estos espacios está dado por

$$\phi \mapsto [\phi]_B^C \in M_{n\times m}(F), \forall \phi \in \mathrm{Hom}(F^m, F^n).$$

***Palabras clave:*** *dimensión de la suma de subespacios, dimensión del espaico cociente, teorema rank + nulidad, dimensión del espacio de transformaciones lineales, multiplicación de matrices, matriz invertible, coordenadas, matriz asociada a una transformación lineal, matriz del cambio de base.*

## 6.4. Ejercicios

**Ejercicio 6.62.** Sean $V$ y $W$ espacios vectoriales sobre $F$ de dimensión finita. Sea $\phi : V \to W$ una transformación lineal cualquiera y sea $\phi_0 : V \to W$ la transformación cero (es decir, $\phi_0(v) = \vec{0}_W$, $\forall v \in V$). Usa el **Teorema Rank + Nulidad** para demostrar lo siguiente:

a) Si $\dim(W) = 1$ y $\phi \neq \phi_0$, entonces $\phi : V \to W$ es sobreyectiva.

b) Si $\dim(V) = 1$ y $\phi \neq \phi_0$, entonces $\phi : V \to W$ es inyectiva.

c) Si $\dim(V) = \dim(W)$ y $\phi : V \to W$ es inyectiva, entonces $V \cong W$.

d) Si $\phi : V \to W$ es inyectiva, entonces $\dim(V) \leq \dim(W)$.

e) Si $\phi : V \to W$ es sobreyectiva, entonces $\dim(V) \geq \dim(W)$.

f) Si $\phi : V \to W$ es biyectiva, entonces $\dim(V) = \dim(W)$.

**Ejercicio 6.63.** Encuentra todos los valores de $\alpha \in \mathbb{R}$ tales que

$$B = \{(1, \alpha, -\alpha), (1 - \alpha, 0, 1), (0, 1, 1)\}$$

sea una base de $\mathbb{R}^3$.

**Ejercicio 6.64.** Encuentra todos los valores de $\alpha \in \mathbb{R}$ tales que el endomorfismo $\phi : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ definido por

$$\phi(x_1, x_2, x_3) = (\alpha x_1, x_1 + (\alpha - 1)x_2, 2x_1 + 3x_2 - (\alpha - 2)x_3)$$

sea un automorfismo.

**Ejercicio 6.65.** Encuentra la extensión lineal $\tau : \mathbb{R}^2 \to \mathbb{R}^2$ de la función $f : B \longrightarrow \mathbb{R}^2$, donde $B = \{(1, 0), (0, 1)\}$, y

$$\begin{aligned} f(1,0) &= (1,1), \\ f(0,1) &= (0,0). \end{aligned}$$

**Ejercicio 6.66.** Encuentra las coordenadas un vector arbitrario con respecto a las siguientes bases:

1. $B = \{(2, 1), (1, 2)\} \subseteq \mathbb{R}^2$.
2. $B = \{(1, 0, 0), (1, 1, 0), (1, 1, 1)\} \subseteq \mathbb{R}^3$.

**Ejercicio 6.67.** Encuentra el dominio, el codominio y la regla de las transformaciones lineales definidas por las siguientes matrices con entradas reales:

1. $A_1 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$.
2. $A_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 2 & 2 \end{bmatrix}$.
3. $A_3 = \begin{bmatrix} 1 & 0 & \alpha \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, donde $\alpha \in \mathbb{R}$.

**Ejercicio 6.68.** Encuentra las matrices con respecto a las bases canónicas de las siguientes transformaciones lineales:

1. $\phi : \mathbb{R}^2 \to \mathbb{R}^2$, definida por $\phi(x_1, x_2) = (x_1 + x_2, x_1 - x_2)$.
2. $\varphi : \mathbb{R}^3 \to \mathbb{R}$, definida por $\varphi(x_1, x_2, x_3) = 2x_1 + 3x_2 - x_3$.
3. $\tau : \mathbb{R}^3 \to \mathbb{R}^3$, definida por $\tau(x_1, x_2, x_3) = (x_2, x_3, x_1)$.
4. $\theta : \mathbb{R}^2 \to \mathbb{R}^2$, definida por $\theta(x_1, x_2) = (\alpha x_1 + \beta x_2, \gamma x_1 + \delta x_2)$, donde $\alpha, \beta, \gamma, \delta \in \mathbb{R}$.
5. $\sigma : \mathbb{R}^3 \to \mathbb{R}^2$, definida por $\sigma(x_1, x_2, x_3) = (x_1 + 2x_2 - x_3, x_2 + x_3)$.

**Ejercicio 6.69.** Con respecto a las transformaciones de los puntos (3.) y (5.) del ejercicio anterior, encuentra $\sigma \circ \tau$ y $[\sigma][\tau]$, y comprueba que $[\sigma \circ \tau] = [\sigma][\tau]$.

**Ejercicio 6.70.** Considera las bases $B = \{(1, -1), (1, 1)\}$ y $C = \{(0, 1), (1, 1)\}$ de $\mathbb{R}^2$, y el endomorfismo $\phi : \mathbb{R}^2 \to \mathbb{R}^2$ definido por

$$\phi(x_1, x_2) = (x_1 - x_2, x_1 + 2x_2).$$

1. Encuentra las matrices $[\phi]_B^B$ y $[\phi]_C^C$.
2. Encuentra la matriz $P$ del cambio de base $B$ a $C$.
3. Comprueba que $[\phi]_C^C = P\,[\phi]_B^B\,P^{-1}$.

**Ejercicio 6.71.** Sea $F$ un campo. Demuestra que la función $\Phi : \mathrm{Hom}(\mathbb{F}^m, \mathbb{F}^n) \to M_{n\times m}(F)$ definida por $\Phi(\tau) = [\tau]$, $\forall \tau \in \mathrm{Hom}(F^m, F^n)$ es un isomorfismo de espacios vectoriales.

**Ejercicio 6.72.** Sea $\tau : \mathbb{R}^m \to \mathbb{R}^n$ una transformación lineal. Demuestra lo siguiente:

1. $\tau$ es inyectiva si y sólo si las columnas de $[\tau]$ son linealmente independientes.
2. $\tau$ es sobreyectiva si y sólo si las columnas de $[\tau]$ generan a $\mathbb{R}^n$.
3. $\tau$ es biyectiva si y sólo si las columnas de $[\tau]$ son una base de $\mathbb{R}^n$

# 7

# Teoría de Matrices

En este capítulo abordaremos algunos conceptos y resultados propios de la teoría de matrices como los determinantes, las matrices elementales y el Teorema Fundamental de Matrices Invertibles. Asumimos que el lector ya ha tenido cierta experiencia operativa con las matrices, así que tomamos un enfoque más teórico y conceptual.

## 7.1. Determinante de una matriz

Existen varias definiciones equivalentes del determinante de una matriz. En esta sección adoptaremos la definición conocida como *expansión de Laplace* o *expansión por cofactores*.

Si $A \in M_{n\times n}(F)$, denotemos por $A_{i,j}$ a la matriz de $(n-1)\times(n-1)$ obtenida después de haber eliminado el renglón $i$ y la columna $j$ de $A$. La matriz $A_{i,j}$ se llama una *menor* de $A$.

**Definición 7.1 (determinante).** Sea $A=(a_{i,j})$ una matriz de $n\times n$. Para cualquier $j=1,\ldots,n$ fijo, el *determinante* de $A$ *a lo largo de la columna* $j$ es

$$\begin{aligned}\det(A) &= \sum_{i=1}^{n}(-1)^{i+j}a_{i,j}\det(A_{i,j})\\ &= a_{1,j}\det(A_{1,j}) - a_{2,j}\det(A_{2,j}) + \ldots + (-1)^{n+j}\,a_{n,j}\det(A_{n,j}).\end{aligned}$$

Observemos que la definición anterior es recursiva, en el sentido de que para calcular el determinante de una matriz de $n\times n$ es necesario saber cómo calcular el determinante de una matriz de $(n-1)\times(n-1)$. Este problema no es tan grave: después de algunos pasos, tendremos que calcular el determinante de una matriz de $2\times 2$, el cual ya se ha definido previamente. De hecho, también es posible empezar por definir el determinante de una matriz de $1\times 1$, $A=(a)$, como $\det(A)=a$, y entonces el determinante de una matriz de $2\times 2$ se obtiene usando la Definición 7.1.

**Observación 7.2.** La Definición 7.1 depende de la columna $j$ a lo largo de la cual se haga la expansión. Sin embargo, es posible demostrar que el resultado

del determinante no cambiará, independientemente de la columna a lo largo de la cual se haga la expansión (ver [9, Corolario 4.17]). Por el Teorema 7.6, la expansión del determinante también puede hacerse a lo largo de cualquier renglón, en lugar de cualquier columna.

**Ejemplo 7.3.** Sea

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

Calculamos el determinante de $A$ a lo largo de la columna 1,

$$\begin{aligned} \det(A) &= 1\det\begin{pmatrix} 1 & \frac{1}{2} \\ 0 & 1 \end{pmatrix} - (-2)\det\begin{pmatrix} 0 & -2 \\ 0 & 1 \end{pmatrix} + \frac{1}{2}\det\begin{pmatrix} 0 & -2 \\ 1 & \frac{1}{2} \end{pmatrix} \\ &= 1\left(1\cdot 1 - \frac{1}{2}\cdot 0\right) + 2\left(0\cdot 1 + 2\cdot 0\right) + \frac{1}{2}\left(0\cdot\frac{1}{2} + 2\cdot 1\right) \\ &= 1 + 0 + \left(\frac{1}{2}\right)2 \\ &= 2. \end{aligned}$$

Por otro lado, el determinante de $A$ a lo largo de la columna 2 es

$$\begin{aligned} \det(A) &= -0\det\begin{pmatrix} -2 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{pmatrix} + 1\det\begin{pmatrix} 1 & -2 \\ \frac{1}{2} & 1 \end{pmatrix} - 0\det\begin{pmatrix} 1 & -2 \\ -2 & \frac{1}{2} \end{pmatrix} \\ &= 1 + 2\left(\frac{1}{2}\right) \\ &= 2. \end{aligned}$$

**Ejemplo 7.4.** Consideremos el caso general de una matriz de $3\times 3$:

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

El determinante de $A$ a lo largo de la columna 1 es

$$\begin{aligned} \det(A) = a_{1,1}\det\begin{pmatrix} a_{2,2} & a_{2,3} \\ a_{3,2} & a_{3,3} \end{pmatrix} &- a_{2,1}\det\begin{pmatrix} a_{1,2} & a_{1,3} \\ a_{3,2} & a_{3,3} \end{pmatrix} \\ &+ a_{3,1}\det\begin{pmatrix} a_{1,2} & a_{1,3} \\ a_{2,2} & a_{2,3} \end{pmatrix} \\ = a_{1,1}\left(a_{2,2}a_{3,3} - a_{2,3}a_{3,2}\right) &- a_{2,1}\left(a_{1,2}a_{3,3} - a_{1,3}a_{3,2}\right) \\ &+ a_{3,1}\left(a_{1,2}a_{2,3} - a_{1,3}a_{2,2}\right) \\ = a_{1,1}a_{2,2}a_{3,3} + a_{1,3}a_{2,1}a_{3,2} &+ a_{1,2}a_{2,3}a_{3,1} - a_{1,1}a_{2,3}a_{3,2} \\ &- a_{1,2}a_{2,1}a_{3,3} - a_{1,3}a_{2,2}a_{3,1}. \end{aligned}$$

Invitamos al lector a calcular el determinante a lo largo de las columnas 2 y 3 para comprobar que se obtiene la misma fórmula.

Una definición alternativa del determinante de una matriz $A = (a_{i,j})$ es la siguiente

$$\det(A) = \sum_{\sigma \in S_n} sign(\sigma)\, a_{1,\sigma(1)} a_{2,\sigma(2)} ... a_{n,\sigma(n)}$$

donde $S_n$ es el grupo de todas las *permutaciones* (funciones biyectivas) del conjunto $\{1, ..., n\}$, y $sign(\sigma) \in \{1, -1\}$ se conoce como el *signo* de la permutación. En general, el signo de una *transposición* (una permutación que intercambia dos números y fija todos los demás) es $-1$, y el signo de la composición de $k$ transposiciones es $(-1)^k$.

**Teorema 7.5 (Propiedades del Determinante).** Sea $n \geq 1$.

**(1)** $\det(I_n) = 1$.

**(2)** Si $A = (a_{i,j}) \in T_{n\times n}(F)$ es triangular superior, triangular inferior, o diagonal, entonces $\det(A) = a_{1,1} a_{2,2} ... a_{n,n}$.

**Demostración.** Demostraremos cada punto.

**(1)** Demostraremos este resultado por inducción sobre $n \in \mathbb{N}$. Si $n = 2$, entonces

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

así que $\det(I_2) = 1$. Supongamos que $\det(I_n) = 1$. Entonces, por definición de determinante,

$$\det(I_{n+1}) = 1 \det(I_n) = 1.$$

**(2)** Nuevamente usaremos inducción sobre $n \in \mathbb{N}$. Si $A \in T_{2,2}(F)$, entonces

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} \\ 0 & a_{2,2} \end{pmatrix},$$

así que

$$\det(A) = a_{1,1} a_{2,2}.$$

Supongamos que $\det(D) = d_{1,1} ... d_{n,n}$ para toda $D = (d_{i,j}) \in T_{n\times n}(F)$. Sea $A \in T_{(n+1)\times(n+1)}(F)$. Por definición,

$$\det(A) = a_{1,1} \det(A_{1,1}),$$

donde $A_{1,1} \in T_{n\times n}(F)$. Por hipótesis de inducción,

$$\det(A_{1,1}) = a_{2,2} ... a_{n+1,n+1}.$$

Esto demuestra que $\det(A) = a_{1,1} a_{2,2} ... a_{n+1,n+1}$. La demostración es similar para las matrices triangulares inferiores.

□

Los siguientes resultados son importantes, y sus demostraciones pueden consultarse en [7].

**Teorema 7.6 (determinante de la transpuesta).** Para cualquier $A \in M_{n\times n}(F)$, se cumple que $\det\left(A^T\right) = \det(A)$.

**Teorema 7.7 (determinante del producto).** Sean $A, D \in M_{n\times n}(F)$. Entonces,

$$\det(AD) = \det(A)\det(D).$$

**Ejemplo 7.8.** Consideremos el caso de las matrices de $2 \times 2$. Sean

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \text{ y } D = \begin{pmatrix} e & f \\ g & h \end{pmatrix}.$$

Entonces

$$\det(A) = ad - bc \text{ y } \det(D) = eh - fg.$$

El producto de los determinantes es:

$$\det(A)\det(D) = (ad - bc)(eh - fg).$$

Ahora, el producto de las matrices es

$$AD = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} ae + bg & af + bh \\ ce + dg & cf + dh \end{pmatrix}.$$

Por lo tanto,

$$\begin{aligned} \det(AD) &= (ae + bg)(cf + dh) - (af + bh)(ce + dg) \\ &= aecf + aedh + bgcf + bgdh - afce - afdg - bhce - bhdg \\ &= aedh + bgcf - bhce - afdg \\ &= ad(eh - fg) + cb(fg - eh) \\ &= (ad - cb)(eh - fg) \\ &= \det(A)\det(D). \end{aligned}$$

**Teorema 7.9 (determinante de una matriz invertible).** Una matriz $A \in M_{n\times n}(F)$ es invertible si y sólo si $\det(A) \neq 0$.

**Demostración.** Demostraremos cada implicación.

($\Longrightarrow$) Supongamos que $A$ es invertible. Entonces existe una matriz $A^{-1} \in M_{n\times n}(F)$ tal que $AA^{-1} = I_n$. Por el Teorema 7.7,

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

Si $\det(A) = 0$, es imposible que $\det(A)\det\left(A^{-1}\right) = 1$. Por lo tanto, $\det(A) \neq 0$

($\Longleftarrow$) Ver Teorema 3.25 en [7].

□

**Corolario 7.10.** Si $A \in M_{n\times n}(F)$ es invertible, entonces $\det\left(A^{-1}\right) = \frac{1}{\det(A)}$.

**Demostración.** Este resultado se deduce de la igualdad $\det(A)\det\left(A^{-1}\right) = 1$. □

**Corolario 7.11.** Sea $A \in M_{n\times n}(F)$. Entonces $A$ es invertible si y sólo si $A^T$ es invertible.

**Demostración.** Ejercicio 7.39. □

**Corolario 7.12.** Sea $F$ un campo y $n \geq 1$. Entonces,

$$GL_n(F) := \{A \in M_{n\times n}(F) : \det(A) \neq 0\}.$$

**Observación 7.13.** Por el Corolario 7.10, una matriz con determinante 1 tiene una inversa con determinante 1. Además, el producto de dos matrices con determinante 1 es una matriz con determinante 1. Por lo tanto, el conjunto

$$SL_n(F) := \{A \in M_{n\times n}(F) : \det(A) = 1\},$$

es un subgrupo de $GL_n(F)$, al cual se le conoce como el *grupo especial lineal de grado $n$ sobre $F$*.

## 7.2. Matrices elementales

Sea $A \in M_{n\times n}(F)$ una matriz cuadrada. Denotamos las filas de $A$ como $R_i$, $i = 1, ..., n$.

**Definición 7.14 (matrices elementales).** Los tres tipos de operaciones elementales de fila son:

(OE1) *Multiplicar por escalar*: consiste en reemplazar la fila $R_i$ por $\alpha R_i$, donde $\alpha \in F$, $\alpha \neq 0$. Esta operación corresponde a multiplicar $A$ por la izquierda por la matriz

$$E_i(\alpha) = \begin{pmatrix} 1 & & & & & \\ & 1 & & & & \\ & & \ddots & & & \\ & & & \alpha & & \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix},$$

donde todos los elementos fuera de la diagonal son 0.

(OE2) *Intercambiar*: consiste en intercambiar las filas $R_i$ y $R_j$. Esta operación corresponde a multiplicar $A$ por la izquierda por la matriz

$$E_{i,j} = \begin{pmatrix} 1 & & & & & & & \\ & 1 & & & & & & \\ & & \ddots & & & & & \\ & & & 0 & & 1 & & \\ & & & & \ddots & & & \\ & & & 1 & & 0 & & \\ & & & & & & \ddots & \\ & & & & & & & 1 \end{pmatrix},$$

donde todos los elementos fuera de la diagonal son 0, excepto $a_{i,j} = a_{j,i} = 1$.

(OE3) *Sumar filas*: consiste en reemplazar la fila $R_i$ por $R_i + \alpha R_j$, donde $\alpha \in F$. Esta operación corresponde a multiplicar $A$ por la izquierda por la matriz

$$S_{i,j}(\alpha) = \begin{pmatrix} 1 & & & & & \\ & 1 & & & & \\ & & \ddots & & & \\ & & & 1 & & \\ & & & & \ddots & \\ & \alpha & & & & \\ & & & & & 1 \end{pmatrix},$$

donde todos los elementos fuera de la diagonal son 0 excepto $a_{i,j} = \alpha$.

Las matrices $E_i(\alpha)$, $E_{i,j}$ y $S_{i,j}$ definidas anteriormente se llaman *matrices elementales*.

Con el objetivo de familiarizar al lector con la definición anterior, en el Ejercicio 7.36 pedimos caluclar algunas matrices elementales y multiplicarlas por una matriz arbitraria.

**Ejemplo 7.15.** Consideremos la matriz

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

Multiplicar la primera fila de $A$ por 2 es equivalente a multiplicar por la matriz elemental $E_1(2)$:

$$E_1(2)\, A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 0 & -2 \\ -1 & 2 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 4 & 2 \\ 2 & 0 & -2 \\ -1 & 2 & 0 \end{pmatrix}.$$

Intercambiar la segunda y tercera fila de $A$ es equivalente a multiplicar por la matriz elemental $E_{2,3}$:

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

Reemplazar la segunda fila por la suma de las dos primeras filas de $A$ es equivalente a multiplicar por la matriz elemental $S_{2,1}(2)$:

$$S_{2,1}(2)\, A = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 0 & -2 \\ -1 & 2 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 1 \\ 4 & 4 & 0 \\ -1 & 2 & 0 \end{pmatrix}.$$

**Teorema 7.16 (inversas de matrices elementales).** Las matrices elementales son invertibles, y sus inversas son las siguientes:

1. $E_i(\alpha)^{-1} = E_i\left(\frac{1}{\alpha}\right)$,
2. $(E_{i,j})^{-1} = E_{j,i}$,
3. $(S_{i,j}(\alpha))^{-1} = S_{i,j}(-\alpha)$.

**Demostración.** Ejercicio.7.37. □

**Teorema 7.17 (determinantes de matrices elementales).** Los determinantes de las matrices elementales son los siguientes:

1. $\det(E_r(\alpha)) = \alpha$,
2. $\det(E_{r,s}) = -1$,
3. $\det(S_{r,s}) = 1$.

**Demostración.** Ejercicio 7.43 □

**Observación 7.18.** Por el teorema anterior, intercambiar dos filas de una matriz cambian el signo del determinante. Como el determinante de una matriz es igual al de su transpuesta, lo mismo sucede si intercambiamos dos columnas. Por lo tanto, la Definición 7.1 de determinante puede aplicarse a cualquier fila y columna de la matriz (no necesariamente a la primera columna).

**Definición 7.19 (forma escalonada reducida).** Decimos que una matriz cuadrada es *escalonada reducida* si

1. Todas las filas cero están en el fondo de la matriz.
2. En cualquier fila distinta de cero, el primer elemento distinto de cero es 1. Este elemento se llama el *pivote* de la fila.

3. Para cualesquiera dos filas consecutivas, el pivote de la fila inferior está a la derecha del pivote de la fila superior.

4. Si una columna contiene un pivote, entonces todos los demás elementos de la columna son 0.

**Ejemplo 7.20.** Las siguientes matrices cuadradas son escalonadas reducidas:

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

**Definición 7.21 (equivalencia por filas).** Sean $A, B \in M_{n\times n}(F)$. Decimos que $A$ y $B$ son *equivalentes por filas* si es posible obtener $B$ a partir de $A$ haciendo una serie de operaciones elementales de fila. En otras palabras, $A$ y $B$ son equivalentes por filas si

$$B = E_1 E_2 ... E_r A$$

para algunas matrices elementales $E_i$.

**Teorema 7.22 (equivalencia por filas).** La relación de equivalencia por filas entre matrices es una relación de equivalencia.

**Demostración.** Ejercicio 7.38. □

**Teorema 7.23 (forma escalonada reducida).** Sea $A \in M_{n\times n}(F)$. Entonces $A$ es equivalente por filas a una única matriz escalonada reducida.

**Ejemplo 7.24.** Sea

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

Observemos que

$$B = E_2\left(\frac{1}{2}\right) S_{3,2}(-2)\, A$$

donde $B$ es la siguiente matriz escalonada reducida:

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

Entonces, $A$ y $B$ son equivalentes por filas.

El proceso de multiplicar a $A$ por matrices elementales para obtener su forma escalonada reducida se conoce como el *método de Gauss-Jordan* (o *eliminación Gaussiana*). Si $A$ es cualquier matriz, encontrar la matriz escalonada reducida de $A$ equivale a resolver el sistema de ecuaciones lineal homogéneo $Ax = \mathbf{0}$, donde $x = (x_1, \ldots, x_n)^T \in F^n$ es una variable, y ambos $x$ y $\mathbf{0}$ son vistos como *vectores columna* (matrices de $n \times 1$).

**Teorema 7.25 (matrices invertibles y sistemas de ecuaciones).** Sea $A \in M_{n\times n}(F)$. Las siguientes afirmaciones son equivalentes:

**(1)** $A$ es invertible.

**(2)** El sistema de ecuaciones $Ax = b$ tiene solución única para toda $b \in F^n$.

**(3)** El sistema de ecuaciones $Ax = \mathbf{0}$ tiene solución única.

**(4)** $A$ es equivalente por filas a la matriz identidad.

**Demostración.** Demostraremos cada implicación.

**( 1 ) $\Rightarrow$ ( 2)** Supongamos que $A$ es invertible. Entonces, $a := A^{-1}b \in F^n$ es una solución del sistema ya que $A(A^{-1}b) = b$. Si $c \in F^n$ es otra solución, entonces, despejando de $Ac = b$, obtenemos que $c = A^{-1}b = a$. Por lo tanto, el sistema tiene solución única.

**( 2 ) $\Rightarrow$ ( 3)** El punto (3) es un caso particular del punto (2), así que obviamente (2) implica (3).

**( 3) $\Rightarrow$ ( 4 )** Supongamos que el sistema $Ax = \mathbf{0}$ tiene solución única. Por reducción al absurdo, supongamos que $A$ no es equivalente por filas a la matriz identidad. Por el Teorema 7.23, $A$ es equivalente por filas a una única matriz escalonada reducida $B \neq I_n$; es decir,

$$B = E_1 ..., E_r A,$$

donde $E_i$ son matrices elementales. Multiplicando el sistema $Ax = \mathbf{0}$ por estas matrices elementales obtenemos que

$$\begin{aligned} E_1 ... E_r Ax &= E_1 ... E_r \mathbf{0} \\ Bx &= \mathbf{0}, \end{aligned}$$

ya que el producto de cualquier matriz por $\mathbf{0}$ es igual a $\mathbf{0}$. Como $B$ es distinta de la identidad, entonces el último renglón de $B$ debe tener todas las entradas iguales a 0. Esto implica que el sistema $Bx = \mathbf{0}$ tiene soluciones infinitas ya que hay al menos una variable $x_n$ libre. Como todas las soluciones de $Bx = \mathbf{0}$ también son soluciones de $Ax = \mathbf{0}$, esto contradice que $Ax = \mathbf{0}$ tenga solución única. Por lo tanto, $A$ es equivalente por filas a la matriz identidad.

**( 4 ) $\Rightarrow$ ( 1 )** Supongamos que $A$ es equivalente por filas a la matriz identidad. Entonces

$$I_n = E_1 ... E_r A$$

para algunas matrices elementales $E_i$. Por definición, esto significa que $A$ es una matriz invertible, donde $A^{-1} = E_1 ... E_r$.

□

Supongamos que $A \in M_{n\times n}(F)$ es una matriz inverible. En la demostración del teorema anterior vimos que $A^{-1} = E_1...E_r$, donde $E_i$ son las matrices elementales necesarias para escribir a $A$ en su forma escalonada reducida. Por lo tanto, un método para encontrar la inversa de $A$ consiste en escribir la matriz ampliada

$$(A \mid I_n)$$

y aplicar el método de Gauss-Jordan para encontrar

$$\left(I_n \mid A^{-1}\right).$$

Claramente, si no es posible escribir la matriz identidad del lado izquierdo de esta matriz ampliada, esto significa que la matriz $A$ no es invertible (Teorema 7.25).

## 7.3. Teorema Fundamental de Matrices Invertibles

Si $A \in M_{n\times n}(F)$ definimos el kernel de $A$ como el conjunto

$$\ker(A) := \left\{v \in F^n : Av^T = \mathbf{0}\right\},$$

y el rango de $A$ como

$$A(F^n) := \{Av : v \in F^n\}.$$

Además, definimos

$$\begin{aligned} rk(A) &:= \dim(A(F^n)), \\ nul(A) &:= \dim(\ker(A)). \end{aligned}$$

**Teorema 7.26 (teorema fundamental de matrices invertibles).** Sea $A \in M_{n\times n}(F)$. Las siguientes afirmaciones son equivalentes:

**(1)** $A$ es invertible.

**(2)** $\det(A) \neq 0$.

**(3)** $rk(A) = n$.

**(4)** $nul(A) = 0$.

**(5)** Las columnas de $A$ forman una base de $F^n$.

**(6)** Las filas de $A$ forman una base de $F^n$.

**(7)** El sistema de ecuaciones $Ax = b$ tiene solución única para toda $b \in F^n$.

**(8)** El sistema de ecuaciones $Ax = \mathbf{0}$ tiene solución única.

**(9)** $A$ es equivalente por filas a la matriz identidad.

**(10)** El endomorfismo definido por $A$ es un automorfismo.

**Demostración.** La equivalencia entre (1) y (2) quedó demostrada en el Teorema 7.9, mientras que las equivalencias entre (1), (7), (8) y (9) quedaron demostradas en el Teorema 7.25. Por el Teorema 6.53, $A$ es invertible si y sólo si la transformación lineal $\phi : F^n \to F^n$, $\phi(v) = Av^T$, es un automorfismo. Por lo tanto, $nul(\phi) = nul(A) = 0$; por el Teorema 6.5, esto es equivalente a $rk(\phi) = rk(A) = n$. En el Teorema 5.47 demostramos que $\phi$ es un automorfismo si y sólo si las columnas de $A$ son una base de $F^n$.

Finalmente, sólo queda demostrar que las filas de $A$ son una base de $F^n$ si y sólo si sus columnas son una base de $F^n$. Como $\det(A) = \det\left(A^T\right)$, $A$ es invertible si y sólo si $A^T$ es invertible. Entonces, las filas de $A$ son una base de $F^n$ si y sólo si las filas de $A^T$ son una base de $F^n$. El teorema queda demostrado porque las filas de $A^T$ son las columnas de $A$. □

**Ejemplo 7.27.** Demostrar que

$$B = \{(-1,1,1),(1,-1,1),(1,1,-1)\}$$

es una base de $\mathbb{R}^3$ es equivalente a demostrar que la matriz

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

tiene determinante distinto de cero. Calculamos que

$$\det(A) = 4 \neq 0,$$

y por lo tanto, $B$ es una base de $\mathbb{R}^3$.

**Ejemplo 7.28.** Consideremos la transformación lineal $\phi : \mathbb{R}^2 \to \mathbb{R}^2$ definida como

$$\phi(x_1, x_2) = (\alpha x_1 + (2-\alpha)x_2, x_1 + \alpha x_2)$$

donde $\alpha \in \mathbb{R}$. Encontraremos todos los valores de $\alpha$ tales que $\phi$ es un automorfismo. Sabemos que $\phi$ es un automorfismo si y sólo si la matriz

$$M_\phi = \begin{pmatrix} \alpha & (2-\alpha) \\ 1 & \alpha \end{pmatrix}$$

es invertible, lo cual se cumple si y sólo si

$$\det(M_\phi) = \alpha^2 + \alpha - 2 \neq 0.$$

Como $\alpha^2 + \alpha - 2 = 0$ precisamente cuando $\alpha = 1$ o $\alpha = -2$ , entonces la función $\phi$ es un automorfismo siempre que $\alpha \in \mathbb{R} \setminus \{1, -2\}$.

**Ejemplo 7.29.** Sea $\alpha \in \mathbb{R}$, y consideremos el conjunto

$$B = \{(1,0,0,0), (1,\alpha,1-\alpha,0), (0,0,\alpha,1), (0,0,0,1)\}.$$

Encontraremos todos los valores de $\alpha$ tales que $B$ sea una base de $\mathbb{R}^4$. Consideremos la matriz cuyas columnas son los vectores de $B$:

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

Entonces,

$$\det(A) = \det\left(\begin{pmatrix} \alpha & 0 & 0 \\ 1-\alpha & \alpha & 0 \\ 0 & 1 & 1 \end{pmatrix}\right) = \alpha^2$$

Por lo tanto, $\det(A) = 0$ si y sólo si $\alpha = 0$. Esto implica que $B$ es una base de $\mathbb{R}^4$ si y sólo si $\alpha \neq 0$.

***Palabras clave:*** *determinante, matrices elementales, forma escalonada reducida, teorema fundamental de matrices invertibles.*

## 7.4. Ejercicios de Teoría de Matrices

**Ejercicio 7.30.** Considera las siguientes matrices con entradas en $\mathbb{R}$:

$$A = \begin{pmatrix} 0 & 1 & 4 \\ 3 & 5 & 0 \\ 1 & 2 & 4 \end{pmatrix}, B = \begin{pmatrix} 1 & 3 & 0 \\ 1 & 2 & 0 \end{pmatrix}, D = \begin{pmatrix} 4 & 1 \\ 3 & 0 \\ 1 & 2 \end{pmatrix}.$$

1. Calcula los productos $AD$, $AA$ y $BA$. ¿Es posible calcular los productos $DA$ y $AB$?
2. Si $v = (2, 4, 1) \in \mathbb{R}^3$,y $w = (1, 1) \in \mathbb{R}^2$, calcula $Av$, $Dw$, $Bv$ y $ADw$. ¿Es posible calcular $Dv$ y $Bw$?
3. Si $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$ y $e_3 = (0, 0, 1)$, calcula $Ae_1 + Ae_2 + Ae_3$.
4. Escribe la matriz identidad $I_3$ y calcula $AI_3$, $BI_3$ y $I_3D$. ¿Es posible calcular $I_3B$ y $DI_3$?

**Ejercicio 7.31.** Sean $A, B \in M_{3\times 3}(\mathbb{Z}_3)$ matrices definidas como:

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

Calcula las siguientes matrices:

1. $A + D$ y $A + 2D$.
2. $A^T$, $D^T$ y $(A + 2D)^T$.
3. $AD$ y $(AD)^T$.

**Ejercicio 7.32.** Demuestra el Teorema 6.20.

**Ejercicio 7.33.** Considera los conjuntos $T_{n\times n}(F)$ de matrices triangulares superiores y $D_{n\times n}(F)$ de matrices diagonales. Demuestra que $T_{n\times n}(F)$ y $D_{n\times n}(F)$ son subespacios de $M_{n\times n}(F)$. Encuentra bases para estos subespacios y escribe sus respectivas dimensiones.

**Ejercicio 7.34.** Encuentra el determinante y, en caso de que exista, la inversa, de cada una de las siguientes matrices de $2 \times 2$.

1. $\begin{pmatrix} 2 & -1 \\ 1 & \frac{1}{2} \end{pmatrix}$ en $M_{2\times 2}(\mathbb{R})$
2. $\begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}$ en $M_{2\times 2}(\mathbb{R})$.
3. $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ en $M_{2\times 2}(\mathbb{R})$.

4. $\begin{pmatrix} 2 & 4 \\ 4 & 3 \end{pmatrix}$ en $M_{2\times 2}(\mathbb{Z}_5)$.

5. $\begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$ en $M_{2\times 2}(\mathbb{Z}_3)$.

**Ejercicio 7.35.** Encuentra todos los valores $\alpha \in \mathbb{R}$ que hagan que la matriz $\begin{pmatrix} \alpha & \alpha - 1 \\ 3 & \alpha \end{pmatrix}$ sea invertible.

**Ejercicio 7.36.** Consideremos una matriz arbitraria $A = (a_{i,j}) \in M_{3\times 3}(\mathbb{R})$. Usando la notación de la Definición 7.14, escribe explícitamente las siguientes matrices elementales

$$E_2(5), \quad E_{1,2}, \quad \text{y} \quad S_{2,3}\left(\frac{1}{2}\right),$$

y calcula los siguientes productos

$$E_2(5)A, \quad E_{1,2}A, \quad \text{y} \quad S_{2,3}\left(\frac{1}{2}\right)A.$$

**Ejercicio 7.37.** Demuestra el Teorema 7.16.

**Ejercicio 7.38.** Demuestra que la equivalencia por filas es una relación de equivalencia.

**Ejercicio 7.39.** Sea $A \in M_{n\times n}(F)$. Usando el Teorema 7.9, demuestra que $A$ es invertible si y sólo si $A^T$ es invertible.

**Ejercicio 7.40.** Calcula el determinante de las siguientes matrices y determina si son matrices invertibles. En caso de ser invertibles, encuentra la matriz inversa.

1. $$A_1 = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ -1 & 0 & 1 \end{pmatrix}.$$

2. $$A_2 = \begin{pmatrix} 2 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 70 & -3 \end{pmatrix}.$$

3. $$A_3 = \begin{pmatrix} 1 & 3 & 5 \\ 3 & -3 & 3 \\ 0 & -1 & -1 \end{pmatrix}$$

4. $$A_4 = \begin{pmatrix} 1 & 0 & 2 & 3 \\ -2 & -1 & 2 & 0 \\ 1 & 0 & 3 & -4 \\ 0 & 1 & 0 & -1 \end{pmatrix}.$$

5.

$$A_5 = \begin{pmatrix} 1 & 7 & -6 & 0 \\ 0 & -2 & 2 & 3 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

**Ejercicio 7.41.** Sea

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

Encuentra todos los valores $\alpha \in \mathbb{R}$ tales que $A$ es una matriz invertible.

**Ejercicio 7.42.** Demuestra que si $A = (a_{i,j})$ es una matriz triangular inferior, entonces $\det(A) = a_{1,1}a_{2,2}...a_{n,n}$.

**Ejercicio 7.43.** Encuentra los determinantes de las matrices elementales, justificando tu respuesta.

# 8

# Autovalores y autovectores

## 8.1. Autovalores y autovectores de una matriz

En este capítulo trataremos el concepto de similitud entre matrices en $M_{n\times n}(F)$. Explicaremos cómo la similitud define una relación de equivalencia y como las clases de equivalencia se componen de matrices que definen a un mismo endomorfismo. Además, encontraremos las propiedades que no cambian entre dos matrices similares. Resumiendo, determinaremos todas las clases de equivalencia definidas bajo similitud e identificaremos los invariantes necesarios y suficientes que las determinan.

**Definición 8.1 (similitud de matrices).** Sean $A_1, A_2 \in M_{n\times n}(F)$ dos matrices. Decimos que $A_1$ es *similar*, o *conjugada*, a $A_2$ (y escribimos $A_1 \sim A_2$) si existe una matriz invertible $P \in M_{n\times n}(F)$ tal que

$$A_1 = P^{-1}A_2P.$$

**Lema 8.2 (similitud de matrices).** Si $A_1, A_2 \in M_{n\times n}(F)$ matrices similares, entonces existe un endomorfismo $\tau : F^n \to F^n$ y dos bases $B_1$ y $B_2$ de $F^n$ tales que

$$A_1 = [\tau]_{B_1}^{B_1} \text{ y } A_2 = [\tau]_{B_2}^{B_2}.$$

**Demostración.** Como $A_1 \sim A_2$, existe una matriz invertible $P \in M_{n\times n}(F)$ tal que $A_1 = P^{-1}A_2P$. Sea $B_1$ la base canónica de $F^n$ y sea $B_2$ el conjunto de vectores columna de la matriz $P$. Por el Teorema Fundamental de Matrices Invertibles, $B_2$ también es una base de $F^n$. Observemos que $P$ es precisamente la matriz del cambio de base de $B_1$ a $B_2$. Sea $\tau : F^n \to F^n$ el endomorfismo tal que $A_1 = [\tau]_{B_1}^{B_1}$. Ahora, despejando de $A_1 = P^{-1}A_2P$ y usando el Teorema del Cambio de Base, obtenemos

$$A_2 = PA_1P^{-1} = P[\tau]_{B_1}^{B_1}P^{-1} = [\tau]_{B_2}^{B_2}.$$

□

**Observación 8.3.** El resultado anterior, combinado con el Teorema del Cambio de Base, puede interpretarse de la siguiente manera: dos matrices son similares

si y solo si representan al mismo endomorfismo respecto a dos bases de $F^n$. Por tal motivo, a partir de ahora nos enfocaremos en estudiar matrices ya que los resultados obtenidos son fácilmente traducidos al lenguaje de endomorfismos.

**Lema 8.4 (similitud como relación de equivalencia).** La similitud de matrices es una relación de equivalencia.

**Demostración.** Ejercicio 8.37.
□

Ahora quisiéramos determinar representantes para todas las clases de equivalencia de matrices definidas por la similitud. Con este propósito identificaremos aquellas propiedades comunes entre dos matrices similares, para ello introduciremos primero el siguiente concepto.

**Definición 8.5 (autovalores y autovectores de una matriz).** Sea $A \in M_{n\times n}(F)$ una matriz y $\lambda \in F$ un escalar arbitrario. Decimos que $\lambda$ es un autovalor de $A$ si existe un vector $\mathbf{v} \in F^n$, $\mathbf{v} \neq \mathbf{0}$, tal que

$$A\mathbf{v} = \lambda\mathbf{v}.$$

En este caso diremos que el vector $\mathbf{v}$ es un $\lambda$-*autovector* de la matriz $A$.

El siguiente resultado determina un método para encontrar los autovalores asociados a una matriz.

**Proposición 8.6 (autovalores).** Sea $A \in M_{n\times n}(F)$ una matriz. Un escalar $\lambda \in F$ es un autovalor de $A$ si y solo si $\det(\lambda I_n - A) = 0$.

**Demostración.**

(⇒) Supongamos que $\lambda$ es un autovalor asociado a la matriz $A$. Entonces por definición, existe un vector $\mathbf{v} \in F^n$ no nulo tal que $A\mathbf{v} = \lambda\mathbf{v}$. Es decir, $\lambda I_n\mathbf{v} - A\mathbf{v} = (\lambda I_n - A)\mathbf{v} = \mathbf{0}$. En consecuencia, $\mathbf{v} \in \ker(A - \lambda I_n)$, lo que implica que $A$ no es invertible. Por el Teorema Fundamental de Matrices Invertibles, $\det(\lambda I_n - A) = 0$.

(⇐) Supongamos que $\lambda \in F$ es un escalar tal que $\det(\lambda I_n - A) = 0$. Entonces, la matriz $\lambda I_n - A$ es una matriz no invertible, y $\ker(\lambda I_n - A) \neq \{\mathbf{0}\}$; esto es, existe un vector $\mathbf{v} \in F^n$ no nulo tal que $(\lambda I_n - A)\mathbf{v} = 0$. Esto implica que $\lambda$ y $\mathbf{v}$ cumplen $\lambda I_n\mathbf{v} = A\mathbf{v}$. Por lo tanto, $\lambda$ es autovalor de la matriz $A$ y $\mathbf{v}$ es un $\lambda$-autovector de $A$.

□

Como hemos visto en el resultado anterior, para determinar los autovalores de la matriz $A$, es necesario encontrar los escalares $\lambda \in k \setminus \{0\}$ para los cuales $\det(\lambda I_n - A) = 0$. Así, en lugar de tomar $\lambda$, consideramos una indeterminada $x$. De este modo $\det(xI_n - A)$ esa ahora un polinomio que denotamos por $p_A(x)$.

**Definición 8.7 (Polinomio característico).** Sea $A \in M_{n\times n}(F)$ una matriz. El polinomio definido como $p_A(x) := \det(xI_n - A)$ es llamado el *polinomio característico de* $A$.

Afirmamos que, si la matriz $A$ es una matriz cuadrada $n \times n$, entonces $p_A(x)$ es un polinomio de grado $n$. Para probar esta afirmación, lo hacemos por inducción sobre $n$. Si $n = 1$, entonces $A = (a_{1,1})$ y $\det(xI_n - A) = x - a_{1,1}$, lo cual muestra nuestra afirmación. Suponga cierto para el caso $n - 1$ y probemos el caso $n + 1$.

Así, si $A$ es una matriz $n + 1 \times n + 1$, entonces

$$\begin{aligned} \det(xI_{n+1} - A) = \quad & (x - a_{1,1}) \det(xI_{n+1} - A)_{1,1} + \\ & \textstyle\sum_{j=2}^{n+1} (-1)^j a_{1,j} \det(xI_{n+1} - A)_{1,j} \end{aligned} \tag{8.1}$$

donde $(xI_{n+1} - A)_{1,j}$ denota a la matriz menor de $(xI_{n+1} - A)$. Notemos que, por hipótesis de inducción $\det(xI_n - A)_{1,1}$ es un polinomio de grado $n$ y $\det(xI_n - A)_{1,j}$ es un polinomio grado $n-2$, para todo $j = 2, 3, \ldots n+1$. De esta manera, $p_A(x)$ es un polinomio grado $n + 1$, obteniendo así que nuestra afirmación es correcta.

Recordemos que la *traza* de una matriz $A = (a_{ij})$ se define como la suma de las entradas en la diagonal: $\operatorname{tr}(A) := \sum_{i=1}^{n} a_{ii}$. De la ecuación (8.1) podemos notar que $p_A(x)$ es un polinomio mónico. Mas aún, haciendo algunos cálculos extra podemos ver que si $A$ es una matriz $n \times n$, entonces

$$p_A(x) = x^n - \operatorname{tr}(A)\ x^{n-1} + \cdots + (-1)^n \det(A)$$

La demostración formal de este hecho la haremos para el caso $n = 2$ dejando como ejercicio para el lector interesado verificar el caso general.

**Lema 8.8 (polinomio característico, caso $2 \times 2$).** Sea $A = (a_{i,j}) \in M_{2\times 2}(F)$. Entonces,

$$p_A(x) = x^2 - \operatorname{tr}(A)x + \det(A).$$

**Demostración.** Calculamos el polinomio característico de $A$:

$$\begin{aligned} \det(xI_2 - A) &= \det \begin{pmatrix} x - a_{1,1} & -a_{1,2} \\ -a_{2,1} & x - a_{2,2} \end{pmatrix} \\ &= (x - a_{1,1})(x - a_{2,2}) - a_{1,2}a_{2,1}, \\ &= x^2 - (a_{1,1} + a_{2,2})x + (a_{1,1}a_{2,2} - a_{1,2}a_{2,1}) \\ &= x^2 - \operatorname{tr}(A)x + \det(A). \end{aligned}$$

□

Por la Proposición 8.6 sabemos que $\lambda \in F$ es un autovalor de $A$ si, y solo si, $\lambda$ es raíz del polinomio característico $p_A(x)$ (es decir, $p_A(\lambda) = 0$). De lo anterior y con base en el Teorema Fundamental del Álgebra podemos concluir que una matriz $A \in M_{n\times n}(F)$ tiene a lo mas $n$ autovalores distintos.

**Ejemplo 8.9.** Dada la matriz

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

verificaremos que el polinomio característico de $A$ tiene la forma

$$p_A(x) = x^2 - \mathrm{tr}(A)x + \det(A).$$

El primer paso es calcular el polinomio característico de $A$:

$$\begin{aligned} \det(xI_2 - A) &= \begin{vmatrix} x-3 & -2 \\ -1 & x-2 \end{vmatrix}, \\ &= (x-3)(x-2) - 2, \\ &= x^2 - 5x + 4. \end{aligned}$$

De esto podemos verificar que $\mathrm{tr}(A) = 5$ y $\det(A) = 4$, cumpliendo así la primer afirmación. Ahora, para determinar los autovalores, calculamos las raíces del polinomio característico $p_A(x) = x^2 - 5x + 4 = (x-4)(x-1)$. Por lo tanto, los autovalores de $A$ son: $\lambda_1 = 4$ y $\lambda_2 = 1$.

**Ejemplo 8.10.** Determinaremos los autovalores asociados a la matriz

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

y comprobaremos que el negativo de la traza de $B$ corresponde al coeficiente del término lineal en el polinomio característico. Siguiendo los mismos pasos que en el ejemplo anterior se tiene que

$$\begin{aligned} \det(xI_3 - B) &= \begin{vmatrix} x-2 & -1 & -1 \\ -2 & x-3 & -2 \\ -3 & -3 & x-4 \end{vmatrix} \\ &= x^3 - 9x^2 + 15x - 7. \end{aligned}$$

Por lo tanto, los autovalores de $B$ son $\lambda_1 = 7,\ \lambda_2 = \lambda_3 = 1$. Por último, vemos que efectivamente $9 = -\mathrm{tr}(B)$ y $\det(B) = 7$.

Por otro lado, no todas las matrices tienen autovalores asociados. Un ejemplo de ello lo encontramos a continuación.

**Ejemplo 8.11.** Sea $A \in M_{2\times 2}(\mathbb{R})$ una matriz definida como

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

En este caso, $p_A(x) = x^2 + 1$ y por lo tanto el polinomio no tiene raíz en $\mathbb{R}$.

Retomando nuestro problema principal, queremos encontrar las propiedades algebraicas que distinguen a dos matrices similares. Mas aún, determinar las propiedades necesarias y suficientes que caracterizan a cada clase de equivalencia dada por la similitud. En este sentido, se tiene el siguiente resultado.

**Proposición 8.12 (polinomios característicos de matrices similares).** Sean $A_1, A_2 \in M_{n\times n}(F)$ dos matrices. Si $A_1 \sim A_2$, entonces

$$p_{A_1}(x) = p_{A_2}(x).$$

**Demostración.** Si $A_1 \sim A_2$, entonces existe una matriz invertible $P \in M_{n\times n}(F)$ tal que $A_1 = P^{-1}A_2P$. Substituyendo la relación anterior y aplicando propiedades de los determinantes, obtenemos

$$\begin{aligned}
\det(xI_n - A_1) &= \det(xI_n - P^{-1}A_2P) \\
&= \det(P^{-1}(xI_n - A_2)P), \\
&= \det(P^{-1})\ \det(xI_n - A_2)\ \det(P), \\
&= \det(xI_n - A_2)\ \frac{\det(P)}{\det(P)}, \\
&= \det(xI_n - A_2).
\end{aligned}$$

De ello concluimos que $p_{A_1}(x) = p_{A_2}(x)$. □

**Corolario 8.13.** Dos matrices similares tienen exactamente los mismos autovalores.

El polinomio característico de una matriz no determina completamente la relación de similitud. A continuación presentamos un ejemplo de dos matrices $A_1, A_2 \in M_{n\times n}(F)$ tales que $A_1 \not\sim A_2$, pero $p_{A_1}(x) = p_{A_2}(x)$.

**Ejemplo 8.14.** La única matriz similar a la matriz identidad $I_n$ es ella misma ya que para cualquier matriz invertible $P \in M_{n\times n}(F)$ se tiene

$$P^{-1}I_nP = I_n.$$

Sea $I_2$ la matriz identidad $2 \times 2$ y consideramos

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

En este caso, $I_2 \not\sim A$, pero $p_{I_2}(x) = p_A(x)$.

**Definición 8.15 (autoespacio).** Sea $A \in M_{n\times n}(F)$ una matriz y sea $\lambda$ un autovalor de $A$. El *$\lambda$-autoespacio de $A$* es el conjunto

$$S_\lambda(A) := \{\mathbf{v} \in F^n | A\mathbf{v} = \lambda\mathbf{v}\}.$$

**Observación 8.16.** El $\lambda$-autoespacio de $A$ es igual al conjunto de todos los $\lambda$-autovectores de $A$ unión con el vector cero.

**Definición 8.17 (subespacio invariante).** Sea $A \in M_{n\times n}(F)$ una matriz y sea $T$ un subespacio de $F^n$. Decimos que $T$ es un *subespacio invariante bajo $A$* si para toda $t \in T$ se cumple que $At \in T$.

**Lema 8.18 (autoespacio es subespacio invariante).** Sea $A \in M_{n\times n}(F)$ una matriz y sea $\lambda$ un autovalor de $A$. El $\lambda$-autoespacio $S_\lambda(A)$ es un subespacio de $F^n$ invariante bajo $A$.

**Demostración.** Un cálculo sencillo, el cual se deja como ejercicio para el lector, demuestra que $S_\lambda(A) = \ker(\lambda I_n - A)$ (ver Ejercicio 8.40). Por lo tanto, $S_\lambda(A)$ es un subespacio de $F^n$. Para demostrar que es invariante bajo $A$, sea $\mathbf{v} \in S_\lambda(A)$. Entonces, $A\mathbf{v} = \lambda\mathbf{v} \in S_\lambda(A)$, ya que $S_\lambda(A)$ es cerrado bajo la multiplicación de escalares. Esto demuestra la invarianza bajo $A$. □

**Definición 8.19 (multiplicidades de autovalores).** Sea $A \in M_{n\times n}(F)$ una matriz y sea $\lambda$ un autovalor de $A$.

- La *multiplicidad geométrica* de $\lambda$, denotada por $\gamma_\lambda(A)$, es la dimensión del autoespacio $S_\lambda(A)$.
- La *multiplicidad algebraica* de $\lambda$, denotada por $\mu_\lambda(A)$, es el mayor número natural $k$ tal que $(x-\lambda)^k$ es un factor del polinomio característico $p_A(x)$.

**Observación 8.20.** Sea $\lambda$ un autovalor asociado a la matriz $A \in M_{n\times n}(F)$. Si $\gamma_\lambda(A) = m$, entonces $A$ tiene $m$ $\lambda$-autovectores linealmente independientes. Esto es claro ya que el $\lambda$-autoespacio tiene dimensión $m$.

**Ejemplo 8.21.** Sea $A \in M_{3\times 3}(F)$ una matriz definida como

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

entonces el polinomio característico es: $p_A(x) = x^3 - 5x^2 - x + 5$ y los autovalores son $\lambda_1 = 1,\ \lambda_2 = -1,\ \lambda_3 = 5$. Utilizando el autovalor $\lambda_1 = 1$, tendremos que $S_{\lambda_1}(A) = \ker(I_3 - A)$. Realizando los cálculos correspondientes se verifica que $S_{\lambda_1}(A)$ está dado por:

$$S_{\lambda_1}(A) := \langle(-2,1,0)\rangle = \{(-2t,t,0) : t \in F\}.$$

Por consiguiente, el espacio vectorial $S_{\lambda_1}(A)$ es de dimensión uno y $\gamma_A(1) = 1$. El lector puede verificar de manera análoga que los autovalores $\lambda_2 = -1$ y $\lambda_3 = 5$ tienen la misma multiplicidad geométrica. La multiplicidad algebraica de cada uno de estos autovalores también es uno porque $p_A(x) = (x-1)(x+1)(x-5)$.

**Ejemplo 8.22.** Consideremos el endomorfismo $\phi : \mathbb{R}^2 \to \mathbb{R}^2$ definido como

$$\phi(x_1, x_2) = (x_1 + x_2, -x_2).$$

Entonces,

$$[\phi] = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}.$$

El polinomio característico de $\phi$ es

$$\begin{aligned} p_\phi(x) &= \det\left(\begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} - x\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right) \\ &= \det\begin{pmatrix} 1-x & 1 \\ 0 & -1-x \end{pmatrix} \\ &= (1-x)(-1-x). \end{aligned}$$

Por lo tanto, los autovalores de $\phi$ son $\lambda_1 = 1$ y $\lambda_2 = -1$. Ahora, los autoespacios de $\phi$ son

$$\begin{aligned} S_1(\phi) &= \{(x_1, x_2) : \phi(x_1, x_2) = (x_1, x_2)\} = \langle (1, 0)\rangle, \\ S_{-1}(\phi) &= \{(x_1, x_2) : \phi(x_1, x_2) = -(x_1, x_2)\} = \langle (1, -2)\rangle. \end{aligned}$$

Con esto, obtenemos que

$$\gamma_1(\phi) = \mu_1(\phi) = 1 \text{ y } \gamma_{-1}(\phi) = \mu_{-1}(\phi) = 1$$

**Ejemplo 8.23.** Sea

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

El polinomio característico de $A$ es

$$\begin{aligned} p_A(x) &= \det\left(\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} - \begin{pmatrix} x & 0 & 0 \\ 0 & x & 0 \\ 0 & 0 & x \end{pmatrix}\right) \\ &= \det\begin{pmatrix} 1-x & 0 & 0 \\ 0 & -x & 1 \\ 0 & 1 & -x \end{pmatrix} \\ &= (1-x)\left(x^2-1\right) \\ &= -(x-1)^2(x+1). \end{aligned}$$

Por lo tanto, 1 y $-1$ son los autovalores de $A$. Los autoespacios de $A$ son

$$\begin{aligned} S_1(A) &= \ker(I_3 - A) = \langle (1, 0, 0), (0, 1, 1)\rangle, \\ S_{-1}(A) &= \ker(-I_3 - A) = \langle (0, 1, -1)\rangle. \end{aligned}$$

En consecuencia,

$$\gamma_1(A) = \mu_1(A) = 2 \text{ y } \gamma_{-1}(A) = \mu_{-1}(A) = 1.$$

**Teorema 8.24 (multiplicidades).** Sea $\lambda \in F$ un autovalor de la matriz $A \in M_{n\times n}(F)$. Entonces,

$$1 \leq \gamma_\lambda(A) \leq \mu_\lambda(A) \leq n.$$

**Demostración.** Es claro que ambas multiplicidades deben estar entre 1 y $n$. Sea $\gamma_\lambda(A) = k$. Por definición, el autoespacio $S_\lambda(A)$ tiene dimensión $k$, así que sea $B = b_1, \ldots, b_k$ una base de $S_\lambda(A)$. Extendamos $B$ a una base $C = \{b_1, \ldots, b_n\}$ de $F^n$. Si $A = [\tau]$ entonces la matriz $A' = [\tau]_C^C$ debe tener la forma

$$A' = \begin{pmatrix} \lambda I_k & M \\ \mathbf{0} & N \end{pmatrix},$$

donde $\mathbf{0}$ es una matriz de $(n-k)\times k$, $M$ es una matriz de $k \times (n-k)$ y $N$ es una matriz de $(n-k)\times(n-k)$. Como $A$ y $A'$ son matrices similares (porque representan al mismo endomorfismo en bases distintas), entonces:

$$\begin{aligned} p_A(x) &= p_{A'}(x) \\ &= \det(xI_n - A') \\ &= \det \begin{pmatrix} \lambda(x-\lambda)I_k & M \\ \mathbf{0} & xI_{n-k} - N \end{pmatrix} \\ &= (x-\lambda)^k \det(xI_{n-k} - N). \end{aligned}$$

Esto demuestra que la multiplicidad algebraica $\lambda$ es al menos $k$.
□

**Definición 8.25 (defectivo, simple y semisimple).** Sea $\lambda \in F$ un autovalor de la matriz $A \in M_{n\times n}(F)$.

- Decimos que $\lambda$ es *defectivo*, o *defectuoso*, si $\gamma_\lambda(A) < \mu_\lambda(A)$.
- Decimos que $\lambda$ es *simple* si $\gamma_\lambda(A) = \mu_\lambda(A) = 1$.
- Decimos que $\lambda$ es *semisimple* si $\gamma_\lambda(A) = \mu_\lambda(A) > 1$.

Una matriz $A$ es *semisimple* si todos sus autovalores son simples o semisimples. En otras palabras, la matriz $A$ es semisimple si no tiene autovalores defectivos.

**Ejemplo 8.26.** Sea $A \in M_{3\times 3}(\mathbb{R})$ una matriz definida como

$$A = \begin{pmatrix} 1 & 4 & 3 \\ -2 & 3 & 5 \\ 2 & 2 & 0 \end{pmatrix}.$$

Entonces el polinomio característico es:

$$p_A(x) = x^3 - 4x^2 - 5x = x(x-5)(x+1).$$

Así los autovalores de $A$ son $\lambda_1 = 0$, $\lambda_2 = 5$ y $\lambda_3 = -1$, todos de multiplicidad algebraica 1. Por el Teorema 8.24, las multiplicidades geométricas de los autovalores son todas 1, así que $A$ es semisimple. Los autoespacios de $A$ son:

$$\begin{aligned} S_{\lambda_1}(A) &= \{w \in \mathbb{R}^3 | w = (t, -t, t),\ t \in \mathbb{R}\} = \langle (1, -1, 1) \rangle, \\ S_{\lambda_2}(A) &= \{w \in \mathbb{R}^3 | w = (13t, 7t, 8t),\ t \in \mathbb{R}\} = \langle (13, 7, 8) \rangle, \\ S_{\lambda_3}(A) &= \{w \in \mathbb{R}^3 | w = (t, -2t, 3t),\ t \in \mathbb{R}\} = \langle (1, -2, 3) \rangle. \end{aligned}$$

**Ejemplo 8.27.** Sea

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

El polinomio característico de $A$ es

$$\begin{aligned} p_A(x) &= \det \begin{pmatrix} 2-x & 0 \\ 2 & 2-x \end{pmatrix} \\ &= (2-x)(2-x), \end{aligned}$$

por lo tanto 2 es el único autovalor de $A$ con $\mu_A(2) = 2$. Ahora,

$$\begin{aligned} S_2(A) &= \{(x_1, x_2) : A(x_1, x_2) = 2(x_1, x_2)\} \\ &= \{(x_1, x_2) : (2x_1, 2x_1 + 2x_2) = (2x_1, 2x_2)\} \\ &= \{(0, x_2) : x_2 \in \mathbb{R}\} = \langle (0, 1) \rangle. \end{aligned}$$

Así, $\gamma_2(A) = 1$ y $\mu_2(A) = 2$. Por lo tanto, 2 es un autovalor defectivo y $A$ no es semisimple.

**Teorema 8.28 (autovectores linealmente independientes).** Sean $\{\lambda_1, \lambda_2, \ldots, \lambda_r\}$ autovalores distintos de $A \in M_{n\times n}(F)$ y $\{\mathbf{v_1}, \ldots, \mathbf{v_r}\}$ un conjunto de vectores en $F^n$ tal que, para cada $i$, $\mathbf{v_i}$ es un $\lambda_i$-autovector de $A$. Entonces, el conjunto $\{\mathbf{v_1}, \ldots, \mathbf{v_r}\}$ es linealmente independiente.

**Demostración.** La demostración es por inducción sobre $r$. El caso $r = 1$ está claro. Supongamos que el teorema es cierto para $r - 1$ y demostraremos que también es cierto para $r$. Consideremos una combinación lineal

$$\alpha_1 \mathbf{v_1} + \cdots + \alpha_r \mathbf{v_r} = \mathbf{0}, \tag{8.2}$$

donde $\alpha_i \in F$. Aplicando la matriz $A$ en ambos lados de (8.2) obtenemos

$$\begin{aligned} A(\alpha_1 \mathbf{v_1} + \cdots + \alpha_r \mathbf{v_r}) &= A\mathbf{0} \\ \alpha_1 A\mathbf{v_1} + \cdots + \alpha_r A\mathbf{v_r} &= \mathbf{0} \\ \lambda_1 \alpha_1 \mathbf{v_1} + \cdots + \lambda_r \alpha_r \mathbf{v_r} &= \mathbf{0}. \end{aligned}$$

Multiplicando por $\lambda_r$ ahora de ambos lados en (8.2) obtenemos

$$\lambda_r \alpha_1 \mathbf{v_1} + \cdots + \lambda_r \alpha_r \mathbf{v_r} = \mathbf{0}.$$

Por lo tanto,

$$(\lambda_r - \lambda_1)\alpha_1\mathbf{v_1} + \cdots + (\lambda_r - \lambda_i)\alpha_i\mathbf{v_i} + \cdots + (\lambda_r - \lambda_{r-1})\alpha_{r-1}\mathbf{v_{r-1}} = \mathbf{0}.$$

Notemos que la igualdad anterior solo involucra $r-1$ vectores, así que por hipótesis de inducción $(\lambda_r - \lambda_i)\alpha_i = 0$ para toda $i = 1, \ldots r-1$. Como $\lambda_r \neq \lambda_i$ por hipótesis del teorema, tenemos que $\alpha_i = 0$ para toda $i = 1, \ldots, r-1$. Substituyendo esto en (8.2), tenemos que $\alpha_r\mathbf{v_r} = \mathbf{0}$ así que $\alpha_r = 0$. Esto demuestra que el conjunto $\{\mathbf{v_1}, \ldots, \mathbf{v_r}\}$ es linealmente independiente. □

**Corolario 8.29.** Sea $A \in M_{n\times n}(F)$ una matriz con $n$ autovalores distintos. Entonces, existe una base de $F^n$ formada por autovectores de $A$.

## 8.2. Matrices y endomorfismos diagonalizables

Sea $D \in M_{n\times n}(F)$ una matriz diagonal, es decir, una matriz de la forma

$$\begin{pmatrix} d_{11} & 0 & \cdots & 0 \\ 0 & d_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & d_{nn} \end{pmatrix},$$

El polinomio característico de $D$ es

$$p_D(x) = \prod_{i=1}^{n}(x - d_{ii})$$

y por lo tanto los autovalores de $D$ son $d_{11}, d_{22}, \ldots, d_{nn}$ (posiblemente repetidos). Además, para toda $i = 1, \ldots, n$ se cumple que el autovalor $d_{ii}$ tiene a $e_i$ como autovector asociado, donde $e_i$ es el $i$-ésimo elemento de la base canónica de $F^n$. De este modo, la información de los valores y autovectores de una matriz diagonal $D$ se obtiene directamente de sus elementos en la diagonal. Esto implica que la transformación lineal asociada a $D$ tiene propiedades algebraicas y geométricas sencillas de entender. Por ejemplo, si consideramos a $D$ como transformación lineal, entonces $D$ envía al vector canónico $e_i$ al múltiplo $d_{ii}e_i$. En consecuencia, la matriz $D$ envía los ejes coordenados en sí mismos, comprimiéndolos o estirándolos según el autovalor $d_{ii}$.

Otro tipo de matrices relativamente sencillas de entender son las triangulares superiores ya que en este caso los autovalores también coinciden con los elementos de su diagonal. Estas matrices las trataremos con más cuidado en la sección 9.4.

**Definición 8.30 (matriz diagonalizable).** Sea $A \in M_{n\times n}(F)$ una matriz. Decimos que $A$ es *diagonalizable* si es similar a una matriz diagonal. Es decir, si existe una matriz invertible $P$ tal que $D = P^{-1}AP$ es una matriz diagonal.

**Lema 8.31 (autovalores de una matriz diagonalizable).** Si $A$ es similar a la matriz diagonal $D = (d_{i,i})$, entonces $\{d_{1,1}, \ldots, d_{n,n}\}$ es el conjunto de autovalores de $A$.

**Demostración.** Como $A$ es similar a $D$, por el Corolario 8.13, los autovalores de las matrices $A$ y $D$ son exactamente los mismos. El lema queda demostrado porque los autovalores de $D$ son $\{d_{1,1}, \ldots, d_{n,n}\}$. □

**Teorema 8.32 (matriz diagonalizable).** La matriz $A \in M_{n\times n}(F)$ es diagonalizable si y solo si existe una base de $F^n$ formada por autovectores de $A$.

**Demostración.**

($\Rightarrow$) Supongamos que $A$ es diagonalizable. Por definición, existe una matriz invertible $P \in M_{n\times n}(F)$ tal que $D = P^{-1}AP$, donde $D$ es una matriz diagonal

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

Consideremos el conjunto de columnas de $P$, es decir, el conjunto $B = \{Pe_1, \ldots, Pe_n\}$, donde $e_i$ es el $i$-ésimo vector de la base canónica de $F^n$. Por el Teorema Fundamental de Matrices Invertibles, $B$ es una base de $F^n$. Además, los elementos de $B$ son autovectores de $A$ porque

$$A(Pe_i) = PDP^{-1}(Pe_i) = PDe_i = P\lambda_i e_i = \lambda_i(Pe_i).$$

Esto demuestra la primera implicación.

($\Leftarrow$) Supongamos que $B = \{v_1, \ldots, v_n\}$ es una base de $F^n$ tal que $v_i$ es un $\lambda_i$-autovector de $A$. Sea $P \in M_{n\times n}$ la matriz cuyas columnas son los vectores de la base $B$. Por el Teorema Fundamental de Matrices Invertibles, $P$ es invertible. Finalmente demostraremos que $D := P^{-1}AP$ es una matriz diagonal. Observemos que la $i$-ésima columna de $D$ es

$$D(e_i) = P^{-1}AP(e_i) = P^{-1}Av_i = P^{-1}\lambda_i v_i = \lambda_i P^{-1}v_i = \lambda_i e_i.$$

Por lo tanto, cada columna de $D$ es $(0, \ldots, \lambda_i, \ldots, 0)^T$, lo que demuestra que $D$ es diagonal.

□

**Observación 8.33.** Si $A \in M_{n\times n}(F)$ tiene $n$ autovalores distintos $\lambda_1, \ldots, \lambda_n$, por el Teorema 8.28, existe una base de $F^n$ formada por autovectores de $A$. Por lo tanto, por el Teorema 8.32, $A$ es diagonalizable, y de hecho es similar a la matriz diagonal $D = (d_{ii})$, con $d_{ii} = \lambda_i$. En conclusión, si $A \in M_{n\times n}(F)$ es una matriz con $n$ autovalores distintos, entonces $A$ es diagonalizable. Sin embargo, esta condición es suficiente mas no es necesaria. En otras palabras, existen matrices con autovalores repetidos que son diagonalizables, un ejemplo de ello es la matriz identidad. Por tal motivo es necesario dar condiciones más precisas para determinar si $A$ es diagonalizable.

**Teorema 8.34 (definiciones equivalentes de matriz diagonalizable).** Sea $A \in M_{n\times n}(F)$. Las siguientes afirmaciones son equivalentes:

**(i)** $A$ es diagonalizable.

**(ii)** Existe una base de $F^n$ formada por autovectores de $A$.

**(iii)** $F^n = S_{\lambda_1}(A) \oplus \cdots \oplus S_{\lambda_k}(A)$, donde $\lambda_1, ..., \lambda_k$ son los distintos autovalores de $A$.

**(iv)** $\gamma_{\lambda_1}(A) + ... + \gamma_{\lambda_k}(A) = n$, donde $\lambda_1, ..., \lambda_k$ son los distintos autovalores de $A$

**Demostración.** La equivalencia entre **(i)** y **(ii)** quedó establecida en el Teorema 8.32. El resto de las equivalencias quedan como ejercicio. □

Una consecuencia del Teorema Fundamental del Álgebra es que cualquier polinomio $p(x) \in \mathbb{C}[x]$ (con coeficientes en los números complejos) de grado $n$ puede factorizarse como $p(x) = a(x-r_1)^{m_1} \dots (x-r_k)^{m_k}$ donde $a$ es el coeficiente principal de $p(x)$ (i.e. el coeficiente de $x^n$), $r_1, \dots, r_k \in \mathbb{C}$ son las distintas raíces de $p(x)$, y $n = m_1 + \cdots + m_k$. Usaremos este hecho para demostrar el siguiente teorema.

**Teorema 8.35 (matriz compleja diagonalizable).** Una matriz $A \in M_{n\times n}(\mathbb{C})$ es diagonalizable si y solo si $A$ es semisimple.

**Demostración.** Recordemos que, por definición, $A$ es semisimple si y solo si $\gamma_\lambda(A) = \mu_\lambda(A)$ para todo autovalor $\lambda \in \mathbb{C}$ de $A$. Sea $p_A(x) \in \mathbb{C}[x]$ el polinomio característico de $A$. Observemos que $p_A(x)$ tiene grado $n$ y su coeficiente principal es 1. Por el Teorema Fundamental del Álgebra, tenemos que

$$p(x) = (x - \lambda_1)^{m_1} \dots (x - \lambda_k)^{m_k},$$

donde $\lambda_1, \dots, \lambda_k$ son los distintos autovalores de $A$ y $m_i = \mu_{\lambda_i}(A)$, por definición de multiplicidad algebraica. Además, se cumple que

$$\mu_{\lambda_1}(A) + ... + \mu_{\lambda_k}(A) = n. \tag{8.3}$$

Ahora demostraremos ambas implicaciones:

**(⇒)** Supongamos que $A$ es diagonalizable. Por el Teorema 8.34, $\sum_{i=1}^k \gamma_{\lambda_i}(A) = n$. Recordemos que $\gamma_{\lambda_i}(A) \leq \mu_{\lambda_i}(A)$ para todo $\lambda_i$. Sin embargo, por (8.3), tenemos que $\sum_{i=1}^k \gamma_{\lambda_i}(A) = \sum_{i=1}^k \mu_{\lambda_i}(A)$, lo que implica que $\gamma_{\lambda_i}(A) = \mu_{\lambda_i}(A)$ para todo $\lambda_i$. Por lo tanto, $A$ es semisimple.

**(⇐)** Si $A$ es semisimple, $\gamma_{\lambda_i}(A) = \mu_{\lambda_i}(A)$ para todo $\lambda_i$, y $\sum_{i=1}^k \gamma_{\lambda_i}(A) = \sum_{i=1}^k \mu_{\lambda_i}(A) = n$. Por el Teorema 8.34, $A$ es diagonalizable.

□

El teorema anterior no es válido para campos donde el Teorema Fundamental del Álgebra no se cumple (es decir, campos que no son *algebraicamente cerrados*), como es el caso de $\mathbb{R}$.

**Ejemplo 8.36.** Sea $A \in M_{3\times 3}(\mathbb{R})$ una matriz definida como

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

El polinomio característico de $A$ es

$$p_A(x) = \det(xI_3 - A) = x^3 - x^2 + x - 1 = (x-1)(x^2+1).$$

El único autovalor real de $A$ es $\lambda_1 = 1$, y tenemos que $\gamma_{\lambda_1}(A) = \mu_{\lambda_1}(A) = 1$. Esto demuestra que $A$ es semisimple. Sin embargo, $A$ no es diagonalizable porque no existen 3 $\lambda_1$-autovectores linealmente independientes.

Por otro lado, si consideramos a $A$ como una matriz con entradas en $\mathbb{C}$, entonces $A$ tiene 3 autovalores distintos: $\lambda_1 = 1$, $\lambda_2 = i$ y $\lambda_3 = -i$. Luego, en este caso, $A$ es semisimple y diagonalizable, de acuerdo al teorema anterior.

***Palabras clave:*** *similitud de matrices, autovalores y autovectores, polinomio característico, autoespacio, subespacio invariante, multiplicidades geométrica y algebraica de autovalores, autovalores defectivos, simples y semisimples, matrices*

## 8.3. Ejercicios

**Ejercicio 8.37.** Demuestra que la relación de similitud de matrices es una relación de equivalencia.

**Ejercicio 8.38.** Demuestra que si $A = (a_{i,j})$ es una matriz triangular superior, entonces los eigenvalores de $A$ son los elementos de la diagonal.

**Ejercicio 8.39.** Encuentra los eigenvalores y eigenvectores de cada una de las siguientes matrices. En cada caso, encuentra también las multiplicidades geométrica y algebraica de cada eigenvalor, y determina si las matrices son diagonalizables.

1. $A_1 = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$.
2. $A_2 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$.
3. $A_3 = \begin{bmatrix} 2 & 2 \\ 1 & 0 \end{bmatrix}$.
4. $A_4 = \begin{bmatrix} 2 & -3 & 7 \\ 0 & -2 & 3 \\ 0 & 0 & 0 \end{bmatrix}$.
5. $A_5 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$.

**Ejercicio 8.40.** Sea $A \in M_{n\times n}(F)$. Demuestra que $S_\lambda(A) = \ker(\lambda I_n - A)$.

**Ejercicio 8.41.** Sea $f$ un endomorfismo de $\mathbb{R}^2$. Determina si $f$ es diagonalizable en los siguientes casos.

1. $f(x, y) = (5x + 3y, -6x - 4y)$
2. $f(x, y) = (2y, y)$
3. $f(x, y) = (5x - 3y, 6x + 4y)$

**Ejercicio 8.42.** Sea $\phi : F^n \to F^n$ un endomorfismo, y sean $\lambda_1, ..., \lambda_k$ los autovalores de $\phi$. Demuestra lo siguiente:

1. $\phi$ es diagonalizable si y sólo si $\gamma_{\lambda_1}(\phi) + ... + \gamma_{\lambda_k}(\phi) = n$.
2. $\phi$ diagonalizable si y sólo si $F^n = S_{\lambda_1}(\phi) \oplus .... \oplus S_{\lambda_k}(\phi)$,
3. Si $\phi$ tiene $n$ autovalores distintos, entonces $\phi$ es diagonalizable.

**Ejercicio 8.43.** Sea $A \in M_{n\times n}(F)$ y sea

$$p_A(x) = x^n + a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \cdots + a_1x + a_0$$

el polinomio característico de $A$. Demuestra lo siguiente:

1. $a_0 = (-1)^n \det(A)$ (*Sugerencia:* recuerda que $\det(\alpha A) = \alpha^n \det(A)$ para cualquier $\alpha \in \mathbb{C}$).
2. Si $n = 2$ o $n = 3$, comprueba que $a_{n-1} = -\mathrm{tr}(A)$. (Esto es algo que se cumple para cualquier $n$, pero es más sencillo demostrarlo usando las herramientas del próximo capítulo).

**Ejercicio 8.44.** Sea $A \in M_{n\times n}(\mathbb{C})$, y sean $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ los autovalores de $A$. Demuestra que

$$\det(A) = \lambda_1 \ldots \lambda_n.$$

(*Sugerencia*: usa el hecho de que cualquier polinomio mónico $p(x) \in \mathbb{C}[x]$ puede factorizarse como $p(x) = (x-r_1)\ldots(x-r_n)$, donde $n = \deg(p(x))$ y $r_1, \ldots, r_n \in \mathbb{C}$ son las raíces de $p(x)$).

**Ejercicio 8.45.** Sea $D \in M_{n\times n}(F)$ una matriz diagonal. Demuestra que todo elemento en la diagonal determina es autovalor de $D$. Demuestra que los elementos $\{d_{ii}\}_{i=1}^n$ corresponden al conjunto de autovalores para $D$.

**Ejercicio 8.46.** Sea $\{\lambda_i\}_{i=1}^n$ un conjunto de autovalores de $A$ distintos entre si. Suponga que $\{v_i\}_{i=1}^n$ un conjunto de vectores tales que $v_i$ es un autovector asociado a $\lambda_i$, para cada $i = 1, \ldots, n$. Demuestra que el conjunto $\{v_i\}_{i=1}^n$ es linealmente independiente.

**Ejercicio 8.47.** Determina dos matrices no equivalentes $A, B \in M_{n\times n}(F)$ tales que $p_A(\lambda) = p_B(\lambda)$

# 9

# Forma Canónica de Jordan

En el Capítulo 8, establecimos la similitud de matrices como una relación de equivalencia. De esta manera, toda matriz $A$ está contenida en una clase de equivalencia $[A]$.

Además, establecimos condiciones para determinar si una matriz $A$ es diagonalizable. En otras palabras establecemos condiciones para determinar si la clase $[A]$ tiene un representante diagonal. Sin embargo, no toda matriz $A$ es diagonalizable, y en este caso será necesario determinar al mejor representante posible para la clase $[A]$. El Teorema de la Forma Canónica de Jordan es la respuesta a nuestra búsqueda del mejor representante, ya que establece que toda matriz cuadrada con entradas complejas es similar a una matriz triangular superior en la *forma de Jordan*.

Para lograr nuestro objetivo, en la Sección 9.1. introducimos el concepto de polinomio mínimo y estudiamos sus propiedades básicas. En particular, demostramos el Teorema de Cayley-Hamilton, el cual establece que cualquier matriz cuadrada es una raíz de su polinomio característico. En la Sección 9.2, demostramos que si el polinomio característico de una matriz se factoriza en polinomios lineales, entonces dicha matriz es similar a una matriz triangular. En la Sección 9.3, estudiamos endomorfismos nilpotentes que son todos aquellos que se anulan después de elevarlos a alguna potencia.

En la Sección 9.4, estudiamos matrices complejas, las cuales sabemos que siempre son similares a una matriz triangular, ya que todo polinomio complejo se factoriza en polinomios lineales (lo cual es una consecuencia del Teorema Fundamental del Álgebra). Más aún, el Teorema de la Forma Canónica de Jordan nos garantiza que esta matriz triangular tiene una forma "especial", llamada la *forma de Jordan.* Finalmente, en la Sección 9.5., estudiamos varios casos de matrices no diagonalizables y describimos cómo es posible encontrar su Forma Canónica de Jordan.

## 9.1. Teorema de Cayley-Hamilton

Dada una matriz $A \in M_{n\times n}(F)$ no cero, consideramos sus potencias

$$I, A, A^2, \ldots, A^{n^2}.$$

Estas matrices son elementos del espacio vectorial $M_{n\times n}(F)$ el cual tiene dimensión $n^2$. Ahora, ya que tenemos $n^2+1$ elementos en un espacio vectorial de dimensión $n^2$, entonces la lista $I, A, A^2, \ldots, A^{n^2}$ es linealmente dependiente, es decir, existe una combinación lineal

$$\alpha_{n^2}\mathbf{A}^{n^2} + \alpha_{n^2-1}\mathbf{A}^{n^2-1} + \cdots + \alpha_2\mathbf{A}^2 + \alpha_1\mathbf{A} + \alpha_0\mathbf{I} = \mathbf{0}, \tag{9.1}$$

donde algunos coeficientes $\alpha_i$ son distintos de cero. Notemos que, podría pasar que $A^m = 0$ para algún $m$, $1 \leq m \leq n$ o que $A^m = A^s$, con $1 \leq m, s \leq n$. En cualquier caso, la lista en cuestión sigue siendo linealmente dependiente (ver Observación 5.9).

Cabe señalar que la ecuación (9.1) también puede ser considerada como un polinomio en una variable, para ello sustituimos la matriz $A$ por la variable $\mathbf{X}$ y obtenemos

$$\alpha_{n^2}\mathbf{X}^{n^2} + \alpha_{n^2-1}\mathbf{X}^{n^2-1} + \cdots + \alpha_2\mathbf{X}^2 + \alpha_1\mathbf{X} + \alpha_0\mathbf{I}. \tag{9.2}$$

De este modo el polinomio (9.2) se anula en $A$, debido a (9.1). Por lo tanto, para toda matriz $A \in M_{n\times n}(F)$ existe un polinomio $p(\mathbf{X})$ que se anula en $A$. Esto nos lleva de manera natural a las siguientes preguntas ¿Es el polinomio (9.2) el único polinomio que se anula en $A$? En caso de existir varios polinomios que se anulan en $A$ ¿Cuál es el polinomio de grado mínimo que anula en $A$? Para responder a estas preguntas, iniciamos con la siguiente definición.

**Definición 9.1 (polinomio mínimo).** Sea $A \in M_{n\times n}(F)$ una matriz. El polinomio mínimo de $A$ es el polinomio mónico de menor grado que se anula en $A$ y lo denotaremos como $m_A(\lambda)$.

**Proposición 9.2 (existencia del polinomio mínimo).** Sea $A \in M_{n\times n}(F)$, entonces el polinomio mínimo de $A$ existe y es único.

**Demostración.** La existencia del polinomio mínimo se obtiene a partir de la existencia del polinomio (9.2) (ver, Ejercicio 9.48). Para demostrar la unicidad del polinomio mínimo, suponemos la existencia de dos polinomios minimales para $A$, digamos $m_A(\lambda)$ y $n_A(\lambda)$. Al ser ambos polinomios mónicos y del mismo grado se tiene que su diferencia $m_A(A) - n_A(A)$ es un polinomio de grado más pequeño el cual se anula en $A$. Sin embargo, esto contradice el hecho de que $m_A(\lambda)$ y $n_A(\lambda)$ sean minimales. □

**Proposición 9.3 (matrices similares y sus polinomios mínimos).** Si $A, B \in M_{n\times n}(F)$ son dos matrices tales que $A \sim B$, entonces el polinomio mínimo de $A$ y el polinomio mínimo de $B$ coinciden.

**Demostración.** Ya que $A$ y $B$ son similares, existe una matriz invertible $P$ tal que $A = P^{-1}BP$. Sea $m_B(\lambda)$ el polinomio minimal de $B$, entonces $m_B(A) = m_B(P^{-1}BP) = P^{-1}m_B(B)P = 0$ (ver Ejercicio 9.48). Lo anterior implica que $\deg m_A(\lambda) \leq \deg m_B(\lambda)$. De manera análoga tenemos que $m_A(B) = 0$ y por lo tanto $\deg m_B(\lambda) \leq \deg m_A(\lambda)$. En consecuencia, $\deg m_B(\lambda) = \deg m_A(\lambda)$ y por la unicidad del polinomio mínimo concluimos que $m_A(\lambda) = m_B(\lambda)$. □

**Proposición 9.4.** Para cualquier matriz $A \in M_{n\times n}(F)$ los polinomios característico y mínimo tienen las mismas raíces salvo multiplicidades.

**Demostración.** De la Proposición 8.6 sabemos que un escalar es autovalor de $A$ si, y solo si, es raíz de su polinomio característico. Así, basta demostrar que $t \in F$ es raíz del polinomio mínimo si, y solo si, $t$ es autovalor de $A$.

($\Rightarrow$) Sea pues $t$ una raíz de $m_A(\lambda)$, es decir, $m_A(t) = 0$. Entonces el polinomio $m_A(\lambda)$ tiene un factor $(\lambda - t)$ y por lo tanto

$$m_A(\lambda) = (\lambda - t)q(\lambda)$$

siendo $q(\lambda)$ un polinomio tal que $\deg q(\lambda) < \deg m_A(\lambda)$. Como $m_A(\lambda)$ es el polinomio de grado más pequeño que se anula en $A$, tenemos que $q(A) \neq \mathbf{0}$, i.e., $q(A)$ no es la matriz cero. Entonces, existe un vector $\mathbf{v} \in F^n$ tal que $\mathbf{v} \notin Ker(q(A))$, equivalentemente $q(A)\mathbf{v} \neq \mathbf{0}$. De esta manera, si hacemos $\mathbf{w} := q(A)\mathbf{v}$, tenemos las siguientes igualdades

$$(A - tI_n)\mathbf{w} = (A - tI_n)q(A)\mathbf{v} = m_A(A)\mathbf{v} = \mathbf{0},$$

por lo que $\mathbf{w}$ es un autovector de $A$ asociado a $t$.

($\Leftarrow$) Ahora suponemos que $t$ es un autovalor de $A$ y demostraremos que $t$ es raíz del polinomio $m_A(\lambda)$. Para ello veamos que como $t$ es autovalor de $A$, existe un vector $\mathbf{v}$ no nulo tal que $A\mathbf{v} = tI_n\mathbf{v}$. Ahora, si escribimos $m_A(\lambda) = \sum_{i=0}^{d} a_i\lambda^i$, entonces $m_A(A) = \sum_{i=0}^{d} a_iA^i$ obteniendo así las siguientes igualdades:

$$\begin{aligned} m_A(A)\mathbf{v} &= \left(\sum_{i=1}^{d} a_iA^i\right)\mathbf{v}, \\ &= \sum_{i=1}^{d} a_iA^i\mathbf{v}, \\ &= \left(\sum_{i=1}^{d} a_it^i\right)\cdot\mathbf{v}, \\ &= m_A(t)\cdot\mathbf{v} \end{aligned}$$

Finalmente, como $m_A(A) = \mathbf{0}$ es la matriz cero, entonces $m_A(A)\mathbf{v} = \mathbf{0}$ y por las igualdades anteriores $m_A(t)\cdot\mathbf{v} = \mathbf{0}$. Ahora bien, como $\mathbf{v}$ es un vector no nulo, concluimos que $m_A(t) = 0$, lo que implica que $t$ es raíz del polinomio minimal. □

Nuestro siguiente paso será mostrar que toda matriz $A$ es raíz de su polinomio característico.

**Teorema 9.5 (Cayley-Hamilton).** Sea $A \in M_{n\times n}(F)$ una matriz, Entonces el polinomio característico $p_A(\lambda)$ se anula en $A$, esto es $p_A(A) = 0$.

**Demostración.** Denotemos por $p_A(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0$ al polinomio característico de $A$. Sea $(\lambda I - A)$ la matriz característica de $A$ y $C$ la matriz adjunta de $(\lambda I - A)$. Por las propiedades de la adjunta (ver, Ejercicio 9.44), sabemos que

$$(\lambda I - A)C = |\lambda I - A|I. \tag{9.3}$$

Para el siguiente paso, consideramos las entradas de la matriz característica como polinomios de grado a lo más uno en la variable $\lambda$. De esta manera, ya que las entradas de $C$ se obtienen a partir de los determinantes de las menores de $(\lambda I - A)$, entonces las entradas de $C$ son polinomios de grado a lo más $n-1$ en la variable $\lambda$. Esto implica que la matriz adjunta puede ser escrita como un polinomio en la variable $\lambda$ donde sus coeficientes son matrices de $n \times n$, esto es:

$$C = C_{n-1}\lambda^{n-1} + C_{n-2}\lambda^{n-2} + \cdots + C_1\lambda + C_0.$$

Ahora, de la ecuación (9.3) obtenemos

$$(\lambda I - A)C = (\lambda I - A)(C_{n-1}\lambda^{n-1} + C_{n-2}\lambda^{n-2} + \cdots + C_1\lambda + C_0) = |\lambda I - A|I.$$

Ya que $|\lambda I - A| = p_A(\lambda)$, se tiene

$$(\lambda I - A)(C_{n-1}\lambda^{n-1} + C_{n-2}\lambda^{n-2} + \cdots + C_1\lambda + C_0) = (\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0)I.$$

Por consiguiente, tenemos las siguientes igualdades:

$$\begin{aligned} -AC_0 &= a_0 I \\ C_0 - AC_1 &= a_1 I \\ C_1 - AC_2 &= a_2 I \\ &\vdots \\ C_{n-2} - AC_{n-1} &= a_{n-1} I \\ C_{n-1} &= I. \end{aligned}$$

Multiplicando la primer igualdad por $I$, la segunda por $A$, la tercera por $A^2$ y así sucesivamente obtenemos:

$$\begin{aligned} -AC_0 &= a_0 I \\ AC_0 - A^2C_1 &= a_1 A \\ A^2C_1 - A^3C_2 &= a_2 A^2 \\ &\vdots \\ A^{n-1}C_{n-2} - A^nC_{n-1} &= a_{n-1}A^{n-1} \\ A^nC_{n-1} &= A^n. \end{aligned}$$

Sumando todas las igualdades obtenemos que

$$0 = A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 I.$$

Por lo tanto $p_A(A) = 0$, lo cual demuestra nuestra afirmación.
□

Claramente, el polinomio característico y minimal no son los únicos polinomios que se anulan en la matriz $A$. Un ejemplo de un tal polinomio es $q(x) := m_A(x) + p_A(x)$. Mas aún, si $f, g \in F[x]$ son dos polinomios que se anulan en $A$, entonces $f + g, \alpha f$ y $hf$ son polinomios que también se anulan en $A$ para cualquier escalar $\alpha \in F$ y cualquier polinomio $h \in F[x]$. Sin embargo todos los polinomios que se anulan en $A$ tienen algo en común, todos son múltiplos del polinomio mínimal tal y como se demuestra en el siguiente resultado.[1]

**Lema 9.6.** Si $f \in F[x]$ es un polinomio que se anula en $A$, entonces $f$ es de la forma $f = m_A \cdot h$ para algún polinomio $h \in F[X]$.

**Demostración.** Si $f$ es el polinomio mínimo, entonces no hay nada que probar. Así suponemos que $f$ es un polinomio que se anula en $A$ y que no es un múltiplo del polinomio mínimo. De esta manera, $\deg(f) > \deg(m_A)$. Aplicando el algoritmo de la división para polinomios obtenemos que existen dos polinomios $h, r \in F[x]$ tales que

$$f = m_A \cdot h + r$$

y $\deg(r) < \deg(m_A)$. Para finalizar la demostración probaremos que $r = 0$. Supongamos que $r \neq 0$, en este caso como $\deg(r) < \deg(m_A)$, entonces $r(A) \neq \mathbf{0}$. Por otro lado,

$$f(A) = m_A(A) \cdot h(A) + r(A) = \mathbf{0}$$

y como los polinomios $f$ y $m_A$ sí se anulan en $A$, entonces $r(A) = \mathbf{0}$. Sin embargo, esto es una contradicción y con esto concluimos que la única opción es $r = 0$.
□

## 9.2. Endomorfismos triangulables

Como vimos en el Capítulo 8, existen endomorfismos (matrices) que son digonalizables y la ventaja de estos endomorfismos es que podemos describir fácilmente sus propiedades, su valores y sus vectores propios. Sin embargo, no todo endomorfismo es diagonalizable así que buscaremos una manera más general de expresar los endomorfismos. De esta manera, si la matriz asociada a un endomorfismo es triangular superior (inferior) entonces el endomorfismo puede ser descrito de manera similar al caso de los endomorfismos diagonalizables. Así que nuestro siguiente paso es determinar cuando un endomorfismo puede ser representado por una matriz triangular.

[1] Esta observación puede expresarse de la siguiente manera: el conjunto de polinomios que se anulan en la matriz $A$ determina un ideal en $F[x]$ y ya que el anillo de polinomios en una variable es un dominio de ideales principales, entonces está generado por un solo polinomio el cual resulta ser el polinomio minimal. Por esta razón, todo polinomio que se anula en $A$ es múltiplo de $m_A$.

**Definición 9.7.** Decimos que un endomorfismo $\varphi : V \to V$ es triangulable si existe una base $B$ de $V$ tal que $[\varphi]_B^B$ es triangular. Similarmente para matrices $A \in M_{n\times n}(F)$.

Observe que por definición todo endomorfismo diagonalizable es triangulable. Por tal situación buscaremos un resultado general al dado para diagonalizable. Para ello necesitaremos las siguientes consideraciones.

**Proposición 9.8.** Sea $\varphi : V \to V$ un endomorfismo y $U$ un subespacio invariante de $V$. Entonces $\varphi$ determina por restricción un endomorfismo $\varphi_U : U \to U$ y a su vez determina un endomorfismo $\varphi_{V/U} : V/U \to V/U$ en el cociente.

**Demostración.** El endomorfismo $\varphi_U : U \to U$ es simplemente la restricción de $\varphi$ al subespacio invariante y como tal está definido como $\varphi_U(\mathbf{u}) := \varphi(\mathbf{u})$ esto determina inmediatamente un endomorfismo. Por otra parte, el endomorfismo $\varphi_{V/U} : V/U \to V/U$ se define de la siguiente manera: $\varphi_{V/U}([\mathbf{v}]) := [\varphi(\mathbf{v})]$, para todo $[\mathbf{v}] \in V/U$. Es un ejercicio para el lector verificar que $\varphi_{V/U}$ es un endomorfismo.
□

Bajo las hipotesis de la proposición anterior se tiene el siguiente resultado.

**Proposición 9.9.** Todo polinomio que se anula en $\varphi$, también se anula en $\varphi_{V/U}$.

**Demostración.** Sea $q \in F[x]$ un polinomio que se anula en $\varphi$ y $[\mathbf{v}] \in V/U$ un vector cualquiera. Aplicando la Proposición 9.8 se obtiene $q(\varphi_{V/U})([\mathbf{v}]) = [q(\varphi)(\mathbf{v})] = [\mathbf{0}]$ y, en consecuencia, el polinomio $q$ también se anula en $\varphi_{V/U}$ obteniendo nuestro resultado.
□

**Corolario 9.10.** El polinomio mínimo de $\varphi_{V/U}$ divide al polinomio mínimo de $\varphi$.

Si consideramos un endomorfismo $\varphi : V \to V$ y un subespacio invariante $U$, como en la Proposición 9.8. Entonces es posible determinar una base $B = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_r}, \mathbf{v_{r+1}}, \ldots \mathbf{v_n}\}$ de $V$, tal que $B_U := \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_r}\}$ es base para $U$ y $B_{V/U} = \{[\mathbf{v_{r+1}}], \ldots [\mathbf{v_n}]\}$ una base para el espacio cociente. Así, en términos de la base $B$ tenemos:

$$[\varphi]_B^B = \begin{pmatrix} M_U & M_2 \\ 0 & M_3 \end{pmatrix}$$

donde $M_U$ es una matriz $r \times r$ y corresponde a la matriz de $[\varphi_U]_{B_U}^{B_U}$. Para determinar esto observe que para cada $i = 1, \ldots, r$ se cumple

$$\varphi_U(\mathbf{v_i}) = \sum_{j=1}^{n} a_{ji}\mathbf{v_j},$$

con $a_{ji} = 0$ si $j > r$. De esta manera tenemos que $M_U = (a_{ij})_{i,j=1}^{r}$.

**Proposición 9.11.** Sea $\varphi : V \to V$ un endomorfismo, $U$ un subespacio invariante de $V$ y $q \in F[x]$ un polinomio. Entonces los endomorfismos $q(\varphi_U)$ y $q(\varphi)$ coinciden en $U$.

**Demostración.** Si $q$ es un polinomio constante, entonces la igualdad se cumple trivialmente. Por lo que supondremos que $q$ es un polinomio no constante y denotamos $q = \sum_{i=0}^{r} a_i x^i$. Sea $\mathbf{u} \in U$, entonces

$$q(\varphi_U)(\mathbf{u}) = \sum_{i=0}^{r} a_i \varphi_U^i(\mathbf{u}) = \sum_{i=0}^{r} a_i \varphi^i(\mathbf{u}) = q(\varphi)(\mathbf{u}).$$

y como $\mathbf{u}$ es arbitrario en $U$, entonces $q(\varphi_U) = q(\varphi)$ en $U$. □

**Corolario 9.12.** $p_\varphi(\varphi_U) = \mathbf{0}$.

**Corolario 9.13.** El polinomio minimal $m_{\varphi_U}$ divide al polinomio minimal $m_\varphi$.

**Demostración.** Por propiedad del polinomio mínimo, $m_\varphi(\varphi) = \mathbf{0}$. Así de la proposición anterior $m_\varphi(\varphi_U) = m_\varphi(\varphi) = \mathbf{0}$. Y por el Lema 9.6, $m_{\varphi_U}$ divide a $m_\varphi$.

□

Ahora que hemos determinado algunas propiedades de los endomorfismos, entonces determinaremos cuando un endomorfismo es triangulable, compare el siguiente resultado con el Teorema 8.32.

**Teorema 9.14.** Sea $\varphi : V \to V$ un endomorfismo tal que su polinomio caracteristico se factoriza en polinomios lineales. Entonces el endomorfismo $\varphi$ es triangulable.

**Demostración.** Demostraremos el teorema mediante inducción sobre la dimensión de $V$. Si $\dim V = 1$, entonces no hay nada que demostrar pues la matriz asociada a $\varphi$ es $1 \times 1$ y por lo tanto es triangular. Ahora suponga cierto para $\dim V = n - 1$ y probaremos para el caso $\dim V = n$.

Como el polinomio característico se factoriza mediante polinomios lineales, entonces $p_\varphi(x) = (x - \lambda_1)^{r_1} \cdots (x - \lambda_s)^{r_s}$. De esta manera, $\lambda_1$ es un valor propio asociado a $\varphi$ y sea $\mathbf{v_1}$ un vector propio asociado a $\lambda_1$. Defina $U := \langle \mathbf{v_1} \rangle$, de esta manera $U$ es un subespacio invariante y por la Proposición 9.8, se induce un endomorfismo $\varphi_{V/U} : V/U \to V/U$. Ahora de la Proposición 9.9 sabemos que $p_\varphi$ se anula en $\varphi_{V/U}$. Y aplicando el Lema 9.6 se tiene que el polinomio mínimo $m_{\varphi_{V/U}}$ divide a $p_\varphi$ y en consecuencia $m_{\varphi_{V/U}}$ se factoriza en polinomios lineales. Ahora, de lo anterior y la Proposición 9.4 tenemos que el polinomio $\varphi_{V/U}$ se factoriza por polinomios lineales y como $\dim \varphi_{V/U} = n - 1$ entonces la hipotesis de inducción implica que $\varphi_{V/U}$ es triangulable. Es decir, existe una base $B' = \{[\mathbf{v_2}], [\mathbf{v_3}], \ldots, [\mathbf{v_n}]\}$ de $V/U$ tal que $[\varphi_{V/U}]_{B'}^{B'}$ es triangular. Mas aún, $B := \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ es una base para $V$ y afirmamos que $[\varphi]_B^B$ es triangular superior. Para verificar esta afirmación vea que si $[\varphi_{V/U}]_{B'}^{B'} = (b_{ij})$ es una matriz triangular de $(n-1) \times (n-1)$, entonces

$$[\varphi]_B^B = \begin{pmatrix} \lambda_1 & M \\ \mathbf{0} & [\varphi_{V/U}]_{B'}^{B'} \end{pmatrix}$$

donde $M$ es una submatriz de $1 \times (n-1)$ y $\mathbf{0}$ representa una submatriz de $(n-1) \times 1$. □

## 9.3. Endomorfismos nilpotentes

Como antes, $U$ denotará un $F-$espacio vectorial y $\varphi : U \to U$ un endomorfismo. Observe que la composición $\varphi \circ \varphi$ es nuevamente un endomorfismo de $U$, el cual denotaremos por $\varphi^2$. En general tendremos que $\varphi^r$ denota al endomorfismo obtenido de la composición

$$\underbrace{\varphi \circ \varphi \circ \cdots \circ \varphi}_{r-veces}.$$

**Definición 9.15.** Sea $\varphi : U \to U$ un endomorfismo, decimos que $\varphi$ es nilpotente si $\varphi^r = 0$, para algún $r > 0$. Además diremos que el índice de nilpotencia de $\varphi$ es $r$ si $\varphi^{r-1} \neq 0$.

**Ejemplo 9.16.** Si consideramos el endomorfismo $\varphi : F^3 \to F^3$ definido por $\varphi(x, y, z) = (y, z, 0)$. Entonces $\varphi$ es un endomorfismo nilpotente con índice de nilpotencia igual a tres, esto ya que si $(x, y, z) \in F^3$ cualquiera, entonces tenemos:

$$(\varphi \circ \varphi \circ \varphi)(x, y, z) = (\varphi \circ \varphi)(\varphi(x, y, z)) = \varphi(\varphi(y, z, 0)) = \varphi(z, 0, 0) = (0, 0, 0).$$

**Ejemplo 9.17.** Sea $\psi : F^3 \to F^3$ definido por $\psi(x, y, z) = (0, z, 0)$. Entonces $\psi$ es un endomorfismo nilpotente con índice de nilpotencia igual a dos, pues como en el ejemplo anterior se tiene:

$$(\psi \circ \psi)(x, y, z) = \psi(\psi(x, y, z)) = \psi(0, z, 0)) = (0, 0, 0).$$

De manera similar, si $A \in M_{n \times n}$ es nilpotente entonces $A^r = \mathbf{0}$ para algún $r > 0$. Más aún, por las propiedades del determinante se tiene que $|A| = 0$.

**Proposición 9.18.** Si $\varphi : U \to U$ es un endomorfismo nilpotente, entonces cero es el único autovalor para $\varphi$.

**Demostración.** Sea $\varphi$ un endomorfismo nilpotente con índice de nilpotencia $r$ y suponga que $t \in F$ es un autovalor de $\varphi$. Entonces, existe un vector $\mathbf{u} \in U$ distinto de cero tal que

$$\varphi(\mathbf{u}) = t\mathbf{u}$$

y así $\varphi^n(\mathbf{u}) = t^n\mathbf{u} = 0$. Ello implica que $t^n = 0$ y por lo tanto $t = 0$. □

**Corolario 9.19.** Si $\varphi : U \to U$ es un endomorfismo nilpotente con índice de nilpotencia $r$, entonces el polinomio mínimo de $\varphi$ es $m_\varphi(\lambda) = \lambda^r$.

**Demostración.** Ejercicio 9.47. □

**Proposición 9.20.** Sea $\varphi : U \to U$ un endomorfismo nilpotente con índice de nilpotencia $r$ y $\mathbf{u} \in U$ un vector tal que $\varphi^{r-1}(\mathbf{u}) \neq 0$. Entonces el conjunto

$$\{\varphi^{r-1}(\mathbf{u}), \varphi^{r-2}(\mathbf{u}), \ldots, \varphi(\mathbf{u}), \mathbf{u}\},$$

es un conjunto linealmente independiente.

**Demostración.** Para demostrar la afirmación consideramos una combinación lineal del vector cero, digamos

$$a_{r-1}\varphi^{r-1}(\mathbf{u}) + a_{r-2}\varphi^{r-2}(\mathbf{u}) + \cdots + a_1\varphi(\mathbf{u}) + a_0\mathbf{u} = \mathbf{0} \tag{9.4}$$

y probaremos que $a_i = 0$, para todo $i = 0, \ldots, r-1$.

Con este fin aplicamos $\varphi^{r-1}$ a ambos lados de la ecuación (9.4) y obtenemos lo siguiente:

$$\varphi^{r-1}(a_{r-1}\varphi^{r-1}(\mathbf{u}) + a_{r-2}\varphi^{r-2}(\mathbf{u}) + \cdots + a_1\varphi(\mathbf{u}) + a_0\mathbf{u} = \varphi^{r-1}(\mathbf{0})$$

Tomando en cuenta que $\varphi^r = 0$, tenemos que $\varphi^{2r-2} = \varphi^{2r-1} = \cdots = \varphi^r = 0$ y en consecuencia la ecuación anterior queda de la siguiente manera:

$$a_0\varphi^{r-1}(\mathbf{u}) = \mathbf{0}.$$

Ahora, ya que $\varphi^{r-1}(\mathbf{u}) \neq \mathbf{0}$, entonces $a_0 = 0$. De manera análoga, si aplicamos $\varphi^{r-2}$ a la ecuación (9.4) entonces podremos concluir que $a_1 = 0$. Siguiendo este procedimiento podemos determinar que la combinación lineal es la trivial. En conclusión, el conjunto es linealmente independiente. □

Sea $\{\varphi^{r-1}(\mathbf{u}), \varphi^{r-2}(\mathbf{u}), \ldots, \varphi(\mathbf{u}), \mathbf{u}\}$ el conjunto dado en la Proposición 9.20, denotaremos por $W$ al espacio generado por este conjunto; esto es:

$$W := \langle \varphi^{r-1}(\mathbf{u}), \varphi^{r-2}(\mathbf{u}), \ldots, \varphi(\mathbf{u}), \mathbf{u} \rangle.$$

El conjunto $W$ es por construcción un subespacio $\varphi$-invariante, esto es, $\varphi(W) \subset W$. De esta manera, podemos considerar la restriccíón de $\varphi$ a $W$

$$\varphi|_W : W \to W$$

el cual es un endomorfismo de $W$ (ver, Ejercicio 1) y como el conjunto $\beta = \{\varphi^{r-1}(\mathbf{u}), \varphi^{r-2}(\mathbf{u}), \ldots, \varphi(\mathbf{u}), \mathbf{u}\}$ es una base para $W$, entonces $\dim(W) = r$ y $[\varphi|_W]^\beta_\beta$ es una matriz $r \times r$.

**Proposición 9.21.** Sea $\varphi : U \to U$ un endomorfismo nilpotente con índice de nilpotencia $r$ y $\mathbf{u} \in U$ tal que $\varphi^{r-1}(\mathbf{u}) \neq \mathbf{0}$. Si $\beta = \{\varphi^{r-1}(\mathbf{u}), \varphi^{r-2}(\mathbf{u}), \ldots, \varphi(\mathbf{u}), \mathbf{u}\}$, y $W \subset U$ es el subespacio generado por $\beta$, entonces

$$[\varphi|_W]^\beta_\beta = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ 0 & 0 & 0 & 0 & \ddots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & 1 \\ 0 & 0 & 0 & 0 & \ldots & 0 \end{pmatrix}$$

**Demostración.** La demostración es un ejercicio para el lector (ver, Ejercicio 9.50).
□

La forma que tiene la matriz $[\varphi|_W]_\beta^\beta$ es conocida como bloque de Jordan, este tipo de matrices serán estudiadas en el siguiente capítulo. Note además que el vector $\varphi^{r-1}(\mathbf{u})$ de la base $\beta$ es un autovector del endomorfismo $\varphi|_W$, (ver Ejercicio 2).

## 9.4. Forma Canónica de Jordan

En esta sección consideraremos principalmente matrices complejas, es decir, elementos de $M_{n\times n}(\mathbb{C})$. Como todo polinomio con coeficientes complejos se factoriza en polinomios lineales (lo cual es una consecuencia del Teorema Fundamental del Álgebra), sabemos por el Teorema 9.14 que toda matriz $A \in M_{n\times n}(\mathbb{C})$ es similar a una matriz triangular superior. En esta sección estudiaremos que esta matriz triangular superior tiene de hecho una forma especial.

**Definición 9.22 (bloque de Jordan).** Una matriz $A \in M_{n\times n}(F)$ es un bloque de Jordan si cumple las siguientes condiciones:

1. $A$ es triangular superior.
2. Todos los elementos en la diagonal son iguales.
3. Las entradas $a_{ii+1}$ (las que se encuentran sobre la diagonal superior) son iguales a uno, i.e., $a_{ii+1} = 1$ para toda $i = 1, \ldots, n-1$.

Así una matriz de $n \times n$ en bloque de Jordan es de la forma

$$\begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ 0 & 0 & \lambda & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{pmatrix}, \tag{9.5}$$

a esta matriz la denotaremos por $BJ_n(\lambda)$.

**Ejemplo 9.23.** Las siguientes matrices son algunos ejemplos de bloques de Jordan.

$$BJ_2(1) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad B_3(2) = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}, \quad BJ_4(3) = \begin{pmatrix} 3 & 1 & 0 & 0 \\ 0 & 3 & 1 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

Note que la matriz identidad $I_n$, con $n > 1$, no es bloque de Jordan debido a que las entradas $a_{ii+1} = 0$.

Al considerar el bloque de Jordan

$$BJ_2(0) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$

podemos comprobar que $(BJ_2(0))^2$ es la matriz cero y por lo tanto es una matriz nilpotente. En general, las matrices bloque de Jordan $BJ_n(0)$ son matrices nilpotentes (ver, Ejercicio 9.49).

**Proposición 9.24.** Si $A$ es un bloque de Jordan de la forma $BJ_n(\lambda)$, entonces $\mu_\lambda(A) = n$ y $\gamma_\lambda(A) = 1$.

**Demostración.** La demostración es un ejercicio para el lector (ver, Ejercicio 9.51). □

**Definición 9.25 (forma de Jordan).** Decimos que una matriz en $M_{n\times n}(F)$ está en forma de Jordan si la diagonal está compuesta por bloques de Jordan y las demás entradas son cero. Esto es, si tiene la siguiente estructura

$$\begin{pmatrix} BJ_{r_1}(\lambda_{i_1}) & 0 & 0 & \cdots & 0 \\ 0 & BJ_{r_2}(\lambda_{i_2}) & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & BJ_{r_s}(\lambda_{i_s}) \end{pmatrix}, \tag{9.6}$$

con $n = r_1 + r_2 + \cdots + r_s$.

**Ejemplos:** Las siguientes matrices se encuentran en su forma de Jordan.

1. $\begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, los bloques de Jordan son $BJ_2(2)$ y $BJ_1(1)$.

2. $\begin{pmatrix} -3 & 0 & 0 & 0 \\ 0 & -3 & 0 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 3 \end{pmatrix}$, los bloques de Jordan son $BJ_1(-3)$, $BJ_1(-3)$ y $BJ_2(3)$.

3. $I_n = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$, el bloque $BJ_1(1)$ aparece $n$ veces en la diagonal.

El siguiente teorema hace que la definición de la forma de Jordan sea particularmente interesante.

**Teorema 9.26 (Teorema de la Forma Canónica de Jordan).** Cualquier matriz $A \in M_{n\times n}(\mathbb{C})$ es similar a una matriz en la forma de Jordan.

Si una matriz se encuentra en su forma de Jordan, entonces la matriz es triangular superior. Así toda la información correspondiente a los autovalores, su multiplicidad algebraica y geométrica, queda completamente determinada por los elementos de la diagonal. El número de veces que aparece un mismo valor en la diagonal indica su multiplicidad algebraica. Por otra parte el número de bloques de Jordan indica el número de autovectores asociados y con ello su multiplicidad geométrica. Esto lo haremos evidente en el siguiente resultado.

**Proposición 9.27.** Sea $A \in M_{n\times n}(F)$ una matriz en forma de Jordan y suponga que los bloques de Jordan de $A$ son $\{BJ_{r_1}(\lambda_1), BJ_{r_2}(\lambda_2), \ldots, BJ_{r_s}(\lambda_s)\}$. Entonces se cumplen las siguientes afirmaciones:

1. Los autovalores de $A$ son $\{\lambda_j\}_{j=1}^{s}$.
2. Si $\lambda_i \neq \lambda_i$ siempre que $i \neq j$, entonces la multiplicidad algebraica de $\lambda_j$ es igual a $r_j$.
3. Si $\lambda_j = \lambda$ para todo $j$, entonces $\mu_\lambda(A) = r_1 + \cdots r_s = n$ y $\gamma_\lambda(A) = s$.

**Demostración.** (1) Ya que $A$ se encuentra en su forma de Jordan, entonces $A$ es triangular superior. De esta manera, los elementos en la diagonal son $\{\lambda_j\}_{j=1}^{s}$.
(2) Esta afirmación es una consecuencia de la Proposición 9.24.
(3) Para probarlo note que $A - \lambda I$ es una matriz de Jordan tal que sus bloques son de la forma $\{BJ_{r_1}(0), BJ_{r_2}(0), \ldots, BJ_{r_s}(0)\}$. Además, $\mathbf{v}$ es un vector propio de $A$ si y solo si $(A-\lambda I)\mathbf{v} = \mathbf{0}$. Ahora bien, si hacemos $\mathbf{v} = (x_1, \ldots x_n)$, entonces

$$(A - \lambda I)\mathbf{v} = (x_2, \ldots, x_{r_1}, 0, x_{r_1+2}, \ldots, x_{r_2}, 0, x_{r_2+2}, \ldots, x_{r_s-1}, 0) = \mathbf{0}$$

De esta manera $\mathbf{v}$ es un vector propio de $A$ si, y solo si, es de la forma

$$\mathbf{v} = (x_1, 0, \ldots, 0, x_{r_1}, 0, \ldots, x_{s-1}, 0, \ldots, x_s)$$

de lo cual concluimos que $\gamma_\lambda(A) = s$.
□

**Corolario 9.28.** Si $A$ es una matriz que está en forma de Jordan y $\lambda$ es un autovalor de $A$ el cual aparece en exactamente $m$ bloques, digamos $BJ_{r_1}(\lambda), \ldots, BJ_{r_m}(\lambda)$. Entonces $\mu_\lambda(A) = r_1 + \cdots + r_m$ y $\gamma_\lambda(A) = m$.

Por otra parte, si una matriz $A$ no se encuentra en su forma de Jordan, es posible calcular una matriz $B$ tal que $A \sim B$ y $B$ está en su forma de Jordan.

**Proposición 9.29.** Sea $A \in M_{n\times n}(F)$ una matriz con autovalores $\{\lambda_i\}_{i=1}^s$. Suponga que para cada $i$, $\mu_{\lambda_i}(A) = \gamma_{\lambda i}(A)$ y $\sum_{i=1}^s \gamma_{\lambda i}(A) = n$. Entonces existe una matriz $C$ tal que $CAC^{-1}$ es diagonal y por lo tanto es una matriz en forma de Jordan.

**Demostración.** Suponga que $\mu_{\lambda_i} = n_i$, entonces como la multiplicidad algebraica y geométrica coinciden, se tiene que $\dim(S_{\lambda_i}) = n_i$, para cada $i$. Sea $\{\mathbf{v}_{ij}\}_{j=1}^{n_i}$ una base para $S_{\lambda_i}$ y por lo tanto es un conjunto linealmente independiente de vectores propios asociados a $\lambda_i$. Por otro lado sabemos que $\Sigma_{i=1}^s n_i = n$. Por lo tanto el conjunto $\{\mathbf{v}_{ij}\}_{i,j=1}^{n,n_i}$ determina una base para $F^n$.

De esta forma, si denotamos por $C$ a la matriz cuyas primeras $n_1$ columnas corresponden a los autovectores $\{\mathbf{v}_{1j}\}_{j=1}^{n_1}$, las siguientes $n_2$ columnas corresponden a los autovectores $\{\mathbf{v}_{2j}\}_{j=1}^{n_2}$ y así sucesivamente, i.e.,

$$C = (\mathbf{v}_{11}\ \mathbf{v}_{12}\cdots\ \mathbf{v}_{1n_1}\ \mathbf{v}_{21}\ \mathbf{v}_{22}\cdots\mathbf{v}_{2n_2}\cdots\mathbf{v}_{s-1n_{s-1}}\ \mathbf{v}_{s1}\ \mathbf{v}_{s2}\cdots\mathbf{v}_{sn_s}).$$

Entonces la matriz C es invertible y $C^{-1}AC$ es una matriz diagonal de la forma

$$\begin{pmatrix} \lambda_1 I_{n_1} & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 I_{n_2} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_{n_s} I_{n_s} \end{pmatrix}.$$

□

**Observación 9.30.** Si en la Proposición 9.29 eliminamos la hipótesis $\sum_{i=1}^s \gamma_{\lambda i(A)} = n$, entonces la afirmación sería falsa (ver, Ejemplo 8.36). Sin embargo, la hipótesis sobre la multiplicidad algebraica puede ser eliminada cuando el campo $F$ es algebraicamente cerrado. Para ello compare la Proposición 9.29 con la Proposición 8.35.

## 9.5. Forma Canónica de Jordan de Matrices No Diagonalizables

Recordemos que una matriz compleja es diagonalizable si y solo si es semisimple, es decir, si no tiene autovalores defectivos. Si $\lambda$ es un autovalor defectivo de $A$, entonces para determinar una base de $S_\lambda$ es necesario un proceso especial. Veamos el siguiente ejemplo para explicar mejor esta situación.

**Ejemplo 9.31.** Tomamos la matriz

$$A = \begin{pmatrix} 0 & 1 & 2 \\ -5 & -3 & -7 \\ 1 & 0 & 0 \end{pmatrix}.$$

En este caso, la matriz $A$ tiene un único autovalor el cual es $\lambda = -1$ y tiene multiplicidad algebraica 3. Ahora para calcular los autovectores consideramos la matriz característica

$$(-1)I - A = \begin{pmatrix} -1 & -1 & -2 \\ 5 & 2 & 7 \\ -1 & 0 & -1 \end{pmatrix}.$$

Se verifica que el autovalor $-1$ tiene un único autovector linealmente independiente, digamos

$$v = \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix}.$$

Así, la multiplicidad geométrica es uno, i.e., $\mu_\lambda(A) = 1$ y de esta manera $A$ tiene un único autovalor el cual es defectivo. Esto implica que la matriz no es diagonalizable (ver, Proposición 8.32). En consecuencia, necesitamos un método que nos permita determinar su forma de Jordan.

### 9.5.1. Un solo autovalor defectivo y un solo autovector asociado

Del primer caso que trataremos se derivará el procedimiento para los casos subsecuentes, por lo que deberemos poner mayor énfasis a los detalles.

En primer lugar estudiaremos el caso en que la matriz $A \in M_{n\times n}(\mathbb{C})$ tiene un solo autovalor defectivo $\lambda$ y su espacio asociado $S_\lambda$ tiene dimensión uno.

En este caso, el polinomio característico de $A$ tiene la forma

$$p_A(\lambda) = (\lambda - a)^n$$

siendo $a$ el único autovalor de $A$. Además, ya que existe un único autovector asociado tenemos que $Nul(aI - A) = 1$ y supondremos que $\langle v\rangle = \ker(aI - A)$.

Por otra parte, es posible demostrar que bajo estas condiciones la matriz caracteristica $aI - A$ es nilpotente (ver, Ejercicio 9.52) y como el polinomio minimal de $A$ tiene los mismos factores que el polinomio característico, entonces $m_A(\lambda) = (\lambda - a)^m$ para algún $m$, con $m \leq n$. El siguiente resultado afirma que con las condiciones descritas anteriormente se cumple que $m = n$.

**Proposición 9.32.** Sea $A \in M_{n\times n}(F)$ una matriz cuyo polinomio característico es $p_A(\lambda) = (\lambda - a)^n$. Suponga además que el autovalor $a$ tiene un único autovector asociado. Entonces el polinomio minimal de $A$ es $m_A(\lambda) = (\lambda - a)^n$.

**Demostración.** Dado que $A$ tiene un solo autovector asociado al autovalor $a$, entonces $Nul(aI - A) = 1$. Por la Proposición 6.5, $rank(aI - A) = n - 1$. Ahora bien, sabemos que

$$\ker(aI - A) \subseteq \ker(aI - A)^2 \subseteq \ker(aI - A)^3 \subseteq \cdots \subseteq \ker(aI - A)^n$$

y así $Nul(aI - A)^m \leq m$ para toda $m$ tal que $1 \leq m \leq n$. Ello implica que para que $(aI - A)^m = 0$, entonces $m = n$. Por lo tanto el polinomio característico y minimal coinciden. □

**Corolario 9.33.** Si $A \in M_{n\times n}(F)$ es una matriz con un solo autovalor y un solo autovector. Entonces la matriz característica $aI - A$ es nilpotente con índice de nilpotencia $n$.

Consideremos una matriz $A$ con las hipótesis de la Proposición 9.32. Tomando en cuenta la Proposición 9.20, tenemos que existe una base $\beta$ de $F^n$ tal que

$$\beta = \{(A-aI)^{n-1}(\mathbf{u}), (A-aI)^{n-2}(\mathbf{u}), \ldots, (A-aI)(\mathbf{u}), \mathbf{u}\}$$

y la matriz $[A-aI]^{\beta}_{\beta}$ tiene la forma:

$$[A-aI]^{\beta}_{\beta} = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ 0 & 0 & 0 & 0 & \ddots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & 1 \\ 0 & 0 & 0 & 0 & \ldots & 0 \end{pmatrix},$$

y en consecuencia tenemos

$$[A]^{\beta}_{\beta} = BJ_n(a) = \begin{pmatrix} a & 1 & 0 & 0 & \cdots & 0 \\ 0 & a & 1 & 0 & \cdots & 0 \\ 0 & 0 & a & 1 & \cdots & 0 \\ 0 & 0 & 0 & a & \ddots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & 1 \\ 0 & 0 & 0 & 0 & \ldots & a \end{pmatrix}. \tag{9.7}$$

Una manera de determinar la base $\beta$ es la siguiente:

Si $A$ es como antes y $v$ el único autovector de $A$ asociado al autovalor $a$. Entonces determinamos vectores $\mathbf{u_1}, \mathbf{u_2}, \ldots \mathbf{u_{n-1}} \in F^n$ tales que

$$\begin{aligned} (A-aI)\mathbf{u_1} &= \mathbf{v}, \\ (A-aI)\mathbf{u_2} &= \mathbf{u_1}, \\ &\vdots \\ (A-aI)\mathbf{u_{n-2}} &= \mathbf{u_{n-3}}, \\ (A-aI)\mathbf{u_{n-1}} &= \mathbf{u_{n-2}}. \end{aligned} \tag{9.8}$$

Así la base buscada es

$$\beta = \{\mathbf{v}, \mathbf{u_1}, \mathbf{u_2}, \ldots \mathbf{u_{n-2}}, \mathbf{u_{n-1}}\}. \tag{9.9}$$

Es un ejercicio verificar que la base $\beta$ cumple las propiedades deseadas, ver Ejercicio 9.53. A los vectores $\mathbf{u_1}, \mathbf{u_2}, \ldots \mathbf{u_{n-2}}, \mathbf{u_{n-1}}$ obtenidos mediante el procedimiento anterior se les conoce como **autovectores generalizados de v**.

**Ejemplo 9.34.** Considere la matriz $A := \begin{pmatrix} 3 & -2 \\ 8 & -5 \end{pmatrix}$. Entonces, el polinomio característico para $A$ es $p_A(\lambda) = (\lambda+1)^2$ y así el único autovalor es $\lambda = -1$. El siguiente paso es determinar los autovectores asociados, para ello resolvemos el sistema $(A-(-1)I)\mathbf{x} = \mathbf{0}$, es decir:

$$\begin{aligned} 4x - 2y &= 0 \\ 8x - 4y &= 0. \end{aligned}$$

En este caso la solución está dada por $4x = 2y$ por lo que un autovector es $\mathbf{v_1} = (1,2)^t$. Esto implica que tenemos un único autovalor defectivo y un único autovector linealmente independiente. Por lo que buscaremos el autovector generalizados de $v$ mediante la solución del sistema:

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

En este caso la solución está dada por $4x = 1 + 2y$ y así un vector generalizado es $\mathbf{v_{11}} = (3/4, 1)^t$ y la base buscada es:

$$\beta = \left\{ \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 3/4 \\ 1 \end{pmatrix} \right\}.$$

Es un ejercicio para el lector comprobar que

$$[A-(-1)I]_\beta = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad y \quad [A]_\beta = \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix}$$

Note además que si

$$N = \begin{pmatrix} 1 & 3/4 \\ 2 & 1 \end{pmatrix},$$

entonces $N^{-1}AN = BJ_2(-1)$.

### 9.5.2. Un único autovalor defectivo

Cuando consideramos una matriz $A$ con un solo autovalor, entonces su polinomio característico tiene la forma $p_A(\lambda) = (\lambda - a)^n$ y si suponemos que $A$ tiene dos autovectores linealmente independientes, tenemos que $\text{nul}(A - aI) = 2$. Así que en este caso buscaremos la base $\beta$ en la cual matriz $[A-aI]_\beta^\beta$ es una matriz de Jordan. Para ello es necesario encontrar una base adecuada para $\ker(A-aI)$,

la cual sabemos que se compone de dos vectores, y completaremos dicha base utilizando la Proposición 5.29.

Supongamos que $\{\mathbf{v_1}, \mathbf{v_2}\}$ es una base para $\ker(A - aI)$, entonces la base $\beta$ buscada es de la forma:

$$\beta = \{\mathbf{v_1}, \mathbf{v_{11}}, \mathbf{v_{12}}, \ldots, \mathbf{v_{1n-2}}, \mathbf{v_2}\},$$

donde los vectores $\mathbf{v_{11}}, \mathbf{v_{12}}, \ldots, \mathbf{v_{1n-2}}$, son obtenidos con el procedimiento descrito en (9.8), esto es:

$$\begin{aligned}
(A - aI)\mathbf{v_{11}} &= \mathbf{v_1}, \\
(A - aI)\mathbf{v_{12}} &= \mathbf{v_{11}}, \\
&\vdots \\
(A - aI)\mathbf{v_{1n-3}} &= \mathbf{v_{1n-4}}, \\
(A - aI)\mathbf{v_{1n-2}} &= \mathbf{v_{1n-3}}.
\end{aligned} \tag{9.10}$$

Para aclarar ideas presentamos el siguiente:

**Ejemplo 9.35.** Para este caso consideraremos la matriz

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

Para esta matriz el polinomio característico es $p_A(\lambda) = -(\lambda - 1)^3$, por lo que el único autovalor de $A$ es $\lambda = 1$. Ahora resolvemos el sistema $(A - I)\mathbf{x} = \mathbf{0}$ para encontrar los autovectores, esto es:

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

La solución para este sistema se obtiene de resolver $2x = y - z$ por lo que $nul(A - I) = 2$ y obtendremos dos autovectores linealmente independientes. Dos de los cuales son:

$$\mathbf{v_1} = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \quad y \quad \mathbf{v_2} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

Nuestro siguiente paso es calcular el vector generalizado de $v_1$, para ello necesitamos una solución del sistema no homogéneo $(A - I)\mathbf{x} = \mathbf{v_1}$, a saber:

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

Por ejemplo, una solución es: $\mathbf{v_{11}} = (1,1,0)^t$. De esta manera nuestra base es

$$\beta = \left\{ \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\}.$$

Además tenemos que

$$[A-I]_\beta = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad y \quad [A]_\beta = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

siendo esta última la forma canónica de Jordan asociada a $A$.

**Observación 9.36.** Es necesario señalar que en ocasiones el sistema $(A-aI)\mathbf{x} = \mathbf{v_1}$ es inconsistente (no tiene solución) y en consecuencia no es posible encontrar el vector $\mathbf{v_{11}}$. En este caso consideramos el vector $v_2$ para determinar la base $\beta$, la cual tendrá la forma:

$$\beta = \{\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_{21}}, \mathbf{v_{22}}, \ldots, \mathbf{v_{2n-2}}\}$$

y los vectores $\mathbf{v_{21}}, \mathbf{v_{22}}, \ldots, \mathbf{v_{2n-2}}$ se obtienen de manera similar a (9.10), esto es, mediante las siguientes igualdades:

$$\begin{aligned} (A-aI)\mathbf{v_{21}} &= \mathbf{v_2}, \\ (A-aI)\mathbf{v_{22}} &= \mathbf{v_{21}}, \\ &\vdots \\ (A-aI)\mathbf{v_{2n-3}} &= \mathbf{v_{2n-4}}, \\ (A-aI)\mathbf{v_{2n-2}} &= \mathbf{v_{2n-3}}. \end{aligned} \tag{9.11}$$

Por último, podría suceder que tanto el sistema $(A-aI)\mathbf{x} = \mathbf{v_1}$ así como el sistema $(A-aI)\mathbf{x} = \mathbf{v_2}$ son inconsistentes y por lo tanto es necesario buscar una base distinta para $\ker(A-aI)$.

Para fijar ideas consideramos el caso $n = 3$ con $A-aI$ nilpotente con índice de nilpotencia dos. Suponga además que los autovectores obtenidos son $\{\mathbf{v_1}, \mathbf{v_2}\}$ y que los sistemas $(A+I)\mathbf{x} = \mathbf{v_1}$ y $(A+I)\mathbf{x} = \mathbf{v_2}$ son inconsistentes.

Ahora, como mencionamos anteriormente en este caso es necesario considerar una nueva base para $\ker(A+I)$. En este caso la estrategia es la siguiente:

Recordando la Proposición 9.32, la base la obtenemos considerando las ecuaciones (9.10), vea también (9.11). En este caso el primer sistema que debemos considerar es:

$$(A+I)\mathbf{x} = \alpha_1\mathbf{v_1} + \alpha_2\mathbf{v_2}, \quad \alpha_i \in F.$$

Los escalares $\alpha_1$ y $\alpha_2$ los tomaremos tales que el sistema sea consistente. Entonces, obtenemos el vector $\mathbf{u}$ tal que

$$(A+I)\mathbf{u} = \alpha_1\mathbf{v_1} + \alpha_2\mathbf{v_2},$$

y la base buscada es:

$$\beta := \{\alpha_1\mathbf{v_1} + \alpha_2\mathbf{v_2}, \mathbf{u}, \mathbf{v_i}\}$$

donde $\mathbf{v_i}$ es alguno de los autovectores iniciales, elegido de tal manera que $\beta$ sea base. Además la forma canónica de Jordan asociada a $A$ es

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

En general lo que buscamos es un vector $u \in \ker(A+I)^2 \setminus \ker(A+I)$, y la base es $\beta = \{(A+I)(\mathbf{u}), \mathbf{u}, \mathbf{v_i}\}$ con $\mathbf{v_i}$ linealmente independiente a $(A+I)(\mathbf{u})$ y $\mathbf{u}$.

**Ejemplo 9.37.** Tomamos la matriz

$$A = \begin{pmatrix} -1 & -3 & -9 \\ 0 & 5 & 18 \\ 0 & -2 & -7 \end{pmatrix}.$$

Para esta matriz tenemos $p_A(\lambda) = -(\lambda+1)^3$ y $m_A(\lambda) = (\lambda+1)^2$ y por tal motivo $A+I$ es nilpotente con índice de nilpotencia 2. Además $A+I$ es tal que $\ker(A+I)^2 = F^3$ y $\ker(A+I) = S_{-1}$.

Al calcular los autovectores de $A$ obtenemos que $\lambda = -1$ es un autovalor de multiplicidad geométrica 2 y multiplicidad algebraica 3.

Los autovectores obtenidos son:

$$\mathbf{v_1} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{v_2} = \begin{pmatrix} 0 \\ -3 \\ 1 \end{pmatrix},$$

y para estos vectores los sistemas $(A+I)\mathbf{x} = \mathbf{v_1}$ y $(A+I)\mathbf{x} = \mathbf{v_2}$ son inconsistentes, tal como lo puede verificar el lector.

Procederemos en consecuencia a resolver el sistema $(A+I)\mathbf{x} = \alpha_1\mathbf{v_1} + \alpha_2\mathbf{v_2}$, para $\alpha_i \in F$. Esto es,

$$\begin{aligned} -3y - 9z &= \alpha_1 \\ 6y + 18z &= -3\alpha_2 \\ -2y - 6z &= \alpha_2 \end{aligned}$$

y para que este sistema sea consistente es necesario que se cumpla $-2\alpha_1 = -3\alpha_2$, por lo que tomamos $\alpha_1 = 3$ y $\alpha_2 = 2$. De esta manera, la ecuación a resolver es $-3y - 9z = 3$ de la cual obtenemos que

$$y = -\frac{3+9z}{3}.$$

En consecuencia, un vector solución del sistema es

$$u = \begin{pmatrix} 0 \\ -4 \\ 1 \end{pmatrix}$$

el cual es el vector generalizado correspondiente al vector $3\mathbf{v_1}+2\mathbf{v_2} = (-3,6,2)^t$, la base buscada es:

$$\beta = \left\{ \begin{pmatrix} 3 \\ -6 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ -4 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\}$$

y la forma canónica de Jordan asociada a la matriz $A$ es:

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

A continuación veremos el procedimiento a seguir para el caso general, esto es: $A$ tiene varios autovalores defectivos.

### 9.5.3. Varios autovalores defectivos

En el caso general, contamos con una matriz $A \in M_{n\times n}(F)$ con $s$ autovalores distintos por lo que el polinomio característico es de la forma

$$p_A(\lambda) = (\lambda - a_1)^{n_1}(\lambda - a_2)^{n_2} \cdots (\lambda - a_s)^{n_s}$$

con $n_1 + n_2 + \cdots + n_s = n$. Por otra parte, ya que el polinomio mínimo y el polinomio característico tiene los mismos factores, entonces el polinomio mínimo tiene la siguiente forma:

$$m_A(\lambda) = (\lambda - a_1)^{r_1}(\lambda - a_2)^{r_2} \cdots (\lambda - a_s)^{r_s}.$$

Utilizando lo anterior tenemos una manera expresar la matriz $A$ como una matriz diagonal por bloques, tal que para cada autovalor $a_i$ se tiene un bloque

de la forma

$$MJ(a_i) = \begin{pmatrix} BJ_{t_1}(a_i) & 0 & 0 & \cdots & 0 \\ 0 & BJ_{t_2}(a_i) & 0 & \cdots & 0 \\ 0 & 0 & BJ_{t_3}(a_i) & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \cdots & BJ_{t_{s_i}}(a_i) \end{pmatrix}$$

donde $t_1 = r_1$ y $t_2, t_3, \ldots, t_{s_i} \leq r_i$ con $t_1 + t_2 + \cdots + t_{s_i} = n_i$. Además se cumple que en la base $\beta$ la matriz $A$ se describe como:

$$[A]_\beta = \begin{pmatrix} MJ(a_1) & 0 & 0 & \cdots & 0 \\ 0 & MJ(a_2) & 0 & \cdots & 0 \\ 0 & 0 & MJ(a_3) & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \cdots & MJ(a_s) \end{pmatrix}. \tag{9.12}$$

En general se cumple el siguiente resultado

**Teorema 9.38.** Sea $A \in M_{n\times n}(F)$ una matriz tal que

$$p_A(\lambda) = (\lambda - a_1)^{n_1}(\lambda - a_2)^{n_2} \cdots (\lambda - a_s)^{n_s}$$

y

$$m_A(\lambda) = (\lambda - a_1)^{r_1}(\lambda - a_2)^{r_2} \cdots (\lambda - a_s)^{r_s}.$$

Entonces $A$ tiene una matriz canónica de Jordan asociada de la forma (9.12).

Para fijar ideas consideraremos el siguiente par de ejemplos:

**Ejemplo 9.39.** En este ejemplo trabajaremos con una matriz con dos autovalores distintos y ambos defectivos. En este caso tratamos cada valor defectivo como en el procedimiento 9.10.

Tomamos la matriz $A$ dada por

$$\begin{pmatrix} 3 & -2 & 4 \\ 2 & -4 & 14 \\ 1 & -3 & 9 \end{pmatrix}$$

y determinemos su Forma Canónica de Jordan asociada. Para ello, el primer paso será calcular los autovalores. Con este fin tenemos que

$$A - \lambda I = \begin{pmatrix} 3-\lambda & -2 & 4 \\ 2 & -4-\lambda & 14 \\ 1 & -3 & 9-\lambda \end{pmatrix}$$

Por lo que el polinomio característico es: $p_A(\lambda) = -(\lambda-2)(\lambda-3)^2$. Es un ejercicio para el lector comprobar que $m_A = (\lambda-2)(\lambda-3)^2$. De lo anterior vemos que los autovalores de la matriz son $\lambda_1 = 2$ y $\lambda_2 = 3$, este último con multiplicada algebraica dos. Además del polinomio minimal observamos que la Forma Canónica de Jordan Asociada a la matriz $A$ tiene un bloque de la forma $BJ_1(3)$ y un bloque del a forma $BJ_2(3)$. Teniendo de esta manera que la matriz de Jordan es:

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

A continuación verificaremos esta afirmación calculando los autovectores y la base $\beta$ correspondiente.

Primero consideramos $\lambda_1 = 2$

$$A - 2I = \begin{pmatrix} 1 & -2 & 4 \\ 2 & -6 & 14 \\ 1 & -3 & 7 \end{pmatrix}$$

y el sistema de ecuaciones obtenido es:

$$\begin{aligned} x - 2y + 4z &= 0 \\ 2x - 6y + 14z &= 0 \\ x - 3y + 7z &= 0 \end{aligned}$$

obteniendo como solución al vector $\mathbf{v_1} = (2, 3, 1)^t$.

Ahora realizamos el mismo procedimiento para el autovalor $\lambda_2 = 3$. Consideramos la matriz

$$A - 3I = \begin{pmatrix} 0 & -2 & 4 \\ 2 & -7 & 14 \\ 1 & -3 & 6 \end{pmatrix}$$

de donde el sistema de ecuaciones es:

$$\begin{aligned} -2y + 4z &= 0 \\ 2x - 7y + 14z &= 0 \\ x - 3y + 6z &= 0 \end{aligned}$$

y la solución obtenida es $\mathbf{v_2} = (0, 2, 1)^t$. Note que este autovalor es defectivo y en consecuencia la matriz $A$ no es diagonalizable. El siguiente paso es determinar el vector generalizado correspondiente y con ello la matriz de Jordan asociada. Para ello seguimos el procedimiento dado en (9.10), es decir resolveremos el

sistema $(A-3I)\mathbf{v_{2,g}} = \mathbf{v_2}$ explícitamente tendremos:

$$\begin{pmatrix} 0 & -2 & 4 \\ 2 & -7 & 14 \\ 1 & -3 & 6 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}$$

De donde la solución es $\mathbf{v_{2,g}} = (1,2,1)^t$ el cual es el autovector generalizado para $\lambda_2 = 3$. Por lo tanto, la base

$$\beta = \left\{ \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \right\}$$

es tal que

$$[A]_\beta^\beta = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{pmatrix}.$$

la cual es la matriz de Jordan asociada a $A$.

**Ejemplo 9.40.** Para $A$ dada por

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

se cumple que $p_A(\lambda) = (\lambda-2)^2(\lambda-4)^3$. Por otro lado, se verifica que $m_A(\lambda) = p_A(\lambda)$ y así la matriz de Jordan Asociada a $A$ tiene un bloque $BJ_2(2)$ y $BJ_3(4)$. En consecuencia, la matriz de Jordan es

$$J = \begin{pmatrix} 2 & 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 4 & 1 & 0 \\ 0 & 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & 0 & 4 \end{pmatrix}.$$

Para verificar la información tendremos que calcular los valores y autovectores.

Para $\lambda_1 = 2$ tenemos

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

El sistema obtenido es:

$$\begin{aligned} y - z &= 0 \\ 2z + \kappa &= 0 \\ 2z + \kappa &= 0 \\ 2\kappa + \omega &= 0 \\ \omega &= 0. \end{aligned}$$

Siendo una solución $v_1 = (1, 0, 0, 0, 0)^t$. El siguiente paso es calcular el vector generalizado para $\mathbf{v_1}$. Para ello seguimos el procedimiento dado en (9.10). Esto es, resolvemos el sistema $(A - 2I)\mathbf{x} = \mathbf{v_1}$ el cual es

$$\begin{aligned} y - z &= 1 \\ 2z + \kappa &= 0 \\ 2z + \kappa &= 0 \\ 2\kappa + \omega &= 0 \\ \omega &= 0, \end{aligned}$$

y donde una solución es $\mathbf{v_{12}} = (1, 1, 0, 0, 0)$.

Ahora para el autovalor $\lambda_2 = 4$ se tiene

$$A - 4I = \begin{pmatrix} -2 & 1 & 1 & 0 & 0 \\ 0 & -2 & 2 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

y el sistema asociado es:

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

obteniendo como solución $\mathbf{v_2} = (1, 1, 1, 0, 0)^t$.

De este modo, es necesario determinar los distintos autovectores generalizados para $\mathbf{v_2}$ y para ello utilizaremos el procedimiento dado en (9.10). Es decir,

vamos a resolver $(A-4I)\mathbf{v_{21}} = \mathbf{v_2}$, $(A-4I)\mathbf{v_{22}} = \mathbf{v_{21}}$, $(A-4I)\mathbf{v_{23}} = \mathbf{v_{22}}$. Y así la base sería.

$$\beta = \{\mathbf{v_1}, \mathbf{v_{11}}, \mathbf{v_2}, \mathbf{v_{21}}, \mathbf{v_{22}}\}$$

Entonces resolvemos el sistema $(A-4I)\mathbf{x} = \mathbf{v_2}$, esto es

$$\begin{aligned} -2x + y + z &= 1 \\ -2y + 2z + \kappa &= 1 \\ \kappa &= 1 \\ \omega &= 0 \end{aligned}$$

y la solución al sistema es: $\mathbf{v_{21}} = (1,1,1,1,0)^t$, este es nuestro primer vector generalizado para $\mathbf{v_2}$.

El siguiente paso es resolver el sistema $(A-4I)\mathbf{x} = \mathbf{v_{21}}$. Obteniendo el sistema

$$\begin{aligned} -2x + y + z &= 1 \\ -2y + 2z + \kappa &= 1 \\ \kappa &= 1 \\ \omega &= 1 \end{aligned}$$

y la solución al sistema es: $\mathbf{v_{22}} = (1,1,1,1,1)^t$.

**Corolario 9.41.** Sea $A \in M_{n\times n}(F)$ una matriz tal que

$$p_A(\lambda) = (\lambda - a_1)^{n_1}(\lambda - a_2)^{n_2} \cdots (\lambda - a_s)^{n_s}$$

y

$$m_A(\lambda) = (\lambda - a_1)(\lambda - a_2) \cdots (\lambda - a_s).$$

Entonces $A$ es diagonalizable.

A continuación mostramos un ejemplo de este caso.

**Ejemplo 9.42.** Tomemos a $A$ como la matriz definida por

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

Para este caso el polinomio característico es: $p_A(\lambda) = (\lambda-1)(\lambda-2)^2$, por lo que los autovalores son: $a_1 = 1$ y $a_2 = 2$ de multiplicidad algebraica 1 y dos respectivamente. Por otro lado su polinomio minimal es: $m_A(\lambda) = (\lambda-1)(\lambda-2)$. Así, por lo descrito anteriormente tendremos que la forma Canónica de Jordan $J$ asociada a la matriz $A$ tiene un bloque del tipo $BJ_1(1)$ y un bloque de tipo

$BJ_1(2)$ además de tener un tercer bloque de la forma $BJ_t(2)$ con $t \leq 1$. En conclusión $J$ es diagonal, tal y como comprobaremos a continuación. Para ello determinaremos los autovectores asociados:

Para $\lambda = 1$ tendremos que

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

obteniendo el sistema

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

y donde una solución es:

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

Ahora consideramos el autovalor $\lambda = 2$ y tendremos:

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

y el sistema correspondiente es:

$$x - y - z = 0$$

por lo que la solución es:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} y + z \\ y \\ z \end{pmatrix} = y \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + z \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

En consecuencia, los autovectores obtenidos determinan una base y así la matriz es diagonalizable tal y como se había establecido. Además se tiene que si definimos como $N$ a la matriz cuyas columnas corresponden a los autovectores

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

Entonces se cumple la siguiente igualdad

$$J = N^{-1} A N,$$

donde

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

### 9.5.4. Forma Canónica de Jordan para endomorfismos

Consideremos $\varphi : U \to U$ un endomorfismo para un espacio vectorial $U$ sobre $F$. Suponga que $U$ es $n$- dimensional. Sea $\alpha = \{\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_n}\}$ una base para $U$, entonces el endomorfismo $\varphi$ tiene una matriz asociada $[\varphi]_\alpha^\alpha$.

Ahora bien, si hacemos $A := [\varphi]_\alpha^\alpha$ entonces definimos $p_\varphi(\lambda) = p_A(\lambda)$. Note que el polinomio característico del endomorfismo $\varphi$ no depende de la base de la matriz asociada $A$ y por lo tanto no depende de la base $\beta$, vea Ejercicio 9.54.

De manera similar definimos los autovalores asociados al endomorfismo $\varphi$ como las raíces del polinomio característico de $p_A(\lambda)$. Similarmente, si $\lambda_0$ es autovalor para $\varphi$ entonces $\mathbf{u} \in U$ es un autovector para $\varphi$ asociado a $\lambda_0$ si $(\varphi - \lambda_0 Id_U)\mathbf{u} = \mathbf{0}$. Mas aún, $\mathbf{u}$ es un autovector de $\varphi$ si $([\varphi]_\alpha^\alpha - \lambda_0 I)[\mathbf{u}]_\alpha^\alpha = \mathbf{0}$. Por tal motivo el estudio realizado para el caso de las matrices, aplica similarmente para cualquier endomorfismo $\varphi$.

**Ejemplo 9.43.** Sea $\varphi : \mathbb{R}^3 \to \mathbb{R}^3$ un endomorfismo definido por

$$\varphi(x, y, z) = (-y, 4x + 4y, 3z).$$

Determinaremos los valores y autovectores asociados a $\varphi$ así como una base $\beta$ para $\mathbb{R}^3$ tal que $[\varphi]_\beta^\beta$ es una matriz en Forma Canónica de Jordan.

Primero, observamos que

$$[\varphi]_e^e = \begin{pmatrix} 0 & 4 & 0 \\ -1 & 4 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

Entonces $p_\varphi(\lambda) = -(\lambda - 2)^2(\lambda - 3)$. Calculemos los autovectores asociados. Para $\lambda_1 = 2$

$$([\varphi]_e^e - 2I) = \begin{pmatrix} -2 & 4 & 0 \\ -1 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

obteniendo el sistema

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

y la solución es $(x, y, z) = (2y, y, z)^t$ y así $\mathbf{v_1} = (2, 1, 0)^t$. Por lo tanto el autovalor $\lambda_1 = 2$ es defectivo y buscamos su autovalor generalizado. Para ello resolvemos

$$\begin{aligned} -2x + 4y &= 2 \\ -x + 2y &= 1 \\ z &= 0 \end{aligned}$$

de lo cual $\mathbf{v_{11}} = (3, 2, 0)^t$.

Ahora hacemos lo correspondiente con $\lambda_2 = 3$

$$([\varphi]_e^e - 3I) = \begin{pmatrix} -3 & 4 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

y el sistema a resolver es

$$\begin{aligned} -3x + 4y &= 0 \\ -x + y &= 0. \end{aligned}$$

A partir del sistema vemos que la solución es $(0, 0, z)$ por lo que $\mathbf{v_2} = (0, 0, 1)$.

En conclusión, la forma canónica de Jordan asociada a $\varphi$ es

$$[\varphi]_\beta^\beta = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

donde la base es

$$\beta = \left\{ \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \right\}$$

Por otra parte, de la forma canónica de Jordan obtenida podemos concluir que el polinomio minimal asociado a $\varphi$ es

$$m_\varphi(\lambda) = -p_\varphi(\lambda) = (\lambda - 2)^2(\lambda - 3)$$

tal y como el lector puede verificarlo directamente.

***Palabras clave:*** *polinomio mínimo, teorema de Cayley-Hamilton, endomorfismo triangulable, endomorfismo nilpotente, bloque de Jordan, forma canónica de Jordan.*

## 9.6. Ejercicios

**Ejercicio 9.44.** Demuestra que si $A \in M_{n\times n}(F)$ y $Adj(A)$ denota la matriz adjunta asociada a la matriz $A$, entonces demuestre que $(A)Adj(A) = |A|I$.

**Ejercicio 9.45.** Sea $q(x) \in F[x]$ y $A, B \in M_{n\times n}(F)$ dos matrices similares, es decir $A = C^{-1}BC$ para una matriz invertible $C$. Demuestre que $q(A) = C^{-1}q(B)C$.

**Ejercicio 9.46.** Demuestre que si $\beta = \{\varphi^{r-1}(u), \varphi^{r-2}(u), \ldots, \varphi(u), u\}$ es la base dada en la Proposición 9.20 y $W$ el subespacio vectorial generado por $\beta$. Entonces demuestra que:

1. La restricción de $\varphi$ al subespacio vectorial $W$ es un endomorfismo de $W$.
2. El vector $\varphi^{r-1}(u)$ es un autovector del endomorfismo $\varphi|_W$.

**Ejercicio 9.47.** Demuestra que si $\varphi : U \to U$ es un endomorfismo nilpotente con índice de nilpotencia $r$, entonces $m_\varphi(\lambda) = \lambda^r$.

**Ejercicio 9.48.** Dada la existencia del polinomio 9.2 demuestra que existe un polinomio de grado mínimo que se anula en $A$.

**Ejercicio 9.49.** Demuestra que el bloque de Jordan $BJ_n(0)$ es una matriz nilpotente.

**Ejercicio 9.50.** Demuestra la Proposición 9.21.

**Ejercicio 9.51.** Demuestra la Proposición 9.24.

**Ejercicio 9.52.** Demuestra que si $A \in M_{n\times n}(F)$ tiene un único autovalor $a$, entonces $aI - A$ es nilpotente.

**Ejercicio 9.53.** Demuestre que con la base $\beta$ obtenida en la Ecuación 9.9 la matriz $[A]^\beta_\beta$ corresponde a la matriz de la Ecuación 9.7.

**Ejercicio 9.54.**

Sea $\varphi : U \to U$ un endomorfismo para un espacio vectorial $U$ sobre $F$. Suponga que $U$ es $n$- dimensional. Sean $\alpha$ y $\alpha'$ dos bases para $U$. Demuestra que el polinomio característico para $\varphi$ no depende de la base elegida.

**Ejercicio 9.55.** Sean $A, B \in M_{n\times n}(F)$ dos matrices similares, es decir $A = N^{-1}BN$ para alguna matriz invertible $N$. Demuestra que $A^m = N^{-1}B^mN$.

**Ejercicio 9.56.** Demuestra que si $A \sim B$, entonces $m_A(\lambda) = m_B(\lambda)$. Ayuda: Utiliza el hecho de que $p_A(\lambda) = p_B(\lambda)$ y que el polinomio mínimo divide al polinomio característico, además del Ejercicio 9.55 .

**Ejercicio 9.57.** Sea $V$ un espacio vectorial de dimensión $n$ sobre un campo $F$. Demuestra que un operador $T : V \to V$ es invertible si, y solo si, no tiene a $\lambda = 0$ como valor propio asociado. Concluye que el polinomio característico de $T$ tiene término constante distinto de cero.

**Ejercicio 9.58.** Demuestra que si $B$ una matriz invertible entonces $B^{-1}$ se puede escribir como un polinomio en términos de $B$. Ayuda: Utiliza el Ejercicio 9.57 y el Teorema de Cayley Hamilton.

**Ejercicio 9.59.** Sean $a, b, c \in \mathbb{R}$ tales que $a \neq c$. Determina los valores de $b$ para los cuales la siguiente matriz es diagonalizable.

$$\begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

**Ejercicio 9.60.** Sea $A$ la matriz definida por

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

Calcula lo siguiente:

1. Polinomio característico.
2. Polinomio minimal.
3. Encuentra una matriz invertible tal que $N^{-1}AN$ sea diagonal.

**Ejercicio 9.61.**

Sea $A = \begin{pmatrix} -1 & -3 & -9 \\ 0 & 5 & 18 \\ 0 & -2 & -7 \end{pmatrix}$

1. Determina la forma canónica de Jordan de la matriz.
2. Describe el procedimiento que debes seguir para calcular la matriz invertible $M$ tal que $MAM^{-1} = J$
3. Calcula la matriz $M$ mencionada en el inciso anterior.

# 10

# Introducción a la Teoría de Códigos

La *teoría de códigos* es el estudio de métodos matemáticos para la transmisión de datos de forma eficiente y confiable. Se relaciona con disciplinas como la teoría de la información, las matemáticas puras, la ingeniería electrónica y las ciencias computacionales, y algunos de sus usos principales son la compresión de datos, criptografía, corrección de errores y, recientemente, códigos en red.

El escenario abstracto consiste en un *emisor* que envía un *mensaje* a un *receptor*. Debido a que en cualquier canal de transmisión existe la posibilidad de el mensaje sea distorsionado (por efectos externos), el objetivo es diseñar un método que permita detectar o corregir los errores en la transmisión. Llamamos *ruido* la distorsión que altera al mensaje en el canal de transmisión.

Un *codificador* es una función que añade redundancia a los mensajes. Un mensaje con redundancia se llama *palabra código*, y la colección de todas éstas es el *código*. El *procesador de errores* es el que se encarga de analizar las palabras código para detectar las alteraciones.

Inspirados en la comunicación digital moderna, podemos suponer que el mensaje a transmitir es una *cadena de bits*, es decir, una cadena de ceros y unos. La transmisión de *códigos en bloque*, consiste en dividir los mensajes en subcadenas de bits de longitud fija, las cuales son llamadas los *bloques* del mensaje. El codificador añade bits de redundancia a cada uno de estos bloques, los cuales son transmitidos y analizados independientemente por el procesador de errores.

## 10.1. Definiciones básicas

**Definición 10.1 (código lineal binario).** Un *código lineal binario* es un *subespacio* del *espacio vectorial* $\mathbb{Z}_2^n$.

Como en este capítulo el campo subyacente de nuestros espacios vectoriales será $\mathbb{Z}_2$, los únicos escalares son 0 y 1, por lo que la multiplicación escalar es una operación trivial:

$$0v = 0, \quad \text{y} \quad 1v = v, \quad \forall v \in \mathbb{Z}_2^n.$$

Por lo tanto, la Definición 10.1 puede reformularse de la siguiente manera: un **código lineal binario** es un subconjunto $\mathbf{C} \subseteq \mathbb{Z}_2^n$ tal que:

$$\text{para cualquier } w_1, w_2 \in \mathbf{C}, \text{ la suma } w_1 + w_2 \text{ también pertenece a } \mathbf{C}.$$

A los elementos de $\mathbf{C}$ se les llama **palabras código**. En estas notas usamos simplemente la palabra **código** para referirnos a un código lineal binario.

**Ejemplo 10.2.** Sea $\mathbf{C} = \{(0,0,0),(1,1,1)\} \subseteq \mathbb{Z}_2^3$. Es fácil verificar que $\mathbf{C}$ es un subespacio:

$$\begin{aligned}(0,0,0)+(0,0,0) &= (0,0,0) \in \mathbf{C},\\ (0,0,0)+(1,1,1) &= (1,1,1) \in \mathbf{C},\\ (1,1,1)+(1,1,1) &= (0,0,0) \in \mathbf{C}.\end{aligned}$$

Por lo tanto, $\mathbf{C}$ es un código.

**Ejemplo 10.3.** Sea $\mathbf{C} = \{(0,0,0),(1,1,1),(1,0,0)\} \subseteq \mathbb{Z}_2^3$. En este caso,

$$(1,1,1)+(1,0,0) = (0,1,1) \notin \mathbf{C}.$$

Por lo tanto, $\mathbf{C}$ **no** es un código.

**Ejemplo 10.4.** Sea $\mathbf{C}$ el subconjunto de $\mathbb{Z}_2^4$ de todos los vectores que tienen un número par de unos. Explícitamente:

$$\mathbf{C} = \left\{ \begin{array}{llll} (0,0,0,0), & (1,1,0,0), & (1,0,1,0), & (1,0,0,1), \\ (0,1,1,0), & (0,1,0,1), & (0,0,1,1), & (1,1,1,1) \end{array} \right\}.$$

Aunque es bastante laborioso, es posible demostrar con cálculos directos que $\mathbf{C}$ es un subespacio: es necesario comprobar que $w_1+w_2 \in \mathbf{C}$ para toda $w_1, w_2 \in \mathbf{C}$. En general, el subconjunto de todos los vectores de $\mathbb{Z}_2^n$ que tienen un número par de unos es un código; demostrar esto se deja como ejercicio.

**Definición 10.5 (Longitud, rango y tasa).** Sea $\mathbf{C} \subseteq \mathbb{Z}_2^n$ un código.

1. La *longitud* de $\mathbf{C}$ es la dimensión de $\mathbb{Z}_2^n$.
2. El *rango* de $\mathbf{C}$ es la dimensión de $\mathbf{C}$.
3. La *tasa* de $\mathbf{C}$ es la razón del rango entre la longitud de $\mathbf{C}$.

**Ejemplo 10.6.** Sea $\mathbf{C} = \{(0,0,0),(1,1,1)\} \subseteq \mathbb{Z}_2^3$. La longitud de $\mathbf{C}$ es $3 = \dim(\mathbb{Z}_2^3)$. Debido a que $\{(1,1,1)\}$ es una base de $\mathbf{C}$, el rango de $\mathbf{C}$ es $1 = \dim(\mathbf{C})$. Por lo tanto, la tasa de $\mathbf{C}$ es $\frac{1}{3}$.

**Ejemplo 10.7.** Sea $\mathbf{C}$ el subconjunto de $\mathbb{Z}_2^4$ de todos los vectores que tienen un número par de unos. La longitud de $\mathbf{C}$ es $4 = \dim(\mathbb{Z}_2^4)$. Demostraremos que

$$B = \{(1,0,0,1),(0,1,0,1),(0,0,1,1)\}$$

es una base de $\mathbf{C}$ comprobando primero que todo vector distinto de cero en $\mathbf{C}$ puede escribirse como la suma de vectores en $B$:

$$(1,1,0,0) = (1,0,0,1) + (0,1,0,1), \qquad (1,0,1,0) = (1,0,0,1) + (0,0,1,1),$$
$$(1,0,0,1) = (1,0,0,1), \qquad (0,1,1,0) = (0,1,0,1) + (0,0,1,1),$$
$$(0,1,0,1) = (0,1,0,1), \qquad (0,0,1,1) = (0,0,1,1),$$
$$(1,1,1,1) = (1,0,0,1) + (0,1,0,1) + (0,0,1,1).$$

Además, el conjunto $B$ es linealmente independiente porque si

$$\alpha_1(1,0,0,1) + \alpha_2(0,1,0,1) + \alpha_3(0,0,1,1) = (0,0,0,0),$$

entonces $\alpha_1 = 0$, $\alpha_2 = 0$ y $\alpha_3 = 0$. Esto demuestra que el rango de $\mathbf{C}$ es $3 = \dim(\mathbf{C})$, y, por lo tanto, la tasa de $\mathbf{C}$ es $\frac{3}{4}$.

## 10.2. Matrices generadoras y verificadoras

**Definición 10.8 (matriz generadora).** Sea $\mathbf{C}$ un código de longitud $n$ y rango $k$. La *matriz* de $n \times k$ definida como

$$G = \begin{pmatrix} \uparrow & \uparrow & & \uparrow \\ b_1 & b_2 & \ldots & b_k \\ \downarrow & \downarrow & & \downarrow \end{pmatrix}$$

donde las columnas $\{b_1, b_k, ..., b_k\}$ forman una base de $\mathbf{C}$, se llama *matriz generadora* de $C$. La matriz generadora de $\mathbf{C}$ representa una transformación lineal

$$T_G : \mathbb{Z}_2^k \longrightarrow \mathbb{Z}_2^n$$

cuya imagen coincide con $\mathbf{C}$.

Sea $\mathbf{C}$ un código de longitud $n$ y rango $k$. Por definición, una matriz generadora de $\mathbf{C}$ es una matriz $G$ de $n \times k$ cuyas columnas son una base de $\mathbf{C}$. Decimos que $G$ está en **forma estándar** si sus primeros $k$ renglones son la base canónica de $\mathbb{Z}_2^k$; es decir,

$$G = \begin{pmatrix} I_k \\ M \end{pmatrix},$$

donde $I_k$ es la matriz identidad de $k \times k$ y $M$ es una matriz de $(n-k) \times k$.

Recordemos que una matriz generadora $G$ de $\mathbf{C}$ representa una transformación lineal $\mathbb{Z}_2^k \to \mathbb{Z}_2^n$, donde $n$ y $k$ son la longitud y el rango de $\mathbf{C}$, respectivamente. Esta situación se puede interpretar de la siguiente forma. Supongamos que los vectores de $\mathbb{Z}_2^k$ son **mensajes** que queremos transmitir. Decimos que **codificamos** un mensaje $v \in \mathbb{Z}_2^k$ al multiplicarlo por la matriz generadora $G$. Observemos que el vector $Gv$ resultante de la multiplicación es una palabra código de $\mathbf{C}$.

**Observación 10.9.** Los vectores mensaje tienen $k$ coordenadas porque son elementos de $\mathbb{Z}_2^k$ mientras que las palabras código tienen $n$ coordenadas porque son elementos de $\mathbf{C} \subseteq \mathbb{Z}_2^n$. A las $n-k$ coordenadas extra en cada palabra código las llamamos **bits de redundancia**.

**Ejemplo 10.10 (Triple Repetición).** Consideremos el código $\mathbf{C} = \{(0,0,0),(1,1,1)\}$ de longitud $n = 3$ y rango $k = 1$. Como $\{(1,1,1)\}$ es la única base de $\mathbf{C}$, la matriz generadora de $\mathbf{C}$ es

$$G = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Esta matriz representa una transformación lineal $\mathbb{Z}_2 \to \mathbb{Z}_2^3$. En esta situación, los vectores mensaje son elementos de $\mathbb{Z}_2 = \{0,1\}$ (cada vector tiene sólo una coordenada). A continuación codificamos los posibles mensajes:

$$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}(1) = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \in \mathbf{C}$$

$$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}(0) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \in \mathbf{C}.$$

Observemos que cada mensaje se codifica como una palabra código haciendo la **triple repetición** del contenido del mensaje.

**Ejemplo 10.11 (Bit de Paridad).** Sea $\mathbf{C}$ el subconjunto de $\mathbb{Z}_2^4$ de todos los vectores que tienen un número par de unos. Recordemos que

$$B = \{(1,0,0,1),(0,1,0,1),(0,0,1,1)\}$$

es una base de $\mathbf{C}$. Por lo tanto, una matriz generadora de $\mathbf{C}$ es

$$G = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

De hecho, $G$ está en su forma estándar porque los primeros 3 renglones forman a la matriz identidad. Debido a que $\mathbf{C}$ tiene rango $k = 3$, los vectores mensaje son elementos de $\mathbb{Z}_2^3$. Ilustramos cómo se codifican algunos mensajes con los

siguientes ejemplos:

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

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

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

En general, cualquier mensaje $(x_1, x_2, x_3) \in \mathbb{Z}_2^3$ se codifica como

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_1 + x_2 + x_3 \end{pmatrix}.$$

Como podemos observar, las palabra código correspondiente a un mensaje es igual al mensaje más un bit de redundancia (que toma valor igual a la suma de las coordenadas) al que llamamos el **bit de paridad**. En esencia, el bit de paridad es 0 o 1 dependiendo si el mensaje tiene un número par o impar de unos, respectivamente.

**Definición 10.12 (Matriz Verificadora).** Sea $\mathbf{C}$ un código de longitud $n$ y rango $k$. Una *matrix verificadora*, o de *comprobación*, de $\mathbf{C}$ es una matriz $H$ de $(n-k) \times n$ con la siguiente propiedad:

$$Hw = \vec{\mathbf{0}} \text{ si y sólo si } w \in \mathbf{C}.$$

Una matriz verificadora tiene la forma estándar si sus últimas $(n-k)$ columnas son la base canónica de $\mathbb{Z}_2^{n-k}$.

**Observación 10.13.** Una matriz verificadora $H$ define una transformación lineal

$$H : \mathbb{Z}_2^n \longrightarrow \mathbb{Z}_2^{n-k},$$

cuyo kernel coincide con $\mathbf{C}$; es decir, $\ker(H) = \mathbf{C}$.

**Teorema 10.14 (Relación con la matriz generadora estándar).** Sea $\mathbf{C}$ un código de longitud $n$ y rango $k$. La matriz

$$G = \begin{pmatrix} I_k \\ M \end{pmatrix}$$

es una matriz generadora estándar de $\mathbf{C}$ si y sólo si

$$H = \begin{pmatrix} M & I_{n-k} \end{pmatrix}$$

es una matriz verificadora estándar de $\mathbf{C}$.

**Ejemplo 10.15 (Bit de paridad).** Sea $\mathbf{C}$ el subconjunto de $\mathbb{Z}_2^4$ de todos los vectores que tienen un número par de unos y consideremos su matriz generadora estándar

$$G = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

En este caso, $M$ es la siguiente matriz de $1 \times 3$:

$$M = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}.$$

Por lo tanto, la matriz verificadora de $\mathbf{C}$ es

$$H = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix},$$

ya que en este caso $I_{n-k}$ es la matriz identidad de $1 \times 1$. Observemos que, para cualquier $(x_1, x_2, x_3, x_4) \in \mathbb{Z}_2^4$ tenemos que

$$\begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = (x_1 + x_2 + x_3 + x_4) = 0 \text{ si y sólo si } \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \in \mathbf{C},$$

ya que la suma $x_1 + x_2 + x_3 + x_4$ es cero si y sólo si $(x_1, x_2, x_3, x_4)$ tiene un número par de unos.

## 10.3. Detección y corrección de errores

**Definición 10.16 (Peso).** Si $v \in \mathbb{Z}_2^n$ es cualquier vector, el *peso*, denotado como $|v|$, es el número de unos en las coordenadas de $v$.

**Ejemplo 10.17.** Consideremos el peso de los siguientes vectores:

1. $|(0, 0, 0, 0)| = 0$.
2. $|(1, 0, 1, 1, 0, 0)| = 3$.

3. $|(1,1,1)| = 3$.
4. $|(0,0,1,0,0,0,1)| = 2$.

**Definición 10.18 (Distancia de Hamming).** Sean $u$ y $v$ dos vectores de $\mathbb{Z}_2^n$. La *distancia de Hamming* entre $u$ y $v$ se define como el peso de la suma $u+v$; en símbolos:

$$d(u,v) = |u+v|.$$

**Ejemplo 10.19.** Si $u = (1,1,0)$ y $v = (1,0,1)$, entonces

$$d(u,v) = |u+v| = |\,(1,1,0) + (1,0,1)\,| = |(0,1,1)| = 2.$$

**Teorema 10.20 (Distancia de Hamming).** La distancia de Hamming cumple las siguientes propiedades para toda $u, v, w \in \mathbb{Z}_2^n$:

**(1) Es no negativa**: $d(u,v) \geq 0$, con igualdad si y sólo si $u = v$.

**(2) Es simétrica:** $d(u,v) = d(v,u)$.

**(3) Cumple la desigualdad del triángulo**: $d(u,v) \leq d(u,w) + d(w,v)$.

**Demostración.**

**(1)** El número de unos de en $u+v$ es una cantidad no negativa, así que $d(u,v) = |u+v| \geq 0$. Además,

$$d(u,v) = 0 \;\Leftrightarrow\; u+v \text{ no tiene unos} \;\Leftrightarrow\; u+v = \vec{\mathbf{0}} \;\Leftrightarrow\; u = v.$$

**(2)** Esa propiedad se deduce de la conmutatividad de la suma de vectores:

$$d(u,v) = |u+v| = |v+u| = d(v,u).$$

**(3)** Supongamos que $u = (u_1, \ldots, u_n)$, $v = (v_1, \ldots, v_n)$ y $w = (w_1, \ldots, w_n)$. Observemos que $d(u,v)$ es igual al número de coordenadas donde $u$ y $v$ difieren. Esto lo escribimos formalmente como

$$d(u,v) = |\{i : u_i \neq v_i\}|\,.$$

Como $\{i : u_i \neq v_i\}$ es la unión disjunta de $\{i : u_i \neq v_i \text{ y } u_i \neq w_i\}$ y $\{i : u_i \neq v_i \text{ y } u_i = w_i\}$, tenemos que

$$d(u,v) = |\{i : u_i \neq v_i \text{ y } u_i \neq w_i\}| + |\{i : u_i \neq v_i \text{ y } u_i = w_i\}|\,.$$

Ahora,

$$\begin{aligned} |\{i : u_i \neq v_i \text{ y } u_i \neq w_i\}| &\leq |\{i : u_i \neq w_i\}| = d(u,w), \\ |\{i : u_i \neq v_i \text{ y } u_i = w_i\}| &\leq |\{i : w_i \neq v_i\}| = d(w,v). \end{aligned}$$

Por lo tanto, $d(u,v) \leq d(u,w) + d(w,v)$.

□

Sea $\mathbf{C} \subseteq \mathbb{Z}_2^n$ un código. Cuando se transmite una palabra código $w \in \mathbf{C}$ a través de un canal, y se produce **un error en el $i$-ésimo bit**, el vector que se recibe es

$$w' = w + e_i \in \mathbb{Z}_2^n$$

donde $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)$. Observemos que la distancia entre $w$ y $w'$ es

$$d(w, w') = |w + w'| = |w + w + e_i| = |\vec{\mathbf{0}} + e_i| = |e_i| = 1.$$

Similarmente, podemos demostrar que si se producen dos errores en bits distintos, la distancia entre la palabra código enviada y el vector recibido será 2. En general, tenemos la siguiente proposición.

**Proposición 10.21.** Sea $\mathbf{C} \subseteq \mathbb{Z}_2^n$ un código. Supongamos que se transmite una palabra código $w \in \mathbf{C}$ a través de un canal, y se producen $t \in \mathbb{N}$ errores en bits distintos. Entonces, la distancia entre $w$ y el vector que se recibe $w'$ es igual a $t$.

De acuerdo a estas observaciones, un código será más eficiente para detectar errores si sus palabras código se encuentran lejanas entre sí. Por ejemplo, si dos palabras código $v, w \in \mathbf{C}$ tienen distancia 1, basta con que se cometa un sólo error en el bit adecuado para que se transmita $w$ y se reciba $v$; en esta situación, es imposible que el destinatario identifique si se ha producido un error o no, ya que podría pensarse que se tansmitió correctamente la palabra código $v$.

Estas consideraciones sobre la distancia entre palabras código nos lleva a la siguiente definición.

**Definición 10.22 (Distancia Mínima).** Definimos la distancia mínima de un código $\mathbf{C}$ como

$$d(\mathbf{C}) = \text{mín}\{d(v, w) : v, w \in \mathbf{C}, v \neq w\}.$$

En general, mientras más grande sea la distancia mínima de un código, mayor será su capacidad para detectar errores. El siguiente teorema proporciona una forma equivalente y más sencilla de calcular la distancia mínima de un código.

**Teorema 10.23 (Distancia Mínima).** Sea $\mathbf{C}$ un código. La distancia mínima de $\mathbf{C}$ es igual a su *peso mínimo*:

$$d(\mathbf{C}) = \text{mín}\left\{|v| : v \in C, v \neq \vec{\mathbf{0}}\right\}.$$

**Ejemplo 10.24 (Triple repetición).** La distancia mínima del código $\mathbf{C} = \{(0,0,0), (1,1,1)\}$ es

$$d(\mathbf{C}) = \text{mín}\left\{|v| : v \in C, v \neq \vec{\mathbf{0}}\right\} = \text{mín}\{|(1,1,1)|\} = |(1,1,1)| = 3.$$

**Ejemplo 10.25 (Bit de paridad).** Sea $\mathbf{C}$ el subconjunto de $\mathbb{Z}_2^4$ de todos los vectores que tienen un número par de unos (es decir, el conjunto de vectores en $\mathbb{Z}_2^4$ con peso par). Entonces,

$$\mathbf{C} = \left\{ \begin{array}{cccc} (0,0,0,0), & (1,1,0,0), & (1,0,1,0), & (1,0,0,1), \\ (0,1,1,0), & (0,1,0,1), & (0,0,1,1), & (1,1,1,1) \end{array} \right\}.$$

Por inspección, vemos que el peso mínimo de una palabra código distinta de cero es 2. Por lo tanto,

$$d(\mathbf{C}) = 2.$$

Ahora estudiaremos dos teoremas sobre la detección y corrección de errores.

**Teorema 10.26 (Detección de Errores).** Sea $\mathbf{C}$ un código. Existe un procesador de errores para $\mathbf{C}$ que **detecte** todos los errores en $t \in \mathbb{N}$ bits distintos si y sólo si $d(\mathbf{C}) \geq t + 1$.

**Demostración.** Si $d(\mathbf{C}) \geq t + 1$, entonces el siguiente detector de errores $D$ puede detectar todos los errores en $t$ bits:

$$D(w') = \begin{cases} \text{No error}, & \text{si } w' \in \mathbf{C}, \\ \text{Error}, & \text{si } w' \notin \mathbf{C}, \end{cases}$$

donde $w'$ es el vector recibido. Si $H$ es la matriz verificadora de $\mathbf{C}$, el detector de errores $D$ también puede escribirse de la siguiente manera:

$$D(w') = \begin{cases} \text{No error}, & \text{si } Hw' = \vec{\mathbf{0}}, \\ \text{Error}, & \text{si } Hw' \neq \vec{\mathbf{0}}. \end{cases}$$

Si $d(\mathbf{C}) \leq t$, entonces existen dos palábras código $w, v \in \mathbf{C}$ tales que $d(w, v) \leq t$: así, al cometer $t$ errores o menos, es posible que se transmita $w$ y se reciba $v$, por lo cual es imposible que $D$ detecte el error. □

**Teorema 10.27 (Corrección de Errores).** Sea $\mathbf{C}$ un código. Existe un procesador de errores para $\mathbf{C}$ que **corrige** todos los errores en $t \in \mathbb{N}$ bits distintos si y sólo si $d(\mathbf{C}) \geq 2t + 1$.

**Demostración.** Si $d(\mathbf{C}) \geq 2t+1$, entonces el siguiente procesador de errores $P$ puede corregir todos los errores en $t$ bits:

$$P(w') = \begin{cases} w, & \text{si existe } w \in \mathbf{C} \text{ tal que } d(w, w') \leq t, \\ \text{Error}, & \text{en otro caso}, \end{cases}$$

donde $w'$ es el vector recibido. □

Los detectores y correctores de errores en los siguientes ejemplos son ilustraciones particulares de los detectores y correctores de errores descritos en los párrafos anteriores.

**Ejemplo 10.28 (Triple repetición).** Como el código $\mathbf{C} = \{(0,0,0), (1,1,1)\}$ tiene distancia mínima 3, existe un procesador de errores $D$ para $\mathbf{C}$ que detecta todos los errores en 2 bits distintos, pero no en 3 bits distintos. Este procesador de errores $D$ funciona como sigue: para cualquier vector $w' \in \mathbb{Z}_2^3$ recibido,

$$D(w') = \begin{cases} \text{No error} & \text{si } |w'| = 0 \text{ o } |w'| = 3, \\ \text{Error} & \text{si } |w'| = 1 \text{ o } |w'| = 2. \end{cases}$$

Además, existe un procesador de errores $P$ para $\mathbf{C}$ que corrige todos los errores en 1 bit que funciona como sigue: para cualquier $w' \in \mathbb{Z}_2^3$ recibido,

$$P(w') = \begin{cases} w = (1,1,1) & \text{si } d((1,1,1), w') \leq 1, \\ w = (0,0,0) & \text{si } d((0,0,0), w') \leq 1. \end{cases}$$

Equivalentemente, podemos escribir este procesador de errores como

$$P(w') = \begin{cases} w = (1,1,1) & \text{si } |w'| \geq 2, \\ w = (0,0,0) & \text{si } |w'| \leq 1. \end{cases}$$

Por ejemplo, si se recibe el vector $w' = (1,0,0)$, el procesador de errores asume que ocurrió un error en el primer bit y lo corrige como $w = (0,0,0)$. Por otro lado, si se recibe el vector $w' = (1,0,1)$, el procesador de errores asume que ocurrió un error en el segundo bit y lo corrige como $w = (1,1,1)$. Si ocurren errores en dos bits distintos, el procesador de errores corregirá el vector recibido equivocadamente. Observemos que en este caso el procesador $P$ no tiene la opción de "Error" porque siempre existe una palabra código $w \in \mathbf{C}$ tal que $d(w, w') \leq 1$.

**Ejemplo 10.29 (Bit de paridad).** Sea $\mathbf{C}$ el subconjunto de $\mathbb{Z}_2^4$ de todos los vectores con peso par. Como $d(\mathbf{C}) = 2$, existe un procesador de errores $D$ para $\mathbf{C}$ que detecta errores en 1 bit. Este procesador de errores funciona como sigue: para cualquier $w' \in \mathbb{Z}_2^3$ recibido,

$$D(w') = \begin{cases} \text{No error} & \text{si } |w'| \text{ es par}, \\ \text{Error} & \text{si } |w'| \text{ es impar}. \end{cases}$$

Este procesador no puede detectar errores en 2 bits distintos. Por ejemplo, si se transmite la palabra código $(1,0,1,0)$ y se producen errores en el primer y segundo bit, se recibe el vector $(0,1,1,0)$ el cual tiene peso par, por lo que $D$ no reporta error. Por otro lado, no existe ningún procesador de errores para $\mathbf{C}$ que corrija errores.

## 10.4. Códigos de Hamming

Los **códigos de Hamming** son una familia importante de códigos lineales binarios con varias propiedades excepcionales. Para cada número natural $r \geq 2$, el siguiente algoritmo describe la construcción de la matriz verificadora del código de Hamming $\text{Ham}(r)$ el cual tiene longitud $\mathbf{n = 2^r - 1}$ y rango $\mathbf{k = 2^r - r - 1}$.

**Definición 10.30 (algoritmo del código de Hamming). Matriz Verificadora de** $\text{Ham}(r)$:

- **Entrada**: Un número natural $r \geq 2$.

- **Procedimiento**: Construir una matriz $H_r$ cuyas columnas son todos los vectores distintos de cero del espacio $\mathbb{Z}_2^r$. La matriz $H_r$ debe tener $r$ renglones y $2^r - 1$ columnas.

- **Salida**: Una matriz verificadora $H_r$ de Ham$(r)$.

**Ejemplo 10.31 (Ham (3)).** Construiremos el código de Hamming con $r = 3$, el cual tiene longitud $n = 2^3 - 1 = 7$ y rango $k = 2^3 - 3 - 1 = 4$. El espacio vectorial $\mathbb{Z}_2^3$ tiene 7 vectores distintos de cero:

$$\begin{array}{llll} (1,0,0), & (0,1,0), & (0,0,1), & \\ (0,1,1), & (1,0,1), & (1,1,0), & (1,1,1). \end{array}$$

De acuerdo al Algoritmo 10.30, estos vectores forman las columnas de una matriz verificadora $H_3$ de Ham(3). Para construir esta matriz en su forma estándar, escribimos la base canónica en las últimas tres columnas de $H_3$; el orden del resto de las columnas de $H_3$ puede elegirse de manera arbitraria. Por ejemplo, la siguiente es una matriz verificadora $H_3$:

$$H_3 = \begin{bmatrix} 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \end{bmatrix}.$$

En este caso, la matriz $M$ del Teorema 10.14 que relaciona a la matriz verificadora con la generadora es la siguiente

$$M = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix}.$$

Por lo tanto, una matriz generadora $G_3$ de Ham (3) en forma estándar es

$$G_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix}$$

Por definición de matriz generadora, el conjunto

$$B = \{(1,0,0,0,1,1,0), (0,1,0,0,0,1,1), (0,0,1,0,1,1,1), (0,0,0,1,1,0,1)\} \subseteq \mathbb{Z}_2^7$$

es una base de Ham(3).

Por ejemplo, si queremos transmitir el mensaje $(1,0,0,1) \in \mathbb{Z}_2^4$, debemos

multiplicar por la matriz generadora

$$G\begin{pmatrix}1\\0\\0\\1\end{pmatrix}=\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\\1&0&1&1\\1&1&1&0\\0&1&1&1\end{bmatrix}\begin{pmatrix}1\\0\\0\\1\end{pmatrix}=\begin{pmatrix}1\\0\\0\\1\\0\\1\\1\end{pmatrix}.$$

Podemos comprobar que $(1,0,0,1,0,1,1) \in \mathbb{Z}_2^7$ efectivamente pertenece al código multiplicando por la matriz verificadora

$$Hw^T=\begin{bmatrix}1&0&1&1&1&0&0\\1&1&1&0&0&1&0\\0&1&1&1&0&0&1\end{bmatrix}\begin{pmatrix}1\\0\\0\\1\\0\\1\\1\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}.$$

**Ejemplo 10.32 (Ham(4)).** Construiremos el código de Hamming con $r = 4$, el cual tiene longitud $n = 2^4 - 1 = 15$ y rango $k = 2^4 - 4 - 1 = 11$. El espacio vectorial $\mathbb{Z}_2^4$ tiene 15 vectores distintos de cero:

$$\begin{array}{lllll}(1,0,0,0), & (0,1,0,0), & (0,0,1,0), & (0,0,0,1), & (1,1,0,0),\\ (1,0,1,0), & (1,0,0,1), & (0,1,1,0), & (0,1,0,1), & (0,0,1,1),\\ (1,1,1,0), & (1,1,0,1), & (1,0,1,1), & (0,1,1,1), & (1,1,1,1).\end{array}$$

De acuerdo al Algoritmo 10.30, estos vectores forman las columnas de una matriz verificadora $H_4$ de Ham(4). Para construir esta matriz en su forma estándar, escribimos la base canónica en las últimas tres columnas de $H_4$; el orden del resto de las columnas de $H_4$ puede elegirse de manera arbitraria. Por ejemplo, la siguiente es una matriz verificadora $H_4$:

$$H_4=\begin{bmatrix}1&1&1&0&0&0&1&1&1&0&1&1&0&0&0\\1&0&0&1&1&0&1&1&0&1&1&0&1&0&0\\0&1&0&1&0&1&1&0&1&1&1&0&0&1&0\\0&0&1&0&1&1&0&1&1&1&1&0&0&0&1\end{bmatrix}.$$

En este caso, la matriz $M$ del Teorema 10.14 que relaciona a la matriz verificadora con la generadora es la siguiente

$$M=\begin{bmatrix}1&1&1&0&0&0&1&1&1&0&1\\1&0&0&1&1&0&1&1&0&1&1\\0&1&0&1&0&1&1&0&1&1&1\\0&0&1&0&1&1&0&1&1&1&1\end{bmatrix}.$$

Por lo tanto, una matriz generadora $G_4$ de $\mathrm{Ham}\,(4)$ en forma estándar es

$$G_4 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

**Teorema 10.33.** Para cualquier $r \geq 2$, la distancia mínima del código $\mathrm{Ham}\,(r)$ es

$$d\,(\mathrm{Ham}(r)) = 3.$$

El código $\mathrm{Ham}\,(r)$ puede corregir errores de un bit de la siguiente manera. Supongamos que enviamos la palabra código $w \in \mathrm{Ham}\,(3)$ a través de un canal con ruido y recibimos $w' = w + e_i$, donde $e_i$ representa un error en el $i$-ésimo bit; es decir, $e_i = (0, ..., 1, ..., 0)$. ¿Cómo podemos recuperar $w$ a partir de $w'$? Usaremos la matriz verificadora $H = H_r$:

$$Hw' = H\,(w + e_i) = Hw + He_i = 0 + He_i = He_i.$$

Como, $He_i$ es igual a la $i$-ésima columna de $H$, así que $Hw'$ también es igual a la $i$-ésima columna de $H$. Esto nos permite identificar que el error de bit ocurrió en la $i$-ésima posición.

**Ejemplo 10.34.** Por ejemplo, supongamos que, en lugar de recibir $w = (1, 0, 0, 1, 0, 1, 1) \in Ham\,(3)$, se produce un error de bit y recibimos $w' = (1, 0, 1, 1, 0, 1, 1)$. Usamos la matriz verificadora para corregir $w'$:

$$H\,(w')^T = \begin{pmatrix} 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \\ 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

El vector $(1, 1, 1)$ corresponde a la tercera columna de la matriz verificadora, lo que implica que el error de bit ocurrió en el tercer bit. Por lo tanto podemos corregir $(1, 0, \mathbf{1}, 1, 0, 1, 1)$ a $(1, 0, \mathbf{0}, 1, 0, 1, 1)$.

**Ejemplo 10.35.** Supongamos que queremos transmitir el mensaje

$$v = (1,0,1,0,1,0,1,0,1,0,1)$$

por un canal con ruido y queremos añadir redundancia usando el código Ham(4). Multiplicamos la matriz generadora para obtener la palabra código a transmitir:

$$w = \begin{bmatrix} 1&0&0&0&0&0&0&0&0&0&0 \\ 0&1&0&0&0&0&0&0&0&0&0 \\ 0&0&1&0&0&0&0&0&0&0&0 \\ 0&0&0&1&0&0&0&0&0&0&0 \\ 0&0&0&0&1&0&0&0&0&0&0 \\ 0&0&0&0&0&1&0&0&0&0&0 \\ 0&0&0&0&0&0&1&0&0&0&0 \\ 0&0&0&0&0&0&0&1&0&0&0 \\ 0&0&0&0&0&0&0&0&1&0&0 \\ 0&0&0&0&0&0&0&0&0&1&0 \\ 0&0&0&0&0&0&0&0&0&0&1 \\ 1&1&1&0&0&0&1&1&1&0&1 \\ 1&0&0&1&1&0&1&1&0&1&1 \\ 0&1&0&1&0&1&1&0&1&1&1 \\ 0&0&1&0&1&1&0&1&1&1&1 \end{bmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}.$$

Supongamos que en lugar de recibir $w$ se produce un error en el sexto bit y se recibe

$$w' = (1,0,1,0,1,1,1,0,1,0,1,1,0,1,0).$$

Para corregir el error, multiplicamos $w'$ por la matriz verificadora:

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

Como $(0,0,1,1)$ es igual a la sexta columna de la matriz verificadora, esto indica que se produjo un error en el sexto bit.

***Palabras clave:*** *código lineal binario, longitud, rango, tasa, matriz generadora, matriz verificadora, peso, distancia de Hamming, distancia mínima, código de Hamming.*

## 10.5. Ejercicios

**Ejercicio 10.36.** Determina si los siguientes conjuntos son códigos lineales binarios. En caso de que lo sean, calcula su longitud, rango, tasa, distancia mínima y matriz generadora y verificadora estándar.

1. $\{(0,0),(1,0),(0,1),(1,1,1)\}$.
2. $\{(0,0,0),(1,0,0)\}$.
3. $\{(1,0,0),(0,1,0),(1,1,0)\}$.
4. $\{(0,0,0,0),(0,1,0,1),(0,0,1,0),(1,0,0,0),(1,1,1,1)\}$.
5. $\{(0,0,0,0,0),(1,1,0,0,0),(1,0,1,0,0),(0,1,1,0,0)\}$.
6. $\{(0,0,0,0),(0,1,1,1),(1,0,0,0),(1,1,1,1)\}$.
7. $\{(0,0,0,0),(1,1,0,0),(0,0,1,1)\}$.
8. $\{(0,0,0,0),(1,0,1,1),(1,1,0,0),(0,1,1,1),(0,0,0,1),(1,0,1,0),(1,1,0,1),(0,1,1,0)\}$.
9. $\{(0,0,0,0),(1,1,1,1),(0,1,0,0),(0,1,0,1),(1,0,1,1),(1,0,1,0),(0,0,0,1),(1,1,1,0)\}$.
10. $\{(0,0,0,0),(1,1,1,1),(0,0,0,1),(1,0,0,0),(1,0,0,1),(0,1,0,0),(1,1,0,0),(0,1,0,0)\}$.

**Ejercicio 10.37.** Sean $v, w \in \mathbb{Z}_2^n$. Supongamos que $v = (v_1, \ldots, v_n)$, y consideremos el siguiente conjunto de índices: $I(v) = \{i : v_i = 1\}$. Definimos el peso de $v$ como $p(v) = |I(v)|$. Demuestra lo siguiente:

1. $p(v+w) = |(I(v) \cup I(w)) \setminus (I(v) \cap I(w))|$.
2. $p(v+w) = p(v) + p(w) - 2|I(v) \cap I(w)|$.
3. El conjunto $S = \{v \in \mathbb{Z}_2^n : p(v) \text{ es par}\}$ es un subespacio de $\mathbb{Z}_2^n$.

**Ejercicio 10.38.** Encuentra todas las palabras código de Ham(3). ¿Cuantas hay? Encuentra dos palabras código cuya distancia entre sí sea igual a 3.

**Ejercicio 10.39.** Usando Ham(3), supongamos que el receptor recibe el mensaje $(1,0,0,1,0,1,0)$. Identifica si ocurrió un error en la transmisión. Suponiendo que ocurrió sólo un error de bit, corrige el mensaje recibido.

# Bibliografía


[1] Axler, Sheldon, *Linear Algebra Done Right*, Tercera Edición, Undergraduate Texts in Mathematics, Springer International Publishing, 2015.

[2] Castillo Pérez, A., Castillo Ramírez, A., De la Cruz García, E. L., Hernández Magdaleno, A. M., *Conjuntos y Números*, Editorial Universitaria, Centro Universitario de Ciencias Exactas e Ingenierías, Universidad de Guadalajara, 2014.

[3] Castillo Ramírez, Alonso, *Un curso en teoría de grupos*, Publicaciones Electrónicas, Sociedad Matemática Mexicana, Serie Textos, Vol. 23, 2021.

[4] Hernández Magdaleno, A. M., Castillo Ramírez, A., *Álgebra Moderna: Anillos y Campos*, Editorial Universitaria, Centro Universitario de Ciencias Exactas e Ingenierías, Universidad de Guadalajara, 2012.

[5] Lluis-Puebla, Emilio, *Álgebra Lineal, Álgebra Multilineal y K-Teoría Algebraica Clásica*, Publicaciones Electrónicas **9**, Sociedad Matemática Mexicana, 2008.

[6] Roman, S., *Advanced Linear Algebra*, Tercera Edición, Graduate Texts in Mathematics **135**, Springer, 2008.

[7] Rose, H.E., *Linear Algebra: A Pure Mathematical Apporach.* Birkhäuser, 2002.

[8] Pretzel, O., *Error-Correcting Codes and Finite Fields*, Oxford University Press, 1992.

[9] Zaldívar, Felipe, *Introducción al álgebra lineal*, Papirhos, IM-UNAM, México, 2019.