Gabriela Jeronimo, Juan Sabia y Susana Tesauri

Álgebra lineal

Buenos Aires, agosto de 2008

Prefacio

El álgebra lineal es una herramienta básica para casi todas las ramas de la matemática así como para disciplinas afines tales como la física, la ingeniería y la computación, entre otras. Estas notas, basadas en la materia Álgebra Lineal destinada a alumnos de la Licenciatura en Ciencias Matemáticas y del Profesorado en Matemáticas de la Facultad de Ciencias Exactas y Naturales de la Universidad de Buenos Aires, que hemos dictado varias veces, pretenden, entre tantos buenos textos de álgebra lineal existentes, ser sólo una introducción básica al tema que se ajusta a los contenidos curriculares del curso y, al mismo tiempo, una guía de estudios para los alumnos.

Las notas no presuponen ningún conocimiento previo de álgebra lineal, aunque sí de algunas propiedades básicas de polinomios a coeficientes en un cuerpo y de números complejos, y en algunos ejercicios se utilizan estructuras que provienen de la aritmética elemental. Se comienza con las definiciones básicas de estructuras algebraicas necesarias para definir la noción de espacio vectorial, para seguir con la noción de subespacio, sistema de generadores e independencia lineal. Después de dar una breve introducción al tema de las matrices a coeficientes en un cuerpo, se definen y estudian las transformaciones lineales, el espacio dual y la teoría de determinantes. La diagonalización de matrices y la forma de Jordan de automorfismos en espacios de dimensión finita se desarrollan a continuación, seguidas del estudio de espacios con producto interno reales y complejos. El capítulo de variedades lineales puede verse como una aplicación del álgebra lineal a la geometría afín. Finalmente, se da una breve introducción a la teoría de formas bilineales.

Gabriela Jeronimo, Juan Sabia y Susana Tesauri

Índice General

Capítulo 1

Espacios vectoriales

En diversos conjuntos conocidos, por ejemplo los de vectores en el plano o en el espacio ($\mathbb{R}^2$ y $\mathbb{R}^3$), o también el de los polinomios ($\mathbb{R}[X]$), sabemos sumar sus elementos y multiplicarlos por números. Todos estos conjuntos comparten una cierta "estructura", que está dada por esa suma y ese producto por números, a la que llamaremos espacio vectorial. En este capítulo presentaremos la noción de espacio vectorial y estudiaremos algunas propiedades básicas que poseen los conjuntos con dicha estructura.

1.1 Espacios vectoriales y subespacios

1.1.1 Preliminares

La noción de espacio vectorial requiere de dos conjuntos: un conjunto K (los escalares) y otro conjunto V (los vectores). Estos dos conjuntos deben satisfacer ciertas propiedades, que esencialmente se refieren a que los elementos de V se puedan sumar entre sí y multiplicar por elementos de K.

Comenzaremos dando algunas definiciones previas para poder después presentar la definición precisa de espacio vectorial.

Definición 1.1 Sea A un conjunto no vacío. Una *operación* (o *ley de composición interna* u *operación binaria*) de A es una función $* : A \times A \to A$.

Notación. $*(a, b) = c$ se escribe $a * b = c$.

Ejemplos.

- $+ : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$, tal que $+(a, b) = a + b$, es una operación de $\mathbb{N}$.
- Como la resta, $-(a, b) = a - b$, no es una función de $\mathbb{N} \times \mathbb{N}$ en $\mathbb{N}$, entonces no es una operación de $\mathbb{N}$.

- La suma $+$, el producto $\cdot$ y la resta $-$ son operaciones de $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ y $\mathbb{C}$.

No nos interesaremos por operaciones cualesquiera, sino que trabajaremos con operaciones que posean algunas propiedades. Entre las propiedades que analizaremos se encuentran las siguientes:

Definición 1.2 (Propiedades básicas) Sea $* : A \times A \to A$ una operación.

i) $*$ se dice *asociativa* si $(a * b) * c = a * (b * c) \quad \forall\, a, b, c \in A$.

ii) Se dice que $*$ tiene *elemento neutro* si $\exists\, e \in A$ tal que $e * a = a * e = a$ para cada $a \in A$.

   (Observar que si $*$ tiene elemento neutro, éste es único, ya que si e y e' son elementos neutros, $e' = e * e' = e$.)

iii) Si $*$ tiene elemento neutro e, se dice que todo elemento tiene *inverso* para $*$ si $\forall\, a \in A$, $\exists\, a' \in A$ tal que $a * a' = a' * a = e$.

iv) $*$ se dice *conmutativa* si $a * b = b * a \quad \forall\, a, b \in A$.

Se pueden estudiar las características que comparten los conjuntos con una operación que satisface algunas de estas propiedades. Una noción útil es la de grupo, que definimos a continuación.

Definición 1.3 Sea A un conjunto, y sea $*$ una operación en A que satisface las propiedades i), ii) y iii) de la definición anterior. Entonces $(A, *)$ se llama un *grupo*. Si además $*$ cumple iv), se dice que $(A, *)$ es un *grupo abeliano* o *conmutativo*.

Ejemplos.

- $(\mathbb{N}, +)$ no es un grupo: se puede probar que no tiene elemento neutro.
- $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$ y $(\mathbb{C}, +)$ son grupos abelianos.
- $(\mathbb{Z}, \cdot)$ no es un grupo: se puede probar que sólo 1 y -1 tienen inverso multiplicativo.
- $(\mathbb{Q} - \{0\}, \cdot)$, $(\mathbb{R} - \{0\}, \cdot)$ y $(\mathbb{C} - \{0\}, \cdot)$ son grupos abelianos.
- $A = \{f : \mathbb{R} \to \mathbb{R}\}$, $* = \circ$ (composición de funciones). Entonces $(A, *)$ no es un grupo: las únicas funciones con inversa para $\circ$ son las biyectivas.
- $S_{\mathbb{R}} = \{f : \mathbb{R} \to \mathbb{R} \,/\, f \text{ es biyectiva }\}$, $* = \circ$. Entonces $(S_{\mathbb{R}}, \circ)$ es un grupo.
- C un conjunto, $\mathcal{P}(C) = \{S \subseteq C\}$. Se define la operación $\triangle : \mathcal{P}(C) \times \mathcal{P}(C) \to \mathcal{P}(C)$, llamada diferencia simétrica, de la siguiente forma:

  $$A \triangle B = (A \cup B) - (A \cap B).$$

  Entonces $(\mathcal{P}(C), \triangle)$ es un grupo abeliano.

A partir de la definición de grupo pueden probarse propiedades que poseen todos los conjuntos con esa estructura. Por ejemplo:

- *Sea $(G, *)$ un grupo. Entonces para cada $a \in G$ existe un único inverso para a.*

  Sea e el elemento neutro de $(G, *)$. Supongamos que b y c son inversos de a. Entonces

$$b = e * b = (c * a) * b = c * (a * b) = c * e = c.$$

Notación. Si G es un grupo abeliano y la operación se nota $+$, el elemento neutro se notará 0 y, para cada $a \in G$, el inverso de a se notará $-a$. (En otros casos, el elemento neutro se nota 1 y el inverso de a se nota a^{-1}.)

La siguiente definición que daremos se refiere a conjuntos en los cuales hay dos operaciones relacionadas entre sí.

Definición 1.4 Sea A un conjunto y sean $+$ y $\cdot$ operaciones de A. Se dice que $(A, +, \cdot)$ es un *anillo* si

i) $(A, +)$ es un grupo abeliano

ii) $\cdot$ es asociativa y tiene elemento neutro

iii) Valen las propiedades distributivas: Para $a, b, c \in A$,
   - $a \cdot (b + c) = a \cdot b + a \cdot c$
   - $(b + c) \cdot a = b \cdot a + c \cdot a$

Además, si $\cdot$ es conmutativa, se dice que $(A, +, \cdot)$ es un *anillo conmutativo.*

Notación. Cuando quede claro cuáles son las operaciones $+$ y $\cdot$, para referirnos al anillo $(A, +, \cdot)$, escribiremos simplemente A. Al elemento neutro del producto se lo notará 1.

Ejemplos.

- $(\mathbb{Z}, +, \cdot)$, $(\mathbb{Q}, +, \cdot)$, $(\mathbb{R}, +, \cdot)$ y $(\mathbb{C}, +, \cdot)$ son anillos conmutativos.
- $(\mathbb{Z}_n, +, \cdot)$ es una anillo conmutativo.
- Si $(A, +, \cdot)$ es un anillo conmutativo, entonces $(A[X], +, \cdot)$ es un anillo conmutativo con las operaciones usuales de polinomios.
- Si C es un conjunto, $(\mathcal{P}(C), \triangle, \cap)$ es un anillo conmutativo.
- $\{f : \mathbb{R} \to \mathbb{R}\}$ con las operaciones usuales de suma y producto de funciones es un anillo conmutativo.

Al igual que en el caso de los grupos, también pueden probarse propiedades que poseen todos los anillos:

- *Sea* $(A, +, \cdot)$ *un anillo, y sea* 0 *el elemento neutro de* $+$*. Entonces* $0 \cdot a = 0$, $\forall\, a \in A$.

  Se tiene que
  $$0 \cdot a = (0 + 0) \cdot a = 0 \cdot a + 0 \cdot a.$$
  Si b es el inverso aditivo de $0 \cdot a$, resulta
  $$0 = 0 \cdot a + b = (0 \cdot a + 0 \cdot a) + b = 0 \cdot a + (0 \cdot a + b) = 0 \cdot a.$$
  Luego, $0 \cdot a = 0$.

En un anillo cualquiera no es cierto que $a \cdot b = 0 \Rightarrow a = 0$ o $b = 0$. Por ejemplo, en $\mathbb{Z}_4$, se tiene que $2 \cdot 2 = 0$, pero $2 \neq 0$.

Definición 1.5 Un anillo conmutativo $(A, +, \cdot)$ se llama un *dominio de integridad* o *dominio íntegro* si $a \cdot b = 0 \;\Rightarrow\; a = 0 \text{ o } b = 0$.

Ejemplos.

- $(\mathbb{Z}, +, \cdot)$, $(\mathbb{Q}, +, \cdot)$, $(\mathbb{R}, +, \cdot)$ y $(\mathbb{C}, +, \cdot)$ son dominios de integridad.
- Si A es un dominio de integridad, entonces $A[X]$ es un dominio de integridad.
- $\mathbb{Z}_p$ es un dominio de integridad $\iff$ p es primo.

La siguiente definición resume las propiedades que debe satisfacer uno de los conjuntos involucrados en la definición de un espacio vectorial.

Definición 1.6 Sea K un conjunto, y sean $+$ y $\cdot$ operaciones de K. Se dice que $(K, +, \cdot)$ es un *cuerpo* si $(K, +, \cdot)$ es un anillo conmutativo y todo elemento no nulo de K tiene inverso multiplicativo. Es decir:

i) $(K, +)$ es un grupo abeliano,

ii) $(K - \{0\}, \cdot)$ es un grupo abeliano, y

iii) vale la propiedad distributiva de $\cdot$ con respecto a $+$.

Ejemplos.

- $(\mathbb{Q}, +, \cdot)$, $(\mathbb{R}, +, \cdot)$ y $(\mathbb{C}, +, \cdot)$ son cuerpos
- $(\mathbb{Z}_p, +, \cdot)$ es un cuerpo $\iff$ p es primo.

- Se define $\mathbb{Q}[\sqrt{2}] = \left\{ \sum_{i=0}^{n} a_i(\sqrt{2})^i \,/\, a_i \in \mathbb{Q},\, n \in \mathbb{N}_0 \right\}$. Veamos que $(\mathbb{Q}[\sqrt{2}], +, \cdot)$ es un cuerpo.

  Usando que $\mathbb{Q}[\sqrt{2}] \subset \mathbb{R}$, se puede probar fácilmente que $(\mathbb{Q}[\sqrt{2}], +, \cdot)$ es un anillo conmutativo.

  Observamos que $\mathbb{Q}[\sqrt{2}] = \{a + b\sqrt{2} : a, b \in \mathbb{Q}\}$. En efecto, para cada $k \in \mathbb{N}$, se tiene que $(\sqrt{2})^{2k} = 2^k$ y $(\sqrt{2})^{2k+1} = 2^k\sqrt{2}$ y entonces, todo elemento de la forma $\sum_{i=0}^{n} a_i(\sqrt{2})^i$ con $a_i \in \mathbb{Q}$ y $n \in \mathbb{N}_0$ puede escribirse como $a + b\sqrt{2}$ con $a, b \in \mathbb{Q}$. Recíprocamente, es claro que todo elemento de la forma $a + b\sqrt{2}$ con $a, b \in \mathbb{Q}$ pertenece a $\mathbb{Q}[\sqrt{2}]$.

  Veamos ahora que todo elemento no nulo tiene inverso.

  Sea $a + b\sqrt{2} \neq 0$. Entonces $(a + b\sqrt{2})(a - b\sqrt{2}) = a^2 - 2b^2 \neq 0$ (pues $a, b \in \mathbb{Q}$), de donde

$$(a + b\sqrt{2})^{-1} = \frac{a}{a^2 - 2b^2} + \frac{-b}{a^2 - 2b^2}\sqrt{2}.$$

También en el caso de los cuerpos se pueden probar propiedades generales. Por ejemplo:

- *Todo cuerpo* $(K, +, \cdot)$ *es un dominio de integridad.*

  Tenemos que probar que $a \cdot b = 0 \Rightarrow a = 0$ o $b = 0$. Supongamos que $a \cdot b = 0$. Si $a = 0$, ya está. Si $a \neq 0$, entonces existe a^{-1} tal que $a \cdot a^{-1} = a^{-1} \cdot a = 1$. Entonces

$$a^{-1} \cdot (a \cdot b) = a^{-1} \cdot 0 \;\Rightarrow\; (a^{-1} \cdot a) \cdot b = 0 \;\Rightarrow\; 1 \cdot b = 0 \;\Rightarrow\; b = 0.$$

Para poder completar la definición de espacio vectorial necesitamos definir una clase especial de funciones que se aplican a elementos de dos conjuntos distintos:

Definición 1.7 Sean A y B dos conjuntos. Una *acción* de A en B es una función $\cdot : A \times B \to B$.
Notación: $\cdot\,(a, b) = a \cdot b$

Estamos ahora en condiciones de dar la definición de espacio vectorial.

1.1.2 Espacios vectoriales

Definición 1.8 Sea $(K, +, \cdot)$ un cuerpo. Sea V un conjunto no vacío, sea $+$ una operación en V y sea $\cdot$ una acción de K en V. Se dice que $(V, +, \cdot)$ es un K-*espacio vectorial* si se cumplen las siguientes condiciones:

i) $(V, +)$ es un grupo abeliano.

ii) La acción $\cdot : K \times V \to V$ satisface:

(a) $a \cdot (v + w) = a \cdot v + a \cdot w \quad \forall\, a \in K;\ \forall\, v, w \in V.$

(b) $(a + b) \cdot v = a \cdot v + b \cdot v \quad \forall\, a, b \in K;\ \forall\, v \in V.$

(c) $1 \cdot v = v \quad \forall\, v \in V.$

(d) $(a \cdot b) \cdot v = a \cdot (b \cdot v) \quad \forall\, a, b \in K;\ \forall\, v \in V.$

Los elementos de V se llaman *vectores* y los elementos de K se llaman *escalares*. La acción $\cdot$ se llama *producto por escalares*.

Nótese que el símbolo $\cdot$ se usa tanto para la acción de K en V como para el producto en K, pero esto no debería generar confusión puesto que en el primer caso estará aplicado a un elemento de K y otro de V, mientras que en el segundo, a dos elementos de K.

En lo que sigue, K denotará un cuerpo. Si $(V, +, \cdot)$ es un K-espacio vectorial y la operación $+$ de V y la acción $\cdot$ de K en V quedan claras del contexto, diremos simplemente que V es un K-espacio vectorial.

Hay propiedades que se cumplen en cualquier espacio vectorial. A continuación mostramos algunas de ellas.

Sea V un K-espacio vectorial. Entonces:

1. *$0 \cdot v = 0$ para todo $v \in V$. (Observar que el elemento 0 que aparece en el miembro izquierdo de la igualdad es el elemento neutro de K, mientras que el de la derecha es el vector $0 \in V$.)*

2. *$(-1) \cdot v = -v$ para todo $v \in V$. (Recuérdese que $-v$ denota al inverso aditivo de v).*

Demostración.

1. Se tiene que
$$0 \cdot v = (0 + 0) \cdot v = 0 \cdot v + 0 \cdot v.$$
Sea w el inverso aditivo de $0 \cdot v$. Entonces
$$0 = 0 \cdot v + w = (0 \cdot v + 0 \cdot v) + w = 0 \cdot v + (0 \cdot v + w) = 0 \cdot v + 0 = 0 \cdot v$$

2. Vemos que
$$v + (-1) \cdot v = (-1) \cdot v + v = (-1) \cdot v + 1 \cdot v = (-1 + 1) \cdot v = 0 \cdot v = 0.$$
Luego, $(-1) \cdot v$ es el inverso aditivo de v, es decir $(-1) \cdot v = -v$.

Ejemplos. En lo que sigue K es un cuerpo.

1. K es un K-espacio vectorial.

2. Sea $K^n = \{(x_1, \dots, x_n) \,/\, x_i \in K\}$. Se definen

   $$\begin{array}{lll} + & : & K^n \times K^n \to K^n, \quad (x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n) \\ \cdot & : & K \times K^n \to K^n, \quad \lambda \cdot (x_1, \dots, x_n) = (\lambda x_1, \dots, \lambda x_n) \end{array}$$

   Entonces K^n es un K-espacio vectorial.

3. Una matriz de n filas y m columnas es un arreglo de $n \times m$ números ubicados en n filas y m columnas.

   Sea $K^{n\times m} = \{A \;/\; A$ es una matriz de n filas y m columnas de elementos en $K\}$. Observamos que un elemento A de $K^{n\times m}$ es de la forma

   $$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1m} \\ A_{21} & A_{22} & \cdots & A_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ A_{n1} & A_{n2} & \cdots & A_{nm} \end{pmatrix}.$$

   Si $A \in K^{n\times m}$, denotaremos por A_{ij} al elemento ubicado en la intersección de la fila i y la columna j de A.

   Se definen

   $$\begin{array}{lll} + & : & K^{n\times m} \times K^{n\times m} \to K^{n\times m}, \ (A + B)_{ij} = A_{ij} + B_{ij} \quad (1 \le i \le n, 1 \le j \le m) \\ \cdot & : & K \times K^{n\times m} \to K^{n\times m}, \ (\lambda \cdot A)_{ij} = \lambda \cdot A_{ij} \quad (1 \le i \le n, 1 \le j \le m) \end{array}$$

   Entonces $K^{n\times m}$ es un K-espacio vectorial.

4. Sea Z un conjunto no vacío. Se considera $K^Z = \{f : Z \to K \;/\; f$ es función $\}$ y se definen

   $$\begin{array}{lll} + & : & K^Z \times K^Z \to K^Z, \ (f + g)(x) = f(x) + g(x) \quad \forall\, x \in Z, \\ \cdot & : & K \times K^Z \to K^Z, \ (\lambda \cdot f)(x) = \lambda \cdot f(x) \quad \forall\, x \in Z. \end{array}$$

   Entonces K^Z es un K-espacio vectorial.

5. $K[X]$, el conjunto de polinomios en la variable X a coeficientes en K, es un K-espacio vectorial con la suma usual de polinomios y la multiplicación usual de polinomios por una constante.

6. $\mathbb{R}$ es un $\mathbb{Q}$-espacio vectorial; $\mathbb{C}$ es un $\mathbb{R}$-espacio vectorial y un $\mathbb{Q}$-espacio vectorial.

7. $\mathbb{Q}[\sqrt{2}]$ es un $\mathbb{Q}$-espacio vectorial.

1.1.3 Subespacios

Dentro de un K-espacio vectorial V, hay subconjuntos que heredan la estructura de V, es decir, que son también espacios vectoriales con la misma operación, el mismo elemento neutro y la misma acción que V. En esta sección, comenzaremos el estudio de los subconjuntos con esta propiedad.

Definición 1.9 Sea V un K-espacio vectorial. Un subconjunto $S \subseteq V$ no vacío se dice un *subespacio de* V si la suma y el producto por escalares (de V) son una operación y una acción en S que lo convierten en un K-espacio vectorial.

Ejemplo. Caractericemos todos los subespacios de $\mathbb{R}^2$:

- $S = \{(0,0)\}$ es un subespacio.
- Supongamos que S es un subespacio y que contiene algún elemento v no nulo. Entonces, para todo $\lambda \in \mathbb{R}$, $\lambda.v \in S$. Si éstos son todos los elementos de S, entonces S es un subespacio (que, gráficamente, resulta ser una recta que pasa por el origen).
- Con la notación del punto anterior, si S contiene algún elemento que no es de la forma $\lambda.v$, digamos v', contiene también a todos los múltiplos de v'. Luego, S contiene a las dos rectas L y L' que pasan por el origen y cuyas direcciones son v y v' respectivamente. Es claro (usando la regla del paralelogramo) que cualquier punto en $\mathbb{R}^2$ es suma de un elemento de L más uno de L', luego pertenece a S. En consecuencia, $S = \mathbb{R}^2$.

Observamos que, dado un K-espacio vectorial V y un subconjunto S de V, para determinar si S es un subespacio de V según la Definición 1.9 debemos verificar la validez de una gran cantidad de propiedades (todas las involucradas en la definición de espacio vectorial). La siguiente proposición nos provee una caracterización de los subespacios en términos de sólo tres propiedades, a partir de las cuales se deducen todas las demás.

Proposición 1.10 *Sea V un K-espacio vectorial y sea $S \subseteq V$. Entonces S es un subespacio de V si y sólo si valen las siguientes condiciones:*

i) $0 \in S$

ii) $v, w \in S \Longrightarrow v + w \in S$

iii) $\lambda \in K, v \in S \Longrightarrow \lambda \cdot v \in S$

Demostración.

($\Rightarrow$) Es inmediato verificar que si S es un subespacio de V se cumplen i), ii) e iii).

($\Leftarrow$) La condición i) asegura que S es no vacío.

Por ii), $+$ es una operación de S y por iii), $\cdot$ es una acción.

La asociatividad y conmutatividad de la suma se deducen de la validez de las mismas para V, el elemento neutro de la suma $0 \in S$ por i), y la existencia de inverso aditivo se deduce de que dado $v \in S$, $-v = (-1) \cdot v$, que pertenece a S por iii).

Las propiedades de la acción en la definición de espacio vectorial se deducen también de su validez en V. □

Observamos que la condición i) en la proposición anterior puede ser reemplazada por *i')* $S \neq \emptyset$.

Es decir, las condiciones i), ii), iii) son equivalentes a i'), ii), iii). La demostración de este hecho queda como ejercicio.

Ejemplos. Sea V un K-espacio vectorial.

1. $\{0\}$ es un subespacio de V.
2. V es un subespacio de V.
3. Si $v \in V$, $S = \{\lambda \cdot v \,/\, \lambda \in K\}$ es un subespacio de V:
   - i) $0 = 0 \cdot v \in S$.
   - ii) Si $\lambda \cdot v$, $\mu \cdot v \in S$, entonces $\lambda \cdot v + \mu \cdot v = (\lambda + \mu) \cdot v \in S$.
   - iii) Si $\lambda \cdot v \in S$ y $\alpha \in K$, entonces $\alpha \cdot (\lambda \cdot v) = (\alpha \cdot \lambda) \cdot v \in S$.

   Este subespacio se denomina el *subespacio generado por* v y se nota $S = < v >$.
4. Sean $v_1, \ldots, v_n \in V$.
   Entonces $S = \{\alpha_1.v_1 + \cdots + \alpha_n.v_n : \alpha_i \in K,\ 1 \leq i \leq n\}$ es un subespacio de V:
   - i) $0 = 0.v_1 + \cdots + 0.v_n \in S$.
   - ii) Si $v, w \in S$, $v = \alpha_1.v_1 + \cdots + \alpha_n.v_n$, $w = \beta_1.v_1 + \cdots + \beta_n.v_n$, entonces $v + w = (\alpha_1 + \beta_1).v_1 + \cdots + (\alpha_n + \beta_n).v_n \in S$.
   - iii) Si $\lambda \in K$ y $v = \alpha_1.v_1 + \cdots + \alpha_n.v_n \in S$, entonces $\lambda.v = (\lambda.\alpha_1).v_1 + \cdots + (\lambda.\alpha_n).v_n \in S$.

   El subespacio S que hemos definido se llama el *subespacio generado por* $v_1, \ldots, v_n$ y se nota $S = < v_1, \ldots, v_n >$.

Si V es un K-espacio vectorial, tiene sentido considerar las operaciones de unión e intersección entre subespacios de V (que son subconjuntos de V). Una pregunta que surge es si estas operaciones preservan la estructura de subespacio. Como veremos a continuación, esto vale en el caso de la intersección de subespacios, pero no para la unión.

Proposición 1.11 *Sea V un K-espacio vectorial, y sean S y T subespacios de V. Entonces $S \cap T$ es un subespacio de V.*

Demostración.

- i) $0 \in S \cap T$ puesto que $0 \in S$ y $0 \in T$.
- ii) Sean $v, w \in S \cap T$. Entonces $v \in S$, $v \in T$, $w \in S$ y $w \in T$. Como $v, w \in S$ y S es un subespacio, entonces $v + w \in S$. Análogamente, $v + w \in T$. Luego, $v + w \in S \cap T$.

iii) Sean $\lambda \in K$ y $v \in S \cap T$. Entonces $v \in S$ y $v \in T$. Como $\lambda \in K$, $v \in S$ y S es un subespacio, entonces $\lambda \cdot v \in S$. Análogamente, $\lambda \cdot v \in T$. Luego, $\lambda \cdot v \in S \cap T$. $\square$

En forma análoga a lo hecho en la demostración de la proposición anterior, se prueba que la intersección de cualquier familia de subespacios de un K-espacio vectorial V es un subespacio de V.

Observación 1.12 Si V es un K-espacio vectorial, S y T subespacios de V, entonces $S \cup T$ no es necesariamente un subespacio de V.

En efecto, consideremos en $\mathbb{R}^2$ los subespacios $S = < (1,0) >$ y $T = < (0,1) >$.

Observamos que $(1,0) \in S$ y $(0,1) \in T$; luego, ambos pertenecen a $S \cup T$. Pero $(1,0) + (0,1) = (1,1) \notin S \cup T$, puesto que $(1,1) \notin S$ y $(1,1) \notin T$.

Concluimos esta sección exhibiendo algunos ejemplos de subespacios de distintos K-espacios vectoriales.

Ejemplos.

1. Sean $a_1, \ldots, a_n \in K$ fijos. Sea $S = \{(x_1, \ldots, x_n) \in K^n : a_1x_1 + \cdots a_nx_n = 0\}$. Es fácil verificar que S es un subespacio de K^n.

2. $S = \left\{ (x_1, \ldots, x_n) \in K^n : \left\{ \begin{array}{c} a_{11}x_1 + \cdots + a_{1n}x_n = 0 \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = 0 \end{array} \right. \right\}$ es un subespacio de K^n, pues $S = \bigcap\limits_{i=1}^{m} S_i$, donde $S_i = \{(x_1, \ldots, x_n) \in K^n : a_{i1}x_1 + \cdots + a_{in}x_n = 0\}$ $(1 \le i \le m)$ y cada S_i es un subespacio de K^n.

3. Sean $V = K[X]$ y $n \in \mathbb{N}$ fijo. Se tiene que $K_n[X] = \{f \in K[X] \ / \ f = 0 \text{ o } \mathrm{gr}(f) \le n\}$ es un subespacio de V:

   i) $0 \in K_n[X]$.

   ii) Sean $f, g \in K_n[X]$. Si $f = 0$ o $g = 0$ es claro que $f + g \in S$. Si $f + g = 0$, entonces $f + g \in S$. Si no, $\mathrm{gr}(f + g) \le \max(\mathrm{gr}(f), \mathrm{gr}(g)) \le n$, y por lo tanto $f + g \in S$.

   iii) Sean $\lambda \in K$ y $f \in K_n[X]$. Si $\lambda = 0$ o $f = 0$, entonces $\lambda.f = 0 \in K_n[X]$. Si no, $\mathrm{gr}(\lambda.f) = \mathrm{gr}(f)$, de donde $\lambda.f \in K_n[X]$.

   Observar que el conjunto $\{f \in K[X] \ / \ f = 0 \text{ o } \mathrm{gr}(f) \ge n\}$, para $n \in \mathbb{N}$ fijo, no es un subespacio de $K[X]$. Por ejemplo: $f = X^n$ y $g = -X^n + 1$ pertenecen a dicho conjunto, pero $f + g = 1$ no.

1.1.4 Sistemas de generadores

El objetivo de esta sección es mostrar cómo pueden describirse todos los elementos de un K-espacio vectorial V a partir de ciertos subconjuntos de elementos de V.

De la definición de K-espacio vectorial vemos que una forma de obtener nuevos elementos de V a partir de los elementos de un subconjunto $G \subseteq V$ es considerando sumas finitas de múltiplos por escalares de elementos de G. Surge entonces la noción de combinación lineal:

Definición 1.13 Sea V un K-espacio vectorial, y sea $G = \{v_1, \dots, v_r\} \subseteq V$. Una *combinación lineal de* G es un elemento $v \in V$ tal que $v = \sum_{i=1}^{r} \alpha_i . v_i$ con $\alpha_i \in K$ para cada $1 \le i \le r$.

Ejemplos.

1. Sea $G = \{(1,2),(3,4)\} \subseteq \mathbb{R}^2$. Una combinación lineal de G es un vector $v = \alpha.(1,2) + \beta.(3,4)$ con $\alpha, \beta \in \mathbb{R}$.
2. Sea $G = \{1, X, \dots, X^n\} \subseteq \mathbb{R}_n[X]$. Una combinación lineal de G es $\sum_{i=0}^{n} \alpha_i X^i$ con $\alpha_i \in \mathbb{R}$ para cada $0 \le i \le n$.

La definición de combinación lineal se extiende al caso de subconjuntos no necesariamente finitos del espacio vectorial considerado:

Definición 1.14 Sea V un K-espacio vectorial, sea I un conjunto de índices y sea $G = \{v_i \,/\, i \in I\} \subset V$. Una *combinación lineal de* G es un elemento $v \in V$ tal que $v = \sum_{i \in I} \alpha_i . v_i$ donde $\alpha_i = 0$ salvo para finitos $i \in I$.

Ejemplos.

1. Sea $G = \{X^i \,/\, i \in \mathbb{N}_0\} \subseteq \mathbb{R}[X]$. Una combinación lineal de G es $\sum_{i=0}^{\infty} \alpha_i X^i$ donde $\alpha_i \in \mathbb{R}$ y $\alpha_i = 0$ salvo para finitos valores de $i \in \mathbb{N}_0$.
2. Sea $G = \{(\alpha, 0) : \alpha \in \mathbb{R}\} \subseteq \mathbb{R}^2$. Una combinación lineal de G es $\sum_{\alpha \in \mathbb{R}} \beta_\alpha . (\alpha, 0)$ tal que $\beta_\alpha \in \mathbb{R}$ y $\beta_\alpha = 0$ salvo para finitos $\alpha \in \mathbb{R}$.

Dado un espacio vectorial V, considerando las combinaciones lineales de los elementos de ciertos subconjuntos de V, podemos obtener cualquier elemento del espacio vectorial en cuestión. Como se verá en los ejemplos, en muchos casos esto nos permitirá describir conjuntos infinitos (como por ejemplo $\mathbb{R}^2$) utilizando finitos elementos del espacio.

Definición 1.15 Sea V un K-espacio vectorial y sea $G \subseteq V$. Se dice que G es un *sistema de generadores de* V (y se nota $< G > = V$) si todo elemento de V es una combinación lineal de G.

Ejemplos.

1. $\mathbb{R}^2 = < (1,0), (0,1) >$, pues $\forall\, x = (\alpha, \beta) \in \mathbb{R}^2$, $x = \alpha.(1,0) + \beta.(0,1)$.
2. $K^n = < (1,0\ldots,0), (0,1,0,\ldots,0), \ldots, (0,\ldots,0,1,0,\ldots,0), \ldots, (0,\ldots,0,1) >$.
3. $K^{n\times m} = < E^{ij} >_{\substack{1\le i\le n\\ 1\le j\le m}}$ donde $(E^{ij})_{kl} = \begin{cases} 1 & \text{si } k = i \text{ y } j = l \\ 0 & \text{si no} \end{cases}$
4. $K[X] =< X^i >_{i\in\mathbb{N}_0}$.
5. Si $G \subseteq K[X]$ tal que para cada $i \in \mathbb{N}_0$, existe $f_i \in G$ con $\text{gr}(f_i) = i$, entonces $K[X] = < G >$:

   Es claro que $0 \in < G >$. Veamos, por inducción en $\text{gr}(g)$, que $g \in < G >$ para cada $g \in K[X]$.

   Si $\text{gr}(g) = 0$, entonces $g \in K$, y como existe $f_0 \in G$ con $\text{gr}(f_0) = 0$ (es decir, $f_0 \in K - \{0\}$), se tiene que $g = \dfrac{g}{f_0}.f_0 \in < G >$.

   Sea $n > 0$ y supongamos que todo polinomio de grado menor que n y el polinomio nulo pertenecen a $< G >$. Sea $g \in K[X]$ con $\text{gr}(g) = n$. Por hipótesis, existe $f_n \in G$ con $\text{gr}(f_n) = n$. Si $g = \sum\limits_{j=0}^{n} a_j X^j$ y $f_n = \sum\limits_{j=0}^{n} b_j X^j$, consideramos $\widetilde{g} = g - \frac{a_n}{b_n} f_n$. Observamos que $\widetilde{g} = 0$ o $\text{gr}(\widetilde{g}) < n$. Por hipótesis inductiva, $\widetilde{g} \in < G >$, es decir $\widetilde{g} = \sum\limits_{f\in G} c_f.f$ con $c_f = 0$ salvo para finitos f. En consecuencia,

$$g = \widetilde{g} + \frac{a_n}{b_n} f_n = \sum_{f\in G,\, f\neq f_n} c_f.f + \Big(c_{f_n} + \frac{a_n}{b_n}\Big) f_n \;\in\; < G >.$$

1.2 Sistemas de ecuaciones lineales

Hemos visto que un conjunto del tipo

$$S = \left\{ (x_1, \ldots, x_m) \in K^m : \left\{ \begin{array}{c} a_{11}x_1 + \cdots + a_{1m}x_m = 0 \\ \vdots \\ a_{n1}x_1 + \cdots + a_{nm}x_m = 0 \end{array} \right. \right\}$$

es un subespacio de K^m. Surge entonces la cuestión de describir estos conjuntos. Esto puede hacerse, por ejemplo, encontrando un sistema de generadores del subespacio S.

Más en general, estudiaremos el problema de dar una descripción del conjunto de soluciones de un sistema de ecuaciones de la forma

$$\left\{ \begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m & = & b_1 \\ & \vdots & \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m & = & b_n \end{array} \right.$$

donde $a_{ij} \in K$ para todo $1 \le i \le n$ y $1 \le j \le m$, y $b_i \in K$ para todo $1 \le i \le n$, a los que llamaremos sistemas de n ecuaciones lineales en m incógnitas.

1.2.1 Sistemas lineales homogéneos

Un primer tipo de sistemas de ecuaciones que estudiaremos son los que tienen todas las ecuaciones igualadas a 0.

Definición 1.16 Un *sistema lineal homogéneo* de n ecuaciones con m incógnitas a coeficientes en un cuerpo K es un sistema del tipo

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m & = & 0 \\ \vdots & & \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m & = & 0 \end{cases}$$

donde $a_{ij} \in K$ para cada $1 \leq i \leq n$, $1 \leq j \leq m$.

Notación. La matriz $A \in K^{n\times m}$ definida por $A_{ij} = a_{ij}$ se llama la *matriz asociada al sistema.*

Observación 1.17 El conjunto de las soluciones de un sistema lineal homogéneo con m incógnitas es un subespacio de K^m (ver Ejemplo 2 en la página 10).

Resolver un sistema de este tipo significará dar un sistema de generadores para el subespacio de las soluciones.

El método que daremos para la resolución de sistemas de ecuaciones lineales consiste en transformar el sistema dado, por medio de ciertas operaciones, en otro que tenga el mismo conjunto de soluciones, pero cuya resolución sea más simple. Aparece entonces la noción de sistemas equivalentes:

Definición 1.18 Dos sistemas lineales homogéneos se dicen *equivalentes* si sus conjuntos de soluciones son iguales.

Ejemplo. Los siguientes sistemas lineales homogéneos a coeficientes en $\mathbb{R}$ son equivalentes:

$$\begin{cases} x+y+z & = & 0 \\ y+z & = & 0 \end{cases} \qquad \begin{cases} x & = & 0 \\ y+z & = & 0 \end{cases}$$

1.2.2 Método de triangulación

Algunos sistemas de ecuaciones lineales son muy fáciles de resolver:

Ejemplo. Consideremos el siguiente sistema lineal homogéneo en $\mathbb{R}^3$:

$$\begin{cases} 2x_1 & + & 3x_2 & - & x_3 & = & 0 \\ & - & x_2 & + & x_3 & = & 0 \\ & & & & 5x_3 & = & 0 \end{cases}$$

Este sistema tiene como única solución a $(0,0,0)$: De la tercera ecuación, resulta que $x_3 = 0$. Teniendo en cuenta que $x_3 = 0$, de la segunda ecuación se deduce que $x_2 = 0$. Finalmente, reemplazando $x_2 = x_3 = 0$ en la primera ecuación, se obtiene que $x_1 = 0$.

Análogamente, será más fácil obtener las soluciones de cualquier sistema lineal que se encuentre en esta forma "triangular", es decir, de la forma

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n + \cdots + a_{1m}x_m & = & 0 \\ \qquad a_{22}x_2 + \cdots + a_{2n}x_n + \cdots + a_{2m}x_m & = & 0 \\ \qquad\qquad \vdots & & \\ \qquad\qquad a_{nn}x_n + \cdots + a_{nm}x_m & = & 0 \end{cases}$$

La idea de lo que sigue es ver cómo puede obtenerse, dado un sistema lineal arbitrario, un sistema de este tipo equivalente al dado.

La siguiente proposición caracteriza ciertas operaciones que producen sistemas equivalentes. En estas operaciones se basa el método de eliminación de Gauss (o método de triangulación) que utilizaremos para resolver sistemas lineales.

Proposición 1.19 *Dado un sistema lineal homogéneo de ecuaciones, los siguientes cambios en las ecuaciones dan lugar a sistemas equivalentes:*

1. *Intercambiar dos ecuaciones de lugar.*
2. *Multiplicar una ecuación por una constante no nula.*
3. *Reemplazar una ecuación por ella misma más un múltiplo de otra.*

Demostración.

1. Si vemos al conjunto de soluciones del sistema como la intersección de los conjuntos de soluciones de cada una de las ecuaciones que lo integran, intercambiar dos ecuaciones corresponde a intercambiar dos conjuntos en la intersección. Como la intersección es conmutativa, el conjunto que resulta es el mismo.
2. Sea $x = (x_1, \ldots, x_m) \in K^m$ una solución de

$$(*) \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m & = & 0 \\ \vdots & & \\ a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{im}x_m & = & 0 \\ \vdots & & \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m & = & 0 \end{cases}$$

Al multiplicar la i-ésima ecuación por $\lambda \in K$, $\lambda \neq 0$, resulta el sistema

$$(**) \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m & = & 0 \\ \vdots & & \\ \lambda a_{i1}x_1 + \lambda a_{i2}x_2 + \cdots + \lambda a_{im}x_m & = & 0 \\ \vdots & & \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m & = & 0 \end{cases}$$

Es claro que x es solución de todas las ecuaciones que no fueron modificadas. Además

$$\lambda a_{i1}x_1 + \lambda a_{i2}x_2 + \cdots + \lambda a_{im}x_m = \lambda(a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{im}x_m) = \lambda.\, 0 = 0.$$

Luego, x es solución de $(**)$.

Recíprocamente, multiplicando la i-ésima ecuación de $(**)$ por $\frac{1}{\lambda}$ se obtiene $(*)$, de donde, con el mismo razonamiento que antes, se deduce que si x es solución de $(**)$ también lo es de $(*)$.

3. Se demuestra en forma análoga. $\square$

Observación 1.20 Si A es la matriz asociada a un sistema lineal homogéneo H, efectuar las operaciones de la proposición anterior sobre las ecuaciones de H equivale a hacerlo sobre las filas de A.

Como consecuencia de esta observación, para resolver un sistema lineal trabajaremos con la matriz asociada al sistema, en lugar de hacerlo con las ecuaciones. Al aplicar en las matrices las operaciones dadas en la Proposición 1.19 estaremos obteniendo matrices cuyos sistemas lineales asociados son equivalentes al original.

El siguiente teorema nos asegura que, por medio de las operaciones permitidas siempre puede obtenerse un sistema triangular equivalente al dado. Más aún, de la demostración se desprende un algoritmo para realizar esta tarea.

Teorema 1.21 *Sea H un sistema lineal homogéneo de n ecuaciones con m incógnitas. Entonces, aplicando los cambios descriptos en la Proposición 1.19, puede obtenerse un sistema lineal homogéneo H' cuya matriz B es triangular superior, es decir, tal que $B_{ij} = 0$ si $i > j$.*

Demostración. Procedemos por inducción en n, la cantidad de ecuaciones del sistema.

Si $n = 1$ no hay nada que probar.

Supongamos que vale para n y consideremos un sistema lineal de $n + 1$ ecuaciones

$$\begin{cases} a_{11}x_1 + \cdots + a_{1m}x_m & = & 0 \\ \qquad\vdots & & \\ a_{n1}x_1 + \cdots + a_{nm}x_m & = & 0 \\ a_{n+1\,1}x_1 + \cdots + a_{n+1m}x_m & = & 0 \end{cases}$$

Si $m = 1$, es claro que el resultado vale. Supongamos $m > 1$.

Primer caso: Si $a_{i1} = 0$ para cada $1 \le i \le n + 1$. Entonces la matriz del sistema es de la forma

$$\begin{pmatrix} 0 & a_{12} & \cdots & a_{1m} \\ \vdots & \vdots & & \vdots \\ 0 & a_{n+1\,2} & \cdots & a_{n+1\,m} \end{pmatrix} = \begin{pmatrix} 0 & c \\ \bar{0} & M \end{pmatrix}$$

donde $\bar{0}$ denota una columna de ceros y $c \in K^{1\times(m-1)}$, $M \in K^{n\times(m-1)}$.

Segundo caso: Existe j, $1 \leq j \leq n+1$, con $a_{1j} \neq 0$. Eventualmente intercambiando las ecuaciones 1 y j, podemos suponer que $a_{11} \neq 0$. Multiplicando la primera ecuación por $\frac{1}{a_{11}}$ y aplicando operaciones de tipo 3. en las otras resulta

$$\begin{pmatrix} 1 & \frac{a_{12}}{a_{11}} & \cdots & \frac{a_{1m}}{a_{11}} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n+1\,1} & a_{n+1\,2} & \cdots & a_{n+1\,m} \end{pmatrix} \overset{F_i - a_{i1}F_1}{\longrightarrow} \begin{pmatrix} 1 & c \\ & \\ \bar{0} & M \end{pmatrix}$$

con $c \in K^{1\times(m-1)}$ y $M \in K^{n\times(m-1)}$.

Entonces, en cualquier caso, aplicando las operaciones descriptas en la Proposición 1.19 al sistema dado, puede obtenerse un sistema cuya matriz asociada es de la forma

$$A = \begin{pmatrix} a & c \\ & \\ \bar{0} & M \end{pmatrix} \quad \text{con } M \in K^{n\times(m-1)} \ \text{ y } \ a = 1 \text{ ó } a = 0.$$

Sea H_M el sistema cuya matriz asociada es M. Por hipótesis inductiva, aplicando operaciones permitidas puede obtenerse un sistema equivalente a H_M cuya matriz M' es triangular superior. Aplicando esas mismas operaciones en la matriz A se obtiene

$$B = \begin{pmatrix} a & c \\ & \\ \bar{0} & M' \end{pmatrix} \qquad \text{con } a = 1 \text{ ó } a = 0,$$

que es triangular superior. □

Ejemplo. Resolver el siguiente sistema lineal homogéneo en $\mathbb{R}^4$:

$$\begin{cases} 2x_2 - x_3 + x_4 = 0 \\ 3x_1 + x_2 + 10x_3 + 5x_4 = 0 \\ x_1 + 3x_3 + x_4 = 0 \end{cases}$$

La matriz asociada al sistema de ecuaciones es

$$A = \begin{pmatrix} 0 & 2 & -1 & 1 \\ 3 & 1 & 10 & 5 \\ 1 & 0 & 3 & 1 \end{pmatrix}.$$

El primer paso del método de Gauss consiste en colocar en el lugar A_{11} un elemento no nulo. Para eso permutamos las filas 1 y 3 de la matriz (podría usarse también la fila 2). Se obtiene

$$\begin{pmatrix} 1 & 0 & 3 & 1 \\ 3 & 1 & 10 & 5 \\ 0 & 2 & -1 & 1 \end{pmatrix}.$$

A continuación debemos realizar operaciones de fila de manera de conseguir que los restantes elementos de la primera columna de la matriz sean ceros. Si F_i denota la i-ésima fila de la matriz, haciendo $F_2 - 3F_1$ resulta

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

Pasamos ahora a la segunda columna de la matriz. El elemento ubicado en la fila 2 columna 2 de la matriz es un 1, con lo que sólo resta conseguir un 0 en la fila 3 columna 2. Para eso efectuamos $F_3 - 2F_2$:

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

Esta matriz se encuentra en forma triangular. El sistema asociado

$$\begin{cases} x_1 + 3x_3 + x_4 = 0 \\ x_2 + x_3 + 2x_4 = 0 \\ -3x_3 - 3x_4 = 0 \end{cases}$$

es equivalente al original.

De la tercera ecuación deducimos que si $X = (x_1, x_2, x_3, x_4)$ es solución del sistema, entonces $x_3 = -x_4$. Reemplazando en la segunda ecuación y despejando x_2 se obtiene $x_2 = -x_4$. Finalmente, de la primera ecuación se deduce que $x_1 = 2x_4$. Además es claro que cualquier X que cumple estas condiciones es solución de la ecuación.

En consecuencia, las soluciones del sistema son todos los vectores en $\mathbb{R}^4$ de la forma $X = (2x_4, -x_4, -x_4, x_4) = x_4(2, -1, -1, 1)$, es decir, el conjunto de las soluciones del sistema es el subespacio

$$S = < (2, -1, -1, 1) >.$$

1.2.3 Cantidad de soluciones de un sistema homogéneo

Una consecuencia inmediata del Teorema 1.21 es la siguiente:

Observación 1.22 Sea H un sistema lineal homogéneo de n ecuaciones con m incógnitas. Supongamos que $n > m$. Entonces, por el teorema anterior, el sistema es equivalente a uno cuya matriz es triangular superior. Luego, las últimas filas de su matriz asociada son nulas y en consecuencia vemos que existe un sistema H' de n ecuaciones con n incógnitas cuyo conjunto de soluciones coincide con el de H (basta considerar las n primeras ecuaciones del sistema obtenido).

Si H es un sistema lineal homogéneo con m incógnitas, es claro que $0 \in K^m$ es una solución de H. Ésta se llama la solución trivial del sistema. En muchos casos nos interesará saber si el sistema tiene alguna solución distinta de 0 (a las que llamaremos soluciones no triviales). El siguiente resultado nos dice que en el caso de un sistema con menos ecuaciones que incógnitas esto siempre sucede.

Teorema 1.23 *Sea H un sistema lineal homogéneo de n ecuaciones con m incógnitas. Supongamos que $n < m$. Entonces existe $x \in K^m$, $x \neq 0$, que es solución del sistema H.*

Demostración. Por inducción en la cantidad n de ecuaciones de H.

Si $n = 1, m \geq 2$: Entonces $H : a_{11}x_1 + a_{12}x_2 \cdots + a_{1m}x_m = 0$. Si $a_{11} = 0$, entonces $(1, 0, \ldots, 0)$ es solución del sistema y si $a_{11} \neq 0$, entonces $(\frac{-a_{12}}{a_{11}}, 1, 0, \ldots, 0)$ es solución.

Supongamos que el resultado vale para sistemas con n ecuaciones y sea H un sistema de $n + 1$ ecuaciones con m incógnitas, $n + 1 < m$.

Triangulando la matriz del sistema, resulta que es equivalente a una de la forma

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ 0 & & B & \end{pmatrix},$$

donde $B \in K^{n\times(m-1)}$, y $m - 1 > n$.

Por lo tanto, el sistema cuya matriz asociada es B está en las condiciones de la hipótesis inductiva. Luego, existe $(x_1, \ldots, x_{m-1}) \neq 0$ que es solución del sistema asociado a B.

- Si $a_{11} = 0$, entonces $(1, 0, \ldots, 0)$ es solución del sistema original.
- Si $a_{11} \neq 0$, entonces $\Big(-\frac{1}{a_{11}}.\big(\sum\limits_{i=2}^{m} a_{1i}x_{i-1}\big), x_1, \ldots, x_{m-1}\Big)$ es una solución no nula del sistema. □

El siguiente teorema se refiere a la existencia de soluciones no triviales para sistemas homogéneos con igual cantidad de ecuaciones que incógnitas. Teniendo en cuenta la observación hecha la comienzo de esta sección, esto resuelve el problema en el caso general.

Teorema 1.24 *Sea H un sistema lineal homogéneo de n ecuaciones y n incógnitas. Sea H' un sistema equivalente a H cuya matriz B es triangular superior. Entonces H tiene solución única si y sólo si $B_{ii} \neq 0 \;\forall\, 1 \leq i \leq n$.*

Demostración.

($\Leftarrow$) Supongamos que $B = \begin{pmatrix} B_{11} & \cdots & B_{1n} \\ 0 & \ddots & \vdots \\ \cdots & 0 & B_{nn} \end{pmatrix}$ con $B_{ii} \neq 0 \;\forall\, 1 \leq i \leq n$.

Entonces, la última ecuación del sistema H' es $B_{nn}x_n = 0$ y, como $B_{nn} \neq 0$, resulta que $x_n = 0$. Reemplazando en la ecuación anterior x_n por 0, queda $B_{n-1\,n-1}x_{n-1} = 0$, de donde $x_{n-1} = 0$.

Siguiendo de este modo, para cada $k = n - 2, \ldots, 1$ de la k-ésima ecuación se obtiene $x_k = 0$.

($\Rightarrow$) Supongamos que $B_{11} \neq 0, \ldots, B_{ii} \neq 0$ y $B_{i+1\,i+1} = 0$, o sea

$$\begin{pmatrix} B_{11} & & & \cdots & & B_{1n} \\ 0 & \ddots & & & & \vdots \\ \vdots & \ddots & B_{ii} & B_{i\,i+1} & \cdots & B_{in} \\ 0 & & 0 & 0 & & \\ \vdots & & \vdots & \vdots & M & \\ 0 & \cdots & 0 & 0 & & \end{pmatrix}$$

Es claro que $(1, 0, \ldots, 0)$ es solución del sistema cuya matriz asociada es $\begin{pmatrix} \overline{0} & M \end{pmatrix}$, o sea $x_{i+1} = 1, \ldots, x_n = 0$.

De la i-ésima ecuación se despeja $x_i = \frac{-B_{i\,i+1}}{B_{ii}}$.

Se sigue así para calcular los valores de todas las variables. Se obtiene una solución de H' de la forma $(x_1, \ldots, x_i, 1, 0, \ldots, 0)$, que es una solución no nula del sistema. $\square$

Ejemplo. Hallar todos los valores de $k \in \mathbb{R}$ para los cuales el sistema homogéneo cuya matriz asociada es $\begin{pmatrix} 1 & 2 & k-1 \\ 2 & -k+1 & 1 \\ k+1 & -4 & 1 \end{pmatrix}$ tiene solución única.

En primer término aplicamos el método de eliminación de Gauss para obtener un sistema triangular equivalente al dado:

$$\begin{pmatrix} 1 & 2 & k-1 \\ 2 & -k+1 & 1 \\ k+1 & -4 & 1 \end{pmatrix} \quad \underset{F_3-(k+1)F_1}{\overset{F_2-2F_1}{\longrightarrow}} \quad \begin{pmatrix} 1 & 2 & k-1 \\ 0 & -k-3 & -2k+3 \\ 0 & -2k-6 & -k^2+2 \end{pmatrix}$$

$$\overset{F_3-2F_2}{\longrightarrow} \quad \begin{pmatrix} 1 & 2 & k-1 \\ 0 & -k-3 & -2k+3 \\ 0 & 0 & -k^2+4k-4 \end{pmatrix}$$

Por el teorema anterior, el sistema tiene solución única si y sólo si $-k-3 \neq 0$ y $-k^2+4k-4 \neq 0$, es decir, para todo $k \in \mathbb{R} - \{-3, 2\}$.

1.2.4 Sistemas lineales no homogéneos.

Para terminar, estudiaremos sistemas de ecuaciones lineales en el caso general, es decir, cuando las ecuaciones que integran el sistema no están necesariamente igualadas a 0.

Definición 1.25 Un sistema de ecuaciones lineales

$$H : \left\{ \begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m & = & b_1 \\ \vdots & & \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m & = & b_n \end{array} \right.$$

se dice *no homogéneo* si existe i, $1 \leq i \leq n$, con $b_i \neq 0$.

La matriz $A = (a_{ij})$ se dice la matriz asociada al sistema.

Llamaremos *sistema homogéneo asociado a H* a

$$\left\{ \begin{array}{lcl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m & = & 0 \\ & \vdots & \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m & = & 0 \end{array} \right.$$

En el caso de un sistema lineal no homogéneo el conjunto de soluciones no es un subespacio (es claro que 0 no es solución). Sin embargo, el conjunto de soluciones de un sistema no homogéneo está íntimamente relacionado con el subespacio de soluciones del sistema homogéneo asociado.

Proposición 1.26 *Sea H un sistema lineal no homogéneo con soluciones. Sea S el conjunto de soluciones del sistema homogéneo asociado a H y sea p una solución particular de H. Entonces, el conjunto M de soluciones de H es $M = S + p = \{s + p : s \in S\}$.*

Demostración. Sea H el sistema

$$\left\{ \begin{array}{lcl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m & = & b_1 \\ & \vdots & \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m & = & b_n \end{array} \right.$$

($\subseteq$) Sea $z \in M$. Se tiene que $z = (z - p) + p$. Luego, para probar que $z \in S + p$, basta ver que $z - p = (z_1 - p_1, \ldots, z_m - p_m) \in S$, es decir, que es solución del sistema homogéneo asociado a H.

Sea i, $1 \leq i \leq n$. Entonces

$$\begin{aligned} a_{i1}(z_1 - p_1) + \cdots + a_{im}(z_m - p_m) &= (a_{i1}z_1 + \cdots + a_{im}z_m) - (a_{i1}p_1 + \cdots + a_{im}p_m) \\ &= b_i - b_i = 0 \end{aligned}$$

puesto que z y p son ambas soluciones de H. Luego, $z - p \in S$.

($\supseteq$) Sea $y \in S + p$. Entonces $y = s + p$ con $s \in S$.

Para cada $1 \leq i \leq n$,

$$\begin{aligned} a_{i1}y_1 + \cdots + a_{im}y_m &= a_{i1}(s_1 + p_1) + \cdots + a_{im}(s_m + p_m) = \\ &= (a_{i1}s_1 + \cdots + a_{im}s_m) + (a_{i1}p_1 + \cdots + a_{im}p_m) = 0 + b_i = b_i, \end{aligned}$$

puesto que p es solución de H y s es solución del sistema homogéneo asociado a H.

En consecuencia, y es solución de H, es decir, $y \in M$. $\square$

Ejemplo. Resolver el siguiente sistema de ecuaciones lineales en $\mathbb{R}^4$:

$$\begin{cases} 2x_2 - x_3 + x_4 = 0 \\ 3x_1 + x_2 + 10x_3 + 5x_4 = 3 \\ x_1 + 3x_3 + x_4 = 1 \end{cases}$$

Por la proposición anterior, para obtener todas las soluciones del sistema basta conocer una solución particular y el conjunto de soluciones del sistema homogéneo asociado.

Vemos que $p = (1, 0, 0, 0)$ es una solución particular del sistema.

Por otro lado, en un ejemplo anterior (página 16) hemos visto que el conjunto de soluciones del sistema homogéneo asociado es $S = <(2, -1, -1, 1)>$.

En consecuencia, el conjunto de soluciones del sistema es $<(2, -1, -1, 1)> + (1, 0, 0, 0)$.

Sin embargo, el resultado que relaciona las soluciones de un sistema no homogéneo con las del homogéneo asociado es más que nada teórico: dado un sistema de ecuaciones lineales no homogéneo, es poco probable que conozcamos una solución particular sin resolverlo. La resolución de un sistema lineal no homogéneo, al igual que en el caso homogéneo, puede realizarse triangulando una matriz adecuada como mostramos en el siguiente ejemplo (comparar con el ejemplo de la página 16).

Ejemplo. Resolver el siguiente sistema lineal no homogéneo en $\mathbb{R}^4$:

$$\begin{cases} 2x_2 - x_3 + x_4 = 2 \\ 3x_1 + x_2 + 10x_3 + 5x_4 = 1 \\ x_1 + 3x_3 + x_4 = -2 \end{cases}$$

Consideraremos la siguiente matriz formada por la matriz del sistema homogéneo asociado al sistema a la que le agregamos como última columna los escalares solución de cada ecuación (lo separamos con una línea para recordar que esos escalares son los que aparecen del otro lado de los iguales):

$$(A \mid b) = \left(\begin{array}{cccc|c} 0 & 2 & -1 & 1 & 2 \\ 3 & 1 & 10 & 5 & 1 \\ 1 & 0 & 3 & 1 & -2 \end{array}\right).$$

El método de resolución es similar al de los sistemas homogéneos. Utilizamos el método de Gauss para triangular la matriz que está a la izquierda de la línea *pero realizando las operaciones en toda la fila, inclusive en los elementos a la derecha de la línea*: el método de Gauss se basa en intercambiar y operar con ecuaciones, así que para no cambiar las soluciones debemos trabajar con ambos miembros de las ecuaciones (en el caso homogéneo, esto no era necesario porque siempre los segundos miembros daban cero). Entonces, triangulando con las mismas operaciones que en el ejemplo de la página 16, obtenemos

$$\left(\begin{array}{cccc|c} 0 & 2 & -1 & 1 & 2 \\ 3 & 1 & 10 & 5 & 1 \\ 1 & 0 & 3 & 1 & -2 \end{array}\right) \longrightarrow \left(\begin{array}{cccc|c} 1 & 0 & 3 & 1 & -2 \\ 3 & 1 & 10 & 5 & 1 \\ 0 & 2 & -1 & 1 & 2 \end{array}\right) \longrightarrow \left(\begin{array}{cccc|c} 1 & 0 & 3 & 1 & -2 \\ 0 & 1 & 1 & 2 & 7 \\ 0 & 2 & -1 & 1 & 2 \end{array}\right) \longrightarrow$$

$$\longrightarrow \left(\begin{array}{cccc|c} 1 & 0 & 3 & 1 & -2 \\ 0 & 1 & 1 & 2 & 7 \\ 0 & 0 & -3 & -3 & -12 \end{array}\right) \longrightarrow \left(\begin{array}{cccc|c} 1 & 0 & 3 & 1 & -2 \\ 0 & 1 & 1 & 2 & 7 \\ 0 & 0 & 1 & 1 & 4 \end{array}\right).$$

Esta matriz se encuentra en forma triangular y su sistema no homogéneo asociado

$$\begin{cases} x_1 + 3x_3 + x_4 = -2 \\ x_2 + x_3 + 2x_4 = 7 \\ x_3 + x_4 = 4 \end{cases}$$

es equivalente al original.

De la tercera ecuación deducimos que si $X = (x_1, x_2, x_3, x_4)$ es solución del sistema, entonces $x_3 = 4 - x_4$. Reemplazando en la segunda ecuación y despejando x_2 se obtiene $x_2 = 3 - x_4$. Finalmente, de la primera ecuación se deduce que $x_1 = -14 + 2x_4$. Además es claro que cualquier X que cumple estas condiciones es solución del sistema. Luego, las soluciones del sistema son todos los vectores en $\mathbb{R}^4$ de la forma

$$X = (2x_4 - 14, -x_4 + 3, -x_4 + 4, x_4) = x_4(2, -1, -1, 1) + (-14, 3, 4, 0),$$

es decir, el conjunto de las soluciones del sistema es el subespacio $S = < (2, -1, -1, 1) >$ (solución del sistema homogéneo asociado) más la solución particular $(-14, 3, 4, 0)$.

Este procedimiento para resolver sistemas lineales no homogéneos motiva la siguiente definición:

Definición 1.27 Dado un sistema de ecuaciones lineales no homogéneo

$$H : \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m & = & b_1 \\ \vdots & & \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m & = & b_n \end{cases},$$

se llama *matriz ampliada asociada al sistema H* a la matriz

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1m} & b_1 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} & b_n \end{array}\right).$$

A diferencia de los sistemas homogéneos, los sistemas no homogéneos pueden no tener soluciones. El método descripto, que triangula la matriz ampliada asociada al sistema, nos muestra que en estos casos no hay solución particular posible:

Ejemplo. Resolver el siguiente sistema lineal no homogéneo en $\mathbb{R}^4$:

$$\begin{cases} x_1 + 2x_2 + x_3 - x_4 = 2 \\ 3x_1 + x_2 - x_3 - x_4 = 7 \\ 5x_1 - 3x_3 - x_4 = 5 \end{cases}$$

Triangulando la matriz ampliada asociada al sistema, tenemos

$$\left(\begin{array}{cccc|c} 1 & 2 & 1 & -1 & 2 \\ 3 & 1 & -1 & -1 & 7 \\ 5 & 0 & -3 & -1 & 5 \end{array}\right) \longrightarrow \left(\begin{array}{cccc|c} 1 & 2 & 1 & -1 & 2 \\ 0 & -5 & -4 & 2 & 1 \\ 5 & 0 & -3 & -1 & 5 \end{array}\right) \longrightarrow$$

$$\longrightarrow \left(\begin{array}{cccc|c} 1 & 2 & 1 & -1 & 2 \\ 0 & -5 & -4 & 2 & 1 \\ 0 & -10 & -8 & 4 & -5 \end{array}\right) \longrightarrow \left(\begin{array}{cccc|c} 1 & 2 & 1 & -1 & 2 \\ 0 & -5 & -4 & 2 & 1 \\ 0 & 0 & 0 & 0 & -7 \end{array}\right).$$

Esto significa que una solución $X = (x_1, x_2, x_3, x_4)$ del sistema debe satisfacer la última ecuación, es decir $0.x_1 + 0.x_2 + 0.x_3 + 0.x_4 = -7$, lo que es un absurdo. Por lo tanto el sistema en cuestión no tiene soluciones.

1.3 Independencia lineal y bases

En la Sección 1.1.4 introdujimos la noción de sistema de generadores de un K-espacio vectorial V. Un espacio vectorial puede tener distintos sistemas de generadores y además dos sistemas de generadores de un mismo espacio vectorial pueden tener distinta cantidad de elementos.

En esta sección veremos que para cualquier sistema de generadores G de un K-espacio vectorial V que cumpla cierta propiedad adicional, que llamaremos independencia lineal, la cantidad de elementos de G estará fija. Esto nos llevará a definir la noción de dimensión de un espacio vectorial.

1.3.1 Independencia lineal

Una cuestión que surge al considerar un sistema de generadores de un K-espacio vectorial V es la de hallar sistemas de generadores que sean minimales respecto de la inclusión, es decir, tal que ningún subconjunto propio sea también un sistema de generadores de V. Los siguientes resultados caracterizan a los conjuntos con esta propiedad.

Proposición 1.28 *Sean V un K-espacio vectorial, S un subespacio de V y $\{v_1, \dots, v_n\} \subseteq V$. Entonces $< v_1, \dots, v_n > \subseteq S \iff v_i \in S \;\forall 1 \le i \le n$.*

Demostración.

($\Rightarrow$) Para cada $1 \le i \le n$,

$$v_i = 0.v_1 + \cdots + 0.v_{i-1} + 1.v_i + 0.v_{i+1} + \cdots + 0.v_n \in < v_1, \dots, v_n > \subseteq S,$$

de donde $v_i \in S$.

($\Leftarrow$) Como $v_1, \dots, v_n \in S$ y S es un subespacio, entonces $\sum_{i=1}^{n} \alpha_i v_i \in S \;\forall \alpha_i \in K$. Luego, $< v_1, \dots, v_n > \subseteq S$. □

Corolario 1.29 *Sea V un K-espacio vectorial, y sea $\{v_1, \dots, v_n, v_{n+1}\} \subseteq V$. Entonces $< v_1, \dots, v_n, v_{n+1} > = < v_1, \dots, v_n > \iff v_{n+1} \in < v_1, \dots, v_n >$.*

Demostración.

($\Rightarrow$) Se tiene $< v_1, \dots, v_n, v_{n+1} > \subseteq < v_1, \dots, v_n >$. Entonces, por la proposición anterior, $v_{n+1} \in < v_1, \dots, v_n >$.

($\Leftarrow$) Por hipótesis, $v_{n+1} \in < v_1, \dots, v_n >$. Además $v_i \in < v_1, \dots, v_n > \forall\, 1 \le i \le n$. Entonces, $< v_1, \dots, v_n, v_{n+1} > \subseteq < v_1, \dots, v_n >$.

Por otro lado, $v_i \in < v_1, \dots, v_{n+1} > \forall\, 1 \le i \le n$, y entonces vale

$$< v_1, \dots, v_n, v_{n+1} > \supseteq < v_1, \dots, v_n >.$$

Luego $< v_1, \dots, v_n, v_{n+1} > = < v_1, \dots, v_n >$. □

Introducimos ahora la noción de independencia lineal.

Definición 1.30 Sea V un K-espacio vectorial y sea $\{v_\alpha\}_{\alpha \in I}$ una familia de vectores de V. Se dice que $\{v_\alpha\}_{\alpha \in I}$ es *linealmente independiente* (l.i.) si

$$\sum_{\alpha \in I} a_\alpha . v_\alpha = 0 \Rightarrow a_\alpha = 0 \quad \forall\, \alpha \in I.$$

Si $\{v_\alpha\}_{\alpha \in I}$ no es linealmente independiente, se dice que es *linealmente dependiente* (l.d.).

Aunque, a diferencia de un conjunto, una *familia* puede contener elementos repetidos, en lo que sigue hablaremos indistintamente de familias o conjuntos de vectores, entendiendo que pueden ocurrir repeticiones.

La noción de independencial lineal está íntimamente relacionada con la minimalidad de un sistema de generadores. Más precisamente:

Observación 1.31 Sea V un K-espacio vectorial y sean $v_1, \dots, v_n \in V$. Entonces el conjunto $\{v_1, \dots, v_n\}$ es linealmente independiente si y sólo si

$$< v_1, \dots, v_n > \ne < v_1, \dots, \widehat{v_i}, \dots, v_n > \forall\, 1 \le i \le n.$$

(Notación: $< v_1, \dots, \widehat{v_i}, \dots, v_n >$ denota el subespacio generado por $\{v_1, \dots, v_n\} - \{v_i\}$.)

Demostración.

($\Rightarrow$) Supongamos que $< v_1, \dots, \widehat{v_i}, \dots, v_n > = < v_1, \dots, v_n >$. En particular

$$v_i \in < v_1, \dots, \widehat{v_i}, \dots, v_n >,$$

es decir, existen $\alpha_j \in K$ $(j \neq i)$ tales que $v_i = \sum\limits_{\substack{j=1 \\ j\neq i}}^{n} \alpha_j v_j$. Entonces

$$0 = \sum_{j=1}^{i-1} \alpha_j v_j + (-1)v_i + \sum_{j=i+1}^{n} \alpha_j v_j,$$

de donde $\{v_1, \dots, v_n\}$ no es linealmente independiente.

($\Leftarrow$) Si $\{v_1, \dots, v_n\}$ es linealmente dependiente, existen $\alpha_1, \dots, \alpha_n \in K$ no todos nulos, tales que $\sum\limits_{i=1}^{n} \alpha_i v_i = 0$. Sin pérdida de generalidad, supongamos que $\alpha_n \neq 0$. Entonces

$$v_n = -\sum_{i=1}^{n-1} \frac{\alpha_i}{\alpha_n}.v_i \ \in \ < v_1, \dots, v_{n-1} >.$$

Luego, $< v_1, \dots, v_n > = < v_1, \dots, v_{n-1} >$. $\square$

Ejemplos. Decidir si los siguientes conjuntos son linealmente independientes.

1. En $\mathbb{R}^3$, $\{(1,0,1), (1,-1,0), (0,0,1)\}$.

   Sean $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$ tales que

   $$\alpha_1(1,0,1) + \alpha_2(1,-1,0) + \alpha_3(0,0,1) = (0,0,0).$$

   Comparando coordenada a coordenada resulta que $\alpha_1, \alpha_2, \alpha_3$ son solución del sistema de ecuaciones

   $$\begin{cases} \alpha_1 + \alpha_2 &= 0 \\ \quad -\alpha_2 &= 0 \\ \alpha_1 + \alpha_3 &= 0 \end{cases}$$

   Es fácil ver que este sistema tiene como única solución a la trivial.

   Luego, el conjunto $\{(1,0,1), (1,-1,0), (0,0,1)\}$ es linealmente independiente.

2. En $\mathbb{R}[X]$, $\{X^i : i \in \mathbb{N}_0\}$.

   Sean $\alpha_i \in \mathbb{R}$ $(i \in \mathbb{N}_0)$ tales que $\alpha_i = 0$ para casi todo $i \in \mathbb{N}_0$ y $\sum\limits_{i\in\mathbb{N}_0} \alpha_i X^i = 0$.

   Para que el elemento $\sum\limits_{i\in\mathbb{N}_0} \alpha_i X^i$ de $\mathbb{R}[X]$ sea el polinomio nulo, todos sus coeficientes deben ser 0. Luego, $\alpha_i = 0$ para todo $i \in \mathbb{N}_0$, de donde el conjunto $\{X^i : i \in \mathbb{N}_0\}$ es linealmente independiente.

La siguiente proposición nos permitirá obtener otro método para decidir si un conjunto de vectores en K^n es linealmente independiente.

Proposición 1.32 *Sea V un K-espacio vectorial. Entonces:*

1. $\{v_1, \dots, v_i, \dots, v_j, \dots, v_n\} \subseteq V$ *es l.i.* $\iff$ $\{v_1, \dots, v_j, \dots, v_i, \dots, v_n\} \subseteq V$ *es l.i.*

2. $\{v_1, \dots, v_i, \dots, v_n\} \subseteq V$ *es l.i.* $\iff$ $\{v_1, \dots, \lambda v_i, \dots, v_n\} \subseteq V$ *es l.i. para* $\lambda \in K - \{0\}$.

3. $\{v_1, \dots, v_i, \dots, v_j, \dots, v_n\} \subseteq V$ *es l.i.* $\iff$ $\{v_1, \dots, v_i + \lambda v_j, \dots, v_j, \dots, v_n\} \subseteq V$ *es l.i. para* $\lambda \in K$.

Demostración.

1. Se deduce del hecho que en un conjunto no interesa el orden de sus elementos.

2. Supongamos que $\{v_1, \dots, v_i, \dots, v_n\}$ es linealmente independiente.

   Sean $\alpha_1, \dots, \alpha_n \in K$ tales que $\alpha_1 v_1 + \dots + \alpha_i(\lambda v_i) + \dots + \alpha_n v_n = 0$. Entonces se tiene que $\alpha_j = 0$ para cada $j \neq i$ y que $\alpha_i.\lambda = 0$. Puesto que $\lambda \neq 0$, resulta que también $\alpha_i = 0$.

   Luego, el conjunto $\{v_1, \dots, \lambda v_i, \dots, v_n\}$ es linealmente independiente.

   Esto prueba la equivalencia, puesto que para demostrar la otra implicación basta multiplicar el i-ésimo vector del conjunto por $\frac{1}{\lambda}$.

3. Supongamos que $\{v_1, \dots, v_i, \dots, v_j, \dots, v_n\}$ es linealmente independiente.

   Sean $\alpha_1, \dots, \alpha_n \in K$ tales que

$$\begin{aligned} 0 &= \alpha_1 v_1 + \dots + \alpha_i(v_i + \lambda v_j) + \dots + \alpha_j v_j + \dots + \alpha_n v_n \\ &= \alpha_1 v_1 + \dots + \alpha_i v_i + \dots + (\alpha_i \lambda + \alpha_j) v_j + \dots + \alpha_n v_n. \end{aligned}$$

   La independencia lineal de $\{v_1, \dots, v_i, \dots, v_j, \dots, v_n\}$ implica que

$$\alpha_1 = \dots = \alpha_i = \dots = \alpha_i \lambda + \alpha_j = \dots = \alpha_n = 0,$$

   de donde $\alpha_k = 0$ para todo $1 \leq k \leq n$.

   En consecuencia, el conjunto $\{v_1, \dots, v_i + \lambda v_j, \dots, v_j, \dots, v_n\}$ es linealmente independiente.

   La otra implicación se deduce de ésta observando que el conjunto $\{v_1, \dots, v_n\}$ se obtiene de $\{v_1, \dots, v_i + \lambda v_j, \dots, v_j, \dots, v_n\}$ cambiando el i-ésimo vector $v_i + \lambda v_j$ por $(v_i + \lambda v_j) + (-\lambda) v_j = v_i$. □

Como consecuencia de la proposición anterior, para decidir si un subconjunto de vectores $\{v_1, \dots, v_r\}$ de K^n es linealmente independiente podemos proceder como sigue:

- Considerar la matriz A cuyas filas son los vectores $v_1, \dots, v_r$.

- Triangular la matriz A.

- Si la matriz obtenida tiene alguna fila nula, el conjunto es linealmente dependiente. De lo contrario, es linealmente independiente.

  En efecto, en cada paso de la triangulación, lo que se hace es cambiar el conjunto de vectores por otro conjunto como en 1., 2. o 3. de la proposición anterior. Luego, el nuevo conjunto de vectores será l.i. si y sólo si el anterior era l.i. Si alguna fila de la matriz obtenida es nula, es decir, uno de los vectores del conjunto de vectores obtenido es el 0, es claro que el conjunto es l.d. Por otro lado, si ninguna fila de la matriz triangular superior es nula, es fácil ver que el conjunto de vectores obtenido es l.i.

1.3.2 Bases y dimensión

Introducimos ahora el concepto de base de un espacio vectorial.

Definición 1.33 Sea V un K-espacio vectorial. Una familia $\{v_\alpha\}_{\alpha\in I}$ se llama una *base* del espacio vectorial V si $\{v_\alpha\}_{\alpha\in I}$ es una familia linealmente independiente de V que satisface $< v_\alpha >_{\alpha\in I} = V$.

Ejemplos.

1. En K^n, $B = \{e_1, \dots, e_n\}$, donde $(e_i)_i = 1$ y $(e_i)_j = 0$ si $j \neq i$, es una base, llamada la *base canónica de* K^n.
2. En $K^{n\times m}$, $B = \{E^{ij} \ / \ 1 \le i \le n, 1 \le j \le m\}$ es una base.
3. En $K[X]$, $B = \{X^i \ / \ i \in \mathbb{N}_0\}$ es una base.

Dos sistemas de generadores cualesquiera de un K-espacio vectorial V pueden tener distinta cantidad de elementos. Esto no sucede en el caso de dos bases y lo demostraremos para espacios vectoriales finitamente generados, lo que nos permitirá definir la dimensión de un espacio vectorial finitamente generado como la cantidad de elementos de una base cualquiera.

Teorema 1.34 *Sea V un K-espacio vectorial. Supongamos que $< v_1, \dots, v_r > = V$ y que $\{w_1, \dots, w_s\} \subseteq V$ es una familia linealmente independiente. Entonces $s \le r$.*

Demostración. Como $V = < v_1, \dots, v_r >$, para cada $1 \le i \le s$, existen $\alpha_{ij} \in K$ $(1 \le j \le r)$ tales que $w_i = \sum\limits_{j=1}^{r} \alpha_{ij} v_j$. Consideremos el siguiente sistema de r ecuaciones y s incógnitas:

$$\sum_{h=1}^{s} \alpha_{hj} x_h = 0 \qquad 1 \le j \le r. \tag{1.1}$$

Sea $(\beta_1, \dots, \beta_s)$ una solución del sistema. Entonces

$$\sum_{h=1}^{s} \beta_h w_h = \sum_{h=1}^{s} \beta_h \Big(\sum_{j=1}^{r} \alpha_{hj} v_j \Big) = \sum_{h=1}^{s} \Big(\sum_{j=1}^{r} \beta_h \alpha_{hj} v_j \Big) =$$

$$= \sum_{j=1}^{r} \Big(\sum_{h=1}^{s} \beta_h \alpha_{hj} v_j \Big) = \sum_{j=1}^{r} \Big(\sum_{h=1}^{s} \beta_h \alpha_{hj} \Big) v_j = 0.$$

Dado que $\{w_1, \dots, w_s\}$ es linealmente independiente, debe ser $(\beta_1, \dots, \beta_s) = 0$.

En consecuencia, el sistema (1.1) tiene solución única, de donde se deduce que la cantidad de ecuaciones del sistema es mayor o igual que el número de variables, es decir $r \geq s$. $\square$

Corolario 1.35 *Sea V un K-espacio vectorial, y sean B_1 y B_2 dos bases de V. Si $B_1 = \{w_1, \dots, w_n\}$ y $B_2 = \{v_1, \dots, v_m\}$, entonces $n = m$.*

Demostración. Por el teorema anterior

- B_1 sistema de generadores de V y B_2 conjunto linealmente independiente $\Longrightarrow n \geq m$.
- B_2 sistema de generadores de V y B_1 conjunto linealmente independiente $\Longrightarrow m \geq n$.

Luego, $n = m$. $\square$

Definición 1.36 Sea V un K-espacio vectorial y sea $B = \{v_1, \dots, v_n\}$ una base de V. Diremos entonces que n es la *dimensión* de V (como espacio vectorial sobre K). En este caso, diremos que V es un K-espacio vectorial de *dimensión finita*, para distinguirlo de los espacios vectoriales que no admiten una base con finitos elementos. Por convención, la dimensión de $\{0\}$ es 0.

Notación. Si n es la dimensión del K-espacio vectorial V, escribimos $n = \dim_K V$, o simplemente $\dim V$ si el cuerpo K queda claro por el contexto.

Una propiedad de las bases es que cualquier vector del espacio vectorial considerado se puede expresar como combinación lineal de los elementos de la base de manera única. Como veremos más adelante, aplicando esta propiedad se trabajará en un K-espacio vectorial de dimensión n arbitrario como si fuese K^n.

Proposición 1.37 *Sea V un K-espacio vectorial de dimensión finita. Sea $\{v_1, \dots, v_n\}$ una base de V. Entonces para cada $x \in V$ existen únicos $\alpha_1, \dots, \alpha_n \in K$ tales que $x = \sum\limits_{i=1}^{n} \alpha_i v_i$.*

Demostración. La existencia se deduce de que, por ser una base de V, $\{v_1, \dots, v_n\}$ es un sistema de generadores de V.

Supongamos que $\sum\limits_{i=1}^{n} \alpha_i v_i = \sum\limits_{i=1}^{n} \beta_i v_i$, entonces $\sum\limits_{i=1}^{n} (\alpha_i - \beta_i) v_i = 0$. Como $\{v_1, \dots, v_n\}$ es un conjunto linealmente independiente, $\alpha_i - \beta_i = 0 \ \forall\, 1 \leq i \leq n$. Luego, $\alpha_i = \beta_i \ \forall 1 \leq i \leq n$, lo que prueba la unicidad. $\square$

La siguiente proposición muestra cómo hallar una base de un K-espacio vectorial de dimensión finita V a partir de cualquier sistema de generadores finito de V y cómo completar un subconjunto linealmente independiente arbitrario de V a una base.

Proposición 1.38 *Sea V un K-espacio vectorial de dimensión finita.*

i) Sea $\{v_1, \dots, v_s\}$ un sistema de generadores de V. Entonces existe un subconjunto $G \subseteq \{v_1, \dots, v_s\}$ que es una base de V.

ii) Sea $\{w_1, \dots, w_r\}$ un conjunto linealmente independiente de V. Entonces existen elementos $w_{r+1}, \dots, w_n \in V$ tales que $\{w_1, \dots, w_r, w_{r+1}, \dots, w_n\}$ es una base de V.

Demostración.

i) Si $\{v_1, \dots, v_s\}$ es linealmente independiente, entonces es una base de V.

Si no es linealmente independiente, alguno de los vectores del conjunto es combinación lineal de los otros. Supongamos que $v_s \in < v_1, \dots, v_{s-1} >$. Consideramos ahora $\{v_1, \dots, v_{s-1}\}$, que es un sistema de generadores de V, y procedemos inductivamente.

ii) Sea $B = \{z_1, \dots, z_n\}$ una base de V.

Sea $G_0 = < w_1, \dots, w_r >$. Consideramos

$$G_1 := \begin{cases} \{w_1, \dots, w_r, z_1\} & \text{si } z_1 \notin < G_0 > \\ \{w_1, \dots, w_r\} & \text{si } z_1 \in < G_0 >. \end{cases}$$

Se procede inductivamente para $2 \le i \le n$, es decir,

$$G_i := \begin{cases} G_{i-1} \cup \{z_i\} & \text{si } z_i \notin < G_{i-1} > \\ G_{i-1} & \text{si } z_i \in < G_{i-1} >. \end{cases}$$

Observar que $\{w_1, \dots, w_r\} \subseteq G_i \ \forall\, 1 \le i \le n$.

Además, para cada $1 \le i \le n$, $< z_1, \dots, z_i > \subseteq < G_i >$, y G_i es un conjunto linealmente independiente. En particular, $V = < z_1, \dots, z_n > \subseteq < G_n >$ y G_n es linealmente independiente. Luego, G_n es una base de V. □

Ejemplos.

1. Extraer una base de $S = < (1, -1, 7, 3), (2, 1, -1, 0), (3, 1, 1, 1) >$ del sistema de generadores dado.

   Observamos que el sistema de generadores dado es linealmente dependiente. En efecto,

$$\begin{pmatrix} 1 & -1 & 7 & 3 \\ 2 & 1 & -1 & 0 \\ 3 & 1 & 1 & 1 \end{pmatrix} \quad \begin{matrix} F_2 - 2F_1 \\ \rightarrow \\ F_3 - 3F_1 \end{matrix} \quad \begin{pmatrix} 1 & -1 & 7 & 3 \\ 0 & 3 & -15 & -6 \\ 0 & 4 & -20 & -8 \end{pmatrix}$$

$$\begin{matrix} \\ \rightarrow \\ F_3 - \frac{4}{3}F_2 \end{matrix} \quad \begin{pmatrix} 1 & -1 & 7 & 3 \\ 0 & 3 & -15 & -6 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Como por la triangulación anterior se ve simultáneamente que $\{(1,-1,7,3),(2,1,-1,0)\}$ es un conjunto l.i. y que $\{(1,-1,7,3),(2,1,-1,0),(3,1,1,1)\}$ es un conjunto l.d, resulta que $(3,1,1,1) \in <(1,-1,7,3),(2,1,-1,0)>$.

Luego, $\{(1,-1,7,3),(2,1,-1,0)\}$ es un sistema de generadores de S. Como además es linealmente independiente, es una base de S.

2. Extender el conjunto linealmente independiente $\{(1,1,0,0),(1,-1,1,0)\}$ a una base de $\mathbb{R}^4$.

   Consideremos la base canónica de $\mathbb{R}^4$,

   $$E = \{(1,0,0,0),(0,1,0,0),(0,0,1,0),(0,0,0,1)\}.$$

   Con la notación utilizada en la demostración de la proposición anterior:

   Se tiene $G_1 := \{(1,1,0,0),(1,-1,1,0),(1,0,0,0)\}$, que es linealmente independiente.

   Ahora, $(0,1,0,0) \in <G_1>$, puesto que $(0,1,0,0) = (1,1,0,0)-(1,0,0,0)$ y entonces $G_2 := G_1$.

   El conjunto $G_2 \cup \{(0,0,1,0)\}$ es linealmente independiente. Consideramos entonces $G_3 := G_2 \cup \{(0,0,1,0)\}$. Este conjunto, formado por cuatro vectores linealmente independientes de $\mathbb{R}^4$, debe poder extenderse a una base de $\mathbb{R}^4$, que tendrá 4 elementos puesto que $\dim \mathbb{R}^4 = 4$; luego, ya es una base de $\mathbb{R}^4$.

Como consecuencia de la proposición anterior, se obtienen los siguientes resultados sobre subespacios de un K-espacio vectorial de dimensión finita.

Observación 1.39 Si V es un K-espacio vectorial de dimensión finita y $S \subseteq V$, entonces S es de dimensión finita.

(Notar que si S tuviese una base con infinitos elementos, podríamos obtener $\dim V + 1$ elementos l.i. en S y por lo tanto en V. Este conjunto podría extenderse a una base de V con más de $\dim V$ elementos, lo que es un absurdo.)

Proposición 1.40 *Sean S y T subespacios de un K-espacio vectorial V de dimensión finita. Entonces:*

i) $S \subseteq T \;\Rightarrow\; \dim S \le \dim T$.

ii) $S \subseteq T$ *y* $\dim S = \dim T \;\Rightarrow\; S = T$.

Demostración.

i) Sea $\{s_1,\dots,s_r\}$ una base de S y sea $n = \dim T$. Como $S \subseteq T$, se tiene que $\{s_1,\dots,s_r\} \subseteq T$, y además es un conjunto linealmente independiente. Luego, puede extenderse a una base de T, y en consecuencia, $\dim S = r \le n = \dim T$.

ii) Siguiendo el razonamiento de la demostración de *i*), al extender una base $\{s_1,\dots,s_r\}$ de S a una de T, como $\dim S = \dim T$, no se agrega ningún vector. Luego $S = <s_1,\dots,s_r> = T$. □

Observar que el ítem ii) de la proposición anterior nos facilita la verificación de la igualdad entre dos subespacios.

Ejemplo. Sean S y T los subespacios de $\mathbb{R}^3$: $S = < (1, -k^2 + 1, 2), (k + 1, 1 - k, -2) >$ y $T = \{x \in \mathbb{R}^3 : x_1 + x_2 + x_3 = 0\}$. Hallar todos los valores de $k \in \mathbb{R}$ para los cuales $S = T$.

En primer lugar, veamos para qué valores de $k \in \mathbb{R}$ se tiene que $S \subset T$:

- $(1, -k^2 + 1, 2) \in T \iff 1 + (-k^2 + 1) + 2 = 0 \iff k = \pm 2$
- $(k + 1, 1 - k, -2) \in T$ para todo $k \in \mathbb{R}$.

Luego, $S \subset T$ si y sólo si $k = -2$ o $k = 2$.

Finalmente, para cada uno de estos valores de k, basta ver si $\dim S = \dim T$. Observar que $\dim T = 2$ (una base de T es $\{(-1, 1, 0), (-1, 0, 1)\}$).

- Si $k = -2$, $S = < (1, -3, 2), (-1, 3, -2) > = < (1, -3, 2) >$, de donde $\dim S = 1$.
- Si $k = 2$, $S = < (1, -3, 2), (3, -1, -2) >$ y, como $\{(1, -3, 2), (3, -1, -2)\}$ es l.i. y por lo tanto una base de S, se tiene que $\dim S = 2$.

Concluimos que $S = T$ si y sólo si $k = 2$.

1.4 Suma de subespacios

Dados dos subespacios S y T de un K-espacio vectorial V la unión $S \cup T$ en general no es un subespacio de V, porque no contiene necesariamente a todos los elementos de la forma $s + t$ con $s \in S$ y $t \in T$, y un subespacio que contenga a S y a T debe contener a todos estos elementos. Esto da lugar a la noción de suma de subespacios.

1.4.1 Subespacio suma

Definición 1.41 Sea V un K-espacio vectorial, y sean S y T subespacios de V. Se llama *suma de S y T* al conjunto $S + T = \{v \in V \ / \ \exists x \in S, y \in T \text{ tales que } v = x + y\} = \{x + y \ / \ x \in S, y \in T\}$.

La siguiente proposición muestra que la suma de dos subespacios es, en efecto, un subespacio que contiene a ambos, y da una caracterización de este conjunto en términos de sistemas de generadores de los subespacios considerados.

Proposición 1.42 *Sea V un K-espacio vectorial, y sean S y T subespacios de V. Entonces:*

i) $S + T$ *es un subespacio de* V.

ii) $S + T$ *es el menor subespacio (con respecto a la inclusión) que contiene a* $S \cup T$.

iii) Si $\{v_i\}_{i \in I}$ *es un sistema de generadores de* S *y* $\{w_j\}_{j \in J}$ *es un sistema de generadores de* T, $\{v_i\}_{i \in I} \cup \{w_j\}_{j \in J}$ *es un sistema de generadores de* $S + T$.

Demostración.

i) $0 = 0 + 0 \in S + T$, pues $0 \in S$, $0 \in T$.

Sean $v, v' \in S + T$. Existen $x, x' \in S$, $y, y' \in T$ tales que $v = x + y$, $v' = x' + y'$. Entonces $v + v' = (x + y) + (x' + y') = (x + x') + (y + y')$, y como S y T son subespacios $x + x' \in S$, $y + y' \in T$. Luego, $v + v' \in S + T$.

Sea $v \in S + T$ y sea $\lambda \in K$. Existen $x \in S, y \in T$ tales que $v = x + y$. Entonces, $\lambda.v = \lambda.(x + y) = \lambda.x + \lambda.y$. Como $\lambda \in K$, $x \in S$ y S es un subespacio, resulta que $\lambda.x \in S$. Análogamente, $\lambda.y \in T$. Luego $\lambda.v \in S + T$.

En consecuencia, $S + T$ es un subespacio de V.

ii) Sea W un subespacio de V tal que $S \cup T \subseteq W$.

Sea $v \in S + T$. Entonces $v = x + y$ con $x \in S$, $y \in T$. Como $S \subseteq S \cup T \subseteq W$, entonces $x \in W$; y como $T \subseteq S \cup T \subseteq W$, entonces $y \in W$. En consecuencia $v = x + y \in W$, puesto que W es un subespacio.

Luego, $S + T \subseteq W$.

iii) Sea $v \in S + T$, $v = x + y$ con $x \in S$, $y \in T$. Dado que $\{v_i\}_{i \in I}$ es un sistema de generadores de S, existen $\alpha_i \in K$ $(i \in I)$, con $\alpha_i = 0$ salvo para finitos $i \in I$, tales que $x = \sum\limits_{i \in I} \alpha_i v_i$. De la misma manera, existen $\beta_j \in K$ $(j \in J)$, con $\beta_j = 0$ salvo para finitos $j \in J$, tales que $y = \sum\limits_{j \in J} \beta_j w_j$. Luego

$$v = \sum_{i \in I} \alpha_i v_i + \sum_{j \in J} \beta_j w_j$$

resulta una combinación lineal de $\{v_i\}_{i \in I} \cup \{w_j\}_{j \in J} \subseteq S + T$. $\square$

Ejemplo. Sean S y T los subespacios de $\mathbb{R}^4$

$$S = < (1, 1, 0, 1), (2, 3, 1, 1) > \quad \text{y} \quad T = < (0, 0, 1, 1), (1, 2, 2, 1) >.$$

Hallar una base de $S + T$.

Por la proposición anterior, podemos obtener un sistema de generadores de $S+T$ mediante la unión de un sistema de generadores de S y un sistema de generadores de T. Entonces

$$S + T = < (1, 1, 0, 1), (2, 3, 1, 1), (0, 0, 1, 1), (1, 2, 2, 1) >.$$

Ahora extraemos una base del sistema de generadores hallado. Se tiene:

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

Esta triangulación muestra simultáneamente que el conjunto $\{(1,1,0,1),(2,3,1,1),(0,0,1,1),(1,2,2,1)\}$ es l.d. y que el conjunto $\{(1,1,0,1),(2,3,1,1),(0,0,1,1)\}$ es l.i. Por lo tanto, $(1,2,2,1) \in \ <(1,1,0,1),(2,3,1,1),(0,0,1,1)>$ y $\{(1,1,0,1),(2,3,1,1),(0,0,1,1)\}$ es una base de $S+T$.

Si S y T son dos subespacios de dimensión finita de un K-espacio vectorial V, el siguiente teorema relaciona las dimensiones de los subespacios S, T, $S\cap T$ y $S+T$.

Teorema 1.43 (Teorema de la dimensión para la suma de subespacios.) *Sea V un K-espacio vectorial. Sean S y T subespacios de V de dimensión finita. Entonces*

$$\dim(S+T) = \dim S + \dim T - \dim(S\cap T).$$

Demostración. Sean $s = \dim S$, $t = \dim T$ y $r = \dim(S\cap T)$.

Si $s = 0$, o sea $S = \{0\}$, se tiene que $S+T = T$ y $S\cap T = \{0\}$ y la igualdad vale. Análogamente se ve que vale si $t = 0$.

Sea $\{v_1,\dots,v_r\}$ una base de $S\cap T$ (si $r = 0$, consideramos simplemente el conjunto vacío).

Sean $w_{r+1},\dots,w_s \in S$ tales que $\{v_1,\dots,v_r,w_{r+1},\dots,w_s\}$ es una base de S, y sean $u_{r+1},\dots,u_t \in T$ tales que $\{v_1,\dots,v_r,u_{r+1},\dots,u_t\}$ es una base de T.

Veamos que $\{v_1,\dots,v_r,w_{r+1},\dots,w_s,u_{r+1},\dots,u_t\}$ es una base de $S+T$:

Es claro que es un sistema de generadores de $S+T$. Veamos que es un conjunto linealmente independiente. Supongamos que

$$\sum_{i=1}^{r} \alpha_i v_i + \sum_{j=r+1}^{s} \beta_j w_j + \sum_{k=r+1}^{t} \gamma_k u_k = 0.$$

Entonces $\sum\limits_{i=1}^{r} \alpha_i v_i + \sum\limits_{j=r+1}^{s} \beta_j w_j = -\sum\limits_{k=r+1}^{t} \gamma_k u_k$. Además,

$$\sum_{i=1}^{r} \alpha_i v_i + \sum_{j=r+1}^{s} \beta_j w_j \in S \quad \text{y} \quad -\sum_{k=r+1}^{t} \gamma_k u_k \in T,$$

de donde $-\sum\limits_{k=r+1}^{t} \gamma_k u_k \in S\cap T$. Luego, existen $\delta_1,\dots,\delta_r \in K$ tales que

$$-\sum_{k=r+1}^{t} \gamma_k u_k = \sum_{\ell=1}^{r} \delta_\ell . v_\ell \quad \text{o, equivalentemente,} \quad \sum_{k=r+1}^{t} \gamma_k u_k + \sum_{\ell=1}^{r} \delta_\ell . v_\ell = 0.$$

Pero $\{v_1,\dots,v_r,u_{r+1},\dots,u_t\}$ es una base de T, en particular, un conjunto linealmente independiente. Luego, $\gamma_k = 0 \ \forall\, r+1 \le k \le t$ y $\delta_\ell = 0 \ \forall\, 1 \le \ell \le r$. Entonces

$$\sum_{i=1}^{r} \alpha_i . v_i + \sum_{j=r+1}^{s} \beta_j . w_j = 0,$$

y como $\{v_1, \ldots, v_r, w_{r+1}, \ldots, w_s\}$ es una base de S, resulta que $\alpha_i = 0$ para todo $1 \leq i \leq r$ y $\beta_j = 0$ para todo $r + 1 \leq j \leq s$.

Luego

$$\dim(S + T) = r + (s - r) + (t - r) = s + t - r = \dim S + \dim T - \dim(S \cap T). \qquad \square$$

1.4.2 Suma directa

Un caso de especial importancia de suma de subespacios se presenta cuando $S \cap T = \{0\}$.

Definición 1.44 Sea V un K-espacio vectorial, y sean S y T subespacios de V. Se dice que *V es suma directa de S y T*, y se nota $V = S \oplus T$, si:

1. $V = S + T$,
2. $S \cap T = \{0\}$.

Ejemplo. Sean $S = \{x \in \mathbb{R}^3 : x_1 + x_2 + x_3 = 0\}$ y $T = < (1, 1, 1) >$. Se tiene que $\dim S = 2$, $\dim T = 1$ y $S \cap T = \{0\}$. Entonces $\dim(S + T) = 3$, de donde $S + T = \mathbb{R}^3$.

Luego, $\mathbb{R}^3 = S \oplus T$.

Proposición 1.45 *Sea V un K-espacio vectorial. Sean S y T subespacios de V tales que $V = S \oplus T$. Entonces, para cada $v \in V$, existen únicos $x \in S$ e $y \in T$ tales que $v = x + y$.*

Demostración.
Existencia: Como $V = S + T$, para cada $v \in V$ existen $x \in S$, $y \in T$ tales que $v = x + y$.
Unicidad: Supongamos que $v = x + y$ y $v = x' + y'$ con $x, x' \in S$, $y, y' \in T$. Entonces $x - x' = y - y'$ y $x - x' \in S$, $y - y' \in T$, luego $x - x' \in S \cap T = \{0\}$. En consecuencia $x - x' = y - y' = 0$, de donde $x = x'$, $y = y'$. $\square$

La Proposición 1.42 establece que dados dos subespacios S y T de un espacio vectorial, la unión de un sistema de generadores de S y un sistema de generadores de T es un sistema de generadores de $S + T$. Esto no vale en el caso de dos bases: la unión de una base de S y una de T puede ser un conjunto linealmente dependiente. Sin embargo, la propiedad es válida en el caso en que los subespacios estén en suma directa:

Proposición 1.46 *Sea V un K-espacio vectorial. Sean S y T subespacios de V. Sean B_S y B_T bases de S y T respectivamente. Son equivalentes:*

i) $V = S \oplus T$

ii) $B = B_S \cup B_T$ *es una base de* V.

Observamos que en la condición ii), B es la *familia* obtenida mediante la unión de las familias B_S y B_T.

Demostración. Supongamos que $B_S = \{v_i\}_{i\in I}$ y $B_T = \{w_j\}_{j\in J}$.

$i) \Rightarrow ii)$ Dado que B_S y B_T son sistemas de generadores de S y T respectivamente, entonces $B = B_S \cup B_T$ es un sistema de generadores de $V = S \oplus T$. Por otro lado, si

$$\underbrace{\sum_{i\in I} \alpha_i v_i}_{\in S} + \underbrace{\sum_{j\in J} \beta_j w_j}_{\in T} = 0,$$

como también se tiene $0 = 0 + 0$ con $0 \in S$ y $0 \in T$, por la proposición anterior $\sum_{i\in I} \alpha_i v_i = \sum_{j\in J} \beta_j w_j = 0$. La independencia lineal de B_S y B_T implica que $\alpha_i = 0$ $\forall\, i \in I$ y $\beta_j = 0\ \forall\, j \in J$. Luego, B es linealmente independiente.

$ii) \Rightarrow i)$ Como $B = B_S \cup B_T$ es una base de V, para cada $v \in V$ existen $\alpha_i \in K$, $i \in I$, y $\beta_j \in K$, $j \in J$, casi todos nulos, tales que $v = \sum\limits_{i\in I} \alpha_i v_i + \sum_{j\in J} \beta_j w_j$ y por lo tanto $v = x + y$ con $x = \sum\limits_{i\in I} \alpha_i v_i \in S$ e $y = \sum_{j\in J} \beta_j w_j \in T$. Luego $V = S + T$.

Si $v \in S \cap T$, se tiene que $v = \sum\limits_{i\in I} \alpha_i v_i = \sum\limits_{j\in J} \beta_j w_j$, de donde $\sum\limits_{i\in I} \alpha_i v_i + \sum\limits_{j\in J} (-\beta_j) w_j = 0$, y por la independencia lineal de B, resulta que $\alpha_i = 0\ \forall\, i \in I$ y $\beta_j = 0\ \forall\, j \in J$, de donde $v = 0$ y $S \cap T = \{0\}$. □

Definición 1.47 Sea V un K-espacio vectorial y sea $S \subseteq V$ un subespacio de V. Diremos que T es un *complemento de* S si $S \oplus T = V$.

Ejemplos.

1. Hallar un complemento de $\mathbb{R}_n[X]$ en $\mathbb{R}[X]$.

   Buscamos un subespacio S de $\mathbb{R}[X]$ tal que $\mathbb{R}[X] = \mathbb{R}_n[X] \oplus S$, es decir, $\mathbb{R}[X] = \mathbb{R}_n[X] + S$ y $\mathbb{R}[X] = \mathbb{R}_n[X] \cap S = \{0\}$.

   Se tiene que $\mathbb{R}_n[X] = < 1, X, \dots, X^n >$.

   Consideremos $S = < X^{n+1}, \dots, X^j, \dots > = < X^i >_{i\geq n+1}$.

   Es claro que $\mathbb{R}_n[X] + S = \mathbb{R}[X]$.

   Si $f \in \mathbb{R}_n[X] \cap S$, entonces $f = 0$ o $\mathrm{gr}(f) \leq n$, y además $f = \sum\limits_{i=n+1}^{h} a_i X^i$ con $a_i \in \mathbb{R}$.

   Luego, $f = 0$.

   En consecuencia, $\mathbb{R}[X] = \mathbb{R}_n[X] \oplus S$.

2. Sea $S = \{f \in \mathbb{R}[X] \ /\ f(1) = 0\}$. Hallar un complemento de S en $\mathbb{R}[X]$.

   Vemos que $S = < (X-1)X^i >_{i\in\mathbb{N}_0}$. Sea $T = < 1 >$.

   Dado $f \in \mathbb{R}[X]$, $f = \big(f - f(1)\big) + f(1)$ y $f - f(1) \in S$, $f(1) \in T$. Entonces, $S + T = \mathbb{R}[X]$.

   Sea $f \in S \cap T$. Como $f \in S$, se tiene que $f = (X-1)g$ para algún $g \in \mathbb{R}[X]$ y como $f \in T$, $f = 0$ o $\mathrm{gr}(f) = 0$. Luego $f = 0$.

   Por lo tanto $S \oplus T = \mathbb{R}[X]$.

1.5 Ejercicios

Ejercicio 1.

i) Representar gráficamente en el plano los siguientes vectores:
$(-1,1)\,;\quad (2,3)\,;\quad (-1,1)+(2,3)\,;\quad \frac{1}{2}.(-1,1)+\frac{3}{2}.(2,3)$

ii) Sean $v,w\in\mathbb{R}^2$. Interpretar geométricamente $-v\,,\ 3.v\,,\ \frac{1}{3}.v\,,\ v+w\,,\ v-w$.

iii) Sean $v=(3,1)\,,\ w=(2,4)\in\mathbb{R}^2$. Representar gráficamente los conjuntos:
$S_1=\{r.v\,/\,r\in\mathbb{R}\}$
$S_2=\{r.v\,/\,r\in\mathbb{R}_{\geq 1}\}$
$S_3=\{r.v+s.w\,/\,r,s\in\mathbb{R}\}$
$S_4=\{r.v+s.w\,/\,r,s\in\mathbb{R}\,,\ 0\leq r,s\leq 1\}$
$S_5=\{r.v+s.w\,/\,r,s\in\mathbb{R}\,,\ 0\leq r,s\leq 1,\ r+s=1\}$

Ejercicio 2. Probar en cada caso que el conjunto V con la suma y el producto por escalares definidos es un espacio vectorial sobre K.

i) $V=\mathrm{K}^{\mathbb{N}}=\{(a_i)_{i\in\mathbb{N}}=(a_1,a_2,\ldots,a_n,\ldots)/a_i\in K\ \forall\, i\in\mathbb{N}\}$, el conjunto de todas las sucesiones de elementos de K (donde K es un cuerpo cualquiera).
$+\ :\ (a_i)_{i\in\mathbb{N}}+(b_i)_{i\in\mathbb{N}}=(a_i+b_i)_{i\in\mathbb{N}}$
$.\ :\ k.(a_i)_{i\in\mathbb{N}}=(k.a_i)_{i\in\mathbb{N}}$

ii) X es un conjunto, $V=\mathcal{P}(X)$, $K=\mathbb{Z}_2$.
$+\ :\ B+C=B\triangle C$
$.\ :\ 0.B=\emptyset,\quad 1.B=B$

iii) $V=\mathbb{R}_{>0}$, $K=\mathbb{Q}$.
$\oplus\ :\ a\oplus b=a.b$
$\otimes\ :\ \frac{m}{n}\otimes a=\sqrt[n]{a^m}$

Ejercicio 3. Sea V un espacio vectorial sobre K, $k\in K$, $v\in V$. Probar las siguientes afirmaciones:

i) $k.\vec{0}=\vec{0}$

ii) $-(-v)=v$

iii) $k.v=\vec{0}\Rightarrow k=0$ ó $v=\vec{0}$

iv) $-\vec{0}=\vec{0}$

Ejercicio 4.

i) Sea $v\in\mathbb{R}^2$ un vector fijo. Se define la función $f_v:\mathbb{R}^2\to\mathbb{R}^2$ de la siguiente forma:
$$f_v(x,y)=(x,y)+v$$
Interpretar geométricamente el efecto de f_v sobre el plano (f_v se llama la *traslación en* v).

ii) Probar que $\mathbb{R}^2$ es un $\mathbb{R}$-espacio vectorial con la suma $+_{(2,1)}$ y el producto por escalares $\cdot_{(2,1)}$ definidos de la siguiente forma:

$$\begin{array}{rcl} (x,y) +_{(2,1)} (x',y') & = & (x+x'-2, y+y'-1) \\ r \cdot_{(2,1)} (x,y) & = & r.(x-2, y-1) + (2,1) \end{array}$$

(Este espacio se notará $\mathbb{R}^2_{(2,1)}$ para distinguirlo de $\mathbb{R}^2$ con la suma y el producto usual. La notación se basa en que el $(2,1)$ resulta el neutro de la suma $+_{(2,1)}$).

iii) Interpretar geométricamente $+_{(2,1)}$ y $\cdot_{(2,1)}$, teniendo en cuenta que:

$$\begin{array}{rcl} (x,y) +_{(2,1)} (x',y') & = & f_{(2,1)}\big(f_{(-2,-1)}(x,y) + f_{(-2,-1)}(x',y')\big) \\ r \cdot_{(2,1)} (x,y) & = & f_{(2,1)}\big(r.f_{(-2,-1)}(x,y)\big) \end{array}$$

Ejercicio 5. Sea $S = \{f \in \mathbb{R}[X] \,/\, f(1) = f(2)\}$.

i) Verificar que la suma usual de polinomios es una operación en S (es decir: $f,\, g \in S \Rightarrow f + g \in S$)

ii) Verificar que el producto usual de un número real por un polinomio es una acción de $\mathbb{R}$ en S (es decir: $r \in \mathbb{R}, f \in S \Rightarrow r.f \in S$)

iii) Probar que $(S, +, .)$ es un $\mathbb{R}$-espacio vectorial. (Si se minimiza el trabajo sólo deberá verificarse una propiedad para esto. Comparar i), ii) y iii) con el criterio para decidir si un subconjunto es un subespacio.)

Ejercicio 6.

i) Encontrar un subconjunto no vacío de $\mathbb{R}^2$ que sea cerrado para la suma y para la resta pero no para la multiplicación por escalares.

ii) Encontrar un subconjunto no vacío de $\mathbb{R}^2$ que sea cerrado para la multiplicación por escalares pero no para la suma.

Ejercicio 7. Decidir cuáles de los siguientes subconjuntos son subespacios de V como K-espacio vectorial:

i) $S_1 = \{a.i \,/\, a \in \mathbb{R}\} \quad V = \mathbb{C} \quad K = \mathbb{R}$ ó $K = \mathbb{C}$

ii) $S_2 = \{f \in K[X] \,/\, f'(1) = 0\} \quad V = K[X]$

iii) $S_3 = \{M \in K^{n\times n} \,/\, M_{ij} = -M_{ji}\ \forall\, i,j\,\} \quad V = K^{n\times n}$

iv) $S_4 = \{f \in C^\infty(\mathbb{R}) \,/\, f'' + 3f' = 0\} \quad V = C^\infty(\mathbb{R}) \quad K = \mathbb{R}$

v) $S_5 = \{v \in \mathbb{R}^2_{(2,1)} \,/\, x + y = 3\} \quad V = \mathbb{R}^2_{(2,1)} \quad K = \mathbb{R}$

vi) $S_6 = \{(a_i)_{i\in\mathbb{N}} \in K^{\mathbb{N}} \,/\, a_1 = 0\} \quad V = K^{\mathbb{N}}$

vii) $S_7 = \{(a_i)_{i\in\mathbb{N}} \in K^{\mathbb{N}} \,/\, \exists\, k \in \mathbb{N}$ tal que $a_r = 0\ \forall\, r \geq k\} \quad V = K^{\mathbb{N}}$

viii) $S_8 = \{(a_i)_{i\in\mathbb{N}} \in K^{\mathbb{N}} \,/\, a_1.a_2 = 0\} \quad V = K^{\mathbb{N}}$

Ejercicio 8. Sean S y T subespacios de un K-espacio vectorial V. Probar que $S \cup T$ es un subespacio de $V \iff S \subseteq T$ ó $T \subseteq S$.

Ejercicio 9. Encontrar un sistema de generadores para los siguientes K-espacios vectoriales:

i) $S = \{(x, y, z) \in \mathbb{R}^3 \,/\, x + y - z = 0\}$, $K = \mathbb{R}$

ii) $K_n[X] = \{f \in K[X] \,/\, f = 0 \text{ ó } \mathrm{gr}(f) \leq n\}$

iii) $\mathbb{C}^{n \times n}$, $K = \mathbb{R}$

iv) $\mathcal{P}(\{a, b, c\})$, $K = \mathbb{Z}_2$

Ejercicio 10. Decidir cuáles de las siguientes afirmaciones son verdaderas y cuáles falsas.

i) Sea V un K-espacio vectorial y sean $v\,,\, w \in V$, $k \in K$.
Entonces $< v\,,\, w > = < v\,,\, w + k.v >$.

ii) Sean $v_1\,,\, v_2\,,\, v_3\,,\, v_4\,,\, w \in \mathbb{R}^7$ tales que $< v_1\,,\, v_2\,,\, w > = < v_3\,,\, v_4\,,\, w >$.
Entonces $< v_1\,,\, v_2 > = < v_3\,,\, v_4 >$.

Ejercicio 11. Resolver los siguientes sistemas de ecuaciones lineales ($K = \mathbb{R}$).

i) $\left\{ \begin{array}{rcl} x_1 + x_2 - 2x_3 + x_4 & = & 0 \\ 3x_1 - 2x_2 + x_3 + 5x_4 & = & 0 \\ x_1 - x_2 + x_3 + 2x_4 & = & 0 \end{array} \right.$

ii) $\left\{ \begin{array}{rcl} x_1 + x_2 - 2x_3 + x_4 & = & -2 \\ 3x_1 - 2x_2 + x_3 + 5x_4 & = & 3 \\ x_1 - x_2 + x_3 + 2x_4 & = & 2 \end{array} \right.$

iii) $\left\{ \begin{array}{rcl} x_1 + x_2 + x_3 - 2x_4 + x_5 & = & 1 \\ x_1 - 3x_2 + x_3 + x_4 + x_5 & = & 0 \\ 3x_1 - 5x_2 + 3x_3 + 3x_5 & = & 0 \end{array} \right.$

iv) $\left\{ \begin{array}{rcl} x_1 + x_2 + x_3 + x_4 & = & 2 \\ x_1 + 3x_2 + 2x_3 + 4x_4 & = & 0 \\ 2x_1 + x_3 - x_4 & = & 6 \end{array} \right.$

¿Cambia algo si $K = \mathbb{Q}$? ¿Y si $K = \mathbb{C}$?

Ejercicio 12.

i) Resolver los siguientes sistemas y comparar los conjuntos de soluciones ($K = \mathbb{R}$).

a) $\{x + 2y - 3z = 4$

b) $\left\{ \begin{array}{rcl} x + 2y - 3z & = & 4 \\ x + 3y + z & = & 11 \end{array} \right.$

c) $\left\{ \begin{array}{rcl} x + 2y - 3z & = & 4 \\ x + 3y + z & = & 11 \\ 2x + 5y - 4z & = & 13 \end{array} \right.$

ii) Interpretar geométricamente los conjuntos de soluciones obtenidos.

Ejercicio 13. Resolver los siguientes sistemas no homogéneos. Considerar en cada uno de ellos el sistema homogéneo asociado $(A.x = 0)$.

i) $\begin{cases} x_1 - x_2 + x_3 &= 2 \\ -x_1 + 2x_2 + x_3 &= -1 \\ -x_1 + 4x_2 + 5x_3 &= 1 \end{cases}$ ii) $\begin{cases} x_1 - x_2 + x_3 &= 1 \\ -x_1 + 2x_2 + x_3 &= 1 \\ -x_1 + 4x_2 + 5x_3 &= 4 \end{cases}$

iii) $\begin{cases} x_1 - x_2 - x_3 &= 2 \\ 2x_1 + x_2 - 2x_3 &= 1 \\ x_1 + 4x_2 + x_3 &= 1 \end{cases}$ iv) $\begin{cases} x_1 - x_2 - x_3 &= \alpha \\ 2x_1 + x_2 - 2x_3 &= \beta \\ x_1 + 4x_2 + x_3 &= \gamma \quad (\alpha, \beta, \gamma \in \mathbb{R}) \end{cases}$

Ejercicio 14. Dado el sistema:

$$\begin{cases} 2x_1 - x_2 + x_3 &= \alpha_1 \\ 3x_1 + x_2 + 4x_3 &= \alpha_2 \\ -x_1 + 3x_2 + 2x_3 &= \alpha_3 \end{cases}$$

Determinar los valores de $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$ para los cuales el sistema admite solución.

Ejercicio 15. Resolver según los valores de a y b en $\mathbb{R}$

i) $\begin{cases} (5-a)x_1 - 2x_2 - x_3 &= 1 \\ -2x_1 + (2-a)x_2 - 2x_3 &= 2 \\ -x_1 - 2x_2 + (5-a)x_3 &= b \end{cases}$ ii) $\begin{cases} ax + y + z &= 1 \\ x + ay + z &= a \\ x + y + az &= a^2 \end{cases}$

Ejercicio 16. Determinar todos los $k \in \mathbb{R}$ para que cada uno de los siguientes sistemas tenga solución única.

i) $\begin{cases} x_1 + kx_2 + x_3 &= 0 \\ 2x_1 + x_3 &= 0 \\ 2x_1 + kx_2 + kx_3 &= 0 \end{cases}$ ii) $\begin{cases} x_1 + (k-1)x_2 &= 0 \\ x_1 + (3k-4)x_2 + kx_3 &= 0 \\ x_1 + (k-1)x_2 + \frac{k}{2}x_3 &= 0 \end{cases}$

Ejercicio 17. Determinar los números reales k para los cuales el sistema

$$\begin{cases} kx_1 + x_2 &= 0 \\ x_1 + kx_2 &= 0 \\ k^3x_1 + x_2 + k^3x_3 + kx_4 &= 0 \\ x_1 + k^2x_2 + kx_3 + kx_4 &= 0 \end{cases}$$

tiene alguna solución no trivial y, para esos k, resolverlo.

Ejercicio 18. Determinar para qué valores de $k \in \mathbb{R}$ cada uno de los siguientes sistemas tiene solución única, no tiene solución o tiene infinitas soluciones.

i) $\begin{cases} x_1 + kx_2 - x_3 &= 1 \\ -x_1 + x_2 + k^2x_3 &= -1 \\ x_1 + kx_2 + (k-2)x_3 &= 2 \end{cases}$ ii) $\begin{cases} kx_1 + 2x_2 + kx_3 &= 1 \\ kx_1 + (k+4)x_2 + 3kx_3 &= -2 \\ -kx_1 - 2x_2 + x_3 &= 1 \\ (k+2)x_2 + (3k+1)x_3 &= -1 \end{cases}$

Ejercicio 19.

i) Resolver el siguiente sistema en $\mathbb{C}^2$:

$$\begin{cases} (1-i)x_1 - ix_2 &= 0 \\ 2x_1 + (1-i)x_2 &= 0 \end{cases}$$

ii) Resolver en $\mathbb{C}^3$ el sistema $A.x = 0$ donde

$$A = \begin{pmatrix} i & -(1+i) & 0 \\ 1 & -2 & 1 \\ 1 & 2i & -1 \end{pmatrix}$$

Ejercicio 20. Resolver los siguientes sistemas:

i) en $\mathbb{Z}_5$: $\begin{cases} x_1 + 2x_2 + 2x_3 + x_4 &= 4 \\ 2x_1 + 3x_3 + x_4 &= 2 \\ 4x_2 + 2x_3 + 4x_4 &= 1 \end{cases}$

ii) en $\mathbb{Z}_7$: $\begin{cases} x + z &= 2 \\ 2y + z &= 6 \\ x + 3y &= 0 \end{cases}$

iii) en $\mathbb{Z}_3$: $\begin{cases} x + y + z &= 1 \\ 2x + y + 2z &= 0 \\ x + z &= 2 \end{cases}$

Ejercicio 21. Encontrar un sistema a coeficientes reales cuya solución general sea:

$$(1,1,0) + \lambda(1,2,1), \ \lambda \in \mathbb{R}.$$

Ejercicio 22. Sean $A \in K^{m\times n}$, $b \in K^{m\times 1}$.

i) Si el sistema $A.x = 0$ tiene solución única, probar que el sistema $A.x = b$ tiene a lo sumo una solución. Dar ejemplos de los distintos casos que se puedan presentar.

ii) ¿Vale la recíproca de i)?

Ejercicio 23. Encontrar un sistema de generadores para cada uno de los siguientes espacios vectoriales sobre K:

i) $S_1 = \{(x,y,z) \in \mathbb{R}^3 \,/\, x + y - z = 0 \ ; \ x - y = 0\}$, $K = \mathbb{R}$

ii) $S_2 = \left\{ (x,y,z) \in (\mathbb{Z}_7)^3 \,/\, \begin{cases} x + z = 0 \\ 2y + z = 0 \\ x + 3y = 0 \end{cases} \right\}$, $K = \mathbb{Z}_7$

iii) $S_3 = \{A \in \mathbb{Q}^{3\times 3} \,/\, A_{ij} = -A_{ji} \ \forall\, i, j\,\}$, $\ K = \mathbb{Q}$

iv) $S_4 = \{f \in \mathbb{R}_4[X] \,/\, f(1) = 0 \text{ y } f(2) = f(3)\}$, $\ K = \mathbb{R}$

v) $S_5 = \{(a_n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}} \,/\, a_i = 0 \ \forall\, i \geq 5\,;\, a_1 + 2a_2 - a_3 = 0\,;\, a_2 + a_4 = 0\}$, $\ K = \mathbb{R}$

vi) $S_6 = \{f \in C^\infty(\mathbb{R}) \,/\, f''' = 0\}$, $\ K = \mathbb{R}$

Ejercicio 24. Sea V un $\mathbb{R}$-espacio vectorial y sean $v_1\,,\, v_2\,,\, v_3 \ \in V$. Probar que si $v_1 + 3v_2 - v_3 = 0 = 2v_1 - v_2 - v_3$ entonces $< v_1\,,\, v_2\,,\, v_3 > = < v_3 >$.

Ejercicio 25. Determinar si $v \in S$ en cada uno de los siguientes casos:

i) $v = (1, 2, -1)$, $\quad S = < (1, 3, 2)\,,\, (2, 0, 1)\,,\, (1, 1, 1) > \ \subseteq \mathbb{R}^3$

ii) $v = (1, 0, -1, 3)$, $\quad S = < (1, 0, 1, 0)\,,\, (2, 1, 0, 1)\,,\, (0, 1, 0, -2) > \ \subseteq \mathbb{R}^4$

Ejercicio 26. Sea $S = < (1, -1, 2, 1), (3, 1, 0, -1), (1, 1, -1, -1) > \ \subseteq \mathbb{R}^4$.

i) Determinar si $(2, 1, 3, 5) \in S$.

ii) Determinar si $\{x \in \mathbb{R}^4 / x_1 - x_2 - x_3 = 0\} \subseteq S$.

iii) Determinar si $S \subseteq \{x \in \mathbb{R}^4 / x_1 - x_2 - x_3 = 0\}$.

Ejercicio 27. Hallar un sistema de generadores para $S \cap T$ como subespacio de V en cada uno de los siguientes casos:

i) $V = \mathbb{R}^3$, $\quad S = \{(x, y, z)/3.x - 2.y + z = 0\} \qquad T = \{(x, y, z)/x + z = 0\}$

ii) $V = \mathbb{R}^3$, $\quad S = \{(x, y, z)/3.x - 2.y + z = 0\} \qquad T = < (1, 1, 0), (5, 7, 3) >$

iii) $V = \mathbb{R}^3$, $\quad S = < (1, 1, 3), (1, 3, 5), (6, 12, 24) > \qquad T = < (1, 1, 0), (3, 2, 1) >$

iv) $V = \mathbb{R}^{3\times 3}$, $\quad S = \{(x_{ij}) \,/\, x_{ij} = x_{ji} \ \forall\, i, j\} \qquad T = \{(x_{ij}) \,/\, x_{11} + x_{12} + x_{13} = 0\}$

v) $V = \mathbb{R}[X]$, $\quad S = \{f \in \mathbb{R}[X] \,/\, f(1) = 0\} \qquad T = < 1\,,\, X,\, X^2,\, X^3 + 2X^2 - X,\, X^5 >$

vi) $V = \mathbb{R}[X]$, $\quad S = \{f \in \mathbb{R}[X] \,/\, f(1) = 0\} \qquad T = \{f \in \mathbb{R}[X] \,/\, f'(1) = f''(1) = 0\}$

Ejercicio 28. Decidir si las siguientes sucesiones de vectores son linealmente independientes sobre K.

i) $(1 - X)^3$, $(1 - X)^2$, $1 - X$, 1 en $K[X]$

ii) $(1, 2, 3)\,,\, (2, 3, 1)\,,\, (1, 1, 4)\,,\, (5, 1, 1)$ en $\mathbb{R}^3$

iii) $(1, 4, -1, 3)\,,\, (2, 1, -3, -1)\,,\, (0, 2, 1, -5)$ en $\mathbb{Q}^4$

iv) $(1 - i, i)\,,\, (2, -1 + i)$ en $\mathbb{C}^2$, para $K = \mathbb{R}$ y $K = \mathbb{C}$

v) $(3+\sqrt{2}, 1+\sqrt{2})$, $(7, 1+2\sqrt{2})$ en $\mathbb{R}^2$, para $K=\mathbb{Q}$ y $K=\mathbb{R}$

vi) $f(x)=1$, $g(x)=x$ en $\mathbb{R}^{\mathbb{R}}$

vii) $f(x)=\operatorname{sen}(x)$, $g(x)=\cos(x)$ en $\mathbb{R}^{\mathbb{R}}$

viii) $f(x)=e^x$, $g(x)=x$ en $\mathbb{R}^{\mathbb{R}}$

ix) $u=(1,0,1,0,1,\ldots)$, $v=(0,1,0,1,0,\ldots)$, $w=(1,1,0,1,1,0,\ldots)$ en $\mathbb{R}^{\mathbb{N}}$

Ejercicio 29. Hallar todos los $k\in\mathbb{R}$ para los cuales $\{v_1\,,\, v_2\,,\, v_3\}\subset V$ es un conjunto linealmente independiente en los siguientes casos:

i) $\{(1,2,k)\,,\,(1,1,1)\,,\,(0,1,1-k)\}\subset\mathbb{R}^3$

ii) $\{(k,1,0)\,,\,(3,-1,2)\,,\,(k,2,-2)\}\subset\mathbb{R}^3$

iii) $\{k.X^2+X\,,\, X^2-k\,,\, k^2.X\}\subset\mathbb{R}[X]$

iv) $\left\{\begin{pmatrix}1 & k\\ -1 & 2\end{pmatrix}, \begin{pmatrix}k & 1\\ 0 & 2k\end{pmatrix}, \begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}\right\}\subset\mathbb{R}^{2\times 2}$

Ejercicio 30. Sean $v_1,\ldots,v_n\in\mathbb{R}^n$. Probar que $\{v_1,\ldots,v_n\}$ son linealmente independientes sobre $\mathbb{R}$ $\iff$ $\{v_1,\ldots,v_n\}$ son linealmente independientes sobre $\mathbb{C}$.

Ejercicio 31. En cada uno de los siguientes casos hallar una base del subespacio de soluciones del sistema lineal homogéneo $A.x=0$ ($K=\mathbb{R}$).

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

ii) $A=\begin{pmatrix}0 & 5 & 3\\ 1 & -1 & 2\\ 2 & 3 & 1\end{pmatrix}$

iii) $A=\begin{pmatrix}3 & -1 & 0 & 1 & 2\\ -1 & 0 & 4 & -1 & 0\\ 3 & 1 & 1 & 0 & 1\\ 2 & 0 & 0 & 3 & 1\end{pmatrix}$

Ejercicio 32. Completar los siguientes conjuntos linealmente independientes a una base del K-espacio vectorial V indicado.

i) $\{(1,1,1,1)\,,\,(0,2,1,1)\}$, $V=\mathbb{R}^4$, $K=\mathbb{R}$

ii) $\{X^3-2X+1\,,\, X^3+3X\}$, $V=\mathbb{R}_3[X]$, $K=\mathbb{R}$

iii) $\left\{\begin{pmatrix}1 & 1\\ i & 1\end{pmatrix}, \begin{pmatrix}0 & i\\ 1 & 1\end{pmatrix}, \begin{pmatrix}0 & 2\\ 1 & 1\end{pmatrix}\right\}$, $V=\mathbb{C}^{2\times 2}$, $K=\mathbb{R}$ y $K=\mathbb{C}$

Ejercicio 33. Extraer una base de S de cada uno de los siguientes sistemas de generadores.

i) $S = < (1,1,2)\,,\,(1,3,5)\,,\,(1,1,4)\,,\,(5,1,1) > \;\subseteq \mathbb{R}^3,\;\; K = \mathbb{R}$

ii) $S = < X^2+2X+1\,,\,X^2+3X+1\,,\,X+2 > \;\subseteq \mathbb{R}[X],\;\; K = \mathbb{R}$

iii) $S = \left\langle \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & i \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & i \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \right\rangle \subseteq \mathbb{C}^{2\times 2},\;\; K = \mathbb{R}$ y $K = \mathbb{C}$

Ejercicio 34.

i) Sea $B = \{f_0, f_1, f_2, \ldots, f_n, \ldots\}$, donde cada $f_i \in K[X]$ es un polinomio de grado exactamente i. Probar que B es una base de $K[X]$.

ii) ¿Es $\{(1,0,0,0,0,\ldots);(0,1,0,0,0,\ldots);(0,0,1,0,0,\ldots);(0,0,0,1,0,\ldots);\ldots\}$ una base de $K^{\mathbb{N}}$?

Ejercicio 35. Hallar una base y la dimensión de los siguientes K-espacios vectoriales:

i) $< (1,4,-2,1)\,,\,(1,-3,-1,2)\,,\,(3,-8,-2,7) > \;\subseteq \mathbb{R}^4,\;\; K = \mathbb{R}$

ii) $\left\langle \begin{pmatrix} 1 & -5 \\ -4 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ -1 & 5 \end{pmatrix}, \begin{pmatrix} 2 & -4 \\ -5 & 7 \end{pmatrix}, \begin{pmatrix} 1 & -7 \\ -5 & 1 \end{pmatrix} \right\rangle \subseteq \mathbb{Q}^{2\times 2},\;\; K = \mathbb{Q}$

iii) $\mathbb{C},\;\; K = \mathbb{R}$ y $K = \mathbb{C}$

iv) $\{f \in \mathbb{R}[X] \,/\, f = 0$ ó $\mathrm{gr}(f) \leq 3$ y $f(2) = f(-1)\},\;\; K = \mathbb{R}$

v) $\mathcal{P}(\{a,b,c\}),\;\; K = \mathbb{Z}_2$

vi) $\{f \in \mathbb{Q}[X] \,/\, f = 0$ ó $\mathrm{gr}(f) \leq 3$ y $(x^2-2) \mid f\},\;\; K = \mathbb{Q}$

vii) $\{(a_n)_{n\in\mathbb{N}} \in K^{\mathbb{N}} \,/\, a_i = a_j \;\forall\, i,j\,\}$

Ejercicio 36. Hallar la dimensión del $\mathbb{R}$-espacio vectorial S para cada $k \in \mathbb{R}$ en los siguientes casos:

i) $S = < (1,k,1)\,,\,(-1,k,1)\,,\,(0,1,k) >$

ii) $S = \{x \in \mathbb{R}^3 / A.x = 0\}$ siendo $A \in \mathbb{R}^{3\times 3}$, $A = \begin{pmatrix} 1 & -k & -1 \\ -1 & 1 & k^2 \\ 1 & k & k-2 \end{pmatrix}$

Ejercicio 37. Hallar todos los $b \in \mathbb{R}$ para los cuales el $\mathbb{R}$-espacio vectorial de soluciones del sistema:

$$\left\{ \begin{array}{rcl} 3x_1 + (b-6)x_2 + 5bx_3 & = & 0 \\ x_1 + (b-2)x_2 + (b^2+4b)x_3 & = & 0 \\ x_1 - 2x_2 + bx_3 & = & 0 \end{array} \right.$$

i) tenga dimensión 1.

ii) tenga dimensión 2.

Ejercicio 38. Sean S y T los subespacios de $\mathbb{R}^4$

$$S = < (1,2,1,0), (2,1,0,1) > \quad \text{y} \quad T = \{x \in \mathbb{R}^4 / x_1 - 3x_2 - 2x_3 = 0\}.$$

Hallar un subespacio U de $\mathbb{R}^4$ tal que $\dim U = 2$ y $S \cap T \subset U \subset T$.

Ejercicio 39. Determinar todos los $k \in \mathbb{R}$ para los cuales

$$< (-2,1,6),\ (3,0,-8) > = < (1,k,2k),\ (-1,-1,k^2-2),\ (1,1,k) >.$$

Ejercicio 40. Se considera el $\mathbb{Q}$-espacio vectorial $V \subset \mathbb{R}$ generado por $\{1,\ \sqrt{2},\ \sqrt{3},\ \sqrt{6}\}$.

i) Utilizando un argumento de dimensión probar que existe un polinomio $f \in \mathbb{Q}[X]$ con $\mathrm{gr}(f) \leq 4$ que se anula en el punto $\psi = \sqrt{2} + \sqrt{3}$. Hallar un tal f.

ii) Calcular $\dim_{\mathbb{Q}} V$.

Ejercicio 41. En cada uno de los siguientes casos caracterizar $S + T \subseteq V$ y determinar si la suma es directa.

i) $V = K^{n\times n}$, $S = \{A \in K^{n\times n} \,/\, A_{ij} = A_{ji}\ \forall\, i,j\}$, $T = \{A \in K^{n\times n} \,/\, A_{ij} = -A_{ji}\ \forall\, i,j\}$

ii) $V = \mathbb{R}^3$, $S = < (1,1,1) >$, $T = < (2,-1,1), (3,0,2) >$

iii) $V = \mathbb{R}[X]$, $S = \{f \in \mathbb{R}[X] \,/\, f = 0 \text{ ó } \mathrm{gr}(f) \leq 3\}$, $T = \{f \in \mathbb{R}[X] \,/\, \mathrm{mult}(4,f) \geq 4\}$

iv) $V = \mathbb{R}^{2\times 3}$, $S = \{A \in \mathbb{R}^{2\times 3} \,/\, A_{11} + A_{21} = 0,\ 3A_{22} - 2A_{11} = A_{13} + A_{23}\}$,
$T = \left\langle \begin{pmatrix} 2 & 0 & 3 \\ 2 & -1 & 1 \end{pmatrix}, \begin{pmatrix} 2 & 3 & 0 \\ -2 & 2 & 1 \end{pmatrix} \right\rangle$

Ejercicio 42. Determinar todos los $k \in \mathbb{R}$ para los cuales $S \cap T = < (0,1,1) >$, siendo

$$S = \{x \in \mathbb{R}^3 / x_1 + x_2 - x_3 = 0\} \quad \text{y} \quad T = < (1,k,2), (-1,2,k) >.$$

Ejercicio 43. Para cada S dado hallar $T \subseteq V$ tal que $S \oplus T = V$.

i) $S = < (1,2,-1,3)\,,\ (2,3,-2,1)\,,\ (0,1,0,7) >$, $V = \mathbb{R}^4$

ii) $S = \{A \in \mathbb{R}^{n\times n} \,/\, \mathrm{tr}(A) = 0\}$, $V = \mathbb{R}^{n\times n}$

iii) $S = < 3, 1 + X^2 >$, $V = \mathbb{R}_4[X]$

Ejercicio 44. Dado $S = \{(x_1,x_2,x_3,x_4) \in \mathbb{R}^4 \,/\, x_1 - x_2 + 2x_4 = 0\,,\ x_2 + x_3 - x_4 = 0\}$, hallar dos vectores $v_3\,,\ v_4$ de $\mathbb{R}^4$ tales que para toda elección de una base $\{v_1\,,\ v_2\}$ de S, $\{v_1\,,\ v_2\,,\ v_3\,,\ v_4\}$ sea una base de $\mathbb{R}^4$.

Ejercicio 45. Decidir si las siguientes afirmaciones son verdaderas o falsas. Justificar.

i) S, T subespacios de $\mathbb{R}^3$, $\dim S = \dim T = 2 \Rightarrow \exists v \neq 0$ tal que $v \in S \cap T$.

ii) S, T, W subespacios de $\mathbb{R}^{11}$, $\dim S = \dim T = \dim W = 4 \Rightarrow \dim(S \cap T \cap W) \geq 1$.

Ejercicio 46. Sea V un K-espacio vectorial y sean S, T y U subespacios de V.

i) Probar que $(S \cap T) + (S \cap U) \subseteq S \cap (T + U)$.

ii) Mostrar que, en general, la inclusión anterior es estricta.

iii) Probar que, si $U \subseteq S$, entonces vale la igualdad en i).

Ejercicio 47. Sean S, T y U subespacios de un K-espacio vectorial V tales que

$$S \cap T = S \cap U, \qquad S + T = S + U \qquad \text{y} \qquad T \subseteq U.$$

Probar que $T = U$.

Ejercicio 48. Sea V un K-espacio vectorial de dimensión n y sea T un hiperplano de V (es decir, un subespacio de dimensión $n - 1$).

i) Probar que $\forall\, v \notin T,\; T \oplus < v > = V$.

ii) Si S es un subespacio de V tal que $S \not\subseteq T$, probar que $S + T = V$. Calcular $\dim(S \cap T)$.

iii) Si S y T son dos hiperplanos distintos, deducir $\dim(S \cap T)$.

Ejercicio 49. Sea $V = \mathbb{R}^{\mathbb{R}}$.

i) Sean $S = \{f \in V \,/\, f(x) = f(-x)\ \forall\, x \in \mathbb{R}\}$ y $T = \{f \in V \,/\, f(-x) = -f(x)\ \forall\, x \in \mathbb{R}\}$ (S es el conjunto de funciones pares y T el conjunto de funciones impares). Probar que S y T son subespacios de V y que $S \oplus T = V$.

ii) Sean $U = \{f \in V \,/\, f(0) = 0\}$ y $W = \{f \in V \,/\, f \text{ es constante}\}$. Probar que U y W son subespacios de V y que $U \oplus W = V$.

Ejercicio 50.

i) Sea $S = \{(u_n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}} \,/\, u_{n+2} = u_{n+1} + u_n\ \ \forall\, n \in \mathbb{N}\}$. Probar que S es un subespacio de $\mathbb{R}^{\mathbb{N}}$. Calcular su dimensión.

ii) Encontrar una base de S formada por sucesiones $(u_n)_{n\in\mathbb{N}}$ que, $\forall\, n \in \mathbb{N}$, verifiquen $u_n = u^{n-1}$ para algún $u \in \mathbb{R}$.

iii) Usando ii), encontrar una fórmula para el término general de la sucesión de Fibonacci:

$$\begin{cases} F_1 = 1 \\ F_2 = 1 \\ F_{n+2} = F_{n+1} + F_n \quad \forall\, n \geq 1 \end{cases}$$

Capítulo 2

Matrices

En el capítulo anterior hemos utilizado matrices para la resolución de sistemas de ecuaciones lineales y hemos visto que, para $n, m \in \mathbb{N}$, el conjunto de las matrices de n filas y m columnas con coeficientes en un cuerpo K es un K-espacio vectorial. A continuación estudiaremos más en detalle estos conjuntos de matrices, así como también ciertas matrices particulares que nos serán de utilidad.

2.1 Definiciones y propiedades

Comenzaremos recordando algunas definiciones y propiedades estudiadas en el capítulo anterior.

Sean $n, m \in \mathbb{N}$. El conjunto de las matrices de n filas y m columnas con coeficientes en un cuerpo K es

$$K^{n\times m} = \left\{ \begin{pmatrix} a_{11} & \dots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nm} \end{pmatrix} \ / \ a_{ij} \in K \ \forall \, 1 \le i \le n, \, 1 \le j \le m \right\}.$$

Para definir una matriz en $K^{n\times m}$ basta especificar, para cada $1 \le i \le n$ y cada $1 \le j \le m$, qué elemento de K se halla en el lugar ij (correspondiente a la intersección de la fila i y la columna j) de la matriz.

Ejemplo. Sean $n, m \in \mathbb{N}$, y sean $1 \le k \le n$, $1 \le l \le m$. Se define la matriz $E^{kl} \in K^{n\times m}$ como

$$(E^{kl})_{ij} = \begin{cases} 1 & \text{si } i = k, \, j = l \\ 0 & \text{si no} \end{cases}$$

Estas matrices se llaman las *matrices canónicas* de $K^{n\times m}$.

Una primera observación que debemos hacer se refiere a cómo determinar si dos matrices (de las mismas dimensiones) son iguales:

Observación 2.1 Sean $A, B \in K^{n\times m}$. Entonces $A = B$ si y sólo si $A_{ij} = B_{ij}$ para cada $1 \le i \le n, 1 \le j \le m$.

Podemos definir una operación (suma) en $K^{n\times m}$ y una acción de K en $K^{n\times m}$ que transforman a este conjunto en un K-espacio vectorial:

Definición 2.2 Se definen la *suma de matrices* y el *producto por escalares* como

$$\begin{aligned} + &: K^{n\times m} \times K^{n\times m} \to K^{n\times m},\ (A+B)_{ij} = A_{ij} + B_{ij} \qquad (1 \le i \le n, 1 \le j \le m) \\ \cdot &: K \times K^{n\times m} \to K^{n\times m},\ (\lambda \cdot A)_{ij} = \lambda \cdot A_{ij} \qquad (1 \le i \le n, 1 \le j \le m). \end{aligned}$$

Es fácil verificar que $(K^{n\times m}, +, \cdot)$ es un K-espacio vectorial.

Definiremos ahora un producto que, dadas dos matrices A y B con coeficientes en K tales que *la cantidad de columnas de A es igual a la cantidad de filas de B*, calcula una nueva matriz C.

Definición 2.3 Sean $A \in K^{n\times m}$ y $B \in K^{m\times r}$. Se define el *producto de A por B* como la matriz $C \in K^{n\times r}$ tal que

$$C_{ij} = \sum_{k=1}^{m} A_{ik} B_{kj} \qquad 1 \le i \le n,\ 1 \le j \le r.$$

Analizaremos ahora algunas propiedades del producto de matrices y su relación con la suma de matrices.

Proposición 2.4 *Propiedades del producto de matrices:*

1. *Propiedad asociativa: dadas $A \in K^{n\times m}$, $B \in K^{m\times r}$ y $C \in K^{r\times s}$, se tiene que $(A.B).C = A.(B.C)$.*
2. *Para cada $n \in \mathbb{N}$, sea $I_n \in K^{n\times n}$ definida por $(I_n)_{ij} = \begin{cases} 1 & si\ \ i = j \\ 0 & si\ \ i \ne j \end{cases}$. Entonces, si $A \in K^{n\times m}$, se verifica: $I_n.A = A.I_m = A$.*

   La matriz I_n se denomina matriz identidad *de $K^{n\times n}$.*
3. *Propiedades distributivas:*
   - *(a) Si $A \in K^{n\times m}$ y $B, C \in K^{m\times r}$, entonces $A.(B + C) = A.B + A.C$.*
   - *(b) Si $A, B \in K^{n\times m}$ y $C \in K^{m\times r}$, entonces $(A + B).C = A.C + B.C$.*

Demostración.

1. Observemos en primer lugar que si $A \in K^{n\times m}$, $B \in K^{m\times r}$ y $C \in K^{r\times s}$, entonces $(A.B).C \in K^{n\times s}$ y $A.(B.C) \in K^{n\times s}$.

Para cada $1 \le i \le n$, $1 \le j \le s$, se tiene:

$$\big((A.B).C\big)_{ij} = \sum_{\alpha=1}^{r} (A.B)_{i\alpha} C_{\alpha j} = \sum_{\alpha=1}^{r} \Big(\sum_{\beta=1}^{m} A_{i\beta} B_{\beta\alpha} \Big) C_{\alpha j} = \sum_{\alpha=1}^{r} \Big(\sum_{\beta=1}^{m} A_{i\beta} B_{\beta\alpha} C_{\alpha j} \Big) =$$

$$\sum_{\beta=1}^{m} \Big(\sum_{\alpha=1}^{r} A_{i\beta} B_{\beta\alpha} C_{\alpha j} \Big) = \sum_{\beta=1}^{m} A_{i\beta} \Big(\sum_{\alpha=1}^{r} B_{\beta\alpha} C_{\alpha j} \Big) = \sum_{\beta=1}^{m} A_{i\beta} (B.C)_{\beta j} = \big(A.(B.C)\big)_{ij}.$$

2. Sean $1 \le i \le n$, $1 \le j \le m$. Se tiene que

$$(I_n.A)_{ij} = \sum_{k=1}^{n} (I_n)_{ik} A_{kj} = 1.A_{ij} = A_{ij}.$$

   De la misma manera, $(A.I_m)_{ij} = A_{ij}$ para cada $1 \le i \le n, \ 1 \le j \le m$.

3. Queda como ejercicio. □

Observemos que, en particular, el producto de matrices está definido para cualquier par de matrices en $K^{n\times n}$ y, por lo tanto, se tiene una operación "producto" en $K^{n\times n}$ para cada $n \in \mathbb{N}$. De la proposición anterior se deduce:

Proposición 2.5 $(K^{n\times n}, +, \cdot)$ *es un anillo.*

Si bien el producto de matrices comparte muchas de sus propiedades con el producto usual de números reales, hay propiedades que verifica éste que no son válidas para el producto de matrices:

Observación 2.6 Dos de las propiedades que no se cumplen para el producto de matrices son las siguientes:

- El producto de matrices *no* es conmutativo. Aún en el caso de matrices cuadradas, en el que siempre se pueden calcular $A.B$ y $B.A$, en general se tiene que $A.B \ne B.A$. Por ejemplo, para $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ y $B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ se tiene que

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

- El hecho que $A.B = 0$ *no* implica que $A = 0$ o $B = 0$. En el ejemplo anterior, $A \ne 0$, $B \ne 0$, pero $A.B = 0$.

El conjunto $K^{n\times n}$ resulta ser a la vez un anillo y un K-espacio vectorial. La noción que engloba a los conjuntos con estas características es la siguiente:

Definición 2.7 Sea K un cuerpo y sea A un conjunto con dos operaciones, $+$ y $\cdot$, y una acción $\cdot_K$ de K en A tales que

1. $(A, +, \cdot)$ es un anillo
2. $(A, +, \cdot_K)$ es un K-espacio vectorial
3. $(\lambda \cdot_K X) \cdot Y = \lambda \cdot_K (X \cdot Y) = X \cdot (\lambda \cdot_K Y) \quad \forall \lambda \in K \ \forall X, Y \in A$

Se dice entonces que A es una K-*álgebra*.

Observación 2.8 $(K^{n\times n}, +, \cdot_K, \cdot)$ es una K-álgebra.

Observamos que el producto de matrices nos permite escribir un sistema lineal de n ecuaciones con m incógnitas

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m &= b_1 \\ \vdots & \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m &= b_n \end{cases}$$

en la forma

$$A.x = b,$$

donde $A \in K^{n\times m}$ es la matriz asociada al sistema, $x \in K^{m\times 1}$ se define como $x_{i1} = x_i$ (matriz de una columna cuyos elementos son las incógnitas del sistema), y $b \in K^{n\times 1}$ se define como $b_{j1} = b_j$ (matriz de una columna cuyos elementos son los resultados a los que están igualadas las ecuaciones). De esta manera, un sistema lineal puede verse como una única ecuación con una única incógnita x, pero que involucra matrices en lugar de escalares.

El hecho que la solución en K de la ecuación $a.x = b$ con $a, b \in K$, $a \neq 0$, se obtiene haciendo simplemente $x = a^{-1}b$, nos lleva a pensar que el sistema lineal $Ax = b$ podría resolverse análogamente como $x = A^{-1}b$ en caso de disponer de una matriz A^{-1} que sea una inversa de A para el producto de matrices. Éste será el tema a estudiar en la próxima sección.

Concluimos esta sección introduciendo dos nociones que nos serán de utilidad en lo sucesivo:

Definición 2.9 Sea $A \in K^{n\times m}$. Se llama matriz *transpuesta* de A, y se nota A^t, a la matriz $A^t \in K^{m\times n}$ definida por $(A^t)_{ij} = A_{ji}$ para cada $1 \le i \le m$, $1 \le j \le n$.

Definición 2.10 Sea $A \in K^{n\times n}$. Se llama *traza* de la matriz A, y se nota $tr(A)$, al escalar $tr(A) = \sum_{i=1}^{n} A_{ii}$.

2.2 Matrices inversibles

No es cierto que todo elemento no nulo de $K^{n\times n}$ tenga inverso con respecto al producto. Por ejemplo: $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \in K^{2\times 2}$ no tiene inversa. En efecto, $A.B \neq I_2$ para toda matriz $B \in K^{2\times 2}$, puesto que $(A.B)_{22} = 0 \neq (I_2)_{22}$ para toda matriz $B \in K^{2\times 2}$.

En esta sección nos ocuparemos de las matrices que sí tienen inversa y veremos también cómo hallar la inversa de una matriz en el caso en que ésta exista.

Definición 2.11 Una matriz $A \in K^{n\times n}$ se dice *inversible* si existe una matriz $B \in K^{n\times n}$ tal que $A.B = B.A = I_n$.

Observemos que la matriz B de la definición es única. En efecto, si $A.B = B.A = I_n$ y $A.C = C.A = I_n$, entonces

$$B = I_n.B = (C.A).B = C.(A.B) = C.I_n = C.$$

Notación. $B = A^{-1}$.

Para cada $n \in \mathbb{N}$ consideraremos el conjunto de todas las matrices inversibles en $K^{n\times n}$:

$$GL(n, K) = \{A \in K^{n\times n} \ / \ A \text{ es inversible}\}.$$

Nos interesa estudiar la estructura de este conjunto.

Proposición 2.12 *Para cada $n \in N$, se verifican las siguientes propiedades:*

1. *Si $A, B \in GL(n, K)$, entonces $A.B \in GL(n, K)$. Más aún, $(A.B)^{-1} = B^{-1}A^{-1}$. En particular, el producto de matrices $\cdot$ es una operación en $GL(n, K)$.*
2. *$I_n \in GL(n, K)$.*
3. *Si $A \in GL(n, K)$, entonces $A^{-1} \in GL(n, K)$.*

Demostración.

1. Sean $A, B \in GL(n, K)$. Entonces existen A^{-1} y B^{-1}. Se tiene que

$$(A.B).(B^{-1}.\,A^{-1}) = I_n \quad \text{y} \quad (B^{-1}.\,A^{-1}).(A.B) = I_n.$$

   Entonces $A.B$ es inversible y $(A.B)^{-1} = B^{-1}.\,A^{-1}$.
2. Es consecuencia de que $I_n.I_n = I_n$.
3. De la definición de inversa se deduce inmediatamente que si $A \in GL(n, K)$, entonces $(A^{-1})^{-1} = A$ y por lo tanto $A^{-1} \in GL(n, K)$. □

De la proposición anterior y la asociatividad del producto de matrices se deduce que:

Proposición 2.13 $(GL(n, K), \cdot)$ *es un grupo, que se denomina el* grupo lineal general (n, K).

Para concluir esta sección, veremos un método para determinar si una matriz en $K^{n\times n}$ es inversible y, en caso de serlo, encontrar su inversa. Lo describimos en el siguiente ejemplo:

Ejemplo. Hallar, si es posible, A^{-1} siendo $A \in \mathbb{R}^{3\times 3}$ la matriz

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

Buscamos $B \in \mathbb{R}^{3\times 3}$ tal que $A.B = B.A = I_3$. Si $B = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$, debe ser

$$A.\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Esta igualdad se traduce en los tres sistemas de ecuaciones siguientes:

$$A.\begin{pmatrix} a \\ d \\ g \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad A.\begin{pmatrix} b \\ e \\ h \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \text{ y} \qquad A.\begin{pmatrix} c \\ f \\ i \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

que podemos resolver simultáneamente:

$$\left(\begin{array}{ccc|c|c|c} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 2 & -1 & 0 & 1 & 0 \\ 2 & 1 & 1 & 0 & 0 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 2 & -1 & 0 & 1 & 0 \\ 0 & -1 & 1 & -2 & 0 & 1 \end{array}\right) \longrightarrow$$

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & -1 & 2 & 0 & -1 \\ 0 & 0 & 1 & -4 & 1 & 2 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 3 & -1 & -1 \\ 0 & 1 & 0 & -2 & 1 & 1 \\ 0 & 0 & 1 & -4 & 1 & 2 \end{array}\right)$$

Entonces $B = \begin{pmatrix} 3 & -1 & -1 \\ -2 & 1 & 1 \\ -4 & 1 & 2 \end{pmatrix}$ verifica la igualdad $A.B = I_3$.

Observemos que, si buscamos una matriz C tal que $B.C = I_3$, bastaría con hacer los pasos anteriores, pero a la inversa, con lo que obtendríamos la matriz A.

Luego, $A.B = B.A = I_3$, es decir $A^{-1} = \begin{pmatrix} 3 & -1 & -1 \\ -2 & 1 & 1 \\ -4 & 1 & 2 \end{pmatrix}$.

Cómo decidir si una matriz es inversible y hallar su inversa:

- Dada $A \in K^{n\times n}$, se arma una matriz en $K^{n\times 2n}$ cuyas primeras n columnas corresponden a la matriz A y cuyas últimas n columnas están formadas por los n vectores de la base canónica de K^n.

Esto corresponde a plantear la ecuación $A.B = I_n$ con $B \in K^{n\times n}$, subdividirla en n sistemas lineales $A.B_i = e_i$, $1 \le i \le n$, igualando columna a columna, y escribir la matriz ampliada de los n sistemas.

- Si al triangular la matriz no aparece ningún cero en la diagonal, se pueden resolver los sistemas que resultan (que tienen solución única) y hallar entonces la inversa de A.

  Para esto se puede proceder como en el ejemplo anterior: al no aparecer ceros en la diagonal, se continúa aplicando operaciones elementales sobre las filas de la matriz de manera que en las primeras n columnas quede formada la matriz I_n. Entonces A^{-1} es la matriz que aparece en las últimas n columnas (esto puede probarse de la misma manera que se hizo en el ejemplo).

- Si al triangular la matriz aparece un cero en la diagonal, la matriz A no es inversible. En efecto, la presencia de un cero en la diagonal al triangular implica que el sistema homogéneo cuya matriz es A tiene solución no trivial, es decir, existe $x_0 \in K^n$ no nulo tal que $A.x_0 = 0$. Si A fuese inversible, multiplicando por A^{-1} resultaría $x_0 = A^{-1}.A.x_0 = 0$, contradiciendo que $x_0 \neq 0$.

2.3 Matrices elementales

El método de triangulación que hemos visto para la resolución de sistemas de ecuaciones lineales se basa en la aplicación de ciertas operaciones elementales (ver Proposición 1.19) a las ecuaciones o, equivalentemente, a las filas de la matriz del sistema. Como veremos a continuación, cada una de estas operaciones puede verse como la multiplicación a izquierda de la matriz del sistema por una matriz conveniente.

A cada operación de filas en una matriz de $n \times n$, le asociaremos la matriz que se obtiene al aplicarle dicha operación a la matriz identidad I_n. Las matrices obtenidas de esta forma se denominan *matrices elementales*. Tendremos entonces tres familias de matrices elementales, correspondientes a los tres tipos de operaciones de filas permitidas. A continuación damos las definiciones precisas y estudiamos el comportamiento de las matrices elementales con respecto al producto.

Comenzamos definiendo las matrices que corresponden a la operacion *"Intercambiar dos ecuaciones"*.

1. Sean $1 \le i, j \le n$. Se define $P^{ij} \in K^{n\times n}$ como

   $$P^{ij} = I_n - E^{ii} - E^{jj} + E^{ij} + E^{ji}.$$

   Observamos que P^{ij} es la matriz que resulta al intercambiar las filas i y j en la matriz I_n.

   Es fácil verificar que, dada $B \in K^{n\times n}$ el producto $P^{ij}B$ es la matriz que resulta al intercambiar en la matriz B las filas i y j.

   En particular, $P^{ij}.P^{ij} = I_n$, es decir que P^{ij} es inversible y ${P^{ij}}^{-1} = P^{ij}$.

Ahora introducimos las matrices elementales asociadas a la operación *"Multiplicar una ecuación por una constante no nula."*

2. Sea $a \in K$, $a \neq 0$, y sea $1 \leq i \leq n$. Se define $M_i(a) \in K^{n\times n}$ como

$$M_i(a) = I_n + (a-1).E^{ii}.$$

Observamos que $M_i(a)$ es la matriz que se obtiene al mutiplicar por a la i-ésima fila de la matriz I_n.

Dada $B \in K^{n\times n}$ se tiene que

$$\big(M_i(a).B\big)_{kj} = \begin{cases} B_{kj} & \text{si } k \neq i \\ a.B_{kj} & \text{si } k = i, \end{cases}$$

es decir, $M_i(a).B$ es la matriz que resulta al multiplicar por a la i-ésima fila de B.

En particular, $M_i(a).M_i(a^{-1}) = M_i(a^{-1}).M_i(a) = I_n$, de donde $M_i(a) \in GL(n,K)$ y $(M_i(a))^{-1} = M_i(a^{-1})$.

Finalmente, la tercera de las familias de matrices elementales es la que representa la operación *"Reemplazar una ecuación por ella misma más un múltiplo de otra."*

3. Sea $a \in K$ y sean $i \neq j$ con $1 \leq i, j \leq n$. Se define $T^{ij}(a) \in K^{n\times n}$ como

$$T^{ij}(a) = I_n + a.E^{ij},$$

la matriz que se obtiene de la matriz I_n al sumarle a la i-ésima fila, a por la fila j.

Si $B \in K^{n\times n}$, entonces

$$\big(T^{ij}(a).B\big)_{kl} = \begin{cases} B_{kl} & \text{si } k \neq i \\ B_{il} + a.B_{jl} & \text{si } k = i, \end{cases}$$

o sea que $T^{ij}(a).B$ es la matriz que se obtiene de B al sumarle a la i-ésima fila, a por la fila j.

En particular, se tiene que $T^{ij}(a).T^{ij}(-a) = T^{ij}(-a).T^{ij}(a) = I_n$, con lo que $T^{ij}(a) \in GL(n,K)$ y $(T^{ij}(a))^{-1} = T^{ij}(-a)$.

De las propiedades de las matrices elementales que hemos visto, se deduce que triangular una matriz mediante operaciones sobre sus filas es multiplicarla a izquierda por matrices elementales. En forma totalmente análoga, puede verse que multiplicar una matriz a derecha por matrices elementales corresponde a efectuar operaciones sobre las columnas de la matriz.

Observación 2.14 Sea $A \in K^{n\times n}$. Entonces existen matrices elementales $E_1, \ldots, E_r \in K^{n\times n}$ tales que $E_r \ldots E_1.A$ es triangular superior. Si además, $E_r \ldots E_1.A$ *no* tiene ceros en la diagonal, existen matrices elementales $E_{r+1}, \ldots, E_s$ tales que $E_s \ldots E_{r+1}.E_r \ldots E_1.A = I_n$. En consecuencia, A es producto de matrices elementales: $A = E_1^{-1} \ldots E_s^{-1}$, y $A^{-1} = E_s \ldots E_1$.

En particular, esta observación nos dice que si por medio de la aplicación de operaciones elementales a las filas de la matriz A obtenemos la matriz identidad I, entonces aplicando las mismas operaciones en las filas de I obtendremos A^{-1}.

Por otro lado, nos da un teorema de estructura para $GL(n, K)$: así como el Teorema Fundamental de la Aritmética en $\mathbb{Z}$ dice que todo número entero no nulo es producto de enteros *primos*, la observación anterior nos dice que toda matriz en $GL(n, K)$ es producto de matrices *elementales*.

2.4 Coordenadas

Dado un K-espacio vectorial V de dimensión n y fijada una base B de V, mediante el concepto de coordenadas de un vector en la base B podremos "identificar" cada elemento de V con un vector en K^n y trabajar entonces con elementos de K^n.

2.4.1 Coordenadas de un vector en una base

Definición 2.15 Sea V un K-espacio vectorial de dimensión finita y sea $B = \{v_1, \dots, v_n\}$ una base de V. Dado $x \in V$, existen únicos $\alpha_1, \dots, \alpha_n \in K$ tales que $x = \alpha_1 v_1 + \dots + \alpha_n v_n$ (ver Proposición 1.37). El vector $(\alpha_1, \dots, \alpha_n) \in K^n$ se llama el *vector de coordenadas de x en la base B* y será denotado por $(x)_B$.

Ejemplos.

i) Sea $V = \mathbb{R}_4[X]$ y sea $B = \{1, X, X^2, X^3, X^4\}$ base de V.
Las coordenadas de $X^3 + 3X^2 - 1$ en la base B son $(X^3 + 3X^2 - 1)_B = (-1, 0, 3, 1, 0)$.
Sea $B' = \{X^4, X^3, X^2, X, 1\}$. Entonces $(X^3 + 3X^2 - 1)_{B'} = (0, 1, 3, 0, -1)$.

ii) Sea $V = \mathbb{R}^3$ y sea $E = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ la base canónica. Entonces para cada $(x, y, z) \in \mathbb{R}^3$, se tiene que $(x, y, z)_E = (x, y, z)$.

iii) Sea $V = \mathbb{R}^3$ y sea $B = \{(1, 1, 1), (1, 1, 0), (1, 0, 0)\}$. Para cada $(x, y, z) \in \mathbb{R}^3$,

$$(x, y, z) = z.(1, 1, 1) + (y - z).(1, 1, 0) + (x - y).(1, 0, 0).$$

Entonces, $(x, y, z)_B = (z, y - z, x - y)$.

Observemos que el vector de coordenadas en la base B de un elemento de $\mathbb{R}^3$ se obtiene de su vector de coordenadas en la base canónica multiplicando éste por una matriz apropiada: Si $v \in \mathbb{R}^3$ tiene coordenadas (x, y, z) en la base canónica E, entonces

$$((v)_B)^t = \begin{pmatrix} z \\ y - z \\ x - y \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 0 \end{pmatrix}}_{C(E,B)} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = C(E, B).\left((v)_E\right)^t.$$

2.4.2 Cambios de base

Dadas dos bases de un mismo K-espacio vectorial V de dimensión finita, cada elemento de V tiene asociados dos vectores de coordenadas (generalmente distintos), uno en cada una de las bases. Con la ayuda de cierta matriz, llamada de cambio de base, se pueden obtener las coordenadas de un vector con respecto a una base de V a partir de las coordenadas del vector en otra base.

Definición 2.16 Sea V un K-espacio vectorial de dimensión n, y sean $B_1 = \{v_1, \dots, v_n\}$ y $B_2 = \{w_1, \dots, w_n\}$ bases de V. Para cada $1 \le j \le n$, sean $\alpha_{ij} \in K$ $(1 \le i \le n)$ tales que $v_j = \sum\limits_{i=1}^{n} \alpha_{ij} w_i$. Se llama *matriz de cambio de base de* B_1 *a* B_2, y se nota $C(B_1, B_2) \in K^{n\times n}$, a la matriz definida por $(C(B_1, B_2))_{ij} = \alpha_{ij}$ para cada $1 \le i, j \le n$.

En otros términos, la matriz de cambio de base $C(B_1, B_2) \in K^{n\times n}$ es la matriz cuya j-ésima columna son las coordenadas en la base B_2 del j-ésimo vector de la base B_1, para cada $1 \le j \le n$.

Ejemplo. Sea $V = \mathbb{R}^3$. Consideremos las bases $B_1 = E = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ y $B_2 = \{(1, 1, 1), (1, 1, 0), (1, 0, 0)\}$. Para construir la matriz $C(B_1, B_2) \in \mathbb{R}^{3\times 3}$, comenzamos por escribir los elementos de B_1 como combinación lineal de los de B_2:

$$\begin{aligned}
(1, 0, 0) &= 0.(1, 1, 1) + 0.(1, 1, 0) + 1.(1, 0, 0)\\
(0, 1, 0) &= 0.(1, 1, 1) + 1.(1, 1, 0) + (-1).(1, 0, 0)\\
(0, 0, 1) &= 1.(1, 1, 1) + (-1)(1, 1, 0) + 0.(1, 0, 0)
\end{aligned}$$

Entonces, la matriz de cambio de base es:

$$C(B_1, B_2) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 0 \end{pmatrix}.$$

(Comparar con el Ejemplo iii) de la sección anterior.)

La proposición siguiente muestra que la matriz de cambio de base cumple la propiedad que hemos mencionado al comienzo de esta sección.

Proposición 2.17 *Sea V un K-espacio vectorial de dimensión finita, y sean B_1 y B_2 bases de V. Entonces, para cada $x \in V$,*

$$C(B_1, B_2).((x)_{B_1})^t = ((x)_{B_2})^t.$$

Demostración. Sean $B_1 = \{v_1, \dots, v_n\}$ y $B_2 = \{w_1, \dots, w_n\}$. Supongamos que, para cada $1 \le j \le n$, $v_j = \sum\limits_{i=1}^{n} \alpha_{ij} w_i$, con $\alpha_{ij} \in K$ para cada $1 \le i \le n$; es decir, $(C(B_1, B_2))_{ij} = \alpha_{ij}$ $(1 \le i, j \le n)$.

Sea $x \in V$. Si $x = \sum\limits_{k=1}^{n} a_k v_k$, entonces para cada $1 \le h \le n$,

$$\Big(C(B_1, B_2).((x)_{B_1})^t\Big)_h = \sum_{r=1}^{n} \alpha_{hr} a_r.$$

Si $b_h = \sum\limits_{r=1}^{n} \alpha_{hr} a_r$ para cada $1 \le h \le n$, por la unicidad de las coordenadas en una base, para probar que $(x)_{B_2} = (b_1, \ldots, b_n)$ basta ver que $x = \sum\limits_{h=1}^{n} b_h w_h$. Ahora,

$$\begin{aligned}\sum_{h=1}^{n} b_h w_h &= \sum_{h=1}^{n} \Big(\sum_{r=1}^{n} \alpha_{hr} a_r\Big) w_h = \sum_{h=1}^{n} \Big(\sum_{r=1}^{n} \alpha_{hr} a_r w_h\Big) = \\ &= \sum_{r=1}^{n} \Big(\sum_{h=1}^{n} \alpha_{hr} a_r w_h\Big) = \sum_{r=1}^{n} a_r \Big(\sum_{h=1}^{n} \alpha_{hr} w_h\Big) = \sum_{r=1}^{n} a_r v_r = x,\end{aligned}$$

que es lo que queríamos probar. □

Una pregunta que surge es la de la unicidad de la matriz de cambio de base: dadas dos bases B_1 y B_2, la matriz $C(B_1, B_2)$ que hemos definido transforma coordenadas en la base B_1 en coordenadas en la base B_2. ¿Existirá alguna otra matriz en $K^{n\times n}$ con esta misma propiedad? El resultado que probamos a continuación nos asegura que no.

Proposición 2.18 *Sean $A, A' \in K^{n\times n}$. Si $A.x = A'.x$ para todo $x \in K^n$, entonces $A = A'$.*

Demostración. Sea $E = \{e_1, \ldots, e_n\}$ la base canónica de K^n. Por hipótesis, $A.e_j = A'.e_j$ para cada $1 \le j \le n$. Pero

$$(A.e_j)_i = \sum_{h=1}^{n} A_{ih}(e_j)_h = A_{ij} \qquad \text{y} \qquad (A'.e_j)_i = \sum_{h=1}^{n} A'_{ih}(e_j)_h = A'_{ij}$$

para cada $1 \le i \le n$, de donde $A_{ij} = A'_{ij}$ para todo $1 \le i, j \le n$. Luego, $A = A'$. □

De las proposiciones anteriores se desprende:

Observación 2.19 Dadas dos bases B_1 y B_2 de un espacio vectorial V de dimensión n, la matriz $C(B_1, B_2)$ es la única matriz en $K^{n\times n}$ que verifica $C(B_1, B_2)((x)_{B_1})^t = ((x)_{B_2})^t$ para todo $x \in V$.

Esta observación dice que si una matriz A verifica $A.((x)_{B_1})^t = ((x)_{B_2})^t$ para todo $x \in V$, entonces necesariamente $A = C(B_1, B_2)$. Utilizando este resultado, es fácil probar las igualdades que enunciamos a continuación.

Corolario 2.20 *Sea V un K-espacio vectorial de dimensión finita, y sean B_1, B_2 y B_3 bases de V. Entonces:*

1. $C(B_1, B_3) = C(B_2, B_3).C(B_1, B_2)$.

2. $C(B_2, B_1) = C(B_1, B_2)^{-1}$.

Para terminar, probaremos algunos resultados que relacionan matrices inversibles con cambios de base.

Proposición 2.21 *Sea $A \in GL(n, K)$. Existen bases B_1, B_2 de K^n tales que $A = C(B_1, B_2)$.*

Demostración. Supongamos que $A_{ij} = a_{ij}$ para cada $1 \le i \le n$, $1 \le j \le n$.

Sea $B_2 = E = \{e_1, \dots, e_n\}$, la base canónica de K^n, y sea $B_1 = \Big\{ \sum\limits_{i=1}^{n} a_{i1}.e_i, \dots, \sum\limits_{i=1}^{n} a_{in}.e_i \Big\}$.

Veamos que B_1 es una base de K^n, para lo cual basta ver que B_1 es un conjunto linealmente independiente. Supongamos que $\sum\limits_{j=1}^{n} \alpha_j \Big(\sum\limits_{i=1}^{n} a_{ij}.e_i \Big) = 0$. Entonces

$$0 = \sum_{j=1}^{n} \Big(\sum_{i=1}^{n} \alpha_j a_{ij} e_i \Big) = \sum_{i=1}^{n} \Big(\sum_{j=1}^{n} \alpha_j a_{ij} e_i \Big) = \sum_{i=1}^{n} \Big(\sum_{j=1}^{n} \alpha_j a_{ij} \Big) e_i,$$

de donde $\sum\limits_{j=1}^{n} a_{ij}\alpha_j = 0$ para cada $1 \le i \le n$, o equivalentemente,

$$A. \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = 0.$$

Como A es inversible, esto implica que $\alpha_1 = \cdots = \alpha_n = 0$. Luego B_1 es linealmente independiente y, en consecuencia, una base de $K^{n \times n}$.

Es claro que $C(B_1, E) = A$. $\square$

Proposición 2.22 *Sea $A \in GL(n, K)$ y sea B una base de K^n. Entonces:*

i) *Existe una base B_1 de K^n tal que $A = C(B_1, B)$.*

ii) *Existe una base B_2 de K^n tal que $A = C(B, B_2)$.*

Demostración.

$i)$ Se prueba en forma análoga a la proposición anterior, reemplazando la base canónica E por la base B dada.

$ii)$ Por la parte $i)$, dadas $A^{-1} \in GL(n, K)$ y la base B de K^n, existe una base B_2 de K^n tal que $A^{-1} = C(B_2, B)$. En consecuencia, $A = C(B_2, B)^{-1} = C(B, B_2)$. $\square$

2.5 Ejercicios

Ejercicio 1. Probar que los siguientes conjuntos son subespacios de $K^{n\times n}$ y calcular su dimensión.

i) $S_1 = \{A \in K^{n\times n} / A = A^t\}$ (matrices simétricas)

ii) $S_2 = \{A \in K^{n\times n} / A = -A^t\}$ (matrices antisimétricas)

iii) $S_3 = \{A \in K^{n\times n} / A_{ij} = 0 \text{ si } i > j\}$ (matrices triangulares superiores)

iv) $S_4 = \{A \in K^{n\times n} / A_{ij} = 0 \text{ si } i \neq j\}$ (matrices diagonales)

v) $S_5 = \{A \in K^{n\times n} / A_{ij} = 0 \text{ si } i \neq j \text{ y } A_{11} = A_{22} = ... = A_{nn}\}$ (matrices escalares)

vi) $S_6 = \{A \in K^{n\times n} / tr(A) = 0\}$

Ejercicio 2. Sean S_1 , S_2 , S_5 y S_6 los subespacios del ejercicio anterior.

i) Probar que $S_1 \oplus S_2 = K^{n\times n}$ si $2 \neq 0$ en K.

ii) Probar que $S_5 \oplus S_6 = K^{n\times n}$ si $K = \mathbb{Q}$, $\mathbb{R}$ ó $\mathbb{C}$.

Ejercicio 3. Sean m, n y $r \in \mathbb{N}$. Probar:

i) Si $A \in K^{m\times n}$, $B \in K^{n\times r}$ con $B = (b_{ij})$ y, para $1 \leq j \leq r$, $B_j = \begin{pmatrix} b_{1j} \\ \vdots \\ b_{nj} \end{pmatrix}$ (la columna j-ésima de B), entonces $A.B = (\, A.B_1 \mid \ldots \mid A.B_r)$ (es decir, $A.B_j$ es la columna j-ésima de $A.B$).

ii) *(Multiplicación de matrices por bloques.)*
Sean $A,\, A' \in K^{n\times n}$; $B,\, B' \in K^{n\times m}$; $C,\, C' \in K^{m\times n}$ y $D,\, D' \in K^{m\times m}$.
Sean $M,\, M' \in K^{(n+m)\times(n+m)}$ definidas por $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ y $M' = \begin{pmatrix} A' & B' \\ C' & D' \end{pmatrix}$.
Entonces $M.M' = \begin{pmatrix} A.A' + B.C' & A.B' + B.D' \\ C.A' + D.C' & C.B' + D.D' \end{pmatrix}$.

Ejercicio 4.

i) Probar que, $\forall\, n \in \mathbb{N}$, $n \geq 2$, el producto de matrices en $K^{n\times n}$ no es conmutativo.

ii) Caracterizar el conjunto $\{A \in K^{n\times n} / A.B = B.A \;\; \forall\, B \in K^{n\times n}\}$.

iii) Sea $A \in K^{n\times n}$. Probar que el conjunto S de todas las matrices que conmutan con A es un subespacio de $K^{n\times n}$. Probar que $I_n \in S$ y que $A^j \in S \;\; \forall\, j \in \mathbb{N}$.

iv) Sea $A \in K^{n\times n}$ con $n \geq 2$. Probar que el conjunto $\{I_n, A, A^2, A^3, \dots, A^{n^2-1}\}$ es linealmente dependiente.

v) Dar condiciones necesarias y suficientes sobre A y $B \in K^{n\times n}$ para que

   a) $(A+B)^2 = A^2 + 2AB + B^2$

   b) $A^2 - B^2 = (A-B).(A+B)$

vi) Probar que si A y $B \in K^{n\times n}$ no necesariamente vale $A^2.B^2 = (A.B)^2$

Ejercicio 5. Sean A, B y $C \in K^{n\times n}$ $(n \geq 2)$. Mostrar la falsedad de las siguientes afirmaciones:

i) $A.B = 0 \;\Rightarrow\; A = 0$ ó $B = 0$

ii) $A.B = A.C$ y $A \neq 0 \;\Rightarrow\; B = C$

iii) $A.B = 0 \;\Rightarrow\; B.A = 0$

iv) $A^j = 0 \;\Rightarrow\; A = 0$

v) $A^2 = A \;\Rightarrow\; A = 0$ ó $A = I_n$

Ejercicio 6. Sea $A \in K^{n\times n}$. Probar que el conjunto $T = \{B \in K^{n\times n} / A.B = 0\}$ es un subespacio de $K^{n\times n}$. Si $S \subset K^n$ es el subespacio de soluciones del sistema homogéneo cuya matriz asociada es A, probar que $\dim T = n.\dim S$.

Ejercicio 7. Sean $A, A' \in K^{m\times n}$; $B \in K^{n\times r}$; $D, D' \in K^{n\times n}$; $\alpha \in K$. Probar:

i) $(A+A')^t = A^t + (A')^t$

ii) $(\alpha.A)^t = \alpha.A^t$

iii) $(A.B)^t = B^t.A^t$

iv) $tr(D+D') = tr(D) + tr(D')$

v) $tr(\alpha.D) = \alpha.tr(D)$

vi) $tr(D.D') = tr(D'.D)$

Ejercicio 8. Sean A y $B \in K^{n\times n}$.

i) Probar que si A y B son triangulares superiores, $A.B$ es triangular superior.

ii) Probar que si A y B son diagonales, $A.B$ es diagonal.

iii) Probar que si A es estrictamente triangular superior (es decir, $A_{ij} = 0$ si $i \geq j$), $A^n = 0$.

Ejercicio 9. Sea $A \in K^{n\times n}$.

i) Probar que $A.A^t$ y $A^t.A$ son simétricas. Encontrar un ejemplo donde $A.A^t \neq A^t.A$.

ii) El producto de dos matrices simétricas, ¿es una matriz simétrica?

iii) Si $K = \mathbb{R}$, probar que $A = 0 \iff A.A^t = 0 \iff tr(A.A^t) = 0$.

Ejercicio 10. Sea $A = \begin{pmatrix} 4 & -1 \\ 8 & -2 \end{pmatrix} \in \mathbb{R}^{2\times 2}$.

i) Hallar b y $c \in \mathbb{R}$ tales que $A^2 + b.A + c.I_2 = 0$.

ii) Calcular $A^n \ \forall\, n \in \mathbb{N}$.

Ejercicio 11. Sea $A \in K^{2\times 2}$ con $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ y sea $\Delta = a.d - b.c$. Probar que, si $\Delta \neq 0$, $A \in GL(2, K)$ y $A^{-1} = \dfrac{1}{\Delta} . \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$.

Ejercicio 12. Sea $A \in GL(n, K)$ y $B\,,\ C \in K^{n\times m}$. Probar:

i) $A.B = A.C \ \Rightarrow\ B = C$

ii) $A.B = 0 \ \Rightarrow\ B = 0$

Ejercicio 13. Decidir si cada una de las siguientes afirmaciones es verdadera o falsa.

i) $A\,,\ B \in GL(n, K) \ \Rightarrow\ A + B \in GL(n, K)$

ii) $A \in GL(n, K) \iff A^t \in GL(n, K)$

iii) $tr(A) = 0 \ \Rightarrow\ A \notin GL(n, K)$

iv) A nilpotente (es decir, $\exists\, j \in \mathbb{N} \,/\, A^j = 0$) $\ \Rightarrow\ A \notin GL(n, K)$

Ejercicio 14. Sea $A \in K^{m\times n}$ y sea $b \in K^m$. Sea $H = \{x \in K^n \,/\, A.x = b\}$. Probar:

i) Si $C \in GL(m, K)$, entonces $H = \{x \in K^n \,/\, (C.A).x = C.b\}$.

ii) Si $m = n$ y $A \in GL(n, K)$, entonces H tiene un solo elemento. ¿Cuál es? (Notar que esto significa que si A es inversible, cualquier sistema lineal cuya matriz asociada sea A tiene solución única).

Ejercicio 15.

i) Sea $A = T^{12}(1) \in \mathbb{R}^{2\times 2}$. Calcular A^{20} y $20.A$.

ii) Calcular $(P^{ij})^{15}$ y $(P^{ij})^{16}$.

iii) Sea $B = M_3(2) \in \mathbb{R}^{4\times 4}$. Calcular B^{20} y $20.B$.

Ejercicio 16. Averiguar si las siguientes matrices son inversibles y en caso afirmativo exhibir sus inversas.

i) $A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$

iv) $A = \begin{pmatrix} 2 & 1 & 3 & 1 & 2 \\ 0 & 5 & -1 & 8 & 2 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}$

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

v) $A = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix}$

iii) $A = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta & 0 \\ \operatorname{sen}\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$

Escribir las que sean inversibles como producto de matrices elementales.

Ejercicio 17. Sea $A \in K^{n\times n}$ y sea $b \in K^n$.

i) Probar que el sistema $A.x = b$ tiene solución única $\iff A \in GL(n, K)$.

ii) Probar que $A \in GL(n, K) \iff$ las filas de A son linealmente independientes $\iff$ las columnas de A son linealmente independientes.

Ejercicio 18. Sea $A \in K^{n\times n}$. Probar que:

$$\exists\, B \in K^{n\times n} / \, B.A = I_n \iff A \in GL(n, K).$$

Deducir que $\exists\, B \in K^{n\times n} / \, A.B = I_n \iff A \in GL(n, K)$.

Ejercicio 19. Encontrar las coordenadas de $v \in V$ respecto de la base B en los siguientes casos:

i) $V = K^n$; $v = (x_1, \dots, x_n)$ y B la base canónica

ii) $V = \mathbb{R}^3$; $v = (1, 2, -1)$ y $B = \{(1, 2, -1), (0, 1, 1), (0, 0, 2)\}$

iii) $V = \mathbb{R}^3$; $v = (1, -1, 2)$ y $B = \{(1, 2, -1), (2, 1, 3), (1, 3, 2)\}$

iv) $V = \mathbb{R}^3$; $v = (x_1, x_2, x_3)$ y $B = \{(1, 2, -1), (2, 1, 3), (1, 3, 2)\}$

v) $V = \mathbb{R}_3[X]$; $v = 2X^2 - X^3$ y $B = \{3\,,\, 1 + X\,,\, X^2 + 5\,,\, X^3 + X^2\}$

vi) $V = \mathbb{R}^{2\times 2}$; $v = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ y $B = \left\{ \begin{pmatrix} 1 & 3 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 1 & 4 \\ 3 & 2 \end{pmatrix}, \begin{pmatrix} 0 & 2 \\ 1 & -1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 2 & 5 \end{pmatrix} \right\}$

Ejercicio 20. Calcular $C(B, B')$ en los siguientes casos:

i) $V = \mathbb{R}^2,\ B = \{(1,1),(1,2)\}\,,\ B' = \{(-1,3),(2,5)\}$

ii) $V = \mathbb{R}^3,\ B = \{(1,1,0),(0,1,1),(1,0,1)\}\,,\ B' = \{(-1,1,1),(2,0,1),(1,-1,3)\}$

iii) $V = \mathbb{R}^3,\ B = \{(1,1,0),(0,1,1),(1,0,1)\}\,,\ B' = \{(0,1,1),(1,0,1),(1,1,0)\}$

iv) $V = \mathbb{R}_2[X],\ B = \{3\,,\,1+X\,,\,X^2\}\,,\ B' = \{1\,,\,X+3\,,\,X^2+X\}$

v) $V = \mathbb{R}^4,\ B = \{v_1, v_2, v_3, v_4\}\,,\ B' = \{v_3, v_1, v_4, v_2\}$

vi) $V = \mathbb{R}^{2\times 2},\ B = \{E^{11},\, E^{12},\, E^{21},\, E^{22}\}\,,$
$B' = \left\{ \begin{pmatrix} 1 & 3 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 1 & 4 \\ 3 & 2 \end{pmatrix}, \begin{pmatrix} 0 & 2 \\ 1 & -1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 2 & 5 \end{pmatrix} \right\}$

Ejercicio 21. Dado $v \in V$ y las bases B y B', hallar las coordenadas de v respecto de B y utilizando la matriz de cambio de base, las coordenadas de v respecto de B'.

i) $v = (2,3)$ y B, B' como en el Ejercicio 20, i)

ii) $v = (-1,5,6)$ y B, B' como en el Ejercicio 20, ii)

iii) $v = (-1,5,6)$ y B, B' como en el Ejercicio 20, iii)

iv) $v = 2.v_1 + 3.v_2 - 5.v_3 + 7.v_4$ y B, B' como en el Ejercicio 20, v)

v) $v = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ y B, B' como en el Ejercicio 20, vi)

Ejercicio 22. Sea V un K-espacio vectorial de dimensión n y sean B, B' y B'' bases de V.

i) Probar que $C(B, B'') = C(B', B'').\, C(B, B')$.

ii) Deducir que $C(B, B') \in GL(n, K)$ con $C(B, B')^{-1} = C(B', B)$.

Ejercicio 23. Dadas la matriz $M = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}$ y la base $B = \{v_1, v_2, v_3\}$ de K^3, hallar una base B' tal que $M = C(B, B')$.

Ejercicio 24. Dadas la matriz $M = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}$ y la base $B' = \{v_1, v_2, v_3\}$ de K^3, hallar una base B tal que $M = C(B, B')$.

Capítulo 3

Transformaciones lineales

Las transformaciones lineales son las funciones con las que trabajaremos en Álgebra Lineal. Se trata de funciones entre K-espacios vectoriales que son compatibles con la estructura (es decir, con la operación y la acción) de estos espacios.

3.1 Definiciones, ejemplos y propiedades básicas

En esta sección introduciremos la noción de transformación lineal, así como también ciertas nociones básicas asociadas a estas funciones.

3.1.1 Transformaciones lineales

Definición 3.1 Sean $(V, +_V, \cdot_V)$ y $(W, +_W, \cdot_W)$ dos K-espacios vectoriales. Una función $f : V \to W$ se llama una *transformación lineal (u homomorfismo, o simplemente morfismo) de* V *en* W si cumple:

i) $f(v +_V v') = f(v) +_W f(v') \quad \forall\, v, v' \in V.$

ii) $f(\lambda \cdot_V v) = \lambda \cdot_W f(v) \quad \forall\, \lambda \in K, \forall\, v \in V.$

Observación 3.2 Si $f : V \to W$ es una transformación lineal, entonces $f(0_V) = 0_W$.

En efecto, puesto que $f(0_V) = f(0_V + 0_V) = f(0_V) + f(0_V)$, entonces

$$\begin{aligned} 0_W &= f(0_V) + (-f(0_V)) = \Big(f(0_V) + f(0_V)\Big) + (-f(0_V)) = \\ &= f(0_V) + \Big(f(0_V) + (-f(0_V))\Big) = f(0_V) + 0_W = f(0_V). \end{aligned}$$

Ejemplos.

1. Sean V y W dos K-espacios vectoriales. Entonces $0 : V \to W$, definida por $0(x) = 0_W$ $\forall\, x \in V$, es una transformación lineal.

2. Si V es un K-espacio vectorial, $id : V \to V$ definida por $id(x) = x$ es una transformación lineal.

3. Sea $A \in K^{m\times n}$. Entonces $f_A : K^n \to K^m$ definida por $f_A(x) = (A.x^t)^t$ es una transformación lineal.

4. $f : K[X] \to K[X]$, $f(P) = P'$ es una transformación lineal.

5. $F : \mathcal{C}(\mathbb{R}) \to \mathbb{R}$, donde $\mathcal{C}(\mathbb{R}) = \{f : \mathbb{R} \to \mathbb{R} \mid f \text{ es continua}\}$, $F(g) = \int_0^1 g(x)\, dx$ es una transformación lineal.

Como hemos mencionado al comienzo, las transformaciones lineales respetan la estructura de K-espacio vectorial. Esto hace que en algunos casos se respete la estructura de subespacio, por ejemplo en las imágenes y pre-imágenes de subespacios por transformaciones lineales:

Proposición 3.3 *Sea $f : V \to W$ una transformación lineal. Entonces:*

1. *Si S es un subespacio de V, entonces $f(S)$ es un subespacio de W.*

2. *Si T es un subespacio de W, entonces $f^{-1}(W)$ es un subespacio de V.*

Demostración.

1. Sea $S \subseteq V$ un subespacio y consideremos $f(S) = \{w \in W \,/\, \exists\, s \in S, f(s) = w\}$.

   (a) $0_W \in f(S)$, puesto que $f(0_V) = 0_W$ y $0_V \in S$.

   (b) Sean $w, w' \in f(S)$. Entonces existen $s, s' \in S$ tales que $w = f(s)$ y $w' = f(s')$. Luego $w + w' = f(s) + f(s') = f(s + s') \in f(S)$, puesto que $s + s' \in S$.

   (c) Sean $\lambda \in K$ y $w \in f(S)$. Existe $s \in S$ tal que $w = f(s)$. Entonces $\lambda\cdot w = \lambda\cdot f(s) = f(\lambda \cdot s) \in f(S)$, puesto que $\lambda \cdot s \in S$.

2. Sea T un subespacio de W y consideremos $f^{-1}(T) = \{v \in V \,/\, f(v) \in T\}$.

   (a) $0_V \in f^{-1}(T)$, puesto que $f(0_V) = 0_W \in T$.

   (b) Sean $v, v' \in f^{-1}(T)$. Entonces $f(v), f(v') \in T$ y, por lo tanto, $f(v + v') = f(v) + f(v') \in T$. Luego $v + v' \in f^{-1}(T)$.

   (c) Sean $\lambda \in K$, $v \in f^{-1}(T)$. Entonces $f(v) \in T$ y, en consecuencia, $f(\lambda\cdot v) = \lambda\cdot f(v) \in T$. Luego $\lambda \cdot v \in f^{-1}(T)$. □

De la Definición 3.1 se deduce inmediatamente que una transformación lineal preserva combinaciones lineales. Veremos que, debido a esto, una transformación lineal queda unívocamente determinada por los valores que toma en los elementos de una base cualquiera de su dominio. Comenzamos con un ejemplo.

Ejemplo. Hallar, si es posible, una transformación lineal $f : \mathbb{R}^2 \to \mathbb{R}^2$ que verifique $f(1,1) = (0,1)$ y $f(1,0) = (2,3)$.

Dado $(x_1, x_2) \in \mathbb{R}^2$ se tiene que $(x_1, x_2) = x_2(1,1) + (x_1 - x_2)(1,0)$. Entonces, si f verifica lo pedido, debe ser

$$\begin{aligned} f(x_1, x_2) &= x_2.f(1,1) + (x_1 - x_2).f(1,0) = x_2.(0,1) + (x_1 - x_2).(2,3) \\ &= (2x_1 - 2x_2, 3x_1 - 2x_2). \end{aligned}$$

Además, es fácil ver que esta función es una transformación lineal y que vale $f(1,1) = (0,1)$ y $f(1,0) = (2,3)$.

Luego, $f(x_1, x_2) = (2x_1 - 2x_2, 3x_1 - 2x_2)$ es la única transformación lineal que satisface lo pedido.

La construcción realizada en el ejemplo puede hacerse en general. Por simplicidad, lo probaremos para el caso en que el dominio de la transformación lineal es un K-espacio vectorial de dimensión finita.

Proposición 3.4 *Sean V y W dos K-espacios vectoriales, V de dimensión finita. Sea $B = \{v_1, \dots, v_n\}$ una base de V y sean $w_1, \dots, w_n \in W$ vectores arbitrarios. Entonces existe una única transformación lineal $f : V \to W$ tal que $f(v_i) = w_i$ para cada $1 \le i \le n$.*

Demostración.

Existencia. Dado $v \in V$ existen únicos $\alpha_1, \dots, \alpha_n \in K$ tales que $v = \sum\limits_{i=1}^{n} \alpha_i v_i$, es decir, $(\alpha_1, \dots, \alpha_n) = (v)_B$ es el vector de coordenadas de v en la base B. Definimos

$$f(v) = \sum_{i=1}^{n} \alpha_i w_i.$$

(Observar que no hay ambigüedad en la definición de f por la unicidad de $\alpha_1, \dots, \alpha_n$.)

Veamos que f es una transformación lineal:

Sean $v, v' \in V$. Supongamos que $v = \sum\limits_{i=1}^{n} \alpha_i v_i$ y $v' = \sum\limits_{i=1}^{n} \alpha'_i v_i$. Entonces

$$v + v' = \sum_{i=1}^{n} \alpha_i v_i + \sum_{i=1}^{n} \alpha'_i v_i = \sum_{i=1}^{n} (\alpha_i + \alpha'_i) v_i,$$

y, en consecuencia,

$$f(v + v') = \sum_{i=1}^{n} (\alpha_i + \alpha'_i) w_i = \sum_{i=1}^{n} \alpha_i w_i + \sum_{i=1}^{n} \alpha'_i w_i = f(v) + f(v').$$

De manera análoga se prueba que $f(\lambda v) = \lambda f(v)\ \forall\, \lambda \in K,\ \forall\, v \in V$.

Unicidad. Supongamos que f y g son dos transformaciones lineales de V en W tales que $f(v_i) = w_i$ y $g(v_i) = w_i$ para cada $1 \le i \le n$. Entonces, dado $v \in V$, si $v = \sum\limits_{i=1}^{n} \alpha_i v_i$, por la linealidad de f y g se tiene que

$$f(v) = \sum_{i=1}^{n} \alpha_i f(v_i) = \sum_{i=1}^{n} \alpha_i g(v_i) = g(v).$$

Luego, $f(v) = g(v)$ para todo $v \in V$, de donde $f = g$. □

Observación 3.5 Con una demostración análoga a la de la proposición anterior se prueba que, si V y W son dos K-espacios vectoriales (V no necesariamente de dimensión finita), $B = \{v_i : i \in I\}$ una base de V y $\{w_i : i \in I\} \subset W$, existe una única transformación lineal $f : V \to W$ tal que $f(v_i) = w_i\ \forall\, i \in I$.

Teniendo en cuenta que las transformaciones lineales son funciones entre conjuntos, tiene sentido estudiar la validez de las propiedades usuales de funciones: inyectividad, suryectividad y biyectividad. Las transformaciones lineales que verifican alguna de estas propiedades reciben nombres particulares:

Definición 3.6 Sean V y W dos K-espacios vectoriales, y sea $f : V \to W$ una transformación lineal. Se dice que:

1. f es un *monomorfismo* si f es inyectiva.
2. f es un *epimorfismo* si f es suryectiva.
3. f es un *isomorfismo* si f es biyectiva.

En algunos casos, consideraremos transformaciones lineales de un K-espacio vectorial en sí mismo:

Definición 3.7 Sea V un K-espacio vectorial. Una transformación lineal $f : V \to V$ se llama un *endomorfismo* de V. Si f es un endomorfismo que es además un isomorfismo, entonces se dice que es un *automorfismo*.

3.1.2 Núcleo e imagen de una transformación lineal

A una transformación lineal $f : V \to W$ podemos asociarle un subespacio de V, llamado su núcleo, que de alguna manera mide el tamaño de la pre-imagen por f de un elemento de su imagen. En particular, conocer este subespacio nos permitirá determinar si f es inyectiva.

Definición 3.8 Sean V y W dos K-espacios vectoriales, y sea $f : V \to W$ una transformación lineal. Se llama *núcleo de* f al conjunto $\mathrm{Nu}(f) = \{v \in V \ /\ f(v) = 0\} = f^{-1}(\{0\})$.

Observamos que si $f : V \to W$ es una transformación lineal, $\mathrm{Nu}(f)$ es un subespacio de V, puesto que es la pre-imagen por f del subespacio $\{0\} \subset W$ (ver Proposición 3.3).

Ejemplo. Sea $f : \mathbb{R}^3 \to \mathbb{R}^2$ la transformación lineal definida por $f(x_1, x_2, x_3) = (x_1, x_2)$. Entonces

$$\begin{aligned} \mathrm{Nu}(f) &= \{(x_1, x_2, x_3) \in \mathbb{R}^3 : f(x_1, x_2, x_3) = 0\} \\ &= \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 = x_2 = 0\} \\ &= <(0, 0, 1)> . \end{aligned}$$

La siguiente proposición nos da una manera de determinar si una transformación lineal es un monomorfismo considerando simplemente su núcleo.

Proposición 3.9 *Sea $f : V \to W$ una transformación lineal. Entonces*

$$f \text{ es monomorfismo} \iff \mathrm{Nu}(f) = \{0\}$$

Demostración.

($\Rightarrow$) Si f es un monomorfismo, entonces es una función inyectiva. En particular, existe a lo sumo un elemento $v \in V$ tal que $f(v) = 0$. Puesto que $f(0) = 0$, debe ser $v = 0$. Luego, $\mathrm{Nu}(f) = \{0\}$.

($\Leftarrow$) Sean $v, v' \in V$. Supongamos que $f(v) = f(v')$. Entonces $f(v - v') = f(v) - f(v') = 0$, con lo que $v - v' \in \mathrm{Nu}(f)$ y por lo tanto, la hipótesis $\mathrm{Nu}(f) = \{0\}$ implica que $v - v' = 0$, es decir, $v = v'$. Luego f es inyectiva. □

Otro conjunto importante asociado a una transformación lineal es su *imagen*. Recordamos que si $f : V \to W$, su imagen se define como $\mathrm{Im}(f) = \{w \in W \,/\, \exists\, v \in V, f(v) = w\}$. De la Proposición 3.3 se desprende que la imagen de una transformación lineal $f : V \to W$ resulta ser un subespacio de W.

Ejemplo. Hallar la imagen de la transformación lineal $f : \mathbb{R}^3 \to \mathbb{R}^3$ definida como $f(x_1, x_2, x_3) = (x_1 - x_2, -x_1 + x_2, 2x_1 - 2x_2 + x_3)$.

Por definición,

$$\begin{aligned} \mathrm{Im}(f) &= \{y \in \mathbb{R}^3 \,/\, \exists\, x \in \mathbb{R}^3,\ f(x) = y\} \\ &= \{y \in \mathbb{R}^3 \,/\, \exists\, (x_1, x_2, x_3) \in \mathbb{R}^3,\ (x_1 - x_2, x_1 - x_2, 2x_1 - 2x_2 + x_3) = y\}. \end{aligned}$$

Entonces, un elemento de y pertenece a $\mathrm{Im}(f)$ si y sólo si es de la forma

$$\begin{aligned} y &= (x_1 - x_2, -x_1 + x_2, 2x_1 - 2x_2 + x_3) \\ &= (x_1, -x_1, 2x_1) + (-x_2, x_2, -2x_2) + (0, 0, x_3) \\ &= x_1.(1, -1, 2) + x_2.(-1, 1, -2) + x_3.(0, 0, 1). \end{aligned}$$

Luego, $\mathrm{Im}(f) = < (1,-1,2), (-1,1,-2), (0,0,1) > = < (1,-1,2), (0,0,1) >$.

Otra manera de calcular la imagen de f, teniendo en cuenta que es una transformación lineal, es la siguiente:

Consideremos un sistema de generadores de $\mathbb{R}^3$, por ejemplo la base canónica $\{e_1, e_2, e_3\}$. Para cada $x \in \mathbb{R}^3$ se tiene que $x = x_1.e_1 + x_2.e_2 + x_3.e_3$, de donde resulta que

$$f(x) = x_1.f(e_1) + x_2.f(e_2) + x_3.f(e_3).$$

Luego,

$$\begin{aligned}\mathrm{Im}(f) &= \{f(x) : x \in \mathbb{R}^3\} = < f(e_1), f(e_2), f(e_3) > = \\ &= < (1,-1,2), (-1,1,-2), (0,0,1) > = < (1,-1,2), (0,0,1) >.\end{aligned}$$

La proposición siguiente generaliza el segundo de los procedimientos utilizados en el ejemplo anterior para el cálculo de la imagen de una transformación lineal.

Proposición 3.10 *Sean V y W dos K-espacios vectoriales y sea $f : V \to W$ una transformación lineal. Entonces, si $\{v_i : i \in I\}$ es un sistema de generadores de V, $\{f(v_i) : i \in I\}$ es un sistema de generadores de* $\mathrm{Im}(f)$.

Demostración. Por definición, $\mathrm{Im}(f) = \{w \in W \ / \ \exists\, v \in V, f(v) = w\} = \{f(v) : v \in V\}$.

Si $\{v_i : i \in I\}$ es un sistema de generadores de V, para cada $v \in V$, existen $i_1, \dots, i_n \in I$ y elementos $\alpha_{i_j} \in K$ tales que $v = \sum\limits_{j=1}^{n} \alpha_{i_j} v_{i_j}$. Luego

$$f(v) = \sum_{j=1}^{n} \alpha_{i_j} f(v_{i_j}) \ \in \ < \{f(v_i) : i \in I\} >.$$

Esto prueba que $\mathrm{Im}(f) \subseteq \ < \{f(v_i) : i \in I\} >$. Es claro que vale la otra inclusión, ya que $f(v_i) \in \mathrm{Im}(f)$ para cada $i \in I$.

Luego, $\mathrm{Im}(f) = < \{f(v_i) : i \in I\} >$. □

Corolario 3.11 *Sean V y W dos K-espacios vectoriales y sea $f : V \to W$ una transformación lineal. Si V es de dimensión finita, entonces* $\mathrm{Im}(f)$ *también lo es y se tiene que* $\dim(\mathrm{Im}(f)) \leq \dim V$.

Corolario 3.12 *Si $f : V \to W$ es un epimorfismo, y $\{v_i : i \in I\}$ es un sistema de generadores de V, entonces $\{f(v_i) : i \in I\}$ es un sistema de generadores de W.*

Ejemplo. Sean S y T los subespacios de $\mathbb{R}^3$ definidos por $S = \{x \in \mathbb{R}^3 \,/\, x_1 - x_2 = 0\}$ y $T = \{x \in \mathbb{R}^3 \,/\, x_3 = 0\}$. Hallar una transformación lineal $f : \mathbb{R}^3 \to \mathbb{R}^3$ tal que $f(S) = T$.

Sabemos que para definir una transformación lineal $f : \mathbb{R}^3 \to \mathbb{R}^3$ basta con especificar los valores que toma sobre los elementos de una base de $\mathbb{R}^3$.

Consideramos entonces una base de S, por ejemplo $\{(1,1,0),(0,0,1)\}$, y la extendemos a una base de $\mathbb{R}^3$, por ejemplo $\{(1,1,0),(0,0,1),(1,0,0)\}$. Teniendo en cuenta que $T = <(1,0,0),(0,1,0)>$, definimos:

$$f(1,1,0) = (1,0,0), \quad f(0,0,1) = (0,1,0), \quad f(1,0,0) = (0,0,1).$$

Entonces $f(S) = < f(1,1,0), f(0,0,1) > = < (1,0,0),(0,1,0) > = T$.

Observemos que si $f : V \to W$ es un epimorfismo y $\{v_i : i \in I\}$ es una base de V, entonces $\{f(v_i) : i \in I\}$ no es necesariamente una base de $\mathrm{Im}(f)$: Por el corolario anterior, es un sistema de generadores, pero podría no ser un conjunto linealmente independiente, como puede verse en el ejemplo presentado en la página 69.

Esto es consecuencia de que una transformación lineal arbitraria no preserva independencia lineal. En la proposición siguiente veremos que esto sí es válido para el caso de monomorfismos. Sin embargo, si $f : V \to W$ no es un monomorfismo, existe $v \in V$, $v \neq 0$, tal que $f(v) = 0$, con lo cual $\{v\} \subset V$ es un conjunto linealmente independiente, pero $\{f(v)\} = \{0\} \subset W$ no lo es.

Proposición 3.13 *Sean V y W dos K-espacios vectoriales y sea $f : V \to W$ un monomorfismo. Entonces, si $\{v_i : i \in I\} \subset V$ es un conjunto linealmente independiente, $\{f(v_i) : i \in I\} \subset W$ es un conjunto linealmente independiente.*

Demostración. Supongamos que una combinación lineal de $\{f(v_i) : i \in I\}$ satisface $\sum\limits_{i \in I} \alpha_i f(v_i) = 0$. Como f es una transformación lineal, entonces $f\Big(\sum\limits_{i \in I} \alpha_i v_i\Big) = 0$, y como es un monomorfismo, debe ser $\sum\limits_{i \in I} \alpha_i v_i = 0$. La independencia lineal de $\{v_i : i \in I\}$ implica que $\alpha_i = 0 \ \forall\, i \in I$. □

Corolario 3.14 *Si $f : V \to W$ es un monomorfismo y $B = \{v_i : i \in I\}$ es una base de V, entonces $\{f(v_i) : i \in I\}$ es una base de $\mathrm{Im}(f)$. En particular, si V es un K-espacio vectorial de dimensión finita,* $\dim(\mathrm{Im}(f)) = \dim V$.

Teniendo en cuenta que un isomorfismo es una transformación lineal que es a la vez un epimorfismo y un monomorfismo, de los Corolarios 3.12 y 3.14 se deduce:

Corolario 3.15 *Sean V y W dos K-espacios vectoriales y sea $f : V \to W$ un isomorfismo. Entonces para toda base B de V, $f(B)$ es una base de W. En particular, si V es de dimensión finita, W también lo es y* $\dim V = \dim W$.

3.1.3 Composición de transformaciones lineales

La composición de funciones usual puede realizarse, en particular, entre dos transformaciones lineales. El resultado es, en este caso, una nueva transformación lineal.

Proposición 3.16 *Sean V, W y Z K-espacios vectoriales. Sean $f : V \to W$ y $g : W \to Z$ transformaciones lineales. Entonces $g \circ f : V \to Z$ es una transformación lineal.*

Demostración. Sean $v, v' \in V$. Entonces

$$g \circ f(v + v') = g\big(f(v + v')\big) = g\big(f(v) + f(v')\big) = g(f(v)) + g(f(v')) = g \circ f(v) + g \circ f(v').$$

Análogamente, si $\lambda \in K$ y $v \in V$, se tiene que

$$g \circ f(\lambda \cdot v) = g(f(\lambda \cdot v)) = g(\lambda \cdot f(v)) = \lambda \cdot g(f(v)) = \lambda \cdot (g \circ f(v)). \qquad \square$$

Finalmente, analizamos las propiedades de la función inversa de una transformación lineal biyectiva (es decir, un isomorfismo).

Proposición 3.17 *Sean V y W dos K-espacios vectoriales y sea $f : V \to W$ una transformación lineal. Si f es un isomorfismo, entonces $f^{-1} : W \to V$ es una transformación lineal (que resulta ser un isomorfismo).*

Demostración. Sean $w, w' \in W$. Como f es un isomorfismo, existen únicos $v, v' \in V$ tales que $w = f(v)$ y $w' = f(v')$. Entonces

$$f^{-1}(w + w') = f^{-1}(f(v) + f(v')) = f^{-1}(f(v + v')) = v + v' = f^{-1}(w) + f^{-1}(w').$$

Dados $w \in W$ y $\lambda \in K$, existe un único $v \in V$ tal que $w = f(v)$. Entonces

$$f^{-1}(\lambda \cdot w) = f^{-1}(\lambda \cdot f(v)) = f^{-1}(f(\lambda \cdot v)) = \lambda \cdot v = \lambda \cdot (f^{-1}(w)).$$

Luego, f^{-1} es una transformación lineal. Es claro que es biyectiva. $\square$

3.2 Espacios vectoriales de dimensión finita

Al estudiar espacios vectoriales de dimensión finita en los capítulos anteriores, dijimos que podríamos trabajar en un K-espacio vectorial arbitrario de dimensión n "como si fuese" K^n simplemente considerando vectores de coordenadas. La noción de isomorfismo nos permite formalizar esta idea.

Proposición 3.18 *Sea V un K-espacio vectorial de dimensión n. Entonces existe un isomorfismo $f : V \to K^n$.*

Demostración. Sea $B = \{v_1, \dots, v_n\}$ una base de V.

Dado $x \in V$, existe únicos $x_1, \dots, x_n \in K$ tales que $x = \sum_{i=1}^{n} x_i v_i$. Definimos

$$f : V \to K^n, \quad f(x) = (x_1, \dots, x_n).$$

Veamos que f es una transformación lineal:

Sean $x, y \in V$. Si $x = \sum\limits_{i=1}^{n} x_i v_i$ e $y = \sum\limits_{i=1}^{n} y_i v_i$, entonces $x + y = \sum\limits_{i=1}^{n} (x_i + y_i) v_i$. Luego

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

En forma análoga se prueba que $f(\lambda.x) = \lambda.f(x)$ para cada $\lambda \in K$ y cada $x \in V$.

Es claro que si $f(x) = 0$, entonces $x = 0$, de donde f es un monomorfismo.

Finalmente, dado $(x_1, \dots, x_n) \in K^n$, consideramos $x = \sum\limits_{i=1}^{n} x_i v_i \in V$. Se tiene que $f(x) = (x_1, \dots, x_n)$. Luego, f es un epimorfismo.

En consecuencia, f es un isomorfismo. $\square$

Ejemplo. Sea $K_3[X] = \{P \in K[X] \ / \ P = 0 \text{ o } \mathrm{gr}(P) \leq 3\}$, que es K-espacio vectorial de dimensión 4.

Un isomorfismo $f : K_3[X] \to K^4$ puede definirse como sigue:

Si $P = \sum\limits_{i=0}^{3} a_i X^i$, entonces $f(P) = (a_0, a_1, a_2, a_3)$, lo que corresponde a considerar en la demostración anterior la base $B = \{1, X, X^2, X^3\}$ de $K_3[X]$.

Observar que, teniendo en cuenta que la aplicación f definida en la demostración de la Proposición 3.18 es tomar coordenadas en la base B, esto nos permite trabajar con coordenadas en una base en el siguiente sentido:

i) $\{w_1, \dots, w_s\}$ es linealmente independiente en $V \iff \{f(w_1), \dots, f(w_s)\}$ es linealmente independiente en K^n.

ii) $\{w_1, \dots, w_r\}$ es un sistema de generadores de $V \iff \{f(w_1), \dots, f(w_r)\}$ es un sistema de generadores de K^n.

iii) $\{w_1, \dots, w_n\}$ es una base de $V \iff \{f(w_1), \dots, f(w_n)\}$ es una base de K^n.

Por ejemplo, para decidir si $\{X^2 - X + 1, \, X^2 - 3.X + 5, \, 2.X^2 + 2.X - 3\}$ es una base de $\mathbb{R}_2[X]$, bastará ver si $\{(1, -1, 1), \, (1, -3, 5), \, (2, 2, -3)\}$ es una base de $\mathbb{R}^3$ para lo que se puede usar el método de triangulación.

3.3 Teorema de la dimensión

El siguiente resultado relaciona las dimensiones del núcleo y de la imagen de una transformación lineal con la de su dominio.

Teorema 3.19 (Teorema de la dimensión para transformaciones lineales) *Sean V y W dos K-espacios vectoriales, V de dimensión finita, y sea $f : V \to W$ una transformación lineal. Entonces*

$$\dim V = \dim(\mathrm{Nu}(f)) + \dim(\mathrm{Im}(f)).$$

Demostración. Sean $n = \dim V$ y $r = \dim(\mathrm{Nu}(f))$.

Si $r = n$, entonces $f \equiv 0$ y $\dim(\mathrm{Im}(f)) = 0$. Por lo tanto, el teorema vale.

Si $r = 0$, entonces f es un monomorfismo. En este caso, si B es una base de V, entonces $f(B)$ es una base de $\mathrm{Im}(f)$. Luego $\dim(\mathrm{Im}(f)) = \dim V$ (ver Corolario 3.14), y el teorema vale.

Supongamos ahora que $0 < r < n$. Sea $\{v_1, \dots, v_r\}$ una base de $\mathrm{Nu}(f)$. Sean $v_{r+1}, \dots, v_n$ en V tales que $\{v_1, \dots, v_r, v_{r+1}, \dots, v_n\}$ es una base de V.

Veamos que entonces $\{f(v_{r+1}), \dots, f(v_n)\}$ es una base de $\mathrm{Im}(f)$, de donde se deduce inmediatamente el teorema:

- Puesto que $\{v_1, \dots, v_r, v_{r+1}, \dots, v_n\}$ es una base de V, se tiene que

$$\mathrm{Im}(f) = < f(v_1), \dots, f(v_r), f(v_{r+1}), \dots, f(v_n) > = < f(v_{r+1}), \dots, f(v_n) >,$$

  pues $f(v_i) = 0$ para $1 \le i \le r$.

- Sean $\alpha_{r+1}, \dots, \alpha_n \in K$ tales que $\sum\limits_{i=r+1}^{n} \alpha_i f(v_i) = 0$. Entonces $f\Big(\sum\limits_{i=r+1}^{n} \alpha_i v_i\Big) = 0$, es decir, $\sum\limits_{i=r+1}^{n} \alpha_i v_i \in \mathrm{Nu}(f)$. Como $\{v_1, \dots, v_r\}$ es una base de $\mathrm{Nu}(f)$, existen $\alpha_1, \dots, \alpha_r \in K$ tales que

$$\sum_{i=r+1}^{n} \alpha_i v_i = \sum_{i=1}^{r} \alpha_i v_i \iff \sum_{i=1}^{r} (-\alpha_i) v_i + \sum_{i=r+1}^{n} \alpha_i v_i = 0$$

  Como $\{v_1, \dots, v_n\}$ es un conjunto linealmente independiente, $\alpha_i = 0$ para cada $1 \le i \le n$. En particular, $\alpha_i = 0$ para $i = r+1, \dots, n$. Luego, $\{f(v_{r+1}), \dots, f(v_n)\}$ es un conjunto linealmente independiente. □

Como consecuencia de este resultado se prueba que si una transformación lineal entre dos espacios vectoriales de dimensión n es inyectiva (resp. suryectiva), entonces también es suryectiva (resp. inyectiva):

Corolario 3.20 *Sean V y W dos K-espacios vectoriales de dimensión n y sea $f : V \to W$ una transformación lineal. Son equivalentes:*

1. *f es un isomorfismo.*
2. *f es un monomorfismo.*
3. *f es un epimorfismo.*

Demostración.

$(1. \Rightarrow 2.)$ Por definición.

(2. $\Rightarrow$ 3.) Por el teorema de la dimensión, $n = \dim V = \dim(\mathrm{Nu}(f)) + \dim(\mathrm{Im}(f))$, y como f es un monomorfismo, $\dim(\mathrm{Nu}(f)) = 0$. Entonces $\dim(\mathrm{Im}(f)) = n = \dim W$, de donde $\mathrm{Im}(f) = W$.

(3. $\Rightarrow$ 1.) Por el teorema de la dimensión, y teniendo en cuenta que f es un epimorfismo, se tiene que $n = \dim V = \dim(\mathrm{Nu}(f)) + \dim(\mathrm{Im}(f)) = \dim(\mathrm{Nu}(f)) + n$. Esto implica que $\dim(\mathrm{Nu}(f)) = 0$, con lo cual, $\mathrm{Nu}(f) = \{0\}$ y f es un monomorfismo. Siendo epimorfismo y monomorfismo, resulta que f es un isomorfismo. □

A diferencia de lo que sucede para muchos de los resultados que hemos demostrado, en el corolario anterior la hipótesis de que los espacios vectoriales sean de dimensión *finita* es esencial. El resultado no vale para transformaciones lineales definidas en espacios de dimensión infinita:

Ejemplo. Sea $V = K[X]$.

1. Sea $f : K[X] \to K[X]$, $f(P) = P'$, que es una transformación lineal.
   - f es epimorfismo: Sea $Q = \sum\limits_{i=0}^{n} a_i X^i$. Entonces $f\Big(\sum\limits_{i=1}^{n+1} \frac{a_i}{i} X^i\Big) = Q$.
   - f no es monomorfismo: $f(1) = 0$, pero $1 \neq 0$.
2. Sea $g : K[X] \to K[X]$, $g(P) = X.P$.
   - g es monomorfismo: Si $f(P) = X.P = 0$, entonces $P = 0$.
   - g no es epimorfismo: $1 \notin \mathrm{Im}(f)$.

3.4 Proyectores

Definición 3.21 Sea V un K-espacio vectorial. Una transformación lineal $f : V \to V$ se llama un *proyector* si $f \circ f = f$.

Proposición 3.22 *Sea V un K-espacio vectorial, y sea $f : V \to V$ una transformación lineal. Entonces f es un proyector si y sólo si $f(x) = x$ para cada $x \in \mathrm{Im}(f)$.*

Demostración.

($\Rightarrow$) Supongamos que f es un proyector. Sea $x \in \mathrm{Im}(f)$. Entonces existe $v \in V$ tal que $x = f(v)$. Luego, $f(x) = f(f(v)) = f \circ f(v) = f(v) = x$.

($\Leftarrow$) Sea $v \in V$. Entonces $f(v) \in \mathrm{Im}(f)$ y, por hipótesis, $f(f(v)) = f(v)$, es decir, $f \circ f(v) = f(v)$. Como esto vale para cada $v \in V$, resulta que $f \circ f = f$. □

Proposición 3.23 *Sea V un K-espacio vectorial y sea $f : V \to V$ un proyector. Entonces* $\mathrm{Nu}(f) \oplus \mathrm{Im}(f) = V$.

Demostración. En primer lugar, veamos que $\mathrm{Nu}(f) \cap \mathrm{Im}(f) = \{0\}$: Sea $x \in \mathrm{Nu}(f) \cap \mathrm{Im}(f)$. Como $x \in \mathrm{Im}(f)$, por la proposición anterior, $f(x) = x$. Pero $x \in \mathrm{Nu}(f)$, de donde $f(x) = 0$. Luego, $x = 0$.

Veamos ahora que $\mathrm{Nu}(f) + \mathrm{Im}(f) = V$: Sea $x \in V$. Entonces $x = (x - f(x)) + f(x)$ y se tiene que $f(x - f(x)) = f(x) - f \circ f(x) = f(x) - f(x) = 0$, con lo que $x - f(x) \in \mathrm{Nu}(f)$ y $f(x) \in \mathrm{Im}(f)$. □

Proposición 3.24 *Sea V un K-espacio vectorial, y sean S y T subespacios de V tales que $S \oplus T = V$. Entonces existe un único proyector $f : V \to V$ tal que* $\mathrm{Nu}(f) = S$, $\mathrm{Im}(f) = T$.

Demostración. Como $V = S \oplus T$, para cada $x \in V$, existen únicos $s \in S$ y $t \in T$ tales que $x = s + t$. Entonces, si $f : V \to V$ es un proyector tal que $\mathrm{Nu}(f) = S$, $\mathrm{Im}(f) = T$, se tiene que $f(x) = f(s + t) = f(s) + f(t) = 0 + t = t$, donde la penúltima igualdad es consecuencia de que f es un proyector y $t \in \mathrm{Im}(f)$ (ver Proposición 3.22).

Consideremos entonces la función $f : V \to V$ definida por

$$f(x) = t \quad \text{si } x = s + t \text{ con } s \in S,\ t \in T.$$

Observamos que f es una transformación lineal:

- Si $x, x' \in V$ tales que $x = s + t$, $x' = s' + t'$, con $s, s' \in S$ y $t, t' \in T$, entonces $x + x' = (s + s') + (t + t')$ con $s + s' \in S$ y $t + t' \in T$ (puesto que S y T son subespacios de V) y, por lo tanto, $f(x + x') = t + t' = f(x) + f(x')$.
- Si $\lambda \in K$ y $x \in V$, $x = s + t$ con $s \in S$, $t \in T$, entonces $\lambda.x = (\lambda.s) + (\lambda.t)$ con $\lambda.s \in S$, $\lambda.t \in T$. Luego $f(\lambda.x) = \lambda.t = \lambda.f(x)$.

Además, f es un proyector: Si $x \in V$ y $x = s + t$ con $s \in S$, $t \in T$, entonces $f \circ f(x) = f(f(s + t)) = f(t) = f(0 + t) = f(x)$.

Es claro que $\mathrm{Im}(f) = T$. Veamos que $\mathrm{Nu}(f) = S$: Si $x \in \mathrm{Nu}(f)$ y $x = s + t$ con $s \in S$, $t \in T$, entonces $0 = f(x) = t$, con lo cual, $x = s \in S$. Por otro lado, si $s \in S$, entonces $s = s + 0$ con $s \in S$, $0 \in T$ y, por lo tanto, $f(s) = 0$.

Luego, la función f que hemos definido es el único proyector $f : V \to V$ con $\mathrm{Nu}(f) = S$, $\mathrm{Im}(f) = T$. □

3.5 Representación matricial

Uno de los primeros ejemplos de transformaciones lineales que hemos visto son aquéllas de la forma $f : K^n \to K^m$, $f(x) = A.x$ con $A \in K^{m\times n}$ (cuando quede claro por el contexto, suprimiremos el signo de ^{t}, escribiendo $A.x$ en lugar de $(A.x^t)^t$).

En esta sección veremos que toda transformación lineal entre espacios vectoriales de dimensión finita puede representarse de esta manera. Para esto, utilizaremos de manera fundamental el hecho de que fijada una base de un K-espacio vectorial V de dimensión finita n, se tiene un isomorfismo entre V y K^n tomando coordenadas en dicha base.

En esta sección todos los espacios vectoriales considerados serán de dimensión finita.

3.5.1 Matriz de una transformación lineal

Si V y W son K-espacios vectoriales de dimensión n y m respectivamente, una transformación lineal $f : V \to W$ queda unívocamente determinada por los n vectores de W que son los valores de f en una base cualquiera de V. Además, fijada una base de W, estos n vectores quedan determinados por medio de sus vectores de coordenadas en K^m. Se define entonces una matriz asociada a f que contiene toda esta información.

Definición 3.25 Sean V y W dos K-espacios vectoriales de dimensión finita. Sean $B_1 = \{v_1, \dots, v_n\}$ una base de V y $B_2 = \{w_1, \dots, w_m\}$ una base de W. Sea $f : V \to W$ una transformación lineal. Supongamos que $f(v_j) = \sum\limits_{i=1}^{m} \alpha_{ij} w_i$ $(1 \le j \le n)$. Se llama *matriz de f en las bases B_1, B_2*, y se nota $|f|_{B_1B_2}$, a la matriz en $K^{m\times n}$ definida por $(|f|_{B_1B_2})_{ij} = \alpha_{ij}$ para cada $1 \le i \le m$, $1 \le j \le n$.

Notación. Si $f : V \to V$ y $B_1 = B_2 = B$, notaremos $|f|_B = |f|_{BB}$.

Ejemplo. Sea $f : \mathbb{R}^3 \to \mathbb{R}^2$, $f(x_1, x_2, x_3) = (x_1 + 2x_2 - x_3, x_1 + 3x_2)$, y sean B_1 y B_2 las bases canónicas de $\mathbb{R}^3$ y $\mathbb{R}^2$ respectivamente. Se tiene que

$$f(1, 0, 0) = (1, 1), \quad f(0, 1, 0) = (2, 3), \quad f(0, 0, 1) = (-1, 0).$$

Entonces $|f|_{B_1B_2} = \begin{pmatrix} 1 & 2 & -1 \\ 1 & 3 & 0 \end{pmatrix}$.

Observación 3.26 Si consideramos la transformación lineal asociada a una matriz $A \in K^{n\times m}$, $f_A : K^m \to K^n$ definida por $f_A(x) = A.x$, entonces, a partir de la definición anterior, la matriz de f_A en las bases canónicas E y E' de K^m y K^n respectivamente resulta ser $|f_A|_{EE'} = A$.

Observación 3.27 Sea V un K-espacio vectorial de dimensión n, y sean B_1 y B_2 bases de V. Entonces $|id_V|_{B_1B_2} = C(B_1, B_2)$, la matriz de cambio de base de B_1 a B_2 (ver Definición 2.16).

Mediante el uso de las matrices introducidas en la Definición 3.25 y de vectores de coordenadas, toda transformación lineal puede representarse como la multiplicación por una matriz fija.

Proposición 3.28 *Sean V y W dos K-espacios vectoriales de dimensión finita, y sea $f : V \to W$ una transformación lineal. Si B_1 y B_2 son bases de V y W respectivamente, entonces para cada $x \in V$,*

$$|f|_{B_1B_2}.(x)_{B_1} = (f(x))_{B_2}.$$

Demostración. Supongamos que $B_1 = \{v_1, \ldots, v_n\}$.

Sea $x \in V$ y sea $(x)_{B_1} = (x_1, \ldots, x_n)$, es decir, $x = \sum_{i=1}^{n} x_i v_i$.

Para cada $1 \leq i \leq n$, sea C_i la i-ésima columna de $|f|_{B_1B_2}$. Por definición, $C_i = (f(v_i))_{B_2}$. Entonces

$$\begin{aligned} |f|_{B_1B_2}.\,(x)_{B_1} &= x_1.C_1 + \cdots + x_n.C_n = \\ &= x_1.(f(v_1))_{B_2} + \cdots + x_n.(f(v_n))_{B_2} = \\ &= \Big(\sum_{i=1}^{n} x_i f(v_i)\Big)_{B_2} = (f(x))_{B_2}. \end{aligned}$$ □

3.5.2 Matriz de la composición y cambios de bases

La composición de dos transformaciones lineales "se traduce" como la multiplicación de sus matrices.

Proposición 3.29 *Sean V, W y U tres K-espacios vectoriales de dimensión finita. Sean B_1, B_2 y B_3 bases de V, W y U respectivamente. Sean $f : V \to W$ y $g : W \to U$ transformaciones lineales. Entonces*

$$|g \circ f|_{B_1B_3} = |g|_{B_2B_3}.|f|_{B_1B_2}.$$

Demostración. Sean $n = \dim V$, $m = \dim W$ y $r = \dim U$. Entonces $|g|_{B_2B_3} \in K^{r\times m}$ y $|f|_{B_1B_2} \in K^{m\times n}$, con lo que $|g|_{B_2B_3}.|f|_{B_1B_2} \in K^{r\times n}$. Además $|g \circ f|_{B_1B_3} \in K^{r\times n}$.

Para cada $x \in V$ se tiene que

$$|g|_{B_2B_3}.|f|_{B_1B_2}.(x)_{B_1} = |g|_{B_2B_3}.(f(x))_{B_2} = g(f(x))_{B_3} = (g \circ f(x))_{B_3} = |g \circ f|_{B_1B_3}.(x)_{B_1}$$

Luego, $|g|_{B_2B_3}.|f|_{B_1B_2} = |g \circ f|_{B_1B_3}$ (ver Proposición 2.18). □

Corolario 3.30 *Sean V y W dos K-espacios vectoriales de dimensión finita, y sean B_1 y B_2 bases de V y W respectivamente. Sea $f : V \to W$ un isomorfismo. Entonces $|f^{-1}|_{B_2B_1} = (|f|_{B_1B_2})^{-1}$.*

Demostración. Se deduce inmediatamente aplicando la proposición anterior a $f^{-1} \circ f = id_V$ y $f \circ f^{-1} = id_W$. □

Concluimos esta sección estudiando cómo se puede obtener a partir de la matriz de una transformación lineal $f : V \to W$ en un par de bases B_1 y B_2 de V y W respectivamente, la matriz de la misma transformación lineal en cualquier otro par de bases B_1' y B_2' de dichos espacios.

Proposición 3.31 *Sean V y W dos K-espacios vectoriales de dimensión finita. Sean B_1, B_1' bases de V y B_2, B_2' bases de W. Entonces*

$$|f|_{B_1'B_2'} = C(B_2, B_2').|f|_{B_1B_2}.C(B_1', B_1).$$

Demostración. Se tiene que $f = id_W \circ f \circ id_V$. Aplicando dos veces el resultado dado en la Proposición 3.29 y el hecho que la matriz de la transformación lineal identidad en un par de bases coincide con la matriz de cambio de base entre las mismas, se obtiene

$$\begin{aligned} |f|_{B_1'B_2'} &= |id_W \circ f \circ id_V|_{B_1'B_2'} = |id_W \circ f|_{B_1B_2'}|id_V|_{B_1'B_1} = \\ &= |id_W|_{B_2B_2'}.|f|_{B_1B_2}.|id_V|_{B_1'B_1} = C(B_2, B_2').|f|_{B_1B_2}.C(B_1', B_1), \end{aligned}$$

que es lo que se quería probar. □

3.6 Rango de una matriz

Utilizando la relación entre matrices y transformaciones lineales introduciremos un nuevo invariante asociado a una matriz: su rango.

3.6.1 Rango columna y rango fila

Sean V y W dos K-espacios vectoriales tales que $\dim V = m$ y $\dim W = n$, y sean B_1 y B_2 bases de V y W respectivamente. Sea $f : V \to W$ una transformación lineal. Consideremos la matriz de f en las bases B_1 y B_2 dada por sus columnas:

$$|f|_{B_1B_2} = (C_1 \mid \ldots \mid C_m) \in K^{n\times m}.$$

Si $B_1 = \{v_1, \ldots, v_m\}$, entonces $\mathrm{Im}(f) = < f(v_1), \ldots, f(v_m) >$. Tomando coordenadas en la base B_2 se obtiene un subespacio $T \subseteq K^n$ dado por $T = < (f(v_1))_{B_2}, \ldots, (f(v_m))_{B_2} > = < C_1, \ldots, C_m >$. Como tomar coordenadas en una base es un isomorfismo, se tiene que

$$\dim(\mathrm{Im}(f)) = \dim < C_1, \ldots, C_m >.$$

Esto motiva la siguiente definición:

Definición 3.32 Sea $A \in K^{n\times m}$. Se llama *rango columna de* A, y se nota $\mathrm{rg}_C(A)$, a la dimensión del subespacio de K^n generado por las columnas de A, es decir, si $A = (C_1 \mid \cdots \mid C_m)$, entonces $\mathrm{rg}_C(A) = \dim < C_1, \ldots, C_m >$.

Mediante el cálculo del rango columna de una matriz A es posible obtener la dimensión del subespacio de soluciones del sistema lineal homogéneo asociado a A:

Observación 3.33 Sea $A \in K^{n\times m}$ y sea $S = \{x \in K^m \,/\, A.x = 0\}$. Entonces $\dim S = m - \mathrm{rg}_C(A)$.

En efecto, consideramos la transformación lineal asociada a A, $f_A : K^m \to K^n$ definida por $f_A(x) = A.x$. Entonces $A = |f_A|_{EE'}$ (donde E y E' son las bases canónicas de K^m y K^n respectivamente) y $S = \mathrm{Nu}(f_A)$. Entonces

$$\dim S = \dim(\mathrm{Nu}(f_A)) = m - \dim(\mathrm{Im}(f_A)) = m - \mathrm{rg}_C(A).$$

Ejemplo. Sea $A \in \mathbb{R}^{3\times 3}$, $A = \begin{pmatrix} 1 & -2 & 3 \\ -1 & 2 & 1 \\ 1 & -2 & 4 \end{pmatrix}$, y sea $S = \{x \in \mathbb{R}^3 \,/\, A.x = 0\}$. Entonces

$$\dim S = 3 - \mathrm{rg}_C(A) = 3 - \dim < (1,-1,1), (-2,2,-2), (3,1,4) > = 3 - 2 = 1.$$

Teniendo en cuenta el subespacio generado por las filas de una matriz en lugar del generado por sus columnas puede darse la siguiente definición de rango fila análoga a la de rango columna.

Definición 3.34 Sea $A \in K^{n\times m}$. Se define el *rango fila de* A, y se nota $\mathrm{rg}_F(A)$, como la dimensión del subespacio de K^m generado por las filas de A. Es decir, si $A = \begin{pmatrix} F_1 \\ \vdots \\ F_n \end{pmatrix}$, entonces $\mathrm{rg}_F(A) = \dim < F_1, \ldots, F_n >$.

Observación 3.35 Sea $A \in K^{n\times m}$. Entonces $\mathrm{rg}_F(A) = \mathrm{rg}_C(A^t)$.

Nuestro siguiente objetivo es mostrar que el rango fila y el rango columna de una matriz coinciden. Para hacer esto nos basaremos en la observación anterior. Primero mostraremos que el rango columna de una matriz A no cambia si se la multiplica a izquierda o derecha por matrices inversibles.

Lema 3.36 *Sea $A \in K^{n\times m}$. Sean $C \in GL(n,K)$ y $D \in GL(m,K)$. Entonces*

$$\mathrm{rg}_C(A) = \mathrm{rg}_C(C.A.D).$$

Demostración. Sea $f_A : K^m \to K^n$ la transformación lineal inducida por la multiplicación a izquierda por A. Si E y E' son las bases canónicas de K^m y K^n respectivamente, se tiene que $|f_A|_{EE'} = A$ y por lo tanto, $\mathrm{rg}_C(A) = \dim(\mathrm{Im}(f_A))$.

Por la Proposición 2.22, puesto que $D \in GL(m,K)$, existe una base B_1 de K^m tal que $D = C(B_1, E)$, y como $C \in GL(n,K)$, existe una base B_2 de K^n tal que $C = C(E', B_2)$.

Entonces

$$C.A.D = C(E', B_2).|f_A|_{EE'}.C(B_1, E) = |f_A|_{B_1B_2},$$

de donde $\mathrm{rg}_C(C.A.D) = \dim(\mathrm{Im}(f_A)) = \mathrm{rg}_C(A)$. □

Ahora veremos que multiplicando a A por matrices inversibles convenientes se puede obtener una matriz tal que su rango y el de su transpuesta son fáciles de comparar.

Lema 3.37 *Sea $A \in K^{n\times m} - \{0\}$. Entonces existen $k \in \mathbb{N}$, $1 \le k \le \min\{n,m\}$, y matrices $C \in GL(n,K)$ y $D \in GL(m,K)$ tales que*

$$(C.A.D)_{ij} = \begin{cases} 0 & si\ i \ne j \\ 1 & si\ i = j \le k \\ 0 & si\ i = j > k \end{cases}$$

Demostración. Consideremos la transformación lineal $f_A : K^m \to K^n$ inducida por la multiplicación a izquierda por A. Sea $\{v_1, \dots, v_s\}$ una base de $\mathrm{Nu}(f_A)$ y sean $w_1, \dots, w_{m-s} \in K^m$ tales que

$$B_1 = \{w_1, \dots, w_{m-s}, v_1, \dots, v_s\}$$

es una base de K^m (si $\mathrm{Nu}(f_A) = \{0\}$, $s = 0$ y se toma una base B_1 cualquiera de K^m).

Entonces $\{f_A(w_1), \dots, f_A(w_{m-s})\}$ es una base de $\mathrm{Im}(f_A)$ y puede extenderse a una base de K^n. Sean $z_1, \dots, z_{n-m+s} \in K^n$ tales que

$$B_2 = \{f_A(w_1), \dots, f_A(w_{m-s}), z_1, \dots, z_{n-m+s}\}$$

es una base de K^n.

Se tiene que

$$(|f_A|_{B_1B_2})_{ij} = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \leq m-s \\ 0 & \text{si } i = j > m-s \end{cases}$$

Observamos que

$$|f_A|_{B_1B_2} = C(E', B_2).|f_A|_{EE'}.C(B_1, E) = C.A.D,$$

donde $C = C(E', B_2) \in GL(n, K)$ y $D = C(B_1, E) \in GL(m, K)$. □

Proposición 3.38 *Sea* $A \in K^{n\times m}$. *Entonces* $\mathrm{rg}_C(A) = \mathrm{rg}_F(A)$.

Demostración. Es claro que el resultado vale si $A = 0$. Dada $A \in K^{n\times m} - \{0\}$, por el lema anterior, existen matrices $C \in GL(n, K)$, $D \in GL(m, K)$ y $k \in \mathbb{N}$, tales que

$$(C.A.D)_{ij} = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \leq k \\ 0 & \text{si } i = j > k \end{cases}$$

Por el Lema 3.36 se tiene que $\mathrm{rg}_C(A) = \mathrm{rg}_C(C.A.D)$, y es claro que $\mathrm{rg}_C(C.A.D) = k$.

Por otro lado, transponiendo se obtiene

$$(D^t.A^t.C^t)_{ij} = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \leq k \\ 0 & \text{si } i = j > k \end{cases}$$

con $D^t \in GL(m, K)$ y $C^t \in GL(n, K)$, de donde $\mathrm{rg}_C(A^t) = \mathrm{rg}_C(D^t.A^t.C^t) = k$.

En consecuencia

$$\mathrm{rg}_F(A) = \mathrm{rg}_C(A^t) = \mathrm{rg}_C(D^t.A^t.C^t) = k = \mathrm{rg}_C(A).$$ □

Definición 3.39 Sea $A \in K^{n\times m}$. Al número $\mathrm{rg}_C(A) = \mathrm{rg}_F(A)$ lo llamaremos el *rango de la matriz* A, y lo notaremos $\mathrm{rg}(A)$.

La Observación 3.33 puede ahora reescribirse utilizando la noción de rango de una matriz.

Proposición 3.40 *Sea* $A \in K^{n\times m}$ *y sea* $S = \{x \in K^m \,/\, A.x = 0\}$*. Entonces* $\dim S = m - \mathrm{rg}(A)$.

Esto significa que la dimensión del espacio de soluciones de un sistema lineal homogéneo es igual a la cantidad de incógnitas menos la cantidad de ecuaciones independientes.

3.6.2 Equivalencia de matrices

Definición 3.41 Sean $A, B \in K^{n\times m}$. Se dice que A es *equivalente* a B, y se nota $A \equiv B$, si existen matrices $C \in GL(n, K)$ y $D \in GL(m, K)$ tales que $A = C.B.D$.

Es inmediato verificar que $\equiv$ es una relación de equivalencia.

Como hemos visto en la sección anterior, si dos matrices son equivalentes entonces tienen el mismo rango. A continuación veremos que la recíproca de esta propiedad también es cierta. En consecuencia, el rango resulta ser un invariante que nos permite determinar fácilmente si dos matrices son equivalentes.

Proposición 3.42 *Sean* $A, B \in K^{n\times m}$*. Entonces* $A \equiv B \iff \mathrm{rg}(A) = \mathrm{rg}(B)$.

Demostración.

($\Rightarrow$) Es consecuencia del Lema 3.36.

($\Leftarrow$) Supongamos que $\mathrm{rg}(A) = \mathrm{rg}(B) = k$. Entonces existen matrices $C_1, C_2 \in GL(n, K)$ y $D_1, D_2 \in GL(m, K)$ tales que

$$(C_1.A.D_1)_{ij} = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \leq k \\ 0 & \text{si } i = j > k \end{cases} \quad \text{y} \quad (C_2.B.D_2)_{ij} = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \leq k \\ 0 & \text{si } i = j > k \end{cases}$$

En consecuencia, $C_1.A.D_1 = C_2.B.D_2$, de donde

$$A = (C_1^{-1}.C_2).B.(D_2.D_1^{-1}) = C.B.D$$

con $C = C_1^{-1}.C_2 \in GL(n, K)$ y $D = D_2.D_1^{-1} \in GL(m, K)$.

Luego, $A \equiv B$. □

Finalmente, la siguiente proposición caracteriza matrices equivalentes por medio de transformaciones lineales: dos matrices son equivalentes si y sólo si son las matrices de una misma transformación lineal en distintas bases.

Proposición 3.43 *Sean* $A, B \in K^{n\times m}$*. Entonces* $A \equiv B$ *si y sólo si existe una transformación lineal* $f : K^m \to K^n$ *y bases* B_1, B_1' *de* K^m *y* B_2, B_2' *de* K^n *tales que* $|f|_{B_1B_2} = A$ *y* $|f|_{B_1'B_2'} = B$.

Demostración. La validez de ($\Leftarrow$) se deduce de la proposición anterior, teniendo en cuenta que $\mathrm{rg}(A) = \dim(\mathrm{Im}(f)) = \mathrm{rg}(B)$.

Veamos que vale la otra implicación. Consideremos la transformación lineal $f : K^m \to K^n$ definida por $f(x) = B.x$. Entonces $B = |f|_{EE'}$, donde E y E' son las bases canónicas de K^m y K^n respectivamente.

Por definición, si $A \equiv B$, existen matrices inversibles C y D tales que $A = C.B.D$. Sea B_1 base de K^m tal que $D = C(B_1, E)$ y sea B_2 base de K^n tal que $C = C(E', B_2)$. Entonces

$$A = C.B.D = C(E', B_2)|f|_{EE'}C(B_1, E) = |f|_{B_1B_2}. \qquad \square$$

3.7 Espacios vectoriales de transformaciones lineales

Fijados dos K-espacios vectoriales V y W, tiene sentido considerar el conjunto de todas las transformaciones lineales de V en W. En esta sección, estudiaremos la estructura de estos conjuntos de transformaciones lineales.

Definición 3.44 *Sean V y W dos K-espacios vectoriales. Definimos*

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

Definimos ahora una operación en $\mathrm{Hom}_K(V, W)$ y una acción de K en $\mathrm{Hom}_K(V, W)$ que lo convierten en un K-espacio vectorial:

Suma. Dadas $f, g \in \mathrm{Hom}_K(V, W)$ se define $f + g$ como

$$(f + g)(x) = f(x) + g(x) \qquad \forall\, x \in V.$$

Veamos que $f + g \in \mathrm{Hom}_K(V, W)$, con lo cual $+$ resulta un operación en $\mathrm{Hom}_K(V, W)$:

- Es claro que $f + g : V \to W$.
- $f + g$ es una transformación lineal:

  Para cada $x, y \in V$, se tiene que

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

  Por otro lado, para cada $\mu \in K$ y cada $x \in V$ vale

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

Es fácil verificar que $(\mathrm{Hom}_K(V, W), +)$ es un grupo abeliano.

Producto por escalares. Dados $f \in \mathrm{Hom}_K(V, W)$ y $\lambda \in K$ se define $(\lambda \cdot f) : V \to W$ como

$$(\lambda \cdot f)(x) = \lambda \cdot f(x) \qquad \forall\, x \in V.$$

Veamos que $\lambda \cdot f \in \mathrm{Hom}_K(V, W)$, y por lo tanto, $\cdot$ es una acción de K en $\mathrm{Hom}_K(V, W)$:

- Por definición, $\lambda \cdot f : V \to W$.
- $\lambda \cdot f$ es una transformación lineal:
  Para todo par de elementos $x, y \in V$:

$$\begin{aligned}(\lambda \cdot f)(x+y) &= \lambda \cdot (f(x+y)) = \lambda \cdot (f(x)+f(y)) = \lambda \cdot f(x) + \lambda \cdot f(y) = \\ &= (\lambda \cdot f)(x) + (\lambda \cdot f)(y).\end{aligned}$$

  Para todo $\mu \in K$ y todo $x \in V$:

$$\begin{aligned}(\lambda \cdot f)(\mu \cdot x) &= \lambda \cdot (f(\mu \cdot x)) = \lambda \cdot (\mu \cdot f(x)) = (\lambda \cdot \mu) \cdot f(x) = \\ &= \mu \cdot (\lambda \cdot f(x)) = \mu \cdot \big((\lambda \cdot f)(x)\big).\end{aligned}$$

Además se cumplen las siguientes propiedades: Si $\lambda, \mu \in K$ y $f, g \in \mathrm{Hom}_K(V, W)$,

i) $\lambda \cdot (f+g) = \lambda \cdot f + \lambda \cdot g$

ii) $(\lambda + \mu) \cdot f = \lambda \cdot f + \mu \cdot f$

iii) $1 \cdot f = f$

iv) $(\lambda \cdot \mu) \cdot f = \lambda \cdot (\mu \cdot f)$

En consecuencia:

Proposición 3.45 $(\mathrm{Hom}_K(V, W), +, \cdot\,)$ *es un K-espacio vectorial.*

En el caso en que ambos V y W son K-espacios vectoriales de dimensión finita, $\dim V = n$ y $\dim W = m$, este K-espacio vectorial resulta ser isomorfo a $K^{m \times n}$.

Proposición 3.46 *Sean V y W dos K-espacios vectoriales de dimensión finita, con $\dim V = n$ y $\dim W = m$. Sean B y B' bases de V y W respectivamente. Entonces la función $T : \mathrm{Hom}_K(V, W) \to K^{m \times n}$ definida por $T(f) = |f|_{BB'}$ es un isomorfismo.*

Demostración. Supongamos que $B = \{v_1, \ldots, v_n\}$ y $B' = \{w_1, \ldots, w_m\}$.

- T es una transformación lineal:
  Sean $f, g \in \mathrm{Hom}_K(V, W)$. Por definición, $T(f+g) = |f+g|_{BB'}$. Observemos que la j-ésima columna de esta matriz es

$$\big((f+g)(v_j)\big)_{B'} = \big(f(v_j) + g(v_j)\big)_{B'} = (f(v_j))_{B'} + (g(v_j))_{B'},$$

  es decir, es la suma de las j-ésimas columnas de $|f|_{BB'}$ y $|g|_{BB'}$.
  Luego, $|f+g|_{BB'} = |f|_{BB'} + |g|_{BB'}$ o, equivalentemente, $T(f+g) = T(f) + T(g)$.
  En forma análoga se prueba que $T(\lambda \cdot f) = \lambda \cdot T(f)$.

- T es un isomorfismo:

  T es monomorfismo: Sea $f \in \mathrm{Hom}_K(V,W)$ tal que $T(f) = 0$, es decir, $|f|_{BB'} = 0$. Entonces, $\mathrm{Im}(f) = \{0\}$, de donde $f \equiv 0$.

  T es epimorfismo: Sea $A \in K^{m\times n}$. Consideramos $f_A : V \to W$ definida por $\big(f_A(x)\big)_{B'} = \big(A.(x)^t_B\big)^t$ para cada $x \in V$.

  Se tiene que $f_A \in \mathrm{Hom}_K(V,W)$ y $T(f_A) = |f_A|_{BB'} = A$. □

3.8 Ejercicios

Ejercicio 1. Determinar cuáles de las siguientes aplicaciones son lineales.

i) $f : \mathbb{R}^3 \to \mathbb{R}^2,\ f(x_1,x_2,x_3) = (x_2 - 3.x_1 + \sqrt{2}.x_3\ ,\ x_1 - \frac{1}{2}.x_2)$

ii) $f : \mathbb{R}^2 \to \mathbb{R}^3,\ f(x_1,x_2) = (x_1 - x_2\ ,\ 2.x_2\ ,\ 1 + x_1)$

iii) $f : \mathbb{R}^3 \to \mathbb{R}^3,\ f(x_1,x_2,x_3) = (2.x_1 - 7.x_3\ ,\ 0\ ,\ 3.x_2 + 2.x_3)$

iv) $f : \mathbb{R}^2 \to \mathbb{R}^2,\ f(x_1,x_2) = (x_1 + x_2\ ,\ |x_1|)$

v) $f : \mathbb{C} \to \mathbb{C}\,,\ f(z) = i.z$ (considerando a $\mathbb{C}$ como $\mathbb{R}$-espacio vectorial y como $\mathbb{C}$-espacio vectorial)

vi) $f : \mathbb{C} \to \mathbb{C}\,,\ f(z) = i.Im(z)$ (considerando a $\mathbb{C}$ como $\mathbb{R}$-espacio vectorial y como $\mathbb{C}$-espacio vectorial)

vii) $f : \mathbb{C} \to \mathbb{C}\,,\ f(z) = \overline{z}$ (considerando a $\mathbb{C}$ como $\mathbb{R}$-espacio vectorial y como $\mathbb{C}$-espacio vectorial)

viii) $f : \mathbb{R}^{2\times 2} \to \mathbb{R}\,,\ f\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}.a_{22} - a_{12}.a_{21}$

ix) $f : \mathbb{R}^{2\times 3} \to \mathbb{R}^3,\ f\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} = (3.a_{13} - a_{23}\ ,\ a_{11} + 2.a_{22} - a_{23}\ ,\ a_{22} - a_{12})$

x) $f : \mathbb{R}^{2\times 2} \to \mathbb{R}^{2\times 3},\ f\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} a_{22} & 0 & a_{12} + a_{21} \\ 0 & a_{11} & a_{22} - a_{11} \end{pmatrix}$

xi) $f : \mathbb{C}^{2\times 2} \to \mathbb{C}^{2\times 2},\ f\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} a_{11} & \overline{a_{12}} \\ \overline{a_{21}} & a_{22} \end{pmatrix}$ (considerando a $\mathbb{C}^{2\times 2}$ como $\mathbb{R}$-espacio vectorial y como $\mathbb{C}$-espacio vectorial)

Ejercicio 2. Interpretar geométricamente las siguientes aplicaciones lineales $f : \mathbb{R}^2 \to \mathbb{R}^2$.

i) $f(x,y) = (x,0)$

ii) $f(x,y) = (0,y)$

iii) $f(x, y) = (x, -y)$

iv) $f(x, y) = (\frac{1}{2}.(x + y), \frac{1}{2}.(x + y))$

v) $f(x, y) = (x.\cos t - y.\mathrm{sen}\, t\,,\ x.\mathrm{sen}\, t + y.\cos t)$

Ejercicio 3.

i) Encontrar una función $f : V \to V$ (para un K-espacio vectorial V conveniente) que cumpla $f(v + w) = f(v) + f(w)$ para cualquier par de vectores $v\,,\ w \in V$ pero que no sea una transformación lineal.

ii) Encontrar una función $f : V \to V$ (para un K-espacio vectorial V conveniente) que cumpla $f(k.v) = k.f(v)$ para cualquier escalar $k \in K$ y cualquier vector $v \in V$ pero que no sea una transformación lineal.

Ejercicio 4. Probar la linealidad de las siguientes aplicaciones:

i) $tr : K^{n\times n} \to K$

ii) $t : K^{n\times m} \to K^{m\times n},\ t(A) = A^t$

iii) $f : K^{n\times m} \to K^{r\times m},\ f(A) = B.A$ donde $B \in K^{r\times n}$

iv) $\delta : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R}),\ \delta(f) = f'$

v) $\epsilon_\alpha : K[X] \to K,\ \epsilon_\alpha(f) = f(\alpha)$ donde $\alpha \in K$

vi) $s : K^{\mathbb{N}} \to K^{\mathbb{N}},\ s(\{a_i\}_{i\in\mathbb{N}}) = (0, a_1, a_2, \ldots, a_n, \ldots)$

Ejercicio 5.

i) Probar que existe una única transformación lineal $f : \mathbb{R}^2 \to \mathbb{R}^2$ tal que $f(1, 1) = (-5, 3)$ y $f(-1, 1) = (5, 2)$. Para dicha f, determinar $f(5, 3)$ y $f(-1, 2)$.

ii) ¿Existirá una transformación lineal $f : \mathbb{R}^2 \to \mathbb{R}^2$ tal que $f(1, 1) = (2, 6)$, $f(-1, 1) = (2, 1)$ y $f(2, 7) = (5, 3)$?

iii) Sean $f,\ g : \mathbb{R}^3 \to \mathbb{R}^3$ transformaciones lineales tales que

$$f(1, 0, 1) = (1, 2, 1),\quad f(2, 1, 0) = (2, 1, 0),\quad f(-1, 0, 0) = (1, 2, 1),$$
$$g(1, 1, 1) = (1, 1, 0),\quad g(3, 2, 1) = (0, 0, 1),\quad g(2, 2, -1) = (3, -1, 2).$$

Determinar si $f = g$.

iv) Hallar todos los $a \in \mathbb{R}$ para los cuales exista una transformación lineal $f : \mathbb{R}^3 \to \mathbb{R}^3$ que satisfaga que $f(1, -1, 1) = (2, a, -1)\,,\ f(1, -1, 2) = (a^2, -1, 1)$ y $f(1, -1, -2) = (5, -1, -7)$.

v) Hallar una fórmula para todas las tranformaciones lineales $f : \mathbb{R}_3[X] \to \mathbb{R}^3$ que satisfacen $f(X^3 + 2X^2 - X + 4) = (6, 5, 3)$, $f(3X^2 + 2X - 5) = (0, 0, -3)$, $f(X^3 - 2X^2 + 3X - 2) = (0, -1, 1)$ y $f(2X^3 - 3X^2 + 7) = (6, 4, 7)$.

Ejercicio 6.

i) Calcular bases del núcleo y de la imagen para cada tranformación lineal del ejercicio 1. Decidir, en cada caso, si f es epimorfismo, monomorfismo o isomorfismo. En el caso que sea isomorfismo, calcular f^{-1}.

ii) Clasificar las transformaciones lineales tr , t , δ , ϵ_α y s del ejercicio 4 en epimorfismos, monomorfismos e isomorfismos.

Ejercicio 7. Sean $f : \mathbb{R}^3 \to \mathbb{R}^4$, $f(x_1, x_2, x_3) = (x_1 + x_2, x_1 + x_3, 0, 0)$ y $g : \mathbb{R}^4 \to \mathbb{R}^2$, $g(x_1, x_2, x_3, x_4) = (x_1 - x_2, 2x_1 - x_2)$. Calcular el núcleo y la imagen de f, de g y de $g \circ f$. Decidir si son monomorfismos, epimorfismos o isomorfismos.

Ejercicio 8. Sean $g : V \to V'$ y $f : V' \to V''$ transformaciones lineales. Probar:

i) $\mathrm{Nu}(g) \subseteq \mathrm{Nu}(f \circ g)$.

ii) Si $\mathrm{Nu}(f) \cap \mathrm{Im}(g) = \{0\}$, entonces $\mathrm{Nu}(g) = \mathrm{Nu}(f \circ g)$.

iii) $\mathrm{Im}(f \circ g) \subseteq \mathrm{Im}(f)$.

iv) Si $\mathrm{Im}(g) = V'$, entonces $\mathrm{Im}(f \circ g) = \mathrm{Im}(f)$.

Ejercicio 9.

i) Sean S, $T \subset \mathbb{R}^4$ los subespacios definidos por $S = \{(x_1, x_2, x_3, x_4)/\, x_1 + x_2 + x_3 = 0\}$ y $T = \{(x_1, x_2, x_3, x_4) \,/\, 2.x_1 + x_4 = 0\,,\, x_2 - x_3 = 0\}$.
¿Existirá algún isomorfismo $f : \mathbb{R}^4 \to \mathbb{R}^4$ tal que $f(S) = T$?

ii) ¿Existirá algún monomorfismo $f : \mathbb{R}^3 \to \mathbb{R}^2$?

iii) ¿Existirá algún epimorfismo $f : \mathbb{R}^2 \to \mathbb{R}^3$?

iv) Sean $v_1 = (1, 0, 1, 0)$, $v_2 = (1, 1, 1, 0)$ y $v_3 = (1, 1, 1, 1)$. ¿Existirá alguna transformación lineal $f : \mathbb{R}^2 \to \mathbb{R}^4$ tal que $\{v_1, v_2, v_3\} \subset \mathrm{Im}(f)$?

Ejercicio 10. Determinar si existe (y en caso afirmativo hallar) una transformación lineal $f : \mathbb{R}^3 \to \mathbb{R}^4$ que verifique $\mathrm{Im}(f) = S$ y $\mathrm{Nu}(f) = T$ en los siguientes casos:

i) $S = \{(x_1, x_2, x_3, x_4)/x_1 + x_2 - x_3 + 2.x_4 = 0\}$, $T = < (1, 2, 1) >$

ii) $S = \{(x_1, x_2, x_3, x_4)/x_1 + x_2 = 0,\, x_3 + x_4 = 0\}$, $T = < (1, -2, 1) >$

Ejercicio 11. En cada uno de los siguientes casos encontrar una transformación lineal $f : \mathbb{R}^3 \to \mathbb{R}^3$ que verifique lo pedido:

i) $(1, 1, 0) \in \text{Nu}(f)$ y $\dim(\text{Im}(f)) = 1$

ii) $\text{Nu}(f) \cap \text{Im}(f) = < (1, 1, 2) >$

iii) $f \neq 0$ y $\text{Nu}(f) \subseteq \text{Im}(f)$

iv) $f \neq 0$ y $f \circ f = 0$

v) $f \neq Id$ y $f \circ f = Id$

vi) $\text{Nu}(f) \neq \{0\}$, $\text{Im}(f) \neq \{0\}$ y $\text{Nu}(f) \cap \text{Im}(f) = \{0\}$

Ejercicio 12. En cada uno de los siguientes casos construir un proyector $f : \mathbb{R}^3 \to \mathbb{R}^3$ que cumpla:

i) $\text{Im}(f) = \{(x_1, x_2, x_3)/x_1 + x_2 + x_3 = 0\}$

ii) $\text{Nu}(f) = \{(x_1, x_2, x_3)/x_1 + x_2 + x_3 = 0\}$

iii) $\text{Nu}(f) = \{(x_1, x_2, x_3)/3.x_1 - x_3 = 0\}$ e $\text{Im}(f) = < (1, 1, 1) >$

Ejercicio 13. Sea V un K-espacio vectorial y sea $f : V \to V$ un proyector. Probar que $g = id_V - f$ es un proyector con $\text{Im}(g) = \text{Nu}(f)$ y $\text{Nu}(g) = \text{Im}(f)$.

Ejercicio 14. Sea V un K-espacio vectorial y sea $f : V \to V$ una transformación lineal. Se dice que f es *nilpotente* si $\exists\, s \in \mathbb{N}$ tal que $f^s = 0$.

i) Probar que si f es nilpotente, entonces f no es ni monomorfismo ni epimorfismo.

ii) Si V es de dimensión n probar que f es nilpotente $\iff$ $f^n = 0$.
(Sugerencia: considerar si las inclusiones $\text{Nu}(f^i) \subseteq \text{Nu}(f^{i+1})$ son estrictas o no).

iii) Sea $B = \{v_1, \ldots, v_n\}$ una base de V. Se define la transformación lineal $f : V \to V$ de la siguiente forma:
$$f(v_i) = \begin{cases} v_{i+1} & \text{si } 1 \leq i \leq n-1 \\ 0 & \text{si } i = n \end{cases}$$
Probar que $f^n = 0$ y $f^{n-1} \neq 0$.

iv) Si $V = \mathbb{R}^n$, para cada i, $2 \leq i \leq n$, construir una transformación lineal nilpotente $f : \mathbb{R}^n \to \mathbb{R}^n$ tal que $f^i = 0$ y $f^{i-1} \neq 0$.

Ejercicio 15. Sea $S = < (1, 1, 0, 1), (2, 1, 0, 1) > \subseteq \mathbb{R}^4$.

i) Hallar una transformación lineal $f : \mathbb{R}^4 \to \mathbb{R}^2$ tal que $\text{Nu}(f) = S$.

ii) Hallar ecuaciones para S (usar i)).

iii) Hallar un sistema de ecuaciones lineales cuyo conjunto de soluciones sea
$< (1,1,0,1),(2,1,0,1) > + (0,1,1,2)$.

Ejercicio 16.

i) Sea $S \subseteq K^n$ el conjunto de soluciones de un sistema lineal homogéneo. Encontrar una transformación lineal $f : K^n \to K^n$ tal que $\mathrm{Nu}(f) = S$.

ii) Sea $T \subseteq K^n$ el conjunto de soluciones de un sistema lineal no homogéneo. Encontrar una transformación lineal $f : K^n \to K^n$ y $x \in K^n$ tales que $T = f^{-1}(x)$.

Ejercicio 17. Sea $f : V \to V$ una tranformación lineal y sean B, B' bases de V. Calcular $|f|_{BB'}$ en cada uno de los siguientes casos:

i) $V = \mathbb{R}^3$, $f(x_1,x_2,x_3) = (3.x_1 - 2.x_2 + x_3, 5.x_1 + x_2 - x_3, x_1 + 3.x_2 + 4.x_3)$,
$B = B'$ la base canónica de $\mathbb{R}^3$

ii) $V = \mathbb{R}^3$, $f(x_1,x_2,x_3) = (3.x_1 - 2.x_2 + x_3, 5.x_1 + x_2 - x_3, x_1 + 3.x_2 + 4.x_3)$,
$B = \{(1,2,1),(-1,1,3),(2,1,1)\}$ y $B' = \{(1,1,0),(1,2,3),(-1,3,1)\}$

iii) $V = \mathbb{C}^2$, $f(x_1,x_2) = (2.x_1 - i.x_2, x_1 + x_2)$, $B = B'$ es la base canónica de $\mathbb{C}^2$ como $\mathbb{C}$-espacio vectorial.

iv) $V = \mathbb{C}^2$, $f(x_1,x_2) = (2.x_1 - i.x_2, x_1 + x_2)$, $B = B' = \{(1,0),(0,1),(i,0),(0,i)\}$ considerando a $\mathbb{C}^2$ como $\mathbb{R}$-espacio vectorial.

v) $V = \mathbb{R}_4[X]$, $f(P) = P'$, $B = B' = \{1, X, X^2, X^3, X^4\}$

vi) $V = \mathbb{R}_4[X]$, $f(P) = P'$, $B = B' = \{X^4, X^3, X^2, X, 1\}$

vii) $V = \mathbb{R}_4[X]$, $f(P) = P'$, $B = \{1, X, X^2, X^3, X^4\}$ y $B' = \{X^4, X^3, X^2, X, 1\}$

viii) $V = \mathbb{R}^{2\times 2}$, $f(A) = A^t$, $B = B'$ la base canónica de $\mathbb{R}^{2\times 2}$.

ix) V, f y $B = B'$ como en el ejercicio 14, iii)

Ejercicio 18. Sean $B = \{v_1, v_2, v_3\}$ una base de $\mathbb{R}^3$ y $B' = \{w_1, w_2, w_3, w_4\}$ una base de $\mathbb{R}^4$. Sea $f : \mathbb{R}^3 \to \mathbb{R}^4$ la transformación lineal tal que

$$|f|_{BB'} = \begin{pmatrix} 1 & -2 & 1 \\ -1 & 1 & -1 \\ 2 & 1 & 4 \\ 3 & -2 & 5 \end{pmatrix}$$

i) Hallar $f(3.v_1 + 2.v_2 - v_3)$. ¿Cuáles son sus coordenadas en la base B'?

ii) Hallar una base de $\text{Nu}(f)$ y una base de $\text{Im}(f)$.

iii) Describir el conjunto $f^{-1}(w_1 - 3.w_3 - w_4)$.

Ejercicio 19. Sea V un K-espacio vectorial y $B = \{v_1, v_2, v_3, v_4\}$ una base de V. Sea $f : V \to V$ la transformación lineal tal que

$$|f|_B = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

i) Calcular $|f^{-1}|_B$.

ii) Calcular $f^{-1}(v_1 - 2.v_2 + v_4)$.

Ejercicio 20. En cada uno de los siguientes casos, hallar una matriz $A \in \mathbb{R}^{n\times n}$ para un n adecuado que verifique:

i) $A \neq I_n$ y $A^3 = I_n$.

ii) $A \neq 0$; $A \neq I_n$ y $A^2 = A$.

Ejercicio 21. Sea V un K-espacio vectorial de dimensión finita y sea B una base de V.

i) Sea $tr : \text{Hom}(V, V) \to K$ la aplicación definida por $tr(f) = tr(|f|_B)$. Probar que $tr(f)$ no depende de la base B elegida.

$tr(f)$ se llama la *traza* del endomorfismo f.

ii) Probar que $tr : \text{Hom}(V, V) \to K$ es una transformación lineal.

Ejercicio 22. Sean $B = \{v_1, v_2, v_3\}$, $U = \{v_1 + v_3, v_1 + 2.v_2 + v_3, v_2 + v_3\}$ y $U' = \{w_1, w_2, w_3\}$ bases de $\mathbb{R}^3$, y sea E la base canónica de $\mathbb{R}^3$. Sea $f : \mathbb{R}^3 \to \mathbb{R}^3$ la transformación lineal tal que

$$|f|_{BE} = \begin{pmatrix} 1 & -1 & 3 \\ 2 & 1 & 1 \\ 3 & 2 & 1 \end{pmatrix} \qquad \text{y} \qquad |f|_{UU'} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

Determinar U'.

Ejercicio 23.

i) Sea $f : \mathbb{R}^4 \to \mathbb{R}^4$ la trasformación lineal definida por

$$f(x_1, x_2, x_3, x_4) = (0, x_1, x_1 + x_2, x_1 + x_2 + x_3)$$

y sea $v = (1, 0, 0, 0)$. Probar que $B = \{v, f(v), f^2(v), f^3(v)\}$ es una base de $\mathbb{R}^4$. Calcular $|f|_B$.

ii) Sea V un K-espacio vectorial de dimensión n y sea $f : V \to V$ una tranformación lineal tal que $f^n = 0$ y $f^{n-1} \neq 0$. Probar que existe una base B de V tal que

$$\big(|f|_B\big)_{ij} = \begin{cases} 1 & \text{si } i = j+1 \\ 0 & \text{si no} \end{cases}$$

(Sugerencia: elegir $v_1 \notin \mathrm{Nu}(f^{n-1})$).

Ejercicio 24. Sea V un K-espacio vectorial de dimensión n y sea $f : V \to V$ un proyector. Probar que existe una base B de V tal que

$$\big(|f|_B\big)_{ij} = \begin{cases} 1 & \text{si } i = j \ ; \ i \leq \dim(\mathrm{Im}(f)) \\ 0 & \text{si no} \end{cases}$$

Ejercicio 25. Sea $f : \mathbb{R}^5 \to \mathbb{R}^4$ definida por

$$f(x_1, x_2, x_3, x_4, x_5) = (2.x_1 - x_5, x_2 + 2.x_3, x_1 + x_4 + x_5, -x_1 + x_4 + x_5).$$

Encontrar bases B y B' de $\mathbb{R}^5$ y $\mathbb{R}^4$ respectivamente tales que $|f|_{BB'}$ sea una matriz diagonal.

Ejercicio 26. Sean V y W K-espacios vectoriales, $\dim V = n$ y $\dim W = m$, y $f : V \to W$ una transformación lineal tal que $\dim(\mathrm{Im}(f)) = s$. Probar que existen una base B de V y una base B' de W tal que

$$\big(|f|_{BB'}\big)_{ij} = \begin{cases} 1 & \text{si } i = j \ ; \ i \leq s \\ 0 & \text{si no} \end{cases}$$

Ejercicio 27. Sea $f : \mathbb{R}^3 \to \mathbb{R}^3$ definida por

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

i) Determinar bases B y B' de $\mathbb{R}^3$ tales que

$$|f|_{BB'} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

ii) Si A es la matriz de f en la base canónica, encontrar matrices C, $D \in GL(3, \mathbb{R})$ tales que

$$C.A.D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Ejercicio 28. Calcular el rango de las siguientes matrices:

i) $A = \begin{pmatrix} 2 & 0 & 3 & -1 \\ 1 & -2 & 1 & 0 \\ -1 & 1 & 0 & 1 \end{pmatrix}$ ii) $A = \begin{pmatrix} 0 & 5 & 3 \\ 1 & -1 & 2 \\ 2 & 3 & 1 \end{pmatrix}$

iii) $A = \begin{pmatrix} 3 & -1 & 0 & 1 & 2 \\ -1 & 0 & 4 & -1 & 0 \\ 3 & 1 & 1 & 0 & 1 \\ 2 & 0 & 0 & 3 & 1 \end{pmatrix}$ iv) $A = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{pmatrix}$

Ejercicio 29. Calcular el rango de $A \in \mathbb{R}^{3\times 3}$ para cada $k \in \mathbb{R}$ siendo

$$A = \begin{pmatrix} 1 & -k & -1 \\ -1 & 1 & k^2 \\ 1 & k & k-2 \end{pmatrix}.$$

Ejercicio 30. Sean $A \in K^{m\times n}$, $b \in K^m$. Se considera el sistema $A.x = b$ y sea $(A \mid b)$ su matriz ampliada. Probar que $A.x = b$ tiene solución $\iff \mathrm{rg}(A) = \mathrm{rg}(A \mid b)$.

Ejercicio 31. Sea $A \in K^{m\times n}$, $\mathrm{rg}(A) = s$ y sea $T = \{x \in K^{n\times r} / \, A.x = 0\}$. Calcular la dimensión de T.

Ejercicio 32. Sean $A \in K^{m\times n}$ y $B \in K^{n\times r}$. Probar que $\mathrm{rg}(A.B) \leq \mathrm{rg}(A)$ y $\mathrm{rg}(A.B) \leq \mathrm{rg}(B)$.

Ejercicio 33. Sean A, $D \in \mathbb{R}^{3\times 3}$,

$$A = \begin{pmatrix} 1 & 1 & -1 \\ 2 & -3 & 2 \\ 3 & -2 & 1 \end{pmatrix} \quad \text{y} \quad D = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ -1 & 0 & 1 \end{pmatrix}.$$

i) Determinar C_1, C_2, C_3 y $C_4 \in GL(3, \mathbb{R})$ tales que

$$C_1.A.C_2 = C_3.D.C_4 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

ii) Determinar $f \in \mathrm{Hom}(\mathbb{R}^3, \mathbb{R}^3)$ y bases B, B', B_1 y B_1' de $\mathbb{R}^3$ tales que

$$|f|_{BB'} = A \quad \text{y} \quad |f|_{B_1B_1'} = D$$

Ejercicio 34. Dadas A, $B \in \mathbb{R}^{n\times n}$, decidir si existen matrices P, $Q \in GL(n, \mathbb{R})$ tales que $A = P.B.Q$.

i) $n = 2$; $A = \begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix}$; $B = \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix}$

ii) $n = 2$; $A = \begin{pmatrix} 2 & 3 \\ 4 & 6 \end{pmatrix}$; $B = \begin{pmatrix} 5 & 8 \\ 1 & 2 \end{pmatrix}$

iii) $n = 3$; $A = \begin{pmatrix} 1 & 0 & 5 \\ 2 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix}$; $B = \begin{pmatrix} 3 & 8 & 5 \\ 2 & 2 & 0 \\ 0 & 7 & 0 \end{pmatrix}$

iv) $n = 3$; $A = \begin{pmatrix} 1 & 1 & 0 \\ 2 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix}$; $B = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 0 & 1 \\ 1 & 1 & 3 \end{pmatrix}$

Ejercicio 35. Sean $A, B \in K^{n\times n}$. Se dice que A es *semejante* a B (y se nota $A \sim B$) si existe $C \in GL(n, K)$ tal que $A = C.B.C^{-1}$.

i) Demostrar que $\sim$ es una relación de equivalencia en $K^{n\times n}$.

ii) Probar que dos matrices semejantes son equivalentes. ¿Vale la recíproca?

Ejercicio 36. Sean $A, C \in K^{n\times n}$. Probar que las siguientes afirmaciones son equivalentes:

i) $A \sim C$.

ii) $\exists\, f : K^n \to K^n$ tranformación lineal y bases B y B' de K^n tales que $|f|_B = A$ y $|f|_{B'} = C$

Ejercicio 37.

i) Sean $A, C \in K^{n\times n}$ tales que $A \sim C$. Probar que $tr(A) = tr(C)$.

ii) Sean $A, C \in \mathbb{R}^{3\times 3}$

$$A = \begin{pmatrix} 1 & -1 & 1 \\ 2 & 3 & -5 \\ 4 & 1 & 3 \end{pmatrix} \qquad \text{y} \qquad C = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}.$$

¿Existen $f \in \mathrm{Hom}(\mathbb{R}^3, \mathbb{R}^3)$ y bases B y B' de $\mathbb{R}^3$ tales que $|f|_B = A$ y $|f|_{B'} = C$?

Capítulo 4

Espacio dual

Una de las situaciones en donde se aplica la teoría de espacios vectoriales es cuando se trabaja con espacios de funciones, como vimos al final del capítulo anterior. En este capítulo estudiaremos algunas nociones básicas de ciertos espacios que de alguna forma le dan una estructura a las ecuaciones lineales.

4.1 El espacio dual de un espacio vectorial

Definición 4.1 Sea V un K-espacio vectorial. Se llama *espacio dual de* V, y se lo nota V^*, al K-espacio vectorial

$$V^* = \mathrm{Hom}_K(V, K) = \{f : V \to K \ / \ f \text{ es una transformación lineal}\}.$$

Según vimos en la Sección 3.7, si $\dim V = n$, dadas B una base de V y B' una base de K, se tiene que $\Gamma : \mathrm{Hom}_K(V, K) \to K^{1\times n}$ definida por $\Gamma(f) = |f|_{BB'}$, es un isomorfismo. En consecuencia,

$$\dim(V^*) = \dim(K^{1\times n}) = n = \dim V.$$

Ejemplo. Se consideran las transformaciones lineales $\delta_1, \delta_2, \delta_3$ de $\mathbb{R}^3$ a $\mathbb{R}$ definidas por $\delta_i(x_1, x_2, x_3) = x_i$ para $i = 1, 2, 3$.

$$\begin{aligned}
(\mathbb{R}^3)^* &= \{f : \mathbb{R}^3 \to \mathbb{R} \ / \ f \text{ es transformación lineal } \}\\
&= \{f : \mathbb{R}^3 \to \mathbb{R} \ / \ f(x_1, x_2, x_3) = ax_1 + bx_2 + cx_3 \text{ con } a, b, c \in \mathbb{R}\}\\
&= \{f : \mathbb{R}^3 \to \mathbb{R} \ / \ f = a\,\delta_1 + b\,\delta_2 + c\,\delta_3 \text{ con } a, b, c \in \mathbb{R}\}\\
&= < \delta_1, \delta_2, \delta_3 >
\end{aligned}$$

4.2 Base dual

Sea $E = \{e_1, e_2, e_3\}$ la base canónica de $\mathbb{R}^3$. Las funciones del ejemplo anterior cumplen la condición $\delta_i(e_j) = \delta_{ij}$ (donde δ_{ij} es la función delta de Kronecker, definida por $\delta_{ij} = 1$ si $i = j$ y $\delta_{ij} = 0$ si $i \neq j$). En lo que sigue, fijada una base de cualquier espacio vectorial V de dimensión finita, vamos a ver cómo encontrar una base de V^* que cumpla esta propiedad.

Proposición 4.2 *Sea V un K-espacio vectorial de dimensión n, y sea $B = \{v_1, \ldots, v_n\}$ una base de V. Existe una única base $B^* = \{\varphi_1, \ldots, \varphi_n\}$ de V^* tal que*

$$\varphi_i(v_j) = \begin{cases} 1 & \textit{si } i = j \\ 0 & \textit{si } i \neq j \end{cases}$$

B^* se llama la base dual de *B.*

Demostración. Para cada $1 \leq i \leq n$, sea $\varphi_i : V \to K$ la transformación lineal definida en la base $\{v_1, \ldots, v_n\}$ por:

$$\varphi_i(v_j) = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j. \end{cases}$$

Como $\dim(V^*) = n$, para ver que $\varphi_1, \ldots, \varphi_n \in V^*$ forman una base de V^*, basta verificar que son linealmente independientes. Sean $a_1, \ldots, a_n \in K$ tales que

$$a_1\varphi_1 + \cdots + a_n\varphi_n = 0.$$

Evaluando en v_i, resulta que $0 = a_1\varphi_1(v_i) + \cdots + a_i\varphi_i(v_i) + a_n\varphi_n(v_i) = a_i$ para $i = 1, \ldots, n$.

En consecuencia, $B^* = \{\varphi_1, \ldots, \varphi_n\}$ verifica las condiciones requeridas.

Supongamos que $\{\widetilde{\varphi}_1, \ldots, \widetilde{\varphi}_n\}$ sea otra base que satisface las mismas condiciones. Entonces, para cada $1 \leq i \leq n$, se tiene que

- $\widetilde{\varphi}_i(v_j) = 0 = \varphi_i(v_j)$ si $1 \leq j \leq n$, $j \neq i$,
- $\widetilde{\varphi}_i(v_i) = 1 = \varphi_i(v_i) = 1$,

es decir, $\widetilde{\varphi}_i$ y φ_i son dos transformaciones lineales que coinciden sobre una base. En consecuencia, $\widetilde{\varphi}_i = \varphi_i$ para cada $1 \leq i \leq n$. □

Ejemplos.

1. El ejemplo de la Sección 4.1 muestra que la base dual de la base canónica de $\mathbb{R}^3$ es $E^* = \{\delta_1, \delta_2, \delta_3\}$, donde $\delta_i(x_1, x_2, x_3) = x_i$ para $i = 1, 2, 3$.
2. Sea $V = \mathbb{R}^2$. Consideremos la base $B = \{(1, 1), (1, -1)\}$. Si $B^* = \{\varphi_1, \varphi_2\}$ es la base dual de B, entonces debe cumplir

$$\begin{cases} \varphi_1(1,1) &= 1 \\ \varphi_1(1,-1) &= 0 \end{cases} \quad \text{y} \quad \begin{cases} \varphi_2(1,1) &= 0 \\ \varphi_2(1,-1) &= 1 \end{cases}$$

Puesto que para cada $(x, y) \in \mathbb{R}^2$ vale

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

resulta que $\varphi_1(x, y) = \dfrac{x+y}{2}$ y $\varphi_2(x, y) = \dfrac{x-y}{2}$.

Si B es una base de un K-espacio vectorial V de dimensión finita y B^* es su base dual, es posible calcular fácilmente las coordenadas de un elemento de V en la base B utilizando la base B^*. Recíprocamente, utilizando la base B, es fácil obtener las coordenadas en la base B^* de un elemento de V^*.

Observación 4.3 Sea $B = \{v_1, \ldots, v_n\}$ una base de V y sea $B^* = \{\varphi_1, \ldots, \varphi_n\}$ su base dual.

- Dado $v \in V$, podemos escribir $v = \sum\limits_{i=1}^{n} \alpha_i v_i$, con $\alpha_i \in K$. Entonces, para cada $1 \le j \le n$,

$$\varphi_j(v) = \varphi_j\Big(\sum_{i=1}^{n} \alpha_i v_i\Big) = \sum_{i=1}^{n} \alpha_i \varphi_j(v_i) = \alpha_j.$$

  Luego, $(v)_B = (\varphi_1(v), \ldots, \varphi_n(v))$.

- Dada $\varphi \in V^*$, existen $\beta_i \in K$ tales que $\varphi = \sum\limits_{i=1}^{n} \beta_i \varphi_i$. Entonces, para cada $1 \le j \le n$,

$$\varphi(v_j) = \Big(\sum_{i=1}^{n} \beta_i \varphi_i\Big)(v_j) = \sum_{i=1}^{n} \beta_i \varphi_i(v_j) = \beta_j.$$

  Luego, $(\varphi)_{B^*} = (\varphi(v_1), \ldots, \varphi(v_n))$.

Ejemplo. Sean $B = \{(1, 1), (1, -1)\} \subset \mathbb{R}^2$ y $B^* = \{\varphi_1, \varphi_2\}$, con $\varphi_1(x, y) = \dfrac{x+y}{2}$ y $\varphi_2(x, y) = \dfrac{x-y}{2}$, su base dual (ver Ejemplo 2. en la página 96).

1. Hallar las coordenadas del vector $v = (5, 7) \in \mathbb{R}^2$ en la base B.

   Teniendo en cuenta que B^* es la base dual de B, por la observación anterior resulta que

$$(5, 7)_B = (\varphi_1(5, 7), \varphi_2(5, 7)) = (6, -1).$$

2. Hallar las coordenadas de $\varphi \in (\mathbb{R}^2)^*$ dada por $\varphi(x, y) = 5x + 3y$ en la base B^*.

   Por el segundo ítem de la observación anterior tenemos que:

$$(\varphi)_{B^*} = (\varphi(1, 1), \varphi(-1, 1)) = (8, -2).$$

Hemos visto que toda base de un K-espacio vectorial V de dimensión finita posee una base dual asociada. Recíprocamente, resulta que toda base de V^* es la base dual de una base de V:

Proposición 4.4 *Sea V un K-espacio vectorial de dimensión n, y sea V^* su espacio dual. Sea $B_1 = \{\varphi_1, \dots, \varphi_n\}$ una base de V^*. Entonces existe una única base $B = \{v_1, \dots, v_n\}$ de V que satisface $B^* = B_1$.*

Demostración.
Existencia. Sea $B_2 = \{w_1, \dots, w_n\}$ una base de V y sea B_2^* su base dual.

Para cada $1 \le i \le n$, se tiene que $(\varphi_i)_{B_2^*} = (\varphi_i(w_1), \dots, \varphi_i(w_n))$. Como $\{\varphi_1, \dots, \varphi_n\}$ es un conjunto linealmente independiente y tomar coordenadas es un isomorfismo, resulta que $\{(\varphi_1)_{B_2^*}, \dots, (\varphi_n)_{B_2^*}\} \subset K^n$ es linealmente independiente. En consecuencia, la matriz

$$M = \begin{pmatrix} \varphi_1(w_1) & \cdots & \varphi_1(w_n) \\ \varphi_2(w_1) & \cdots & \varphi_2(w_n) \\ \vdots & & \vdots \\ \varphi_n(w_1) & \cdots & \varphi_n(w_n) \end{pmatrix}$$

es inversible (sus filas son linealmente independientes).

Sea $A = (a_{ij})$ su inversa. Entonces $M.A = I_n$, de donde, para cada $1 \le i, j \le n$,

$$\delta_{ij} = (I_n)_{ij} = (M.A)_{ij} = \sum_{k=1}^{n} M_{ik} A_{kj} = \sum_{k=1}^{n} \varphi_i(w_k) a_{kj} = \varphi_i\Big(\sum_{k=1}^{n} a_{kj} w_k\Big).$$

Para cada $1 \le j \le n$, sea $v_j = \sum\limits_{k=1}^{n} a_{kj} w_k$.

Por construcción, es claro que vale $\varphi_i(v_j) = \delta_{ij}$ para cada $1 \le i, j \le n$. Queda por ver que $\{v_1, \dots, v_n\}$ es una base de V. Como $\dim V = \dim V^* = n$, basta ver que este conjunto es linealmente independiente. Ahora, si $\sum\limits_{j=1}^{n} \alpha_j v_j = 0$, para cada $1 \le i \le n$, se tiene que

$$0 = \varphi_i\Big(\sum_{j=1}^{n} \alpha_j v_j\Big) = \sum_{j=1}^{n} \alpha_j \varphi_i(v_j) = \alpha_i,$$

lo que prueba la independencia lineal.

Unicidad. Supongamos que $B = \{v_1, \dots, v_n\}$ y $B' = \{u_1, \dots, u_n\}$ son dos bases de V tales que $B^* = (B')^* = \{\varphi_1, \dots, \varphi_n\}$. Entonces, para cada $1 \le i \le n$,

$$(u_i)_B = (\varphi_1(u_i), \dots, \varphi_n(u_i)) = e_i = (v_i)_B,$$

de donde $u_i = v_i$. □

Ejemplo. Sea $V = \mathbb{R}_2[X]$. Sean $\varepsilon_0, \varepsilon_1, \varepsilon_2 \in (\mathbb{R}^2[X])^*$ definidas por $\varepsilon_0(P) = P(0)$, $\varepsilon_1(P) = P(1)$ y $\varepsilon_2(P) = P(2)$.

Veamos que $\{\varepsilon_0, \varepsilon_1, \varepsilon_2\}$ es una base de $(\mathbb{R}_2[X])^*$: Como $\dim(\mathbb{R}_2[X])^* = 3$, basta ver que son linealmente independientes. Supongamos que $\alpha_0\varepsilon_0 + \alpha_1\varepsilon_1 + \alpha_2\varepsilon_2 = 0$. Entonces para cada $P \in \mathbb{R}_2[X]$, se tiene que $(\alpha_0\varepsilon_0 + \alpha_1\varepsilon_1 + \alpha_2\varepsilon_2)(P) = 0$ o, equivalentemente,

$$\alpha_0\varepsilon_0(P) + \alpha_1\varepsilon_1(P) + \alpha_2\varepsilon_2(P) = 0.$$

Tomando $P = (X-1)(X-2)$ y evaluando, resulta que $\alpha_0 = 0$. De la misma manera, para $P = X(X-2)$ y $P = X(X-1)$ se obtiene $\alpha_1 = 0$ y $\alpha_2 = 0$ respectivamente. Luego, $\{\varepsilon_0, \varepsilon_1, \varepsilon_2\}$ es una base de $(\mathbb{R}_2[X])^*$.

Entonces, por la proposición anterior, existe una única base $B = \{P_0, P_1, P_2\}$ de $\mathbb{R}_2[X]$ tal que $B^* = \{\varepsilon_0, \varepsilon_1, \varepsilon_2\}$. Hallemos esta base: El polinomio P_0 debe satisfacer las condiciones

$$\left\{\begin{array}{rcl} P_0(0) & = & 1 \\ P_0(1) & = & 0 \\ P_0(2) & = & 0 \end{array}\right.$$

con lo que $P_0 = \dfrac{(X-1)(X-2)}{(-1)(-2)} = \dfrac{1}{2}(X-1)(X-2)$. Análogamente se calculan los otros dos elementos de la base: $P_1 = -X(X-2)$ y $P_2 = \dfrac{1}{2}X(X-1)$.

Si se quiere hallar un polinomio $P \in \mathbb{R}_2[X]$ tal que $P(0) = \alpha$, $P(1) = \beta$ y $P(2) = \gamma$, basta tener en cuenta que esto equivale a que

$$(P)_{\{P_0, P_1, P_2\}^*} = (\alpha, \beta, \gamma),$$

puesto que $\{P_0, P_1, P_2\}^* = \{\varepsilon_0, \varepsilon_1, \varepsilon_2\}$. Luego, $P = \alpha.P_0 + \beta.P_1 + \gamma.P_2$.

Nótese que este polinomio P hallado es el único polinomio de grado menor o igual que 2 que cumple lo pedido (polinomio interpolador de Lagrange).

4.3 Anulador de un subespacio

En lo que sigue vamos a relacionar los subespacios de V con ciertos subespacios de V^*. Esencialmente, dado un subespacio S de V consideraremos el conjunto de todas las ecuaciones lineales que se anulan en S y veremos que tiene una estructura de subespacio.

Definición 4.5 Sea V un K-espacio vectorial y sea S un subespacio de V. Se llama *anulador* de S al conjunto

$$S^\circ = \{f \in V^* \ / \ f(s) = 0 \ \forall\, s \in S\} = \{f \in V^* \ / \ S \subseteq \mathrm{Nu}(f)\}.$$

Observación 4.6 S° es un subespacio de V^*.

En efecto:

- Es claro que $0 \in S^\circ$.
- Si f, $g \in S^\circ$, entonces $f(s) = 0$ y $g(s) = 0$ para todo $s \in S$, de donde $(f+g)(s) = f(s) + g(s) = 0$ para todo $s \in S$. Luego, $f + g \in S^\circ$.
- Si $\lambda \in K$ y $f \in S^\circ$, entonces $(\lambda \cdot f)(s) = \lambda \cdot f(s) = \lambda \cdot 0 = 0$ para todo $s \in S$, puesto que $f(s) = 0$ para cada $s \in S$. Luego $\lambda \cdot f \in S^\circ$.

En el caso de un K-espacio vectorial de dimensión finita, existe una relación entre las dimensiones de un subespacio y su anulador.

Proposición 4.7 *Sea V un K-espacio vectorial de dimensión n y sea S un subespacio de V. Entonces* $\dim(S^\circ) = n - \dim S$.

Demostración. Sea $\{v_1, \ldots, v_r\}$ una base de S, y sean $v_{r+1}, \ldots, v_n \in V$ tales que $B = \{v_1, \ldots, v_r, v_{r+1}, \ldots, v_n\}$ es una base de V. Sea $B^* = \{\varphi_1, \ldots, \varphi_r, \varphi_{r+1}, \ldots, \varphi_n\} \subset V^*$ la base dual de B. Entonces, para cada $r+1 \le i \le n$, se tiene que $\varphi_i(v_1) = \ldots = \varphi_i(v_r) = 0$ y, por lo tanto, φ_i se anula sobre todo S. En consecuencia, $\{\varphi_{r+1}, \ldots, \varphi_n\} \subseteq S^\circ$.

Como $\{\varphi_{r+1}, \ldots, \varphi_n\}$ es parte de una base, es un conjunto linealmente independiente. Veamos que también es un sistema de generadores de S°:

Sea $g \in S^\circ$. Puesto que B^* es una base de V^*, existen $\alpha_1, \ldots, \alpha_n \in K$ tales que $g = \sum_{i=1}^{n} \alpha_i \varphi_i$. Por la Observación 4.3, para cada $1 \le i \le n$, se tiene que $\alpha_i = g(v_i)$. Además, como $g \in S^\circ$ y $\{v_1, \ldots, v_r\}$ es una base de S, $g(v_i) = 0$ para cada $1 \le i \le r$. En consecuencia, $\alpha_i = 0$ para cada $1 \le i \le r$, y por lo tanto $g \in < \varphi_{r+1}, \ldots, \varphi_n >$.

Luego, $\{\varphi_{r+1} \ldots, \varphi_n\}$ es una base de S°, de donde

$$\dim S + \dim S^\circ = n. \qquad \square$$

La demostración de la proposición anterior nos da una manera de calcular el anulador de un subespacio:

Ejemplo. Sea $S = < (1, 1, 1), (1, 2, 1) > \subseteq \mathbb{R}^3$. Hallar una base de S°.

Consideramos una base B de $\mathbb{R}^3$ que extienda a una base de S, por ejemplo

$$B = \{(1, 1, 1), (1, 2, 1), (1, 0, 0)\}.$$

Si $B^* = \{\varphi_1, \varphi_2, \varphi_3\}$ es la base dual de B, de la demostración anterior deducimos que $\{\varphi_3\}$ es una base de S°. A partir de las condiciones $\varphi_3(1, 1, 1) = 0$, $\varphi_3(1, 2, 1) = 0$, $\varphi_3(1, 0, 0) = 1$ obtenemos que $\varphi_3(x, y, z) = x - z$.

En la siguiente proposición veremos cómo recuperar un subespacio a partir de su anulador.

Proposición 4.8 *Sea V un K-espacio vectorial de dimensión n y sea S un subespacio de V. Entonces*

$$\{x \in V \,/\, f(x) = 0 \ \forall f \in S^\circ\} = S.$$

Demostración. Sea $T = \{x \in V \ / \ f(x) = 0 \ \forall f \in S^\circ\}$. Veamos que $T = S$.

($\supseteq$) Si $x \in S$, para cada $f \in S^\circ$ se tiene que $f(x) = 0$. Luego, $x \in T$.

($\subseteq$) Supongamos que existe $x \in T$ tal que $x \notin S$.

Sea $\{v_1, \dots, v_r\}$ una base de S. Entonces $\{v_1, \dots, v_r, x\}$ es linealmente independiente. Sean $v_{r+2}, \dots, v_n \in V$ tales que $B = \{v_1, \dots, v_r, x, v_{r+2}, \dots, v_n\}$ es una base de V. Si $B^* = \{\varphi_1, \dots, \varphi_r, \varphi_{r+1}, \dots, \varphi_n\}$ es la base dual de B, se tiene que $\varphi_{r+1}(v_1) = \ldots = \varphi_{r+1}(v_r) = 0$, de donde $\varphi_{r+1} \in S^\circ$.

Como $x \in T$, debe ser $\varphi_{r+1}(x) = 0$, lo que contradice que, por construcción, vale $\varphi_{r+1}(x) = 1$.

Luego, $T \subseteq S$. $\square$

Este resultado nos da otra forma de encontrar ecuaciones para un subespacio:

Ejemplo. Sea $S = < (1, 1, 1), (1, 2, 1) > \subseteq \mathbb{R}^3$. Hallar ecuaciones para S.

En el ejemplo anterior vimos que $S^\circ = < \varphi_3 > \subseteq (\mathbb{R}^3)^*$, donde $\varphi_3(x, y, z) = x - z$. Entonces, por la Proposición 4.8,

$$\begin{aligned} S &= \{(x, y, z) \in \mathbb{R}^3 \ / \ f(x, y, z) = 0 \ \forall f \in S^\circ\} \\ &= \{(x, y, z) \in \mathbb{R}^3 \ / \ \lambda.(x - z) = 0 \ \forall \lambda \in \mathbb{R}\} \\ &= \{(x, y, z) \in \mathbb{R}^3 \ / \ x - z = 0\}. \end{aligned}$$

Más en general, podemos enunciar el siguiente resultado.

Observación 4.9 Sea V un K-espacio vectorial de dimensión finita y sea S un subespacio de V. Sea $\{\varphi_1, \dots, \varphi_r\}$ una base de S°. Entonces

$$S = \{v \in V \ / \ \varphi_1(v) = 0 \wedge \ldots \wedge \varphi_r(v) = 0\} = \bigcap_{i=1}^{r} \mathrm{Nu}(\varphi_i).$$

Para terminar, vamos a ver cómo se comporta el anulador con la suma y la intersección de subespacios.

Proposición 4.10 *Sea V un K-espacio vectorial de dimensión n. Sean S y T subespacios de V. Entonces:*

1. $(S + T)^\circ = S^\circ \cap T^\circ$.
2. $(S \cap T)^\circ = S^\circ + T^\circ$.

Demostración.

1. Sea $f \in V^*$. Se tiene que:

$$\begin{aligned} f \in (S+T)^\circ &\iff f(s+t)=0 \quad \forall\, s \in S,\ \forall\, t \in T \\ &\iff f(s)=0 \quad \forall\, s \in S \quad \wedge \quad f(t)=0 \quad \forall\, t \in T \\ &\iff f \in S^\circ \cap T^\circ \end{aligned}$$

2. Sea $f \in S^\circ + T^\circ$. Entonces $f = f_S + f_T$, con $f_S \in S^\circ$ y $f_T \in T^\circ$. Para cada $x \in S \cap T$, se tiene que $f(x) = f_S(x) + f_T(x) = 0 + 0 = 0$. Luego, $f \in (S \cap T)^\circ$. Por lo tanto, $S^\circ + T^\circ \subseteq (S \cap T)^\circ$.

   Además, por el teorema de la dimensión para la suma de subespacios, la Proposición 4.7 y el hecho que $S^\circ \cap T^\circ = (S+T)^\circ$, resulta que

$$\begin{aligned} \dim(S^\circ + T^\circ) &= \dim S^\circ + \dim T^\circ - \dim(S^\circ \cap T^\circ) \\ &= \dim S^\circ + \dim T^\circ - \dim(S+T)^\circ \\ &= (n - \dim S) + (n - \dim T) - (n - \dim(S+T)) \\ &= n - (\dim S + \dim T - \dim(S+T)) \\ &= n - \dim(S \cap T) \\ &= \dim(S \cap T)^\circ. \end{aligned}$$

   En consecuencia, $(S \cap T)^\circ = S^\circ + T^\circ$. □

4.4 Ejercicios

Ejercicio 1. Sea $S \subseteq (\mathbb{R}^3)^*$ el subespacio $S = \{\varphi \in (\mathbb{R}^3)^* \,/\, \varphi(1,-1,2) = 0\}$. Encontrar una base de S.

Ejercicio 2. Dada la base B del K-espacio vectorial V, hallar su base dual en cada uno de los siguientes casos:

i) $V = \mathbb{R}^2,\ B = \{(1,-1),(2,0)\}$

ii) $V = \mathbb{R}^3,\ B = \{(1,-1,0),(0,1,0),(0,0,1)\}$

iii) $V = \mathbb{R}_3[X],\ B = \{-X+2,\ X-1,\ X^2 - 3.X + 2,\ X^3 - 3.X^2 + 2.X\}$

Ejercicio 3. Sea $B' = \{\varphi_1, \varphi_2, \varphi_3\}$ la base de $(\mathbb{R}^3)^*$ definida por

$$\varphi_1(x_1,x_2,x_3) = x_1 + x_2 \qquad \varphi_2(x_1,x_2,x_3) = x_1 + x_3 \qquad \varphi_3(x_1,x_2,x_3) = x_2 + x_3$$

Hallar la base B de $\mathbb{R}^3$ tal que $B' = B^*$.

Ejercicio 4. Sean f_1 , f_2 y $f_3 \in (\mathbb{R}_2[X])^*$ las siguientes formas lineales:

$$f_1(p) = \int_0^1 p(x)\,dx \qquad f_2(p) = \int_0^2 p(x)\,dx \qquad f_3(p) = \int_{-1}^0 p(x)\,dx$$

i) Probar que $\{f_1,\, f_2,\, f_3\}$ es una base de $(\mathbb{R}_2[X])^*$.

ii) Hallar una base B de $\mathbb{R}_2[X]$ tal que $B^* = \{f_1,\, f_2,\, f_3\}$.

Ejercicio 5. Sea V un K-espacio vectorial de dimensión n.

i) Sean $\varphi_1, \varphi_2 \in V^* - \{0\}$. Demostrar que $\mathrm{Nu}(\varphi_1) = \mathrm{Nu}(\varphi_2) \iff \{\varphi_1, \varphi_2\}$ es linealmente dependiente.

ii) Sean φ_i $(1 \le i \le r)$ formas lineales en V^* y sea $\varphi \in V^*$ tales que

$$\varphi_1(x) = \varphi_2(x) = \ldots = \varphi_r(x) = 0 \;\Rightarrow\; \varphi(x) = 0.$$

Probar que $\varphi \in <\varphi_1, \ldots, \varphi_r>$.

iii) Sean φ_i $(1 \le i \le n)$ formas lineales en V^*. Probar que

$$\{\, \varphi_1, \ldots, \varphi_n \,\} \text{ es base de } V^* \iff \bigcap_{i=1}^{n} \mathrm{Nu}(\varphi_i) = 0.$$

Ejercicio 6. Sea $\varphi \in (\mathbb{R}^3)^*$ definida por $\varphi(x_1, x_2, x_3) = 2.x_1 + 3.x_2 - x_3$ y sea $E^* = \{\delta_1, \delta_2, \delta_3\} \subseteq (\mathbb{R}^3)^*$ la base dual de la canónica.

i) Calcular las coordenadas de φ en E^*.

ii) Calcular las coordenadas de φ en la base $B^* = \{\delta_1 + \delta_2 + \delta_3, \delta_1 + \delta_2, \delta_1\}$.

iii) Sea $S \subseteq \mathbb{R}^3$ el subespacio $S = \{(x_1, x_2, x_3) \in \mathbb{R}^3 /\, 2.x_1 + 3.x_2 - x_3 = 0\}$ y sea $B \subset \mathbb{R}^3$ la base $B = \{(0, 0, 1), (0, 1, -1), (1, -1, 0)\}$. Encontrar una ecuación para S en la base B.

(Sugerencia: notar que B^* es la base dual de B y no hacer ninguna cuenta.)

Ejercicio 7. Sea $B \subset \mathbb{R}^2$ la base $B = \{(1, 1), (1, -1)\}$. Encontrar las coordenadas de la base dual de B en la base dual de la canónica.

Ejercicio 8. Sean B y B_1 las bases de $\mathbb{R}^3$ definidas por $B = \{(1, 1, 0), (1, 0, 1), (0, 1, 1)\}$ y $B_1 = \{(1, 1, -1), (1, -1, 1), (-1, 1, 1)\}$. Si $\varphi \in (\mathbb{R}^3)^*$ tiene coordenadas $(1, -3, 2)$ respecto de B^*, calcular sus coordenadas respecto de B_1^*.

Ejercicio 9. Hallar una base de $S^\circ \subseteq V^*$ en los siguientes casos:

i) $V = \mathbb{R}^3$ y $S = <(1, -1, 2), (2, 1, 3), (1, 5, 0)>$

ii) $V = \mathbb{R}^4$ y $S = <(1, 1, -1, 1), (2, -1, 3, 1)>$

iii) $V = \mathbb{R}^3$ y $S = \left\{ (x_1, x_2, x_3) \in \mathbb{R}^3 \,/\, \begin{cases} x_1 + x_3 = 0 \\ 2.x_1 - x_2 + x_3 = 0 \end{cases} \right\}$

iv) $V = \mathbb{R}^4$ y $S = \left\{ (x_1, x_2, x_3, x_4) \in \mathbb{R}^4 \,/\, \begin{cases} x_1 - x_2 + x_3 + x_4 = 0 \\ 2.x_1 + x_2 - 2.x_3 + 3.x_4 = 0 \\ 4.x_1 - x_2 + 5.x_4 = 0 \end{cases} \right\}$

Ejercicio 10. Sea $B = \begin{pmatrix} 2 & -2 \\ -1 & 1 \end{pmatrix} \in \mathbb{R}^{2\times 2}$ y sea $W = \{A \in \mathbb{R}^{2\times 2} / \, A.B = 0\}$. Sea $f \in W^\circ$ tal que $f(I_2) = 0$ y $f\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = 3$. Calcular $f(B)$.

Ejercicio 11. Para los siguientes subespacios S y T de V, determinar una base de $(S+T)^\circ$ y una base de $(S \cap T)^\circ$.

i) $V = \mathbb{R}^4$, $S = < (1, 1, -1, 1), (2, -1, 3, 1) >$, $T = < (2, -4, 8, 0), (-1, 1, 2, 3) >$

ii) $V = \mathbb{R}^4$, $S = \left\{ (x_1, x_2, x_3, x_4) \,/\, \begin{cases} x_1 - x_3 = 0 \\ x_1 + x_2 + x_4 = 0 \end{cases} \right\}$, $T = < (2, 1, 3, 1) >$

iii) $V = \mathbb{R}^3$, $S = \left\{ (x_1, x_2, x_3) \,/\, \begin{cases} x_1 - 2.x_2 + x_3 = 0 \\ 3.x_2 - 2.x_3 = 0 \end{cases} \right\}$,
$T = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 / \, 2.x_1 - x_2 = 0\}$

Ejercicio 12. Sea V un K-espacio vectorial de dimensión finita y sean S y T subespacios tales que $V = S \oplus T$. Probar que $V^* = S^\circ \oplus T^\circ$.

Ejercicio 13. Sea V un $\mathbb{Z}_p$-espacio vectorial de dimensión n. Probar que

$$\#\{S \subseteq V \text{ subespacio} \,/\, \dim(S) = 1\} = \#\{S \subseteq V \text{ subespacio} \,/\, \dim(S) = n - 1\}.$$

Calcular dicho número.

Ejercicio 14. Sea $tr : K^{n\times n} \to K$ la forma lineal traza y dado $a \in K^{n\times n}$ se define $f_a : K^{n\times n} \to K$ como $f_a(x) = tr(a.x)$.

i) Probar que $f_a \in (K^{n\times n})^* \ \forall a \in K^{n\times n}$.

ii) Probar que $f_a(x) = 0 \ \forall x \in K^{n\times n} \ \Rightarrow \ a = 0$.

iii) Se define $\gamma : K^{n\times n} \to (K^{n\times n})^*$ como $\gamma(a) = f_a$. Probar que γ es un isomorfismo.

iv) Sea $f : \mathbb{R}^{2\times 2} \to \mathbb{R}$ definida por:

$$f\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = 3.a_{11} - 2.a_{12} + 5.a_{22}.$$

Encontrar una matrix $a \in \mathbb{R}^{2\times 2}$ tal que $\gamma(a) = f$.

Ejercicio 15. Sea $\varphi \in (K^{n\times n})^*$ tal que $\varphi(a.b) = \varphi(b.a)\ \forall\, a,\, b \in K^{n\times n}$. Probar que $\exists\, \alpha \in K$ tal que $\varphi = \alpha.tr$. Deducir que si $\varphi(a.b) = \varphi(b.a)\ \forall\, a,\, b \in K^{n\times n}$ y $\varphi(I_n) = n$ entonces $\varphi = tr$.

Ejercicio 16. Sean $\alpha_0, \dots, \alpha_n \in K$, $\alpha_i \neq \alpha_j$ si $i \neq j$. Para cada i, $0 \leq i \leq n$, se define $\epsilon_{\alpha_i} : K_n[X] \to K$ como $\epsilon_{\alpha_i}(P) = P(\alpha_i)$.

i) Probar que $B_1 = \{\epsilon_{\alpha_0}, \dots, \epsilon_{\alpha_n}\}$ es una base de $(K_n[X])^*$.

ii) Sea $B = \{P_0, \dots, P_n\}$ la base de $K_n[X]$ tal que $B^* = B_1$. Probar que el polinomio

$$P = \sum_{i=0}^{n} \beta_i.P_i$$

es el único polinomio en $K[X]$ de grado menor o igual que n tal que, $\forall\, i$, $0 \leq i \leq n$, $P(\alpha_i) = \beta_i$. Este polinomio se llama el *polinomio interpolador de Lagrange.*

iii) Probar que existen números reales $a_0, \dots, a_n$ tales que, para todo $P \in \mathbb{R}_n[X]$,

$$\int_0^1 P(x)\, dx = \sum_{i=0}^{n} a_i.P(\alpha_i).$$

Hallar a_0, a_1 y a_2 en el caso en que $n = 2$, $\alpha_0 = 1$, $\alpha_1 = \frac{1}{2}$ y $\alpha_2 = 0$.

Ejercicio 17. Sean V y W K-espacios vectoriales de dimensión finita y sea $f : V \to W$ una transformación lineal. Se define la función $f^t : W^* \to V^*$ de la siguiente manera:

$$f^t(\varphi) = \varphi \circ f \qquad \forall\, \varphi \in W^*.$$

f^t se llama la función *transpuesta* de f.

i) Probar que f^t es una tranformación lineal.

ii) Probar que $(\mathrm{Im}(f))^\circ = \mathrm{Nu}(f^t)$ y que $\mathrm{Im}(f^t) = (\mathrm{Nu}(f))^\circ$.

iii) Sean $V = \mathbb{R}^2$ y $W = \mathbb{R}^3$ y sea $f(x_1, x_2) = (2.x_1 - x_2, 3x_1, x_1 - 2.x_2)$.
Si $B = \{(1,2), (1,3)\}$ y $B_1 = \{(1,1,1), (1,1,0), (1,0,0)\}$, calcular $|f|_{BB_1}$ y $|f^t|_{B_1^* B^*}$.

iv) Si B y B_1 son bases de V y W respectivamente, probar que

$$|f^t|_{B_1^* B^*} = (|f|_{BB_1})^t.$$

Ejercicio 18. Sea V un $\mathbb{C}$-espacio vectorial. Sean $f\,,\, g \in V^*$ tales que $f.\, g \in V^*$. Probar que $f = 0$ ó $g = 0$.

Ejercicio 19. Sea V un $\mathbb{C}$-espacio vectorial de dimensión finita. Sean $v_1, \dots, v_n \in V$ vectores no nulos. Probar que existe una forma lineal $f \in V^*$ tal que $f(v_i) \neq 0\ \forall\, i$, $1 \leq i \leq n$.

Capítulo 5

Determinantes

Los determinantes aparecieron originalmente al tratar de resolver sistemas de ecuaciones lineales. A continuación vamos a dar una definición precisa de determinante y a relacionarlo, entre otras cosas, con la inversibilidad de matrices.

En el caso de matrices en $K^{2\times 2}$, sabemos que $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ es inversible si y sólo si $ad - bc \neq 0$ (ver Ejercicio 11, Sección 2.5). Este escalar $ad - bc$ se llama el determinante de la matriz A. Para $n > 2$, el determinante de una matriz en $K^{n\times n}$ es también un escalar que se calcula a partir de los coeficientes de la matriz.

5.1 Definición y ejemplos básicos

Existen distintas definiciones de determinante. La definición que daremos introduce al determinante como una función particular de $K^{n\times n}$ en K.

5.1.1 Funciones multilineales alternadas

En esta sección daremos la definición y estudiaremos algunas propiedades básicas de funciones multilineales y alternadas, las que nos permitirán definir el determinante.

Notación. Dada una matriz $A \in K^{n\times n}$ cuyas columnas son $A_1, \dots, A_n$ escribiremos $A = (\, A_1 \mid A_2 \mid \dots \mid A_n \,)$.

Definición 5.1 Una función $f : K^{n\times n} \to K$ se dice *multilineal alternada (por columnas)* si para cada $1 \le i \le n$:

i) $f(A_1 \mid \dots \mid A_i + A_i' \mid \dots \mid A_n) = f(A_1 \mid \dots \mid A_i \mid \dots \mid A_n) + f(A_1 \mid \dots \mid A_i' \mid \dots \mid A_n)$.

ii) $f(\, A_1 \mid \dots \mid \lambda.A_i \mid \dots \mid A_n \,) = \lambda \cdot f(\, A_1 \mid \dots \mid A_i \mid \dots \mid A_n \,)$ para todo $\lambda \in K$.

iii) $f(\, A_1 \mid \ldots \mid \underbrace{A_i}_{col.\, i} \mid \ldots \mid \underbrace{A_i}_{col.\, j} \mid \ldots \mid A_n\,) = 0$ para cada $j \neq i$, $1 \leq j \leq n$ (es decir, si la matriz A tiene dos columnas iguales, $f(A) = 0$).

Notar que los ítems i) y ii) dicen que si $f : K^{n\times n} \to K$ es multilineal, para cada columna i, una vez fijados los valores de las columnas restantes, se obtiene una función lineal en la columna i.

Ejemplos.

1. $f : K^{1\times 1} \to K$ es multilineal alternada si y sólo si f es lineal.
2. $f : K^{2\times 2} \to K$ definida por $f\begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$ es multilineal alternada:

i) $f\begin{pmatrix} a+a' & b \\ c+c' & d \end{pmatrix} = (a+a')d - b(c+c') = ad - bc + a'd - bc' =$

$$= f\begin{pmatrix} a & b \\ c & d \end{pmatrix} + f\begin{pmatrix} a' & b \\ c' & d \end{pmatrix}.$$

Análogamente se prueba la aditividad en la segunda columna.

ii) $f\begin{pmatrix} \lambda a & b \\ \lambda c & d \end{pmatrix} = \lambda ad - b\lambda c = \lambda(ad - bc) = \lambda \cdot f\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ para todo $\lambda \in K$, y lo mismo vale en la segunda columna.

iii) $f\begin{pmatrix} a & a \\ b & b \end{pmatrix} = ab - ba = 0.$

En la siguiente proposición damos algunas de las propiedades básicas que verifica toda función multilineal alternada.

Proposición 5.2 *Sea $f : K^{n\times n} \to K$ multilineal alternada. Entonces*

i) $f(\, A_1 \mid \ldots \mid \underbrace{\vec{0}}_{col.\, i} \mid \ldots \mid A_n\,) = 0 \quad \forall\, 1 \leq i \leq n.$

ii) $f(\, A_1 \mid \ldots \mid A_i \mid \ldots \mid A_j \mid \ldots \mid A_n\,) = -f(\, A_1 \mid \ldots \mid \underbrace{A_j}_{col.\, i} \mid \ldots \mid \underbrace{A_i}_{col.\, j} \mid \ldots \mid A_n\,).$

iii) $f(\, A_1 \mid \ldots \mid \underbrace{A_i + \alpha A_j}_{col.\, i} \mid \ldots \mid A_j \mid \ldots \mid A_n\,) = f(\, A_1 \mid \ldots \mid A_i \mid \ldots \mid A_j \mid \ldots \mid A_n\,).$

iv) Si $A_i = \sum\limits_{\substack{j=1 \\ j\neq i}}^{n} \alpha_j A_j$, *entonces* $f(\, A_1 \mid \ldots \mid A_i \mid \ldots \mid A_n\,) = 0.$

Demostración.

i) $f(A_1 \mid \dots \mid \vec{0} \mid \dots \mid A_n) = f(A_1 \mid \dots \mid 0.\vec{0} \mid \dots \mid A_n) = 0 \cdot f(A_1 \mid \dots \mid \vec{0} \mid \dots \mid A_n) = 0.$

ii) Sea $A \in K^{n\times n}$. Consideremos la matriz que en las columnas i y j tiene la suma de las columnas i y j de A. Por la definición de función multilineal alternada se tiene que $f(\, A_1 \mid \dots \mid A_i + A_j \mid \dots \mid A_i + A_j \mid \dots \mid A_n\,) = 0$ y, por lo tanto, vale:

$$\begin{aligned}
0 &= f(\, A_1 \mid \dots \mid A_i \mid \dots \mid A_i + A_j \mid \dots \mid A_n\,) + \\
&\qquad\qquad + f(\, A_1 \mid \dots \mid A_j \mid \dots \mid A_i + A_j \mid \dots \mid A_n\,) \\
&= f(\, A_1 \mid \dots \mid A_i \mid \dots \mid A_i \mid \dots \mid A_n\,) + f(\, A_1 \mid \dots \mid A_i \mid \dots \mid A_j \mid \dots \mid A_n\,) + \\
&\quad + f(\, A_1 \mid \dots \mid A_j \mid \dots \mid A_i \mid \dots \mid A_n\,) + f(\, A_1 \mid \dots \mid A_j \mid \dots \mid A_j \mid \dots \mid A_n\,) \\
&= f(\, A_1 \mid \dots \mid A_i \mid \dots \mid A_j \mid \dots \mid A_n\,) + f(\, A_1 \mid \dots \mid A_j \mid \dots \mid A_i \mid \dots \mid A_n\,),
\end{aligned}$$

de donde

$$f(\, A_1 \mid \dots \mid A_i \mid \dots \mid A_j \mid \dots \mid A_n\,) = -f(\, A_1 \mid \dots \mid A_j \mid \dots \mid A_i \mid \dots \mid A_n\,).$$

iii) Se tiene que

$$\begin{aligned}
&f(\, A_1 \mid \dots \mid A_i + \alpha A_j \mid \dots \mid A_j \mid \dots \mid A_n\,) = \\
&= f(\, A_1 \mid \dots \mid A_i \mid \dots \mid A_j \mid \dots \mid A_n\,) + \alpha f(\, A_1 \mid \dots \mid A_j \mid \dots \mid A_j \mid \dots \mid A_n\,) = \\
&= f(\, A_1 \mid \dots \mid A_i \mid \dots \mid A_j \mid \dots \mid A_n\,),
\end{aligned}$$

donde la primera igualdad es consecuencia de i) y ii) de la Definición 5.1, y la tercera se deduce de iii) de dicha definición.

iv) Si $A_i = \sum\limits_{\substack{j=1\\ j\neq i}}^{n} \alpha_j A_j$, entonces

$$\begin{aligned}
f(\, A_1 \mid \dots \mid A_i \mid \dots \mid A_n\,) &= f(\, A_1 \mid \dots \mid \overbrace{\sum_{\substack{j=1\\ j\neq i}}^{n} \alpha_j A_j}^{col.\,i} \mid \dots \mid A_n\,) \\
&= \sum_{\substack{j=1\\ j\neq i}}^{n} \alpha_j f(\, A_1 \mid \dots \mid \underbrace{A_j}_{col.\, i} \mid \dots \mid A_n\,) = 0.
\end{aligned}$$

□

Podemos ahora caracterizar fácilmente *todas* las funciones multilineales alternadas de $K^{2\times 2}$ en K y ver cómo puede definirse el determinante en este caso particular.

Ejemplo. Hallar todas las funciones multilineales alternadas $f : K^{2\times 2} \to K$.

Si $f : K^{2\times 2} \to K$ es multilineal alternada, entonces:

$$\begin{aligned}
f\begin{pmatrix} a & b \\ c & d \end{pmatrix} &= f\begin{pmatrix} a & b \\ 0 & d \end{pmatrix} + f\begin{pmatrix} 0 & b \\ c & d \end{pmatrix} && \text{(Def. 5.1 i) en primera columna)} \\
&= f\begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix} + f\begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix} + f\begin{pmatrix} 0 & b \\ c & 0 \end{pmatrix} + f\begin{pmatrix} 0 & 0 \\ c & d \end{pmatrix} \\
& && \text{(Def. 5.1 i) en segunda columna)} \\
&= f\begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix} + f\begin{pmatrix} 0 & b \\ c & 0 \end{pmatrix} && \text{(Prop. 5.2 i))} \\
&= ad.\, f\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + cb.\, f\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} && \text{(Def. 5.1 ii))} \\
&= (ad - cb).\, f\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} && \text{(Prop. 5.2 ii))}
\end{aligned}$$

Por otro lado, toda función de este tipo es multilineal alternada (comparar con el Ejemplo 2 de la página 108). Luego, todas las funciones multilineales alternadas $f : K^{2\times 2} \to K$ son de la forma $f\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \alpha\,(ad - bc)$ con $\alpha \in K$.

Vemos entonces que una forma posible de definir una función de $K^{2\times 2}$ en K que coincide con lo que conocemos como determinante en dicha situación es como la única función f multilineal alternada tal que $f(I_2) = 1$.

En la próxima sección, generalizaremos lo que hemos visto en el ejemplo anterior para matrices en $K^{2\times 2}$ a matrices en $K^{n\times n}$ para $n \in \mathbb{N}$ arbitrario.

5.1.2 Existencia y unicidad del determinante

El siguiente lema que relaciona una función multilineal alternada definida en $K^{(n+1)\times(n+1)}$ con otras definidas en $K^{n\times n}$ será la base para un argumento recursivo que nos permitirá probar la unicidad de la función determinante.

Lema 5.3 *Sea $f : K^{(n+1)\times(n+1)} \to K$ una función multilineal alternada tal que $f(I_{n+1}) = \alpha$. Sea i con $1 \le i \le n+1$. Se define $f_i : K^{n\times n} \to K$ como*

$$f_i(A) = f\begin{pmatrix}
a_{11} & \dots & a_{1\,i-1} & 0 & a_{1\,i} & \dots & a_{1n} \\
\vdots & & \vdots & \vdots & \vdots & & \vdots \\
a_{i-1\,1} & \dots & a_{i-1\,i-1} & 0 & a_{i-1\,i} & \dots & a_{i-1\,n} \\
0 & \dots & 0 & 1 & 0 & \dots & 0 \\
a_{i1} & \dots & a_{i\,i-1} & 0 & a_{i\,i} & \dots & a_{i\,n} \\
\vdots & & \vdots & \vdots & \vdots & & \vdots \\
a_{n1} & \dots & a_{n\,i-1} & 0 & a_{n\,i} & \dots & a_{nn}
\end{pmatrix} \qquad \textit{si } A = (a_{jl})_{1\le j,\, l\le n}.$$

Entonces f_i es una función multilineal alternada y $f_i(I_n) = \alpha$.

Demostración. De la definición de f_i y del hecho que f es multilineal alternada, se deduce fácilmente que f_i es multilineal alternada.

Además, $f_i(I_n) = f(I_{n+1}) = \alpha$. $\square$

Veamos cómo puede utilizarse este resultado para hallar una función multilineal alternada de $K^{3\times 3}$ en K con un determinado valor sobre I_3 conociendo las funciones multilineales alternadas de $K^{2\times 2}$ en K:

Ejemplo. Hallar $f : K^{3\times 3} \to K$ multilineal alternada tal que $f(I_3) = 1$.

Supongamos que f satisface lo pedido. Entonces

$$\begin{aligned} f\begin{pmatrix} a & b & c \\ d & e & k \\ g & h & i \end{pmatrix} &= a.\, f\begin{pmatrix} 1 & b & c \\ 0 & e & k \\ 0 & h & i \end{pmatrix} + d.\, f\begin{pmatrix} 0 & b & c \\ 1 & e & k \\ 0 & h & i \end{pmatrix} + g.\, f\begin{pmatrix} 0 & b & c \\ 0 & e & k \\ 1 & h & i \end{pmatrix} \\ &= a.\, f\begin{pmatrix} 1 & 0 & 0 \\ 0 & e & k \\ 0 & h & i \end{pmatrix} + d.\, f\begin{pmatrix} 0 & b & c \\ 1 & 0 & 0 \\ 0 & h & i \end{pmatrix} + g.\, f\begin{pmatrix} 0 & b & c \\ 0 & e & k \\ 1 & 0 & 0 \end{pmatrix} \\ &= a.\, f\begin{pmatrix} 1 & 0 & 0 \\ 0 & e & k \\ 0 & h & i \end{pmatrix} - d.\, f\begin{pmatrix} b & 0 & c \\ 0 & 1 & 0 \\ h & 0 & i \end{pmatrix} + g.\, f\begin{pmatrix} b & c & 0 \\ e & k & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= a.\, f_1\begin{pmatrix} e & k \\ h & i \end{pmatrix} - d.\, f_2\begin{pmatrix} b & c \\ h & i \end{pmatrix} + g.\, f_3\begin{pmatrix} b & c \\ e & k \end{pmatrix}. \end{aligned}$$

Por el lema anterior, $f_1, f_2, f_3 : K^{2\times 2} \to K$ son funciones multilineales alternadas que en la identidad I_2 valen 1. Pero ya vimos, en un ejemplo anterior, que hay una única función que cumple esto, a saber, $f_0 : K^{2\times 2} \to K$, definida por $f_0\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \alpha\,\delta - \beta\,\gamma$. En consecuencia

$$f\begin{pmatrix} a & b & c \\ d & e & k \\ g & h & i \end{pmatrix} = a.(ei - hk) - d.(bi - hc) + g.(bk - ce).$$

Además, esta f cumple lo pedido con lo cual resulta que es la única función multilineal alternada tal que $f(I_3) = 1$.

El siguiente teorema nos permite definir la noción de determinante en general.

Teorema 5.4 *Sea $\alpha \in K$. Para cada $n \in \mathbb{N}$, existe una única función multilineal alternada $f : K^{n\times n} \to K$ tal que $f(I_n) = \alpha$.*

Definición 5.5 La única función multilineal alternada $f : K^{n\times n} \to K$ tal que $f(I_n) = 1$ se llama el *determinante de orden n*.

Demostración. Dada $A \in K^{(n+1)\times(n+1)}$, notaremos $A(i|j) \in K^{n\times n}$ a la matriz que se obtiene al suprimir la fila i y la columna j de A.

Existencia. Por inducción en n.

Para $n = 1$, definimos $f : K^{1\times 1} \to K$, $f(x) = \alpha.x$, que es multilineal alternada y cumple $f(1) = \alpha$.

Supongamos que existe $g : K^{n\times n} \to K$ multilineal alternada tal que $g(I_n) = \alpha$. Definimos $f : K^{(n+1)\times(n+1)} \to K$ como

$$f(A) = \sum_{i=1}^{n+1} (-1)^{i+1} a_{1i}.\, g(A(1|i)) \qquad \text{si } A = (a_{j\ell})_{1\le j,\ell\le n+1}.$$

Veamos que f es multilineal alternada y que $f(I_n) = \alpha$.

i) Sean $A = (\, A_1 \mid \dots \mid A_k \mid \dots \mid A_{n+1} \,)$, $A' = (\, A_1 \mid \dots \mid A'_k \mid \dots \mid A_{n+1} \,)$ y $\widetilde{A} = (\, A_1 \mid \dots \mid A_k + A'_k \mid \dots \mid A_{n+1} \,)$. Entonces

$$\begin{aligned} f(\widetilde{A}) &= \sum_{\substack{i=1\\ i\ne k}}^{n+1} (-1)^{i+1}\, a_{1i}\, g(\widetilde{A}(1|i)) + (-1)^{k+1}\, (a_{1k} + a'_{1k})\, g(\widetilde{A}(1|k)) \\ &= \sum_{\substack{i=1\\ i\ne k}}^{n+1} (-1)^{i+1} a_{1i}\, g(A(1|i)) + \sum_{\substack{i=1\\ i\ne k}}^{n+1} (-1)^{i+1} a_{1i}\, g(A'(1|i)) + \\ &\quad + (-1)^{k+1}\, a_{1k}\, g(A(1|k)) + (-1)^{k+1}\, a'_{1k}\, g(A'(1|k)) \\ &= f(A) + f(A'). \end{aligned}$$

ii) Sea $A = (\, A_1 \mid \dots \mid A_k \mid \dots \mid A_{n+1} \,)$ y sea $\widetilde{A} = (\, A_1 \mid \dots \mid \lambda A_k \mid \dots \mid A_{n+1} \,)$. Entonces

$$\begin{aligned} f(\widetilde{A}) &= \sum_{\substack{i=1\\ i\ne k}}^{n+1} (-1)^{i+1}\, a_{1i}\, g(\widetilde{A}(1|i)) + (-1)^{k+1}\, \lambda\, a_{1k}\, g(\widetilde{A}(1|k)) \\ &= \sum_{\substack{i=1\\ i\ne k}}^{n+1} (-1)^{i+1}\, a_{1i}\, \lambda\, g(A(1|i)) + (-1)^{k+1}\, \lambda\, a_{1k}\, g(A(1|k)) \\ &= \lambda \cdot f(A). \end{aligned}$$

iii) Sea $A = (\, A_1 \mid \dots \mid A_j \mid \dots \mid A_j \mid \dots \mid A_{n+1} \,)$, donde la k-ésima columna coincide con la j-ésima ($k > j$). Entonces

$$f(A) = \sum_{\substack{i=1\\ i\ne k,j}}^{n+1} (-1)^{i+1}\, a_{1i}\, g(A(1|i)) + (-1)^{j+1}\, a_{1j}\, g(A(1|j)) + (-1)^{k+1}\, a_{1j}\, g(A(1|k)).$$

Observamos que para cada $i \neq j, k$, la matriz $A(1|i)$ tiene dos columnas iguales, con lo que $g(A(1|i)) = 0$. Además $A(1|j)$ y $A(1|k)$ sólo difieren en la posición de una columna: la j-ésima columna de $A(1|k)$ es la $(k-1)$-ésima columna de $A(1|j)$. En consecuencia, $A(1|j)$ puede obtenerse a partir de $A(1|k)$ mediante $k-1-j$ intercambios de columnas y por lo tanto, $g(A(1|k)) = (-1)^{k-1-j} g(A(1|j))$. Luego

$$\begin{aligned} f(A) &= (-1)^{j+1}\, a_{1j}\, g(A(1|j)) + (-1)^{k+1}\, a_{1j}\, (-1)^{k-1-j} g(A(1|j)) \\ &= ((-1)^{j+1} + (-1)^{2k-j})\, a_{1j}\, g(A(1|j)) = 0. \end{aligned}$$

Esto prueba que f es multilineal alternada.

Calculemos $f(I_{n+1})$. Se tiene que

$$f(I_{n+1}) = \sum_{i=1}^{n+1} (-1)^{i+1}\, (I_{n+1})_{1i}.\, g(I_{n+1}(1|i)) = (-1)^2.1.\, g(I_n) = \alpha.$$

Unicidad. Por inducción en n.

Para $n = 1$, basta tener en cuenta que $f : K^{1\times 1} \to K$ es multilineal alternada si y sólo si es lineal. Por la condición $f(1) = \alpha$, resulta que la única función con las condiciones requeridas es $f(x) = \alpha.x$.

Supongamos que hay una única $g : K^{n\times n} \to K$ multilineal alternada con $g(I_n) = \alpha$.

Consideremos $f : K^{(n+1)\times(n+1)} \to K$ multilineal alternada tal que $f(I_{n+1}) = \alpha$. Sea $A = (a_{ij}) \in K^{(n+1)\times(n+1)}$. Por linealidad en la primer columna, se tiene que

$$f(A) = \sum_{i=1}^{n+1} a_{i1} \,.\, f \begin{pmatrix} 0 & a_{12} & \dots & a_{1n+1} \\ \vdots & \vdots & & \vdots \\ 0 & a_{i-1\,2} & \dots & a_{i-1\,n+1} \\ 1 & a_{i2} & \dots & a_{i\,n+1} \\ 0 & a_{i+1\,2} & \dots & a_{i+1\,n+1} \\ \vdots & \vdots & & \vdots \\ 0 & a_{n+1\,2} & \dots & a_{n+1\,n+1} \end{pmatrix}.$$

Restándole a la columna j la primer columna multiplicada por a_{ij} para $j = 2, \dots, n+1$, por la Proposición 5.2 iii) tenemos que

$$f(A) = \sum_{i=1}^{n+1} a_{i1} \,.\, f \begin{pmatrix} 0 & a_{12} & \dots & a_{1n+1} \\ \vdots & \vdots & & \vdots \\ 0 & a_{i-1\,2} & \dots & a_{i-1\,n+1} \\ 1 & 0 & \dots & 0 \\ 0 & a_{i+1\,2} & \dots & a_{i+1\,n+1} \\ \vdots & \vdots & & \vdots \\ 0 & a_{n+1\,2} & \dots & a_{n+1\,n+1} \end{pmatrix}$$

$$= \sum_{i=1}^{n+1} (-1)^{i-1} a_{i1} \, . \, f \begin{pmatrix} a_{12} & \dots & a_{1i} & 0 & a_{1i+1} & \dots & a_{1n+1} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i-1\,2} & \dots & a_{i-1\,i} & 0 & a_{i-1\,i+1} & \dots & a_{i-1\,n+1} \\ 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ a_{i+1\,2} & \dots & a_{i+1\,i} & 0 & a_{i+1\,i+1} & \dots & a_{i+1\,n+1} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n+1\,2} & \dots & a_{n+1\,i} & 0 & a_{n+1\,i+1} & \dots & a_{n+1\,n+1} \end{pmatrix}$$

$$= \sum_{i=1}^{n+1} (-1)^{i+1} a_{i1} \, . \, f_i(A(i|1)).$$

Por el Lema 5.3, $f_i : K^{n\times n} \to K$ es una función multilineal alternada y $f_i(I_n) = \alpha$, luego debe ser $f_i = g$. Por lo tanto, $f(A)$ es necesariamente

$$f(A) = \sum_{i=1}^{n+1} (-1)^{i+1} a_{i1} \, . \, g(A(i|1)),$$

de donde f es única. □

De la demostración anterior se deduce inmediatamente el siguiente resultado:

Corolario 5.6 *Sea $A \in K^{n\times n}$. Si para cada $r \in \mathbb{N}$, $\det : K^{r\times r} \to K$ denota la función determinante de orden r (quedando en claro por el contexto de qué función determinante se trata), entonces*

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+1} a_{i1} \, . \, \det(A(i|1)) = \sum_{i=1}^{n} (-1)^{i+1} a_{1i} \, . \, \det(A(1|i)).$$

Las fórmulas recursivas para el cálculo del determinante de una matriz dadas en el corolario anterior se llaman el desarrollo del determinante por la primera columna y por la primera fila respectivamente.

Ejemplo. Calcular $\det(A)$, siendo $A = \begin{pmatrix} 2 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix}$.

Utilizando la fórmula del desarrollo por la primera fila del Corolario 5.6 obtenemos:

$$\begin{aligned} \det \begin{pmatrix} 2 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix} &= (-1)^2 . \, 2 . \det \begin{pmatrix} 1 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 1 & 0 \end{pmatrix} + (-1)^5 . \, 1 . \det \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix} \\ &= 2. \left((-1)^2 .1. \det \begin{pmatrix} -1 & 0 \\ 1 & 0 \end{pmatrix} + (-1)^4 .1. \det \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix} \right) + \end{aligned}$$

$$+ (-1)\left((-1)^3.1.\det\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}\right)$$
$$= \quad 2.(0+1) + (-1)(-1) = 3.$$

5.2 Propiedades del determinante

En esta sección estudiaremos algunas propiedades básicas de los determinantes que facilitan su cálculo.

5.2.1 Determinante de la transpuesta de una matriz

La identidad enunciada en el Corolario 5.6 nos permite deducir el siguiente resultado:

Proposición 5.7 *Sea* $A \in K^{n\times n}$. *Entonces* $\det(A) = \det(A^t)$.

Demostración. Probaremos la igualdad por inducción en n:

Para $n = 1$, no hay nada que probar.

Supongamos que vale para n y sea $A = (a_{ij}) \in K^{(n+1)\times(n+1)}$. Entonces

$$\begin{aligned} \det(A^t) &= \sum_{i=1}^{n+1} (-1)^{i+1} (A^t)_{1i} \det(A^t(1|i)) = \sum_{i=1}^{n+1} (-1)^{i+1} a_{i1} \det(A(1|i)^t) \\ &= \sum_{i=1}^{n+1} (-1)^{i+1} a_{i1} \det(A(i|1)) = \det(A). \end{aligned}$$ □

Observación 5.8 La definición de función multilineal alternada podría haberse dado en términos de las filas de las matrices, en lugar de respecto de las columnas, y se hubiese obtenido la misma función determinante.

5.2.2 Matrices triangulares

Un caso en el que es muy fácil calcular determinantes es el de las matrices triangulares. Lo veremos para matrices triangulares superiores, pero el mismo resultado vale para una matriz triangular inferior.

Proposición 5.9 *Sea* $A = (a_{ij}) \in K^{n\times n}$ *una matriz triangular superior. Entonces* $\det(A) = \prod_{i=1}^{n} a_{ii}$.

Demostración. Probaremos la validez de la fórmula por inducción en n:

Para $n = 1$, no hay nada que hacer.

Supongamos que vale para n y sea $A = (a_{ij}) \in K^{(n+1)\times(n+1)}$ una matriz triangular superior. Entonces

$$\det(A) = \sum_{i=1}^{n+1} (-1)^{i+1} a_{i1} \det(A(i|1)) = a_{11} \det(A(1|1)),$$

puesto que por nuestra hipótesis sobre A, se tiene que $a_{i1} = 0$ para cada $i \geq 2$.

Ahora, la matriz $A(1|1) \in K^{n\times n}$ es triangular superior y entonces, por hipótesis inductiva, $\det(A(1|1)) = \prod\limits_{j=1}^{n} (A(1|1))_{jj} = \prod\limits_{i=2}^{n+1} a_{ii}$. En consecuencia,

$$\det(A) = a_{11} \det(A(1|1)) = \prod_{i=1}^{n+1} a_{ii},$$

que es lo que se quería probar. □

Observación 5.10 Dado que el determinante es una función multilineal alternada por filas (ver Observación 5.8), podemos calcular el determinante de una matriz triangulándola, teniendo en cuenta el cambio del determinante de acuerdo a la operación elemental efectuada:

- $\det \begin{pmatrix} F_1 \\ \vdots \\ F_i \\ \vdots \\ F_j \\ \vdots \\ F_n \end{pmatrix} = -\det \begin{pmatrix} F_1 \\ \vdots \\ F_j \\ \vdots \\ F_i \\ \vdots \\ F_n \end{pmatrix}$ "*Intercambiar dos filas*" cambia el signo del determinante.

- $\det \begin{pmatrix} F_1 \\ \vdots \\ \lambda F_i \\ \vdots \\ F_n \end{pmatrix} = \lambda \det \begin{pmatrix} F_1 \\ \vdots \\ F_i \\ \vdots \\ F_n \end{pmatrix}$ "*Multiplicar una fila por una constante no nula*" multiplica el determinante por esa constante.

- $\det \begin{pmatrix} F_1 \\ \vdots \\ F_i + \lambda F_j \\ \vdots \\ F_j \\ \vdots \\ F_n \end{pmatrix} = \det \begin{pmatrix} F_1 \\ \vdots \\ F_i \\ \vdots \\ F_j \\ \vdots \\ F_n \end{pmatrix}$ "*Sumarle a una fila un múltiplo de otra*" no modifica el determinante.

El determinante de la matriz triangular obtenida es el producto de los elementos de su diagonal.

Ejemplo. Calcular $\det(A)$, siendo $A = \begin{pmatrix} 2 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix}$.

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

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

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

5.2.3 Desarrollo del determinante por una fila o columna

Veremos ahora fórmulas para el determinante análogas a las del Corolario 5.6 en las cuales la primera columna (o fila) de la matriz es reemplazada por cualquier otra columna o fila.

Sea $A \in K^{n\times n}$, $A = (a_{ij})$. Sean $A_1, \dots, A_n$ las columnas de A. Observemos que se puede ubicar la j-ésima columna de A en el lugar de la primera, sin modificar el orden de las restantes, por medio de $j-1$ intercambios de columnas. Entonces

$$\begin{aligned} \det(A) &= (-1)^{j-1}\det(\, A_j \mid A_1 \mid A_2 \mid \dots \mid A_{j-1} \mid A_{j+1} \mid \dots \mid A_n\,) \\ &= (-1)^{j-1}\Big(\sum_{i=1}^{n}(-1)^{1+i}a_{ij}\,\det(A(i|j))\Big) \\ &= \sum_{i=1}^{n}(-1)^{i+j}a_{ij}\,\det(A(i|j)), \end{aligned}$$

lo que da una fórmula para el desarrollo del determinante de una matriz por la j-ésima columna para $1 \le j \le n$ arbitrario. Usando que $\det(A) = \det(A^t)$, se prueba una fórmula para el desarrollo por la i-ésima fila para cualquier $1 \le i \le n$.

Hemos demostrado el siguiente resultado:

Proposición 5.11 *Sea $A = (a_{ij}) \in K^{n\times n}$. Entonces*

$$\begin{aligned} \det(A) &= \sum_{i=1}^{n}(-1)^{i+j}a_{ij}\,\det(A(i|j)) \qquad \textit{para todo } j \textit{ con } 1 \le j \le n \\ \det(A) &= \sum_{j=1}^{n}(-1)^{i+j}a_{ij}\,\det(A(i|j)) \qquad \textit{para todo } i \textit{ con } 1 \le i \le n. \end{aligned}$$

Ejemplo. Calcular $\det(A)$, siendo $A = \begin{pmatrix} 1 & -2 & 1 & 2 \\ 1 & 0 & 2 & 1 \\ 1 & 0 & 2 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix}$.

Desarrollando el determinante por la segunda columna, se obtiene:

$$\det(A) = (-1)^{1+2}. (-2). \det \begin{pmatrix} 1 & 2 & 1 \\ 1 & 2 & 0 \\ 1 & 1 & 0 \end{pmatrix} = 2. \det \begin{pmatrix} 1 & 2 & 1 \\ 1 & 2 & 0 \\ 1 & 1 & 0 \end{pmatrix},$$

y desarrollando por la tercera columna el determinante del miembro derecho,

$$\det(A) = 2. (-1)^{1+3}. 1. \det \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} = 2. (1. 1 - 2. 1) = -2.$$

5.2.4 Determinante del producto de matrices

En esta sección estudiaremos cómo se comporta el determinante con respecto al producto de matrices. Para esto, probaremos en primer lugar un resultado sobre funciones multilineales alternadas que nos será de utilidad.

Proposición 5.12 *Sea $f : K^{n\times n} \to K$ multilineal alternada tal que $f(I_n) = \alpha$. Entonces* $f = \alpha. \det$.

Demostración. Como consecuencia de que $\det : K^{n\times n} \to K$ es una función multilineal alternada, se tiene que $\alpha. \det : K^{n\times n} \to K$ es multilineal alternada. Además, $(\alpha. \det)(I_n) = \alpha$.

Por la unicidad de las funciones multilineales alternadas (ver Teorema 5.4), resulta que $f = \alpha. \det$. □

Proposición 5.13 *Sean $A, B \in K^{n\times n}$. Entonces* $\det(A.B) = \det(A). \det(B)$.

Demostración. Se define $f : K^{n\times n} \to K$ como $f(X) = \det(A.X)$. Nuestro objetivo es probar que para cada $B \in K^{n\times n}$ se tiene que $f(B) = \det(A). \det(B)$.

Observamos que la función f es multilineal alternada y calculamos su valor en I_n:

i) Para i con $1 \le i \le n$,

$$\begin{aligned} f(\, X_1 \mid \ldots \mid X_i + X_i' \mid \ldots \mid X_n\,) &= \det\big(A. (\, X_1 \mid \ldots \mid X_i + X_i' \mid \ldots \mid X_n\,)\big) = \\ &= \det(\, AX_1 \mid \ldots \mid AX_i + AX_i' \mid \ldots \mid AX_n\,) = \\ &= \det(\, AX_1 \mid \ldots \mid AX_i \mid \ldots \mid AX_n\,) + \det(\, AX_1 \mid \ldots \mid AX_i' \mid \ldots \mid AX_n\,) = \\ &= f(\, X_1 \mid \ldots \mid X_i \mid \ldots \mid X_n\,) + f(\, X_1 \mid \ldots \mid X_i' \mid \ldots \mid X_n\,). \end{aligned}$$

ii) Para $\lambda \in K$, $1 \le i \le n$,

$$\begin{aligned} f(\, X_1 \mid \ldots \mid \lambda X_i \mid \ldots \mid X_n\,) &= \det\big(A.(\, X_1 \mid \ldots \mid \lambda X_i \mid \ldots \mid X_n\,)\big) = \\ &= \det\big(\, AX_1 \mid \ldots \mid A.\lambda X_i \mid \ldots \mid AX_n\,\big) = \lambda \det\big(\, AX_1 \mid \ldots \mid AX_i \mid \ldots \mid AX_n\,\big) = \\ &= \lambda f(\, X_1 \mid \ldots \mid X_i \mid \ldots \mid X_n\,). \end{aligned}$$

iii) Para $1 \le i \le n$,

$$f(\, X_1 \mid \dots \mid X_i \mid \dots \mid X_i \mid \dots \mid X_n \,) = \det\big(A.(\, X_1 \mid \dots \mid X_i \mid \dots \mid X_i \mid \dots \mid X_n \,)\big) =$$
$$= \det\big(\, AX_1 \mid \dots \mid AX_i \mid \dots \mid AX_i \mid \dots \mid AX_n \,\big) = 0.$$

iv) $f(I_n) = \det(A.I_n) = \det(A)$.

Por la proposición anterior, resulta que $f = \det(A).\det$.
Luego, para cada $B \in K^{n\times n}$ se tiene que $\det(A.B) = f(B) = \det(A).\det(B)$. □

5.3 Determinantes y matrices inversibles

El objetivo de esta sección es estudiar la relación entre determinantes e inversibilidad de matrices. Probaremos que una matriz $A \in K^{n\times n}$ es inversible si y sólo si su determinante es no nulo. A continuación, mostraremos que los determinantes pueden ser utilizados también para la resolución de sistemas de ecuaciones lineales.

5.3.1 Inversibilidad de matrices

El siguiente resultado, cuya demostración se basa en la fórmula para el determinante de un producto de matrices vista en la sección anterior, caracteriza la inversibilidad de matrices por medio de determinantes:

Proposición 5.14 *Sea $A \in K^{n\times n}$. Entonces A es inversible si y sólo si* $\det A \ne 0$.

Demostración.

($\Rightarrow$) Supongamos que $A \in K^{n\times n}$ es inversible. Entonces existe una matriz $B \in K^{n\times n}$ tal que $A.B = I_n$. Aplicando el resultado de la Proposición 5.13 se obtiene que

$$1 = \det(I_n) = \det(A.B) = \det(A).\det(B),$$

de donde se deduce que $\det(A) \ne 0$.

($\Leftarrow$) Si $\det(A) \ne 0$, entonces las columnas de A son linealmente independientes (ver Proposición 5.2 iv)) y, por lo tanto, A es inversible (ver Ejercicio 17, Sección 2.5). □

5.3.2 Adjunta de una matriz

Dada una matriz $A \in K^{n\times n}$, podemos asociarle una matriz, cuyos elementos se calculan a partir de determinantes de submatrices de A que, en el caso en que A sea inversible, nos permitirá obtener la inversa de A.

Definición 5.15 Sea $A = (a_{ij}) \in K^{n\times n}$. Se llama *adjunta de* A, y se nota $\mathrm{adj}(A)$, a la matriz $\mathrm{adj}(A) \in K^{n\times n}$ definida por

$$(\mathrm{adj}(A))_{ij} = (-1)^{i+j}\det(A(j|i)).$$

Ejemplo. Sea $A = \begin{pmatrix} 1 & 5 & 2 \\ 2 & 1 & 1 \\ 1 & 2 & 0 \end{pmatrix}$. Entonces la adjunta de A es la matriz

$$\mathrm{adj}(A) = \begin{pmatrix} +\det\begin{pmatrix} 1 & 1 \\ 2 & 0 \end{pmatrix} & -\det\begin{pmatrix} 5 & 2 \\ 2 & 0 \end{pmatrix} & +\det\begin{pmatrix} 5 & 2 \\ 1 & 1 \end{pmatrix} \\ -\det\begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} & +\det\begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix} & -\det\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \\ +\det\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} & -\det\begin{pmatrix} 1 & 5 \\ 1 & 2 \end{pmatrix} & +\det\begin{pmatrix} 1 & 5 \\ 2 & 1 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} -2 & 4 & 3 \\ 1 & -2 & 3 \\ 3 & 3 & -9 \end{pmatrix}$$

Si calculamos $A.\,\mathrm{adj}(A)$, tenemos que

$$A.\,\mathrm{adj}(A) = \begin{pmatrix} 9 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{pmatrix},$$

de donde se deduce que $A^{-1} = \frac{1}{9}.\,\mathrm{adj}(A)$. Teniendo en cuenta que $\det(A) = 9$, resulta que $A^{-1} = \frac{1}{\det(A)}.\,\mathrm{adj}(A)$.

En el ejemplo anterior obtuvimos una relación entre la matriz A, su adjunta y su determinante. La siguiente proposición muestra que lo mismo sucede en general.

Proposición 5.16 *Sea* $A \in K^{n\times n}$. *Entonces* $A.\,\mathrm{adj}(A) = \det(A).\,I_n$. *Luego, si* $\det(A) \neq 0$, *se tiene que* $A^{-1} = \frac{1}{\det(A)}.\,\mathrm{adj}(A)$.

Demostración. Sea $A \in K^{n\times n}$, $A = (a_{ij})$. Entonces

$$(A.\,\mathrm{adj}(A))_{k\ell} = \sum_{i=1}^{n} a_{ki}.\,(\mathrm{adj}(A))_{i\,\ell} = \sum_{i=1}^{n} a_{ki}\,(-1)^{i+\ell}\det(A(\ell|i)).$$

Si $k = \ell$, entonces $(A.\,\mathrm{adj}(A))_{\ell\,\ell} = \sum\limits_{i=1}^{n}(-1)^{i+\ell}a_{\ell\,i}\det(A(\ell|i)) = \det(A)$, puesto que la sumatoria resulta ser el desarrollo de $\det(A)$ por la ℓ-ésima fila.

Por otro lado, si $k \neq \ell$, se tiene que $(A.\operatorname{adj}(A))_{k\ell} = \sum_{i=1}^{n} (-1)^{i+\ell} a_{ki} \det(A(\ell|i)) = 0$, puesto que se trata del desarrollo por la ℓ-ésima fila del determinante de la matriz $A^{(k\ell)}$ definida por

$$(A^{(k\ell)})_{ij} = \begin{cases} a_{ij} & \text{si } i \neq \ell \\ a_{kj} & \text{si } i = \ell \end{cases}$$

que tiene dos filas iguales.

La segunda parte de la proposición se deduce inmediatamente de la igualdad $A.\operatorname{adj}(A) = \det(A).I_n$. □

5.3.3 Regla de Cramer

Por último en esta sección presentaremos la regla de Cramer, que permite obtener la (única) solución de un sistema lineal no homogéneo cuya matriz asociada es inversible por medio de determinantes.

Proposición 5.17 (Regla de Cramer) *Sea $A = (a_{ij}) \in K^{n\times n}$ una matriz inversible, y sea $b \in K^{n\times 1}$. Entonces la (única) solución del sistema lineal $A.x = b$ está dada por*

$$x_i = \frac{\det \begin{pmatrix} a_{11} & \dots & a_{1\,i-1} & b_1 & a_{1\,i+1} & \dots & a_{1n} \\ a_{21} & \dots & a_{2\,i-1} & b_2 & a_{2\,i+1} & \dots & a_{2n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \dots & a_{n\,i-1} & b_n & a_{n\,i+1} & \dots & a_{nn} \end{pmatrix}}{\det A} \qquad \textit{para } i = 1, \dots, n.$$

Demostración. Multiplicando la ecuación $A.x = b$ por $\operatorname{adj}(A)$, se tiene que

$$\operatorname{adj}(A).A.x = \operatorname{adj}(A).\, b.$$

Por la proposición anterior y el hecho que A es inversible, $\operatorname{adj}(A).A = A.\operatorname{adj}(A) = \det(A).\, I_n$. En consecuencia,

$$\det(A).x = \operatorname{adj}(A).\, b.$$

Sea $1 \leq i \leq n$. Entonces, de la igualdad anterior, resulta que

$$\begin{aligned} \det(A).x_i &= (\operatorname{adj}(A).\, b)_{i1} = \sum_{j=1}^{n} (-1)^{i+j} \det(A(j|i))\, b_j = \sum_{j=1}^{n} (-1)^{i+j} b_j \det(A(j|i)) = \\ &= \det \begin{pmatrix} a_{11} & \dots & a_{1i-1} & b_1 & a_{1i+1} & \dots & a_{1n} \\ a_{21} & \dots & a_{2i-1} & b_2 & a_{2i+1} & \dots & a_{2n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \dots & a_{ni-1} & b_n & a_{ni+1} & \dots & a_{nn} \end{pmatrix}, \end{aligned}$$

de donde se deduce la fórmula del enunciado de la proposición. □

La regla de Cramer en general no se utiliza para resolver sistemas de ecuaciones lineales, pero es útil para derivar resultados teóricos sobre las soluciones de esta clase de sistemas.

Ejemplo. Sea $A \in \mathbb{Z}^{n\times n}$ tal que $\det(A) = \pm 1$ y sea $b \in \mathbb{Z}^n$. Entonces el sistema lineal $A.x = b$ tiene solución en $\mathbb{Z}^n$.

Sea $x_0 \in \mathbb{Q}^n$ la solución del sistema $A.x = b$. Por la regla de Cramer, sabemos que cada coordenada de x_0 se obtiene como el cociente entre el determinante de una matriz cuyos coeficientes son coeficientes de A y de b, y el determinante de la matriz A. Como tanto los coeficientes de A como los de b son números enteros, el determinante que aparece en cada numerador es un número entero y, puesto que $\det(A) = \pm 1$, el cociente resulta entero. Luego $x_0 \in \mathbb{Z}^n$.

5.4 Cálculo de algunos determinantes

Ejemplo. Calcular el determinante de la matriz $A \in K^{n\times n}$ definida por:

$$A = \begin{pmatrix} t & 0 & 0 & \dots & 0 & 0 & a_0 \\ -1 & t & 0 & \dots & 0 & 0 & a_1 \\ 0 & -1 & t & \dots & 0 & 0 & a_2 \\ 0 & 0 & -1 & \dots & 0 & 0 & a_3 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & -1 & t & a_{n-2} \\ 0 & 0 & 0 & \dots & 0 & -1 & t + a_{n-1} \end{pmatrix}$$

Probaremos, inductivamente, que $\det(A) = t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0$.

Para $n = 2$:

$$\det \begin{pmatrix} t & a_0 \\ -1 & t + a_1 \end{pmatrix} = t(t + a_1) + a_0 = t^2 + a_1 t + a_0.$$

Supongamos que vale para toda matriz de este tipo en $K^{n\times n}$. Entonces, dada una matriz de esta forma en $K^{(n+1)\times(n+1)}$, desarrollando el determinante por la primera fila, y aplicando luego la hipótesis inductiva resulta que

$$\det \begin{pmatrix} t & 0 & 0 & \dots & 0 & 0 & a_0 \\ -1 & t & 0 & \dots & 0 & 0 & a_1 \\ 0 & -1 & t & \dots & 0 & 0 & a_2 \\ 0 & 0 & -1 & \dots & 0 & 0 & a_3 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & -1 & t & a_{n-1} \\ 0 & 0 & 0 & \dots & 0 & -1 & t + a_n \end{pmatrix} =$$

$$= t.\det \begin{pmatrix} t & 0 & 0 & \dots & 0 & a_1 \\ -1 & t & 0 & \dots & 0 & a_2 \\ 0 & -1 & t & \dots & 0 & a_3 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & t & a_{n-1} \\ 0 & 0 & 0 & \dots & -1 & t + a_n \end{pmatrix} + (-1)^{n+2} a_0 (-1)^n =$$

$$= t.(t^n + a_n t^{n-1} + \dots + a_1) + a_0 = t^{n+1} + a_n t^n + \dots + a_1 t + a_0.$$

Ejemplo. Dados $k_1, \dots, k_n \in K$ se define la matriz de *Vandermonde*:

$$V(k_1, k_2, \dots, k_n) = \begin{pmatrix} 1 & 1 & \dots & \dots & 1 \\ k_1 & k_2 & \dots & \dots & k_n \\ k_1^2 & k_2^2 & \dots & \dots & k_n^2 \\ \dots & \dots & \dots & \dots & \dots \\ k_1^{n-1} & k_2^{n-1} & \dots & \dots & k_n^{n-1} \end{pmatrix} \in K^{n \times n}.$$

Entonces $\det\big(V(k_1, k_2, \dots, k_n)\big) = \prod\limits_{1 \le i < j \le n} (k_j - k_i)$.

Vamos a probarlo por inducción en n:

Para $n = 2$,

$$\det\big(V(k_1, k_2)\big) = \det \begin{pmatrix} 1 & 1 \\ k_1 & k_2 \end{pmatrix} = k_2 - k_1,$$

y por lo tanto, la fórmula vale.

Supongamos ahora que vale para n y calculemos $\det\big(V(k_1, k_2, \dots, k_n, k_{n+1})\big)$. Se tiene que

$$\det\big(V(k_1, k_2, \dots, k_{n+1})\big) = \det \begin{pmatrix} 1 & 1 & \dots & \dots & 1 \\ k_1 & k_2 & \dots & \dots & k_{n+1} \\ k_1^2 & k_2^2 & \dots & \dots & k_{n+1}^2 \\ \dots & \dots & \dots & \dots & \dots \\ k_1^n & k_2^n & \dots & \dots & k_{n+1}^n \end{pmatrix}.$$

Para $i = n, n-1, \dots, 2$ a la i-ésima fila de esta matriz le restamos k_1 veces la fila $i - 1$ y obtenemos:

$$\det\big(V(k_1, k_2, \dots, k_{n+1})\big) =$$

$$= \det \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 0 & k_2 - k_1 & k_3 - k_1 & \dots & k_{n+1} - k_1 \\ 0 & k_2^2 - k_1 k_2 & k_3^2 - k_1 k_3 & \dots & k_{n+1}^2 - k_1 k_{n+1} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & k_2^n - k_1 k_2^{n-1} & k_3^n - k_1 k_3^{n-1} & \dots & k_{n+1}^n - k_1 k_{n+1}^{n-1} \end{pmatrix}$$

$$= \det \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & k_2 - k_1 & k_3 - k_1 & \dots & k_{n+1} - k_1 \\ 0 & (k_2 - k_1)k_2 & (k_3 - k_1)k_3 & \dots & (k_{n+1} - k_1)k_{n+1} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & (k_2 - k_1)k_2^{n-1} & (k_3 - k_1)k_3^{n-1} & \dots & (k_{n+1} - k_1)k_{n+1}^{n-1} \end{pmatrix}$$

$$= \prod_{j=2}^{n+1}(k_j - k_1).\ \det \begin{pmatrix} 1 & 1 & \dots & 1 \\ k_2 & k_3 & \dots & k_{n+1} \\ \dots & \dots & \dots & \dots \\ k_2^{n-1} & k_3^{n-1} & \dots & k_{n+1}^{n-1} \end{pmatrix}$$

$$= \prod_{j=2}^{n+1}(k_j - k_1)\ .\ \prod_{2\leq i<j\leq n+1}(k_j - k_i)$$

$$= \prod_{1\leq i<j\leq n+1}(k_j - k_i).$$

Observación 5.18 Del ejemplo anterior se deduce que, si $k_1, \dots, k_n \in K$ son escalares distintos, la matriz $V(k_1, \dots, k_n) \in K^{n\times n}$ es inversible, pues su determinante es no nulo.

La matriz de Vandermonde se relaciona con el siguiente problema de interpolación: dados $\alpha_0, \dots, \alpha_n \in K$ escalares *distintos*, y $\beta_0, \dots, \beta_n \in K$ escalares arbitrarios, hallar un polinomio $P = \sum\limits_{i=0}^{n} a_i X^i \in K[X]$ de grado menor o igual que n tal que $P(\alpha_i) = \beta_i$ para cada $0 \leq i \leq n$. Estas condiciones dan lugar a un sistema de ecuaciones lineales para los coeficientes de P:

$$(V(\alpha_0, \dots, \alpha_n))^t.x = \beta^t,$$

donde x_i $(0 \leq i \leq n)$ representa el coeficiente de X^i en P y $\beta = (\beta_0, \dots, \beta_n) \in K^{n+1}$.

Ahora, siendo $(V(\alpha_0, \dots, \alpha_n))^t$ una matriz inversible (por ser $\alpha_0, \dots, \alpha_n$ escalares distintos), este sistema tiene solución única $(a_0, \dots, a_n) \in K^{n+1}$, y el polinomio $P = \sum\limits_{i=0}^{n} a_i X^i$ cuyos coeficientes son las coordenadas de esta solución resulta ser el único polinomio de grado menor o igual que n que satisface $P(\alpha_i) = \beta_i$ para cada $0 \leq i \leq n$ (polinomio interpolador de Lagrange).

5.5 Rango de una matriz y determinante

De acuerdo a lo que hemos visto previamente, para decidir si una matriz es inversible, basta verificar si su determinante es no nulo. En esta sección veremos que, aún en el caso de una matriz A no inversible, es posible determinar el rango de A calculando determinantes de submatrices de A.

Definición 5.19 Sea $A \in K^{n\times m}$ y sean $1 \leq r \leq n, 1 \leq s \leq m$. Una *submatriz de A en $K^{r\times s}$* es una matriz $B \in K^{r\times s}$ que se obtiene suprimiendo $n - r$ filas y $m - s$ columnas de A.

Ejemplo. Sea $A \in \mathbb{R}^{3\times 3}$, $A = \begin{pmatrix} 2 & 1 & 3 \\ 2 & 2 & 1 \\ 1 & 4 & 1 \end{pmatrix}$.

- Una submatriz de A de 2×3 es, por ejemplo, $\begin{pmatrix} 2 & 2 & 1 \\ 1 & 4 & 1 \end{pmatrix}$, que se obtiene al suprimir la primera fila de A.
- Una submatriz de A de 2×2 es, por ejemplo, $\begin{pmatrix} 2 & 3 \\ 2 & 1 \end{pmatrix}$, que se obtiene al suprimir la tercera fila y la segunda columna de A.

La relación entre rango y submatrices con determinante no nulo viene dada por la siguiente proposición.

Proposición 5.20 *Sea $A \in K^{n\times m}$. Son equivalentes:*

i) $\mathrm{rg}(A) \geq r$.

ii) Existe $B \in K^{r\times r}$ submatriz de A con $\det(B) \neq 0$.

Demostración.

i) $\Rightarrow$ *ii)* Si $\mathrm{rg}(A) \geq r$, entonces existen r filas $F_{i_1}, \dots, F_{i_r}$ $(i_1 < \dots < i_r)$ de A que son linealmente independientes. Consideramos la submatriz A' de A formada por dichas filas. Se tiene que $A' \in K^{r\times m}$ y $\mathrm{rg}(A') = r$. Esto último implica que A' tiene r columnas $C_{j_1}, \dots, C_{j_r}$ $(j_1 < \dots < j_r)$ linealmente independientes.

Sea $B \in K^{r\times r}$ la submatriz de A' cuyas columnas son $C_{j_1}, \dots, C_{j_r}$. Es claro que B es una submatriz de A y, como sus columnas son linealmente independientes, $\det(B) \neq 0$.

ii) $\Rightarrow$ *i)* Supongamos que $B \in K^{r\times r}$ es una submatriz de A con determinante no nulo. Entonces, las columnas de B son linealmente independientes.

Consideremos la submatriz $A' \in K^{r\times m}$ de A que resulta de suprimir las mismas filas que para obtener B (pero sin suprimir ninguna columna). Entonces las columnas de B son algunas de las columnas de A', con lo que $\mathrm{rg}(A') \geq \mathrm{rg}(B) = r$ y, por lo tanto, $\mathrm{rg}(A') = r$, pues A' tiene r filas.

Finalmente, observemos que las filas de A son algunas de las filas de A', de donde $\mathrm{rg}(A) \geq \mathrm{rg}(A') = r$. □

De esta proposición se desprende el siguiente resultado que permite obtener el rango de una matriz estudiando los rangos de sus submatrices cuadradas.

Observación 5.21 Sea $A \in K^{n\times m}$ una matriz que posee una submatriz de $r \times r$ inversible, pero no posee ninguna submatriz de $(r + 1) \times (r + 1)$ inversible. Entonces $\mathrm{rg}(A) = r$.

5.6 Otra fórmula para el determinante

Concluimos este capítulo dando una fórmula alternativa para el determinante.

Dada $A = (a_{ij}) = \begin{pmatrix} F_1 \\ F_2 \\ \vdots \\ F_n \end{pmatrix} \in K^{n\times n}$, usando que el determinante es una función multilineal alternada por filas tenemos que

$$\det(A) = \sum_{1\le i_1 \le n} a_{1i_1} \det \begin{pmatrix} e_{i_1} \\ F_2 \\ \vdots \\ F_n \end{pmatrix}$$

donde e_{i_1} es el i_1-ésimo vector de la base canónica de K^n. Repitiendo el procedimento para todas las filas tenemos que

$$\det(A) = \sum_{1\le i_1, i_2, \dots, i_n \le n} a_{1i_1} a_{2i_2} \dots a_{ni_n} \det \begin{pmatrix} e_{i_1} \\ e_{i_2} \\ \vdots \\ e_{i_n} \end{pmatrix}.$$

Dado que, si una matriz tiene dos filas iguales, su determinante es cero, en la suma podemos quedarnos sólo con los determinantes de las matrices cuyas filas son los n vectores distintos de la base canónica, eventualmente cambiados de orden. Para facilitar la notación, daremos la siguiente

Definición 5.22 Sea $I_n = \{1, 2, \dots, n\} \subseteq \mathbb{N}$. Una *permutación* de I_n es una función $\sigma : I_n \to I_n$ biyectiva. El conjunto de todas las permutaciones de I_n se nota S_n.

Luego,

$$\det(A) = \sum_{\sigma\in S_n} a_{1\sigma(1)} a_{2\sigma(2)} \dots a_{n\sigma(n)} \det \begin{pmatrix} e_{\sigma(1)} \\ e_{\sigma(2)} \\ \vdots \\ e_{\sigma(n)} \end{pmatrix}.$$

Como el determinante de la matriz $\begin{pmatrix} e_{\sigma(1)} \\ e_{\sigma(2)} \\ \vdots \\ e_{\sigma(n)} \end{pmatrix}$ sólo depende de la permutación σ y siempre da 1 o -1 (ya que la matriz se obtiene intercambiando filas de la matriz I_n), podemos definir el *signo de la permutación* σ (que notaremos $\mathrm{sg}(\sigma)$) como dicho determinante. Usando esta notación, tenemos:

Proposición 5.23 *Sea* $A = (a_{ij}) \in K^{n\times n}$. *Entonces*

$$\det(A) = \sum_{\sigma \in S_n} \mathrm{sg}(\sigma)\, a_{1\sigma(1)} a_{2\sigma(2)} \dots a_{n\sigma(n)}.$$

Observación 5.24 Para calcular el signo de una permutación $\sigma \in S_n$ basta considerar la matriz $\begin{pmatrix} e_{\sigma(1)} \\ e_{\sigma(2)} \\ \vdots \\ e_{\sigma(n)} \end{pmatrix}$ y contar cuántos intercambios de filas se deben realizar para conseguir la matriz I_n. Si el número de intercambios es r, el signo de la permutación será $(-1)^r$.

Nota. La definición de signo de una permutación puede darse independientemente de la definición de determinante. De esta forma, la identidad de la Proposición 5.23 nos da una función de $K^{n\times n}$ en K que puede probarse que es multilineal alternada y que en la identidad vale 1 y por lo tanto es la función determinante. Esto nos daría una definición no inductiva del determinante independiente de nuestra definición anterior.

5.7 Ejercicios

Ejercicio 1. Calcular el determinante de las siguientes matrices:

i) $\begin{pmatrix} -3 & 2 \\ 4 & 5 \end{pmatrix}$ ii) $\begin{pmatrix} 2 & -2 \\ -1 & 1 \end{pmatrix}$ iii) $\begin{pmatrix} 1 & 2 & 5 \\ -3 & 0 & -1 \\ 1 & -4 & -2 \end{pmatrix}$

iv) $\begin{pmatrix} 2 & -1 & 3 \\ -1 & 1 & -2 \\ 4 & -1 & 5 \end{pmatrix}$ v) $\begin{pmatrix} 2 & 3 & -2 & 5 \\ 4 & -5 & 0 & 6 \\ 2 & 0 & -1 & 7 \\ 6 & 3 & -4 & 8 \end{pmatrix}$ vi) $\begin{pmatrix} 5 & 4 & -2 & 5 \\ 2 & -3 & 0 & 6 \\ 0 & 0 & 2 & 0 \\ -4 & 3 & 3 & 8 \end{pmatrix}$

Ejercicio 2. Calcular el determinante de las matrices elementales definidas en la Sección 2.3.

Ejercicio 3. Calcular el determinante de $A \in K^{n\times n}$ siendo

$$A = \begin{pmatrix} 0 & 0 & \dots & 0 & a_1 \\ 0 & 0 & \dots & a_2 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & a_{n-1} & \dots & 0 & 0 \\ a_n & 0 & \dots & 0 & 0 \end{pmatrix}.$$

Ejercicio 4.

i) Si $A \in K^{n\times n}$, $B \in K^{m\times m}$ y $C \in K^{n\times m}$, sea $M \in K^{(n+m)\times(n+m)}$ la matriz de bloques definida por $M = \begin{pmatrix} A & C \\ 0 & B \end{pmatrix}$. Probar que $\det(M) = \det(A).\det(B)$.

ii) Sean $r_1, r_2, \dots, r_n \in \mathbb{N}$ y para cada i, $1 \le i \le n$, sea $A_i \in K^{r_i\times r_i}$. Se considera la matriz de bloques

$$M = \begin{pmatrix} A_1 & 0 & 0 & \dots & 0 \\ 0 & A_2 & 0 & \dots & 0 \\ 0 & 0 & A_3 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & A_n \end{pmatrix}.$$

Calcular $\det(M)$.

Ejercicio 5. Calcular los determinantes de las siguientes matrices:

i) $\begin{pmatrix} 1 & 2 & 3 & \dots & n \\ -1 & 0 & 3 & \dots & n \\ -1 & -2 & 0 & \dots & n \\ \dots & \dots & \dots & \dots & \dots \\ -1 & -2 & -3 & \dots & 0 \end{pmatrix}$ ii) $\begin{pmatrix} x & a & a & \dots & a \\ a & x & a & \dots & a \\ a & a & x & \dots & a \\ \dots & \dots & \dots & \dots & \dots \\ a & a & a & \dots & x \end{pmatrix}$

iii) $\begin{pmatrix} 0 & 1 & 1 & \dots & 1 & 1 \\ 1 & 0 & x & \dots & x & x \\ 1 & x & 0 & \dots & x & x \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & x & x & \dots & 0 & x \\ 1 & x & x & \dots & x & 0 \end{pmatrix}$

Ejercicio 6. Calcular inductivamente el determinante de $A \in \mathbb{R}^{n\times n}$:

$$A = \begin{pmatrix} 2 & 1 & 0 & 0 & \dots & \dots & 0 \\ 1 & 2 & 1 & 0 & \dots & \dots & 0 \\ 0 & 1 & 2 & 1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & 1 & 2 & 1 & 0 \\ 0 & \dots & \dots & 0 & 1 & 2 & 1 \\ 0 & \dots & \dots & \dots & 0 & 1 & 2 \end{pmatrix}.$$

Ejercicio 7. (Cálculo alternativo para el determinante de la matriz de Vandermonde.) Dada la matriz

$$V(k_1, k_2, \dots, k_n) = \begin{pmatrix} 1 & 1 & \dots & \dots & 1 \\ k_1 & k_2 & \dots & \dots & k_n \\ k_1^2 & k_2^2 & \dots & \dots & k_n^2 \\ \dots & \dots & \dots & \dots & \dots \\ k_1^{n-1} & k_2^{n-1} & \dots & \dots & k_n^{n-1} \end{pmatrix},$$

probar que $\det\big(V(k_1, k_2, \ldots, k_n)\big) = \prod_{1\le i<j\le n} (k_j - k_i)$ de la siguiente forma: Sin perder generalidad se puede suponer que $k_i \neq k_j$ si $i \neq j$. Si se considera el determinante de $V(k_1, k_2, \ldots, k_{n-1}, X)$ como polinomio en X, probar que $k_1, \ldots, k_{n-1}$ son sus raíces y factorizarlo.

Ejercicio 8. Calcular los siguientes determinantes:

i) $\begin{pmatrix} 1+a & 1+b & 1+c & 1+d \\ 1+a^2 & 1+b^2 & 1+c^2 & 1+d^2 \\ 1+a^3 & 1+b^3 & 1+c^3 & 1+d^3 \\ 1+a^4 & 1+b^4 & 1+c^4 & 1+d^4 \end{pmatrix}$ ii) $\begin{pmatrix} 1 & 1 & 1 & 1 \\ a^2 & b^2 & c^2 & d^2 \\ a^3 & b^3 & c^3 & d^3 \\ a^4 & b^4 & c^4 & d^4 \end{pmatrix}$

Ejercicio 9. Sea $A = (a_{ij}) \in \mathbb{R}^{3\times 3}$ tal que $A.\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 7 \end{pmatrix}$. Si $\det(A) = 3$, calcular el determinante de la matriz

$$\begin{pmatrix} a_{12} & a_{22} & a_{32} \\ 1 & 2 & 7 \\ a_{11} + 2a_{13} & a_{21} + 2a_{23} & a_{31} + 2a_{33} \end{pmatrix}.$$

Ejercicio 10. Dadas las matrices $A,\ B \in \mathbb{R}^{2\times 2}$

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

probar que no existe ninguna matriz $C \in GL(2, \mathbb{R})$ tal que $A.C = C.B$. ¿Y si no se pide que C sea inversible?

Ejercicio 11. Sea $A \in \mathbb{R}^{3\times 3}$ la matriz $A = \begin{pmatrix} 0 & 1 & 2 \\ 0 & 1 & 2 \\ 0 & 2 & 3 \end{pmatrix}$ y sea $B \in \mathbb{R}^{3\times 3}$, $B = (b_{ij})$, una matriz tal que $\det(A + B) = \det(A - B)$. Probar que B es inversible si y sólo si $b_{11} \neq b_{21}$.

Ejercicio 12.

i) Sea $A \in \mathbb{R}^{4\times 4}$ la matriz

$$A = \begin{pmatrix} a & b & c & d \\ b & -a & d & -c \\ c & -d & -a & b \\ d & c & -b & -a \end{pmatrix}.$$

Probar que el sistema $A.x = 0$ tiene solución única si y sólo si a, b, c y d no son todos iguales a cero.

ii) Analizar la validez de la afirmación anterior si $A \in \mathbb{C}^{4\times 4}$.

Ejercicio 13. Sea $A \in K^{n\times n}$ y sea $\lambda \in K$. Probar que existe $x \in K^n$, $x \neq 0$, tal que $A.x = \lambda.x$ si y sólo si $\det(A - \lambda.I_n) = 0$.

Ejercicio 14. Sean $\alpha_1, \dots, \alpha_n \in \mathbb{R}$, todos distintos y no nulos. Probar que las funciones $e^{\alpha_1 x}, \dots, e^{\alpha_n x}$ son linealmente independientes sobre $\mathbb{R}$. Deducir que $\mathbb{R}^{\mathbb{R}}$ no tiene dimensión finita.

Sugerencia: Derivar $n-1$ veces la función $\sum\limits_{i=1}^{n} c_i e^{\alpha_i x}$.

Ejercicio 15. Calcular el determinante, la adjunta y la inversa de cada una de las siguientes matrices:

$$\text{i)} \begin{pmatrix} 2 & 3 \\ 5 & 1 \end{pmatrix} \qquad \text{ii)} \begin{pmatrix} 2 & -3 & 3 \\ -5 & 4 & 0 \\ 0 & -2 & 2 \end{pmatrix} \qquad \text{iii)} \begin{pmatrix} -1 & 1 & 6 & 5 \\ 1 & 1 & 2 & 3 \\ -1 & 2 & 5 & 4 \\ 2 & 1 & 0 & 1 \end{pmatrix} \qquad \text{iv)} \begin{pmatrix} \cos & 0 & -\text{sen} \\ 0 & 1 & 0 \\ \text{sen} & 0 & \cos \end{pmatrix}$$

Ejercicio 16. Sea A una matriz inversible. Calcular $\det\big(\operatorname{adj} A\big)$. ¿Qué pasa si A no es inversible?

Ejercicio 17.

i) Resolver los siguientes sistemas lineales sobre $\mathbb{Q}$ empleando la regla de Cramer:

$$\text{a)} \begin{cases} 3.x_1 - x_2 = -3 \\ x_1 + 7.x_2 = 4 \end{cases}$$

$$\text{b)} \begin{cases} 3.x_1 - 2.x_2 + x_3 = 0 \\ -x_1 + x_2 + 2.x_3 = 1 \\ 2.x_1 + x_2 + 4.x_3 = 2 \end{cases}$$

$$\text{c)} \left\{ \begin{array}{rcl} x_1 + x_2 + x_3 + x_4 & = & 0 \\ -x_1 + 2.x_2 - 4.x_3 + x_4 & = & 1 \\ x_1 - x_2 - x_3 - x_4 & = & 4 \\ 5.x_1 + x_2 - 3.x_3 + 2.x_4 & = & 0 \end{array} \right.$$

ii) Resolver el siguiente sistema lineal sobre $\mathbb{Z}_7$ empleando la regla de Cramer:

$$\left\{ \begin{array}{rcl} 3x + y + 2z & = & 1 \\ x + z & = & 6 \\ 2x + 2y + z & = & 3 \end{array} \right.$$

Ejercicio 18. Sea $A \in \mathbb{R}^{3\times 3}$ la matriz $A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$. Se sabe que

$$\det \begin{pmatrix} 1 & b & c \\ 2 & e & f \\ 5 & h & i \end{pmatrix} = 0, \quad \det \begin{pmatrix} a & 2 & c \\ d & 4 & f \\ g & 10 & i \end{pmatrix} = 0, \quad \text{y} \quad \det \begin{pmatrix} a & b & -1 \\ d & e & -2 \\ g & h & -5 \end{pmatrix} = 0.$$

Calcular $\det A$.

Ejercicio 19.

i) Sea $A \in K^{3\times 3}$ no inversible tal que $A_{11}.A_{33} - A_{13}.A_{31} \neq 0$. Calcular la dimensión de $S = \{x \in K^3 / \, A.x = 0\}$.

ii) Sea $A \in K^{n\times n}$ no inversible tal que $\mathrm{adj}(A) \neq 0$. Calcular $\mathrm{rg}(A)$ y $\mathrm{rg}(\mathrm{adj}(A))$.

Ejercicio 20.

i) Sea $A = (a_{ij}) \in K^{6\times 6}$. ¿Con qué signos aparecen los siguientes productos en $\det(A)$?

   a) $a_{23}.a_{31}.a_{42}.a_{56}.a_{14}.a_{65}$

   b) $a_{32}.a_{43}.a_{14}.a_{51}.a_{66}.a_{25}$

ii) Sea $A = (a_{ij}) \in K^{5\times 5}$. Elegir todos los posibles valores de j y de k tales que el producto $a_{1j}.a_{32}.a_{4k}.a_{25}.a_{53}$ aparezca en $\det(A)$ con signo $+$

iii) Sea $A = (a_{ij}) \in K^{4\times 4}$. Escribir todos los términos de $\det(A)$ que tengan al factor a_{23} y signo $+$

iv) Sin calcular el determinante, calcular los coeficientes de X^4 y de X^3 en

$$\det\begin{pmatrix} 2.X & X & 1 & 2 \\ 1 & X & 1 & -1 \\ 3 & 2 & X & 1 \\ 1 & 1 & 1 & X \end{pmatrix}.$$

v) Sin calcular el determinante, calcular el coeficiente de a^6 y el de b^6 en

$$\det\begin{pmatrix} 1 & b & a & 1 & 1 & a \\ 1 & 1 & b & 1 & a & 1 \\ 1 & 1 & 1 & a & b & 1 \\ a & 1 & 1 & 1 & 1 & b \\ 1 & 1 & a & b & 1 & a \\ b & a & 1 & 1 & 1 & 1 \end{pmatrix}.$$

Ejercicio 21. Sean A, B, C, $D \in K^{n\times n}$. Sea $M \in K^{2n\times 2n}$ la matriz de bloques

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}.$$

Probar que si $A \in GL(n, K)$, $\det(M) = \det(A.D - A.C.A^{-1}.B)$. Si además $A.C = C.A$ entonces $\det(M) = \det(A.D - C.B)$.

Capítulo 6

Diagonalización

En este capítulo empezaremos a estudiar la estructura de los endomorfismos de un espacio vectorial de dimensión finita.

6.1 Nociones básicas

Dada una transformación lineal $f : K^n \to K^n$, y dos bases B_1 y B_2 de K^n se tiene que

$$|f|_{B_1} = C(B_2, B_1)|f|_{B_2} C(B_1, B_2) = C(B_2, B_1)|f|_{B_2} C(B_2, B_1)^{-1},$$

y por lo tanto, existe una matriz $C \in GL(n, K)$ tal que $|f|_{B_1} = C. |f|_{B_2}. C^{-1}$. Recíprocamente, si $A, B \in K^{n\times n}$ son tales que existe una matriz $C \in GL(n, K)$ tal que $A = C.B.C^{-1}$, definiendo $f : K^n \to K^n$ como $f(x) = A.x$ y considerando $B_1 = E$ la base canónica de K^n y B_2 la base de K^n formada por las columnas de C, resulta que

$$A = |f|_{B_1} \quad \text{y} \quad B = C^{-1}.A.C = C(E, B_2)|f|_E\, C(B_2, E) = |f|_{B_2}.$$

Esto da lugar a la siguiente definición (ya introducida en el Ejercicio 35 de la Sección 3.8):

Definición 6.1 Sean $A, B \in K^{n\times n}$. Se dice que A y B son *semejantes*, y se nota $A \sim B$, si existe una matriz $C \in GL(n, K)$ tal que $A = C.B.C^{-1}$.

Por lo tanto, se demostró la siguiente propiedad (que es lo propuesto por el Ejercicio 36 de la Sección 3.8):

Proposición 6.2 Sean $A, B \in K^{n\times n}$. Entonces $A \sim B$ si y sólo si existen una transformación lineal $f : K^n \to K^n$ y bases B_1 y B_2 de K^n tales que $|f|_{B_1} = A$ y $|f|_{B_2} = B$.

Por lo que vimos, una misma transformación lineal da lugar a matrices semejantes si calculamos sus matrices en distintas bases. Es por eso que, en lo que sigue, estudiaremos la semejanza de matrices. El primer problema que consideraremos es el de determinar si una matriz es semejante a una matriz diagonal.

Definición 6.3 Una matriz $A \in K^{n\times n}$ se dice *diagonalizable* si existe una matriz $C \in GL(n, K)$ tal que $C.A.C^{-1}$ es una matriz diagonal.

En otras palabras, una matriz diagonalizable es una matriz que es semejante a una matriz diagonal. La noción correspondiente para transformaciones lineales es la siguiente:

Definición 6.4 Sea V un K-espacio vectorial de dimensión finita, y sea $f : V \to V$ una transformación lineal. Se dice que f es *diagonalizable* o *diagonal* si existe una base B de V tal que $|f|_B$ es diagonal.

Teniendo en cuenta que la semejanza de matrices es una relación de equivalencia (ver Ejercicio 35 de la Sección 3.8) deducimos que:

Observación 6.5 Sea V un K-espacio vectorial de dimensión finita y sea $f : V \to V$ una transformación lineal. Entonces f es diagonalizable si y sólo si $|f|_B$ es diagonalizable para toda base B de V.

6.1.1 Autovalores y autovectores

Sea V un K-espacio vectorial de dimensión n y sea $f : V \to V$ una transformación lineal diagonalizable. Luego, existe una base $B = \{v_1, \dots, v_n\}$ de V tal que $|f|_B$ es diagonal:

$$|f|_B = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix}.$$

Entonces, para cada $1 \le i \le n$, $f(v_i) = \lambda_i v_i$.

Recíprocamente, si para una base $B = \{v_1, \dots, v_n\}$ de V y $\lambda_1, \dots, \lambda_n \in K$ se cumple que $f(v_i) = \lambda_i v_i$ para cada $1 \le i \le n$, la matriz $|f|_B$ es diagonal y, en consecuencia, f es diagonalizable.

Esto nos lleva a la siguiente definición:

Definición 6.6 Sea V un K-espacio vectorial, y sea $f : V \to V$ una transformación lineal. Se dice que $v \in V$, $v \neq 0$, es un *autovector* de f si existe $\lambda \in K$ tal que $f(v) = \lambda.v$. El elemento $\lambda \in K$ se llama un *autovalor* de f.

Usando estas definiciones, el razonamiento anterior se puede reescribir de esta forma:

Proposición 6.7 *Sea V un K-espacio vectorial de dimensión n y sea $f : V \to V$ una transformación lineal. Entonces f es diagonalizable si y sólo si existe una base B de V formada por autovectores de f.*

La mismas nociones se pueden definir para matrices: Dada $A \in K^{n\times n}$, se le puede asociar una transformación lineal $f_A : K^n \to K^n$ definida por $f_A(x) = A.x$. Notar que $|f_A|_E = A$, donde E es la base canónica de K^n. Entonces $v \in K^n$, $v \neq 0$, es un autovector de f_A de autovalor λ si y sólo si $A.v = \lambda.v$.

Definición 6.8 Sea $A \in K^{n\times n}$. Se dice que $v \in K^n$, $v \neq 0$, es un *autovector* de A si existe $\lambda \in K$ tal que $A.v = \lambda.v$. El elemento $\lambda \in K$ que verifica la condición anterior se llama un *autovalor* de A.

Podemos dar también un resultado análogo a la Proposición 6.7 para matrices:

Proposición 6.9 *Sea $A \in K^{n\times n}$. Entonces A es diagonalizable si y sólo si existe una base B de K^n formada por autovectores de A.*

Ejemplos.

1. Decidir si $A = \begin{pmatrix} 2 & 3 \\ 2 & 1 \end{pmatrix} \in \mathbb{R}^{2\times 2}$ es diagonalizable.

   En virtud de la proposición anterior, basta buscar los autovectores de A, es decir, los vectores $x = (x_1, x_2) \in \mathbb{R}^2$ tales que $(x_1, x_2) \neq (0, 0)$ y $A.\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \lambda.x_1 \\ \lambda.x_2 \end{pmatrix}$ para algún $\lambda \in \mathbb{R}$.

   Para esto, buscaremos en primer término los elementos $\lambda \in \mathbb{R}$ para los cuales el sistema $A.x = \lambda.x$ tiene solución no trivial (autovalores de A) y después, para cada uno de los valores hallados, los vectores $(x_1, x_2) \in \mathbb{R}^2$ que son soluciones del sistema lineal correspondiente.

   Observamos que
   $$A.x = \lambda.x \iff (\lambda I_2 - A).x = 0 \iff \begin{pmatrix} \lambda - 2 & -3 \\ -2 & \lambda - 1 \end{pmatrix}.\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

   Este sistema homogéneo tiene solución no trivial si y sólo si el determinante de su matriz asociada es 0, o sea, si y sólo si $\lambda^2 - 3\lambda - 4 = 0$. Luego, los autovalores de A son $\lambda = -1$ y $\lambda = 4$.

   Busquemos ahora los autovectores correspondientes a cada autovalor:

   Para $\lambda = -1$, queda el sistema
   $$\begin{pmatrix} -3 & -3 \\ -2 & -2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$
   cuyo conjunto de soluciones es $< (1, -1) >$. Luego el conjunto de los autovectores asociados a $\lambda = -1$ es $< (1, -1) > - \{(0, 0)\}$.

   Para $\lambda = 4$, el sistema es
   $$\begin{pmatrix} 2 & -3 \\ -2 & 3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

cuyo conjunto de soluciones es $< (3,2) >$. Luego el conjunto de los autovectores asociados a $\lambda = 4$ es $< (3,2) > - \{(0,0)\}$.

En consecuencia, A es diagonalizable, puesto que $B = \{(1,-1),(3,2)\}$ es una base de $\mathbb{R}^2$ formada por autovectores de A. Más aún, si $C = C(E,B)$ se tiene que

$$C.\,A.\,C^{-1} = \begin{pmatrix} -1 & 0 \\ 0 & 4 \end{pmatrix}.$$

2. Decidir si $A = \begin{pmatrix} 3 & 0 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}$ es diagonalizable en $\mathbb{R}^{3\times 3}$.

   Busquemos los autovalores de A, es decir, los valores de $\lambda \in \mathbb{R}$ para los cuales el sistema $A.x = \lambda.x$ tiene solución no trivial o, equivalentemente, el sistema $(\lambda.I_3 - A).x = 0$ tiene solución no trivial. Pero esto vale si y sólo si $\det(\lambda.I_3 - A) = 0$, es decir $(\lambda - 3)^3 = 0$. Luego, $\lambda = 3$ es el único autovalor de A.

   Si A fuese diagonalizable, existiría $C \in GL(n,K)$ tal que

$$C.A.C^{-1} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix} \iff A = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

   Luego, A no es diagonalizable.

6.1.2 Polinomio característico

Como vimos en la sección anterior, un método para determinar si una matriz es diagonalizable consiste en buscar sus autovalores y luego ver si se puede armar una base de autovectores.

Sea $A \in K^{n\times n}$ y sea $\lambda \in K$. Se tiene que:

$$\begin{array}{lcl} \lambda \text{ es autovalor de } A & \iff & \exists\, x \in K^n - \{0\} \text{ tal que } A.x = \lambda.x. \\ & \iff & \text{El sistema } A.x = \lambda.x \text{ tiene solución no trivial.} \\ & \iff & \text{El sistema } (\lambda.I_n - A).x = 0 \text{ tiene solución no trivial.} \\ & \iff & \det(\lambda.I_n - A) = 0. \end{array}$$

(Comparar con el Ejercicio 13 de la Sección 5.7.)

Definición 6.10 Sea $A \in K^{n\times n}$. Se llama *polinomio característico de* A, y se nota $\mathcal{X}_A$, al polinomio $\mathcal{X}_A = \det(X.I_n - A) \in K[X]$.

Si $A \in K^{n\times n}$, $\mathcal{X}_A$ resulta ser un polinomio mónico de grado n (notar que en la matriz $X.I_n - A$, sólo aparece n veces X y que el signo del término $(X - a_{11})\dots(X - a_{nn})$ en el determinante es 1). Por lo anterior, tenemos:

Proposición 6.11 *Sea $A \in K^{n\times n}$ y sea $\lambda \in K$. Entonces λ es autovalor de A si y sólo si λ es raíz del polinomio característico de A.*

Ejemplo. Decidir si $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ es diagonalizable en $\mathbb{Q}^{2\times 2}, \mathbb{R}^{2\times 2}$ y $\mathbb{C}^{2\times 2}$.

Los autovalores de A son las raíces del polinomio

$$\mathcal{X}_A = \det \begin{pmatrix} X & -1 \\ 1 & X \end{pmatrix} = X^2 + 1.$$

Como este polinomio no tiene raíces en $\mathbb{Q}$ ni en $\mathbb{R}$, resulta que A no es diagonalizable en $\mathbb{Q}^{2\times 2}$ ni en $\mathbb{R}^{2\times 2}$.

Considerada como matriz en $\mathbb{C}^{2\times 2}$, los autovalores de A son i y $-i$, y los autovectores asociados son $< (1, i) > - \{(0, 0)\}$ y $< (-1, i) > - \{(0, 0)\}$. Como $B = \{(1, i), (-1, i)\}$ es una base de $\mathbb{C}^2$ formada por autovectores de A, entonces A es diagonalizable en $\mathbb{C}^{2\times 2}$.

Queremos definir el polinomio característico asociado a un endomorfismo de un espacio vectorial de dimensión finita. Para eso, veremos que dos matrices semejantes tienen el mismo polinomio característico.

Proposición 6.12 *Sea $A \in K^{n\times n}$ y sea $C \in GL(n, K)$. Entonces $\mathcal{X}_{C.A.C^{-1}} = \mathcal{X}_A$.*

Demostración. Se tiene que

$$\begin{aligned} \mathcal{X}_{C.A.C^{-1}} &= \det(X.I_n - C.A.C^{-1}) = \det(C.X.I_n.C^{-1} - C.A.C^{-1}) \\ &= \det(C.(X.I_n - A).C^{-1}) = \det(X.I_n - A) = \mathcal{X}_A. \end{aligned}$$ □

Definición 6.13 Sea V un K-espacio vectorial de dimensión finita, y sea $f : V \to V$ una transformación lineal. Se define el *polinomio característico* de f como $\mathcal{X}_f = \mathcal{X}_{|f|_B}$, donde B es una base cualquiera de V.

Como en el caso de matrices, se tiene que:

Observación 6.14 Sea V un K-espacio vectorial de dimensión finita. Sea $f : V \to V$ una transformación lineal y sea $\lambda \in K$. Entonces λ es autovalor de f si y sólo si λ es raíz de $\mathcal{X}_f$.

6.2 Una caracterización de matrices diagonalizables

6.2.1 Suma directa de subespacios

Para lo que sigue, vamos a necesitar generalizar la noción de suma directa de dos subespacios de un K-espacio vectorial (ver Sección 1.4.2) al caso de cualquier cantidad finita de subespacios.

Definición 6.15 Sea V un K-espacio vectorial y sean $S_1, S_2, \ldots, S_r$ subespacios de V. Se define la *suma* de $S_1, S_2, \ldots, S_r$ como

$$W = S_1 + S_2 + \cdots + S_r = \{s_1 + \cdots + s_r \,/\, s_i \in S_i,\ 1 \le i \le r\}.$$

Es fácil ver que W es un subespacio de V.

Definición 6.16 Sea V un K-espacio vectorial y sean $S_1, S_2, \ldots, S_r$ subespacios de V. Se dice que $S_1, S_2, \ldots, S_r$ están en *suma directa* si, para cada $w \in W = S_1 + S_2 + \cdots + S_r$ existen *únicos* $s_i \in S_i$, $1 \le i \le r$, tales que $w = s_1 + \ldots + s_r$. En este caso se dice que W es la *suma directa de los subespacios* $S_1, \ldots, S_r$, y se nota

$$W = S_1 \oplus S_2 \oplus \cdots \oplus S_r = \bigoplus_{i=1}^{r} S_i.$$

Vamos a dar una definición equivalente de la suma directa de varios subespacios:

Proposición 6.17 *Sea V un K-espacio vectorial y sean $S_1, S_2, \ldots, S_r$ subespacios de V. Son equivalentes:*

i) $W = \bigoplus\limits_{i=1}^{r} S_i$.

ii) $W = S_1 + \ldots + S_r$ *y para cada* $1 \le j \le r$, *vale*

$$S_j \ \cap\ (S_1 + S_2 + \cdots + S_{j-1} + S_{j+1} + \cdots + S_r) = \{0\}.$$

Demostración.

i) $\Rightarrow$ *ii)* Sea $1 \le j \le r$. Sea $x \in S_j \cap (S_1 + S_2 + \cdots + S_{j-1} + S_{j+1} + \cdots + S_r)$. Entonces

$$\begin{aligned} x &= 0 + \cdots + 0 + x + 0 + \cdots + 0, \\ x &= s_1 + \cdots + s_{j-1} + 0 + s_{j+1} + \cdots + s_r. \end{aligned}$$

Por la unicidad de la escritura en la suma directa, resulta que $x = 0$.

ii) $\Rightarrow$ *i)* Por hipótesis, existen $s_1, \ldots, s_r$ con $s_i \in S_i$ para cada $1 \le i \le r$ tales que $w = \sum\limits_{i=1}^{r} s_i$.

Supongamos que $w = \sum\limits_{i=1}^{r} s_i = \sum\limits_{i=1}^{r} s_i'$ con $s_i, s_i' \in S_i$ para cada $1 \le i \le r$.

Entonces, para cada $1 \le j \le r$, se tiene que

$$s_j - s_j' = \sum_{\substack{i=1 \\ i \ne j}}^{r} (s_i' - s_i).$$

Como $s_j - s_j' \in S_j$ y $\sum\limits_{\substack{i=1 \\ i \ne j}}^{r} (s_i' - s_i)$ es una suma donde cada $s_i' - s_i \in S_i$, de la hipótesis se deduce que $s_j - s_j' = 0$. Luego, $s_j = s_j'$. □

Como en la suma directa de dos subespacios, en este caso también es posible obtener una base del espacio suma uniendo bases de cada uno de los sumandos. La demostración de este resultado es análoga a la de la Proposición 1.46.

Proposición 6.18 *Sea V un K-espacio vectorial y sean $S_1, S_2, \dots, S_r$ subespacios de V. Para cada $1 \le i \le r$, sea B_i una base de S_i. Son equivalentes:*

i) $W = \bigoplus\limits_{i=1}^{r} S_i$.

ii) $B = B_1 \cup B_2 \cup \dots \cup B_r$ *es una base de* W.

Observamos que en la condición ii), B es la *familia* obtenida mediante la unión de las familias $B_1, B_2, \dots, B_r$.

6.2.2 Espacios de autovectores y diagonalización

Dado un autovalor λ de una matriz $A \in K^{n\times n}$, el conjunto de los autovectores de autovalor λ no es un subespacio de K^n, puesto que, por definición, 0 no es un autovector de A. Sin embargo, podemos considerar el siguiente subespacio que consiste en agregar el vector 0 a ese conjunto:

Definición 6.19 Sea $A \in K^{n\times n}$ y sea λ un autovalor de A. Se define

$$E_\lambda = \{v \in K^n \,/\, A.v = \lambda.v\} = \{v \in K^n \,/\, (\lambda\, I_n - A).v = 0\}.$$

Observamos que E_λ es un subespacio de K^n, puesto que es el conjunto de soluciones de un sistema lineal homogéneo.

Proposición 6.20 *Sea $A \in K^{n\times n}$ y sean $\lambda_1, \dots, \lambda_r$ autovalores distintos de A. Entonces $E_{\lambda_1}, \dots, E_{\lambda_r}$ están en suma directa.*

Demostración. Lo probaremos por inducción en la cantidad r de autovalores considerados.

Para $r = 2$: Sean λ_1 y λ_2 autovalores distintos de A. Si $v \in E_{\lambda_1} \cap E_{\lambda_2}$, se tiene que $A.v = \lambda_1.v$ y $A.v = \lambda_2.v$, de donde $(\lambda_1 - \lambda_2).v = 0$. Como $\lambda_1 - \lambda_2 \neq 0$, resulta que $v = 0$. Luego, $E_{\lambda_1} \cap E_{\lambda_2} = \{0\}$ y la suma es directa.

Supongamos ahora que el resultado vale para el caso de r autovalores distintos, y sean $\lambda_1, \dots, \lambda_r, \lambda_{r+1}$ autovalores distintos de A.

Debemos probar que para cada $1 \le i \le r+1$, $E_{\lambda_i} \cap \bigoplus\limits_{\substack{j=1\\ j\neq i}}^{r+1} E_{\lambda_j} = \{0\}$. Sin pérdida de generalidad, supongamos que $i = r+1$.

Sea $v \in E_{\lambda_{r+1}} \cap \bigoplus\limits_{j=1}^{r} E_{\lambda_j}$. Entonces, existen $v_j \in E_{\lambda_j} (1 \le j \le r)$ tales que $v = v_1 + \ldots + v_r$. Multiplicando esta igualdad por la matriz A, tenemos

$$\lambda_{r+1} v = \lambda_1 v_1 + \ldots + \lambda_r v_r,$$

pero si la multiplicamos por λ_{r+1}, se tiene

$$\lambda_{r+1}v = \lambda_{r+1}v_1 + \ldots + \lambda_{r+1}v_r.$$

Restando miembro a miembro,

$$0 = (\lambda_1 - \lambda_{r+1})v_1 + \ldots + (\lambda_r - \lambda_{r+1})v_r.$$

Como por hipótesis inductiva, los subespacios $E_{\lambda_j}(1 \le j \le r)$ están en suma directa, el vector cero se escribe de forma única como suma de ceros. Luego, $(\lambda_j - \lambda_{r+1})v_j = 0$ para cada $1 \le j \le r$ y, por lo tanto, $v_j = 0$ para cada $1 \le j \le r$, con lo cual $v = 0$. $\square$

Ya vimos que todo autovalor λ de una matriz A es raíz de su polinomio característico. Ahora veremos que siempre existe una relación entre la dimensión de E_λ y la multiplicidad de λ como raíz de $\mathcal{X}_A$.

Proposición 6.21 *Sea $A \in K^{n\times n}$ y sea $\lambda \in K$. Sea r la multiplicidad de λ como raíz de $\mathcal{X}_A$ (es decir, $\mathcal{X}_A = (X-\lambda)^r P$ con $P(\lambda) \neq 0$) y sea $E_\lambda = \{x \in K^n \,/\, A.x = \lambda x\}$. Entonces* $\dim(E_\lambda) \le r$.

Demostración. Sea $f_A : K^n \to K^n$ la transformación lineal definida por $f_A(x) = A.x$. Sea $s = \dim(E_\lambda)$ y sea $\{v_1, \ldots, v_s\}$ una base de E_λ. Sean $v_{s+1}, \ldots, v_n \in K^n$ tales que $B = \{v_1, \ldots, v_s, v_{s+1}, \ldots, v_n\}$ es una base de K^n. Se tiene que

$$|f_A|_B = \left(\begin{array}{cc} \overbrace{\begin{matrix} \lambda & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda \end{matrix}}^{s\times s} & N \\ \mathbf{0} & M \end{array} \right),$$

de donde

$$\begin{aligned} \mathcal{X}_{f_A} &= \det \left(\begin{array}{cc} \overbrace{\begin{matrix} X-\lambda & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & X-\lambda \end{matrix}}^{s\times s} & -N \\ \mathbf{0} & XI_{n-s} - M \end{array} \right) \\ &= (X-\lambda)^s \det(XI_{n-s} - M) \\ &= (X-\lambda)^s Q. \end{aligned}$$

Por hipótesis, $\mathcal{X}_A = (X-\lambda)^r P$ con $P \in K[X]$ tal que $P(\lambda) \neq 0$. Entonces

$$(X-\lambda)^s Q = \mathcal{X}_{f_A} = \mathcal{X}_A = (X-\lambda)^r P \quad \text{ con } P(\lambda) \neq 0,$$

de donde $s \le r$, es decir, $\dim(E_\lambda) \le \text{mult}(\lambda, \mathcal{X}_A)$. □

El siguiente teorema establece condiciones necesarias y suficientes sobre los subespacios E_λ asociados a los autovalores de una matriz para que ésta sea diagonalizable.

Teorema 6.22 *Sea $A \in K^{n\times n}$ y sean $\lambda_1, \dots, \lambda_r$ todos los autovalores de A en K, con $\lambda_i \neq \lambda_j$ si $i \neq j$. Son equivalentes:*

i) *A es diagonalizable en $K^{n\times n}$.*

ii) $\bigoplus_{i=1}^{r} E_{\lambda_i} = K^n$.

iii) $\mathcal{X}_A = (X-\lambda_1)^{a_1} \dots (X-\lambda_r)^{a_r}$ *y* $a_i = \dim E_{\lambda_i}$ *para cada* $1 \le i \le r$.

Demostración.

i) $\Rightarrow$ *ii)* Si A es diagonalizable en $K^{n\times n}$, existe una base $B = \{v_1, \dots, v_n\}$ de K^n formada por autovectores de A. Para cada $v_j \in B$, existe i, $1 \le i \le r$, tal que v_j es autovector de A de autovalor λ_i (porque $\lambda_1, \dots, \lambda_r$ son *todos* los autovalores de A), es decir $v_j \in E_{\lambda_i}$ para algún $1 \le i \le r$, lo que implica que $v_j \in \bigoplus_{i=1}^{r} E_{\lambda_i}$.

En consecuencia, $K^n = < v_1, \dots, v_n > = \bigoplus_{i=1}^{r} E_{\lambda_i}$.

ii) $\Rightarrow$ *iii)* Por la proposición anterior, para cada $1 \le i \le r$, $\dim E_{\lambda_i} \le \text{mult}(\lambda_i, \mathcal{X}_A)$. Si $K^n = \bigoplus_{i=1}^{r} E_{\lambda_i}$ se tiene que

$$n = \dim K^n = \sum_{i=1}^{r} \dim E_{\lambda_i} \le \sum_{i=1}^{r} \text{mult}(\lambda_i, \mathcal{X}_A) \le \text{gr}(\mathcal{X}_A) = n.$$

Luego, en la cadena anterior, son todas igualdades.

En particular:

- La igualdad $\sum_{i=1}^{r} \text{mult}(\lambda_i, \mathcal{X}_A) = \text{gr}(\mathcal{X}_A)$ implica que $\mathcal{X}_A$ se puede factorizar como producto de polinomios de grado 1 en $K[X]$: si $a_i = \text{mult}(\lambda_i, \mathcal{X}_A)$ $(1 \le i \le r)$, entonces $\mathcal{X}_A = (X-\lambda_1)^{a_1} \dots (X-\lambda_r)^{a_r}$.

- Como $\dim E_{\lambda_i} \le \text{mult}(\lambda_i, \mathcal{X}_A)$ para cada $1 \le i \le r$, de la igualdad $\sum_{i=1}^{r} \dim E_{\lambda_i} = \sum_{i=1}^{r} \text{mult}(\lambda_i, \mathcal{X}_A)$ se deduce que $\dim E_{\lambda_i} = \text{mult}(\lambda_i, \mathcal{X}_A)$ para cada $1 \le i \le r$.

iii) $\Rightarrow$ *i)* Para cada $1 \leq i \leq r$ sea B_i una base de E_{λ_i}. Por la Proposición 6.18, $B = \bigcup_{i=1}^{r} B_i$ es una base de $\bigoplus_{i=1}^{r} E_{\lambda_i} \subseteq K^n$. Ahora,

$$\#B = \sum_{i=1}^{r} \#B_i = \sum_{i=1}^{r} \dim E_{\lambda_i} = \sum_{i=1}^{r} a_i = \mathrm{gr}(\mathcal{X}_A) = n,$$

de donde $\dim \left(\bigoplus_{i=1}^{r} E_{\lambda_i} \right) = \dim K^n$. Luego $\bigoplus_{i=1}^{r} E_{\lambda_i} = K^n$ y entonces B es una base (formada por autovectores de A) de K^n. En consecuencia, A es diagonalizable. $\square$

Ejemplo. Decidir si $A = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix} \in \mathbb{R}^{4\times 4}$ es diagonalizable.

Calculamos $\mathcal{X}_A = (X-1)^3(X-2)$. Para el autovalor 1 se tiene que

$$E_1 = \{x \in \mathbb{R}^4 \,/\, (I_4 - A).x = 0\} = <(0,1,0,0), (0,0,1,0)>,$$

de donde $\dim E_1 = 2 < 3 = \mathrm{mult}(1, \mathcal{X}_A)$. El teorema anterior implica entonces que A no es diagonalizable.

6.3 Polinomios minimales

En lo que sigue, a cada matriz le asociaremos un polinomio. Este polinomio, entre otras propiedades, nos dará un nuevo criterio para decidir si la matriz es diagonalizable.

6.3.1 Polinomio minimal de una matriz

Sea $P \in K[X]$, $P = a_0 + a_1 X + \cdots + a_r X^r$. Dada $A \in K^{n\times n}$ definimos

$$P(A) = a_0 I_n + a_1 A + \cdots + a_r A^r \;\in\; K^{n\times n}.$$

Observemos que si $P, Q \in K[X]$ y $A \in K^{n\times n}$, entonces $(P+Q)(A) = P(A) + Q(A)$ y $(P.Q)(A) = P(A).Q(A)$.

Análogamente, si V es un K-espacio vectorial de dimensión finita y $f : V \to V$ es una transformación lineal, definimos

$$P(f) = a_0 id_V + a_1 f + \cdots + a_r f^r \;\in\; \mathrm{Hom}_K(V,V),$$

donde, para $k \in \mathbb{N}$, $f^k = \underbrace{f \circ f \circ \ldots \circ f}_{k \text{ veces}}$ es la composición.

Dada una matriz $A \in K^{n\times n}$ nos interesa considerar polinomios $P \in K[X]$ que anulen a A, es decir, tales que $P(A) = 0$. El siguiente resultado asegura que para cualquier matriz existe un polinomio no nulo con esta propiedad.

Lema 6.23 *Sea $A \in K^{n\times n}$. Existe un polinomio $P \in K[X]$, $P \neq 0$, tal que $P(A) = 0$.*

Demostración. Consideremos el conjunto $\{I_n, A, A^2, \ldots, A^{n^2}\} \subseteq K^{n\times n}$. Este conjunto es linealmente dependiente, puesto que tiene $n^2 + 1$ elementos y $\dim(K^{n\times n}) = n^2$. Luego, existen $a_0, a_1, \ldots, a_{n^2} \in K$ no todos nulos, tales que $\sum_{i=0}^{n^2} a_i A^i = 0$. Sea $P \in K[X]$ el polinomio $P = \sum_{i=0}^{n^2} a_i X^i$. Entonces $P \neq 0$ y $P(A) = 0$. □

Para cada matriz, distinguimos un polinomio particular entre todos los polinomios que la anulan: el de grado mínimo y mónico. Veamos que para toda matriz existe un polinomio con estas propiedades y que es único:

Existencia. Sea $H = \{\text{gr}(P) : P \in K[X],\ P \neq 0,\ P(A) = 0\} \subseteq \mathbb{N}$. Por el Lema 6.23, $H \neq \emptyset$. Luego, H tiene primer elemento r. Entonces existe un polinomio $Q \neq 0$ de grado r tal que $Q(A) = 0$ y Q es mónico (si no lo fuera, bastaría dividir por su coeficiente principal).

Unicidad. Supongamos que existe un polinomio $Q' \in K[X]$ mónico, $Q' \neq Q$, tal que $\text{gr}(Q') = r$ y $Q'(A) = 0$. Entonces $(Q' - Q)(A) = 0$ y $\text{gr}(Q' - Q) < r$, puesto que Q y Q' son ambos mónicos, lo que contradice que r es el primer elemento de H.

Definición 6.24 Sea $A \in K^{n\times n}$. Llamaremos *polinomio minimal* de A, y lo notaremos m_A, al polinomio mónico de grado mínimo que anula a A.

Ejemplo. Sea $A = \begin{pmatrix} -1 & 0 \\ 1 & -1 \end{pmatrix}$. Calcular m_A.

Puesto que $\{I_2, A\}$ es linealmente independiente, no existe $P \in \mathbb{R}[X]$ tal que $\text{gr}(P) = 1$ y $P(A) = 0$.

Buscamos entonces $P = X^2 + aX + b$ que anule a A, para lo cual estudiamos la independecia lineal de $\{I_2, A, A^2\}$.

$$\begin{aligned} A^2 + aA + bI_2 = 0 \quad &\Longleftrightarrow \quad \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix} + a \begin{pmatrix} -1 & 0 \\ 1 & -1 \end{pmatrix} + b \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 0 \\ &\Longleftrightarrow \quad \begin{cases} 1 - a + b = 0 \\ -2 + a = 0 \end{cases} \quad \Longleftrightarrow \quad a = 2,\ b = 1. \end{aligned}$$

Luego, $m_A = X^2 + 2X + 1$. (Observar que, en este caso, m_A coincide con $\mathcal{X}_A$.)

Dada una matriz $A \in K^{n\times n}$, el conjunto de todos los polinomios que anulan a A puede caracterizarse a partir de su polinomio minimal m_A.

Proposición 6.25 *Sea* $A \in K^{n\times n}$ *y sea* $P \in K[X]$. *Entonces* $P(A) = 0$ *si y sólo si* m_A *divide a* P.

Demostración.

($\Leftarrow$) Si $m_A \mid P$, existe $Q \in K[X]$ tal que $P = Q.\, m_A$. Luego

$$P(A) = (Q.\, m_A)(A) = Q(A).\, m_A(A) = Q(A).\, 0 = 0.$$

($\Rightarrow$) Por el algoritmo de división en $K[X]$, existen polinomios $Q, R \in K[X]$ tales que $P = Q.\, m_A + R$ con $R = 0$ o $\mathrm{gr}(R) < \mathrm{gr}(m_A)$. Se tiene que

$$0 = P(A) = Q(A).\, m_A(A) + R(A) = Q(A).\, 0 + R(A) = R(A).$$

Como m_A es de grado mínimo entre los polinomios que anulan a A, no puede ser $\mathrm{gr}(R) < \mathrm{gr}(m_A)$ y, por lo tanto, $R = 0$. □

Como sucede en el caso del polinomio característico, vemos que dos matrices semejantes tienen el mismo polinomio minimal. Para eso, analizamos en primer lugar la relación entre los polinomios que anulan a dos matrices semejantes. Utilizaremos el siguiente hecho:

Observación 6.26 Si $A, B \in K^{n\times n}$ y $C \in GL(n, K)$ son matrices tales que $A = C.\, B.\, C^{-1}$ entonces $A^k = C.\, B^k.\, C^{-1}$ para todo $k \in \mathbb{N}$.

En efecto, es claro que la igualdad vale para $k = 1$ y, que si vale para $k \in \mathbb{N}$, entonces $A^{k+1} = A.\, A^k = C.\, B.\, C^{-1}.\, C.\, B^k.\, C^{-1} = C.\, B^{k+1}.\, C^{-1}$.

Lema 6.27 *Sean* $A, B \in K^{n\times n}$, $A \sim B$. *Entonces* $P(A) \sim P(B)$ *para todo* $P \in K[X]$. *En particular,* $P(A) = 0$ *si y sólo si* $P(B) = 0$.

Demostración. Sea $P \in K[X]$, $P = \sum\limits_{i=0}^{r} a_i X^i$. Si $A \sim B$, existe $C \in GL(n, K)$ tal que $A = C.B.C^{-1}$, y entonces

$$\begin{aligned} P(A) &= P(C.\, B.\, C^{-1}) = \sum_{i=0}^{r} a_i.(C.\, B.\, C^{-1})^i \\ &= \sum_{i=0}^{r} a_i.\, C.\, B^i.\, C^{-1} = C.\Big(\sum_{i=0}^{r} a_i B^i\Big).\, C^{-1} = C.\, P(B).\, C^{-1}. \end{aligned}$$

Luego, $P(A) \sim P(B)$. □

Proposición 6.28 *Sean* $A, B \in K^{n\times n}$, $A \sim B$. *Entonces* $m_A = m_B$.

Demostración. Por el lema anterior y la Proposición 6.25,

- $m_A(A) = 0 \Rightarrow m_A(B) = 0 \Rightarrow m_B \mid m_A$.
- $m_B(B) = 0 \Rightarrow m_B(A) = 0 \Rightarrow m_A \mid m_B$.

Puesto que m_A y m_B son ambos mónicos, resulta que $m_A = m_B$. □

Este resultado implica que si $f : V \to V$ es una transformación lineal definida en un K-espacio vectorial V de dimensión finita, y consideramos la matriz de f en dos bases de V distintas, los polinomios minimales de estas dos matrices coinciden. Esto nos permite dar la siguiente definición de polinomio minimal para transformaciones lineales:

Definición 6.29 Sea V un K-espacio vectorial de dimensión finita, y sea B una base de V. Sea $f : V \to V$ una transformación lineal. Se define el *polinomio minimal de f* como $m_f = m_{|f|_B}$.

En la Proposición 6.11 vimos que las raíces del polinomio característico de una matriz son sus autovalores. El siguiente resultado muestra que lo mismo vale para el polinomio minimal.

Proposición 6.30 *Sea $A \in K^{n\times n}$, y sea m_A el polinomio minimal de A. Sea $\lambda \in K$. Entonces λ es autovalor de A si y sólo si λ es raíz de m_A.*

Demostración.

($\Rightarrow$) Sea λ un autovalor de A. Por el algoritmo de división en $K[X]$ existen $Q \in K[X]$ y $R \in K$ tales que $m_A = Q.\,(X - \lambda) + R$. Entonces

$$0 = m_A(A) = Q(A).(A - \lambda I_n) + R.\,I_n.$$

Como λ es autovalor de A, existe $v \in K^n$, $v \neq 0$, tal que $A.v = \lambda.v$. Se tiene que

$$0 = Q(A).(A - \lambda I_n).v + R.v = Q(A).(Av - \lambda\, v) + R.v = Q(A).\,0 + R.v = R.v,$$

es decir, $R.v = 0$. Como $v \neq 0$, debe ser $R = 0$.

En consecuencia, $m_A = Q.\,(X - \lambda)$, y entonces λ es raíz de m_A.

($\Leftarrow$) Sea $\lambda \in K$ una raíz de m_A. Entonces $m_A = (X - \lambda).Q$ y, por lo tanto, $0 = m_A(A) = (A - \lambda I_n).Q(A)$.

Observamos que $Q(A) \neq 0$, puesto que $\mathrm{gr}(Q) = \mathrm{gr}(m_A) - 1$. En consecuencia, existe $w \in K^n$ tal que $Q(A).w \neq 0$. Sea $v = Q(A).w$. Entonces

$$(A - \lambda I_n).v = (A - \lambda I_n).Q(A).w = 0.w = 0,$$

de donde λ es un autovalor de A. □

6.3.2 Polinomio minimal de un vector

Sea $A \in K^{n\times n}$ y sea $v \in K^n$. Dado $P \in K[X]$, definimos $P(v) = P(A).v$. Diremos que P anula a v si $P(v) = 0$.

Observación 6.31 Sea $A \in K^{n\times n}$ y sea m_A el polinomio minimal de A. Entonces, para cada $v \in K^n$, se tiene que

$$m_A(v) = m_A(A).v = 0.v = 0.$$

Luego, para cada $v \in K^n$ existe un polinomio mónico que anula a v.

Definición 6.32 Sea $A \in K^{n\times n}$ y sea $v \in K^n$. El *polinomio minimal de* v, que notaremos m_v, es el polinomio mónico de grado mínimo que anula a v.

La existencia y unicidad de un polinomio que cumple las condiciones de la definición se prueban, a partir de la Observación 6.31, en forma análoga a lo hecho para el polinomio minimal de una matriz.

Ejemplos.

1. Sea $A \in K^{n\times n}$ y sea $v \in K^n$ un autovector de A de autovalor λ. Entonces $m_v = X - \lambda$.

2. Sea $A = \begin{pmatrix} -1 & 0 \\ 1 & -1 \end{pmatrix} \in \mathbb{R}^{2\times 2}$, y sea $e_1 = (1, 0)$. Calcular m_{e_1}.

   Comenzamos buscando un polinomio mónico $P = X + b \in \mathbb{R}[X]$ de grado 1 tal que $P(e_1) = 0$: Observamos que

   $$P(e_1) = 0 \iff (A + bI_2).e_1 = 0 \iff A.e_1 + b.e_1 = 0 \iff (-1, 1) + (b, 0) = (0, 0),$$

   pero esto no vale para ningún valor de $b \in \mathbb{R}$. Luego, $\mathrm{gr}(m_{e_1}) \geq 2$.

   Buscamos entonces un polinomio $P \in \mathbb{R}[X]$ mónico de grado 2: $P = X^2 + aX + b$. En este caso

   $$\begin{aligned} P(e_1) = 0 &\iff (A^2 + aA + bI_2).e_1 = 0 \\ &\iff A^2e_1 + a.Ae_1 + b.e_1 = 0 \\ &\iff (1, -2) + a(-1, 1) + b(1, 0) = (0, 0) \\ &\iff \begin{cases} 1 - a + b = 0 \\ -2 + a = 0 \end{cases} \\ &\iff a = 2,\ b = 1. \end{aligned}$$

   Luego, el polinomio minimal de e_1 es $P = X^2 + 2X + 1$.

Cómo hallar el polinomio minimal de un vector:

Sea $A \in K^{n\times n}$ y sea $v \in K^n$. Si $m_v = X^m + a_{m-1}X^{m-1} + \cdots + a_2X^2 + a_1X + a_0$ es el polinomio minimal de v, entonces

$$A^m.v + a_{m-1}A^{m-1}.v + \cdots + a_2A^2.v + a_1A.v + a_0.v = 0.$$

Además, como m_v es de grado mínimo entre los polinomios que satisfacen $P(v) = 0$, se tiene que $\{v, A.v, \ldots, A^{m-1}.v\}$ es un conjunto linealmente independiente.

Luego, para hallar el polinomio minimal de v, un primer paso consiste en buscar el mínimo $m \in \mathbb{N}$ tal que $\{v, A.v, \ldots, A^m.v\}$ es linealmente dependiente. Si $A^m.v = -a_0.v - a_1.A.v - \cdots - a_{m-1}.A^{m-1}.v$, entonces $m_v = X^m + a_{m-1}X^{m-1} + \cdots + a_1X + a_0$.

Al igual que para el polinomio minimal de una matriz (ver Proposición 6.25), fijada una matriz A, todo polinomio que anula a un vector dado resulta ser múltiplo de su polinomio minimal.

Proposición 6.33 *Sean $A \in K^{n\times n}$, $v \in K^n$ y $P \in K[X]$. Entonces $P(v) = 0$ si y sólo si m_v divide a P. En particular, m_v divide a m_A.*

Demostración. Dado $P \in K[X]$ se tiene que $P = Q.\, m_v + R$ con $Q, R \in K[X]$, $R = 0$ o $\mathrm{gr}(R) < \mathrm{gr}(m_v)$. En consecuencia,

$$P(v) = P(A).\, v = Q(A).\, m_v(A).\, v + R(A).\, v = Q(A).\, 0 + R(v) = R(v),$$

de donde $P(v) = 0$ si y sólo si $R(v) = 0$.

Como m_v es de grado mínimo entre los polinomios que anulan a v, resulta que $R(v) = 0$ si y sólo si $R = 0$, es decir, si y sólo si $m_v \mid P$. □

La siguiente proposición muestra cómo calcular el polinomio minimal de una matriz $A \in K^{n\times n}$ a partir de los polinomios minimales de los vectores de una base de K^n.

Proposición 6.34 *Sea $A \in K^{n\times n}$ y sea $B = \{v_1, \ldots, v_n\}$ una base de K^n. Entonces $m_A = \mathrm{mcm}\{m_{v_i} : i = 1, \ldots, n\}$ (mcm denota el mínimo común múltiplo).*

Demostración. Sea $P = \mathrm{mcm}\{m_{v_i} : i = 1, \ldots, n\}$.

Por la proposición anterior, $m_{v_i} \mid m_A$ para cada $1 \le i \le n$. Luego, $P \mid m_A$.

Por otro lado, como $m_{v_i} \mid P$ para cada $1 \le i \le n$, se tiene que $P(A).v_i = 0$ para cada $1 \le i \le n$. Sea $v \in K^n$ y sean $\alpha_1, \ldots, \alpha_n \in K$ tales que $v = \sum\limits_{i=1}^{n} \alpha_i v_i$. Entonces

$$P(A).v = P(A)\Big(\sum_{i=1}^{n} \alpha_i.v_i\Big) = \sum_{i=1}^{n} \alpha_i P(A).v_i = 0.$$

En consecuencia, se tiene que $P(A) \in K^{n\times n}$ satisface $P(A).v = 0 \ \forall\, v \in K^n$ y, por lo tanto, $P(A) = 0$. Luego, $m_A \mid P$.

Como m_A y P son dos polinomios mónicos que se dividen mutuamente, $m_A = P$. $\square$

Ejemplo. Calcular m_A para $A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$.

Consideremos $E = \{e_1, e_2, e_3\}$ la base canónica de $\mathbb{R}^3$. Por la proposición anterior, $m_A = \mathrm{mcm}\{m_{e_1}, m_{e_2}, m_{e_3}\}$.

Busquemos entonces los polinomios minimales de cada uno de los vectores de E:

m_{e_1}: Vemos que $\{e_1, A.e_1\} = \{e_1, e_1 + e_2\}$ es un conjunto linealmente independiente, pero $\{e_1, A.e_1, A^2.e_1\} = \{e_1,\ e_1 + e_2,\ e_1 + 2e_2\}$ no lo es.

Además, $0 = (e_1 + 2e_2) - 2.(e_1 + e_2) + e_1 = A^2.e_1 - 2.A.e_1 + e_1$, de donde $m_{e_1} = X^2 - 2X + 1 = (X-1)^2$.

m_{e_2}: Como $A.e_2 = e_2$, entonces $m_{e_2} = X - 1$.

m_{e_3}: Como $A.e_3 = 2.e_3$, entonces $m_{e_3} = X - 2$.

Luego, $m_A = \mathrm{mcm}\{(X-1)^2, X-1, X-2\} = (X-1)^2(X-2)$.

6.3.3 Teorema de Hamilton-Cayley

El teorema de Hamilton-Cayley, cuya demostración damos a continuación, establece que para cualquier matriz A, el polinomio característico de A anula a A.

Teorema 6.35 (Teorema de Hamilton-Cayley) *Sea $A \in K^{n\times n}$ y sea $\mathcal{X}_A$ el polinomio característico de A. Entonces $m_A \mid \mathcal{X}_A$ (lo que es equivalente a que $\mathcal{X}_A(A) = 0$).*

Demostración. Sea $f_A : K^n \to K^n$ la transformación lineal definida por $f_A(x) = A.x$.

Sea $v \in K^n$. Supongamos que $\{v, f_A(v), \dots, f_A^k(v)\}$ es un conjunto linealmente independiente y que $f_A^{k+1}(v) = (-a_k)f_A^k(v) + \cdots + (-a_1)f_A(v) + (-a_0)$. Entonces $m_v = X^{k+1} + a_kX^k + \cdots + a_1X + a_0$.

Extendemos el conjunto $\{v, f_A(v), \dots, f_A^k(v)\}$ a una base de K^n: sean $w_{k+2}, \dots, w_n \in K^n$ tales que

$$B = \{v, f_A(v), \dots, f_A^k(v), w_{k+2}, \dots, w_n\}$$

es una base de K^n. Se tiene que

$$|f_A|_B = \left(\begin{array}{ccccc|c} 0 & 0 & \dots & 0 & -a_0 & \\ 1 & 0 & & 0 & -a_1 & \\ 0 & 1 & & \vdots & \vdots & M \\ \vdots & & \ddots & 0 & -a_{k-1} & \\ 0 & \dots & & 1 & -a_k & \\ & & 0 & & & N \end{array} \right).$$

Sabemos que $\mathcal{X}_A = \mathcal{X}_{f_A} = \mathcal{X}_{|f_A|_B}$, con lo que

$$\begin{aligned} \mathcal{X}_A &= \det \left(\begin{array}{ccccc|c} X & 0 & \dots & 0 & a_0 & \\ -1 & X & & 0 & a_1 & \\ 0 & -1 & & \vdots & \vdots & -M \\ \vdots & & \ddots & X & a_{k-1} & \\ 0 & \dots & & -1 & X+a_k & \\ & & 0 & & & X.\,I_{n-k-1}-N \end{array} \right) \\ &= \det \begin{pmatrix} X & 0 & \dots & 0 & a_0 \\ -1 & X & & 0 & a_1 \\ 0 & -1 & & \vdots & \vdots \\ \vdots & & \ddots & X & a_{k-1} \\ 0 & \dots & & -1 & X+a_k \end{pmatrix} \det(X.\,I_{n-k-1}-N). \end{aligned}$$

Por lo calculado en el primer ejemplo de la Sección 5.4, obtenemos que

$$\begin{aligned} \mathcal{X}_A &= (X^{k+1} + a_k X^k + \dots + a_1 X + a_0).\, \det(X.\,I_{n-k-1}-N) \\ &= m_v.\, \det(X.\,I_{n-k-1}-N). \end{aligned}$$

Por lo tanto, $m_v \mid \mathcal{X}_A$ para $v \in K^n$ arbitrario.

Sea $E = \{e_1, \dots, e_n\}$ la base canónica de K^n. Por lo anterior, $m_{e_i} \mid \mathcal{X}_A$ para cada $1 \le i \le n$. En consecuencia, $m_A = \mathrm{mcm}\{m_{e_1}, \dots, m_{e_n}\}$ divide a $\mathcal{X}_A$. □

A partir del Teorema de Hamilton-Cayley podemos deducir algunos resultados relacionados con el polinomio minimal de una matriz.

Observación 6.36 Sea $A \in K^{n\times n}$. Entonces:

1. $\mathrm{gr}(m_A) \le n$.

2. Si $\text{gr}(m_A) = n$, entonces $m_A = \mathcal{X}_A$.
3. Si existe $v \in K^n$ tal que $\text{gr}(m_v) = n$, entonces $m_v = m_A = \mathcal{X}_A$.

En el ejemplo que sigue, mostramos cómo puede utilizarse el hecho de conocer un polinomio que anula a una matriz (en este caso, su polinomio característico) para calcular las potencias de la matriz.

Ejemplo. Sea $A = \begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix} \in \mathbb{R}^{2\times 2}$. Calcular A^n para cada $n \in \mathbb{N}$.

Calculamos el polinomio característico de A: $\mathcal{X}_A = (X-1)^2$. Por el Teorema de Hamilton-Cayley, $\mathcal{X}_A(A) = 0$, es decir, $A^2 - 2.A + I_2 = 0$.

La idea para calcular A^n para cada $n \in \mathbb{N}$ es ver que cada potencia de A es combinación lineal de I_2 y A (observar que la igualdad dada por el teorema de Hamilton-Cayley dice que $A^2 = 2A - I_2$), encontrar esta combinación lineal y usarla para dar una fórmula cerrada para los coeficientes de A^n en función de n.

Sea $n \in \mathbb{N}$. Por el algoritmo de división, existen $P(X) \in \mathbb{R}[X]$ y $a_n, b_n \in \mathbb{R}$ tales que

$$X^n = (X-1)^2 P(X) + a_n.X + b_n. \tag{6.1}$$

Evaluando esta igualdad en $X = 1$, se obtiene

$$1 = 0.\,P(1) + a_n + b_n \iff a_n + b_n = 1.$$

Derivando la identidad (6.1) resulta que

$$n.X^{n-1} = 2(X-1)P + (X-1)^2 P' + a_n,$$

y evaluando esta igualdad en $X = 1$, se deduce que $a_n = n$. En consecuencia, $b_n = 1 - n$.

Luego, para cada $n \in \mathbb{N}$ se tiene que

$$X^n = (X-1)^2 P(X) + nX + 1 - n,$$

de donde

$$A^n = n.A + (1-n)I_2 = n.\begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix} + (1-n)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1-n & -n \\ n & n+1 \end{pmatrix}.$$

Aplicando el Teorema de Hamilton-Cayley, también podemos calcular A^{-1} como combinación lineal de I_2 y A. Basta observar que

$$A^2 - 2A + I_2 = 0 \iff I_2 = 2A - A^2 \iff I_2 = A(2I_2 - A) = (2I_2 - A)A,$$

lo que implica que $A^{-1} = 2I_2 - A$.

El procedimiento para calcular A^{-1} del ejemplo anterior puede generalizarse para hallar la inversa de cualquier matriz en $GL(n, K)$:

Observación 6.37 Sea $A \in GL(n, K)$. Entonces $A^{-1} \in < I_n, A, A^2, \dots, A^{n-1} >$.

Supongamos que $\mathcal{X}_A = X^n + a_{n-1}X^{n-1} + \cdots + a_1 X + a_0$. Entonces, por el Teorema de Hamilton-Cayley,

$$A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 I_n = 0.$$

Observemos que $a_0 = \mathcal{X}_A(0) = \det(0.I_n - A) = \det(-A) = (-1)^n \det(A) \neq 0$, puesto que A es inversible. En consecuencia,

$$\begin{aligned} I_n &= \frac{-1}{a_0} . (A^n + a_{n-1}A^{n-1} + \cdots + a_2 A^2 + a_1 A) \\ &= A. \Big(\frac{-1}{a_0} . (A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_2 A + a_1 I_n) \Big) \\ &= \Big(\frac{-1}{a_0} . (A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_2 A + a_1 I_n) \Big) . A, \end{aligned}$$

de donde $A^{-1} = \dfrac{-1}{a_0} . (A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_2 A + a_1 I_n)$.

6.3.4 Un criterio de diagonalización usando el polinomio minimal

En primer lugar, notemos que vale lo siguiente:

Observación 6.38 Sea $A \in K^{n\times n}$ y sea $\mathcal{X}_A \in K[X]$ el polinomio característico de A. Si $\mathcal{X}_A$ se factoriza linealmente en $K[X]$ y tiene todas sus raíces simples, entonces A es diagonalizable.

En efecto, si $\mathcal{X}_A = (X - \lambda_1) \dots (X - \lambda_n)$, con $\lambda_i \neq \lambda_j$ para cada $i \neq j$, A tiene n autovalores distintos. Además, para cada $1 \leq i \leq n$, existe $v_i \in K^n$ autovector de A de autovalor λ_i. Puesto que $E_{\lambda_1}, \dots, E_{\lambda_n}$ están en suma directa, $\{v_1, \dots, v_n\}$ resulta una base de K^n formada por autovectores de A, y por lo tanto, A es diagonalizable.

La recíproca de este resultado no es cierta, es decir, $\mathcal{X}_A$ puede tener raíces múltiples y, de todas formas, A ser diagonalizable. Basta considerar, por ejemplo, la matriz $A = I_n$.

Sin embargo, es posible dar una condición necesaria y suficiente para la diagonalización, considerando la factorización del polinomio minimal de la matriz.

Proposición 6.39 *Sea $A \in K^{n\times n}$. Entonces A es diagonalizable en $K^{n\times n}$ si y sólo si m_A tiene todas sus raíces en K y son simples.*

Demostración.

($\Rightarrow$) Supongamos que A es diagonalizable. Sean $\lambda_1, \dots, \lambda_r$ los autovalores de A, con $\lambda_i \neq \lambda_j$ si $i \neq j$ y sea $\{v_1, \dots, v_n\}$ una base de autovectores de A. Si m_v denota el minimal del vector v para la matriz A, se tiene que:

$$\begin{aligned} m_A &= \text{mcm}\{m_{v_1}, \dots, m_{v_n}\} \\ &= \text{mcm}\{X - \lambda_1, \dots, X - \lambda_1, \dots, X - \lambda_r, \dots, X - \lambda_r\} \\ &= (X - \lambda_1)(X - \lambda_2) \dots (X - \lambda_r) \end{aligned}$$

En consecuencia, m_A tiene todas las raíces en K y son simples.

($\Leftarrow$) Supongamos que $m_A = (X - \lambda_1) \dots (X - \lambda_r)$ con $\lambda_i \neq \lambda_j$ si $i \neq j$. Entonces $\lambda_1, \dots, \lambda_r$ son todos los autovalores de A en K.

Veamos que $K^n = \bigoplus_{i=1}^{r} E_{\lambda_i}$, donde $E_{\lambda_i} = \{x \in K^n : A.x = \lambda_i.x\}$.

Sea $v \in K^n - \{0\}$. Consideremos el subespacio

$$S = < v, A.v, A^2.v, \dots, A^m.v, \dots > \subseteq K^n.$$

Supongamos que $\{v, A.v, \dots, A^k.v\}$ es un conjunto linealmente independiente, pero $\{v, A.v, \dots, A^k.v, A^{k+1}.v\}$ es linealmente dependiente.

Luego $A^j.v \in < v, A.v, \dots, A^k.v >$ para todo $j \in \mathbb{N}$ y $B_S = \{v, A.v, A^2.v, \dots, A^k.v\}$ resulta ser una base de S. Además, $\text{gr}(m_v) = k + 1$, o sea, m_v es de la forma

$$m_v = X^{k+1} + a_k X^k + \dots + a_1 X + a_0.$$

Por la construcción de S, si $x \in S$ resulta que $A.x \in S$. Sea $f_A : S \to S$ la transformación lineal definida por $f_A(x) = A.x$. Se tiene que

$$|f_A|_{B_S} = \begin{pmatrix} 0 & 0 & \dots & 0 & -a_0 \\ 1 & 0 & & \vdots & -a_1 \\ 0 & 1 & \ddots & & \vdots \\ \vdots & \vdots & \ddots & 0 & -a_{k-1} \\ 0 & 0 & \dots & 1 & -a_k \end{pmatrix}.$$

Observamos que $\mathcal{X}_{f_A} = \det(XI_{k+1} - |f_A|_{B_S}) = X^{k+1} + a_k X^k + \dots + a_1 X + a_0 = m_v$.

Puesto que $m_v \mid m_A$ y, por hipótesis, m_A tiene todas sus raíces en K y son simples, resulta que $\mathcal{X}_{f_A} = m_v$ tiene todas sus raíces en K, son simples y son algunos de los autovalores de A. Por la Observación 6.38 concluimos que f_A es diagonalizable sobre S. Además, si $\mathcal{X}_{f_A} = (X - \lambda_{i_1}) \dots (X - \lambda_{i_{k+1}})$, como $v \in S$, existen $v_{i_1}, \dots, v_{i_{k+1}}$ autovectores de f_A (donde v_{i_j} es un autovector de autovalor λ_{i_j}) tales que $v = v_{i_1} + \dots + v_{i_{k+1}}$. Pero si v_{i_j} es un autovector de f_A de autovalor λ_{i_j}, entonces es un autovector de A con el mismo autovalor. En consecuencia $v \in \bigoplus_{i=1}^{r} E_{\lambda_i}$.

Como $v \in K^n$ era arbitrario, resulta que $K^n = \bigoplus_{i=1}^{r} E_{\lambda_i}$. La equivalencia dada por el Teorema 6.22 dice entonces que A es diagonalizable. $\square$

A continuación presentamos un ejemplo en el que mostramos cómo puede aplicarse este resultado para determinar si una matriz es diagonalizable.

Ejemplo. Sea $A \in \mathbb{C}^{n \times n}$ tal que $A^k = I_n$ para algún $k \in \mathbb{N}$. Entonces A es diagonalizable.

Por hipótesis, $A^k - I_n = 0$, con lo que el polinomio $X^k - 1$ anula a A. En consecuencia, $m_A \mid X^k - 1$. Como $X^k - 1$ tiene todas sus raíces en $\mathbb{C}$ y son simples, entonces m_A también. Luego, A es diagonalizable.

Notar que si $A \in \mathbb{R}^{n\times n}$ satisface $A^k = I_n$, A no es necesariamente diagonalizable sobre $\mathbb{R}$. Por ejemplo, $A = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ cumple $A^3 = I_3$, pero A no es diagonalizable, puesto que $m_A = (X-1)(X^2+X+1)$, que no tiene todas sus raíces en $\mathbb{R}$.

6.4 Subespacios invariantes

Dada una transformación lineal $f : V \to V$ donde V es un K-espacio vectorial de dimensión finita, una posible manera de estudiar f consiste en descomponer el espacio V como suma directa de subespacios $V = \bigoplus_{i=1}^{r} S_i$ y analizar la restricción de f a cada uno de estos subespacios S_i. Ahora, para que esto sea posible, es necesario que la restricción de f a cada S_i sea una transformación lineal de S_i en S_i, es decir, que la imagen por f del subespacio esté incluida en el mismo.

Definición 6.40 Sea V un K-espacio vectorial y sea $f : V \to V$ una transformación lineal. Un subespacio $S \subseteq V$ se dice *invariante por* f (o f*-invariante*) si $f(S) \subseteq S$.

Ejemplos.

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

   Si $E = \{e_1, e_2, e_3, e_4\}$ es la base canónica de $\mathbb{R}^4$, algunos subespacios invariantes por f_A son: $< e_1, e_2 >$, $< e_2 >$, $< e_3 >$, $< e_4 >$, $< e_3, e_4 >$, $< e_1, e_2, e_4 >$.

2. Hallar todos los subespacios invariantes por f_A siendo $A = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} \in \mathbb{R}^{2\times 2}$.

   (a) Subespacios invariantes de dimensión 0: $S = \{0\}$.

   (b) Subespacios invariantes de dimensión 2: $S = \mathbb{R}^2$.

   (c) Subespacios invariantes de dimensión 1: Observamos que $S = < v >$ es un subespacio de dimensión 1 invariante por f_A si y sólo si $v \neq 0$ y $A.v \in < v >$ o, equivalentemente, $v \neq 0$ y $A.v = \lambda.v$ para algún $\lambda \in \mathbb{R}$.

   Luego, un subespacio $S = < v >$ de dimensión 1 es f_A-invariante si y sólo si v es autovector de A.

   Es fácil ver entonces que el único subespacio f_A-invariante de dimensión 1 es $S = < (0,1) >$.

En la siguiente proposición se prueban algunas propiedades sobre subespacios invariantes.

Proposición 6.41 *Sea V un K-espacio vectorial. Sea $f : V \to V$ una transformación lineal. Entonces:*

i) $\mathrm{Nu}(f)$ *e* $\mathrm{Im}(f)$ *son subespacios f-invariantes de V.*

ii) S es un subespacio f-invariante de V de dimensión 1 *si y sólo si $S = < v >$ con $v \in V$ un autovector de f.*

iii) Si S y T son subespacios f-invariantes de V, entonces $S \cap T$ y $S + T$ son subespacios f-invariantes de V.

Demostración.

i) Se tiene que $f(\mathrm{Nu}(f)) = \{0\} \subseteq \mathrm{Nu}(f)$, con lo que $\mathrm{Nu}(f)$ es invariante por f.

Además, es claro que $f(\mathrm{Im}(f)) \subseteq \mathrm{Im}(f)$, de donde $\mathrm{Im}(f)$ es invariante por f.

ii) Sea S un subespacio f-invariante de V de dimensión 1. Entonces $S = < v >$ para algún $v \in V$, $v \neq 0$, tal que $f(v) \in < v >$. Esta última condición implica que $f(v) = \lambda. v$ para algún $\lambda \in K$, y siendo $v \neq 0$, resulta que v es un autovector de f.

Recíprocamente, si $S = < v >$ con v un autovector de f, como $v \neq 0$, entonces $\dim S = 1$, y como $f(v) = \lambda. v$ para algún $\lambda \in K$, resulta que $f(S) \subseteq S$. Luego, S es un subespacio f-invariante de V de dimensión 1.

iii) Sean S y T subespacios de V invariantes por f. Entonces $f(S \cap T) \subseteq f(S) \subseteq S$, puesto que S es f-invariante. Análogamente, $f(S \cap T) \subseteq T$. En consecuencia, $f(S \cap T) \subseteq S \cap T$, de donde $S \cap T$ es invariante por f.

Para $S + T$, teniendo en cuenta que f es una transformación lineal y que S y T son invariantes por f, se tiene que $f(S + T) \subseteq f(S) + f(T) \subseteq S + T$ y, por lo tanto, $S + T$ es invariante por f. □

Dada una transformación lineal $f : V \to V$ y un subespacio S de V invariante por f, la restricción de f a S, que notaremos $f_{|S}$ resulta ser una transformación lineal de S en S. En lo que sigue, analizamos la relación entre los polinomios minimales y característicos de esta restricción $f_{|S}$ y los de f.

Proposición 6.42 *Sea V un K-espacio vectorial de dimensión finita, sea $f : V \to V$ una transformación lineal y sea $S \subseteq V$ un subespacio invariante por f. Sea $f_{|S} : S \to S$ la restricción. Entonces:*

i) $m_{f_{|S}} \mid m_f$.

ii) $\mathcal{X}_{f_{|S}} \mid \mathcal{X}_f$.

Demostración. Sean $n = \dim V$ y $s = \dim S$. Sea $B_S = \{v_1, \ldots, v_s\}$ una base de S y sean $v_{s+1}, \ldots, v_n \in V$ tales que $B = \{v_1, \ldots, v_s, v_{s+1}, \ldots, v_n\}$ es una base de V.

i) Sabemos que

$$\begin{aligned} m_{f_{|S}} &= \operatorname{mcm}\{m_{v_1}, \ldots, m_{v_s}\} \\ m_f &= \operatorname{mcm}\{m_{v_1}, \ldots, m_{v_s}, m_{v_{s+1}}, \ldots, m_{v_n}\}. \end{aligned}$$

Ahora, como $m_{v_i} \mid m_f$ para cada $1 \leq i \leq s$, el mcm de estos polinomios también lo divide, es decir, $m_{f_{|S}} \mid m_f$.

ii) Para la base B considerada, se tiene que

$$|f|_B = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix} \in K^{n\times n} \qquad \text{con } A = |f_{|S}|_{B_S} \in K^{s\times s}.$$

Luego,

$$\begin{aligned} \mathcal{X}_f = \mathcal{X}_{|f|_B} &= \det \begin{pmatrix} X.I_s - A & -B \\ 0 & X.I_{n-s} - C \end{pmatrix} \\ &= \det(X.I_s - A).\det(X.I_{n-s} - C) \ = \ \mathcal{X}_{f_{|S}}.Q, \end{aligned}$$

con lo que $\mathcal{X}_{f_{|S}} \mid \mathcal{X}_f$. □

Si f es una transformación lineal definida en un espacio V de dimensión finita y el espacio V puede descomponerse como suma directa de subespacios f-invariantes, entonces existe una base de V en la que la matriz de f tiene forma diagonal por bloques. Más precisamente:

Observación 6.43 Sea V un K-espacio vectorial de dimensión finita y sea $f : V \to V$ una transformación lineal. Sean S y T subespacios de V invariantes por f tales que $S \oplus T = V$.

Supongamos que $\dim(S) = s > 0$, $\dim(T) = t > 0$. Sean $B_S = \{v_1, \ldots, v_s\}$ y $B_T = \{w_1, \ldots, w_t\}$ bases de S y T respectivamente. Entonces

$$B = B_S \cup B_T = \{v_1, \ldots, v_s, w_1, \ldots, w_t\}$$

es una base de V y

$$|f|_B = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix},$$

donde $A_1 \in K^{s\times s}$ y $A_2 \in K^{t\times t}$. Más aún, si $f_{|S} : S \to S$ y $f_{|T} : T \to T$ son las restricciones, se tiene que $A_1 = |f_{|S}|_{B_S}$ y $A_2 = |f_{|T}|_{B_T}$.

Esto nos lleva a la siguiente definición:

Definición 6.44 Sea V un K-espacio vectorial y sea $f : V \to V$ una transformación lineal. Sea $S \subseteq V$ un subespacio f-invariante. Un *complemento invariante para S* es un subespacio T de V tal que T es f-invariante y $S \oplus T = V$.

Observamos que dada una transformación lineal $f : V \to V$ y un subespacio S de V invariante por f no siempre existe un complemento invariante para S.

Por ejemplo, sea $f : \mathbb{R}^2 \to \mathbb{R}^2$ definida por $f(x, y) = (0, x)$. Entonces $S = <(0, 1)>$ es f-invariante ($S = \mathrm{Nu}(f)$), pero no admite un complemento invariante. En efecto, si $S \oplus T = \mathbb{R}^2$, entonces $\dim T = 1$. Luego, si T es f-invariante, $T = <v>$ con $v \in \mathbb{R}^2$ un autovector de f. Ahora, el conjunto de autovectores de f es $S - \{0\}$, con lo que $v \in S$, contradiciendo que T es un complemento para S.

En el caso en que V sea suma directa de subespacios f-invariantes, podemos relacionar los polinomios característico y minimal de f con los de sus restricciones a estos subespacios:

Proposición 6.45 *Sea V un K-espacio vectorial de dimensión finita y sea $f : V \to V$ una transformación lineal. Sean S y T subespacios f-invariantes de V tales que $S \oplus T = V$. Entonces:*

i) $\mathcal{X}_f = \mathcal{X}_{f_{|S}} . \mathcal{X}_{f_{|T}}$

ii) $m_f = \mathrm{mcm}(m_{f_{|S}}, m_{f_{|T}})$

Demostración.

i) Se deduce inmediatamente de la Observación 6.43.

ii) Sea $P = \mathrm{mcm}(m_{f_{|S}}, m_{f_{|T}})$. Puesto que S y T son subespacios f-invariantes de V, por la Proposición 6.42, se tiene que $m_{f_{|S}} \mid m_f$ y $m_{f_{|T}} \mid m_f$. Luego $P \mid m_f$.

Por otro lado, por la Observación 6.43, si B_S y B_T son bases de S y T respectivamente, y $B = B_S \cup B_T$, entonces
$$|f|_B = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}$$
con $A_1 = |f_{|S}|_{B_S}$ y $A_2 = |f_{|T}|_{B_T}$.

Como $m_{f_{|S}} \mid P$ y $m_{f_{|T}} \mid P$, resulta que $P(A_1) = 0$ y $P(A_2) = 0$. Entonces, operando por bloques,
$$P(|f|_B) = \begin{pmatrix} P(A_1) & 0 \\ 0 & P(A_2) \end{pmatrix} = 0,$$
de donde $m_f \mid P$.

Luego $P = m_f$, puesto que P y m_f son dos polinomios mónicos que se dividen mutuamente. □

6.5 Ejercicios

Ejercicio 1. Calcular el polinomio característico, los autovalores y los autovectores de la matriz A en cada uno de los siguientes casos ($a \in \mathbb{R}$):

(Analizar por separado los casos $K = \mathbb{R}$ y $K = \mathbb{C}$)

i) $A = \begin{pmatrix} 1 & 3 \\ -3 & -1 \end{pmatrix}$ ii) $A = \begin{pmatrix} 1 & 3 \\ 3 & -1 \end{pmatrix}$ iii) $A = \begin{pmatrix} 0 & a \\ -a & 0 \end{pmatrix}$

iv) $A = \begin{pmatrix} 0 & 2 & 1 \\ -2 & 0 & 3 \\ -1 & -3 & 0 \end{pmatrix}$ v) $A = \begin{pmatrix} 3 & 1 & 0 \\ -4 & -1 & 0 \\ 4 & -8 & -2 \end{pmatrix}$ vi) $A = \begin{pmatrix} a & 1 & 1 \\ 1 & a & 1 \\ 1 & 1 & a \end{pmatrix}$

vii) $A = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}$ viii) $A = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ ix) $A = \begin{pmatrix} 1 & 0 & 0 \\ a & 1 & 0 \\ 0 & a & 1 \end{pmatrix}$

Ejercicio 2. Para cada una de las matrices A del ejercicio anterior, sea U una base de K^n y sea $f : K^n \to K^n$ la tranformación lineal tal que $|f|_U = A$. Decidir si es posible encontrar una base B de K^n tal que $|f|_B$ sea diagonal. En caso afirmativo, calcular $C(U, B)$.

Ejercicio 3. Sea $f : \mathbb{R}^3 \to \mathbb{R}^3$ la transformación lineal definida por:

$$f(x, y, z) = (-x - 2.y + 2.z, -y, -x - 3.y - 4.z)$$

i) Encontrar una base B de $\mathbb{R}^3$ tal que $|f|_B$ sea diagonal.

ii) Calcular $\begin{pmatrix} -1 & -2 & 2 \\ 0 & -1 & 0 \\ -1 & -3 & -4 \end{pmatrix}^n$, $\forall n \in \mathbb{N}$. (Sugerencia: ver Observación 6.26.)

iii) ¿Existe una matriz $P \in \mathbb{R}^{3\times 3}$ tal que $P^2 = \begin{pmatrix} -1 & -2 & 2 \\ 0 & -1 & 0 \\ -1 & -3 & -4 \end{pmatrix}$?

Ejercicio 4.

i) Sea $A = \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \in K^{2\times 2}$. Determinar todos los a, b y $c \in K$ para los que A es diagonalizable.

ii) Probar que toda matriz $A \in \mathbb{C}^{2\times 2}$ es diagonalizable o bien es semejante a una matriz del tipo $\begin{pmatrix} \alpha & 0 \\ 1 & \alpha \end{pmatrix}$.

Ejercicio 5. Diagonalizar las matrices $A \in \mathbb{R}^{n\times n}$ y $B \in \mathbb{R}^{6\times 6}$ encontrando sus autovectores:

$$A = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & 1 & 1 & \dots & 1 \\ 1 & 1 & 1 & \dots & 1 \\ \dots & \dots & \dots & \dots & \dots \\ 1 & 1 & 1 & \dots & 1 \end{pmatrix} \qquad \text{y} \qquad B = \begin{pmatrix} 2 & 1 & 2 & 1 & 2 & 1 \\ 1 & 2 & 1 & 2 & 1 & 2 \\ 2 & 1 & 2 & 1 & 2 & 1 \\ 1 & 2 & 1 & 2 & 1 & 2 \\ 2 & 1 & 2 & 1 & 2 & 1 \\ 1 & 2 & 1 & 2 & 1 & 2 \end{pmatrix}$$

Ejercicio 6. Se sabe que la matriz $A \in \mathbb{R}^{2\times 2}$ tiene a $(1, -1)$ como autovector de autovalor $\sqrt{2}$ y, además, $\mathcal{X}_A \in \mathbb{Q}[X]$. Decidir si A es diagonalizable en $\mathbb{R}^{2\times 2}$. ¿Es A única?

Ejercicio 7.

i) Sea $A \in \mathbb{R}^{3\times 3}$ diagonalizable con $tr(A) = -4$. Calcular los autovalores de A, sabiendo que los autovalores de $A^2 + 2.A$ son -1, 3 y 8.

ii) Sea $A \in \mathbb{R}^{4\times 4}$ tal que $\det(A) = 6$; 1 y -2 son autovalores de A y -4 es autovalor de la matriz $A - 3.I_4$. Hallar los restantes autovalores de A.

Ejercicio 8. Sea $A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \in \mathbb{R}^{2\times 2}$.

i) Probar que, para todo $n \in \mathbb{N}$, $A^n.\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} F_n \\ F_{n+1} \end{pmatrix}$ donde F_i es el i-ésimo término de la sucesión de Fibonacci (es decir, $F_0 = 0$, $F_1 = 1$ y $F_{i+1} = F_i + F_{i-1}$).

ii) Encontrar una matriz $P \in GL(2, \mathbb{R})$ tal que $P.A.P^{-1}$ sea diagonal.

iii) Hallar la fórmula general para el término F_n, $\forall\, n \in \mathbb{N}_0$ (comparar con el Ejercicio 50 de la Sección 1.5).

iv) Se define la sucesión $\{a_n\}_{n\in\mathbb{N}_0}$ de la siguiente manera:

$$\begin{cases} a_0 = 0\,,\ a_1 = 1\,,\ a_2 = 1 \\ a_{n+3} = 6.a_{n+2} - 11.a_{n+1} + 6.a_n \quad \forall\, n \in \mathbb{N}_0 \end{cases}$$

Hallar una fórmula general para el término a_n, $\forall\, n \in \mathbb{N}_0$.

Ejercicio 9. Resolver el siguiente sistema de ecuaciones diferenciales

$$\begin{cases} x'(t) = 6x(t) + 2y(t) \\ y'(t) = 2x(t) + 3y(t) \end{cases}$$

con condiciones iniciales $x(0) = 3$, $y(0) = -1$.

Sugerencia: Hallar una matriz $C \in GL(2, \mathbb{R})$ tal que $C^{-1}\begin{pmatrix} 6 & 2 \\ 2 & 3 \end{pmatrix} C$ sea diagonal y hacer el cambio de variables $\begin{pmatrix} u(t) \\ v(t) \end{pmatrix} = C^{-1}.\begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$.

Ejercicio 10. Sea $A \in K^{n\times n}$. Probar que A y A^t tienen los mismos autovalores. Dar un ejemplo en el que los autovectores sean distintos.

Ejercicio 11. Sea $\delta : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$ la transformación lineal derivación. Mostrar que todo número real es un autovalor de δ y exhibir un autovector correspondiente.

Ejercicio 12. Analizar la validez de las siguientes afirmaciones:

i) $A \in \mathbb{R}^{n\times n}$ inversible $\Rightarrow 0$ no es autovalor de A.

ii) $A \in \mathbb{R}^{n\times n}$ inversible, x autovector de $A \Rightarrow x$ autovector de A^{-1}.

iii) $A \in \mathbb{R}^{n\times n}$ con n impar $\Rightarrow A$ admite un autovalor real.

Ejercicio 13.

i) Sea $f : K^n \to K^n$ un proyector con $\dim(\mathrm{Im}(f)) = s$. Calcular $\mathcal{X}_f$. ¿Es f diagonalizable?

ii) Sea K un cuerpo incluido en $\mathbb{C}$ y sea $f : K^n \to K^n$ un morfismo nilpotente no nulo. Calcular $\mathcal{X}_f$. ¿Es f diagonalizable?

Ejercicio 14. Sea $A \in \mathbb{R}^{n\times n}$ que verifica $A^2 + I_n = 0$. Probar que A es inversible, que no tiene autovalores reales y que n debe ser par.

Ejercicio 15. Sea V un K-espacio vectorial de dimensión finita y sea $f : V \to V$ una transformación lineal tal que $\dim(\mathrm{Im}(f)) = 1$. Probar que f es diagonalizable si y sólo si $\mathrm{Nu}(f) \cap \mathrm{Im}(f) = \{0\}$.

Ejercicio 16. Sea $D \in K^{n\times n}$ una matriz inversible y diagonal. Sea $f : K^{n\times n} \to K^{n\times n}$ la transformación lineal definida por $f(A) = D^{-1}.A.D$. Hallar los autovalores y los autovectores de f y probar que es diagonalizable.

Ejercicio 17. Sea $f : \mathbb{C}^n \to \mathbb{C}^n$ una transformación lineal. Probar que existe una base B de $\mathbb{C}^n$ tal que $|f|_B$ es triangular superior.

Ejercicio 18. Sea $A \in \mathbb{C}^{n\times n}$ y sean $\lambda_1, \dots, \lambda_n$ las raíces de $\mathcal{X}_A$ contadas con multiplicidad.

i) Probar que $\det(A) = \prod_{i=1}^{n} \lambda_i$.

ii) Probar que $tr(A) = \sum_{i=1}^{n} \lambda_i$.

Ejercicio 19. Sean $A \in K^{m\times n}$ y $B \in K^{n\times m}$.

i) Probar que las matrices $\begin{pmatrix} A.B & 0 \\ B & 0 \end{pmatrix}$ y $\begin{pmatrix} 0 & 0 \\ B & B.A \end{pmatrix}$ en $K^{(m+n)\times(m+n)}$ son semejantes.

ii) Deducir que, si $n = m$, $\mathcal{X}_{A.B} = \mathcal{X}_{B.A}$.

Ejercicio 20. Dadas las matrices $A \in \mathbb{C}^{2\times 2}$ y los polinomios $P \in \mathbb{C}[X]$, calcular $P(A)$ para:

i) $A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$, a) $P = X - 1$, b) $P = X^2 - 1$, c) $P = (X-1)^2$

ii) $A = \begin{pmatrix} i & 0 \\ 1 & -i \end{pmatrix}$, $P = X^3 - i.X^2 + 1 + i$

Ejercicio 21. Sea $A \in \mathbb{R}^{n\times n}$. Probar que el minimal de A como matriz real y el minimal de A como matriz compleja coinciden.

Ejercicio 22. Hallar el polinomio minimal de las siguientes matrices (comparar con el polinomio característico):

i) $\begin{pmatrix} 4 & 1 \\ -2 & 1 \end{pmatrix}$ ii) $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ iii) $\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$ iv) $\begin{pmatrix} i & 0 \\ 1 & i \end{pmatrix}$

v) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ vi) $\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$ vii) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ viii) $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$

ix) $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$ x) $\begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 2 & -1 & 0 \\ 3 & 4 & 0 & -1 \end{pmatrix}$ xi) $\begin{pmatrix} 1 & i & 1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$

xii) $\begin{pmatrix} a & 0 & 0 & 0 \\ 0 & a & 0 & 0 \\ 0 & 0 & a & 0 \\ 0 & 0 & 0 & a \end{pmatrix}$ xiii) $\begin{pmatrix} a & 0 & 0 & 0 \\ 1 & a & 0 & 0 \\ 0 & 0 & a & 0 \\ 0 & 0 & 1 & a \end{pmatrix}$ xiv) $\begin{pmatrix} a & 0 & 0 & 0 \\ 1 & a & 0 & 0 \\ 0 & 1 & a & 0 \\ 0 & 0 & 0 & a \end{pmatrix}$ xv) $\begin{pmatrix} a & 0 & 0 & 0 \\ 1 & a & 0 & 0 \\ 0 & 1 & a & 0 \\ 0 & 0 & 1 & a \end{pmatrix}$

Ejercicio 23. Calcular el polinomio minimal para cada una de las siguientes transformaciones lineales:

i) $f : \mathbb{R}_2[X] \to \mathbb{R}_2[X]$, $f(P) = P' + 2.P$

ii) $f : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$, $f(A) = A^t$

Ejercicio 24. Sea $A \in K^{n\times n}$ la matriz

$$A = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & -a_0 \\ 1 & 0 & 0 & \dots & 0 & -a_1 \\ 0 & 1 & 0 & \dots & 0 & -a_2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & -a_{n-2} \\ 0 & 0 & 0 & \dots & 1 & -a_{n-1} \end{pmatrix}$$

Calcular su polinomio minimal y su polinomio característico. Comparar con lo calculado en la Sección 5.4.

Ejercicio 25. Sea $\delta : \mathbb{R}[X] \to \mathbb{R}[X]$ la transformación lineal derivada. Probar que δ no admite ningún polinomio minimal.

Ejercicio 26. Utilizando el Teorema de Hamilton-Cayley:

i) Calcular $A^4 - 4.A^3 - A^2 + 2.A - 5.I_2$ para $A = \begin{pmatrix} 2 & -1 \\ 1 & 3 \end{pmatrix}$.

ii) Calcular A^{1000} para $A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$.

iii) Dada $A = \begin{pmatrix} 1 & 3 \\ -1 & 4 \end{pmatrix}$, expresar a A^{-1} como combinación lineal de A y de I_2.

iv) Dada $A = \begin{pmatrix} 1 & -1 \\ 2 & 5 \end{pmatrix}$, expresar a $(2.A^4 - 12.A^3 + 19.A^2 - 29.A - 37.I_2)^{-1}$ como combinación lineal de A y de I_2.

v) Dada $A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$, calcular A^{-1}, A^3 y A^{-3}.

vi) Calcular $\begin{pmatrix} 2 & 2 & 0 \\ 1 & 2 & 1 \\ 1 & 2 & 1 \end{pmatrix}^n \quad \forall n \in \mathbb{N}$.

Ejercicio 27. Sea V un K-espacio vectorial de dimensión finita y sea $f : V \to V$ una transformación lineal. Probar que f es un isomorfismo si y sólo si el término constante de $\mathcal{X}_f$ es no nulo. En dicho caso, hallar la expresión general de f^{-1} como polinomio en f.

Ejercicio 28. Exhibir una matriz $A \in \mathbb{C}^{n\times n}$ tal que $A^2 + I_n = 0$. Comparar con el Ejercicio 14.

Ejercicio 29.

i) Sea $f : \mathbb{R}^2 \to \mathbb{R}^2$ la transformación lineal definida por $f(x, y) = (x + 3.y, 3.x - 2.y)$. Hallar todos los subespacios de $\mathbb{R}^2$ que sean f-invariantes.

ii) Sea $f_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ la rotación de ángulo θ. Probar que, para todo $\theta \neq k.\pi$ ($k \in \mathbb{Z}$), f_θ no es diagonalizable. Hallar todos los subespacios de $\mathbb{R}^2$ que sean f_θ-invariantes.

iii) Sea $\theta \in \mathbb{R}$ y $g_\theta : \mathbb{C}^2 \to \mathbb{C}^2$ la transformación $\mathbb{C}$-lineal cuya matriz en la base canónica es

$$|g_\theta|_E = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}$$

¿Es g_θ diagonalizable? Hallar todos los subespacios de $\mathbb{C}^2$ que sean g_θ-invariantes.

Ejercicio 30. Sea $f : \mathbb{R}^n \to \mathbb{R}^n$ una tranformación lineal nilpotente tal que $f^n = 0$ y $f^{n-1} \neq 0$. Probar que existe un hiperplano de $\mathbb{R}^n$ que es f-invariante pero que no admite un complemento f-invariante (comparar con el Ejercicio 23. ii) de la Sección 3.8).

Ejercicio 31.

i) Sea V un K-espacio vectorial de dimensión finita y sea $f : V \to V$ una transformación lineal diagonalizable. Si S es un subespacio de V f-invariante, probar que $f : S \to S$ es diagonalizable.

ii) Sean $A, B \in K^{n\times n}$ tales que $A.B = B.A$ y sea $E_\lambda = \{x \in K^n \,/\, A.x = \lambda.x\}$. Probar que E_λ es B-invariante.

iii) Sean $A, B \in K^{n\times n}$ dos matrices diagonalizables tales que $A.B = B.A$. Probar que existe $C \in GL(n, K)$ tal que $C.A.C^{-1}$ y $C.B.C^{-1}$ son diagonales. (Es decir, A y B se pueden diagonalizar simultáneamente.)

Ejercicio 32.

i) Hallar una matriz $A \in \mathbb{C}^{3\times 3}$ tal que $m_A(X) = X^3 - 5X^2 + 6X + 8$. Decidir si A es diagonalizable.

ii) Hallar una matriz $A \in \mathbb{C}^{4\times 4}$ tal que $m_A(X) = X^4 + 4X^3 + 8X^2 + 8X + 4$. Decidir si A es diagonalizable.

Ejercicio 33. Sea $A \in K^{n\times n}$.

i) Probar que si A es nilpotente, entonces existe $k \in \mathbb{N}$ tal que $m_A(X) = X^k$. Calcular todos los autovalores de A.

ii) Si $K = \mathbb{C}$ y el único autovalor de A es el 0, probar que A es nilpotente. ¿Qué pasa si $K = \mathbb{R}$?

Ejercicio 34. Sea $A \in \mathbb{C}^{n\times n}$ una matriz de traza nula. Probar que A es semejante a una matriz que tiene toda la diagonal nula.

Capítulo 7

Forma de Jordan

En este capítulo continuaremos estudiando la estructura de los endomorfismos de un espacio vectorial de dimensión finita.

Veremos que si V es un K-espacio vectorial de dimensión finita y f es un endomorfismo de V, bajo ciertas condiciones, existe una base de V en la cual la matriz de f es de una forma particular que nos permite clasificar los endomorfismos y también trabajar más fácilmente con ellos. Esto, en particular, resuelve el problema de decidir si dos matrices complejas son semejantes o no, o sea si son las matrices de la misma transformación lineal en distintas bases.

Comenzaremos estudiando algunos casos particulares y luego extenderemos los resultados obtenidos al caso general.

7.1 Transformaciones lineales nilpotentes

7.1.1 Definiciones y propiedades básicas

Empezamos con el caso en que la transformación lineal tenga como polinomio característico a una potencia de X (notar que en este caso, usando el Teorema de Hamilton-Cayley, el polinomio minimal de la transformación lineal también es una potencia de X y su único autovalor es el 0). Esto significa que existe una potencia de f que da cero, lo que motiva la siguiente definición.

Definición 7.1 Sea V un K-espacio vectorial. Una transformación lineal $f : V \to V$ se dice *nilpotente* si existe $k \in \mathbb{N}$ tal que $f^k = \underbrace{f \circ f \circ \ldots \circ f}_{k \text{ veces}} = 0$.

Análogamente, se dice que una matriz $A \in K^{n\times n}$ es *nilpotente* si existe $k \in \mathbb{N}$ tal que $A^k = 0$.

Observación 7.2 Si V es un K-espacio vectorial de dimensión finita y $f : V \to V$ una transformación lineal, entonces f es nilpotente si y sólo si para cualquier base B de V, $|f|_B$

es una matriz nilpotente.

Definición 7.3 Sea $f : V \to V$ una transformación lineal nilpotente. Se define el *índice de nilpotencia* de f como $\min\{j \in \mathbb{N} \ / \ f^j = 0\}$.

Análogamente, se define el índice de nilpotencia de una matriz nilpotente $A \in K^{n\times n}$ como $\min\{j \in \mathbb{N} \ / \ A^j = 0\}$.

Lema 7.4 *Sea V un K-espacio vectorial de dimensión finita y sea $f : V \to V$ una transformación lineal. Entonces f es nilpotente de índice k si y sólo si $m_f = X^k$.*

Demostración. Si f es nilpotente de índice k, se tiene que $f^k = 0$ y $f^{k-1} \neq 0$. La primera condición implica que el polinomio X^k anula a f, y en consecuencia, $m_f \mid X^k$. Luego, $m_f = X^j$ para algún $j \le k$. Ahora, como $f^{k-1} \neq 0$, resulta que $m_f = X^k$.

Recíprocamente, es claro que si $m_f = X^k$, entonces $f^k = 0$ y $f^{k-1} \neq 0$, con lo cual f es nilpotente de índice k. □

Notar que, con las mismas hipótesis del lema anterior, como el grado del polinomio minimal de f es siempre menor o igual que la dimensión n de V, tendremos que f es nilpotente si y sólo si $f^n = 0$ (comparar con el Ejercicio 14 de la Sección 3.8).

Proposición 7.5 *Sea V un K-espacio vectorial de dimensión n y sea $f : V \to V$ una transformación lineal nilpotente de índice k. Entonces*

$$\{0\} \subset \mathrm{Nu}(f) \subset \mathrm{Nu}(f^2) \subset \ldots \subset \mathrm{Nu}(f^k) = V$$

y todas las inclusiones son estrictas.

Demostración. Siendo k el índice de nilpotencia de f, se tiene que $f^k = 0$, de donde $\mathrm{Nu}(f^k) = V$. Además, es claro que valen las inclusiones. Veamos que son estrictas.

En primer lugar, observamos que si $\mathrm{Nu}(f^i) = \mathrm{Nu}(f^{i+1})$ para algún $i \in \mathbb{N}_0$, entonces $\mathrm{Nu}(f^{i+1}) = \mathrm{Nu}(f^{i+2})$: Si $v \in \mathrm{Nu}(f^{i+2})$, se tiene que $f^{i+2}(v) = 0$, de donde $f^{i+1}(f(v)) = 0$. Luego, $f(v) \in \mathrm{Nu}(f^{i+1})$ y, como por hipótesis $\mathrm{Nu}(f^{i+1}) = \mathrm{Nu}(f^i)$, entonces $f(v) \in \mathrm{Nu}(f^i)$. Esto dice que $f^{i+1}(v) = f^i(f(v)) = 0$, es decir, $v \in \mathrm{Nu}(f^{i+1})$.

Luego, si $\mathrm{Nu}(f^{i_0}) = \mathrm{Nu}(f^{i_0+1})$ para algún $i_0 \in \mathbb{N}_0$, se tiene que $\mathrm{Nu}(f^i) = \mathrm{Nu}(f^{i+1})$ para todo $i \ge i_0$. Pero como el índice de nilpotencia de f es k, $\mathrm{Nu}(f^{k-1}) \neq V = \mathrm{Nu}(f^k)$, y en consecuencia debe ser $i_0 \ge k$. □

Notación. Para cada matriz $A \in K^{n\times n}$, notaremos $\mathrm{Nu}(A)$ al núcleo de la transformación lineal $f_A : K^n \to K^n$ definida por $f_A(x) = A.x$, es decir, al conjunto $\{x \in K^n : A.x = 0\}$.

Con esta notación, como consecuencia de la Proposición 7.5 se tiene que si $A \in K^{n\times n}$ es una matriz nilpotente de índice k, entonces

$$\{0\} \subset \mathrm{Nu}(A) \subset \mathrm{Nu}(A^2) \subset \ldots \subset \mathrm{Nu}(A^k) = K^n$$

y todas las inclusiones son estrictas.

7.1.2 Existencia de forma de Jordan para una transformación lineal nilpotente

Comenzamos estudiando un caso particular de endomorfismos nilpotentes: los de índice de nilpotencia máximo (o sea, aquéllos en los que el polinomio minimal coincide con el característico):

Sea V un K-espacio vectorial de dimensión n y sea $f : V \to V$ una transformación lineal nilpotente de índice n (es decir, $m_f = \mathcal{X}_f = X^n$).

Sea $v \in V$ tal que $f^{n-1}(v) \neq 0$, es decir $v \in V - \mathrm{Nu}(f^{n-1})$ (existe un elemento con esta propiedad porque por hipótesis $f^{n-1} \neq 0$). Como m_v divide a $m_f = X^n$, resulta que $m_v = X^k$ con $1 \leq k \leq n$. Como $f^{n-1}(v) \neq 0$, resulta que $m_v = X^n$ y, por lo tanto, el conjunto

$$B = \{v, f(v), f^2(v), \dots, f^{n-1}(v)\}$$

es una base de V. Además,

$$|f|_B = \begin{pmatrix} 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix}.$$

Este tipo de matrices aparecerá en nuestra clasificación y les daremos un nombre:

Definición 7.6 Sea $J \in K^{n \times n}$. Se dice que J es un *bloque de Jordan nilpotente* si

$$J = \begin{pmatrix} 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix}.$$

A continuación probaremos que para cualquier transformación lineal nilpotente $f : V \to V$, donde V es un K-espacio vectorial de dimensión n, es posible encontrar una base donde la matriz esté formada únicamente por bloques de Jordan nilpotentes ubicados sobre la diagonal y ceros en los demás lugares.

Teorema 7.7 *Sea V un K-espacio vectorial de dimensión n, y sea $f : V \to V$ una transformación lineal nilpotente de índice k. Entonces existe una base B de V tal que*

$$|f|_B = \begin{pmatrix} J_1 & 0 & \dots & 0 \\ 0 & J_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & J_r \end{pmatrix}$$

donde, para cada $1 \leq i \leq r$, $J_i \in K^{n_i \times n_i}$ es un bloque de Jordan nilpotente y $k = n_1 \geq n_2 \geq \dots \geq n_r$.

Para la demostración de este teorema, usaremos el siguiente resultado técnico:

Lema 7.8 *Sea V un K-espacio vectorial de dimensión finita. Sea $f : V \to V$ una transformación lineal y sea $i \in \mathbb{N}$. Sea $\{v_1, \dots, v_r\} \subseteq V$ un conjunto linealmente independiente tal que* $\mathrm{Nu}(f^i) \cap < v_1, \dots, v_r > = \{0\}$. *Entonces $\{f(v_1), \dots, f(v_r)\}$ es linealmente independiente y* $\mathrm{Nu}(f^{i-1}) \cap < f(v_1), \dots, f(v_r) > = \{0\}$.

Demostración. Supongamos que $v = \alpha_1 f(v_1) + \cdots + \alpha_r f(v_r) \in \mathrm{Nu}(f^{i-1})$. Entonces

$$0 = f^{i-1}(v) = f^i(\alpha_1 v_1 + \cdots + \alpha_r v_r),$$

de donde $\alpha_1 v_1 + \cdots + \alpha_r v_r \in \mathrm{Nu}(f^i)$. Como $\mathrm{Nu}(f^i) \cap < v_1, \dots, v_r > = \{0\}$, resulta que $\alpha_1 v_1 + \cdots + \alpha_r v_r = 0$, y como $v_1, \dots, v_r$ son linealmente independientes, $\alpha_1 = \ldots = \alpha_r = 0$.

Luego, $v = 0$.

De la misma demostración se deduce que si $\alpha_1 f(v_1) + \cdots + \alpha_r f(v_r) = 0$, entonces $\alpha_1 = \ldots = \alpha_r = 0$, con lo cual $\{f(v_1), \dots, f(v_r)\}$ es linealmente independiente. $\square$

Demostración del Teorema 7.7. Como f es nilpotente de índice k se tiene que

$$\{0\} \subsetneq \mathrm{Nu}(f) \subsetneq \mathrm{Nu}(f^2) \subsetneq \ldots \subsetneq \mathrm{Nu}(f^{k-1}) \subsetneq \mathrm{Nu}(f^k) = V.$$

Lo que haremos a continuación es construir conjuntos de vectores en $\mathrm{Nu}(f^j)$ recursivamente, comenzando en $j = k$ hasta $j = 1$, utilizando el lema anterior y de forma tal que la unión de esos conjuntos sea una base de V.

Sea B_{k-1} una base de $\mathrm{Nu}(f^{k-1})$, y sea

$$C_k = \{v_1^{(k)}, \dots, v_{r_k}^{(k)}\} \subset \mathrm{Nu}(f^k) = V$$

un conjunto linealmente independiente tal que $B_{k-1} \cup C_k$ es una base de $\mathrm{Nu}(f^k) = V$. Por construcción, C_k es un conjunto linealmente independiente y

$$\mathrm{Nu}(f^{k-1}) \oplus < C_k > = \mathrm{Nu}(f^k) = V.$$

Para fijar ideas, hagamos el paso siguiente de la recursión:

Por el lema anterior $f(C_k) \subset \mathrm{Nu}(f^{k-1})$ es un conjunto linealmente independiente tal que $\mathrm{Nu}(f^{k-2}) \cap f(C_k) = \{0\}$. Sea B_{k-2} una base de $\mathrm{Nu}(f^{k-2})$. Completamos el conjunto linealmente independiente $B_{k-2} \cup f(C_k)$ a una base de $\mathrm{Nu}(f^{k-1})$ con $\{v_1^{(k-1)}, \dots, v_{r_{k-1}}^{(k-1)}\}$. Luego, si llamamos

$$C_{k-1} = f(C_k) \cup \{v_1^{(k-1)}, \dots, v_{r_{k-1}}^{(k-1)}\} \subset \mathrm{Nu}(f^{k-1}),$$

tenemos que C_{k-1} es un conjunto linealmente independiente y vale

$$\mathrm{Nu}(f^{k-2}) \oplus < C_{k-1} > = \mathrm{Nu}(f^{k-1}).$$

Notar que, por lo tanto,

$$\mathrm{Nu}(f^{k-2}) \oplus < C_{k-1} > \oplus < C_k > = V.$$

Pasemos ahora al paso j-ésimo de la recursión:

Sea j con $1 \le j < k$. Supongamos construidos los conjuntos linealmente independientes $C_{j+1} \subset \mathrm{Nu}(f^{j+1}), \dots, C_k \subset \mathrm{Nu}(f^k)$ tales que $f(C_h) \subset C_{h-1}\ \forall\, j+2 \le h \le k$, $\mathrm{Nu}(f^j) \oplus < C_{j+1} > = \mathrm{Nu}(f^{j+1})$ y

$$\mathrm{Nu}(f^j) \oplus < C_{j+1} > \oplus \dots \oplus < C_k > = V.$$

Por el lema anterior, $f(C_{j+1}) \subset \mathrm{Nu}(f^j)$ es un conjunto linealmente independiente y $\mathrm{Nu}(f^{j-1}) \cap < f(C_{j+1}) > = \{0\}$. Consideremos una base B_{j-1} de $\mathrm{Nu}(f^{j-1})$. Entonces $B_{j-1} \cup f(C_{j+1}) \subset \mathrm{Nu}(f^j)$ es un conjunto linealmente independiente y, por lo tanto, existen $v_1^{(j)}, \dots, v_{r_j}^{(j)} \in \mathrm{Nu}(f^j)$ tales que $B_{j-1} \cup f(C_{j+1}) \cup \{v_1^{(j)}, \dots, v_{r_j}^{(j)}\}$ es una base de $\mathrm{Nu}(f^j)$. Sea

$$C_j = f(C_{j+1}) \cup \{v_1^{(j)}, \dots, v_{r_j}^{(j)}\} \subset \mathrm{Nu}(f^j).$$

Es claro que $C_j \subset \mathrm{Nu}(f^j)$ es un conjunto linealmente independiente (puesto que, por construcción, es un subconjunto de una base de $\mathrm{Nu}(f^j)$) y que

$$\mathrm{Nu}(f^{j-1}) \oplus < C_j > = \mathrm{Nu}(f^j).$$

Por lo tanto,

$$\mathrm{Nu}(f^{j-1}) \oplus < C_j > \oplus \dots \oplus < C_k > = V.$$

Al terminar la recursión tendremos que

$$< C_1 > \oplus \dots \oplus < C_k > = V,$$

y como cada conjunto C_j para cada $1 \le j \le k$ es linealmente independiente, resulta que $\bigcup\limits_{j=1}^{k} C_j$ es una base de V.

Consideremos la base B de V obtenida reordenando esta base como sigue:

$$\begin{aligned} B \;=\; \{ & v_1^{(k)}, f(v_1^{(k)}), \dots, f^{k-1}(v_1^{(k)}), \dots, v_{r_k}^{(k)}, f(v_{r_k}^{(k)}), \dots, f^{k-1}(v_{r_k}^{(k)}), \dots, \\ & v_1^{(j)}, f(v_1^{(j)}), \dots, f^{j-1}(v_1^{(j)}), \dots, v_{r_j}^{(j)}, f(v_{r_j}^{(j)}), \dots, f^{j-1}(v_{r_j}^{(j)}), \dots, v_1^{(1)}, \dots, v_{r_1}^{(1)}\}. \end{aligned}$$

Se puede verificar que $|f|_B$ tiene la forma del enunciado del Teorema. □

Definición 7.9 Una matriz $A \in K^{n \times n}$ se dice una *forma de Jordan nilpotente* si

$$A = \begin{pmatrix} J_1 & 0 & \dots & 0 \\ 0 & J_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & J_r \end{pmatrix}$$

con $J_i \in K^{n_i \times n_i}$ bloques de Jordan nilpotentes $(1 \le i \le r)$ y $n_1 \ge n_2 \ge \dots \ge n_r$.

El Teorema 7.7 nos dice entonces que para todo endomorfismo nilpotente $f : V \to V$, donde V es un K-espacio vectorial de dimensión finita, existe una base B de V tal que $|f|_B$ es una forma de Jordan nilpotente. A una tal base B la llamaremos una *base de Jordan para* f y a la matriz $|f|_B$, una *forma de Jordan para* f.

Aplicando este teorema a la transformación lineal $f_A : K^n \to K^n$ asociada a una matriz $A \in K^{n\times n}$, obtenemos el análogo para matrices:

Teorema 7.10 *Sea $A \in K^{n\times n}$ una matriz nilpotente. Entonces A es semejante a una forma de Jordan nilpotente.*

A una base B de K^n tal que $|f_A|_B = J_A$ es una forma de Jordan nilpotente la llamaremos una *base de Jordan para* A, y a la matriz J_A una *forma de Jordan para* A.

Ejemplo. Hallar una forma de Jordan y una base de Jordan para $A \in \mathbb{R}^{6\times 6}$, siendo

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

Como $\mathcal{X}_A = X^6$, A es una matriz nilpotente.

Calculemos m_A, que será de la forma $m_A = X^k$ para algún k con $1 \le k \le 6$. Como

$$A^2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \qquad \text{y} \qquad A^3 = 0,$$

resulta que $m_A = X^3$.

Sea $E = \{e_1, e_2, e_3, e_4, e_5, e_6\}$ la base canónica de $\mathbb{R}^6$. Entonces,

$$B_1 = \{e_3, e_4, e_6\}, \quad B_2 = \{e_3, e_4, e_6, e_2, e_5\} \quad \text{y} \quad B_3 = \{e_3, e_4, e_6, e_2, e_5, e_1\}$$

son bases de $\mathrm{Nu}(A)$, $\mathrm{Nu}(A^2)$ y $\mathrm{Nu}(A^3)$ respectivamente.

Construimos una base de Jordan para A siguiendo la demostración del Teorema 7.7 (donde la transformación lineal nilpotente que consideramos es $f_A : \mathbb{R}^6 \to \mathbb{R}^6$): Tenemos que

$$\{0\} \subsetneq \mathrm{Nu}(A) \subsetneq \mathrm{Nu}(A^2) \subsetneq \mathrm{Nu}(A^3) = \mathbb{R}^6.$$

Extendemos la base B_2 de $\mathrm{Nu}(A^2)$ a una base de $\mathrm{Nu}(A^3) = \mathbb{R}^6$, por ejemplo, agregando el vector e_1 que completa B_3. Consideramos $A.e_1 = (0, 1, -1, 0, -1, 1) \in \mathrm{Nu}(A^2)$. Se tiene:

$$\begin{array}{ccccccccc} \{0\} & \subsetneq & \mathrm{Nu}(A) & \subsetneq & \mathrm{Nu}(A^2) & \subsetneq & \mathrm{Nu}(A^3) & = \mathbb{R}^6 \\ & & & & A.e_1 & & e_1 & \end{array}$$

Ahora consideramos la base B_1 de $\mathrm{Nu}(A)$, tomamos el conjunto $B_1 \cup \{A.e_1\} \subset \mathrm{Nu}(A^2)$, y extendemos este conjunto a una base de $\mathrm{Nu}(A^2)$. Para esto podemos elegir, por ejemplo, el vector $e_5 \in \mathrm{Nu}(A^2)$. Multiplicando por A los vectores $A.e_1$ y e_5 se obtiene el conjunto linealmente independiente $\{A^2e_1, A.e_5\} = \{(0,0,-1,0,0,1),(0,0,0,1,0,-1)\} \subset \mathrm{Nu}(A^2)$:

$$\begin{array}{ccccccccc}
\{0\} & \subsetneq & \mathrm{Nu}(A) & \subsetneq & \mathrm{Nu}(A^2) & \subsetneq & \mathrm{Nu}(A^3) & = \mathbb{R}^6 \\
 & & A^2.e_1 & & A.e_1 & & e_1 & \\
 & & A.e_5 & & e_5 & & &
\end{array}$$

Finalmente, extendemos el conjunto $\{A^2.e_1, A.e_5\}$ a una base de $\mathrm{Nu}(A)$, por ejemplo, con el vector e_3. Obtenemos:

$$\begin{array}{ccccccccc}
\{0\} & \subsetneq & \mathrm{Nu}(A) & \subsetneq & \mathrm{Nu}(A^2) & \subsetneq & \mathrm{Nu}(A^3) & = \mathbb{R}^6 \\
 & & A^2.e_1 & & A.e_1 & & e_1 & \\
 & & A.e_5 & & e_5 & & & \\
 & & e_3 & & & & &
\end{array}$$

Entonces, una base de Jordan para A es

$$\begin{aligned}
B &= \{e_1,\, A.e_1,\, A^2.e_1,\, e_5,\, A.e_5,\, e_3\} \\
&= \{(1,0,0,0,0,0),(0,1,-1,0,-1,1),(0,0,-1,0,0,1),(0,0,0,0,1,0), \\
&\quad (0,0,0,1,0,-1),(0,0,1,0,0,0)\}.
\end{aligned}$$

y una forma de Jordan de A es $J_A = |f_A|_B$, es decir:

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

7.1.3 Unicidad de la forma de Jordan nilpotente. Semejanza

A continuación probaremos que la forma de Jordan de una transformación lineal nilpotente es única. Este resultado, junto con la existencia de forma de Jordan para matrices nilpotentes probada en la sección anterior, nos permitirá resolver el problema de decidir si dos matrices nilpotentes son semejantes o no.

La demostración de la unicidad consiste en mostrar que la cantidad de bloques de Jordan de cada tamaño que aparecen en una forma de Jordan de una transformación lineal f está unívocamente determinada por f.

Comenzamos probando un resultado auxiliar.

Lema 7.11 Sea $J \in K^{m\times m}$ un bloque de Jordan nilpotente. Entonces $\mathrm{rg}(J^i) = m - i$ para cada $1 \le i \le m$.

Demostración. Se puede verificar inductivamente que $J^i = (e_{i+1}^t \mid \ldots \mid e_m^t \mid 0 \mid \ldots \mid 0)$, donde e_j denota el j-ésimo vector de la base canónica de K^n (es decir que al elevar un bloque de Jordan nilpotente a la i, los unos bajan $i-1$ lugares).

En consecuencia, $\text{rg}(J^i) = \dim < e_{i+1}, \ldots, e_m > = m - i$. □

Este resultado nos permite calcular la cantidad de bloques de cada tamaño que aparecen en una forma de Jordan usando los rangos de las potencias de la matriz.

Proposición 7.12 *Sea $A \in K^{n\times n}$ una forma de Jordan nilpotente de índice k. Entonces el bloque de Jordan más grande que aparece en A es de tamaño $k \times k$. Además, para cada $0 \le i \le k-1$ la cantidad de bloques de Jordan nilpotentes de tamaño mayor que i que aparecen en A es*

$$b_i = \text{rg}(A^i) - \text{rg}(A^{i+1}).$$

En particular, la cantidad de bloques de Jordan que aparecen en A es $b_0 = n - \text{rg}(A) = \dim(\text{Nu}(A))$.

Demostración. Como el índice de nilpotencia de A es k, se tiene que $m_A = X^k$.

Sean $J_1, \ldots, J_r$ los bloques de Jordan que aparecen en A con $J_\ell \in K^{n_\ell \times n_\ell}$ para cada $1 \le \ell \le r$ (ver Definición 7.9). Entonces

$$m_A = \text{mcm}\{m_{J_1}, \ldots, m_{J_r}\} = \text{mcm}\{X^{n_1}, \ldots, X^{n_r}\} = X^{n_1}.$$

Luego, $n_1 = k$, es decir, el bloque de Jordan más grande que aparece en A es de $k \times k$.

Para cada $1 \le i \le k$, sean c_i la cantidad de bloques de Jordan nilpotentes de tamaño $i \times i$ que aparecen en A y b_i la cantidad de bloques de tamaño mayor que i.

Si A está formada por r bloques de Jordan nilpotentes, resulta que $\text{rg}(A) = n - r$ ya que este rango es la suma de los rangos de los distintos bloques de Jordan. En consecuencia, la cantidad de total de bloques de Jordan que forman A es $b_0 = n - \text{rg}(A)$.

Sea $1 \le i \le k-1$. Observamos, por el lema anterior, que para un bloque de Jordan $J \in K^{j\times j}$ se tiene que $J^i = 0$ si $j \le i$ o $\text{rg}(J^i) = j - i$ si $j > i$. Además, $\text{rg}(A^i)$ es la suma de los rangos de los bloques que aparecen en la diagonal. En consecuencia

$$\text{rg}(A^i) - \text{rg}(A^{i+1}) = \sum_{j=i+1}^{k} c_j.(j-i) - \sum_{j=i+2}^{k} c_j.(j-(i+1)) = \sum_{j=i+1}^{k} c_j = b_i. \qquad \square$$

Tenemos entonces el siguiente resultado:

Corolario 7.13 *Sea $A \in K^{n\times n}$ una forma de Jordan nilpotente de índice k. Entonces, para cada $1 \le i \le k$, la cantidad de bloques de Jordan nilpotentes de tamaño $i \times i$ que aparecen en A es*

$$c_i = \text{rg}(A^{i+1}) - 2\,\text{rg}(A^i) + \text{rg}(A^{i-1}).$$

Demostración. Observamos que $c_k = b_{k-1} = \text{rg}(A^{k-1}) - \text{rg}(A^k) = \text{rg}(A^{k+1}) - 2\,\text{rg}(A^k) + \text{rg}(A^{k-1})$, puesto que $A^k = 0$.

Sea $1 \leq i \leq k-1$. Entonces

$$\begin{aligned} c_i &= b_{i-1} - b_i \\ &= (\text{rg}(A^{i-1}) - \text{rg}(A^i)) - (\text{rg}(A^i) - \text{rg}(A^{i+1})) \\ &= \text{rg}(A^{i+1}) - 2\,\text{rg}(A^i) + \text{rg}(A^{i-1}). \end{aligned}$$ □

Ejemplo. Decidir si existe una matriz $A \in \mathbb{R}^{15\times 15}$ tal que $\text{rg}(A) = 10$, $\text{rg}(A^4) = 3$ y $\text{rg}(A^5) = 0$.

Si $\text{rg}(A^5) = 0$ y $\text{rg}(A^4) = 3$, entonces $A^5 = 0$ y $A^4 \neq 0$, de donde A es nilpotente de índice 5. Entonces A es semejante a una forma de Jordan nilpotente J_A cuyo bloque de tamaño más grande es de 5×5.

Ahora, por la Proposición 7.12, J_A tiene $\text{rg}(A^4) - \text{rg}(A^5) = 3$ bloques de 5×5 y, como $J_A \in \mathbb{R}^{15\times 15}$ éstos son los únicos bloques que aparecen.

Pero la cantidad de bloques de Jordan en J_A debe ser $15 - \text{rg}(A) = 15 - 10 = 5$, contradiciendo lo anterior.

Luego, no existe una matriz $A \in \mathbb{R}^{15\times 15}$ que satisfaga las condiciones del enunciado.

El Corolario 7.13 nos permite probar el siguiente resultado:

Lema 7.14 *Sean J y J' formas de Jordan nilpotentes. Si $J \sim J'$, entonces $J = J'$.*

Demostración. Según lo que hemos visto, las cantidades de bloques de cada tamaño que aparecen en una forma de Jordan nilpotente sólo dependen de los rangos de sus potencias. Por otro lado, si $J \sim J'$, entonces para cada $1 \leq i \leq k$, se tiene que $\text{rg}(J^i) = \text{rg}((J')^i)$. En consecuencia, la cantidad de bloques de Jordan de cada tamaño es la misma en J que en J'. Luego, $J = J'$. □

A partir de este lema, se deduce la unicidad de la forma de Jordan en el caso nilpotente:

Teorema 7.15 *Sea V un K-espacio vectorial de dimensión n. Sea $f : V \to V$ una transformación lineal nilpotente. Entonces existe una única forma de Jordan nilpotente $J \in K^{n\times n}$ tal que $|f|_B = J$ para alguna base B de V.*

Demostración. Si B y B' son bases de V tales que $|f|_B = J$ y $|f|_{B'} = J'$ con J y J' formas de Jordan nilpotentes, entonces $J \sim J'$ y, por el lema anterior, $J = J'$. □

Por lo tanto, se obtiene este resultado sobre semejanza de matrices nilpotentes:

Teorema 7.16 *Sean $A, B \in K^{n\times n}$ matrices nilpotentes. Sean J y J' formas de Jordan nilpotentes tales que $A \sim J$ y $B \sim J'$. Entonces*

$$A \sim B \iff J = J'.$$

Demostración.

($\Rightarrow$) Si $A \sim B$, como por hipótesis $A \sim J$ y $B \sim J'$, teniendo en cuenta que $\sim$ es una relación de equivalencia, resulta que $J \sim J'$. Luego, por el Lema 7.14, $J = J'$.

($\Leftarrow$) Si $J = J'$, siendo $A \sim J$, $B \sim J'$ y por ser $\sim$ una relación de equivalencia, se deduce que $A \sim B$. □

Ejemplos.

1. Sean $A, B \in \mathbb{R}^{4\times 4}$ dos matrices tales que $m_A = m_B = X^3$. Probar que $A \sim B$.

   Por el teorema anterior, el problema es equivalente a probar que A y B tienen la misma forma de Jordan. Luego, basta ver que existe una única forma de Jordan nilpotente $J \in \mathbb{R}^{4\times 4}$ tal que $m_J = X^3$.

   Si $m_J = X^3$, entonces J tiene (al menos) un bloque de Jordan nilpotente de 3×3. Como $J \in \mathbb{R}^{4\times 4}$, la única posibilidad es entonces que J esté formada por un bloque de 3×3 y otro de 1×1, es decir

$$J = \left(\begin{array}{ccc|c} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & \boxed{0} \end{array}\right).$$

2. Sean $A, B \in \mathbb{R}^{7\times 7}$ las matrices

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

   Observamos que $\mathcal{X}_A = X^7 = \mathcal{X}_B$, $m_A = X^3 = m_B$ y $\mathrm{rg}(A) = \mathrm{rg}(B)$, pero $A \not\sim B$ puesto que son dos formas de Jordan nilpotentes distintas.

7.2 Caso general

En esta sección generalizaremos lo anterior para endomorfismos en un K-espacio vectorial de dimensión finita cuyos polinomios minimales se factorizan linealmente (es decir, con todas sus raíces en K).

7.2.1 Forma de Jordan de una transformación lineal

Primero veremos un caso particular, en el que el polinomio minimal del endomorfismo se factoriza linealmente pero tiene una única raíz.

Sea V un K-espacio vectorial de dimensión n y sea $f : V \to V$ una transformación lineal tal que $m_f = (X - \lambda)^k$ para algún $k \le n$. Se tiene entonces que $(f - \lambda.\, id_V)^k = 0$ y $(f - \lambda.\, id_V)^{k-1} \ne 0$, con lo cual $f - \lambda.\, id_V$ es nilpotente de índice k.

Por el Teorema 7.7, existe una base B de V tal que $|f - \lambda.\, id_V|_B \in K^{n\times n}$ es una forma de Jordan nilpotente, es decir,

$$|f - \lambda.\, id_V|_B = \begin{pmatrix} J_1 & 0 & \dots & 0 \\ 0 & J_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & J_r \end{pmatrix}$$

donde, para cada $1 \le i \le r$, $J_i \in K^{n_i\times n_i}$ es un bloque de Jordan nilpotente y $k = n_1 \ge n_2 \ge \ldots \ge n_r$.

Observamos que $|f|_B = |f - \lambda.\, id_V|_B + |\lambda.\, id_V|_B = |f - \lambda.\, id_V|_B + \lambda.\, I_n$, de donde

$$|f|_B = \begin{pmatrix} J(\lambda, n_1) & 0 & \dots & 0 \\ 0 & J(\lambda, n_2) & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & J(\lambda, n_r) \end{pmatrix}$$

donde, para cada $1 \le i \le r$,

$$J(\lambda, n_i) = \begin{pmatrix} \lambda & 0 & \dots & 0 & 0 \\ 1 & \lambda & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & \lambda \end{pmatrix} \in K^{n_i\times n_i}$$

y $k = n_1 \ge n_2 \ge \ldots \ge n_r$.

Esto motiva la siguiente definición:

Definición 7.17 Sea $\lambda \in K$. Se llama *bloque de Jordan asociado al autovalor* λ de tamaño n a la matriz $J(\lambda, n) \in K^{n\times n}$

$$J(\lambda, n) = \begin{pmatrix} \lambda & 0 & \dots & 0 & 0 \\ 1 & \lambda & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & \lambda \end{pmatrix}.$$

La idea de la forma de Jordan general es, como en el caso nilpotente, encontrar una base donde la matriz del endomorfismo considerado tenga una forma particular (bloques de Jordan en la diagonal).

La demostración de la existencia de forma de Jordan se basa en el siguiente lema, que permite descomponer el espacio en suma directa de subespacios invariantes, en cada uno de los cuales la transformación lineal tiene un solo autovalor.

Lema 7.18 *Sea V un K espacio vectorial de dimensión finita. Sea $f : V \to V$ una transformación lineal tal que $m_f = P.\,Q$ con $(P, Q) = 1$. Entonces:*

- $\mathrm{Nu}(P(f))$ *y* $\mathrm{Nu}(Q(f))$ *son subespacios invariantes por f;*
- $V = \mathrm{Nu}(P(f)) \oplus \mathrm{Nu}(Q(f))$*;*
- $m_{f_{|\mathrm{Nu}(P(f))}} = P$ *y* $m_{f_{|\mathrm{Nu}(Q(f))}} = Q$.

Demostración.

- $\mathrm{Nu}(P(f))$ y $\mathrm{Nu}(Q(f))$ son invariantes por f:

  Sea $P = \sum\limits_{i=0}^{r} a_i X^i$ y sea $x \in \mathrm{Nu}(P(f))$. Entonces $P(f)(x) = 0$. Aplicando f se obtiene $f(P(f)(x)) = 0$. Luego
  $$0 = f\Big(\sum_{i=0}^{r} a_i f^i(x)\Big) = \sum_{i=0}^{r} a_i f^{i+1}(x) = \Big(\sum_{i=0}^{r} a_i f^i\Big)(f(x)),$$
  de donde $f(x) \in \mathrm{Nu}(P(f))$. Por lo tanto, $\mathrm{Nu}(P(f))$ es invariante por f.

  De la misma manera, $\mathrm{Nu}(Q(f))$ es invariante por f.

- $V = \mathrm{Nu}(P(f)) \oplus \mathrm{Nu}(Q(f))$:

  Puesto que $(P, Q) = 1$, existen $R, S \in K[X]$ tales que $1 = R.P + S.Q$, de donde
  $$id_V = R(f) \circ P(f) + S(f) \circ Q(f).$$
  Sea $x \in \mathrm{Nu}(P(f)) \cap \mathrm{Nu}(Q(f))$. Entonces
  $$x = id_V(x) = R(f)\big(P(f)(x)\big) + S(f)\big(Q(f)(x)\big) = R(f)(0) + S(f)(0) = 0.$$
  Luego, $\mathrm{Nu}(P(f)) \cap \mathrm{Nu}(Q(f)) = \{0\}$.

  Por otro lado, para cada $x \in V$ se tiene que
  $$x = (R(f) \circ P(f))(x) + (S(f) \circ Q(f))(x).$$
  Ahora, teniendo en cuenta que $Q(f) \circ R(f) = (Q.\,R)(f) = (R.\,Q)(f) = R(f) \circ Q(f)$, resulta que
  $$\begin{aligned} Q(f)((R(f) \circ P(f))(x)) &= (Q(f) \circ R(f) \circ P(f))(x) = R(f)\Big((Q(f) \circ P(f))(x)\Big) = \\ &= R(f)(m_f(f)(x)) = R(f)(0) = 0, \end{aligned}$$

de donde $(R(f) \circ P(f))(x) \in \mathrm{Nu}(Q(f))$. Análogamente, $(S(f) \circ Q(f))(x) \in \mathrm{Nu}(P(f))$. En consecuencia, $\mathrm{Nu}(P(f)) + \mathrm{Nu}(Q(f)) = V$.

- $m_{f_{|\mathrm{Nu}(P(f))}} = P$ y $m_{f_{|\mathrm{Nu}(Q(f))}} = Q$:

  Sean f_1 y f_2 las restricciones de f a $\mathrm{Nu}(P(f))$ y $\mathrm{Nu}(Q(f))$ respectivamente. Como $V = \mathrm{Nu}(P(f)) \oplus \mathrm{Nu}(Q(f))$, se tiene que $m_f = \mathrm{mcm}(m_{f_1}, m_{f_2})$.

  Si $P = \sum\limits_{i=0}^{r} a_i X^i$, para cada $x \in \mathrm{Nu}(P(f))$, se tiene que

$$P(f_1)(x) = \Big(\sum_{i=0}^{r} a_i f_1^i\Big)(x) = \sum_{i=0}^{r} a_i f_1^i(x) = \sum_{i=0}^{r} a_i f^i(x) = P(f)(x) = 0,$$

  con lo cual $m_{f_1} \mid P$. Análogamente, $m_{f_2} \mid Q$.

  Como P y Q son coprimos, resulta que m_{f_1} y m_{f_2} también lo son y, por lo tanto,

$$P.Q = m_f = \mathrm{mcm}(m_{f_1}, m_{f_2}) = m_{f_1}.m_{f_2}$$

  de donde $m_{f_1} = P$ y $m_{f_2} = Q$. □

Definición 7.19 Diremos que $J \in K^{n\times n}$ es una *matriz de Jordan* o una *forma de Jordan* si

$$J = \begin{pmatrix} J_1 & 0 & \dots & 0 \\ 0 & J_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & J_s \end{pmatrix}$$

donde, para cada $1 \le i \le s$, J_i es de la forma

$$J_i = \begin{pmatrix} J(\lambda_i, n_1^{(i)}) & 0 & \dots & 0 \\ 0 & J(\lambda_i, n_2^{(i)}) & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & J(\lambda_i, n_{r_i}^{(i)}) \end{pmatrix} \qquad \text{con } n_1^{(i)} \ge \dots \ge n_{r_i}^{(i)}$$

y $\lambda_i \ne \lambda_j$ para $i \ne j$, *o sea, cada J_i está formada por (varios) bloques de Jordan de autovalor λ_i ubicados sobre la diagonal.*

Demostremos ahora el resultado principal de esta sección:

Teorema 7.20 *Sea V un K-espacio vectorial de dimensión finita. Sea $f : V \to V$ una transformación lineal tal que m_f se factoriza linealmente sobre K. Entonces existe una base B de V tal que $|f|_B$ es una forma de Jordan.*

Con la notación anterior, a una base B con la propiedad del teorema la llamaremos una *base de Jordan para* f y a la matriz $|f|_B$ una *forma de Jordan para* f.

Demostración. Probaremos el teorema por inducción en $n = \dim V$.

Para $n = 1$, no hay nada que hacer.

Supongamos que el teorema vale para toda transformación lineal definida en un K-espacio vectorial de dimensión $m < n$ y sea $f : V \to V$ una transformación lineal definida en un K-espacio vectorial V de dimensión n.

Si $m_f = (X - \lambda)^k$, estamos en el caso analizado al comienzo de esta sección, para el que el teorema vale.

Supongamos entonces que f tiene al menos dos autovalores distintos. Si λ_1 es uno de los autovalores de f, entonces $m_f = (X - \lambda_1)^{k_1}. Q$ con $\text{gr}(Q) \geq 1$ y $((X - \lambda_1)^{k_1}, Q) = 1$. Por el Lema 7.18, $\text{Nu}\big((f - \lambda_1.\, id_V)^{k_1}\big)$ y $\text{Nu}(Q(f))$ son subespacios f-invariantes de V y

$$V = \text{Nu}((f - \lambda_1.\, id_V)^{k_1}) \oplus \text{Nu}(Q(f)).$$

Además, como λ_1 es autovalor de f pero no el único, $\{0\} \subset \text{Nu}\big((f - \lambda_1.id_V)^{k_1}\big) \subset V$ y las inclusiones son estrictas. En particular, $0 < \dim(\text{Nu}((f - \lambda_1.\, id_V)^{k_1})) < \dim V = n$. Entonces también vale $0 < \dim \text{Nu}(Q(f)) < \dim V = n$.

Consideremos las restricciones de f a $\text{Nu}\big((f - \lambda_1.\, id_V)^{k_1}\big)$ y $\text{Nu}(Q(f))$:

$$f_1 : \text{Nu}((f - \lambda_1.\, id_V)^{k_1}) \to \text{Nu}((f - \lambda_1.\, id_V)^{k_1}) \qquad \text{y} \qquad f_2 : \text{Nu}(Q(f)) \to \text{Nu}(Q(f)).$$

Por hipótesis inductiva, existen una base B_1 de $\text{Nu}\big((f - \lambda_1.I)^{k_1}\big)$ y una base B_2 de $\text{Nu}(Q(f))$, tales que $|f_1|_{B_1}$ y $|f_2|_{B_2}$ son formas de Jordan. Entonces, tomando $B = B_1 \cup B_2$ obtenemos una base de V tal que

$$|f|_B = \begin{pmatrix} |f_1|_{B_1} & 0 \\ 0 & |f_2|_{B_2} \end{pmatrix}.$$

Observamos que, de acuerdo al Lema 7.18, $m_{f_1} = (X - \lambda_1)^{k_1}$ y $m_{f_2} = Q$. Entonces $|f_1|_{B_1}$ está formada por bloques de Jordan de autovalor λ_1 y, como λ_1 no es raíz de Q, $|f_2|_{B_2}$ está formada por bloques de Jordan de autovalor λ con $\lambda \neq \lambda_1$. En consecuencia, $|f|_B$ es una forma de Jordan. □

Observemos que, con las notaciones del teorema anterior, si $m_f = \prod_{i=1}^{r} (X - \lambda_i)^{k_i}$ con $\lambda_i \neq \lambda_j$ si $i \neq j$, la demostración nos permite dar una forma constructiva de obtener una forma de Jordan de f. Como

$$V = \text{Nu}((f - \lambda_1.\, id_V)^{k_1}) \oplus \cdots \oplus \text{Nu}((f - \lambda_r.\, id_V)^{k_r}),$$

podemos obtener una forma de Jordan para cada una de las restricciones de f a estos subespacios invariantes $\text{Nu}((f - \lambda_i.\, id_V)^{k_i})$ y las bases de Jordan B_i correspondientes. Entonces $B = B_1 \cup \cdots \cup B_r$ resulta ser una base de V y $|f|_B$ resulta una forma de Jordan de f.

El resultado del teorema se puede enunciar también para matrices complejas.

Teorema 7.21 *Sea $A \in \mathbb{C}^{n\times n}$. Entonces A es semejante a una forma de Jordan.*

A una base B de K^n tal que $|f_A|_B$ es una forma de Jordan, la llamaremos una *base de Jordan para A*, y a la matriz $|f_A|_B$ una *forma de Jordan para A*.

Demostración. Consideremos la transformación lineal $f_A : \mathbb{C}^n \to \mathbb{C}^n$ definida por $f_A(x) = A.x$. Observamos que $m_{f_A} = m_A \in \mathbb{C}[X]$ se factoriza linealmente en $\mathbb{C}[X]$. Por el teorema anterior, existe entonces una base de $\mathbb{C}^n$ tal que la matriz $J_A = |f_A|_B$ es una forma de Jordan, semejante a A. □

Ejemplo. Hallar una forma de Jordan semejante a A y una base de Jordan para A, siendo

$$A = \begin{pmatrix} 1 & 0 & 0 & 2 \\ 2 & -1 & 0 & 2 \\ 2 & 0 & -1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix} \in \mathbb{C}^{4\times 4}.$$

Se tiene que $\mathcal{X}_A = (X-1)^2(X+1)^2$, luego los autovalores de A son 1 y -1.

Calculemos $m_A = \text{mcm}\{m_{e_1}, m_{e_2}, m_{e_3}, m_{e_4}\}$, donde $\{e_1, e_2, e_3, e_4\}$ es la base canónica de $\mathbb{C}^4$:

Puesto que $A.e_1 = e_1 + 2e_2 + 2e_3$ (luego $\{e_1, A.e_1\}$ es l.i.) y $A^2e_1 = e_1$, se tiene que $m_{e_1} = X^2 - 1$.

Por otro lado, $A.e_2 = -e_2$, con lo cual $m_{e_2} = X + 1$. De la misma manera, $m_{e_3} = X + 1$.

Finalmente, para e_4 tenemos que $A.e_4 = 2e_1 + 2e_2 + 2e_3 + e_4$ (y entonces $\{e_4, A.e_4\}$ es l.i.) y $A^2e_4 = 4e_1 + 4e_2 + 4e_3 + e_4 = 2.\,A.e_4 - e_4$. Luego $m_{e_4} = X^2 - 2X + 1$.

En consecuencia, $m_A = \text{mcm}\{X^2 - 1, X + 1, X^2 - 2X + 1\} = (X-1)^2(X+1)$.

Sabemos entonces que

$$\mathbb{C}^4 = \text{Nu}((A-I)^2) \oplus \text{Nu}(A+I),$$

y si $f_1 : \text{Nu}((A-I)^2) \to \text{Nu}((A-I)^2)$ y $f_2 : \text{Nu}(A+I) \to \text{Nu}(A+I)$ son las restricciones de f_A a $\text{Nu}((A-I)^2)$ y $\text{Nu}(A+I)$ respectivamente, una forma de Jordan para A es

$$J_A = \begin{pmatrix} J_1 & 0 \\ 0 & J_2 \end{pmatrix}$$

donde J_1 y J_2 son formas de Jordan de f_1 y f_2 respectivamente. Más aún, si B_1 y B_2 son bases de Jordan para f_1 y f_2, entonces $B = B_1 \cup B_2$ es una base de Jordan para A.

- Base y forma de Jordan de $f_1 : \text{Nu}((A-I)^2) \to \text{Nu}((A-I)^2)$.

  Se tiene que $m_{f_1} = (X-1)^2$, luego $f_1 - id_{\text{Nu}((A-I)^2)}$ es nilpotente de índice 2. Además

$$A - I = \begin{pmatrix} 0 & 0 & 0 & 2 \\ 2 & -2 & 0 & 2 \\ 2 & 0 & -2 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{y} \quad (A-I)^2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ -4 & 4 & 0 & 0 \\ -4 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

de donde $\mathrm{Nu}(A-I) = <(1,1,1,0)>$ y $\mathrm{Nu}((A-I)^2) = <(1,1,1,0),(0,0,0,1)>$. Consideramos el vector $e_4 = (0,0,0,1)$ que extiende una base de $\mathrm{Nu}(A-I)$ a una de $\mathrm{Nu}((A-I)^2)$. Obtenemos

$$\begin{array}{ccccc} \{0\} & \subsetneq & \mathrm{Nu}(A-I) & \subsetneq & \mathrm{Nu}((A-I)^2) \\ & & (A-I).e_4 & & e_4 \end{array}$$

Luego, una base de Jordan para f_1 es $B_1 = \{e_4, (A-I).e_4\} = \{(0,0,0,1),(2,2,2,0)\}$ y su forma de Jordan es

$$|f_1|_{B_1} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

- Base y forma de Jordan de $f_2 : \mathrm{Nu}(A+I) \to \mathrm{Nu}(A+I)$.

  Sabemos que $m_{f_2} = X-1$, luego f_2 es diagonalizable. Se tiene que

$$A+I = \begin{pmatrix} 2 & 0 & 0 & 2 \\ 2 & 0 & 0 & 2 \\ 2 & 0 & 0 & 2 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

  Luego, una base de $\mathrm{Nu}(A+I)$ (que será también una base de Jordan para f_2) es $B_2 = \{(0,1,0,0),(0,0,1,0)\}$ y la forma de Jordan de f_2 es

$$|f_2|_{B_2} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}.$$

En consecuencia, una base de Jordan para A es

$$B = B_1 \cup B_2 = \{(0,0,0,1),(2,2,2,0),(0,1,0,0),(0,0,1,0)\}$$

y una forma de Jordan semejante a A es

$$J_A = \left(\begin{array}{cc|c|c} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ \hline 0 & 0 & -1 & 0 \\ \hline 0 & 0 & 0 & -1 \end{array}\right).$$

7.2.2 Unicidad de la forma de Jordan

Veremos ahora que la forma de Jordan asociada a una transformación lineal $f : V \to V$, donde V es un K-espacio vectorial de dimensión finita, es única salvo por el orden en que aparecen los bloques de Jordan correspondientes a autovalores distintos.

Teorema 7.22 *Sea V un K espacio vectorial de dimensión n, y sea $f : V \to V$ una transformación lineal tal que m_f se factoriza linealmente sobre K. Entonces existe una única forma de Jordan $J \in K^{n\times n}$ (salvo por el orden de sus bloques) tal que para alguna base B de V, $|f|_B = J$.*

Demostración. Sea $B = \{v_1, \dots, v_n\}$ una base de V tal que $|f|_B$ es una forma de Jordan $J \in K^{n\times n}$,

$$J = \begin{pmatrix} J_1(\lambda_1) & 0 & \dots & 0 \\ 0 & J_2(\lambda_2) & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & J_s(\lambda_s) \end{pmatrix},$$

donde para cada $1 \le i \le s$, $J_i(\lambda_i) \in K^{d_i\times d_i}$ denota la matriz formada por todos los bloques de Jordan de autovalor λ_i que aparecen en J y $\lambda_i \ne \lambda_j$ si $i \ne j$.

Se tiene que $\mathcal{X}_f = \mathcal{X}_J = \prod\limits_{i=1}^{s} (X - \lambda_i)^{d_i}$. Entonces $\lambda_1, \dots, \lambda_s$ son los autovalores de f y, para cada $1 \le i \le s$, $d_i = \text{mult}(\lambda_i, \mathcal{X}_f)$.

Ahora, para cada $1 \le i \le s$, se tiene que $m_{J_i(\lambda_i)} = (X - \lambda_i)^{k_i}$ para algún $1 \le k_i \le d_i$. Entonces $m_f = m_J = \prod\limits_{i=1}^{s} (X - \lambda_i)^{k_i}$, con lo que, para cada $1 \le i \le s$, $k_i = \text{mult}(\lambda_i, m_f)$.

Sea λ uno de los autovalores de f y sea $k = \text{mult}(\lambda, m_f)$. Sin pérdida de generalidad, supongamos que $\lambda = \lambda_1$. Observamos que

$$J - \lambda.\, I_n = \begin{pmatrix} J_1(0) & 0 & \dots & 0 \\ 0 & J_2(\lambda_2 - \lambda) & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & J_s(\lambda_s - \lambda) \end{pmatrix},$$

de donde

$$(J - \lambda.\, I_n)^k = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & (J_2(\lambda_2 - \lambda))^k & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & (J_s(\lambda_s - \lambda))^k \end{pmatrix}$$

Puesto que $(J_i(\lambda_i - \lambda))^k$ es inversible para cada $2 \le i \le s$, entonces $\text{Nu}\big((J - \lambda I_n)^k\big) = \, < e_1, \dots, e_{d_1} >$. Teniendo en cuenta que $J = |f|_B$, resulta que $J - \lambda.I_n = |f - \lambda.id_V|_B$, con lo que

$$\text{Nu}((f - \lambda.\, id_V)^k) = \, < v_1, \dots, v_{d_1} >.$$

Consideremos la restricción

$$f_1 = (f - \lambda.\, id_V)_{|_{\text{Nu}((f-\lambda.\, id_V)^k)}} : \text{Nu}((f - \lambda.\, id_V)^k) \to \text{Nu}((f - \lambda.\, id_V)^k).$$

La restricción f_1 resulta una transformación lineal nilpotente y, por lo tanto, tiene una única forma de Jordan nilpotente J_1 asociada.

Sea $B_1 = \{v_1, \dots, v_{d_1}\}$, que como vimos, es una base de $\text{Nu}\big((f - \lambda.\, id_V)^k\big)$. Observamos que

$$|f_1|_{B_1} = \Big| f_{|_{\text{Nu}(f-\lambda.\, id_V)^k}} \Big|_{B_1} - \lambda.I_{d_1} = J_1(0).$$

Como $J_1(0)$ es una forma de Jordan nilpotente, debe ser *la* forma de Jordan J_1 de f_1.

En consecuencia, $J_1(\lambda_1) = J_1 + \lambda_1.I_{d_1}$ está unívocamente determinada por f. (Notar que el subespacio invariante $\mathrm{Nu}((f - \lambda.\, id_V)^k)$ y la restricción $f_1 = (f - \lambda.\, id_V)_{|_{\mathrm{Nu}((f-\lambda.\, id_V)^k)}}$ sólo dependen de f y no de una base.)

Haciendo lo mismo para cada λ_i con $1 \le i \le s$, resulta que, $J_i(\lambda_i) \in K^{d_i \times d_i}$ satisface:

$$J_i(\lambda_i) = J_i + \lambda_i.\, I_{d_i},$$

donde $d_i = \mathrm{mult}(\lambda_i, \mathcal{X}_f)$ y, si $k_i = \mathrm{mult}(\lambda_i, m_f)$, J_i es la forma de Jordan nilpotente de la restricción

$$(f - \lambda_i.\, id_V)_{|_{\mathrm{Nu}((f-\lambda_i.\, id_V)^{k_i})}} : \mathrm{Nu}((f - \lambda_i.\, id_V)^{k_i}) \to \mathrm{Nu}((f - \lambda_i.\, id_V)^{k_i}).$$

Por lo tanto, la forma de Jordan de f está unívocamente determinada por f (salvo el orden de los bloques correspondientes a los distintos autovalores de f). $\square$

El hecho que todo polinomio se factorice linealmente en $\mathbb{C}[X]$ nos permite demostrar el siguiente resultado sobre semejanza de matrices en $\mathbb{C}^{n\times n}$.

Teorema 7.23 *Sean $A, B \in \mathbb{C}^{n\times n}$, y sean J_A y J_B las formas de Jordan de A y B respectivamente. Entonces*

$$A \sim B \iff J_A = J_B \quad \textit{(salvo el orden de los bloques).}$$

Demostración.

($\Rightarrow$) Sabemos que $A \sim J_A$ y $B \sim J_B$. Si $A \sim B$, como $\sim$ es una relación de equivalencia, resulta que $J_A \sim J_B$. Entonces existe una transformación lineal $f : K^n \to K^n$ y bases B_1 y B_2 de K^n tales que $|f|_{B_1} = J_A$ y $|f|_{B_2} = J_B$.

Por el teorema anterior, la forma de Jordan de una transformación lineal es única salvo el orden de los bloques correspondientes a los distintos autovalores. Luego $J_A = J_B$ salvo el orden de los bloques.

($\Leftarrow$) Se deduce inmediatamente de que $\sim$ es una relación de equivalencia. $\square$

Ejemplo. Sean $A, B \in \mathbb{C}^{3\times 3}$. Probar que $A \sim B \iff \mathcal{X}_A = \mathcal{X}_B$ y $m_A = m_B$. ¿Vale el mismo resultado para matrices en $\mathbb{C}^{4\times 4}$?

Ya sabemos que vale ($\Rightarrow$). Probemos la otra implicación. Por el Teorema 7.23, A y B son semejantes si tienen la misma forma de Jordan (salvo el orden de los distintos autovalores). Luego, basta ver que la forma de Jordan de una matriz en $\mathbb{C}^{3\times 3}$ queda unívocamente determinada por su polinomio característico y su polinomio minimal.

Sea $J \in \mathbb{C}^{3\times 3}$ una forma de Jordan. Entonces $\mathcal{X}_J$ es un polinomio mónico de grado 3 en $\mathbb{C}[X]$, luego puede escribirse de alguna de las siguientes formas:

(i) $\mathcal{X}_J = (X - \lambda_1)(X - \lambda_2)(X - \lambda_3)$ con $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{C}$ y $\lambda_i \neq \lambda_j$ si $i \neq j$.

(ii) $\mathcal{X}_J = (X - \lambda_1)^2(X - \lambda_2)$ con $\lambda_1, \lambda_2 \in \mathbb{C}$ y $\lambda_1 \neq \lambda_2$.

(iii) $\mathcal{X}_J = (X - \lambda)^3$ con $\lambda \in \mathbb{C}$.

Para cada una de las opciones anteriores, veremos que existe una única J para cada polinomio minimal posible.

(i) Si $\mathcal{X}_J = (X - \lambda_1)(X - \lambda_2)(X - \lambda_3)$ con $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{C}$ y $\lambda_i \neq \lambda_j$ si $i \neq j$, J es diagonalizable y, por lo tanto, tiene un bloque de 1×1 para cada autovalor. Luego (salvo el orden de los autovalores)

$$J = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}.$$

(ii) Si $\mathcal{X}_J = (X - \lambda_1)^2(X - \lambda_2)$ con $\lambda_1, \lambda_2 \in \mathbb{C}$ y $\lambda_1 \neq \lambda_2$, entonces $m_J = (X - \lambda_1)(X - \lambda_2)$ o $m_J = \mathcal{X}_J$.

- Si $m_J = (X - \lambda_1)(X - \lambda_2)$, J es diagonalizable y, por lo tanto, cada bloque en J es de 1×1. Luego
$$J = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{pmatrix},$$
salvo el orden de los autovalores.

- Si $m_J = (X - \lambda_1)^2(X - \lambda_2)$, entonces J tiene un bloque de 2×2 con autovalor λ_1 y uno de 1×1 con autovalor λ_2. Luego (salvo el orden de los autovalores),
$$J = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 1 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{pmatrix}.$$

(iii) Si $\mathcal{X}_J = (X - \lambda)^3$ con $\lambda \in \mathbb{C}$, entonces $m_J = (X - \lambda)^k$ para $k = 1, 2$ o 3.

- Si $m_J = (X - \lambda)$, entonces J es diagonalizable y sólo tiene bloques de 1×1, es decir,
$$J = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix}.$$

- Si $m_J = (X - \lambda)^2$, entonces el bloque más grande de J es de 2×2, luego sólo puede tener un bloque de 2×2 y uno de 1×1:
$$J = \begin{pmatrix} \lambda & 0 & 0 \\ 1 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix}.$$

- Si $m_J = (X-\lambda)^3$, entonces J tiene un bloque de 3×3 y, por lo tanto
$$J = \begin{pmatrix} \lambda & 0 & 0 \\ 1 & \lambda & 0 \\ 0 & 1 & \lambda \end{pmatrix}.$$

El resultado no vale en $\mathbb{C}^{4\times 4}$. Por ejemplo, las matrices
$$A = \begin{pmatrix} 0&0&0&0\\ 1&0&0&0\\ 0&0&0&0\\ 0&0&1&0 \end{pmatrix} \qquad \text{y} \qquad B = \begin{pmatrix} 0&0&0&0\\ 1&0&0&0\\ 0&0&0&0\\ 0&0&0&0 \end{pmatrix}$$
verifican $\mathcal{X}_A = \mathcal{X}_B = X^4$ y $m_A = m_B = X^2$, pero $A \not\sim B$, porque son dos formas de Jordan distintas.

7.3 Aplicación: Cálculo de las potencias de una matriz

En diversas aplicaciones, por ejemplo en la resolución de sistemas de ecuaciones diferenciales lineales de primer orden, surge la necesidad de calcular las potencias de una matriz dada. En esta sección veremos que esto puede hacerse utilizando la forma de Jordan de la matriz.

En primer lugar, observamos que es posible calcular las potencias de una matriz a partir de las potencias de una matriz semejante:

Si $A \sim B$, existe $C \in GL(n,K)$ tal que $A = C.\,B.\,C^{-1}$. Entonces, para cada $k\in\mathbb{N}$, $A^k = C.\,B^k.\,C^{-1}$ (ver Observación 6.26).

A partir de esta observación vemos que, si una matriz $A \in K^{n\times n}$ es diagonalizable, entonces se puede calcular fácilmente A^k para cada $k\in\mathbb{N}$: Basta hallar $C\in GL(n,K)$ y $D \in K^{n\times n}$ diagonal tales que $A = C.\,D.\,C^{-1}$ y tener en cuenta que
$$D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix} \quad\Longrightarrow\quad D^k = \begin{pmatrix} \lambda_1^k & 0 & \dots & 0 \\ 0 & \lambda_2^k & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n^k \end{pmatrix} \quad \forall\, k\in\mathbb{N}.$$
Utilizando la igualdad $A^k = C.\,D^k.\,C^{-1}$ se obtienen las potencias de A.

Consideremos ahora el caso de una matriz $A\in K^{n\times n}$ que no sea diagonalizable. Si
$$A \sim M = \begin{pmatrix} M_1 & 0 & \dots & 0 \\ 0 & M_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & M_r \end{pmatrix} \qquad \text{con } M_i \in K^{n_i\times n_i}\ (1\le i\le r)$$

existe $C \in GL(n,K)$ tal que $A = C.M.C^{-1}$. Entonces, para cada $k \in \mathbb{N}$, $A^k = C.\,M^k.\,C^{-1}$ y, multiplicando por bloques, resulta que

$$M^k = \begin{pmatrix} M_1^k & 0 & \dots & 0 \\ 0 & M_2^k & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & M_r^k \end{pmatrix}.$$

Por lo visto en las secciones anteriores, si m_A se factoriza linealmente en K, se puede hallar una matriz M (la forma de Jordan de A) en la que cada M_i es un bloque de Jordan de autovalor λ_i para algún $\lambda_i \in K$. Luego, para calcular las potencias de A basta poder calcular $J(\lambda,m)^k$.

Se puede probar inductivamente (ver Lema 7.11) que

$$J(0,m)^k = \begin{pmatrix} 0_{k\times(m-k)} & 0_{k\times k} \\ & \\ I_{m-k} & 0_{(m-k)\times k} \end{pmatrix}.$$

Si $\lambda \neq 0$, escribimos $J(\lambda,m) = \lambda.\,I_m + J(0,m)$. Puesto que $\lambda.\,I_m$ y $J(0,m)$ conmutan, podemos calcular las potencias de $\lambda.\,I_m + J(0,m)$ aplicando la fórmula del binomio de Newton:

$$\begin{aligned} J(\lambda,m)^k &= (\lambda.\,I_m + J(0,m))^k = \sum_{i=0}^{k} \binom{k}{i} (\lambda I_m)^{k-i}.\,J(0,m)^i \\ &= \sum_{i=0}^{k} \binom{k}{i} \lambda^{k-i}.\begin{pmatrix} 0_{i\times(m-i)} & 0_{i\times i} \\ I_{m-i} & 0_{(m-i)\times i} \end{pmatrix} \\ &= \sum_{i=0}^{k} \begin{pmatrix} 0_{i\times(m-i)} & 0_{i\times i} \\ \binom{k}{i}\lambda^{k-i}.\,I_{m-i} & 0_{(m-i)\times i} \end{pmatrix} \\ &= \begin{pmatrix} \lambda^k & 0 & \dots & \dots & 0 \\ \binom{k}{1}\lambda^{k-1} & \lambda^k & \ddots & & \vdots \\ \binom{k}{2}\lambda^{k-2} & \binom{k}{1}\lambda^{k-1} & \lambda^k & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \binom{k}{m-1}\lambda^{k-(m-1)} & \dots & \binom{k}{2}\lambda^{k-2} & \binom{k}{1}\lambda^{k-1} & \lambda^k \end{pmatrix}. \end{aligned}$$

(Consideramos $\binom{k}{h} = 0$ si $h > k$.)

Esto nos da una fórmula para el cálculo de las potencias de un bloque de Jordan de autovalor λ. Usando las observaciones anteriores podemos calcular las potencias de A.

7.4 Ejercicios

Ejercicio 1. Dadas las matrices A y A' en $K^{n\times n}$

$$A = \begin{pmatrix} 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix} \qquad \text{y} \qquad A' = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}$$

i) Probar que ambas son nilpotentes y que A es semejante a A'.

ii) Dar bases B y B' de $\mathbb{R}_{n-1}[X]$ tal que la matriz de la derivación en la base B sea A y en la base B' sea A'.

iii) Sea B una base de K^n y sea $f : K^n \to K^n$ tal que $|f|_B = A$. Probar que no existen subespacios propios f-invariantes S y T de K^n tales que $K^n = S \oplus T$.

Ejercicio 2. Sean A_i $(1 \le i \le 6)$ matrices en $\mathbb{C}^{8\times 8}$ nilpotentes tales que $m_{A_i} = X^3$ $(1 \le i \le 6)$. ¿Es cierto que necesariamente dos de estas matrices son semejantes?

Ejercicio 3. Sean $A, B \in \mathbb{C}^{6\times 6}$ son matrices nilpotentes tales que $m_A = m_B$ y $\text{rg}(A) = \text{rg}(B)$. Probar que A y B son semejantes. ¿Es cierto esto en $\mathbb{C}^{7\times 7}$?

Ejercicio 4. Hallar la forma y una base de Jordan de la matriz $A \in \mathbb{C}^{9\times 9}$:

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

Ejercicio 5. Hallar la forma y una base de Jordan de la matriz $A = (a_{ij}) \in \mathbb{C}^{n\times n}$ donde

$$a_{ij} = \begin{cases} 0 & \text{si } i \le j \\ 1 & \text{si } i > j \end{cases}$$

Ejercicio 6.

i) Decidir si existe $A \in \mathbb{C}^{8\times 8}$ nilpotente tal que $\text{rg}(A) = 6$, $\text{rg}(A^2) = 4$, $\text{rg}(A^3) = 3$, $\text{rg}(A^4) = 1$ y $\text{rg}(A^5) = 0$ simultáneamente. En caso afirmativo, exhibir una.

ii) Decidir si existe $A \in \mathbb{R}^{16\times 16}$ tal que $m_A(X) = X^5$, $\text{rg}(A) = 9$, $\text{rg}(A^2) = 5$, $\text{rg}(A^3) = 3$, $\text{rg}(A^4) = 1$ y $\text{rg}(A^5) = 0$ simultáneamente. En caso afirmativo, exhibir una.

Ejercicio 7. Sea $f : \mathbb{C}^7 \to \mathbb{C}^7$ una transformación lineal y sea B una base de $\mathbb{C}^7$ tal que

$$|f|_B = \begin{pmatrix} 2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 3 \end{pmatrix}.$$

i) Hallar $\mathcal{X}_f$ y m_f.

ii) Sea λ un autovalor de f tal que $\text{mult}(\lambda, \mathcal{X}_f) = m$. Se definen $E_\lambda = \{v \in \mathbb{C}^7 / \, f(v) = \lambda.v\}$ y $V_\lambda = \{v \in \mathbb{C}^7 / \, (\lambda.\, Id - f)^m(v) = 0\} = \text{Nu}\big((\lambda.\, Id - f)^m\big)$.

¿Para qué autovalores λ de f se tiene que $E_\lambda = V_\lambda$?

iii) Para cada autovalor λ de f, ¿cuál es la menor potencia k tal que $V_\lambda = \text{Nu}\big((\lambda.\, Id - f)^k\big)$?

iv) Si λ es un autovalor de f, se nota f_λ a la restricción de f a V_λ. Calcular $\dim(\text{Im}(f_\lambda))$ y $\dim(\text{Im}(f_\lambda^2))$ para cada λ.

Ejercicio 8. Sea V un K-espacio vectorial, sea $f : V \to V$ una transformación lineal y sea $P \in K[X]$.

i) Probar que $\text{Nu}(P(f))$ e $\text{Im}(P(f))$ son subespacios invariantes por f.

ii) Probar que si un autovalor λ de f es raíz de P, entonces $E_\lambda \subseteq \text{Nu}(P(f))$.

iii) Probar que si un autovalor λ de f no es raíz de P, entonces $E_\lambda \subseteq \text{Im}(P(f))$.

Ejercicio 9. Hallar la forma y una base de Jordan de $A \in \mathbb{C}^{n\times n}$ en cada uno de los siguientes casos:

$$\begin{pmatrix} 1 & 1 & -1 \\ -3 & -3 & 3 \\ -2 & -2 & 2 \end{pmatrix} \; ; \; \begin{pmatrix} 3 & 0 & 8 \\ 3 & -1 & 6 \\ -2 & 0 & -5 \end{pmatrix} \; ; \; \begin{pmatrix} -4 & 2 & 10 \\ -4 & 3 & 7 \\ -3 & 1 & 7 \end{pmatrix} \; ; \; \begin{pmatrix} -2 & 8 & 6 \\ -4 & 10 & 6 \\ 4 & -8 & -4 \end{pmatrix}$$

$$\begin{pmatrix} 1 & -1 & 2 \\ 3 & -3 & 6 \\ 2 & -2 & 4 \end{pmatrix} \; ; \; \begin{pmatrix} 3 & 1 & 0 & 0 \\ -4 & -1 & 0 & 0 \\ 7 & 1 & 2 & 1 \\ -17 & -6 & -1 & 0 \end{pmatrix} \; ; \; \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Ejercicio 10. Sea $A \in \mathbb{R}^{4\times 4}$ la matriz

$$A = \begin{pmatrix} 3 & 0 & 8 & a \\ 3 & -1 & 6 & 0 \\ -2 & 0 & -5 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

i) Para cada $a \in \mathbb{R}$, calcular $\mathcal{X}_A$, m_A y hallar la forma de Jordan de A.

ii) Para $a = 2$, hallar una base de Jordan para A.

Ejercicio 11. Sea $V \subseteq C^\infty(\mathbb{R})$ el subespacio $V = < e^x, x.\,e^x, x^2.e^x, e^{2x} >$. Sea $\delta : V \to V$ la transformación lineal definida por $\delta(f) = f'$. Hallar la forma y una base de Jordan para δ.

Ejercicio 12. Sean $A\,,\, B \in \mathbb{C}^{\,4\times 4}$ las matrices

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

Decidir si A y B son semejantes.

Ejercicio 13. Sean $A, B \in \mathbb{C}^{5\times 5}$ tales que $\mathcal{X}_A = \mathcal{X}_B = (X-1)^3.(X-3)^2$ y $m_A = m_B$. Decidir si, necesariamente, A es semejante a B.

Ejercicio 14. Encontrar todas las formas de Jordan posibles de la matriz $A \in \mathbb{C}^{n\times n}$ en cada uno de los siguientes casos:

i) $\mathcal{X}_A(X) = (X-2)^4(X-3)^2$; $\quad m_A(X) = (X-2)^2(X-3)^2$

ii) $\mathcal{X}_A(X) = (X-7)^5$; $\quad m_A(X) = (X-7)^2$

iii) $\mathcal{X}_A(X) = (X-2)^7$; $\quad m_A(X) = (X-2)^3$

iv) $\mathcal{X}_A(X) = (X-3)^4(X-5)^4$; $\quad m_A(X) = (X-3)^2(X-5)^2$

Ejercicio 15. Sea $A \in \mathbb{C}^{15\times 15}$ una matriz con autovalores λ_1, λ_2 y λ_3 y que cumple, simultáneamente:

$$\begin{array}{llll} \mathrm{rg}(A-\lambda_1.I) = 13, & \mathrm{rg}(A-\lambda_1.I)^2 = 11, & \mathrm{rg}(A-\lambda_1.I)^3 = 10, & \mathrm{rg}(A-\lambda_1.I)^4 = 10, \\ \mathrm{rg}(A-\lambda_2.I) = 13, & \mathrm{rg}(A-\lambda_2.I)^2 = 11, & \mathrm{rg}(A-\lambda_2.I)^3 = 10, & \mathrm{rg}(A-\lambda_2.I)^4 = 9, \\ \mathrm{rg}(A-\lambda_3.I) = 13, & \mathrm{rg}(A-\lambda_3.I)^2 = 12, & \mathrm{rg}(A-\lambda_3.I)^3 = 11. & \end{array}$$

Hallar su forma de Jordan.

Ejercicio 16. Dar la forma de Jordan de una matriz $A \in \mathbb{C}^{14\times 14}$ que verifica, simultáneamente:

$m_A = (X-\lambda_1)^2.(X-\lambda_2).(X-\lambda_3)^2.(X-\lambda_4)^3$ (con $\lambda_i \neq \lambda_j$ si $i \neq j$),

$\mathrm{rg}(A-\lambda_1.I) = 11$, $\mathrm{rg}(A-\lambda_1.I)^2 = 10$, $\mathrm{rg}(A-\lambda_3.I) = 12$, $\mathrm{rg}(A-\lambda_3.I)^2 = 10$ y

$\mathrm{rg}(A-\lambda_4.I) = 13$.

Ejercicio 17. Hallar la forma de Jordan de la matriz $A \in \mathbb{C}^{n\times n}$:

$$\begin{pmatrix} 2 & 3 & 4 & 5 & \dots & n+1 \\ 0 & 2 & 3 & 4 & \dots & n \\ 0 & 0 & 2 & 3 & \dots & n-1 \\ 0 & 0 & 0 & 2 & \dots & n-2 \\ & \vdots & & & & \vdots \\ 0 & 0 & 0 & 0 & \dots & 2 \end{pmatrix}$$

Ejercicio 18. Sean $x, y \in \mathbb{C}^n$ y $A \in \mathbb{C}^{n\times n}$, $A = (a_{ij})$ con $a_{ij} = x_i.y_j$.

i) Calcular todos los autovalores y autovectores de A.

ii) Calcular las posibles formas de Jordan de A.

Ejercicio 19. Sea $A \in \mathbb{C}^{n\times n}$. Probar que A y A^t son semejantes.

Ejercicio 20. Sea $A \in \mathbb{C}^{6\times 6}$ una matriz tal que $m_A = X^6$ y sea $\{v_1, v_2, v_3, v_4, v_5, v_6\}$ una base de Jordan para A. Calcular la forma y una base de Jordan para las matrices A^2, A^3, A^4 y A^5.

Ejercicio 21. Dada la matriz $A = \begin{pmatrix} 5 & 1 & 4 \\ -1 & 3 & -1 \\ 0 & 0 & 1 \end{pmatrix}$, encontrar $B \in \mathbb{Q}^{3\times 3}$ tal que $B^2 = A$.

Ejercicio 22. Sean α, $\beta \in \mathbb{R}$. Se define la sucesión $\{a_n\}_{n\in\mathbb{N}_0}$ de la siguiente manera:

$$\begin{cases} a_0 = \alpha\,,\ a_1 = \beta \\ a_{n+2} = 4.a_{n+1} - 4.a_n \quad \forall\, n \in \mathbb{N}_0 \end{cases}$$

Hallar una fórmula general para el término a_n, $\forall\, n \in \mathbb{N}_0$.

Sugerencia: Ver el razonamiento del Ejercicio 8 de la Sección 6.5 e intentar modificarlo convenientemente utilizando el cálculo de potencias de una matriz de la Sección 7.3.

Ejercicio 23. Resolver el siguiente sistema de ecuaciones diferenciales

$$\begin{cases} x_1'(t) &=& 3\,x_1(t) - x_2(t) \\ x_2'(t) &=& x_1(t) + x_2(t) \\ x_3'(t) &=& -x_2(t) + 2\,x_3(t) \end{cases}$$

con condiciones iniciales $x_1(0) = 1$, $x_2(0) = 2$, $x_3(0) = 1$.

Sugerencia: Ver Ejercicio 9 de la Sección 6.5.

Capítulo 8

Espacios vectoriales con producto interno

En este capítulo, se generalizarán las nociones geométricas de distancia y perpendicularidad, conocidas en $\mathbb{R}^2$ y en $\mathbb{R}^3$, a otros espacios vectoriales. Sólo se considerarán espacios vectoriales sobre $\mathbb{R}$ o sobre $\mathbb{C}$.

8.1 Producto interno

Algunas nociones geométricas en $\mathbb{R}^2$ y en $\mathbb{R}^3$ pueden definirse a partir del producto escalar. La definición que sigue es una generalización del producto escalar a otros espacios vectoriales.

8.1.1 Definición y ejemplos

Definición 8.1 Sea V un espacio vectorial sobre $\mathbb{R}$ (respectivamente $\mathbb{C}$). Un *producto interno sobre* V es una función $\Phi : V \times V \to \mathbb{R}$ (respectivamente $\mathbb{C}$) que cumple:

i) Para cada $\alpha \in \mathbb{R}$ (respectivamente $\mathbb{C}$), y $v, w, z \in V$

- $\Phi(v + w, z) = \Phi(v, z) + \Phi(w, z)$
- $\Phi(\alpha.v, z) = \alpha.\,\Phi(v, z)$

ii) $\Phi(v, w) = \overline{\Phi(w, v)} \quad \forall\, v, w \in V$.
(Notar que esta condición implica que para cada $v \in V$, $\Phi(v, v) = \overline{\Phi(v, v)}$, es decir que $\Phi(v, v) \in \mathbb{R}$.)

iii) $\Phi(v, v) > 0$ si $v \neq 0$.

Notación. Si Φ es un producto interno, escribiremos $\Phi(v, w) = \langle v, w\rangle$.

Definición 8.2 A un espacio vectorial real (respectivamente complejo) provisto de un producto interno se lo llama un *espacio euclídeo* (respectivamente *espacio unitario*).

Observación 8.3 De las condiciones i) y ii) de la definición de producto interno se deduce que si $\Phi : V \times V \to \mathbb{R}$ (respectivamente $\mathbb{C}$) es un producto interno, para cada $\alpha \in \mathbb{R}$ (respectivamente $\mathbb{C}$), y $v, w, z \in V$ vale:

$$\begin{aligned} \Phi(v, w + z) &= \Phi(v, w) + \Phi(v, z), \\ \Phi(v, \alpha.w) &= \overline{\alpha}\,.\Phi(v, w). \end{aligned}$$

Ejemplos. Se puede comprobar que las funciones Φ definidas a continuación son productos internos sobre los espacios vectoriales correspondientes:

- Producto interno canónico en $\mathbb{R}^n$:
$$\Phi((x_1, \dots, x_n), (y_1, \dots, y_n)) = x_1y_1 + \cdots + x_ny_n.$$
- Producto interno canónico en $\mathbb{C}^n$:
$$\Phi((x_1, \dots, x_n), (y_1, \dots, y_n)) = x_1\overline{y}_1 + \cdots + x_n\overline{y}_n.$$
- Dada $B \in \mathbb{C}^{m\times n}$, denotamos por $B^* \in \mathbb{C}^{n\times m}$ a la matriz transpuesta conjugada de B, es decir, a la matriz definida por $(B^*)_{ij} = \overline{B_{ji}}$. Se define $\Phi : \mathbb{C}^{m\times n} \times \mathbb{C}^{m\times n} \to \mathbb{C}$ como
$$\Phi(A, B) = \operatorname{tr}(A.B^*).$$
- Si $a < b \in \mathbb{R}$ y $C[a, b] = \{f : [a, b] \to \mathbb{R} \mathrel{/} f \text{ continua}\}$, se define $\Phi : C[a, b] \times C[a, b] \to \mathbb{R}$ como
$$\Phi(f, g) = \int_a^b f(x)g(x)\; dx.$$

Dado un espacio vectorial V es posible definir distintos productos internos sobre V. En el ejemplo siguiente veremos una familia de productos internos en $\mathbb{R}^2$.

Ejemplo. Sea $\Phi : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ definida por

$$\Phi((x_1, x_2), (y_1, y_2)) = x_1y_1 - x_1y_2 - x_2y_1 + \alpha.\, x_2y_2$$

Hallar todos los valores de $\alpha \in \mathbb{R}$ para los cuales Φ es un producto interno.

Es inmediato verificar que, para cualquier $\alpha \in \mathbb{R}$ se cumplen las condiciones i) y ii) de la definición de producto interno. Veamos para qué valores de α se cumple la condición iii). Se tiene que

$$\begin{aligned} \Phi((x_1, x_2), (x_1, x_2)) &= x_1^2 - 2x_1x_2 + \alpha x_2^2 \\ &= x_1^2 - 2x_1x_2 + x_2^2 + (\alpha - 1)x_2^2 \\ &= (x_1 - x_2)^2 + (\alpha - 1)x_2^2 \end{aligned}$$

De esta igualdad se deduce que $\Phi(v, v) > 0 \;\forall\, v \neq 0 \iff \alpha > 1$.

En consecuencia, Φ es un producto interno si y sólo si $\alpha > 1$.

8.1.2 Norma de un vector

La noción que sigue generaliza la de longitud de un vector en $\mathbb{R}^2$ o $\mathbb{R}^3$.

Definición 8.4 Sea $(V, \langle, \rangle)$ un espacio vectorial sobre $\mathbb{R}$ (respectivamente $\mathbb{C}$) con producto interno y sea $v \in V$. Se define la *norma* de v asociada a $\langle, \rangle$ (y se nota $\|v\|$) como $\|v\| = \langle v, v\rangle^{\frac{1}{2}}$.

Proposición 8.5 *(Propiedades de la norma.) Sea* $(V, \langle, \rangle)$ *un espacio vectorial con producto interno.*

i) Para cada $v \in V$, $\|v\| \geq 0$, *y* $\|v\| = 0$ *si y sólo si* $v = 0$.

ii) Sean $\alpha \in \mathbb{R}$ *(respectivamente* $\mathbb{C}$*) y* $v \in V$. *Entonces* $\|\alpha . v\| = |\alpha| . \|v\|$.

iii) Desigualdad de Cauchy-Schwartz. *Si* $v, w \in V$, *entonces*

$$|\langle v, w\rangle| \leq \|v\| . \|w\|.$$

iv) Desigualdad triangular. *Si* $v, w \in V$, *entonces*

$$\|v + w\| \leq \|v\| + \|w\|.$$

Demostración. Las propiedades *i)* e *ii)* se deducen inmediatamente de la definición de norma.

iii) Si $w = 0$, no hay nada que hacer. Supongamos entonces que $w \neq 0$. Se tiene que

$$\begin{aligned} 0 &\leq \left\langle v - \frac{\langle v, w\rangle}{\|w\|^2}\, w,\ v - \frac{\langle v, w\rangle}{\|w\|^2}\, w \right\rangle \\ &= \left\langle v, v - \frac{\langle v, w\rangle}{\|w\|^2}\, w \right\rangle - \frac{\langle v, w\rangle}{\|w\|^2} \left\langle w, v - \frac{\langle v, w\rangle}{\|w\|^2}\, w \right\rangle \\ &= \langle v, v\rangle - \frac{\overline{\langle v, w\rangle}}{\|w\|^2} \langle v, w\rangle - \frac{\langle v, w\rangle}{\|w\|^2} \langle w, v\rangle + \frac{\langle v, w\rangle}{\|w\|^2}\, \overline{\frac{\langle v, w\rangle}{\|w\|^2}}\, \langle w, w\rangle \\ &= \|v\|^2 - \frac{|\langle v, w\rangle|^2}{\|w\|^2} \end{aligned}$$

Esto implica que $|\langle v, w\rangle|^2 \leq \|v\|^2 \|w\|^2$, de donde se obtiene la desigualdad buscada.

iv) En primer lugar, observamos que

$$\begin{aligned} \|v + w\|^2 &= \langle v + w,\ v + w\rangle \\ &= \langle v, v\rangle + \langle v, w\rangle + \overline{\langle v, w\rangle} + \langle w, w\rangle \\ &= \|v\|^2 + 2\,\mathrm{Re}\langle v, w\rangle + \|w\|^2. \end{aligned} \tag{8.1}$$

Entonces, teniendo en cuenta que $\text{Re}\langle v,w\rangle \le |\langle v,w\rangle|$ y aplicando la desigualdad de Cauchy-Schwartz, resulta que

$$\begin{aligned}\|v+w\|^2 &\le \|v\|^2 + 2\,|\langle v,w\rangle| + \|w\|^2\\ &\le \|v\|^2 + 2\,\|v\|\|w\| + \|w\|^2\\ &= (\|v\|+\|w\|)^2,\end{aligned}$$

de donde se deduce que $\|v+w\| \le \|v\| + \|w\|$. □

La desigualdad de Cauchy-Schwartz vista en ciertos espacios vectoriales con producto interno nos permite obtener distintas propiedades:

Ejemplo. Sean $f, g \in C[a,b]$. Entonces

$$\left|\int_a^b f(x)g(x)\,dx\right| \le \left(\int_a^b f^2(x)\ dx\right)^{\frac{1}{2}} \left(\int_a^b g^2(x)\ dx\right)^{\frac{1}{2}}.$$

En general, se puede definir una norma en un espacio vectorial V sobre $\mathbb{R}$ o $\mathbb{C}$ como una función $\|\ \| : V \to \mathbb{R}$ que cumpla las condiciones i), ii) y iv) de la proposición anterior. Una norma cualquiera en este sentido puede no provenir de un producto interno (es decir, puede no haber ningún producto interno tal que la norma asociada sea $\|\ \|$). Un ejemplo de esto es la norma infinito en $\mathbb{R}^n$, definida por $\|(x_1,\dots,x_n)\|_\infty = \max\{|x_1|,\dots,|x_n|\}$.

Dada una norma, se puede decidir si proviene o no de un producto interno, ya que éste se puede recuperar a partir de su norma asociada:

Proposición 8.6 *(Identidades de polarización.)*

i) *Sea $(V, \langle, \rangle)$ un $\mathbb{R}$-espacio vectorial con producto interno. Entonces para cada $v, w \in V$ vale:*
$$\langle v,w\rangle = \frac{1}{4}\|v+w\|^2 - \frac{1}{4}\|v-w\|^2.$$

ii) *Sea $(V, \langle, \rangle)$ un $\mathbb{C}$-espacio vectorial con producto interno. Entonces para cada $v, w \in V$ vale:*
$$\langle v,w\rangle = \frac{1}{4}\|v+w\|^2 - \frac{1}{4}\|v-w\|^2 + \frac{i}{4}\|v+iw\|^2 - \frac{i}{4}\|v-iw\|^2.$$

Demostración.

i) Si V es un $\mathbb{R}$-espacio vectorial, entonces $\|v+w\|^2 = \|v\|^2 + 2\langle v,w\rangle + \|w\|^2$ y $\|v-w\|^2 = \|v\|^2 - 2\langle v,w\rangle + \|w\|^2$. Entonces
$$\frac{1}{4}\|v+w\|^2 - \frac{1}{4}\|v-w\|^2 = \frac{1}{2}\langle v,w\rangle - \left(\frac{-1}{2}\right)\langle v,w\rangle = \langle v,w\rangle.$$

ii) En forma análoga a lo hecho en *i)*, usando la identidad (8.1), resulta que si V es un $\mathbb{C}$-espacio vectorial, entonces

$$\frac{1}{4}\|v+w\|^2 - \frac{1}{4}\|v-w\|^2 = \mathrm{Re}\langle v, w\rangle.$$

Por otro lado, $\|v+iw\|^2 = \|v\|^2 + 2\,\mathrm{Re}\langle v, iw\rangle + \|iw\|^2 = \|v\|^2 + 2\,\mathrm{Im}\langle v, w\rangle + \|w\|^2$ y, similarmente, $\|v-iw\|^2 = \|v\|^2 - 2\,\mathrm{Im}\langle v, w\rangle + \|w\|^2$, lo que implica que

$$\frac{i}{4}\|v+iw\|^2 - \frac{i}{4}\|v-iw\|^2 = i\,\mathrm{Im}\langle v, w\rangle.$$

La identidad del enunciado se obtiene haciendo $\langle v, w\rangle = \mathrm{Re}\langle v, w\rangle + i\,\mathrm{Im}\langle v, w\rangle$. $\square$

Una norma cualquiera estará asociada a un producto interno si y sólo si la función que se obtiene mediante estas identidades resulta un producto interno. En lo que sigue, sólo se considerarán normas asociadas a productos internos.

8.1.3 Distancia entre vectores

A partir de la definición de norma de un vector, podemos definir la noción de distancia entre dos vectores.

Definición 8.7 Sea V un $\mathbb{R}$-(o $\mathbb{C}$-) espacio vectorial con producto interno $\langle , \rangle$. Se define $d : V \times V \to \mathbb{R}$ como $d(v, w) = \|v - w\|$.

Utilizando las propiedades de la norma se puede verificar que la función d satisface las siguientes propiedades:

i) $d(v, w) \geq 0 \ \forall\, v, w \in V$.

ii) $d(v, w) = 0 \iff v = w$.

iii) $d(v, w) = d(w, v) \ \forall\, v, w \in V$.

iv) $d(v, z) \leq d(v, w) + d(w, z) \ \forall\, v, w, z \in V$.

Dados vectores $v, w \in V$ se dice que $d(v, w)$ es la *distancia* entre v y w.

Dado un conjunto no vacío cualquiera X, puede definirse una distancia en X como cualquier función $d : X \times X \to \mathbb{R}$ que cumpla las propiedades i), ii), iii) y iv) anteriores. Hay distancias en espacios vectoriales que no provienen de ninguna norma. En lo que sigue, sólo trabajaremos con distancias asociadas a normas asociadas a productos internos.

8.1.4 Ángulo entre dos vectores

Sea $(V, \langle, \rangle)$ un espacio euclídeo. La desigualdad de Cauchy-Schwartz establece que para todo par de vectores $v, w \in V$ se tiene que $|\langle v, w\rangle| \leq \|v\|.\|w\|$. Si $v, w \neq 0$, resulta entonces que

$$-1 \leq \frac{\langle v, w\rangle}{\|v\|.\|w\|} \leq 1.$$

Esto nos permite introducir la siguiente noción de ángulo entre dos vectores:

Definición 8.8 Sea $(V, \langle, \rangle)$ un espacio euclídeo y sean $v, w \in V$ no nulos. Se define el *ángulo entre v y w* como el único número real $\alpha \in [0, \pi]$ tal que $\cos(\alpha) = \dfrac{\langle v, w\rangle}{\|v\|.\|w\|}$.

Observamos que si α es el ángulo entre v y w, entonces

$$\begin{aligned} \|v+w\|^2 &= \langle v+w, v+w\rangle = \|v\|^2 + 2\langle v, w\rangle + \|w\|^2 \\ &= \|v\|^2 + 2.\cos(\alpha).\,\|v\|.\|w\| + \|w\|^2, \end{aligned}$$

que es la fórmula conocida como *teorema del coseno*.

8.1.5 Matriz de un producto interno

Si V es un espacio vectorial de dimensión finita con un producto interno $\langle, \rangle$, fijada una base de V, vamos a construir una matriz asociada al producto interno y a dicha base.

Definición 8.9 Sea V un $\mathbb{R}$-(respectivamente $\mathbb{C}$-) espacio vectorial de dimensión finita, sea $\langle, \rangle$ un producto interno sobre V y sea $B = \{v_1, \ldots, v_n\}$ una base de V. Se define la *matriz del producto interno $\langle, \rangle$ en la base B* como la matriz $A \in \mathbb{R}^{n\times n}$ (respectivamente $A \in \mathbb{C}^{n\times n}$) tal que

$$A_{ij} = \langle v_i, v_j\rangle \qquad \forall\, 1 \leq i, j \leq n.$$

Notación. Escribiremos $|\langle, \rangle|_B$ para denotar la matriz del producto interno $\langle, \rangle$ en la base B.

Observación 8.10 Si A es la matriz de un producto interno, entonces $A_{ij} = \overline{A_{ji}}\ \ \forall\, i, j$.

Sin embargo, la condición $A_{ij} = \overline{A_{ji}}\ \forall\, i, j$ no es suficiente para que A sea la matriz de un producto interno. Por ejemplo, $A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$ no puede ser la matriz de un producto interno en una base, ya que si v es el primer vector de la base, sería $\langle v, v\rangle = 0$.

Ejemplo. Para el producto interno en $\mathbb{R}^2$ definido por

$$\langle (x_1, x_2), (y_1, y_2)\rangle = x_1y_1 - x_1y_2 - x_2y_1 + \alpha\, x_2y_2 \qquad (\alpha > 1)$$

(ver página 190) y E la base canónica de $\mathbb{R}^2$, resulta

$$|\langle,\rangle|_E = \begin{pmatrix} 1 & -1 \\ -1 & \alpha \end{pmatrix}.$$

La matriz de un producto interno en una base B nos permite calcular el producto interno entre cualquier par de vectores:

Proposición 8.11 *Sea V un $\mathbb{R}$-(o $\mathbb{C}$-) espacio vectorial de dimensión finita y sea $\langle,\rangle$ un producto interno sobre V. Sea B una base de V. Entonces, para cada $v, w \in V$, se tiene que*

$$\langle v, w\rangle = (v)_B.\, |\langle,\rangle|_B.\, \overline{(w)_B^t}.$$

Demostración. Sea $B = \{v_1, \dots, v_n\}$. Supongamos que $v = \sum\limits_{i=1}^{n} \alpha_i v_i$ y $w = \sum\limits_{i=1}^{n} \beta_i v_i$. Entonces

$$\langle v, w\rangle = \Big\langle \sum_{i=1}^{n} \alpha_i v_i, \sum_{j=1}^{n} \beta_j v_j \Big\rangle = \sum_{i=1}^{n} \alpha_i \Big\langle v_i, \sum_{j=1}^{n} \beta_j v_j \Big\rangle = \sum_{i=1}^{n} \alpha_i \Big(\sum_{j=1}^{n} \overline{\beta_j} \langle v_i, v_j\rangle \Big).$$

Por otro lado, teniendo en cuenta que $(v)_B = (\alpha_1, \dots, \alpha_n)$ y $(w)_B = (\beta_1, \dots, \beta_n)$, resulta que

$$(v)_B.\, |\langle,\rangle|_B.\, \overline{(w)_B^t} = \sum_{i=1}^{n} \alpha_i \Big(|\langle,\rangle|_B \overline{(w)_B^t} \Big)_{i1} = \sum_{i=1}^{n} \alpha_i \Big(\sum_{j=1}^{n} \langle v_i, v_j\rangle \overline{\beta_j} \Big).$$

Luego $\langle v, w\rangle = (v)_B.\, |\langle,\rangle|_B.\, \overline{(w)_B^t}$. □

8.2 Ortogonalidad

En esta sección generalizaremos la noción conocida de perpendicularidad en $\mathbb{R}^2$ y algunas propiedades que se basan en esta noción, a espacios vectoriales con producto interno arbitrarios.

8.2.1 Conjuntos ortogonales y ortonormales

Definición 8.12 Sea $(V, \langle,\rangle)$ un espacio vectorial con producto interno. Dos vectores $v, w \in V$ se dicen *ortogonales* (o perpendiculares) si $\langle v, w\rangle = 0$.

Observación 8.13 *(Teorema de Pitágoras.)* Si $v, w \in V$ son vectores ortogonales, entonces $\|v + w\|^2 = \|v\|^2 + \|w\|^2$.

Definición 8.14 Sea $(V, \langle,\rangle)$ un espacio vectorial con producto interno. Se dice que $\{v_1, \dots, v_r\} \subseteq V$ es un conjunto *ortogonal* si $\langle v_i, v_j\rangle = 0\ \forall\, i \neq j$. El conjunto se dice *ortonormal* si es ortogonal y $\|v_i\| = 1$ para cada $1 \leq i \leq r$.

Ejemplos.

1. En $\mathbb{R}^n$ (o $\mathbb{C}^n$) con el producto interno canónico, la base canónica es un conjunto ortonormal:
   - $\langle e_i, e_j\rangle = 0$ si $i \neq j$.
   - $\|e_i\|^2 = \langle e_i, e_i\rangle = 1$ para cada $1 \le i \le n$.
2. En $\mathbb{R}^2$ con el producto interno canónico, el conjunto $\{(1,1),(1,-1)\}$ es un conjunto ortogonal, pues $\langle (1,1),(1,-1)\rangle = 0$.

   Este conjunto no es ortonormal, ya que $\|(1,1)\| = \sqrt{2} \neq 1$ y $\|(1,-1)\| = \sqrt{2} \neq 1$. A partir de este conjunto podemos hallar uno que es ortonormal dividiendo cada uno de los vectores por su norma: $\left\{\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), \left(\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}\right)\right\}$.

Si $(V, \langle,\rangle)$ es un espacio vectorial con producto interno, la matriz de $\langle,\rangle$ en una base ortogonal (u ortonormal) de V es particularmente simple:

Observación 8.15 Si $(V, \langle,\rangle)$ es un espacio vectorial de dimensión n con producto interno, entonces $B = \{v_1, \ldots, v_n\}$ es una base de V ortogonal para $\langle,\rangle$ si y sólo si

$$|\langle,\rangle|_B = \begin{pmatrix} \langle v_1, v_1\rangle & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \langle v_n, v_n\rangle \end{pmatrix}.$$

En particular, B es una base ortonormal de V si y sólo si $|\langle,\rangle|_B = I_n$.

Como consecuencia, si $B = \{v_1, \ldots, v_n\}$ es una base ortonormal de V se puede calcular fácilmente el producto interno entre dos vectores (y, en particular, la norma de un vector) a partir de sus coordenadas en la base B: si $v = \sum\limits_{i=1}^{n} \alpha_i\, v_i$ y $w = \sum\limits_{i=1}^{n} \beta_i\, v_i$, entonces:

$$\langle v, w\rangle = \sum_{i=1}^{n} \alpha_i\, \overline{\beta_i} \quad \text{ y } \quad \|v\| = \Big(\sum_{i=1}^{n} |\alpha_i|^2\Big)^{\frac{1}{2}}.$$

Proposición 8.16 *Sea $(V, \langle,\rangle)$ un espacio vectorial con producto interno. Sea $\{v_1, \ldots, v_r\}$ un conjunto ortogonal de V con $v_i \neq 0$ para cada $1 \le i \le r$. Entonces $\{v_1, \ldots, v_r\}$ es linealmente independiente.*

Demostración. Supongamos que $\sum\limits_{i=1}^{r} \alpha_i v_i = 0$. Entonces para cada $1 \le j \le r$,

$$0 = \langle 0, v_j\rangle = \Big\langle \sum_{i=1}^{r} \alpha_i v_i, v_j \Big\rangle = \sum_{i=1}^{r} \alpha_i \langle v_i, v_j\rangle = \alpha_j.\, \|v_j\|^2,$$

y como $v_j \neq 0$, resulta que $\alpha_j = 0$.

En consecuencia, $\{v_1, \ldots, v_r\}$ es linealmente independiente. $\square$

Si se tiene una base ortogonal de un subespacio, las coordenadas en esta base de cualquier vector del subespacio pueden encontrarse fácilmente usando el producto interno:

Proposición 8.17 *Sea $(V, \langle, \rangle)$ un espacio vectorial con producto interno. Sea $\{v_1, \ldots, v_r\} \subseteq V$ un conjunto ortogonal tal que $v_i \neq 0$ para cada $1 \leq i \leq r$ y sea $v \in < v_1, \ldots, v_r >$. Entonces*

$$v = \sum_{j=1}^{r} \frac{\langle v, v_j \rangle}{\|v_j\|^2} . v_j.$$

Demostración. Si $v = \sum_{i=1}^{r} \alpha_i \, v_i$, para cada $1 \leq j \leq r$ se tiene que

$$\langle v, v_j \rangle = \Big\langle \sum_{i=1}^{r} \alpha_i v_i, v_j \Big\rangle = \sum_{i=1}^{r} \alpha_i \langle v_i, v_j \rangle = \alpha_j \langle v_j, v_j \rangle = \alpha_j \, \|v_j\|^2,$$

de donde se deduce, teniendo en cuenta que $v_j \neq 0$, que $\alpha_j = \dfrac{\langle v, v_j \rangle}{\|v_j\|^2}$. $\square$

Corolario 8.18 *Sea $(V, \langle, \rangle)$ un espacio vectorial con producto interno y sea $\{v_1, \ldots, v_r\}$ un conjunto ortonormal de V. Entonces, para cada $v \in < v_1, \ldots, v_r >$, se tiene que*

$$v = \sum_{i=1}^{r} \langle v, v_i \rangle . \, v_i.$$

Finalmente, combinando este resultado con la Observación 8.15, se obtiene:

Corolario 8.19 *Sea $(V, \langle, \rangle)$ un espacio vectorial con producto interno y sea $\{v_1, \ldots, v_r\}$ un conjunto ortonormal de V. Entonces:*

i) Para $v, w \in < v_1, \ldots, v_r >$, $\langle v, w \rangle = \sum_{i=1}^{r} \langle v, v_i \rangle . \overline{\langle w, v_i \rangle}$.

ii) Para cada $v \in < v_1, \ldots, v_r >$, $\|v\| = \Big(\sum_{i=1}^{r} |\langle v, v_i \rangle|^2 \Big)^{\frac{1}{2}}$.

En lo que sigue, intentaremos encontrar bases ortonormales en un espacio vectorial de dimensión finita con producto interno. Comenzaremos haciendo esto en un ejemplo.

Ejemplo. Se considera en $\mathbb{R}^2$ el producto interno definido por

$$\langle (x_1, x_2), (y_1, y_2) \rangle = x_1 y_1 - x_1 y_2 - x_2 y_1 + \alpha \, x_2 y_2 \qquad (\alpha > 1)$$

(ver página 190). Hallar una base de $\mathbb{R}^2$ ortonormal para este producto interno.

Elegimos un vector en $\mathbb{R}^2$, por ejemplo, $(1,0)$. Buscamos un vector ortogonal a éste para el producto interno dado, es decir, un vector $(y_1, y_2) \in \mathbb{R}^2$ tal que

$$0 = \langle (1,0), (y_1, y_2) \rangle = y_1 - y_2,$$

por ejemplo, $(y_1, y_2) = (1,1)$.

Entonces $\{(1,0), (1,1)\}$ es una base ortogonal de $\mathbb{R}^2$ y, por lo tanto, basta normalizar (es decir, dividir por la norma) cada uno de los vectores de esta base. Se tiene que

$$\begin{aligned} \|(1,0)\| &= 1, \\ \|(1,1)\| &= \langle (1,1), (1,1) \rangle^{\frac{1}{2}} = (1 - 1 - 1 + \alpha)^{\frac{1}{2}} = \sqrt{\alpha - 1}. \end{aligned}$$

Luego, $B = \left\{ (1,0), \left(\frac{1}{\sqrt{\alpha-1}}, \frac{1}{\sqrt{\alpha-1}} \right) \right\}$ es una base ortonormal de $\mathbb{R}^2$ para el producto interno dado.

La proposición siguiente asegura que todo espacio vectorial de dimensión finita con producto interno tiene una base ortonormal. Más aún, su demostración da un procedimiento recursivo, conocido como el método de ortonormalización de Gram-Schmidt, que permite obtener una base ortonormal del espacio a partir de una base cualquiera del mismo.

Proposición 8.20 (Método de ortonormalización de Gram-Schmidt) *Sea $(V, \langle , \rangle)$ un espacio vectorial con producto interno y sea $\{v_1, \dots, v_n\}$ una base de V. Existe un base ortonormal $B = \{w_1, \dots, w_n\}$ de V tal que*

$$< v_1, \dots, v_k > = < w_1, \dots, w_k > \qquad \forall\, 1 \le k \le n.$$

Demostración. Se construirá una base ortogonal $\{z_1, \dots, z_n\}$ de V que cumpla:

$$< z_1, \dots, z_k > = < v_1, \dots, v_k > \qquad \forall\, 1 \le k \le n.$$

Normalizando los vectores de esta base se obtendrá la base ortonormal buscada.

Construiremos los vectores de la base recursivamente:

- Tomamos $z_1 = v_1$, que satisface $< z_1 > = < v_1 >$.
- Buscamos $z_2 \in V$ con $\langle z_2, z_1 \rangle = 0$ y tal que $< z_1, z_2 > = < v_1, v_2 >$. Esta segunda condición vale si y sólo si z_2 es de la forma $z_2 = a.v_1 + b.v_2$ con $b \neq 0$. Podemos suponer entonces que $b = 1$, es decir, $z_2 = v_2 + a.v_1$ y buscar a de manera que se cumpla la primera condición:

  $$0 = \langle z_2, z_1 \rangle = \langle v_2 + av_1, v_1 \rangle = \langle v_2, v_1 \rangle + a.\langle v_1, v_1 \rangle,$$

  lo que implica que

  $$a = \frac{-\langle v_2, v_1 \rangle}{\|v_1\|^2}.$$

Luego, el vector

$$z_2 = v_2 - \frac{\langle v_2, v_1\rangle}{\|v_1\|^2}.v_1 = v_2 - \frac{\langle v_2, z_1\rangle}{\|z_1\|^2}.z_1$$

satisface las condiciones requeridas.

- Supongamos construidos $z_1, \dots, z_r \in V$ tales que

  (i) $\langle z_i, z_j\rangle = 0$ si $i \neq j$.

  (ii) $< z_1, \dots, z_k > = < v_1, \dots, v_k > \quad \forall\, 1 \leq k \leq r$.

  Consideramos el vector

$$z_{r+1} = v_{r+1} - \sum_{i=1}^{r} \frac{\langle v_{r+1}, z_i\rangle}{\|z_i\|^2}.z_i. \tag{8.2}$$

  Se tiene que:

  a) $< z_1, \dots, z_r, z_{r+1} > = < z_1, \dots, z_r, v_{r+1} > = < v_1, \dots, v_r, v_{r+1} >$,

  b) para cada $j \leq r$

$$\begin{aligned} \langle z_{r+1}, z_j\rangle &= \left\langle v_{r+1} - \sum_{i=1}^{r} \frac{\langle v_{r+1}, z_i\rangle}{\|z_i\|^2}.z_i, z_j \right\rangle \\ &= \langle v_{r+1}, z_j\rangle - \sum_{i=1}^{r} \frac{\langle v_{r+1}, z_i\rangle}{\|z_i\|^2}.\langle z_i, z_j\rangle \\ &= \langle v_{r+1}, z_j\rangle - \frac{\langle v_{r+1}, z_j\rangle}{\|z_j\|^2}.\langle z_j, z_j\rangle \\ &= 0. \end{aligned}$$

  Luego, z_{r+1} satisface las condiciones requeridas.

De esta manera, al concluir el n-ésimo paso se obtiene una base ortogonal $\{z_1, \dots, z_n\}$ de V que además satisface $< z_1, \dots, z_k > = < v_1, \dots, v_k >$ para cada $1 \leq k \leq n$.

Finalmente, para cada $1 \leq i \leq n$ consideramos el vector $w_i = \frac{z_i}{\|z_i\|}$. Luego el conjunto $B = \{w_1, \dots, w_n\}$ resulta una base ortonormal de V que cumple lo pedido. □

Corolario 8.21 *Sea $(V, \langle, \rangle)$ un espacio vectorial de dimensión finita con producto interno, y sea S un subespacio de V, $S \neq \{0\}$. Entonces existe una base ortonormal de V que contiene una base ortonormal de S.*

Demostración. Sea $\{s_1, \dots, s_r\}$ una base de S. Existen $v_{r+1}, \dots, v_n \in V$ tales que $B = \{s_1, \dots, s_r, v_{r+1}, \dots, v_n\}$ es una base de V. Aplicando el proceso de Gram-Schmidt a B se obtiene una base ortonormal $B' = \{w_1, \dots, w_r, w_{r+1}, \dots, w_n\}$ de V que satisface

$$< w_1, \dots, w_r > = < s_1, \dots, s_r > = S.$$

En consecuencia, $\{w_1, \dots, w_r\}$ es una base ortonormal de S incluida en la base ortonormal B' de V. $\square$

Mostramos ahora el método de Gram-Schmidt en un ejemplo:

Ejemplo. Dada la base $B = \{(1, 0, i), (1, 1, 2 + i), (0, 0, 1)\}$ de $\mathbb{C}^3$, ortonormalizarla usando el método de Gram-Schmidt.

Notaremos $v_1 = (1, 0, i)$, $v_2 = (1, 1, 2 + i)$ y $v_3 = (0, 0, 1)$.

- Comenzamos tomando z_1 como el primer vector de la base: $z_1 = (1, 0, i)$.
- Construimos ahora z_2 aplicando la fórmula (8.2) para $r = 1$:

$$\begin{aligned} z_2 &= v_2 - \frac{\langle v_2, z_1\rangle}{\|z_1\|^2} . z_1 = (1, 1, 2 + i) - \frac{\langle (1, 1, 2 + i), (1, 0, i)\rangle}{\|(1, 0, i)\|^2} . (1, 0, i) \\ &= (1, 1, 2 + i) - (1 - i). (1, 0, i) \;=\; (i, 1, 1). \end{aligned}$$

- Finalmente, hallamos z_3 aplicando nuevalmente la fórmula (8.2) para $r = 2$:

$$\begin{aligned} z_3 &= v_3 - \frac{\langle v_3, z_1\rangle}{\|z_1\|^2} .z_1 - \frac{\langle v_3, z_2\rangle}{\|z_2\|^2} .z_2 \\ &= (0, 0, 1) - \frac{\langle (0, 0, 1), (1, 0, i)\rangle}{\|(1, 0, i)\|^2} . (1, 0, i) - \frac{\langle (0, 0, 1), (i, 1, 1)\rangle}{\|(i, 1, 1)\|^2} . (i, 1, 1) \\ &= (0, 0, 1) + \frac{i}{2} . (1, 0, i) - \frac{1}{3} . (i, 1, 1) \\ &= \Big(\frac{i}{6}, \frac{-1}{3}, \frac{1}{6}\Big). \end{aligned}$$

El conjunto $\{z_1, z_2, z_3\}$ es una base ortogonal de $\mathbb{C}^3$. Dividiendo cada vector por su norma, obtenemos

$$\begin{aligned} w_1 &= \frac{z_1}{\|z_1\|} = \frac{1}{\sqrt{2}} . (1, 0, i) = \Big(\frac{1}{\sqrt{2}}, 0, \frac{i}{\sqrt{2}}\Big) \\ w_2 &= \frac{z_2}{\|z_2\|} = \frac{1}{\sqrt{3}} . (i, 1, 1) = \Big(\frac{i}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\Big) \\ w_3 &= \frac{z_3}{\|z_3\|} = \sqrt{6}. \Big(\frac{i}{6}, \frac{-1}{3}, \frac{1}{6}\Big) = \Big(\frac{\sqrt{6}i}{6}, \frac{-\sqrt{6}}{3}, \frac{\sqrt{6}}{6}\Big) \end{aligned}$$

tales que $\{w_1, w_2, w_3\}$ es una base ortonormal de $\mathbb{C}^3$ que cumple $< v_1 > \; = \; < w_1 >$ y $< v_1, v_2 > = < w_1, w_2 >$.

8.2.2 Complemento ortogonal

Definición 8.22 Sea $(V, \langle, \rangle)$ un espacio vectorial con producto interno, y sea $S \subseteq V$ un conjunto. Se define el *complemento ortogonal* de S como

$$S^{\perp} = \{v \in V : \langle v, s\rangle = 0 \quad \forall\, s \in S\}.$$

Observación 8.23 $S^{\perp}$ es un subespacio de V:

i) $0 \in S^{\perp}$, pues $\langle 0, s\rangle = 0$ para todo $s \in S$.

ii) Sean $v, w \in S^{\perp}$. Para cada $s \in S$, se tiene que $\langle v, s\rangle = 0$ y $\langle w, s\rangle = 0$, con lo que $\langle v + w, s\rangle = \langle v, s\rangle + \langle w, s\rangle = 0 + 0 = 0$. Luego, $v + w \in S^{\perp}$.

iii) Si $v \in S^{\perp}$ y $\lambda \in \mathbb{R}$ (o $\mathbb{C}$), entonces para cada $s \in S$ vale $\langle \lambda. v, s\rangle = \lambda.\langle v, s\rangle = \lambda. 0 = 0$. Luego, $\lambda. v \in S^{\perp}$.

Ejemplos.

1. En $\mathbb{R}^2$ con el producto interno canónico:

$$\begin{aligned} \{(1,1)\}^{\perp} &= \{(x,y) \in \mathbb{R}^2 \,/\, \langle (x,y), (1,1)\rangle = 0\} \\ &= \{(x,y) \in \mathbb{R}^2 \,/\, x + y = 0\} \\ &= < (1,-1) > . \end{aligned}$$

2. En $\mathbb{C}^3$ con el producto interno canónico:

$$\begin{aligned} < (1, i, 1+i) >^{\perp} &= \{(x,y,z) \in \mathbb{C}^3 \,/\, \langle (x,y,z), (\alpha, \alpha i, \alpha(1+i)\rangle = 0 \;\forall\, \alpha \in \mathbb{C}\} \\ &= \{(x,y,z) \in \mathbb{C}^3 \,/\, \overline{\alpha}\,(x.1 + y(-i) + z(1-i)) = 0 \;\forall\, \alpha \in \mathbb{C}\} \\ &= \{(x,y,z) \in \mathbb{C}^3 \,/\, x - iy + (1-i)z = 0\} \\ &= < (i,1,0), (i-1,0,1) >. \end{aligned}$$

En el caso en que S es un *subespacio* de V, se tiene el siguiente resultado, cuya demostración provee un método para hallar $S^{\perp}$.

Proposición 8.24 *Sea $(V, \langle, \rangle)$ un espacio vectorial de dimensión finita con producto interno, y sea $S \subseteq V$ un subespacio. Entonces:*

i) $S \cap S^{\perp} = \{0\}$.

ii) $\dim(S) + \dim(S^{\perp}) = \dim V$.

En consecuencia, $S \oplus S^{\perp} = V$.

Demostración.

i) Sea $x \in S \cap S^{\perp}$. Como $x \in S^{\perp}$, para cada $s \in S$, se tiene que $\langle x, s\rangle = 0$. En particular, tomando $s = x \in S$, resulta que $\|x\|^2 = \langle x, x\rangle = 0$, de donde $x = 0$.

ii) Sea $\{s_1, \dots, s_r\}$ una base de S. Existen $v_{r+1}, \dots, v_n \in V$ tales que

$$\{s_1, \dots, s_r, v_{r+1}, \dots, v_n\}$$

es una base de V. Aplicando el proceso de Gram-Schmidt a esta base se obtiene una base ortonormal de V, $B = \{w_1, \dots, w_r, w_{r+1}, \dots, w_n\}$, que además cumple

$$< w_1, \dots, w_r > = < s_1, \dots, s_r > = S.$$

Sea $j > r$. Veamos que $w_j \in S^{\perp}$. Dado $s \in S$, existen escalares $\alpha_1, \dots, \alpha_r$ tales que $s = \sum\limits_{i=1}^{r} \alpha_i w_i$. Entonces

$$\langle w_j, s\rangle = \Big\langle w_j, \sum_{i=1}^{r} \alpha_i w_i \Big\rangle = \sum_{i=1}^{r} \overline{\alpha_i}\langle w_j, w_i\rangle = 0,$$

ya que, como la base B es ortogonal y $j > r$, se tiene que $\langle w_j, w_i\rangle = 0$ para cada $1 \le i \le r$. En consecuencia, $w_j \in S^{\perp}$.

Se tiene entonces que $w_{r+1}, \dots, w_n \in S^{\perp}$, con lo que $< w_{r+1}, \dots, w_n > \subseteq S^{\perp}$ (pues $S^{\perp}$ es un subespacio) y, por lo tanto,

$$\dim(S^{\perp}) \ge \dim(< w_{r+1}, \dots, w_n >) = n - r = n - \dim(S),$$

es decir, $\dim(S) + \dim(S^{\perp}) \ge n$. Por otro lado, como $S \cap S^{\perp} = \{0\}$,

$$\dim(S) + \dim(S^{\perp}) = \dim(S + S^{\perp}) \le \dim(V) = n.$$

Entonces $\dim(S) + \dim(S^{\perp}) = \dim(V)$. Más aún, de la demostración se deduce también que $S^{\perp} = < w_{r+1}, \dots, w_n >$. □

Ejemplo. Sea $S = < (1, 0, i), (1, 1, 2 + i) > \subseteq \mathbb{C}^3$. En el ejemplo de la página 200, hallamos una base ortonormal $\{w_1, w_2, w_3\}$ de $\mathbb{C}^3$ con la propiedad de que $\{w_1, w_2\}$ es una base de S. Entonces, la demostración del ítem *ii)* de la proposición anterior nos dice que $S^{\perp} = < w_3 > = \Big\langle \Big(\frac{\sqrt{6}i}{6}, \frac{-\sqrt{6}}{3}, \frac{\sqrt{6}}{6} \Big) \Big\rangle$.

Proposición 8.25 *Sea $(V, \langle, \rangle)$ un espacio vectorial de dimensión finita con producto interno, y sea S un subespacio de V. Entonces $(S^{\perp})^{\perp} = S$.*

Demostración. Por definición, $(S^{\perp})^{\perp} = \{v \in V \,/\, \langle v, t\rangle = 0 \; \forall\, t \in S^{\perp}\}$.

Veamos que $S \subseteq (S^{\perp})^{\perp}$: Sea $s \in S$. Para cada $t \in S^{\perp}$ se tiene que $\langle s, t\rangle = \overline{\langle t, s\rangle} = 0$, de donde se deduce que $s \in (S^{\perp})^{\perp}$.

Por otro lado, por la proposición anterior, $\dim((S^\perp)^\perp) = \dim S$, y por lo tanto, vale la igualdad $S = (S^\perp)^\perp$. $\square$

Ejemplo. Hallar el complemento ortogonal del subespacio

$$S = \left\{(x_1, x_2, x_3, x_4) \in \mathbb{C}^4 \ / \ \begin{cases} x_1 + ix_2 + x_3 - x_4 = 0 \\ (1-i)x_2 + x_3 = 0 \end{cases} \right\}$$

para el producto interno canónico.

Observamos que la condición $x_1+ix_2+x_3-x_4 = 0$ puede reescribirse, utilizando el producto interno canónico de $\mathbb{C}^4$, en la forma $\langle (x_1, x_2, x_3, x_4), (1, -i, 1, -1)\rangle = 0$. Análogamente, la ecuación $(1-i)x_2 + x_3 = 0$ puede reescribirse como $\langle (x_1, x_2, x_3, x_4), (0, 1+i, 1, 0)\rangle = 0$.

Concluimos entonces que $S = < (1, -i, 1, -1), (0, 1 + i, 1, 0) >^\perp$ y, por lo tanto, aplicando la proposición anterior resulta que

$$S^\perp = (< (1, -i, 1, -1), (0, 1 + i, 1, 0) >^\perp)^\perp = < (1, -i, 1, -1), (0, 1 + i, 1, 0) >.$$

8.2.3 Proyección ortogonal

Dado un subespacio S de un espacio vectorial V de dimensión finita con producto interno, como $S \oplus S^\perp = V$, se puede considerar el proyector $p_S : V \to V$ cuya imagen es S y cuyo núcleo es $S^\perp$ (ver Sección 3.4).

Definición 8.26 Sea $(V, \langle, \rangle)$ un espacio vectorial de dimensión finita con producto interno y sea $S \subseteq V$ un subespacio. Se define la *proyección ortogonal sobre* S como la transformación lineal $p_S : V \to V$ que satisface:

$$\begin{array}{rcl} p_S(s) & = & s \quad \forall s \in S \\ p_S(t) & = & 0 \quad \forall t \in S^\perp \end{array}$$

Observamos que si $B = \{v_1, \dots, v_r, v_{r+1}, \dots, v_n\}$ es una base ortonormal de V tal que $\{v_1, \dots, v_r\}$ es una base de S y $\{v_{r+1}, \dots, v_n\}$ una base de $S^\perp$, la proyección ortogonal sobre S es la única transformación lineal $p_S : V \to V$ que satisface:

$$p_S(v_i) = v_i \quad \forall\, 1 \le i \le r \qquad \text{y} \qquad p_S(v_i) = 0 \quad \forall\, r + 1 \le i \le n.$$

En consecuencia, para cada $v \in V$, recordando que $v = \sum_{i=1}^{n} \langle v, v_i\rangle.\, v_i$, resulta que

$$p_S(v) = p_S\Big(\sum_{i=1}^{n} \langle v, v_i\rangle.\, v_i\Big) = \sum_{i=1}^{n} \langle v, v_i\rangle.\, p_S(v_i) = \sum_{i=1}^{r} \langle v, v_i\rangle.\, v_i,$$

lo que nos da una expresión para $p_S(v)$ en términos de los vectores de la base ortonormal de S.

Ejemplo. Sea S el subespacio de $\mathbb{C}^3$ dado por $S = < (1,0,i),(1,1,2+i) >$. Hallar la proyección ortogonal $p_S : \mathbb{C}^3 \to \mathbb{C}^3$.

De acuerdo a lo calculado en el ejemplo de la página 200, el conjunto

$$B_S = \Big\{\Big(\frac{1}{\sqrt{2}}, 0, \frac{i}{\sqrt{2}}\Big), \Big(\frac{i}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\Big)\Big\}$$

es una base ortonormal de S. Entonces, para cada $x = (x_1, x_2, x_3) \in \mathbb{C}^3$, vale

$$\begin{aligned} p_S(x) &= \Big\langle x, \Big(\frac{1}{\sqrt{2}}, 0, \frac{i}{\sqrt{2}}\Big)\Big\rangle . \Big(\frac{1}{\sqrt{2}}, 0, \frac{i}{\sqrt{2}}\Big) + \Big\langle x, \Big(\frac{i}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\Big)\Big\rangle . \Big(\frac{i}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\Big) \\ &= \Big(\frac{x_1 - ix_3}{2}, 0, \frac{ix_1 + x_3}{2}\Big) + \Big(\frac{x_1 + ix_2 + ix_3}{3}, \frac{-ix_1 + x_2 + x_3}{3}, \frac{-ix_1 + x_2 + x_3}{3}\Big) \\ &= \Big(\frac{5}{6}\, x_1 + \frac{i}{3}\, x_2 - \frac{i}{6}\, x_3,\ \frac{-i}{3}\, x_1 + \frac{1}{3}\, x_2 + \frac{1}{3}\, x_3,\ \frac{i}{6}\, x_1 + \frac{1}{3}\, x_2 + \frac{5}{6}\, x_3\Big). \end{aligned}$$

Notar que si $\{w_1, \dots, w_r\}$ es una base ortogonal de S, de la fórmula hallada más arriba para p_S se desprende que $p_S(v) = \sum\limits_{i=1}^{r} \dfrac{\langle v, w_i\rangle}{\|w_i\|^2} w_i$ para cada $v \in V$.

Observación 8.27 *Sea V un espacio vectorial de dimensión finita con producto interno y sea S un subespacio de V. Entonces $p_S + p_{S^\perp} = id_V$.*

En efecto, si $B = \{v_1, \dots, v_r, v_{r+1}, \dots, v_n\}$ es una base ortonormal de V tal que $\{v_1, \dots, v_r\}$ es una base de S y $\{v_{r+1}, \dots, v_n\}$ es una base de $S^\perp$, para cada $v \in V$ se tiene que $p_S(v) = \sum\limits_{i=1}^{r} \langle v, v_i\rangle . v_i$ y $p_{S^\perp}(v) = \sum\limits_{i=r+1}^{n} \langle v, v_i\rangle . v_i$. Entonces $p_S(v) + p_{S^\perp}(v) = \sum\limits_{i=1}^{n} \langle v, v_i\rangle . v_i = v$.

8.2.4 Distancia de un punto a un subespacio

A continuación nos concentraremos en el estudio del siguiente problema: dado un subespacio S de un espacio vectorial V de dimensión finita con producto interno y un punto $p \in V$, encontrar, si es posible, el punto de S que se encuentra a menor distancia de p. Por ejemplo, si S es una recta en $\mathbb{R}^2$, dicho punto es el único $s \in S$ tal que el segmento de s a p es perpendicular a S. En lo que sigue, generalizamos esta idea a un espacio vectorial y un subespacio arbitrarios.

Comenzamos introduciendo el concepto de distancia de un punto a un conjunto.

Definición 8.28 Sea $(V, \langle , \rangle)$ un espacio vectorial con producto interno y sea $S \subseteq V$. Para cada $v \in V$ se define la *distancia de v a S* como

$$d(v, S) = \inf\{d(v, s) : s \in S\} = \inf\{\|v - s\| : s \in S\}.$$

La proyección ortogonal nos permitirá resolver el problema enunciado al comienzo y calcular la distancia de un punto a un subespacio.

Proposición 8.29 *Sea $(V, \langle , \rangle)$ un espacio vectorial de dimensión finita con producto interno y sea $S \subseteq V$ un subespacio. Entonces para cada $v \in V$, $d(v, S) = \|v - p_S(v)\|$. En otras palabras, el punto de S a menor distancia de v es $p_S(v)$.*

Demostración. Sea $B = \{v_1, \dots, v_r, v_{r+1}, \dots, v_n\}$ una base ortonormal de V tal que $\{v_1, \dots, v_r\}$ es una base de S.

Sea $v \in V$. Se tiene que $v = \sum_{i=1}^{n} \langle v, v_i\rangle .\, v_i$ y $p_S(v) = \sum_{i=1}^{r} \langle v, v_i\rangle .\, v_i$. Por otro lado, para cada $s \in S$, vale $s = \sum_{i=1}^{r} \langle s, v_i\rangle . v_i$. Entonces

$$v - s = \sum_{i=1}^{n} \langle v, v_i\rangle .\, v_i - \sum_{i=1}^{r} \langle s, v_i\rangle .\, v_i = \sum_{i=1}^{r} \langle v - s, v_i\rangle .\, v_i + \sum_{i=r+1}^{n} \langle v, v_i\rangle .\, v_i$$

y, en consecuencia,

$$\|v - s\|^2 = \sum_{i=1}^{r} |\langle v - s, v_i\rangle|^2 + \sum_{i=r+1}^{n} |\langle v, v_i\rangle|^2 \geq \sum_{i=r+1}^{n} |\langle v, v_i\rangle|^2.$$

Además, la igualdad vale si y sólo si

$$\sum_{i=1}^{r} |\langle v - s, v_i\rangle|^2 = 0 \iff |\langle v - s, v_i\rangle| = 0 \ \forall\, 1 \leq i \leq r \iff \langle s, v_i\rangle = \langle v, v_i\rangle \ \forall\, 1 \leq i \leq r,$$

es decir, para

$$s = \sum_{i=1}^{r} \langle s, v_i\rangle . v_i = \sum_{i=1}^{r} \langle v, v_i\rangle .\, v_i = p_S(v).$$

Luego, el punto de S a menor distancia de v es $p_S(v)$ y $d(v, S) = \|v - p_S(v)\|$. □

Como consecuencia de la Observación 8.27 y de la proposición anterior se deduce:

Observación 8.30 Sea V un espacio vectorial de dimensión finita con producto interno y sea S un subespacio de V. Entonces, para cada $v \in V$, vale $d(v, S) = \|p_{S^\perp}(v)\|$.

Ejemplo. Sea $S = \{x \in \mathbb{R}^3 \ / \ 2x_1 + 2x_2 - x_3 = 0\}$. Hallar la distancia de $(1, -1, 2)$ a S y el punto de S más cercano a $(1, 1, 2)$.

Sabemos que $d((1, -1, 2), S) = \|p_{S^\perp}(1, -1, 2)\|$. Calculamos entonces $p_{S^\perp}(1, -1, 2)$.

En primer lugar, observemos que $S = \{x \in \mathbb{R}^3 : \langle x, (2, 2, -1)\rangle = 0\} = < (2, 2, -1) >^\perp$, de donde $S^\perp = < (2, 2, -1) >$. Luego, $\{(2, 2, -1)\}$ es una base (ortogonal) de $S^\perp$. Entonces

$$p_{S^\perp}(1, -1, 2) = \frac{\langle (1, -1, 2), (2, 2, -1)\rangle}{\|(2, 2, -1)\|^2} .\, (2, 2, -1) = \frac{-2}{9} .\, (2, 2, -1) = \Big(\frac{-4}{9}, \frac{-4}{9}, \frac{2}{9}\Big),$$

y

$$d((1,-1,2),S) = \|p_{S^\perp}(1,-1,2)\| = \left\|\left(\frac{-4}{9}, \frac{-4}{9}, \frac{2}{9}\right)\right\| = \frac{2}{3}.$$

El punto de S más cercano a $(1,-1,2)$ es $p_S(1,-1,2)$, que podemos calcular como

$$p_S(1,-1,2) = (1,-1,2) - p_{S^\perp}(1,-1,2) = (1,-1,2) - \left(\frac{-4}{9}, \frac{-4}{9}, \frac{2}{9}\right) = \left(\frac{13}{9}, \frac{-5}{9}, \frac{16}{9}\right).$$

8.3 Endomorfismos en espacios vectoriales con producto interno

8.3.1 Adjunta de una transformación lineal

En lo que sigue, le asociaremos a cada endomorfismo f de un espacio vectorial V de dimensión finita con producto interno, otra transformación lineal $f^* : V \to V$.

Definición 8.31 Sea $(V, \langle , \rangle)$ un espacio vectorial con producto interno y sea $f : V \to V$ una transformación lineal. Se llama *adjunta de* f, y se nota f^*, a una transformación lineal $f^* : V \to V$ tal que

$$\langle f(v), w\rangle = \langle v, f^*(w)\rangle \quad \forall\, v, w \in V.$$

Ejemplo. Consideremos $\mathbb{C}^2$ con el producto interno canónico. Sea $f : \mathbb{C}^2 \to \mathbb{C}^2$ la transformación lineal definida por $f(x,y) = (x+iy, 2x-(1+i)y)$. Se tiene que

$$\begin{aligned}
\langle f(x,y),(z,w)\rangle &= \langle (x+iy, 2x-(1+i)y),(z,w)\rangle \\
&= (x+iy)\,\overline{z} + (2x-(1+i)y)\,\overline{w} \\
&= x\,(\overline{z}+2\overline{w}) + y\,(i\overline{z}-(1+i)\overline{w}) \\
&= x\,\overline{(z+2w)} + y\,\overline{(-iz+(-1+i)w)} \\
&= \langle (x,y),(z+2w, -iz+(-1+i)w)\rangle.
\end{aligned}$$

Entonces, la función $f^* : \mathbb{C}^2 \to \mathbb{C}^2$ definida por $f^*(z,w) = (z+2w, -iz+(-1+i)w)$ satisface $\langle f(x,y),(z,w)\rangle = \langle (x,y), f^*(z,w)\rangle$ para todo par de vectores (x,y), (z,w) en $\mathbb{C}^2$.

Observar que, si E es la base canónica de $\mathbb{C}^2$, entonces

$$|f|_E = \begin{pmatrix} 1 & i \\ 2 & -1-i \end{pmatrix} \quad \text{y} \quad |f^*|_E = \begin{pmatrix} 1 & 2 \\ -i & -1+i \end{pmatrix},$$

y, por lo tanto, $|f^*|_E$ es la matriz transpuesta y conjugada de $|f|_E$.

El siguiente resultado prueba la existencia y unicidad de la adjunta para endomorfismos en espacios vectoriales de dimensión finita con producto interno.

Proposición 8.32 *Sea $(V, \langle , \rangle)$ un espacio vectorial de dimensión finita con producto interno y sea $f : V \to V$ una transformación lineal. Entonces existe una única transformación lineal $f^* : V \to V$ que satisface $\langle f(v), w\rangle = \langle v, f^*(w)\rangle$ $\forall\, v, w \in V$.*

Demostración.

Unicidad. Supongamos que $f^* : V \to V$ y $\widetilde{f^*} : V \to V$ son dos transformaciones lineales que verifican la propiedad del enunciado.

Fijemos $w \in V$. Para cada $v \in V$, se tiene que

$$\langle f(v), w\rangle = \langle v, f^*(w)\rangle \quad \text{y} \quad \langle f(v), w\rangle = \langle v, \widetilde{f^*}(w)\rangle.$$

Entonces $\langle v, f^*(w)\rangle - \langle v, \widetilde{f^*}(w)\rangle = 0$ para todo $v \in V$ o, equivalentemente,

$$\langle v,\, f^*(w) - \widetilde{f^*}(w)\rangle = 0 \quad \forall\, v \in V.$$

En particular, tomando $v = f^*(w) - \widetilde{f^*}(w)$ resulta que $\langle f^*(w) - \widetilde{f^*}(w),\, f^*(w) - \widetilde{f^*}(w)\rangle = 0$, lo que implica que $f^*(w) - \widetilde{f^*}(w) = 0$. Luego, $f^*(w) = \widetilde{f^*}(w)$.

Como esto vale para cada $w \in V$, concluimos que $f^* = \widetilde{f^*}$.

Existencia. Sea $\{v_1, \dots, v_n\}$ una base ortonormal de V. Si existe $f^* : V \to V$ con las condiciones del enunciado, para cada $w \in V$ debe cumplirse

$$f^*(w) = \sum_{i=1}^{n} \langle f^*(w), v_i\rangle\, v_i = \sum_{i=1}^{n} \overline{\langle v_i, f^*(w)\rangle}\, v_i = \sum_{i=1}^{n} \overline{\langle f(v_i), w\rangle}\, v_i = \sum_{i=1}^{n} \langle w, f(v_i)\rangle\, v_i.$$

Definimos entonces $f^* : V \to V$ como $f^*(w) = \sum\limits_{i=1}^{n} \langle w, f(v_i)\rangle\, v_i$.

Veamos que la función f^* que definimos es una transformación lineal que cumple la propiedad del enunciado:

- f^* es una transformación lineal:

  – Para $w, w' \in V$, se tiene que

$$\begin{aligned} f^*(w + w') &= \sum_{i=1}^{n} \langle w + w', f(v_i)\rangle\, v_i = \sum_{i=1}^{n} (\langle w, f(v_i)\rangle + \langle w', f(v_i)\rangle)\, v_i \\ &= \sum_{i=1}^{n} \langle w, f(v_i)\rangle\, v_i + \sum_{i=1}^{n} \langle w', f(v_i)\rangle\, v_i = f^*(w) + f^*(w'). \end{aligned}$$

  – Para $\lambda \in \mathbb{C}$ (o $\mathbb{R}$), $w \in V$, vale

$$f^*(\lambda\, w) = \sum_{i=1}^{n} \langle \lambda w, f(v_i)\rangle\, v_i = \sum_{i=1}^{n} \lambda\, \langle w, f(v_i)\rangle\, v_i = \lambda\, f^*(w).$$

- Para todo par de vectores $v, w \in V$, vale $\langle f(v), w\rangle = \langle v, f^*(w)\rangle$:

  Sean $v, w \in V$. Se tiene que $v = \sum_{i=1}^{n} \langle v, v_i\rangle\, v_i$ y entonces $f(v) = \sum_{i=1}^{n} \langle v, v_i\rangle\, f(v_i)$. Observamos que

$$\langle f(v), w\rangle \;=\; \Big\langle \sum_{i=1}^{n} \langle v, v_i\rangle f(v_i),\, w\Big\rangle = \sum_{i=1}^{n} \langle v, v_i\rangle\, \langle f(v_i), w\rangle.$$

  Por otro lado,

$$\begin{aligned} \langle v, f^*(w)\rangle &= \Big\langle \sum_{i=1}^{n} \langle v, v_i\rangle\, v_i,\, \sum_{j=1}^{n} \langle w, f(v_j)\rangle\, v_j \Big\rangle \;=\; \sum_{i=1}^{n} \langle v, v_i\rangle \Big\langle v_i,\, \sum_{j=1}^{n} \langle w, f(v_j)\rangle\, v_j \Big\rangle \\ &= \sum_{i=1}^{n} \langle v, v_i\rangle \Big(\sum_{j=1}^{n} \overline{\langle w, f(v_j)\rangle} \langle v_i, v_j\rangle \Big) \;=\; \sum_{i=1}^{n} \langle v, v_i\rangle \overline{\langle w, f(v_i)\rangle} \\ &= \sum_{i=1}^{n} \langle v, v_i\rangle\, \langle f(v_i), w\rangle. \end{aligned}$$

  Concluimos entonces que $\langle f(v), w\rangle = \langle v, f^*(w)\rangle$. □

A partir de la matriz de un endomorfismo $f : V \to V$ en una base ortonormal de V, puede obtenerse fácilmente la matriz de su adjunta en la misma base:

Proposición 8.33 *Sea $(V, \langle , \rangle)$ un espacio vectorial de dimensión finita con producto interno y sea $f : V \to V$ una transformación lineal. Sea B una base ortonormal de V. Entonces $|f^*|_B = (|f|_B)^*$.*

Demostración. Supongamos que $B = \{v_1, \dots, v_n\}$ es la base ortonormal dada de V. Entonces, para cada $1 \le i, j \le n$,

$$(|f^*|_B)_{ij} = \langle f^*(v_j), v_i\rangle = \overline{\langle v_i, f^*(v_j)\rangle} = \overline{\langle f(v_i), v_j\rangle} = (\overline{|f|_B})_{ji} = ((|f|_B)^*)_{ij},$$

de donde concluimos que $|f^*|_B = (|f|_B)^*$. □

Este resultado puede utilizarse para hallar la adjunta de una transformación lineal:

Ejemplo. Sea $f : \mathbb{C}^3 \to \mathbb{C}^3$ la transformación lineal definida por

$$f(x, y, z) = (x + iy - iz, (2 + i)x + iy + z, (1 + i)y + 2z).$$

Hallar f^* para el producto interno canónico de $\mathbb{C}^3$.

Consideremos la base canónica E de $\mathbb{C}^3$, que es una base ortonormal para el producto interno canónico. Se tiene que

$$|f|_E = \begin{pmatrix} 1 & i & -i \\ 2+i & i & 1 \\ 0 & 1+i & 2 \end{pmatrix}.$$

Por la proposición anterior,

$$|f^*|_E = (|f|_E)^* = \begin{pmatrix} 1 & 2-i & 0 \\ -i & -i & 1-i \\ i & 1 & 2 \end{pmatrix}.$$

Luego, $f^*(x, y, z) = (x + (2-i)y, -ix - iy + (1-i)z, ix + y + 2z)$.

8.3.2 Transformaciones autoadjuntas y matrices hermitianas

En esta sección estudiaremos una clase particular de transformaciones lineales en espacios con producto interno: las transformaciones lineales $f : V \to V$ cuya adjunta coincide con f.

Definición 8.34 Sea $(V, \langle, \rangle)$ un espacio vectorial de dimensión finita con producto interno y sea $f : V \to V$ una transformación lineal. Se dice que f es *autoadjunta* si $f^* = f$.

Esta definición puede reescribirse en términos de la transformación lineal y el producto interno del espacio considerado:

Observación 8.35 Una transformación lineal $f : V \to V$ es autoadjunta si y sólo si

$$\langle f(x), y\rangle = \langle x, f(y)\rangle \qquad \forall\, x, y \in V.$$

En lo que sigue veremos que la matriz de una transformación lineal autoadjunta en una base ortonormal tiene cierta estructura particular. Más precisamente, las matrices involucradas serán del siguiente tipo:

Definición 8.36 Una matriz $A \in \mathbb{R}^{n\times n}$ se dice *simétrica* si $A_{ij} = A_{ji}\ \forall\, 1 \le i, j \le n$ o, equivalentemente, si $A = A^t$. Una matriz $A \in \mathbb{C}^{n\times n}$ se dice *hermitiana* si $A_{ij} = \overline{A_{ji}}\ \forall\, 1 \le i, j \le n$ o, equivalentemente, si $A = A^*$.

Proposición 8.37 *Sea $(V, \langle, \rangle)$ un espacio vectorial de dimensión finita con producto interno y sea $f : V \to V$ una transformación lineal. Son equivalentes:*

i) *f es autoadjunta.*

ii) *$\forall\, B$ base ortonormal de V, $|f|_B$ es hermitiana.*

iii) *$\exists\, B$ base ortonormal de V tal que $|f|_B$ es hermitiana.*

Demostración. La equivalencia de los tres enunciados se deduce de la Proposición 8.33 y la definición de transformación lineal autoadjunta. □

Diagonalización de transformaciones lineales autoadjuntas

A continuación nos concentraremos en el estudio de la diagonalización de transformaciones lineales autoadjuntas. Probaremos que si $f : V \to V$ es una transformación lineal autoadjunta, entonces f es diagonalizable. Más aún, veremos que existe una base ortonormal de V formada por autovectores de f y que todos sus autovalores son reales.

Comenzamos con el estudio de los autovalores:

Proposición 8.38 *Sea $(V, \langle , \rangle)$ un espacio vectorial de dimensión finita con producto interno. Sea $f : V \to V$ una transformación lineal autoadjunta. Entonces, el polinomio característico de f tiene todas sus raíces en $\mathbb{R}$.*

Demostración. Consideraremos por separado los casos en que f está definida en un $\mathbb{C}$-espacio vectorial o en un $\mathbb{R}$-espacio vectorial.

- Si V es un $\mathbb{C}$-espacio vectorial:

  Sea $\lambda \in \mathbb{C}$ una raíz de $\mathcal{X}_f$. Entonces, λ es un autovalor de f. Sea $v \in V$ un autovector de f de autovalor $\lambda \in \mathbb{C}$. Se tiene que

  $$\lambda \,.\, \|v\|^2 = \lambda \,.\, \langle v, v\rangle = \langle \lambda v, v\rangle = \langle f(v), v\rangle = \langle v, f(v)\rangle = \langle v, \lambda v\rangle = \bar{\lambda} \,.\, \|v\|^2.$$

  Como $\|v\| \neq 0$, resulta que $\lambda = \bar{\lambda}$, es decir, $\lambda \in \mathbb{R}$.

- Si V es un $\mathbb{R}$-espacio vectorial:

  Sea B una base ortonormal de V y sea $A = |f|_B \in \mathbb{R}^{n\times n}$. Como f es autoadjunta, A es una matriz simétrica.

  Si $\lambda \in \mathbb{C}$ es raíz del polinomio característico de f, entonces $\lambda \in \mathbb{C}$ es un autovalor de A considerada como matriz en $\mathbb{C}^{n\times n}$. Entonces existe $x \in \mathbb{C}^n$ autovector de A asociado al autovalor λ. Si consideramos $\mathbb{C}^n$ con el producto interno canónico, entonces

  $$\lambda \,.\, \|x\|^2 = \langle \lambda x, x\rangle = \langle Ax, x\rangle = \langle x, Ax\rangle = \langle x, \lambda x\rangle = \bar{\lambda} \,.\, \|x\|^2,$$

  de donde $\lambda = \bar{\lambda}$ y, por lo tanto $\lambda \in \mathbb{R}$. □

Esta proposición nos dice, en particular:

Observación 8.39 Sea $A \in \mathbb{C}^{n\times n}$ una matriz hermitiana. Entonces todas las raíces del polinomio característico de A son reales.

Probamos ahora un resultado sobre diagonalización de transformaciones lineales autoadjuntas:

Teorema 8.40 *Sea $(V, \langle , \rangle)$ un espacio vectorial de dimensión finita con producto interno. Sea $f : V \to V$ una transformación lineal autoadjunta. Entonces existe una base ortonormal B de V tal que $|f|_B$ es diagonal real.*

Demostración. Por inducción en $n = \dim V$:

Para $n = 1$, el resultado es inmediato.

Sea $n = \dim V > 1$ y supongamos que la propiedad vale para transformaciones lineales autoadjuntas definidas en espacios de dimensión $n - 1$.

Sea $f : V \to V$ una transformación lineal autoadjunta.

Por la proposición anterior, existe $\lambda \in \mathbb{R}$ autovalor de f. Sea v un autovector asociado al autovalor λ y sea $w = \frac{v}{\|v\|}$, que es un autovector de norma 1 asociado al mismo autovalor.

Sea $S = < w >^{\perp}$. Se tiene que S es un subespacio de V con $\dim S = n - 1$. Además, S es f-invariante, puesto que para cada $x \in S$ se tiene que

$$\langle f(x), w\rangle = \langle x, f(w)\rangle = \langle x, \lambda w\rangle = \lambda\langle x, w\rangle = 0,$$

de donde $f(x) \in < w >^{\perp} = S$.

Consideremos S con el producto interno obtenido por la restricción a S del producto interno de V, y la transformación lineal $f_{|_S} : S \to S$, que resulta autoadjunta. Por hipótesis inductiva, existe una base ortonormal B' de S tal que $|f_{|_S}|_{B'}$ es diagonal real.

Sea $B = \{w\} \cup B'$. Entonces, B es una base ortonormal de V y

$$|f|_B = \begin{pmatrix} \lambda & 0 \\ 0 & |f_{|_S}|_{B'} \end{pmatrix}$$

es una matriz diagonal real. □

A partir de este teorema es posible deducir un resultado análogo sobre diagonalización de matrices hermitianas:

Observación 8.41 Sea $A \in \mathbb{C}^{n\times n}$ una matriz hermitiana. Entonces, si se considera $\mathbb{C}^n$ con el producto interno canónico, la transformación lineal $f_A : \mathbb{C}^n \to \mathbb{C}^n$, definida por $f_A(x) = A.x$, es autoadjunta y, si E es la base canónica de $\mathbb{C}^n$, $|f_A|_E = A$.

Por la proposición anterior, existe una base ortonormal B de $\mathbb{C}^n$ tal que $|f_A|_B = D$ es diagonal real. Entonces

$$C(B, E)^{-1}. A. C(B, E) = D.$$

Además, como E y B son bases ortonormales de $\mathbb{C}^n$, la matriz de cambio de base satisface:

$$(C(B, E)^{-1})_{ij} = C(E, B)_{ij} = \langle e_j, v_i\rangle = \overline{\langle v_i, e_j\rangle} = \overline{C(B, E)_{ji}} = (C(B, E)^*)_{ij}$$

para todo $1 \le i, j \le n$. En consecuencia, $C(B, E)^{-1} = C(B, E)^*$.

Análogamente, si $A \in \mathbb{R}^{n\times n}$ es una matriz simétrica, la transformación lineal $f_A : \mathbb{R}^n \to \mathbb{R}^n$ es autoadjunta para el producto interno canónico de $\mathbb{R}^n$ y, por lo tanto, existe una base ortonormal B de $\mathbb{R}^n$ tal que $|f_A|_B = D \in \mathbb{R}^{n\times n}$ es diagonal. En este caso, la matriz de cambio de base que da la igualdad $C(B, E)^{-1}. A. C(B, E) = D$ cumple: $C(B, E)^{-1} = C(B, E)^t$.

Esto nos lleva a la siguiente definición:

Definición 8.42 Una matriz $U \in \mathbb{C}^{n\times n}$ se dice *unitaria* si es inversible y $U^{-1} = U^*$. Una matriz $O \in \mathbb{R}^{n\times n}$ se dice *ortogonal* si es inversible y $O^{-1} = O^t$.

Utilizando esta definición, la Observación 8.41 puede resumirse como sigue:

Corolario 8.43 *Sea $A \in \mathbb{C}^{n\times n}$ una matriz hermitiana. Entonces existe una matriz unitaria $C \in \mathbb{C}^{n\times n}$ tal que $C^*. A. C$ es diagonal real.*

Sea $A \in \mathbb{R}^{n\times n}$ una matriz simétrica. Entonces existe una matriz ortogonal $O \in \mathbb{R}^{n\times n}$ tal que $O^t. A. O$ es diagonal.

El hecho que la matriz de cambio de base de una base ortonormal B de $\mathbb{C}^n$ (respectivamente $\mathbb{R}^n$) a la base canónica de $\mathbb{C}^n$ (respectivamente $\mathbb{R}^n$) es una matriz unitaria (respectivamente ortogonal) vale también para dos bases ortonormales cualesquiera:

Observación 8.44 Sea $(V, \langle , \rangle)$ un espacio vectorial de dimensión finita con producto interno y sean B y B' bases ortonormales de V. Entonces $C(B, B')$ es una matriz unitaria (u ortogonal).

Supongamos que $B = \{v_1, \dots, v_n\}$ y $B' = \{w_1, \dots, w_n\}$. Entonces, para cada $1 \le i, j \le n$, vale

$$(C(B, B')^{-1})_{ij} = C(B', B)_{ij} = \langle w_j, v_i\rangle = \overline{\langle v_i, w_j\rangle} = \overline{C(B, B')_{ji}} = (C(B, B')^*)_{ij},$$

de donde $C(B, B')^{-1} = C(B, B')^*$, es decir, $C(B, B')$ es una matriz unitaria.

8.3.3 Transformaciones unitarias y ortogonales

En esta sección estudiaremos los endormorfismos de un espacio vectorial $(V, \langle , \rangle)$ que preservan el producto interno y, en particular, las distancias entre vectores. El siguiente teorema caracteriza dichos endomorfismos.

Teorema 8.45 *Sea $(V, \langle , \rangle)$ un espacio vectorial de dimensión finita con producto interno y sea $f : V \to V$ una transformación lineal. Son equivalentes:*

i) Existe una base ortonormal B de V tal que $f(B)$ es una base ortonormal de V.

ii) $\langle f(v), f(w)\rangle = \langle v, w\rangle \ \forall\, v, w \in V$.

iii) Para toda base ortonormal B de V, $f(B)$ es una base ortonormal de V.

iv) $\|f(v)\| = \|v\| \ \forall\, v \in V$.

v) $f^* \circ f = f \circ f^* = id_V$.

Definición 8.46 Sea $(V, \langle , \rangle)$ un espacio vectorial con producto interno. Una transformación lineal $f : V \to V$ que cumple las condiciones equivalentes del Teorema 8.45 se dice *unitaria* si V es un $\mathbb{C}$-espacio vectorial, u *ortogonal* si V es un $\mathbb{R}$-espacio vectorial.

Demostración del Teorema 8.45. Probaremos que $i) \Rightarrow ii) \Rightarrow iii) \Rightarrow i)$, $ii) \Leftrightarrow iv)$ y $ii) \Leftrightarrow v)$.

$i) \Rightarrow ii)$ Sea $B = \{v_1, \ldots, v_n\}$ una base ortonormal de V tal que $f(B) = \{f(v_1), \ldots, f(v_n)\}$ es también una base ortonormal de V. Entonces, si $v, w \in V$ con $v = \sum\limits_{i=1}^{n} \alpha_i v_i$ y $w = \sum\limits_{i=1}^{n} \beta_i v_i$, como B y $f(B)$ son bases ortonormales de V, se tiene que

$$\begin{aligned} \langle v, w\rangle &= \Big\langle \sum_{i=1}^{n} \alpha_i v_i, \sum_{j=1}^{n} \beta_j v_j \Big\rangle = \sum_{i=1}^{n} \alpha_i \overline{\beta_i} \\ \langle f(v), f(w)\rangle &= \Big\langle \sum_{i=1}^{n} \alpha_i f(v_i), \sum_{j=1}^{n} \beta_j f(v_j) \Big\rangle = \sum_{i=1}^{n} \alpha_i \overline{\beta_i}, \end{aligned}$$

de donde $\langle v, w\rangle = \langle f(v), f(w)\rangle$.

$ii) \Rightarrow iii)$ Sea $B = \{v_1, \ldots, v_n\}$ una base ortonormal de V. Consideramos el conjunto $f(B) = \{f(v_1), \ldots, f(v_n)\}$. Se tiene que

$$\begin{aligned} \langle f(v_i), f(v_i)\rangle &= \langle v_i, v_i\rangle = 1 \quad \forall\, 1 \le i \le n \\ \langle f(v_i), f(v_j)\rangle &= \langle v_i, v_j\rangle = 0 \quad \forall\, 1 \le i, j \le n,\ i \ne j. \end{aligned}$$

Luego, $f(B)$ es una base ortonormal de V.

$iii) \Rightarrow i)$ No hay nada que probar.

$ii) \Rightarrow iv)$ Para cada $v \in V$, se tiene que $\|f(v)\| = \langle f(v), f(v)\rangle^{\frac{1}{2}} = \langle v, v\rangle^{\frac{1}{2}} = \|v\|$, donde la segunda igualdad es consecuencia de la hipótesis *ii)*.

$iv) \Rightarrow ii)$ Sean $v, w \in V$. Aplicando la primera de las identidades de polarización (ver Proposición 8.6) resulta, si V es un espacio vectorial real,

$$\begin{aligned} \langle f(v), f(w)\rangle &= \frac{1}{4}\|f(v) + f(w)\|^2 - \frac{1}{4}\|f(v) - f(w)\|^2 \\ &= \frac{1}{4}\|f(v + w)\|^2 - \frac{1}{4}\|f(v - w)\|^2 \\ &= \frac{1}{4}\|v + w\|^2 - \frac{1}{4}\|v - w\|^2 \\ &= \langle v, w\rangle. \end{aligned}$$

En el caso en que V es un espacio vectorial complejo, la igualdad se prueba análogamente utilizando la segunda de las identidades de polarización.

$ii) \Rightarrow v)$ Sea $v \in V$. Para cada $w \in V$, se tiene que

$$\langle f^* \circ f(v), w\rangle = \langle f^*(f(v)), w\rangle = \langle f(v), f(w)\rangle = \langle v, w\rangle,$$

lo que implica que $\langle f^* \circ f(v) - v,\, w\rangle = 0$. Tomando $w = f^* \circ f(v) - v$, resulta que $\langle f^* \circ f(v) - v, f^* \circ f(v) - v\rangle = 0$, de donde concluimos que $f^* \circ f(v) = v$.

En consecuencia, $f^* \circ f = id_V$. Como V es de dimensión finita, esto implica que también vale $f \circ f^* = id_V$.

$v) \Rightarrow ii)$ Para $v, w \in V$, se tiene que $\langle f(v), f(w)\rangle = \langle v, f^*(f(w))\rangle = \langle v,\, f^* \circ f(w)\rangle = \langle v,\, id_V(w)\rangle = \langle v, w\rangle$. □

La proposición siguiente nos da la relación entre transformaciones lineales unitarias (ortogonales) y matrices unitarias (ortogonales).

Proposición 8.47 *Sea $(V, \langle , \rangle)$ un espacio vectorial de dimensión finita con producto interno y sea B una base ortonormal de V. Sea $f : V \to V$ una trasformación lineal. Entonces*

$$f \text{ es unitaria (ortogonal)} \iff |f|_B \text{ es unitaria (ortogonal).}$$

Demostración. Haremos la demostración para transformaciones lineales y matrices unitarias (el otro caso es totalmente análogo). Sea $n = \dim V$.

($\Rightarrow$) Si f es unitaria, entonces $f^* \circ f = f \circ f^* = id_V$. Para una base ortonormal B de V se tiene que

$$I_n = |f^* \circ f|_B = |f^*|_B \cdot |f|_B = (|f|_B)^* \cdot |f|_B.$$

En consecuencia, $|f|_B$ es inversible y $(|f|_B)^{-1} = (|f|_B)^*$, con lo que es una matriz unitaria.

($\Leftarrow$) Si $|f|_B$ es unitaria, $|f|_B^{-1} = (|f|_B)^*$, de donde resulta que $|f^* \circ f|_B = |f \circ f^*|_B = I_n$ y, por lo tanto, $f^* \circ f = f \circ f^* = id_V$. □

8.3.4 Clasificación de transformaciones ortogonales

Para terminar, estudiaremos más en detalle las transformaciones ortogonales en un espacio euclídeo V. Veremos que dada $f : V \to V$ ortogonal, existe una base ortonormal de V en la que la matriz de f tiene cierta estructura particular. Esto nos permitirá clasificar las transformaciones ortogonales.

A lo largo de esta sección, V denotará un espacio euclídeo, es decir, un espacio vectorial real de dimensión finita con producto interno.

En primer lugar, probamos dos resultados acerca de los autovalores de una transformación ortogonal y de sus subespacios invariantes.

Lema 8.48 *Sea $f : V \to V$ una transformación lineal ortogonal y sea $\lambda \in \mathbb{R}$ un autovalor de f. Entonces $\lambda = \pm 1$.*

Demostración. Si $\lambda \in \mathbb{R}$ es autovalor de f, existe $v \in V$, $v \neq 0$, tal que $f(v) = \lambda.v$. Como f es ortogonal vale $\|f(v)\| = \|v\|$. Entonces

$$\|v\| = \|f(v)\| = \|\lambda.v\| = |\lambda|.\|v\|,$$

de donde resulta que $|\lambda| = 1$ puesto que $v \neq 0$. Luego, $\lambda = \pm 1$. □

Lema 8.49 *Sea $f : V \to V$ una transformación lineal ortogonal. Sea $S \subseteq V$ un subespacio f-invariante. Entonces $S^\perp$ es f-invariante.*

Demostración. Sea $x \in S^\perp$. Veamos que $f(x) \in S^\perp$: Sea $s \in S$. Se tiene que $f_{|S} : S \to S$ es una transformación lineal ortogonal, lo que implica que es un isomorfismo. Entonces existe $\widetilde{s} \in S$ tal que $s = f(\widetilde{s})$ y, en consecuencia,

$$\langle f(x), s\rangle = \langle f(x), f(\widetilde{s})\rangle = \langle x, \widetilde{s}\rangle = 0.$$

En conclusión, $\langle f(x), s\rangle = 0$ para todo $s \in S$. Luego, $f(x) \in S^\perp$. □

Clasificación de transformaciones ortogonales en un espacio de dimensión 2

En lo que sigue, daremos una caracterización para las trasformaciones lineales ortogonales definidas en espacios vectoriales de dimensión 2. En particular, clasificaremos las transformaciones ortogonales en $\mathbb{R}^2$.

Sea V un espacio euclídeo con $\dim V = 2$ y sea $f : V \to V$ una transformación lineal ortogonal.

Sea $B = \{v_1, v_2\}$ una base ortonormal de V. Entonces $\{f(v_1), f(v_2)\}$ es una base ortonormal de V y, si

$$f(v_1) = \alpha v_1 + \beta v_2 \quad \text{y} \quad f(v_2) = \alpha' v_1 + \beta' v_2,$$

resulta que $\{(\alpha, \beta), (\alpha', \beta')\}$ es una base ortonormal de $\mathbb{R}^2$, es decir,

$$\|(\alpha, \beta)\| = \|(\alpha', \beta')\| = 1 \quad \text{y} \quad \alpha\alpha' + \beta\beta' = 0.$$

De estas condiciones se desprende que $\alpha^2+\beta^2 = 1$ y que $(\alpha', \beta') = (-\beta, \alpha)$ o $(\alpha', \beta') = (\beta, -\alpha)$. En consecuencia, la matriz de f en la base B es de alguna de las dos formas siguientes:

$$(1)\ \ |f|_B = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix} \qquad \text{o} \qquad (2)\ \ |f|_B = \begin{pmatrix} \alpha & \beta \\ \beta & -\alpha \end{pmatrix}.$$

(1) En este caso se tiene que $\mathcal{X}_f = (X - \alpha)^2 + \beta^2 = X^2 - 2\alpha X + 1$.
Si $\alpha = \pm 1$, entonces $f = \pm id_V$. Si no, $\mathcal{X}_f$ no tiene raíces reales.
Por otro lado, como $\|(\alpha, \beta)\| = 1$, existe $\theta \in [0, 2\pi)$ tal que $\alpha = \cos\theta$, $\beta = \operatorname{sen}\theta$. Luego

$$|f|_B = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}.$$

Eventualmente cambiando la base $\{v_1, v_2\}$ por $\{v_1, -v_2\}$ se puede tomar $\theta \in [0, \pi]$.

(2) Como $|f|_B$ es simétrica, existe una base ortonormal B' de V tal que $|f|_{B'}$ es diagonal. Puesto que $\mathcal{X}_f = (X-\alpha)(X+\alpha)-\beta^2 = X^2-1 = (X-1)(X+1)$ se puede tomar B' tal que
$$|f|_{B'} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$
Observamos que si $B' = \{w_1, w_2\}$ entonces $f(w_1) = w_1$ y $f(w_2) = -w_2$.

En el caso en que $V = \mathbb{R}^2$, clasificaremos las transformaciones lineales ortogonales en los siguientes tipos:

Definición 8.50 i) Sea $f : \mathbb{R}^2 \to \mathbb{R}^2$ una transformación lineal ortogonal. Se dice que f es una *rotación* si $\det(f) = 1$.

ii) Sea $f : \mathbb{R}^2 \to \mathbb{R}^2$ una transformación ortogonal. Sea $H \subseteq \mathbb{R}^2$ un subespacio de dimensión 1. Se dice que f es una *simetría respecto de* H si $f_{|_H} = id_H$ y $f_{|_{H^\perp}} = -id_{H^\perp}$.

De nuestro análisis anterior se deduce que:

Observación 8.51 Sea $f : \mathbb{R}^2 \to \mathbb{R}^2$ una transformación lineal ortogonal. Entonces f es una rotación o f es una simetría.

Vimos que si f es una transformación ortogonal, existe una base ortonormal $B = \{v_1, v_2\}$ de $\mathbb{R}^2$ tal que
$$|f|_B = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix} \quad \text{o} \quad |f|_B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

En el primer caso, $\det(|f|_B) = 1$, con lo cual, f es una rotación.

En el segundo caso, $f(v_1) = v_1$ y $f(v_2) = -v_2$. Además, $< v_1 >^\perp = < v_2 >$. Luego, f es una simetría respecto del subespacio $< v_1 >$.

Ejemplos.

1. Hallar la simetría $f : \mathbb{R}^2 \to \mathbb{R}^2$ respecto de la recta L de ecuación $x + y = 0$.

   Tenemos que $L = < (1,-1) >$ y $L^\perp = < (1,1) >$. Definimos $f : \mathbb{R}^2 \to \mathbb{R}^2$ en la base $\{(1,-1), (1,1)\}$ como sigue:
$$\begin{aligned} f(1,-1) &= (1,-1) \\ f(1,\ 1) &= (-1,-1). \end{aligned}$$
   Entonces $f_{|_L} = id_L$ y $f_{|_{L^\perp}} = -id_{L^\perp}$, lo que dice que f es la simetría respecto de la recta L.

   Se tiene que $f(x,y) = (-y,-x)$ para cada $(x,y) \in \mathbb{R}^2$.

2. Hallar una rotación $f : \mathbb{R}^2 \to \mathbb{R}^2$ tal que $f(2,1) = (1,2)$.

   En primer lugar, observemos que $\|f(2,1)\| = \|(1,2)\| = \sqrt{5} = \|(2,1)\|$ (recordar que una transformación ortogonal debe cumplir $\|f(v)\| = \|v\|$ para cada vector v).

   Vemos a definir f en una base ortonormal de $\mathbb{R}^2$. Para esto, consideramos el vector $\frac{(2,1)}{\|(2,1)\|} = (\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}})$ y construimos una base ortonormal de $\mathbb{R}^2$ que lo contiene:
$$B = \Big\{ \Big(\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}} \Big), \Big(\frac{-1}{\sqrt{5}}, \frac{2}{\sqrt{5}} \Big) \Big\}.$$
   La condición $f(2,1) = (1,2)$ equivale a $f(\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}) = (\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}})$. Escribimos este vector como combinación lineal de la base B:
$$\Big(\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}} \Big) = \frac{4}{5} \Big(\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}} \Big) + \frac{3}{5} \Big(\frac{-1}{\sqrt{5}}, \frac{2}{\sqrt{5}} \Big).$$
   Esto implica que la primera columna de la matriz $|f|_B$ de una transformación lineal f que verifique las condiciones del enunciado es $(\frac{4}{5}, \frac{3}{5})^t$. Teniendo en cuenta la estructura de la matriz de una rotación en una base ortonormal, concluimos que debe ser
$$|f|_B = \begin{pmatrix} \frac{4}{5} & \frac{-3}{5} \\ \frac{3}{5} & \frac{4}{5} \end{pmatrix}.$$
   En consecuencia, debe ser
$$f\Big(\frac{-1}{\sqrt{5}}, \frac{2}{\sqrt{5}} \Big) = \frac{-3}{5} \Big(\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}} \Big) + \frac{4}{5} \Big(\frac{-1}{\sqrt{5}}, \frac{2}{\sqrt{5}} \Big) = \Big(\frac{-2}{\sqrt{5}}, \frac{1}{\sqrt{5}} \Big).$$
   Consideramos entonces la transformación lineal $f : \mathbb{R}^2 \to \mathbb{R}^2$ definida como:
$$\begin{array}{rcl} f\Big(\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}} \Big) & = & \Big(\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}} \Big) \\ f\Big(\frac{-1}{\sqrt{5}}, \frac{2}{\sqrt{5}} \Big) & = & \Big(\frac{-2}{\sqrt{5}}, \frac{1}{\sqrt{5}} \Big). \end{array}$$
   Observamos que f es una transformación ortogonal, puesto que la imagen por f de una base ortonormal es una base ortonormal. Además, f es una rotación, porque $\det(|f|_B) = 1$. La condición $f(2,1) = (1,2)$ se deduce de la definición en el primer vector de la base.

Clasificación de transformaciones ortogonales en un espacio de dimensión 3

En lo que sigue, daremos una caracterización de las transformaciones ortogonales definidas en un espacio euclídeo V de dimensión 3.

En el caso en que $V = \mathbb{R}^3$, para hacer esta clasificación nos interesarán particularmente las transformaciones lineales ortogonales que introducimos en la definición siguiente.

Definición 8.52 i) Sea $f : \mathbb{R}^3 \to \mathbb{R}^3$ una transformación lineal ortogonal. Se dice que f es una *rotación* si $\det(f) = 1$.

ii) Sea $f : \mathbb{R}^3 \to \mathbb{R}^3$ una transformación ortogonal. Sea $H \subseteq \mathbb{R}^3$ un subespacio de dimensión 2. Se dice que f es una *simetría respecto de H* si $f_{|_H} = id_H$ y $f_{|_{H^\perp}} = -id_{H^\perp}$.

Sea V un espacio euclídeo con $\dim V = 3$ y sea $f : V \to V$ una transformación lineal ortogonal.

Se tiene que $\mathcal{X}_f \in \mathbb{R}[X]$ y $\text{gr}\mathcal{X}_f = 3$. Entonces, $\mathcal{X}_f$ tiene una raíz en $\mathbb{R}$, es decir, f tiene un autovalor real λ. Por el Lema 8.48, $\lambda = 1$ o $\lambda = -1$.

- Si $\lambda = 1$ es autovalor de f:

  Sea v_1 un autovector de f asociado al autovalor 1 con $\|v_1\| = 1$ (si $\|v_1\| \neq 1$, lo dividimos por su norma). Entonces $S = < v_1 >$ es un subespacio invariante por f. Por el Lema 8.49, $S^\perp$ es f-invariante.

  Consideremos $S^\perp$ con el producto interno inducido por el producto interno de V. Entonces $f_{|_{S^\perp}} : S^\perp \to S^\perp$ es una transformación ortogonal en un espacio euclídeo con $\dim S^\perp = 2$. En consecuencia, existe una base ortonormal $B_1 = \{v_2, v_3\}$ de $S^\perp$ tal que vale alguna de las dos igualdades siguientes:

  (1) $|f_{|_{S^\perp}}|_{B_1} = \begin{pmatrix} \cos\theta & -\text{sen}\,\theta \\ \text{sen}\,\theta & \cos\theta \end{pmatrix}$, con $\theta \in [0, \pi]$,

  (2) $|f_{|_{S^\perp}}|_{B_1} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

  Sea $B = \{v_1, v_2, v_3\}$. Entonces B es una base ortonormal de V y vale alguna de las dos igualdades siguientes:

  (1) $|f|_B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\text{sen}\,\theta \\ 0 & \text{sen}\,\theta & \cos\theta \end{pmatrix}$, con $\theta \in [0, \pi]$,

  (2) $|f|_B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$.

  En particular, vemos que si $V = \mathbb{R}^3$:

  – Si la matriz $|f|_B$ es como en (1), entonces f es una rotación. En este caso, al subespacio $< v_1 >$ se lo llama el *eje* de la rotación.

  – Si la matriz $|f|_B$ es como en (2), entonces f es una simetría respecto del subespacio $H = < v_1, v_2 >$.

- Si $\lambda = 1$ no es autovalor de f, entonces $\lambda = -1$ lo es. Sea v_1 un autovector de f de autovalor -1 con $\|v_1\| = 1$, y sea $S = < v_1 >$. De manera análoga a lo hecho en el caso anterior, se considera $S^\perp$, que es un subespacio f-invariante de dimensión 2, y la restricción $f_{|_{S^\perp}} : S^\perp \to S^\perp$ resulta una transformación lineal ortogonal. Como 1 no es autovalor de f, existe una base ortonormal $B_1 = \{v_2, v_3\}$ de $S^\perp$ tal que

$$|f_{|_{S^\perp}}|_{B_1} = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}, \quad \text{con } \theta \in (0, \pi].$$

Entonces $B = \{v_1, v_2, v_3\}$ es una base ortonormal de V tal que

$$|f|_B = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos\theta & -\operatorname{sen}\theta \\ 0 & \operatorname{sen}\theta & \cos\theta \end{pmatrix}, \quad \text{con } \theta \in (0, \pi].$$

Vemos que en el caso en que $V = \mathbb{R}^3$, f es una rotación compuesta con una simetría, puesto que:

$$|f|_B = \underbrace{\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{\text{simetría}} . \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\operatorname{sen}\theta \\ 0 & \operatorname{sen}\theta & \cos\theta \end{pmatrix}}_{\text{rotación}} .$$

Resumimos los resultados obtenidos sobre transformaciones lineales ortogonales en $\mathbb{R}^3$:

Observación 8.53 Sea $f : \mathbb{R}^3 \to \mathbb{R}^3$ una transformación lineal ortogonal. Entonces f es una rotación o una simetría o una composición de una simetría y una rotación.

Ejemplo. Definir una rotación $f : \mathbb{R}^3 \to \mathbb{R}^3$ tal que $f(1, 1, 0) = (0, 1, 1)$ y tal que el eje de la rotación sea ortogonal a $(1, 1, 0)$ y $(0, 1, 1)$.

Sea $H = < (1, 1, 0), (0, 1, 1) >$ el subespacio de $\mathbb{R}^3$ de dimensión 2 que contiene a $(1, 1, 0)$ y a $(0, 1, 1)$.

Consideremos $H^\perp = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 + x_2 = 0, x_2 + x_3 = 0\} = < (1, -1, 1) >$. Queremos definir f de manera que $H^\perp$ sea el eje de la rotación, es decir, que $f_{|_{H^\perp}} = id_{H^\perp}$.

En primer lugar, construimos una base ortonormal de $\mathbb{R}^3$ que contenga una base de $H^\perp$ y una base de H:

$$B = \Big\{ \underbrace{\Big(\frac{1}{\sqrt{3}}, \frac{-1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\Big)}_{\text{base de } H^\perp}, \underbrace{\Big(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\Big), \Big(\frac{-1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}\Big)}_{\text{base de } H} \Big\}.$$

Definimos f en los primeros vectores de la base como sigue:

$$(a)\ f\Big(\frac{1}{\sqrt{3}}, \frac{-1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\Big) = \Big(\frac{1}{\sqrt{3}}, \frac{-1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\Big)$$

$$(b)\ f\Big(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\Big) = \Big(0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\Big) = \frac{1}{2}\Big(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\Big) + \frac{\sqrt{3}}{2}\Big(\frac{-1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}\Big)$$

De esta manera, tendremos que la matriz de f en la base B es de la forma

$$|f|_B = \begin{pmatrix} 1 & 0 & * \\ 0 & \frac{1}{2} & * \\ 0 & \frac{\sqrt{3}}{2} & * \end{pmatrix}$$

Para que la matriz tenga la estructura del ítem (1) de la página 218, definimos

$$(c) \quad f\Big(\frac{-1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}\Big) = \frac{-\sqrt{3}}{2}\Big(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\Big) + \frac{1}{2}\Big(\frac{-1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}\Big).$$

En conclusión, $f : \mathbb{R}^3 \to \mathbb{R}^3$ es la transformación lineal definida por (a), (b) y (c). La condición (b) implica que $f(1,1,0) = (0,1,1)$. Además, para la base ortonormal B considerada, se tiene que

$$|f|_B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{-\sqrt{3}}{2} \\ 0 & \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix},$$

de donde se desprende que f es una rotación.

Clasificación general de transformaciones lineales ortogonales

A continuación generalizamos lo hecho en espacios de dimensión 2 y 3 al caso general de una transformación lineal ortogonal definida en un espacio euclídeo V de dimensión finita arbitraria.

Teorema 8.54 *Sea $(V, \langle , \rangle)$ un espacio euclídeo de dimensión finita. Sea $f : V \to V$ una transformación lineal ortogonal. Entonces existe una base ortonormal B de V tal que*

$$|f|_B = \begin{pmatrix} I_{n_1} & & & & & & \\ & -I_{n_2} & & & & & \\ & & \cos\theta_1 & -\operatorname{sen}\theta_1 & & & \\ & & \operatorname{sen}\theta_1 & \cos\theta_1 & & & \\ & & & & \ddots & & \\ & & & & & \cos\theta_r & -\operatorname{sen}\theta_r \\ & & & & & \operatorname{sen}\theta_r & \cos\theta_r \end{pmatrix} \qquad (8.3)$$

con $n_1, n_2 \in \mathbb{N}_0$ y $\theta_i \in (0, \pi)$ para cada $1 \le i \le r$ (donde todos los coeficientes de la matriz cuyos valores no están indicados son cero).

Demostración. Por inducción en $n = \dim V$.

Ya vimos que el resultado vale para $\dim V = 2$.

Sea $n > 2$ y supongamos que el resultado vale para transformaciones ortogonales definidas en espacios de dimensión menor que n. Sea V un espacio euclídeo tal que $\dim V = n$ y sea $f : V \to V$ una transformación lineal ortogonal. Por el Lema 8.48, sabemos que los posibles autovalores de f son 1 y -1.

- Si $\lambda = 1$ es autovalor de f:

  Sea v_1 una autovector de f de autovalor 1 tal que $\|v_1\| = 1$ (si $\|v_1\| \neq 1$, lo dividimos por su norma). Entonces $S = < v_1 >$ es un subespacio de V invariante por f, de donde $S^\perp$ también es invariante por f.

  Consideremos $f_1 = f_{|_{S^\perp}} : S^\perp \to S^\perp$, que es una transformación lineal ortogonal definida en un espacio euclídeo de dimensión $n-1$ (considerando en $S^\perp$ la restricción del producto interno de V). Por hipótesis inductiva, existe una base ortonormal B_1 de $S^\perp$ tal que $|f_1|_{B_1}$ es de la forma (8.3).

  Tomando $B = \{v_1\} \cup B_1$, se obtiene una base ortonormal de V tal que

  $$|f|_B = \begin{pmatrix} 1 & 0_{1\times(n-1)} \\ 0_{(n-1)\times 1} & |f_1|_{B_1} \end{pmatrix},$$

  que es de la forma (8.3).

- Si $\lambda = 1$ no es autovalor de f, pero $\lambda = -1$ sí lo es:

  Tomando un autovector v_1 de autovalor -1 con $\|v_1\| = 1$ y procediendo como en el caso anterior, resulta que existe una base ortonormal B de V tal que la matriz $|f|_B$ es de la forma (8.3) con $n_1 = 0$ (es decir, sin unos en la diagonal).

- Si f no tiene autovalores reales:

  En este caso, el polinomio minimal de f tiene la forma $m_f = P_1.P_2 \dots P_r$, donde para cada $1 \le i \le r$, $P_i \in \mathbb{R}[X]$ es un polinomio irreducible de grado 2. Se tiene que

  $$0 = m_f(f) = P_1(f) \circ P_2(f) \circ \dots \circ P_r(f).$$

  Entonces, si definimos $Q = P_2 \dots P_r \in \mathbb{R}[X]$, vale $Q(f) = P_2(f) \circ \dots \circ P_r(f)$ y, por lo tanto,

  $$(P_1(f) \circ Q(f))(v) = 0 \quad \forall\, v \in V.$$

  Como $Q \mid m_f$ y $Q \neq m_f$, existe $w \in V$ tal que $Q(f)(w) \neq 0$. Sea $v = Q(f)(w)$ y sea $S = < v, f(v) >$. Observamos que S es f-invariante, puesto que $P_1(f)(v) = 0$ y $\mathrm{gr}(P_1) = 2$. Además, v no es autovector de f puesto que f no tiene autovalores, con lo que $\dim(S) = 2$.

  Consideremos la restricción $f_1 = f_{|_S} : S \to S$, que es una transformación lineal ortogonal sin autovalores reales definida sobre un espacio euclídeo de dimensión 2. Entonces existe una base ortonormal B_S de S tal que

  $$|f_1|_{B_S} = \begin{pmatrix} \cos\theta_1 & -\operatorname{sen}\theta_1 \\ \operatorname{sen}\theta_1 & \cos\theta_1 \end{pmatrix} \quad \text{con } \theta_1 \in (0, \pi).$$

Dado que S es f-invariante y f es ortogonal, entonces $S^\perp$ es f-invariante. Sea $f_2 = f_{|_{S^\perp}} : S^\perp \to S^\perp$ la restricción, que es una transformación ortogonal definida sobre un espacio euclídeo con $\dim S^\perp = n-2$. Por hipótesis inductiva, existe una base ortonormal $B_{S^\perp}$ de $S^\perp$ tal que la matriz $|f_2|_{B_{S^\perp}}$ tiene la forma (8.3). Además, como f no tiene autovalores reales, f_2 tampoco los tiene, y por lo tanto en esta matriz no aparecen 1 ni -1 en la diagonal. Luego

$$|f_2|_{B_{S^\perp}} = \begin{pmatrix} \cos\theta_2 & -\operatorname{sen}\theta_2 & & & \\ \operatorname{sen}\theta_2 & \cos\theta_2 & & & \\ & & \ddots & & \\ & & & \cos\theta_r & -\operatorname{sen}\theta_r \\ & & & \operatorname{sen}\theta_r & \cos\theta_r \end{pmatrix}$$

con $\theta_i \in (0, \pi)$ para todo $2 \le i \le r$.

En consecuencia, tomando $B = B_S \cup B_{S^\perp}$ se obtiene una base ortonormal de V tal que

$$|f|_B = \begin{pmatrix} |f_1|_{B_S} & 0 \\ 0 & |f_2|_{B_{S^\perp}} \end{pmatrix} = \begin{pmatrix} \cos\theta_1 & -\operatorname{sen}\theta_1 & & & \\ \operatorname{sen}\theta_1 & \cos\theta_1 & & & \\ & & \ddots & & \\ & & & \cos\theta_r & -\operatorname{sen}\theta_r \\ & & & \operatorname{sen}\theta_r & \cos\theta_r \end{pmatrix}$$

con $\theta_i \in (0, \pi)$ para cada $1 \le i \le r$. □

8.4 Ejercicios

Ejercicio 1. Sea V un espacio vectorial y sea $\langle , \rangle$ un producto interno sobre V. Probar:

i) $\langle x, y + z\rangle = \langle x, y\rangle + \langle x, z\rangle$

ii) $\langle x, cy\rangle = \bar{c} \,.\, \langle x, y\rangle$

iii) $\langle x, y\rangle = \langle x, z\rangle \ \forall\, x \in V \ \Rightarrow \ y = z$

Ejercicio 2. Sea $\big(V, \langle\, , \rangle\big)$ un espacio vectorial con producto interno. Probar que $|\langle x, y\rangle| = \|x\|.\|y\|$ si y sólo si $\{x, y\}$ es un conjunto linealmente dependiente.

Ejercicio 3. Sea V un espacio vectorial. Demostrar que la suma de dos productos internos sobre V es un producto interno sobre V.

Ejercicio 4. Sea V un espacio vectorial con producto interno y sea d la distancia asociada. Demostrar que:

i) $d(x, y) \ge 0$

ii) $d(x,y) = 0 \iff x = y$

iii) $d(x,y) = d(y,x)$

iv) $d(x,y) \leq d(x,z) + d(z,y)$

Ejercicio 5. Determinar si las siguientes funciones son o no productos internos. En caso afirmativo encontrar su matriz en la base canónica del espacio correspondiente.

i) $\Phi : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$, $\Phi(x,y) = 2.x_1.y_1 + 3.x_2.y_1 - x_2.y_2 + 3.x_1.y_2$

ii) $\Phi : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$, $\Phi(x,y) = x_1.y_1 + x_2.y_1 + 2.x_2.y_2 - 3.x_1.y_2$

iii) $\Phi : K^2 \times K^2 \to K$, $\Phi(x,y) = 2.x_1.y_1 + x_2.y_2 - x_1.y_2 - x_2.y_1$, con $K = \mathbb{R}$ y $K = \mathbb{C}$

iv) $\Phi : \mathbb{C}^2 \times \mathbb{C}^2 \to \mathbb{C}$, $\Phi(x,y) = 2.x_1.\overline{y}_1 + x_2.\overline{y}_2 - x_1.\overline{y}_2 - x_2.\overline{y}_1$

v) $\Phi : \mathbb{C}^2 \times \mathbb{C}^2 \to \mathbb{C}$, $\Phi(x,y) = 2.x_1.\overline{y}_1 + (1+i).x_1.\overline{y}_2 + (1+i).x_2.\overline{y}_1 + 3.x_2.\overline{y}_2$

vi) $\Phi : \mathbb{C}^2 \times \mathbb{C}^2 \to \mathbb{C}$, $\Phi(x,y) = x_1.\overline{y}_1 - i.x_1.\overline{y}_2 + i.x_2.\overline{y}_1 + 2.x_2.\overline{y}_2$

vii) $\Phi : K^3 \times K^3 \to K$, $\Phi(x,y) = 2.x_1.\overline{y}_1 + x_3.\overline{y}_3 - x_1.\overline{y}_3 - x_3.\overline{y}_1$, con $K = \mathbb{R}$ y $K = \mathbb{C}$

viii) $\Phi : K^3 \times K^3 \to K$, $\Phi(x,y) = 3.x_1.\overline{y}_1 + x_2.\overline{y}_1 + 2.x_2.\overline{y}_2 + x_1.\overline{y}_2 + x_3.\overline{y}_3$, con $K = \mathbb{R}$ y $K = \mathbb{C}$

Ejercicio 6. Sea $A \in \mathbb{R}^{2\times 2}$. Sea $\Phi : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ definida por $\Phi(x,y) = y.A.x^t$. Probar que Φ es un producto interno sobre $\mathbb{R}^2$ si y sólo si $A = A^t$, $A_{11} > 0$ y $\det(A) > 0$.

Ejercicio 7. Determinar para qué valores de a y b en $\mathbb{R}$ es

$$\Phi(x,y) = a.x_1.y_1 + b.x_1.y_2 + b.x_2.y_1 + b.x_2.y_2 + (1+b).x_3.y_3$$

un producto interno en $\mathbb{R}^3$.

Ejercicio 8. Probar que las siguientes funciones definen productos internos sobre los espacios vectoriales considerados:

i) $\langle , \rangle : K^{n\times n} \times K^{n\times n} \to K$, $\langle A, B\rangle = tr(A.B^*)$, con $K = \mathbb{R}$ y $K = \mathbb{C}$

ii) $\langle , \rangle : C[0,1] \times C[0,1] \to \mathbb{R}$, $\langle f, g\rangle = \int_0^1 f(x).g(x)\, dx$

iii) $\langle , \rangle : K^n \times K^n \to K$, $\langle x, y\rangle = \overline{y}.\, Q^*.Q.\, x^t$, con $K = \mathbb{R}$ y $K = \mathbb{C}$
donde $Q \in K^{n\times n}$ es una matriz inversible.

iv) $\langle , \rangle_T : V \times V \to K$, $\langle x, y\rangle_T = \langle T(x), T(y)\rangle$, con $K = \mathbb{R}$ y $K = \mathbb{C}$
donde V y W son espacios vectoriales sobre K, $\langle\, ,\, \rangle$ es un producto interno sobre W y $T : V \to W$ es un monomorfismo.

Ejercicio 9. Restringir el producto interno del item ii) del ejercicio anterior a $\mathbb{R}_n[X]$ y calcular su matriz en la base $B = \{1, X, \dots, X^n\}$.

Ejercicio 10.

i) Sea $\Phi : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ definida por $\Phi(x, y) = x_1.y_1 - 2.x_1.y_2 - 2.x_2.y_1 + 6.x_2.y_2$.

  a) Probar que Φ es un producto interno.

  b) Encontrar una base de $\mathbb{R}^2$ que sea ortonormal para Φ.

ii) Encontrar una base de $\mathbb{C}^2$ que sea ortonormal para el producto interno definido en el Ejercicio 5. vi).

Ejercicio 11. Sea V un espacio vectorial de dimensión n y sea $B = \{v_1, \dots, v_n\}$ una base de V.

i) Probar que existe un único producto interno en V para el cual B resulta ortonormal.

ii) Hallarlo en los casos

  a) $V = \mathbb{R}^2$ y $B = \{(1, 1), (2, -1)\}$

  b) $V = \mathbb{C}^2$ y $B = \{(1, i), (-1, i)\}$

  c) $V = \mathbb{R}^3$ y $B = \{(1, -1, 1), (1, 1, 0), (0, 1, 1)\}$

  d) $V = \mathbb{C}^3$ y $B = \{(1, i, 1), (0, 0, 1), (0, 1, i)\}$

Ejercicio 12. Hallar el complemento ortogonal de los siguientes subespacios de V:

i) $V = \mathbb{R}^3$, $\quad S_1 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \,/\, 2.x_1 - x_2 = 0\}$ para el producto interno canónico.

ii) $V = \mathbb{R}^3$, $\quad S_2 = < (1, 2, 1) >$

  a) Para el producto interno canónico.

  b) Para el producto interno definido por $\langle x, y\rangle = x_1.y_1 + 2.x_2.y_2 + x_3.y_3 - x_1.y_2 - x_2.y_1$.

iii) $V = \mathbb{C}^3$, $\quad S_3 = < (i, 1, 1), (-1, 0, i) >$
para el producto interno $\langle , \rangle_T$ definido en el Ejercicio 8. iv) con $T : \mathbb{C}^3 \to \mathbb{C}^3$

$$T(x) = \begin{pmatrix} i & -1+i & 0 \\ 1 & i & 0 \\ 1 & i+1 & i \end{pmatrix} . x^t \quad \text{y} \quad \langle , \rangle \text{ el producto interno canónico sobre } \mathbb{C}^3.$$

iv) $V = \mathbb{C}^4$, $\quad S_4 = \left\{ (x_1, x_2, x_3, x_4) \in \mathbb{C}^4 \,/\, \begin{cases} x_1 + 2i.x_2 - x_3 + (1+i).x_4 = 0 \\ x_2 + (2-i).x_3 + x_4 = 0 \end{cases} \right\}$
para el producto interno $\langle x, y\rangle = x_1.\overline{y}_1 + 2.x_2.\overline{y}_2 + x_3.\overline{y}_3 + 3.x_4.\overline{y}_4$.

v) $V = \mathbb{R}^4$, $\quad S_5 = < (1, 1, 0, -1), (-1, 1, 1, 0), (2, -1, 1, 1) >$
para el producto interno canónico.

Ejercicio 13.

i) Hallar bases ortonormales para los subespacios del ejercicio anterior para cada uno de los productos internos considerados.

ii) Definir explícitamente las proyecciones ortogonales sobre cada uno de dichos subespacios.

iii) Hallar el punto de S_5 más cercano a $(0, 1, 1, 0)$.

Ejercicio 14. Sean S_1, S_2 y S_3 los subespacios de $\mathbb{R}^4$ definidos por

$$S_1 : \begin{cases} x_1 + x_2 - 2.x_3 = 0 \\ x_1 - x_2 = 0 \\ x_2 - x_4 = 0 \end{cases} \qquad S_2 : \begin{cases} x_1 + x_2 - x_3 + x_4 = 0 \\ 2.x_1 + 2.x_4 = 0 \end{cases} \qquad S_3 : \{2.x_1 + x_2 + 2.x_3 = 0$$

Encontrar una base ortonormal $\{v_1, v_2, v_3, v_4\}$ de $\mathbb{R}^4$ tal que $v_i \in S_i$ $(i = 1, 2, 3)$. ¿Por qué este problema tiene solución?

Ejercicio 15. Se define $\langle , \rangle : \mathbb{R}_n[X] \times \mathbb{R}_n[X] \to \mathbb{R}$ como

$$\langle f, g \rangle = \sum_{k=0}^{n} f\Big(\frac{k}{n}\Big).g\Big(\frac{k}{n}\Big).$$

i) Probar que $\langle , \rangle$ es un producto interno.

ii) Para $n = 2$, calcular $< X >^{\perp}$.

Ejercicio 16.

i) Se considera $\mathbb{C}^{n\times n}$ con el producto interno $\langle A, B \rangle = tr(A.B^*)$. Hallar el complemento ortogonal del subespacio de las matrices diagonales.

ii) Se considera $\mathbb{R}_3[X]$ con el producto interno $\langle f, g \rangle = \int_{-1}^{1} f(x).g(x)\, dx$. Aplicar el proceso de Gram-Schmidt a la base $\{1, X, X^2, X^3\}$. Hallar el complemento ortogonal del subespacio $S = < 1 >$.

iii) Se considera $C[-1, 1]$ con el producto interno $\langle f, g \rangle = \int_{-1}^{1} f(x).g(x)\, dx$. Hallar el polinomio de grado menor o igual que 3 más próximo a la función $f(x) = \operatorname{sen}(\pi x)$.

Sugerencia: Observar que basta considerar el subespacio $S = < 1, x, x^2, x^3, \operatorname{sen}(\pi x) >$.

iv) Se considera $C[0, \pi]$ con el producto interno $\langle f, g \rangle = \int_0^{\pi} f(x).g(x)\ dx$.

a) Aplicar el proceso de Gram-Schmidt a la base $B = \{1, \cos x, \operatorname{sen} x\}$.

b) Sea S el subespacio de $C[0, \pi]$ generado por B. Hallar el elemento de S más próximo a la función $f(x) = x$.

Ejercicio 17. Sea V un espacio vectorial con producto interno $\langle , \rangle$. Sea $W \subseteq V$ un subespacio de dimensión finita de V. Probar que si $x \notin W$, entonces existe $y \in V$ tal que $y \in W^{\perp}$ y $\langle x, y\rangle \neq 0$.

Ejercicio 18. *Cálculo de volúmenes.* Consideremos $\mathbb{R}^n$ con el producto interno canónico $\langle , \rangle$.

El área del paralelogramo $P(v_1, v_2)$ que definen dos vectores v_1 y v_2 linealmente independientes en $\mathbb{R}^n$ se puede calcular con la fórmula "base por altura", o sea, $\|v_1\|. \|p_{<v_1>^{\perp}}(v_2)\|$.

El volumen del paralelepípedo $P(v_1, v_2, v_3)$ que definen tres vectores v_1, v_2, v_3 linealmente independientes en $\mathbb{R}^n$ sería "área de la base por altura", o sea,

$$\|v_1\|. \|p_{<v_1>^{\perp}(v_2)}\|. \|p_{<v_1,v_2>^{\perp}}(v_3)\|.$$

Si esto se generaliza a k vectores linealmente independientes en $\mathbb{R}^n$, el volumen del paralelepípedo $P(v_1, \ldots, v_k)$ sería

$$\|v_1\|. \|p_{<v_1>^{\perp}}(v_2)\|. \|p_{<v_1,v_2>^{\perp}}(v_3)\| \ldots \|p_{<v_1,\ldots,v_{k-1}>^{\perp}}(v_k)\|.$$

Se define entonces recursivamente el volumen del paralelepípedo $P(v_1, \ldots, v_k)$ definido por los vectores linealmente independientes $v_1, \ldots, v_k \in \mathbb{R}^n$ como:

$$\begin{cases} \operatorname{vol}(P(v_1)) = \|v_1\| \\ \operatorname{vol}(P(v_1, \ldots, v_k)) = \operatorname{vol}(P(v_1, \ldots, v_{k-1})). \|p_{<v_1,\ldots,v_{k-1}>^{\perp}}(v_k)\| \quad \text{para } k \geq 2. \end{cases}$$

Vamos a probar que el volumen del paralelepípedo definido por los vectores linealmente independientes $v_1, \ldots, v_n$ en $\mathbb{R}^n$ es igual a $|\det(A)|$, donde $A \in \mathbb{R}^{n\times n}$ es la matriz cuyas columnas son los vectores $v_1, \ldots, v_n$.

i) Dados $v_1, \ldots, v_k \in \mathbb{R}^n$ se define $G(v_1, \ldots, v_k) \in \mathbb{R}^{k\times k}$ como $G(v_1, \ldots, v_k)_{ij} = \langle v_i, v_j\rangle$. Probar que:

   a) Si $v_k \in < v_1, \ldots, v_{k-1} >$, entonces $\det(G(v_1, \ldots, v_k)) = 0$.
   b) Si $v_k \in < v_1, \ldots, v_{k-1} >^{\perp}$, entonces $\det(G(v_1, \ldots, v_k)) = \det(G(v_1, \ldots, v_{k-1})). \|v_k\|^2$.
   c) $\det(G(v_1, \ldots, v_k)) = \det(G(v_1, \ldots, v_{k-1})). \|p_{<v_1,\ldots,v_{k-1}>^{\perp}}(v_k)\|^2$.

ii) Probar que, si $v_1, \ldots, v_k$ son vectores linealmente independientes,

$$(\operatorname{vol}(P(v_1, \ldots, v_k)))^2 = \det(G(v_1, \ldots, v_k)).$$

iii) Sean $v_1, \ldots, v_n \in \mathbb{R}^n$ linealmente independientes y sea $A \in \mathbb{R}^{n\times n}$ la matriz cuyas columnas son los vectores $v_1, \ldots, v_n$. Probar que $G(v_1, \ldots, v_n) = A^t. A$. Deducir que $\operatorname{vol}(P(v_1, \ldots, v_n)) = |\det(A)|$.

iv) Calcular el área del paralelogramo definido por los vectores $(2, 1)$ y $(-4, 5)$ en $\mathbb{R}^2$. Calcular el volumen del paralelepípedo definido por $(1, 1, 3)$, $(1, 2, -1)$ y $(1, 4, 1)$ en $\mathbb{R}^3$.

v) Sea $f : \mathbb{R}^n \to \mathbb{R}^n$ un isomorfismo. Si $v_1, \dots, v_n \in \mathbb{R}^n$ son linealmente independientes, probar que
$$\mathrm{vol}(P(f(v_1), \dots, f(v_n))) = |\det f|.\,\mathrm{vol}(P(v_1, \dots, v_n)).$$

Ejercicio 19. Calcular f^* para cada una de las transformaciones lineales siguientes:

i) $f : \mathbb{R}^2 \to \mathbb{R}^2, \quad f(x_1, x_2) = (3.x_1 + x_2, -x_1 + x_2)$

ii) $f : \mathbb{C}^3 \to \mathbb{C}^3, \quad f(x_1, x_2, x_3) = (2.x_1 + (1-i).x_2, x_2 + (3+2i).x_3, x_1 + i.x_2 + x_3)$

iii) $B = \{(1, 2, -1), (1, 0, 0), (0, 1, 1)\}, \quad f : \mathbb{R}^3 \to \mathbb{R}^3$ tal que
$$|f|_B = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$$

iv) $f : \mathbb{R}_2[X] \to \mathbb{R}_2[X], \quad f(p) = p' \qquad$ (donde $\langle p, q \rangle = \int_0^1 p(x).q(x)\,dx$).

v) $P \in GL(n, \mathbb{C}), \quad f : \mathbb{C}^{n \times n} \to \mathbb{C}^{n \times n},\ f(A) = P^{-1}.A.P \qquad$ (donde $\langle A, B \rangle = tr(A.B^*)$).

vi) $\mu_f : \mathbb{R}[X] \to \mathbb{R}[X], \quad \mu_f(p) = f.p \qquad$ donde $f \in \mathbb{R}[X]$ y $\langle p, q \rangle = \int_0^1 p(x).q(x)dx$

Ejercicio 20. Sea $(V, \langle, \rangle)$ un espacio vectorial de dimensión finita con producto interno. Sean f_1 y f_2 endomorfismos de V y sea k un escalar. Probar:

i) $(f_1 + f_2)^* = f_1^* + f_2^*$

ii) $(k.f_1)^* = \overline{k}.f_1^*$

iii) $(f_1 \circ f_2)^* = (f_2)^* \circ (f_1)^*$

iv) Si f_1 es un isomorfismo, entonces f_1^* es un isomorfismo y $(f_1^*)^{-1} = (f_1^{-1})^*$

v) $((f_1)^*)^* = f_1$

vi) $f_1^* \circ f_1 = 0 \Rightarrow f_1 = 0$

Ejercicio 21. Sea $(V, \langle, \rangle)$ un espacio vectorial de dimensión finita con producto interno y sea $f : V \to V$ una tranformación lineal. Probar que $\mathrm{Im}(f^*) = (\mathrm{Nu}(f))^\perp$.

Ejercicio 22. Sea $f : \mathbb{R}^3 \to \mathbb{R}^3$ la transformación lineal definida por
$$f(x, y, z) = (-x - 3y - 2z, 4x + 6y + 2z, -3x - 3y).$$
Hallar un producto interno $\langle, \rangle : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$ tal que f sea autoadjunta para $\langle, \rangle$.

Ejercicio 23. Sea $(V, \langle, \rangle)$ un espacio vectorial de dimensión finita con producto interno y sea S un subespacio de V. Probar que la proyección ortogonal $P : V \to V$ sobre S es autoadjunta. Calcular sus autovalores.

Ejercicio 24.

i) En cada uno de los siguientes casos, encontrar una matriz $O \in \mathbb{R}^{n\times n}$ ortogonal tal que $O.A.O^t$ sea diagonal:

$$A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \qquad A = \begin{pmatrix} 1 & 3 \\ 3 & -1 \end{pmatrix} \qquad A = \begin{pmatrix} 5 & 0 & -2 \\ 0 & 7 & -2 \\ -2 & -2 & 6 \end{pmatrix}$$

ii) En cada uno de los siguientes casos, encontrar una matriz $U \in \mathbb{C}^{n\times n}$ unitaria tal que $U.A.U^*$ sea diagonal:

$$A = \begin{pmatrix} 4 & 1 & i & 0 \\ 1 & 3 & 2i & 1 \\ -i & -2i & 3 & i \\ 0 & 1 & -i & 2 \end{pmatrix} \qquad A = \begin{pmatrix} 2 & -1 & -i & 0 \\ -1 & 2 & -i & 0 \\ i & i & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$$

Ejercicio 25. Encontrar una base ortonormal B de $\mathbb{R}^2$ tal que $|f|_B$ y $|g|_B$ sean diagonales si las matrices de f y de g en la base canónica son:

$$\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}$$

Sugerencia: ver el Ejercicio 31 de la Sección 6.5.

Ejercicio 26. Sea $(V, \langle , \rangle)$ un $\mathbb{C}$-espacio vectorial de dimensión finita con producto interno y sea $f : V \to V$ una transformación lineal.

Definición: Se dice que f es *normal* si $f \circ f^* = f^* \circ f$.

i) Probar que si f admite una base ortonormal de autovectores, entonces f es normal.

ii) Probar que si f es normal valen las siguientes afirmaciones:

a) $\|f(v)\| = \|f^*(v)\| \quad \forall\, v \in V$. En particular, $\mathrm{Nu}(f) = \mathrm{Nu}(f^*)$.

b) $\forall\, \lambda \in \mathbb{C},\ f - \lambda.\, id_V$ es normal.

c) Si v es un autovector de f de autovalor λ, entonces v es un autovector de f^* de autovalor $\overline{\lambda}$.

d) $E_\lambda = \{v \in V \,/\, f(v) = \lambda.v\}$ es f^*-invariante.

iii) Probar que si f es normal, entonces admite una base ortonormal de autovectores.

Sugerencia: observar que $(E_\lambda)^\perp$ es f-invariante y f^*-invariante.

iv) Deducir de lo anterior que las matrices unitarias son diagonalizables sobre $\mathbb{C}$. Encontrar un ejemplo de matriz ortogonal que *no* sea diagonalizable sobre $\mathbb{R}$.

Ejercicio 27. Hallar la matriz en la base canónica de las siguientes transformaciones ortogonales:

i) $f : \mathbb{R}^2 \to \mathbb{R}^2$, rotación de ángulo $\frac{\pi}{3}$.

ii) $f : \mathbb{R}^2 \to \mathbb{R}^2$, simetría respecto de la recta de ecuación $x_1 - x_2 = 0$.

iii) $f : \mathbb{R}^3 \to \mathbb{R}^3$, simetría respecto del plano de ecuación $x_1 + x_2 - x_3 = 0$.

iv) $f : \mathbb{R}^3 \to \mathbb{R}^3$, rotación de ángulo $\frac{\pi}{4}$ y eje $< (1, 0, 1) >$.

Ejercicio 28. Sea $f : \mathbb{R}^3 \to \mathbb{R}^3$ la transformación lineal cuya matriz en la base canónica es

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

Decidir si f es una rotación, una simetría o una composición de una rotación y una simetría. Encontrar la rotación, la simetría o ambas.

Ejercicio 29. Sea $f : \mathbb{R}^3 \to \mathbb{R}^3$ la transformación lineal cuya matriz en la base canónica es

$$\begin{pmatrix} \frac{4}{9} & \frac{8}{9} & -\frac{1}{9} \\ -\frac{4}{9} & \frac{1}{9} & -\frac{8}{9} \\ -\frac{7}{9} & \frac{4}{9} & \frac{4}{9} \end{pmatrix}.$$

i) Probar que f es una rotación.

ii) Hallar $g : \mathbb{R}^3 \to \mathbb{R}^3$ tal que $g \circ g = f$.

Ejercicio 30. Una función $f : \mathbb{R}^n \to \mathbb{R}^n$ se llama *isometría* si verifica que

$$d(x, y) = d(f(x), f(y)) \qquad \forall\, x,\, y \in \mathbb{R}^n.$$

i) Probar que si $f : \mathbb{R}^n \to \mathbb{R}^n$ es una isometría tal que $f(0) = 0$, f resulta una transformación lineal y además f es ortogonal.

ii) Deducir que $f : \mathbb{R}^n \to \mathbb{R}^n$ es una isometría si y sólo si existen $g : \mathbb{R}^n \to \mathbb{R}^n$ transformación ortogonal y $v \in \mathbb{R}^n$ tales que $f(x) = g(x) + v$, $\forall\, x \in \mathbb{R}^n$.

Capítulo 9

Variedades lineales

Al considerar los subespacios de $\mathbb{R}^2$, vimos que éstos son el conjunto $\{(0,0)\}$, el espacio $\mathbb{R}^2$ y las rectas que pasan por el origen. Ahora, en algunos contextos, por ejemplo para resolver problemas geométricos, puede ser útil trabajar con rectas que no pasan por el origen, es decir, que no son subespacios.

En este capítulo estudiaremos una noción que generaliza a la de subespacios de un K-espacio vectorial: las variedades lineales. Así como los subespacios aparecen como conjuntos de soluciones de sistemas de ecuaciones lineales homogéneas, las variedades lineales pueden verse como conjuntos de soluciones de ecuaciones lineales no homogéneas.

9.1 Nociones básicas

9.1.1 Variedades lineales

Las variedades lineales en un K-espacio vectorial V pueden definirse como sigue a partir de los subespacios de V.

Definición 9.1 Sea V un K-espacio vectorial. Una *variedad lineal* $M \subseteq V$ es un conjunto de la forma $M = \{s + p \;/\; s \in S\}$, donde S es un subespacio de V y $p \in V$.

Notación. $M = S + p$.

Podemos entonces pensar a la variedad lineal $M = S + p \subseteq V$ como el subespacio S "corrido", es decir, que en lugar de pasar por el 0 pasa por un punto p.

Observación 9.2 Sea V un K-espacio vectorial, y sea $p \in V$. Si $M = S + p \subseteq V$ (con S un subespacio de V y $p \in V$) es una variedad lineal, se le puede dar a V otra estructura de K-espacio vectorial (donde p pasa a tener el papel del 0 de V) tal que M resulta ser un subespacio de V con esa nueva estructura:

Se definen

$$\begin{array}{rcl} +_p & : & V \times V \to V\,, \quad v +_p w = v + w - p \\ \cdot_p & : & K \times V \to V\,, \quad \lambda \cdot_p v = \lambda \cdot (v - p) + p \end{array}$$

Entonces $(V, +_p, \cdot_p)$ es un K-espacio vectorial, p es el elemento neutro de $+_p$ (con lo cual cumple la función de ser el "nuevo origen" de V) y M es un subespacio de $(V, +_p, \cdot_p)$ (comparar con el Ejercicio 4 de la Sección 1.5).

La siguiente proposición muestra que una variedad lineal en un K-espacio vectorial V determina unívocamente un subespacio de V.

Proposición 9.3 *Sea V un K-espacio vectorial y sea $M \subseteq V$ una variedad lineal. Supongamos que existen $p, p' \in V$ y subespacios S y S' de V tales que $M = S + p$ y $M = S' + p'$. Entonces $S = S'$ y $p - p' \in S$.*

Demostración. En primer lugar, veamos que bajo las hipótesis de la proposición, $p - p' \in S'$ y $p - p' \in S$.

- Como $p \in M = S' + p'$, existe $s' \in S'$ tal que $p = s' + p'$. Entonces $p - p' = s' \in S'$.
- Por otro lado, se tiene que $p' \in M = S + p$, con lo que existe $s \in S$ tal que $p' = s + p$. Luego, $p - p' = -s \in S$.

Veamos ahora que $S = S'$:

Sea $s \in S$. Entonces $s + p \in M = S' + p'$, y por lo tanto, existe $s' \in S'$ tal que $s + p = s' + p'$. En consecuencia $s = s' + (p' - p) \in S'$. Luego, $S \subseteq S'$.

Análogamente se prueba la otra inclusión. □

Este resultado nos permite introducir una noción de dimensión para variedades lineales.

Definición 9.4 Sea V un K-espacio vectorial y sea $M \subseteq V$ una variedad lineal. Supongamos que $M = S + p$, donde S es un subespacio de V y $p \in V$. Entonces S se llama el *subespacio asociado a M*. Si S es de dimensión finita, se define la *dimensión de M* como $\dim(M) = \dim(S)$.

Si bien el subespacio S asociado a una variedad lineal M está unívocamente determinado por M, para cualquier punto $p \in M$ resulta que $M = S + p$:

Observación 9.5 Si V es un K-espacio vectorial, $M = S + p \subseteq V$ (con S un subespacio de V y $p \in V$) es una variedad lineal y $m \in M$, entonces $M = S + m$.

Como $m \in M$, existe $s' \in S$ tal que $m = s' + p$.

($\subseteq$) Sea $s + p \in M = S + p$. Entonces

$$s + p = s + p - m + m = s + p - (s' + p) + m = (s - s') + m \in S + m.$$

($\supseteq$) Sea $s + m \in S + m$. Entonces $s + m = s + (s' + p) = (s + s') + p \in S + p = M$.

A continuación damos algunos ejemplos de variedades lineales en distintos espacios vectoriales.

Ejemplos.

1. Los subespacios de un K-espacio vectorial V son variedades lineales: $S = S + \vec{0}$.
2. Un conjunto formado por un punto de un K-espacio vectorial es una variedad lineal.
3. Sea $A \in K^{m\times n}$ y sea $b \in K^{m\times 1}$. Entonces
$$\{x \in K^n \ / \ A.x = b\} = S_{\text{hom}} + p,$$
donde $S_{\text{hom}} \subset K^n$ es el subespacio de soluciones del sistema homogéneo asociado $A.x = 0$ y p es una solución particular del sistema $A.x = b$, es una variedad lineal.
(Observar que si el sistema es incompatible, el conjunto es vacío y entonces no es una variedad lineal.)
4. Se considera en $K[X]$ el conjunto
$$\begin{aligned} M &= \{P \in K[X] \ / \ P(0) = 1\} = \{P \in K[X] \ / \ P = \sum_{i=1}^{n} a_i X^i + 1\} \\ &= \{P \in K[X] \ / \ P(0) = 0\} + 1. \end{aligned}$$
Como $\{P \in K[X] \ / \ P(0) = 0\}$ es un subespacio de $K[X]$, entonces M es una variedad lineal.
5. En $\mathcal{C}^\infty(\mathbb{R})$, sea
$$M = \{f \in \mathcal{C}^\infty(\mathbb{R}) \ / \ f'' = \text{sen}(x)\} = \{f \in \mathcal{C}^\infty(\mathbb{R}) \ / \ (f + \text{sen}(x))'' = 0\}$$
Sea $S = \{f \in \mathcal{C}^\infty(\mathbb{R}) \ / \ f'' = 0\}$, que es un subespacio de $\mathcal{C}^\infty(\mathbb{R})$. Observamos que $f \in M$ si y sólo si $f + \text{sen}(x) \in S$. Luego, $M = S - \text{sen}(x)$ es una variedad lineal.

9.1.2 Algunas variedades lineales particulares

Las nociones de recta y plano conocidas en $\mathbb{R}^2$ o $\mathbb{R}^3$, así como también algunas propiedades básicas de estos conjuntos, se generalizan a K-espacios vectoriales arbitrarios:

Definición 9.6 Sea V un K-espacio vectorial.

i) Una *recta* en V es una variedad lineal de dimensión 1, es decir, $L = < v > + p$, con $v, p \in V$, $v \neq \vec{0}$.

ii) Un *plano* en V es una variedad lineal de dimensión 2, es decir, $\Pi = < v, w > + p$, con $\{v, w\} \subset V$ linealmente independiente

iii) Si $\dim V = n$, un *hiperplano* de V es una variedad lineal de dimensión $n - 1$.

Observación 9.7 Sea V un K-espacio vectorial y sean $p \neq q \in V$. Entonces existe una única recta $L \subseteq V$ tal que $p \in L$ y $q \in L$.

Demostración. Sea $L = < p - q > + q$. Se tiene que L es una recta, puesto que como $p \neq q$, $\dim(< p - q >) = 1$. Además:

- $q \in L$, puesto que $q = 0 + q$.
- $p \in L$, puesto que $p = 1.(p - q) + q$.

Sea L' una recta tal que $p \in L'$ y $q \in L'$. Entonces existe un subespacio S con $\dim S = 1$ tal que $L' = S + q$ y $L' = S + p$. Por lo tanto, $p - q \in S$ y, como $p - q \neq 0$ y $\dim S = 1$, resulta que $S = < p - q >$. Luego $L' = < p - q > + q = L$. □

Observación 9.8 Sea V un K-espacio vectorial. Dados $x, y, z \in V$ no alineados (es decir, que no pertenecen a una misma recta), existe un único plano Π tal que $x, y, z \in \Pi$.

Demostración. Sea $\Pi = < y - x,\ z - x > + x$. Se tiene que $x \in \Pi$, $y = (y - x) + x \in \Pi$ y $z = (z - x) + x \in \Pi$.

Como por hipótesis, x, y, z no están alineados, debe ser $\dim(< y - x,\ z - x >) = 2$ (si no, Π sería una recta con $x, y, z \in \Pi$ o un solo punto). Luego, Π es un plano con $x, y, z \in \Pi$.

Supongamos que $\Pi' = S + p$ es un plano con $x, y, z \in \Pi'$. Entonces $\Pi' = S + x$, $\Pi' = S + y$, $\Pi' = S + z$. Luego, $y - x \in S$ y $z - x \in S$. En consecuencia, $S = < y - x,\ z - x >$, de donde $\Pi' = \Pi$. □

En general, dada una cantidad finita de vectores en un K-espacio vectorial V, existe una variedad lineal que los contiene. Esto da lugar a la siguiente definición:

Definición 9.9 Sea V un K-espacio vectorial y sean $a_0, a_1, \ldots, a_n \in V$. Se llama *variedad lineal generada por* $a_0, \ldots, a_n$ a la variedad lineal $M \subseteq V$ más chica tal que $a_0 \in M, \ldots, a_n \in M$ (es decir, si $M' \subset V$ es una variedad lineal con $a_0 \in M', \ldots, a_n \in M'$, entonces $M \subset M'$).

El resultado siguiente caracteriza la variedad lineal generada por un conjunto finito de vectores de un K-espacio vectorial. En particular, establece que dados $n + 1$ vectores existe una variedad lineal de dimensión menor o igual que n que los contiene.

Proposición 9.10 *Sea V un K-espacio vectorial y sean $a_0, a_1, \ldots, a_n \in V$. Entonces, la variedad lineal generada por $a_0, a_1, \ldots, a_n$ es $M = < a_1 - a_0, \ldots, a_n - a_0 > + a_0$. Observar que* $\dim(M) \leq n$.

Demostración. Es claro que $a_i \in M$ para cada $0 \leq i \leq n$. Veamos que M es la menor variedad lineal (con respecto a la inclusión) con esta propiedad.

Supongamos que $M' = S + a_0$ verifica que $a_i \in M'$ para cada $0 \leq i \leq n$. Entonces $a_i - a_0 \in S$ para cada $1 \leq i \leq n$, de donde $< a_1 - a_0,\ a_2 - a_0, \ldots,\ a_n - a_0 > \subseteq S$, y en consecuencia $M \subseteq M'$.

Luego, la variedad lineal más chica que contiene a $a_0, \ldots, a_n$ es M. □

9.1.3 Otra forma de describir variedades lineales

En el Ejemplo 3 de la página 233 vimos que el conjunto de las soluciones de un sistema lineal compatible de ecuaciones con n incógnitas es una variedad lineal en K^n. Por otro lado, sabemos que todo subespacio de K^n es el conjunto de las soluciones de un sistema lineal homogéneo $Ax = 0$, lo que implica que una variedad lineal en K^n es el conjunto de las soluciones de un sistema lineal (no necesariamente homogéneo).

Esto dice que las variedades lineales de K^n son los conjuntos de soluciones de sistemas de ecuaciones lineales.

Mediante el uso de coordenadas, es posible dar una descripción del mismo tipo para variedades lineales en un K-espacio vectorial de dimensión finita arbitrario:

Sea V un K-espacio vectorial de dimensión n y sea B una base de V. Sea $M = S + p$ una variedad lineal en V, con S un subespacio de V y $p \in V$. Denotemos por S_B al subespacio de K^n formado por las coordenadas de los vectores de S en la base B. Entonces

$$v \in M \iff (v)_B \in S_B + (p)_B \text{ en } K^n.$$

Por lo tanto, $\{(v)_B \in K^n : v \in M\}$ es una variedad lineal en K^n y, entonces, es el conjunto de las soluciones de un sistema lineal (no homogéneo si no es un subespacio).

Ejemplo. Sea $M = \{P \in \mathbb{R}_2[X] \ / \ P(2) = 1\}$, que es una variedad lineal en $\mathbb{R}_2[X]$. Consideremos la base $B = \{1, X, X^2\}$. Entonces

$$P \in M \iff P(2) = 1 \iff (P)_B = (a, b, c) \ \text{ y } \ a + 2b + 4c = 1.$$

9.2 Intersección y suma de variedades lineales

9.2.1 Intersección de variedades lineales

A diferencia de lo que sucede para subespacios, la intersección de dos variedades lineales puede ser el conjunto vacío. La siguiente proposición muestra que si esta intersección es no vacía, entonces es una variedad lineal.

Proposición 9.11 *Sea V un K-espacio vectorial y sean M_1 y M_2 variedades lineales en V. Entonces $M_1 \cap M_2 = \emptyset$ o $M_1 \cap M_2$ es una variedad lineal.*

Demostración. Supongamos que $M_1 \cap M_2 \neq \emptyset$. Entonces existe $p \in M_1 \cap M_2$, con lo cual $M_1 = S_1 + p$ y $M_2 = S_2 + p$ con S_1 y S_2 subespacios de V.

Veamos que $M_1 \cap M_2 = (S_1 \cap S_2) + p$:

($\subseteq$) Sea $q \in M_1 \cap M_2$. Entonces existen $s_1 \in S_1$ y $s_2 \in S_2$ tales que $q = s_1 + p = s_2 + p$. En consecuencia, $s_1 = s_2$, es decir $q = s + p$ con $s \in S_1 \cap S_2$.

($\supseteq$) Si $q = s + p$ con $s \in S_1 \cap S_2$, como $s \in S_1$, se tiene que $q \in M_1$ y como $s \in S_2$, entonces $q \in M_2$. Luego $q \in M_1 \cap M_2$. □

9.2.2 Variedades lineales paralelas y alabeadas

A continuación estudiaremos las distintas situaciones que se pueden presentar para que la intersección de dos variedades lineales sea el conjunto vacío. Por ejemplo, esto sucede en $\mathbb{R}^2$ en el caso de dos rectas paralelas no coincidentes.

Definición 9.12 Sea V un K-espacio vectorial y sean M_1 y M_2 variedades lineales en V tales que $M_1 = S_1 + p_1$ y $M_2 = S_2 + p_2$. Se dice que M_1 y M_2 son *paralelas*, y se nota $M_1 \parallel M_2$, si $S_1 \subseteq S_2$ o $S_2 \subseteq S_1$.

Observación 9.13 De la definición anterior se deduce que:

- Un punto es paralelo a cualquier variedad lineal.
- Si $M_1 = S_1 + p_1$ y $M_2 = S_2 + p_2$ son variedades lineales de la misma dimensión en un K-espacio vectorial V, entonces $M_1 \parallel M_2$ si y sólo si $S_1 = S_2$.

Proposición 9.14 *Sea V un K-espacio vectorial con $\dim(V) \geq 2$. Sean L_1 y L_2 dos rectas en V. Son equivalentes:*

i) Existe un plano $\Pi \subseteq V$ tal que $L_1 \subseteq \Pi$ y $L_2 \subseteq \Pi$.

ii) $L_1 \cap L_2 \neq \emptyset$ o $L_1 \parallel L_2$.

Demostración.

ii) $\Rightarrow$ *i)* Analizaremos por separado los casos a) $L_1 \cap L_2 \neq \emptyset$ y b) $L_1 \parallel L_2$.

a) Supongamos que $L_1 \cap L_2 \neq \emptyset$. Sea $p \in L_1 \cap L_2$. Entonces $L_1 = S_1 + p = < v_1 > + p$ y $L_2 = S_2 + p = < v_2 > + p$, con $v_1, v_2 \in V$ no nulos.

   Si v_1 y v_2 son linealmente independientes, consideramos $\Pi = < v_1, v_2 > + p$, que es un plano y contiene a L_1 y a L_2.

   Si v_1 y v_2 son linealmente dependientes, sea $w \in V$ tal que $\{v_1, w\}$ es linealmente independiente. Entonces $\Pi = < v_1, w > + p$ es un plano que contiene a $L_1 = L_2$.

b) Si $L_1 \cap L_2 = \emptyset$ y $L_1 \parallel L_2$, entonces $L_1 = < v > + p_1$ y $L_2 = < v > + p_2$ para algún $v \in V$ no nulo.

   Sea $\Pi = < v, p_1 - p_2 > + p_2$. Vemos que $\dim(< v, p_1 - p_2 >) = 2$, puesto que si fuera $p_1 - p_2 = \lambda.v$, se tendría $p_1 = \lambda.v + p_2 \in L_1 \cap L_2$, contradiciendo la hipótesis. Luego, Π es un plano, y contiene a ambas rectas.

i) $\Rightarrow$ *ii)* Sea Π un plano con $L_1 \subseteq \Pi$ y $L_2 \subseteq \Pi$. Supongamos que $L_1 = < v_1 > + p_1$, $L_2 = < v_2 > + p_2$ y $\Pi = S + p_1$, con $v_1, v_2 \in V$ no nulos, S un subespacio de dimensión 2 de V, y $p_1, p_2 \in V$.

Dado que $v_1 + p_1$, p_2 y $v_2 + p_2$ pertenecen a Π, se tiene que

$$\begin{array}{rl} v_1 \in S & \text{pues } \exists\, s \in S : v_1 + p_1 = s + p_1, \text{ de donde } v_1 = s, \\ p_2 - p_1 \in S & \text{pues } \exists\, s' \in S : p_2 = s' + p_1, \text{ de donde } p_2 - p_1 = s', \\ v_2 \in S & \text{pues } \exists\, s'' \in S : v_2 + p_2 = s'' + p_1, \text{ de donde } v_2 = s'' - (p_2 - p_1), \end{array}$$

Como $\dim S = 2$, existe una combinación lineal no trivial

$$a.\, v_1 + b.\, (p_2 - p_1) + c.\, v_2 = 0.$$

Si $b = 0$, resulta que $< v_1 > = < v_2 >$ con lo que $L_1 \parallel L_2$.
Si $b \neq 0$, entonces $\dfrac{c}{b}\, v_2 + p_2 = \dfrac{-a}{b}\, v_1 + p_1$, de donde $L_1 \cap L_2 \neq \emptyset$. □

Es posible que dos variedades lineales no sean paralelas, pero tampoco tengan intersección no vacía (por ejemplo, $L_1 = < (1, 0, 0) > + (0, 0, 1)$ y $L_2 = < (0, 1, 0) > + (0, 0, 2)$ son dos rectas en $\mathbb{R}^3$ que no son paralelas ni se cortan). Esto da lugar a la siguiente definición:

Definición 9.15 Sea V un K-espacio vectorial. Dos variedades lineales M_1 y M_2 de V se dicen *alabeadas* si $M_1 \cap M_2 = \emptyset$ y $M_1 \nparallel M_2$.

Ejemplos.

1. $L_1 = < (1, 0, 0) > + (0, 0, 1)$ y $L_2 = < (0, 1, 0) > + (0, 0, 2)$ son dos rectas alabeadas en $\mathbb{R}^3$.
2. Los planos definidos en $\mathbb{R}^4$ por

   $$\Pi_1 = \{(x, y, z, w) \in \mathbb{R}^4 : x = 1,\, y = 1\} \quad \text{y} \quad \Pi_2 = \{(x, y, z, w) \in \mathbb{R}^4 : x = 2,\, z = 3\}$$

   son alabeados.

   (Se puede probar que, si V es un K-espacio vectorial tal que existen dos planos Π_1 y Π_2 en V que son alabeados, entonces $\dim(V) \geq 4$.)

En el ejemplo que sigue, estudiaremos cómo determinar si dos rectas en $\mathbb{R}^3$ se intersecan, son paralelas o son alabeadas a partir de ecuaciones que las definen.

Ejemplo. En $\mathbb{R}^3$, consideremos dos rectas L_1 y L_2 definidas por

$$L_1 : \left\{ \begin{array}{rcl} ax + by + cz & = & d \\ ex + fy + gz & = & h \end{array} \right. \qquad \text{y} \qquad L_2 : \left\{ \begin{array}{rcl} a'x + b'y + c'z & = & d' \\ e'x + f'y + g'z & = & h' \end{array} \right.$$

donde $\dim < (a, b, c), (d, e, f) > = 2$ y $\dim < (a', b', c'), (d', e', f') > = 2$. Se tiene que

$$L_1 \cap L_2 : \left\{ \begin{array}{rcl} ax + by + cz & = & d \\ ex + fy + gz & = & h \\ a'x + b'y + c'z & = & d' \\ e'x + f'y + g'z & = & h'. \end{array} \right.$$

Sean
$$A = \begin{pmatrix} a & b & c \\ e & f & g \\ a' & b' & c' \\ e' & f' & g' \end{pmatrix} \qquad \text{y} \qquad B = \left(\begin{array}{ccc|c} a & b & c & d \\ e & f & g & h \\ a' & b' & c' & d' \\ e' & f' & g' & h' \end{array}\right).$$

Observamos que:

i) Si $\mathrm{rg}(A) = 2$ y $\mathrm{rg}(B) = 2$, el sistema de ecuaciones que define $L_1 \cap L_2$ es compatible y su conjunto de soluciones es una recta. Entonces $L_1 = L_2$.

ii) Si $\mathrm{rg}(A) = 2$ y $\mathrm{rg}(B) = 3$, el sistema es incompatible, pero el conjunto de soluciones del sistema homogéneo asociado es un subespacio de dimensión 1. Luego, $L_1 \parallel L_2$ y $L_1 \cap L_2 = \emptyset$.

iii) Si $\mathrm{rg}(A) = 3$ y $\mathrm{rg}(B) = 3$, el sistema tiene solución única, de donde $L_1 \cap L_2 \neq \emptyset$ y $L_1 \nparallel L_2$.

iv) Si $\mathrm{rg}(A) = 3$ y $\mathrm{rg}(B) = 4$, entonces el sistema es incompatible, con lo que $L_1 \cap L_2 = \emptyset$ y el sistema homogéneo asociado tiene solución única, lo que implica que $L_1 \nparallel L_2$. Luego, L_1 y L_2 son alabeadas.

9.2.3 Suma de variedades lineales

Para concluir esta sección, introducimos el concepto de suma de variedades lineales. Al igual que en el caso de subespacios, dadas variedades lineales M_1 y M_2 en un K-espacio vectorial V, la idea es construir la menor variedad lineal en V que contiene a M_1 y a M_2 simultáneamente.

Definición 9.16 Sea V un K-espacio vectorial y sean M_1 y M_2 variedades lineales en V. Si $M_1 = S_1 + p_1$ y $M_2 = S_2 + p_2$ con S_1 y S_2 subespacios de V y $p_1, p_2 \in V$, se define la *variedad lineal suma* de M_1 y M_2, que notaremos $M_1 \vee M_2$, como
$$M_1 \vee M_2 = (S_1 + S_2 + < p_1 - p_2 >) + p_2.$$

Algunas observaciones respecto de la definición:

Observación 9.17 Se puede probar que:

1. Si $M_1 = S_1 + p_1 = S_1 + p'_1$ y $M_2 = S_2 + p_2 = S_2 + p'_2$, con S_i subespacio de V y $p_i, p'_i \in V$ para $i = 1, 2$, entonces $(S_1 + S_2 + < p'_1 - p'_2 >) + p'_2 = (S_1 + S_2 + < p_1 - p_2 >) + p_2$.

   Es decir, la definición de $M_1 \vee M_2$ no depende de las descripciones de M_1 y M_2.

2. $M_1 \subseteq M_1 \vee M_2$ y $M_2 \subseteq M_1 \vee M_2$.

3. Si $M_1 = S_1 + p_1$ y $M_2 = S_2 + p_2$, entonces $p_1 - p_2 \in S_1 + S_2$ si y sólo si $M_1 \cap M_2 \neq \emptyset$, en cuyo caso $M_1 \vee M_2 = (S_1 + S_2) + p_2$.

En efecto, si $p_1 - p_2 \in S_1 + S_2$ existen $s_1 \in S_1$ y $s_2 \in S_2$ tales que $p_1 - p_2 = s_1 + s_2$, y entonces $-s_1 + p_1 = s_2 + p_2 \in M_1 \cap M_2$. Recíprocamente, si $M_1 \cap M_2 \neq \emptyset$ y $q \in M_1 \cap M_2$, entonces existen $s_1 \in S_1$ y $s_2 \in S_2$ tales que $s_1 + p_1 = q = s_2 + p_2$, de donde $p_1 - p_2 = -s_1 + s_2 \in S_1 + S_2$.

El análogo del Teorema 1.43 para variedades lineales es el siguiente resultado:

Teorema 9.18 (Teorema de la dimensión para la suma de variedades lineales.) *Sea V un K espacio vectorial y sean M_1 y M_2 variedades lineales de V de dimensión finita. Entonces:*

i) Si $M_1 \cap M_2 \neq \emptyset$,

$$\dim(M_1 \vee M_2) = \dim(M_1) + \dim(M_2) - \dim(M_1 \cap M_2).$$

ii) Si $M_1 \cap M_2 = \emptyset$, $M_1 = S_1 + p_1$ y $M_2 = S_2 + p_2$, con S_1, S_2 subespacios de V y $p_1, p_2 \in V$,

$$\dim(M_1 \vee M_2) = \dim(M_1) + \dim(M_2) - \dim(S_1 \cap S_2) + 1.$$

Demostración. Si $M_1 = S_1 + p_1$ y $M_2 = S_2 + p_2$, por las definiciones de variedad suma y de dimensión de una variedad lineal, se tiene que

$$\dim(M_1 \vee M_2) = \dim(S_1 + S_2 + < p_1 - p_2 >).$$

i) Si $M_1 \cap M_2 \neq \emptyset$ y $p \in M_1 \cap M_2$, entonces $M_1 \cap M_2 = (S_1 \cap S_2) + p$. Además, por lo visto en la observación anterior, $p_1 - p_2 \in S_1 + S_2$ y entonces $M_1 \vee M_2 = (S_1 + S_2) + p_2$. Entonces, aplicando el teorema de la dimensión para subespacios, tenemos que

$$\begin{aligned} \dim(M_1 \vee M_2) &= \dim(S_1 + S_2) \\ &= \dim(S_1) + \dim(S_2) - \dim(S_1 \cap S_2) \\ &= \dim(M_1) + \dim(M_2) - \dim(M_1 \cap M_2). \end{aligned}$$

ii) Si $M_1 \cap M_2 = \emptyset$, entonces $p_1 - p_2 \notin S_1 + S_2$, con lo que

$$\begin{aligned} \dim(M_1 \vee M_2) &= \dim(S_1 + S_2 + < p_1 - p_2 >) \\ &= \dim(S_1 + S_2) + 1 \\ &= \dim(S_1) + \dim(S_2) - \dim(S_1 \cap S_2) + 1 \\ &= \dim(M_1) + \dim(M_2) - \dim(S_1 \cap S_2) + 1. \end{aligned}$$ □

9.3 Variedades lineales en espacios con producto interno

Para terminar, estudiaremos variedades lineales en espacios euclídeos (es decir, $\mathbb{R}$-espacios vectoriales con producto interno). El hecho de que en el espacio esté definido un producto interno permite extender las nociones de perpendicularidad, ángulo y distancia a variedades lineales, lo que a su vez posibilita el estudio de problemas geométricos.

9.3.1 Ortogonalidad de variedades lineales

Definición 9.19 Sea $(V, \langle, \rangle)$ un espacio euclídeo. Sean $M_1 = S_1 + p_1$ y $M_2 = S_2 + p_2$ (con S_1, S_2 subespacios de V y $p_1, p_2 \in V$) variedades lineales en V. Se dice que M_1 y M_2 son *ortogonales* si $S_1 \perp S_2$, es decir, si $\forall\, s_1 \in S_1$, $s_2 \in S_2$, se tiene que $\langle s_1, s_2 \rangle = 0$.

Si $M = S + p$ es una variedad lineal en un espacio euclídeo V de dimensión finita y $q \in V$, se puede considerar la variedad lineal ortogonal a M de dimensión máxima que pasa por q (por ejemplo, si $q = 0$, esta variedad lineal es $S^\perp$):

Definición 9.20 Sea $(V, \langle, \rangle)$ un espacio euclídeo de dimensión finita, sea $M \subseteq V$ una variedad lineal y sea $q \in V$. El *complemento ortogonal a* M *por el punto* q es la variedad lineal $M_q^\perp = S^\perp + q$, donde S es el subespacio de V asociado a M.

Escribiremos $M^\perp$ para denotar al complemento ortogonal a M por $q = 0$.

Ejemplos.

1. Sea $L = <(1,2,3)> + (1,5,4) \subseteq \mathbb{R}^3$. Hallar $L^\perp_{(1,1,2)}$.

   Por la definición, se tiene que

$$\begin{aligned} L^\perp_{(1,1,2)} &= \{(x,y,z) \in \mathbb{R}^3 \,/\, \langle (x,y,z),(1,2,3)\rangle = 0\} + (1,1,2) \\ &= \{(x,y,z) \in \mathbb{R}^3 \,/\, x + 2y + 3z = 0\} + (1,1,2) \\ &= \{(x,y,z) \in \mathbb{R}^3 \,/\, x + 2y + 3z = 9\}. \end{aligned}$$

2. Hallar $\Pi^\perp_{(1,0,1)}$ siendo $\Pi = \{(x,y,z) \in \mathbb{R}^3 \,/\, 3x - 2y + z = 7\}$.

   Se tiene que

$$\begin{aligned} \Pi &= \{(x,y,z) \in \mathbb{R}^3 \,/\, 3x - 2y + z = 0\} + (0,0,7) \\ &= \{(x,y,z) \in \mathbb{R}^3 \,/\, \langle (x,y,z),(3,-2,1)\rangle = 0\} + (0,0,7) \\ &= <(3,-2,1)>^\perp + (0,0,7). \end{aligned}$$

   Entonces $\Pi^\perp_{(1,0,1)} = \big(<(3,-2,1)>^\perp\big)^\perp + (0,0,1) = <(3,-2,1)> + (0,0,1)$.

9.3.2 Ángulo entre rectas y planos

Utilizando el concepto de ángulo entre vectores introducido en la Sección 8.1.4 podemos definir el ángulo entre dos rectas en un espacio euclídeo como sigue:

Definición 9.21 Sea $(V, \langle, \rangle)$ un espacio euclídeo y sean $L_1 = <v_1> + p_1$ y $L_2 = <v_2> + p_2$, con $v_1, v_2 \in V$ no nulos, dos rectas en V. Se define el *ángulo entre* L_1 *y* L_2 como el (único) número real comprendido entre 0 y $\frac{\pi}{2}$ que coincide con el ángulo entre v_1 y v_2 o con el ángulo entre $-v_1$ y v_2.

Esta noción nos permite a su vez definir el ángulo entre una recta y un plano, y el ángulo entre dos planos, en un espacio euclídeo de dimensión 3:

Definición 9.22 Sea $(V, \langle , \rangle)$ un espacio euclídeo con $\dim V = 3$.

- Sean L una recta y Π un plano en V. Si α es el ángulo entre las rectas L y $\Pi^{\perp}$, se define el *ángulo entre* L *y* Π como $\frac{\pi}{2} - \alpha$.
- Sean Π_1 y Π_2 planos en V. Se define el *ángulo entre* Π_1 *y* Π_2 como el ángulo entre las rectas $\Pi_1^{\perp}$ y $\Pi_2^{\perp}$.

9.3.3 Distancia de un punto a una variedad lineal

Definición 9.23 Sea V un espacio euclídeo de dimensión finita. Sea M una variedad lineal en V y sea $q \in V$. Se define la *distancia de* q *a* M como

$$d(q, M) = \inf\{d(q, z) \ / \ z \in M\}.$$

Aplicando los resultados vistos en la Sección 8.2.4, podemos probar que si $q \in V$ y M es una variedad lineal en V, existe un punto $q' \in M$ tal que $d(q, M) = d(q, q')$ y dar una fórmula para su cálculo:

Observación 9.24 Con la notación de la definición anterior, si $M = S + p$ con S un subespacio de V y $p \in V$, entonces

$$\begin{aligned} d(q, M) &= \inf\{d(q, z) \,/\, z \in M\} = \inf\{d(q, s + p) \,/\, s \in S\} \\ &= \inf\{\|q - (s + p)\| \,/\, s \in S\} = \inf\{\|q - p - s\| \,/\, s \in S\} = d(q - p,\, S) \\ &= d(q - p,\, p_S(q - p)) = \|p_{S^{\perp}}(q - p)\|, \end{aligned}$$

donde p_S y $p_{S^{\perp}}$ denotan las proyecciones ortogonales sobre S y $S^{\perp}$ respectivamente.

Notar que lo que se hizo es "trasladar el problema al 0", o sea, restarle p a todos los puntos de la variedad y a q, y entonces calcular la distancia de un vector a un subespacio.

De las igualdades de la observación anterior se deduce que $d(q, M) = \|q - (p_S(q - p) + p)\|$ y, como $p_S(q - p) + p \in M$, concluimos que éste es el punto de M más cercano a q.

Finalmente observamos que se puede probar que, si $M_q^{\perp}$ es el complemento ortogonal a M por q, entonces $M_q^{\perp} \cap M = \{p_S(q - p) + p\}$.

Ejemplo. Calcular la distancia de $(1, 2, 4)$ a $M = < (1, 1, 0), (0, 1, 0) > + (5, 1, 2)$.

Se tiene que $S = < (1, 1, 0), (0, 1, 0) >$ es el subespacio asociado a M y $S^{\perp} = < (0, 0, 1) >$. Por lo tanto, para cada $(x, y, z) \in \mathbb{R}^3$ vale $p_{S^{\perp}}(x, y, z) = (0, 0, z)$. De acuerdo a la observación anterior,

$$d((1, 2, 3), M) = \|p_{S^{\perp}}((1, 2, 4) - (5, 1, 2))\| = \|(0, 0, 2)\| = 2.$$

9.3.4 Distancia entre variedades lineales

Para terminar, estudiaremos la noción de distancia entre variedades lineales. Comenzamos con un ejemplo:

Ejemplo. Sean L_1 y L_2 las rectas definidas en $\mathbb{R}^3$ como

$$L_1 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \,/\, x_1 = 1, x_2 = 1\} \;\text{ y }\; L_2 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \,/\, x_1 = 2, x_3 = 4\}.$$

La distancia $d(L_1, L_2)$ entre L_1 y L_2 puede definirse como el ínfimo de las distancias $d(m_1, m_2)$ con $m_1 \in L_1$, $m_2 \in L_2$. Entonces

$$\begin{aligned} d(L_1, L_2) &= \inf\{d(m_1, m_2) \,/\, m_1 \in L_1, m_2 \in L_2\} \\ &= \inf\{d((1, 1, \alpha), (2, \beta, 4)) : \alpha, \beta \in \mathbb{R}\} \\ &= \inf\{\sqrt{1 + (1-\beta)^2 + (\alpha - 4)^2} : \alpha, \beta \in \mathbb{R}\} \\ &= 1. \end{aligned}$$

Más aún, el conjunto posee mínimo, el cual se alcanza para $\alpha = 4$, $\beta = 1$. En conclusión, $d(L_1, L_2)$ coincide con la distancia entre los puntos $(1, 1, 4) \in L_1$ y $(2, 1, 4) \in L_2$.

En lo que sigue, veremos que lo que sucede en el ejemplo (es decir, que la distancia entre L_1 y L_2 es la mínima de las distancias entre un punto de L_1 y un punto de L_2) se da también para variedades lineales arbitrarias en espacios euclídeos de dimensión finita. Damos entonces la siguiente definición:

Definición 9.25 Sea $(V, \langle , \rangle)$ un espacio euclídeo de dimensión finita. Sean M_1 y M_2 variedades lineales en V. Se define la *distancia entre* M_1 *y* M_2 como

$$d(M_1, M_2) = \inf\{d(m_1, m_2) \;/\; m_1 \in M_1,\; m_2 \in M_2\}.$$

Veamos que, dadas dos variedades lineales M_1 y M_2 es un espacio euclídeo V, existen $m_1 \in M_1$ y $m_2 \in M_2$ tales que $d(m_1, m_2) = d(M_1, M_2)$:

Supongamos que $M_1 = S_1 + p_1$ y $M_2 = S_2 + p_2$, con S_1 y S_2 subespacios de V y $p_1, p_2 \in V$.

Consideremos el subespacio $S = S_1 + S_2$. Como $S \oplus S^\perp = V$, existen únicos $v \in S$ y $u \in S^\perp$ tales que

$$p_1 - p_2 = v + u.$$

A continuación mostraremos que $d(M_1, M_2) = \|u\|$.

i) En primer lugar, veamos que el elemento $u \in S^\perp$ no depende de los puntos $p_1 \in M_1$ y $p_2 \in M_2$, es decir, que si $M_1 = S_1 + p_1'$ y $M_2 = S_2 + p_2'$, entonces $p_1' - p_2' = v' + u$ con $v' \in S$:

   Como $p_1' \in M_1 = S_1 + p_1$ y $p_2' \in M_2 = S_2 + p_2$, existen $s_1' \in S_1$ y $s_2' \in S_2$ tales que $p_1' = s_1' + p_1$ y $p_2' = s_2' + p_2$. Entonces

   $$p_1' - p_2' = s_1' + p_1 - s_2' - p_2 = s_1' - s_2' + (p_1 - p_2) = s_1' - s_2' + v + u = v' + u,$$

   donde $v' = s_1' - s_2' + v \in S$.

ii) Veamos ahora que $d(x,y) \geq \|u\|$ para cada $x \in M_1$, $y \in M_2$:

Por i), para $x \in M_1$, $y \in M_2$ se tiene que $x - y = v_{xy} + u$, para algún $v_{xy} \in S$. Entonces

$$\|x - y\|^2 = \|v_{xy}\|^2 + \|u\|^2 \geq \|u\|^2$$

y, en consecuencia, $d(x,y) \geq \|u\|$.

iii) Finalmente veamos que existen $m_1 \in M_1$ y $m_2 \in M_2$ tales que $d(m_1, m_2) = \|u\|$:

Sea $v \in S$ como al comienzo, es decir, tal que $p_1 - p_2 = v + u$. Como $v \in S = S_1 + S_2$, existen $s_1 \in S_1$ y $s_2 \in S_2$ tales que $v = s_1 + s_2$.

Sean $m_1 = -s_1 + p_1 \in M_1$ y $m_2 = s_2 + p_2 \in M_2$. Entonces

$$\begin{aligned} d(m_1, m_2) &= \|m_1 - m_2\| = \| - s_1 + p_1 - p_2 - s_2\| \\ &= \|(p_1 - p_2) - (s_1 + s_2)\| = \|v + u - v\| = \|u\|. \end{aligned}$$

Observamos que, como consecuencia de ii) y iii), resulta que

$$d(M_1, M_2) = \|u\| = d(m_1, m_2),$$

con $m_1 \in M_1$ y $m_2 \in M_2$ los puntos definidos en iii) y $u = p_{S^\perp}(p_1 - p_2)$, donde $p_{S^\perp}$ es la proyección ortogonal sobre $S^\perp$.

Hemos demostrado entonces el siguiente resultado:

Proposición 9.26 *Sea $(V, \langle , \rangle)$ un espacio euclídeo de dimensión finita y sean $M_1 = S_1 + p_1$ y $M_2 = S_2 + p_2$, con S_1, S_2 subespacios de V y $p_1, p_2 \in V$, variedades lineales en V. Entonces $d(M_1, M_2) = \|p_{(S_1+S_2)^\perp}(p_1 - p_2)\|$.*

Ejemplo. Hallar la distancia entre las rectas L_1 y L_2 en $\mathbb{R}^3$, siendo

$$L_1 = < (1,2,1) > + (1,7,2) \quad \text{y} \quad L_2 = < (2,1,1) > + (0,0,1).$$

Consideremos el subespacio $S = < (1,2,1), (2,1,1) >$, que es la suma de los subespacios asociados a L_1 y L_2.

Se tiene que

$$S^\perp = \left\{ (x,y,z) \in \mathbb{R}^3 : \left\{ \begin{array}{l} x + 2y + z = 0 \\ 2x + y + z = 0 \end{array} \right. \right\} = < (-1,-1,3) >.$$

Buscamos $u = p_{S^\perp}((1,7,2) - (0,0,1))$:

$$u = p_{S^\perp}(1,7,1) = \frac{\langle (1,7,1), (-1,-1,3) \rangle}{\|(-1,-1,3)\|^2} (-1,-1,3) = \Big(\frac{5}{11}, \frac{5}{11}, \frac{-15}{11}\Big).$$

En consecuencia,

$$d(L_1, L_2) = \left\| \Big(\frac{5}{11}, \frac{5}{11}, \frac{-15}{11}\Big) \right\| = \frac{5\sqrt{11}}{11}.$$

9.4 Ejercicios

Ejercicio 1. Sea V un K-espacio vectorial y sean $p,\, q \in V$. Si M es la recta que pasa por p y q, probar la validez de las siguientes afirmaciones:

i) Si $M' \subseteq V$ es una variedad lineal tal que $p,\, q \in M'$ entonces $M \subseteq M'$.

ii) $M = \{\lambda.q + \mu.p \;/\; \lambda,\, \mu \in K\,;\, \lambda + \mu = 1\}$

Ejercicio 2. Probar que cada uno de los siguientes conjuntos son variedades lineales y calcular su dimensión:

i) $M_1 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \,/\, 2.x_1 - x_3 = 1 \text{ y } x_2 + x_3 = -2\}$

ii) $M_2 = \{(1, 2, 3)\} \subseteq \mathbb{R}^3$

iii) $M_3 = \{P \in \mathbb{Q}_3[X] \,/\, P'(2) = 1\}$

iv) $M_4 = \{A \in \mathbb{C}^{2\times 2} \,/\, tr(A) = 5\}$

Ejercicio 3.

i) Sea $L \subseteq \mathbb{R}^3$ la recta que pasa por los puntos $(2, -1, 0)$ y $(1, 3, -1)$. Hallar una variedad lineal M de dimensión 2 que contenga a L. ¿Es M única?

ii) Sea $\Pi = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \,/\, 2.x_1 - x_2 + x_3 = 1\}$ y sea $L = <(0, 1, 1)> + (1, 1, 0)$. Hallar una variedad lineal $M \subseteq \mathbb{R}^3$ de dimensión 2 tal que $M \cap \Pi = L$.

Ejercicio 4. Determinar la dimensión de la variedad lineal

$$M = \big\{(x_1, x_2, x_3) \in \mathbb{R}^3 \,/\, x_1 - x_2 + 3.x_3 = 0,\ 2.x_1 + x_2 + x_3 = 1,\ -x_1 + x_2 + a.x_3 = 0\big\}$$

de acuerdo a los distintos valores de $a \in \mathbb{R}$.

Ejercicio 5. Hallar ecuaciones implícitas para las siguientes variedades lineales:

i) $M = <(1, 2, 1), (2, 0, 1)> + (1, 1, 1) \subseteq \mathbb{R}^3$.

ii) $M \subseteq \mathbb{R}^4$ la mínima variedad que contiene a $(1, 1, 2, 0)$, $(2, 1, 1, 0)$ y $(-1, 0, 4, 1)$.

Ejercicio 6. Sea $L = <(2, 1, 1)> + (0, -1, 1) \subseteq \mathbb{R}^3$.

i) Hallar un plano Π tal que $0 \in \Pi$ y $L \subseteq \Pi$.

ii) ¿Existirá un plano Π' tal que $L \subseteq \Pi'$, $0 \in \Pi'$ y $(0, 0, 1) \in \Pi'$ simultáneamente?

Ejercicio 7.

i) Encontrar en $\mathbb{R}^3$ dos rectas alabeadas que pasen por $(1, 2, 1)$ y $(2, 1, 1)$ respectivamente.

ii) Encontrar en $\mathbb{R}^4$ dos planos alabeados que pasen por $(1, 1, 1, 0)$ y $(0, 1, 1, 1)$ respectivamente.

iii) ¿Hay planos alabeados en $\mathbb{R}^3$? Más generalmente, si V es un K-espacio vectorial de dimensión n y M_1 y M_2 son variedades lineales alabeadas en V, ¿qué se puede decir de sus dimensiones?

Ejercicio 8.

i) Sea $S =< (2, -3) > \subseteq \mathbb{R}^2$. Hallar una recta $L \parallel S$ tal que $(1, -1) \in L$. Graficar.

ii) Sea $L_1 =< (2, 1, 0) > + (0, 0, 1)$. Hallar una recta $L_2 \parallel L_1$ que pase por el punto $(-1, 3, 0)$.

iii) Si L_1 y L_2 son las variedades de ii), hallar un plano $\Pi \subseteq \mathbb{R}^3$ tal que $L_1 \subseteq \Pi$ y $L_2 \subseteq \Pi$ simultáneamente. ¿Es Π único?

iv) Con las notaciones anteriores, hallar un plano $\Pi' \subseteq \mathbb{R}^3$ tal que $\Pi \cap \Pi' = L_1$.

Ejercicio 9. En cada uno de los siguientes casos, decidir si las variedades lineales M_1 y M_2 se cortan, son paralelas o alabeadas. En cada caso, hallar $M_1 \cap M_2$, $M_1 \vee M_2$ y calcular todas las dimensiones:

i) $M_1 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \,/\, x_1 + x_2 - x_3 = 1\}$
$M_2 =< (1, 0, 1) > + (0, 0, -3)$

ii) $M_1 =< (1, 2, 1, 0), (1, 0, 0, 1) > + (1, 2, 2, -1)$
$M_2 =< (1, 0, 1, 1), (2, 2, 1, 0) > + (-1, 4, 2, -3)$

iii) $M_1 = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 \,/\, x_1 - x_2 - 1 = x_3 + x_4 = 0\}$
$M_2 = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 \,/\, x_1 - x_2 = x_2 + x_3 + x_4 - 1 = 0\}$

Ejercicio 10. Sean

$$M_1 =< (1, 1, 1) > + (0, 2, 0) \;\text{ y }\; M_2 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \,/\, x_1 + x_2 - x_3 = x_1 - x_2 + x_3 = 1\}.$$

i) Hallar planos Π_1 y Π_2 de $\mathbb{R}^3$ tales que $M_1 \subseteq \Pi_1$, $M_2 \subseteq \Pi_2$ y $\Pi_1 \parallel \Pi_2$ simultáneamente.

ii) Hallar $M_1 \cap M_2$ y $M_1 \vee M_2$ y calcular sus dimensiones.

Ejercicio 11. Sean

$$\begin{aligned} L_1 &= \left\{(x_1,x_2,x_3)\in\mathbb{R}^3/x_1-3.x_3=0\,,\ x_2-x_3=-2\right\} \quad \text{y} \\ L_2 &= \left\{(x_1,x_2,x_3)\in\mathbb{R}^3/x_1-6.x_3=1\,,\ x_2+2.x_3=0\right\}. \end{aligned}$$

Hallar una recta $L\subseteq\mathbb{R}^3$ que pase por el punto $(1,0,2)$ y corte a L_1 y a L_2.

Ejercicio 12. Sean $A=(1,1,2)$ y $B=(2,0,2)$. Sea $\Pi=\{(x_1,x_2,x_3)\in\mathbb{R}^3\,/\,x_1+x_2=2\}$. Hallar $C\in\Pi$ tal que A, B y C formen un triángulo equilátero. ¿La solución es única?

Ejercicio 13. Se consideran en $\mathbb{R}^2$ las rectas dadas por las ecuaciones $L_1: x_2=0$, $L_2: x_2=\alpha$ y $L_3: x_2=\beta$, con α y β dos números no nulos y distintos entre sí. Sean L y L' dos rectas transversales a L_1, L_2 y L_3. Probar que

$$\frac{d(L_1\cap L,L_2\cap L)}{d(L_2\cap L,L_3\cap L)}=\frac{d(L_1\cap L',L_2\cap L')}{d(L_2\cap L',L_3\cap L')}.$$

Este enunciado es una versión del *Teorema de Thales*.

Ejercicio 14. Dado el triángulo PQR, se llama *mediana correspondiente al vértice* P a la recta que pasa por dicho vértice y por el punto medio del lado $\overline{QR}$.

Se considera en $\mathbb{R}^2$ el triángulo cuyos vértices son $P=(0,0)$, $Q=(c,0)$ y $R=(a,b)$.

i) Probar que sus tres medianas se cortan en un punto M.

ii) Probar que si $d(M,P)=d(M,Q)=d(M,R)$, el triángulo PQR es equilátero.

Ejercicio 15. Un *paralelogramo* es un cuadrilátero tal que sus lados opuestos son paralelos.

i) Probar que si el cuadrilátero dado en $\mathbb{R}^2$ por los puntos $(0,0)$, (a,b), (c,d) y $(e,0)$ es un paralelogramo, el punto de intersección de sus diagonales es el punto medio de cada una de ellas.

ii) Bajo las mismas hipótesis de i), probar que si las diagonales son perpendiculares, los cuatro lados son iguales.

Ejercicio 16. Sean A_1, A_2 y A_3 en $\mathbb{R}^3$ tres puntos no alineados. Probar que el conjunto

$$S=\{x\in\mathbb{R}^3\,/\,d(x,A_1)=d(x,A_2)=d(x,A_3)\}$$

es una recta ortogonal al plano que contiene a A_1, A_2 y A_3. Calcular S en el caso $A_1=(1,-1,0)$, $A_2=(0,1,1)$ y $A_3=(1,1,2)$.

Ejercicio 17. Sea $P:\mathbb{R}^3\to\mathbb{R}^3$ la proyección ortogonal, para el producto interno canónico, sobre el subespacio $S=\{(x_1,x_2,x_3)\in\mathbb{R}^3\,/\,2.x_1-x_2=0\}$.

i) Encontrar una recta $L \subset \mathbb{R}^3$ tal que $P(L) = (1, 2, 1)$. ¿Es única?

ii) Encontrar una recta $L_1 \subset \mathbb{R}^3$ tal que $P(L_1) = L_2$ siendo $L_2 : \begin{cases} 2.x_1 - x_2 = 0 \\ x_1 - x_3 = 0 \end{cases}$.

¿Es única?

Ejercicio 18. Hallar en $\mathbb{R}^n$ el complemento ortogonal a M que pasa por A, la proyección ortogonal de A sobre M y $d(A, M)$ en los siguientes casos:

i) $n = 2, \quad M : x_1 - x_2 = 2, \quad A = (2, 3)$

ii) $n = 3, \quad M : \begin{cases} 3.x_1 + x_3 = 1 \\ x_1 - x_2 = -1 \end{cases}, \quad A = (1, 0, 0)$

iii) $n = 4, \quad M : \begin{cases} x_1 - x_2 + x_3 = 1 \\ 2.x_1 - 3.x_4 = 2 \end{cases}, \quad A = (0, 2, 0, -1)$

Ejercicio 19. Dado en $\mathbb{R}^2$ el triángulo de vértices $A = (2, -3)$, $B = (8, 5)$ y $C = (14, 11)$, hallar la longitud de la altura que pasa por el vértice A.

Ejercicio 20. Se consideran en $\mathbb{R}^2$ los puntos $O = (0, 0)$, $P = (a, b)$ y $Q = (c, d)$. Dichos puntos forman un triángulo isósceles con base $\overline{PQ}$. Probar que la altura correspondiente a la base corta a ésta en su punto medio.

Ejercicio 21. Sean en $\mathbb{R}^3$ los puntos $A_1 = (1, -1, 0)$ y $A_2 = (1, 1, 1)$. Encontrar tres hiperplanos H tales que $d(A_1, H) = d(A_2, H)$.

Ejercicio 22.

i) Calcular el ángulo entre las rectas de $\mathbb{R}^2$ definidas por $L_1 : x_1 - x_2 = 1$ y $L_2 : x_1 + x_2 = 3$.

ii) Hallar una recta L_3 tal que $\text{Ang}(L_1, L_2) = \text{Ang}(L_2, L_3)$ y $L_1 \cap L_2 \in L_3$.

Ejercicio 23. Sea $L \subset \mathbb{R}^3$ la recta $L = < (1, -1, 1) > + \ (2, 1, 0)$. Encontrar un plano $\Pi \subset \mathbb{R}^3$ tal que $(2, 1, 0) \in \Pi$ y $\text{Ang}(L, \Pi) = \frac{\pi}{4}$.

Ejercicio 24. Hallar la distancia entre M_1 y M_2 en los siguientes casos:

i) $M_1 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \,/\, x_1 - 2.x_2 + x_3 = 1\}$
$M_2 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \,/\, x_1 - 2.x_2 + x_3 = 3\}$

ii) $M_1 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \,/\, x_1 + x_2 = 1 \,,\, x_1 - x_3 = 0\}$
$M_2 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \,/\, x_1 + x_2 + x_3 = 0 \,,\, x_3 = 1\}$

iii) $M_1 = < (1, -1, 0), (2, 1, 1) > + \ (1, 0, 0)$
$M_2 = \{(3, 0, 1)\}$

iv) $M_1 = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 \,/\, x_1 - x_2 + x_3 = -2 \,,\, x_2 - 2.x_4 = 2\}$
$M_2 = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 \,/\, x_1 + x_2 + x_3 = 0 \,,\, x_2 - 2.x_4 = -8 \,,\, x_1 - x_2 + x_4 = 5\}$

Ejercicio 25. Probar que si M_1 y M_2 son variedades lineales de $\mathbb{R}^n$ con $\dim M_1 \leq \dim M_2$ y $M_1 \parallel M_2$, entonces $d(M_1, M_2) = d(P, M_2)$ para todo $P \in M_1$.

Ejercicio 26. Sea en $\mathbb{R}^2$ la recta L que pasa por los puntos $(2, -1)$ y $(5, 3)$. Determinar una recta $L' \parallel L$ tal que $d(L, L') = 2$.

Ejercicio 27. Sean en $\mathbb{R}^3$ la recta $L = <(1, 1, 2)>$ y el punto $P = (1, 0, -2)$. Encontrar un plano H ortogonal a L tal que $d(P, H) = \sqrt{6}$.

Ejercicio 28. Sean en $\mathbb{R}^3$ la recta $L = <(1, 2, -2)> + (0, 2, 0)$ y el punto $P = (1, 2, 2)$. Encontrar ecuaciones implícitas de una recta L' ortogonal a L tal que $d(P, L') = 3$ y $L \cap L' = \emptyset$. ¿Es única?

Ejercicio 29. Sean

$$M_1 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \,/\, 2.x_1 - x_2 + x_3 = 1\} \text{ y } M_2 = (1, 1, 1) + <(0, 1, 1), (1, 0, -2)>.$$

Hallar un plano H tal que $M_i \parallel H$ $(i = 1, 2)$ y $d(P_1, H) = d(P_2, H)$.

Ejercicio 30. Sea $L = <(3, 0, -4)> + (1, -1, 0)$. Encontrar una recta L' alabeada con L, tal que $d(L, L') = 2$.

Ejercicio 31.

i) Construir una rotación $f : \mathbb{R}^3 \to \mathbb{R}^3$ tal que $f(M_1) = M_2$ en cada uno de los siguientes casos:

   a) $M_1 = \{(1, 2, -1)\}$, $\quad M_2 = \{(-1, 2, 1)\}$

   b) $M_1 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \,/\, x_1 - x_2 = 2,\ x_3 = 1\}$
      $M_2 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \,/\, x_1 - 2.x_2 = 1,\ 3.x_2 - x_3 = -4\}$

   c) $M_1 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \,/\, x_1 - x_2 + x_3 = 3\}$
      $M_2 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \,/\, x_1 - x_2 + x_3 = -3\}$

ii) Encontrar M_1 y M_2 variedades lineales de $\mathbb{R}^3$ de igual dimensión tales que no haya ninguna rotación $f : \mathbb{R}^3 \to \mathbb{R}^3$ que cumpla $f(M_1) = M_2$.

Ejercicio 32. Sean en $\mathbb{R}^3$ los planos Π_1 y Π_2 definidos por las ecuaciones:

$$\Pi_1 : x_2 - x_3 = 1 \qquad \text{y} \qquad \Pi_2 : x_2 + x_3 = -1.$$

Definir una transformación ortogonal $f : \mathbb{R}^3 \to \mathbb{R}^3$ tal que $f(\Pi_1) = \Pi_2$ y $f(\Pi_2) = \Pi_1$.

Ejercicio 33. Sea $k \in \mathbb{R}$ y sean Π_1 y Π_2 los planos en $\mathbb{R}^3$ definidos por

$$\Pi_1 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 / \, x_1 - x_2 + 2x_3 = k\} \quad \text{y} \quad \Pi_2 = <(1, 0, 1), (0, 1, 2)> + (1, -1, 1).$$

Determinar k para que exista una simetría $f : \mathbb{R}^3 \to \mathbb{R}^3$ tal que $f(\Pi_1) = \Pi_2$. Para ese valor de k hallar dicha simetría y calcular $f(\Pi_2)$.

Capítulo 10

Formas bilineales

Las formas bilineales son un caso particular de funciones multilineales que tienen diversas aplicaciones. En este capítulo sólo daremos una breve introducción al tema para poder demostrar una caracterización de las formas bilineales simétricas reales definidas positivas que se utiliza en el cálculo de extremos de funciones multivariadas.

10.1 Definición y ejemplos

Definición 10.1 Sea V un K-espacio vectorial. Una función $\Phi : V \times V \to K$ es una *forma bilineal* si:

i) $\Phi(v + v', w) = \Phi(v, w) + \Phi(v', w) \;\; \forall\, v, v', w \in V$.

ii) $\Phi(\lambda v, w) = \lambda.\, \Phi(v, w) \;\; \forall\, \lambda \in K, \forall\, v, w \in V$.

iii) $\Phi(v, w + w') = \Phi(v, w) + \Phi(v, w') \;\; \forall\, v, w, w' \in V$.

iv) $\Phi(v, \lambda.w) = \lambda.\, \Phi(v, w) \;\; \forall\, \lambda \in K, \forall\, v, w \in V$.

Notar que el nombre bilineal deriva del hecho que, cuando se fija cualquiera de las variables, la función resulta lineal en la otra.

Ejemplos. Se puede probar fácilmente que:

1. Los productos internos reales son formas bilineales.
2. Sea V un K espacio vectorial y sean $f_1 : V \to K$ y $f_2 : V \to K$ transformaciones lineales. Entonces $\Phi : V \times V \to K$ definida por $\Phi(v, w) = f_1(v).\, f_2(w)$ es una forma bilineal.
3. Sea $A \in K^{n\times n}$. Entonces $\Phi : K^n \times K^n \to K$, $\Phi(x, y) = x.\, A.\, y^t$, es una forma bilineal.

4. Sea $\Phi : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ definida por $\Phi(x, y) = x_1y_1 + x_2y_1 + x_1y_2 + 3x_2y_2$. Vemos que
$$\begin{aligned}\Phi(x,y) &= x_1.(y_1+y_2) + x_2.(y_1+3y_2) = (x_1,x_2).\begin{pmatrix} y_1+y_2 \\ y_1+3y_2\end{pmatrix}\\ &= (x_1,x_2)\begin{pmatrix} 1 & 1 \\ 1 & 3\end{pmatrix}\begin{pmatrix} y_1 \\ y_2\end{pmatrix},\end{aligned}$$
de donde Φ resulta ser una forma bilineal como las del ejemplo anterior.

10.2 Matriz de una forma bilineal

Al igual que las transformaciones lineales y los productos internos, fijando una base, las formas bilineales pueden caracterizarse por una matriz.

Definición 10.2 Sea V un K-espacio vectorial de dimensión finita y sea $B = \{v_1, \dots, v_n\}$ una base de V. Sea $\Phi : V \times V \to K$ una forma bilineal. Se define la *matriz de* Φ en la base B como
$$(|\Phi|_B)_{ij} = \Phi(v_i, v_j) \qquad \forall\, 1 \le i, j \le n.$$

Ejemplo. Para la forma bilineal Φ del ejemplo 4 anterior, si E es la base canónica de $\mathbb{R}^2$, se tiene
$$|\Phi|_E = \begin{pmatrix} \Phi(e_1,e_1) & \Phi(e_1,e_2) \\ \Phi(e_2,e_1) & \Phi(e_2,e_2)\end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 3\end{pmatrix}.$$

La matriz de una forma bilineal en una base sirve, como en el caso de las transformaciones lineales, para poder calcular el valor de la forma si conocemos las coordenadas de los vectores en dicha base.

Proposición 10.3 *Sea V un K-espacio vectorial de dimensión finita y sea $B = \{v_1, \dots, v_n\}$ una base de V. Sea $\Phi : V \times V \to K$ una forma bilineal. Entonces, para cada $x, y \in V$,*
$$\Phi(x, y) = (x)_B.\, |\Phi|_B.\, (y)_B^t.$$

Demostración. Sean $x, y \in V$. Supongamos que $x = \sum\limits_{i=1}^{n} \alpha_i v_i$ e $y = \sum\limits_{j=1}^{n} \beta_j v_j$. Entonces
$$\Phi(x,y) = \Phi\Big(\sum_{i=1}^n \alpha_i v_i, \sum_{j=1}^n \beta_j v_j\Big) = \sum_{i=1}^n \alpha_i \Big(\Phi(v_i, \sum_{j=1}^n \beta_j v_j)\Big) = \sum_{i=1}^n \alpha_i \Big(\sum_{j=1}^n \Phi(v_i,v_j)\beta_j\Big).$$
Por otro lado
$$\begin{aligned}(x)_B.\,|\Phi|_B.\,(y)_B^t &= (\alpha_1,\dots,\alpha_n).|\Phi|_B.\begin{pmatrix}\beta_1\\ \vdots \\ \beta_n\end{pmatrix} = \sum_{i=1}^n \alpha_i \left(|\Phi|_B \begin{pmatrix}\beta_1\\ \vdots \\ \beta_n\end{pmatrix}\right)_{i1}\\ &= \sum_{i=1}^n \alpha_i \Big(\sum_{j=1}^n (|\Phi|_B)_{ij}\beta_j\Big) = \sum_{i=1}^n \alpha_i \Big(\sum_{j=1}^n \Phi(v_i,v_j)\beta_j\Big).\end{aligned}$$

En consecuencia, $\Phi(x, y) = (x)_B . |\Phi|_B . (y)^t_B$. $\square$

Observación 10.4 Si $A \in K^{n\times n}$ es una matriz tal que $\forall\, x, y \in V$ vale $(x)_B.A.(y)^t_B = \Phi(x, y)$, entonces $A_{ij} = \Phi(v_i, v_j)$. En efecto,

$$\Phi(v_i, v_j) = (v_i)_B . A. (v_j)^t_B = e_i . A. e^t_j = A_{ij}.$$

Las matrices de cambio de bases nos permiten calcular la matriz de una forma bilineal en distintas bases.

Proposición 10.5 *Sea V un K-espacio vectorial de dimensión finita y sean B_1 y B_2 bases de V. Sea $\Phi : V \times V \to K$ una forma bilineal. Entonces*

$$C(B_2, B_1)^t . |\Phi|_{B_1} . C(B_2, B_1) = |\Phi|_{B_2}.$$

Demostración. Sea $A = C(B_2, B_1)^t.|\Phi|_{B_1}.C(B_2, B_1)$. Por la observación anterior, basta verificar que $(x)_{B_2}.A.(y)^t_{B_2} = \Phi(x, y)$ para todo par de vectores $x, y \in V$. Se tiene que

$$\begin{aligned}(x)_{B_2}.A.(y)^t_{B_2} &= (x)_{B_2}.C(B_2, B_1)^t.|\Phi|_{B_1}.C(B_2, B_1).(y)^t_{B_2} = \\ &= \Big(C(B_2, B_1).(x)^t_{B_2}\Big)^t.|\Phi|_{B_1} . (y)^t_{B_1} = (x)_{B_1}|\Phi|_{B_1}(y)^t_{B_1} = \Phi(x, y).\end{aligned}$$ $\square$

Notar que la matriz de cambio de base de la izquierda aparece transpuesta, lo que es natural por la forma de calcular la forma bilineal en un par de vectores, conocidas sus coordenadas.

Ejemplo. Sea $\Phi : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ la forma bilineal definida por

$$\Phi(x, y) = x_1y_1 + x_1y_2 + x_2y_1 + 3x_2y_2.$$

Si E es la base canónica de $\mathbb{R}^2$, vimos que $|\Phi|_E = \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix}$. Calcular $|\Phi|_B$ para $B = \{(1, 1), (1, 2)\}$.

Este cálculo puede hacerse de dos formas:

1. A partir de la definición de matriz de una forma bilineal:

$$|\Phi|_B = \begin{pmatrix} \Phi((1, 1), (1, 1)) & \Phi((1, 1), (1, 2)) \\ \Phi((1, 2), (1, 1)) & \Phi((1, 2), (1, 2)) \end{pmatrix} = \begin{pmatrix} 6 & 10 \\ 10 & 17 \end{pmatrix}.$$

2. Usando matrices de cambio de base:

$$\begin{aligned} |\Phi|_B &= C(B, E)^t . |\Phi|_E . C(B, E) \\ &= \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 6 & 10 \\ 10 & 17 \end{pmatrix}.\end{aligned}$$

10.3 Formas bilineales simétricas

10.3.1 Definiciones y propiedades básicas

Las formas bilineales que nos interesan son las que no varían cuando se cambia el orden de los vectores en los que se las calcula. Estas formas se llaman simétricas.

Definición 10.6 Sea V un K-espacio vectorial. Una forma bilineal $\Phi : V \times V \to K$ se dice *simétrica* si $\Phi(x, y) = \Phi(y, x)\ \forall\, x, y \in V$.

Ejemplo. Un producto interno real es una forma bilineal simétrica.

A las formas bilineales simétricas les corresponden matrices simétricas:

Proposición 10.7 *Sea V un K-espacio vectorial de dimensión finita y sea $\Phi : V \times V \to K$ una forma bilineal. Sea B una base de V. Entonces*

$$\Phi \text{ es simétrica} \iff |\Phi|_B \text{ es simétrica.}$$

Demostración.

($\Rightarrow$) Supongamos que Φ es una forma bilineal simétrica. Entonces

$$(|\Phi|_B)_{ij} = \Phi(v_i, v_j) = \Phi(v_j, v_i) = (|\Phi|_B)_{ji}.$$

($\Leftarrow$) Sean $x, y \in V$. Entonces

$$\begin{aligned} \Phi(x, y) &= (x)_B . |\Phi|_B . (y)_B^t = \Big((x)_B . |\Phi|_B . (y)_B^t\Big)^t \\ &= (y)_B . |\Phi|_B^t . (x)_B^t = (y)_B . |\Phi|_B . (x)_B^t = \Phi(y, x). \end{aligned}$$ □

A cada forma bilineal simétrica, se le asocia un subespacio núcleo de la siguiente forma:

Definición 10.8 Sea V un K-espacio vectorial y sea $\Phi : V \times V \to K$ una forma bilineal simétrica. Se define el *núcleo de* Φ como $\mathrm{Nu}(\Phi) = \{x \in V \ / \ \Phi(x, y) = 0\ \forall\, y \in V\}$. En el caso en que $\mathrm{Nu}(\Phi) \neq \{0\}$ se dice que Φ es una forma bilineal simétrica *degenerada*.

Ejemplo. Sea $\Phi : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ la forma bilineal simétrica definida por

$$\Phi(x, y) = (x_1, x_2) \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

Se tiene que

$$\begin{aligned} x \in \mathrm{Nu}(\Phi) &\iff (x_1 + 2x_2, 2x_1 + 4x_2) \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = 0 \quad \forall (y_1, y_2) \\ &\iff \begin{cases} x_1 + 2x_2 = 0 \\ 2x_1 + 4x_2 = 0 \end{cases} \iff x_1 + 2x_2 = 0. \end{aligned}$$

Luego, $\mathrm{Nu}(\Phi) = < (-2, 1) >$.

Observación 10.9 Sea $\Phi : V \times V \to K$ una forma bilineal simétrica. Entonces:

1. $\mathrm{Nu}(\Phi)$ es un subespacio de V.
2. Si $\dim V = n$ y B es una base de V, entonces

$$\begin{aligned}\mathrm{Nu}(\Phi) &= \{x \in V \ / \ (x)_B . |\Phi|_B . (y)_B^t = 0 \ \forall y \in V\} = \{x \in V \ / \ (x)_B . |\Phi|_B = 0\} \\ &= \{x \in V \ / \ |\Phi|_B^t . (x)_B^t = 0\} = \{x \in V \ / \ |\Phi|_B . (x)_B^t = 0\} \\ &= \{x \in V \ / \ (x)_B \in \mathrm{Nu}(|\Phi|)_B\}.\end{aligned}$$

   En consecuencia, $\dim \mathrm{Nu}(\Phi) = n - \mathrm{rg}(|\Phi|_B)$.

A partir de este resultado, se define una noción de rango de una forma bilineal simétrica:

Definición 10.10 Sea V un K-espacio vectorial de dimensión finita. Sea $\Phi : V \times V \to K$ una forma bilineal simétrica. Se define el *rango de* Φ como

$$\mathrm{rg}(\Phi) = \dim V - \dim \mathrm{Nu}(\Phi).$$

Notar que, por la observación anterior, si B es cualquier base de V, entonces $\mathrm{rg}(\Phi) = \mathrm{rg}(|\Phi|_B)$.

10.3.2 Diagonalización de formas bilineales simétricas

A continuación, se mostrará un proceso inductivo algorítmico que permite diagonalizar formas bilineales simétricas en cuerpos arbitrarios donde $2 \neq 0$.

Proposición 10.11 *Sea K un cuerpo tal que $2 \neq 0$ en K. Sea V un K-espacio vectorial de dimensión finita, y sea $\Phi : V \times V \to K$ una forma bilineal simétrica. Entonces existe una base B de V tal que $|\Phi|_B$ es diagonal.*

Demostración. Sea $B_0 = \{v_1, \dots, v_n\}$ una base de V. Supongamos que

$$|\Phi|_{B_0} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{12} & & & \\ \vdots & & M & \\ a_{1n} & & & \end{pmatrix}$$

donde $M \in K^{(n-1)\times(n-1)}$ es una matriz simétrica.

1. Si $a_{11} \neq 0$:

$$\begin{pmatrix} 1 & 0 & \dots & 0 \\ -\frac{a_{12}}{a_{11}} & 1 & & \vdots \\ \vdots & & \ddots & 0 \\ -\frac{a_{1n}}{a_{11}} & 0 & \dots & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{12} & & & \\ \vdots & & M & \\ a_{1n} & & & \end{pmatrix} \begin{pmatrix} 1 & -\frac{a_{12}}{a_{11}} & \dots & -\frac{a_{1n}}{a_{11}} \\ 0 & 1 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & 1 \end{pmatrix} =$$

$$= \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & M' & \\ 0 & & & \end{pmatrix} \qquad \text{con } M' \text{ simétrica.}$$

Ahora se sigue diagonalizando M'.

2. Si $a_{1i} = 0 \ \forall\, 1 \leq i \leq n$:

   Entonces $|\Phi|_{B_0} = \begin{pmatrix} 0 & \dots & 0 \\ \vdots & M & \\ 0 & & \end{pmatrix}$ y basta diagonalizar M.

3. Si $a_{11} = 0$ y existe i con $2 \leq i \leq n$ tal que $a_{1i} \neq 0$, pueden pasar dos cosas:

i) Si $a_{ii} \neq 0$, se multiplica a izquierda y a derecha por la matriz elemental P^{1i} y se obtiene la matriz simétrica $P^{1i}.|\Phi|_{B_0}.P^{1i}$ tal que $(P^{1i}.|\Phi|_{B_0}.P^{1i})_{11} = a_{ii}$, con lo que estamos en el primer caso.

ii) Si $a_{ii} = 0$, sea $C^{(i)}$ la matriz definida por

$$C^{(i)}_{k\ell} = \begin{cases} 1 & \text{si } k = \ell \text{ o } k = i, \ell = 1 \\ 0 & \text{si no} \end{cases}$$

Entonces $((C^{(i)})^t.|\Phi|_{B_0}.C^{(i)})_{11} = 2.a_{1i}$ y la matriz es simétrica, con lo que también estamos en el primer caso, pues $2 \neq 0$ en K. □

Ejemplos.

1. Sea $\Phi : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ la forma bilineal simétrica tal que

$$|\Phi|_E = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Hallar una base B de $\mathbb{R}^3$ tal que la matriz $|\Phi|_B$ sea diagonal.

Siguiendo el algoritmo que demuestra la proposición anterior, estamos en el caso en que $a_{11} = 0$ pero $a_{22} \neq 0$. Multiplicando a izquierda y a derecha por la matriz P^{12} obtenemos

$$P^{12}.|\Phi|_E.P^{12} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Estamos ahora en condiciones de aplicar el primer paso del algoritmo, ya que tenemos $(P^{12}.|\Phi|_E.P^{12})_{11} = 1 \neq 0$. Luego

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

Ahora, nos basta diagonalizar el bloque $\begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix}$, que nuevamente satisface las condiciones del paso 1, así que,

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

Por lo tanto, podemos resumir la operaciones efectuadas de la siguiente forma:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.P^{12}.|\Phi|_E.P^{12}.\begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} =$$

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

Por lo tanto, la base B de $\mathbb{R}^3$ que estamos buscando debe satisfacer que

$$C(B,E) = P^{12}.\begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & -1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$$

y por lo tanto, la base B resulta ser $B = \{(0,1,0),(1,-1,0),(1,-1,1)\}$.

2. Sea $\Phi : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ la forma bilineal simétrica tal que

$$|\Phi|_E = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 0 & 0 \\ 2 & 0 & 0 \end{pmatrix}.$$

Hallar una base B de $\mathbb{R}^3$ tal que la matriz $|\Phi|_B$ sea diagonal.

En este caso, $a_{11} = 0$ pero $a_{22} = 0$. Como estamos en el caso 3. ii) del algoritmo anterior, multiplicamos por las matrices $C^{(2)}$ y $(C^{(2)})^t$ a derecha y a izquiera respectivamente y obtenemos

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.|\Phi|_E.\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 2 \\ 1 & 0 & 0 \\ 2 & 0 & 0 \end{pmatrix}.$$

Luego, como estamos en el primer caso del algoritmo,

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

Diagonalizando ahora el bloque de 2×2 queda que

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

Como antes, si B es la base buscada, la matriz de cambio de base $C(B,E)$ resulta

$$C(B,E) = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -\frac{1}{2} & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -\frac{1}{2} & 0 \\ 1 & \frac{1}{2} & -2 \\ 0 & 0 & 1 \end{pmatrix},$$

con lo que $B = \{(1,1,0), (-\frac{1}{2}, \frac{1}{2}, 0), (0,-2,1)\}$.

10.4 Formas bilineales simétricas reales

10.4.1 Clasificación

En lo que sigue daremos una clasificación de formas bilineales simétricas reales que se utiliza, entre otras cosas, para la clasificación de extremos locales de funciones multivariadas:

Definición 10.12 Sea V un $\mathbb{R}$-espacio vectorial. Una forma bilineal simétrica $\Phi : V \times V \to \mathbb{R}$ se dice:

i) *Definida positiva* si $\Phi(x,x) > 0 \; \forall \, x \neq 0$.

ii) *Semidefinida positiva* si $\Phi(x,x) \geq 0 \; \forall \, x \in V$.

iii) *Definida negativa* si $\Phi(x,x) < 0 \; \forall \, x \neq 0$.

iv) *Semidefinida negativa* si $\Phi(x,x) \leq 0 \; \forall \, x \in V$.

v) *Indefinida* si no vale ninguna de las condiciones anteriores.

Ejemplos.

i) $\Phi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $\Phi(x,y) = \langle x, y \rangle$ es definida positiva.

ii) $\Phi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $\Phi(x,y) = x_1 y_1$ es semidefinida positiva.

iii) $\Phi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $\Phi(x,y) = -\langle x, y \rangle$ es definida negativa.

iv) $\Phi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $\Phi(x,y) = -x_1 y_1$ es semidefinida negativa.

v) $\Phi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ $(n \geq 2)$, $\Phi(x,y) = x_1 y_1 - x_2 y_2$ es indefinida.

Todas las formas bilineales simétricas reales pueden diagonalizarse de una forma especial:

Teorema 10.13 *Sea V un $\mathbb{R}$-espacio vectorial de dimensión n y sea $\Phi : V \times V \to \mathbb{R}$ una forma bilineal simétrica. Entonces existe una base B de V tal que $|\Phi|_B$ es diagonal con 1, -1 y 0 en la diagonal, es decir,*

$$(|\Phi|_B)_{ij} = \begin{cases} 1 & si\ 1 \le i \le r,\ j = i \\ -1 & si\ r+1 \le i \le s,\ j = i \\ 0 & en\ otro\ caso \end{cases} \qquad (*)$$

Demostración. Sea B_1 una base de V. Como Φ es simétrica, $|\Phi|_{B_1}$ es simétrica. En consecuencia, es diagonalizable. Más aún, existe una matriz ortogonal O tal que

$$O^t.\,|\Phi|_{B_1}.\,O = \begin{pmatrix} \alpha_1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & \alpha_r & & & & & & \\ & & & \beta_{r+1} & & & & & \\ & & & & \ddots & & & & \\ & & & & & \beta_s & & & \\ & & & & & & 0 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 0 \end{pmatrix} \qquad \text{con } \alpha_i > 0,\ \beta_i < 0.$$

Sea $A \in \mathbb{R}^{n\times n}$ la matriz definida por

$$A_{ij} = \begin{cases} \frac{1}{\sqrt{\alpha_i}} & \text{si } 1 \le i \le r,\ j = i \\ \frac{1}{\sqrt{-\beta_i}} & \text{si } r+1 \le i \le s,\ j = i \\ 1 & \text{si } i = j > s \\ 0 & \text{en otro caso} \end{cases}$$

Es fácil ver que $A.O^t.\,|\Phi|_{B_1}.\,O.A$ tiene la forma del enunciado del teorema.

Puesto que O y A son ambas inversibles, entonces $O.A$ es inversible. En consecuencia, existe una base B tal que $O.A = C(B, B_1)$. Además, $A.O^t = A^t.O^t = (O.A)^t = C(B, B_1)^t$.

Por lo tanto

$$|\Phi|_B = C(B, B_1)^t.\,|\Phi|_{B_1}.\,C(B, B_1) = A.O^t.\,|\Phi|_{B_1}.\,O.A$$

es de la forma $(*)$. □

La cantidad de 1, la de -1 y la de 0 que aparecen en la diagonal cuando se diagonaliza una forma bilineal simétrica real son invariantes asociados a la forma bilineal.

Teorema 10.14 *Sea $\Phi : V \times V \to \mathbb{R}$ una forma bilineal simétrica. Sean*

$$\begin{aligned} B_1 &= \{v_1, \dots, v_r, v_{r+1}, \dots, v_s, v_{s+1}, \dots, v_n\} \\ B_2 &= \{w_1, \dots, w_t, w_{t+1}, \dots, w_\ell, w_{\ell+1}, \dots, w_n\} \end{aligned}$$

bases de V tales que

$$(|\Phi|_{B_1})_{ij} = \begin{cases} 1 & si\ 1 \leq i \leq r,\ j = i \\ -1 & si\ r+1 \leq i \leq s,\ j = i \\ 0 & en\ otro\ caso \end{cases}$$

$$(|\Phi|_{B_2})_{ij} = \begin{cases} 1 & si\ 1 \leq i \leq t,\ j = i \\ -1 & si\ t+1 \leq i \leq \ell,\ j = i \\ 0 & en\ otro\ caso \end{cases}$$

Entonces $r = t$ y $s = \ell$.

Demostración. Sabemos que $\operatorname{rg}(|\Phi|_B) = \dim V - \dim \operatorname{Nu}(\Phi)$ para cualquier base B de V. Entonces $s = \operatorname{rg}(|\Phi|_{B_1}) = \operatorname{rg}(|\Phi|_{B_2}) = \ell$. Más aún, $\operatorname{Nu}(\Phi) = < v_{s+1}, \ldots, v_n > = < w_{\ell+1}, \ldots, w_n >$.

Consideremos

- V^+ un subespacio de V de dimensión máxima tal que $\Phi_{|_{V^+ \times V^+}}$ es definida positiva.
- V^- un subespacio de V de dimensión máxima tal que $\Phi_{|_{V^- \times V^-}}$ es definida negativa.

Veamos que $V^+ \oplus V^- \oplus \operatorname{Nu}(\Phi)$:

i) $V^+ \cap (V^- + \operatorname{Nu}(\Phi)) = \{0\}$:
Supongamos que $x \in V^+$ y $x = x_1 + x_2$ con $x_1 \in V^-$ y $x_2 \in \operatorname{Nu}(\Phi)$. Entonces

$$\Phi(x, x) = \Phi(x_1 + x_2, x_1 + x_2) = \Phi(x_1, x_1) + 2\,\Phi(x_1, x_2) + \Phi(x_2, x_2).$$

Como $x_1 \in V^-$, $\Phi(x_1, x_1) \leq 0$, y como $x_2 \in \operatorname{Nu}(\Phi)$, $\Phi(x_1, x_2) = \Phi(x_2, x_2) = 0$. Entonces $\Phi(x, x) \leq 0$. Pero $x \in V^+$, con lo que $\Phi(x, x) > 0$ o $x = 0$.

En consecuencia, $x = 0$.

ii) $V^- \cap (V^+ + \operatorname{Nu}(\Phi)) = \{0\}$:
Se prueba de la misma manera que el caso anterior.

iii) $\operatorname{Nu}(\Phi) \cap (V^+ + V^-) = \{0\}$:
Supongamos que $x \in \operatorname{Nu}(\Phi)$ y $x = x_1 + x_2$ con $x_1 \in V^+$, $x_2 \in V^-$. Entonces

$$\begin{aligned} 0 &= \Phi(x, x_1) = \Phi(x_1 + x_2, x_1) = \Phi(x_1, x_1) + \Phi(x_2, x_1), \\ 0 &= \Phi(x, x_2) = \Phi(x_1 + x_2, x_2) = \Phi(x_1, x_2) + \Phi(x_2, x_2), \end{aligned}$$

y como $\Phi(x_1, x_2) = \Phi(x_2, x_1)$, resulta que $\Phi(x_1, x_1) = \Phi(x_2, x_2)$.

Puesto que $x_1 \in V^+$ y $x_2 \in V^-$ se tiene que $\Phi(x_1, x_1) \geq 0$ y $\Phi(x_2, x_2) \leq 0$, con lo cual debe ser

$$\Phi(x_1, x_1) = \Phi(x_2, x_2) = 0.$$

Luego, $x_1 = x_2 = 0$ y entonces $x = 0$.

Observamos que si $S = < v_1, \ldots, v_r >$, entonces $\Phi_{|S\times S}$ es definida positiva, puesto que

$$\Phi\Big(\sum_{i=1}^{r} \alpha_i v_i, \sum_{j=1}^{r} \alpha_j v_j\Big) = \sum_{i=1}^{r} \alpha_i \sum_{j=1}^{r} \alpha_j \Phi(v_i, v_j) = \sum_{i=1}^{r} \alpha_i^2 > 0$$

si algún $\alpha_i \neq 0$. Como V^+ tiene dimensión máxima entre los subespacios en los cuales la restricción de Φ es definida positiva, $\dim V^+ \geq \dim S = r$.

De la misma manera resulta que $\dim V^- \geq \dim < v_{r+1}, \ldots, v_s > = s - r$.

En consecuencia se tiene que

$$n \geq \dim(V^+ \oplus V^- \oplus \mathrm{Nu}\Phi) = \dim V^+ + \dim V^- + \dim \mathrm{Nu}(\Phi) \geq r + (s-r) + (n-s) = n,$$

de donde se desprende que $V^+ \oplus V^- \oplus \mathrm{Nu}\Phi = V$, $\dim V^+ = r$ y $\dim V^- = s - r$.

Considerando la base B_2 se obtiene $\dim V^+ = t$ y $\dim V^- = \ell - t$.

Por lo tanto, $r = t$. □

Observación 10.15 Los subespacios V^+ y V^- que aparecen en la demostración anterior no son únicos.

Por ejemplo, para la forma bilineal $\Phi : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ tal que $|\Phi|_E = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ podría tomarse $V^+ = < (1, 0) >$ o $V^+ = < (1, 1) >$.

La clasificación de formas bilineales introducida al comienzo de esta sección puede efectuarse a partir de las cantidades de 1, de -1 y de 0 en la matriz diagonal dada por el Teorema 10.13:

Observación 10.16 Sea V es un $\mathbb{R}$-espacio vectorial de dimensión n, y sea $\Phi : V \times V \to \mathbb{R}$ una forma bilineal simétrica. Si B es una base de V tal que $|\Phi|_B$ es de la forma $(*)$, entonces:

i) Φ es definida positiva $\iff r = n$.

ii) Φ es semidefinida positiva $\iff r < n$ y $s = r$.

iii) Φ es definida negativa $\iff r = 0$ y $s = n$.

iv) Φ es semidefinida negativa $\iff r = 0$ y $s < n$.

v) Φ es indefinida $\iff 0 < r < s$.

La suma de la cantidad de 1 y la de -1 en una matriz de la forma $(*)$ es el rango de la forma bilineal Φ. Definimos a continuación otro invariante que, junto con el rango, permite clasificar formas bilineales simétricas reales:

Definición 10.17 Sea V un espacio vectorial real y sea $\Phi : V \times V \to \mathbb{R}$ una forma bilineal simétrica tal que para una base B

$$(|\Phi|_B)_{ij} = \begin{cases} 1 & \text{si } 1 \le i \le r,\ j = i \\ -1 & \text{si } r+1 \le i \le s,\ j = i \\ 0 & \text{si no} \end{cases}$$

Se define la *signatura de* Φ como la diferencia entre la cantidad de 1 y la cantidad de -1 que aparecen en $|\Phi|_B$, es decir

$$\text{signatura}(\Phi) = \dim V^+ - \dim V^- = 2r - s.$$

10.4.2 Formas bilineales definidas positivas

El hecho de que una forma bilineal simétrica real sea definida positiva se relaciona con los signos de los autovectores de su matriz en una base.

Proposición 10.18 *Sea V un $\mathbb{R}$-espacio vectorial de dimensión n y sea $\Phi : V \times V \to \mathbb{R}$ una forma bilineal simétrica. Sea B una base de V. Entonces Φ es definida positiva si y sólo si todos los autovalores de $|\Phi|_B \in \mathbb{R}^{n\times n}$ son positivos.*

Demostración. Como A es una matriz simétrica, existe una matriz ortogonal $O \in \mathbb{R}^{n\times n}$ tal que

$$O^t. A. O = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix},$$

donde $\lambda_1, \ldots, \lambda_n$ son los autovalores de A.

Sea $B' = \{v_1, \ldots, v_n\}$ una base de V tal que $O = C(B', B)$. Entonces

$$|\Phi|_{B'} = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}.$$

($\Rightarrow$) Para cada $1 \le i \le n$, $\lambda_i = \Phi(v_i, v_i)$. Si Φ es definida positiva, resulta $\lambda_i > 0 \ \forall\, 1 \le i \le n$.

($\Leftarrow$) Supongamos que $\lambda_1, \ldots, \lambda_n$ son positivos. Entonces, para cada $x \in V$, si $x = \sum\limits_{i=1}^{n} x_i v_i$, se tiene que

$$\Phi(x, x) = (x_1, \ldots, x_n). \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix} . \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \sum_{i=1}^{n} \lambda_i x_i^2 \ge 0,$$

y vale $\Phi(x, x) = 0$ si y sólo si $x_i = 0 \ \forall\, 1 \le i \le n$, es decir, si y sólo si $x = 0$.

En consecuencia, Φ es definida positiva. $\square$

Finalmente, estamos en condiciones de demostrar el criterio que permite decidir si una forma bilineal simétrica real es definida positiva mediante el cálculo de los menores principales de su matriz en una base.

Teorema 10.19 (Criterio de Sylvester) *Sea V un $\mathbb{R}$-espacio vectorial de dimensión n y sea $\Phi : V \times V \to \mathbb{R}$ una forma bilineal simétrica. Sea B una base de V. Entonces Φ es definida positiva si y sólo si todos los menores principales de $|\Phi|_B \in \mathbb{R}^{n\times n}$ son positivos.*

Demostración. Sea $B = \{v_1, \dots, v_n\}$ y denotemos por $A = (a_{ij})_{1\le i,j\le n} \in \mathbb{R}^{n\times n}$ a la matriz de Φ en la base B.

($\Rightarrow$) Por inducción en n.

Es claro que vale para $n = 1$.

Supongamos que vale para $n - 1$ y que V es un $\mathbb{R}$-espacio vectorial de dimensión n.

Como Φ es definida positiva, $a_{11} = \Phi(v_1, v_1) > 0$. Entonces, A se puede triangular, en el sentido de las formas bilineales, como sigue:

$$\begin{pmatrix} 1 & 0 & \dots & 0 \\ -\frac{a_{12}}{a_{11}} & 1 & & \vdots \\ \vdots & & \ddots & 0 \\ -\frac{a_{1n}}{a_{11}} & 0 & \dots & 1 \end{pmatrix} . A . \begin{pmatrix} 1 & -\frac{a_{12}}{a_{11}} & \dots & -\frac{a_{1n}}{a_{11}} \\ 0 & 1 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & 1 \end{pmatrix} = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & M & \\ 0 & & & \end{pmatrix}.$$

Es fácil ver que estas operaciones *no* cambian los menores principales. Por lo tanto, el i-ésimo menor principal de A se obtiene multiplicando el $(i-1)$-ésimo menor principal de M por a_{11}.

Ahora, si $S = < v_2 - \frac{a_{12}}{a_{11}}. v_1, \dots, v_n - \frac{a_{1n}}{a_{11}}. v_1 >$, entonces M es la matriz (simétrica) de la forma bilineal $\Phi_{|S\times S} : S \times S \to \mathbb{R}$ (que se obtiene al restringir Φ a $S \times S$) en la base $B_S = \{v_2 - \frac{a_{12}}{a_{11}}. v_1, \dots, v_n - \frac{a_{1n}}{a_{11}}. v_1\}$. Como $\Phi_{|S\times S}$ es definida positiva y $\dim S = n - 1$, por hipótesis inductiva, los menores principales de M son positivos.

Como además $a_{11} > 0$, resulta que todos los menores principales de A son positivos.

($\Leftarrow$) Por inducción en n.

Para $n = 1$ no hay nada que hacer.

Supongamos que $n > 1$ y que el resultado vale para $n - 1$.

El primer menor principal de A es positivo, es decir, $a_{11} > 0$. Entonces

$$\begin{pmatrix} 1 & 0 & \dots & 0 \\ -\frac{a_{12}}{a_{11}} & 1 & & \vdots \\ \vdots & & \ddots & 0 \\ -\frac{a_{1n}}{a_{11}} & 0 & \dots & 1 \end{pmatrix} . A . \begin{pmatrix} 1 & -\frac{a_{12}}{a_{11}} & \dots & -\frac{a_{1n}}{a_{11}} \\ 0 & 1 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & 1 \end{pmatrix} = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & M & \\ 0 & & & \end{pmatrix}.$$

Observar que todos los menores principales de M son positivos, puesto que al multiplicarlos por a_{11} se obtienen menores principales de A, que son positivos.

Si $S = < v_2 - \frac{a_{12}}{a_{11}}. v_1, \dots, v_n - \frac{a_{1n}}{a_{11}}. v_1 >$ y $B_S = \{v_2 - \frac{a_{12}}{a_{11}}. v_1, \dots, v_n - \frac{a_{1n}}{a_{11}}. v_1\}$, entonces $M = |\Phi_{|_{S\times S}}|_{B_S}$. Por hipótesis inductiva, $\Phi_{|_{S\times S}}$ es definida positiva. Entonces existe una base $B' = \{w_2, \dots, w_n\}$ de S tal que $|\Phi_{|_{S\times S}}|_{B'} = I_{n-1}$ (ver Observación 10.16). En consecuencia,

$$|\Phi|_{\{v_1, w_2, \dots, w_n\}} = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 \end{pmatrix} \qquad \text{con } a_{11} > 0.$$

Luego, Φ es definida positiva. □

10.5 Ejercicios

Ejercicio 1. Probar que las siguientes funciones son formas bilineales:

i) $\Phi : K^n \times K^n \to K$ definida por $\Phi(x, y) = x.A.y^t$ donde $A \in K^{n\times n}$.

ii) $\Phi : V \times V \to K$ definida por $\Phi(v, w) = f_1(v).f_2(w)$ donde V es un K-espacio vectorial y $f_1, f_2 \in V^*$.

iii) $\Phi : K^{m\times n} \times K^{m\times n} \to K$ definida por $\Phi(A, B) = \mathrm{tr}(A^t.C.B)$ donde $C \in K^{m\times m}$.

Ejercicio 2. Determinar si las siguientes funciones son o no formas bilineales. En caso afirmativo calcular su matriz en la base canónica correspondiente y determinar si la forma bilineal es simétrica:

i) $\Phi : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$, $\Phi(x, y) = 2.x_1.y_1 + 3.x_2.y_1 - x_2.y_2 + 3.x_1.y_2$

ii) $\Phi : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$, $\Phi(x, y) = -x_1.y_1 - x_2.y_1 + 2.x_2.y_2 + 2.x_1.y_2$

iii) $\Phi : \mathbb{C}^2 \times \mathbb{C}^2 \to \mathbb{C}$, $\Phi(x, y) = (1 + i).x_1.y_1 + x_2.y_1 + (1 - i).x_2.y_2 - 3.x_1.y_2$

iv) $\Phi : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$, $\Phi(x, y) = x_1^2 + x_2.y_1 + x_1.y_2 - x_2^2$

v) $\Phi : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$, $\Phi(x, y) = 2.x_1.y_1 + x_3.y_3 - x_1.y_3 - x_3.y_1$

vi) $\Phi : \mathbb{C}^3 \times \mathbb{C}^3 \to \mathbb{C}$, $\Phi(x, y) = x_1.y_1 + (2 + i).x_2.y_1 + 2.x_2.y_2 + (2 + i).x_1.y_2 + x_1.y_3 + x_3.y_1 - x_3.y_3$

vii) $\Phi : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$, $\Phi(x, y) = (3x_1 + x_2 - x_3).(4y_2 + 2y_3)$

Ejercicio 3.

i) Para las formas bilineales sobre $\mathbb{R}^3$ del ejercicio anterior, calcular su matriz en la base $\{(1,2,4),(2,-1,0),(-1,2,0)\}$.

ii) Para las formas bilineales simétricas del ejercicio anterior calcular su núcleo.

Ejercicio 4. Hallar una forma bilineal Φ simétrica en $\mathbb{R}^3$ tal que $\mathrm{Nu}(\Phi) =< (1,2,-1) >$ y $\Phi((1,1,1),(1,1,1)) < 0$. Calcular la matriz de Φ en la base canónica.

Ejercicio 5. Para cada una de las formas bilineales reales siguientes hallar una base ortonormal tal que la matriz de la forma bilineal en dicha base sea diagonal y exhibir la matriz de la forma bilineal en esta base. Calcular signatura y rango, decidir si es degenerada o no, definida (positiva o negativa), semidefinida (positiva o negativa) o indefinida.

i) $\Phi : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ tal que $|\Phi|_E = \begin{pmatrix} 5 & 0 & -2 \\ 0 & 7 & -2 \\ -2 & -2 & 6 \end{pmatrix}$

ii) $\Phi : \mathbb{R}^4 \times \mathbb{R}^4 \to \mathbb{R}$ definida por $\Phi(x,y) = 2.x_1.y_1 + 2.x_1.y_3 + 2.x_3.y_1 - x_3.y_3 - x_4.y_4$.

Ejercicio 6. Para cada una de las formas bilineales simétricas reales dadas en la base canónica por las matrices siguientes, hallar una base B tal que la matriz de la forma bilineal en dicha base sea diagonal con 1, -1 y 0 en la diagonal. Calcular signatura y rango, decidir si es degenerada o no, definida (positiva o negativa), semidefinida (positiva o negativa) o indefinida.

i) $\begin{pmatrix} 1 & -2 & 3 \\ -2 & 6 & -9 \\ 3 & -9 & 4 \end{pmatrix}$

ii) $\begin{pmatrix} 1 & 1 & -2 & -3 \\ 1 & 2 & -5 & -1 \\ -2 & -5 & 6 & 9 \\ -3 & -1 & 9 & 11 \end{pmatrix}$

Ejercicio 7. Sea $A \in \mathbb{R}^{n\times n}$ una matriz simétrica. Probar que la forma bilineal que tiene a A como matriz en la base canónica es definida negativa si y sólo si los signos de los menores principales de A van alternándose comenzando por un signo menos.

Bibliografía

[1] Kenneth Hoffman, Ray A. Kunze, *Álgebra lineal.* Prentice-Hall Hispanoamericana, México, 1973.

[2] Serge Lang, *Linear Algebra.* 3rd. ed. Undergraduate texts in mathematics, Springer, New York, 1987.

[3] Ángel Larotonda, *Álgebra lineal y geometría.* 2a. ed. Eudeba, Buenos Aires, 1977.

[4] Elon Lages Lima, *Algebra Linear.* R. de J. Instituto de Matemática Pura e Aplicada. CNPq, 1998.

[5] Seymour Lipschutz, *Álgebra lineal.* 2a. ed. McGraw-Hill, 1992.

[6] Carl Meyer, *Matrix analysis and applied linear algebra.* SIAM, Philadelphia, 2000.

[7] Ben Noble, D. Daniel James, *Algebra lineal aplicada.* Prentice-Hall Hispanoamericana, México, 1989.

[8] Gilbert Strang, *Álgebra lineal y sus aplicaciones.* Fondo Educativo Interamericano, México, 1982.

Índice Alfabético

Gabriela Jeronimo

Departamento de Matemática, Facultad de Ciencias Exactas y Naturales,
Universidad de Buenos Aires

CONICET - Argentina

jeronimo@dm.uba.ar

Juan Sabia

Departamento de Ciencias Exactas, Ciclo Básico Común,
Universidad de Buenos Aires

CONICET - Argentina

jsabia@dm.uba.ar

Susana Tesauri

Departamento de Matemática, Facultad de Ciencias Exactas y Naturales,
Universidad de Buenos Aires

stesauri@dm.uba.ar