

# Álgebra lineal

Ángel del Río Mateos

February 27, 2024

Estos son los apuntes de la asignatura de Álgebra del primer curso del Grado en Física de la Universidad de Murcia. Hay que leerlos con cuidado porque no están revisados con lo que contienen muchas erratas y seguro que algunos errores.

# Contenidos

<b>1</b>	<b>Cuerpos, números complejos y polinomios</b>	<b>5</b>
1.1	Operaciones y cuerpos . . . . .	5
1.2	El cuerpo de los números complejos . . . . .	10
1.3	Polinomios . . . . .	11
1.4	Raíces de polinomios . . . . .	15
1.5	Problemas . . . . .	22
<b>2</b>	<b>Espacios Vectoriales</b>	<b>25</b>
2.1	Definición y ejemplos de espacios vectoriales . . . . .	25
2.2	Aplicaciones lineales y matrices . . . . .	29
2.3	Subespacios . . . . .	36
2.4	Dependencia e independencia lineal . . . . .	43
2.5	Problemas . . . . .	48
<b>3</b>	<b>Gauss-Jordan, ecuaciones lineales, determinantes</b>	<b>53</b>
3.1	El Método de Gauss-Jordan . . . . .	53
3.2	Sistemas de ecuaciones lineales . . . . .	60
3.3	Espacios fila, columna y nulo de una matriz . . . . .	65
3.4	Determinantes . . . . .	72
3.5	Problemas . . . . .	88
<b>4</b>	<b>Bases y dimensión</b>	<b>97</b>
4.1	Bases y dimensión . . . . .	97
4.2	Suma e intersección de subespacios vectoriales . . . . .	106
4.3	Variedades Lineales . . . . .	114
4.4	Problemas . . . . .	119
<b>5</b>	<b>Aplicaciones Lineales</b>	<b>125</b>
5.1	Definición y Ejemplos . . . . .	125
5.2	Matriz asociada a una aplicación lineal . . . . .	129
5.3	Cambio de base . . . . .	134
5.4	Núcleo e Imagen . . . . .	139
5.5	Problemas . . . . .	142

<b>6</b>	<b>Diagonalización y formas de Jordan</b>	<b>151</b>
6.1	Endomorfismos y matrices diagonalizables . . . . .	151
6.2	¿Cuándo una matriz o un endomorfismo es diagonalizable . . . . .	160
6.3	Aplicación 1: Potencias de matrices y sucesiones recurrentes . . . . .	163
6.4	Aplicación 2: Ecuaciones diferenciales . . . . .	165
6.5	Formas de Jordan . . . . .	169
6.6	Demostración del Teorema de Jordan . . . . .	181
6.7	Aplicación 2. Ecuaciones diferenciales (continuación) . . . . .	185
6.8	Forma de Jordan real . . . . .	187
6.9	Problemas . . . . .	191
<b>7</b>	<b>Espacios euclídeos y unitarios</b>	<b>201</b>
7.1	Producto escalar y hermítico . . . . .	201
7.2	Norma y ángulos . . . . .	206
7.3	Ortogonalidad . . . . .	209
7.4	Método de Gram-Schmidt . . . . .	215
7.5	Subespacio ortogonal . . . . .	217
7.6	Aplicación: Ajuste por mínimos cuadrados . . . . .	223
7.7	Diagonalización ortogonal . . . . .	229
7.8	Producto vectorial . . . . .	231
7.9	. . . . .	233
	<b>Índice terminológico</b>	<b>239</b>

# Capítulo 1

## Cuerpos, números complejos y polinomios

### 1.1 Operaciones y cuerpos

Si  $A$  y  $B$  son dos conjuntos entonces el *producto cartesiano*  $A \times B$  de  $A$  por  $B$  es el conjunto formado por los pares ordenados  $(a, b)$  con  $a \in A$  y  $b \in B$ :

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

Más generalmente si  $A_1, \dots, A_n$  son  $n$  conjuntos entonces el producto cartesiano  $A_1 \times \dots \times A_n$  es el conjunto formado por las listas  $(a_1, \dots, a_n)$  con  $a_i \in A_i$  para cada  $i = 1, \dots, n$ .

$$A_1 \times \dots \times A_n = \{(a_1, \dots, a_n) : a_1 \in A_1, \dots, a_n \in A_n\}.$$

En el caso en que todos los  $A_i$  sean iguales, es decir  $A_i = A$  para todo  $i$ , este producto cartesiano se denota  $A^n$ , o sea

$$A^n = \{(a_1, \dots, a_n), a_1, \dots, a_n \in A\}.$$

Una *operación binaria* o simplemente operación, es una regla que asocia a dos elementos  $a$  y  $b$  un tercer elemento. Más formalmente, una operación binaria es una aplicación

$$\begin{aligned} A \times B &\rightarrow C \\ (a, b) &\mapsto a * b \end{aligned}$$

En el caso en que  $A = B = C$  decimos que se trata de una operación binaria interna en  $A$  o simplemente operación en  $A$ .

Los ejemplos más obvios de operaciones binarias se obtienen con la suma o el producto de números. Por ejemplo, la suma y el producto definen operaciones binarias internas en los cinco

conjuntos numéricos básicos:

$$\begin{aligned}\mathbb{N} &= \{0, 1, 2, \dots\}, \text{ números naturales,} \\ \mathbb{Z} &= \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}, \text{ números enteros,} \\ \mathbb{Q} &= \left\{ \frac{a}{b} : a, b \in \mathbb{Z}, b \neq 0 \right\}, \text{ números racionales,} \\ \mathbb{R}, & \quad \text{números reales,} \\ \mathbb{C} &= \{a + bi : a, b \in \mathbb{R}\}, \text{ números complejos.}\end{aligned}$$

También tenemos operaciones de suma y producto en los productos cartesianos de los conjuntos numéricos. Por ejemplo, la suma de dos listas de números de la misma longitud viene dada por

$$(a_1, \dots, a_n) + (b_1, \dots, b_n) = (a_1 + b_1, \dots, a_n + b_n).$$

Todos estos son ejemplos de operaciones internas. Ejemplos de operaciones no internas son el producto de números por listas de números:

$$\alpha(a_1, \dots, a_n) = (\alpha a_1, \dots, \alpha a_n)$$

o el producto escalar de vectores:

$$(a_1, \dots, a_n) \cdot (b_1, \dots, b_n) = a_1 b_1 + \dots + a_n b_n.$$

Otros ejemplos de operaciones son la suma y producto de polinomios o matrices.

Una operación  $*$  se dice que es

- *asociativa* si  $a * (b * c) = (a * b) * c$  para todo  $a, b, c$ ;
- *conmutativa* si  $a * b = b * a$  para todo  $a, b$ ;
- tiene elemento *neutro* si hay un elemento  $e$  tal que  $e * a = a * e = a$ .
- Si la operación  $*$  tiene neutro  $e$  entonces un *simétrico* de un elemento  $a$  respecto de  $*$  es otro elemento  $b$  tal que  $a * b = b * a = e$ . Este simétrico puede existir o no.
- Si  $\circ$  es otra operación decimos que  $\circ$  es *distributiva* respecto de  $*$  si  $a \circ (b * c) = (a \circ b) * (a \circ c)$  y  $(b * c) \circ a = (b \circ a) * (c \circ a)$  para todo  $a, b$  y  $c$ .

En realidad en las definiciones anteriores habría que ser algo más preciso para que tengan sentido las operaciones que aparecen. Es decir decimos que la operación tiene la propiedad si las operaciones que aparecen en la propiedad tienen sentido y se verifica lo que dice la propiedad.

Por ejemplo, la suma y producto en los cinco conjuntos numéricos son asociativas y conmutativas, tienen neutro, que es 0 para la suma y 1 para el producto. El simétrico respecto de la suma será el opuesto pero en  $\mathbb{N}$  el único elemento que tiene opuesto en  $\mathbb{N}$  es el 0. El simétrico respecto del producto es el inverso. En  $\mathbb{N}$  el único elemento que tiene inverso es el 1 y en  $\mathbb{Z}$  los únicos que tienen inverso son 1 y  $-1$ . Sin embargo, en  $\mathbb{Q}$ ,  $\mathbb{R}$  y  $\mathbb{C}$  todos los elementos tienen opuesto, y todos menos el cero tienen inverso.

Vamos con una de las definiciones más importantes de esta asignatura:

**Definición 1.1.1** *Un cuerpo es un conjunto  $K$  con dos operaciones internas que denotaremos  $+$  y  $\cdot$  que satisfacen las siguientes propiedades, a los que nos referiremos como axiomas de cuerpo:*

- $+$  y  $\cdot$  son **asociativas**.
- $+$  y  $\cdot$  son **conmutativas**.
- $\cdot$  es **distributiva** respecto de  $+$ .
- $+$  y  $\cdot$  tienen **neutro** que denotamos  $0$  y  $1$  respectivamente y los llamamos cero y uno de  $K$ .
- **Elemento opuesto:** Todo elemento  $a$  de  $K$  tiene simétrico respecto de  $+$  que llamamos opuesto de  $a$  y denotamos  $-a$ .
- **Elemento inverso:** Todo elemento  $a$  de  $K$  con  $a \neq 0$  tiene simétrico respecto de  $\cdot$ . Lo llamamos inverso de  $a$  y lo denotamos  $a^{-1}$ .

Los racionales  $\mathbb{Q}$ , reales  $\mathbb{R}$  y complejos  $\mathbb{C}$  son cuerpos con la suma y producto habituales. Otro cuerpo es  $\mathbb{Z}_2 = \{0, 1\}$  con la suma y el producto dados por las siguientes tablas:

$$\begin{array}{c|cc} + & 0 & 1 \\ \hline 0 & 0 & 1 \\ 1 & 1 & 0 \end{array} \qquad \begin{array}{c|cc} \cdot & 0 & 1 \\ \hline 0 & 0 & 0 \\ 1 & 0 & 1 \end{array} \qquad (1.1)$$

Observa que para hacer una suma o producto en  $\mathbb{Z}_2$  primero hacemos la suma y el producto como siempre y ponemos 0 si el resultado es par y 1 si el resultado es impar. O sea simplemente estamos modelando las conocidas reglas:

$$\begin{aligned} \text{Par} + \text{Par} &= \text{Impar} + \text{Impar} = \text{Par}, & \text{Par} + \text{Impar} &= \text{Impar}, \\ \text{Par} \cdot \text{Par} &= \text{Par} \cdot \text{Impar} = \text{Par}, & \text{Impar} \cdot \text{Impar} &= \text{Impar}. \end{aligned}$$

Podemos generalizar el cuerpo con dos elementos  $\mathbb{Z}_2$  de la siguiente manera: Sea  $n$  un número entero positivo y sea  $\mathbb{Z}_n = \{0, 1, 2, \dots, n-1\}$ . Para sumar y multiplicar en  $\mathbb{Z}_n$  primero hacemos la operación como si fueran números enteros, después dividimos por  $n$  y nos quedamos con el resto. Por ejemplo, las siguientes son operaciones en  $\mathbb{Z}_{24}$ :

$$12 + 11 = 23, \quad 20 + 8 = 4, \quad 5 \cdot 7 = 11, \quad 10 \cdot 7 = 22, \quad 3\dot{7} + 14 = 17$$

Podéis pensar en  $\mathbb{Z}_{24}$  como el mundo de las horas del día: Si ahora son las 12 dentro de 11 horas serán las 23; pero si son las 20, dentro de 8 horas serán las 4. A las doce de la noche (las 0 horas) comienzo a estudiar el examen que tengo mañana a las 11. Como tengo 5 temas y tardo 7 horas en estudiar cada tema, si no duermo ni esta noche ni mañana, habré acabado justo para entrar en el examen.

Observa que  $\mathbb{Z}_2$  es un cuerpo, pero  $\mathbb{Z}_{24}$  no lo es. Le pasa lo mismo que a  $\mathbb{Z}$  cumple todas las propiedades para ser cuerpo salvo que algunos elementos no tienen inverso. De hecho  $\mathbb{Z}_n$

es cuerpo si y solo si  $n$  es primo. Por tanto tenemos muchos cuerpos finitos:  $\mathbb{Z}_2, \mathbb{Z}_3, \mathbb{Z}_5, \mathbb{Z}_7, \dots$ . Por ejemplo, aquí tenéis la tablas de sumar y multiplicar en  $\mathbb{Z}_5$ :

+	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

·	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

Observad que en  $\mathbb{Z}_5$  se tiene que  $-1 = 4$  y  $-2 = 3$  y que  $\mathbb{Z}_5$  es un cuerpo pues todo elemento diferente de 0 tiene inverso:  $1^{-1} = 1$ ,  $2^{-1} = 3$ ,  $3^{-1} = 2$ ,  $4^{-1} = 4$ .

Lo importante que quiero que recordéis de los cuerpos es que casi todo lo que vamos a estudiar en esta asignatura vale para cualquier cuerpo y está basado en las propiedades que hemos puesto en la definición de cuerpo: Asociativa, conmutativa, distributiva, cero y opuestos, uno e inversos. Si os resulta complicado pensar en abstracto pensad en los números reales cuando estemos hablando de cuerpos. Pero tened en cuenta que las propiedades de las que hablemos servirán para todos los cuerpos y que uno de los cuerpos fundamentales en la física es el de los números complejos.

Vamos a ver ahora algunas propiedades elementales de los cuerpos:

**Proposición 1.1.2** *Si  $K$  es un cuerpo y  $a, b, c \in K$  entonces se verifican las siguientes propiedades:*

- (1)  $K$  solo tiene un cero y un uno.
- (2) Si  $a + b = a + c$  entonces  $b = c$ . En particular, si  $a + b = a$  entonces  $b = 0$  y si  $a + b = 0$  entonces  $b = -a$ .
- (3)  $-(-a) = a$ .
- (4) Si  $ab = ac$  y  $a \neq 0$  entonces  $b = c$ . En particular, si  $a \neq 0$  y  $ab = a$  entonces  $b = 1$  y si  $ab = 1$  entonces  $b = a^{-1}$ .
- (5)  $ab = 0$  si y solo si  $a = 0$  ó  $b = 0$ .
- (6)  $a(-b) = (-a)b = -(ab)$  y  $(-a)(-b) = ab$ .
- (7)  $a(b - c) = ab - ac$  (donde, por definición  $b - c = b + (-c)$ ).
- (8)  $-(a + b) = (-a) + (-b)$ .
- (9) Si  $a \neq 0$  entonces  $(a^{-1})^{-1} = a$  y  $\frac{b+c}{a} = \frac{b}{a} + \frac{c}{a}$ , donde por definición  $\frac{a}{c} = ac^{-1}$ .
- (10) Si  $ab \neq 0$  entonces  $(ab)^{-1} = a^{-1}b^{-1}$ .

**Demostración.** (1) Si  $0$  y  $0'$  fueran dos ceros entonces  $0 = 0 + 0' = 0'$ . Similarmente, si  $1$  y  $1'$  fueran dos unos entonces  $1 = 1 \cdot 1' = 1'$ . (2) Si  $a + b = a + c$  entonces  $b = 0 + b = ((-a) + a) + b = (-a) + (a + b) = (-a) + (a + c) = ((-a) + a) + c = 0 + c = c$ .

(3)  $(-a) + a = 0 = (-a) + (-(-a))$  y usando (2) deducimos que  $a = -(-a)$ .

(4) Si  $ab = ac$  y  $a \neq 0$  entonces  $b = 1b = (a^{-1}a)b = a^{-1}(ab) = a^{-1}(ac) = (a^{-1}a)c = 1c = c$ .

(5)  $0b + 1b = (0 + 1)b = 1b$  y usando (2) deducimos que  $0b = 0$ . Esto demuestra una implicación. Para demostrar la otra supongamos que  $ab = 0$  y que  $a \neq 0$ . Entonces, por lo que acabamos de ver tenemos que  $a0 = 0 = ab$ . Usando (4) deducimos que  $b = 0$ .

(6) De (5) tenemos que  $a(-b) + ab = a((-b) + b) = a0 = 0 = 0b = ((-a) + a)b = (-a)b + ab$ . De (2) se deduce que  $a(-b) = -(ab) = (-a)b$ .

(7) Usando (6) tenemos que  $a(b - c) = a(b + (-c)) = ab + a(-c) = ab + (-ac) = ab - ac$ .

(8) Aplicando (6) tenemos  $-(a + b) = (-1)(a + b) = (-1)a + (-1)b = (-a) + (-b)$ .

(9) Si  $a \neq 0$  entonces  $aa^{-1} = 1 \neq 0$ . Luego  $a^{-1} \neq 0$  (por (5)) y  $a^{-1}a = 1 = a^{-1}((a^{-1})^{-1})$ . Usando (2) deducimos que  $a = (a^{-1})^{-1}$ . Por otro lado  $\frac{a+b}{c} = (a+b)c^{-1} = ac^{-1} + bc^{-1} = \frac{a}{c} + \frac{b}{c}$ .

(10)  $(ab)(a^{-1}b^{-1}) = (aa^{-1})(bb^{-1}) = 1 \cdot 1 = 1 = (ab)(ab)^{-1}$ . Aplicando (4) deducimos que  $a^{-1}b^{-1} = (ab)^{-1}$ . ■

Los enteros satisfacen casi todas las condiciones para ser cuerpo, de hecho solo les falta una, a saber no todo elemento distinto de  $0$  tiene inverso. Lo mismo le pasa a todos los conjuntos  $\mathbb{Z}_n$ . De hecho el  $1$  y el  $-1$  son los únicos números enteros cuyos inversos son también enteros y en  $\mathbb{Z}_n$  los elementos invertibles son los número entre  $1$  y  $n$  que son coprimos con  $n$ . En particular  $\mathbb{Z}_n$  es un cuerpo si y solo si  $n$  es primo.

**Definición 1.1.3** *Un anillo conmutativo es un conjunto con dos operaciones que denotamos  $+$  y  $\cdot$  que cumple todos los axiomas de cuerpo excepto que no todos los elementos distintos de cero tienen inverso.*

*Un anillo es un conjunto con dos operaciones que cumple todos los axiomas de anillo conmutativo excepto que el producto puede no ser conmutativo.*

Como en el caso de cuerpos, al neutro de la suma de un anillo lo llamamos *cero* y lo denotamos  $0$ ; al neutro del producto lo llamamos *uno*; el simétrico de un elemento  $a$  respecto de la suma lo llamamos *opuesto* y lo denotamos  $-a$ . Un elemento  $a$  del anillo se dice que es *invertible* si tiene inverso y en tal caso lo denotamos  $a^{-1}$ .

**Ejemplos 1.1.4** Todos los cuerpos son anillos conmutativos,  $\mathbb{Z}$  y todos los  $\mathbb{Z}_n$  son anillos. Sin embargo  $\mathbb{Z}$  no es anillo y  $\mathbb{Z}_n$  es cuerpo si y solo si  $p$  es primo. Los únicos elementos invertibles de  $\mathbb{Z}$  son  $1$  y  $-1$ . Los elementos invertibles de  $\mathbb{Z}_n$  son los números enteros  $a$  que cumplen:  $1 \leq a \leq n - 1$  y  $\text{mcd}(a, n) = 1$ .

Los polinomios y las matrices nos van a servir para construir otros anillos pero eso ya lo veremos con más calma en las secciones 1.3 y 2.1.

**Observación 1.1.5** Muchas de las propiedades de la Proposición 1.1.2 se verifican también en los anillos conmutativos. Más concretamente se verifican (2), (3), (6), (7), (8). También se verifica que  $0a = 0$  en todos los anillos. Por otro lado si suponemos que  $a$  es invertible entonces se verifican las propiedades (4), (5), (9). Finalmente, si  $a$  y  $b$  son elementos invertibles

de un anillo conmutativo entonces también lo es  $ab$  y se verifica (10). ¡Atención! Si  $A$  no es conmutativo y  $a$  y  $b$  son invertibles entonces  $(ab)^{-1} = b^{-1}a^{-1}$ .

Vamos a ver un ejemplo de un anillo en el que fallan todas las propiedades que pueden fallar. Se trata de  $\mathbb{Z}_4 = \{0, 1, 2, 3\}$ . Obsérvese que en este anillo  $3 \cdot 3 = 1$ , pero  $2 \cdot 2 = 0$ . Por tanto, los elementos invertibles de  $\mathbb{Z}_4$  son 1 y 3 y tenemos  $1^{-1} = 1$  y  $3^{-1} = 3$ . Además de que  $2 \cdot 2 = 0$  y  $2 \neq 0$  también se verifica  $2 \cdot 3 = 2$  mientras que  $3 \neq 1$ .

## 1.2 El cuerpo de los números complejos

En esta sección vamos a revisar las propiedades fundamentales de los números complejos. Los números reales se identifican con los puntos de una recta. Vamos a representar los números complejos como puntos de un plano de forma que el complejo  $z = a + bi$  lo ubicamos en el punto del plano de coordenadas  $(a, b)$ . A menudo lo vemos como el vector anclado en el 0 con vértice en el punto  $(a, b)$ . Los números reales  $a$  y  $b$  se llaman respectivamente *parte real* y *parte imaginaria* de  $a + bi$  y los denotamos respectivamente  $\operatorname{Re}(a + bi)$  e  $\operatorname{Im}(a + bi)$ .

La suma y producto de números complejos viene dada por

$$\begin{aligned}(a_1 + b_1i) + (a_2 + b_2i) &= (a_1 + a_2) + (b_1 + b_2)i \\ (a_1 + b_1i)(a_2 + b_2i) &= (a_1a_2 - b_1b_2) + (a_1b_2 + a_2b_1)i.\end{aligned}$$

El *conjugado* de un número complejo  $z = a + bi$  (con  $a, b \in \mathbb{R}$ ) es

$$\bar{z} = a - bi,$$

su *módulo* es

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

y su *argumento*, denotado  $\arg(z)$  es el ángulo que forma el vector que lo representa con la semirecta real positiva ( $\alpha$  en la Figura 1.1).

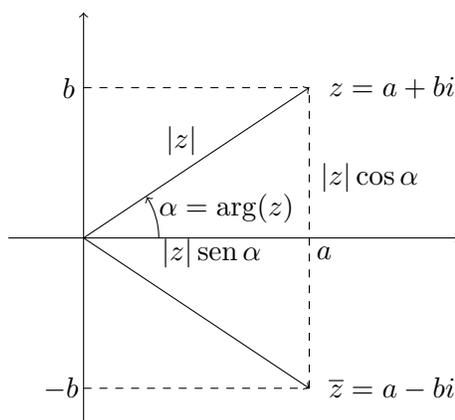


Figure 1.1:

La siguiente proposición recopila las propiedades fundamentales de estos conceptos:

**Proposición 1.2.1** Si  $y, z \in \mathbb{C}$  entonces se verifican las siguientes propiedades:

- (1)  $z \in \mathbb{R}$  si y solo si  $z = \bar{z}$ .
- (2)  $\overline{y + z} = \bar{y} + \bar{z}$  e  $\overline{yz} = \bar{y} \bar{z}$ .
- (3)  $z\bar{z} = |z|^2 \geq 0$ ,  $z + \bar{z} = 2\operatorname{Re}(z) \in \mathbb{R}$  y  $z - \bar{z} = 2\operatorname{Im}(z)i$ .
- (4)  $z = 0$  si y solo si  $|z| = 0$ .
- (5) Si  $z \neq 0$  entonces  $z^{-1} = \frac{\bar{z}}{|z|^2}$ .
- (6) Si  $\alpha = \arg(z)$  entonces  $z = |z|(\cos(\alpha) + i \operatorname{sen}(\alpha))$ .
- (7)  $|yz| = |y| |z|$ .
- (8)  $\arg(yz) = \arg(y) + \arg(z)$ .

**Demostración.** Las 6 primeras propiedades se demuestran de forma rutinaria.

(7) Usando (2) y (3) se tiene que  $|yz|^2 = yz\bar{y}\bar{z} = yz\bar{y}\bar{z} = |y|^2|z|^2 = (|y| |z|)^2$  y por tanto  $|yz| = |y||z|$ .

(8) Por otro lado si  $\alpha = \arg(y)$  y  $\beta = \arg(z)$  entonces de (6) y (7) tenemos que

$$\begin{aligned} yz &= |y|(\cos \alpha + i \operatorname{sen} \alpha)|z|(\cos \beta + i \operatorname{sen} \beta) \\ &= |yz|((\cos \alpha \cos \beta - \operatorname{sen} \alpha \operatorname{sen} \beta) + i(\cos \alpha \operatorname{sen} \beta + \operatorname{sen} \alpha \cos \beta)) \\ &= |yz|(\cos(\alpha + \beta) + i \operatorname{sen}(\alpha + \beta)) \end{aligned}$$

y usando de nuevo (6) y (7) deducimos que el argumento de  $yz$  es  $\alpha + \beta$ . ■

De la Proposición 1.2.1.(5) se tiene que si  $x$  e  $y$  son números complejos e  $y \neq 0$ . Entonces

$$\frac{x}{y} = \frac{x\bar{y}}{|y|^2}.$$

Vemos esto de otra manera que nos da una fórmula práctica para dividir números complejos:

$$\begin{aligned} \frac{a + bi}{c + di} &= \frac{(a + bi)(c - di)}{(c + di)(c - di)} = \frac{(a + bi)(c - di)}{|c + di|^2} = \frac{(ac + bd) + (bc - ad)i}{c^2 + d^2} \\ &= \frac{ac + bd}{c^2 + d^2} + \frac{bc - ad}{c^2 + d^2}i. \end{aligned}$$

## 1.3 Polinomios

En esta sección  $K$  es un cuerpo y  $X$  es una indeterminada. Un *polinomio* con coeficientes en  $K$  es una expresión de la forma

$$a_0 + a_1X + a_2X^2 + \cdots + a_nX^n$$

con  $a_0, a_1, \dots, a_n \in K$ . Vamos a escribir el polinomio de forma compacta usando el símbolo  $\sum$  de sumatorio:

$$\sum_{i=0}^n a_i X^i = a_0 + a_1 X + a_2 X^2 + \dots + a_n X^n.$$

A veces no pondremos los límites del sumatorio y solo escribiremos  $\sum_i a_i X^i$ .

El sumando  $a_i X^i$  se llama *monomio* de grado  $i$  y el número  $a_i$  se llama *coeficiente* del monomio de grado  $i$ . El coeficiente  $a_0$  se llama *término independiente* del polinomio.

Vamos a usar letras mayúsculas para denotar los polinomios e intentaremos usar la misma letra en minúscula para los coeficientes:

$$A = \sum_i a_i X^i$$

Si  $A \neq 0$ , entonces el grado de  $A$ , denotado  $\text{gr}(A)$  es el mayor de los grados de sus monomios con coeficiente diferente de 0. Por ejemplo, para el polinomio anterior

$$\text{gr}(A) = \max\{i : a_i \neq 0\}.$$

Por convenio  $\text{gr}(0) = -\infty$ . En particular, los elementos de  $K$  diferentes de 0 son los polinomios de grado 0 y los elementos de  $K$  son los polinomios de grado menor o igual que 0. Si  $A$  tiene grado  $n \neq \infty$  entonces  $a_n$  se llama *coeficiente principal* de  $A$ . Para remarcar la incógnita a veces escribimos  $A(X)$  en lugar de  $A$ .

El conjunto de los polinomios con coeficientes en  $K$  se denota  $K[X]$ .

Definimos una suma y producto de polinomios de la siguiente manera:

$$(a_0 + a_1 X + \dots + a_n X^n) + (b_0 + b_1 X + \dots + b_n X^n) = (a_0 + b_0) + (a_1 + b_1)X + \dots + (a_n + b_n)X^n.$$

$$(a_0 + a_1 X + \dots + a_n X^n)(b_0 + b_1 X + \dots + b_n X^n) = (a_0 b_0) + (a_0 b_1 + a_1 b_0)X + (a_0 b_2 + a_1 b_1 + a_2 b_0)X^2 \dots$$

Con la notación compacta:

$$\left( \sum_i a_i X^i \right) + \left( \sum_i b_i X^i \right) = \sum_i (a_i + b_i) X^i,$$

$$\left( \sum_i a_i X^i \right) \left( \sum_i b_i X^i \right) = \sum_i \left( \sum_{j+k=i} a_j b_k \right) X^i.$$

O sea la suma se obtiene sumando los coeficientes de los términos del mismo grado, y el coeficiente del término de grado  $i$  de  $(\sum_i a_i X^i)(\sum_i b_i X^i)$  se obtiene sumando los productos  $a_j b_k$  con  $j + k = i$ .

**Proposición 1.3.1** *Si  $K$  es un cuerpo entonces  $K[X]$  es un anillo conmutativo con la suma y producto de polinomios y se verifican las siguientes propiedades para  $A, B \in K[X]$ :*

- (1)  $\text{gr}(A) \leq 0 \Leftrightarrow A \in K$ .
- (2)  $\text{gr}(A + B) \leq \max(\text{gr}(A), \text{gr}(B))$ . Además  $\text{gr}(A + B) < \max(\text{gr}(A), \text{gr}(B))$  si y solo si  $\text{gr}(A) = \text{gr}(B)$  y los términos principales de  $A$  y  $B$  son opuestos.
- (3)  $\text{gr}(AB) = \text{gr}(A) + \text{gr}(B)$ .
- (4)  $A$  es invertible en  $K[X]$  si y solo si  $\text{gr}(A) = 0$ .

**Demostración.** Comprobar que  $K[X]$  es un anillo conmutativo es rutinario y el lector debería hacerlo como ejercicio. Simplemente identificaremos el cero y el uno, que son el 0 y 1 del cuerpo; y el opuesto de  $A = \sum_i a_i X^i$ , que es  $-A = \sum_i (-a_i) X^i$ . (1), (2) y (3) son también sencillos de demostrar. Para demostrar (4) observemos que si  $\text{gr}(A) = 0$  entonces  $A = a_0$  es un elemento de  $K$  diferente de 0 y tendrá un inverso  $a_0^{-1}$  en  $K$  y en  $K[X]$ . Recíprocamente si  $A$  es invertible entonces  $AB = 1$  para otro polinomio  $B$ . Entonces, de (3) se deduce que  $0 = \text{gr}(1) = \text{gr}(A) + \text{gr}(B)$  y como  $\text{gr}(A), \text{gr}(B) \geq 0$  se deduce que  $\text{gr}(A) = 0$ . ■

**Corolario 1.3.2** Sean  $A, B, C \in K[X]$ . Entonces

- (1)  $AB = 0$  si y solo si  $A = 0$  ó  $B = 0$ .
- (2) Si  $AB = AC$  y  $A \neq 0$  entonces  $B = C$ .

**Demostración.** (1) Está claro que si  $A = 0$  ó  $B = 0$  entonces  $AB = 0$ . Recíprocamente, supongamos que  $A \neq 0$ ,  $B \neq 0$ ,  $a$  y  $b$  son los coeficientes de los términos principales de  $A$  y  $B$ ,  $n = \text{gr}(A)$  y  $m = \text{gr}(B)$ . Entonces el coeficiente del monomio de grado  $m + n$  de  $AB$  es  $ab \neq 0$ . Por tanto  $AB \neq 0$ . Esto demuestra (1).

(2) Si  $AB = AC$  entonces  $A(B - C) = AB - AC = 0$ . Si  $A \neq 0$  entonces, de (1) deducimos que  $B - C = 0$ , o sea  $B = C$ . ■

**Observación 1.3.3** Se pueden definir polinomios con coeficientes en un anillo conmutativo y una suma y un producto de dichos polinomios. Esto nos proporciona otros anillos conmutativos pero eso no tiene demasiado interés para nosotros.

Sea  $A = \sum_i a_i X^i \in K[X]$  y sea  $\alpha \in K$ . Entonces definimos la *evaluación* de  $A$  en  $\alpha$  como

$$A(\alpha) = \sum_i a_i \alpha^i.$$

Decimos que  $\alpha$  es una *raíz* de  $A$  si  $A(\alpha) = 0$ .

**Proposición 1.3.4** Si  $A, B \in \mathbb{K}[X]$  y  $\alpha \in K$  entonces

$$(A + B)(\alpha) = A(\alpha) + B(\alpha) \quad \text{y} \quad (AB)(\alpha) = A(\alpha)B(\alpha).$$

En particular  $\alpha$  es raíz de  $AB$  si y solo si  $\alpha$  es raíz de  $A$  o de  $B$ .

Aunque el conjunto  $\mathbb{Z}$  de los números enteros no sea un cuerpo, pues la mayoría de los elementos no son invertibles, estamos acostumbrados a hacer divisiones de números enteros con resto. O sea dados dos números enteros  $a$  y  $b$ , con  $b \neq 0$ , existen dos números enteros  $q$  y  $r$  que cumplen

$$a = bq + r \quad \text{y} \quad 0 \leq r < |b|.$$

De hecho solo hay dos números  $q$  y  $r$  que cumplen esto y se llaman cociente y resto de la división de  $a$  entre  $b$ . Por ejemplo, el cociente y el resto de dividir 27 entre 7 son 3 y 6 pues  $27 = 3 \cdot 7 + 6$ . El cociente y el resto de dividir  $-27$  entre 7 son  $-4$  y 1 pues  $(-4)7 + 1 = -27$ . El cociente y el resto de dividir 27 entre  $-7$  son  $(-3)$  y 6 pues  $(-3)(-7) + 6 = 27$ .

Este mismo tipo de propiedad se verifica también para polinomios con coeficientes en un cuerpo:

**Proposición 1.3.5** *Dados dos polinomios  $A, B \in K[X]$  con  $B \neq 0$ , existen dos polinomios únicos  $Q$  y  $R$  tales que*

$$A = BQ + R \quad \text{y} \quad \text{gr}(R) < \text{gr}(B).$$

Los polinomios  $Q$  y  $R$  de la proposición anterior se llaman ¿cómo no? *cociente y resto* de la división de  $A$  entre  $B$ . Vamos a explicar ahora el procedimiento para calcular el cociente y el resto de la división de un polinomio  $A = \sum_{i=0}^n a_i X^i$  entre otro polinomio  $B = \sum_{i=0}^m b_i X^i$ . Suponemos que  $A$  tiene grado  $n$  y  $B$  tiene grado  $m$ , con lo que  $a_n \neq 0$  y  $b_m \neq 0$ .

Si  $n < m$  entonces el cociente es 0 y el resto es  $A$ .

Si  $n \geq m$ , entonces el primer paso del algoritmo consiste en poner  $a_n b_m^{-1} X^{n-m}$  en el cociente de la división y multiplicar esto por el dividendo como hacemos en las multiplicaciones entre enteros. Después restamos esto al dividendo, al igual que en las divisiones entre enteros:

$$\begin{array}{r} a_n X^n + a_{n-1} X^{n-1} + \dots \\ -(a_n X^n + a_n b_m^{-1} b_{m-1} X^{n-1} + \dots) \\ \hline c_{n-1} X^{n-1} + c_{n-2} X^{n-2} + \dots \end{array} \quad \begin{array}{r} | b_m X^m + b_{m-1} X^{m-1} + \dots \\ a_n b_m^{-1} X^{n-m} \\ \hline \end{array}$$

Obsérvese que el polinomio  $\sum_{i=0}^{n-1} c_i X^i$  tiene grado menor que el dividendo anterior. El procedimiento se repite con este nuevo polinomio haciendo el papel de dividendo hasta que el polinomio obtenido tenga grado menor que  $m$ . Veamos un paso más:

$$\begin{array}{r} a_n X^n + a_{n-1} X^{n-1} + \dots \\ -(a_n X^n + a_n b_m^{-1} b_{m-1} X^{n-1} + \dots) \\ \hline c_{n-1} X^{n-1} + c_{n-2} X^{n-2} + \dots \\ -(c_{n-1} X^{n-1} + c_{n-1} b_m^{-1} b_{m-1} X^{n-2} + \dots) \\ \hline d_{n-2} X^{n-2} + d_{n-3} X^{n-3} + \dots \end{array} \quad \begin{array}{r} | b_m X^m + b_{m-1} X^{m-1} + \dots \\ a_n b_m^{-1} X^{n-m} + c_{n-1} b_m^{-1} X^{n-m-1} \\ \hline \end{array}$$

Veámoslo con un ejemplo concreto:

$$\begin{array}{r} X^4 + 3X^3 - X + 2 \\ -(X^4 - \frac{1}{2}X) \\ \hline 3X^3 - \frac{1}{2}X + 2 \\ -(3X^3 - \frac{3}{2}) \\ \hline -\frac{1}{2}X + \frac{7}{2} \end{array} \quad \begin{array}{r} | 2X^3 - 1 \\ \frac{1}{2}X + \frac{3}{2} \\ \hline \end{array}$$

Por tanto el cociente de dividir  $A = X^4 + 3X^2 - X + 2$  entre  $B = 2X^3 - 1$  es  $Q = \frac{1}{2}X + \frac{3}{2} = \frac{1}{2}(X + 3)$  y el resto es  $R = -\frac{1}{2}X + \frac{7}{2} = \frac{1}{2}(7 - X)$ . Esto lo podemos comprobar calculando

$$\begin{aligned} QB + R &= \frac{1}{2}(X + 3)(2X^3 - 1) + \frac{1}{2}(7 - X) \\ &= \frac{1}{2}(2X^4 + 6X^3 - X - 3 + 7 - X) = X^4 + 3X^3 - X + 2 = A \end{aligned}$$

## 1.4 Raíces de polinomios

La Regla de Ruffini proporciona un método rápido para calcular el cociente y el resto de la división de un polinomio entre  $X - \alpha$ . La siguiente proposición muestra cómo funciona:

**Proposición 1.4.1 (Regla de Ruffini)** Sean  $A = a_0 + a_1X + a_2X^2 + \dots + a_nX^n \in K[X]$  con  $a_n \neq 0$  y  $\alpha \in K$ . Entonces los coeficientes del cociente  $Q = q_0 + q_1X + \dots + q_{n-1}X^{n-1}$  y el resto  $r$  de la división de  $A$  entre  $X - \alpha$  son el resultado de las siguientes operaciones:

$$q_{n-1} = a_n, \quad q_i = a_{i+1} + \alpha q_{i+1}, \quad r = a_0 + \alpha q_0.$$

Esto queda expresado esquemáticamente en la siguiente tabla en la que los términos de la última fila se obtienen sumando los de las dos primeras:

$a_n$	$a_{n-1}$	$a_{n-2}$	$\dots$	$a_2$	$a_1$	$a_0$
$\alpha$	$\alpha q_{n-1}$	$\alpha q_{n-2}$	$\dots$	$\alpha q_2$	$\alpha q_1$	$\alpha q_0$
	$q_{n-1}$	$q_{n-2}$	$q_{n-3}$	$\dots$	$q_1$	$q_0$   $r$

Otra forma de calcular el resto la proporciona el siguiente Teorema:

**Proposición 1.4.2 (Teorema del Resto)** Si  $P \in K[X]$  y  $\alpha \in K$ , entonces el resto de dividir un polinomio  $A$  entre  $X - \alpha$  es  $P(\alpha)$ .

**Demostración.** Si  $Q$  y  $R$  son el cociente y el resto de dividir  $A$  entre  $X - \alpha$ , entonces  $\text{gr}(R) \leq 0$ , con lo que  $R \in K$  y  $A = BQ + R$ . Lo primero implica que  $R(\alpha) = R$ . Aplicando la Proposición 1.3.4 tenemos

$$A(\alpha) = Q(\alpha)(\alpha - \alpha) + R(\alpha) = R.$$

■

Si  $A, B \in K[X]$  entonces decimos que  $B$  divide a  $A$  o que  $A$  es múltiplo de  $A$  si  $A = BC$  para algún polinomio  $C \in K[X]$ . Claramente el único múltiplo del polinomio 0 es el propio 0. En caso en que  $B \neq 0$  entonces  $B$  divide a  $A$  si y solo si el resto de dividir  $A$  entre  $B$  es 0.

Como consecuencia del Teorema del Resto tenemos el siguiente:

**Corolario 1.4.3 (Teorema de Ruffini)** Si  $A \in K[X]$  y  $\alpha \in K$  entonces  $\alpha$  es raíz de  $A$  si y solo si  $X - \alpha$  divide a  $A$ .

Sean  $A \in K[X]$  con  $A \neq 0$  y  $\alpha \in K$ . Supongamos que  $\alpha$  es raíz de  $A$ . Por el Teorema de Ruffini  $X - \alpha$  divide a  $A$ . Por otro lado si  $n$  es un entero positivo tal que  $(X - \alpha)^n$  divide a  $A$  entonces  $n$  es como mucho el grado de  $A$ , por la Proposición 1.3.1(3). La *multiplicidad* de  $\alpha$  como raíz de  $A$  es el mayor entero positivo  $n$  tal que  $(X - \alpha)^n$  divide a  $A$ .

**Proposición 1.4.4** Sean  $A \in K[X] \setminus \{0\}$ ,  $\alpha \in A$  y  $m$  un entero no negativo. Entonces las siguientes son equivalentes:

- (1)  $m$  es la multiplicidad de  $\alpha$  como raíz de  $A$ .
- (2)  $(X - \alpha)^m$  divide a  $A$  y  $(X - \alpha)^{m+1}$  no divide a  $A$ .
- (3) Existe un polinomio  $B$  tal que  $A = (X - \alpha)^m B$  y  $B(\alpha) \neq 0$ .

**Demostración.** La equivalencia entre (1) y (2) es obvia.

(2) implica (3). Si se verifica (2) entonces  $A = (X - \alpha)^m B$  para algún  $B \in K[X]$ . Si  $B(\alpha) = 0$  entonces por el Teorema de Ruffini  $B = (X - \alpha)C$  para algún  $C$  y por tanto  $A = (X - \alpha)^{m+1}C$ , en contra de la hipótesis. Por tanto se verifica (3).

(3) implica (2). Supongamos que  $A = (X - \alpha)^m B$  con  $B \in K[X]$  y  $B(\alpha) \neq 0$ . Si  $(X - \alpha)^{m+1}$  divide a  $A$  entonces  $(X - \alpha)^{m+1}C$  para algún  $C \in K[X]$ . Por tanto  $(X - \alpha)^{m+1}C = A = (X - \alpha)^m B$  y aplicando el Corolario 1.3.2 deducimos que  $(X - \alpha)C = B$  lo que implica que  $B(\alpha) = 0$ , que nos lleva a una contradicción. Por tanto  $(X - \alpha)^m$  divide a  $A$  pero  $(X - \alpha)^{m+1}$  no divide a  $A$ . ■

Por el Teorema de Ruffini eso pasa si y solo si existe un polinomio  $B$  tal que  $A = (X - \alpha)^m B$  y  $B(\alpha) \neq 0$ . En el caso en el que  $\alpha$  no sea raíz de  $A$  diremos que la multiplicidad de  $\alpha$  como raíz de  $A$  es 0.

**Corolario 1.4.5** Si  $A \in K[X]$  con  $A \neq 0$  y sea  $n = \text{gr}(A)$ . Entonces la suma de las multiplicidades de las raíces de  $A$  es menor o igual que  $n$ . En particular,  $A$  tiene a lo sumo  $n$  raíces.

**Demostración.** Sean  $\alpha_1, \dots, \alpha_k$  raíces de  $A$  y sea  $m_i$  la multiplicidad de  $\alpha_i$  como raíz de  $A$ . Entonces cada  $(X - \alpha_i)^{m_i}$  divide a  $A$ . Por tanto  $A = (X - \alpha_1)^{m_1} A_1$  para un polinomio  $A_1$ . Como  $\alpha_2 \neq \alpha_1$ ,  $\alpha_2$  es raíz de  $A_1$  y la multiplicidad de  $\alpha_2$  como raíz de  $A_1$  es la misma que como raíz de  $A$ . Luego  $A = (X - \alpha_1)^{m_1} (X - \alpha_2)^{m_2} A_2$  para otro polinomio  $A_2$ . Siguiendo el razonamiento de esta manera terminamos teniendo que  $A = (X - \alpha_1)^{m_1} \dots (X - \alpha_k)^{m_k} B$  para algún polinomio  $k$ . Por tanto  $n = \text{gr}(A) = m_1 + \dots + m_k + \text{gr}(B) \geq m_1 + \dots + m_k$ . ■

Un polinomio de  $K[X]$  diferente de 0 se dice que es *completamente factorizable* en  $K[X]$  si es de la forma

$$a(X - \alpha_1) \cdots (X - \alpha_n)$$

con  $a, \alpha_1, \dots, \alpha_n \in K$ .

**Proposición 1.4.6** Las siguientes condiciones son equivalentes para un polinomio  $A$  de  $K[X]$  distinto de 0:

- (1)  $A$  es completamente factorizable en  $K[X]$ .
- (2)  $A$  es producto de polinomios de grado a lo sumo 1.
- (3) La suma de las multiplicidades de las raíces de  $A$  es igual al grado de  $A$ .

**Demostración.** (2) implica (1). Supongamos que  $A$  es producto de polinomios de grado  $\leq 1$ . Es decir  $A = a_1 \cdots a_k (b_{11}X + b_{10}) \cdots (b_{n1}X + b_{n0})$  con cada  $a_i$  con  $1 \leq i \leq k$  y cada  $b_{j1}$  en  $K \setminus \{0\}$  y cada  $b_{j0}$  en  $K$  con  $1 \leq j \leq n$ . Sean  $a = a_1 \cdots a_k b_{11} \cdots b_{n1}$  y  $\alpha_i = -b_{i1}^{-1} b_{i0}$  para todo  $i = 1, \dots, n$ . Entonces  $A = a(X - \alpha_1) \cdots (X - \alpha_n)$ . Por tanto  $K$  es completamente factorizable.

(1) implica (3). Supongamos que  $A = a(X - \alpha_1) \cdots (X - \alpha_n)$  con  $a, \alpha_1, \dots, \alpha_n \in K$ . Agrupando los  $\alpha_i$  iguales podemos reescribir esto como  $A = a(X - \beta_1)^{m_1} \cdots (X - \beta_k)^{m_k}$  con  $\beta_1, \dots, \beta_k$  distintos dos a dos. Comparando los grados deducimos que  $n = m_1 + \cdots + m_k$ . Si  $\gamma$  es una raíz de  $A$  entonces  $0 = a(\gamma - \beta_i)^{m_1} \cdots (\gamma - \beta_k)^{m_k}$  y por tanto  $\gamma = \beta_i$  para algún  $i$ . Es decir,  $\beta_1, \dots, \beta_k$  son las raíces de  $A$ . Además para cada  $i = 1, \dots, k$  tenemos  $A = (X - \beta_i)^{m_i} B_i$ , con  $B_i$  el producto de  $a$  y los factores  $(X - \beta_j)^{m_j}$  con  $j \neq i$ . Como los  $\beta_i$  son distintos dos a dos  $B_i(\beta_i) \neq 0$ . Aplicando la Proposición 1.4.4 deducimos que  $m_i$  es la multiplicidad de  $\alpha_i$  como raíz de  $A$ . Por tanto la igualdad  $n = m_1 + \cdots + m_k$  asegura precisamente que se cumple (3).

(3) implica (2). Sean  $\alpha_1, \dots, \alpha_k$  las raíces de  $A$  y sea  $m_i$  la multiplicidad de  $\alpha_i$  como raíz de  $A$ . Eso implica que  $A = (X - \alpha_1)^{m_1} \cdots (X - \alpha_k)^{m_k} B$  para algún polinomio  $B$ . La hipótesis significa que  $n = m_1 + \cdots + m_k$ . Por tanto,  $n = \text{gr}(A) = m_1 + \cdots + m_k + \text{gr}(B) = n + \text{gr}(B)$  con lo que  $\text{gr}(B) = 0$ . Luego  $A$  es producto de polinomios de grado  $\leq 1$ . ■

Todo polinomio de grado 1 tiene una raíz pues si  $a \neq 0$  entonces  $-ba^{-1}$  es raíz de  $aX + b$ . Sin embargo los polinomios de grado 2 no tienen necesariamente raíces. Por ejemplo,  $\sqrt{2}$  no es un número racional lo que implica que el polinomio  $X^2 - 2$  que está en  $\mathbb{Q}[X]$ , no tiene raíces en  $\mathbb{Q}$ , pero sí que tiene raíces en  $\mathbb{R}$ . Por eso necesitamos un cuerpo más grande que los racionales. Por otro lado el polinomio  $X^2 + 1$  que también está en  $\mathbb{Q}[X]$  no tiene ninguna raíz en  $\mathbb{R}$ , aunque  $i = \sqrt{-1}$  es una raíz de  $X^2 + 1$ . Por eso necesitamos un cuerpo todavía más grande que los reales y esa es la razón de la existencia de los números complejos. De esta manera todo número real  $a$  tiene una raíz cuadrada que será real si  $a \geq 0$  e imaginaria si  $a < 0$ .

El siguiente teorema muestra que este proceso de ampliación no es necesario con el cuerpo de los números complejos.

**Teorema 1.4.7 (Teorema Fundamental del Álgebra)** Sea  $A \in \mathbb{C}[X]$  con  $\text{gr}(A) > 0$ . Entonces  $A$  tiene una raíz en  $\mathbb{C}$ .

**Corolario 1.4.8** Todo polinomio no nulo  $A$  de  $\mathbb{C}[X]$  es completamente factorizable en  $\mathbb{C}[X]$ . Más aún, si  $\alpha_1, \dots, \alpha_k$  son las raíces de  $A$  y  $m_i$  es la multiplicidad de  $\alpha_i$  como raíz de  $A$  entonces

$$A = a(X - \alpha_1)^{m_1} \cdots (X - \alpha_k)^{m_k}.$$

Esta expresión se llama factorización de  $A$  en  $\mathbb{C}[X]$ .

**Demostración.** Sea  $A \in \mathbb{C}[X]$  con  $A \neq 0$ , sean  $\alpha_1, \dots, \alpha_k$  las raíces de  $A$  y sea  $m_i$  la multiplicidad de  $\alpha_i$  como raíz de  $A$ . Entonces  $A = (X - \alpha_1)^{m_1} \cdots (X - \alpha_k)^{m_k} B$  para algún polinomio  $B$ . Si  $\text{gr}(B) > 0$  entonces por el Teorema Fundamental del Álgebra,  $B$  tiene una raíz  $\beta$ . En realidad  $\beta$  es raíz de  $A$  con lo que  $\beta = \alpha_i$  para algún  $i$  y  $(X - \alpha_i)^{m_i+1}$  divide a  $A$  en contra de que  $m_i$  es la multiplicidad de  $\alpha_i$  como raíz de  $A$ . Concluimos que  $B$  tiene grado 0 y por tanto  $A$  completamente factorizable. ■

Encontrar la factorización en  $\mathbb{C}[X]$  de un polinomio es equivalente a encontrar sus raíces y las multiplicidades de cada una de ellas. Eso no es una tarea fácil. La siguiente proposición proporciona una herramienta para al menos calcular las raíces racionales para polinomios con coeficientes enteros. De hecho también se puede utilizar para calcular las raíces racionales para polinomios con coeficientes racionales.

**Proposición 1.4.9** *Sea  $A$  es un polinomio no nulo con coeficientes enteros y sean  $a$  el término independiente y  $b$  el coeficiente del término principal. Entonces todas las raíces racionales de  $A$  son de la forma  $\frac{u}{v}$  con  $u$  un divisor de  $a$  y  $v$  un divisor de  $b$ .*

**Ejemplo 1.4.10** Vamos a ver cómo factorizar algunos polinomios como producto de polinomios complejos de grado 1, o lo que es lo mismo vamos a ver cómo calcular sus raíces y las multiplicidades de cada una de ellas. La demostración del Corolario 1.4.8 nos indica el procedimiento que vamos a ilustrar factorizando el polinomio  $A = X^4 + 6X^3 + 13X^2 + 12X + 4$ . Por la Proposición 1.4.9 las únicas posibles racionales son 1, -1, 2, -2, 4 y -4. Está claro que 1 no puede ser raíz pues todos los coeficientes son positivos. Probamos con -1 usando la Regla de Ruffini:

$$\begin{array}{r|rrrrr} & 1 & 6 & 13 & 12 & 4 \\ -1 & & -1 & -5 & -8 & -4 \\ \hline & 1 & 5 & 8 & 4 & 0 \end{array}$$

Como el resto es 0, efectivamente -1 es raíz de  $A$  y tenemos que  $A = (X+1)(X^3+5X^2+8X+4)$ . Ahora podemos intentar ver si el factor de grado 3 es divisible por  $X+1$  aplicando de nuevo la regla de Ruffini:

$$\begin{array}{r|rrrr} & 1 & 5 & 8 & 4 \\ -1 & & -1 & -4 & -4 \\ \hline & 1 & 4 & 4 & 0 \end{array}$$

Por tanto  $A = (X+1)^2(X^2+4X+4)$ . Podríamos seguir usando la Regla de Ruffini pero como sabemos resolver ecuaciones de segundo grado no nos hace falta pues las raíces de  $X^2+4X+4$  son

$$X = \frac{-4 \pm \sqrt{16-16}}{2} = -2.$$

O sea la única raíz de  $X^2+4X+4$  es -2 con multiplicidad 2 pues en efecto  $(X+2)^2 = X^2+4X+4$ . Concluimos que la factorización del polinomio original es

$$X^4 + 6X^3 + 13X^2 + 12X + 4 = (X+1)^2(X+2)^2.$$

**Ejemplo 1.4.11** Vamos ahora a factorizar  $A = X^3 - \frac{1}{2}X^2 + X - \frac{1}{2}$ . Para empezar ponemos  $A = \frac{1}{2}(2X^3 - X^2 + 2X - 1)$  y de esta manera procedemos a factorizar  $B = 2X^3 - X^2 + 2X - 1$  para el cuál podemos usar la Proposición 1.4.9. Las raíces racionales de este polinomio habrán de ser 1, -1,  $\frac{1}{2}$  ó  $-\frac{1}{2}$ . Probando con la Regla de Ruffini observamos que ni 1 ni -1 son raíces:

$$\begin{array}{r|rrrr} 1 & 2 & -1 & 2 & -1 \\ & & 2 & 1 & 3 \\ \hline & 2 & 1 & 3 & 2 \end{array} \quad \begin{array}{r|rrrr} -1 & 2 & -1 & 2 & -1 \\ & & -2 & 3 & 5 \\ \hline & 2 & -3 & 5 & 4 \end{array}$$

Sin embargo  $\frac{1}{2}$  sí que es una raíz:

$$\begin{array}{r|rrrr} \frac{1}{2} & 2 & -1 & 2 & -1 \\ & & 1 & 0 & 1 \\ \hline & 2 & 0 & 2 & 0 \end{array}$$

Por tanto  $A = \frac{1}{2}(X - \frac{1}{2})(2X^2 + 2) = (X - \frac{1}{2})(X^2 + 1)$ . Resolviendo la ecuación  $X^2 + 1 = 0$  cuyas raíces son  $i$  y  $-i$  obtenemos

$$X^3 - \frac{1}{2}X^2 + X - \frac{1}{2} = \left(X - \frac{1}{2}\right)(X - i)(X + i).$$

En los dos ejemplos que hemos puesto hemos conseguido factorizar el polinomio dado buscando primero raíces racionales y al final resolviendo una ecuación de segundo grado. Por desgracia esto no es siempre posible. De hecho a menudo el cálculo de las raíces de un polinomio es casi imposible a mano.

**Ejemplo 1.4.12** Queremos calcular las raíces del polinomio

$$X^4 + X^3 - X^2 - 2X - 2$$

Después de probar con las cuatro posibles raíces racionales,  $\pm 1$  y  $\pm 2$  llegamos a la conclusión de que no tiene ninguna raíz racional. Sin embargo sabemos que tiene raíces complejas. En estos casos podemos usar el comando `solve` de Maxima de la siguiente forma:

`solve(X^4+X^3-X^2-2*X-2);`

$$\left[X = -\sqrt{2}, X = \sqrt{2}, X = -\frac{\sqrt{3}\%i + 1}{2}, X = \frac{\sqrt{3}\%i - 1}{2}\right]$$

Teniendo en cuenta que `%i` representa en Maxima el número complejo  $i = \sqrt{-1}$  obtenemos que el polinomio tiene dos raíces reales  $\pm\sqrt{2}$  y dos raíces complejas:  $\frac{-1 \pm \sqrt{3}i}{2}$ . Todas ellas tendrán que tener multiplicidad 1. Podemos comprobar el resultado de la siguiente forma:

`a:2^(1/2)$ b:(-1+3^(1/2)*%i)/2$ c:(-1-3^(1/2)*%i)/2$  
expand((X-a)*(X+a)*(X-b)*(X-c));`

$$X^4 + X^3 - X^2 - 2X - 2$$

que proporciona el polinomio original.

Veamos otro ejemplo:

**Ejemplo 1.4.13** Vamos ahora a calcular las multiplicidades de las raíces del siguiente polinomio:

$$X^{10} + 11X^9 + 46X^8 + 83X^7 - 4X^6 - 391X^5 - 1090X^4 - 1695X^3 - 1701X^2 - 1080X - 324$$

Para ello realizamos el siguiente cálculo con Maxima:

```
p:X^10+11*X^9+46*X^8+83*X^7-4*X^6-391*X^5-1090*X^4-1695*X^3-1701*X^2-1080*X-324;
```

```
solve(p);
```

```
multiplicities;
```

$$X^{10} + 11X^9 + 46X^8 + 83X^7 - 4X^6 - 391X^5 - 1090X^4 - 1695X^3 - 1701X^2 - 1080X - 324 \quad (p)$$

$$\left[ X = -\frac{\sqrt{13}-1}{2}, X = \frac{\sqrt{13}+1}{2}, X = -1, X = -\frac{\sqrt{7}i+1}{2}, X = \frac{\sqrt{7}i-1}{2}, X = -3 \right]$$

$$[1, 1, 1, 2, 2, 3]$$

El output obtenido nos está diciendo que el polinomio tiene las raíces y multiplicidades que se indican en la siguiente tabla:

raíz :	-1	-3	$\frac{1+\sqrt{13}}{2}$	$\frac{1-\sqrt{13}}{2}$	$\frac{-1+\sqrt{7}i}{2}$	$\frac{-1-\sqrt{7}i}{2}$
multiplicidad :	1	3	1	1	2	2

Podemos comprobar el resultado obtenido con la siguiente cuenta:

```
a:(1+sqrt(13))/2$ b:(1-sqrt(13))/2$ c:(-1+sqrt(7)*%i)/2$ d:conjugate(c)$
```

```
expand((X+1)*(X+3)^3*(X-a)*(X-b)*(X-c)^2*(X-d)^2);
```

$$X^{10} + 11X^9 + 46X^8 + 83X^7 - 4X^6 - 391X^5 - 1090X^4 - 1695X^3 - 1701X^2 - 1080X - 324$$

Si  $A = \sum_i a_i X^i$  es un polinomio con coeficientes complejos entonces ponemos  $\bar{A} = \sum_i \bar{a}_i X^i$ . Como consecuencia de las propiedades de la conjugación de números complejos que vimos en la Proposición 1.2.1 se tiene que si  $\alpha \in \mathbb{C}$  entonces  $\overline{A(\bar{\alpha})} = A(\alpha)$  con lo que  $\alpha$  es raíz de  $A$  si y solo si  $\bar{\alpha}$  es raíz de  $\bar{A}$ . Más aún, si  $B$  es otro polinomio con coeficientes complejos entonces

$$\overline{AB} = \bar{A} \bar{B}$$

**Proposición 1.4.14** Si  $A \in \mathbb{C}[X]$  y  $\alpha \in \mathbb{C}$  entonces la multiplicidad de  $\alpha$  como raíz de  $A$  es la misma que la de  $\bar{\alpha}$  como raíz de  $\bar{A}$ .

**Demostración.** Sea  $n$  la multiplicidad de  $\alpha$  como raíz de  $A$ . Entonces existe un polinomio  $B \in K[X]$  tal que  $A = (X - \alpha)^n B$  y  $B(\alpha) \neq 0$ . Luego  $\bar{A} = (X - \bar{\alpha})^n \bar{B}$  y  $\bar{B}(\bar{\alpha}) = \overline{B(\alpha)} \neq 0$ . Por tanto  $n$  es la multiplicidad de  $\bar{\alpha}$  como raíz de  $\bar{B}$ . ■

**Corolario 1.4.15** Si  $A \in \mathbb{R}[X]$  entonces  $\alpha$  y  $\bar{\alpha}$  tienen las mismas multiplicidades como raíces de  $A$ . Además  $A$  es un producto de polinomios de  $\mathbb{R}[X]$  de grado a lo sumo 2.

**Demostración.** La primera parte es consecuencia inmediata de la Proposición 1.4.14. Para ver la segunda tomamos la factorización de  $A$  en  $\mathbb{C}[X]$ . Separando las raíces reales  $\alpha_1, \dots, \alpha_r$  por un lado y las que no lo son emparejadas por conjugados  $\beta_1, \bar{\beta}_1, \dots, \beta_s, \bar{\beta}_s$  tenemos

$$A = a(X - \alpha_1)^{m_1} \cdots (X - \alpha_r)^{m_r} (X - \beta_1)^{n_1} (X - \bar{\beta}_1)^{n_1} \cdots (X - \beta_s)^{n_s} (X - \bar{\beta}_s)^{n_s}.$$

Obsérvese que cada  $X - \beta_i$  tienen el mismo exponente que  $X - \bar{\beta}_i$  por la Proposición 1.4.14. Entonces cada  $Q_i = (X - \beta_i)(X - \bar{\beta}_i) = X^2 - 2\operatorname{Re}(\beta_i)X + |\beta_i|^2$  es un polinomio en  $\mathbb{R}[X]$  de grado 2 y

$$A = a(X - \alpha_1)^{m_1} \cdots (X - \alpha_r)^{m_r} Q_1^{n_1} \cdots Q_s^{n_s}.$$

■

**Ejemplo 1.4.16** En el Ejemplo 1.4.12 hemos obtenido

$$P = X^4 + X^3 - X^2 - 2X - 2 = (X - 2)(X + 2) \left( X - \frac{-1 + \sqrt{3}i}{2} \right) \left( X - \frac{-1 + \sqrt{3}i}{2} \right).$$

Obsérvese que los dos últimos factores son de la forma  $X - \alpha$  y  $X - \bar{\alpha}$ . Por tanto su producto será  $X^2 - 2(\alpha + \bar{\alpha}X + |\alpha|^2) = X^2 + X + 1$ . Por tanto

$$P = (X - 2)(X + 2)(X^2 + X + 1)$$

De forma si  $P$  es el polinomio del Ejemplo 1.4.13 y ponemos

$$\alpha_1 = \frac{1 + \sqrt{13}}{2}, \quad \alpha_2 = \frac{1 - \sqrt{13}}{2} \quad \text{y} \quad \beta = \frac{-1 + \sqrt{7}i}{2}$$

entonces  $\alpha_1$  y  $\alpha_2$  son reales de multiplicidad 1, mientras que  $\beta$  y  $\bar{\beta}$  son raíces complejas conjugadas de multiplicidad 2. Como  $2\operatorname{Re}(\beta) = -1$  y  $|\beta|^2 = \frac{1}{4} + \frac{7}{4} = 2$  tenemos

$$P = (X + 1)(X + 3)^3(X - \alpha_1)(X - \alpha_2)(X + X + 2)^2$$

lo que podemos comprobar con Maxima:

`expand((X+1)*(X+3)^3*(X-a)*(X-b)*(X^2+X+2)^2);`

$$X^{10} + 11X^9 + 46X^8 + 83X^7 - 4X^6 - 391X^5 - 1090X^4 - 1695X^3 - 1701X^2 - 1080X - 324$$

## 1.5 Problemas

**1.1** Sea  $K$  el conjunto formado por los números complejos cuya parte real e imaginaria es un número racional, es decir, los elementos de  $K$  son de la forma  $a + bi$  con  $a, b \in \mathbb{Q}$ . Demostrar

- (1) Si  $x, y \in K$ , entonces  $x + y, xy \in K$ .
- (2)  $0, 1 \in K$ .
- (3) Si  $x \in K$ , entonces  $-x \in K$  y, si además  $x \neq 0$ , entonces  $x^{-1} \in K$ .

Deducir que  $K$  es un cuerpo.

**1.2** Demostrar que el conjunto formado por los números reales de la forma  $a + b\sqrt{2}$  es un cuerpo con la suma y producto habituales. (Indicación: Seguir las pautas del ejercicio anterior.)

**1.3** Construye las tablas de sumar y multiplicar de  $\mathbb{Z}_6, \mathbb{Z}_7$  y  $\mathbb{Z}_{24}$ , calcula el opuesto de cada elemento, identifica los elementos que son invertibles y para cada uno de ellos calcula sus inverso.

**1.4** Demostrar que en  $\mathbb{Z}_n$  el opuesto de  $a$  es  $n - a$ .

**1.5** Definimos en el conjunto de los números reales la siguiente suma y producto (que denotamos  $\oplus$  y  $\otimes$  para distinguirlas de la suma y producto habituales que las denotamos como siempre):

$$a \oplus b = a + b + 1, \quad a \otimes b = a + b + ab.$$

Demostrar que con estas operaciones  $\mathbb{R}$  es un cuerpo identificando quién sería el cero, el uno, los opuestos y los inversos. Demostrar que esto funciona también si en lugar de comenzar con los números reales hubiéramos comenzado con cualquier cuerpo.

**1.6** Demostrar que si  $a$  es un elemento invertible de un anillo y  $b$  y  $c$  son dos elementos distintos del mismo anillo entonces  $ab \neq ac$ . Dar un ejemplo de un anillo que no sea un cuerpo en el que esto pasa para todos los elementos del anillo distinto de cero aunque no sean invertibles.

**1.7** Demostrar las seis primeras propiedades del Teorema 1.2.1.

**1.8** Sea  $z$  un número complejo distinto de 0 y sean  $r = |z|$  y  $\alpha = \arg(z)$ . Sea  $s = \sqrt[n]{r}$ , sea  $n$  un entero positivo y para cada entero  $j$  sea

$$x_j = s \left( \cos \frac{\alpha + 2j\pi}{n} + i \operatorname{sen} \frac{\alpha + 2j\pi}{n} \right)$$

Demostrar:

- (1) Para cada entero  $i$  se verifica que  $x_j^n = z$ .
- (2)  $x_k = x_j$  si y solo si  $n$  divide a  $k - j$ .

- (3) Si  $w$  es un número complejo tal que  $w^n = z$  entonces  $w = x_j$  para algún entero  $j$ .
- (4) Calcular todos los números complejos  $z$  que satisfacen alguna de las siguientes igualdades y en todos los casos representa las soluciones que satisfacen cada una de las ecuaciones en el plano complejo.

$$z^2 = 1, \quad z^3 = 1, \quad z^4 = 1, \quad z^5 = 1, \quad z^4 = 4, \quad z^2 = i, \quad z^3 = 1 + i.$$

**1.9** Completar la demostración de la Proposición 1.3.1.

**1.10** Factorizar los siguientes polinomios en  $\mathbb{C}[X]$ :  $X^2 - 9$ ,  $X^2 + 1$ ,  $X^2 + 9$ ,  $X^2 + 2$ ,  $X^3 - 7X + 6$ ,  $X^4 - 1$ ,  $X^4 - 2$ ,  $X^2 + i$ .

**1.11** Sean  $A, B$  polinomios no nulos de  $K[X]$  y  $\alpha \in K$ . Demostrar que  $\alpha$  es raíz de  $AB$  si y solo si es raíz de  $A$  o de  $B$ . Sean  $m$  la multiplicidad de  $\alpha$  como raíz de  $A$  y  $n$  la multiplicidad de  $B$  como raíz de  $B$ . Demostrar que la multiplicidad de  $\alpha$  como raíz de  $AB$  es  $m + n$ .

**1.12** Realizar las siguientes divisiones de polinomios:  $X^5 - 3X^2 + 1 : X^3 - 1$ ,  $X^6 - 1 : X^2 - 1$ ,  $X^4 + i : X^2 + 1$ ,  $X^4 - iX + 2 : X^3 + 2i$ .

**1.13** Calcula las raíces del siguiente polinomio y sus multiplicidades usando Maxima, escribir cada uno de estos polinomios como producto de polinomios con coeficientes reales de grado al menos dos y comprueba los resultados obtenidos:

$$\begin{aligned} &X^7 + 10X^6 + 33X^5 + 42X^4 + 36X^3 + 66X^2 + 80X + 32, \\ &X^5 + 4X^4 + 5X^3 + 5X^2 + 4X + 1, \\ &X^8 + 5X^7 + 10X^6 + 15X^5 + 18X^4 + 15X^3 + 10X^2 + 5X + 1 \end{aligned}$$



## Capítulo 2

# Espacios Vectoriales

En todo este capítulo  $K$  es un cuerpo, cuyos elementos llamaremos escalares.

### 2.1 Definición y ejemplos de espacios vectoriales

El siguiente es el concepto más importante de todo el curso.

**Definición 2.1.1** *Un espacio vectorial sobre  $K$  es un conjunto  $V$ , cuyos elementos llamaremos vectores, junto con dos operaciones:*

**Suma de vectores:** *Cada dos vectores  $v, w \in V$  dan lugar a otro vector  $v + w \in V$ .*

**Producto de escalares por vectores:** *Con un escalar  $\alpha \in K$  y un vector  $v \in V$  tenemos otro vector  $\alpha v \in V$ .*

*De forma que se cumplan las siguientes condiciones para  $x, y, z \in V$  y  $\alpha, \beta \in K$ .*

- (1) *Asociativa de la Suma:*  $x + (y + z) = (x + y) + z$ .
- (2) *Conmutativa de la Suma:*  $x + y = y + x$ .
- (3) *Vector Cero:* *Existe un elemento  $\bar{0}$  en  $V$ , que llamamos vector cero de  $V$ , tal que  $x + \bar{0} = x$ .*
- (4) *Vector Opuesto:* *Existe un elemento  $-x$  en  $V$ , que llamamos opuesto de  $x$ , tal que  $x + (-x) = \bar{0}$ .*
- (5) *Asociativa del Producto por Escalares:*  $\alpha(\beta x) = (\alpha\beta)x$
- (6) *Distributiva respecto de la Suma de Escalares:*  $(\alpha + \beta)x = \alpha x + \beta x$ .
- (7) *Distributiva respecto de la Suma de Vectores:*  $\alpha(x + y) = \alpha x + \alpha y$ .
- (8)  $1x = x$ .

A menudo distinguiremos los escalares de los vectores usando letras griegas para los escalares y letras latinas para los vectores aunque no siempre seremos consistentes con esta notación.

**Ejemplo 2.1.2 (El espacio vectorial  $K^n$ )** Sea  $n$  un entero positivo y  $V = K^n$ , o sea los elementos de  $K^n$  son  $n$ -uplas  $(a_1, \dots, a_n)$  con cada  $a_i \in K$ . A partir de la suma de  $K$  definimos una suma en  $K^n$  mediante la siguiente regla:

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n).$$

A partir del producto en  $K$  definimos un producto de escalares por vectores mediante la siguiente regla:

$$\alpha(x_1, \dots, x_n) = (\alpha x_1, \dots, \alpha x_n).$$

Es fácil ver que estas operaciones satisfacen los axiomas de espacio vectorial con lo que  $K^n$  es un espacio vectorial sobre  $K$ .

**Ejemplo 2.1.3 (El espacio vectorial  $M_{n,m}(K)$ )** Una *matriz* con  $n$  filas y  $m$  columnas, o simplemente matriz  $n \times m$ , es una tabla formada por  $nm$  elementos de  $K$  dispuestos en  $n$  filas y  $m$  columnas, o sea una configuración de la siguiente forma en la que todos los  $a_{ij}$  están en  $K$ .

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix}$$

Si  $n = m$  entonces decimos que se trata de una *matriz cuadrada*.

Vamos a denotar por  $M_{n \times m}(K)$  el conjunto de las matrices con  $n$  filas y  $m$  columnas. También usamos la notación  $M_n(K) = M_{n \times n}(K)$ , para el caso de matrices cuadradas. Además, con el fin de escribir lo menos posible vamos a denotar una matriz como la anterior de forma compacta poniendo  $(a_{ij})$ . Si queremos enfatizar el número  $n$  de filas y  $m$  de columnas a veces denotaremos la matriz como  $(a_{ij})_{n \times m}$ .

Podemos pensar en una matriz como un elemento de  $K^{mn}$ , o sea una lista de  $mn$  elementos de  $K$ , solo que los ubicamos en una tabla en lugar de en una lista. Este punto de vista nos sirve para darnos cuenta de que la siguiente suma de matrices y producto por escalares por matrices hace de  $M_{n,m}(K)$  un espacio vectorial sobre  $K$ :

$$(a_{ij}) + (b_{ij}) = (a_{ij} + b_{ij}), \quad \alpha(a_{ij}) = (\alpha a_{ij}).$$

**Ejemplo 2.1.4 (El espacio vectorial de las sucesiones  $K^\infty$ )** Sea  $V$  el conjunto de las sucesiones  $(x_1, x_2, \dots)$  de elementos de  $K$ . Escribimos estas sucesiones de forma compacta poniendo  $(x_i)$  e imitamos la suma y el producto de escalares de  $K^n$  y  $M_{m,n}(K)$  definiendo la siguiente suma y producto por escalares:

$$(x_i) + (y_i) = (x_i + y_i), \quad \alpha(x_i) = (\alpha x_i).$$

Esto proporciona un nuevo espacio vectorial.

**Ejemplo 2.1.5 (El espacio vectorial de los polinomios)** El conjunto  $K[X]$  de los polinomios con coeficientes en  $K$  es otro espacio vectorial sobre  $K$  con la suma de polinomios y el producto de elementos de  $K$  por polinomios tal como se realizan en el anillo  $K[X]$ .

**Ejemplo 2.1.6 (Producto de espacios vectoriales)** Si  $V$  y  $W$  son dos espacios vectoriales entonces el producto cartesiano  $V \times W$  es otro espacio vectorial con las siguientes operaciones:

$$(v_1, w_1) + (v_2, w_2) = (v_1 + v_2, w_1 + w_2), \quad \alpha(v, w) = (\alpha v, \alpha w).$$

Esto se puede generalizar de forma obvia al producto de una lista de espacios vectoriales  $V_1, \dots, V_n$ . Incluso a un producto cartesiano de una familia infinita de espacios vectoriales, es decir supongamos que  $I$  es un conjunto y para cada  $i \in I$  tenemos un espacio vectorial  $V_i$  sobre  $K$ . El producto  $\prod_{i \in I} V_i$  está formado por  $I$ -uplas  $(v_i)_{i \in I}$  de forma que cada  $v_i \in V_i$  y las siguientes operaciones definen en  $\prod_{i \in I} V_i$  una estructura de espacio vectorial sobre  $K$ :

$$(v_i)_{i \in I} + (w_i)_{i \in I} = (v_i + w_i)_{i \in I}, \quad \alpha(v_i)_{i \in I} = (\alpha v_i)_{i \in I}.$$

**Ejemplo 2.1.7 (El espacio vectorial de las aplicaciones de un conjunto al cuerpo  $K$ )**

Sea  $A$  un conjunto no vacío y sea  $K^A$  el conjunto de las aplicaciones  $f : A \rightarrow K$ . Si  $f, g : A \rightarrow K$  son dos aplicaciones de un conjunto  $A$  en el cuerpo  $K$ , entonces definimos la suma de  $f$  y  $g$  como la aplicación  $f + g : A \rightarrow K$  dada por

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

y si  $\alpha \in K$ , entonces se define la aplicación  $\alpha f : A \rightarrow K$  por la regla

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

Esto dota al conjunto de las aplicaciones de  $A$  en  $K$  de una estructura de espacio vectorial sobre  $K$  que denotaremos  $K^A$ .

Por ejemplo, supongamos que  $A = \mathbb{R}^+ = \{x \in \mathbb{R} : x \geq 0\}$  y  $K = \mathbb{R}$ ,  $f, g : \mathbb{R}^+ \rightarrow \mathbb{R}$  son las funciones raíz cuadrada y coseno, es decir

$$f(x) = \sqrt{x}, \quad g(x) = \cos x.$$

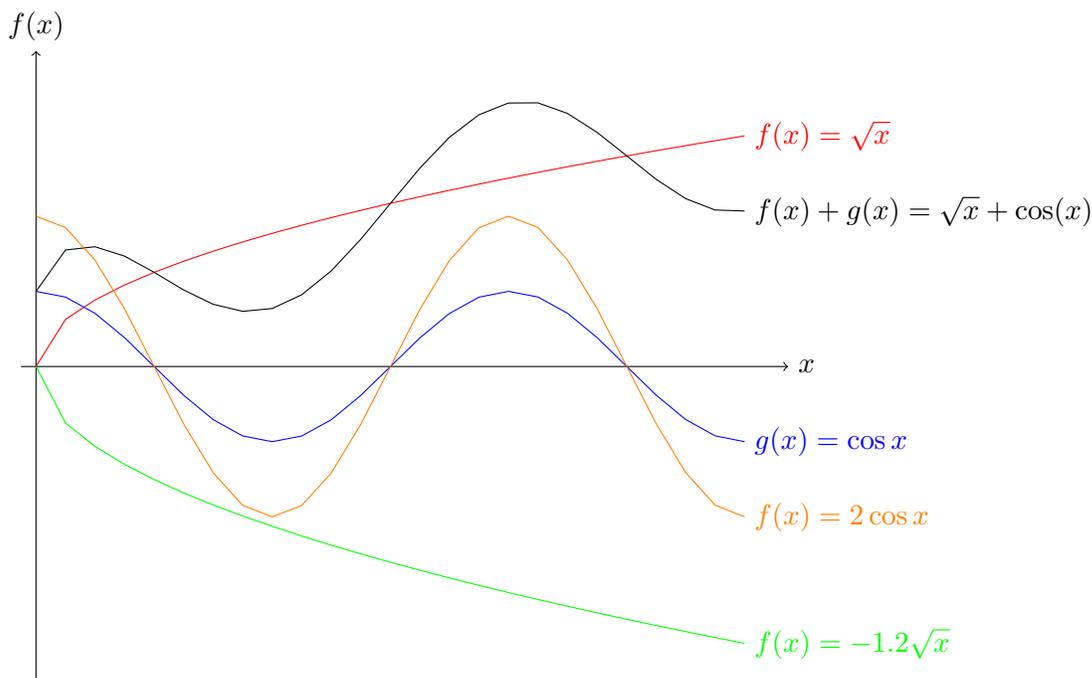
Entonces  $f + g$  es la función dada por

$$(f + g)(x) = \sqrt{x} + \cos x$$

y  $-3f$  y  $4g$  son las funciones dadas por

$$(-3f)(x) = -3\sqrt{x} \quad (4g)(x) = 4\cos x.$$

En la siguiente figura se han representado las funciones  $f$  (rojo),  $g$  (azul),  $f + g$  (negro),  $-1.2f$  (verde) y  $2g$  (naranja) en el intervalo  $[0, 3\pi]$ .



En física los elementos de este espacio vectorial aparecen cada vez que tenemos asociado a los puntos de un ambiente físico un valor. Por ejemplo, podemos considerar como ambiente físico una habitación de forma que  $A$  serían los puntos de la habitación y la aplicación que asocia a cada punto de la habitación la temperatura en dicho punto es un vector del espacio vectorial  $\mathbb{R}^A$ . Seguro que te puedes imaginar otros muchos contextos físicos en los que aparecen estos “vectores”.

**Ejemplo 2.1.8 (El espacio vectorial de las aplicaciones a un espacio vectorial)** Sea  $V$  un espacio vectorial sobre  $K$  y sea  $A$  un conjunto arbitrario. Sea  $V^A$  el conjunto de las aplicaciones de  $A$  a  $V$ . Copiando las definiciones del Ejemplo 2.1.7 podemos definir una suma y un producto por escalares en  $V^A$ . O sea si  $f, g \in V^A$  y  $\alpha \in K$  entonces definimos  $f + g, \alpha f \in V^A$  con las siguientes reglas:

$$(f + g)(a) = f(a) + g(a) \quad \text{y} \quad (\alpha f)(a) = \alpha f(a) \quad (a \in A).$$

Esto nos proporciona un nuevo espacio vectorial.

En el mundo de la física los elementos de este espacio vectorial aparecen por ejemplo en la forma de campos de vectores. Imaginemos, por ejemplo, los puntos de un campo magnético como conjunto  $A$  y como vector el elemento de  $(\mathbb{R}^3)^A$  que a cada punto de  $A$  le asocia el vector del campo magnético en dicho punto.

**Proposición 2.1.9** Si  $V$  es un espacio vectorial sobre  $K$ ,  $x, y, z \in V$  y  $\alpha \in K$ , entonces:

- (1) Si  $x + y = x + z$ , entonces  $y = z$ .
- (2)  $\alpha(-x) = -\alpha x = (-\alpha)x$ .

(3)  $\alpha x = \bar{0}$  si y sólo si  $\alpha = 0$  ó  $x = \bar{0}$ .

**Demostración.** (1) Se demuestra sumando  $-x$  en la igualdad:

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

(2) Aplicando (1) a la igualdad  $\alpha x + \alpha(-x) = \alpha(x + (-x)) = \alpha 0 = 0 = \alpha x + (-\alpha x)$ , se obtiene la primera propiedad y la segunda se hace de forma análoga.

(3) Aplicando (1) a la igualdad  $0x + 0x = (0 + 0)x = 0x = 0 + 0x$  se deduce que  $\alpha x = 0$ . Análogamente, aplicando 1 a la igualdad  $\alpha\bar{0} + \alpha\bar{0} = \alpha(\bar{0} + \bar{0}) = \alpha\bar{0} = 0 + \alpha\bar{0}$  deducimos que  $\alpha x = 0$ . Esto demuestra que si  $\alpha = 0$  ó  $x = \bar{0}$ , entonces  $\alpha x = \bar{0}$ .

Recíprocamente, supongamos que  $\alpha x = 0$ . Si  $\alpha \neq 0$ , entonces multiplicando por  $\alpha^{-1}$  tenemos

$$x = 1x = (\alpha^{-1}\alpha)x = \alpha^{-1}(\alpha x) = \alpha^{-1}\bar{0} = \bar{0},$$

donde la última igualdad la acabamos de ver en el párrafo anterior. ■

A partir de ahora, salvo que haya problemas de confusión el vector cero  $\bar{0}$ , lo denotaremos simplemente 0.

## 2.2 Aplicaciones lineales y matrices

**Definición 2.2.1** Sean  $V$  y  $W$  dos espacios vectoriales sobre  $K$ . Una aplicación  $f : V \rightarrow W$  se dice que es lineal si para cada  $\alpha \in K$  y cada  $v_1, v_2 \in V$  se verifican las siguientes igualdades:

$$f(v_1 + v_2) = f(v_1) + f(v_2), \quad f(\alpha v_1) = \alpha f(v_1).$$

**Ejemplos 2.2.2** Dado  $a \in K$  consideramos la aplicación

$$\begin{aligned} M_a : K &\rightarrow K \\ x &\mapsto ax \end{aligned}$$

Comprobamos que  $M_a$  es lineal:

$$M_a(x_1 + x_2) = a(x_1 + x_2) = ax_1 + ax_2 = M_a(x_1) + M_a(x_2),$$

$$M_a(\alpha x) = a\alpha x = \alpha M_a(x).$$

En realidad todas las aplicaciones lineales de  $K \rightarrow K$  son de este tipo pues si  $f : K \rightarrow K$  es una aplicación lineal y  $f(1) = a$  entonces para todo  $x \in K$  tenemos

$$f(x) = f(x1) = xf(1) = xa = M_a(f).$$

Si representamos la gráfica de  $M_a$  para el caso en que  $K = \mathbb{R}$  observamos que obtenemos una línea que pasa por el origen (Figura 2.1).

Vamos ahora a considerar aplicaciones lineales  $K \rightarrow K^n$ : Repitiendo el mismo argumento observamos que si  $f(1) = (a_1, \dots, a_n) = a$  entonces

$$f(x) = f(x1) = xf(1) = x(a_1, \dots, a_n) = M_a(x)$$

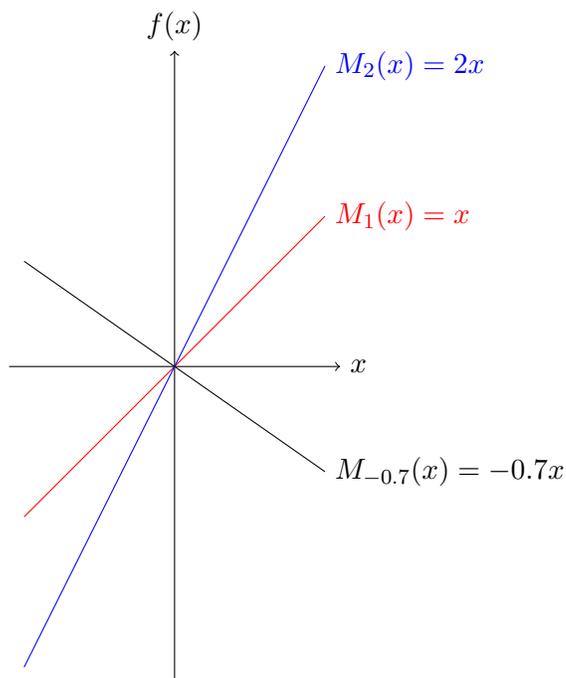


Figure 2.1:

Donde  $M_a(x) = ax$ . La única diferencia es que ahora  $a$  es un vector. Obsérvese que la imagen de  $M_a$  está formada por todos los elementos de la forma  $\alpha a$  con  $\alpha \in K$ . Si  $K = \mathbb{R}$  esta imagen es la recta (línea) que pasa por el origen y tiene dirección  $a$  (Figura 2.2). De ahí viene el nombre de aplicación lineal.

La siguiente proposición es evidente.

**Proposición 2.2.3** *Las siguientes condiciones son equivalentes para una aplicación  $f : V \rightarrow W$  entre dos espacios vectoriales sobre  $K$ :*

(1)  $f$  es lineal

(2) Para todo  $\alpha_1, \alpha_2 \in K$  y todo  $v_1, v_2 \in V$ ,

$$f(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 f(v_1) + \alpha_2 f(v_2).$$

(3) Para todo  $\alpha_1, \alpha_2, \dots, \alpha_n \in K$  y todo  $v_1, v_2, \dots, v_n \in V$ ,

$$f(\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n) = \alpha_1 f(v_1) + \alpha_2 f(v_2) + \dots + \alpha_n f(v_n).$$

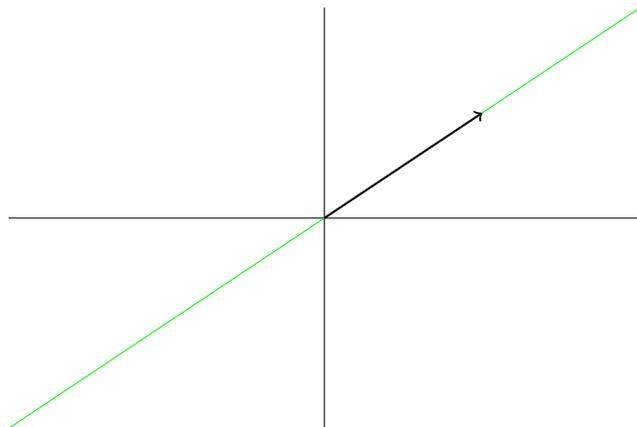


Figure 2.2:

**Ejemplo 2.2.4** Sea  $f : K^2 \rightarrow K^3$ , dada por  $f(x, y) = (2x, x - y, x + 3y)$ . Vamos a comprobar que  $f$  es una aplicación lineal. En efecto, si  $\alpha_1, \alpha_2 \in K$  y  $v_1 = (x_1, y_1), v_2 = (x_2, y_2) \in K^2$ , entonces

$$\begin{aligned}
 f(\alpha_1 v_1 + \alpha_2 v_2) &= f(\alpha_1(x_1, y_1) + \alpha_2(x_2, y_2)) \\
 &= f(\alpha_1 x_1 + \alpha_2 x_2, \alpha_1 y_1 + \alpha_2 y_2) \\
 &= (2(\alpha_1 x_1 + \alpha_2 x_2), (\alpha_1 x_1 + \alpha_2 x_2) - (\alpha_1 y_1 + \alpha_2 y_2), \\
 &\quad (\alpha_1 x_1 + \alpha_2 x_2) + 3(\alpha_1 y_1 + \alpha_2 y_2)) \\
 &= (2\alpha_1 x_1 + 2\alpha_2 x_2, \alpha_1(x_1 - y_1) + \alpha_2(x_2 - y_2), \\
 &\quad \alpha_1(x_1 + 3y_1) + \alpha_2(x_2 + 3y_2)) \\
 &= \alpha_1(2x_1, x_1 - y_1, x_1 + 3y_1) + \alpha_2(2x_2, x_2 - y_2, x_2 + 3y_2) \\
 &= \alpha_1 f(v_1) + \alpha_2 f(v_2).
 \end{aligned}$$

Más generalmente es fácil ver que toda aplicación  $f : K^2 \rightarrow K^3$  de la forma

$$f(x_1, x_2) = (a_{11}x_1 + a_{12}x_2, a_{21}x_1 + a_{22}x_2, a_{31}x_1 + a_{32}x_2)$$

es una aplicación lineal. En realidad esto se puede ver de forma muy fácil escribiendo los elementos de  $K^2$  y  $K^3$  como columnas. Si ponemos

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}$$

entonces

$$f \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \\ a_{31}x_1 + a_{32}x_2 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

En la última igualdad hemos usado la multiplicación de matrices que nos enseñaron en el instituto. Vamos a recordar esto. Obsérvese que para poder sumar dos matrices tienen que

tener el mismo número de filas y el mismo número de columnas. Lo que vamos a exigir a dos matrices para poder multiplicarlas es diferente. Ahora vamos a exigir que el número de columnas de la primera sea igual al de filas de la segunda y multiplicamos de la siguiente forma las matrices  $(a_{ij}) \in M_{m,t}(K)$  y  $(b_{ij}) \in M_{t,n}(K)$  para obtener la matriz  $(c_{ij}) \in M_{m,n}(K)$ :

$$(a_{ij})(b_{ij}) = (c_{ij}) \quad \text{con} \quad c_{ij} = \sum_{k=1}^t a_{ik}b_{kj} = a_{i1}b_{1j} + \cdots + a_{it}b_{tj}.$$

Un *vector fila* es una matriz con 1 fila y un *vector columna* es una matriz con una columna. Nosotros identificaremos los elementos de  $K^n$  tanto con los vectores fila como con los vectores columna de longitud  $n$ . Solo algunas veces nos interesará distinguir entre vectores fila y columna, sobre todo cuando estemos considerándolos como matrices.

Podemos pensar en una matriz  $n \times m$  como una lista de  $n$  vectores fila de longitud  $m$  puestos de arriba a abajo, o como una lista de  $m$  vectores columna de longitud  $n$  puestos de izquierda a derecha. Denotamos la fila  $i$ -ésima como  $a_i$  y la columna  $i$ -ésima como  $a_i$ :

$$(a_{ij}) = \begin{pmatrix} a_{1.} \\ \vdots \\ a_{n.} \end{pmatrix} = ( a_{.1} \quad \dots \quad a_{.m} ).$$

Llamamos entrada  $(i, j)$  de la matriz  $(a_{i,j})$  al elemento  $a_{ij}$  que está en la fila  $i$ -ésima y la columna  $j$ -ésima.

Por ejemplo, para poder multiplicar un vector fila por un vector columna ambos tendrán que tener la misma longitud y la regla de multiplicar es la siguiente:

$$(x_1, \dots, x_t) \begin{pmatrix} y_1 \\ \vdots \\ y_t \end{pmatrix} = x_1y_1 + \cdots + x_t y_t = \sum_{i=1}^t x_i y_i.$$

Obsérvese que esta multiplicación de vectores fila por vectores columnas es en realidad la esencia del producto en general pues pensando la primera matriz como una lista de  $n$  vectores fila de longitud  $t$  y el segundo como una lista de vectores columna, también de longitud  $t$  tenemos que es producto viene dado por

$$\begin{pmatrix} a_{1.} \\ \vdots \\ a_{n.} \end{pmatrix} ( b_{.1} \quad \dots \quad b_{.m} ) = \begin{pmatrix} a_{1.}b_{.1} & a_{1.}b_{.2} & \dots & a_{1.}b_{.m} \\ a_{2.}b_{.1} & a_{2.}b_{.2} & \dots & a_{2.}b_{.m} \\ \dots & \dots & \dots & \dots \\ a_{n.}b_{.1} & a_{n.}b_{.2} & \dots & a_{n.}b_{.m} \end{pmatrix},$$

O sea para multiplicar dos matrices multiplicamos los vectores fila del primero por los vectores columna del segundo y ponemos en la entrada  $(i, j)$  el resultado de multiplicar la fila  $i$  del primero por la columna  $j$  del segundo.

Si  $(a_{ij}) \in M_n(K)$  entonces llamamos diagonal de  $A$  a la parte de la matriz formada por las entradas  $a_{ii}$ . Una *matriz diagonal* es una matriz cuadrada en la que las entradas fuera de la

diagonal son todas 0. Las representamos así:

$$\text{diag}(a_1, \dots, a_n) = \begin{pmatrix} a_1 & 0 & \dots & 0 \\ 0 & a_2 & \dots & 0 \\ \dots & \dots & \ddots & \dots \\ 0 & 0 & \dots & a_n \end{pmatrix} = \begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{pmatrix}.$$

En la última forma hemos mostrado otro convenio que vamos a usar a menudo que consiste en no escribir partes de la matriz en las que solo aparezcan ceros.

Hay dos tipos de matrices que van a representar un papel importante. Un tipo está formado por las matrices que tiene todas las entradas iguales a 0, que representaremos como 0 o  $0_{n \times m}$  si queremos enfatizar el número de filas y columnas. El otro es el de las que llamamos *matrices identidad* que son matrices diagonales que tienen 1 en todas las entradas de la diagonal o sea las matrices cuadradas de la forma

$$I_n = \text{diag}(1, 1, \dots, 1) = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}.$$

A menudo simplemente escribiremos  $I$  sin especificar el tamaño de la matriz identidad.

La siguiente proposición recopila las propiedades fundamentales del producto de matrices:

**Proposición 2.2.5** *Se verifican las siguientes propiedades para todas las matrices  $A, B, C$  en las que para cada una de las propiedades suponemos que la matrices tienen las dimensiones adecuadas para que se puedan realizar las operaciones que aparecen:*

- (1) **Asociativa:**  $A(BC) = (AB)C$ .
- (2) **Distributiva:**  $A(B + C) = AB + AC$  y  $(B + C)A = BA + CA$ .
- (3) **Matriz Identidad:**  $IA = AI = A$ .
- (4) **Conmutativa con escalares:** Si  $\alpha \in K$  entonces  $\alpha A = A\alpha$ .

**Demostración.** Se trata de meras comprobaciones. Haremos la más engorrosa que es la asociativa y dejamos que el lector haga las otras como ejercicio. Para ello ponemos  $A = (a_{ij})$ ,  $B = (b_{ij})$  y  $C = (c_{ij})$ . Entonces ponemos  $BC = (d_{ij})$  y  $AB = (e_{ij})$ ,  $A(BC) = (x_{ij})$  y  $(AB)C = (y_{ij})$ . Tenemos que comprobar que  $x_{ij} = y_{ij}$ . Para ellos primero observamos que usando las propiedades conmutativa y distributiva en  $K$  se verifican las siguientes igualdades para cualesquiera elementos  $\alpha, \beta_i$  y  $\gamma_{ij}$  de  $K$ :

$$\alpha \sum_i \beta_i = \sum_i \alpha \beta_i \quad \text{y} \quad \sum_i \sum_j \gamma_{ij} = \sum_j \sum_i \gamma_{ij}. \quad (2.1)$$

Aplicando estas igualdades varias veces tenemos

$$\begin{aligned} x_{ij} &= \sum_k a_{ik} c_{kj} = \sum_k a_{ij} \left( \sum_l b_{kl} c_{lj} \right) = \sum_k \sum_l a_{ik} b_{kl} c_{lj} = \sum_l \sum_k a_{ik} b_{kl} c_{lj} \\ &= \sum_l \left( \sum_k a_{ik} b_{kl} \right) c_{lj} = \sum_l d_{il} c_{lj} = y_{ij}. \end{aligned}$$

■

El producto de matrices no es conmutativo. En primer lugar para que podamos realizar los productos de las matrices  $A = (a_{ij})_{n,m}$  y  $B = (b_{ij})_{k,l}$  en las dos formas posibles  $AB$  y  $BA$  es necesario que  $m = k$  y  $l = n$  pero incluso aunque esto pase para que  $AB$  y  $BA$  puedan ser iguales es necesario que además  $m = n$ , o sea que  $A$  y  $B$  sean cuadradas del mismo tamaño. Ni siquiera en ese caso está garantizada la igualdad de  $AB$  y  $BA$  como muestra el siguiente ejemplo:

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

Obsérvese que si  $n$  es un entero positivo entonces el conjunto  $M_n(K)$  es un anillo en el que el cero es la matriz  $0$ , el uno es la matriz identidad  $I_n$ , el opuesto de  $A$  es la matriz opuesta.

Diremos que una matriz  $A$  es *invertible* si y solo si existe otra matriz  $B$  tal que  $AB = BA = I_n$ . Obsérvese que todas las matrices invertibles han de ser cuadradas porque si  $A$  no es cuadrada no existe ninguna matriz  $B$  para la que los productos  $AB$  y  $BA$  tengan sentido. La siguiente proposición es fácil de demostrar:

**Proposición 2.2.6** Sean  $A, B \in M_n(K)$  y  $X, Y \in M_{n,m}(K)$ .

- (1) Si  $A$  es invertible y  $AX = AY$  entonces  $X = Y$ .
- (2) Si  $A$  y  $B$  son invertibles entonces  $AB$  es invertible y  $(AB)^{-1} = B^{-1}A^{-1}$ .

Todo este interludio sobre matrices tiene la intención de introducir la siguiente proposición en la que miramos los elementos de  $K^n$  y  $K^m$  como vectores columna:

**Proposición 2.2.7** Las aplicaciones lineales de  $K^n$  en  $K^m$  son las de la forma:

$$\begin{aligned} M_A : K^n &\rightarrow K^m \\ v &\mapsto M_A(v) = Av \end{aligned}$$

con  $A \in M_{m,n}(K)$ .

**Demostración.** En primer lugar, usando la Proposición 2.2.5, es fácil ver que  $A \in M_{m,n}(K)$  entonces  $M_A$  es una aplicación lineal:

$$M_A(\alpha u + \beta v) = A(\alpha u + \beta v) = A\alpha u + A\beta v = \alpha Au + \beta Av = \alpha M_A(u) + \beta M_A(v).$$

Recíprocamente, sea  $f : K^n \rightarrow K^m$  una aplicación lineal arbitraria y consideremos los siguientes elementos de  $K^n$ :

$$\begin{aligned} e_1 &= (1, 0, 0, \dots, 0) \\ e_2 &= (0, 1, 0, \dots, 0) \\ e_3 &= (0, 0, 1, \dots, 0) \\ &\vdots \\ e_n &= (0, 0, 0, \dots, 1) \end{aligned}$$

Nos referiremos al conjunto  $\{e_1, \dots, e_n\}$  como la *base canónica* de  $K^n$ . Supongamos que las imágenes por  $f$  de los elementos  $e_1, \dots, e_n$  de la base canónica de  $K^n$  son

$$f(e_1) = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, \dots, f(e_n) = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}.$$

Construimos una matriz  $A = (f(e_1), \dots, f(e_n)) = (a_{ij})$  poniendo estos vectores en columna. Entonces, viendo  $x = (x_1, \dots, x_n)$  como vector columna tenemos

$$\begin{aligned} f(x) &= f(x_1 e_1 + \dots + x_n e_n) = f(e_1)x_1 + \dots + f(e_n)x_n \\ &= \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} x_1 + \dots + \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} x_n = \begin{pmatrix} a_{11}x_1 + \dots + a_{1n}x_n \\ a_{21}x_1 + \dots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n \end{pmatrix} \\ &= (a_{ij}) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = Ax. \end{aligned}$$

Por tanto  $f = M_A$ . ■

Recordemos que si  $f : X \rightarrow Y$  y  $g : Y \rightarrow Z$  son dos aplicaciones, entonces se llama *aplicación composición* a la dada por

$$g \circ f : X \rightarrow Z, (g \circ f)(x) = g(f(x)).$$

La aplicación identidad  $1_X : X \rightarrow X$  de un conjunto  $X$  es la dada por  $1_X(x) = x$  para todo  $x$ . Si  $f : A \rightarrow B$  es una aplicación biyectiva entonces la *aplicación inversa* de  $f$  es la aplicación  $f^{-1} : B \rightarrow A$  que asocia cada elemento  $y \in Y$  con el único elemento  $x \in X$  tal que  $f(x) = y$ . Es decir

$$f(x) = y \Leftrightarrow f^{-1}(y) = x.$$

Está claro que se verifica

$$f \circ f^{-1} = 1_B \quad \text{y} \quad f^{-1} \circ f = 1_A.$$

De hecho, si  $f : A \rightarrow B$  es una aplicación, entonces  $f$  es invertible si y sólo si existe una aplicación  $g : B \rightarrow A$  tal que  $f \circ g = 1_B$  y  $g \circ f = 1_A$  y en tal caso  $g = f^{-1}$ .

Es fácil ver que la composición de dos aplicaciones lineales es otra aplicación lineal. Lo bueno de haber definido el producto de matrices como lo hemos definido es que se verifican las siguiente fórmula para matrices  $A$  y  $B$  multiplicables:

$$M_{AB} = M_A \circ M_B, \quad M_{I_n} = 1_{K^n}. \quad (2.2)$$

Además  $A$  es una matriz invertible si y solo si  $M_A$  es una aplicación lineal invertible y en tal caso

$$M_A^{-1} = M_{A^{-1}}. \quad (2.3)$$

**Ejemplo 2.2.8 (Rotaciones en el plano)** Sea  $0 \leq \alpha \leq 2\pi$  y sea  $R_\alpha : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  la *rotación* de ángulo  $\alpha$  en sentido antihorario alrededor del origen  $O = (0, 0)$ , es decir  $R_\alpha$  transforma cada punto del plano en el resultado de rotarlo un ángulo  $\alpha$  alrededor del origen como se muestra en la Figura 2.3: Como se trata de una aplicación lineal de  $\mathbb{R}^2$  en  $\mathbb{R}^2$  tendrá que tener la

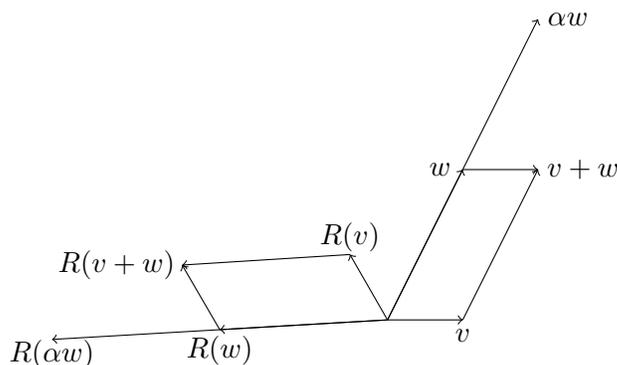


Figure 2.3:

forma  $M_A$  para alguna matriz  $A \in M_2(\mathbb{R})$ . La demostración de la Proposición 2.2.7 nos dice cómo obtener esta matriz  $A$ . Para ello basta poner en columna las imágenes de los vectores  $(1, 0)$  y  $(0, 1)$  que claramente son los vectores  $R(e_1) = (\cos \alpha, \sin \alpha)$  y  $R(e_2) = (-\sin \alpha, \cos \alpha)$  (Figura 2.4). Por tanto para obtener las coordenadas del punto que se obtiene al girar el punto  $(x, y)$  un ángulo  $\alpha$  basta multiplicar por la matriz  $\begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}$ .

## 2.3 Subespacios

En el resto del capítulo  $V$  será un espacio vectorial sobre  $K$ .

**Definición 2.3.1** Sea  $L$  una lista  $v_1, v_2, \dots, v_n$  de elementos de  $V$ . Una combinación lineal de  $L$  es una expresión del tipo

$$\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n$$

con  $\alpha_1, \alpha_2, \dots, \alpha_n \in K$ . Los  $\alpha_i$  se llaman coeficientes de la combinación lineal.

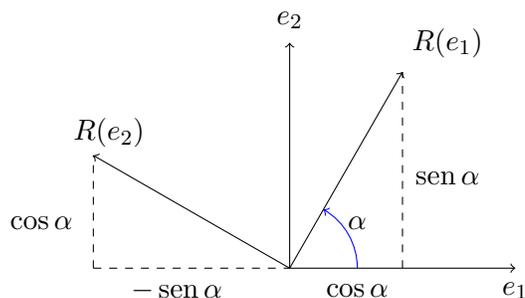


Figure 2.4:

Sea ahora  $S$  un subconjunto de  $V$ , no necesariamente finito. Un elemento  $v$  de  $V$  se dice que es combinación lineal de  $S$  si se puede expresar como combinación lineal de una lista finita de elementos de  $S$ .

El conjunto de las combinaciones lineales de  $S$  lo denotamos por  $\langle S \rangle$  y lo llamamos subespacio generado por  $S$ . En el caso en que  $S = \{v_1, \dots, v_n\}$  escribimos  $\langle v_1, \dots, v_n \rangle$  en lugar de  $\langle \{v_1, \dots, v_n\} \rangle$ .

Un conjunto o lista  $S$  de elementos de  $V$  se dice que es un sistema generador si  $\langle S \rangle = V$ , es decir, si todo elemento de  $V$  es combinación lineal de  $S$ .

Obsérvese que en la definición anterior hemos hablado a veces de listas y a veces de conjuntos. La diferencia es que en una lista los elementos pueden repetirse.

**Ejemplos 2.3.2** (1) Sea  $S = \{(1, 0, -1), (1, 1, 0), (0, 1, 1)\}$ . Entonces

$$2(1, 0, -1) + (1, 1, 0) - (0, 1, 1),$$

$$3(1, 0, -1) + 0(1, 1, 0) + 0(0, 1, 1)$$

y

$$0(1, 0, -1) + 3(1, 1, 0) - 3(0, 1, 1)$$

son tres combinaciones lineales de  $S$ . El resultado de hacer las operaciones es en los tres casos igual al vector  $(3, 0, -3)$ . Esto explica porque hemos sido tan puntillosos distinguiendo entre combinación lineal como expresión y vector que es combinación lineal.

(2) La aplicación constante  $f(x) = 1$  es una combinación lineal de las aplicaciones  $\sin^2$  y  $\cos^2$  pues recordemos la fórmula trigonométrica

$$\sin^2 x + \cos^2 x = 1.$$

Otra aplicación que se puede poner como combinación lineal de  $\sin^2$  y  $\cos^2$  es la aplicación  $f$  dada por  $f(x) = \cos(2x)$  ya que

$$\cos(2x) = \cos^2 x - \sin^2 x.$$

**Ejemplos 2.3.3** (1)  $\langle (1, 0, 0), (0, 1, 0) \rangle = \{x(1, 0, 0) + y(0, 1, 0) : x, y \in K\} = \{(x, y, 0) : x, y \in K\}$ .

- (2) En el espacio vectorial de los polinomios,  $\langle X, X^2 \rangle = \{a_1X + a_2X^2 : a_1, a_2 \in K\}$ , que es el conjunto formado por los polinomios de grado menor o igual que 2 con término independiente nulo.
- (3) Si  $K = \mathbb{R}$  y  $v$  es un elemento no nulo de  $\mathbb{R}^n$ , entonces  $\langle v \rangle$  es la recta que pasa por el origen con dirección  $v$ . Si ahora  $P$  es otro elemento de  $\mathbb{R}^n$  entonces

$$P + \langle v \rangle = \{P + \alpha v : \alpha \in \mathbb{R}\}$$

es la recta que pasa por  $P$  y es paralela a  $v$ . Aprovecharemos esto para trasladar a los espacios vectoriales el concepto de recta: Si  $P, v \in V$  con  $v \neq 0$  entonces el conjunto

$$P + \langle v \rangle = \{P + \alpha v : \alpha \in K\}$$

se llama recta que pasa por  $P$  y es paralela a  $V$ .

- (4) Si  $v$  y  $w$  son dos vectores no paralelos de  $\mathbb{R}^n$ , entonces

$$\langle v, w \rangle = \{\alpha v + \beta w : \alpha, \beta \in \mathbb{R}\}$$

que es un plano que pasa por el origen paralelo a los vectores  $v$  y  $w$ . Si además es  $P$  es otro elemento de  $\mathbb{R}^n$

$$P + \langle v, w \rangle = \{P + x : x \in \langle v, w \rangle\}$$

es el plano que pasa por  $P$  y es paralelo a  $\langle v, w \rangle$ .

Más adelante, veremos cómo generalizar esto para definir planos en espacios vectoriales arbitrarios.

- (5) Si  $s \in S$ , entonces  $0 = 0s$  y, por tanto, si  $S \neq \emptyset$  entonces  $0 \in \langle S \rangle$ . De hecho vamos a convenir que  $\langle \emptyset \rangle = \{0\}$ , con lo que  $0 \in \langle S \rangle$  para todo subconjunto  $S$  de  $V$ .

**Definición 2.3.4** Sea  $V$  un espacio vectorial sobre un cuerpo  $K$ . Un subconjunto  $W$  de  $V$  decimos que es un subespacio de  $V$  si las operaciones de suma y producto por escalares en  $V$  también son operaciones en  $W$  y con ellas  $W$  es un espacio vectorial.

Obsérvese que si  $W$  es un subespacio y  $w \in W$  entonces  $-w = (-1)w \in W$  por la Proposición 2.1.9. La siguiente proposición muestra diferentes alternativas para verificar si un subconjunto de un espacio vectorial es un subespacio.

**Proposición 2.3.5** Sea  $W$  un subconjunto de un espacio vectorial  $V$  sobre  $K$ . Entonces las siguientes condiciones son equivalentes:

- (1)  $W$  es un subespacio de  $V$ .
- (2)  $0 \in W$  y para todo  $x, y \in W$  y todo  $\alpha \in K$ ,  $\alpha x, x + y \in W$ .

(3)  $W \neq \emptyset$  y para todo  $x, y \in W$  y todo  $\alpha \in K$ ,  $\alpha x, x + y \in W$ .

(4)  $W \neq \emptyset$  y para todo  $x, y \in W$  y todo  $\alpha, \beta \in K$ ,  $\alpha x + \beta y \in W$ .

(5)  $W \neq \emptyset$  y las combinaciones lineales de elementos de  $W$  son elementos de  $W$ .

**Demostración.** (1) implica (5) Como  $W$  es subespacio al multiplicar elementos de  $W$  por escalares debemos obtener otros elementos de  $W$  y al sumarlos también. Por tanto todas las combinaciones lineales de elementos de  $W$  están en  $W$ . Además,  $W$  tendrá que tener un cero con lo que  $W \neq \emptyset$ .

(5) implica (4); (4) implica (3) y (3) implica (2) son obvios.

(2) implica (1) Finalmente, supongamos que  $W$  satisface (2). Entonces la suma de  $V$  y el producto por escalares proporciona una suma y un producto por escalares en  $W$  que tiene un 0. Además, si  $x \in W$  entonces  $-x = (-1)x \in W$  lo que nos asegura que todo elemento de  $W$  tiene opuesto. El resto de los axiomas de espacio vectorial son igualdades que se verifican para todos los elementos de  $V$  y por tanto también para los de  $W$ . Luego la suma de vectores de  $V$  y el producto por escalares define en  $W$  una estructura de espacio vectorial, con lo que  $W$  es un subespacio de  $V$ . ■

**Ejemplos 2.3.6** (1) El conjunto  $W$  de los elementos  $(x_1, x_2, x_3, x_4)$  de  $K^4$  que satisfacen la siguiente ecuación

$$x_1 + x_2 - 7x_4 = 0$$

es un subespacio de  $K^4$ . En efecto,  $W \neq \emptyset$  pues, por ejemplo,  $\bar{0} = (0, 0, 0, 0) \in W$ . Si  $x = (x_1, x_2, x_3, x_4)$  e  $y = (y_1, y_2, y_3, y_4)$  son dos elementos de  $W$ , entonces se verifican

$$x_1 + x_2 - 7x_4 = 0 \quad \text{y} \quad y_1 + y_2 - 7y_4 = 0.$$

Si  $\alpha, \beta \in K$ , entonces

$$(\alpha x_1 + \beta y_1) + (\alpha x_2 + \beta y_2) - 7(\alpha x_4 + \beta y_4) = \alpha(x_1 + x_2 - 7x_4) + \beta(y_1 + y_2 - 7y_4) = \alpha \cdot 0 + \beta \cdot 0 = 0.$$

De aquí deducimos que  $W$  es un subespacio.

(2) Más generalmente, dados  $a_1, a_2, \dots, a_n \in K$ , el subconjunto  $W$  de  $K^n$  formados por los elementos  $x = (x_1, x_2, \dots, x_n) \in K^n$  que satisfacen una ecuación del tipo

$$a_1x_1 + a_2x_2 + \dots + a_nx_n = 0$$

es un subespacio. Una ecuación de este tipo se dice que es una *ecuación homogénea* con coeficientes en  $K$ .

Más generalmente aún, las soluciones de un sistema de ecuaciones lineales homogéneas con coeficientes en  $K$  forman un subespacio de  $K^n$ .

(3) El conjunto de los polinomios de grado  $\leq n$  con coeficientes en  $K$  es un subespacio vectorial del espacio vectorial de todos los polinomios con coeficientes en  $K$ .

- (4) Sea  $W$  el conjunto de las funciones  $f : \mathbb{R} \rightarrow \mathbb{R}$  periódicas de periodo  $2\pi$ , es decir que satisfacen la siguiente igualdad para todo  $x \in \mathbb{R}$ :

$$f(x + 2\pi) = f(x).$$

Por ejemplo, las funciones trigonométricas  $\sin$ ,  $\cos$  y  $\tan x$  son elementos de  $W$ . Entonces  $W$  es un subespacio del espacio vectorial  $\mathbb{R}^{\mathbb{R}}$ .

- (5) En todo espacio vectorial hay dos subespacios al menos: El más grande,  $V$ , y el más pequeño,  $\{0\}$ . Para simplificar este último lo denotaremos por  $0$ .
- (6) Sea  $f : V \rightarrow W$  una aplicación lineal de espacios vectoriales sobre  $K$ . Recordemos que la imagen de  $f$  es el conjunto

$$f(V) = \{f(v) : v \in V\}$$

Vemos que  $f(V)$  es un subespacio de  $W$  comprobando que si  $\alpha_1, \alpha_2 \in K$  y  $v_1, v_2 \in V$  entonces

$$\alpha_1 f(v_1) + \alpha_2 f(v_2) = f(\alpha_1 v_1 + \alpha_2 v_2) \in f(V).$$

Más generalmente si  $V_1$  es un subespacio de  $V$  entonces  $f(V_1) = \{f(v) : v \in V_1\}$  es un subespacio de  $W$ . La comprobación es similar.

Por otro lado el siguiente conjunto que se llama *núcleo* de  $f$  es un subespacio de  $V$ :

$$N(f) = \{v \in V : f(v) = 0\}$$

En efecto, si  $v_1, v_2 \in N(f)$  entonces

$$f(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 f(v_1) + \alpha_2 f(v_2) = \alpha_1 0 + \alpha_2 0 = 0$$

y por tanto  $\alpha_1 v_1 + \alpha_2 v_2 \in N(f)$ .

**Proposición 2.3.7** *Sea  $S$  un subconjunto de un espacio vectorial  $V$ . Entonces  $\langle S \rangle$  es el menor (respeto de la inclusión) subespacio de  $V$  que contiene a  $S$ , es decir se verifican las siguientes condiciones:*

- (1)  $S \subseteq \langle S \rangle$ .
- (2)  $\langle S \rangle$  es un subespacio de  $V$ .
- (3) Si  $W$  es un subespacio de  $V$  que contiene a  $S$ , entonces también contiene a  $\langle S \rangle$ .

**Demostración.** (1) es evidente pues si  $s \in S$ , entonces  $s = 1s \in \langle S \rangle$ .

(2)  $0 \in \langle S \rangle$  (Ejemplos 2.3.3) y obviamente una combinación lineal de combinaciones lineales de  $S$  es otra combinación lineal de  $S$ . Por la Proposición 2.3.5, esto prueba que  $\langle S \rangle$  es un subespacio de  $V$ .

- (3) Es una consecuencia inmediata de la Proposición 2.3.5. ■

**Ejemplos 2.3.8** (1) Los conjuntos  $\{(1, 0), (0, 1)\}$  y  $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$  son conjuntos generadores de  $K^2$  y  $K^3$  respectivamente. Más generalmente la base canónica  $e_1, e_2, \dots, e_n$  de  $K^n$  que ya apareció en la demostración de la Proposición 2.2.7, es un sistema generador de  $K^n$ :

- (2) Otro sistema generador de  $K^3$  es  $\{(1, 0, 1), (1, 1, 0), (1, 1, 1)\}$ . Para ver esto intentemos escribir un elemento genérico  $(x, y, z)$  como combinación lineal de estos tres vectores, es decir intentemos encontrar escalares  $\alpha, \beta, \gamma$  tal que

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

Haciendo operaciones nos encontramos con un sistema de ecuaciones lineales como el siguiente.

$$\begin{aligned} \alpha + \beta + \gamma &= x \\ \beta + \gamma &= y \\ \alpha + \gamma &= z \end{aligned} \tag{2.4}$$

donde las incógnitas son  $\alpha, \beta$  y  $\gamma$ . Resolviendo el sistema nos encontramos con la siguiente solución:

$$\begin{aligned} \alpha &= x - y \\ \beta &= x - z \\ \gamma &= -x + y + z \end{aligned}$$

Efectivamente, podemos comprobar que

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

y esto prueba que el conjunto dado es un sistema generador.

Otra forma de haberlo hecho habría sido considerando la matriz del sistema (2.4):

$$\left( \begin{array}{ccc|c} 1 & 1 & 1 & x \\ 0 & 1 & 1 & y \\ 1 & 0 & 1 & z \end{array} \right)$$

Obsérvese que la matriz de coeficientes tiene determinante 1. Por tanto, se trata de un sistema de Cramer, luego tiene solución.

- (3) El conjunto  $\{1, X, X^2\}$  es un sistema generador del conjunto de los polinomios de grado menor o igual que 2 con coeficientes en  $K$ . Más generalmente  $\{1, X, X^2, \dots, X^n\}$  es un sistema generador del conjunto de los polinomios de grado  $\leq n$  con coeficientes en  $K$ ; y  $\{1, X, X^2, \dots\}$  es un sistema generador del conjunto de todos los polinomios con coeficientes en  $K$ . Nos referiremos a este conjunto como *base canónica*.
- (4) En realidad todo subconjunto  $S$  de  $V$  es un sistema generador de algún espacio vectorial. En concreto del subespacio  $\langle S \rangle$  generado por él. Por ejemplo,  $\emptyset$  es un sistema generador del subespacio  $0$  y  $\{v\}$  es un sistema generador de la recta con dirección  $v$  que contiene a  $0$ .

Como consecuencia inmediata de la Proposición 2.3.7 se tienen los dos siguientes corolarios.

**Corolario 2.3.9** Sean  $S$  y  $S'$  dos subconjuntos de un espacio vectorial  $V$ . Entonces

- (1)  $\langle S \rangle \subseteq \langle S' \rangle$  si y sólo si  $S \subseteq \langle S' \rangle$ , es decir, si todos los elementos de  $S$  son combinación lineal de los elementos de  $S'$ .
- (2)  $\langle S \rangle = \langle S' \rangle$  si y sólo si  $S \subseteq \langle S' \rangle$  y  $S' \subseteq \langle S \rangle$ . Es decir, dos subconjuntos de  $V$  generan el mismo subespacio si y sólo si los elementos de cada uno de ellos son combinación lineal de los elementos del otro.
- (3) Supongamos que  $S$  es un sistema generador de  $V$ . Entonces  $S'$  es sistema generador de  $V$  si y sólo si todo elemento de  $S$  es combinación lineal de los elementos de  $S'$ .
- (4) Si  $S' = S \cup \{v\}$ , entonces  $\langle S' \rangle = \langle S \rangle$  si y sólo si  $v \in \langle S \rangle$ .

**Ejemplo 2.3.10** El conjunto  $\{(1, 2), (2, 1), (3, -3)\}$  genera el mismo subespacio que  $\{(1, 2), (2, 1)\}$  pues  $3(2, 1) - 3(1, 2) = (3, -3)$ . Si además  $3 \neq 0$  entonces  $3^{-1}(2(2, 1) - (1, 2)) = (1, 0)$  y  $3^{-1}(2(1, 2) - (1, 2)) = (0, 1)$  con lo que  $\{(1, 2), (2, 1)\}$  es un sistema generador de  $K^2$ .

Vamos a explicar por qué hemos puesto la hipótesis  $3 \neq 0$ . La idea es que nadie nos ha dicho que nuestro cuerpo contenga el número 3. Sin embargo podemos suponer que se refiere a  $1 + 1 + 1$ , donde 1 es el uno de  $K$ . Pero en  $\mathbb{Z}_3$  se tiene que  $1 + 1 + 1 = 0$ . En general, si  $n \in \mathbb{N}$  podemos suponer que  $n$  representa el elemento que se obtiene al sumar  $n$  veces el uno de  $K$ . Pero como hemos visto no siempre los números naturales representan distintos elementos de  $K$ . Por un momento vamos a suponer que  $K = \mathbb{Z}_3$  y vamos a calcular  $\langle (1, 2), (2, 1) \rangle$  en ese caso. Como  $K$  solo tiene 3 elementos este subespacio tiene a lo sumo 9 elementos. Pero en realidad en este caso  $(1, 2) + (2, 1) = (0, 0)$ , con lo que en este caso  $\langle (1, 2), (2, 1) \rangle = \langle (1, 2) \rangle = \{(0, 0), (1, 2), (2, 1)\}$ .

**Convenio:** Para evitar las complicaciones que hemos visto en el Ejemplo 2.3.10 a partir de ahora supondremos que  $K$  contiene todos los números naturales, y por tanto todos los números enteros como elementos distintos. Por tanto,  $K$  contiene al cuerpo  $\mathbb{Q}$  de los números racionales. En realidad solo nos hace falta para simplificar algunas cosas mínimas en las cuentas. De hecho todos los resultados que veremos son ciertos sin esta hipótesis.

**Corolario 2.3.11** Sea  $S = \{v_1, v_2, \dots, v_n\}$  una lista de elementos de un espacio vectorial y sea  $S'$  una lista obtenida a partir de  $S$  mediante uno de los siguientes cambios.

- (1)  $S'$  se obtiene reordenando los elementos de  $S$ . Por ejemplo  $S' = \{v_2, v_1, v_3, \dots, v_n\}$ .
- (2)  $S'$  se obtiene multiplicando uno de los elementos de  $S$  por un escalar no nulo y dejando los demás igual. Por ejemplo,  $S' = \{\alpha v_1, v_2, v_3, \dots, v_n\}$  con  $\alpha \neq 0$ .
- (3)  $S'$  se obtiene sumando a un elemento de  $S$  el resultado de multiplicar otro por un escalar. Por ejemplo,  $S' = \{v_1 + \alpha v_2, v_2, v_3, \dots, v_n\}$ .

Entonces  $\langle S' \rangle = \langle S \rangle$ . Por tanto  $S$  es un sistema generador precisamente si lo es  $S'$ .

La primera de las construcciones del corolario anterior justifica hablar de que un conjunto de vectores sea un sistema generador o de que unos vectores formen un sistema generador, entendiendo que eso quiere decir que lo es la lista formada por ellos, sin repetición y en cualquiera de los órdenes posibles será un sistema generador.

## 2.4 Dependencia e independencia lineal

Obsérvese que dos vectores  $v_1$  y  $v_2$  de  $\mathbb{R}^n$  son paralelos si uno de ellos es el resultado de multiplicar el otro por un escalar. Todos los vectores de  $\mathbb{R}$  son paralelos mientras que en  $\mathbb{R}^2$  podemos encontrar dos vectores no paralelos  $v_1$  y  $v_2$ . Incluso podemos encontrar un tercer vector que no sea paralelo a ninguno de los dos anteriores, pero este tercer vector en realidad es combinación lineal de  $v_1$  y  $v_2$ . Esto sugiere la siguiente definición.

**Definición 2.4.1** *Sea  $S$  una lista de elementos de un espacio vectorial  $V$  sobre  $K$ . Decimos que  $S$  es linealmente dependiente o ligado si uno de los vectores es combinación lineal de los demás y decimos que es linealmente independiente o libre en caso contrario.*

*Decimos que  $S$  es una base de  $V$  si es linealmente independiente y sistema generador de  $V$ .*

**Ejemplos 2.4.2** (1) La lista  $S = \{(1, 0, -1), (1, 1, 0), (0, 1, 1)\}$  es linealmente dependiente pues

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

(2) Las listas  $\{(1, 0), (0, 1)\}$  y  $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$  de  $K^2$  y  $K^3$  son evidentemente linealmente independientes. Más generalmente la base canónica  $\{e_1, e_2, \dots, e_n\}$  de  $K^n$  es linealmente independiente.

(3) Supongamos que  $S$  tiene un sólo elemento  $v$ . En tal caso entendemos que el conjunto formado por “los demás” es el conjunto vacío  $\emptyset$ . De acuerdo con el convenio que hemos establecido en el Ejemplo 2.3.3.(4), el vector  $v$  es combinación lineal de  $\emptyset$  precisamente si  $v = 0$ . Por tanto  $\{v\}$  es linealmente dependiente si y solo si  $v = 0$ .

(4) Si una lista contiene el vector  $0$ , entonces es linealmente dependiente pues el vector  $0$  es combinación lineal de cualquier lista de vectores, incluso de la lista vacía.

La siguiente proposición recoge algunas propiedades obvias de los conceptos de dependencia e independencia lineal.

**Proposición 2.4.3** *Sean  $S \subseteq T$  dos subconjuntos de un espacio vectorial  $V$ .*

- (1) *Si  $S$  tiene el vector  $0$  entonces  $S$  es linealmente dependiente.*
- (2) *Si  $S$  tiene un vector repetido entonces  $S$  es linealmente dependiente.*
- (3) *Si  $S$  es linealmente dependiente entonces  $T$  es linealmente dependiente.*
- (4) *Si  $T$  es linealmente independiente entonces  $S$  es linealmente independiente.*

Observemos que en la Definición 2.4.1 hemos hablado de listas y no de conjuntos. Recordemos que la diferencia entre una lista y un conjunto es que en una lista los elementos se pueden repetir mientras que eso no pasa en un conjunto. Por ejemplo  $\{0, 1, 1\}$  es una lista pero no un conjunto. Si una lista de vectores tiene un elemento repetido, entonces es linealmente dependiente pues  $x = 1x$ .

Otra diferencia entre una lista y un conjunto es que el orden en que se escriban los elementos del conjunto no importa pero sí que importa en una lista. Por ejemplo  $\{1, 2\}$  y  $\{2, 1\}$  representan el mismo conjunto pero dos listas diferentes. Sin embargo en lo que respecta a dependencia lineal esto no es un problema pues dos listas con los mismos elementos en diferente orden son ambas sistemas linealmente dependientes o ambas sistemas linealmente independientes.

En realidad la Definición 2.4.1 es un poco incómoda pues para comprobar si una lista es un sistema linealmente dependiente o independiente, supuestamente hay que ver si cada uno de los vectores de la lista es combinación lineal o no de los demás. Vamos a explicar un método mejor para ver si una lista de vectores es linealmente dependiente o independiente. La idea básica nos la proporciona el primer ejemplo de la lista anterior. Vimos que  $S = \{(1, 0, -1), (1, 1, 0), (0, 1, 1)\}$  es linealmente dependiente pues

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

De hecho esta lista ya apareció en el primero de los ejemplos de 2.3.2. En ese ejemplo veíamos que el vector  $(3, 0, -3)$  es combinación lineal de los elementos de esa lista de tres formas diferentes:

$$\begin{aligned} (3, 0, -3) &= 2(1, 0, -1) + (1, 1, 0) - (0, 1, 1) \\ &= 3(1, 0, -1) \\ &= 3(1, 1, 0) - 3(0, 1, 1) \end{aligned}$$

El único vector que es combinación lineal de cualquier lista de vectores es el vector 0. Una forma de escribirlo como combinación lineal de  $v_1, v_2, \dots, v_n$  es elegir todos los coeficientes 0:

$$0 = 0v_1 + 0v_2 + \dots + 0v_n.$$

Es decir para toda lista  $\{v_1, v_2, \dots, v_n\}$  la ecuación

$$0 = X_1v_1 + X_2v_2 + \dots + X_nv_n \tag{2.5}$$

siempre tiene al menos una solución  $X_1 = X_2 = \dots = X_n = 0$ .

**Proposición 2.4.4** *Las siguientes condiciones son equivalentes para una lista de vectores  $S = \{v_1, v_2, \dots, v_n\}$  de un espacio vectorial  $V$ :*

- (1)  $S$  es linealmente independiente.
- (2) La única forma de escribir 0 como combinación lineal de  $S$  es con todos los coeficientes 0, es decir, la única solución de la ecuación (2.5) es  $X_1 = X_2 = \dots = X_n = 0$ .
- (3) Si  $v \in \langle S \rangle$  entonces el conjunto de soluciones de la ecuación

$$v = X_1v_1 + X_2v_2 + \dots + X_nv_n \tag{2.6}$$

tiene un único elemento.

**Demostración.** (1) implica (2) Si (2) es falso, entonces existe una solución  $(\alpha_1, \alpha_2, \dots, \alpha_n)$  de la ecuación (2.5), con algún  $\alpha_i \neq 0$ . Reordenando los  $v_n$  si es necesario podemos suponer que  $\alpha_1 \neq 0$ . Entonces

$$v_1 = -\frac{\alpha_2}{\alpha_1}v_2 - \dots - \frac{\alpha_n}{\alpha_1}v_n$$

y por tanto  $S$  es linealmente dependiente.

Este método de demostración se llama Método de los Contrarrecíprocos. Para demostrar que una afirmación  $P$  implica otra  $Q$ , se demuestra que la negación de  $Q$  implica la negación de  $P$ .

(2) implica (3). Volvemos a dar una demostración por el Método de los Contrarrecíprocos, es decir supongamos que (3) es falso, lo que quiere decir que existe  $v \in \langle S \rangle$  para el que la ecuación (2.6) tiene dos soluciones diferentes. Sean  $(\alpha_1, \alpha_2, \dots, \alpha_n)$  y  $(\beta_1, \beta_2, \dots, \beta_n)$  dos soluciones diferentes. Entonces

$$(\alpha_1 - \beta_1)v_1 + \dots + (\alpha_n - \beta_n)v_n = (\alpha_1v_1 + \dots + \alpha_nv_n) - (\beta_1v_1 + \dots + \beta_nv_n) = v - v = 0.$$

y  $(\alpha_1 - \beta_1, \alpha_2 - \beta_2, \dots, \alpha_n - \beta_n) \neq (0, 0, \dots, 0)$  por hipótesis.

(3) implica (1) Una vez más utilizamos el Método de los Contrarrecíprocos. Si  $S$  es linealmente dependiente, entonces uno de sus elementos es combinación lineal de los demás. Reordenándolos si es necesario podemos suponer que es  $v_1$ . Pongamos

$$v_1 = \alpha_2v_2 + \alpha_3v_3 + \dots + \alpha_nv_n$$

Entonces  $(1, 0, 0, \dots, 0)$  y  $(0, \alpha_2, \alpha_3, \dots, \alpha_n)$  son dos soluciones diferentes de la ecuación (2.6) para  $v = v_1 \in \langle S \rangle$ . ■

**Ejemplos 2.4.5** (1) Los monomios  $1, X, X^2, X^3, \dots, X^n$  forman un conjunto linealmente independiente pues una combinación lineal de ellos es un polinomio  $\alpha_0 + \alpha_1X + \alpha_2X^2 + \dots + \alpha_nX^n$  que sólo puede ser 0 si todos los coeficientes son 0.

(2) Vamos a ver que las tres funciones reales trigonométricas  $\text{sen}, \text{cos}$  y  $\text{tan}$ , definidas en el intervalo  $(-\frac{\pi}{2}, \frac{\pi}{2})$  forman un conjunto linealmente independiente. Supongamos que

$$\alpha \text{sen} + \beta \text{cos} + \gamma \text{tan} = 0$$

con  $\alpha, \beta, \gamma \in \mathbb{R}$ . Eso significa que para todo número real  $x$  se tiene

$$\alpha \text{sen}(x) + \beta \text{cos}(x) + \gamma \text{tan}(x) = 0.$$

Por ejemplo, sustituyendo por  $x = 0, \frac{\pi}{6}$  y  $\frac{\pi}{3}$  obtenemos el siguiente sistema de ecuaciones lineales homogéneas:

$$\begin{aligned} \frac{1}{2}\alpha + \frac{\sqrt{3}}{2}\beta + \frac{1}{\sqrt{3}}\gamma &= 0 \\ \frac{\sqrt{3}}{2}\alpha + \frac{1}{2}\beta + \sqrt{3}\gamma &= 0 \end{aligned}$$

Resolviéndolo obtenemos que  $\alpha = \beta = \gamma = 0$ , lo que muestra que efectivamente  $\{\text{sen}, \text{cos}, \text{tan}\}$  es linealmente independiente. Otra forma de resolverlo es observar que la matriz de coeficientes tiene determinante diferente de cero, lo que implica que el sistema tiene una única solución que por supuesto será  $\alpha = \beta = \gamma = 0$ .

A menudo diremos que un conjunto de vectores es linealmente independientes o que unos vectores son linealmente independientes para entender que la lista formada por ellos, sin repetición y en cualquiera de los órdenes posibles es linealmente independiente. Esto está justificado por la primera de las construcciones de la siguiente proposición.

**Proposición 2.4.6** *Sea  $S = \{v_1, v_2, \dots, v_n\}$  una lista de elementos de un espacio vectorial y sea  $S'$  una lista obtenida a partir de  $S$  mediante uno de los siguientes cambios.*

- (1)  *$S'$  se obtiene reordenando los elementos de  $S$ . Por ejemplo  $S' = \{v_2, v_1, v_3, \dots, v_n\}$ .*
- (2)  *$S'$  se obtiene multiplicando uno de los elementos de  $S$  por un escalar no nulo y dejando los demás igual. Por ejemplo,  $S' = \{\alpha v_1, v_2, v_3, \dots, v_n\}$  con  $\alpha \neq 0$ .*
- (3)  *$S'$  se obtiene sumando a un elemento de  $S$  el resultado de multiplicar otro por un escalar. Por ejemplo,  $S' = \{v_1 + \alpha v_2, v_2, v_3, \dots, v_n\}$ .*

*Entonces  $S'$  es linealmente independiente (dependiente) si y sólo si lo es  $S$ .*

**Demostración.** Obsérvese que si  $S'$  se obtiene a partir de  $S$  mediante una de las tres operaciones citadas, entonces  $S$  se puede obtener a partir de  $S'$  mediante otra operación del mismo tipo. Más concretamente si la primera operación transforma  $S$  en  $S'$ , la misma operación transforma  $S'$  en  $S$ ; si la primera operación transforma  $S$  en  $S'$ , el escalar usado es  $\alpha$  multiplicado por el  $i$ -ésimo vector de  $S$ , entonces la operación usando el escalar  $\alpha^{-1}$  en el vector  $i$ -ésimo de  $S'$  transforma  $S'$  en  $S$ ; finalmente si la operación utilizada es la tercera con el escalar  $\alpha$  y los vectores de las posiciones  $i$  y  $j$  la operación del mismo tipo, con el escalar  $-\alpha$  con los vectores de  $S'$  en la misma posición, transforma  $S'$  en  $S$ .

Por tanto para demostrar el Teorema basta demostrar que si  $S$  es linealmente independiente y  $S'$  se ha obtenido de una de las tres formas dadas, entonces  $S'$  también es linealmente independiente. El caso en que  $S'$  se obtenga reordenando los elementos de  $S$  ya lo hemos visto en los párrafos que siguen a los Ejemplos 2.4.2.

Supongamos ahora que  $S'$  se obtenga mediante la segunda operación. Podemos suponer que  $S' = \{\alpha v_1, v_2, v_3, \dots, v_n\}$ , con  $\alpha \neq 0$ . Si

$$\beta_1(\alpha v_1) + \beta_2 v_2 + \dots + \beta_n v_n = 0,$$

entonces de la dependencia lineal de  $S$  se obtiene que

$$\beta_1 \alpha = \beta_2 = \dots = \beta_n = 0$$

y como  $\alpha \neq 0$ , entonces  $\beta_1 = 0$ . Por tanto  $S'$  es linealmente independiente.

Supongamos ahora que  $S'$  se ha obtenido mediante por el tercero de los métodos y podemos suponer que  $S' = \{v_1 + \alpha v_2, v_2, v_3, \dots, v_n\}$ . Si

$$\beta_1(v_1 + \alpha v_2) + \beta_2 v_2 + \dots + \beta_n v_n = 0,$$

entonces

$$\beta_1 v_1 + (\beta_1 \alpha + \beta_2) v_2 + \beta_3 v_3 + \dots + \beta_n v_n = 0.$$

y como  $S$  es linealmente independiente, entonces

$$\beta_1 = \beta_1\alpha + \beta_2 = \beta_3 = \dots = \beta_n = 0$$

y despejando se ve que también  $\beta_2 = 0$ . ■

**Ejemplo 2.4.7** Vamos a ver si el conjunto

$$S = \{(1, 2, -3, 0), (-2, 8, 7, 1), (-3, -6, 1, -5), (-3, 6, 2, -4)\}$$

es linealmente dependiente o independiente. Mediante sucesivas operaciones como las explicadas en la Proposición 2.4.6 vamos pasando de  $S$  a otros conjuntos que van a ser sistemas linealmente dependientes precisamente si lo es  $S$ .

$$\begin{array}{r}
 (1, 2, -3, 0) \\
 (-2, 8, 7, 1) \\
 (-3, -6, 1, -5) \\
 (-3, 6, 2, -4) \\
 \hline
 (1, 2, -3, 0) \\
 (0, 12, 1, 1) \quad \text{Sumando 2 veces la primera} \\
 (0, 0, -8, -5) \quad \text{Sumando 3 veces la primera} \\
 (0, 12, -7, -4) \quad \text{Sumando 3 veces la primera} \\
 \hline
 (1, 2, -3, 0) \\
 (0, 12, 1, 1) \\
 (0, 0, -8, -5) \\
 (0, 0, -8, -5) \quad \text{Restando la segunda} \\
 \hline
 \end{array}$$

Luego  $S$  es linealmente independiente si y sólo lo es

$$S' = \{(1, 2, -3, 0), (0, 12, 1, 1), (0, 0, -8, -5), (0, 0, -8, -5)\}.$$

Pero  $S'$  es linealmente dependiente porque tiene dos vectores iguales. Concluimos que  $S$  es linealmente dependiente.

**Proposición 2.4.8** Si  $\{v_1, v_2, \dots, v_n\}$  es linealmente independiente y  $w \notin \langle v_1, v_2, \dots, v_n \rangle$  entonces  $\{v_1, v_2, \dots, v_n, w\}$  es linealmente independiente.

**Demostración.** Supongamos que  $\{v_1, v_2, \dots, v_n\}$  es linealmente independiente y  $w \notin \langle v_1, v_2, \dots, v_n \rangle$ . Supongamos que

$$\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n + \beta w = 0.$$

Si  $\beta \neq 0$ , entonces

$$w = \left(-\frac{\alpha_1}{\beta}\right) v_1 + \left(-\frac{\alpha_2}{\beta}\right) v_2 + \dots + \left(-\frac{\alpha_n}{\beta}\right) v_n \in \langle v_1, v_2, \dots, v_n \rangle,$$

en contra de la hipótesis. Por tanto  $\beta = 0$ , con lo que

$$\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n = 0$$

y, como  $\{v_1, v_2, \dots, v_n\}$  es linealmente independiente,  $\alpha_1 = \alpha_2 = \dots = \alpha_n = 0$ . ■

## 2.5 Problemas

- 2.1** (1) Comprobar que  $\mathbb{C}$  es un espacio vectorial sobre  $\mathbb{R}$ , donde la suma en  $\mathbb{C}$  y el producto de elementos de  $\mathbb{R}$  por elementos de  $\mathbb{C}$  son los habituales.
- (2) Comprobar que  $\mathbb{R}$  es un espacio vectorial sobre  $\mathbb{Q}$ , con suma y producto por escalares habituales.
- (3) Más generalmente, sean  $K$  y  $K'$  dos cuerpos de forma que  $K \subseteq K'$  y las sumas y productos de elementos de  $K$  de la misma forma tanto si los consideramos como elementos de  $K$  como de  $K'$ . Demostrar que  $K'$  es un espacio vectorial sobre  $K$  donde para la suma de vectores y productos de escalares por vectores se usan la suma y multiplicación del cuerpo  $K'$ .

**2.2** Vamos a considerar polinomios con monomios de grado impar. Es decir, sea  $X$  una indeterminada y sea  $V$  el conjunto de las expresiones  $A = \sum_{i \in \mathbb{Z}} a_i X^i$  donde cada  $a_i \in K$  y existen enteros  $n \leq m$  tales que  $a_i = 0$  si  $i < n$  ó  $j = m$ . O sea

$$A = \sum_{i=n}^m a_i X^i.$$

Definimos la suma y el producto por escalares como

$$\left(\sum_{i \in \mathbb{Z}} a_i X^i\right) + \left(\sum_{i \in \mathbb{Z}} b_i X^i\right) = \sum_{i \in \mathbb{Z}} (a_i + b_i) X^i, \quad \alpha \sum_{i \in \mathbb{Z}} a_i X^i = \sum_{i \in \mathbb{Z}} (\alpha a_i) X^i.$$

Demostrar que  $V$  es un espacio vectorial.

**2.3** Consideremos el conjunto  $M_{\mathbb{N}}(K)$  de las matrices con infinitas filas y columnas  $(a_{ij})_{i,j \in \mathbb{N}}$  con  $a_{ij} \in K$  para todo  $i, j$ . Definir una suma y un producto por escalares que convierta este conjunto en un espacio vectorial sobre  $K$ . ¿Se te ocurre como hacer algo similar cambiando  $\mathbb{N}$  por un conjunto arbitrario?

**2.4** Estudiar cuáles de las siguientes aplicaciones son lineales:

- (1)  $f : K^3 \rightarrow K^3$  dada por  $f(x, y, z) = (x + 1, y + 2, z + 3)$ .
- (2)  $f : K^3 \rightarrow K^2$  dada por  $f(x, y, z) = (xy, z)$ .
- (3)  $f : K^2 \rightarrow K^3$  dada por  $f(x, y) = \left(\frac{x}{2}, \frac{x+y}{2}, \frac{x-y}{2}\right)$ .
- (4)  $f : K^2 \rightarrow K$  dada por  $f(x, y) = x + y^2$ .

**2.5** Decidir cuáles de los siguientes subconjuntos de  $K^3$  son sistemas generadores y cuáles son linealmente independientes.

- (1)  $\{(1, -1, 9), (2, -1, 0)\}$

- (2)  $\{(1, 2, 0), (1, 2, -1), (0, 1, -1)\}$ .  
 (3)  $\{(1, -1, 1), (1, 0, -1), (2, \frac{1}{2}, -3), (1, 2, -5)\}$ .

**2.6** Decidir cuáles de los siguientes subconjuntos de  $K^4$  son subespacios y en cada caso razona la conclusión:

- (1)  $\{(x_1, x_2, 0, x_3) : x_1, x_2, x_3 \in K\}$ .  
 (2)  $\{(x_1, 1, x_2, x_3) : x_1, x_2, x_3 \in K\}$ .  
 (3)  $\{(x_1, x_2, x_1^2, x_2^2) : x_1, x_2 \in K\}$ .  
 (4)  $\{(x_1, x_2, x_3, x_4) : x_1 + 2x_2 + 3x_3 + 4x_4 = 0 \in K\}$ .  
 (5)  $\{(x_1, x_2, x_3, x_4) : x_1 + 2x_2 + 3x_3 + 4x_4 = 1 \in K\}$ .

**2.7** Decidir cuáles de los siguientes subconjuntos de  $K[X]$  son subespacios y en cada caso razona la conclusión:

- (1) Para  $n \in \mathbb{N}$ ,  $K_n[X] =$  conjunto de los polinomios de grado menor o igual que  $n$ .  
 (2) Polinomios de la forma  $a + bX + bX^2 + aX^3$  con  $a, b \in K$ .  
 (3)  $\{A \in K[X] : A(\alpha) = 0\}$ , con  $\alpha \in K$ .  
 (4)  $\{A \in K[X] : A(\alpha) = \beta\}$ , con  $\alpha, \beta \in K$ .

**2.8** ¿Cuáles de las siguientes aplicaciones son lineales?

- (1)  $E_\alpha : K[X] \rightarrow K$ , con  $E_\alpha(A) = A(\alpha)$ , donde  $\alpha$  es un elemento de  $\alpha$ .  
 (2)  $f_A : K \rightarrow K$ , con  $f_A(x) = A(x)$ , donde  $A \in K[X]$ .  
 (3)  $E_a : K^A \rightarrow K$ , con  $E_a(f) = f(a)$  para  $A$  un conjunto y  $a \in A$ .  
 (4)  $E_a : V^A \rightarrow V$ , con  $E_a(f) = f(a)$  para  $A$  un conjunto,  $V$  un espacio vectorial sobre  $K$  y  $a \in A$ .

**2.9** Para cada dos de los siguientes subespacios de  $K^4$ , decidir si uno está incluido en el otro y si son iguales.

- (1)  $W_1 = \langle (1, 1, 0, 0), (1, 0, -1, 0) \rangle$ .  
 (2)  $W_2 = \langle (2, 1, -1, 0), (0, 1, 1, 1), (-1, 2, 3, 0) \rangle$   
 (3)  $W_3$  el conjunto de soluciones del siguiente sistema de ecuaciones lineales

$$\begin{array}{rccccrc} X_1 & - & X_2 & + & X_3 & + & X_4 & = & 0 \\ X_1 & & & & + & X_3 & + & X_4 & = & 0 \\ & & X_2 & + & & & + & 2X_4 & = & 0 \end{array}$$

**2.10** Demostrar que el conjunto  $W$  de las funciones  $f : \mathbb{R} \rightarrow \mathbb{R}$  diferenciables dos veces que satisfacen la siguiente ecuación diferencial

$$f''(x) = f(x)$$

es un subespacio de  $\mathbb{R}^{\mathbb{R}}$  y comprobar que  $\langle f_1, f_2 \rangle \subseteq W$ , donde

$$f_1(x) = e^x \quad \text{y} \quad f_2(x) = e^{-x}.$$

(De hecho se verifica la igualdad.)

**2.11** Sea  $\alpha$  un número real fijado. Demostrar que el conjunto  $W$  de las funciones  $f : \mathbb{R} \rightarrow \mathbb{R}$  diferenciables dos veces que satisfacen la siguiente ecuación diferencial

$$f''(x) + \alpha^2 f(x) = 0,$$

es un subespacio de  $\mathbb{R}^{\mathbb{R}}$  y comprobar que  $\langle f_1, f_2 \rangle \subseteq W$ , donde

$$f_1(x) = \text{sen}(\alpha x) \quad \text{y} \quad f_2(x) = \text{cos}(\alpha x).$$

(De hecho se verifica la igualdad.)

**2.12** Sea  $S = \{v_1, v_2, v_3\}$  una lista de elementos de un espacio vectorial  $V$  y sea  $T = \{2v_1 + v_2, 2v_2 + v_3, 2v_3 + v_1\}$ . Demostrar

- (1)  $S$  es un sistema generador si y sólo si lo es  $T$ .
- (2)  $S$  es linealmente independiente si y sólo si lo es  $T$ .

**2.13** Sea  $S = \{v_1, v_2, \dots, v_n\}$  un subconjunto de un espacio vectorial  $V$  y  $T = \{v_1, v_1 + 2v_2, v_1 + v_2 + 3v_3, \dots, v_1 + v_2 + v_3 + \dots + nv_n\}$ . Demostrar

- (1)  $S$  es un sistema generador si y sólo si lo es  $T$ .
- (2)  $S$  es linealmente independiente si y sólo si lo es  $T$ .

**2.14** Sea  $V$  el espacio vectorial de las funciones  $f : [0, 1] \rightarrow \mathbb{R}$ . Consideremos los siguientes elementos de  $V$ :

$$f_1(x) = \text{sen}^2(x), \quad f_2(x) = \text{cos}^2, \quad f_3(x) = \text{sen}(x) \text{cos}(x),$$

$$f_4(x) = 1, \quad f_5(x) = \text{sen}(2x), \quad f_6(x) = \text{cos}(2x).$$

Demostrar

- (1)  $\langle f_1, f_2, f_3 \rangle = \langle f_4, f_5, f_6 \rangle = \langle f_1, f_5, f_6 \rangle$ .
- (2)  $\{f_1, f_2, f_3\}$ ,  $\{f_4, f_5, f_6\}$  y  $\{f_1, f_5, f_6\}$  son los tres linealmente independientes.

**2.15** Sea  $A$  un subconjunto de  $\mathbb{R}$  y consideremos las funciones  $f_n : A \rightarrow \mathbb{R}$  dadas por  $f_n(x) = x^n$  como subconjunto del espacio vectorial  $\mathbb{R}^A$  del Ejemplo 2.1.7. Demostrar que el conjunto  $\{f_n : n \in \mathbb{N}\}$  es linealmente independiente si y solo si  $A$  es infinito. Indicación: Si  $A = \{a_1, \dots, a_n\}$  considera el polinomio  $(X - a_1)(X - a_2) \cdots (X - a_n)$ .

**2.16** ¿Cuáles de los siguientes polinomios pertenecen al subespacio  $W$ , del espacio vectorial  $K[X]$  de los polinomios con coeficientes en  $K$ , generado por los polinomios  $1 + X + X^2$  y  $1 + X^2 + X^4$ ?

- (a)  $X$ , (b)  $X - X^3$ , (c)  $X^4 - X$ , (d)  $X^5$ .

Dar un conjunto de ecuaciones que sirvan para determinar si un polinomio genérico  $a_0 + a_1X + a_2X^2 + a_3X^3 + a_4X^4$  está en  $W$ .

**2.17** Decidir cuáles de los siguientes subconjuntos de  $\mathbb{R}^2$  es un subespacio:

- (1) El conjunto de los puntos  $(x, y)$  que satisfacen la siguiente ecuación  $3x = 2y$ .
- (2) El conjunto de los puntos  $(x, y)$  que satisfacen la siguiente ecuación  $3x - 2y = 1$ .
- (3) El conjunto de los puntos del cuadrado  $[0, 1] \times [0, 1]$ , es decir los puntos  $(x, y)$  tales que  $0 \leq x, y \leq 1$ .
- (4) El conjunto de vectores paralelos al vector  $(1, -1)$ .
- (5) El conjunto de vectores paralelos al vector  $(0, 0)$ .
- (6) Los puntos de una circunferencia de radio 1, centrada en el origen, es decir los puntos  $(x, y)$  que satisfacen la ecuación  $x^2 + y^2 = 1$ .

**2.18** Decidir cuáles de los siguientes subconjuntos de  $K[X]$  son subespacios:

- (1) El conjunto de los polinomios con término independiente igual a 1.
- (2) El conjunto de los polinomios con término independiente igual a 0.
- (3) El conjunto de los polinomios  $P$  tales que  $P(3) = 0$ .
- (4) El conjunto de los polinomios  $P$  tales que  $P(-3) = 1$ .

**2.19** Decidir cuáles de los siguientes subconjuntos del espacio vectorial  $\mathbb{R}^{\mathbb{R}}$ , formado por las aplicaciones  $f : \mathbb{R} \rightarrow \mathbb{R}$ , son subespacios.

- (1) El conjunto de las aplicaciones continuas en 0.
- (2) El conjunto de las aplicaciones continuas en 1.
- (3) El conjunto de las aplicaciones continuas en  $\mathbb{R}$ .
- (4) El conjunto de las aplicaciones derivables en  $\mathbb{R}$ .

- (5) El conjunto de las aplicaciones  $f$  tales que  $f(0) = 1$ .
- (6) El conjunto de las aplicaciones  $f$  tales que  $f(0) = 0$ .
- (7) El conjunto de las aplicaciones  $f$  tales que  $f(0) = f(1)$ .
- (8) El conjunto de las aplicaciones  $f$  tales que  $f(0) = f(1)^2$ .
- (9) El conjunto de las aplicaciones derivables  $f$  en  $\mathbb{R}$  que satisfacen la siguiente ecuación diferencial

$$f''(x) - 3f'(x) + 2f(x) = 0$$

- (10) El conjunto de las aplicaciones derivables  $f$  en  $\mathbb{R}$  que satisfacen la siguiente ecuación diferencial

$$f''(x) + f'(x) + 2x = 0$$

## Capítulo 3

# Eliminación de Gauss, ecuaciones lineales y determinantes

### 3.1 El Método de Gauss-Jordan

En el Ejemplo 2.3.10 hemos visto que un conjunto  $S$  puede generar el mismo subespacio vectorial que un subconjunto  $T \subset S$  con  $T \neq S$ . En el Corolario 2.3.11 hemos visto una serie de operaciones que podemos hacer a un conjunto de vectores que no cambia el subespacio vectorial generado. También hemos visto en la Proposición 2.4.6 que las mismas operaciones no cambian el carácter de dependencia o independencia lineal de una lista. Esta es la base del conocido como *Método de Gauss-Jordan* o de *Eliminación de Gauss*, que nos va a servir para muchas cosas como ya veremos.

El Método de Gauss-Jordan se basa en aplicar las operaciones del Corolario 2.3.11 a las filas de una matriz  $A$  hasta conseguir otra matriz que satisfaga las condiciones de la siguiente definición:

**Definición 3.1.1** Sea  $A = (a_{ij}) \in M_{n \times m}(K)$  y sean enteros  $0 \leq r \leq n$  y  $1 \leq j_1 < j_2 < \dots < j_r \leq m$ . Decimos que  $(a_{ij})$  es escalonada de rango  $r$  con pivotes situados en las columnas  $j_1, \dots, j_r$  si se verifican las siguientes propiedades (ver la Figura 3.1):

(E1)  $a_{ij_i} = 1$  para todo  $i = 1, \dots, r$ . Nos referiremos a estos unos como pivotes. Obsérvese que los pivotes están situados como los peldaños de una escalera que desciende de izquierda a derecha. No todos los peldaños tienen por qué tener el mismo paso.

(E2) Las  $n - r$  filas inferiores son 0. O sea las filas sin pivote son cero.

(E3) Si  $i \leq r$  y  $j < j_i$  entonces  $a_{ij} = 0$ . O sea, en una fila con pivote todas las entradas a la izquierda del pivote son 0.

Si  $A$  cumple además la siguiente condición entonces decimos que la matriz es fuertemente escalonada.

(E4) Si  $i \leq r$  y  $j \neq i$  entonces  $a_{j_i} = 0$ . O sea, la única entrada no nula de una columna que tenga un pivote es precisamente el pivote.

$$\left( \begin{array}{cccccc} 1 & * & * & * & * & * \\ & & 1 & * & * & * \\ & & & 1 & * & * \\ & & & & & & \\ & & & & & & \end{array} \right) \quad \left( \begin{array}{cccccc} 1 & * & * & & * & * \\ & & & 1 & * & * \\ & & & & & 1 & * \\ & & & & & & \\ & & & & & & \end{array} \right)$$

Figure 3.1: Una matriz escalonada  $5 \times 7$  con pivotes en las columnas 2, 4 y 5 y una matriz fuertemente escalonada con pivotes en las columnas 1, 4 y 6. Las entradas vacías representan ceros.

**Definición 3.1.2** Llamamos operaciones elementales fila a las siguientes modificaciones de una matriz  $(a_{ij})$ :

- (1) Intercambiar dos filas.
- (2) Multiplicar una fila por un escalar diferente de 0.
- (3) Sumar a una fila el resultado de multiplicar otra por un escalar.

Las operaciones elementales columna se definen de la misma manera pero las operaciones se hacen con las columnas:

Para  $A$  una matriz vamos a usar la siguiente notación para las operaciones elementales.

- $E_{i,j}A$  = Resultado de intercambiar las filas  $i$  y  $j$  de  $A$ .
- $AE_{i,j}$  = Resultado de intercambiar las columnas  $i$  y  $j$  de  $A$ .
- $E_{\alpha \cdot i}A$  = Resultado de multiplicar por  $\alpha$  la fila  $i$ -ésima de  $A$ .
- $AE_{\alpha \cdot i}$  = Resultado de multiplicar por  $\alpha$  la columna  $i$ -ésima de  $A$ .
- $E_{i,\alpha,j}A$  = Resultado de sumar a la fila  $i$  la fila  $j$  multiplicada por  $\alpha$ .
- $AE_{i,\alpha,j}$  = Resultado de sumar a la columna  $j$  la columna  $i$  multiplicada por  $\alpha$ .

**Atención:** Los papeles de  $i$  y  $j$  en las dos últimas están intercambiados.

**Ejemplo 3.1.3** Si

$$A = \begin{pmatrix} 0 & 1 & 2 & 0 \\ -1 & -3 & i & -1 \\ 5 & 0 & 1 & 6 \end{pmatrix}$$

entonces

$$\begin{aligned} E_{1,3}A &= \begin{pmatrix} 5 & 0 & 1 & 6 \\ -1 & -3 & i & -1 \\ 0 & 1 & 2 & 0 \end{pmatrix} \\ AE_{3,2} &= \begin{pmatrix} 0 & -3 & 2 & 0 \\ -1 & 9 & i & -1 \\ 5 & 0 & 1 & 6 \end{pmatrix} \\ E_{1,i,2}A &= \begin{pmatrix} -i & 1-3i & 1 & -i \\ -1 & 9 & i & -1 \\ 5 & 0 & 1 & 6 \end{pmatrix} \end{aligned}$$

**Teorema 3.1.4 (Método de Gauss-Jordan)** *Toda matriz se puede convertir en una matriz fuertemente escalonada después de aplicarle un número finito de operaciones elementales fila.*

El proceso de convertir una matriz en una matriz escalonada mediante operaciones elementales fila se llama *Método de Gauss* y el de convertirla en una matriz fuertemente escalonada se llama *Método de Gauss-Jordan*.

Vamos a ilustrar el teorema convirtiendo la siguiente matriz en una matriz escalonada:

$$A = \begin{pmatrix} 0 & 3 & 3 & 0 \\ 1 & 1 & 2 & 2 \\ 2 & 5 & 7 & 4 \\ -1 & -1 & -2 & 7 \end{pmatrix} \quad (3.1)$$

Empezamos intercambiando las dos primeras filas para ver el primer 1 de la segunda fila como un pivote:

$$A_1 = E_{1,2}A = \begin{pmatrix} 1 & 1 & 2 & 2 \\ 0 & 3 & 3 & 0 \\ 2 & 5 & 7 & 4 \\ -1 & -1 & -2 & 7 \end{pmatrix}$$

Ahora ponemos ceros debajo de ese pivote sumando a cada fila la primera multiplicada por el opuesto de lo que esa fila tiene debajo del pivote. Es decir, a la segunda fila no le hacemos nada, a la tercera le restamos la primera multiplicada por 2 y a la última le sumamos la primera:

$$A_2 = E_{4,1,1}E_{3,-2,1}A_1 = \begin{pmatrix} 1 & 1 & 2 & 2 \\ 0 & 3 & 3 & 0 \\ 0 & 3 & 3 & 0 \\ 0 & 0 & 0 & 9 \end{pmatrix}$$

A partir de aquí nos olvidamos de la primera fila. Obtenemos un segundo pivote multiplicando la segunda fila por el inverso de su primera entrada diferente de 0, o sea por  $1/3$ . De esa forma obtenemos

$$A_3 = E_{\frac{1}{3},2}A_2 = \begin{pmatrix} 1 & 1 & 2 & 2 \\ 0 & 1 & 1 & 0 \\ 0 & 3 & 3 & 0 \\ 0 & 0 & 0 & 9 \end{pmatrix}$$

Ahora usamos ese nuevo pivote para obtener ceros debajo de él con el mismo procedimiento con el que conseguimos ceros debajo del primer pivote. O sea a la tercera fila le restamos la segunda multiplicada por 3 y tenemos:

$$A_4 = E_{3,-3,2}A_3 = \begin{pmatrix} 1 & 1 & 2 & 2 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 9 \end{pmatrix}$$

Ha aparecido una fila de ceros que la bajaremos para abajo intercambiándola por la última.

$$A_5 = E_{3,4}A_4 = \begin{pmatrix} 1 & 1 & 2 & 2 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 9 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Solo falta conseguir un último pivote en la tercera fila multiplicándola por  $1/9$ , o sea dividiendo sus entradas por 9 con lo que obtenemos una matriz escalonada de rango 3 con pivotes situados en las columnas 1, 2 y 4.

$$A_6 = E_{\frac{1}{9}}A_5 = \begin{pmatrix} 1 & 1 & 2 & 2 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Esto completa el Método de Gauss y continuamos con el de Gauss-Jordan para convertir la matriz escalonada en fuertemente escalonada. Para ello vamos usando los pivotes de derecha a izquierda para poner ceros por encima de ellos con el mismo procedimiento que hemos usado para poner ceros debajo. Comenzamos restando a la primera fila la tercera multiplicada por 2:

$$A_7 = E_{1,-2,3} = \begin{pmatrix} 1 & 1 & 2 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Después restamos la primera fila a la segunda y ya tendremos una matriz fuertemente escalonada:

$$A_8 = E_{1,-1,2} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

**Definición 3.1.5** Llamamos matrices elementales a las matrices que se obtienen al aplicar una operación elemental a la matriz identidad. O sea las matrices elementales en  $K^{n,n}$  son las matrices de los siguientes tipos con  $i \neq j$  y  $\alpha \in K$  y  $\alpha \neq 0$ :

$$E_{i,j} = E_{i,j}I = IE_{i,j} \begin{pmatrix} I_{i-1} & & & & \\ & & & & 1 \\ & & & & \\ & & I_{j-i-1} & & \\ & 1 & & & \\ & & & & I_{n-j} \end{pmatrix} \quad E_{\alpha,i} = E_{\alpha,i}I = IE_{\alpha,i} \begin{pmatrix} I_{i-1} & & & & \\ & & & & \\ & & & & \alpha \\ & & & & \\ & & & & I_{n-i} \end{pmatrix}$$

$$E_{i,\alpha,j} = E_{i,\alpha,j}I = IE_{i,\alpha,j} = \left( \begin{array}{c|c|c|c|c} I_{i-1} & & & & \\ \hline & 1 & & \alpha & \\ \hline & & I_{j-i-1} & & \\ \hline & & & 1 & \\ \hline & & & & I_{n-j} \end{array} \right)$$

En la anterior definición hemos usado otra notación que utilizaremos a menudo que consiste en construir una matriz pegando otras matrices más pequeñas que llamamos cajas. La separación entre las distintas cajas las marcamos por líneas horizontales y verticales.

Obsérvese que  $E_{i,j}$  es la matriz que se obtiene al intercambiar las filas  $i$  y  $j$ ;  $E_{\alpha,i}$  se obtiene multiplicando la fila  $i$  de la matriz identidad por  $\alpha$ ; y  $E_{i,\alpha,j}$  se obtiene sumando en la matriz identidad la fila  $j$  multiplicada por  $\alpha$  a la fila  $i$ . Además en las matrices  $E_{i,j}$  y  $E_{i,\alpha,j}$  que se han representado  $i < j$ . Para  $j < i$  tendrían la forma:

$$E_{i,j} = \left( \begin{array}{c|c|c|c|c} I_{j-1} & & & & \\ \hline & & & 1 & \\ \hline & & I_{i-j-1} & & \\ \hline & 1 & & & \\ \hline & & & & I_{n-i} \end{array} \right) \quad E_{i,\alpha,j} = \left( \begin{array}{c|c|c|c|c} I_{j-1} & & & & \\ \hline & 1 & & & \\ \hline & & I_{i-j-1} & & \\ \hline & \alpha & & 1 & \\ \hline & & & & I_{n-i} \end{array} \right)$$

La *transpuesta* de una matriz  $A = (a_{ij})_{n \times m}$  es la matriz  $A^T = (a_{ji})_{m \times n}$  es decir es la matriz en la que hemos convertido las filas en columnas y viceversa manteniendo el orden. Por ejemplo:

$$\begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix}^T = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} 2 & -1 & 3 \\ 0 & -1 & i \end{pmatrix}^T = \begin{pmatrix} 2 & 0 \\ -1 & -1 \\ 3 & i \end{pmatrix}.$$

Obsérvese que  $x$  es un vector fila entonces  $x^T$  es un vector columna. Sin embargo recuérdese que estamos identificando los elementos de  $K^n$  con vectores fila y con vectores columna de longitud  $n$  y solo ocasionalmente los distinguimos, cuando los consideremos como matrices. En tal caso si  $x$  es un vector fila, entonces  $x^T$  será el correspondiente vector columna.

Una matriz se dice que es *simétrica* si es igual a su transpuesta.

**Proposición 3.1.6** *Se cumplen las siguientes propiedades para dos matrices  $A$  y  $B$ :*

- (1) Si  $A \in M_{m,n}(K)$  entonces  $A^T \in M_{n,m}(K)$ .
- (2)  $(A^T)^T = A$ .
- (3)  $(A + B)^T = A^T + B^T$  y  $(AB)^T = B^T A^T$ .
- (4) Si  $A$  es invertible entonces  $A^T$  es invertible y  $(A^{-1})^T = (A^T)^{-1}$ .
- (5) Si  $A$  es simétrica entonces es cuadrada.
- (6) Las matrices elementales  $E_{i,j}$  y  $E_{\alpha,i}$  son simétricas y para el tercer tipo de matrices elementales tenemos  $E_{i,\alpha,j}^T = E_{j,\alpha,i}$ .

**Proposición 3.1.7** *El resultado de multiplicar por la izquierda una matriz elemental a una matriz  $A$  es igual al resultado de aplicar a  $A$  la misma operación elemental fila que se usó para obtener la matriz elemental.*

*El resultado de multiplicar por la derecha una matriz elemental a una matriz  $A$  es igual a aplicar a  $A$  la misma operación elemental columna que se aplicó para obtener la matriz elemental.*

En otras palabras la notación  $E_{i,j}A$ ,  $E_{\alpha,i}A$ ,  $E_{i,\alpha,j}A$  la podemos interpretar bien como una multiplicación de matrices o como el resultado de aplicar una operación elemental fila a  $A$  pues ambas cosas dan el mismo resultado. Análogamente,  $AE_{i,j}$ ,  $AE_{\alpha,i}$ ,  $AE_{i,\alpha,j}$  se pueden interpretar como producto de matrices o como el resultado de aplicar una operación elemental columna a  $A$ . Fíjate en que  $E_{ij}$  y  $E_{\alpha,i}$  son matrices simétricas. Sin embargo  $E_{i,\alpha,j}^T = E_{j,\alpha,i}$  y esta es la consecuencia del intercambio de papeles entre  $i$  y  $j$  en la tercera operación elemental fila y columna.

Concluimos esta sección con una primera aplicación del Métodos de Gauss-Jordan que nos permite decidir si una matriz es invertible y en su caso calcular la inversa por Eliminación de Gauss.

**Teorema 3.1.8** *Las siguientes propiedades son equivalentes para una matriz  $A$  con  $n$  filas.:*

- (1)  *$A$  es producto de matrices elementales.*
- (2)  *$A$  es invertible.*
- (3) *Cualquiera de las matrices escalonadas que se obtienen al aplicar el Método de Gauss a  $A$  tiene rango  $n$ .*
- (4) *La matriz fuertemente escalonada que se obtiene al aplicar el Método de Gauss-Jordan a  $A$  es la matriz identidad.*
- (5) *Al aplicar el Método de Gauss-Jordan a la matriz  $(A|I)$  se obtiene una matriz de la forma  $(I|U)$ .*

*En tal caso  $U = A^{-1}$ .*

**Demostración.** (1) implica (2). Todas las matrices elementales son invertibles pues  $E_{i,j}E_{j,i} = I$ ,  $E_{i,\alpha,j}E_{i,-\alpha,j} = I$  y si  $\alpha \neq 0$  entonces  $E_{\alpha,i}E_{\alpha^{-1},i} = I$ . Como el producto de matrices invertibles es invertible (Proposición 2.2.6) se tiene que todo producto de matrices elementales es invertible.

(2) implica (3) Supongamos que  $A$  es invertible. Como aplicar una operación elemental es lo mismo que multiplicar por una matriz elemental, el resultado de aplicar el Método de Gauss a  $A$  equivale a multiplicar  $A$  por la izquierda por una matriz  $U$  que es un producto de matrices elementales. Entonces  $UA$  es una matriz escalonada invertible. Si el rango de  $UA$  es menor que  $n$  entonces la última fila de  $UA$  es 0. Entonces la última fila de  $(UA)(UA)^{-1} = I$  también es 0 pero eso no es verdad. Por tanto, el rango de  $UA$  es  $n$ .

(3) implica (4). Supongamos que  $A$  verifica (3) y sea  $E$  una matriz fuertemente escalonada que se obtiene cuando aplicamos el Método de Gauss-Jordan a  $A$ . Entonces  $E$  tiene rango  $n$ ,

con lo que tiene  $n$  pivotes, que obviamente están en la diagonal. Como además  $E$  es fuertemente escalonada, debajo y encima de cada pivote solo hay ceros con lo que  $E$  es la matriz identidad.

(4) implica (5). Supongamos que la matriz identidad es el resultado de aplicar el Método de Gauss-Jordan a la matriz  $A$ . Como el resultado de hacer una operación elemental a una matriz es el resultado de multiplicarla por una matriz elemental se tiene que  $I = UA$  donde  $U$  es un producto de matrices elementales. Sea  $B = (A|I)$ . Como  $U$  es un producto de matrices elementales,  $UB$  es el resultado de aplicar el Método de Gauss a  $B$  pero  $UB = (UA|U) = (I|U)$ .

(5) implica (1) Supongamos que una matriz de la forma  $(I|U)$  es el resultado de aplicar el Método de Gauss a la matriz  $(A|I)$ . El mismo argumento del párrafo anterior muestra que  $(I|U) = V(A|I) = (VA, V)$  para una matriz  $V$  que es el producto de matrices elementales. Entonces  $U = V$  y  $UA = I$ . Como  $U$  es invertible,  $A$  es invertible y  $A^{-1} = U$ , un producto de matrices elementales. Como la inversa de una matriz elemental es elemental,  $A$  es también el producto de matrices elementales. ■

**Ejemplo 3.1.9** *Vamos a calcular la inversa de la siguiente matriz*

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

utilizando el comando `echelon` de Maxima que proporciona una forma escalonada obtenida aplicando el Método de Gauss a una matriz<sup>1</sup>. Para ello comenzamos construyendo la matriz  $(A|I)$ :

```
b:matrix([3,1,1,4,1,0,0,0],[0,1,3,4,0,1,0,0],[4,-1,2,4,0,0,1,0],[2,1,4,1,0,0,0,1])$
e:echelon(b);
```

$$\begin{pmatrix} 1 & \frac{1}{3} & \frac{1}{3} & \frac{4}{3} & \frac{1}{3} & 0 & 0 & 0 \\ 0 & 1 & 3 & 4 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \frac{24}{23} & -\frac{4}{23} & \frac{7}{23} & \frac{3}{23} & 0 \\ 0 & 0 & 0 & 1 & \frac{6}{125} & \frac{24}{125} & \frac{7}{125} & -\frac{23}{125} \end{pmatrix} \quad (e)$$

Por desgracia Maxima no tiene un comando para obtener la forma fuertemente escalonada pero lo podemos hacerlo usando Maxima en varios pasos:

```
e[1]:e[1]-4/3*e[4]$ e[2]:e[2]-4*e[4]$ e[3]:e[3]-24/23*e[4]$ e;
```

$$\begin{pmatrix} 1 & 0 & 0 & -\frac{4}{3} & \frac{3}{23} & -\frac{13}{25} & -\frac{2}{73} & \frac{28}{75} \\ 0 & 1 & 0 & -4 & \frac{36}{125} & -\frac{106}{125} & -\frac{83}{125} & \frac{112}{125} \\ 0 & 0 & 1 & -\frac{24}{23} & -\frac{788}{2875} & -\frac{277}{2875} & \frac{39}{2875} & \frac{48}{125} \\ 0 & 0 & 0 & 1 & \frac{6}{125} & \frac{24}{125} & \frac{7}{125} & -\frac{23}{125} \end{pmatrix}$$

<sup>1</sup>En la literatura en inglés una matriz escalonada se dice que tiene “echelon form”

$e[1]:e[1]-1/3*e[3]$   $e[2]:e[2]-3*e[3]$   $e$ ;

$$\begin{pmatrix} 1 & \frac{1}{3} & 0 & \frac{68}{69} & \frac{9}{12} & -\frac{7}{69} & -\frac{1}{23} & 0 \\ 0 & 1 & 0 & \frac{20}{23} & \frac{12}{23} & \frac{2}{23} & -\frac{9}{23} & 0 \\ 0 & 0 & 1 & \frac{24}{23} & -\frac{4}{23} & \frac{7}{23} & \frac{3}{23} & 0 \\ 0 & 0 & 0 & 1 & \frac{6}{125} & \frac{24}{125} & \frac{7}{125} & -\frac{23}{125} \end{pmatrix}$$

$e[1]:e[1]-1/3*e[2]$   $e$ ;

$$\begin{pmatrix} 1 & 0 & 0 & \frac{16}{23} & \frac{5}{23} & -\frac{3}{23} & \frac{2}{23} & 0 \\ 0 & 1 & 0 & \frac{20}{23} & \frac{12}{23} & \frac{2}{23} & -\frac{9}{23} & 0 \\ 0 & 0 & 1 & \frac{24}{23} & -\frac{4}{23} & \frac{7}{23} & \frac{3}{23} & 0 \\ 0 & 0 & 0 & 1 & \frac{6}{125} & \frac{24}{125} & \frac{7}{125} & -\frac{23}{125} \end{pmatrix}$$

Por tanto

$$A^{-1} = \frac{1}{125} \begin{pmatrix} 23 & -33 & 6 & 16 \\ 60 & -10 & -55 & 20 \\ -28 & 13 & 9 & 24 \\ 6 & 24 & 7 & -23 \end{pmatrix}$$

Lo que podemos comprobar multiplicando esta matriz por  $A$ :

$c:1/125*matrix([23,-33,6,16],[60,-10,-55,20],[-28,13,9,24],[6,24,7,-23])$   $a.c$ ;

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

## 3.2 Sistemas de ecuaciones lineales

Una *ecuación lineal* con coeficientes en  $K$  es una expresión del tipo

$$a_1X_1 + \cdots + a_nX_n = b$$

donde  $a_1, a_2, \dots, a_n, b \in K$  y  $X_1, \dots, X_n$  son indeterminadas. Una solución de la ecuación lineal anterior es un vector  $(x_1, \dots, x_n) \in K^n$  que cumple  $a_1x_1 + \cdots + a_nx_n = b$ .

Un *sistema de ecuaciones lineales* es una lista finita de ecuaciones lineales:

$$\left. \begin{array}{cccccc} a_{11}X_1 & + & \cdots & + & a_{1n}X_n & = & b_1 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m1}X_1 & + & \cdots & + & a_{mn}X_n & = & b_m \end{array} \right\}$$

Una solución de un sistema de ecuaciones lineales es un vector que es solución de todas las ecuaciones que lo forman.

Un sistema de ecuaciones lineales se dice que es *compatible* si tiene al menos una solución. En caso contrario se dice que es *incompatible*.

Un sistema de ecuaciones lineales compatible se dice que es *determinado* si tiene exactamente una solución. En caso contrario, o sea si tiene más de una solución, se dice que es *indeterminado*.

Dos sistemas de ecuaciones lineales son *equivalentes* si tienen el mismo conjunto de soluciones.

La siguiente matriz se llama *matriz del sistema de ecuaciones lineales*:

$$(A|b) = \left( \begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{array} \right).$$

La última columna de esta matriz se llama *columna de términos independientes* y el resto de la matriz *matriz de los coeficientes* de la matriz.

La siguiente proposición se sigue de forma inmediata de la definición de multiplicación de matrices:

**Proposición 3.2.1** *Sea  $(A|b) \in M_{m,n+1}(K)$  la matriz de un sistema de ecuaciones lineales y sean  $a_1, \dots, a_n$  las columnas de  $A$ . Entonces las siguientes condiciones son equivalentes para un elemento  $x \in K^n$ :*

- (1)  $x$  es solución del sistema con matriz  $(A|b)$ .
- (2)  $Ax = b$ .
- (3)  $b = x_1 a_{.1} + \cdots + x_n a_{.n}$ .

Por tanto podemos ver un sistema de ecuaciones lineales como una ecuación  $AX = b$  donde  $X$  representa un vector columna cuya longitud  $n$  es igual al número de incógnitas del sistema de ecuaciones inicial,  $A \in M_{m,n}(K)$  es la matriz de coeficientes que tiene tantas filas como el número de ecuaciones y tantas columnas como el número de incógnitas y  $b$  es el vector columna de términos independientes cuya longitud es el número de ecuaciones. También lo podemos ver como una ecuación  $x_1 a_{.1} + \cdots + x_n a_{.n} = b$ , con  $a_1, \dots, a_n, b \in K^m$  y resolverlo equivale a primero decidir si  $b \in \langle a_{.1}, \dots, a_{.n} \rangle$  y, en tal caso, encontrar todas las formas de escribir  $b$  como combinación lineal de  $a_{.1}, \dots, a_{.n}$ .

La proposición siguiente tiene demostración obvia:

**Proposición 3.2.2** *Si  $U$  es una matriz invertible entonces los sistemas  $AX = b$  y  $(UA)X = (Ub)$  tienen las mismas soluciones. En particular, una operación elemental en un sistema lo transforma en otro equivalente. Por tanto, todo sistema es equivalente a cualquiera de los que tienen como matriz de coeficientes la que resulta de aplicar el Método de Gauss a la matriz del sistema original.*

Resolver un sistema de ecuaciones lineales  $AX = b$  consiste en encontrar el conjunto de soluciones y por la proposición anterior si  $(E|c) = (e_{i,j}|c_j)$  es una de las matrices escalonadas que se obtienen al aplicar el Método de Gauss a  $(A|b)$  entonces el sistema inicial  $AX = b$  es equivalente al sistema  $EX = c$ . Pero un sistema así es muy fácil de resolver. En efecto, si la

última columna de  $(E|c)$  tiene un pivote, entonces una de las ecuaciones del sistema es  $0 = 1$  con lo que el sistema es incompatible. Supongamos por el contrario que la última columna de la matriz  $(E|c)$  no tiene ningún pivote y sean  $j_1 < j_2 < \dots < j_r$  las columnas en las que aparecen los pivotes. Entonces cada ecuación del sistema es de la forma

$$X_{j_i} + e_{i,j_i+1}X_{j_i+1} + \dots + a_{i,n}X_n = c_i$$

con lo que en una solución, cada una de las  $r$  coordenadas  $x_{j_i}$  están determinadas por las posteriores. De esta manera podemos elegir valores arbitrarios  $x_j$  para cada una de las  $n - r$  variables  $X_j$  con  $j \notin \{j_1, \dots, j_r\}$ . A partir de estos valores determinamos las otras coordenadas  $x_{j_1}, x_{j_2}, \dots, x_{j_r}$ , comenzando por la última:

$$x_{j_r} = c_r - (a_{r,j_r+1}x_{j_r+1} + \dots + a_{r,n}x_n);$$

después la penúltima  $x_{j_{r-1}}$ ; y en general cada una de las  $x_{j_i}$  en función de las posteriores con la siguiente fórmula:

$$x_{j_i} = c_i - (a_{i,j_i+1}x_{j_i+1} + \dots + a_{i,n}x_n).$$

De hecho si  $(E|c)$  es fuertemente escalonada el sistema con matriz  $(E|c)$  tiene la forma

$$X_{j_i} + \sum_{k \neq j_i} e_{i,k}X_k = c_i.$$

Por tanto podemos elegir las  $n - r$  variables  $X_k$  con  $k \notin \{j_1, \dots, j_r\}$  de forma arbitraria, y cada una de estas elecciones proporciona una solución.

**Ejemplo 3.2.3** Ilustremos esto resolviendo el siguiente sistema de ecuaciones:

$$\left. \begin{array}{rcl} & 3Y & + 3Z = 0 \\ X & + Y & + 2Z = 2 \\ 2X & + 5Y & + 7Z = 4 \\ X & - Y & - 2Z = 7 \end{array} \right\}$$

La matriz del sistema resulta ser precisamente la que aparece en (3.1) a la que ya le aplicamos el Método de Gauss y obtuvimos la siguiente matriz escalonada:

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

Como uno de los pivotes aparece en la última columna el sistema no tiene solución.

**Ejemplo 3.2.4** Vamos ahora a resolver el siguiente sistema

$$\left. \begin{array}{rcl} 2X_1 & - 6X_2 & + 2X_3 & & = 4 \\ & X_2 & - 3X_3 & + X_4 & = 2 \\ X_1 & - 4X_2 & + 4X_3 & - X_4 & = 0 \\ X_1 & - X_2 & & & - 2X_4 = 7 \end{array} \right\}$$

La matriz del sistema es

$$\left( \begin{array}{cccc|c} 2 & -6 & 2 & 0 & 4 \\ 0 & 1 & -3 & 1 & 2 \\ 1 & -4 & 4 & -1 & 0 \\ 1 & -1 & 0 & -2 & 7 \end{array} \right)$$

Con el comando `echelon` de Maxima obtenemos:

```
a:matrix([2,-6,2,0,4],[0,1,-3,1,2],[1,-4,4,-1,0],[1,-1,0,-2,7])$
e:echelon(a);
```

$$\left( \begin{array}{ccccc} 1 & -3 & 1 & 0 & 2 \\ 0 & 1 & -3 & 1 & 2 \\ 0 & 0 & 1 & -\frac{4}{5} & \frac{1}{5} \\ 0 & 0 & 0 & 0 & 0 \end{array} \right) \quad (e)$$

que nos proporciona una forma escalonada de la matriz del sistema:

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

Por tanto el sistema original es equivalente al siguiente:

$$\left. \begin{array}{rcl} X_1 - 3X_2 + X_3 & = & 2 \\ X_2 - 3X_3 + X_4 & = & 2 \\ X_3 - \frac{4}{5}X_4 & = & \frac{1}{5} \end{array} \right\}$$

Como los pivotes están en las tres primeras columnas, el sistema tiene solución y la cuarta coordenada  $x_4$  de una solución se puede elegir arbitrariamente y poner las otras en función de esa:

$$\begin{aligned} x_3 &= \frac{1}{5}(1 + 4x_4) \\ x_2 &= 2 + 3x_3 - x_4 = 2 + \frac{3}{5}(1 + 4x_4) - x_4 = \frac{1}{5}(13 + 7x_4) \\ x_1 &= 2 + 3x_2 - x_3 = 2 + \frac{3}{5}(13 + 7x_4) - \frac{1}{5}(1 + 4x_4) = \frac{1}{5}(48 + 17x_4) \end{aligned}$$

Por tanto las soluciones del sistema son los vectores de la forma

$$\left( \frac{1}{5}(48 + 17t), \frac{1}{5}(13 + 7t), \frac{1}{5}(1 + 4t), t \right).$$

Aunque también se podrían dar en la forma

$$(48 + 17t, 13 + 7t, 1 + 4t, 5t) = (48, 13, 1, 0) + (17, 7, 4, 5)t.$$

$$\begin{cases} x_1 = 48 & 17t \\ x_2 = 13 & + 7t \\ x_3 = 1 & + 4t \\ x_4 = & 5t \end{cases}$$

Otra forma de haber obtenido lo mismo consistiría en haber continuado hasta obtener una matriz fuertemente escalonada.

$e[1]:e[1]-e[3]$   $e[2]:e[2]+3*e[3]$   $e$ ;

$$\begin{pmatrix} 1 & -3 & 0 & \frac{4}{5} & \frac{9}{5} \\ 0 & 1 & 0 & -\frac{7}{5} & \frac{13}{5} \\ 0 & 0 & 1 & -\frac{4}{5} & \frac{1}{5} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$e[1]:e[1]+3*e[2]$   $e$ ;

$$\begin{pmatrix} 1 & 0 & 0 & -\frac{17}{5} & \frac{48}{5} \\ 0 & 1 & 0 & -\frac{7}{5} & \frac{13}{5} \\ 0 & 0 & 1 & -\frac{4}{5} & \frac{1}{5} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

que nos proporciona la siguiente forma fuertemente escalonada:

$$\left( \begin{array}{cccc|c} 1 & 0 & 0 & -\frac{17}{5} & \frac{48}{5} \\ 0 & 1 & 0 & -\frac{7}{5} & \frac{13}{5} \\ 0 & 0 & 1 & -\frac{4}{5} & \frac{1}{5} \\ 0 & 0 & 0 & 0 & 0 \end{array} \right)$$

Esta es la matriz del siguiente sistema

$$\left. \begin{array}{l} X_1 \\ X_2 \\ X_3 \end{array} \begin{array}{l} - \\ - \\ - \end{array} \begin{array}{l} \frac{17}{5} X_4 \\ \frac{7}{5} X_4 \\ \frac{4}{5} X_4 \end{array} = \begin{array}{l} \frac{48}{5} \\ \frac{13}{5} \\ \frac{1}{5} \end{array} \right\}$$

Poniendo  $X_4 = 5t$  y despejando las tres primeras variables obtenemos el mismo resultado.

Obsérvese que si  $n$  es el número de variables de un sistema y  $r$  es el rango de la matriz escalonada obtenida entonces las soluciones quedan en función de  $n - r$  variables. Por ejemplo, si partimos una matriz arbitraria  $A$  y consideramos el sistema de ecuaciones  $AX = 0$  entonces el sistema es compatible, pues al menos el vector  $0$  es solución, y el número de variables que determinan las demás es  $n - r$  donde  $r$  es el rango de la cualquiera de las matrices escalonadas que se obtienen al aplicar el Método de Gauss a la matriz  $A$ . Eso tiene la siguiente consecuencia:

**Proposición 3.2.5** *Todas las formas escalonadas que se obtienen al aplicar el Método de Gauss a una matriz  $A$  tienen el mismo rango.*

**Definición 3.2.6** *Llamamos rango de una matriz  $A$  al de cualquiera de las matrices escalonadas que se obtienen al aplicar el Método de Gauss. Denotamos por  $r(A)$  al rango de  $A$ .*

Al aplicar el Método de Gauss a la matriz  $(A|b)$  de un sistema obtenemos una matriz escalonada  $(E|t)$  y  $E$  será una de las formas escalonadas de  $A$ . Por tanto  $A$  tiene el mismo rango que  $E$  y  $(A|b)$  tiene el mismo rango que  $(E|t)$ . El sistema es incompatible si y solo si  $t$  tiene un pivote o lo que es lo mismo cuando  $(E|t)$  tiene más rango que  $E$ , que es cuando  $r(A) < r(A|b)$ .

Supongamos que el sistema es compatible y que los pivotes en una de la matrices escalonadas obtenidas al aplicar de Gauss están en las columnas  $j_1, \dots, j_r$ . Si  $n$  es el número de incógnitas, entonces las soluciones del sistema se puede obtener eligiendo arbitrariamente las  $n-r$  variables  $X_j$  con  $j \neq j_i$  para todo  $i = 1, \dots, r$  lo cual determina el valor de las otras variables. En tal caso decimos que el sistema tiene  $n-r$  grados de libertad. El sistema es compatible determinado si y solo si tiene 0-grados de libertad o sea si  $n = r$ . Resumiendo:

**Teorema 3.2.7 (Rouché-Frobenius)** *Consideremos un sistema de ecuaciones lineales  $AX = b$  con  $n$  incógnitas. Entonces*

- (1) *El sistema es incompatible si y solo si  $r(A) < r(A|b)$ .*
- (2) *El sistema es compatible determinado si y solo si  $r(A) = r(A|b) = n$ .*
- (3) *El sistema es compatible indeterminado si y solo si  $r(A) = r(B) < n$ .*

### 3.3 Espacios fila, columna y nulo de una matriz

**Definición 3.3.1** *Dada una matriz  $A \in M_{m,n}(K)$  llamamos espacio fila de  $A$  al subespacio vectorial  $F(A)$  de  $K^n$  generado por las filas de  $A$ . Análogamente se llama espacio columna de  $A$  al subespacio vectorial  $C(A)$  de  $K^m$  generado por las columnas de  $A$ . Finalmente se llama espacio nulo de  $A$  al conjunto  $N(A) = \{v \in K^n : Av = 0\}$ .*

Obviamente el espacio fila de una matriz es el espacio columna de su transpuesta y viceversa.

**Proposición 3.3.2** *Sea  $A \in M_{m,n}(K)$ . Entonces la imagen y el núcleo de la aplicación lineal  $M_A : K^n \rightarrow K^m$  son precisamente el espacio columna y el núcleo de  $A$  respectivamente.*

**Demostración.** Claramente  $x \in N(M_A)$  si y solo si  $Ax = 0$  si y solo si  $x \in N(A)$ . Por otro lado si  $A = (a_{ij})$  entonces

$$M_A(K^n) = \{M_A x : x \in K^n\} = \{Ax : x \in K^n\}$$

Pero si  $x = (x_1, \dots, x_n)$  entonces

$$Ax = A \begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix} = x_1 a_{.1} + x_2 a_{.2} + \dots + x_n a_{.n}$$

Por tanto

$$M_A(K^n) = \langle a_{.1}, \dots, a_{.n} \rangle$$

que es el espacio columna de  $A$ . ■

Como consecuencia de la Proposición 3.2.1 y de la Proposición 2.4.4 tenemos el siguiente

**Corolario 3.3.3** *Un sistema de ecuaciones lineales  $AX = b$  es compatible si y solo si  $b$  pertenece al espacio columna de  $A$  si y solo si  $b$  pertenece a la imagen de la aplicación  $M_A$ .*

*El sistema compatible es determinado si y solo si las columnas de la matriz de coeficientes son linealmente independientes.*

Si tenemos una lista  $v_1, \dots, v_m$  de  $K^n$ , los podemos poner como las filas de una matriz  $A = (a_{ij})$ , es decir  $a_i = v_i$  y tendremos que  $\langle v_1, \dots, v_n \rangle$  es el espacio fila de  $A$ . Si las ponemos como columnas de una matriz entonces  $\langle v_1, \dots, v_n \rangle$  será el espacio columna de dicha matriz.

Como consecuencia del Corolario 2.3.11 todas las matrices que tengan la misma forma escalonada tienen el mismo espacio fila de donde se deducen muchas consecuencias.

**Proposición 3.3.4** *Si  $A$  es una matriz y  $E$  es una matriz escalonada que se obtiene al aplicar el Método de Gauss a  $A$  entonces  $F(A) = F(E)$  y  $N(A) = N(E)$ .*

**Demostración.** La igualdad  $F(A) = F(E)$  es consecuencia del Corolario 2.3.11. Por otro lado, por la Proposición 3.1.7 se tiene que  $E = UA$  con  $U$  un producto de matrices elementales. Además por el Teorema 3.1.8, la matriz  $U$  es invertible y eso implica que  $Ax = 0$  si y solo si  $Ex = 0$ . O sea  $N(A) = N(E)$ . ■

Recordemos la definición de base.

**Definición 3.3.5** *Una lista  $B$  de elementos de un espacio vectorial  $V$  se dice que es una base si es linealmente independiente y sistema generador.*

**Ejemplos 3.3.6** (1) Ya conocemos la base canónica  $\{e_1, e_2, \dots, e_n\}$  de  $K^n$  (Ejemplos 2.3.8 que obviamente es una base de  $K^n$ ).

(2) Sea  $B = \{v_1, v_2, \dots, v_n\}$  una base de  $V$  y consideremos los siguientes elementos de  $V$ :

$$f_1 = v_1, f_2 = v_1 + v_2, f_3 = v_1 + v_2 + v_3, \dots, f_n = v_1 + v_2 + v_3 + \dots + v_n.$$

Vamos a ver que  $\{f_1, \dots, f_n\}$  es otra base de  $K^n$ .

Empecemos viendo que es linealmente independiente. Supongamos que

$$\alpha_1 f_1 + \alpha_2 f_2 + \dots + \alpha_n f_n = 0.$$

Entonces

$$\begin{aligned} 0 &= \alpha_1 v_1 + \alpha_2 (v_1 + v_2) + \alpha_3 (v_1 + v_2 + v_3) + \dots + \alpha_n (v_1 + v_2 + v_3 + \dots + v_n) \\ &= (\alpha_1 + \alpha_2 + \alpha_3 + \dots + \alpha_n) v_1 + (\alpha_2 + \alpha_3 + \dots + \alpha_n) v_2 + \dots + \alpha_n v_n. \end{aligned}$$

Como sabemos que  $\{v_1, v_2, \dots, v_n\}$  es linealmente independiente, tenemos que

$$\begin{array}{rcccccc} \alpha_1 & + & \alpha_2 & + & \alpha_3 & + & \cdots & + & \alpha_{n-1} & + & \alpha_n & = & 0 \\ & & \alpha_2 & + & \alpha_3 & + & \cdots & + & \alpha_{n-1} & + & \alpha_n & = & 0 \\ & & & & \alpha_3 & + & \cdots & + & \alpha_{n-1} & + & \alpha_n & = & 0 \\ & & & & & & & & & & & & \vdots \\ & & & & & & & & \alpha_{n-1} & + & \alpha_n & = & 0 \\ & & & & & & & & & & \alpha_n & = & 0 \end{array}$$

La matriz de coeficientes de este sistema es triangular y todos los elementos de la diagonal son iguales a 1. Eso implica que es una matriz invertible, por la Proposición 3.4.6, y por tanto el sistema tiene una única solución, que solo puede ser  $\alpha_1 = \cdots = \alpha_n = 0$ . Esto demuestra que  $\{f_1, \dots, f_n\}$  es linealmente independiente.

Vamos a ver ahora que  $\{w_1, \dots, w_n\}$  es un sistema generador. Es decir queremos ver que  $\langle w_1, w_2, \dots, w_n \rangle = V$ . Como  $\{v_1, v_2, \dots, v_n\}$  es un sistema generador, aplicando el Corolario 2.3.9 basta demostrar que cada  $v_i$  es combinación lineal de los  $w_i$  lo que efectivamente se verifica pues  $v_1 = w_1$  y  $v_i = w_i - w_{i-1}$  para todo  $i \geq 2$ .

- (3) El espacio vectorial  $M_{m,n}(K)$  de las matrices con  $m$  filas y  $n$  columnas con entradas en  $K$  tiene una base muy similar a la base canónica de  $K^n$  formada por las matrices que tienen una entrada 1 y las demás 0. Ésta se llama *base canónica* de  $M_{m,n}(K)$  de  $M_{m,n}(K)$ .

Por ejemplo, la base canónica de  $M_{2,3}(K)$  está formada por las matrices

$$\begin{array}{l} e_{11} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad e_{12} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad e_{13} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \\ e_{21} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad e_{22} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad e_{23} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \end{array}$$

- (4) El conjunto formado por los monomios  $1, X, X^2, \dots$  es una base del espacio vectorial  $K[X]$  de los polinomios con coeficientes en  $K$  que también hemos llamado base canónica de  $K[X]$ .

Supongamos que queremos calcular una base del espacio fila de una matriz  $A$ . Sea  $E$  una matriz fuertemente escalonada obtenida al aplicar el Método de Gauss a  $A$ . Por la Proposición 3.3.4 tenemos que  $F(A) = F(E)$ . Sea  $U$  la matriz obtenida quitando a  $E$  las columnas que no tengan pivotes y las filas nulas. Esta matriz tiene la siguiente forma:

$$U = \begin{pmatrix} 1 & * & * & \cdots & * \\ & 1 & * & \cdots & * \\ & & \ddots & \ddots & \vdots \\ & & & \ddots & * \\ & & & & 1 \end{pmatrix}$$

Es obvio que las filas de esta matriz son linealmente independientes, con lo que también lo son las de  $E$ . Por tanto, si  $v_1, \dots, v_m$  son las filas  $A$  entonces de las filas no nulas de  $E$  forman una base de  $\langle v_1, \dots, v_n \rangle$ .

Por otro lado, si  $E$  tiene un vector 0 entonces las filas de  $E$  son linealmente dependientes y por tanto también lo serán los vectores  $v_1, \dots, v_n$ . Sin embargo, si ninguna de las filas de  $E$  son cero entonces por lo dicho arriba las filas de  $E$  son linealmente independientes y por tanto también lo son  $v_1, \dots, v_n$ .

Por tanto tenemos los dos siguientes algoritmos:

**Algoritmo 3.3.7 (Base de un subespacio generado por vectores de  $K^n$ : Gauss)**

Entrada: Una lista  $v_1, \dots, v_m$  de vectores de  $K^n$ .

- Construir la matriz  $A$  formada poniendo los vectores  $v_1, \dots, v_m$  como filas.
- $E :=$  Matriz escalonada obtenida al aplicar el Método de Gauss a la matriz  $A$ .

Salida: Los vectores no nulos de  $E$  forman una base de  $\langle v_1, \dots, v_n \rangle$ .

Está claro que el mismo algoritmo nos sirve para calcular una base del espacio fila de una matriz y aplicando dicho algoritmo a la transpuesta de la matriz obtenemos una base del espacio columna.

**Algoritmo 3.3.8 (Dependencia o independencia de vectores de  $K^n$ : Gauss)**

Entrada: Una lista  $v_1, \dots, v_m$  de vectores de  $K^n$ .

- Construir la matriz  $A$  formada poniendo los vectores  $v_1, \dots, v_m$  como filas.
- $E :=$  Matriz escalonada obtenida al aplicar el Método de Gauss a la matriz  $A$ .

Salida: Si  $E$  tiene una fila nula entonces  $\{v_1, \dots, v_n\}$  es linealmente dependiente. En caso contrario es linealmente independiente.

Vamos a ilustrar los Algoritmos 3.3.7 y 3.3.8 con un ejemplo.

**Ejemplo 3.3.9** Consideremos los cuatro vectores de  $K^5$ :

$$(3, -2, 0, 1, 2), (0, 2, 6, 2, 7), (1, 0, 2, 1, 3), (1, -1, 1, 2, 5)$$

Construimos una matriz poniendo estos vectores como fila y le aplicamos el Método de Gauss

$$A = \begin{pmatrix} 3 & -2 & 0 & 1 & 2 \\ 0 & 2 & 6 & 2 & 7 \\ 1 & 0 & 2 & 1 & 3 \\ 1 & -1 & -1 & 2 & 5 \end{pmatrix}, \quad A_1 = E_{1,3}A = \begin{pmatrix} 1 & 0 & 2 & 1 & 3 \\ 0 & 2 & 6 & 2 & 7 \\ 3 & -2 & 0 & 1 & 2 \\ 1 & -1 & -1 & 2 & 5 \end{pmatrix},$$

$$A_2 = E_{4,1,1}E_{3,-3,1}A_1 = \begin{pmatrix} 1 & 0 & 2 & 1 & 3 \\ 0 & 2 & 6 & 2 & 7 \\ 0 & -2 & -6 & -2 & -7 \\ 0 & -1 & -3 & 1 & 2 \end{pmatrix},$$

$$A_3 = E_{(-1) \cdot 2} E_{3,4} E_{2,4} E_{3,1,2} A_2 = \begin{pmatrix} 1 & 0 & 2 & 1 & 3 \\ 0 & 1 & 3 & -1 & -2 \\ 0 & 2 & 6 & 2 & 11 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$A_4 = E_{\frac{1}{4} \cdot 3} E_{2,3,1} A_3 = \begin{pmatrix} 1 & 0 & 2 & 1 & 3 \\ 0 & 1 & 3 & -1 & -2 \\ 0 & 0 & 0 & 1 & \frac{11}{4} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Estos cálculos nos indican que los cuatro vectores originales son linealmente dependientes y que los vectores

$$(1, 0, 2, 1, 3), (0, 1, 1, -1, -2), (0, 0, 0, 4, 11)$$

forman una base del espacio vectorial generado por los cuatro vectores originales. Obsérvese que el último vector lo hemos multiplicado por 4 para liberarnos de los denominadores.

Supongamos ahora que queremos calcular una base del espacio nulo de  $A$ . En realidad los elementos de  $A$  son las soluciones del sistema lineal homogéneo  $AX = 0$ . Por tanto basta resolver el sistema con el Método de Gauss a la matriz  $A$ , pues la última columna de la matriz del sistema  $(A|0)$  se mantendrá todo el tiempo igual a 0. Si  $A$  tiene  $n$  columnas y rango  $r$  entonces el conjunto tiene  $n - r$  grados de libertad y de ahí nos saldrá una base con  $n - r$  elementos.

**Ejemplo 3.3.10** Consideremos la matriz

$$A = \begin{pmatrix} 3 & -2 & 0 & 1 & 2 \\ 0 & 2 & 6 & 2 & 7 \\ 1 & 0 & 2 & 1 & 3 \\ 1 & -1 & -1 & 2 & 5 \end{pmatrix}$$

que ya apareció en el Ejemplo 3.3.9. Habíamos obtenido la siguiente forma escalonada suya aplicando el Método de Gauss

$$A_4 = \begin{pmatrix} 1 & 0 & 2 & 1 & 3 \\ 0 & 1 & 3 & -1 & -2 \\ 0 & 0 & 0 & 1 & \frac{11}{4} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Continuamos aplicando el Método de Gauss-Jordan para obtener una forma fuertemente escalonada:

$$A_5 = E_{1,-1,3} E_{2,1,3} A_4 = \begin{pmatrix} 1 & 0 & 2 & 0 & 1/4 \\ 0 & 1 & 3 & 0 & 3/4 \\ 0 & 0 & 0 & 1 & \frac{11}{4} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Los pivotes están en las columnas 1, 2 y 4. Obsérvese que el espacio nulo de la matriz anterior (que es el mismo que el de la matriz  $A$ ) viene dado por las ecuaciones:

$$\begin{aligned} x_1 &= -2x_3 - \frac{1}{4}x_5 \\ x_2 &= -3x_3 - \frac{3}{4}x_5 \\ x_4 &= -\frac{11}{4}x_5 \end{aligned}$$

La base que nos proporciona el algoritmo se obtiene asignando a  $(x_3, x_5)$  los dos vectores de la base canónica de  $K^2$ , o sea, asignando primero  $x_3 = 1$  y  $x_5 = 0$  obtenemos el primero de los siguientes vectores, y asignando  $x_3 = 0$  y  $x_5 = 1$ , obtenemos el segundo:

$$(-2, -3, 1, 0, 0), \quad \left(-\frac{1}{4}, -\frac{3}{4}, 0, -\frac{11}{4}, 1\right).$$

Estos dos vectores forman una base de  $N(A)$ , pero multiplicar alguno de los vectores por un número distinto de 0 no cambia el que siga siendo una base. Por tanto multiplicamos el segundo por 4 para librarnos de los denominadores. De forma que nos quedamos con la base:

$$(-2, -3, 1, 0, 0), \quad (-1, -3, 0, -11, 4).$$

Según vimos en la Proposición 3.3.2 si  $A \in M_{m,n}(A)$  entonces la imagen de  $M_A$  es el espacio columna de  $A$  que es el espacio fila de la matriz transpuesta  $A^T$ . Por tanto, para calcular la imagen de  $M_A$ , podemos usar el Algoritmo 3.3.7 aplicado a  $A^T$ .

**Ejemplo 3.3.11** Consideremos ahora la aplicación lineal  $M_A : K^5 \rightarrow K^4$  con  $A$  la matriz del Ejemplo anterior. Aplicamos el Método de Gauss a su transpuesta:

$$B = A^T = \begin{pmatrix} 3 & 0 & 1 & 1 \\ -2 & 2 & 0 & -1 \\ 0 & 6 & 2 & -1 \\ 1 & 2 & 1 & 2 \\ 2 & 7 & 3 & 5 \end{pmatrix}.$$

`b:matrix([3,0,1,1], [-2,2,0,-1],[0,6,2,-1],[1,2,1,2],[2,7,3,5])$ e:echelon(b);`

$$\begin{pmatrix} 1 & 0 & \frac{1}{3} & \frac{1}{3} \\ 0 & 1 & \frac{1}{3} & -\frac{1}{6} \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad (\text{e})$$

Las filas de la matriz obtenida forman una base del espacio fila de  $A^T$ , que es el espacio columna de  $A$ , que a su vez es la imagen de  $M_A$ . Para quitar denominadores multiplicamos la primera por 3 y la segunda por 6 y de esta forma hemos obtenido una base de la imagen de  $M_A$ :

$$(3, 0, 1, 1), \quad (0, 6, 2, -2), \quad (0, 0, 0, 1)$$

**Algoritmo 3.3.12 (Ecuaciones del subespacio generado por vectores de  $K^n$ : Gauss)**

Entrada: Una lista  $v_1, \dots, v_m$  de vectores de  $K^n$ .

- Construir una matriz  $(A|x)$  donde  $A = (v_1^T, \dots, v_m^T)$ , la matriz formada poniendo en columna los vectores  $v_1, \dots, v_m$  y  $x$  es un vector de indeterminadas  $x_1, \dots, x_n$ .

- Aplicar el Método de Gauss a la matriz  $(A|x)$ . En realidad esto no puede completarse pues la última columna está formada por variables y en el proceso no sabemos qué entradas son cero o no. Por tanto, lo que se hace es aplicar el Método a la matriz  $A$  de forma que la última columna va transformándose de la misma manera que  $A$ . Al final tendremos una matriz

$$\left( \begin{array}{c|c} E & * \\ \hline 0 & y \end{array} \right)$$

donde  $E$  es una matriz escalonada e

$$y = \begin{pmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{k1}x_1 + \cdots + a_{kn}x_n \end{pmatrix}$$

Salida: El sistema de ecuaciones lineales

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= 0 \\ &\vdots \\ a_{k1}x_1 + \cdots + a_{kn}x_n &= 0 \end{aligned}$$

Los vectores que están  $\langle v_1, \dots, v_m \rangle$  son las soluciones de dicho sistema.

**Ejemplo 3.3.13** Vamos a calcular ecuaciones para el subespacio vectorial generado por los vectores

$$(1, 2, 4, 0, 2), (-1, 2, 0, -4, 1), (-1, 6, 4, 0, -1)$$

Ponemos en columna los cinco vectores y le añadimos una columna de indeterminadas:

$$\left( \begin{array}{ccc|c} 1 & -1 & -1 & x_1 \\ 2 & 2 & 6 & x_2 \\ 4 & 0 & 4 & x_3 \\ 0 & 0 & 0 & x_4 \\ 2 & 1 & -1 & x_5 \end{array} \right)$$

Aplicando el Método de Gauss obtenemos lo siguiente:

$$\begin{aligned} &\left( \begin{array}{ccc|c} 1 & -1 & -1 & x_1 \\ 0 & 4 & 8 & x_2 - 2x_1 \\ 0 & 4 & 8 & x_3 - 4x_1 \\ 0 & -1 & -7 & x_5 - x_2 \\ 0 & 0 & 0 & x_4 \end{array} \right) \left( \begin{array}{ccc|c} 1 & -1 & -1 & x_1 \\ 0 & 4 & 8 & x_2 - 2x_1 \\ 0 & 0 & 0 & x_3 - x_2 - 2x_1 \\ 0 & 1 & 7 & x_2 - x_5 \\ 0 & 0 & 0 & x_4 \end{array} \right) \\ &\left( \begin{array}{ccc|c} 1 & -1 & -1 & x_1 \\ 0 & 1 & 7 & x_2 - x_5 \\ 0 & 0 & -20 & 4x_5 - 3x_2 - 2x_1 \\ 0 & 0 & 0 & x_3 - x_2 - 2x_1 \\ 0 & 0 & 0 & x_4 \end{array} \right) \end{aligned}$$

Por tanto, los vectores en el subespacio generado por los vectores iniciales son lo que satisfacen las ecuaciones:

$$x_4 = 0, \quad x_3 - x_2 - 2x_1 = 0.$$

### 3.4 Determinantes

A cada matriz cuadrada  $A$  le vamos a asociar un escalar que llamaremos determinante de  $A$  y que denotaremos  $|A|$  ó  $\det(A)$ . En lugar de dar una fórmula cerrada para el determinante de  $A$  vamos dar una lista de propiedades que lo van a determinar. En concreto el determinante cumple las siguientes propiedades:

(Det1) El determinante de una matriz identidad es 1.

(Det2) Si dos filas de  $A$  son iguales entonces  $\det(A) = 0$ .

(Det3) Si  $\alpha \in F$  entonces  $\det(E_{\alpha,i}A) = \alpha \det(A)$ .

(Det4) Sean  $A, B$  y  $C$  dos matrices que tienen todas las filas iguales menos una y que esa fila en  $A$  es  $a$ , en  $B$  es  $b$  y en  $C$  es  $a + b$ . Entonces  $\det(C) = \det(A) + \det(B)$ . O sea

$$\det((a_1 \ a_2 \ \dots \ a_i + b_i \ \dots \ a_n)^T) = \det((a_1 \ a_2 \ \dots \ a_i \ \dots \ a_n)^T) + \det((a_1 \ a_2 \ \dots \ b_i \ \dots \ a_n)^T).$$

(Det2) nos dice cómo afecta una de las dos operaciones elementales al determinante. Las siguientes fórmulas nos dice cómo afectan las otras dos:

$$\det(E_{i,j}A) = -\det(A). \quad (3.2)$$

$$\det(E_{i,\alpha,j}A) = \det(A). \quad (3.3)$$

Empezamos demostrando (3.3). Supongamos que  $A = (a_{ij})$ . Entonces, aplicando sucesivamente (Det4), (Det3) y (Det1) y obtenemos

$$\begin{aligned} \det(E_{i,\alpha,j}A) &= \det((a_1 \ \dots \ a_i + \alpha a_j \ \dots \ a_j \ \dots \ a_n)^T) \\ &= \det((a_1 \ \dots \ a_i \ \dots \ a_j \ \dots \ a_n)^T) + \det((a_1 \ \dots \ \alpha a_j \ \dots \ a_j \ \dots \ a_n)^T) \\ &= \det((a_1 \ \dots \ a_i \ \dots \ a_j \ \dots \ a_n)^T) + \alpha \det((a_1 \ \dots \ a_j \ \dots \ a_j \ \dots \ a_n)^T) \\ &= \det((a_1 \ \dots \ a_i \ \dots \ a_j \ \dots \ a_n)^T) = \det(A). \end{aligned}$$

Para demostrar (3.2) aplicamos sucesivamente (Det2), (Det4) dos veces y (Det2) otra vez:

$$\begin{aligned} \det(A) + \det(E_{i,j}A) &= \det((a_1 \ \dots \ a_i \ \dots \ a_j \ \dots \ a_n)^T) + \det((a_1 \ \dots \ a_j \ \dots \ a_i \ \dots \ a_n)^T) \\ &= \det((a_1 \ \dots \ a_i \ \dots \ a_j \ \dots \ a_n)^T) + \det((a_1 \ \dots \ a_i \ \dots \ a_i \ \dots \ a_n)^T) + \\ &\quad \det((a_1 \ \dots \ a_j \ \dots \ a_i \ \dots \ a_n)^T) + \det((a_1 \ \dots \ a_j \ \dots \ a_j \ \dots \ a_n)^T) \\ &= \det((a_1 \ \dots \ a_i \ \dots \ a_i + a_j \ \dots \ a_n)^T) + \det((a_1 \ \dots \ a_i + a_j \ \dots \ a_i \ \dots \ a_n)^T) \\ &= \det((a_1 \ \dots \ a_i + a_j \ \dots \ a_i + a_j \ \dots \ a_n)^T) = 0. \end{aligned}$$

Aplicando (Det3) con  $\alpha = 0$  se tiene el siguiente

**Lema 3.4.1** Si  $A$  tiene una fila igual a 0 entonces  $\det(A) = 0$ .

Es fácil calcular los determinantes de las matrices elementales:

$$\begin{aligned}\det(E_{i,j}) &= \det(E_{i,j}I) = -\det(I) = -1, \\ \det(E_{\alpha,i}) &= \det(E_{\alpha,i}I) = \alpha \det(I) = \alpha, \\ \det(E_{i,\alpha,j}) &= \det(I) + \det(A) = 1.\end{aligned}\tag{3.4}$$

Recuerda que siempre que usemos las matrices elementales  $E_{i,j}$  o  $E_{i,\alpha,j}$  se entiende que  $i \neq j$ .

Una vez que sabemos cómo afectan las operaciones elementales combinándolo con (Det1) y el Lema 3.4.1 se puede calcular el determinante aplicando el Método de Gauss-Jordan pues cada uno de los pasos del método nos dice cómo varía el determinante y al final, la matriz fuertemente escalonada es la una matriz identidad o tiene una fila cero. Veamos esto con un ejemplo:

$$\begin{aligned}\begin{vmatrix} 0 & 2 & 1 & 2 \\ 1 & 0 & 3 & -1 \\ -1 & 4 & 0 & 2 \\ -2 & 3 & 5 & 1 \end{vmatrix} &= - \begin{vmatrix} 1 & 0 & 3 & -1 \\ 0 & 2 & 1 & 2 \\ -1 & 4 & 0 & 2 \\ -2 & 3 & 5 & 1 \end{vmatrix} = - \begin{vmatrix} 1 & 0 & 3 & -1 \\ 0 & 2 & 1 & 2 \\ 0 & 4 & 3 & 1 \\ 0 & 3 & 11 & -1 \end{vmatrix} \\ &= -2 \begin{vmatrix} 1 & 0 & 3 & -1 \\ 0 & 1 & \frac{1}{2} & 1 \\ 0 & 4 & 3 & 1 \\ 0 & 3 & 11 & -1 \end{vmatrix} = -2 \begin{vmatrix} 1 & 0 & 3 & -1 \\ 0 & 1 & \frac{1}{2} & 1 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & \frac{19}{2} & -4 \end{vmatrix} \\ &= -2 \begin{vmatrix} 1 & 0 & 3 & -1 \\ 0 & 1 & \frac{1}{2} & 1 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & -\frac{49}{2} \end{vmatrix} = 49 \begin{vmatrix} 1 & 0 & 3 & -1 \\ 0 & 1 & \frac{1}{2} & 1 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 1 \end{vmatrix}\end{aligned}$$

Esto termina el Método de Gauss y ahora deberíamos seguir con el Método de Gauss poniendo ceros encima de la diagonal pero eso nos lo podemos ahorrar pues en esa parte final solo se utiliza la tercera operación elemental que no afecta al determinante como hemos visto en (3.3). Por tanto al final tendríamos

$$\begin{vmatrix} 0 & 2 & 1 & 2 \\ 1 & 0 & 3 & -1 \\ -1 & 4 & 0 & 2 \\ -2 & 3 & 5 & 1 \end{vmatrix} = 49 \det(I) = 49.$$

La otra alternativa que podría haber ocurrido es que tuviéramos menos pivotes que columnas con lo que la última fila sería cero y el determinante sería 0. Eso queda reflejado en el siguiente

**Teorema 3.4.2** *Una matriz cuadrada es invertible si y solo si su determinante es distinto de 0 si y solo si su rango es igual an número de filas (y columnas).*

**Demostración.** Sea  $B$  la matriz que resulta después de aplicar una operación elemental a  $A$ . Entonces  $\det(B) = \alpha \det(A)$  para algún  $\alpha \neq 0$ , pues no estamos usando (Det3) con  $\alpha = 0$ . Por

tanto  $\det(A) = 0$  si y solo si  $\det(B) = 0$ . Luego, si  $E$  es una matriz fuertemente escalonada que se obtiene después de aplicar el Método de Gauss a  $A$  entonces  $\det(A) = 0$  si y solo si  $\det(E) = 0$ . Por el Teorema 3.1.8, si  $n$  es el número de filas,  $A$  es invertible si y solo si  $E = I_n$  si y solo si  $E$  tiene rango  $n$ . En tal caso  $\det(E) = 1$  y por tanto  $\det(A) \neq 0$ . En caso contrario  $E$  tiene una fila igual a 0 y por el Lema 3.4.1,  $\det(E) = 0$  con lo que  $\det(A) = 0$ . ■

**Teorema 3.4.3**  $\det(AB) = \det(A) \det(B)$ .

**Demostración.** Supongamos que  $A$  es invertible. Entonces es un producto de matrices elementales, pongamos  $A = E_1 \dots E_k$  con cada  $E_i$  una matriz elemental. Razonamos por inducción en  $k$ . Si  $k = 0$  entonces  $A = I$  y por tanto  $\det(AB) = \det(B) = \det(I) \det(B) = \det(A) \det(B)$ . Si  $k = 1$  entonces  $A$  es una matriz elemental y el resultado es consecuencia de los axiomas (Det2), (Det3) y (Det4) y las fórmulas en (3.4). Suponiendo que se cumple el resultado para el caso en que  $A$  es producto de menos de  $k$  matrices elementales tenemos que  $\det(E_2 \dots E_k B) = \det(E_2 \dots E_k) \det(B)$ . Aplicando que también se cumple para el caso en que  $A$  es una matriz elemental tenemos que  $\det(AB) = \det(E_1 \dots E_k B) = \det(E_1) \det(E_2 \dots E_k B) = \det(E_1) \det(E_2 E_3 \dots E_k) \det(B) = \det(E_1 E_2 E_3 \dots E_k) \det(B) = \det(A) \det(B)$ .

Supongamos finalmente que  $A$  no es invertible. Sea  $E$  una matriz fuertemente escalonada que se obtiene al aplicar el Método de Gauss a  $A$ . Por los Teoremas 3.4.2 y 3.1.8 se tiene que  $\det(A) = 0$  y la última fila de  $E$  es 0. Entonces la última fila de  $EB$  también tiene una fila igual a cero. Si ahora aplicamos el Método de Gauss a  $EB = UAB$  obtendremos otra matriz fuertemente escalonada  $F = VEB = VUAB$  con  $V$  un producto de matrices elementales y en el proceso de aplicar el Método de Gauss a  $EB$  podemos suponer que no hemos usado la última fila (que es cero) con lo que la última fila de  $F$  también es cero. Además  $VU$  es un producto de matrices elementales y por tanto  $F$  es una matriz fuertemente escalonada obtenida al aplicar el Método de Gauss a  $AB$ . Aplicando una vez más los Teoremas 3.1.8 y 3.4.2 tenemos que  $AB$  no es invertible y  $\det(AB) = 0 = \det(A) \det(B)$ . ■

**Corolario 3.4.4**  $\det(A^T) = \det(A)$ .

**Demostración.** Supongamos primero que  $A$  no es invertible. Por la Proposición 3.1.6 tenemos que  $A^T$  tampoco es invertible con lo que  $\det(A) = 0 = \det(A^T)$ , por el Teorema 3.4.2.

Supongamos que  $A$  es una matriz elemental. Si  $A = E_{i,j}$  ó  $A = E_{\alpha,i}$  entonces  $A$  es una matriz simétrica y por tanto  $\det(A^T) = \det(A)$ . En el caso que resta, o sea si  $A = E_{i,\alpha,j}$  se tiene que  $A^T = E_{j,\alpha,i}$  y tenemos que  $\det(A) = 1 = \det(A^T)$ .

Supongamos ahora que  $A$  es una matriz invertible arbitraria. Entonces  $A$  es un producto de matrices elementales por el Teorema 3.1.8. Pongamos que  $A = E_1 \dots E_k$  con cada  $E_i$  una matriz elemental. Entonces  $A^T = E_k^T \dots E_1^T$ , por la Proposición 3.1.6 y aplicando el Teorema 3.4.3 y que la propiedad se verifica para matrices elementales tenemos:

$$\begin{aligned} \det(A^T) &= \det(A_k^T \dots A_1^T) = \det(A_k^T) \dots \det(A_1^T) \\ &= \det(A_k) \dots \det(A_1) = \det(A_1) \dots \det(A_k) = \det(A_1 \dots A_k) = \det(A). \end{aligned}$$

■

Una consecuencia inmediata del Corolario 3.4.4 es la siguiente.

**Corolario 3.4.5** *Todas las propiedades del determinante que hemos visto para las filas vale para las columnas. Por ejemplo si  $A$  tiene dos columnas iguales entonces  $\det(A) = 0$ ,*

$$\begin{aligned} \det(AE_{i,j}) &= -\det(A) \\ \det(AE_{\alpha,i}) &= \alpha \det(A) \\ \det(AE_{i,\alpha,j}) &= \det(A) \\ \det(a_{.1}a_{.2}, \dots, a_{.i} + b_{.i}, \dots, a_{.n}) &= \det(a_{.1}a_{.2}, \dots, a_{.i}, \dots, a_{.n}) + \det(a_{.1}a_{.2}, \dots, b_{.i}, \dots, a_{.n}) \end{aligned}$$

**Proposición 3.4.6** *El determinante de una matriz triangular es el producto de los elementos de su diagonal.*

**Demostración.** Supongamos primero que todas las entradas en la diagonal de  $A$  son iguales a 1. Entonces  $A$  es escalonada y tiene pivotes en todas las columnas. Eso implica que  $A$  es invertible y por tanto cuando al aplicar el Método de Gauss-Jordan  $A$  obtendremos la matriz identidad. Además como  $A$  ya es escalonada para obtener la forma fuertemente escalonada solo necesitamos usar operaciones elementales del tercer tipo que no cambian el determinante. Por tanto  $\det(A) = \det(I) = 1$  que es el producto de los elementos de la diagonal de  $A$ .

Supongamos ahora que  $A$  es triangular superior y todas las entradas en la diagonal de  $A$  son diferentes de cero. Aplicando (Det3), y lo que hemos visto en el párrafo anterior tenemos

$$\begin{aligned} \begin{vmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ & a_{22} & a_{23} & \dots & a_{2n} \\ & & a_{33} & \dots & a_{3n} \\ & & & \ddots & * \\ & & & & a_{nn} \end{vmatrix} &= a_{11}a_{22} \dots a_{nn} \begin{vmatrix} 1 & \frac{a_{12}}{a_{11}} & \frac{a_{13}}{a_{11}} & \dots & \frac{a_{1n}}{a_{11}} \\ & 1 & \frac{a_{23}}{a_{11}} & \dots & \frac{a_{2n}}{a_{11}} \\ & & 1 & \dots & \frac{a_{3n}}{a_{11}} \\ & & & \ddots & * \\ & & & & 1 \end{vmatrix} \\ &= a_{11}a_{22} \dots a_{nn}. \end{aligned}$$

Para completar el caso en que la matriz sea triangular superior supongamos que  $A$  lo es y que una de las entradas de la diagonal de  $A$  es cero. Supongamos además que la última que es cero aparece en la posición  $i$ -ésima, o sea  $a_{ii} = 0$  pero  $a_{jj} \neq 0$  para todo  $j > i$ . Aplicando el Método de Gauss-Jordan a la submatriz formada por las filas debajo de la fila  $i$ -ésima tenemos lo siguiente:

$$\begin{vmatrix} a_{11} & \dots & a_{1i} & a_{1,i+1} & \dots & a_{1n} \\ & \ddots & * & \dots & \dots & * \\ & & 0 & a_{i,i+1} & \dots & a_{i,n} \\ & & & a_{i+1,i+1} & \dots & a_{i+1,n} \\ & & & & \ddots & * \\ & & & & & a_{nn} \end{vmatrix} = a_{i+1,i+1} \dots a_{nn} \begin{vmatrix} a_{11} & \dots & a_{1i} & a_{1,i+1} & \dots & a_{1n} \\ & \ddots & * & \dots & \dots & * \\ & & 0 & a_{i,i+1} & \dots & a_{i,n} \\ & & & 1 & \dots & * \\ & & & & \ddots & * \\ & & & & & 1 \end{vmatrix}$$

A partir de aquí aplicamos operaciones elementales fila para poner ceros por encima de los pivotes que aparecen en las últimas filas y tendríamos

$$\left| \begin{array}{cccccc} a_{11} & \dots & a_{1i} & a_{1,i+1} & \dots & a_{1n} \\ & \ddots & * & \dots & \dots & * \\ & & 0 & a_{i,i+1} & \dots & a_{i,n} \\ & & & a_{i+1,i+1} & \dots & a_{i+1,n} \\ & & & & \ddots & * \\ & & & & & a_{nn} \end{array} \right| = a_{i+1,i+1} \dots a_{nn} \left| \begin{array}{cccccc} a_{11} & \dots & a_{1i} & \dots & & \\ & \ddots & * & \dots & \dots & * \\ & & 0 & 0 & \dots & 0 \\ & & & 1 & \dots & \\ & & & & \ddots & \\ & & & & & 1 \end{array} \right|$$

Como la matriz que queda tiene una fila de ceros el determinante es cero que es el producto de los elementos de la diagonal pues la diagonal tiene un cero.

Finalmente si  $A$  es una matriz triangular inferior entonces  $A^T$  es una triangular superior. Aplicando el Corolario 3.4.4 y que la proposición se verifica para  $A^T$  tenemos que  $\det(A) = \det(A^T)$  y esto es el producto de los elementos de la diagonal de  $A^T$  que son los mismos que los de  $A$ . ■

Una submatriz de una matriz es una matriz obtenida eliminando varias filas y columnas de la matriz inicial o alternativamente quedándonos con las entradas de la matriz original que están en unas filas y columnas determinadas a priori. Por ejemplo, si partimos de una matriz

$$\begin{pmatrix} 4 & 1 & 0 & -2 \\ 0 & 3 & -3 & 1 \\ -19 & i & 9 & 1+i \end{pmatrix}$$

las siguientes son submatrices suyas:

$$\begin{pmatrix} 4 & 1 & 0 & -2 \\ -19 & i & 9 & 1+i \end{pmatrix} \quad \begin{pmatrix} 1 & -2 \\ 3 & 1 \\ i & 1+i \end{pmatrix} \quad \begin{pmatrix} 0 \\ 1 \\ 1+i \end{pmatrix} \quad \begin{pmatrix} 4 & -2 \\ -19 & 1+i \end{pmatrix}$$

**Notación 3.4.7** Dada una matriz  $A = (a_{ij}) \in M_n(K)$ , vamos a denotar por  $\bar{a}_{ij} \in M_{n-1}(K)$  a la submatriz que se obtiene eliminando en  $A$  la fila  $i$  y la columna  $j$ .

Se llama *menor complementario*  $(i, j)$  de  $A$  a

$$A_{ij} = (-1)^{i+j} \det \bar{a}_{ij}.$$

La matriz

$$\text{adj}(A) = (A_{ij}) \in M_n(K)$$

se llama *matriz adjunta* de  $A$ .

Por ejemplo, si

$$(a_{ij}) = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

entonces

$$\bar{a}_{12} = \begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix} \quad \text{y} \quad \bar{a}_{23} = \begin{pmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{pmatrix}.$$

Obsérvese que los signos que se utilizan para calcular  $A_{ij}$  van alternándose como se muestra en el siguiente dibujo

$$\begin{pmatrix} + & - & + & - \\ - & + & - & + \\ + & - & + & - \\ - & + & - & + \end{pmatrix}.$$

La siguiente propiedad se llama *desarrollo del determinante por filas*.

$$\det(a_{ij}) = a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in}. \quad (3.5)$$

**Demostración.** Pongamos  $A = (a_{ij})$ .

El resultado es consecuencia del Lema 3.4.1 obvio si  $a_i = 0$ .

Supongamos que  $i = 1$ ,  $a_{11} = 1$  y  $a_{1j} = 0$  para todo  $j \neq 1$ . O sea  $A$  tiene la siguiente descomposición en bloques:

$$A = \left( \begin{array}{c|c} 1 & 0 \\ * & \bar{a}_{11} \end{array} \right)$$

y tenemos que demostrar que  $\det(A) = A_{11} = \det(\bar{a}_{11})$ . Para ello miramos  $a_{11} = 1$  como un pivote que usamos para poner cero debajo de él con operaciones elemental del tercer tipo. Como estas operaciones no afectan al determinante podemos suponer que  $a_{i1} = 0$  para todo  $i \neq 1$ . O sea podemos suponer que  $A$  tiene la siguiente forma:

$$A = \left( \begin{array}{c|c} 1 & \\ \hline & \bar{a}_{11} \end{array} \right)$$

Si continuamos con el Método de Gauss-Jordan estamos modificando la matriz inicial de la misma manera que la matriz  $\bar{a}_{ij}$  o sea después de varias operaciones elementales cambiamos  $A$  por una matriz fuertemente escalonada

$$B = \left( \begin{array}{c|c} 1 & \\ \hline & \bar{b}_{11} \end{array} \right)$$

y para un cierto escalar  $\alpha$  se tiene  $\det(A) = \det(B)$  y  $\det(\bar{a}_{11}) = \alpha \det(\bar{b}_{11})$ . Ahora hay dos opciones:  $\bar{b}_{11}$  es la matriz identidad o no. En el primer caso también  $B$  es la matriz identidad y en el segundo caso tanto  $\bar{b}_{11}$  como  $B$  tienen una fila de ceros. En ambos casos  $\det(B) = \det(\bar{b}_{11})$  y por tanto  $\det(A) = \alpha \det(B) = \alpha \det(\bar{b}_{11}) = \det(\bar{a}_{11})$ , como queríamos demostrar en este caso.

Supongamos ahora que  $i = 1$  y  $a_{1j} = 0$  para todo  $j \neq 1$ . O sea  $A$  tiene una descomposición en bloques de la siguiente forma:

$$A = \left( \begin{array}{c|c} a_{11} & 0 \\ * & \bar{a}_{11} \end{array} \right)$$

Aplicando (Det3) y lo que hemos visto en el párrafo anterior tenemos

$$\det(A) = a_{11} \left| \begin{array}{c|c} 1 & 0 \\ * & a_{11} \end{array} \right| = a_{11}A_{11},$$

lo que demuestra el resultado en este caso.

Sigamos manteniendo la hipótesis de que  $i = 1$  y que la fila primera solo tenga una entrada distinta de cero pero ahora admitimos que esa entrada sea cualquiera. O sea  $A$  tiene la forma

$$A = \left( \begin{array}{c|c|c} 0 & a_{1i} & 0 \\ A_1 & b & A_2 \end{array} \right).$$

Intercambiando en  $i - 1$  pasos la fila  $i$  con la anterior, después con la anterior, y así sucesivamente obtendremos una matriz

$$B = \left( \begin{array}{c|c|c} a_{1i} & 0 & 0 \\ b & A_1 & A_2 \end{array} \right) = \left( \begin{array}{c|c} a_{1i} & 0 \\ b & A_{ii} \end{array} \right).$$

Como hemos hecho  $i - 1$  intercambios de columnas y usando que ya sabemos que se cumple la propiedad para  $B$  tenemos que

$$\det(A) = (-1)^{i-1} \det(B) = (-1)^{1+i} a_{1i} A_{1i}.$$

Supongamos ahora que  $i = 1$  sin poner condiciones sobre la primera fila. Aplicando (Det4) y que ya sabemos que se cumple la propiedad para la primera fila con  $a$  lo sumo una entrada distinta de cero tenemos

$$\det(A) = \left| \frac{a_{11}0 \dots 0}{B} \right| + \left| \frac{0a_{12}0 \dots 0}{B} \right| + \dots + \left| \frac{0 \dots 0a_{1n}}{B} \right| = a_{11}A_{11} + \dots + a_{1n}A_{1n}$$

Finalmente podemos subir la fila  $i$  a la primera fila mediante  $i - 1$  intercambios de fila para obtener una matriz  $B$  en la que  $b_{1i} = a_{1i}$  y  $b_{1j} = \overline{a_{1j}}$ . Esto último implica que  $A_{ij} = (-1)^{i+1} B_{1j}$ . Como para pasar de  $A$  a  $B$  hemos realizado  $i - 1$  intercambios de fila y aplicando el resultado para el caso en que la fila es la primera tenemos que

$$\det(A) = (-1)^{i-1} \det(B) = (-1)^{i+1} (b_{11}B_{11} + \dots + b_{1n}B_{1n}) = a_{11}A_{11} + \dots + a_{1n}A_{1n}.$$

■

Aplicando el principio de que para determinantes lo que valga para filas vale para columnas tenemos que también funciona el desarrollo por columnas:

$$\det(a_{ij}) = a_{1i}A_{1i} + a_{2i}A_{2i} + \dots + a_{ni}A_{ni}.$$

Usando el desarrollo por filas se pueden obtener fórmulas para los determinantes de tamaño pequeño partiendo de los de tamaño  $1 \times 1$ :

$$\det(a) = a.$$

Aplicando el Teorema 3.5 obtenemos fórmulas para los determinantes de matriz cuadradas con 2 ó 3 filas:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

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

Esta última fórmula se llama *Regla de Sarrus*.

De hecho se puede usar para calcular cualquier determinante de forma recursiva como en este ejemplo:

$$\begin{aligned} \begin{vmatrix} 0 & 1 & 5 & -5 \\ -2 & -5 & -3 & -4 \\ 3 & 5 & 1 & -1 \\ -6 & -1 & -4 & -3 \end{vmatrix} &= 0 \begin{vmatrix} -5 & -3 & -4 \\ 5 & 1 & -1 \\ -1 & -4 & -3 \end{vmatrix} - 1 \begin{vmatrix} -2 & -3 & -4 \\ 3 & 1 & -1 \\ -6 & -4 & -3 \end{vmatrix} \\ &\quad + 5 \begin{vmatrix} -2 & -5 & -4 \\ 3 & 5 & -1 \\ -6 & -1 & -3 \end{vmatrix} - (-5) \begin{vmatrix} -2 & -5 & -3 \\ 3 & 5 & 1 \\ -6 & -1 & -4 \end{vmatrix} \\ &= 0 \cdot 63 - 1 \cdot (-7) + 5 \cdot (-151) - (-5)(-73) = -1113 \end{aligned}$$

Sin embargo para matrices de más de tres filas suele ser más rápido calcular el determinante utilizando el Método de Gauss-Jordan como ya mostramos al principio de la sección.

**Proposición 3.4.8** *Para toda matriz cuadrada  $A$  tenemos*

$$A \cdot \text{adj}(A)^T = \text{adj}(A)^T \cdot A = |A|I.$$

**Demostración.** Pongamos  $A \in M_n(A)$  y  $\text{adj}(A)^T = (b_{ij})$ . Para demostrar la igualdad  $A \cdot \text{adj}(A) = |A|I$  demostraremos que para todo  $i, j = 1, 2, \dots, n$  se verifica

$$a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{in}b_{nj} = \begin{cases} |A| & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}.$$

La otra igualdad se demuestra de la misma forma.

Obsérvese que  $b_{ij} = A_{ji}$ . De (3.5) tenemos que para cada  $i = 1, 2, \dots, n$ ,

$$a_{i1}b_{1i} + a_{i2}b_{2i} + \dots + a_{in}b_{ni} = a_{i1}A_{i1} + a_{i2}A_{i2} + \dots + a_{in}A_{in} = |A|.$$

Supongamos ahora que  $i \neq j$  y consideremos la matriz  $C$  que se obtiene al sustituir la fila  $j$  de  $A$  por la fila  $i$ . Como  $C = (c_{ij})$  tiene dos filas iguales  $|C| = 0$ . Por otro lado, para todo  $k$  tenemos  $c_{jk} = a_{ik}$  y  $C_{jk} = A_{jk} = b_{kj}$  ya que  $A$  y  $C$  sólo difieren en la fila  $j$ . Entonces

$$a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{in}b_{nj} = c_{j1}C_{j1} + c_{j2}C_{j2} + \dots + c_{jn}C_{jn} = |C| = 0.$$

■

Como consecuencia de la Proposición 3.4.8 tenemos la siguiente fórmula para calcular la inversa de una matriz.

**Corolario 3.4.9** *Si  $A$  es una matriz invertible, entonces*

$$A^{-1} = \frac{\text{adj}(A)^T}{|A|}.$$

**Corolario 3.4.10 (Regla de Cramer)** *Consideremos un sistema de ecuaciones lineales  $Ax = b$  donde  $A \in M_n(K)$  es una matriz invertible. Entonces la única solución de  $A$  es  $x = (x_1, x_2, \dots, x_n)$  donde*

$$x_i = \frac{|B^{(i)}|}{|A|}$$

siendo  $B^{(i)}$  la matriz que se obtiene cambiando la columna  $i$  de  $A$  por  $b$ , es decir

$$B^{(1)} = (b \ a_2 \ \dots \ a_n), B^{(2)} = (a_1 \ b \ \dots \ a_n), \dots, B^{(n)} = (a_1 \ a_2 \ \dots \ b).$$

**Demostración.** Como  $A$  es invertible la única solución del sistema  $Ax = b$  es  $A^{-1}b = \frac{\text{adj}(A)^T b}{|A|}$ .

Por tanto basta ver que si  $B = (b_{kl}) = \text{adj}(A)^T$ , entonces  $Bb = |B^{(i)}|$ .

Observemos que si  $B^{(i)} = (c_{kl})$ , entonces para todo  $j = 1, 2, \dots, n$  tenemos  $B_{ji}^{(i)} = A_{ij}$  y  $c_{ki} = b_i$ , es la  $i$ -ésima coordenada de  $b$ . Desarrollando el determinante  $|B^{(i)}|$  por la columna  $i$  tenemos

$$\begin{aligned} |B^{(i)}| &= c_{1i}B_{1i}^{(i)} + c_{2i}B_{2i}^{(i)} + \dots + c_{ni}B_{ni}^{(i)} \\ &= b_1A_{1i} + b_2A_{2i} + \dots + b_nA_{ni} \end{aligned}$$

que es la  $i$ -ésima coordenada de  $\text{adj}(A)^T b$ . ■

Vamos a ver cómo podemos utilizar los últimos resultados para calcular el rango de una matriz.

**Definición 3.4.11** *Se llama menor de orden  $k$  de una matriz  $A$  al determinante de cualquier submatriz cuadrada de orden  $k$  de  $A$ .*

**Ejemplo 3.4.12** Sea

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}.$$

Los menores de orden 1 de  $A$  son  $a_{11}$ ,  $a_{12}$ ,  $a_{13}$ ,  $a_{21}$ ,  $a_{22}$  y  $a_{23}$ . Los menores de orden 2 son

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21},$$

$$\begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} = a_{11}a_{23} - a_{13}a_{21} \quad \text{y}$$

$$\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} = a_{12}a_{23} - a_{13}a_{22}$$



entonces la matriz obtenida orlando  $B$  con la fila 4 y la columna 2 es la matriz

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{41} & a_{42} & a_{43} \end{pmatrix}.$$

El Teorema 3.4.13 nos permite dar un método para calcular el rango de una matriz:

**Algoritmo 3.4.14 (Cálculo del rango de una matriz usando determinantes)**

Entrada: *Una matriz arbitraria  $A$ .*

- Si  $A = 0$ , entonces la salida del algoritmo es que el rango de  $A$  es 0.
- En caso contrario ponemos
  - $A_1$  la matriz obtenida eliminando las filas y columnas nulas de  $A$ .
  - $r = 1$  y
  - $|M|$  un menor no nulo de  $A$  de orden 1.
- Repetimos lo siguiente hasta que  $M$  tenga tantas filas o tantas columnas como  $A_1$ .
  - Elegimos una fila de  $A_1$  que no corte a  $M$  y vamos calculando los menores obtenidos orlando  $M$  con dicha fila y una a una las columnas de  $A$  que no corten a  $M$ .
  - Si uno de estos menores  $|M_1|$  es distinto de 0, aumentamos  $r$  en una unidad, sustituimos  $M$  por  $M_1$  volvemos al principio del bucle.
  - en caso contrario, o sea si todos los menores que hemos obtenido son cero eliminamos de  $A_1$  la fila utilizada y volvemos al principio del bucle.

Salida:  $r$  que es el rango de la matriz  $A$  y la submatriz  $M$  que proporciona un menor no nulo de tamaño  $r$ .

**Ejemplo 3.4.15** Vamos a calcular el rango de la siguiente matriz en función de los parámetros  $a$ ,  $b$  y  $c$ .

$$A = \begin{pmatrix} -4 & -4 & 0 & 4 & -4 \\ -3 & -3 & -4 & 7 & -15 \\ 5 & -2 & -2 & -10 & 20 \\ -2 & 5 & 1 & 8 & -20 \\ a & a & -a & b & c \end{pmatrix}.$$

Como  $A \neq 0$ , ponemos  $r = 1$  y  $M = (-4)$  que es un menor no nulo de orden 1. Elegimos la primera fila y vamos orlando con las columnas

$$\begin{vmatrix} -4 & -4 \\ -3 & -3 \end{vmatrix} = 0; \quad \begin{vmatrix} -4 & 0 \\ -3 & -4 \end{vmatrix} = 16 \neq 0$$

hasta que encontramos un menor de orden 2 no nulo. Ponemos

$$r = 2, M = \begin{pmatrix} -4 & 0 \\ -3 & -4 \end{pmatrix}.$$

Elegimos la siguiente fila y repetimos el proceso:

$$\begin{vmatrix} -4 & -4 & 0 \\ -3 & -3 & -4 \\ 5 & -2 & -2 \end{vmatrix} = 112 \neq 0.$$

Ponemos por tanto

$$r = 3, M = \begin{pmatrix} -4 & -4 & 0 \\ -3 & -3 & -4 \\ 5 & -2 & -2 \end{pmatrix}.$$

Elegimos ahora la cuarta fila y orlamos con las dos columnas posibles:

$$\begin{vmatrix} -4 & -4 & 0 & 4 \\ -3 & -3 & -4 & 7 \\ 5 & -2 & -2 & -10 \\ -2 & 5 & 1 & 8 \end{vmatrix} = 0, \quad \begin{vmatrix} -4 & -4 & 0 & -4 \\ -3 & -3 & -4 & -15 \\ 5 & -2 & -2 & 20 \\ -2 & 5 & 1 & -20 \end{vmatrix} = 0.$$

Por tanto eliminamos la fila 4 y orlamos con la quinta fila de las dos formas posibles

$$\begin{vmatrix} -4 & -4 & 0 & 4 \\ -3 & -3 & -4 & 7 \\ 5 & -2 & -2 & -10 \\ a & a & -a & b \end{vmatrix} = 112b, \quad \begin{vmatrix} -4 & -4 & 0 & -4 \\ -3 & -3 & -4 & -15 \\ 5 & -2 & -2 & 20 \\ a & a & -a & c \end{vmatrix} = 112(2a + c).$$

Concluimos que

$$r(A) = \begin{cases} 3, & \text{si } b = 2a + c = 0; \\ 4, & \text{en caso contrario.} \end{cases}$$

La Regla de Cramer se puede utilizar en un sentido amplio para resolver sistemas de ecuaciones lineales arbitrarios como muestra el siguiente ejemplo

**Ejemplo 3.4.16** Consideremos el siguiente sistema lineal de ecuaciones

$$\begin{aligned} 4x_1 - 12x_2 - 2x_3 - 20x_4 - 60x_5 &= 8 \\ -4x_1 + 12x_2 - x_3 + 26x_4 + 66x_5 &= -8 \\ -3x_1 + 9x_2 + 3x_3 + 12x_4 + 42x_5 &= -6 \\ 4x_1 - 12x_2 + 3x_3 - 30x_4 - 70x_5 &= 8 \\ 3x_1 - 9x_2 + x_3 - 20x_4 - 50x_5 &= 6 \\ -3x_1 + 9x_2 - 2x_3 + 22x_4 + 52x_5 &= -6 \end{aligned}$$

La matriz ampliada es

$$\left( \begin{array}{ccccc|c} 4 & -12 & -2 & -20 & -60 & 8 \\ -4 & 12 & -1 & 26 & 66 & -8 \\ -3 & 9 & 3 & 12 & 42 & -6 \\ 4 & -12 & 3 & -30 & -70 & 8 \\ 3 & -9 & 1 & -20 & -50 & 6 \\ -3 & 9 & -2 & 22 & 52 & -6 \end{array} \right).$$

Calculando observamos que el rango de la matriz de los coeficientes y el de la ampliada son ambos 2. Por tanto, el sistema es compatible indeterminado con  $5 - 2 = 3$  grados de libertad. Un menor no nulo de orden 2 es el formado por las dos primeras filas y las columnas 1 y 3. Eso implica que las dos primeras ecuaciones tienen el mismo grado de libertad y por tanto podemos eliminar todas las otras ecuaciones y obtener un sistema equivalente. Además pasando las incógnitas que no corresponden a las columnas que aparecen en el menor encontrado podemos ver dicho sistema como uno de Cramer:

$$\begin{aligned} 4x_1 - 2x_3 &= 8 + 12x_2 + 20x_4 + 60x_5 \\ -4x_1 - x_3 &= -8 - 12x_2 - 26x_4 - 66x_5 \end{aligned}$$

al que podemos aplicar la Regla de Cramer:

$$\begin{aligned} x_1 &= \frac{\begin{vmatrix} 8 + 12x_2 + 20x_4 + 60x_5 & -2 \\ -8 - 12x_2 - 26x_4 - 66x_5 & -1 \end{vmatrix}}{\begin{vmatrix} 4 & -2 \\ -4 & -1 \end{vmatrix}} \\ &= -\frac{1}{12} \left( \begin{vmatrix} 8 & -2 \\ -8 & -1 \end{vmatrix} + \begin{vmatrix} 12 & -2 \\ -12 & -1 \end{vmatrix} x_2 + \begin{vmatrix} 20 & -2 \\ -26 & -1 \end{vmatrix} x_4 + \begin{vmatrix} 60 & -2 \\ -66 & -1 \end{vmatrix} x_5 \right) \\ &= 2 + 3x_2 + 6x_4 + 16x_5 \\ \\ x_3 &= \frac{\begin{vmatrix} 4 & 8 + 12x_2 + 20x_4 + 60x_5 \\ -4 & -8 - 12x_2 - 26x_4 - 66x_5 \end{vmatrix}}{\begin{vmatrix} 4 & -2 \\ -4 & -1 \end{vmatrix}} \\ &= -\frac{1}{12} \left( \begin{vmatrix} 4 & 8 \\ -4 & -8 \end{vmatrix} + \begin{vmatrix} 4 & 12 \\ -4 & -12 \end{vmatrix} x_2 + \begin{vmatrix} 4 & 20 \\ -4 & -26 \end{vmatrix} x_4 + \begin{vmatrix} 4 & 60 \\ -4 & -66 \end{vmatrix} x_5 \right) \\ &= 2x_4 + 2x_5. \end{aligned}$$

El ejemplo anterior se puede generalizar a cualquier sistema. Más concretamente, una forma de resolver un sistema de ecuaciones lineales  $AX = b$  es aplicar el siguiente

**Algoritmo 3.4.17 (Resolución de un SELs usando determinantes)**

Entrada: Una matriz  $(A|b) = (a_{ij}|b)$  que representa el sistema de ecuaciones lineales  $AX = b$ .

- (1) Se busca un menor no nulo  $M$  de orden máximo en la matriz de coeficientes  $A$ . El tamaño de ese menor  $M$  será el rango de  $A$ .
- (2) Se busca un menor no nulo más grande en la matriz del sistema  $(A|b)$  orlándolo con la última columna  $b$  y las filas que no cortan a  $M$ . Si existe el rango de  $(A|b)$  será más grande que el de  $M$  y el sistema será incompatible. Es decir si esto pasa el algoritmo proporciona la siguiente Salida: El sistema es incompatible.
- (3) En caso contrario el sistema es compatible. Sean  $j_1, \dots, j_r$  las columnas de  $A$  que cortan a  $M$  y  $k_1, \dots, k_{n-r}$  las otras columnas.

- Descartamos las filas que no cortan a  $M$  y con la matriz que queda construimos un sistema de Cramer

$$M \begin{pmatrix} X_{j_1} \\ \vdots \\ X_{j_k} \end{pmatrix} = b - \sum_{i=1}^{n-r} a_{.k_i} X_{k_i}.$$

- Aplicando la Regla de Cramer o el Método de Gauss-Jordan resolvemos el sistema  $MY = c$  que expresa las variables  $X_{j_1}, \dots, X_{j_r}$  en función de las variables  $X_{k_1}, \dots, X_{k_{n-r}}$ , es decir, en este caso el sistema proporciona la siguiente:

$$\text{Salida : } \begin{cases} X_{j_1} = c_1 + d_{1,1}X_{k_1} + \dots + d_{1,n-r}X_{k_{n-r}} = \frac{\det(N_1)}{\det(M)} \\ \vdots \\ X_{j_r} = c_r + d_{r,1}X_{k_1} + \dots + d_{r,n-r}X_{k_{n-r}} = \frac{\det(N_r)}{\det(M)} \end{cases}$$

donde  $N_i$  es la matriz que se obtiene reemplazando en  $M$  la columna  $i$ -ésima por  $c$ .

También nos resultan útiles los determinantes para discutir sistemas de ecuaciones lineales que dependen de uno o más parámetros.

**Ejemplo 3.4.18** Vamos a discutir el siguiente sistema de ecuaciones lineales:

$$\begin{aligned} x + y + z &= 1 \\ x + ay + z &= a \\ x + y + a^2z &= a^3 \end{aligned}$$

Si calculamos el determinante de la matriz de los coeficientes obtenemos

$$\begin{vmatrix} 1 & 1 & 1 \\ 1 & a & 1 \\ 1 & 1 & a^2 \end{vmatrix} = 1 - a - a^2 + a^3 = (a - 1)^2(a + 1).$$

Por tanto el rango de la matriz de los coeficientes es 3, salvo que  $a = 1$  ó  $a = -1$ . Por tanto, si  $a \neq 1, -1$ , entonces el sistema es compatible determinado y aplicando la Regla de Cramer podemos obtener que la única solución del sistema es

$$x = \frac{a^3 + a}{a^2 - 1}, \quad y = 1, \quad z = -\frac{a^3 + a}{a^2 - 1}.$$

Si  $a = 1$ , entonces el sistema es evidentemente compatible indeterminado y equivalente a la siguiente ecuación

$$x + y + z = 1.$$

Si  $a = -1$ , entonces la matriz de los coeficientes tiene rango 2 siendo un menor no nulo el formado por las dos primeras filas y las dos primeras columnas. Orlando con la última fila y la última columna de la matriz ampliada tenemos

$$\begin{vmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{vmatrix} = 4 \neq 0.$$

Por tanto la matriz ampliada tiene rango 3 y por el Teorema de Rouche-Frobenius, el sistema es incompatible.

Una alternativa para calcular bases de los espacios fila, columna y nulo es modificar ligeramente los algoritmos 3.4.14 y 3.4.17. Más concretamente:

**Algoritmo 3.4.19 (Espacios fila, columna y nulo: Determinantes)**

Entrada: Una matriz  $A \in M_{m,n}(K)$ .

Aplicamos el Algoritmo 3.4.14 que nos proporciona el rango  $r$  de la matriz  $A$  y una submatriz suya  $M$  con determinante diferente de cero.

- Base del espacio fila: Las filas de  $A$  que cortan a  $M$  forman una base del espacio fila de  $A$ .
- Base del espacio columna: Las columnas de  $A$  que cortan a  $M$  forman una base del espacio columna.
- Ecuaciones del espacio fila: Construimos una matriz  $\begin{pmatrix} F \\ X \end{pmatrix} \in M_{r+1,n}(F)$  donde las primeras  $r$  filas son las filas de  $A$  que cortan a  $M$  y  $X = (x_1, \dots, x_n)$ , el vector con las indeterminadas. Orlamos la matriz  $M$  con las  $n - r$  formas posibles y obtenemos  $n - r$  submatrices  $B_1, \dots, B_{n-r} \in M_{r+1}(K)$ , en las que la última fila está formada por  $r + 1$  de las indeterminadas. Las ecuaciones de el espacio fila de  $A$  son

$$|B_1| = 0, \dots, |B_{n-r}| = 0.$$

- Ecuaciones del espacio columna: Construimos una matriz  $\begin{pmatrix} C & X^T \end{pmatrix} \in M_{m,r+1}(F)$  donde las primeras  $r$  columnas son las columnas de  $A$  que cortan a  $M$  y  $X = (x_1, \dots, x_m)$ , el vector con las indeterminadas. Orlamos la matriz  $M$  con las  $m - r$  formas posibles y obtenemos  $m - r$  submatrices  $C_1, \dots, C_{m-r} \in M_{r+1}(K)$ , en las que la última columna está formada por  $r + 1$  de las indeterminadas. Las ecuaciones de el espacio columna de  $A$  son

$$|C_1| = 0, \dots, |C_{m-r}| = 0.$$

- Espacio nulo:

Aplicando el Algoritmo 3.4.17 con entrada  $(A|0)$  proporciona como salida una lista de ecuaciones

$$\begin{aligned} X_{j_1} &= d_{1,1}X_{k_1} + \dots + d_{1,n-r}X_{k_{n-r}} = \frac{\det(N_1)}{\det(M)} \\ &\vdots \\ X_{j_r} &= d_{r,1}X_{k_1} + \dots + d_{r,n-r}X_{k_{n-r}} = \frac{\det(N_r)}{\det(M)} \end{aligned}$$

Una base del espacio nulo está formada por los vectores  $v_1, \dots, v_{n-r}$  donde las coordenadas  $k_1, \dots, k_{n-r}$  de  $v_i$  son las del  $i$ -ésimo vector de la base canónica de  $K^{n-r}$ , o sea la  $k_i$ -ésima coordenada de  $v_i$  es 1 y las otras  $k_j$ -ésimas coordenadas son cero, y las coordenadas  $j_1, \dots, j_r$  se obtienen con las ecuaciones anteriores. O sea la  $j_t$ -ésima coordenada de  $v_i$  es  $d_{t,k_i}$ .

Por ejemplo, si  $(j_1, \dots, j_r) = (1, \dots, r)$  entonces

$$\begin{aligned} v_1 &= (d_{11}, \dots, d_{r,1}, 1, 0, \dots, 0), \\ v_2 &= (d_{12}, \dots, d_{r,2}, 0, 1, \dots, 0), \\ &\vdots \\ v_{n-r} &= (d_{1,n-r}, \dots, d_{r,n-r}, 0, 0, \dots, 1). \end{aligned}$$

**Ejemplo 3.4.20** En el Ejemplo 3.4.15 hemos obtenido el rango de la siguiente matriz

$$A = \begin{pmatrix} -4 & -4 & 0 & 4 & -4 \\ -3 & -3 & -4 & 7 & -15 \\ 5 & -2 & -2 & -10 & 20 \\ -2 & 5 & 1 & 8 & -20 \\ a & a & -a & b & c \end{pmatrix}.$$

en función de los parámetros. Para ello primero hemos obtenido un menor no nulo con las tres primeras filas y las tres primeras columnas. Cuando orlamos este menor con la cuarta fila y todas las columnas siempre obtuvimos menores iguales a 0 lo que implica que la cuarta fila es combinación lineal de las otras 3. Orlando con la última fila observamos que si  $b = 2a + c = 0$ , entonces  $r(A) = 3$ , con lo que las primeras tres filas forman una base del espacio fila y las primeras tres columnas forman una base del espacio columna.

En caso contrario, es decir si  $b \neq 0$  ó  $2a + c \neq 0$  entonces  $r(A) = 4$  y una base del espacio fila se obtiene eliminando la fila 4. Si  $b \neq 0$ , entonces un menor de tamaño 4 distinto de cero se obtiene con las cuatro primeras columnas y eliminando la cuarta fila. Por tanto en este caso las cuatro primeras columnas forman una base del espacio columna. En caso contrario, es decir si  $b = 0$  y  $2a + c \neq 0$  entonces encontramos una base del espacio columna de  $A$  eliminando la columna cuarta pues es el menor de orden 4 formado eliminando la fila 4 y la columna 4 es distinto de 0.

**Algoritmo 3.4.21 (Subespacio generado por vectores de  $K^n$ : Determinantes)**

Entrada: Una lista  $v_1, \dots, v_m$  de vectores de  $K^n$ .

- Construir una matriz  $(A|x)$  donde  $A = (v_1^T, \dots, v_m^T)$ , la matriz formada poniendo en columna los vectores  $v_1, \dots, v_m$  y  $x$  es un vector de indeterminadas  $x_1, \dots, x_m$ .
- Aplicar el Algoritmo 3.4.14 a la matriz  $A$  que proporciona el rango  $r$  de  $A$  y una submatriz  $M$  de  $A$  con determinante no nulo y de tamaño máximo con esta propiedad.
- Construir las  $m - r$  matrices cuadradas  $M_1, \dots, M_{m-r}$  que se obtienen al orlar la matriz  $M$  con una fila adicional y la última columna de indeterminadas:

Salida: El sistema de ecuaciones

$$\begin{aligned} \det(M_1) &= 0 \\ &\dots = 0 \\ \det M_{m-r} &= 0 \end{aligned}$$

Los vectores que están  $\langle v_1, \dots, v_m \rangle$  son las soluciones de dicho sistema.

Obsérvese que al desarrollar los determinantes  $\det(M_i)$  por la última columna se obtienen ecuaciones lineales con lo que las ecuaciones del último algoritmo forman un sistema de ecuaciones lineales.

**Ejemplo 3.4.22** En el Ejemplo 3.3.13 hemos calculado ecuaciones del subespacio generado por los vectores  $(1, 2, 4, 0, 2)$ ,  $(-1, 2, 0, -4, 1)$ ,  $(-1, 6, 4, 0, -1)$  usando el Algoritmo 3.3.12. Vamos hacerlo ahora utilizando el Algoritmo 3.4.21. Para ello ponemos los tres vectores como columnas de una matriz y le añadimos una columna de indeterminadas:

$$\begin{pmatrix} 1 & -1 & -1 & x_1 \\ 2 & 6 & 4 & x_2 \\ 4 & 0 & 4 & x_3 \\ 0 & -4 & 0 & x_4 \\ 2 & 1 & -1 & x_5 \end{pmatrix}$$

Buscamos una submatriz con determinante diferente de cero lo más grande entre las tres primeras columnas. La formada por las dos primeras filas y la última cumple esta propiedad tiene determinante diferente de cero y por tanto cumple la propiedad deseada. Orlando esta matriz de las dos formas posibles tenemos las ecuaciones:

$$\begin{vmatrix} 1 & -1 & -1 & x_1 \\ 2 & 2 & 6 & x_2 \\ 4 & 0 & 4 & x_3 \\ 2 & 1 & -1 & x_5 \end{vmatrix} = 0, \quad \begin{vmatrix} 1 & -1 & -1 & x_1 \\ 2 & 2 & 6 & x_2 \\ 0 & 0 & 0 & x_4 \\ 2 & 1 & -1 & x_5 \end{vmatrix} = 0$$

Desarrollando por la última columna obtenemos las siguientes ecuaciones

$$\begin{vmatrix} 2 & 2 & 6 \\ 4 & 0 & 4 \\ 2 & 1 & -1 \end{vmatrix} x_1 - \begin{vmatrix} 1 & -1 & -1 \\ 4 & 0 & 4 \\ 2 & 1 & -1 \end{vmatrix} x_2 + \begin{vmatrix} 1 & -1 & -1 \\ 2 & 2 & 6 \\ 2 & 1 & -1 \end{vmatrix} x_3 - \begin{vmatrix} 1 & -1 & -1 \\ 2 & 2 & 6 \\ 4 & 0 & 4 \end{vmatrix} x_5 = 0,$$

$$\begin{vmatrix} 2 & 2 & 6 \\ 0 & 0 & 0 \\ 2 & 1 & -1 \end{vmatrix} x_1 - \begin{vmatrix} 1 & -1 & -1 \\ 0 & 0 & 0 \\ 2 & 1 & -1 \end{vmatrix} x_2 + \begin{vmatrix} 1 & -1 & -1 \\ 2 & 2 & 6 \\ 2 & 1 & -1 \end{vmatrix} x_4 - \begin{vmatrix} 1 & -1 & -1 \\ 2 & 2 & 6 \\ 0 & 0 & 0 \end{vmatrix} x_4 = 0$$

Calculando los determinantes tenemos

$$40x_1 + 20x_2 - 20x_3 = 0, \quad -20x_4 = 0.$$

que dividiendo por 20 nos proporcionan las mismas ecuaciones que obtuvimos en el Ejemplo 3.3.13.

## 3.5 Problemas

**3.1** Hallar las matrices elementales 3 por 3 por las que es necesario multiplicar para obtener las siguientes operaciones elementales fila.

- (1) Intercambiar las filas 1 y 2.
- (2) Sumar la primera fila multiplicada por 5 a la segunda.
- (3) Multiplicar la segunda fila por  $\frac{1}{2}$ .

¿Y si se tratara de las correspondientes operaciones elementales columna?

**3.2** Demostrar las siguientes igualdades para matrices elementales.

$$\begin{array}{ll} E_{i,j}^{-1} = E_{i,j} & E_{i,j}^T = E_{i,j} \\ E_{\alpha \cdot i}^{-1} = E(\alpha^{-1} \cdot i) & E_{\alpha \cdot i}^T = E_{\alpha \cdot i} \\ E_{i,\alpha,j}^{-1} = E_{i,-\alpha,j} & E_{i,\alpha,j}^T = E_{j,\alpha,i} \end{array}$$

**3.3** Calcular una forma fuertemente escalonada de las siguientes matrices y de sus transpuestas:

$$\begin{pmatrix} 1 & 2 & -1 & 1 & 1 \\ 3 & -1 & 0 & 1 & -1 \\ 6 & 1 & 0 & 1 & 1 \\ 1 & -2 & 2 & -2 & 0 \\ 11 & 0 & 1 & 1 & 1 \end{pmatrix}, \begin{pmatrix} -24 & 107 & 166 & 309 & -36 \\ 72 & -320 & -496 & -924 & 108 \\ 840 & -3488 & -5296 & -10026 & 1314 \\ -240 & 985 & 1490 & 2829 & -378 \end{pmatrix}$$

$$\begin{pmatrix} 0 & -2 & -4 & 2 & -14 & -44 & -2 & -98 \\ -4 & -3 & -26 & -4 & -127 & -382 & 0 & -782 \\ 1 & -4 & -3 & 0 & -13 & -55 & 3 & -110 \\ 1 & 4 & 13 & -4 & 51 & 153 & -4 & 323 \\ -3 & -2 & -19 & 4 & -79 & -223 & -1 & -474 \\ -2 & 5 & 0 & 1 & 1 & 28 & 3 & 57 \\ 3 & -2 & 11 & -2 & 47 & 119 & 2 & 255 \end{pmatrix}.$$

**3.4** Hallar la inversa, si existe, de las siguientes matrices:

$$\begin{pmatrix} 2 & 5 & -1 \\ 4 & -1 & 2 \\ 6 & 4 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 2 & 0 & 3 \end{pmatrix}, \begin{pmatrix} -3 & 2 & -4 & -3 \\ -2 & 0 & 3 & -2 \\ 4 & -5 & 0 & 4 \\ 0 & 0 & -1 & -2 \end{pmatrix}, \begin{pmatrix} -5 & 3 & -5 & 3 & -2 \\ 1 & -1 & 1 & -6 & -2 \\ -6 & -1 & 5 & 1 & 3 \\ 3 & -3 & 2 & 1 & -1 \\ 5 & -1 & 3 & -2 & 4 \end{pmatrix}.$$

**3.5** ¿Cuáles de los siguientes puntos pertenecen  $\pi = \langle (1, 2, 3, 0), (3, 1, 5, 2) \rangle$ ?:

- (a)  $(1, 0, 3, 1)$ , (b)  $(5, 5, 11, 2)$ , (c)  $(0, 4, -1, -4)$ .

Dar un conjunto de ecuaciones que sirvan para determinar si un punto genérico  $(x_1, x_2, x_3, x_4)$  está en  $\pi$ .

**3.6** Encontrar, si existen, las soluciones de los siguientes sistemas de ecuaciones:

$$\left. \begin{array}{l} 4x_1 - 3x_2 = 10 \\ 8x_1 + x_2 = 10 \end{array} \right\} \qquad \left. \begin{array}{l} 2x - y = 8 \\ 6x - 5y = 32 \end{array} \right\}$$

$$\left. \begin{array}{l} y + z = 6 \\ 3x - y + z = -7 \\ x + y - 3z = -13 \end{array} \right\} \qquad \left. \begin{array}{l} x_1 - 4x_2 + x_3 = 8 \\ 3x_1 - 12x_2 + 5x_3 = 26 \\ 2x_1 - 9x_2 - x_3 = 14 \end{array} \right\}$$

$$\left. \begin{array}{l} y + z = 1 \\ 3x - y + z = 2 \\ x + y - 3z = 3 \end{array} \right\} \qquad \left. \begin{array}{l} x_1 - 3x_2 + x_3 + 2x_4 = 2 \\ x_1 - 2x_2 + 2x_3 + 4x_4 = -1 \\ 2x_1 - 8x_2 - x_3 = 3 \\ 3x_1 - 9x_2 + 4x_3 = 7 \end{array} \right\}$$

$$\left. \begin{array}{l} x - y + 2z = 3 \\ y - 4z = 2 \end{array} \right\} \qquad \left. \begin{array}{l} x_1 - 3x_2 + 2x_3 - x_4 = 8 \\ 3x_1 - 7x_2 + x_4 = 0 \end{array} \right\}$$

$$\left. \begin{array}{l} x - 3y + z = 2 \\ 3x - 8y + 2z = 5 \\ 3x - 7y + z = 1 \end{array} \right\} \qquad \left. \begin{array}{l} x - 2y = 1 \\ 2x - y = 0 \\ 3x - 2y = 2 \end{array} \right\}$$

$$\left. \begin{array}{l} 2x + y + 4z = 0 \\ x + 3y + 2z = 0 \\ 3x - y + 6z = 0 \end{array} \right\}$$

**3.7** El siguiente problema ha sido hallado en un texto babilonio de aproximadamente el año 300 AC (es decir, más de 2000 años de antigüedad). Hemos modificado las unidades de medida para que se entienda.

Hay dos campos cuya área total es de 1500 m<sup>2</sup>. Uno produce grano a razón de 28 kg por m<sup>2</sup> y el otro a razón de 21 kg por m<sup>2</sup>. El total producido en los dos campos es de 38500 kg. ¿Cuál es el tamaño de cada campo?

**3.8** Un típico problema de los antiguos chinos, tomado del texto de la dinastía Han "Nueve capítulos del arte de las matemáticas" (alrededor de 200 AC) plantea:

Hay tres clases de grano. Tres gavillas de la primera clase, dos de la segunda y una de la tercera hacen 39 medidas. Dos de la primera, tres de la segunda y una de la tercera hacen 34 medidas. Y una de la primera, dos de la segunda y tres de la tercera hacen 26 medidas. ¿Cuántas medidas de grano están contenidas en una gavilla de cada clase?

**3.9** Un fabricante tiene a la venta dos mezclas de hierbas tranquilizantes: la mezcla A y la B. Para hacer sus mezclas utiliza dos tipos de hierba: la tipo 1 y la tipo 2.

Para fabricar un kilo de mezcla A utiliza medio kilo de hierba tipo 1 y medio kilo de tipo 2. Para fabricar un kilo de mezcla B, utiliza un tercio de kilo de tipo 1 y dos tercios de kilo de tipo 2.

- (1) ¿Cuántos kilos de mezcla A y B puede fabricar si tiene 70 kilos de hierba de tipo 1 y 80 kilos de hierba de tipo 2, de tal manera que se use toda la hierba de ambos tipos?
- (2) El mismo fabricante, como es muy caprichoso, quiere ahora fabricar dos nuevas mezclas C y D usando las dos anteriores A y B. Para fabricar un kilo de mezcla C utiliza medio kilo de mezcla A y medio kilo de mezcla B. Para fabricar un kilo de mezcla D utiliza  $\frac{4}{5}$  kilo de mezcla A y  $\frac{1}{5}$  kilo de mezcla B. ¿Cuántos kilos de mezcla C y D puede fabricar con los mismos kilos de hierba tipo 1 y tipo 2, también usando toda la hierba?

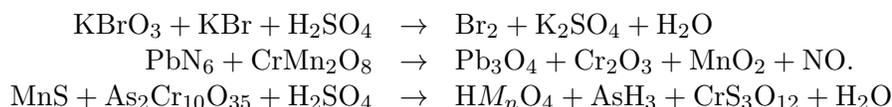
**3.10** Sobre un nuevo compuesto químico desarrollado por un gobierno enemigo nuestros espías averiguaron que tiene tres elementos: uranio, plutonio y manganeso. También lograron averiguar que una muestra de un gramo tiene las siguientes proporciones: la cantidad de plutonio y de manganeso son el doble que la de uranio, y hay el doble de plutonio que de manganeso. ¿Qué cantidad de uranio, plutonio y manganeso tiene un gramo de compuesto?

**3.11** Una empresa quiere comprar tres productos distintos,  $P_1$ ,  $P_2$ ,  $P_3$ . Las condiciones de la compra son:

- Por un problema de almacenamiento, se tienen que comprar exactamente 100 unidades en total,
- Se tiene que comprar al menos una unidad de cada producto.
- El presupuesto es de 100 000 pts.
- Los precios son:  $P_1$  cuesta 50 pts.,  $P_2$  cuesta 1000 pts., y  $P_3$  cuesta 5000 pts.

¿Cuántas unidades de cada producto se pueden comprar?

**3.12** Ajustar las tres siguientes reacciones químicas



**3.13** Sean  $A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$  y  $B = \begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix}$ . Se pide:

- (1) Comprobar que  $AB = BA$ .
- (2) Encontrar todas las matrices  $X \in M_{2,2}(K)$  tales que  $AX = XA$ .

**3.14** Dada  $A = \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix}$ , encontrar todas las matrices  $X = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix}$ , tales que  $AX = 0$ .

**3.15** Sean  $A = \begin{pmatrix} 3 & 2 \\ 2 & 1 \end{pmatrix}$ ,  $B = \begin{pmatrix} 1 & -2 & 0 & 1 \\ 4 & 1 & 3 & 2 \end{pmatrix}$ ,  $C = \begin{pmatrix} -2 & 1 & 3 \\ 3 & 0 & 1 \\ 0 & 4 & -2 \end{pmatrix}$  y  $D = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 0 & 1 \\ 0 & 1 & -2 \end{pmatrix}$ .

Hallar, si existen, las matrices  $X$  que verifican: (a)  $AX = B$ , (b)  $DXC = D$ .

**3.16** Resolver la ecuación de matrices  $ABX^{-1}C = B^T A^T C^{-1}$  donde  $A, B$  y  $C$  son invertibles.

**3.17** Resolver el siguiente sistema matricial,

$$\begin{aligned}(X + Y)C &= X \\ C - Y &= 3I - 2A\end{aligned}$$

donde  $A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ ,  $B = \begin{pmatrix} 4 & 3 \\ 4 & 6 \end{pmatrix}$ ,  $C = \begin{pmatrix} 1 & 2 \\ 3 & 0 \end{pmatrix}$ ,  $D = \begin{pmatrix} -2 & 9 \\ 6 & 6 \end{pmatrix}$ .

**3.18** Calcular los siguientes determinantes:

$$\begin{vmatrix} 5 & 2 & 1 \\ 1 & -1 & 4 \\ 3 & 0 & 2 \end{vmatrix}, \quad \begin{vmatrix} 3 & 2 & 4 \\ 0 & 1 & 2 \\ 1 & 4 & 1 \end{vmatrix}, \quad \begin{vmatrix} 0 & 1 & 4 \\ 2 & 3 & 1 \\ 1 & 4 & 1 \end{vmatrix}, \quad \begin{vmatrix} 2 & 3 & 4 & 6 \\ 2 & 0 & -9 & 6 \\ 4 & 1 & 0 & 2 \\ 0 & 1 & -1 & 0 \end{vmatrix}, \quad \begin{vmatrix} 2 & 0 & -1 & 7 \\ 6 & 1 & 0 & 4 \\ 8 & -2 & 1 & 0 \\ 4 & 1 & 0 & 2 \end{vmatrix}.$$

**3.19** Sea  $A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{pmatrix}$ . Calcular  $\det(A^5)$  sin calcular  $A^5$ .

**3.20** Suponiendo que  $\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = 5$ . Calcular:

$$\begin{vmatrix} d & e & f \\ g & h & i \\ a & b & c \end{vmatrix}, \quad \begin{vmatrix} -a & -b & -c \\ 2d & 2e & 2f \\ -g & -h & -i \end{vmatrix}, \quad \begin{vmatrix} a+d & b+e & c+f \\ d & e & f \\ g & h & i \end{vmatrix}.$$

**3.21** Sabiendo que  $\begin{vmatrix} x & y & z \\ 5 & 0 & 3 \\ 1 & 1 & 1 \end{vmatrix} = 1$ . Calcular:  $\begin{vmatrix} 3x & 3y & 3z \\ 5 & 0 & 3 \\ 1 & 1 & 1 \end{vmatrix} + y \begin{vmatrix} 5x & 5y & 5z \\ 1 & 0 & \frac{3}{5} \\ 1 & 1 & 1 \end{vmatrix}$ .

**3.22** Demostrar que se verifica la igualdad:  $\begin{vmatrix} a & a & b+c \\ a & b & c+a \\ a & c & a+b \end{vmatrix} = 0$

**3.23** Sean  $A, B \in M_n(K)$  tales que  $\det(A) = 2$ ,  $\det(B) = -3$ . Calcular:  $\det(3A)$ ,  $\det(AB)$ ,  $\det(A^T B)$ ,  $\det(A^{-1}B)$ ,  $\det(B^{-1}A)$ .

**3.24** Utilizando determinantes, calcular la inversa (si existe) de las siguientes matrices.

$$\begin{pmatrix} 2 & 0 \\ 1 & -1 \end{pmatrix}, \quad \begin{pmatrix} 2 & 1 & 1 \\ 0 & 1 & 1 \\ -2 & 1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 3 & 0 & 4 \\ -2 & 1 & 1 \\ 3 & 1 & 2 \end{pmatrix}.$$

**3.25** Obtener una fórmula que relacione  $|\operatorname{adj}(A)|$  y  $|A|$ .

**3.26** Calcular todos los valores de  $\det(A)$  en los siguientes casos: (a)  $A = A^2$ , (b)  $AA^T = I$ , (c)  $A^2 = 0$ .

**3.27** Indicar si las siguientes igualdades son verdaderas o falsas, para  $A \in M_{5,5}(K)$ .

$$|A + B| = |A| + |B|, \quad |A + A| = 32|A|, \quad |A + A| = 2|A|.$$

**3.28** Calcular el rango de las siguientes matrices

$$A = \begin{pmatrix} 2 & -1 & 0 & 1 & 3 \\ 1 & 0 & -1 & 2 & 3 \\ 5 & -2 & -1 & 4 & 9 \end{pmatrix}, B = \begin{pmatrix} 5 & 0 & 1 \\ 1 & 3 & 0 \\ 2 & 6 & 0 \\ 3 & -6 & 1 \\ 4 & -3 & 1 \end{pmatrix}, C = \begin{pmatrix} 1 & -1 & 0 & 3 \\ 3 & 1 & 1 & 2 \\ 5 & 3 & 2 & 1 \\ 5 & -1 & 1 & 8 \end{pmatrix}, E = \begin{pmatrix} 1 & 5 & 2 \\ 1 & 1 & 7 \\ 0 & -3 & 4 \end{pmatrix}.$$

Calcular también el rango de  $DA$  y  $BA$ .

**3.29** ¿Cuales de las siguientes afirmaciones son verdad y cuales falsas?

- (1) El rango de una matriz triangular es el número de elementos no nulos en la diagonal.
- (2) Si  $A$  es invertible entonces  $r(A) \neq 0$ .
- (3)  $r(A + B) = r(A) + r(B)$ .
- (4)  $r(AB) = r(A)r(B)$ .
- (5)  $r(AB) \leq r(A)$ .
- (6)  $r(AB) \leq r(B)$ .
- (7) Si  $A$  tiene rango  $k$  y  $E$  es una matriz elemental entonces  $r(EA) = k$ .
- (8) Un sistema de dos ecuaciones con cuatro incognitas siempre es compatible determinado.
- (9) Un sistema de dos ecuaciones con cuatro incognitas siempre es incompatible.

**3.30** Para cada uno de los siguientes determinantes, hallar los valores de  $x$  para los se anulan.

$$\begin{vmatrix} 1-x & 2 \\ 3 & 2-x \end{vmatrix}, \begin{vmatrix} 2-x & 0 & 0 \\ 0 & 1-x & 4 \\ 0 & 1 & 1-x \end{vmatrix}, \begin{vmatrix} x & 1 & 1 \\ 1 & x & 1 \\ 1 & 1 & x \end{vmatrix},$$

$$\begin{vmatrix} 1+x & 1 & 1 & 1 \\ 1 & 1+x & 1 & 1 \\ 1 & 1 & 1+x & 1 \\ 1 & 1 & 1 & 1+x \end{vmatrix}, \begin{vmatrix} 1-x & 0 & 2 \\ 0 & 2-x & 3 \\ 0 & 4 & -x \end{vmatrix}.$$

**3.31** Resolver los siguientes sistemas de ecuaciones lineales usando el Método de Cramer:

$$\left. \begin{array}{l} x_1 - 2x_2 = 1 \\ 3x_1 - 4x_2 = 3 \end{array} \right\} \quad \left. \begin{array}{l} x_1 + 2x_2 - x_3 = -2 \\ 2x_1 + 2x_2 + x_3 = 0 \\ 3x_1 - x_2 + 5x_3 = 1 \end{array} \right\}$$

$$\left. \begin{array}{l} 5x_1 - 2x_2 + x_3 = 1 \\ \quad \quad x_2 + x_3 = 0 \\ x_1 + 6x_2 - x_3 = 4 \end{array} \right\} \quad \left. \begin{array}{l} 6x_1 + x_2 - x_3 = 4 \\ x_1 - x_2 + 5x_4 = -2 \\ -x_1 + 3x_2 + x_3 = 2 \\ x_1 + x_2 - x_3 + 2x_4 = 0 \end{array} \right\}$$

**3.32** Calcular usando determinantes el rango de las siguientes matrices y un menor no nulo de orden máximo.

$$\left( \begin{array}{cccc} 3 & -5 & 10 & -12 \\ 2 & -3 & 6 & -7 \\ 4 & -1 & 2 & 1 \\ -1 & 2 & -4 & 5 \\ 3 & -2 & 4 & -3 \\ 4 & -4 & 8 & -8 \end{array} \right) \quad \left( \begin{array}{cccccc} 0 & -1 & -5 & 4 & 27 & 140 \\ -5 & 3 & 0 & -2 & 14 & 45 \\ 1 & 2 & 13 & 1 & -62 & -274 \\ 0 & 2 & 10 & 0 & -46 & -208 \\ -3 & -1 & -14 & -3 & 71 & 302 \\ 1 & -3 & -12 & 2 & 54 & 255 \end{array} \right)$$

$$\left( \begin{array}{cccccc} 1 & -5 & -5 & 22 & -1 & -64 & -314 \\ 4 & -20 & -4 & 40 & 4 & -96 & -504 \\ -1 & 5 & -5 & 8 & 1 & -56 & -236 \\ 0 & 0 & 0 & 0 & -4 & 16 & 64 \\ 4 & -20 & 1 & 25 & 4 & -36 & -229 \end{array} \right)$$

**3.33** Discutir los siguientes sistemas de ecuaciones lineales en función de los parámetros.

$$\left. \begin{array}{l} x + 2y - z = a \\ x - ay + z = 2 \end{array} \right\} \quad \left. \begin{array}{l} ax - y + 2z = 1 + a \\ x + ay - z = -1 \\ 3x + y + z = a \end{array} \right\} \quad \left. \begin{array}{l} x + 2y = a \\ 3x - y = 1 \\ x + y = a - 3 \\ 2x - y = a + b \end{array} \right\}$$

$$\left. \begin{array}{l} ax + by + z = 1 \\ x + aby + z = b \\ x + by + az = 0 \end{array} \right\} \quad \left. \begin{array}{l} x + ay - z = a \\ 2x - y - bz = b \end{array} \right\} \quad \left. \begin{array}{l} ax + 2z = 2 \\ 5x + 2y = 1 \\ x - 2y + bz = 3 \end{array} \right\}$$

$$\left. \begin{array}{l} ax + by + z = 1 \\ x + by + az = 0 \\ x + aby + z = -1 \\ -x + 2z = -2 \end{array} \right\} \quad \left. \begin{array}{l} x - 2y = 3k + 3m \\ x - y = 2k + 2m + 1 \\ mx + ky = m^2 - k^2 - 6 \\ kx + my = k^2 - m^2 + 6 \end{array} \right\} \quad \left. \begin{array}{l} ax + y + z + t = 1 \\ x + ay + z + t = b \\ x + y + az + t = b^2 \\ x + y + z + at = b^3 \end{array} \right\}$$

$$\left. \begin{array}{l} ax + y + z = a^2 \\ x - y + z = 1 \\ 3x - y - z = 1 \\ 6x - y + z = 3a \end{array} \right\} \quad \left. \begin{array}{l} (a + 1)x + y + z = a^2 + 3a \\ x + (a + 1)y + z = a^3 + 3a^2 \\ x + y + (a + 1)z = a^4 + 3a^3 \end{array} \right\} \quad \left. \begin{array}{l} 3x + 2y + 2z = 1 \\ x + 4y + z = b \\ 2x - 5y + az = -2 \end{array} \right\}$$

**3.34** Calcular el determinante de la siguiente matriz conocida como matriz de Vandermonde:

$$V(a_0, a_1, a_2, \dots, a_n) = \begin{pmatrix} 1 & a_0 & a_0^2 & \dots & a_0^n \\ 1 & a_1 & a_1^2 & \dots & a_1^n \\ 1 & a_2 & a_2^2 & \dots & a_2^n \\ \dots & \dots & \dots & \dots & \dots \\ 1 & a_n & a_n^2 & \dots & a_n^n \end{pmatrix}.$$

(Indicación: Haz primero los casos en que  $n = 1$  y  $n = 2$  y después razona por inducción de forma que comienzas restando a cada columna la anterior multiplicada por  $a_0$ , después desarrollas por la primera fila y después extraes en cada fila factor común.)

demostrar que  $|V(a_0, a_1, a_2, \dots, a_n)| \neq 0$  si y sólo los  $a_i$  son diferentes dos a dos.

**3.35** Calcular bases y ecuaciones para los espacio fila, columna y nulo de las matrices del Problema 3.3.



# Capítulo 4

## Bases y dimensión

### 4.1 Bases y dimensión

Vamos a recordar un poco de notación y terminología de aplicaciones. Dada una aplicación  $f : A \rightarrow B$ , la imagen de  $f$  es el siguiente conjunto

$$\text{Im}(f) = \{f(a) : a \in A\}.$$

Decimos que  $f$  es *suprayectiva* si  $\text{Im}(f) = B$ , o sea si para todo elemento  $b \in B$  existe un  $a \in A$  tal que  $f(a) = b$ .

Decimos que  $f$  es *inyectiva* si  $A$  no tiene dos elementos distintos con la misma imagen por  $f$ , es decir si siempre que se cumpla  $f(a) = f(a')$  con  $a, a' \in A$  se tiene que cumplir necesariamente que  $a = a'$ .

Finalmente se dice que  $f$  es *biyectiva* si es suprayectiva e inyectiva, o lo que es lo mismo si para todo  $b \in B$  existe un único  $a \in A$  tal que  $f(a) = b$ . En tal caso la aplicación que asocia  $b \in B$  con el único  $a \in A$  para el que  $f(a) = b$  se llama inversa de  $f$  y se denota  $f^{-1}$ .

Si  $S = \{v_1, v_2, \dots, v_n\}$  es una lista de elementos de  $V$ , entonces vamos a considerar la aplicación  $C'_S : K^n \rightarrow V$  definida de la siguiente manera:

$$C'_S(\alpha_1, \alpha_2, \dots, \alpha_n) = \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n.$$

Como consecuencia inmediata de las definiciones 2.3.1, 2.4.1 y 3.3.5 tenemos:

**Proposición 4.1.1** *Sea  $S = \{v_1, v_2, \dots, v_n\}$  una lista de elementos de  $V$ .*

- (1)  $\text{Im}(C'_S) = \langle S \rangle$ .
- (2)  $S$  es un sistema generador de  $V$  si y sólo si  $C'_S$  es suprayectiva.
- (3)  $S$  es linealmente independiente de  $V$  si y sólo si  $C'_S$  es inyectiva.
- (4)  $S$  es base si y solo si  $C'_S$  es biyectiva.

Si  $B$  una base finita de  $V$  entonces  $C'_B$  es una aplicación biyectiva y por tanto tiene una aplicación inversa que vamos a denotar por  $C_B$ . Para cada  $v \in V$  denotamos por  $v_B$  a  $C_B(v)$ , es decir  $v_B$  es el único elemento  $(\alpha_1, \alpha_2, \dots, \alpha_n) \in K^n$  tal que

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n.$$

El vector  $v_B$  se llama *vector de coordenadas de  $v$  en la base  $B$* . Si pensamos en los elementos de  $K^n$  como palabras, podemos interpretar una base como un diccionario que asocia a cada vector una palabra, que sería su nombre en ese “idioma”.

La siguiente proposición muestra la utilidad de las coordenadas.

**Proposición 4.1.2** *Sea  $B$  una base de  $V$ . Entonces para todo  $x, w \in V$  y todo  $\alpha, \beta \in K$  se verifica*

$$(\alpha v + \beta w)_B = \alpha v_B + \beta w_B. \quad (4.1)$$

Además, si  $S = \{v_1, v_2, \dots, v_n\}$  es una lista de  $V$  y  $S_B = \{v_{1B}, v_{2B}, \dots, v_{nB}\}$ , entonces

(1)  $S$  es linealmente dependiente (independiente) si y sólo si lo es  $S_B$ .

(2)  $S$  es sistema generador de  $V$  si y sólo si  $S_B$  es un sistema generador de  $K^n$ .

**Demostración.** Supongamos que  $B = \{x_1, \dots, x_n\}$ ,  $v_B = (\mu_1, \dots, \mu_n)$  y  $w_B = (\tau_1, \dots, \tau_n)$  con lo que

$$v = \mu_1 x_1 + \dots + \mu_n x_n \quad \text{y} \quad w = \tau_1 x_1 + \dots + \tau_n x_n.$$

Entonces

$\alpha v + \beta w = \alpha(\mu_1 x_1 + \dots + \mu_n x_n) + \beta(\tau_1 x_1 + \dots + \tau_n x_n) = (\alpha\mu_1 + \beta\tau_1)x_1 + \dots + (\alpha\mu_n + \beta\tau_n)x_n$  y de ahí se deduce la igualdad (4.1).

Por otro lado, como  $C'_B$  es biyectiva se tiene que si  $\mu_1, \dots, \mu_n \in K$  entonces

$$\mu_1 v_1 + \dots + \mu_n v_n = 0 \quad \Leftrightarrow \quad (\mu_1 v_1 + \dots + \mu_n v_n)_B = 0 \quad \Leftrightarrow \quad \mu_1 v_{1B} + \dots + \mu_n v_{nB} = 0.$$

De aquí se deduce que  $S$  es linealmente independiente si y solo si  $S_B$  lo es. Por otro lado si  $x \in V$  entonces  $v = \mu_1 v_1 + \dots + \mu_n v_n$  si y solo si  $v_B = \mu_1 v_{1B} + \dots + \mu_n v_{nB}$ . Esto demuestra que  $S$  es sistema generador de  $V$  si y solo si  $S_B$  es sistema generador de  $K^n$ . ■

Es decir, el diccionario asociado a una base no sólo sirve para denotar los vectores sino también para trasladar cualquier propiedad, problema, etc entre  $V$  y  $K^n$ .

**Ejemplos 4.1.3** (1) Consideremos la base  $B = \{1, X, X^2, X^3\}$  del espacio vectorial  $V$  de los polinomios de grado  $\leq 3$ . Supongamos que queremos ver si el siguiente conjunto es linealmente dependiente o no.

$$S = \{1 + 2X - 3X^2, -2 + 8X + 7X^2 + X^3, -3 - 6X + X^2 - 5X^3, -3 + 6X + 2X^2 - 4X^3\}.$$

Con la notación del Corolario 4.1.2

$$S_B = \{(1, 2, -3, 0), (-2, 8, 7, 1), (-3, -6, 1, -5), (-3, 6, 2, -4)\}.$$

que en el Ejemplo 2.4.7 vimos que es linealmente dependiente. Por tanto  $S$  también es linealmente dependiente.

(2) Consideramos las funciones

$$\begin{aligned} f_1(x) &= \operatorname{sen}(x) + \operatorname{cos}(x) \\ f_2(x) &= 2 \operatorname{sen}(x) + \operatorname{tan}(x) \\ f_3(x) &= 2 \operatorname{cos}(x) - 3 \operatorname{tan}(x). \end{aligned}$$

En el Ejemplo 2.4.5 vimos que  $B = \{\operatorname{sen}, \operatorname{cos}, \operatorname{tan}\}$  son linealmente independiente y por tanto  $B$  es una base de  $\langle B \rangle$ . Por el Corolario 4.1.2,  $S = \{f_1, f_2, f_3\}$  es un sistema linealmente independiente si y sólo si lo es  $S_B = \{(1, 1, 0), (2, 0, 1), (0, 2, -3)\}$ . El lector puede ver que  $S_B$  es linealmente independiente y por tanto, también lo es  $S$ .

Una vez vista la utilidad de las bases se plantea el problema de si existen. La respuesta es afirmativa. En la demostración del siguiente teorema usaremos la siguiente notación para dos conjuntos  $X$  e  $Y$ :

$$X \setminus Y = \{x \in X : y \notin Y\}.$$

**Teorema 4.1.4** Sean  $S, G \subseteq V$  con  $S$  linealmente independiente y  $G$  un sistema generador. Entonces

- (1) Si  $S \subseteq G$ , entonces existe una base  $B$  tal que  $S \subseteq B \subseteq G$ .
- (2)  $S$  está contenido en una base de  $V$ .
- (3)  $G$  contiene una base de  $V$ .
- (4) Todo espacio vectorial tiene una base.

**Demostración.** (1) Para empezar  $S$  no tiene elementos repetidos y eliminando repeticiones en  $G$  podemos suponer que  $G$  tampoco los tiene. Hacemos la demostración suponiendo que el conjunto  $G$  es finito, de hecho vamos a ver cuatro demostraciones. La demostración en el caso en que  $G$  es infinito es más complicada pero más adelante diremos algunas palabras sobre ella.

Primera demostración<sup>1</sup>: Como  $G$  es finito entre los subconjuntos de  $G$  que contengan a  $S$  y sean linealmente independientes habrá uno que tenga cardinal máximo. Elegimos uno así y lo denotamos por  $B$ . Si  $g \in G \setminus B$  entonces  $B \cup \{g\}$  es linealmente dependiente pues tiene más elementos que  $B$ . Aplicando la Proposición 2.4.8 deducimos que  $g \in \langle B \rangle$ . Esto demuestra que todo elemento de  $G$  está en  $B$  y aplicando la Proposición (2.3.9).(1) deducimos que  $V = \langle G \rangle \subseteq \langle B \rangle$  con lo que  $B$  es un conjunto generador de  $V$ . Como  $B$  es linealmente independiente deducimos que  $B$  es una base de  $V$ .

Segunda demostración: Entre los subconjuntos de  $G$  que contienen a  $S$  y son generadores elegimos uno  $B$  con cardinal mínimo. Si  $B$  es linealmente dependiente, uno de los elementos  $v$  de  $B$  es combinación lineal de los demás, lo que no es posible por que implicaría que  $B \setminus \{v\}$  sea también generador con menos elementos que  $B$ . Por tanto  $B$  es linealmente dependiente y por tanto es una base de  $V$ .

---

<sup>1</sup>Tengo que agradecer a Rubén Cárceles, alumno del curso 2020-21, que sugirió esta demostración. Yo normalmente hago en clase las demostraciones tercera y cuarta porque, aunque son menos elegantes, muestran un método más claro de cómo obtener la base buscada a partir de un conjunto linealmente independiente o de uno generador.

Tercera demostración: Razonaremos por inducción en  $k = |G| - |S|$ . Si  $k = 0$  entonces  $S = G$  y por tanto  $S$  es una base que cumple la propiedad que queremos. Supongamos pues que  $k > 0$ . Como estamos razonando por inducción en  $k$  podemos suponer que la propiedad se verifica siempre que los conjuntos  $S$  y  $G$  del enunciado cumplan  $|G| - |S| < k$ . Esto es lo que llamamos la hipótesis de inducción. Si  $G$  es linealmente independiente, entonces  $G$  es una base de  $V$  y habríamos acabado. En caso contrario existen  $v_1, \dots, v_n \in G$  y escalares, no todos 0 tales que

$$\alpha_1 v_1 + \dots + \alpha_n v_n = 0.$$

Podemos suponer sin pérdida de generalidad que todos los  $\alpha_i$  son distintos de cero. Como  $S$  es linealmente independiente algún  $v_i$  no está en  $S$  y podemos suponer sin pérdida de generalidad que  $v_n \notin S$ . Sea  $G_1 = G \setminus \{v_n\}$ . Como  $\alpha_n \neq 0$  tenemos que

$$v_n = -\alpha_n^{-1} \alpha_1 v_1 - \dots - \alpha_n^{-1} \alpha_{n-1} v_{n-1} \in \langle G_1 \rangle$$

y por tanto  $\langle G_1 \rangle = \langle G \rangle = V$ , o sea  $G_1$  es otro sistema generador de  $G$  y contiene a  $S$  pero con un elemento menos. Por tanto  $|G_1| - |S| = k - 1$  y, por la hipótesis de inducción, tenemos que  $G_1$  contiene una base  $B$  que contiene a  $S$  y por tanto  $G$  también.

Cuarta demostración. Utilizamos la misma notación de la demostración anterior. La idea de la tercera demostración era ir eliminando elementos de  $G$  hasta obtener una base. El proceso ahora es el contrario, se trata de añadir elementos a  $S$  hasta conseguir una base. De nuevo argumentamos por inducción en  $k = |G| - |S|$  y si  $k = 0$  está claro que  $S$  es una base. La hipótesis de inducción de nuevo dice que si  $|G| - |S|$  fuera menor que  $k$  entonces entre  $G$  y  $S$  habría una base.

Si  $S$  ya es una base, no hay nada que hacer. En caso contrario  $\langle S \rangle \neq V = \langle G \rangle$  y, por la Proposición 2.3.9, existe un elemento  $v$  de  $G$  que no está en  $\langle S \rangle$ . Claramente este elemento no es ninguno de los de  $S$  y añadiendo  $v$  a  $S$  tenemos un conjunto más grande  $S_1$  que también será linealmente independiente por la Proposición 2.4.8. Luego  $S \subset S_1 \subseteq G_1$  y es a  $S_1$  y  $G$  a los que le podemos aplicar la hipótesis de inducción pues  $|G| - |S_1| = k - 1$ . Por tanto  $G$  contiene una base que contiene a  $S_1$ , y por tanto también a  $S$ .

(2) y (3) son consecuencias inmediatas de (1), ya que todo conjunto está contenido en un sistema generador (por ejemplo,  $V$ ) y contiene uno linealmente independiente (por ejemplo,  $\emptyset$ ).

Finalmente (4) se deduce de (2) cogiendo  $S = \emptyset$  de (3) cogiendo  $G = V$ . ■

En realidad la demostración del Teorema 4.1.4 que hemos hecho solo funciona en el caso en que el espacio vectorial tenga un conjunto generador finito pues en la demostración de (1) hemos supuesto que  $G$  es finito. Para la demostración del caso en que  $G$  no sea finito hay que utilizar un resultado matemático llamado Lema de Zorn que dice que si tenemos un conjunto ordenado  $X$  en el que cada familia creciente  $x_1 \leq x_2 \leq \dots$  tenga una cota superior entonces  $X$  tiene un elemento que no tiene ninguno estrictamente mayor. La idea de la demostración consiste en ver que los subconjuntos linealmente independientes entre  $S$  y  $G$  cumplen esa propiedad con respecto a la inclusión. Aplicando el Lema de Zorn deducimos que entre  $S$  y  $G$  hay un subconjunto linealmente independiente  $T$  que no está contenido propiamente en ninguno que sea linealmente independiente. El argumento que hemos dado muestra que  $G \subseteq \langle T \rangle$  y en conclusión  $T$  es un conjunto generador.

**Ejemplo 4.1.5** Sea  $S = \{(1, 2, -3, 0), (2, 1, 0, 1)\}$ . Vamos a calcular una base de  $K^4$  que contiene a  $S$  siguiendo los pasos de la cuarta demostración del primer apartado del Teorema 4.1.4. Para eso empezamos calculando  $\langle S \rangle$ . Si  $(x, y, z, w) \in \langle S \rangle$ , entonces el siguiente sistema de ecuaciones, en el que las incógnitas son  $\alpha$  y  $\beta$  ha de tener solución

$$\begin{array}{rcl} \alpha & + & 2\beta & = & x \\ 2\alpha & + & \beta & = & y \\ -3\alpha & & & = & z \\ & & & & \beta & = & w \end{array}$$

Despejando  $\alpha$  y  $\beta$  en las dos últimas ecuaciones y sustituyendo en las dos primeras obtenemos que para que el sistema tenga solución se deben satisfacer las dos siguientes ecuaciones homogéneas

$$\begin{array}{rcl} -z & + & 6w & - & 3x & = & 0 \\ -2z & + & 3w & - & 3y & = & 0 \end{array}$$

Por ejemplo el vector  $(1, 0, 0, 0)$  no satisface la primera ecuación y por tanto no pertenece a  $\langle S \rangle$ . De acuerdo con el Corolario 2.3.9, se tiene que  $\langle S \rangle$  está contenido estrictamente en  $\langle S_1 \rangle$ , donde  $S_1 = \{(1, 2, -3, 0), (2, 1, 0, 1), (1, 0, 0, 0)\}$ . Vamos ahora a calcular  $\langle S_1 \rangle$  de forma análoga:

$$\begin{array}{rcl} \alpha & + & 2\beta & + & \gamma & = & x \\ 2\alpha & + & \beta & & & = & y \\ -3\alpha & & & & & = & z \\ & & & & \beta & = & w \end{array}$$

Con lo que obtenemos

$$\begin{array}{rcl} \alpha & = & -\frac{z}{3} \\ \beta & = & w \\ \gamma & = & x - \frac{z}{3} - 2w \end{array}$$

y sustituyendo en la segunda ecuación todavía obtenemos una relación

$$-2z + 3w - 3y = 0$$

que han de cumplir los elementos de  $\langle S_1 \rangle$ . Un elemento que no cumple esta relación es  $(0, 1, 0, 0)$ , con lo que  $\langle S_2 \rangle$ , con  $S_2 = \{(1, 2, -3, 0), (2, 1, 0, 1), (1, 0, 0, 0), (0, 1, 0, 0)\}$ . Ahora dejamos que la lectora compruebe que  $S_2$  sí es un sistema generador de  $K^4$ .

**Ejemplo 4.1.6** Consideremos el siguiente conjunto de polinomios con coeficiente reales

$$G = \{1 + 2X + 4X^2 + 5X^3, X + 3X^2 - 2X^3, 2 + X - X^2 + 16X^3\}.$$

Obviamente  $G$  es un sistema generador de  $W = \langle G \rangle$ . Vamos a calcular una base de  $W$  siguiendo los pasos de la tercera demostración del Teorema 4.1.4. Lo primero que hacemos es pasar el problema a  $\mathbb{R}^4$ , trabajando con las coordenadas en la base  $B = \{1, X, X^2, X^3\}$ . Las coordenadas en la base  $B$  de los polinomios dados son

$$v_1 = (1, 2, 4, 5), \quad v_2 = (0, 1, 3, -2), \quad v_3 = (2, 1, -1, 16).$$

Para ver si estos vectores son linealmente independientes planteamos la ecuación:

$$\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 = 0$$

que nos lleva al siguiente sistema lineal de ecuaciones:

$$\begin{array}{rccccrcr} \alpha_1 & & & & + & 2\alpha_3 & = & 0 \\ 2\alpha_1 & + & \alpha_2 & + & \alpha_3 & = & 0 \\ 4\alpha_1 & + & 3\alpha_2 & - & \alpha_3 & = & 0 \\ 5\alpha_1 & - & 2\alpha_2 & + & 16\alpha_3 & = & 0 \end{array}$$

Resolviendo el sistema nos encontramos que  $(2, -3, -1)$  es una solución, es decir

$$2v_1 - 3v_2 - v_3 = 0$$

de donde deducimos que cualquiera de los vectores es combinación lineal de los otros dos. Por tanto  $\langle v_1, v_2, v_3 \rangle = \langle v_1, v_2 \rangle$ . Como obviamente  $\{v_1, v_2\}$  es linealmente independiente tenemos que  $\{1 + X + 4X^2 + 5X^3, X + 3X^2 - 2X^3\}$  es una base de  $W$ .

En lo que resta de sección vamos a demostrar que todas las bases de un espacio vectorial tienen el mismo número de elementos. Para ello partimos del siguiente

**Lema 4.1.7** *Sean  $S$  un conjunto linealmente independiente y  $G$  un conjunto generador de un espacio vectorial. Para cada elemento  $s$  de  $S$  que no esté en  $G$  existe un elemento  $g$  de  $G$  que no está en  $S$  tal que el conjunto que se obtiene cambiando en  $G$  el elemento  $g$  por  $s$ , es decir  $(G \setminus \{g\}) \cup \{s\}$ , es también generador.*

**Demostración.** Como  $G$  es conjunto generador,  $s \in \langle S \rangle$ , o sea  $s = \alpha_1 v_1 + \dots + \alpha_n v_n$  para  $\alpha_1, \dots, \alpha_n \in K$  y  $v_1, \dots, v_n \in G$ . Podemos suponer  $v_1, \dots, v_n$  son distintos dos a dos y que todos los  $\alpha_i \neq 0$ . Como  $s \notin G$  se tiene que  $s \neq v_i$  para todo  $i$ . Además, como  $S$  es linealmente independiente y  $\{s, v_1, \dots, v_n\}$  no lo es, necesariamente algún  $v_i$  no está en  $S$ . Podemos pues suponer que  $v_1 \notin S$ . Despejando  $v_1$ , que podemos hacer porque  $\alpha_1 \neq 0$ , deducimos que  $v_1 \in \langle s, v_2, \dots, v_n \rangle \subseteq (G \setminus \{v_1\}) \cup \{s\}$ . Luego  $G \subseteq \langle (G \setminus \{v_1\}) \cup \{s\} \rangle$ , con lo que  $(G \setminus \{v_1\}) \cup \{s\}$  es conjunto generador por el Corolario 2.3.9. ■

Aplicando el Lema anterior uno a uno a los elementos de  $S$  que no estén en  $G$ , podemos ir cambiando uno a uno elementos de  $G$  por elementos de  $S$  de forma que vamos obteniendo conjuntos generadores de  $V$ , que tienen número de elementos que  $G$  y cada vez tienen más elementos de  $S$ , hasta que los tengamos todos (en el caso infinito se puede argumentar con el Lema de Zorn). Esto demuestra lo siguiente:

**Teorema 4.1.8** *Sea  $V$  un espacio vectorial y sean  $S, G \subseteq V$  con  $S$  linealmente independiente y  $G$  un sistema generador. Entonces, el número de elementos de  $S$  es menor o igual que el de  $G$ . En particular si  $B$  es una base entonces  $|S| \leq |B| \leq |G|$ .*

En consecuencia tenemos

**Corolario 4.1.9** [Teorema de la Dimensión]. *Todas las bases de un espacio vectorial tienen el mismo número de elementos.*

**Definición 4.1.10** *Se llama dimensión de un espacio vectorial al número de elementos de una (cualquiera) de sus bases. Denotaremos por  $\dim(V)$  a la dimensión del espacio vectorial  $V$ . Un espacio vectorial se dice que tiene dimensión finita si su dimensión es un número natural, es decir si tiene una base finita.*

**Corolario 4.1.11** *Sea  $V$  un espacio vectorial de dimensión  $n$  y sea  $X$  una lista con  $n$  elementos de  $V$ . Entonces las siguientes condiciones son equivalentes.*

- (1)  $X$  es una base.
- (2)  $X$  es linealmente independiente.
- (3)  $X$  es un sistema generador.

**Demostración.** Basta demostrar que tanto (2) como (3) implican (1). Si  $X$  es linealmente independiente, entonces está contenido en una base  $B$  (Teorema 4.1.4). Como  $X$  y  $B$  tienen el mismo número de elementos  $|X| = |B| = n$  y por tanto  $X = B$  y por tanto  $X$  es una base. Si  $X$  es un sistema generador, contiene una base  $B$  y el mismo argumento muestra que  $X$  es una base. ■

Por el Corolario 4.1.11 si ya conocemos la dimensión de  $V$ , digamos  $n$ , y tenemos un subconjunto  $S$  de  $V$ , para demostrar que  $S$  sea una base lo primero que se tiene que cumplir es que  $|S|$  sea igual a  $n$ . Si eso se verifica para demostrar que  $S$  es base no hace falta comprobar que  $S$  es linealmente independiente y sistema generados, pues basta con comprobar una de las dos cosas.

Como una base de un subespacio vectorial de  $V$  es un conjunto linealmente independiente de  $V$  tenemos que

**Corolario 4.1.12** *Si  $W$  es un subespacio vectorial  $V$  de dimensión finita, entonces  $W$  también tiene dimensión finita y  $\dim W \leq \dim V$ . Además  $W = V$  si y sólo si  $\dim W = \dim V$ .*

**Proposición 4.1.13** *Sea  $A \in M_{m,n}(K)$ . Entonces*

$$\dim F(A) = \dim C(A) = r(A) \quad \text{y} \quad \dim F(A) + \dim N(A) = n.$$

**Demostración.** Sea  $r = r(A)$ . Sea  $E$  una matriz fuertemente escalonada obtenida al aplicar a  $A$  el Método de Gauss. Por la Proposición 3.3.4 tenemos que  $F(A) = F(E)$  y  $N(A) = N(E)$ . Además  $E$  tiene exactamente  $r$  filas distintas de cero con lo que  $\dim F(A) = \dim F(E) \leq r$ . Usando la forma escalonada de  $E$  es obvio que las  $r$  filas no nulas de  $E$  son linealmente independientes. Por tanto  $\dim F(A) = r$ .

Aplicando esto a  $A^T$  y usando (3.6) tenemos que  $\dim C(A) = \dim F(A^T) = r(A^T) = r(A)$ . Esto demuestra las igualdades de la izquierda.

Para simplificar la demostración de la igualdad de la derecha suponemos que los pivotes de  $E$  aparecen en las primeras  $r$  columnas. Entonces  $N(A) = N(E)$  que a su vez es el conjunto de soluciones de un sistema de ecuaciones lineales como el siguiente:

$$\left. \begin{array}{rcccccc} X_1 & & & + a_{1,r+1}X_{r+1} & + \dots & + a_{1,n}X_n & = 0 \\ & X_2 & & + a_{2,r+1}X_{r+1} & + \dots & + a_{2,n}X_n & = 0 \\ & & \ddots & & & & \\ & & & X_r & + a_{r,r+1}X_{r+1} & + \dots & + a_{r,n}X_n & = 0 \end{array} \right\}$$

Podemos dar las soluciones de este sistema en coordenadas paramétricas:

$$\left. \begin{array}{rcccccc} X_1 & = & - & a_{1,r+1}t_1 & - & \dots & - & a_{1,n}t_{n-r} \\ X_2 & = & - & a_{2,r+1}t_1 & - & \dots & - & a_{2,n}t_{n-r} \\ & & \vdots & & & & & \\ X_r & = & - & a_{r,r+1}t_1 & - & \dots & - & a_{r,n}t_{n-r} \\ X_{r+1} & = & & t_1 & & & & \\ & & \vdots & & & & & \\ X_n & = & & & & & & t_{n-r} \end{array} \right\}$$

Está claro que el conjunto de soluciones es el subespacio vectorial generado por los siguientes vectores

$$\begin{array}{c} (-a_{1,r+1}, -a_{2,r+1}, \dots, -a_{2,r+1}, 1, 0 \dots, 0), \\ (-a_{1,r+2}, -a_{2,r+2}, \dots, -a_{2,r+2}, 0, 1 \dots, 0), \\ \vdots \\ (-a_{1,n}, -a_{2,n}, \dots, -a_{2,n}, 0, 0 \dots, 1) \end{array}$$

Obviamente esta es una lista de  $n-r$  vectores linealmente independientes y por tanto  $\dim N(A) = \dim N(E) = n - r$ . ■

**Ejemplo 4.1.14** Vamos a calcular una base del subespacio vectorial  $W$  generado por los polinomios  $2 + 4X - 8X^2 - X^3 + 3X^4$ ,  $-1 + 5X - 17X^2 + 3X^3$ ,  $7 - 7X + 35X^2 - 11X^3 + 6X^4$ ,  $-1 + X - 5X^2 + 2X^3 + X^4$ .

Lo primero que hacemos es construir una matriz cuyas filas sean las coordenadas de los polinomios en una base. En este caso elegimos como base  $\{1, X, X^2, X^3, X^4\}$ , con lo que la matriz mencionada es

$$A = \begin{pmatrix} 2 & 4 & -8 & -1 & 3 \\ -1 & 5 & -17 & 3 & 0 \\ 7 & -7 & 35 & -11 & 6 \\ -1 & 1 & -5 & 2 & 1 \end{pmatrix}.$$

Dejamos que la lectora haga las cuentas y obtenga la siguiente forma fuertemente escalonada al aplicarle a  $A$  el Método de Gauss-Jordan:

$$\begin{pmatrix} 1 & 0 & 2 & 0 & 19/3 \\ 0 & 1 & -3 & 0 & 4/3 \\ 0 & 0 & 0 & 1 & 13/3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Por tanto,  $W$  tiene dimensión 3 y una base suya es

$$\left\{ 1 + 2X^2 + \frac{19}{3}X^4, X - 3X^2 + \frac{4}{3}X^4, X^3 + \frac{13}{3}X \right\}.$$

Si queremos obtener una base con coeficientes enteros, podemos multiplicar cada elemento de la base dada por un escalar conveniente de forma que desaparezcan los denominadores. Por ejemplo, la siguiente es otra base de  $W$ :

$$\{3 + 6X^2 + 19X^4, 3X - 9X^2 + 4X^4, 3X^3 + 13X\}.$$

Finalizamos esta sección con dos aplicaciones de la Proposición 4.1.13.

**Corolario 4.1.15** *Las siguientes condiciones son equivalentes para una matriz  $A \in M_{m,n}(K)$ :*

- (1)  $A$  es invertible.
- (2) Las filas de  $A$  forman una base de  $K^n$ .
- (3)  $m = n$  y las filas de  $A$  forman un sistema generador de  $K^n$ .
- (4)  $m = n$  y las filas de  $A$  son linealmente independientes.
- (5)  $m = n =$  dimensión del espacio fila.
- (6)  $m = n = r(A)$ .
- (7) Las columnas de  $A$  forman una base de  $K^m$ .
- (8)  $m = n$  y las columnas de  $A$  forman un sistema generador de  $K^m$ .
- (9)  $m = n$  y las columnas de  $A$  son linealmente independientes.
- (10)  $m = n =$  dimensión del espacio columna.
- (11)  $m = n$  y existe una matriz  $B$  tal que  $BA = I$ .
- (12)  $m = n$  y existe una matriz  $B$  tal que  $AB = I$ .

**Demostración.** (2)  $\Leftrightarrow$  (3)  $\Leftrightarrow$  (4) es consecuencia del Corolario 4.1.11.

(4)  $\Leftrightarrow$  (5) es consecuencia inmediata de la definición de espacio fila.

(5)  $\Leftrightarrow$  (6) es consecuencia de la Proposición 4.1.13.

Esto demuestra que las propiedades entre la (2) y la (6) son equivalentes para la matriz  $A$ . Aplicando esto a la matriz  $A^T$  y teniendo en cuenta que  $r(A^T) = r(A)$  deducimos que las propiedades entre (6) y el (10) también son equivalentes. Por tanto ya sabemos que todas las propiedades entre (2) y (10) son equivalentes. Como además (1) y (6) son equivalentes por el Teorema 3.1.8, deducimos que las (10) primeras propiedades son equivalentes.

Claramente (1) implica (11) y (12).

Supongamos que se cumple (11). Si  $\{e_1, \dots, e_n\}$  son los vectores de la base canónica de  $K^n$ , entonces

$$e_i = b_i A = b_{i1}a_1 + b_{i2}a_2 + \dots + b_{in}a_n.$$

pertenece al espacio fila de  $A$  y por tanto las filas forman un sistema generador de  $K^n$ . Esto demuestra (11)  $\Rightarrow$  (3) con que ya sabemos que las primeras (11) condiciones son equivalentes.

Finalmente si  $A$  verifica (12) entonces  $A^T$  verifica (11) y por tanto  $A^T$  es invertible lo que implica que  $A$  es invertible por el Teorema 3.1.6. Esto demuestra que (12) implica (1) y concluye la demostración. ■

**Corolario 4.1.16** *Todas las matrices fuertemente escalonadas que se obtienen al aplicar el Método de Gauss a una matriz  $A$  son iguales.*

**Demostración.** Sean  $E$  y  $E'$  dos matrices escalonadas obtenidas al aplicar el Método de Gauss a la matriz  $A$ . Por la Proposición 3.3.4 tenemos que  $F(E) = F(A) = F(E')$ . Por tanto las filas no nulas de  $E$  son combinación lineal de las de  $E'$  y viceversa. De hecho queremos ver que son iguales. Ya sabemos que ese número de filas es el rango de  $A$  que denotaremos por  $r$ . Sean pues  $e_1, \dots, e_r$  las filas no nulas de  $E$  y  $e'_1, \dots, e'_r$  las filas no nulas de  $E'$ . Si no fuera cierto que  $e_i = e'_i$  para todo  $i$  habrá un  $i$  máximo para el que  $e_i \neq e'_i$ . O sea  $e'_j = e_j$  para todo  $i < j \leq r$ .

Pongamos  $e'_i = \alpha_1 e_1 + \dots + \alpha_r e_r$ . Sean  $j_1, \dots, j_r$  las columnas de  $E$  que contienen los pivotes de  $E$  y  $j'_1, \dots, j'_r$  las columnas en las que están los pivotes de  $E'$ . Vamos a denotar con  $v_j$  y  $v'_j$  los vectores que se obtienen con las coordenadas  $j_{i+1}, \dots, j_r$  de  $e_j$  y  $e'_j$  respectivamente (en azul en la Figura 4.1. En particular  $v_{i+1}, \dots, v_r$  forman la base canónica y  $v_j = v'_j = 0$  para todo  $j < i$ ). Por tanto  $v'_j = \alpha_1 v_1 + \dots + \alpha_r v_r$ . Por otro lado de  $e'_i = \alpha_1 e_1 + \dots + \alpha_r e_r$  se deduce que  $v'_i = \alpha_1 v_1 + \dots + \alpha_r v_r = \alpha_{i+1} v_{i+1} + \dots + \alpha_r v_r = (\alpha_{i+1}, \dots, \alpha_r)$ . Por tanto  $\alpha_{i+1} = \dots = \alpha_r = 0$ . Luego  $e'_i = \alpha_1 e_1 + \dots + \alpha_i e_i$ .

Supongamos que el pivote de  $e_i$  está en la columna  $j$  y el de  $e'_i$  está en la columna  $j'$ . Podemos suponer sin pérdida de generalidad que  $j \leq j'$  y ahora vamos a fijarnos en las primeras  $j' - 1$  coordenadas de las filas  $e_k$  y  $e'_k$ . Ahora denotamos por  $w_k$  y  $w'_k$  a los vectores formados por las primeras  $j' - 1$  coordenadas de  $e_k$  y  $e'_k$  (en negro en la Figura 4.1). De nuevo  $w'_i = \sum_{l=1}^i \alpha_l w_l$  pero ahora  $0 = w'_i = \alpha_1 w_1 + \dots + \alpha_i w_i$  lo que implica que  $w_1, \dots, w_i$  son linealmente dependientes. Por la forma escalonada de  $E$  y el hecho de que  $j \leq j'$  se tiene que necesariamente  $j = j'$  y  $\alpha_1 = \dots = \alpha_{i-1} = 0$ . Solo nos queda un coeficiente  $\alpha$  diferente de cero:  $\alpha_i$ . Luego  $e'_i = \alpha_i e_i$ , y como tanto  $e_i$  como  $e'_i$  tienen un 1 en la  $j$ -ésima coordenada, necesariamente  $\alpha_i = 1$ , o sea  $e'_i = e_i$  que nos da la contradicción que buscábamos. ■

## 4.2 Suma e intersección de subespacios vectoriales

**Proposición 4.2.1** *Si  $W_1, W_2, \dots, W_n$  son subespacios vectoriales de  $V$ , entonces la intersección*

$$W_1 \cap W_2 \cap \dots \cap W_n = \{x : x \in W_1, x \in W_2, \dots, x \in W_n\}$$

*es otro subespacio vectorial de  $V$ .*

$$\begin{array}{rcl}
e_1 & = & (1, *, 0, \dots, 0, 0, *, 0) & e'_1 & = & (*, *, *, \dots, *, 0, 0, *, 0) \\
e_2 & = & (0, 0, 1, \dots, 0, 0, *, 0) & e'_2 & = & (*, *, *, \dots, *, 0, 0, *, 0) \\
& & & \vdots & & \\
e_i & = & (0, 0, \dots, 1, 0, *, 0) & \neq e'_i & = & (0, 0, \dots, 0, 1, 0, *, 0) \\
e_{i+1} & = & (0, 0, \dots, 0, 1, *, 0) & = e'_{i+1} & = & (0, 0, \dots, 0, 0, 1, *, 0) \\
e_{i+2} & = & (0, 0, \dots, 0, 0, 0, 1) & = e'_{i+2} & = & (0, 0, \dots, 0, 0, 0, 1)
\end{array}$$

Figure 4.1:

**Demostración.** Como 0 está en todos los subespacios vectoriales, en particular está en todos los  $W_i$  y por tanto en su intersección que vamos a denotar  $W$ . Un elemento  $v$  de  $V$  que es combinación lineal de elementos de  $W$  también es combinación lineal de elementos de  $W_i$  para cada  $i$ . Por tanto  $v$  está en todos los  $W_i$  y concluimos que está en  $W$ . Esto prueba que  $W$  es un subespacio vectorial por la Proposición 2.3.5. ■

**Ejemplo 4.2.2** Hemos visto en los Ejemplos 2.3.6 que si  $a_1, a_2, \dots, a_n$  son elementos de  $K$ , entonces el conjunto  $S(E)$  de soluciones de la ecuación

$$E \equiv a_1x_1 + a_2x_2 + \dots + a_nx_n = 0$$

es un subespacio vectorial de  $K^n$ . Una ecuación de este tipo se llama *ecuación lineal homogénea*. Hay una un poco simplona que es el caso en que  $a_1 = a_2 = \dots = a_n = 0$ . En tal caso  $S(E) = V$ . En caso contrario decimos que  $S(E)$  es un *hiperplano vectorial* de  $K^n$ . Por ejemplo, un hiperplano vectorial de  $K^2$  es una recta que pasa por el origen y un hiperplano vectorial de  $K^3$  es un plano que pasa por el origen.

Como un hiperplano vectorial es un subespacio vectorial, entonces la intersección de varios hiperplanos vectoriales es también un subespacio vectorial por la Proposición 4.2.1. De hecho todos los subespacios vectoriales de  $K^n$  son intersecciones de a lo sumo  $n$  hiperplanos vectoriales, es decir todo subespacio vectorial de  $K^n$  es el conjunto de soluciones de un sistema con a lo sumo  $n$  ecuaciones lineales homogéneas. Esto lo veremos con más detenimiento más adelante, pero en este momento nos gustaría hacer la siguiente reflexión. Tenemos dos formas económicas para dar un subespacio vectorial de  $K^n$ : Por extensión, dando un conjunto de generadores, por comprensión dando un conjunto de ecuaciones lineales que lo determinan.

Hemos visto que la intersección de subespacios vectoriales proporciona otros subespacios vectoriales. Sin embargo la unión de dos subespacios vectoriales no es necesariamente un subespacio vectorial. Esto es fácil de ver considerando dos vectores  $v_1$  y  $v_2$  y los subespacios vectoriales  $W_1 = \langle v_1 \rangle$  y  $W_2 = \langle v_2 \rangle$ . Si  $v_1$  y  $v_2$  son linealmente independientes, entonces  $v_2 \notin W_1$  y  $v_1 \notin W_2$ , lo que implica que  $v_1 + v_2 \notin W_1 \cup W_2$  y por tanto  $W_1 \cup W_2$  no es un espacio vectorial.

Recordemos que la intersección de una familia de conjuntos es el mayor conjunto contenido en todos ellos. A la vista de la Proposición 4.2.1, si  $W_1, \dots, W_n$  son subespacios vectoriales de  $V$ , entonces  $W_1 \cap \dots \cap W_n$  es el mayor subespacio contenido en todos los  $W_i$ . Por otro lado la unión de una familia de conjuntos, es el menor conjunto que los contiene a todos. Sin embargo

acabamos de ver que la unión de subespacios vectoriales no es necesariamente un subespacio vectorial. Vamos a ver como obtener el menor subespacio vectorial que contiene a una familia de subespacios.

**Definición 4.2.3** Llamaremos suma de los subespacios vectoriales  $W_1$  y  $W_2$  al siguiente subconjunto de  $V$

$$W_1 + W_2 = \{x_1 + x_2 : x_1 \in W_1, x_2 \in W_2\}.$$

Más generalmente, si  $W_1, W_2, \dots, W_n$  son subespacios vectoriales de  $V$ , entonces llamamos suma de estos subespacios a

$$W_1 + W_2 + \dots + W_n = \{w_1 + w_2 + \dots + w_n : w_1 \in W_1, \dots, w_n \in W_n\}.$$

**Proposición 4.2.4** Si  $W_1, W_2, \dots, W_n$  son subespacios vectoriales de  $V$ , entonces  $W_1 + W_2 + \dots + W_n$  es el menor subespacio vectorial de  $V$  que contiene a todos los  $W_i$ .

**Demostración.** Como  $W_1 + W_2 + \dots + W_n = (W_1 + W_2 + \dots + W_{n-1}) + W_n$ , basta con demostrar la proposición para  $n = 2$ . Como  $0 \in W_1 \subseteq W_1 + W_2$  tenemos que  $W_1 + W_2 \neq \emptyset$ . Sean  $x, y \in W_1 + W_2$  y  $\alpha, \beta \in K$ . Entonces  $x = x_1 + x_2$  e  $y = y_1 + y_2$  para ciertos  $x_1, y_1 \in W_1$  y  $x_2, y_2 \in W_2$  y por tanto  $\alpha x_1 + \beta y_1 \in W_1$  y  $\alpha x_2 + \beta y_2 \in W_2$ . Entonces

$$\alpha x + \beta y = \alpha(x_1 + x_2) + \beta(y_1 + y_2) = (\alpha x_1 + \beta y_1) + (\alpha x_2 + \beta y_2) \in W_1 + W_2.$$

Esto demuestra que  $W_1 + W_2$  es un subespacio vectorial. Como obviamente  $W_1, W_2 \subseteq W_1 + W_2$ , sólo falta demostrar que todo subespacio vectorial que contiene a  $W_1$  y a  $W_2$  también contiene a  $W_1 + W_2$  lo cual es obvio. ■

**Corolario 4.2.5** Si  $W_1 = \langle S_1 \rangle$  y  $W_2 = \langle S_2 \rangle$ , entonces  $W_1 + W_2 = \langle S_1 \cup S_2 \rangle$ .

**Ejemplo 4.2.6** Sea  $W_1 = \langle (1, 2, 0, -2), (3, 1, -2, 4) \rangle$  y sea  $W_2$  el conjunto de las soluciones del siguiente sistema de ecuaciones lineales

$$\begin{array}{rcl} 3X_1 & - & 4X_2 & & = & 0 \\ X_1 & & & + & 2X_3 & = & 0 \end{array}$$

Vamos a calcular los subespacios vectoriales  $W_1 \cap W_2$  y  $W_1 + W_2$ .

Para calcular  $W_1 \cap W_2$  será conveniente determinar qué condiciones han de cumplir los elementos de  $W_1$ . Si  $(x_1, x_2, x_3, x_4)$  es un elemento de  $W_1$ , entonces existen escalares  $\alpha$  y  $\beta$  tales que

$$\begin{array}{rcl} \alpha & + & 3\beta & = & x_1 \\ 2\alpha & + & \beta & = & x_2 \\ & & - & 2\beta & = & x_3 \\ -2\alpha & + & 4\beta & = & x_4 \end{array}$$

Despejando en las ecuaciones primera y tercera obtenemos

$$\alpha = x_1 + \frac{3x_3}{2}, \quad \beta = -\frac{x_3}{2}.$$

Sustituyendo en las otras dos ecuaciones obtenemos que  $W_1$  es el conjunto de soluciones del siguiente conjunto de ecuaciones lineales.

$$\begin{array}{rcccc} 4X_1 & - & 2X_2 & + & 5X_3 & & = & 0 \\ 2X_1 & & & & + & 5X_3 & + & X_4 & = & 0. \end{array}$$

Por tanto  $W_1 \cap W_2$  es el conjunto de soluciones del sistema de ecuaciones.

$$\begin{array}{rcccc} 4X_1 & - & 2X_2 & + & 5X_3 & & = & 0 \\ 2X_1 & & & & + & 5X_3 & + & X_4 & = & 0. \\ 3X_1 & - & 4X_2 & & & & = & 0 \\ X_1 & & & + & 2X_3 & & = & 0 \end{array}$$

O sea  $W_1 \cap W_2$  es el espacio nulo de la matriz

$$\begin{pmatrix} 4 & -2 & 5 & 0 \\ 2 & 0 & 5 & 4 \\ 3 & -4 & 0 & 0 \\ 1 & 0 & 2 & 0 \end{pmatrix}$$

Es fácil ver que esta matriz tiene rango 3 y aplicando o el Algoritmo 3.3.7 o el Algoritmo 3.4.19 obtenemos que  $W_1 \cap W_2$  viene dado por las ecuaciones

$$\begin{array}{l} x_1 = 8x_4 \\ x_2 = 6x_4 \\ x_3 = -4x_4. \end{array}$$

o equivalentemente  $W_1 \cap W_2 = \langle (8, 6, -4, 1) \rangle$  Dicho de otra forma  $W_1 \cap W_2$  tiene dimensión 1 y una base suya es  $\{(8, 6, -4, 1)\}$ .

Vamos ahora a calcular  $W_1 + W_2$ . Si queremos utilizar la Corolario 4.2.5 será conveniente calcular un sistema generador de  $W_2$ . Despejando en las dos ecuaciones que definen  $W_2$  obtenemos que la cuarta coordenada es arbitraria y las otras dos están ligadas por las condiciones:

$$X_2 = \frac{3X_1}{4}, \quad X_3 = -\frac{X_1}{2}$$

Por tanto los elementos de  $W_1 + W_2$  son de la forma

$$\left(x_1, \frac{3x_1}{4}, -\frac{x_1}{2}, x_4\right)$$

o equivalentemente de la forma

$$(4\alpha, 3\alpha, -2\alpha, \beta) = \alpha(4, 3, -2, 0) + \beta(0, 0, 0, 1).$$

Por tanto un sistema generador (y una base) de  $W_2$  es  $\{(4, 3, -2, 0), (0, 0, 0, 1)\}$ . Aplicando la Corolario 4.2.5 tenemos que

$$W_1 + W_2 = \langle (1, 2, 0, -2), (3, 1, -2, 4), (4, 3, -2, 0), (0, 0, 0, 1) \rangle.$$

O sea  $W_1 + W_2$  es el espacio fila de la matriz

$$\begin{pmatrix} 1 & 2 & 0 & -2 \\ 3 & 1 & -2 & 4 \\ 4 & 3 & -2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Aplicando o el Algoritmo 3.3.7 o el Algoritmo 3.4.19 deducimos que

$$\{(1, 2, 0, 0), (0, -5, -2, 0), (0, 0, 0, 1)\}$$

es una base de  $W_1 + W_2$ .

La siguiente fórmula se conoce con el nombre de *Fórmula de Grassman*:

$$\dim(W_1 + W_2) + \dim(W_1 \cap W_2) = \dim W_1 + \dim W_2. \quad (4.2)$$

**Demostración.** Supongamos que  $\dim(W_1) = m_1$ ,  $\dim(W_2) = m_2$  y  $\dim(W_1 \cap W_2) = k$ . Sea  $S = \{u_1, \dots, u_k\}$  una base de  $W_1 \cap W_2$ . Como  $S$  es linealmente independiente de  $W_1$  y de  $W_2$ , del Teorema 4.1.4 sabemos que  $S$  está contenido en una base de  $W_1$  y en una base de  $W_2$ . Sean

$$\begin{aligned} S_1 &= \{u_1, \dots, u_k, v_{k+1}, \dots, v_{m_1}\} \\ S_2 &= \{u_1, \dots, u_k, w_{k+1}, \dots, w_{m_2}\} \end{aligned}$$

bases de  $W_1$  y  $W_2$  respectivamente. Vamos a demostrar que

$$T = \{u_1, \dots, u_k, v_{k+1}, \dots, v_{m_1}, w_{k+1}, \dots, w_{m_2}\}$$

es una base de  $W_1 + W_2$  lo que acabará la demostración pues en tal caso

$$\begin{aligned} \dim(W_1 + W_2) &= k + (m_1 - k) + (m_2 - k) = m_1 + m_2 - k \\ &= \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2). \end{aligned}$$

Que  $T$  es un sistema generador de  $W_1 + W_2$  es consecuencia inmediata de la Corolario 4.2.5. Vamos ahora a ver que  $T$  es linealmente independiente. Supongamos que

$$\sum_{i=1}^k \alpha_i u_i + \sum_{j_1=k+1}^{m_1} \beta_{j_1} v_{j_1} + \sum_{j_2=k+1}^{m_2} \gamma_{j_2} w_{j_2} = 0.$$

Entonces

$$x = \sum_{i=1}^k \alpha_i u_i + \sum_{j_1=k+1}^{m_1} \beta_{j_1} v_{j_1} = - \sum_{j_2=k+1}^{m_2} \gamma_{j_2} w_{j_2} \in W_1 \cap W_2.$$

Como  $S$  es una base de  $W_1 \cap W_2$ , entonces existen  $\lambda_1, \dots, \lambda_k \in K$  tales que

$$x = \sum_{i=1}^k \lambda_i u_i.$$

Por tanto

$$\sum_{i=1}^k \lambda_i u_i + \sum_{j_2=k+1}^{m_2} \gamma_{j_2} w_{j_2} = 0.$$

Aplicando que  $S_2$  es linealmente independiente deducimos que todos los  $\lambda$  y todos los  $\gamma$  son 0. Sustituyendo en la primera ecuación tenemos que

$$\sum_{i=1}^k \alpha_i u_i + \sum_{j_1=k+1}^{m_1} \beta_{j_1} v_{j_1} = 0$$

y, como  $S_1$  es linealmente independiente los  $\alpha$  y los  $\beta$  son 0, lo que acaba la demostración. ■

**Definición 4.2.7** Decimos que dos subespacios vectoriales  $W_1$  y  $W_2$  están en situación de suma directa si  $W_1 \cap W_2 = 0$  y en tal caso la suma  $W_1 + W_2$  la denotamos  $W_1 \oplus W_2$ .

Decimos que  $W_1$  y  $W_2$  son subespacios vectoriales complementarios de  $V$ , o que  $W_2$  es un complemento de  $W_1$  en  $V$  si  $V = W_1 \oplus W_2$ . O sea  $W_1$  y  $W_2$  son subespacios complementarios de  $V$  si y solo si  $W_1 \cap W_2 = 0$  y  $W_1 + W_2 = V$ .

De la Fórmula de Grassman se deduce inmediatamente la siguiente:

$$\dim(W_1 \oplus W_2) = \dim W_1 + \dim W_2. \quad (4.3)$$

**Proposición 4.2.8** Sea  $V$  un espacio vectorial de dimensión finita. Las siguientes condiciones son equivalentes para dos subespacios vectoriales  $W_1$  y  $W_2$  de  $V$ :

- (1)  $W_1$  y  $W_2$  son subespacios complementarios de  $V$ .
- (2)  $W_1 \cap W_2 = 0$  y  $W_1 + W_2 = V$ .
- (3)  $W_1 \cap W_2 = 0$  y  $\dim(V) = \dim W_1 + \dim W_2$ .
- (4)  $V = W_1 + W_2$  y  $\dim(V) = \dim W_1 + \dim W_2$ .
- (5) Para todo  $v \in V$  existen únicos  $w_1 \in W_1$  y  $w_2 \in W_2$  tales que  $v = w_1 + w_2$ .

**Demostración.** (1)  $\Leftrightarrow$  (2) es obvio. (2)  $\Rightarrow$  (3) y (2)  $\Rightarrow$  (4) son consecuencia de (4.3).

(3)  $\Rightarrow$  (2) y (4)  $\Rightarrow$  (2) son consecuencia de la Fórmula de Grassman (4.2).

(2)  $\Rightarrow$  (5). Supongamos que se verifica (2). Entonces para cada elemento  $v \in V$  existen  $w_1 \in W_1$  y  $w_2 \in W_2$  tales que  $v = w_1 + w_2$ . Si  $w'_1 \in W_1$  y  $w'_2 \in W_2$  son otros elementos para los que  $v = w'_1 + w'_2$  entonces  $w_1 - w'_1 = w'_2 - w_2 \in W_1 \cap W_2 = 0$  y por tanto  $w_1 - w'_1 = w'_2 - w_2 = 0$  lo que implica que  $w_1 = w'_1$  y  $w_2 = w'_2$ .

(5)  $\Rightarrow$  (2) Si se cumple (5) entonces  $V = W_1 + W_2$ . Si  $v \in W_1 \cap W_2$  entonces  $v = v + 0 = 0 + v$  y de la unicidad en la propiedad (5) deducimos que  $v = 0$ . Esto demuestra que  $W_1 \cap W_2 = 0$ . ■

Supongamos que  $U$  es un subespacio vectorial de  $V$  y sea  $W$  un complemento de  $U$  en  $V$ . Por la Proposición 4.2.8 para cada  $v \in V$  existe un único  $u \in U$  y un único  $w \in W$  tal que

$v = u + w$ . El elemento  $u$  se llama *proyección* de  $v$  en  $u$  en la *dirección* de  $W$  y lo denotamos  $P_{U,W}(v)$ . Esto define una aplicación

$$\begin{aligned} P_{U,W} : V &\longrightarrow U \\ v &\longmapsto P_{U,W}(v) \end{aligned}$$

que se llama *proyección* de  $V$  en  $U$  en la *dirección* de  $W$ . Claramente

$$v = P_{U,W}(v) + P_{W,U}(v).$$

**Proposición 4.2.9** *Si  $S$  es un conjunto linealmente independiente y  $B$  es una base de  $V$  que contiene a  $S$ , entonces  $B \setminus S$  es una base de un complemento de  $\langle S \rangle$ . En otras palabras*

$$V = \langle S \rangle \oplus \langle B \setminus S \rangle.$$

**Demostración.** Como  $B$  es linealmente también lo son  $S$  y  $B \setminus S$  (Proposición 2.4.3). En particular,  $S$  es una base de  $\langle S \rangle$  y  $B \setminus S$  es una base de  $\langle B \setminus S \rangle$ , con lo que solo tenemos que demostrar que  $\langle B \setminus S \rangle$  es un complemento de  $\langle S \rangle$  en  $V$ . Del Corolario 4.2.5 se deduce que  $\langle S \rangle + \langle B \setminus S \rangle = \langle S \cup (B \setminus S) \rangle = \langle B \rangle = V$ . Como  $\dim \langle S \rangle = |S|$  y  $\dim \langle B \setminus S \rangle = |B \setminus S| = |B| - |S|$  aplicando la Fórmula de Grassman (4.2) tenemos que  $\dim(\langle S \rangle \cap \langle B \setminus S \rangle) = \dim(V) - \dim \langle S \rangle - \dim \langle B \setminus S \rangle = |B| - |S| - (|B| - |S|) = 0$  y por tanto  $\langle S \rangle \cap \langle B \setminus S \rangle = 0$ . Concluimos que  $V = \langle S \rangle \oplus \langle B \setminus S \rangle$ . ■

**Ejemplo 4.2.10** Sea  $V$  el subespacio vectorial de  $K[X]$  generado por los polinomios  $1 + 4X, 1 + 2X^2, 1 + X^3$  y sea  $W$  el conjunto de los polinomios  $P(X) = a_0 + a_1X + a_2X^2 + a_3X^3$  tales que  $P(0) = P(2) = 0$ . Vamos a calcular un complemento  $W$  en  $V$ .

Vamos a utilizar las ideas provenientes del Corolario 4.1.2. Obsérvese que tanto  $V$  como  $W$  están contenidos en el espacio vectorial  $U$  formado por los polinomios de grado  $\leq 3$ . Una base de  $U$  es  $B = \{1, X, X^2, X^3\}$ . Entonces  $V_1 = C_B(V)$  es el subespacio vectorial de  $K^4$  generado por

$$\{v_1 = (1, 4, 0, 0), v_2 = (1, 0, 2, 0), v_3 = (1, 0, 0, 1)\}$$

y  $W_1 = C_B(W)$  es el conjunto de las soluciones del sistema de ecuaciones

$$\begin{aligned} X_1 &= 0 \\ X_1 + 2X_2 + 4X_3 + 8X_4 &= 0 \end{aligned}$$

O equivalentemente

$$\begin{aligned} X_1 &= 0 \\ X_2 &= -2X_3 - 4X_4 \end{aligned}$$

Utilizando esto se puede ver que  $\{w_1 = (0, 2, -1, 0), w_2 = (0, 4, 0, -1)\}$  es una base de  $W_1$ . Ahora podemos ver que  $W_1 \subseteq V_1$ , observando que  $w_1 = \frac{v_1 - v_2}{2}$  y  $w_2 = v_1 - v_3$ . Por tanto  $W \subseteq V$ . Como  $\{w_1, w_2, v_1\}$  es linealmente independiente, entonces es una base de  $V_1$ . De todo esto deducimos que

$$\{C'_B(w_1) = 2X - X^2, C'_B(w_2) = 2X - X^3\}$$

es una base de  $W$  y

$$\{C'_B(w_1) = 2X - X^2, C'_B(w_2) = 2X - X^3, C'_B(v_1) = 1 + 4X\}$$

es una base de  $V$ . Aplicando la Proposición 4.2.9 deducimos que  $\{1 + 4X\}$  es una base de un complemento  $U$  de  $W$  en  $V$ .

Cada elemento de  $V$  tiene una forma única como  $v = \alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3$  con  $\alpha_1, \alpha_2, \alpha_3 \in K$ . Como  $2w_1 = v_1 + v_2$  y  $w_2 = v_1 - v_3$  tenemos que

$$v = \alpha_1 v_1 - \alpha_2 v_1 + \alpha_3 v_1 + \alpha_2(v_1 + v_2) + \alpha_3(v_3 - v_1) = (\alpha_1 - \alpha_2 + \alpha_3)v_1 + 2\alpha_2 w_1 - \alpha_3 w_2.$$

Como  $(\alpha_1 - \alpha_2 + \alpha_3)v_1 \in U$  y  $2\alpha_2 w_1 - \alpha_3 w_2 \in W$  tenemos que

$$P_{W,U}(\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3) = 2\alpha_2 w_1 - \alpha_3 w_2 \quad y \quad P_{U,W}(\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3) = (\alpha_1 - \alpha_2 + \alpha_3)v_1.$$

Vamos a generalizar el concepto de suma directa de dos subespacios vectoriales.

**Proposición 4.2.11** Sean  $W_1, \dots, W_n$  subespacios vectoriales de un espacio vectorial  $V$ . Entonces las siguientes condiciones son equivalentes:

- (1)  $W_i \cap \sum_{j \neq i} W_j = 0$  para todo  $i = 1, \dots, n$ .
- (2) Si  $w_1 + \dots + w_n = 0$  con  $w_1 \in W_1, \dots, w_n \in W_n$  entonces  $w_i = 0$  para todo  $i = 1, \dots, n$ .
- (3) Si  $v_1, w_1 \in W_1, \dots, v_n, w_n \in W_n$  satisfacen  $v_1 + \dots + v_n = w_1 + \dots + w_n$  entonces  $v_i = w_i$  para todo  $i = 1, \dots, n$ .

**Demostración.** (1) implica (2) Supongamos que se verifica (1) y sean  $w_1 \in W_1, \dots, w_n \in W_n$  tales que  $w_1 + \dots + w_n = 0$ . Entonces  $w_i = -\sum_{j \neq i} w_j \in W_i \cap \sum_{j \neq i} W_j = 0$  y por tanto  $w_i = 0$  para todo  $i$ .

(2) implica (3) Supongamos que se verifica (2) y supongamos que  $v_1, w_1 \in W_1, \dots, v_n, w_n \in W_n$  satisfacen  $v_1 + \dots + v_n = w_1 + \dots + w_n$ . Entonces  $\sum_{i=1}^n (v_i - w_i) = 0$  y por hipótesis  $v_i - w_i = 0$ , o sea  $v_i = w_i$ , para todo  $i = 1, \dots, n$ .

(3) implica (1). Supongamos que se verifica (1) y sea  $x \in W_i \cap (\sum_{j \neq i} W_j)$ . Entonces existen  $w_j \in W_j$ , para cada  $j \neq i$  de forma que  $x = \sum_{j \neq i} w_j$ . Si ponemos  $v_i = x$ ,  $v_j = 0$  para todo  $j \neq i$  y  $w_i = 0$  entonces  $\sum_{i=1}^n v_i = \sum_{i=1}^n w_i$ . Por tanto  $v_k = w_k$  para todo  $k$  con lo que en particular  $x = v_i = w_i = 0$ . ■

Dados subespacios vectoriales  $W_1, \dots, W_n$  de un espacio vectorial  $V$  decimos que  $W_1, \dots, W_n$  son *independientes* si se verifican las propiedades de la Proposición 4.2.11. En tal caso denotamos

$$W_1 \oplus \dots \oplus W_n = W_1 + \dots + W_n.$$

Usando inducción (4.3) se deduce que

$$\dim(W_1 \oplus \dots \oplus W_n) = \dim W_1 + \dots + \dim W_n.$$

La siguiente proposición es muy fácil de demostrar:

**Proposición 4.2.12** Si  $W_1, \dots, W_n$  son subespacios independientes de un espacio vectorial y para cada  $i = 1, \dots, n$  tenemos un conjunto linealmente independiente  $X_i$  de  $W_i$  entonces  $X_1 \cup \dots \cup X_n$  es linealmente independiente.

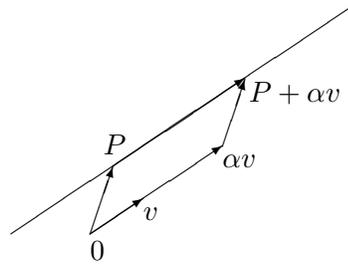
Si cada  $X_i$  es una base de  $W_i$  entonces  $X_1 \cup \dots \cup X_n$  es una base de  $W_1 \oplus \dots \oplus W_n$ .

### 4.3 Variedades Lineales

Dados dos elementos  $P$  y  $v$  de  $K^n$ , con  $v \neq 0$ , llamamos *recta* que pasa por  $P$  y tiene dirección  $v$  al conjunto de los elementos  $X$  de  $K^n$  de la forma

$$P + \alpha v, \text{ con } \alpha \in K.$$

Para convencernos de que esto está relacionado con nuestro concepto de recta vamos a interpretarlo en el caso en que  $K^n = \mathbb{R}^2$ . Los puntos de la forma  $\alpha v$  con  $\alpha \in \mathbb{R}$  son los de la recta que pasa por el  $\bar{0}$  y  $v$  y también son los vectores paralelos a  $v$ . Los puntos de la forma  $P + \alpha v$  son los de la recta que pasa por  $P$  y es paralela a la anterior.



**Ejemplo 4.3.1** La recta del espacio tridimensional que pasa por  $P = (1, 2, 7)$  y es paralela al vector  $(3, -7, 0)$  es la formada por los puntos de la forma

$$X = (1, 2, 7) + \alpha(3, -7, 0), \text{ con } \alpha \in K.$$

Por tanto una recta viene dada por una ecuación del tipo

$$X = P + \alpha v \quad (\alpha \in K) \tag{4.4}$$

que se llama *ecuación vectorial de la recta* que pasa por  $P$  y es paralela al vector  $v$ . Podemos dar una recta con otro tipo de ecuación. Si ponemos  $X = (X_1, X_2, \dots, X_n)$ ,  $P = (P_1, P_2, \dots, P_n)$  y  $v = (v_1, v_2, \dots, v_n)$ , entonces la ecuación (4.4) toma la forma

$$\begin{aligned} X_1 &= P_1 + \alpha v_1 \\ X_2 &= P_2 + \alpha v_2 \\ &\vdots \\ X_n &= P_n + \alpha v_n \end{aligned}$$

La última ecuación se llama *ecuación paramétrica de la recta*. Despejando  $\alpha$  obtenemos

$$\alpha = \frac{X_1 - P_1}{v_1} = \frac{X_2 - P_2}{v_2} = \dots = \frac{X_n - P_n}{v_n}$$

o lo que es lo mismo podemos dar la recta como el conjunto de puntos que satisfacen la siguiente ecuación

$$\frac{X_1 - P_1}{v_1} = \frac{X_2 - P_2}{v_2} = \dots = \frac{X_n - P_n}{v_n} \tag{4.5}$$

conocida como *ecuación continua de la recta*, que en realidad es un conjunto de  $n-1$  ecuaciones. Hay que tener un poco cuidado con la ecuación (4.5), ya que alguno de los denominadores  $v_i$  puede ser 0. En tal caso se interpreta que el numerador también es cero.

**Ejemplo 4.3.2** En el ejemplo anterior hemos calculado la ecuación vectorial de la recta que pasa por  $P = (1, 2, 7)$  y es paralela al vector  $(3, -7, 0)$  como

$$X = (1, 2, 7) + \alpha(3, -7, 0), \text{ con } \alpha \in K.$$

Su ecuación paramétrica es

$$\begin{aligned} X_1 &= 1 + 3\alpha \\ X_2 &= 2 - 7\alpha \\ X_3 &= 7. \end{aligned}$$

La ecuación continua será

$$\frac{X_1 - 1}{3} = \frac{X_2 - 2}{-7} = \frac{X_3 - 7}{0}$$

o equivalentemente

$$-7(X_1 - 1) = 3(X_2 - 2) \quad \text{y} \quad X_3 - 7 = 0$$

Está claro que dos parejas  $(P_1, v_1)$  y  $(P_2, v_2)$  de elementos de  $K^n$  con  $v_1, v_2 \neq 0$  pueden definir la misma recta. La siguiente proposición nos ayuda a determinar cuándo esto es así.

**Proposición 4.3.3** Sean  $L_1$  la recta que pasa por  $P_1$  y tienen dirección  $v_1$  y  $L_2$  la recta que pasa por  $P_2$  y tienen dirección  $v_2$ . Entonces  $L_1 = L_2$  precisamente si los tres vectores  $P_2 - P_1$ ,  $v_1$  y  $v_2$  son paralelos dos a dos.

**Corolario 4.3.4** Dados dos puntos distintos de  $K^n$ , existe una única recta que los contiene.

**Demostración.** Sean  $P_1$  y  $P_2$  dos puntos distintos de  $K^n$ . Es evidente que la recta  $L$  que pasa por  $P_1$  y tiene dirección  $v = P_2 - P_1$  contiene a  $P_1$  y a  $P_2$ . Si  $L'$  es otra recta que contiene a  $P_1 = P_1 + 0(P_2 - P_1)$  y a  $P_2 = P_1 + (P_2 - P_1)$  y tiene por dirección el vector  $v$ , entonces  $L'$  es la recta que pasa por  $P_1$  y tiene por dirección  $v$ . Entonces  $P_2 - P_1$  es paralelo a  $v$  y por tanto los vectores  $P_1 - P_1 = 0$ ,  $P_2 - P_1$  y  $v$  son paralelos dos a dos. ■

Está claro que las ecuaciones vectoriales, paramétricas y continua de la recta que pasa por los puntos  $P = (P_1, P_2, \dots, P_n)$  y  $Q = (Q_1, Q_2, \dots, Q_n)$  son respectivamente

$$\begin{aligned} X &= P + \alpha(Q - P), \\ X_1 &= P_1 + \alpha(Q_1 - P_1) \\ X_2 &= P_2 + \alpha(Q_2 - P_2) \\ &\vdots \\ X_n &= P_n + \alpha(Q_n - P_n) \end{aligned}$$

y

$$\frac{X_1 - P_1}{Q_1 - P_1} = \frac{X_2 - P_2}{Q_2 - P_2} = \dots = \frac{X_n - P_n}{Q_n - P_n}.$$

Dados dos elementos  $v$  y  $w$  de  $K^n$  linealmente independientes y  $P$  otro elemento de  $K^n$ , llamaremos *plano* que pasa por  $P$  con dirección  $\langle v, w \rangle$  al conjunto de puntos  $x$  de  $K^n$  de la forma

$$P + \alpha v + \beta w \text{ con } \alpha, \beta \in K.$$

Entonces la ecuación

$$X = P + \alpha v + \beta w$$

se llama *ecuación vectorial del plano* que pasa por  $P$  y tiene  $v$  y  $w$  por vectores directores y las ecuaciones

$$\begin{aligned} X_1 &= P_1 + \alpha v_1 + \beta w_1 \\ X_2 &= P_2 + \alpha v_2 + \beta w_2 \\ &\vdots \\ X_n &= P_n + \alpha v_n + \beta w_n \end{aligned}$$

se llaman *ecuaciones paramétricas del plano*. Despejando  $\alpha$  y  $\beta$  se obtendrían  $n-2$  ecuaciones sin parámetros que se llaman *ecuaciones implícitas* del plano. En particular, si  $n=3$  entonces un plano viene representado por una única ecuación.

**Ejemplo 4.3.5** Vamos a calcular las ecuaciones del plano de  $\mathbb{R}^3$  que contiene al punto  $(1, 3, -1)$  y tiene por dirección los vectores  $(1, 0, -1)$  y  $(0, 2, -3)$ . La ecuación vectorial de dicho plano es

$$X = (1, 3, -1) + \alpha(1, 0, -1) + \beta(0, 2, -3)$$

y sus ecuaciones paramétricas son

$$\begin{aligned} X_1 &= 1 + \alpha \\ X_2 &= 3 + 2\beta \\ X_3 &= -1 - \alpha - 3\beta \end{aligned}$$

Despejando  $\alpha$  y  $\beta$  en las dos primeras ecuaciones obtenemos

$$\begin{aligned} \alpha &= X_1 - 1 \\ \beta &= \frac{X_2 - 3}{2}. \end{aligned}$$

Sustituyendo en la última ecuación concluimos que el sistema tiene solución precisamente si

$$X_3 = -1 - (X_1 - 1) - 3\frac{X_2 - 3}{2}$$

o equivalentemente

$$2X_1 + 3X_2 + 2X_3 = 9.$$

Podemos generalizar estas construcciones de rectas y planos en  $K^n$  de la siguiente forma. Dados  $k+1$  elementos  $P, v_1, \dots, v_k$  de  $K^n$  se llama *variedad lineal* que pasa por  $P$  y es paralela a  $v_1, v_2, \dots, v_k$  al subconjunto de elementos de  $K^n$  formado por los elementos de la forma

$$P + \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_k v_k$$

con  $\alpha_1, \dots, \alpha_k \in K$ . Interpretamos las rectas como objetos de dimensión 1 y los planos como objetos de dimensión 2. ¿Cuál es la dimensión de una variedad lineal? Obsérvese que una recta en  $K^n$  viene determinada por  $n - 1$  ecuaciones y un plano por  $n - 2$  ecuaciones. ¿Cuántas ecuaciones son necesarias para determinar una variedad lineal en  $K^n$ ? Para responder a estas preguntas necesitamos introducir conceptos nuevos.

**Definición 4.3.6** *Dados  $P \in V$  y un subespacio vectorial  $W$  de  $V$  llamamos variedad lineal que pasa por  $P$  y tiene dirección  $W$  al siguiente subconjunto de  $V$ :*

$$P + W = \{P + w : w \in W\}.$$

*Se llama dimensión de la variedad  $P + W$  a la dimensión de  $W$ .*

**Ejemplos 4.3.7** (1) Una recta de  $K^n$  es una variedad de dimensión 1 de  $K^n$  cuya dirección es el subespacio vectorial  $\langle v \rangle$  generado por un vector no nulo  $v$ . Un plano de  $K^n$  es una variedad lineal cuya dirección es el subespacio vectorial generado por dos vectores no paralelos. Generalizando esto llamamos rectas de  $V$  a las variedades de dimensión 1 de  $V$ .

(2) La variedad lineal de  $K[X]$  que pasa por 1 y tiene dirección  $\langle X, X^2, X^3 \rangle$  está formada por los polinomios de la forma

$$1 + a_1X + a_2X^2 + a_3X^3$$

es decir por los polinomios de grado  $\leq 2$  con término independiente 1.

(3) Sea  $X$  el conjunto de vectores de  $\mathbb{R}^4$  que son soluciones del sistema de ecuaciones siguiente:

$$\begin{array}{rcccccc} x_1 & + & 2x_2 & & - & 3x_4 & + & x_5 & = & 3 \\ 2x_1 & & & - & x_3 & + & 7x_4 & & = & 0 \end{array}$$

Despejando  $x_3$  y  $x_5$  obtenemos

$$\begin{array}{rcccccc} x_3 & = & & 2x_1 & & + & 7x_4 \\ x_5 & = & 3 & - & x_1 & - & 2x_2 & + & 3x_4 \end{array}$$

Si ponemos  $x_1 = \alpha$ ,  $x_2 = \beta$  y  $x_4 = \gamma$ , las soluciones del sistema de ecuaciones son los vectores de la forma

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 3 \end{pmatrix} + \alpha \begin{pmatrix} 1 \\ 0 \\ 2 \\ 0 \\ -1 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ -2 \end{pmatrix} + \gamma \begin{pmatrix} 0 \\ 0 \\ 7 \\ 1 \\ 3 \end{pmatrix}.$$

Es decir

$$X = (0, 3, 0, 0) + \langle (1, 0, 2, 0, -1), (0, 1, 0, 0, -2), (0, 0, 7, 1, 3) \rangle.$$

O sea,  $X$  la variedad de dimensión 3 que pasa por  $(0, 0, 0, 0, 3)$  y tiene dirección  $\langle (1, 0, 2, 0, -1), (0, 1, 0, 0, -2), (0, 0, 7, 1, 3) \rangle$ .

**Proposición 4.3.8** *Sea  $W$  un subespacio vectorial de  $V$  y  $P, Q \in V$ . Sea  $A = P + W$ . Entonces las siguientes condiciones son equivalentes:*

- (1)  $Q \in A$ .
- (2)  $Q - P \in W$ .
- (3)  $A = Q + W$ .
- (4)  $Q + W \subseteq A$ .

**Demostración.** (1) implica (2) Si  $Q \in A$  entonces  $Q = P + w$  para algún  $w \in W$ . Luego  $Q - P = w \in W$ .

(2) implica (3) Supongamos que  $Q - P \in W$ . Si  $R \in Q + W$  entonces existe  $w \in W$  tal que  $R = Q + w$ . Luego  $R = P + (Q - P) + w \in A$ . Esto demuestra que  $Q + W \subseteq A$ . Como además  $P - Q = -(Q - P) \in W$ , la misma demostración muestra que  $A = P + W \subseteq Q + W$ .

(3) implica (4) es obvio.

(4) implica (1) Si  $Q + W \subseteq A$  entonces  $Q \in A$  pues  $Q \in Q + W$ . ■

**Proposición 4.3.9** *Si  $A$  es una variedad lineal de un espacio vectorial  $V$ , entonces su dirección es*

$$D(A) = \{x - y : x, y \in A\}.$$

**Demostración.** Supongamos que  $A$  es la variedad lineal que pasa por  $P$  y tiene como dirección el subespacio vectorial  $W$ . Tenemos que ver que  $W = D(A)$ . Si  $x \in W$ , entonces por definición  $P, P + x \in A$ , con lo que  $x = (P + x) - P \in D(A)$ . Recíprocamente, si  $x \in D(A)$ , entonces existen  $a, b \in A$  tales que  $x = a - b$ . Por tanto existen  $v$  y  $w$  en  $W$  tales que  $a = P + v$  y  $b = P + w$ , con lo que  $x = v - w \in W$ , pues  $W$  es un subespacio vectorial ■

**Corolario 4.3.10** *Sean  $A_1$  y  $A_2$  dos variedades lineales de  $V$  con direcciones  $W_1$  y  $W_2$ . Entonces se verifica una de las dos siguientes condiciones:*

- (1)  $A_1 \cap A_2 = \emptyset$
- (2)  $A_1 \cap A_2$  es una variedad lineal y su dirección es  $W_1 \cap W_2$ .

**Demostración.** Basta con demostrar que si  $A_1 \cap A_2$  contiene un elemento  $P$ , entonces  $A_1 \cap A_2 = P + W_1 \cap W_2$ . Si  $P \in A_1 \cap A_2$ , entonces, por la Proposición 4.3.8 se tiene que  $A_1 = P + W_1$  y  $A_2 = P + W_2$ . Por tanto  $P + W_1 \cap W_2 \subseteq (P + W_1) \cap (P + W_2) = A_1 \cap A_2$ . Recíprocamente, si  $x \in A_1 \cap A_2$ , entonces aplicando la Proposición 4.3.9 a las variedades  $A_1$  y  $A_2$  se deduce que  $x - P \in W_1 \cap W_2$ . Por tanto  $x = P + (x - P) \in P + W_1 \cap W_2$ . ■

Vamos ahora a relacionar las variedades de  $K^n$  con los sistemas de ecuaciones lineales con  $n$  incógnitas. Consideremos un sistema de ecuaciones lineales

$$AX = b. \tag{4.6}$$

Llamamos sistema lineal homogéneo asociado a él al que se obtiene cambiando  $b$  por 0 o sea al sistema

$$AX = 0. \quad (4.7)$$

Está claro que el conjunto de soluciones del sistema lineal homogéneo  $N(A)$ , es el espacio nulo de  $A$ .

**Proposición 4.3.11** *Si  $Ax = b$  es sistema de ecuaciones lineales compatible con  $n$  incógnitas y  $P$  es una solución entonces el conjunto de sus soluciones es la variedad  $P + N(A)$ . Además esta variedad tiene dimensión  $n - r(A)$ .*

*Recíprocamente, toda variedad de  $K^n$  es el conjunto de soluciones de un sistema de ecuaciones lineales compatible.*

**Demostración.** Supongamos que  $P$  es una solución de (4.6). Tenemos que ver que el conjunto de soluciones de (SEL) es la variedad  $P + N(A)$ . En efecto, si  $w \in N(A)$  entonces

$$A(P + w) = AP + Aw = b + 0 = b,$$

es decir  $P + w$  es una solución de (4.6). Recíprocamente, si  $Q$  es una solución de (4.6). Entonces

$$A(Q - P) = AQ - AP = b - b = 0.$$

Por tanto  $w = Q - P \in N(A)$  y  $Q = P + w \in P + W$ .

El que la dimensión de la variedad es  $n - r(A)$  es consecuencia de la Proposición 4.1.13.

Sea ahora  $V$  una variedad de  $K^n$  y sean  $D$  la dirección de  $V$ . Sea  $\{v_1, \dots, v_m\}$  una base de  $D$  y sea  $A$  la matriz cuyas filas son los vectores  $v_1, \dots, v_m$ . Sea ahora  $w_1, \dots, w_r$  una base del espacio nulo de  $D$ . Como  $v_1, \dots, v_m$  son linealmente independientes,  $A$  tiene rango  $m$  y por tanto  $r = n - m$ . Además si  $B$  es la matriz que tiene como columnas  $w_1, \dots, w_r$  entonces  $AB = 0$  pues  $Dw_i = 0$  para todo  $i = 1, \dots, r$ . Por tanto  $B^T A^T = 0$  con lo que  $v_1, \dots, v_m \in N(B^T)$ . Como  $w_1, \dots, w_m$  son linealmente independientes  $B^T$  tiene rango  $r$  y  $n$  columnas con lo que  $N(B^T) = n - r = m$ . Eso implica  $v_1, \dots, v_m$  es una base de  $N(B^T)$  con lo que  $D = N(B^T)$ . Consideremos ahora  $P \in V$  y sea  $b = B^T P^T$ . Eso implica que  $P$  es una solución de  $B^T X^T = b$  y por tanto el conjunto de soluciones de este sistema de ecuaciones lineales es  $P + N(B^T) = P + D = V$ . ■

## 4.4 Problemas

**4.1** Calcular la dimensión de  $\mathbb{C}$  visto como espacio vectorial sobre  $\mathbb{R}$  como en el Ejercicio 2.1 y dar una base de este espacio vectorial.

**4.2** Buscar dos vectores de  $\mathbb{Q}^4$  que, junto con los vectores  $(1, 2, -4, -1)$  y  $(3, -5, 1, -2)$ , formen una base de  $\mathbb{Q}^4$ .

**4.3** Encontrar un subconjunto de

$$\{(1, -1, 2, 3, -2), (1, 1, 0, -1, 2), (0, 1, -1, -2, 2), (4, -1, 5, 2, 1)\}$$

que sea linealmente independiente y de tamaño lo máximo posible y otro que sea una base de  $\mathbb{Q}^5$  que contenga a dicho subconjunto.

**4.4** Encontrar una base de cada uno de los subespacios vectoriales de  $K^4$  formados por las soluciones de cada uno de los siguientes sistemas de ecuaciones lineales homogéneas:

$$(1) X_1 - 2X_2 - 4X_3 - X_4 = 0.$$

$$(2) \begin{cases} X_1 - 2X_2 + 6X_3 & = 0 \\ X_2 - 6X_3 - X_4 & = 0 \end{cases}$$

$$(3) \begin{cases} X_1 - 2X_2 + 6X_3 & = 0 \\ X_2 - 6X_3 - X_4 & = 0 \\ X_1 - X_2 & - X_4 = 0 \end{cases}$$

**4.5** Demostrar que  $B = \{(1, 2, 3), (0, 4, 5), (0, 0, 6)\}$  es una base de  $\mathbb{R}^3$  y calcular las coordenadas de los vectores de la base canónica en esta base.

**4.6** Sea  $\{u, v, w\}$  linealmente independiente de un espacio vectorial sobre  $K$  y sea

$$S = \{x = u + \alpha(v + w) - w, y = \alpha v + w, z = -v + \alpha w\},$$

donde  $\alpha$  es un elemento arbitrario de  $K$ . Demostrar:

- (1) Si  $K = \mathbb{R}$ , entonces  $S$  es linealmente independiente.
- (2) Si  $K = \mathbb{C}$ , determinar para qué valores de  $\alpha$ , es  $S$  linealmente dependiente.

**4.7** Sea  $n$  un número natural y sea  $K_n[X]$  el espacio vectorial formado por los polinomios de grado menor o igual que  $n$  con coeficientes en  $K$ . Sean  $a_0, \dots, a_n$  elementos de  $K$  distintos dos a dos.

- (1) Calcular la dimensión de  $K_n[X]$  y una base suya.
- (2) Demostrar que para cada  $i = 0, 1, \dots, n$  existe un único elemento  $P_i \in K_n[X]$  que verifica

$$P_i(a_j) = \begin{cases} 1, & \text{si } j = i; \\ 0, & \text{si } j \neq i. \end{cases}$$

(Indicación: Usar el Problema 3.34.)

- (3) Demostrar que  $P_0, P_1, \dots, P_n$  es una base de  $K_n[X]$ .
- (4) Demostrar que para toda lista  $b_0, b_1, \dots, b_n$  formada por elementos de  $K$  existe un único  $P \in K_n[X]$  tal que  $P(a_i) = b_i$  para todo  $i = 0, 1, \dots, n$ .

**4.8** Sean  $a_1, \dots, a_n$  números reales distintos. Demostrar que las funciones  $f_1, \dots, f_n$  dadas por  $f_i(x) = e^{a_i x}$  son linealmente independientes. (Indicación: Sustituir  $x$  por 0 en las ecuaciones que se obtienen al derivar sucesivamente en una igualdad  $\alpha_1 e^{a_1 x} + \dots + \alpha_n e^{a_n x} = 0$ . Usar también el Problema 3.34)

**4.9** Demostrar que si  $A \in M_{m,n}(K)$  es una matriz entonces  $C(A) = \{Ax : x \in K^n\}$  y  $F(A) = \{xA : x \in K^m\}$ . Deducir que las siguientes condiciones son equivalentes:

- (1) El sistema de ecuaciones lineales  $AX = b$  es compatible.
- (2)  $b \in C(A)$ .
- (3)  $b$  es combinación lineal de las columnas de  $A$ .

**4.10** Sean  $W_1$  y  $W_2$  dos subespacios vectoriales de un espacio vectorial. Demostrar que  $W_1 \cup W_2$  es un subespacio vectorial si y sólo si  $W_1 \subseteq W_2$  ó  $W_2 \subseteq W_1$ .

**4.11** Consideremos los dos siguientes subespacios vectoriales de  $\mathbb{R}^4$ :

$$W_1 = \langle (3, 1, 3, 1), (-1, 1, -1, 1), (1, 1, 1, 1) \rangle, \quad W_2 = \langle (1, 3, 1, 3), (1, 0, 1, 0), (1, 0, 0, 2) \rangle$$

- (1) Calcular bases y ecuaciones de  $W_1$ ,  $W_2$ ,  $W_1 + W_2$  y  $W_1 \cap W_2$ .
- (2) Calcular un complemento de cada uno de ellos en  $\mathbb{R}^4$ .
- (3) Calcular complementos de  $W_1 \cap W_2$  en  $W_1$  y en  $W_2$ .

**4.12** Sea  $W_1$  el conjunto de los polinomios  $P \in \mathbb{R}[X]$  con coeficientes reales de grado  $\leq 3$  que satisfacen la ecuación  $P(-1) = 0$  y sea  $W_2 = \langle 1 - X, 1 - X^2, 1 - X^3 \rangle$ .

- (1) Calcular bases de  $W_1$ ,  $W_2$ ,  $W_1 \cap W_2$  y  $W_1 + W_2$ .
- (2) Determinar las ecuaciones de estos subespacios en la base canónica  $\{1, X, X^2, X^3\}$  del espacio vectorial  $V$  de los polinomios de grado  $\leq 3$ .
- (3) Calcular un complemento de cada uno de ellos en  $V$ .

**4.13** Sean  $v_1, \dots, v_n$  elementos distintos de 0 de un espacio vectorial. Demostrar que  $v_1, \dots, v_n$  son linealmente independientes si y solo si los subespacios  $\langle v_1 \rangle, \dots, \langle v_n \rangle$  son independientes.

**4.14** Sean

$$W_1 = \left\{ \begin{pmatrix} a & b & b \\ 0 & c & 0 \end{pmatrix} : a, b, c \in K \right\}, \quad W_2 = \left\{ \begin{pmatrix} a-b & 2a & a \\ b & a+2b & 0 \end{pmatrix} : a, b \in K \right\}.$$

Demostrar que  $W_1$  y  $W_2$  son subespacios vectoriales de  $M_{2,3}(K)$  y calcular bases de  $W_1$ ,  $W_2$ ,  $W_1 \cap W_2$  y  $W_1 + W_2$ .

**4.15** Consideramos los dos siguientes subespacios vectoriales de  $\mathbb{R}^{2n}$ :

$$\begin{aligned} W_1 &= \{(x_1, x_2, \dots, x_{2n}) : x_{n+i} = x_i, \text{ para } i = 1, 2, \dots, n\} \\ W_2 &= \{(x_1, x_2, \dots, x_{2n}) : x_{2n+1-i} = x_i, \text{ para } i = 1, 2, \dots, n\}. \end{aligned}$$

Calcular bases y dimensiones de  $W_1$ ,  $W_2$ ,  $W_1 \cap W_2$  y  $W_1 + W_2$ . (Indicación: Empezar con valores pequeños de  $n$  hasta hacerse una idea de cuáles pueden ser las bases y la dimensión y una vez que se tenga una conjetura demostrarla en general.)

**4.16** Calcular las ecuaciones vectoriales, paramétricas, continuas e implícitas de las rectas dadas de la siguiente forma:

- (1) que pasa por  $(2, 4)$  y es paralela al vector  $(5, -2)$ .
- (2) que pasa por  $(4, 1, 0, -1)$  y es paralela al vector  $(1, 0, \frac{3}{4}, -2)$ . (Indicación para simplificar los cálculos: Obsérvese que se pueden eliminar denominadores en los vectores de dirección utilizando que si una recta es paralela a un vector  $v$ , entonces también lo es a cualquier vector paralelo a  $v$ .)
- (3) que pasa por los puntos  $(2, 3, 45)$  y  $(1, 2, 3)$ . (Indicación: Una recta que pasa por dos puntos distintos  $P$  y  $Q$  es paralela al vector  $Q - P$ .)
- (4) que pasa por los puntos  $(0, 1, 3, -1)$  y  $(\frac{2}{3}, -1, \frac{7}{9}, -3)$ .

**4.17** Calcular las ecuaciones vectoriales, paramétricas e implícitas de los planos que se dan de la siguiente forma

- (1) que pasa por  $(1, 0, 0)$  y es paralelo a los vectores  $(1, 0, 1)$  y  $(1, 1, 0)$ .
- (2) que pasa por  $(1, -1, 1, -1)$  y es paralelo a los vectores  $(\frac{1}{3}, 0, 0, \frac{2}{3})$  y  $(\frac{2}{3}, \frac{5}{7}, -2, \frac{1}{6})$ . (La indicación del apartado 2 del Ejercicio anterior es útil aquí también.)
- (3) que pasa por los puntos  $(1, 2, 0)$  y  $(2, 3, 9)$  y es paralela al vector  $(1, 1, 0)$ . (También es útil aquí una de las indicaciones del ejercicio anterior.)
- (4) que pasa por los puntos  $(1, 0, 0)$ ,  $(0, 1, 0)$  y  $(0, 0, 1)$ .

**4.18** Calcular un punto y un vector paralelo a la recta de  $K^3$  dada por las siguientes ecuaciones implícitas:

$$\begin{aligned} 2x_1 - x_2 + x_3 &= 1 \\ 3x_1 + x_2 + x_3 &= 2 \end{aligned}$$

y de la recta de  $K^4$  dada por las siguientes ecuaciones implícitas:

$$\begin{aligned} x_1 + 3x_2 + x_3 &= 1 \\ 7x_1 - x_2 + 4x_3 - x_4 &= 2 \\ 5x_2 + 2x_3 - 4x_4 &= 2 \end{aligned}$$

**4.19** Calcular las ecuaciones paramétricas de los planos dados por las siguientes ecuaciones implícitas:

$$(a) 2x_1 + x_3 = 3, \text{ en } K^3 \quad (b) \begin{cases} x_1 - x_2 + x_3 & = 1 \\ x_2 - x_3 + x_4 & = 1 \\ x_3 - x_4 + x_5 & = 1 \end{cases}, \text{ en } K^5.$$

**4.20** Demostrar que tres puntos  $P = (p_1, p_2, \dots, p_n)$ ,  $Q = (q_1, q_2, \dots, q_n)$  y  $R = (r_1, r_2, \dots, r_n)$  están alineados, es decir pertenecen a la misma recta, si y sólo si se verifican las siguientes ecuaciones

$$\frac{r_1 - p_1}{q_1 - p_1} = \frac{r_2 - p_2}{q_2 - p_2} = \dots = \frac{r_n - p_n}{q_n - p_n},$$

donde si un denominador es 0 interpretamos la ecuación diciendo que el numerador correspondiente también es 0.

Utilizar esto para determinar si los tres puntos que se dan están alineados o no.

- (1)  $(1, 0)$ ,  $(2, 1)$  y  $(3, 2)$ .
- (2)  $(0, 1, 3)$ ,  $(2, 1, 6)$  y  $(-8, 1, -6)$ .
- (3)  $(1, 0, 2, 0, 1)$ ,  $(0, 0, 0, 1, 0)$  y  $(1, 0, -1, 4, 2)$ .

**4.21** Sean  $S$  y  $H$  los subconjunto de  $M_n(K)$  formados por las matrices simétricas y anti-simétricas ( $A \in M_n(K)$  se dice que es antisimétrica si  $A^T = -A$ ). Denotaremos por  $E_{ij}$  la matriz de  $M_n$  que tiene 1 en la entrada  $(i, j)$  y cero en las demás entradas. Demostrar:

- (1)  $S \cap H = 0$ .
- (2) Si  $A \in M_n(K)$ , entonces  $A + A^t \in S$  y  $A - A^t \in H$ .
- (3)  $S \oplus H = M_n(K)$ . (Indicación: Fíjate bien en el apartado anterior.)
- (4)  $\{E_{ij} + E_{ji} : i \leq j\}$  es una base de  $S$ .
- (5)  $\{E_{ij} - E_{ji} : i < j\}$  es una base de  $H$ .
- (6) Calcular las dimensiones de  $S$  y  $H$ .

(Ver la indicación del Problema 4.15.)

**4.22** Calcular las dimensiones y las ecuaciones de las variedades lineales de  $K^5$  que pasan por el punto  $(1, 2, 0, 0, -1)$  y con dirección cada uno de los siguientes subespacios:

- (1)  $W_1 = \langle (1, 0, -1, 0, 1), (-2, 3, -2, 3, -2), (1, 1, 1, 1, 1) \rangle$ .
- (2)  $W_2 =$  conjunto de las soluciones del sistema de ecuaciones lineales

$$\begin{aligned} X_1 + X_2 + X_3 + X_4 + X_5 &= 0 \\ 3X_1 - X_2 - 3X_3 + 7X_4 &= 0 \end{aligned}$$

- (3)  $W_1 \cap W_2$ .

**4.23** Sean  $A \in M_{m,n}(K)$ ,  $P \in K^n$  y  $b = AP$ . Demostrar que si  $N$  es el espacio nulo de la matriz  $A$  entonces la variedad  $P + N$  está formada por los elementos de  $K^n$  que son soluciones del siguiente sistema de ecuaciones lineales  $AX = b$ .

**4.24** Consideremos las dos siguientes matrices

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

Sean  $U$  la variedad que pasa por  $(1, 1, 1, -1)$  y tiene por dirección el espacio fila de  $A$  y  $V$  la variedad que pasa por  $(2, -1, 3, 1)$  y tiene por dirección el espacio nulo de  $B$ . Calcular la dirección de  $U \cap V$  y un sistema de ecuaciones cuyo conjunto de soluciones sea  $U \cap V$ .

**4.25** Consideremos las dos siguientes matrices

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

Sean  $U$  la variedad que pasa por  $(1, 1, -1, 1)$  y tiene por dirección el espacio nulo de  $A$  y  $V$  la variedad que pasa por  $(1, 1, 0, 1)$  y tiene por dirección el espacio nulo de  $B$ . Calcular  $U \cap V$ . ¿Es  $U \cap V$  una variedad? ¿Y si cambiamos  $V$  por la variedad que pasa por  $(1, 1, -1, 1)$  y tiene por dirección el espacio nulo de  $B$ ? ¿Sería  $U \cap V$  una variedad? ¿Cuál es su dirección?

# Capítulo 5

## Aplicaciones Lineales

En todo el capítulo  $K$  es un cuerpo.

### 5.1 Definición y Ejemplos

Recordemos la definición de aplicación lineal que ya apareció en la Sección 2.2 y la proposición que las caracteriza:

**Definición 5.1.1** Sean  $V$  y  $W$  dos espacios vectoriales sobre  $K$ . Una aplicación  $f : V \rightarrow W$  se dice que es lineal si para cada  $\alpha \in K$  y cada  $v_1, v_2 \in V$  se verifican las siguientes igualdades:

$$(1) f(v_1 + v_2) = f(v_1) + f(v_2);$$

$$(2) f(\alpha v_1) = \alpha f(v_1).$$

**Proposición 5.1.2** Las siguientes condiciones son equivalentes para una aplicación  $f : V \rightarrow W$  entre dos espacios vectoriales sobre  $K$ :

(1)  $f$  es lineal

(2) Para todo  $\alpha_1, \alpha_2 \in K$  y todo  $v_1, v_2 \in V$ ,

$$f(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 f(v_1) + \alpha_2 f(v_2).$$

(3) Para todo  $\alpha_1, \alpha_2, \dots, \alpha_n \in K$  y todo  $v_1, v_2, \dots, v_n \in V$ ,

$$f(\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n) = \alpha_1 f(v_1) + \alpha_2 f(v_2) + \dots + \alpha_n f(v_n).$$

**Ejemplos 5.1.3** (1) Sea  $V = K[X]$  el espacio vectorial de los polinomios con coeficientes en  $K$ . Para cada polinomio  $P \in K[X]$  y cada  $a \in K$  sea  $P(a)$  el resultado de sustituir la variable de  $P$  por  $a$ . Por ejemplo, si  $P = X^2 + 1$ , entonces  $P(2) = 2^2 + 1 = 5$  y  $P(-3) = (-3)^2 + 1 = 10$ .

Definimos la aplicación

$$f : K[X] \rightarrow K, f(P) = P(a)$$

Vamos a ver que  $f$  es una aplicación lineal:

$$f(\alpha_1 P_1 + \alpha_2 P_2) = (\alpha_1 P_1 + \alpha_2 P_2)(a) = \alpha_1 P_1(a) + \alpha_2 P_2(a) = \alpha_1 f(P_1) + \alpha_2 f(P_2).$$

- (2) Vamos ahora a suponer que  $K = \mathbb{R}$  y sea  $D : K[X] \rightarrow K[X]$  la aplicación que asocia cada polinomio  $P \in K[X]$  con su derivada  $D(P) = P'$ . De las propiedades de la derivada se ve fácilmente que  $D$  es una aplicación lineal.
- (3) La derivada se puede ver como una aplicación lineal en otro espacio vectorial. Sea  $V = \mathbb{R}^{(a,b)}$  el espacio vectorial  $V$  de las funciones  $f : (a,b) \rightarrow \mathbb{R}$  y sea  $W$  el subconjunto de  $V$  formado por las funciones derivables en el intervalo  $(a,b)$ . Entonces la derivada  $D : W \rightarrow V$  es una aplicación lineal.

**Ejemplo 5.1.4** Sean  $V$  y  $W$  dos espacios vectoriales arbitrarios y sea  $a$  un elemento de  $W$ . Vamos a ver cuándo una aplicación constante  $C_a : V \rightarrow W$  es lineal, donde  $a$  es un elemento de  $W$  y  $C_a(v) = a$  para todo  $v \in V$ . Está claro que si  $a = 0$  entonces  $C_a$  es una aplicación lineal. Recíprocamente, si  $C_a$  es una aplicación lineal, entonces

$$a + a = C_a(0) + C_a(0) = C_a(0 + 0) = C_a(0) = a$$

y por tanto  $a = 0$ . Podíamos haber llegado a esta conclusión utilizando la siguiente Proposición.

**Proposición 5.1.5** Si  $f : V \rightarrow W$  es una aplicación lineal, entonces  $f(0) = 0$ .

**Demostración.**  $f(0) = f(0 + 0) = f(0) + f(0)$  y utilizando la Proposición 2.1.9 deducimos que  $f(0) = 0$ . ■

**Ejemplo 5.1.6** Sea  $V$  un espacio vectorial y  $v \in V$ . Se llama *traslación* determinada por  $v$  a la aplicación  $T_v : V \rightarrow V$  dada por  $T_v(x) = x + v$ , para todo  $x \in V$ . La única traslación que es una aplicación lineal es  $T_0 = 1_V$ . Esto es debido a la Proposición 5.1.5 y a que  $T_v(0) = v$ .

Si  $c \in \mathbb{R}^2$  y  $\alpha \in [0, 2\pi)$ , denotamos por  $R_{c,\alpha}$  a la rotación alrededor de  $c$  de ángulo  $\alpha$ . En este caso  $R_{c,\alpha}$  es una aplicación lineal si y sólo si  $c = 0$  ó  $\alpha = 0$ . Obsérvese que  $R_{c,\alpha} = T_v \circ R_\alpha T_{-v}$ .

Denotaremos por  $\mathcal{L}(V, W)$  el conjunto de las aplicaciones lineales de  $V$  en  $W$ .

**Proposición 5.1.7** Sean  $f, g : V \rightarrow W$  dos aplicaciones lineales y  $\alpha \in K$ . Entonces  $f + g$  y  $\alpha f$  son aplicaciones lineales, es decir  $\mathcal{L}(V, W)$  es un subespacio del espacio vectorial  $W^V$  de las aplicaciones de  $V$  en  $W$ .

**Demostración.** Ejercicio. ■

**Proposición 5.1.8** Si  $f : V \rightarrow W$  y  $g : W \rightarrow U$  son aplicaciones lineales, entonces  $g \circ f$  es una aplicación lineal.

**Demostración.**

$$\begin{aligned}
 (g \circ f)(\alpha_1 v_1 + \alpha_2 v_2) &= g(f(\alpha_1 v_1 + \alpha_2 v_2)) \\
 &= g(\alpha_1 f(v_1) + \alpha_2 f(v_2)) \\
 &= \alpha_1 g(f(v_1)) + \alpha_2 g(f(v_2)) \\
 &= \alpha_1 (g \circ f)(v_1) + \alpha_2 (g \circ f)(v_2).
 \end{aligned}$$

■

**Definición 5.1.9** Un isomorfismo de espacios vectoriales es una aplicación lineal biyectiva. Dos espacios vectoriales se dice que son isomorfos si existe un isomorfismo entre ellos.

**Proposición 5.1.10** Si  $f : V \rightarrow W$  y  $g : W \rightarrow U$  son isomorfismos de espacios vectoriales, entonces  $f^{-1} : W \rightarrow V$  y  $g \circ f$  son también isomorfismos y  $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$ .

**Demostración.** Qué  $f^{-1}$  es biyectiva está claro pero tenemos que comprobar que es una aplicación lineal. Sean  $\alpha_1, \alpha_2 \in K$  y  $w_1, w_2 \in V$ . Entonces

$$\begin{aligned}
 f^{-1}(\alpha_1 w_1 + \alpha_2 w_2) &= f^{-1}(\alpha_1 (f \circ f^{-1})(w_1) + \alpha_2 (f \circ f^{-1})(w_2)) \\
 &= f^{-1}(\alpha_1 f(f^{-1}(w_1)) + \alpha_2 f(f^{-1}(w_2))) \\
 &= f^{-1}(f(\alpha_1 f^{-1}(w_1) + \alpha_2 f^{-1}(w_2))) \\
 &= \alpha_1 f^{-1}(w_1) + \alpha_2 f^{-1}(w_2)
 \end{aligned}$$

Ya sabemos que  $g \circ f$  es lineal y es evidente que  $(f^{-1} \circ g^{-1}) \circ (f \circ g) = 1_V$  y  $(f \circ g) \circ (f^{-1} \circ g^{-1}) = 1_W$ . Por tanto  $f \circ g$  es biyectiva y su inversa es  $f^{-1} \circ g^{-1}$ . ■

La función de los isomorfismos es considerar como esencialmente iguales dos espacios vectoriales isomorfos. En efecto, si  $f : V \rightarrow W$  es un isomorfismo, entonces las propiedades de los elementos de  $V$  pasan a los elementos de  $W$  mediante  $f$ . Más concretamente tenemos propiedades como las siguientes.

**Proposición 5.1.11** Si  $f : V \rightarrow W$  es un isomorfismo y  $X = \{v_1, v_2, \dots, v_n\}$  es una lista de elementos de  $V$  entonces se verifican las siguientes condiciones:

- (1)  $X$  es linealmente independiente si y sólo si  $f(X) = \{f(v_1), f(v_2), \dots, f(v_n)\}$  es linealmente independiente.
- (2)  $X$  es sistema generador de  $V$  si y sólo si  $f(X) = \{f(v_1), f(v_2), \dots, f(v_n)\}$  es sistema generador de  $W$ .
- (3)  $X$  es base de  $V$  si y sólo si  $f(X) = \{f(v_1), f(v_2), \dots, f(v_n)\}$  es base de  $W$ .
- (4)  $\dim V = \dim W$ .
- (5) Si  $V_1$  es un subespacio de  $V$ , entonces  $\dim V_1 = \dim f(V_1)$ .

**Demostración.** Todas las demostraciones de propiedades de este tipo son similares. Por ejemplo, vamos a demostrar la primera.

Supongamos que  $X$  es linealmente independiente. Si

$$\alpha_1 f(v_1) + \alpha_2 f(v_2) + \cdots + \alpha_n f(v_n) = 0$$

entonces

$$f(\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n) = 0 = f(0)$$

y como  $f$  es biyectiva, entonces

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = 0$$

de donde se deduce que  $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$ . Por tanto  $f(X)$  es linealmente independiente.

Una vez que sabemos que las aplicaciones biyectivas transforman sistemas linealmente independiente en sistemas linealmente independientes podemos demostrar el recíproco utilizando esta propiedad para  $f^{-1}$ . En efecto, que si  $f(X)$  es linealmente independiente, entonces  $X = f^{-1}(f(X))$  es linealmente independiente. ■

**Ejemplo 5.1.12** Sea  $A \in M_{m,n}(K)$  y sea  $f_A : K^m \rightarrow K^n$  la aplicación definida en el Ejemplo 2.2.4. Entonces  $f_A$  es un isomorfismo si y sólo si  $A$  es una matriz invertible.

En efecto, es fácil ver que si  $A$  es una matriz invertible, entonces  $f$  es biyectiva ya que en tal caso  $f_A \circ f_{A^{-1}} = 1_{K^n}$  y análogamente  $f_{A^{-1}} \circ f_A = 1_{K^m}$ . Por otro lado, si  $f_A$  es un isomorfismo, entonces la imagen de la base canónica de  $K^m$  es una base de  $K^n$  esto prueba que  $n = m$  y como los elementos de esta base son precisamente las columnas de  $A$ , estas columnas han de formar una base de  $K^n$ , es decir  $A$  es invertible.

**Ejemplo 5.1.13** Sea  $V$  un espacio vectorial sobre  $K$  de dimensión  $n$  arbitrario y sea  $S = \{v_1, v_2, \dots, v_n\}$  una lista de elementos de  $V$ . Entonces la aplicación

$$\begin{aligned} C'_S : K^n &\rightarrow V \\ (\alpha_1, \dots, \alpha_n) &\mapsto \alpha_1 v_1 + \cdots + \alpha_n v_n \end{aligned}$$

es lineal pues

$$\begin{aligned} C'_S(\lambda(\alpha_1, \dots, \alpha_n) + \mu(\beta_1, \dots, \beta_n)) &= C'_S(\lambda\alpha_1 + \mu\beta_1, \dots, \lambda\alpha_n + \mu\beta_n) \\ &= (\lambda\alpha_1 + \mu\beta_1)v_1 + \cdots + (\lambda\alpha_n + \mu\beta_n)v_n \\ &= \lambda(\alpha_1 v_1 + \cdots + \alpha_n v_n) + \mu(\beta_1 v_1 + \cdots + \beta_n v_n) \\ &= \lambda C'_S(\alpha_1, \dots, \alpha_n) + \mu C'_S(\beta_1, \dots, \beta_n). \end{aligned}$$

De hecho esta aplicación ya nos había aparecido que al principio de la Sección 4.1 y habíamos visto que era inyectiva si y solo si  $S$  es linealmente independiente y suprayectiva si y solo si  $S$  es sistema generador (Proposición 4.1.1).

Por tanto  $C'_S$  es biyectiva si y solo si  $S$  es una base y en tal caso su inversa es la aplicación  $C_S$  que asocia cada elemento  $v$  de  $V$  con sus coordenadas  $v_S$  en la base  $S$ . Además  $C_S : V \rightarrow K^n$  es lineal por la Proposición 5.1.11, lo que en realidad ya habíamos visto en Fórmula (4.1).

Como consecuencia de las Proposiciones 5.1.10, 5.1.11 y el Ejemplo 5.1.13 se tiene la siguiente

**Proposición 5.1.14** *Para cada número natural  $n$  se verifica que todos los espacios de dimensión  $n$  son isomorfos a  $K^n$ . Por tanto, dos espacios vectoriales son isomorfos si y sólo si tienen la misma dimensión.*

*Más concretamente si  $B$  es una base de un espacio vectorial de dimensión  $n$  entonces  $C_B : V \rightarrow K^n$  es un isomorfismo de espacios vectoriales. Si  $B_1$  es otra base de otro espacio vectorial  $V_1$  de dimensión  $n$  entonces  $C'_{B_1} \circ C_B : V \rightarrow V_1$  es un isomorfismo de espacios vectoriales.*

## 5.2 Matriz asociada a una aplicación lineal

En esta sección  $V$  y  $W$  son dos espacios vectoriales sobre  $K$ .

La siguiente proposición nos dice que para conocer una aplicación lineal sólo necesitamos saber cómo actúa en los elementos de un sistema generador.

**Proposición 5.2.1** *Sean  $f, g : V \rightarrow W$  dos aplicaciones lineales y  $G$  un sistema generador de  $V$ . Si para todo  $x \in G$  se verifica  $f(x) = g(x)$ , entonces  $f = g$ .*

**Demostración.** Sea  $v$  un elemento arbitrario de  $V$ . Como  $G$  es un sistema generador de  $V$ , entonces existen elementos  $x_1, x_2, \dots, x_n$  de  $G$  y escalares  $\alpha_1, \alpha_2, \dots, \alpha_n \in K$  tales que

$$v = \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n.$$

Por hipótesis  $f(x_i) = g(x_i)$  para todo  $i = 1, 2, \dots, n$  y por tanto

$$\begin{aligned} f(v) &= f(\alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n) = \alpha_1 f(x_1) + \alpha_2 f(x_2) + \dots + \alpha_n f(x_n) \\ &= \alpha_1 g(x_1) + \alpha_2 g(x_2) + \dots + \alpha_n g(x_n) = g(\alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n) \\ &= g(v). \end{aligned}$$

■

Por la proposición anterior si  $B = \{v_1, \dots, v_n\}$  es una base de un espacio vectorial  $V$ , para conocer una aplicación lineal  $f : V \rightarrow W$  solo necesitamos conocer los  $f(v_i)$ . Lo que nos dice la siguiente proposición es que podemos elegir como imagen de cada  $v_i$  el elemento  $w_i$  que queramos. O sea que tomando  $w_1, \dots, w_n \in W$  arbitrarios, va a existir una aplicación lineal que mande cada  $v_i$  al  $w_i$  que hayamos elegido.

**Proposición 5.2.2** *Si  $B = \{v_1, v_2, \dots, v_n\}$  es una base de  $V$  y  $\{w_1, w_2, \dots, w_n\}$  es una lista de elementos de  $W$ , entonces existe una única aplicación lineal  $f : V \rightarrow W$  tal que  $f(v_i) = w_i$  para todo  $i = 1, 2, \dots, n$ .*

**Demostración.** Que como mucho existe una aplicación lineal  $f : V \rightarrow W$  satisfaciendo la propiedad requerida es una consecuencia inmediata de la Proposición 5.2.1. Para ver que existe una observamos que si  $S = \{w_1, \dots, w_n\}$  entonces  $f = C'_S \circ C_B$  es una aplicación lineal, pues

es la composición de dos aplicaciones lineales y  $f(v_i) = C'_S(C_B(v_i)) = C'_S(e_i) = w_i$ , donde  $\{e_1, \dots, e_n\}$  es la base canónica de  $K^n$ . ■

En el Ejemplo 2.2.4 vimos que si  $A \in M_{m,n}(K)$  entonces la aplicación

$$\begin{aligned} f_A : K^m &\rightarrow K^n \\ x &\mapsto f_A(x) = Ax \end{aligned}$$

es una aplicación lineal. Ahora podemos ver que todas las aplicaciones lineales de  $K^m \rightarrow K^n$  son de este tipo.

**Proposición 5.2.3** *Para cada aplicación lineal  $F : K^m \rightarrow K^n$  existe una única matriz  $A$  tal que  $F = f_A$ . Más concretamente la columna  $i$ -ésima de  $A$  es  $f(e_i)^T$ , donde  $\{e_1, e_2, \dots, e_n\}$  es la base canónica de  $V$ .*

**Demostración.** Sea  $F : K^m \rightarrow K^n$  una aplicación lineal y pongamos

$$F(e_i) = \begin{pmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{ni} \end{pmatrix}.$$

Sea  $A = (a_{ij}) \in M_{n,m}$ . Entonces, para todo  $i = 1, 2, \dots, m$  tenemos

$$f_A(e_i) = \begin{pmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{ni} \end{pmatrix} = F(e_i).$$

De la Proposición 5.2.1 se deduce que  $F = f_A$ .

La unicidad de la matriz  $A$  es consecuencia de la cuenta que hemos hecho pues hemos visto que  $f_A(e_i)$  es la  $i$ -ésima columna de  $A$ . Por tanto, si  $F = f_A = f_B$ , con  $A, B \in M_{m,n}(K)$ , entonces las columnas de  $A$  y  $B$  coinciden, es decir  $A = B$ . ■

**Notación 5.2.4** *Recuérdese que si  $B$  es una base de  $V$  y  $v \in B$  entonces  $v_B = C_B(v)$  denota el vector formado por las coordenadas de  $v$  en la base  $B$ . A partir de ahora vamos a considerar  $v_B$  como un vector columna y vamos a generalizar esta notación de la siguiente manera. Para cada lista  $L = \{v_1, \dots, v_m\}$  de elementos de  $V$  ponemos*

$$L_B = (v_{1B}, \dots, v_{mB}) \in M_{n,m}(K).$$

*Es decir  $L_B$  es la matriz formada poniendo en columnas las coordenadas  $v_{iB}$  de los elementos de la lista.*

Con esta notación la matriz  $A$  de la Proposición 5.2.3 la podemos representar como  $f(E_m)_{E_n}$  donde  $E_n$  es la base canónica de  $K^n$ . De forma que tenemos la siguiente fórmula para una aplicación lineal  $f : K^m \rightarrow K^n$ :

$$f(v) = f(E_m)_{E_n} v. \quad (5.1)$$

**Ejemplo 5.2.5** Sea  $R_\alpha$  la rotación en el plano alrededor del origen en sentido antihorario. Si  $B = \{e_1, e_2\}$  es la base canónica, entonces (ver la Figura 5.1)

$$R_\alpha(e_1) = \begin{pmatrix} \cos a \\ \text{sen } a \end{pmatrix}, \quad R_\alpha(e_2) = \begin{pmatrix} -\text{sen } a \\ \cos a \end{pmatrix}.$$

Por tanto

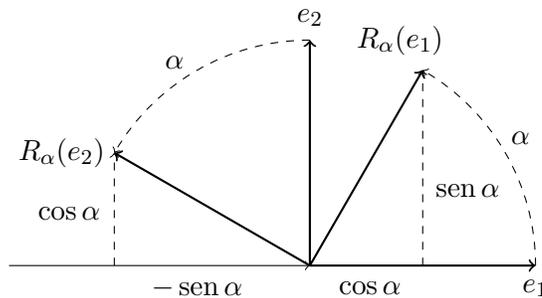


Figure 5.1:

$$f(E_2)_{E_2} = \begin{pmatrix} \cos a & -\text{sen } a \\ \text{sen } a & \cos a \end{pmatrix}.$$

Por ejemplo, si queremos calcular el resultado de girar el triángulo de vértices  $(1, 2), (4, 1), (4, 1)$  podemos poner las coordenadas de los vértices por columnas en una matriz:

$$T = \begin{pmatrix} 1 & 4 & 4 \\ 2 & 1 & 2 \end{pmatrix}$$

y multiplicar por la matriz que produce un giro del ángulo deseado. Por ejemplo, si queremos un giro de  $\pi/3$  radianes, la matriz de giro es

$$A = \begin{pmatrix} \cos \frac{\pi}{3} & -\text{sen } \frac{\pi}{3} \\ \text{sen } \frac{\pi}{3} & \cos \frac{\pi}{3} \end{pmatrix} = \begin{pmatrix} 1/2 & \sqrt{3}/2 \\ -\sqrt{3}/2 & 1/2 \end{pmatrix}$$

Por tanto los vértices del triángulo girado aquellos cuyas coordenadas en la base canónica son las columnas de la siguiente matriz.

$$AT = \begin{pmatrix} 1/2 - \sqrt{3} & 2 - \sqrt{3}/2 & 2 - \sqrt{3} \\ 1 + \sqrt{3}/2 & 1/2 + 2\sqrt{3} & 1 + 2\sqrt{3} \end{pmatrix}.$$

En la Figura 5.2 se representa el triángulo original y su imagen por esta rotación.

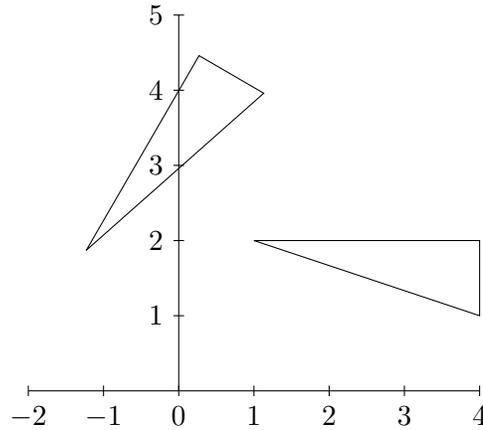


Figure 5.2:

Si combinamos las Proposiciones 5.1.14 y 5.2.3 no podemos por menos que deducir que podemos conocer todas las aplicaciones lineales entre espacios vectoriales de dimensión finita en términos de matrices. En efecto, como todos los espacios vectoriales de dimensión finita son isomorfos a espacios vectoriales de la forma  $K^n$  y conocemos todas las aplicaciones lineales entre espacios vectoriales de esta forma, necesariamente habremos de ser capaces de determinar cómo son las aplicaciones lineales entre espacios vectoriales de dimensión finita.

Más concretamente, sean  $V$  y  $W$  dos espacios vectoriales de dimensión finita y sean

$$B = \{v_1, v_2, \dots, v_m\} \quad \text{y} \quad D = \{w_1, w_2, \dots, w_n\}$$

bases de  $V$  y  $W$  respectivamente. Consideremos los isomorfismos

$$C_B : V \rightarrow K^m, \quad C_D : W \rightarrow K^n$$

que asocian cada vector de  $V$  y  $W$  con sus coordenadas en las bases  $B$  y  $D$  respectivamente. Si  $f : V \rightarrow W$  es una aplicación lineal, entonces de las Proposiciones 5.1.8 y 5.1.10 se deduce que

$$C_D \circ f \circ C_B^{-1} : K^m \rightarrow K^n$$

es una aplicación lineal. Por tanto existe una única matriz en  $M_{m,n}(K)$  que vamos a denotar por  $f(B)_D$  tal que

$$C_D \circ f \circ C_B^{-1} = f_{f(B)_D}.$$

Obsérvese que esto es equivalente a la siguiente fórmula

$$C_D \circ f = f_{f(B)_D} \circ C_B$$

o lo que es lo mismo, para todo  $v \in V$  se verifica

$$f(v)_D = f(B)_D v_B.$$

Además de la Proposición 5.2.3 se deduce que la columna  $i$ -ésima de  $f(B)_D$  está formada por

$$(C_D \circ f \circ C_B^{-1})(e_i)$$

donde  $\{e_1, e_2, \dots, e_m\}$  es la base canónica de  $K^m$ . Obsérvese que  $C_B(v_i) = e_i$ , con lo que

$$(C_D \circ f \circ C_B^{-1})(e_i) = C_D(f(C_B^{-1}(e_i))) = C_D(f(v_i)) = f(v_i)_D.$$

Es decir, las columnas de  $f(B)_D$  son los vectores de coordenadas de los  $f(v_i)$  en la base  $D$ , o sea

$$f(B)_D = (f(v_1)_D, f(v_2)_D, \dots, f(v_m)_D). \quad (5.2)$$

En resumen:

**Teorema 5.2.6** Sean  $V$  y  $W$  dos espacios vectoriales con bases respectivas

$$B = \{v_1, v_2, \dots, v_m\} \quad y \quad D = \{w_1, w_2, \dots, w_n\}.$$

Entonces para cada aplicación lineal  $f : V \rightarrow W$

$$f(v)_D = f(B)_D v_B, \quad \text{para todo } v \in V. \quad (5.3)$$

Además  $f(B)_D$  es la única matriz que satisface (5.3).

La matriz  $f(B)_D$  se llama *matriz asociada a  $f$  en las bases  $B$  y  $D$* .

Un *endomorfismo* de un espacio vectorial  $V$  es una aplicación lineal  $f : V \rightarrow V$ , de  $V$  en si mismo. Si  $f : V \rightarrow V$  es un endomorfismo de  $V$  y  $B$  es una base de  $V$ , entonces ponemos

$$f_B = f(B)_B$$

y a esta matriz la llamamos *matriz asociada a  $f$  en la base  $B$* . En este caso la fórmula (5.3) queda de la siguiente forma:

$$f(v)_B = f_B v_B. \quad (5.4)$$

**Ejemplo 5.2.7** Si  $A \in M_{m,n}(K)$  entonces  $A$  es la matriz asociada a  $f_A$  en las bases canónicas. Esto lo vimos en la Proposición 5.2.3.

**Ejemplo 5.2.8** Sea  $V$  el espacio vectorial de los polinomios de grado  $\leq 4$  con coeficiente reales y consideremos la base  $B = \{1, X, X^2, X^3, X^4\}$  de  $V$ . Sea  $D : V \rightarrow V$  la aplicación derivada. Entonces

$$D(1) = 0, \quad D(X) = 1, \quad D(X^2) = 2X, \quad D(X^3) = 3X^2 \quad y \quad D(X^4) = 4X^3.$$

Luego

$$D(1)_B = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, D(X)_B = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, D(X^2)_B = \begin{pmatrix} 0 \\ 2 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

$$D(X^3)_B = \begin{pmatrix} 0 \\ 0 \\ 3 \\ 0 \\ 0 \end{pmatrix}, D(X^4)_B = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 4 \\ 0 \end{pmatrix}.$$

Poniendo estas columnas en una matriz tenemos

$$D_B = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

**Proposición 5.2.9** Sean  $f : V_1 \rightarrow V_2$  y  $g : V_2 \rightarrow V_3$  aplicaciones lineales y sea  $B_i$  una base de  $V_i$  para cada  $i = 1, 2, 3$ . Entonces la matriz asociada a  $g \circ f$  en las bases  $B_1$  y  $B_3$  es el producto de la matriz asociada a  $g$  en las bases  $B_2$  y  $B_3$ , por la matriz asociada a  $f$  en las bases  $B_1$  y  $B_2$ , o sea:

$$(g \circ f)(B_1)_{B_3} = g(B_2)_{B_3} f(B_1)_{B_2}$$

**Demostración.** Supongamos que  $V_1$  tiene dimensión  $n$  y sean  $x \in K^n$  y  $v = C'_{B_1}(x)$ . Entonces  $x = C_{B_1}(v) = v_{B_1}$ . Aplicando (5.3) con las tres aplicaciones lineales  $g \circ f$ ,  $g$  y  $f$  sucesivamente tenemos

$$\begin{aligned} (g \circ f)(B_1)_{B_3} x &= (g \circ f)(B_1)_{B_3} v_{B_1} = (g \circ f)(v)_{B_3} = g(f(v))_{B_3} \\ &= g(B_2)_{B_3} f(v)_{B_2} = g(B_2)_{B_3} f(B_1)_{B_2} v_{B_1} = (g(B_2)_{B_3} f(B_1)_{B_2}) x \end{aligned}$$

O sea para cada  $x \in K^n$ , el resultado de multiplicar  $(g \circ f)(B_1)_{B_3}$  y  $(g(B_2)_{B_3} f(B_1)_{B_2})$  por  $x$  es el mismo. Si aplicamos esto con  $x$  recorriendo los vectores  $e_1, \dots, e_n$  de la base canónica deducimos que las dos matrices  $(g \circ f)(B_1)_{B_3}$  y  $(g(B_2)_{B_3} f(B_1)_{B_2})$  tienen las mismas columnas, con lo que son iguales. ■

## 5.3 Cambio de base

### Matriz de cambio de base

Cuando introducimos las coordenadas de un vector en una base las comparamos con la forma de llamar a los vectores en un idioma. Aquí el idioma lo fija la base con lo que una base distinta

marca un idioma distinto, es decir un mismo vector  $v$  tiene diferentes coordenadas  $v_B$  y  $v_{B'}$  en dos bases distintas, al igual que dos mismos objetos o conceptos tienen diferentes nombres en dos idiomas distintos. Lo que vamos a ver ahora es como traducir de un idioma a otro, es decir vamos a ver un método eficaz para calcular  $v_{B'}$  a partir de  $v_B$ . Para ello nos va a venir de perlas que  $C_B$  y  $C_{B'}$  sean aplicaciones lineales porque en realidad el traductor es su composición

$$\begin{array}{ccc} K^n & \xrightarrow{C'_B} & V & \xrightarrow{C_{B'}} & K^n \\ v_B & \mapsto & v & \mapsto & v_{B'} \end{array}$$

Pongamos  $B = \{v_1, \dots, v_n\}$  y  $B' = \{v'_1, \dots, v'_n\}$ . Por la Proposición 5.2.3, tenemos que  $C_{B'} \circ C'_B = f_A$  donde  $A = ((C_{B'} \circ C'_B)(e_1)^T, \dots, (C_{B'} \circ C'_B)(e_n)^T) = (C_{B'}(v_1), \dots, C_{B'}(v_n)) = (v_{1B'}, \dots, v_{nB'}) = B_{B'}$ . Esta matriz

$$B_{B'} = (v_{1B'}, \dots, v_{nB'}) \quad (5.5)$$

se llama *matriz de cambio de base* de  $B$  a  $B'$  y el nombre está perfectamente justificado pues es nuestro traductor de las coordenadas en la base  $B'$  a las coordenadas en la base  $B$  ya que por lo que hemos visto tenemos que

$$v_{B'} = B_{B'} v_B. \quad (5.6)$$

Obsérvese que la matriz de cambio de base  $B_{B'}$  se obtiene con las coordenadas de los elementos de  $B$  en la base  $B'$  puestas en columna es decir escribimos

$$v_j = a_{1j}v'_1 + \dots + a_{nj}v'_n \quad (a_{ij} \in K).$$

y ponemos  $B_{B'} = (a_{ij})$ .

Usando lo que vimos en el Ejemplo 2.2.4 sabemos que, como  $C_{B'} \circ C'_B = f_{B_{B'}}$ , es invertible, la matriz  $B_{B'}$  es invertible. Como además  $(C_{B'} \circ C'_B)^{-1} = C_B \circ C'_{B'}$ , tendremos que

$$B_{B'} = B_{B'}^{-1}.$$

O sea, si queremos traducir en la dirección opuesta tendremos que aplicar

$$v_B = B'_{B'} v_{B'} = B_{B'}^{-1} v_{B'}.$$

**Ejemplo 5.3.1** Vamos a calcular una fórmula que nos proporcione las coordenadas de un vector  $(x, y, z)$  en la base  $B = \{(1, 2, 3), (1, -1, 0), (0, 4, -2)\}$ . O sea queremos calcular  $v_B$  para  $v = (x, y, z)$ . Si  $E$  es la base canónica entonces  $v_E = (x, y, z)$  y

$$B_E = \begin{pmatrix} 1 & 1 & 0 \\ 2 & -1 & 4 \\ 3 & 0 & -2 \end{pmatrix}$$

Por tanto  $v_B = E_B v_E = B_E^{-1} v_E$ . En consecuencia necesitamos calcular

$$B_E^{-1} = \frac{1}{18} \begin{pmatrix} 2 & 2 & 4 \\ 16 & -2 & -4 \\ 3 & 3 & -3 \end{pmatrix}.$$

Luego

$$\begin{aligned}(x, y, z)_B &= \frac{1}{18} \begin{pmatrix} 2 & 2 & 4 \\ 16 & -2 & -4 \\ 3 & 3 & -3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \\ &= \frac{1}{18} \begin{pmatrix} 2x + 2y + 4z \\ 16x - 2y - 4z \\ 3x + 3y + 3z \end{pmatrix} = \begin{pmatrix} (x + y + 2z)/9 \\ (8x - y - 2z)/9 \\ (x + y + z)/6 \end{pmatrix}.\end{aligned}$$

### Cambio de base y matrices asociadas a una aplicación lineal

Vamos a ver como están relacionadas las matrices asociadas a una aplicación lineal en dos bases diferentes.

Sea  $f : V \rightarrow W$  una aplicación lineal y sean  $B$  y  $B'$  bases de  $V$  y  $D$  y  $D'$  bases de  $W$ . Utilizando las fórmulas (5.3) y (5.6) tenemos

$$D'_D f(v)_{D'} = f(v)_D = f(B)_D v_B = f(B)_D B'_B v_{B'},$$

para todo  $v \in V$  y por tanto

$$f(v)_{D'} = D'^{-1}_D f(B)_D B'_B v_{B'} = D'^{-1}_D f(B)_D B'_B v_{B'}, \quad (v \in V).$$

Como la fórmula (5.3) caracteriza  $f(B)_D$  deducimos la siguiente fórmula

$$f(B')_{D'} = D'^{-1}_D f(B)_D B'_B = D_{D'} f(B)_D B'_B \quad (5.7)$$

que relaciona las matrices asociadas a la misma aplicación lineal en dos bases diferentes.

En particular, si  $f$  es un endomorfismo de  $V$  y  $B$  y  $B'$  son dos bases de  $V$  entonces

$$f_{B'} = B'^{-1}_B f_B B'_B = B_{B'} f_B B'_B. \quad (5.8)$$

**Ejemplo 5.3.2** Sea  $f : K^3 \rightarrow K^4$  una aplicación lineal de la que conocemos los siguientes datos:

$$f(-1, 2, 3) = (0, 1, 0, 3), \quad f(0, 1, 1) = (-1, 0, -3, 5), \quad f(1, 0, 1) = (-1, -1, 0, 10)$$

y nos gustaría conocer una fórmula general para esta aplicación lineal. Sean  $B = \{e_1, e_2, e_3\}$  y  $D = \{f_1, f_2, f_3, f_4\}$  las bases canónicas de  $R^3$  y  $R^4$  respectivamente y pongamos

$$B' = \{v_1 = (1, 2, 3), v_2 = (0, 1, 1), v_3 = (1, 0, 1)\}.$$

Sea

$$A = B'_B = \begin{pmatrix} -1 & 0 & 1 \\ 2 & 1 & 0 \\ 3 & 1 & 1 \end{pmatrix}.$$

como  $|A| = -2$ ,  $A$  es una matriz invertible y por tanto sus columnas forman una base de  $K^3$ , es decir,  $B'$  es una base de  $K^3$ . De los datos que conocemos podemos calcular  $f(B')_D$  ya que

$$\begin{aligned} f(v_1) &= (0, 1, 0, 3) = f_2 + 3f_4 \\ f(v_2) &= (-1, 0, -3, 5) = -v_1 - 3v_3 + 5v_4 \\ f(v_3) &= (-1, -1, 0, 10) = -v_1 - v_2 + 10v_3 \end{aligned}$$

y, por tanto

$$f(B')_D = \begin{pmatrix} 0 & -1 & -1 \\ 1 & 0 & -1 \\ 0 & -3 & 0 \\ 3 & 5 & 10 \end{pmatrix}.$$

De la ecuación (5.7) tenemos

$$f(B)_D = D_D f(B')_D B_{B'}$$

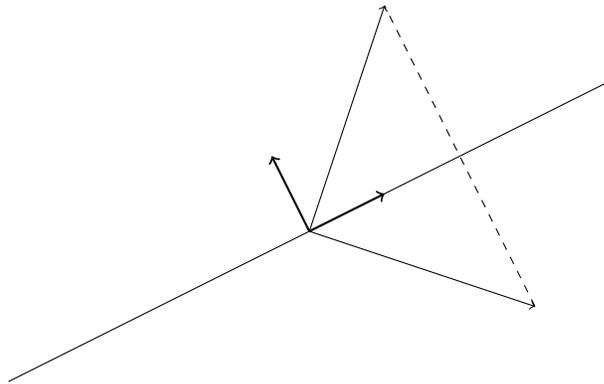
Claramente  $D_D = I$  y  $B'_B = A$ . Por desgracia lo que necesitamos es  $B_{B'} = A^{-1}$ . Calculando esta inversa obtenemos

$$B_{B'} = \frac{1}{2} \begin{pmatrix} -1 & -1 & 1 \\ 2 & 4 & -2 \\ 1 & -1 & 1 \end{pmatrix}$$

y concluimos que

$$f(B)_D = \begin{pmatrix} 0 & -1 & -1 \\ 1 & 0 & -1 \\ 0 & -3 & 0 \\ 3 & 5 & 10 \end{pmatrix} \frac{1}{2} \begin{pmatrix} -1 & -1 & 1 \\ 2 & 4 & -2 \\ 1 & -1 & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} -3 & -3 & 1 \\ -2 & 0 & 0 \\ -6 & -12 & 6 \\ 17 & 7 & 3 \end{pmatrix}$$

**Ejemplo 5.3.3** Sea  $L$  una recta del plano que pasa por el origen y consideremos la aplicación  $S : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  de reflexión en dicha recta. Si elegimos como base  $B = \{v_1, v_2\}$  la formada por



un vector director  $v_1$  de  $L$  y un vector perpendicular  $v_2$ , entonces  $S(v_1) = v_1$  y  $S(v_2) = -v_2$ . Por tanto

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

Supongamos que  $v_1 = (2, 1)$ . Entonces podemos coger como  $v_2 = (-1, 2)$ . Si  $B' = \{e_1, e_2\}$  es la base canónica de  $\mathbb{R}^2$ , entonces

$$U = B_{B'} = \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}$$

Por tanto,

$$S_{B'} = B_{B'} S_B B'_B = \frac{1}{5} \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix} = \frac{1}{5} \begin{pmatrix} 3 & 4 \\ 4 & -3 \end{pmatrix}.$$

Es decir,

$$S \begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{5} \begin{pmatrix} 3 & 4 \\ 4 & -3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{5} \begin{pmatrix} 3x + 4y \\ 4x - 3y \end{pmatrix}.$$

**Ejemplo 5.3.4** Consideremos ahora la aplicación  $G$  que produce un giro de  $\alpha$  grados alrededor de una recta del espacio tridimensional. Por supuesto que esto es ambiguo, ya que hay que decir en que dirección se hace el giro. Supongamos que  $B = \{v_1, v_2, v_3\}$  está formado por tres vectores perpendiculares entre si de forma que  $v_1$  es la dirección del eje de giro y el giro es antihorario mirado desde la parte a la que apunta  $v_1$  y que midiendo ángulos de esta forma el ángulo de  $v_2$  a  $v_3$  sea  $\pi/2$ . Es decir,  $G(v_1) = v_1$  y  $G$  actúa en el plano que contiene a  $v_2$  y  $v_3$  como un giro de  $\alpha$  grados en sentido antihorario. Si  $v_2$  y  $v_3$  tienen la misma longitud, entonces utilizando lo que vimos en el Ejemplo 5.2.5 tendremos

$$G_B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\operatorname{sen} \alpha \\ 0 & \operatorname{sen} \alpha & \cos \alpha \end{pmatrix}.$$

Por ejemplo, supongamos que  $\alpha = \pi/4$  y  $v_1 = (1, 1, 0)$ . Un vector perpendicular a  $v_1$  es  $v_2 = (1, -1, 0)$ . Entonces  $v_1 \times v_2 = (0, 0, -2)$  cumple casi todo lo que queremos: es perpendicular a  $v_1$  y el ángulo entre  $v_2$  y este nuevo vector visto desde  $v_1$  es  $\pi/2$ . Sin embargo no tiene la misma longitud que  $v_2$ . Para conseguir esto último multiplicamos  $v_1 \times v_2$  por el cociente entre la longitud de  $v_1$  (que es  $\sqrt{2}$  y la de  $v_1 \times v_2$  que 2. Así obtenemos el tercer vector  $v_3 = (0, 0, -\sqrt{2})$ . Si  $B'$  es la base canónica de  $\mathbb{R}^3$ , entonces

$$\begin{aligned} G_{B'} &= B_{B'} G_B B'_B \\ &= \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & -\sqrt{2} \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\pi/4) & -\operatorname{sen}(\pi/4) \\ 0 & \operatorname{sen}(\pi/4) & \cos(\pi/4) \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & -\sqrt{2} \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & -\sqrt{2} \end{pmatrix} \frac{1}{2} \begin{pmatrix} 2 & 0 & 0 \\ 0 & \sqrt{2} & -\sqrt{2} \\ 0 & \sqrt{2} & \sqrt{2} \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & -\sqrt{2} \end{pmatrix} \\ &= \frac{1}{4} \begin{pmatrix} 2 + \sqrt{2} & 2 - \sqrt{2} & 2 \\ 2 - \sqrt{2} & 2 + \sqrt{2} & -2 \\ -2 & 2 & 2\sqrt{2} \end{pmatrix}. \end{aligned}$$

Por tanto

$$\begin{aligned} G \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \frac{1}{4} \begin{pmatrix} 2 + \sqrt{2} & 2 - \sqrt{2} & 2 \\ 2 - \sqrt{2} & 2 + \sqrt{2} & -2 \\ -2 & 2 & 2\sqrt{2} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \\ &= \frac{1}{4} \begin{pmatrix} (2 + \sqrt{2})x + (2 - \sqrt{2})y + 2z \\ (2 - \sqrt{2})x + (2 + \sqrt{2})y - 2z \\ -2x + 2y + 2\sqrt{2}z \end{pmatrix}. \end{aligned}$$

## 5.4 Núcleo e Imagen

En esta sección  $f : V \rightarrow W$  es una aplicación lineal entre espacios vectoriales sobre  $K$ .

**Definición 5.4.1** Se llama núcleo de  $f$  al siguiente subconjunto de  $V$ :

$$N(f) = \{v \in V : f(v) = 0\}$$

y se llama imagen de  $f$  al siguiente subconjunto de  $W$ :

$$\text{Im}f = f(V) = \{f(v) : v \in V\}.$$

Llamamos rango de  $f$  a la dimensión de su imagen.

**Proposición 5.4.2** Sea  $f : V \rightarrow W$  una aplicación lineal. Entonces

- (1) El núcleo  $N(f)$  es un subespacio de  $V$ .
- (2) La imagen  $\text{Im}f$  es un subespacio de  $W$ .

**Demostración.** 1. Sean  $v_1, v_2 \in N(f)$  y  $\alpha_1, \alpha_2 \in K$ . Entonces  $f(v_1) = f(v_2) = 0$  y

$$f(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 f(v_1) + \alpha_2 f(v_2) = 0,$$

es decir  $\alpha_1 v_1 + \alpha_2 v_2 \in N(f)$ .

2. Si  $w_1, w_2 \in \text{Im}f$ , entonces existen  $v_1, v_2 \in V$  tales que  $f(v_1) = w_1$  y  $f(v_2) = w_2$ . Si  $\alpha_1, \alpha_2 \in K$ , entonces

$$\alpha_1 w_1 + \alpha_2 w_2 = \alpha_1 f(v_1) + \alpha_2 f(v_2) = f(\alpha_1 v_1 + \alpha_2 v_2) \in \text{Im}f.$$

■

Si  $A \in M_{m,n}(K)$ , entonces está claro que  $N(f_A)$  es el espacio nulo de  $A$ , es decir el conjunto de soluciones del sistema homogéneo de ecuaciones lineales  $AX = 0$ . Vamos ahora a ver como podemos determinar la imagen.

**Proposición 5.4.3** Si  $G = \{v_1, v_2, \dots, v_n\}$  es un sistema generador de  $V$ , entonces  $f(G) = \{f(v_1), f(v_2), \dots, f(v_n)\}$  es un sistema generador de  $\text{Im}f$ .

**Demostración.** Si  $x \in \text{Im}f$ , entonces existe  $v \in V$  tal que  $f(v) = w$ . Como  $G$  es un sistema generador de  $V$ , existen escalares  $\alpha_1, \alpha_2, \dots, \alpha_n \in K$  tales que

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n.$$

Entonces

$$w = f(v) = f(\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n) = \alpha_1 f(v_1) + \alpha_2 f(v_2) + \dots + \alpha_n f(v_n).$$

■

Si  $A \in M_{m,n}(K)$  es una matriz y  $B = \{e_1, e_2, \dots, e_n\}$  es la base canónica de  $K^n$ , entonces  $f(B) = \{f(e_1), f(e_2), \dots, f(e_n)\}$  está formado por las columnas de  $A$ . Por tanto:

**Corolario 5.4.4** *Si  $A$  es una matriz, entonces  $\text{Im}f_A$  es el espacio columna de  $A$ .*

Sean  $B$  es una base de  $V$ ,  $D$  es una base de  $W$  y  $A = f(B)_D$ . Los isomorfismos  $C_B : V \rightarrow K^n$  y  $C_D : W \rightarrow K^n$  asocian  $N(f)$  con  $N(f_A)$  que es el espacio nulo de  $A$  y la imagen  $\text{Im}f$  con la imagen de  $f_A$ , es decir, el espacio columna de  $A$ . Por tanto:

**Corolario 5.4.5** *Si  $f$  es una aplicación lineal y  $A$  es la matriz asociada a  $f$  en un par de bases entonces el rango de  $f$  coincide con el rango de  $A$  y la dimensión de  $N(f)$  con la dimensión de  $N(A)$ .*

**Ejemplo 5.4.6** Sea  $V$  el espacio vectorial de los polinomios de grado menor o igual que 2 y sea  $B = \{v_1, v_2, v_3\}$  la siguiente base de  $V$ :

$$v_1 = 3X + 3X^2, \quad v_2 = -1 + 3X + 2X^2, \quad v_3 = 3 + 7X + 2X^2.$$

Sea  $f$  el endomorfismo de  $V$  cuya matriz asociada en la base  $B$  es

$$f_B = \begin{pmatrix} 1 & 3 & -1 \\ 2 & 0 & 5 \\ 3 & -3 & 11 \end{pmatrix}.$$

Vamos a calcular el núcleo y la imagen de  $f$ .

El núcleo está formado por los polinomios de la forma  $a_0 + a_1X + a_2X^2$  tales que  $a = (a_0, a_1, a_2)$  es un elemento del espacio nulo de  $A$ , es decir una solución del sistema homogéneo  $f_B X = 0$ . Para calcular el espacio nulo de  $A$  usamos la versión del Algoritmo ?? que usa el Método de Gauss, es decir aplicamos el Método de Gauss-Jordan a la matriz  $A$  y obtenemos la siguiente forma fuertemente escalonada de  $A$ :

$$\begin{pmatrix} 1 & 0 & 5/2 \\ 0 & 1 & -7/6 \\ 0 & 0 & 0 \end{pmatrix}$$

Con lo que el espacio nulo está formado por las soluciones del sistema de ecuaciones lineales:

$$\begin{aligned}x_1 &= \frac{5}{2}x_3 \\x_2 &= -\frac{7}{6}x_3\end{aligned}$$

Por tanto el espacio nulo de  $A$  está generado por el vector  $(-15, 7, 6)$  de donde deducimos que el núcleo de  $f$  está generado por

$$-15v_1 + 7v_2 + 6v_3 = 11 + 18X - 19X^2.$$

Para calcular la imagen de  $f$  calculamos el espacio columna de  $A$ . Ya sabemos que el rango de  $A$  es 2 y aplicando el Algoritmo 3.4.14 para calcular el rango de una matriz usando determinantes nos damos cuenta que la submatriz formada por las dos primeras filas y las dos primeras columnas tiene determinante diferente de cero. Eso implica que las dos primeras columnas de  $f_B$  forman una base del espacio nulo y por tanto  $\{f(v_1), f(v_2)\}$  es una base de la imagen de  $f$ . Para obtener las coordenadas de  $f(v_1)$  y  $f(v_2)$  en la base  $B$  usamos la fórmula (5.4)

$$\begin{aligned}f(v_1)_B = f_B(v_1)_B &= \begin{pmatrix} 1 & 3 & -1 \\ 2 & 0 & 5 \\ 3 & -3 & 11 \end{pmatrix} \begin{pmatrix} 0 \\ 3 \\ 3 \end{pmatrix} = \begin{pmatrix} 6 \\ 15 \\ -24 \end{pmatrix}, \\f(v_2)_B = f_B(v_2)_B &= \begin{pmatrix} 1 & 3 & -1 \\ 2 & 0 & 5 \\ 3 & -3 & 11 \end{pmatrix} \begin{pmatrix} -1 \\ 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 6 \\ 8 \\ 10 \end{pmatrix}\end{aligned}$$

Por tanto

$$\text{Im}f = \langle f(v_1) = 6 + 15X - 24X^2, f(v_2) = 6 + 8X + 10X^2 \rangle.$$

Si quisiéramos saber qué debe satisfacer un polinomio  $a_0 + a_1X + a_2X^2$  para estar en la imagen podríamos aplicar el Algoritmo 3.3.12 y obtendríamos que lo que se tendría que cumplir es que el siguiente determinante sea cero:

$$\begin{vmatrix} 1 & 3 & a_0 \\ 2 & 0 & a_1 \\ 3 & -3 & a_2 \end{vmatrix}$$

O sea  $-6a_0 + 12a_1 - 6a_2 = 0$  o equivalentemente

$$a_0 - 2a_1 + a_2 = 0.$$

Combinando el Teorema 4.1.13 y el Corolario 5.4.5 junto con que el número de columnas de la matriz asociada a una aplicación lineal es la dimensión del espacio inicial se tiene lo siguiente:

**Teorema 5.4.7** *Si  $f : V \rightarrow W$  es una aplicación lineal, entonces*

$$\dim N(f) + \dim \text{Im}f = \dim V.$$

Recuérdese que una aplicación  $f : X \rightarrow Y$  es inyectiva si para cada dos elementos distintos  $x_1 \neq x_2$  de  $X$  se verifica  $f(x_1) \neq f(x_2)$  y se dice suprayectiva si todo elemento de  $Y$  es de la forma  $f(x)$  para algún  $x \in X$ . Como consecuencia inmediata del Corolario 5.4.4 tenemos la siguiente caracterización de cuando una aplicación lineal es suprayectiva.

**Corolario 5.4.8** *Sea  $f : V \rightarrow W$  una aplicación lineal. Entonces las siguientes condiciones son equivalentes*

- (1)  $f$  es suprayectiva.
- (2)  $\text{Im} f = W$ .
- (3)  $\dim \text{Im} f = \dim W$ .
- (4) El rango de la matriz asociada a  $f$  en cualquier base coincide con la dimensión de  $W$ .

La siguiente proposición caracteriza cuando una aplicación lineal es inyectiva.

**Proposición 5.4.9** *Sea  $f : V \rightarrow W$  una aplicación lineal. Entonces las siguientes condiciones son equivalentes*

- (1)  $f$  es inyectiva.
- (2)  $N(f) = 0$ .
- (3)  $\dim N(f) = 0$ .
- (4) El rango de la matriz asociada a  $f$  en cualquier base coincide con la dimensión de  $V$ .

**Demostración.** (1) implica (2) Si  $f$  es inyectiva y  $x \in N(f)$ , entonces  $f(x) = 0 = f(0)$  y por tanto  $x = 0$ . Esto prueba que  $N(f) = 0$ .

(2) implica (1) Supongamos que  $N(f) = 0$  y sean  $x \neq y$  elementos distintos de  $f$ . Entonces  $x - y \notin N(f)$ , con lo que  $f(x) - f(y) = f(x - y) \neq 0$ . Por tanto  $f(x) \neq f(y)$ . Es decir  $f$  es inyectiva.

(2) y (3) son evidentemente equivalentes.

La equivalencia entre (3) y (4) es consecuencia del Teorema 5.4.7. ■

## 5.5 Problemas

**5.1** Sea  $f : V \rightarrow V$  una aplicación lineal. Probar que el conjunto  $S = \{v \in V \mid f(v) = v\}$  es un subespacio de  $V$ .

**5.2** Sea  $f : V \rightarrow W$  una aplicación lineal. Demostrar que

- (1) Si  $V_1$  es un subespacio de  $V$  entonces  $f(V_1) = \{f(x) : x \in V_1\}$  es subespacio de  $W$ .

- (2) Si  $W_1$  es un subespacio de  $W$  entonces  $f^{-1}(W_1) = \{x \in V : f(x) \in W_1\}$  es subespacio de  $V$ . (¡Cuidado! Aunque hayamos utilizado la notación  $f^{-1}$  eso no quiere decir que  $f$  sea biyectiva simplemente  $f^{-1}(W_1)$  es la notación que se usa para el conjunto que se ha dado.)
- (3) Si  $w \in W$  entonces  $f^{-1}(w) = \{v \in V : f(v) = w\}$  es un subespacio de  $V$  si y solo si  $w = 0$ .

**5.3** Sea  $V$  un  $K$ -espacio vectorial con base  $B = \{u_1, u_2, u_3\}$  y sea  $D = \{(1, 1, 0), (1, 0, 1), (0, 1, 1)\}$  una base de  $K^3$ . Sea  $f : V \rightarrow K^3$  tal que

$$f(B)_D = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}.$$

Si  $v \in V$  es un vector tal que  $v = 3u_1 + 2u_2 + u_3$ , determinar  $f(v)$ .

**5.4** Encontrar ecuaciones para la aplicación de rotación en el plano real  $\mathbb{R}^2$  de rotación alrededor ángulos  $\pi/2, \pi/6, 3\pi/4$ ; en sentido horario y antihorario.

**5.5** Calcular los vértices de un hexágono regular centrado en el origen en el que uno de los vértices es el punto  $(1, 5)$ .

**5.6** Calcular los vértices de un cuadrado en  $R^3$  centrado en el origen contenido en el plano  $x + y + z = 0$  si uno de los vértices es el punto  $(1, 2, -3)$ .

**5.7** Demostrar que si  $P \in K[X]$  entonces la aplicación  $f_P : K[X] \rightarrow K[X]$  dada por  $S_P(Q(X)) = Q(P(X))$  es lineal. Por ejemplo, si  $P = X^2 - 1$  entonces

$$S_P(X^3 - 2X + 1) = (X^2 - 1)^3 - 2(X^2 - 1) + 1 = X^6 - 3X^4 + 3X^2 - 1 - 2X^2 + 2 + 1 = X^6 - 3X^4 + X^2 + 2.$$

**5.8** Sea  $K_4[X]$  el espacio vectorial de los polinomios de grado menor o igual que 4. Demostrar que la aplicación  $S_{X-1}$  del ejercicio anterior define una aplicación lineal  $K_4[X] \rightarrow K_4[X]$  y calcular la matriz asociada a  $S_{X-1}$  en la base  $\{1, X, X^2, X^3, X^4\}$ . Utilizar la matriz para dar una expresión de  $f(a_0 + a_1X + a_2X^2 + a_3X^3 + a_4X^4)$  como polinomio.

**5.9** Demostrar que la aplicación  $S_{X^2}$  define una aplicación lineal  $K_n[X] \rightarrow K_{2n}[X]$ . Para el caso en que  $n = 3$  dar la matriz asociada a  $S_{X^2}$  en las bases  $\{1, X, X^2, X^3\}$  de  $K_3[X]$  y  $\{1, X, X^2, X^3, X^4, X^5, X^6\}$  de  $K_6[X]$  respectivamente. Utilizar la matriz para dar una expresión de  $f(a_0 + a_1X + a_2X^2 + a_3X^3)$  como polinomio.

**5.10** Sea  $V$  el espacio vectorial generado por las funciones  $f_1, \dots, f_n$  del Problema 4.8. Calcular la matriz asociada a la aplicación lineal derivada en la base formada por  $f_1, \dots, f_n$ .

**5.11** Siendo  $B = \{(1, 1, 1), (1, 1, 0), (1, 0, 0)\}$  una base de  $K^3$ , y  $f(x, y, z) = (x + y, y + z, x)$ , hallar la matriz asociada a  $f$  cuando:

- (1) En el espacio inicial se considera la base canónica y en el final la base  $B$ .
- (2) En el espacio inicial se considera la base  $B$  y en el final la base canónica.
- (3) En ambos se considera la base  $B$ .

**5.12** Encontrar ecuaciones para la aplicación del plano real  $\mathbb{R}^2$  de reflexión en cada una de las siguientes rectas:

- (1)  $x - 2y = 0$ .
- (2)  $\langle (2, -3) \rangle$ .

**5.13** Encontrar ecuaciones para la aplicación del espacio real  $\mathbb{R}^3$  de reflexión en cada una de los siguientes planos:

- (1)  $x + 2y + z = 0$ .
- (2)  $\langle (1, 3, 0), (0, -1, 1) \rangle$ .

**5.14** Encontrar ecuaciones para la aplicación del espacio real  $\mathbb{R}^3$  de rotación alrededor de los ejes y ángulos que se dan.

- (1) Eje =  $\langle (1, 0, -1) \rangle$ , ángulo =  $\pi/4$ .
- (2) Eje =  $\langle (1, 2, 2) \rangle$ , ángulo =  $4\pi/3$ .

**5.15** Sea  $f : V \rightarrow W$ , una aplicación lineal y  $\{v_1, v_2, \dots, v_n\}$  una lista de elementos de  $V$ . ¿Cuáles de las siguientes afirmaciones es cierta y cual es falsa?

- (1) Si  $\{v_1, \dots, v_n\}$  es un sistema generador de  $V$  entonces  $\{f(v_1), \dots, f(v_n)\}$  es sistema generador de  $W$ .
- (2) Si  $\{v_1, \dots, v_n\}$  es un sistema linealmente independiente entonces  $\{f(v_1), \dots, f(v_n)\}$  es sistema linealmente independiente.
- (3) Si  $\{v_1, \dots, v_n\}$  es un sistema linealmente dependiente entonces  $\{f(v_1), \dots, f(v_n)\}$  es sistema linealmente dependiente.
- (4) Si  $f$  es suprayectiva y  $\{v_1, \dots, v_n\}$  es un sistema generador de  $V$  entonces  $\{f(v_1), \dots, f(v_n)\}$  es sistema generador de  $W$ .
- (5) Si  $f$  es inyectiva y  $\{v_1, \dots, v_n\}$  es un conjunto l.i entonces  $\{f(v_1), \dots, f(v_n)\}$  es sistema linealmente independiente.
- (6) Si  $\{f(v_1), \dots, f(v_n)\}$  es un sistema linealmente independiente entonces  $\{v_1, \dots, v_n\}$  es un sistema linealmente independiente.
- (7) Si  $\{f(v_1), \dots, f(v_n)\}$  es un sistema generador de  $W$  entonces  $\{v_1, \dots, v_n\}$  es un sistema generador de  $V$ .

**5.16** Para aquellas aplicaciones del Problema 2.4 que sean lineales, hallar la matriz asociada en las bases canónicas, la dimensión del núcleo y de la imagen y una base del núcleo y una de la imagen.

**5.17** Sea  $f : K^3 \rightarrow K^3$  una aplicación lineal tal que  $f(x, y, z) = (x + y, y + z, z + x)$ . Calcular

- (1) la matriz asociada en las bases canónicas;
- (2) una base del núcleo y una de la imagen y
- (3) una base de  $f(S)$  siendo  $S = \{(x, y, z) \in K^3 \mid 2x - y = 0\}$ .

**5.18** Clasificar las siguientes aplicaciones lineales en el sentido de decidir si son inyectivas y si son suprayectivas.

- (1)  $f : K^3 \rightarrow K^2$  tal que  $f(x, y, z) = (x, y)$ .
- (2)  $f : K^3 \rightarrow K^3$  tal que  $f(x, y, z) = (x + y, x - y, z - x)$ .

Sea  $S = \{(x, y, z) \in K^3 \mid x + y + z = 0\}$  y  $f : K^3 \rightarrow K^2$  tal que  $f(x, y, z) = (x, y)$ .

- (1) Hallar una base de  $S$  (la que se quiera) y la matriz asociada a  $f' : S \rightarrow K^2$  tal que  $f'(x, y, z) = f(x, y, z)$ , de  $(x, y, z) \in S$ ; en la base elegida y la base canónica de  $K^2$ .
- (2) Clasificar  $f'$  y  $f$ .

**5.19** Para las siguientes aplicaciones lineales, calcular la matriz asociada en las bases canónicas, la dimensión del núcleo y de la imagen y una base del núcleo y una de la imagen.

$$\begin{array}{llll} f(1, 0, 0) = (1, 1, 1) & f(1, 1, 0) = (1, 0) & f(1, 2, 1) = (1, 2) \\ \text{(a) } f(0, 1, 0) = (2, 0, 2) & \text{(b) } f(0, 1, 1) = (0, 1) & \text{(c) } f(2, 1, 2) = (2, 1) \\ f(0, 0, 1) = (3, 2, 3) & f(1, 0, 1) = (1, 1) & f(1, 1, 0) = (1, 1) \end{array}$$

**5.20** Sea  $f : K^2 \rightarrow K^3$  tal que  $f(x, y) = (x + y, x - y, 2x)$ .

- (1) Hallar la matriz asociada en las bases canónicas.
- (2) Las ecuaciones implícitas del núcleo y la imagen.
- (3) Sea  $S = \{(x, y) \in K^2 \mid x = y\}$  y  $f' : S \rightarrow K^3$  tal que  $f'(v) = f(v)$  para todo  $v \in S$ . Hallar una base de  $S$  y la matriz asociada a  $f'$  en dicha base y la canónica.
- (4) Clasificar  $f$  y  $f'$ .

**5.21** Sea  $f : K^3 \rightarrow K^3$  aplicación lineal tal que:

$$f(1, 1, 0) = (2, 0, 2), \quad f(2, 1, 1) = (0, 1, 0), \quad f(0, 1, 1) = (a, 1, 1).$$

- (1) Hallar la matriz asociada en las bases canónicas.

(2) Los valores de  $a \in K$ , si existen, para los que  $f$  es un isomorfismo.

**5.22** Sea  $B = \{(1, -1, 1), (2, 1, 3), (0, 1, 0)\}$  una base de  $K^3$  y  $D = \{(1, 1), (1, -1)\}$  base de  $K^2$ . Sea  $f$  la aplicación lineal cuya matriz en las bases anteriores es  $A = \begin{pmatrix} 3 & 1 & 0 \\ 4 & 0 & 1 \end{pmatrix}$ . Hallar la matriz asociada en las bases canónicas y decidir si  $f$  es inyectiva o suprayectiva.

**5.23** Sean  $f : K^2 \rightarrow K^3$  y  $g : K^3 \rightarrow K^2$  aplicaciones lineales tales que

$$\begin{aligned} f(1, 0) &= (2, 0, 1), & f(0, 3) &= (3, 3, 3) \\ g(1, 1, 0) &= (1, 0), & g(0, 1, 1) &= (0, 1), & g(1, 0, 1) &= (1, 1) \end{aligned}$$

Hallar las matrices asociadas de  $f$ ,  $g$  y  $(g \circ f)$  en las bases canónicas y decidir si son inyectivas o suprayectivas.

**5.24** Sea  $V$  un espacio vectorial real con base  $B = \{v_1, v_2, v_3\}$ . Sea  $f : V \rightarrow V$  una aplicación lineal tal que

$$f(v_1) = v_1 + v_2 - v_3, \quad f(v_2) = v_2, \quad f(v_3) = v_1 + v_3.$$

- (1) Hallar la matriz asociada a  $f$  tomando a  $B$  como base inicial y final.
- (2) Dar una base del núcleo y una de la imagen de  $f$ .
- (3) Clasificar  $f$ .
- (4) Si  $g : V \rightarrow V$  es una aplicación lineal tal que

$$g(v_1) = v_3, \quad g(v_2) = v_1 + v_3, \quad g(v_3) = v_1,$$

hallar la matriz asociada a  $(g \circ f)$  en la base  $B$  como base inicial y final.

**5.25** Dar un ejemplo de una aplicación lineal  $f : K^3 \rightarrow K^3$  tal que  $\text{Im} f = \langle (1, 0, 1), (1, 2, 2) \rangle$ .

**5.26** Sea  $f : K^3 \rightarrow K^3$  tal que  $f(x, y, z) = (mx + my, 2my, x + 3z)$ . Decidir para que valores de  $m$  la aplicación  $f$  inyectiva y para cuáles es suprayectiva

**5.27** Demostrar el siguiente enunciado:

Sean  $E_1, E_2$  espacios vectoriales de dimensión finita sobre un cuerpo  $K$ . Entonces:

- (1)  $\dim E_1 \leq \dim E_2 \Leftrightarrow$  existe un aplicación lineal inyectiva  $f : E_1 \rightarrow E_2$ .
- (2)  $\dim E_1 \geq \dim E_2 \Leftrightarrow$  existe un aplicación lineal suprayectiva  $f : E_1 \rightarrow E_2$ .
- (3)  $\dim E_1 = \dim E_2 \Leftrightarrow$  existe un isomorfismo  $f : E_1 \rightarrow E_2$ .

**5.28** Sean  $V$  y  $W$  espacios vectoriales,  $V_1$  un subespacio de  $V$  y  $W_1$  un subespacio de  $W$ . Demostrar que si  $\dim V = \dim V_1 + \dim W_1$  entonces existe una aplicación lineal  $f : V \rightarrow W$  tal que  $N(f) = V_1$  e  $\text{Im} f = W_1$ .

**5.29** Sean  $V$  y  $W$  espacios vectoriales con bases  $B = \{v_1, v_2, v_3, v_4\}$  y  $B' = \{w_1, w_2, w_3, w_4, w_5\}$  respectivamente. Sea  $f : V \rightarrow W$  una aplicación lineal tal que  $v_1 + v_2, v_1 + v_3 \in N(f)$ ,  $f(v_3) = f(v_2 + v_3 + v_4) = w_3$ .

- (1) Calcular la matriz asociada a  $f$  en las bases  $B$  y  $B'$ .
- (2) Calcular bases de  $N(f)$  e  $\text{Im}f$ .

**5.30** Sea  $f : K^4 \rightarrow K^3$  la aplicación lineal cuya matriz respecto de las bases  $B$  y  $B'$  es:

$$f(B)_{B'} = \begin{pmatrix} 1-a & 2a+1 & 2a+2 & -a \\ a & a & 0 & 2a \\ 2 & a+1 & a-1 & 4 \end{pmatrix}$$

donde

$$B = \{(1, 0, 0, 1), (1, 1, 0, 0), (1, 0, 1, 0), (1, 1, 1, 1)\}$$

y

$$B' = \{(1, 1, 0), (1, 0, 1), (1, 1, 1)\}.$$

- (1) Calcular la matriz respecto de las bases canónicas de  $K^4$  y  $K^3$ .
- (2) Calcular las dimensiones de  $N(f)$  e  $\text{Im}f$  en función de los valores de  $a$ .
- (3) Para  $a = 2$ , ampliar las bases de  $N(f)$  e  $\text{Im}f$  a bases de  $K^4$  y  $K^3$  respectivamente, y calcular la matriz de  $f$  respecto de esas nuevas bases.

**5.31** Sea  $f$  una aplicación lineal de  $K^3$  en  $K^3$  verificando:

- (1) El núcleo de  $f$  está formado por los vectores

$$N(f) = \{xu + yv + zw \mid 3x - y + 4z = 0\}$$

donde  $u = (1, 0, 0)$ ,  $v = (1, 1, 0)$ ,  $w = (1, 1, 1)$ .

- (2) El vector  $w = (1, 1, 1)$  verifica que  $f(w) = u + v + w$ .

Calcular la matriz de  $f$  respecto de las bases canónicas y ecuaciones implícitas de  $N(f)$ .

**5.32** Sea  $f : K^3 \rightarrow K^3$  la aplicación lineal que verifica que:

$$f(1, 0, -1) = (0, 1, s^2), \quad f(0, 1, 0) = (-s, -s, -2s), \quad f(1, 0, 0) = (1, 2, 3).$$

- (1) Calcular la dimensión de  $N(f)$  en función de los valores de  $s$ .
- (2) Estudiar para qué valores de  $s$  se verifica que el vector  $(1, -1, 0)$  pertenece a la imagen de  $f$ .

**5.33** Sea  $f : K^3 \rightarrow K^3$  la aplicación lineal cuya matriz respecto de las bases canónicas es:

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

- (1) Calcular las dimensiones del núcleo y de la imagen en función de los valores de  $a$  y  $b$ .
- (2) Decir para qué valores de  $a, b$ , el vector  $(-1, -5, 3b)$  está en la imagen de  $f$ .

**5.34** Sea  $f : K^3 \rightarrow K^3$  la aplicación lineal que verifica

$$f \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \alpha + \beta \\ 1 + \alpha\beta \\ 1 + \beta \end{pmatrix} \quad f \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 + \beta \\ 1 + \alpha\beta \\ \alpha + \beta \end{pmatrix} \quad f \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ \alpha \end{pmatrix},$$

donde  $\alpha, \beta \in K$  son parámetros.

- (1) Calcular la matriz de  $f$  en la base canónica.
- (2) Encontrar valores para  $\alpha$  y  $\beta$  para que se tenga  $\dim N(f) = 1$ , y dar entonces una base de  $N(f)$  y otra de  $\text{Im}f$ .
- (3) Encontrar valores para  $\alpha$  y  $\beta$  para los que se tenga  $\dim N(f) = 2$ , y dar entonces una base de  $N(f)$  y otra de  $\text{Im}f$ .

**5.35** Sea  $f : K^3 \rightarrow K^3$  una aplicación lineal y  $B$  una base de  $K^3$  tal que

$$f_B = \begin{pmatrix} a+1 & 2 & a+1 \\ 4 & a-b+2 & a-b+2 \\ 2 & 1 & 1 \end{pmatrix}.$$

- (1) Calcular dimensión del núcleo y de la imagen de  $f$  y clasificar  $f$  de acuerdo con los valores de  $a$ .
- (2) Determinar si existen bases  $C$  y  $D$  de  $K^3$  de tal manera que  $f(C)_D$  sea triangular. (Si existen las bases hay que escribirlas.)
- (3) Suponiendo (sólo en este apartado) que  $B = \{(1, 1, 0), (1, 0, 1), (0, 1, 1)\}$ , hallar una base de  $N(f)$ .

**5.36** Sea  $f : K^3 \rightarrow K^4$  una aplicación lineal cuya matriz en las bases  $B = \{(1, 1, 0), (1, 2, 1), (0, 1, 2)\}$  de  $K^3$  y canónica  $D$  de  $K^4$  es

$$\begin{pmatrix} 1 & a-b & 2 \\ 2 & 3a-b & 5 \\ 1 & 2a & a-b+5 \\ -1 & 2b & 2a-2b+3 \end{pmatrix}$$

Clasificar la aplicación de acuerdo con los valores de  $a$  y  $b$ . Ahora tomando los valores  $a = 1$  y  $b = -1$

- (1) Dar ecuaciones implícitas del núcleo y la imagen.
- (2) Dada una base de la imagen y extenderla a una base de  $K^4$ .
- (3) Determinar si el vector  $(1, 1, 1, 1)$  pertenece a la imagen de  $f$ .

**5.37** Consideremos las aplicaciones  $S, H : M_n(K) \rightarrow M_n(K)$  dadas por

$$S(A) = \frac{A + A^T}{2}, \quad H(A) = \frac{A - A^T}{2}.$$

- (1) Demostrar que  $S$  y  $H$  son lineales.
- (2) Determinar sus núcleos e imágenes.
- (3) Comprobar que  $S^2 = S$ ,  $H^2 = H$  y  $S + H$  es la aplicación identidad.
- (4) Dar las matrices en las bases canónicas para  $n = 2$ .

**5.38** Sea  $f : V \rightarrow W$  una aplicación lineal y sea  $w$  un elemento de la imagen de  $f$ . Demostrar que  $f^{-1}(w) = \{v \in V : f(v) = w\}$  es una variedad lineal de  $V$  con dirección  $N(f)$ .



## Capítulo 6

# Diagonalización y formas de Jordan

A lo largo de todo el capítulo,  $K$  es un cuerpo y  $V$  es un espacio vectorial de dimensión finita sobre  $K$ .

### 6.1 Endomorfismos y matrices diagonalizables

**Definición 6.1.1** Un endomorfismo  $f$  de un espacio vectorial  $V$  se dice diagonalizable si existe una base  $B$  de  $V$  tal que  $f_B$  es una matriz diagonal. En tal caso decimos que  $B$  diagonaliza  $f$ .

Dos matrices cuadradas  $M$  y  $N$  se dice que son semejantes si existe una matriz invertible  $U$  tal que  $M = U^{-1}NU$ .

Una matriz se dice que es diagonalizable si es semejante a una matriz diagonal. Si  $U$  es una matriz invertible tal que  $U^{-1}MU$  es diagonal decimos que  $U$  diagonaliza  $M$ .

**Proposición 6.1.2** Sea  $f$  un endomorfismo de  $V$  y  $B$  una base de  $V$ . Entonces  $f$  es diagonalizable si y sólo si  $f_B$  es diagonalizable. Además una base  $D$  de  $V$  diagonaliza  $f$  si y sólo si la matriz de cambio de base  $D_B$  diagonaliza  $f_B$ .

**Demostración.** Supongamos que  $f$  es diagonalizable y  $D$  es una base de  $V$  que diagonaliza  $f$ . Entonces  $f_D$  es diagonal y de (5.8) tenemos que

$$f_D = D_B^{-1} f_B D_B$$

de donde deducimos que  $f_B$  es diagonalizable y  $D_B$  diagonaliza  $f_B$ .

Recíprocamente, supongamos que  $f_B$  es diagonalizable y sea  $U$  una matriz invertible que diagonaliza  $f_B$ , de forma que  $U^{-1}f_B U$  es diagonal. Entonces, el conjunto de vectores de  $V$  cuyas coordenadas son las columnas de  $U$  forma una base de  $V$  y  $D_B = U$ . Entonces  $f_D = D_B^{-1} f_B D_B = U^{-1} f_B U$  es diagonal. ■

*Diagonalizar un endomorfismo* consiste encontrar una base que lo diagonalice y *diagonalizar una matriz* en encontrar una matriz que la diagonalice. La Proposición 6.1.2 muestra que estos dos problemas son equivalentes.

**Ejemplo 6.1.3** Consideremos el endomorfismo  $f$  de  $K^3$  dado por

$$f \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 13x - 14y - 10z \\ 6x - 6y - 6z \\ 4x - 8y - z \end{pmatrix}.$$

Si  $B$  es la base canónica, entonces

$$f_B = \begin{pmatrix} 13 & -14 & -10 \\ 6 & -6 & -6 \\ 4 & -8 & -1 \end{pmatrix}.$$

Si

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

entonces

$$U^{-1}AU = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & -3 \end{pmatrix}.$$

Luego  $U$  diagonaliza  $A$  y por tanto

$$D = \left\{ v_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, v_3 = \begin{pmatrix} 3 \\ 2 \\ 2 \end{pmatrix} \right\}$$

diagonaliza  $f$ .

**Definición 6.1.4** Sea  $f : V \rightarrow V$  un endomorfismo. Un vector propio o autovector de  $f$  es un elemento  $v \in V$  tal que  $f(v) = \lambda v$  para algún escalar  $\lambda \in K$ .

Un valor propio o autovalor de  $f$  es un escalar  $\lambda$  tal que existe un  $0 \neq v \in V$  tal que  $f(v) = \lambda v$ . En tal caso se dice que  $\lambda$  es el valor propio del vector propio  $v$  o que  $v$  es un vector propio de  $f$  con valor propio  $\lambda$ .

Sea  $A \in M_n(K)$  una matriz. Un vector propio o autovector de  $M$  es un vector propio de  $f_A$  y un valor propio es un valor propio de  $f_A$ . Es decir, un vector propio de  $A$  es un elemento  $v \in K^n$  tal que  $Av = \lambda v$  para algún escalar  $\lambda$  y si  $v \neq 0$  se dice que  $\lambda$  es un valor propio de  $A$  y  $v$  es un vector propio con valor propio  $\lambda$ .

Obsérvese que el vector  $0$  es un vector propio de cualquier endomorfismo o de cualquier matriz cuadrada, ya que  $f(0) = \lambda 0$  para todo  $\lambda \in K$ . Sin embargo si  $v \neq 0$  es un vector propio de  $f$  y  $f(v) = \lambda v$ , entonces el escalar  $\lambda$  es único, es decir, si  $f(v) = \mu v$  para otro escalar  $\mu$  entonces  $\lambda = \mu$ .

**Ejemplo 6.1.5** Consideremos el endomorfismo de  $\mathbb{R}^4$  dado por  $f(x_1, x_2, x_3) = (-2x_1 + x_2 + 2x_2, 2x_1 - 2x_2 + 3x_3, 3x_1 + 2x_2 + 2x_3)$ . Entonces  $f(1, 1, -1) = (-3, -3, 3) = -3(1, 1, -1)$ . Por tanto  $(1, 1, -1)$  es un vector propio de  $f$  con valor propio  $-3$ .

**Ejemplo 6.1.6** Sea  $V = K^{\mathbb{N}}$  el espacio vectorial de las sucesiones infinitas de elementos del cuerpo  $K$ , o sea los elementos de  $V$  son sucesiones  $(x_n) = (x_1, x_2, \dots)$ . La suma en  $V$  y el producto por escalares es el natural:

$$(x_n) + (y_n) = (x_n + y_n) \quad \alpha(x_n) = (\alpha x_n).$$

Consideremos el endomorfismo de  $f$  dado por  $f(x_1, x_2, \dots) = (x_2, x_3, \dots)$ . Para cualquier sucesión constante  $v = (a, a, \dots)$  se cumple  $f(v) = v$ . Por tanto las sucesiones constantes son vectores propios con valor propio 1.

**Teorema 6.1.7** (1) Sea  $f$  un endomorfismo del espacio vectorial  $V$ . Una base de  $V$  diagonaliza  $f$  si y sólo si está formada por vectores propios de  $V$ . Por tanto  $f$  es diagonalizable si y sólo si  $V$  tiene una base formada por vectores propios de  $f$ .

(2) Sea  $A \in M_n(K)$  una matriz. Una matriz invertible  $U \in M_n(K)$  diagonaliza  $A$  si y sólo si sus columnas son vectores propios de  $V$ . Por tanto  $A$  es diagonalizable si y sólo si  $K^n$  tiene una base formada por vectores propios de  $A$ .

**Demostración.** Como consecuencia de la Proposición 6.1.2 las propiedades (1) y (2) son equivalentes. Por tanto sólo vamos a demostrar (1).

Si  $B = \{v_1, v_2, \dots, v_n\}$  es una base de  $V$  formada por vectores propios de  $V$ , entonces para cada  $i = 1, 2, \dots, n$ ,  $f(v_i) = \lambda_i v_i$  para algún escalar  $\lambda_i$ . Entonces

$$f_B = \begin{pmatrix} \lambda_1 & 0 & 0 \dots & 0 \\ 0 & \lambda_2 & 0 \dots & 0 \\ 0 & 0 & \lambda_3 \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & 0 \dots & \lambda_n \end{pmatrix}.$$

Por tanto  $B$  diagonaliza  $f$ .

Recíprocamente, supongamos que  $B = \{v_1, v_2, \dots, v_n\}$  es una base de  $V$  tal que  $f_B$  es diagonal. Pongamos

$$f_B = \begin{pmatrix} \lambda_1 & 0 & 0 \dots & 0 \\ 0 & \lambda_2 & 0 \dots & 0 \\ 0 & 0 & \lambda_3 \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & 0 \dots & \lambda_n \end{pmatrix}.$$

Entonces  $f(v_i) = \lambda_i v_i$  para cada  $i = 1, 2, \dots, n$ , es decir  $B$  es una base de  $V$  formada por vectores propios de  $f$ . ■

Por tanto para decidir si un endomorfismo o una matriz es diagonalizable tenemos que intentar encontrar una base de vectores propios. Este proceso se llama *diagonalización* de un endomorfismo o una matriz. Como ya hemos observado más arriba los dos problemas, de diagonalizar un endomorfismo o una matriz son en realidad el mismo. Mas concretamente, diagonalizar una matriz  $A$  es lo mismo que diagonalizar el endomorfismo  $f_A$  pues si  $v_1, \dots, v_n$

es una base de vectores propios de  $f_A$  entonces la matriz que tiene  $v_1, \dots, v_n$  como columnas cumple que  $U^{-1}AU$  es diagonal. Recíprocamente, si  $f$  es un endomorfismo de  $V$  y  $B$  es una base de  $V$  entonces para diagonalizar  $f$  basta con diagonalizar la matriz  $B$  pues un elemento  $v$  de  $V$  es un vector propio de  $f$  precisamente si  $v_B$ , su vector de coordenadas en la base  $B$  es un vector propio de  $A$ , con lo cual si sabemos diagonalizar  $f_B$ , es decir conocemos una matriz invertible  $U$  para el que  $U^{-1}f_B U$  es diagonal entonces los vectores que tienen como coordenadas en la base  $B$  las columnas de  $U$  forman una base de vectores propios de  $f$ .

Recordemos que si  $P \in K[X]$  es un polinomio con coeficientes en  $K$  y  $a \in K$ , entonces  $P(a)$  es el elemento de  $K$  que se obtiene al sustituir la indeterminada  $X$  por  $a$  y que un elemento  $a$  de  $K$  se dice que es una raíz de  $P$  si  $P(a) = 0$ . (Ver la Sección 1.3.)

**Definición 6.1.8** Si  $A \in M_n(K)$  es una matriz se llama polinomio característico de  $A$  al polinomio que se obtiene al calcular el siguiente determinante:

$$\chi_A = |XI_n - A|.$$

**Ejemplo 6.1.9** Si

$$A = \begin{pmatrix} -1 & 3 \\ -4 & 6 \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} 2 & -1 & 1 \\ 0 & 2 & 1 \\ -1 & 0 & -1 \end{pmatrix}$$

entonces

$$\chi_A = \begin{vmatrix} X+1 & -3 \\ 4 & X-6 \end{vmatrix} = (X+1)(X-6) + 12 = X^2 - 5X + 6$$

y

$$\chi_B = \begin{vmatrix} X-2 & 1 & -1 \\ 0 & X-2 & -1 \\ 1 & 0 & X+1 \end{vmatrix} = (X-2)^2(X+1) - 1 + 3(X-2) = X^3 - 3X^2 + X + 1.$$

**Lema 6.1.10** Si  $A$  y  $B$  son dos matrices semejantes, entonces  $\chi_A = \chi_B$ .

**Demostración.** Si  $A$  y  $B$  son semejantes, entonces existe una matriz invertible  $U$  tal que  $B = U^{-1}AU$ . Entonces

$$\begin{aligned} \chi_B &= |XI - B| = |XU^{-1}U - U^{-1}AU| = |U^{-1}(XI - A)U| \\ &= |U|^{-1}|XI - A||U| = |XI - A| = \chi_A. \end{aligned}$$

■

Sea  $f$  un endomorfismo de  $V$ . Entonces las matrices asociadas a  $f$  en todas las bases son semejantes y por tanto los polinomios característicos de todas estas matrices son iguales. Por tanto podemos definir el *polinomio característico de un endomorfismo  $f$*  como

$$\chi_f = \chi_{f_B}$$

donde  $B$  es cualquier base de  $V$ .

**Ejemplo 6.1.11** Consideremos el endomorfismo de  $\mathbb{R}^3$  dado por la fórmula

$$f(x, y, z) = (z - y, 3x + 4y - 3z, x + y).$$

La matriz asociada a este endomorfismo en la base canónica es

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

Por tanto

$$\chi_f = \chi_A = \begin{vmatrix} X & 1 & -1 \\ -3 & X-4 & 3 \\ -1 & -1 & X \end{vmatrix} = X^3 - 4X^2 + 5X - 2.$$

**Ejemplo 6.1.12** Consideremos la aplicación  $T$  que asocia a cada  $A \in M_2(K)$  la transpuesta de  $A$ . Fijamos la base  $B = \{e_{11}, e_{12}, e_{21}, e_{22}\}$  donde  $e_{ij}$  es la matriz que tiene un 1 en la entrada  $(i, j)$  y cero en las demás entradas. Como  $T(e_{11}) = e_{11}$ ,  $T(e_{12}) = e_{21}$ ,  $T(e_{21}) = e_{12}$  y  $T(e_{22}) = e_{22}$  la matriz asociada en esta base es

$$T_B = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix}$$

y por tanto

$$\chi_B = \begin{vmatrix} X-1 & & & \\ & X & -1 & \\ & -1 & X & \\ & & & X-1 \end{vmatrix} = (X-1)^2(X^2-1) = (X-1)^3(X+1).$$

Sea  $A$  una matriz cuadrada y sea  $\lambda \in K$  un escalar. Vamos a denotar por  $N(A, \lambda)$  al espacio nulo de  $\lambda I - A$ . Obsérvese que los vectores propios de  $A$  de valor propio  $\lambda$  son los elementos de  $N(A, \lambda)$ . Por tanto

**Proposición 6.1.13** *Las siguientes condiciones son equivalentes para una matriz cuadrada  $A \in M_n(K)$  y un escalar  $\lambda \in K$ .*

- (1)  $\lambda$  es un valor propio de  $A$ .
- (2)  $N(A, \lambda) \neq 0$ .
- (3)  $|\lambda I - A| = 0$ .
- (4)  $\lambda$  es una raíz del polinomio característico  $\chi_A$  de  $A$ .

Análogamente si  $f$  es un endomorfismo de un espacio vectorial  $V$  sobre  $K$  y  $\lambda \in K$  es un escalar, denotamos por  $N(f, \lambda)$  al núcleo de  $\lambda 1_V - f$  y tenemos:

**Proposición 6.1.14** *Las siguientes condiciones son equivalentes para un endomorfismo  $f$  de un espacio vectorial  $V$  sobre  $K$  y un escalar  $\lambda \in K$ .*

- (1)  $\lambda$  es un valor propio de  $f$ .
- (2)  $N(f, \lambda) \neq 0$ .
- (3)  $\lambda$  es una raíz del polinomio característico  $\chi_A$  de  $f$ .

Por tanto para diagonalizar una matriz  $A$  o un endomorfismo  $f$  procedemos de la siguiente forma:

**Algoritmo para diagonalizar una matriz cuadrada  $A$ :**

- (1) Calculamos el polinomio característico  $\chi_A$  de  $A$ .
- (2) Calculamos las raíces  $\lambda_1, \lambda_2, \dots, \lambda_k$  de  $\chi_A$ , que son los valores propios de  $A$  (respectivamente,  $f$ ).
- (3) Para cada  $i = 1, 2, \dots, k$  calculamos una base  $B_i$  de  $N(A, \lambda_i)$ .
- (4) Juntamos todos los  $B_i$  en una lista

$$B = B_1 \cup B_2 \cup \dots \cup B_k.$$

- (5) Si  $B$  tiene tantos elementos como el tamaño de  $A$ , entonces  $A$  es diagonalizable y  $B$  es una base de vectores propios de  $A$ . Por tanto, poniendo los vectores de  $B$  por columnas en una matriz,  $U$  tendremos que  $U^{-1}AU$  es diagonal. La matriz  $U$  se llama a menudo *matriz de paso*.
- (6) En caso contrario, la matriz  $A$  no es diagonalizable.

**Ejemplos 6.1.15** Sean  $A$  y  $B$  las dos matrices del Ejemplo 6.1.9.

Como el polinomio característico de  $A$  es  $X^2 - 5X + 6$ , los valores propios de  $A$  son las raíces de este polinomio, es decir 2 y 3. Ahora calculamos los espacios nulos  $N(A, 2)$  y  $N(A, 3)$  de  $A - 2I$  y  $A - 3I$  respectivamente. Empezamos con el primero que es el espacio nulo de la matriz

$$A - 2I = \begin{pmatrix} -3 & 3 \\ -4 & 4 \end{pmatrix}.$$

Como el rango de esta matriz es 1, su espacio nulo también tiene dimensión 1 y una base de  $N(A, 2)$  es  $\{v_1 = (1, 1)\}$ . Ahora calculamos una base del  $N(A, 3)$ . Como

$$A - 3I = \begin{pmatrix} -4 & 3 \\ -4 & 3 \end{pmatrix}$$

Una base de  $N(A, 3)$  es  $\{v_2 = (3, 4)\}$ . Entonces  $\{v_1, v_2\}$  es una base de vectores propios y poniendo  $v_1$  y  $v_2$  en columna obtenemos la matriz de paso:

$$U = \begin{pmatrix} 1 & 3 \\ 1 & 4 \end{pmatrix}.$$

Comprobamos:

$$U^{-1}AU = \begin{pmatrix} 1 & 3 \\ 1 & 4 \end{pmatrix}^{-1} \begin{pmatrix} -1 & 3 \\ -4 & 6 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 1 & 4 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}.$$

Vamos ahora con la matriz  $B$  cuyo polinomio característico es  $X^3 - 3X^2 + X + 1$ . Las raíces de este polinomio son  $1$ ,  $1 - \sqrt{2}$  y  $1 + \sqrt{2}$ . Tenemos

$$B - I = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & 1 \\ -1 & 0 & -2 \end{pmatrix}, \quad B - (1 - \sqrt{2})I = \begin{pmatrix} 1 + \sqrt{2} & -1 & 1 \\ 0 & 1 + \sqrt{2} & 1 \\ -1 & 0 & -2 + \sqrt{2} \end{pmatrix}$$

$$B - (1 + \sqrt{2})I = \begin{pmatrix} 1 - \sqrt{2} & -1 & 1 \\ 0 & 1 - \sqrt{2} & 1 \\ -1 & 0 & -2 - \sqrt{2} \end{pmatrix}.$$

y entonces

$$N(B, 1) = \langle v_1 = (-2, -1, 1) \rangle,$$

$$N(B, 1 - \sqrt{2}) = \langle v_2 = (-2 + \sqrt{2}, 1 - \sqrt{2}, 1) \rangle,$$

$$N(B, 1 + \sqrt{2}) = \langle v_3 = (-2 - \sqrt{2}, 1 + \sqrt{2}, 1) \rangle.$$

Por tanto  $\{v_1, v_2, v_3\}$  es una base de vectores propios. Luego, si ponemos

$$U = \begin{pmatrix} -2 & -2 + \sqrt{2} & -2 - \sqrt{2} \\ -1 & 1 - \sqrt{2} & 1 + \sqrt{2} \\ 1 & 1 & 1 \end{pmatrix}$$

tendremos

$$U^{-1}AU = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 - \sqrt{2} & 0 \\ 0 & 0 & 1 + \sqrt{2} \end{pmatrix}.$$

#### Algoritmos para diagonalizar un endomorfismo $f$ de $V$ :

- (1) Elegimos una base  $B$  de  $V$ .
- (2) Calculamos  $A = f_B$ .
- (3) Aplicamos el algoritmo anterior a  $A$ .
- (4) Si  $A$  no es diagonalizable, tampoco lo es  $f$ .
- (5) Si  $A$  es diagonalizable y la matriz invertible  $U$  diagonaliza  $A$ , entonces la base  $D$  de  $V$  tal que  $D_B = U$  diagonaliza  $f$ .

**Ejemplo 6.1.16** Sea ahora  $f$  el endomorfismo del Ejemplo 6.1.11. Ya vimos que el polinomio característico de  $f$  es  $\chi_f = X^3 - 4X^2 + 5X - 2$ . Es fácil ver que 1 es una raíz

$$\begin{array}{c|ccc} 1 & 1 & -4 & 5 & -2 \\ & & 1 & -3 & 2 \\ \hline & 1 & -3 & 2 & |0 \end{array}$$

Obtenemos que

$$\chi_A = (X - 1)(X^2 - 3X + 2)$$

y también 1 es una raíz de  $X^2 - 3X + 2$ :

$$\begin{array}{c|ccc} 1 & 1 & -3 & 2 \\ & & 1 & -2 \\ \hline & 1 & -2 & |0 \end{array}$$

Por tanto

$$\chi_A = (X - 1)^2(X - 2).$$

Es decir, los valores propios de  $A$  son 1 y 2. Ahora calculamos  $N(A, 1)$  donde  $A$  es la matriz en la base canónica que ya calculamos en el Ejemplo 6.1.11. La matriz

$$A - I = \begin{pmatrix} -1 & -1 & 1 \\ 3 & 3 & -3 \\ 1 & 1 & -1 \end{pmatrix}$$

tiene rango 1 con lo que  $N(A, 1)$  tiene dimensión 2 y la siguiente es una base suya

$$w_1 = (1, -1, 0) \quad \text{y} \quad w_2 = (1, 0, 1).$$

Ahora calculamos  $N(A, 2)$

$$A - 2I = \begin{pmatrix} -2 & -1 & 1 \\ 3 & 2 & -3 \\ 1 & 1 & -2 \end{pmatrix}.$$

Una base del espacio nulo de esta matriz está formada por

$$w_3 = (-1, 3, 1)$$

Por tanto  $\{w_1, w_2, w_3\}$  es una base vectores propios de  $A$  y por tanto la matriz

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

diagonaliza  $A$ . Echando las cuentas obtenemos

$$f_D = U^{-1}fU = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

**Ejemplo 6.1.17** Vamos ahora a diagonalizar el endomorfismo  $T$  definido por la transposición en  $M_2(K)$ . Ya calculamos su polinomio característico en el Ejemplo 6.1.12:  $\chi_T = (X-1)^3(X+1)$ . Por tanto los valores propios son 1 y  $-1$ . Para calcular  $N(T, 1)$  y  $N(T, -1)$  usamos la matriz  $A = T_B$  asociada en la base  $B = \{e_{11}, e_{12}, e_{21}, e_{22}\}$ . Para calcular  $N(T, 1)$  tenemos que calcular el espacio nulo de la matriz

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

que está formado por los vectores de la forma  $(x_1, x_2, x_2, x_4)$ . Una base suya es por ejemplo

$$(1, 0, 0, 0), \quad (0, 1, 1, 0), \quad (0, 0, 0, 1).$$

Por tanto una base de  $N(T, 1)$  es

$$e_{11}, \quad e_{12} + e_{21}, \quad e_{22}.$$

Ahora calculamos una base de  $N(T, -1)$  para lo cual calculamos una base del espacio nulo de

$$A + I = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

Claramente  $(0, 1, -1, 0)$  es una base de este espacio nulo y por tanto  $e_{12} - e_{21}$  es una base de  $N(T, -1)$ . Por tanto,  $T$  es diagonalizable y la siguiente es una base de vectores propios:

$$e_{11}, \quad e_{12} + e_{21}, \quad e_{22}, \quad e_{12} - e_{21}.$$

Hay una cosa de los algoritmos anteriores que no hemos justificado. ¿Por que estamos tan seguros de que si  $B_1, B_2, \dots, B_k$  son bases de los vectores propios  $\lambda_1, \lambda_2, \dots, \lambda_k$ , entonces  $B_1 \cup B_2 \cup \dots \cup B_k$  es linealmente independiente? Esto es consecuencia de la siguiente proposición.

**Proposición 6.1.18** *Si  $\lambda_1, \lambda_2, \dots, \lambda_k$  son valores propios distintos de un endomorfismo  $f$  de un espacio vectorial entonces  $N(f, \lambda_1), \dots, N(f, \lambda_k)$  son independientes.*

**Demostración.** Razonamos por inducción sobre  $k$ , siendo el caso  $k = 1$  trivial. Supongamos que se verifica la propiedad para  $k - 1$  valores propios.

Sean

$$v_1 \in N(f, \lambda_1), \quad v_2 \in N(f, \lambda_2), \quad \dots, \quad v_k \in N(f, \lambda_k),$$

tales que

$$v_1 + v_2 + \dots + v_k = 0.$$

Entonces

$$\begin{aligned} 0 &= f(v_1 + v_2 + \dots + v_k) - \lambda_k(v_1 + v_2 + \dots + v_k) \\ &= (\lambda_1 - \lambda_k)v_1 + (\lambda_2 - \lambda_k)v_2 + \dots + (\lambda_{k-1} - \lambda_k)v_{k-1} + (\lambda_k - \lambda_k)v_k \\ &= (\lambda_1 - \lambda_k)v_1 + (\lambda_2 - \lambda_k)v_2 + \dots + (\lambda_{k-1} - \lambda_k)v_{k-1} \end{aligned}$$

Pero

$$(\lambda_1 - \lambda_k)v_1 \in N(V, \lambda_1), \quad (\lambda_2 - \lambda_k)v_2 \in N(V, \lambda_2), \dots, \quad (\lambda_{k-1} - \lambda_k)v_{k-1} \in N(V, \lambda_{k-1})$$

y por la Hipótesis de Inducción deducimos que

$$(\lambda_1 - \lambda_k)v_1 = (\lambda_2 - \lambda_k)v_2 = \dots = (\lambda_{k-1} - \lambda_k)v_{k-1} = 0.$$

Como los  $\lambda_i$  son diferentes dos a dos, tenemos que

$$v_1 = v_2 = \dots = v_{k-1} = 0$$

y, por tanto también  $v_k = 0$ . ■

De las Proposiciones 6.1.18 y (4.2.12) tenemos que si  $\lambda_1, \dots, \lambda_k$  son los valores propios de  $f$  y para cada  $i = 1, \dots, k$ , se tiene una base  $B_i$  de  $N(f, \lambda_i)$  entonces  $B_1 \cup B_2 \cup \dots \cup B_k$  es una base de  $N(f, \lambda_1) \oplus N(f, \lambda_2) \oplus \dots \oplus N(f, \lambda_k)$ . Como todos los vectores propios distintos de cero estarán en alguna base de algún  $N(f, \lambda_i)$  esta será la única forma posible de obtener bases de vectores propios. Por tanto

**Corolario 6.1.19** Sean  $\lambda_1, \lambda_2, \dots, \lambda_k$  los valores propios (todos y distintos dos a dos) de un endomorfismo  $f$  de un espacio vectorial  $V$ . Entonces  $f$  es diagonalizable si y sólo si

$$\dim N(f, \lambda_1) + \dim N(f, \lambda_2) + \dots + \dim N(f, \lambda_k) = \dim V.$$

## 6.2 ¿Cuándo una matriz o un endomorfismo es diagonalizable

En contra de lo que podría parecer por los ejemplos de la Sección 6.1 no todos las las matrices (y por tanto tampoco todos los endomorfismos) son diagonalizables. Hay dos obstáculos posibles para que un endomorfismo o una matriz no sea diagonalizable: que no haya suficientes valores propios o que no haya suficientes vectores propios. Vamos a hacer esto más preciso con ejemplos. Empecemos con un ejemplo que no tiene ningún valor propio.

**Ejemplo 6.2.1** Consideremos la siguiente matriz de números reales:

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

Entonces  $\chi_A = X^2 + 1$  que no tiene ninguna raíz real. Por tanto  $A$  no tiene vectores propios en  $\mathbb{R}^2$  y por tanto  $A$  no es diagonalizable.

Si el cuerpo de los escalares es  $\mathbb{C}$ , la patología del Ejemplo 6.2.1 no puede ocurrir. De hecho toda matriz de números reales la podemos considerar como una matriz de números complejos y diagonalizarla como tal.

**Ejemplo 6.2.2** Sea

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

como en el Ejemplo 6.2.1. Entonces  $\chi_A = X^2 + 1 = (X - i)(X + i)$ , con lo que los valores propios de  $A$  son  $i$  y  $-i$ . Calculando los espacios nulos de

$$A - iI = \begin{pmatrix} -i & 1 \\ -1 & -i \end{pmatrix} \quad \text{y} \quad A + iI = \begin{pmatrix} i & 1 \\ -1 & i \end{pmatrix}$$

obtenemos que

$$N(A, i) = \left\langle \left( \begin{pmatrix} 1 \\ i \end{pmatrix} \right) \right\rangle \quad \text{y} \quad N(A, -i) = \left\langle \left( \begin{pmatrix} 1 \\ -i \end{pmatrix} \right) \right\rangle$$

con lo que

$$U = \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}$$

diagonaliza  $A$ :

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

Sin embargo, aún sobre los números complejos hay matrices no diagonalizables.

**Ejemplo 6.2.3** Sea

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

Entonces

$$\chi_A = (X - 1)^2$$

con lo que el único valor propio de  $A$  es 1. Como

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

tiene rango 1, la dimensión de  $N(A, 1)$  es 1. Es decir todos los vectores propios de  $A$  están en un espacio vectorial de dimensión 1, a saber  $\langle (1, 0) \rangle$ . Por tanto  $K^2$  no tiene una base de vectores propios de  $A$ , es decir  $A$  no es diagonalizable.

**Teorema 6.2.4** *Un endomorfismo  $f$  de un espacio vectorial  $V$  de dimensión finita es diagonalizable si y sólo si satisface las siguientes condiciones:*

- (1) *El polinomio característico  $\chi_f$  de  $f$  es completamente factorizable.*
- (2) *Para cada valor propio  $\lambda$  de  $f$ , la dimensión de  $N(f, \lambda)$  coincide con la multiplicidad de  $\lambda$  como raíz de  $\chi_f$ .*

**Demostración.** Si se verifican las dos condiciones, entonces la suma de las dimensiones de los  $N(f, \lambda)$  coincide con la suma de las multiplicidades de los valores propios de  $f$  y esta coincide con la dimensión de  $V$ . Por el Corolario 6.2.4,  $f$  es diagonalizable.

Recíprocamente, supongamos que  $f$  es diagonalizable y sea  $B$  una base de  $V$  formada por vectores propios de  $f$ . Sean  $\lambda_1, \lambda_2, \dots, \lambda_k$  los valores propios de  $V$ . Podemos ordenar  $B$  de forma que los vectores propios con el mismo valor propio estén juntos, es decir podemos suponer que

$$B = \{v_{11}, v_{12}, \dots, v_{1,n_1}, v_{21}, v_{22}, \dots, v_{2,n_2}, \dots, v_{k1}, v_{k2}, \dots, v_{k,n_k}\}$$

donde cada  $v_{ij} \in N(f, \lambda_i)$ . En particular  $\dim N(f, \lambda_i) \geq n_i$ . De la Proposición 6.1.18 tenemos

$$\dim V \geq \dim N(f, \lambda_1) + \dim N(f, \lambda_2) + \dots + \dim N(f, \lambda_k) \geq n_1 + n_2 + \dots + n_k = \dim V$$

y por tanto  $\dim N(f, \lambda_i) = n_i$  para todo  $i = 1, 2, \dots, k$ . Entonces

$$f_B = \left( \begin{array}{c|c|c|c} \lambda_1 I_{n_1} & & & \\ \hline & \lambda_2 I_{n_2} & & \\ \hline & & \ddots & \\ \hline & & & \lambda_k I_{n_k} \end{array} \right)$$

donde los huecos están formados por matrices nulas. El polinomio característico de  $f$  es el polinomio característico de  $f_B$ , es decir

$$\begin{aligned} \chi_f &= |XI - f_B| = \left| \begin{array}{c|c|c|c} (X - \lambda_1)I_{n_1} & & & \\ \hline & (X - \lambda_2)I_{n_2} & & \\ \hline & & \ddots & \\ \hline & & & (X - \lambda_k)I_{n_k} \end{array} \right| \\ &= |X - \lambda_1 I_{n_1}| |X - \lambda_2 I_{n_2}| \cdots |X - \lambda_k I_{n_k}| \\ &= (X - \lambda_1)^{n_1} (X - \lambda_2)^{n_2} \cdots (X - \lambda_k)^{n_k} \end{aligned}$$

Por tanto  $\chi_f$  es completamente reducible y la multiplicidad de cada  $\lambda_i$  es  $n_i$ . ■

Como consecuencia del Teorema 6.2.4 y del Teorema Fundamental del Álgebra (Teorema 1.4.7) tenemos el siguiente corolario.

**Corolario 6.2.5** *Si  $f$  es un endomorfismo de un espacio vectorial de dimensión finita sobre el cuerpo  $\mathbb{C}$  de los números complejos, entonces  $f$  es diagonalizable si y sólo si para cada valor propio  $\lambda$  de  $f$ , la dimensión de  $N(f, \lambda)$  coincide con la multiplicidad de  $\lambda$  en  $\chi_f$ .*

Por supuesto que los mismos resultados valen para matrices.

**Teorema 6.2.6** *Una matriz cuadrada  $A \in M_n(K)$  es diagonalizable si y sólo si satisface las siguientes condiciones:*

- (1) *El polinomio característico  $\chi_A$  de  $A$  es completamente factorizable.*

(2) Para cada valor propio  $\lambda$  de  $A$ , la dimensión de  $N(A, \lambda)$  coincide con la multiplicidad de  $\lambda$  en  $\chi_A$ .

**Corolario 6.2.7** Una matriz cuadrada  $A \in M_n(\mathbb{C})$  es diagonalizable si y sólo si para cada valor propio  $\lambda$  de  $A$ , la dimensión de  $N(A, \lambda)$  coincide con la multiplicidad de  $\lambda$  en  $\chi_A$ .

**Corolario 6.2.8** Si una matriz cuadrada  $A \in M_n(K)$  o un endomorfismo  $f$  de un espacio vectorial de dimensión  $n$  tiene  $n$  valores propios distintos, entonces es diagonalizable.

### 6.3 Aplicación 1: Potencias de matrices y sucesiones recurrentes

Calcular las potencias de una matriz diagonal

$$A = \begin{pmatrix} a_1 & 0 & 0 & \cdots & 0 \\ 0 & a_2 & 0 & \cdots & 0 \\ 0 & 0 & a_3 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & a_n \end{pmatrix}.$$

es muy sencillo. En efecto, para todo número  $k$  tenemos

$$A^k = \begin{pmatrix} a_1^k & 0 & 0 & \cdots & 0 \\ 0 & a_2^k & 0 & \cdots & 0 \\ 0 & 0 & a_3^k & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & a_n^k \end{pmatrix}.$$

Si  $A$  no es diagonal pero si es diagonalizable, entonces también es fácil calcular sus potencias. En efecto, si  $U$  es una matriz que diagonaliza  $A$  y, por tanto

$$D = U^{-1}AU$$

es una matriz diagonal, entonces

$$A = UDU^{-1}$$

y por tanto

$$A^k = (UDU^{-1})(UDU^{-1}) \cdots (UDU^{-1}) = UD^kU^{-1}.$$

**Ejemplo 6.3.1** Consideremos la matriz

$$A = \frac{1}{5} \begin{pmatrix} 193 & -342 & 108 \\ 132 & -233 & 72 \\ 66 & -114 & 31 \end{pmatrix}.$$

Entonces

$$\chi_A = \frac{5X^3 + 9X^2 + 3X - 1}{5} = (X + 1)^2 \left( X - \frac{1}{5} \right),$$

con lo que los valores propios de  $A$  son  $-1$ , con multiplicidad 2 y  $\frac{1}{5}$  con multiplicidad 1. Calculando vectores propios obtenemos una base de  $N(A, -1)$  formada por los vectores

$$(-6, 0, 11) \quad \text{y} \quad (19, 11, 0)$$

y una base de vectores propios de  $N(A, \frac{1}{5})$  formada por el vector

$$(1, 0, 1).$$

Luego

$$U = \begin{pmatrix} -6 & 19 & 3 \\ 0 & 11 & 2 \\ 11 & 0 & 1 \end{pmatrix}$$

y

$$A = U \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & \frac{1}{5} \end{pmatrix} U^{-1}.$$

Por tanto

$$\begin{aligned} A^k &= U \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & \frac{1}{5} \end{pmatrix}^k U^{-1} \\ &= \begin{pmatrix} -6 & 19 & 3 \\ 0 & 11 & 2 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} (-1)^k & 0 & 0 \\ 0 & (-1)^k & 0 \\ 0 & 0 & \frac{1}{5^k} \end{pmatrix} \frac{1}{11} \begin{pmatrix} -11 & 19 & -5 \\ -22 & 39 & -12 \\ 121 & -209 & 66 \end{pmatrix} \\ &= \begin{pmatrix} -32(-1)^k + 33\frac{1}{5^k} & 57(-1)^k - 57\frac{1}{5^k} & -18(-1)^k + 18\frac{1}{5^k} \\ -22(-1)^k + 22\frac{1}{5^k} & 39(-1)^k - 38\frac{1}{5^k} & -12(-1)^k + 12\frac{1}{5^k} \\ -11(-1)^k + 11\frac{1}{5^k} & 19(-1)^k - 19\frac{1}{5^k} & -5(-1)^k + 6\frac{1}{5^k} \end{pmatrix} \end{aligned}$$

Por ejemplo,

$$A^4 = \frac{1}{625} \begin{pmatrix} -19967 & 35568 & -11232 \\ -13728 & 24337 & -7488 \\ -6864 & 11856 & -3119 \end{pmatrix}.$$

Una *sucesión recurrente lineal* de grado  $k$  es una sucesión de números  $(a_n)$  en la que cada término  $a_n$  con  $n > k$  viene determinado de forma lineal por los  $k$  anteriores. Por ejemplo, la sucesión de Fibonacci es la sucesión  $(a_n)$  tal que

$$a_1 = a_2 = 1 \quad \text{y} \quad a_n = a_{n-1} + a_{n-2}, \quad \text{para todo } n \geq 3.$$

Vamos a calcular una fórmula para el término general de la sucesión de Fibonacci. Obsérvese que se verifica la siguiente fórmula

$$\begin{pmatrix} a_{n-1} \\ a_n \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} a_{n-2} \\ a_{n-1} \end{pmatrix}.$$

Por tanto, si ponemos

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

tendremos

$$\begin{pmatrix} a_{n-1} \\ a_n \end{pmatrix} = A^{n-2} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = A^{n-2} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Vamos a calcular  $A^n$ . El polinomio característico de  $A$  es  $X^2 - X - 1 = (X - \frac{1+\sqrt{5}}{2})(X - \frac{1-\sqrt{5}}{2})$ . Por el Corolario, 6.2.8,  $A$  es diagonalizable. Si ponemos  $\alpha = \frac{1+\sqrt{5}}{2}$  y  $\beta = \frac{1-\sqrt{5}}{2}$  entonces se comprueban fácilmente las relaciones:

$$\alpha + \beta = 1, \quad \alpha - \beta = \sqrt{5}, \quad \alpha\beta = -1, \quad \alpha^2 = \alpha + 1, \quad \beta^2 = \beta + 1. \quad (6.1)$$

Si

$$U = \begin{pmatrix} 1 & 1 \\ \alpha & \beta \end{pmatrix}$$

entonces

$$U^{-1}AU = \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix} = J.$$

Luego

$$\begin{aligned} A^n &= UJ^nU^{-1} = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 & 1 \\ \alpha & \beta \end{pmatrix} \begin{pmatrix} \alpha^n & 0 \\ 0 & \beta^n \end{pmatrix} \begin{pmatrix} -\beta & 1 \\ \alpha & -1 \end{pmatrix} \\ &= \frac{1}{\sqrt{5}} \begin{pmatrix} \alpha^{n-1} - \beta^{n-1} & \alpha^n - \beta^n \\ \alpha^n - \beta^n & \alpha^{n+1} - \beta^{n+1} \end{pmatrix} \end{aligned}$$

Utilizando las relaciones (6.1) se deduce fácilmente que

$$\begin{aligned} \begin{pmatrix} a_{n-1} \\ a_n \end{pmatrix} &= A^{n-2} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ &= \frac{1}{\sqrt{5}} \begin{pmatrix} \alpha^{n-3} - \beta^{n-3} & \alpha^{n-2} - \beta^{n-2} \\ \alpha^{n-2} - \beta^{n-2} & \alpha^{n-1} - \beta^{n-1} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \end{aligned}$$

y, por tanto, volviendo a utilizar (6.1) tenemos

$$a_n = \frac{1}{\sqrt{5}}(\alpha^{n-2} - \beta^{n-2} + \alpha^{n-1} - \beta^{n-1}) = \frac{1}{\sqrt{5}}(\alpha^n - \beta^n).$$

## 6.4 Aplicación 2: Ecuaciones diferenciales

Un sistema lineal de ecuaciones diferenciales de primer orden es un sistema de ecuaciones lineales de la forma

$$\begin{aligned} X'_1 &= a_{11}X_1 + a_{12}X_2 + \cdots + a_{1m}X_m \\ X'_2 &= a_{21}X_1 + a_{22}X_2 + \cdots + a_{2m}X_m \\ &\vdots \\ X'_m &= a_{m1}X_1 + a_{m2}X_2 + \cdots + a_{mm}X_m \end{aligned} \quad (6.2)$$

donde los  $a_{ij}$  y los  $b_i$  son números reales. Es decir un sistema de lineal de ecuaciones diferenciales tiene la forma

$$X' = AX$$

donde  $A$  es una matrix y

$$X = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_m \end{pmatrix}$$

es un vector de indeterminadas.

Una solución de (6.2) es una lista de funciones  $f_1, f_2, \dots, f_n : \mathbb{R} \rightarrow \mathbb{R}$  que satisfacen las ecuaciones.

Por ejemplo, si  $n = 1$ , tenemos una única ecuación

$$X' = aX$$

y su solución es

$$X(t) = Ce^{at},$$

donde  $C$  es un número real arbitrario. Si tenemos un sistema lineal de ecuaciones de la forma

$$\begin{aligned} X_1' &= b_1 X_1 \\ X_2' &= b_2 X_2 \\ &\vdots \\ X_m' &= b_m X_m \end{aligned} \tag{6.3}$$

entonces la solución es

$$\begin{aligned} X_1(t) &= C_1 e^{b_1 t} \\ X_2(t) &= C_2 e^{b_2 t} \\ &\vdots \\ X_m(t) &= C_m e^{b_m t} \end{aligned}$$

donde  $C_1, C_2, \dots, C_m$  son números reales arbitrarios. Es decir un sistema lineal de ecuaciones diferenciales con matriz de coeficientes diagonal es fácil de resolver.

Si realizamos un cambio de variable lineal en (6.2), es decir ponemos

$$\begin{aligned} X_1 &= u_{11}Y_1 + u_{12}Y_2 + \dots + u_{1m}Y_m \\ X_2 &= u_{21}Y_1 + u_{22}Y_2 + \dots + u_{2m}Y_m \\ &\vdots \\ X_m &= u_{m1}Y_1 + u_{m2}Y_2 + \dots + u_{mm}Y_m \end{aligned}$$

o en forma compacta

$$X = UY$$

donde  $U$  es una matriz invertible, entonces el sistema toma la forma

$$UY' = AU$$

o lo que es lo mismo

$$Y' = U^{-1}AU Y.$$

Si  $U$  diagonaliza  $A$ , entonces el sistema se habrá convertido en uno del tipo (6.3) que ya sabemos resolver.

**Ejemplo 6.4.1** Consideremos el siguiente sistema de ecuaciones diferenciales:

$$\begin{aligned} X_1' &= 7X_1 - 3X_2 - 7X_3 \\ X_2' &= 2X_1 \quad \quad - 2X_3 \\ X_3' &= 4X_1 - 2X_2 - 4X_3 \end{aligned}$$

La matriz de los coeficientes es

$$A = \begin{pmatrix} 7 & -3 & -7 \\ 2 & 0 & -2 \\ 4 & -2 & -4 \end{pmatrix}$$

cuyo polinomio característico es

$$\chi_A = X^3 - 3X^2 + 2X = X(X-1)(X-2).$$

Luego los valores propios de  $A$  son 0, 1 y 2 y tenemos una matriz

$$U = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 2 & 1 \\ 1 & 0 & 1 \end{pmatrix}$$

que diagonaliza  $A$ . Poniendo  $X = UY$  convertimos el sistema lineal de ecuaciones diferenciales en el siguiente

$$\begin{aligned} Y_1' &= 0 \\ Y_2' &= Y_2 \\ Y_3' &= 2Y_3 \end{aligned}$$

cuya solución general es

$$\begin{aligned} Y_1(t) &= C_1 \\ Y_2(t) &= C_2 e^t \\ Y_3(t) &= C_3 e^{2t}. \end{aligned}$$

Si queremos calcular la solución general del sistema original ponemos

$$\begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} = U \begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \end{pmatrix}$$

es decir

$$\begin{pmatrix} X_1(t) \\ X_2(t) \\ X_3(t) \end{pmatrix} = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 2 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} C_1 \\ C_2 e^t \\ C_3 e^{2t} \end{pmatrix}$$

o sea

$$\begin{aligned} X_1 &= C_1 + C_2 e^t + 2C_3 e^{2t} \\ X_2 &= \quad \quad 2C_2 e^t + C_3 e^{2t} \\ X_3 &= C_1 \quad \quad \quad + C_3 e^{2t}. \end{aligned}$$

Una *ecuación diferencial lineal homogénea* de grado  $n$  es una ecuación del tipo

$$X^{(n)} + a_{n-1}X^{(n-1)} + a_{n-2}X^{(n-2)} + \cdots + a_2X'' + a_1X' + a_0X = 0 \quad (6.4)$$

donde  $X^{(i)}$  representa la derivada  $i$ -ésima de  $X$ . Podemos resolver ecuaciones diferenciales de grado  $n$  convirtiéndolas en sistemas lineales de ecuaciones diferenciales lineales poniendo

$$\begin{aligned} X_0 &= X \\ X_1 &= X' \\ X_2 &= X'' \\ &\vdots \\ X_n &= X^{(n)} \end{aligned}$$

de forma que la ecuación (6.4) se convierte en el siguiente sistema lineal de ecuaciones diferenciales

$$\begin{aligned} X_0' &= && X_1 \\ X_1' &= && X_2 \\ X_2' &= && X_3 \\ &\vdots && \ddots \\ X_{n-1}' &= -a_0X_0 & -a_1X_1 & -a_2X_2 & \cdots & -a_{n-2}X_{n-2} & -a_{n-1}X_{n-1} \end{aligned}$$

**Ejemplo 6.4.2** Consideremos la ecuación diferencial de grado 2

$$X'' - 5X' + 4X = 0.$$

Ponemos

$$X_0 = X, \quad X_1 = X'$$

con lo que obtenemos el siguiente sistema lineal de ecuaciones diferenciales de primer orden:

$$\begin{aligned} X_0' &= && X_1 \\ X_1' &= -4X_0 & +5X_1 \end{aligned}$$

La matriz de los coeficientes es

$$A = \begin{pmatrix} 0 & 1 \\ -4 & 5 \end{pmatrix}$$

cuyo polinomio característico es

$$\chi_A = X^2 - 5X + 4 = (X - 1)(X - 4).$$

La matriz

$$U = \begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix}$$

diagonaliza  $A$  y si ponemos  $X = UY$ , entonces sistema original se convierte en

$$\begin{aligned} Y_0' &= Y_0 \\ Y_1' &= 4Y_1 \end{aligned}$$

con lo que la solución es

$$\begin{aligned} Y_0(t) &= C_0 e^t \\ Y_1(t) &= C_1 e^{4t}. \end{aligned}$$

Entonces

$$\begin{pmatrix} X_0(t) \\ X_1(t) \end{pmatrix} = U \begin{pmatrix} Y_0 \\ Y_1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} C_0 e^t \\ C_1 e^{4t} \end{pmatrix} = \begin{pmatrix} C_0 e^t + C_1 e^{4t} \\ D_0 e^t + 4D_1 e^{4t} \end{pmatrix}.$$

Por tanto, la solución general de la ecuación diferencial original es

$$X(t) = X_0(t) = C_0 e^t + C_1 e^{4t},$$

donde  $C_0$  y  $C_1$  son números reales arbitrarios, o en otras palabras,  $\{e^t, e^{4t}\}$  es una base del conjunto de soluciones.

## 6.5 Formas de Jordan

En la Sección 6.2 hemos visto que hay dos razones que impiden que un endomorfismo sea diagonalizable. Una es que el polinomio característico no sea completamente factorizable y otra que la suma de los espacios de vectores propios de los diferentes valores propios no alcance la dimensión del espacio. Vamos a ver en esta sección que si no falla la primera condición necesaria, es decir, si el polinomio característico es completamente factorizable, entonces se puede encontrar una matriz asociada a  $f$  que si bien no es tan sencilla como una diagonal, es suficientemente sencilla.

El Ejemplo (6.2.3) nos va a servir para mostrar que es lo peor que puede pasar. En dicho ejemplo partíamos de la matriz

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

pero podríamos haber partido de cualquier matriz de la forma

$$A = \begin{pmatrix} a & 0 \\ 1 & a \end{pmatrix}$$

pues el polinomio característico de  $A$  es  $(X - a)^2$ , con lo que  $a$  es el único valor propio y el rango de  $aI - A$  es 1, con lo que  $\dim N(A, a) = 1$ . Más generalmente podríamos haber partido de una matriz del siguiente tipo que se llama *matriz elemental de Jordan con valor propio  $a$  y orden  $n$* <sup>1</sup>

$$J_n(a) = \begin{pmatrix} a & & & & \\ 1 & a & & & \\ & 1 & a & & \\ & & \ddots & \ddots & \\ & & & 1 & a \end{pmatrix} \in M_n(K).$$

---

<sup>1</sup>En multitud de textos los unos de las matrices elementales de Jordan se sitúan sobre la diagonal en lugar de bajo la diagonal. El tratamiento en ambos casos es esencialmente similar. La elección es una cuestión de gustos.

Una *matriz de Jordan* es una matriz formada por varias matrices de Jordan elementales situadas en bloques situados diagonalmente:

$$\begin{pmatrix} J_{k_1}(a_1) & & & \\ & J_{k_2}(a_2) & & \\ & & \ddots & \\ & & & J_{k_r}(a_r) \end{pmatrix}.$$

Observa que una matriz es diagonal es una matriz de Jordan formada con bloques de orden 1.

El teorema fundamental de esta sección es el siguiente:

**Teorema 6.5.1 (Teorema de Jordan)** *Sea  $f$  un endomorfismo de un espacio vectorial  $V$  de dimensión finita. Si el polinomio característico de  $f$  es completamente factorizable entonces  $V$  tiene una base  $B$  para la que  $f_B$  es una matriz de Jordan.*

La misma filosofía usada hasta ahora para pasar de endomorfismos a matrices cuadradas nos dice que el Teorema 6.5.1 es equivalente al siguiente:

**Teorema 6.5.2** *Si  $A$  es una matriz cuadrada cuyo polinomio característico es completamente factorizable entonces  $A$  semejante a una matriz de Jordan.*

Las matrices de Jordan a las que se refieren los Teoremas 6.5.1 y 6.5.2 se llaman *formas de Jordan* del endomorfismo o la matriz respectivamente.

En el resto de la sección vamos a explicar cómo conseguir para un endomorfismo  $f$  la base  $B$  para la que  $f_B$  sea una matriz de Jordan. Para el caso en que  $A$  sea una matriz cuadrada el mismo procedimiento es el que se usa para encontrar una matriz  $U$  para la que  $U^{-1}AU$  es una matriz de Jordan, o sea  $U$  estará formada poniendo en columnas los elementos de  $B$  para  $f = f_A$ , el endomorfismo dado por  $f_A(x) = Ax$ . De nuevo  $U$  se suele llamar *matriz de paso*.

Para comprender cómo tiene que ser la base  $B = \{v_1, \dots, v_n\}$ , vamos a suponer primero que  $f_B$  es una matriz elemental de Jordan, o sea

$$f_B = J_n(a) = \begin{pmatrix} a & & & & \\ 1 & a & & & \\ & 1 & a & & \\ & & \ddots & \ddots & \\ & & & 1 & a \\ & & & & 1 & a \end{pmatrix}.$$

En tal caso tenemos

$$f(v_1) = av_1 + v_2, \quad f(v_2) = av_2 + v_3, \quad \dots, \quad f(v_{n-1}) = av_{n-1} + v_n, \quad f(v_n) = av_n.$$

La última condición quiere decir que  $v_n$  es un vector propio de  $f$  con valor propio  $a$ . De hecho  $a$  es el único valor propio pues

$$\chi_{J_n(a)} = (X - a)^n$$

Las otras condiciones significan lo siguiente:

$$v_2 = f(v_1) - av_1, \quad v_3 = f(v_2) - av_2, \quad \dots, \quad v_n = f(v_{n-1}) - av_{n-1},$$

o lo que es lo mismo

$$v_2 = (f - aI)(v_1), \quad v_3 = (f - aI)(v_2), \quad \dots, \quad v_n = (f - aI)(v_{n-1}).$$

O sea el primer elemento  $v_1$  determina todos los demás:

$$v_{i+1} = (f - aI)^i(v_1), \quad i = 1, \dots, n - 1.$$

Pero además queremos que  $v_1, \dots, v_n$  sean linealmente independientes, y en particular distintos de cero, con lo que  $v_n = (f - aI)^{n-1}(v_1) \neq 0$  pero como  $v_n$  es valor propio de valor propio  $n$  tenemos que

$$(f - aI)^n(v_1) = (f - aI)(v_n) = 0.$$

O sea

$$v_1 \in \ker(f - aI)^n \setminus \ker(f - aI)^{n-1}.$$

Por tanto si queremos encontrar una base  $B$  para la que  $f_B$  sea una matriz elemental de Jordan necesitamos encontrar un elemento de  $\ker(f - a)^n \setminus \ker(f - a)^{n-1}$ . En el caso de que partamos de una matriz eso habrá que sustituirlo por encontrar un elemento de  $N((A - I)^n) \setminus N((A - I)^{n-1})$ .

**Ejemplo 6.5.3** Consideremos la matriz

$$A = \begin{pmatrix} 5 & 7 & -3 \\ 0 & 0 & 1 \\ 1 & -4 & 4 \end{pmatrix}.$$

Calculamos el polinomio característico:

$$\chi_A = \begin{vmatrix} X - 5 & -7 & 3 \\ 0 & X & -1 \\ -1 & 4 & -4 \end{vmatrix} = (X - 3)^3.$$

O sea 3 es el único valor propio. Calculamos las primeras potencias de  $A - 3I$ :

$$A - 3I = \begin{pmatrix} 2 & 7 & -3 \\ 0 & -3 & 1 \\ 1 & -4 & 1 \end{pmatrix}, \quad (A - 3I)^2 = \begin{pmatrix} 1 & 5 & -2 \\ 1 & 5 & -2 \\ 3 & 15 & -6 \end{pmatrix}, \quad (A - 3I)^3 = 0.$$

Por tanto necesitamos calcular un elemento que no esté en el espacio nulo de  $(A - 3I)^2$ , por ejemplo  $v_1 = (1, 0, 0)$ , y calcular

$$v_2 = (A - 3I)v_1 = (2, 0, 1), \quad v_3 = (A - 3I)^2 v_1 = (1, 1, 3).$$

Por tanto la matriz

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

satisface

$$U^{-1}AU = \begin{pmatrix} 3 & & \\ 1 & 3 & \\ & 1 & 3 \end{pmatrix} = J_3(3).$$

Podemos hacernos una idea de por qué pasa lo que pasa calculando las potencias sucesivas de la matriz

$$J_n(a) - aI = J_n(0)$$

Por ejemplo calculemos las potencias de  $J_4(0)$ :

$$J_4(0) = \begin{pmatrix} 0 & & & \\ 1 & 0 & & \\ & 1 & 0 & \\ & & 1 & 0 \end{pmatrix}, \quad J_4(0)^2 = \begin{pmatrix} 0 & & & \\ 0 & 0 & & \\ 1 & 0 & 0 & \\ 0 & 1 & 0 & 0 \end{pmatrix},$$

$$J_4(0)^3 = \begin{pmatrix} 0 & & & \\ 0 & 0 & & \\ 0 & 0 & 0 & \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad J_4(0)^4 = 0.$$

Si ahora tomamos una matriz de Jordan formada por dos matrices elementales de Jordan con el mismo autovalor la cosa es ligeramente diferente: Por ejemplo si

$$A = \left( \begin{array}{c|c} J_3(a) & \\ \hline & J_1(a) \end{array} \right) = \left( \begin{array}{ccc|c} a & & & \\ 1 & a & & \\ & 1 & a & \\ \hline & & & a \end{array} \right)$$

Entonces

$$A - aI = \left( \begin{array}{c|c} J_3(0) & \\ \hline & J_1(0) \end{array} \right) = \left( \begin{array}{ccc|c} 0 & & & \\ 1 & 0 & & \\ & 1 & 0 & \\ \hline & & & 0 \end{array} \right)$$

y

$$(A - aI)^2 = \left( \begin{array}{c|c} J_3(0)^2 & \\ \hline & J_1(0)^2 \end{array} \right) = \left( \begin{array}{ccc|c} 0 & & & \\ 0 & 0 & & \\ 1 & 0 & 0 & \\ \hline & & & 0 \end{array} \right), \quad (A - aI)^3 = 0.$$

Si la base  $B = \{u_1, \dots, u_{k_1}, v_1, v_2, \dots, v_{k_2}\}$  cumple

$$f_B = \left( \begin{array}{c|c} J_{k_1}(a) & \\ \hline & J_{k_2}(a) \end{array} \right)$$

entonces tenemos

$$f(u_1) = au_1 + u_2, \quad f(u_2) = au_2 + u_3, \quad \dots, \quad f(u_{k_1-1}) = au_{k_1-1} + u_{k_1}, \quad f(u_{k_1}) = u_{k_1}$$

y

$$f(v_1) = av_1 + v_2, \quad f(v_2) = av_2 + v_3, \quad \dots, \quad f(v_{k_2-1}) = av_{k_2-1} + v_{k_2}, \quad f(v_{k_2}) = v_{k_2}.$$

O sea  $u_1 \in \ker(f - aI)^{k_1} \setminus \ker(f - aI)^{k_1-1}$  y  $u_2 \in \ker(f - aI)^{k_2} \setminus \ker(f - aI)^{k_2-1}$ . Veamos con ejemplos dos situaciones diferentes que pueden pasar.

**Ejemplo 6.5.4** Consideremos las dos siguientes matrices

$$A = \begin{pmatrix} -5 & 1 & -1 & 0 \\ -9 & 1 & -2 & 1 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & -2 \end{pmatrix} \quad y \quad B = \begin{pmatrix} -2 & 0 & 0 & 0 \\ 1 & -2 & 1 & 1 \\ 1 & 0 & -2 & 0 \\ -1 & 0 & 0 & -2 \end{pmatrix}$$

Ambas tienen el mismo polinomio característico:

$$\chi_A = \begin{vmatrix} X+5 & -1 & 1 & 0 \\ 9 & X-1 & 2 & -1 \\ 0 & 0 & X+2 & 0 \\ 0 & 0 & 0 & X+2 \end{vmatrix} = (X+2)^4,$$

$$\chi_B = \begin{vmatrix} X+2 & 0 & 0 & 0 \\ -1 & X+2 & -1 & -1 \\ -1 & 0 & X+2 & 0 \\ 1 & 0 & 0 & X+2 \end{vmatrix} = (X+2)^4.$$

Calculamos las potencias de  $(A + 2I)$  y  $(B + 2I)$ :

$$A + 2I = \begin{pmatrix} -3 & 1 & -1 & 0 \\ -9 & 3 & -2 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad (A + 2I)^2 = \begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad (A + 2I)^3 = 0$$

$$B + 2I = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}, \quad (B + 2I)^2 = 0.$$

Si partimos de un elemento  $u_1 = (0, 0, 1, 0) \in N(A + 2I)^3 \setminus N(A + 2I)^2$  podemos construir

$$u_2 = (A + 2I)u_1 = (-1, -2, 0, 0), \quad u_3 = (A + 2I)^2 u_1 = (1, 3, 0, 0)$$

Pero eso solo nos da tres vectores que nos van a proporcionar un bloque  $J_3(-2)$ . Eso implica que falta un cuarto vector que necesariamente tendrá que ser un vector propio. Obsérvese

que  $u_3$  un vector propio de  $A$  con valor propio  $-2$  y que  $N(A + 2I)$  tiene dimensión 2. Por tanto necesitamos completar  $u_3$  a una base de  $N(A + 2I)$ , por ejemplo añadiendo el vector  $u_4 = (0, 1, 1, -1)$ . Poniendo los vectores  $u_1, u_2, u_3, u_4$  como columnas de una matriz

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

Obtenemos

$$U^{-1}AU = \left( \begin{array}{cc|cc} -2 & & & \\ 1 & -2 & & \\ & 1 & -2 & \\ \hline & & & -2 \end{array} \right) = \left( \begin{array}{c|c} J_3(-2) & \\ \hline & J_1(-2) \end{array} \right).$$

Sin embargo la matriz  $B$  necesita un tratamiento diferente pues cada elemento  $u \in N((B + 2I)^2) \setminus N(B + 2I)$  nos proporcionará un vector propio  $v = (B + 2I)u$  lo que no nos llega para obtener una base. El truco es tomar dos elementos linealmente independientes  $u_1, u_2$  para los que  $\langle u_1, u_2 \rangle \cap N(B + 2I) = 0$ . Eso implicará que  $K^4 = \langle u_1, u_2 \rangle \oplus N(B + 2)$ , o sea  $\langle u_1, u_2 \rangle$  es un complemento de  $N(B + 2)$ . Para ello calculamos

$$N(B + 2I) \equiv x_1 = x_3 + x_4 = 0$$

y observamos que  $W = \{(x_1, x_2, x_3, x_4) : x_2 = x_3 = 0\} = \langle (1, 0, 0, 0), (0, 0, 1, 0) \rangle$  es un subespacio vectorial de dimensión 2 con  $W \cap N(B + 2I) = 0$ . Tomando

$$\begin{aligned} v_1 = u_1 &= (1, 0, 0, 0), & v_2 &= (B + 2I)u_1 = (0, 1, 1, -1), \\ v_3 = u_2 &= (0, 0, 1, 0), & v_4 &= (B + 2I)v_3 = (0, 1, 0, 0) \end{aligned}$$

tenemos la base que proporciona la forma de Jordan. O sea poniendo

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

tenemos

$$U^{-1}BU = \left( \begin{array}{cc|cc} -2 & & & \\ 1 & -2 & & \\ & & -2 & \\ \hline & & 1 & -2 \end{array} \right) = \left( \begin{array}{c|c} J_2(-2) & \\ \hline & J_2(-2) \end{array} \right).$$

La Figura 6.1 muestra cuál es el proceso de construir la base  $B$  para la que  $f_B$  es una matriz de Jordan para el caso en el que el endomorfismo o la matriz solo tenga un valor propio. En la Figura  $A_\lambda = A - \lambda I$ .

La idea consiste en construir los vectores linealmente independientes iniciales  $u_1, u_2, \dots$ , de forma que si  $k$  es el menor entero positivo para el que  $N(A_\lambda^k) = N(A_\lambda^{k+1})$  (en la figura  $k = 4$ ) entonces:

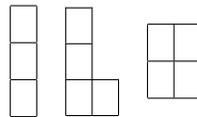
$N((A_\lambda)^4)$	$u_1$	$u_2$	$u_3$					
$N((A_\lambda)^3)$	$A_\lambda u_1$	$A_\lambda u_2$	$A_\lambda u_3$					
$N((A_\lambda)^2)$	$A_\lambda^2 u_1$	$A_\lambda^2 u_2$	$A_\lambda^2 u_3$	$u_4$	$u_5$			
$N(A_\lambda)$	$A_\lambda^3 u_1$	$A_\lambda^3 u_2$	$A_\lambda^2 u_3$	$A_\lambda u_4$	$A_\lambda u_5$	$u_6$	$u_7$	$u_8$

$$\begin{aligned}
 N(A_\lambda^4) &= N(A_\lambda^3) \oplus W_4 \\
 W_4 &= \langle u_1, u_2, u_3 \rangle, \\
 N(A_\lambda^3) &= N(A_\lambda^2) \oplus W_3 \\
 W_3 &= \langle A_\lambda(u_1), A_\lambda(u_2), A_\lambda(u_3) \rangle, \\
 N(A_\lambda^2) &= N(A_\lambda) \oplus W_2, \\
 W_2 &= \langle A_\lambda^2(u_1), A_\lambda^2(u_2), A_\lambda^2(u_3), u_4, u_5 \rangle, \\
 N(A_\lambda) &= \langle A_\lambda^3(u_1), A_\lambda^3(u_2), A_\lambda^3(u_3), A_\lambda^2(u_4), A_\lambda(u_5), u_6, u_7, u_8 \rangle.
 \end{aligned}$$

Figure 6.1: Esquema del cálculo de una base  $N(A_\lambda^k)$  con  $A_\lambda = A - \lambda I$ .

- Los elementos de la primera fila ( $u_1, u_2, u_3$  en la Figura 6.1) forman una base de un complemento  $W_k$  de  $N(A_\lambda^{k-1})$  en  $N(A_\lambda^k)$ .
- Formamos la segunda fila aplicando  $A_\lambda$  a los elementos de la fila superior y les añadimos los siguientes  $u_i$  (en la Figura 6.1,  $u_4$ ) de forma que entre todos ellos forman una base de un complemento  $W_{k-1}$  de  $N(A_\lambda^{k-2})$  en  $N(A_\lambda^{k-1})$ .
- Repetimos el procedimiento  $k$  veces, es decir aplicamos  $A_\lambda$  a los vectores de la fila anterior y les añadimos los siguientes  $u_i$  para que todos los de esa fila formen una base de un complemento  $W_r$  de  $N(A_\lambda^{r-1})$  en  $N(A_\lambda^r)$ .
- En el último paso  $N(A_\lambda^0) = 0$ , con lo que los últimos  $u_i$  se obtienen añadiendo a los que resultan de aplicar  $A_\lambda$  a los de la fila anterior hasta completar una base del espacio  $N(A_\lambda)$  de vectores propios de valor propio  $\lambda$ .

Las formas que toma el esquema de la Figura 6.1 en las tres matrices de los Ejemplos 6.5.3 y 6.5.4 son las siguientes:



En el caso en el que haya más de un valor propio hay que realizar el mismo procedimiento para cada uno de ellos.

Vamos a poner esto en forma de algoritmo:

**Algoritmo para calcular la forma de Jordan de una matriz cuadrada  $A \in M_n(K)$ .**

- (1) Calculamos los valores propios  $\lambda_1, \dots, \lambda_k$  y sus multiplicidades respectivas  $m_1, \dots, m_k$  y comprobamos que  $m_1 + \dots + m_k = n$  (en caso contrario la matriz no tendrá forma de Jordan).

- (2) Para cada valor propio  $\lambda = \lambda_i$  con multiplicidad  $m$  calculamos  $A_\lambda = A - \lambda I$  y vamos calculando los rangos de las sucesivas potencias de  $A_\lambda$  hasta que obtengamos un rango igual a  $n - m$ . Pongamos que estos rangos son

$$r(A_\lambda) > r(A_\lambda^2) > r(A_\lambda^t) = n - m.$$

- (3) Calculamos una base de  $N(A_\lambda^{t-1})$  y la completamos a una base de  $N(A_\lambda^t)$ . Sean  $u_1, \dots, u_k$  los vectores que hemos añadido para pasar de una base de  $N(A_\lambda^{t-1})$  a una base de  $N(A_\lambda^t)$ .

Para fijar ideas a partir de ahora suponemos que vamos a obtener un esquema como en la Figura 6.1. En tal caso tendremos que  $t = 4$  y las dimensiones de los espacios nulos son

$$\begin{aligned} \dim N(A_\lambda) &= 8, \\ \dim N(A_\lambda^2) &= 8 + 5; \\ \dim N(A_\lambda^3) &= 8 + 5 + 3; \\ \dim N(A_\lambda^4) &= 8 + 5 + 3 + 3; \end{aligned}$$

Por tanto  $\dim N(A_\lambda^4) - \dim N(A_\lambda^3) = 3$  y por tanto en nuestro caso  $k = 3$  de forma que  $u_1, u_2, u_3$  son los tres vectores que aparecen en el piso superior de la Figura 6.1.

- (4) Aplicamos  $A_\lambda$  a los vectores de  $B$  con lo que tendremos otros tantos elementos  $A_\lambda(u_1), A_\lambda(u_2), A_\lambda(u_3)$ , que estarán en  $N(A_\lambda^{t-1})$  en este caso en  $N(A_\lambda^3)$ . Añadimos estos vectores a  $B$ . En este caso ya tenemos

$$B = \{u_1, u_2, u_3, A_\lambda(u_1), A_\lambda(u_2), A_\lambda(u_3)\}.$$

Por otro lado los vectores calculados  $A_\lambda(u_1), A_\lambda(u_2), A_\lambda(u_3)$  son linealmente independientes.

- (5) Ahora necesitamos calcular una base de un complemento de  $N(A_\lambda^{t-2})$  en  $N(A_\lambda^{t-1})$  que contenga  $A_\lambda(u_1), A_\lambda(u_2), A_\lambda(u_3)$ .

En este caso  $\dim N(A_\lambda^3) - \dim N(A_\lambda^2) = 3$  con lo que ya tenemos la base  $A_\lambda(u_1), A_\lambda(u_2), A_\lambda(u_3)$  del complemento de  $N(A_\lambda^2)$  en  $N(A_\lambda^3)$  que buscábamos.

- (6) Repetimos el proceso anterior. Es decir a todos los vectores que ya estaban en  $B$  les volvemos a aplicar  $A_\lambda$  y se los añadimos a  $B$ . Algunos de ellos ya estaban. En este caso solo habrá que aplicarlo a  $A_\lambda(u_1), A_\lambda(u_2)$  y  $A_\lambda(u_3)$ . De forma que ahora tenemos

$$B = \{u_1, u_2, u_3, A_\lambda(u_1), A_\lambda(u_2), A_\lambda(u_3), A_\lambda^2(u_1), A_\lambda^2(u_2), A_\lambda^2(u_3)\}.$$

Los vectores añadidos, en este caso  $A_\lambda^2(u_1), A_\lambda^2(u_2), A_\lambda^2(u_3)$ , son linealmente independientes, están en un complemento de  $N(A_\lambda^2)$  en  $N(A_\lambda)$ , y tenemos que conseguir un complemento de  $N(A_\lambda^2)$  en  $N(A_\lambda)$  añadiéndole más vectores a estos. En este caso

$$\dim N(A_\lambda^2) - \dim N(A_\lambda) = 5$$

con lo que hay que encontrar dos vectores para completar la base del complemento. Una forma de hacerlo es primero encontrar una base  $B_2$  de  $N(A_\lambda)^2$  que contenga dichos vectores (en este caso tendrá 13 vectores), buscar una  $B_1$  base de  $N(A_\lambda)$  y buscar dos vectores  $u_4, u_5$  en  $B_2$  de forma  $B_1 \cap \{A_\lambda^2(u_1), A_\lambda^2(u_2), A_\lambda^2(u_3)\}$  sean linealmente independientes (y por tanto una base de  $N(A_\lambda^2)$ ). Por tanto añadimos estos dos vectores a  $B$  y ya tenemos

$$B = \{u_1, u_2, u_3, A_\lambda(u_1), A_\lambda(u_2), A_\lambda(u_3), A_\lambda^2(u_1), A_\lambda^2(u_2), A_\lambda^2(u_3), u_4, u_5\}.$$

- (7) Repetimos una vez más el proceso de forma que ahora  $B$  habrá crecido en cinco vectores más

$$B = \left\{ u_1, u_2, u_3, A_\lambda(u_1), A_\lambda(u_2), A_\lambda(u_3), A_\lambda^2(u_1), A_\lambda^2(u_2), A_\lambda^2(u_3), u_4, u_5, \right. \\ \left. A_\lambda^3(u_1), A_\lambda^3(u_2), A_\lambda^3(u_3), A_\lambda(u_4), A_\lambda(u_5) \right\}.$$

- (8) Los últimos cinco vectores añadidos  $A_\lambda^3(u_1), A_\lambda^3(u_2), A_\lambda^3(u_3), A_\lambda(u_4), A_\lambda(u_5)$  son linealmente independientes y están en  $N(A_\lambda)$ . Solo falta completarlos a una base de  $N(A_\lambda)$  para lo cual hay que añadir tres vectores más que los podremos encontrar entre la base  $B_1$  que calculamos dos pasos atrás.

El siguiente esquema muestra en rojo el número de elementos que hay que añadir en cada paso

$$\begin{aligned} \dim N(A_\lambda^4) &= \dim N(A_\lambda^3) + 3, \\ \dim N(A_\lambda^3) &= \dim N(A_\lambda^2) + 3, \\ \dim N(A_\lambda^2) &= \dim N(A_\lambda) + 3 + 3 + 2, \\ \dim N(A_\lambda) &= 3 + 3 + 2 + 3. \end{aligned}$$

**Ejemplo 6.5.5** Consideremos la matriz

$$A = \begin{pmatrix} 1 & -14 & 6 & -2 & 1 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 1 & 2 & 0 & 0 \\ 0 & -3 & 2 & 1 & 0 \\ 0 & 10 & -3 & 1 & 1 \end{pmatrix}.$$

Primero calculamos el polinomio característico:

$$\chi_A = \begin{vmatrix} X-1 & 14 & -6 & 2 & -1 \\ 0 & X-2 & 0 & 0 & 0 \\ 0 & -1 & X-2 & 0 & 0 \\ 0 & 3 & -2 & X-1 & 0 \\ 0 & -10 & 3 & -1 & X-1 \end{vmatrix} = (X-1)^3(X-2)^2.$$

y después calculamos los espacios  $N(A, 1)$  y  $N(A, 2)$  para lo cual tenemos que calcular los espacios nulos de las matrices

$$A - I = \begin{pmatrix} 0 & -14 & 6 & -2 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & -3 & 2 & 0 & 0 \\ 0 & 10 & -3 & 1 & 0 \end{pmatrix} \quad \text{y} \quad A - 2I = \begin{pmatrix} -1 & -14 & 6 & -2 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & -3 & 2 & -1 & 0 \\ 0 & 10 & -3 & 1 & -1 \end{pmatrix}.$$

Aplicando el Método de Gauss a estas dos matrices obtenemos las siguientes formas fuertemente escalonadas

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Como se ve las matrices  $I - A$  y  $A - 2I$  tienen rango 4 con lo que sus espacios nulos tienen dimensión 1 y de esta forma solo vamos a encontrar dos vectores propios linealmente independientes, uno con valor propio 1 y otro con valor propio 2.

Las cuentas que hemos hecho nos sirven para calcular ecuaciones de  $N(A - I)$  y  $N(A - 2I)$ :

$$N(A - I) \equiv x_2 = x_3 = x_4 = x_5 = 0$$

$$N(A - 2I) \equiv x_1 = x_3 = -x_5, x_2 = 0, x_4 = 2x_5. \quad (6.5)$$

Lo que vamos a hacer ahora es calcular las potencias de  $(I - A)^2$  y  $(A - 2I)^2$ :

$$(A - I)^2 = \begin{pmatrix} 0 & 8 & -1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 \\ 0 & -1 & 2 & 0 & 0 \\ 0 & 4 & -1 & 0 & 0 \end{pmatrix} \quad (A - 2I)^2 = \begin{pmatrix} 1 & 36 & -13 & 5 & -2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 5 & -2 & 1 & 0 \\ 0 & -16 & 5 & -2 & 1 \end{pmatrix}$$

Aplicando el Método de Gauss a estas dos nuevas matrices, de una forma sui géneris con la segunda para simplificar los cálculos, obtenemos las matrices:

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} 1 & -1 & -1 & 0 & 0 \\ 0 & 5 & -2 & 1 & 0 \\ 0 & -6 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Obsérvese que mientras  $A - I$  y  $A - 2I$  tienen rango 4, ahora  $(A - I)^2$  y  $(A - 2I)^2$  tienen rango 3 con lo que sus espacios nulos tienen rango 2 y vienen dados por

$$N((A - I)^2) \equiv x_2 = x_3 = x_4 = 0 \quad (6.6)$$

$$N((A - 2I)^2) \equiv x_1 = x_2 + x_3, x_4 = -5x_2 + 2x_3, x_5 = 6x_2 - x_3 \quad (6.7)$$

Antes de continuar parémonos a reflexionar un poco. Obsérvese que si  $B$  es una matriz cuadrada arbitraria, el espacio nulo de  $B$  está contenido en el espacio nulo de  $B^2$  pero el espacio nulo de  $B^2$  no tiene por qué estar contenido en el de  $B$ . Como las dimensiones de  $N(A - I)$  y  $N(A - 2I)$  son menores que las de  $N((A - I)^2)$  y  $N((A - 2I)^2)$ , en este caso las inclusiones  $N(A - I) \subset N((A - I)^2)$  y  $N(A - 2I) \subset N((A - 2I)^2)$  son estrictas. Más generalmente  $N(B^n) \subsetneq N(B^{n+1})$  y la igualdad se verificará si y solo si  $B^n$  y  $B^{n+1}$  tienen el mismo rango.

Vamos a ver lo que pasa con las terceras potencias:

$$(A - I)^3 = \begin{pmatrix} 0 & 4 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 3 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 0 \\ 0 & 3 & -1 & 0 & 0 \end{pmatrix} \quad (A - 2I)^3 = \begin{pmatrix} -1 & -62 & 22 & -9 & 3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & -5 & 2 & -1 & 0 \\ 0 & 21 & -7 & 3 & -1 \end{pmatrix}$$

Obsérvamos ahora que  $(A - I)^3$  tiene rango 2 mientras que  $(A - I)^2$  tenía rango 3. Esto muestra que  $N((A - I)^3)$  contiene propiamente a  $N((A - I)^2)$ . Sin embargo  $(A - 2I)^3$  tiene rango 3 que es el mismo rango que el de  $(A - 2I)^2$ . Por tanto  $N((A - 2I)^2) = N((A - 2I)^3)$ . Eso implica que el esquema para el autovalor 2 tiene dos filas y como la multiplicidad del 2 como raíz del polinomio característico es también 2 solo vamos a poder obtener dos vectores aquí con lo que el esquema tiene la siguiente forma:

$$\begin{array}{|c|} \hline \square \\ \hline \square \\ \hline \end{array}$$

Para completar este esquema necesitamos un elemento de  $N((A - 2I)^2) \setminus N(A - 2I)$ . Comparando (6.5) y (6.7) podemos poner en el primer piso  $v_4 = (1, 1, 0, -5, 6)$  y por tanto tendremos

$$\begin{array}{|l} \hline v_4 = (1, 1, 0, -5, 6) \\ \hline v_5 = (A - 2I)v_4 = (1, 0, 1, 2, -1) \\ \hline \end{array}$$

Para completar el estudio con el autovalor 1 necesitaremos calcular

$$N((A - I)^3) \equiv x_2 = x_3 = 0 \quad (6.8)$$

y

$$(A - I)^4 = \begin{pmatrix} 0 & 5 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 4 & 1 & 0 & 0 \\ 0 & 3 & 2 & 0 & 0 \\ 0 & 2 & -1 & 0 & 0 \end{pmatrix}$$

Como esta matriz tiene rango 2 que es el mismo que el de  $(A - I)^3$  se tiene que  $N((A - I)^4) = N((A - I)^3)$ . Por tanto el esquema para el valor propio 1 tiene tres pisos y necesitamos elegir un

vector de  $N((A - I)^3) \setminus N((A - I)^2)$ . Comparando (6.6) con (6.8) podemos tomar  $(0, 0, 0, 1, 0)$  y completar el esquema para este autovalor con

$v_1 = (0, 0, 0, 1, 0)$
$v_2 = (A - I)v_3 = (-2, 0, 0, 0, 1)$
$v_3 = (A - I)v_2 = (1, 0, 0, 0, 0)$

Ahora ponemos  $v_1, v_2, v_3, v_4, v_5$  como columnas de una matriz:

$$U = (v_1, v_2, v_3, v_4, v_5) = \begin{pmatrix} 0 & -2 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & -5 & 2 \\ 0 & 1 & 0 & 6 & -1 \end{pmatrix}$$

Tenemos que

$$U^{-1}AU = \left( \begin{array}{ccc|cc} 1 & & & & \\ 1 & 1 & & & \\ & 1 & 1 & & \\ \hline & & & 2 & \\ & & & 1 & 2 \end{array} \right) = \left( \begin{array}{c|c} J_3(1) & \\ \hline & J_2(2) \end{array} \right).$$

Antes de cerrar esta sección veamos un último

**Ejemplo 6.5.6** Vamos a buscar la forma de Jordan de la matriz

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

El polinomio característico de  $A$

$$\chi_A = X^4 + 2X^2 + 1 = (X^2 + 1)^2 = (X - i)^2(X + i)^2$$

Por tanto la los valores propios son  $i$  y  $-i$  ambos con multiplicidad 2. Calculamos

$$A - iI = \begin{pmatrix} 1 - i & -2 & 0 & 3 \\ 1 & -3 - i & 2 & 2 \\ 1 & -2 & 1 - i & 1 \\ 0 & -2 & 2 & 1 - i \end{pmatrix} \quad \text{y} \quad A + iI = \begin{pmatrix} 1 + i & -2 & 0 & 3 \\ 1 & -3 + i & 2 & 2 \\ 1 & -2 & 1 + i & 1 \\ 0 & -2 & 2 & 1 + i \end{pmatrix}.$$

Una cuenta rápida muestra que ambas matrices tienen rango 3 con lo que  $\dim N(A - iI) = \dim N(A + iI) = 1$ , y por tanto  $A$  no es diagonalizable. Calculando los cuadrados de estas dos matrices obtenemos

$$(A - iI)^2 = \begin{pmatrix} -2 - 2i & -2 + 4i & 2 & 2 - 6i \\ -2i & -2 + 6i & -4i & 1 - 4i \\ -2i & 4i & -2 - 2i & 1 - 2i \\ 0 & 4i & -4i & -2 - 2i \end{pmatrix}$$

y

$$(A + iI)^2 = \begin{pmatrix} -2 + 2i & -2 - 4i & 2 & 2 + 6i \\ 2i & -2 - 6i & 4i & 1 + 4i \\ 2i & -4i & -2 + 2i & 1 + 2i \\ 0 & -4i & 4i & -2 + 2i \end{pmatrix}$$

Aplicando el Método de Gauss a las cuatro matrices obtenemos

$$\begin{aligned} N(A - iI) &\equiv x_4 = 0, x_2 = x_3, x_1 = (1 + i)x_3, \\ N(A + iI) &\equiv x_4 = 0, x_2 = x_3, x_1 = (1 - i)x_3, \\ N((A - iI)^2) &\equiv 2x_1 - 3(1 - i)x_2 + (1 - 5i)x_3 = 0, x_4 = (1 + i)(x_2 - x_3), \\ N((A + iI)^2) &\equiv 2x_1 - 3(1 + i)x_2 + (1 + 5i)x_3 = 0, x_4 = (1 - i)(x_2 - x_3), \end{aligned}$$

Tomando

$$\begin{aligned} v_1 &= (-3i, 1 - i, 0, 2) \in N(A - iI)^2 \setminus N(A - iI), \\ v_2 &= (A - iI)v_1 = (1 - i, -i, -i, 0), \\ v_3 &= (3i, 1 + i, 0, 2) \in N(A + iI)^2 \setminus N(A + iI), \\ v_4 &= (A + iI)v_3 = (1 + i, i, i, 0). \end{aligned}$$

construimos la matriz

$$U = \begin{pmatrix} -3i & 1 - i & 3i & 1 + i \\ 1 - i & -i & 1 + i & i \\ 0 & -i & 0 & i \\ 2 & 0 & 2 & 0 \end{pmatrix}$$

que cumple

$$U^{-1}AU = \begin{pmatrix} i & & & \\ 1 & i & & \\ & & -i & \\ & & 1 & -i \end{pmatrix}$$

## 6.6 Demostración del Teorema de Jordan

En la sección anterior hemos explicado el proceso algorítmico del cálculo de la forma de Jordan de un endomorfismo o una matriz, pero no hemos justificado por qué funciona. Por ejemplo, no está obvio por qué los vectores que aparecen en el esquema de la Figura 6.1 son linealmente independientes ni por qué al juntar todos los obtenidos con los diferentes valores propios se obtiene una base. Vamos a justificarlo todo en esta sección.

Fijaremos un endomorfismo  $f$  de un espacio vectorial  $V$  de dimensión  $n$  sobre un cuerpo  $K$  y supongamos que su polinomio característico es completamente factorizable. Supongamos que los valores propios son  $\lambda_1, \dots, \lambda_r$  y que sus multiplicidades son  $m_1, \dots, m_r$ , o sea

$$\chi_f = (X - \lambda_1)^{m_1} \dots (X - \lambda_r)^{m_r}.$$

Por tanto

$$n = \dim V = m_1 + \cdots + m_k.$$

Para cada  $i = 1, \dots, k$  sea  $g_i = f - \lambda_i I$ . Obsérvese que

$$0 \neq \ker g_i \subseteq \ker g_i^2 \subseteq \ker g_i^3 \subseteq \dots$$

pero esta sucesión no puede crecer indefinidamente pues las dimensiones están acotadas por  $n$ . Por tanto existirá un  $k$  tal que  $\ker g_i^k = \ker g_i^{k+1}$ . Pero eso va a implicar que  $\ker g_i^h = \ker g_i^k$  para todo  $h > k$ . Eso es consecuencia del siguiente

**Lema 6.6.1** *Si  $g$  es un endomorfismo y  $\ker g^k = \ker g^{k+1}$  entonces  $\ker g^k = \ker g^h$  para todo  $s > k$ .*

**Demostración.** Razonamos por inducción sobre  $s$ . El primer caso  $h = k + 1$  es la hipótesis por lo que no hay que demostrarlo. Supongamos que  $\ker g^k = \ker g^h$  con  $h > r$ . Si  $x \in \ker g^{h+1}$  entonces  $g^{k+1}(g^{h-k})(x) = g^{h+1}(x) = 0$ , con lo que  $g^{h-k}(x) \in \ker g^{k+1} = \ker g^k$ . Luego  $g^h(x) = g^k(g^{h-k}(x)) = 0$  con lo que  $x \in \ker g^h = \ker g^k$ . Esto demuestra que  $\ker g^{h+1} \subseteq \ker g^k$ . Como la otra inclusión siempre se verifica pues  $k < h$ , tenemos que  $\ker g^{h+1} = \ker g^k$ . ■

Para cada  $i = 1, \dots, k$  fijamos el primer  $k_i$  para el que  $\ker g_i^{k_i} = \ker g_i^{k_i+1}$ . Esta será la altura del esquema de la Figura 6.1 para el autovalor  $\lambda_i$ . Obsérvese que en la notación de dicho esquema  $A_{\lambda_i}$  representa el papel de  $g_i$ . Recordemos que lo que hacíamos en dicho esquema era comenzar tomando  $u_1, \dots, u_r$  elementos linealmente independientes  $\ker g_i^{k_i}$  tales que

$$\langle u_1, \dots, u_r \rangle \cap \ker g_i^{k_i-1} = 0 \quad y \quad r = \dim \ker g_i^{k_i} - \dim \ker g_i^{k_i-1}.$$

El siguiente lema justifica por qué al aplicar  $k_i - 1$  potencias sucesivas de  $g_i$  a estos elementos seguimos obteniendo una lista linealmente independiente.

**Lema 6.6.2** *Sea  $g$  un endomorfismo y sea  $u_1, \dots, u_r$  una lista linealmente independiente de elementos de  $\ker g^k$  con  $\langle u_1, \dots, u_r \rangle \cap \ker g^{k-1} = 0$ . Entonces para todo  $i = 0, \dots, k - 1$  se tiene que  $g^i(u_1), \dots, g^i(u_r)$  es una lista linealmente independiente de elementos de  $\ker g^{k-i}$  con  $\langle g^i(u_1), \dots, g^i(u_r) \rangle \cap \ker g^{k-i-1} = 0$ .*

**Demostración.** Como cada  $u_j \in \ker g^k$  se tiene que  $g^i(u_j) \in \ker g^{k-i}$ . Supongamos que  $\sum_{j=1}^r \alpha_j g^i(u_j) = 0$ . Eso implica que  $g^i(\sum_{j=1}^r \alpha_j u_j) = 0$ , o sea  $\sum_{j=1}^r \alpha_j u_j \in \ker g^i \cap \langle u_1, \dots, u_r \rangle \subseteq \ker g^{k-1} \cap \langle u_1, \dots, u_r \rangle = 0$ . Como  $u_1, \dots, u_r$  es linealmente independiente se tiene que todos los  $\alpha_j$  son 0. Esto demuestra que  $g^i(u_1), \dots, g^i(u_r)$  es linealmente independiente.

Sea ahora  $x \in \langle g^i(u_1), \dots, g^i(u_r) \rangle \cap \ker g^{k-i-1}$ . Entonces  $x = \sum_{j=1}^r \alpha_j g^i(u_j)$  para algunos escalares  $\alpha_j$  y

$$0 = g^{k-i-1}(x) = \sum_{j=1}^r \alpha_j g^{k-1}(u_j).$$

Pero como  $g^{k-1}(u_1), \dots, g^{k-1}(u_r)$  son linealmente independientes tenemos que todos los  $\alpha_j = 0$  y por tanto  $x = 0$ . ■

El Lema 6.6.2, junto con la Fórmula de Grassman, implica que se verifica

$$\ker g_i^{k_i} = \ker g_i^{k_i-1} \oplus \langle u_1, \dots, u_r \rangle.$$

O sea los vectores de la primera fila de la Figura 6.1 forman una base de un complemento de  $\ker g_i^{k_i-1}$  en  $\ker g_i^{k_i}$ . Pero implica mucho más pues el procedimiento explicado en la Figura 6.1 continua ampliando la lista de forma que

$$\begin{aligned} \ker g_i^{k_i-1} &= \ker g_i^{k_i-2} \oplus \langle B(u_1), \dots, B(u_r), u_{r+1}, \dots, u_s \rangle \\ &= \ker g_i^{k_i-2} \oplus \langle B(u_1), \dots, B(u_r) \rangle \oplus \langle u_{r+1}, \dots, u_s \rangle. \end{aligned}$$

Lo que, con el Lema 6.6.2 muestra que

$$u_1, \dots, u_r, B(u_1), \dots, B(u_r), u_{r+1}, \dots, u_s$$

es una base de un complemento de  $\ker g_i^{k_i-2}$  en  $\ker g_i^{k_i-1}$ . El proceso continua de la siguiente manera de forma que los elementos de la fila  $j$ -ésima forman una base de un complemento de  $\ker g_i^{k_i-j-1}$  en  $\ker g_i^{k_i-j}$ . En el último paso tenemos que los elementos de la última fila forman una base de  $\ker g_i$ , el espacio de vectores propios de valor propio  $\lambda_i$ . Juntando todo esto deducimos el siguiente

**Lema 6.6.3** *Los vectores obtenidos por el procedimiento explicado en la Figura 6.1 forman una base de  $\ker g_i^{k_i}$ .*

Para terminar la demostración del Teorem 6.5.1 nos falta demostrar la siguiente igualdad:

$$V = \ker g_1^{k_1} \oplus \dots \oplus \ker g_r^{k_r}.$$

Para conseguir esto basta demostrar que los sumandos son independientes y que la suma de sus dimensiones es  $n$ .

Para conseguir lo primero observamos que los  $g_i$  conmutan entre si pues

$$g_i g_j = (f - \lambda_i I)(f - \lambda_j I) = f^2 - (\lambda_i + \lambda_j)f + \lambda_i \lambda_j I = (f - \lambda_j I)(f - \lambda_i I) = g_j g_i$$

y también conmutan con  $f$  por un argumento similar. Eso sirve para demostrar las siguientes propiedades:

**Lema 6.6.4** (1) *Si  $x \in \ker g_i^h$  entonces  $f(x) \in \ker g_i^h$  y  $g_j(x) \in \ker g_i^h$  para todo  $i$  y  $j$ .*

(2) *Si  $i \neq j$  y  $h, k \geq 1$  entonces  $\ker g_i^h \cap \ker g_j^k = 0$ .*

(3) *Si  $i \neq j$  entonces la restricción de  $g_i$  a  $\ker g_j^h$  es un automorfismo de  $\ker g_i^h$ .*

**Demostración.** (1) Si  $x \in \ker g_i^h$  entonces  $g_i^h(f(x)) = f(g_i^h(x)) = f(0) = 0$  y por tanto  $f(x) \in \ker g_i^h$ . Por tanto  $f(x) \in \ker g_i^h$  y de la misma manera se demuestra que  $g_j(x) \in \ker g_i^h$ .

(2) Sea  $x \in \ker g_i^h \cap \ker g_j^k$ . Tenemos que demostrar que  $x = 0$ . Razonamos por inducción sobre  $h$ . Supongamos primero que  $h = 1$ . Entonces  $f(x) = \lambda_i x$  y por tanto  $g_j(x) = f(x) -$

$\lambda_j x = (\lambda_i - \lambda_j)x$ . Por tanto  $0 = g_j^k(x) = (\lambda_i - \lambda_j)x$  y como  $\lambda_i \neq \lambda_j$  tenemos que  $x = 0$ . Supongamos ahora que  $h > 1$  y sea  $y = g_i(x)$ . Del apartado (1) tenemos que  $y \in \ker g_j^k$  pero además  $y \in \ker g_i^{h-1}$ . Por la hipótesis de inducción tenemos que  $y = 0$ , o sea  $x \in \ker g_i \cap \ker g_j^k$  y por el caso en que  $h = 1$  tenemos que  $x = 0$ .

(3) es una consecuencia inmediata de (1) y (2). ■

**Lema 6.6.5**  $\ker g_1^{k_1}, \dots, \ker g_r^{k_r}$  son independientes.

**Demostración.** Supongamos que  $v_1 + \dots + v_r = 0$  con  $v_i \in \ker g_i^{k_i}$ . Tenemos que demostrar que todos los  $v_i$  son cero y para ello vamos a razonar por inducción sobre  $r$ . El caso en que  $r = 1$  es obvio. Supongamos que  $r > 1$ . Como  $v_r \in \ker g_r^{k_r}$ , aplicando  $g_r^{k_r}$  tenemos

$$g_r^{k_r}(v_1) + \dots + g_r^{k_r}(v_{r-1}) = 0.$$

Por la primera parte del Lema 6.6.4 tenemos que cada uno de los sumandos con  $i = 1, \dots, r-1$  cumple  $g_r^{k_r}(v_i) \in \ker g_i^{k_i}$ . Aplicando la hipótesis de inducción deducimos que  $g_r^{k_r}(v_i) = 0$  para todo  $i$ . Pero entonces  $v_i \in \ker g_i^{k_i} \cap \ker g_r^{k_r} = 0$ , por la segunda parte del Lema 6.6.4. ■

Nos falta demostrar el siguiente

**Lema 6.6.6**  $V = \sum_{i=1}^r \ker g_i^{k_i}$

**Demostración.** Razonamos por reducción al absurdo. Supongamos que no se de la igualdad y sean  $U = \sum_{i=1}^r \ker g_i^{k_i}$  y  $W$  un complemento de  $U$  en  $V$ , o sea  $V = U \oplus W$ . Del Lema 6.6.4 se tiene que  $f(U) \subseteq U$ . Por tanto la restricción de  $f$  a  $U$  define un endomorfismo  $R$  de  $U$ . Por otro lado la restricción de  $f$  a  $W$  define un homomorfismo  $S : W \rightarrow V$ . Si  $B_1$  es una base de  $U$  y  $B_2$  es una base de  $W$  entonces  $B = B_1 \cup B_2$  es una base de  $V$  y la matriz asociada a  $f$  en esta base tiene la forma

$$f_B = \left( \begin{array}{c|c} R_{B_1} & A_1 \\ \hline 0 & A_2 \end{array} \right)$$

donde  $\begin{pmatrix} A_1 \\ A_2 \end{pmatrix}$  es la matriz asociada a  $S$  en las bases  $B_2$  y  $B$ . Entonces

$$\chi_f = \chi_{f_B} = \chi_{R_{B_1}} \chi_{A_2}.$$

Como  $\chi_f$  es completamente factorizable y  $\chi_S$  no es constante,  $\chi_S$  tiene una raíz que será un valor propio de  $f$ , o sea será  $\lambda_i$  para algún  $i$ .

Para cada  $w \in W$  existen únicos  $a(w) \in U$  y  $b(w) \in W$  tales que  $S(w) = f(w) = a(w) + b(w)$  y  $b$  es un endomorfismo de  $w$  para el que  $b_{B_2} = A_2$ . Como  $\lambda_i$  es valor propio de  $A_2$ , también lo es de  $b$  con lo que existe  $w \in W \setminus 0$  tal que  $b(w) = \lambda_i w$ . Por tanto  $g_i(w) = f(w) - \lambda_i w = a(w) - b(w) = a(w) \in U$ . Luego

$$g_i(w) = v_1 + \dots + v_r$$

con  $v_j \in \ker g_j^{k_j}$  para todo  $j$  y por tanto

$$g_i^{k_i+1}(w) = \sum_{j \neq i} g_i^{k_i}(v_j).$$

Del tercer apartado del Lema 6.6.4 tenemos que para todo  $j \neq i$  existe un  $w_j \in \ker g_j^{k_j}$  tal que  $g_i(w_j) = v_j$ . Por tanto

$$g_i^{k_i+1}(w) = \sum_{j \neq i} g_i^{k_i+1}(w_j),$$

con lo que  $w - \sum_{j \neq i} w_j \in \ker g_i^{k_i+1} = \ker g_i^{k_i} \subseteq U$ . Eso implica que  $w \in U \cap W = 0$  en contra de que  $w = 0$ . ■

## 6.7 Aplicación 2. Ecuaciones diferenciales (continuación)

Vamos a ver en esta sección cómo resolver un sistema de ecuaciones lineales  $X' = AX$ , con  $X = (x_1(t), \dots, x_n(t))$  una lista de funciones en la variable  $t$  y  $A$  es una matriz cuadrada no diagonalizable. El planteamiento inicial es similar al que vimos en la Sección 6.4 es decir hacemos un cambio de variable  $Y = UX$  donde  $U$  es una matriz invertible para la que  $U^{-1}AU = J$  es una matriz de Jordan. Eso cambia el sistema original por  $Y' = JY$ . Por tanto podemos suponer que  $A = J$ , una matriz de Jordan y en la práctica basta considerar el caso en que  $J$  es una matriz de Jordan elemental  $J = J_n(a)$ . Eso implica que el sistema de ecuaciones diferenciales lineales a considerar es de la siguiente forma

$$\begin{aligned} x_1' &= ax_1, \\ x_2' &= x_1 + ax_2, \\ x_3' &= x_2 + ax_3, \\ &\vdots \\ x_n' &= x_{n-1} + ax_n. \end{aligned}$$

y la solución es

$$\begin{aligned} x_1 &= C_1 e^{at}, \\ x_2 &= (C_2 + C_1 t) e^{at}, \\ x_3 &= \left( C_3 + C_2 t + C_1 \frac{t^2}{2} \right) e^{at}, \\ &\vdots \\ x_{n-1} &= \left( C_{n-1} + C_{n-2} t + C_{n-3} \frac{t^2}{2} + \cdots + C_2 \frac{t^{n-2}}{(n-2)!} + C_1 \frac{t^{n-1}}{(n-1)!} \right) e^{at}, \\ x_n &= \left( C_n + C_{n-1} t + C_{n-2} \frac{t^2}{2} + \cdots + C_2 \frac{t^{n-1}}{(n-1)!} + C_1 \frac{t^n}{n!} \right) e^{at}, \end{aligned}$$

con  $C_1, \dots, C_n$  constantes. O sea

$$x_i = e^{at} \sum_{j=0}^{i-1} C_{j+1} \frac{t^{i-j}}{(i-j)!}.$$

**Ejemplo 6.7.1** Consideremos el siguiente sistema de ecuaciones diferenciales:

$$\begin{aligned} x_1' &= x_1 - 14x_2 + 6x_3 - 2x_4 + x_5 \\ x_2' &= 2x_2 \\ x_3' &= x_2 + 2x_3 \\ x_4' &= -3x_2 + 2x_3 + x_4 \\ x_5' &= 10x_2 - 3x_3 + x_4 + x_5 \end{aligned}$$

o lo que es lo mismo, si  $X = (x_1, x_2, x_3, x_4, x_5)$  entonces

$$X' = AX$$

donde  $A$  es la matriz del Ejemplo 6.5.5. Por las cuentas que hicimos en dicho ejemplo tenemos

$$U^{-1}AU = \left( \begin{array}{c|c} J_3(1) & \\ \hline & J_2(2) \end{array} \right) \quad \text{con} \quad U = (v_1, v_2, v_3, v_4, v_5) = \begin{pmatrix} 0 & -2 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & -5 & 2 \\ 0 & 1 & 0 & 6 & -1 \end{pmatrix}.$$

Por tanto si  $(y_1, y_2, y_3, z_4, z_5) = UX$ ,  $Y = (y_1, y_2, y_3)$  y  $Z = (z_1, z_2)$  tenemos que

$$Y' = J_3(1)Y \quad \text{y} \quad Z' = J_2(2)$$

Por tanto

$$y_1 = C_1 e^t, \quad y_2 = (C_2 + C_1 t) e^t, \quad y_3 = \left( C_3 + C_2 t + C_1 \frac{t^2}{2} \right) e^t, \quad z_1 = C_4 e^{2t}, \quad z_2 = (C_5 + C_4 t) e^{2t}.$$

Calculamos la inversa de  $U$  para obtener

$$U^{-1} = \begin{pmatrix} 0 & 5 & -2 & 1 & 0 \\ 0 & -6 & 1 & 0 & 1 \\ 1 & -13 & 1 & 0 & 2 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

y obtenemos la solución del sistema original:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 0 & 5 & -2 & 1 & 0 \\ 0 & -6 & 1 & 0 & 1 \\ 1 & -13 & 1 & 0 & 2 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ z_1 \\ z_2 \end{pmatrix}.$$

Luego

$$\begin{aligned}x_1 &= 5y_2 - 2y_2 + z_1 = ((5C_2 - 2C_3) + (5C_1 - 2C_2)t - C_1t^2)e^t + C_4e^{2t} \\x_2 &= -6y_2 + y_3 + z_2 = \left((C_3 - 6C_2) + (C_2 - 6C_1)t + \frac{C_1}{2}t^2\right)e^t + (C_5 + C_4t)e^{2t} \\x_3 &= y_1 - 13y_2 + y_3 + 2z_2 = \left((C_1 - 13C_2 + C_3) + (C_2 - 13C_1)t + \frac{C_1}{2}t^2\right)e^t + 2(C_5 + C_4t)e^{2t} \\x_4 &= y_2 = (C_2 + C_1t)e^t \\x_5 &= y_3 = \left(C_3 + C_2t + C_1\frac{t^2}{2}\right)e^t.\end{aligned}$$

## 6.8 Forma de Jordan real

El Teorema de Jordan (Teorema 6.5.1) junto con el Teorema Fundamental del Álgebra (Teorema 1.4.7) nos garantizan que toda matriz con entradas complejas tendrá una forma de Jordan pero eso no funciona para matrices con entradas reales. Podemos considerar que una matriz con entradas reales es en la práctica una matriz con entradas complejas y tratarla como tal pero si los valores propios no son reales la matriz de Jordan tendrá entradas no reales y lo mismo pasará con la matriz de paso. En esta sección vamos a ver que podemos evitar esto sin salirnos de los números reales cambiando ligeramente las matrices de Jordan. La idea es que si  $A \in M_n(\mathbb{R})$  entonces  $\chi_A$  es un polinomio con coeficientes reales. Eso implica que si  $\lambda$  es un número complejo que es un valor propio de  $A$ , o sea  $\chi_A(\lambda) = 0$  entonces  $\chi_A(\bar{\lambda}) = \overline{\chi_A(\lambda)} = 0$ . O sea  $\bar{\lambda}$  también es un valor propio. De hecho  $\lambda$  y  $\bar{\lambda}$  tienen la misma multiplicidad como raíces de  $\chi_A$  pues  $(X - \lambda)^n$  divide a  $\chi_A$  si y solo si  $(X - \bar{\lambda})^n$  divide a  $\chi_A$ . Además si  $v \in \mathbb{C}^n$  y  $\bar{v}$  denota el vector que se obtiene cambiando sus entradas por sus conjugados entonces  $Av = \lambda v$  si y solo si  $A\bar{v} = \bar{\lambda}\bar{v}$  con lo que  $v$  es vector propio de valor propio  $\lambda$  si y solo si  $\bar{v}$  es vector propio de valor propio  $\bar{\lambda}$ . Esto muestra que la aplicación  $v \mapsto \bar{v}$  es un isomorfismo  $N(A - \lambda I) \rightarrow N(A - \bar{\lambda} I)$ . De hecho también proporciona un isomorfismo de  $N((A - \lambda I)^k)$  a  $N((A - \bar{\lambda} I)^k)$ . Eso muestra que los tamaños de las submatrices que componen la forma de Jordan de  $A$  correspondiente a los autovalores  $\lambda$  y  $\bar{\lambda}$  son los mismos y de hecho si  $v_1, \dots, v_k$  son los vectores de la base que proporcionan unos de las submatrices elementales de Jordan para el valor propio  $\lambda$ , es decir

$$Av_1 = \lambda v_1 + v_2, \quad Av_2 = \lambda v_2 + v_3, \quad \dots, \quad Av_{k-1} = \lambda v_{k-1} + v_k, \quad Av_k = \lambda v_k,$$

entonces

$$A\bar{v}_1 = \bar{\lambda}\bar{v}_1 + \bar{v}_2, \quad A\bar{v}_2 = \bar{\lambda}\bar{v}_2 + \bar{v}_3, \quad \dots, \quad A\bar{v}_{k-1} = \bar{\lambda}\bar{v}_{k-1} + \bar{v}_k, \quad A\bar{v}_k = \bar{\lambda}\bar{v}_k,$$

Si ponemos

$$\lambda = a + bi$$

y

$$v_j = u_j + w_j i \quad \text{con} \quad u_j, w_j \in \mathbb{R}^n.$$

Entonces

$$Au_1 + (Aw_1)i = Av_1 = \lambda v_1 + v_2 = (au_1 - bw_1 + u_2) + (bu_1 + aw_1 + w_2)i$$

o lo que es lo mismo

$$Au_1 = au_1 - bw_1 + u_2 \quad y \quad Aw_1 = bu_1 + aw_1 + w_2.$$

De la misma manera

$$Au_2 = au_2 - bw_2 + u_3 \quad y \quad Aw_2 = bu_2 + aw_2 + w_3$$

y siguiendo de esta manera tenemos

$$Au_j = au_j - bw_j + u_{j+1} \quad y \quad Aw_j = bu_j + aw_j + w_{j+1} \quad \text{para } j = 1, \dots, k-1.$$

y

$$Au_k = au_k - bw_k \quad y \quad Aw_k = bu_k + aw_k.$$

Si además observamos que

$$u_j = \frac{v_j + \bar{v}_j}{2} \quad w_j = \frac{v_j - \bar{v}_j}{2i}$$

nos damos cuenta de que en el caso en que  $\lambda$  no sea real se tiene que

$$\langle v_1, \dots, v_k, \bar{v}_1, \dots, \bar{v}_1, \dots, \bar{v}_k \rangle = \langle u_1, w_1, u_2, w_2, \dots, u_k, w_k \rangle$$

como subespacios vectoriales de  $\mathbb{C}^n$  y multiplicar por  $A$  da un endomorfismo de este espacio vectorial que en la base formada por los  $v_i$  y los  $\bar{v}_i$  toma la forma

$$\left( \begin{array}{c|c} J_k(\lambda) & \\ \hline & J_k(\bar{\lambda}) \end{array} \right)$$

pero en la base formada por los  $u_i$  y  $w_i$  toma la forma

$$JR_k(a, b) = \left( \begin{array}{cccccc} A & & & & & \\ I_2 & A & & & & \\ & I_2 & A & & & \\ & & \ddots & \ddots & & \\ & & & I_2 & A & \\ & & & & I_2 & A \end{array} \right) \quad \text{con } A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}.$$

Una matriz de este tipo se llama *matriz de Jordan real elemental* y una *matriz de Jordan real* es una formada poniendo diagonalmente matrices de Jordan elementales con valores propios reales o matrices de Jordan reales elementales.

En resumen agrupando cajas elementales de Jordan del mismo tamaño con valores propios conjugados y cambiando las parejas  $v_j$  y  $\bar{v}_j$  de la base por las parejas  $u_j$  y  $w_j$  obtenemos otra matriz de paso  $U$  para la cual  $U^{-1}AU$  es una matriz de Jordan real. En el Ejemplo 6.2.2 obtuvimos la diagonalización de la matriz

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

utilizando como base de vectores propios

$$v_1 = (1, i), \quad \bar{v}_1 = (1, -i)$$

Si aplicamos el método que hemos explicado tendríamos que usar la base

$$u_1 = (1, 0), \quad w_1 = (0, 1)$$

o sea la base canónica lo que no le haría nada a la matriz original que efectivamente ya es una matriz de Jordan real  $J_1(0, 1)$ .

**Ejemplo 6.8.1** En el Ejemplo 6.5.6 hemos obtenido la forma de Jordan de la matriz

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

con la matriz de paso

$$U = \begin{pmatrix} -3i & 1-i & 3i & 1+i \\ 1-i & -i & 1+i & i \\ 0 & -i & 0 & i \\ 2 & 0 & 2 & 0 \end{pmatrix}$$

o sea con la base que proporciona la forma de Jordan es  $v_1 = (-3i, 1-i, 0, 2)$ ,  $v_2 = (1-i, -i, -i, 0)$ ,  $\bar{v}_1, \bar{v}_2$ .

Para obtener la forma de Jordan real tomamos como base

$$u_1 = (0, 1, 0, 2), w_1 = (-3, -1, 0, 0), u_2 = (1, 0, 0, 0), w_2 = (-1, -1, -1, 0).$$

Es decir con la matriz de paso

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

obtenemos la matriz de Jordan real

$$V^{-1}AV = \left( \begin{array}{cc|cc} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ \hline 1 & 0 & 0 & 1 \\ 0 & 1 & -1 & 0 \end{array} \right).$$

**Ejemplo 6.8.2** Vamos a calcular la forma de Jordan real de la siguiente matriz:

$$A = \begin{pmatrix} 3 & 5 & 8 & -18 & 31 & 27 \\ -1 & -1 & 6 & 20 & -74 & -69 \\ 0 & 0 & 1 & -4 & 14 & 20 \\ 0 & 0 & 6 & 8 & -30 & -21 \\ 0 & 0 & 1 & 2 & -3 & -2 \\ 0 & 0 & 0 & -1 & 0 & 2 \end{pmatrix}.$$

El polinomio característico es

$$\begin{aligned}\chi_A &= X^6 - 10X^5 + 60X^4 - 204X^3 + 461X^2 - 546X + 338 \\ &= (X - (1 + i))(X - (1 - i))(X - (2 + 3i))^2(X - (2 - 3i))^2.\end{aligned}$$

Calculamos

$$N(A - (1 + i)I) = \langle (-2 - i, 1, 0, 0, 0, 0) \rangle$$

que nos proporciona un vector propio  $v_1 = (-2 - i, 1, 0, 0, 0, 0)$  de valor propio  $1 + i$  y otro vector propio  $\bar{v}_1 = (-2 + i, 1, 0, 0, 0, 0)$  de valor propio  $1 - i$ .

Ahora calculamos

$$N(A - (2 + 3i)I) = \langle (2 - i, 1, -2i, -3i, -1 - i, 1) \rangle$$

que nos proporcionará otros dos vectores propios pero que claramente muestra que la matriz no es diagonalizable. Calculamos pues

$$N((A - (2 + 3i)I)^2) = \langle (3i, 3 - 2i, -1 + i, 1, 0, 0), (-7 - i, 7 + i, -3 - 5i, 0, -1 - i, 1) \rangle.$$

Así tomamos un elemento  $v_2 = (3i, 3 - 2i, -1 + i, 1, 0, 0) \in N((A - (2 + 3i)I)^2) \setminus N(A - (2 + 3i)I)$  y con él calculamos

$$v_3 = (A - (2 + 3i)I)v_2 = (-2 + i, -1, 2i, 3i, 1 + i, -1)$$

Esto nos proporciona una base  $\{v_1, \bar{v}_1, v_2, v_3, \bar{v}_2, \bar{v}_3\}$  con la que podemos calcular la matriz de paso

$$U = \begin{pmatrix} -2 - i & -2 + i & 3i & -2 + i & -3i & -2 - i \\ 1 & 1 & 3 - 2i & -1 & 3 + 2i & -1 \\ 0 & 0 & -1 + i & 2i & -1 - i & -2i \\ 0 & 0 & 1 & 3i & 1 & -3i \\ 0 & 0 & 0 & 1 + i & 0 & 1 - i \\ 0 & 0 & 0 & -1 & 0 & -1 \end{pmatrix}$$

Tomando ahora las partes reales e imaginarias de los vectores por separado obtenemos otra base:

$$\begin{aligned}u_1 &= (-2, 1, 0, 0, 0, 0), & w_1 &= (-1, 0, 0, 0, 0, 0), \\ u_2 &= (0, 3, -1, 1, 0, 0), & w_2 &= (3, -2, 1, 0, 0, 0), \\ u_3 &= (-2, -1, 0, 0, 1, -1), & w_3 &= (1, 0, 2, 3, 1, 0).\end{aligned}$$

O sea la matriz de paso

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

obtenemos la forma de Jordan real de  $A$ :

$$V^{-1}AV = \left( \begin{array}{cc|cc|cc} 1 & 1 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 2 & 3 & 0 & 0 \\ 0 & 0 & -3 & 2 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 & 2 & 3 \\ 0 & 0 & 0 & 1 & -3 & 2 \end{array} \right)$$

## 6.9 Problemas

**6.1** Hallar todas las matrices semejantes con una matriz identidad. ¿Y con las matrices de la forma  $\alpha I$  con  $\alpha$  un escalar?

**6.2** Demostrar que si una matriz solo tiene una matriz semejante entonces es de la forma  $\alpha I$  con  $\alpha$  un escalar.

**6.3** Sea  $A$  una matriz cuadrada con  $n$  filas. Demostrar que  $\chi_A$  tiene grado  $n$ , su término principal es 1, su término independiente es  $(-1)^n \det(A)$  y el coeficiente de  $X^{n-1}$  es el opuesto de la traza de  $A$ .

**6.4** Diagonalizar, si es posible, las siguientes matrices:

$$\begin{pmatrix} 3 & 4 \\ 5 & 2 \end{pmatrix}, \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 1 \\ -1 & 1 & 1 \\ 0 & 3 & 2 \end{pmatrix}, \begin{pmatrix} 3 & 1 & 0 \\ -4 & 1 & 0 \\ 4 & -8 & -2 \end{pmatrix}, \begin{pmatrix} 2 & 5 & -6 \\ 4 & 6 & -9 \\ 3 & 6 & -8 \end{pmatrix}, \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix},$$

$$\begin{pmatrix} 17 & -36 & 54 & -147 & 78 \\ 34 & -69 & 102 & -269 & 142 \\ 10 & -20 & 29 & -65 & 34 \\ -6 & 12 & -18 & 52 & -26 \\ -6 & 12 & -18 & 49 & -23 \end{pmatrix}, \begin{pmatrix} -7 & 16 & -22 & 57 & -30 \\ 1 & -1 & 4 & -8 & 4 \\ -4 & 8 & -11 & 41 & -22 \\ -6 & 12 & -18 & 52 & -26 \\ -6 & 12 & -18 & 49 & -23 \end{pmatrix}.$$

**6.5** Diagonalizar el endomorfismo  $f : R^3 \rightarrow R^3$  definido por  $f(x, y, z) = (x - z, x + 2y + z, 2x + 2y + 3z)$ .

**6.6** Sea  $V = K^{\mathbb{N}}$  el espacio vectorial de las sucesiones infinitas del Ejemplo 6.1.5. Sea  $f$  el endomorfismo de  $V$  dado por  $f(x_1, x_2, \dots) = (x_2, x_3, \dots)$ . Calcular todos los valores y vectores propios de  $f$ . Hacer lo mismo para el endomorfismo  $g(x_1, x_2, \dots) = (0, x_1, x_2, \dots)$ .

**6.7** Sea  $\mathcal{C}(\mathbb{R})$  el espacio vectorial de las funciones continuas de variable real del Ejemplo 2.1.7. Sea  $W$  el subespacio formado por las funciones infinitamente diferenciales y sea  $D : W \rightarrow W$  la aplicación derivada. Demostrar que las funciones exponenciales  $f(x) = a^x$  son vectores propios de  $D$ .

**6.8** Estudiar para que valores de  $a \in \mathbb{R}$  la matriz  $\begin{pmatrix} a & 0 & 0 \\ 1 & 1 & 0 \\ a & 1 & a \end{pmatrix}$  es diagonalizable.

**6.9** Diagonalizar la matriz  $\begin{pmatrix} a & b \\ b & a \end{pmatrix}$  donde  $a$  y  $b \in \mathbb{R}$  no son cero.

**6.10** Sean  $a$  y  $b$  dos números reales que satisfacen  $a^2 + b^2 = 1$ . Estudiar si la siguiente matriz es diagonalizable como matriz de números reales y como matriz de números complejos:

$$A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}.$$

Interpretar geoméricamente la aplicación lineal  $x \mapsto Ax$ .

Hacer lo mismo con las siguientes matrices

$$\begin{pmatrix} 1 & & \\ & a & b \\ & -b & a \end{pmatrix} \quad \begin{pmatrix} -1 & & \\ & a & b \\ & -b & a \end{pmatrix}.$$

**6.11** Sea  $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  una aplicación lineal dada por:  $f(1, 0, 0) = (0, a^2, 0)$ ,  $f(1, 0, 1) = (0, 1 - a^2, 1)$ ,  $f(0, 1, 1) = (-1, 1, 1)$ .

Hallar:

- (1) Matriz asociada en la base canónica.
- (2) Dimensión de la imagen y dimensión del núcleo, según los valores de  $a \in \mathbb{R}$ .
- (3) Valores propios de  $f$ .
- (4) Valores de  $a \in \mathbb{R}$  para los cuales la matriz es diagonalizable.
- (5) Ecuaciones implícitas de  $N(f, \lambda)$  para cada valor propio  $\lambda$  de  $f$ .

**6.12** Sea  $A \in M_3(\mathbb{R})$  una matriz cuyos valores propios son 0 y 1 y

$$N(A, 0) = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : y + z = x = 0 \right\} \quad N(A, 1) = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid x - 2y + z = 0 \right\}.$$

¿Es  $A$  diagonalizable? Hallar la matriz  $A$ .

**6.13** Sea  $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  un endomorfismo que verifica:

- Los valores propios de  $f$  son 2 y  $-1$ .
- $N(f, 2)$  es el conjunto de soluciones del sistema lineal de ecuaciones:  $\left. \begin{array}{l} x + y + z = 0 \\ x - 2y + z = 0 \end{array} \right\}$ .

- $N(f, -1)$  es el conjunto de soluciones del sistema lineal de ecuaciones: 
$$\left. \begin{aligned} x + ay - z &= 0 \\ ax + y - az &= 0 \end{aligned} \right\}$$

- (1) Determinar para qué valores de  $a$  el endomorfismo  $f$  es diagonalizable.
- (2) Calcular una fórmula para  $f(x, y, z)$ , para los valores de  $a$  tal que  $f$  es diagonalizable.

**6.14** Sea  $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & -4 & -1 \end{pmatrix}$ . Calcular  $A^{29}$  y  $A^{30}$ . Probar que para toda  $n \in \mathbb{N}$ ,  $A^{2n} = I$ .

**6.15** Calcular la potencia  $n$ -ésima de las matrices diagonalizables del Ejercicio 6.4.

**6.16** Calcular una fórmula para  $\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}^n$ .

**6.17** Demostrar las siguientes afirmaciones:

- (1) Una matriz es invertible si y sólo si no tiene a 0 como valor propio.
- (2) Si  $\lambda$  es valor propio de una matriz cuadrada  $A$  entonces  $\lambda^n$  es valor propio de  $A^n$ , para  $n \in \mathbb{N}$ .
- (3) Si  $\lambda$  es valor propio de una matriz cuadrada  $A$  entonces  $\lambda^{-1}$  es valor propio de  $A^{-1}$ .
- (4) Si una matriz cuadrada es nilpotente, es decir existe un número natural  $n$  tal que  $A^n = 0$ , y diagonalizable entonces ha de ser  $A = 0$ .

**6.18** Sea  $A$  una matriz cuadrada  $n \times n$  y diagonalizable cuyos valores propios son 1 y  $-1$ . Demostrar que  $A$  verifica:

$$A^m = \begin{cases} I_n & \text{si } m \text{ es par} \\ A & \text{si } m \text{ es impar.} \end{cases}$$

**6.19** Sean  $\{a_n\}$ ,  $\{b_n\}$  y  $\{c_n\}$  tres sucesiones recurrentes de números reales cuyos términos generales verifican las siguientes relaciones:

$$\left. \begin{aligned} a_n &= 4a_{n-1} - 5b_{n-1} - 2c_{n-1} \\ b_n &= 2a_{n-1} - 3b_{n-1} - 2c_{n-1} \\ c_n &= a_{n-1} - b_{n-1} + c_{n-1} \end{aligned} \right\}.$$

Sabiendo que  $a_0 = 1$ ,  $b_0 = -1$  y  $c_0 = 2$ , calcular los términos generales de las sucesiones.

**6.20** Calcular el término general de la sucesión de vectores  $(v_n)$ , donde  $v_0 = \begin{pmatrix} 1 \\ 11 \\ -2 \end{pmatrix}$  y

$$v_{n+1} = \frac{1}{9} \begin{pmatrix} 7 & -2 & 0 \\ -2 & 6 & 2 \\ 0 & 2 & 5 \end{pmatrix} v_n.$$

**6.21** Un curso de álgebra se imparte en dos grupos. Si cada semana abandonan el curso un 10 por ciento del primer grupo y un 20 por ciento del segundo y además  $1/6$  de alumnos de cada grupo se pasa a la otro. ¿Cuántos alumnos quedarán en cada grupo después de tres meses si empezó el curso con 100 alumnos en cada grupo

**6.22** Un país está dividido en tres regiones demográficas. Se ha determinado que cada año el 5 por ciento de los residentes en la region A se cambia a la región B y otro 5 por ciento a la región C. De los residentes de la región B, el 10 por ciento se muda a la A y el 5 por ciento a la C. Finalmente, de la región C, el 15 por ciento se muda a la A y el 10 por ciento a la B. ¿Que porcentaje reside en cada una de las regiones tras un largo plazo?

**6.23** Supongamos que una empresa tiene camiones en tres centros; uno en Sevilla, otro en Barcelona y otro en Madrid. Cada mes la mitad de los que están en Barcelona y en Madrid, van a Sevilla y la otra mitad no se mueve. Los camiones que están en Sevilla se mueven todos, la mitad va a Barcelona y la otra mitad a Madrid. ¿Existe un estado estacionario ¿Cuál?

**6.24** Para estudiar la evolución de una enfermedad en una población dividimos ésta en tres grupos:  $I$  = inmunes,  $E$  = enfermos y  $R$  = de riesgo = resto de la población. En un cierto periodo de tiempo: Se ha observado estadísticamente que se verifican las siguientes transiciones:

- (1) El 20% de la población inmune, permanece en tal estado y el 80% pierde su inmunidad sin enfermar.
- (2) De la población de riesgo el 10% pasa a estar inmune y el 20% enferma, permaneciendo el 70% restante como población de riesgo.
- (3) En dicho periodo de tiempo todos los enfermos sanan adquiriendo el 40% inmunidad y el 60% no.

Calcular la evolución de la población a largo plazo, es decir estudiar la distribución de población en los tres grupos en el infinito. Para ello, partir de una distribución  $(i_0, e_0, r_0)$  inicial, calcular  $(i_n, e_n, r_n)$  después de  $n$  periodos de tiempo, en función de  $(i_0, e_0, r_0)$  y calcular el límite cuando  $n$  tiende a infinito del resultado obtenido.

**6.25** Los ejercicios anteriores son ejemplos de Procesos de Markov en los que se estudia la evolución de la distribución de una población dividida en un número finito de clases en las que hay una matriz de transición que rige las transiciones entre las distintas clases en un cierto periodo de tiempo. Obsérvese que estas matrices están formadas por números positivos y que la suma de las entradas de cada columna ha de ser 1 por ser las probabilidades de pasar de uno de los estados dados a los demás estados. Demostrar que para todas estas matrices 1 es un valor propio y todos los valores propios están entre 0 y 1. Los vectores propios de valor propio 1 corresponden a los estados estacionarios, es decir a las distribuciones que permanecen estables y corresponden a las situaciones en el límite. ¿Por qué?

**6.26** La siguiente tabla representa las transiciones entre seis grupos de población divididos por renta. Cada fila corresponde a la proporción de población que pasó en un periodo de tiempo de uno de los grupos al siguiente. Por ejemplo, 47.3% del grupo 1, se quedó en el mismo grupo, 38% pasó al grupo 2, 7.7% pasó al grupo 3, etc.

	Grupo 1	Grupo 2	Grupo 3	Grupo 4	Grupo 5	Grupo 6
Grupo 1	47.3	38.6	7.7	3.8	0.4	2.2
Grupo 2	5.3	59.4	20.3	9.4	4.4	1.1
Grupo 3	0.6	34.8	37.5	14.8	9.3	3.1
Grupo 4	0.4	14.6	32.3	33.0	14.0	5.7
Grupo 5	0.3	10.8	19.5	29.6	29.0	10.9
Grupo 6	0.3	14.4	25.3	25.0	20.7	14.2

(Tened en cuenta que hay errores de redondeo.) Hallar la distribución en el límite, suponiendo que la matriz de transición permanece constante.

**6.27** Calcular el término general de una sucesión  $(a_n)$  de números cuyos dos primeros términos son  $a_1 = 0$  y  $a_2 = 1$ ; y los siguientes vienen generados por la relación  $a_n = 2a_{n-1} + a_{n-2}$  para  $n \geq 3$ .

**6.28** Calcular el término general de una sucesión  $(a_n)$  de números generados por la relación  $a_n = a_{n-1} - a_{n-2}$  para  $n \geq 3$ . Demostrar que independientemente del valor de  $a_1$  y  $a_2$  se verifica que  $a_{n+6} = a_n$  para todo número  $n$ .

**6.29** Calcular el término general de la sucesión  $(x_n)$  dada por  $x_{n+3} = 2x_{n+2} + x_{n+1} + x_n$ , donde  $x_1 = x_2 = 1$  y  $x_3 = 0$ .

**6.30** Observando el crecimiento de una población de conejos llegamos a la conclusión de que las hembras de conejo recién nacidas no tienen descendencia durante su primer mes de vida, pero que a partir de ahí cada hembra produce una media de dos conejos cada mes, uno de cada sexo. Comenzando con una hembra y un macho recién nacidos, estimar el número de conejos nacidos durante  $k$  meses. Calcular el número de conejos después de 30 meses suponiendo que no muere ningún conejo.

**6.31** Resolver los siguientes sistemas de ecuaciones diferenciales lineales:

$$\left. \begin{array}{l} y'_1 = y_1 - 2y_2 \\ y'_2 = 5y_1 + 4y_2 \end{array} \right\} \left. \begin{array}{l} x' = 2x - z \\ y' = 2y - 2z \\ z' = x - y + 2z \end{array} \right\}$$

Resolver las ecuaciones diferenciales

$$y'' = y' + y \quad y''' = 3y'' - y' - 3y$$

(Indicación: Poner  $y_1 = y$  y  $y_2 = y'$ ,  $y_3 = y'' \dots$  y convertir la ecuación en un sistema de ecuaciones diferenciales lineales.)

**6.32** Consideremos un tubo que suponemos de longitud infinita y que lo hemos dividido en cuatro segmentos, como se indica en la figura



de forma que los segmentos  $S_1$  y  $S_4$  no están limitados a izquierda y derecha respectivamente. En el instante  $t = 0$ , los segmentos finitos contienen respectivamente concentraciones  $x(0)$  e  $y(0)$  de alguna solución química, mientras que la concentración en los segmentos infinitos en cualquier instante es 0. La difusión comienza en el instante 0 y está gobernada por la siguiente ley: en cada tiempo  $t$  la velocidad de difusión entre dos segmentos adyacentes es igual a la diferencia de las concentraciones. Por ejemplo, la concentración  $x(t)$  en el segmento  $S_1$  varía  $0 - x(t)$  en la unidad de tiempo por difusión al segmento exterior, e  $y(t) - x(t)$  por difusión al segmento  $S_2$ . Es decir  $x'(t) = (0 - x(t)) + (y(t) - x(t)) = y(t) - 2x(t)$ . Plantear un sistema de ecuaciones diferenciales, resolverlo e interpretarlo físicamente. (Indicación sobre la interpretación física: Los valores propios van a proporcionar las razones de caída de la concentración en el límite y los vectores propios correspondientes van a proporcionar proporciones entre las concentraciones en los segmentos  $S_2$  y  $S_3$  en el límite. Tened en cuenta que valores negativos de  $x$  e  $y$  no tienen un significado físico.)

**6.33** Repetir el ejercicio anterior dividiendo el tubo en cinco secciones, dos infinitas en los extremos y otras tres finitas de la misma longitud. Después repetir el ejercicio con seis secciones. (Observación: Estas soluciones son aproximaciones discretas a la ecuación del calor:  $\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$ .)

**6.34** La ecuación diferencial

$$X'' + k^2 X = 0, \quad (k \in \mathbb{R}). \quad (6.9)$$

aparece en numerosos fenómenos físicos, el más evidente es el de un oscilador. Si la resolvemos por los métodos habituales, nos encontraremos con que el polinomio característico de la matriz asociada es  $X^2 + k^2$  que para  $k$  positivo, no tiene raíces reales. Si interpretamos  $X$  como una función de variable compleja, los valores propios son  $\pm ki$ , donde  $i$  es la raíz de  $-1$ , y obtendremos el conjunto de soluciones:

$$X = ae^{kit} + be^{-kit}.$$

Una interpretación ingenua nos llevaría a concluir que las soluciones no tienen un significado físico, lo cual es bien conocido ser falso ya que las soluciones de (6.9) son

$$X = a \cos(kt) + b \operatorname{sen}(kt), \quad (a, b \in \mathbb{R}) \quad (6.10)$$

con un significado físico más que evidente. Sin embargo todo gana sentido introduciendo la función de variable compleja

$$e^{x+iy} = e^x(\cos y + i \operatorname{sen} y).$$

Es decir, las soluciones de la ecuación diferencial original son

$$X = p(\cos(kt) + i \operatorname{sen}(kt)) + q(\cos(-kt) + i \operatorname{sen}(-kt)) = (p + q) \cos(kt) + i(p - q) \operatorname{sen}(kt).$$

Si entendemos que  $a$  y  $b$  toman valores complejos arbitrarios, pongamos

$$p = p_1 + p_2 i, \quad q = q_1 + q_2 i$$

tendremos

$$X = [(p_1+p_2)+(q_1+q_2)i] \cos(kt) + i[(p_1+p_2)-(q_1+q_2)i] \sin(kt) = (a \cos(kt) + b \sin(kt)) + i(c \cos(kt) + d \sin(kt))$$

para ciertos números reales  $a, b, c$  y  $d$ . Si nos quedamos con la parte real obtenemos la solución (6.10).

¡Calma que el ejercicio empieza ahora!

- (1) Resolver el sistema de ecuaciones diferenciales  $\left. \begin{aligned} X_1'' &= 6X_1 + 10X_2 \\ X_2'' &= -15X_1 - 19X_2 \end{aligned} \right\}$ . Para ello encontrar una matriz  $U$  que diagonalice la matriz de los coeficientes y hacer el cambio de variable  $X = UY$ , para convertir el sistema de ecuaciones en un sistema de dos ecuaciones del tipo (6.9).

- (2) Resolver de la misma forma el siguiente sistema de ecuaciones diferenciales:

$$X'' = \begin{pmatrix} -1 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -1 \end{pmatrix} X,$$

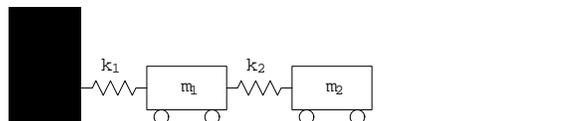
$$\text{con las condiciones iniciales } X(0) = \begin{pmatrix} 2 \\ -2 \\ 0 \end{pmatrix}, \quad X'(0) = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

**6.35** La fuerza que un muelle ejerce sobre un cuerpo es proporcional al desplazamiento del muelle de su posición de equilibrio, por tanto el movimiento de un cuerpo de masa  $m$  sometido a la acción de un muelle viene regido por la siguiente ecuación

$$mX'' + kX = 0$$

donde  $m$  es la masa del cuerpo y  $k$  es la constante de elasticidad del resorte.

**6.36** Consideremos dos carros unidos por un resorte y uno de éstos unido a una pared como se indica en la siguiente figura



Plantear un sistema con dos ecuaciones diferenciales que regulen el movimiento de los dos carros y resolverlo para los siguientes casos:

$k_1$	$k_2$	$m_1$	$m_2$
1	1	2	4
1	2	2	2
1	3	2	1

**6.37** Escribir todos los tipos de matrices de Jordan posibles de orden a lo sumo 4

**6.38** Calcular la forma de Jordan de las siguientes matrices:

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

**6.39** Demostrar que si  $J$  es una matriz elemental de Jordan con autovalor  $a$  entonces  $\dim N(A - aI) = 1$ .

**6.40** Calcular todas las formas de Jordan posible para una matriz cuyo polinomio característico es  $(X + 3)^4(X + 1)^3$ .

**6.41** Demuestra que si  $A$  es una matriz cuadrada entonces  $A$  y  $A^T$  son matrices semejantes. Indicación: Comienza suponiendo que  $A$  es una matriz elemental de Jordan.

**6.42** Sea  $A$  cuyo polinomio característico es de la forma  $(X - \lambda)^n$ . Para cada entero positivo  $i$  sea  $k_i = \dim N((A - \lambda I)^i)$ . También ponemos  $k_0 = 0$ . Demostrar que las siguientes condiciones son equivalentes:

- (1) El esquema de la Figura 6.1 para esta matriz tiene  $m$  pisos.
- (2)  $k_m = n > k_{m-1}$ .
- (3)  $k_{m+1} = k_m > k_{m-1} > \dots > k_2 > k_1$ .

Demostrar que el número de casillas de la fila inferior es  $k_1$ , el de la penúltima fila es  $k_2 - k_1$  y en general el de la  $i$ -ésima fila comenzando a contar por abajo es  $k_i - k_{i-1}$ .

**6.43** Sea  $A$  una matriz cuadrada con entradas reales. Sea  $\lambda$  un número complejo que sea valor propio de  $A$ . Demostrar que  $\bar{\lambda}$  es un valor propio de  $A$ , que  $\lambda$  y  $\bar{\lambda}$  tienen la misma multiplicidad como raíces de  $\chi_A$  y que  $\dim N((A - \lambda I)^i) = \dim N((A - \bar{\lambda} I)^i)$  para todo  $i$ .

**6.44** Sean  $f$  y  $g$  son dos endomorfismos de un espacio vectorial tales que  $g \circ f = f \circ g$ . Demostrar que si  $x \in \ker g$  entonces  $f(x) \in \ker g$  y si  $x \in \text{Im}(g)$  entonces  $f(x) \in \text{Im}(g)$ .

**6.45** Dados un polinomio  $P = p_0 + p_1X + \dots + p_kX^k$ , una matriz cuadrada  $A$  y un endomorfismo  $f$  denotamos

$$P(A) = p_0I + p_1A + \dots + p_kA^k \quad \text{y} \quad P(f) = p_0 + p_1f + \dots + p_kf^k$$

donde se entiende que  $P(f)$  es el endomorfismo dado por  $P(f)(v) = p_0v + p_1f(v) + \dots + p_kf^k(v)$ .

Demostrar las siguientes afirmaciones para polinomios  $P$  y  $Q$ , una matriz  $A$  y un endomorfismo  $f$ .

- (1)  $(P + Q)(A) = P(A) + Q(A)$  y  $(PQ)(A) = P(A)Q(A)$ .
- (2)  $(P + Q)(f) = P(f) + Q(f)$  y  $(PQ)(f) = P(f) \circ Q(f)$ .
- (3) Si  $A = \left( \begin{array}{c|c} A_1 & \\ \hline & A_2 \end{array} \right)$  con  $A_1$  y  $A_2$  matrices cuadradas entonces  $P(A) = \left( \begin{array}{c|c} P(A_1) & \\ \hline & P(A_2) \end{array} \right)$ .
- (4) Si  $P(A) = 0$  y  $P$  divide a  $Q$  entonces  $Q(A) = 0$ .
- (5) Si  $P = X - \lambda$  entonces  $P^n(J_n(\lambda)) = 0$ .
- (6) Si  $A$  es la matriz asociada a  $f$  en una base entonces  $P(A)$  es la matriz asociada a  $P(f)$  en la misma base.
- (7) Si el polinomio característico  $\chi_A$  de  $A$  es completamente factorizable entonces  $P(A) = 0$ . Análogamente, si el polinomio característico  $\chi_f$  de  $f$  es completamente factorizable entonces  $P(f) = 0$  (Teorema de Cayley-Hamilton).
- (8) Si  $\lambda$  es un autovalor de  $f$  entonces  $P(\lambda)$  es autovalor de  $P(f)$ . Usar esto para decir cuáles podrían ser los autovalores de un endomorfismo  $f$  que verificara una de las siguientes condiciones:  $f^2 = 0$ ,  $f^2 = f$ ,  $f^2 = I$ ,  $f^2 = -I$ .

**6.46** Demostrar que una matriz  $A \in M_2(K)$  cumple  $A^2 = I$  si y solo si existen  $a, b, c \in K$  tales que

$$A = a \begin{pmatrix} bc & b^2 \\ -c^2 & -bc \end{pmatrix}.$$

**6.47** Resolver los siguientes sistemas de ecuaciones diferenciales:

$$\begin{cases} x'_1 = x_1 + 2x_3 - 6x_4 \\ x'_2 = x_2 - x_2 + 3x_4 \\ x'_3 = x_3 + 3x_4 \\ x'_4 = 2x_4 \end{cases} \quad \begin{cases} x'_1 = 3x_1 + x_2 + 2x_3 + 5x_4 \\ x'_2 = -x_1 + 2x_2 - 2x_4 \\ x'_3 = -x_2 + 2x_3 + 3x_4 \\ x'_4 = -x_3 - x_4 \end{cases}$$

**6.48** Calcular las formas de Jordan compleja y real de las siguientes matrices:

$$\begin{pmatrix} 3 & -2 & 0 \\ 6 & -1 & -4 \\ 3 & -1 & -1 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & -2 & 0 \\ 2 & -1 & 0 & 2 \\ 1 & 0 & -1 & 0 \\ 0 & -1 & 2 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 \\ -1 & -1 & 1 & -1 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -2 & 0 & 1 \\ 0 & 0 & 0 & 2 & -4 & -4 \end{pmatrix}$$



# Capítulo 7

## Espacios euclídeos y unitarios

En este capítulo vamos dejar de usar un cuerpo arbitrario y solo usaremos los cuerpos  $\mathbb{R}$  de los números reales y  $\mathbb{C}$  de los complejos. Vamos a usar  $\mathbb{K}$  para denotar uno de estos dos cuerpos.

### 7.1 Producto escalar y hermítico

**Definición 7.1.1** Sea  $V$  un espacio vectorial de  $\mathbb{K}$ . Una producto hermítico en  $V$  es una operación

$$\begin{aligned} V \times V &\longrightarrow \mathbb{K} \\ (v, w) &\longmapsto v \cdot w \end{aligned}$$

que satisface las siguientes propiedades para todo  $u, v, w \in V$  y  $\alpha \in \mathbb{K}$ :

(EH1)  $(u + v) \cdot w = u \cdot w + v \cdot w$ .

(EH2)  $(\alpha u) \cdot v = \alpha(u \cdot v)$ .

(EH3)  $v \cdot u = \overline{u \cdot v}$ .

(EH4) Si  $v \neq 0$  entonces  $v \cdot v$  es un número real positivo.

Entre los físicos se suele denotar el producto escalar con la notación de Dirac que algunas veces toma la forma  $\langle u | v \rangle$  y otras la forma  $\langle u, v \rangle$ . Yo, al igual que muchos matemáticos, prefiero usar  $u \cdot v$  por que es más simple.

En el caso en que  $\mathbb{K} = \mathbb{R}$  el producto hermítico se llama *producto escalar*. En algunos textos se separa el tratamiento de producto escalar y producto hermítico pero como tienen muchas propiedades comunes trataremos los dos casos  $\mathbb{K} = \mathbb{R}$  ó  $\mathbb{K} = \mathbb{C}$  a la vez y cuando queramos diferenciarlos lo diremos explícitamente. Otra terminología que se utiliza es la siguiente: Un *espacio euclídeo* es un espacio vectorial sobre  $\mathbb{R}$  con un producto escalar y un *espacio unitario* es un espacio vectorial sobre  $\mathbb{C}$  con un producto hermítico. Nosotros hablaremos de *espacio hermítico* para referirnos a ambos casos a la vez.

De los axiomas se deduce de forma fácil que todo producto hermítico también satisface siguientes igualdades:

$$0 \cdot v = v \cdot 0 = 0, \quad (7.1)$$

$$w \cdot (u + v) = w \cdot u + w \cdot v, \quad (7.2)$$

$$u \cdot (\alpha v) = \bar{\alpha}(u \cdot v) \quad (7.3)$$

En efecto, de las propiedades (EH1) y (EH2) se deduce que para todo  $v \in V$  la aplicación

$$\begin{aligned} V &\rightarrow \mathbb{K} \\ x &\mapsto x \cdot v \end{aligned}$$

es lineal. Por tanto la imagen del 0 es cero, o sea  $0 \cdot v = 0$ . Aplicando ahora (EH3) tenemos que  $v \cdot 0 = \overline{0 \cdot v} = \overline{0} = 0$ . Por otro lado (EH1') y (EH2'):  $w \cdot (u + v) = \overline{(u + v) \cdot w} = \overline{u \cdot w + v \cdot w} = \overline{u \cdot w} + \overline{v \cdot w} = w \cdot u + w \cdot v$  demuestra (7.2) y  $u \cdot (\alpha v) = \overline{(\alpha v) \cdot u} = \overline{\alpha(v \cdot u)} = \overline{\alpha v \cdot u} = \bar{\alpha}(u \cdot v)$  demuestra (7.3).

De (EH4) y (7.1) se deduce que

$$v \cdot v = 0 \quad \Leftrightarrow \quad v = 0.$$

**Ejemplos 7.1.2** (1) Consideremos el siguiente producto en  $\mathbb{K}^n$ :

$$(a_1, \dots, a_n) \cdot (b_1, \dots, b_n) = (a_1, \dots, a_n) \begin{pmatrix} \bar{b}_1 \\ \vdots \\ \bar{b}_n \end{pmatrix} = a_1 \bar{b}_1 + \dots + a_n \bar{b}_n.$$

Esto define un producto hermítico en  $\mathbb{K}^n$ . Nos referiremos a este producto como *producto hermítico estándar* o *producto escalar estándar* si  $\mathbb{K} = \mathbb{R}$ .

Antes de comprobarlo introducimos la siguiente notación para una matriz arbitraria:

$$\overline{(a_{ij})} = (\bar{a}_{ij}).$$

Con esta notación el producto definido está dado por

$$v \cdot w = v \bar{w}^T.$$

Teniendo en cuenta las propiedades de la conjugación se tiene lo siguiente para dos matrices  $A$  y  $B$  que puedan sumarse o multiplicarse:

$$\overline{A^T} = \overline{A}^T, \quad \overline{A + B} = \overline{A} + \overline{B} \quad \text{y} \quad \overline{AB} = \overline{A} \overline{B}.$$

Además, si  $v$  y  $w$  son vectores fila entonces  $wv^T = (vw^T)^T = vw^T$  pues  $wv^T \in \mathbb{K}$ .

(EH1) y (EH2) se comprueban fácilmente como consecuencia de propiedades del producto de matrices:

$$(u + v) \cdot w = (u + v) \bar{w}^T = u \bar{w}^T + v \bar{w}^T = u \cdot w + v \cdot w$$

$$(\alpha u) \cdot v^T = (\alpha u) \overline{w}^T = \alpha(u \overline{w}^T) = \alpha(u \cdot v).$$

(EH3) es consecuencia de las propiedades de la conjugación:

$$w \cdot v = w \overline{v}^T = \overline{v} w^T = \overline{\overline{v} w^T} = \overline{v \cdot w};$$

y (EH4) de que si  $v = (a_1, \dots, a_n)$  entonces

$$v \cdot v = a_1 \overline{a_1} + \dots + a_n \overline{a_n} = |a_1|^2 + \dots + |a_n|^2$$

que es no negativo y solo será cero cuando  $v$  lo sea.

(2) Sea  $A = (a_{ij}) \in M_n(\mathbb{K})$  y definimos el siguiente producto en  $\mathbb{K}^n$ :

$$v \cdot w = v A \overline{w}^T.$$

De nuevo las propiedades (EH1) y (EH2) son consecuencia de las propiedades del producto de matrices. Sin embargo, (EH3) y (EH4) no tienen por qué cumplirse.

Vamos a ver que este producto cumple (EH3) si y solo si  $A^T = \overline{A}$ . Una matriz que cumple esta propiedad se dice que es una matriz *hermítica*. En efecto, si  $A$  es hermítica entonces usando una vez más que  $w A \overline{w}^T \in \mathbb{K}$  tenemos que

$$w \cdot v = w A \overline{v}^T = (w A \overline{v}^T)^T = \overline{v} A^T w^T = \overline{v} \overline{A} w^T = \overline{v A w^T} = \overline{v A \overline{w}^T} = \overline{v \cdot w}.$$

Recíprocamente, supongamos que este producto satisface (EH3) y pongamos  $A = (a_{ij})$ . Si  $e_1, \dots, e_n$  es la base canónica entonces  $e_i A = (a_{i1}, \dots, a_{in})$  y por tanto

$$e_i \cdot e_j = e_i A e_j^T = a_{ij}$$

Por tanto,

$$a_{ji} = e_j \cdot e_i = \overline{e_i \cdot e_j} = \overline{a_{ij}},$$

que obviamente significa que  $A$  es una matriz hermítica.

Para que  $\cdot$  cumpla la propiedad (EH4) se tiene que cumplir que  $v A \overline{v}^T$  sea un número real positivo para todo  $v \in \mathbb{K} \setminus \{0\}$ . En tal caso se dice que  $A$  es *definida positiva*. Por el momento resulta complicado decidir cuándo una matriz cuadrada es definida positiva. Eso lo veremos más adelante. Vamos a considerar por ahora el caso en que  $A$  sea una matriz diagonal

$$A = \text{diag}(a_1, \dots, a_n).$$

Entonces

$$(v_1, \dots, v_n) \cdot (v_1, \dots, v_n) = (v_1, \dots, v_n) A \begin{pmatrix} \overline{v_1} \\ \vdots \\ \overline{v_n}^T \end{pmatrix} = a_1 |v_1|^2 + \dots + a_n |v_n|^2$$

y esto es positivo para todo  $v = (v_1, \dots, v_n) \neq 0$  si y solo si  $a_1, \dots, a_n$  son todos reales positivos. Por tanto una matriz diagonal es definida positiva si y solo si los elementos de su diagonal son números reales positivos.

- (3) Sea  $[a, b]$  un intervalo de la recta real y sea  $\mathcal{C}([a, b])$  el conjunto de las aplicaciones continuas  $[a, b] \rightarrow \mathbb{R}$ . Definimos en  $\mathcal{C}([a, b])$  el siguiente producto

$$f \cdot g = \int_a^b f(x)g(x)dx.$$

Esto define un producto escalar en  $\mathcal{C}([a, b])$ .

Si cambiamos ahora  $\mathcal{C}([a, b])$  por el conjunto de las funciones continuas  $f : [a, b] \rightarrow \mathbb{C}$  podemos definir el siguiente producto que resulta ser un producto hermítico:

$$f \cdot g = \int_a^b f(x)\overline{g(x)}dx.$$

Vamos a ver que si  $\cdot$  es un producto hermítico en un espacio vectorial  $V$  y  $B = \{b_1, \dots, b_n\}$  es una base de  $V$  entonces la siguiente matriz, que vamos a llamar *matriz asociada al producto hermítico*  $\cdot$  en la base  $B$ , determina completamente el producto:

$$A = (b_i \cdot b_j)_{ij}.$$

En efecto, vamos a poner  $a_{ij} = b_i \cdot b_j$  para todo  $1 \leq i, j \leq n$ . Sean  $v$  y  $w$  dos elementos de  $V$  y sean

$$v_B = (v_1, \dots, v_n) \quad \text{y} \quad w_B = (w_1, \dots, w_n).$$

Entonces

$$\begin{aligned} v \cdot w &= (v_1 b_1 + \dots + v_n b_n) \cdot (w_1 b_1 + \dots + w_n b_n) \\ &= (v_1 b_1) \cdot (w_1 b_1 + \dots + w_n b_n) + \dots + (v_n b_n) \cdot (w_1 b_1 + \dots + w_n b_n) \\ &= v_1 (b_1 \cdot (w_1 b_1 + \dots + w_n b_n)) + \dots + v_n (b_n \cdot (w_1 b_1 + \dots + w_n b_n)) \\ &= v_1 (b_1 \cdot (w_1 b_1)) + \dots + v_1 (b_1 \cdot (w_n b_n)) \\ &\quad + \dots + \\ &\quad v_n (b_n \cdot (w_1 b_1)) + \dots + v_n (b_n \cdot (w_n b_n)) \\ &= v_1 \overline{w_1} (b_1 \cdot b_1) + \dots + v_1 \overline{w_n} (b_1 \cdot b_n) \\ &\quad + \dots + \\ &\quad v_n \overline{w_1} (b_n \cdot b_1) + \dots + v_n \overline{w_n} (b_n \cdot b_n) \\ &= v_1 a_{11} \overline{w_1} + \dots + v_1 a_{1n} \overline{w_n} + \dots + v_n a_{n1} \overline{w_1} + \dots + v_n a_{nn} \overline{w_n} \\ &= v_1 a_{11} \overline{w_1} + \dots + v_1 a_{1n} \overline{w_n} + \dots + v_n a_{n1} \overline{w_1} + \dots + v_n a_{nn} \overline{w_n} \\ &= (v_1, \dots, v_n) \begin{pmatrix} a_{11} \overline{w_1} + \dots + v_1 a_{1n} \overline{w_n} \\ \vdots \\ a_{n1} \overline{w_1} + \dots + v_n a_{nn} \overline{w_n} \end{pmatrix} \\ &= (v_1, \dots, v_n) (a_{ij}) \begin{pmatrix} \overline{w_1} \\ \vdots \\ \overline{w_n} \end{pmatrix} = v_B A \overline{w_B}^T. \end{aligned}$$

Hemos pues demostrado que si  $A$  es la matriz asociada a  $\cdot$  en la base  $B$  entonces

$$v \cdot w = v_B A \overline{v_B}^T.$$

Lo que explicamos en el Ejemplo 7.1.2.(2) junto con el cálculo anterior demuestra la siguiente

**Proposición 7.1.3** *Sea  $V$  un espacio vectorial sobre  $\mathbb{K}$  de dimensión  $n$  y sea  $B$  una base de  $V$ . Sea  $A \in M_n(\mathbb{K})$  y definimos el siguiente producto en  $V$ :*

$$v \cdot w = v_B A \overline{w_B}^T.$$

*Entonces  $\cdot$  es un producto hermítico en  $V$  si y solo si  $A$  es una matriz hermítica definida positiva. En tal caso  $A$  es la matriz asociada a  $\cdot$  en la base  $B$ .*

*Recíprocamente todo producto hermítico en  $V$  está dado de esta forma para una matriz hermítica definida positiva.*

Aplicando la Proposición anterior con la base canónica de  $K^n$  tenemos el siguiente

**Corolario 7.1.4** *Todo producto hermítico en  $\mathbb{K}^n$  viene dado por la siguiente fórmula*

$$x \cdot y = x A \overline{y}^T$$

*para alguna matriz hermítica definida positiva  $A$ . En tal caso  $A$  es la matriz asociada a dicho producto en la base canónica.*

Obsérvenos que en el caso en que  $\mathbb{K} = \mathbb{R}$  que la matriz  $A$  sea hermítica es lo mismo que que sea simétrica. Por otro lado la propia definición nos proporciona un método obvio como decidir si una matriz es hermítica pues solo hay que mirar sus entradas. Sin embargo, la definición de definida positiva no muestra una manera clara de cómo decidir si una matriz cuadrada arbitraria es definida positiva. La siguiente definición nos proporciona dicho método para el caso de matrices simétricas reales pero su demostración supera el nivel de esta asignatura:

**Proposición 7.1.5** *Sea  $A \in M_n(\mathbb{R})$  una matriz simétrica. Para cada  $i = 1, \dots, n$  sea  $\Delta_i$  el menor de  $A$  que corta a las primeras  $i$  filas y las primeras  $i$  columnas. Entonces  $A$  es definida positiva si y solo si los menores  $\Delta_1, \dots, \Delta_n$  son todos positivos.*

**Ejemplo 7.1.6** De las siguientes matrices simétricas

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

solo las dos últimas son definidas positivas pues para la primera  $\Delta_1 = 1$  pero  $\Delta_2 = -3$ ; para la segunda, también  $\Delta_1 = 1$ , pero  $\Delta_2 = 0$ ; para la tercera  $\Delta_1 = \Delta_2 = 1$ , pero  $\Delta_3 = -1$ ; para la cuarta  $\Delta_1 = \Delta_2 = \Delta_3 = 1$ ; y para la última  $\Delta_1 = 1$ ,  $\Delta_2 = 3$  y  $\Delta_3 = 1$ .

## 7.2 Norma y ángulos

Sea  $\cdot$  un producto hermítico en un espacio vectorial  $V$  sobre  $\mathbb{K}$ . Para cada  $v \in V$  definimos la *norma* de  $v$  con respecto al producto escalar dado por la siguiente fórmula:

$$\|v\| = +\sqrt{v \cdot v}.$$

Decimos que  $v$  es *unitario* si  $\|v\| = 1$ .

**Proposición 7.2.1** *La norma en un espacio hermítico  $V$  verifica las siguientes propiedades para todo  $u, v \in V$ .*

- (1)  $\|v\| \geq 0$  y la igualdad se verifica si y solo si  $v = 0$ .
- (2)  $\|\alpha v\| = |\alpha| \|v\|$ .
- (3)  $\|u + v\|^2 = \|u\|^2 + \|v\|^2 + 2\operatorname{Re}(u \cdot v)$  y  $\|u - v\|^2 = \|u\|^2 + \|v\|^2 - 2\operatorname{Re}(u \cdot v)$ .  
En el caso en que  $\mathbb{K} = \mathbb{R}$  esto toma la forma  $\|u + v\|^2 = \|u\|^2 + \|v\|^2 + 2(u \cdot v)$  y  $\|u - v\|^2 = \|u\|^2 + \|v\|^2 - 2(u \cdot v)$ .
- (4)  $|u \cdot v| \leq \|u\| \|v\|$  y la igualdad se verifica si y solo si  $\dim \langle u, v \rangle \leq 1$ . (*Desigualdad de Schwarz*).
- (5)  $\|u + v\| \leq \|u\| + \|v\|$  (*Desigualdad Triangular*).

**Demostración.** (1) es consecuencia inmediata de (EH4) y (7.1).

(2) Usando (EH2) y (7.3) tenemos  $\|\alpha v\|^2 = (\alpha v) \cdot (\alpha v) = \alpha \bar{\alpha} (v \cdot v) = |\alpha|^2 \|v\|^2$  de donde se deduce la propiedad tomando raíces cuadradas.

(3) Usando (EH1), (EH3) y (7.2) y la Proposición 1.2.1.(3) tenemos  $\|u + v\|^2 = (u \pm v) \cdot (u \pm v) = (u \cdot u) \pm (v \cdot v) \pm (u \cdot v) + (v \cdot u) = \|u\|^2 + \|v\|^2 + u \cdot v + \bar{v} \cdot \bar{u} = \|u\|^2 + \|v\|^2 + 2\operatorname{Re}(u \cdot v)$ . La demostración con signo menos se hace igual.

(4) Si  $v = 0$  entonces los dos lados de la desigualdad son 0. Si  $v \neq 0$  entonces  $\|v\| > 0$  y

$$\begin{aligned} 0 &\leq \left\| u - \frac{u \cdot v}{\|v\|^2} v \right\|^2 = \left( u - \frac{u \cdot v}{\|v\|^2} v \right) \cdot \left( u - \frac{u \cdot v}{\|v\|^2} v \right) \\ &= u \cdot u - u \cdot \left( \frac{u \cdot v}{\|v\|^2} v \right) - \left( \frac{u \cdot v}{\|v\|^2} v \right) \cdot u + \left( \frac{u \cdot v}{\|v\|^2} v \right) \cdot \left( \frac{u \cdot v}{\|v\|^2} v \right) \\ &= \|u\|^2 - \frac{\bar{u} \cdot \bar{v}}{\|v\|^2} (u \cdot v) - \frac{u \cdot v}{\|v\|^2} (v \cdot u) + \frac{u \cdot v}{\|v\|^2} \frac{\bar{u} \cdot \bar{v}}{\|v\|^2} (v \cdot v) \\ &= \|u\|^2 - 2 \frac{(u \cdot v) \bar{u} \cdot \bar{v}}{\|v\|^2} + \frac{(u \cdot v) \bar{u} \cdot \bar{v}}{\|v\|^4} \|v\|^2 = \|u\|^2 - \frac{|u \cdot v|^2}{\|v\|^2} \end{aligned}$$

Luego  $\|u\|^2 \geq \frac{|u \cdot v|^2}{\|v\|^2}$  y multiplicando por  $\|v\|^2$  y tomando raíces cuadradas tenemos la desigualdad deseada. Además demuestra que se verifica la igualdad si y solo si  $u = \frac{u \cdot v}{\|v\|^2} v$ .

Vamos a demostrar ahora que la igualdad se verifica si y solo si  $\dim \langle u, v \rangle \leq 1$ . Ya hemos visto que si  $v = 0$  entonces  $u \cdot v = \|v\| = 0$  con lo que en este caso efectivamente se da la

igualdad y  $\dim \langle u, v \rangle = \dim \langle u \rangle \leq 1$ . Supongamos que  $v \neq 0$ . Entonces  $\dim \langle u, v \rangle = 1$  si y solo si  $u \in \langle v \rangle$ . Por tanto tenemos que demostrar que  $|u \cdot v| = \|u\| \|v\|$  si y solo si  $u \in \langle v \rangle$ . Ya hemos visto que si  $|u \cdot v| = \|u\| \|v\|$  entonces  $u = \frac{u \cdot v}{\|v\|^2} v \in \langle v \rangle$ . Lo que demuestra una implicación. Recíprocamente, si  $u \in \langle v \rangle$  entonces existe un escalar  $\alpha$  tal que  $u = \alpha v$  y por tanto  $|u \cdot v| = |(\alpha v) \cdot v| = |\alpha| |v \cdot v| = |\alpha| \|v\|^2 = \|u\| \|v\|$ . Esto demuestra la otra implicación.

(5) Obsérvese que si  $z = a + bi$  es un número complejo entonces  $\operatorname{Re}(z) = a \leq \sqrt{a^2 + b^2} = |z|$  y por tanto, usando (3) y la Desigualdad de Schawrtz tenemos que

$$\begin{aligned} \|u + v\|^2 &= \|u\|^2 + \|v\|^2 + 2\operatorname{Re}(u \cdot v) \leq \|u\|^2 + \|v\|^2 + 2|u \cdot v| \\ &\leq \|u\|^2 + \|v\|^2 + 2\|u\| \|v\| = (\|u\| + \|v\|)^2 \end{aligned}$$

y tomando raíces cuadradas obtenemos la desigualdad deseada. ■

Si  $v, w \in V$  entonces la distancia entre  $v$  y  $w$  es

$$d(v, w) = \|v - w\|.$$

La siguiente proposición reúne las propiedades fundamentales de la distancia.

**Proposición 7.2.2** *Las siguientes condiciones se verifican para elementos  $u, v, w$  de un espacio hermítico.*

- (1)  $d(u, v) \geq 0$  y la igualdad se verifica si y solo si  $u = v$ .
- (2)  $d(u, v) = d(v, u)$ .
- (3)  $d(u, w) \leq d(u, v) + d(v, w)$  (Desigualdad triangular).

**Demostración.** (1) es consecuencia inmediata de la Proposición 7.2.1.(1).

(2) Aplicando la Proposición 7.2.1.(1) se tiene que  $d(u, v) = \|u - v\| = \|(-1)(v - u)\| = |1| \|v - u\| = d(v, u)$ .

(3) Aplicando la Proposición 7.2.1.(5) tenemos  $d(u, w) = \|u - w\| = \|(u - v) + (v - w)\| \leq \|u - v\| + \|v - w\| = d(u, v) + d(v, w)$ . ■

En el resto de la sección  $V$  es un espacio euclídeo, o sea un espacio vectorial sobre  $\mathbb{R}$  con un producto euclídeo. Por la Desigualdad de Schwarz tenemos que

$$-1 \leq \frac{u \cdot v}{\|u\| \|v\|} \leq 1$$

lo que implica que existe un único ángulo  $0 \leq \theta \leq \pi$  tal que

$$\cos \theta = \frac{u \cdot v}{\|u\| \|v\|}.$$

En tal caso decimos que  $\theta$  es el ángulo que forman  $u$  y  $v$  y lo denotamos  $\theta = \widehat{u, v}$ . O sea, por definición el *ángulo* que forman dos elementos  $u$  y  $v$  de un espacio hermítico es el único ángulo  $0 \leq \widehat{u, v} \leq \pi$  para el que

$$u \cdot v = \|u\| \|v\| \cos(\widehat{u, v})$$

**Ejemplos 7.2.3** (1) La norma en  $\mathbb{R}^n$  con respecto al producto estándar viene dada por la fórmula habitual:

$$\|(a_1, \dots, a_n)\| = +\sqrt{a_1^2 + \dots + a_n^2}.$$

En consecuencia la distancia viene dada por

$$d((a_1, \dots, a_n), (b_1, \dots, b_n)) = +\sqrt{(a_1 - b_1)^2 + \dots + (a_n - b_n)^2}.$$

En el espacio hermitico en  $\mathbb{C}^n$  con respecto al producto hermitico estándar la norma y la distancia vienen dadas por

$$\|(a_1, \dots, a_n)\| = +\sqrt{|a_1|^2 + \dots + |a_n|^2}$$

y

$$d((a_1, \dots, a_n), (b_1, \dots, b_n)) = +\sqrt{|a_1 - b_1|^2 + \dots + |a_n - b_n|^2}.$$

En ambos casos la Desigualdad de Schwartz toma la siguiente forma:

$$|a_1 b_1 + \dots + a_n b_n|^2 \leq (a_1^2 + \dots + a_n^2)(b_1^2 + \dots + b_n^2).$$

y la desigualdad triangular toma la siguiente forma:

$$\sqrt{|a_1 + b_1|^2 + \dots + |a_n + b_n|^2} = \sqrt{|a_1|^2 + \dots + |a_n|^2} + \sqrt{|b_1|^2 + \dots + |b_n|^2}.$$

(2) La norma en  $\mathcal{C}([a, b])$  viene dada por

$$\|f\| = +\sqrt{\int_a^b f(x)^2 dx},$$

la distancia entre  $f$  y  $g$  es

$$d(f, g) = +\sqrt{\int_a^b (f(x) - g(x))^2 dx}$$

lo cual es el área bajo la curva  $y = (f(x) - g(x))^2$  en el intervalo  $[a, b]$ . Ver la Figura 7.1. Las desigualdades de Schwartz y triangular toman las formas

$$\left| \int_a^b f(x) \overline{g(x)} dx \right|^2 \leq \left( \int_a^b |f(x)|^2 dx \right) \left( \int_a^b |g(x)|^2 dx \right)$$

y

$$\left( \int_a^b f(x) \overline{g(x)} ds \right)^2 \leq \left( \int_a^b f(x)^2 dx \right)^2 + \left( \int_a^b g(x)^2 dx \right)^2.$$

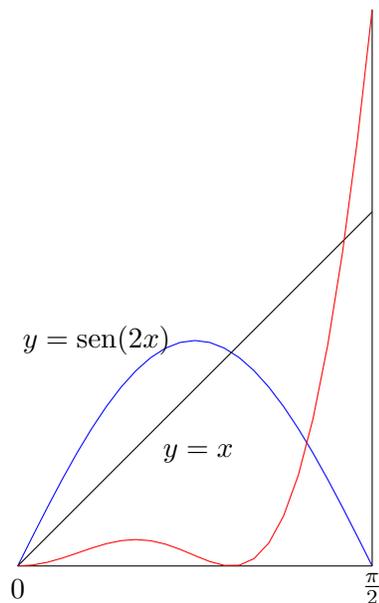


Figure 7.1: La distancia entre  $y = \text{sen}(2x)$  (en azul) y  $y = x$  (en negro), en el intervalo  $[0, \frac{\pi}{2}]$  es la raíz cuadrada del área bajo la curva  $y = (x - \text{sen}(2x))^2$  (en rojo).

### 7.3 Ortogonalidad

En esta sección  $V$  es un espacio hermítico.

Dos elementos de  $u$  y  $v$  de  $V$  decimos que son *ortogonales* si  $u \cdot v = 0$ .

Obsérvese que si se trata de un espacio euclídeo entonces  $u$  y  $v$  son ortogonales si y solo si  $\widehat{u, v} = \frac{\pi}{2}$ . Sin embargo en un espacio hermítico arbitrario no tenemos definido el concepto de ángulo.

Por la Proposición 7.2.1.(3) se tiene que si  $u$  y  $v$  son ortogonales entonces

$$\|u + v\|^2 = \|u\|^2 + \|v\|^2$$

lo cual es lo que conocemos como *Teorema de Pitágoras*. En el caso en que  $\mathbb{K} = \mathbb{R}$  se verifica el recíproco en resumen:

**Teorema 7.3.1 (Teorema de Pitágoras)** *Si  $u$  y  $v$  son elementos ortogonales de un espacio hermítico entonces*

$$\|u + v\|^2 = \|u\|^2 + \|v\|^2.$$

*Si  $u$  y  $v$  son elementos de un espacio euclídeo entonces  $u$  y  $v$  son ortogonales si y solo si  $\|u + v\|^2 = \|u\|^2 + \|v\|^2$ .*

Una lista  $v_1, v_2, \dots$  de elementos de  $V$  se dice que es *ortogonal* si cada para todo  $i \neq j$  se tiene que  $v_i$  y  $v_j$  son ortogonales. Si además los elementos son unitarios se dice que es

un conjunto *ortonormal*. Si  $v \neq 0$  entonces  $\frac{v}{\|v\|}$  es unitario con lo que si  $X$  es un conjunto ortogonal entonces  $\left\{ \frac{x}{\|x\|} : x \in X \right\}$  es ortonormal.

**Ejemplo 7.3.2** Consideremos las funciones  $f_n(x) = \cos nx$  con  $n$  un entero positivo. Vamos a ver que forman una familia ortogonal de elementos del espacio euclídeo  $\mathcal{C}([0, \pi])$ . Tenemos que ver que si  $n \neq m$  entonces la siguiente integral es 0:

$$f_n \cdot f_m = \int_0^\pi \cos(nx) \cos(mx) dx$$

Para calcular la integral usamos las fórmulas del coseno de la suma y la resta:

$$\begin{aligned} \cos(\alpha + \beta) &= \cos \alpha \cos \beta - \operatorname{sen} \alpha \operatorname{sen} \beta \\ \cos(\alpha - \beta) &= \cos \alpha \cos \beta + \operatorname{sen} \alpha \operatorname{sen} \beta \end{aligned}$$

Sumando estas dos igualdades obtenemos

$$\cos \alpha \cos \beta = \frac{\cos(\alpha + \beta) + \cos(\alpha - \beta)}{2}.$$

Por tanto, si  $n \neq m$  tenemos

$$\begin{aligned} \int_0^\pi \cos(nx) \cos(mx) dx &= \frac{1}{2} \int_0^\pi (\cos((n+m)x) + \cos((n-m)x)) \\ &= \frac{1}{2} \left[ \frac{\operatorname{sen}((n+m)x)}{n+m} + \frac{\operatorname{sen}((n-m)x)}{n-m} \right]_0^\pi = 0. \end{aligned}$$

Pues como  $n$  y  $m$  son números enteros, tenemos  $\operatorname{sen}((n+m)\pi) = \operatorname{sen}((n-m)\pi) = 0$ .

Ahora calculamos la norma de cada una de estas funciones:

$$\|f_n\|^2 = \int_0^\pi \cos^2(nx) dx = \frac{1}{2} \int_0^\pi (\cos(2nx) + 1) = \frac{1}{2} \left[ \frac{\operatorname{sen}(2nx)}{2n} + x \right]_0^\pi = \pi.$$

Por tanto tenemos la siguiente lista ortonormal en  $\mathcal{C}([0, 2\pi])$ :

$$\frac{1}{\sqrt{\pi}} \cos(nx) \quad (n \in \mathbb{Z}^{>0} = \{n \in \mathbb{Z} : n > 0\}).$$

**Proposición 7.3.3** *Toda list ortogonal de vectores no nulos es linealmente independiente.*

**Demostración.** Supongamos que  $v_1, \dots, v_n$  es una lista ortogonal de vectores no nulos y supongamos que  $\alpha_1 v_1 + \dots + \alpha_n v_n = 0$ . Entonces

$$0 = v_i \cdot (\alpha_1 v_1 + \dots + \alpha_n v_n) = \alpha_1 v_i \cdot v_1 + \dots + \alpha_i v_i \cdot v_i + \dots + \alpha_n v_i \cdot v_n = \alpha_i \|v_i\|^2$$

Como  $v_i \neq 0$  tenemos que  $\|v_i\| \neq 0$  y por tanto  $\alpha_i = 0$ . ■

**Corolario 7.3.4** *Toda lista ortonormal es linealmente independiente.*

La siguiente proposición nos da una forma de calcular las coordenadas en una base ortogonal.

**Proposición 7.3.5** *Si  $B = \{v_1, \dots, v_n\}$  es una base ortogonal de un espacio hermítico  $V$  entonces para todo  $v \in V$  se verifica*

$$v = \frac{v \cdot v_1}{\|v_1\|^2} v_1 + \dots + \frac{v \cdot v_n}{\|v_n\|^2} v_n$$

o lo que es lo mismo

$$v_B = \left( \frac{v \cdot v_1}{\|v_1\|^2}, \dots, \frac{v \cdot v_n}{\|v_n\|^2} \right)$$

En particular si  $B$  es ortonormal entonces

$$v = (v \cdot v_1)v_1 + \dots + (v \cdot v_n)v_n,$$

o sea,

$$v_B = (v \cdot v_1, \dots, v \cdot v_n).$$

**Demostración.** Supongamos que  $v_B = (a_1, \dots, a_n)$ , o sea  $v = a_1 v_1 + \dots + a_n v_n$ . Entonces

$$(v, v_i) = (a_1 v_1 + \dots + a_n v_n, v_i) = a_1 (v_1 \cdot v_i) + \dots + a_i (v_i \cdot v_i) + \dots + a_n (v_n \cdot v_i) = a_i \|v_i\|^2.$$

Despejando  $a_i$  en la última igualdad obtenemos el resultado deseado. ■

**Corolario 7.3.6** *Si  $B = \{v_1, \dots, v_n\}$  es una base ortonormal de un espacio euclídeo  $V$  y  $v \in V$  entonces*

$$v_B = (\|v\| \cos(\widehat{v, v_1}), \dots, \|v\| \cos(\widehat{v, v_n}))$$

**Demostración.** Es consecuencia de la Proposición 7.3.5 pues

$$v \cdot v_i = \|v\| \|v_i\| \cos(\widehat{v, v_i}) = \|v\| \cos(\widehat{v, v_i}).$$

■

**Corolario 7.3.7** *Si  $B$  es una base ortonormal de un espacio hermítico  $V$  entonces para cada  $u, v \in V$  se tiene*

$$u \cdot v = u_B \overline{v_B}^T.$$

**Demostración.** Pongamos  $B = \{v_1, \dots, v_n\}$ ,  $u_B = (a_1, \dots, a_n)$  y  $v_B = (b_1, \dots, b_n)$ . Entonces

$$u \cdot v = \left( \sum_{i=1}^n a_i v_i \right) \cdot \left( \sum_{i=1}^n b_i v_i \right) = \sum_{1 \leq i, j \leq n} a_i \overline{b_j} v_i \cdot v_j = \sum_{i=1}^n a_i \overline{b_i} = u_B \overline{v_B}^T.$$

■

Obsérvese que el Corolario 7.3.7 nos da las siguientes fórmulas para  $B$  una base ortonormal donde en los términos de la derecha de las igualdades el producto hermítico y la norma se refieren a la del Ejemplo 7.1.2.(1):

$$u \cdot v = u_B \cdot v_B \quad \text{y} \quad \|u\| = \|u_B\|.$$

**Ejemplo 7.3.8** Consideremos un producto hermítico en  $\mathbb{C}^3$  en el que los vectores  $(1, 1 + i, 1)$ ,  $(i, 0, -i)$ ,  $(0, 1, 1)$  forman una base ortonormal. Vamos a obtener la expresión de dicho producto.

En primer lugar observamos que esto es posible pues los vectores dados son linealmente independientes ya que si  $A$  es la matriz formada por los vectores puestos en columna entonces

$$|A| = \begin{vmatrix} 1 & 1+i & 1 \\ i & 0 & -i \\ 0 & 1 & 1 \end{vmatrix} = i \begin{vmatrix} 1 & 1+i & 1 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \end{vmatrix} = i \begin{vmatrix} 1 & 1+i & 2 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{vmatrix} = i(1-i) = 1+i \neq 0.$$

Ahora calculamos las coordenadas  $v_B$  de un vector arbitrario  $v = (x, y, z)$  en esta base. Para ello usamos la fórmula de cambio de base que nos dice que  $v = Av_B$ , o lo que es lo mismo  $x_B = A^{-1}x$ . Calculamos la inversa de  $A$ :

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

Por tanto, interpretando  $x_B$  e  $y_B$  como vectores fila, pero poniéndolo en columna para que quepa en espacio del texto tenemos:

$$\begin{aligned} x_B^T &= \frac{1}{2} \begin{pmatrix} 1+i & -1-i & -2i \\ -1-i & 1-i & 2(1+i) \\ 1+i & i-1 & i-2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} (1+i)x_1 - (1+i)x_2 - 2ix_3 \\ -(1+i)x_1 + (1-i)x_2 + 2(1+i)x_3 \\ (1+i)x_1 + (i-1)x_2 + (i-2)x_3 \end{pmatrix} \end{aligned}$$

Aplicando el Corolario 7.3.7 si  $x = (x_1, x_2, x_3)$  e  $y = (y_1, y_2, y_3)$  tenemos que  $x \cdot y$  es el producto hermítico estándar de  $x_B$  por  $y_B$ ,

$$x_B^T = \frac{1}{2} \begin{pmatrix} (1+i)x_1 - (1+i)x_2 - 2ix_3 \\ -(1+i)x_1 + (1-i)x_2 + 2(1+i)x_3 \\ (1+i)x_1 + (i-1)x_2 + (i-2)x_3 \end{pmatrix}$$

e

$$y_B^T = \frac{1}{2} \begin{pmatrix} (1+i)y_1 - (1+i)y_2 - 2iy_3 \\ -(1+i)y_1 + (1-i)y_2 + 2(1+i)y_3 \\ (1+i)y_1 + (i-1)y_2 + (i-2)y_3 \end{pmatrix}$$

Luego

$$\begin{aligned}
x \cdot y &= x_B \cdot \overline{y_B}^T \\
&= \frac{1}{4} [((1+i)x_1 - (1+i)x_2 - 2ix_3)((1+i)y_1 - (1+i)y_2 - 2iy_3) + \\
&\quad (- (1+i)x_1 + (1-i)x_2 + 2(1+i)x_3)(- (1+i)y_1 + (1-i)y_2 + 2(1+i)y_3) + \\
&\quad ((1+i)x_1 + (i-1)x_2 + (i-2)x_3)((1+i)y_1 + (i-1)y_2 + (i-2)y_3)] \\
&= \frac{1}{4} [((1+i)(1-i) + (-1-i)(-1+i) + (1+i)(1-i))x_1\overline{y_1} + \\
&\quad ((1+i)(-1+i) + (-1-i)(1+i) + (1+i)(-1-i))x_1\overline{y_2} + \\
&\quad ((1+i)2i + (-1-i)2(1-i) + (1+i)(-2-i))x_1\overline{y_3} + \\
&\quad ((-1-i)(1-i) + (1-i)(-1+i) + (i-1)(1-i))x_2\overline{y_1} + \\
&\quad ((-1-i)(-1+i) + (1-i)(1+i) + (i-1)(-1-i))x_2\overline{y_2} + \\
&\quad ((-1-i)2i + (1-i)2(1-i) + (i-1)(-2-i))x_2\overline{y_3} + \\
&\quad (-2i(1-i) + 2(1+i)(-1+i) + (i-2)(1-i))x_3\overline{y_1} + \\
&\quad (-2i(-1+i) + 2(1+i)(1+i) + (i-2)(-1-i))x_3\overline{y_2} + \\
&\quad (-2i2i + 2(1+i)2(1-i) + (i-2)(-2-i))x_3\overline{y_3}] \\
&= \frac{1}{4} [6x_1\overline{y_1} - 2(1+2i)x_1\overline{y_2} + (-7-i)x_1\overline{y_3} + \\
&\quad 2(-1+2i)x_2\overline{y_1} + 6x_2\overline{y_2} + (5-7i)x_2\overline{y_3} + \\
&\quad (-7+i)x_3\overline{y_1} + (5+7i)x_3\overline{y_2} + 17x_3\overline{y_3}]
\end{aligned}$$

Luego

$$(x_1, x_2, x_3) \cdot (y_1, y_2, y_3) = \frac{1}{4} (x_1, x_2, x_3) \begin{pmatrix} 6 & -2-4i & -7-7i \\ -2-4i & 6 & 5-7i \\ -7+i & 5+7i & 17 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}.$$

O sea se trata del producto hermítico del Ejemplo 7.1.2.(2) para la matriz

$$\frac{1}{4} \begin{pmatrix} 6 & -2-4i & -7-7i \\ -2-4i & 6 & 5-7i \\ -7+i & 5+7i & 17 \end{pmatrix}.$$

Una *matriz* cuadrada  $A$  se dice que es *ortogonal* si su transpuesta es su inversa es decir si  $A^T A = I$ . Si  $A$  tiene entradas complejas decimos que  $A$  es unitaria si su transpuesta conjugada es su inversa, o sea si  $\overline{A}^T A = I$ , o equivalentemente si  $A^T \overline{A} = I$ . En particular, si  $A$  tiene entradas reales  $A$  es ortogonal si y solo si es unitaria.

**Proposición 7.3.9** *Si  $B$  es una base ortonormal de un espacio hermítico y  $B'$  es otra base entonces la matriz  $A$  de cambio de base  $B'_B$  es unitaria si y solo si  $B'$  es una base ortonormal. En tal caso  $|\det(A)| = 1$ .*

**Demostración.** Las columnas de  $A$ , que son también las filas de  $A^T$ , están formadas por las coordenadas de  $B'$  en la base  $B$ . Usando esto junto con el Corolario 7.3.7 tenemos que

$$A^T \overline{A} = (v_{iB} \cdot \overline{(v_j)_B}^T)_{i,j} = (v_i \cdot v_j)_{i,j}$$

Esto demuestra la primera parte del corolario. Por otro lado, aplicando las propiedades de los determinantes tenemos que  $1 = \det(A) \det(\overline{A})^T = \det(A) \overline{\det(A)} = \|\det(A)\|^2$  con lo que  $\|\det(A)\| = 1$ . ■

**Corolario 7.3.10** Si  $B$  es una base ortonormal de un espacio euclídeo y  $B'$  es otra base entonces la matriz  $A$  de cambio de base  $B'_B$  es ortogonal si y solo si  $B'$  es una base ortonormal. En tal caso,  $\det(A) = \pm 1$ .

Decimos que dos bases ortonormales  $B$  y  $B'$  de un espacio euclídeo tienen la misma *orientación* si  $\det(B'_B) = 1$ . En caso contrario, es decir si  $\det(B'_B) = -1$ , decimos que tienen diferente orientación.

**Ejemplo 7.3.11** Sea  $B = \{v_1, v_2\}$  la base de  $\mathbb{R}^2$  y sea  $B' = \{w_1, w_2\}$  la base que se obtiene al girar los vectores de  $B$  un ángulo de  $\alpha$  grados alrededor del origen. Entonces

$$v_1 = v_1 \cos \alpha + v_2 \sin \alpha, \quad v_2 = -v_1 \sin \alpha + v_2 \cos \alpha).$$

Con lo que

$$B'_B = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}.$$

es una matriz ortogonal de determinante 1. Por tanto los giros alrededor del origen no cambian la orientación.

Supongamos ahora los vectores de  $B'$  son lo que se obtienen al reflejar los elementos de  $B$  a través de una recta. Supongamos que la recta tiene por dirección el vector  $v$  y sea  $u$  un vector no nulo ortogonal a  $v$ , o sea un generado de la recta ortogonal a la recta original. Entonces  $B_1 = \{u, v\}$  es otra base. La matriz asociada a  $R$  en esta base es

$$R_{B_1} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

que tiene determinante  $-1$ . Por otro lado los vectores  $w_1$  y  $w_2$  son las columnas de la matriz asociada a  $R$  en la base  $B$ , o sea  $B'_B = R_B$ . Pero  $R_B$  y  $R_{B_1}$  tienen el mismo determinante. Por tanto  $\det(B'_B) = -1$ . Por tanto las reflexiones en rectas del plano cambian la orientación.

Esto se puede generalizar a giros alrededor de rectas arbitrarias y reflexiones en hiperplanos, o sea subespacios de dimensión  $n - 1$ . La matriz asociada a un giro en una base adecuada es de la forma

$$\left( \begin{array}{cc|c} \cos \alpha & -\sin \alpha & \\ \sin \alpha & \cos \alpha & \\ \hline & & I_{n-2} \end{array} \right)$$

que tiene determinante 1 y la matriz asociada a una reflexión en un hiperplano en una base adecuada tiene la forma

$$\left( \begin{array}{c|c} I_{n-1} & \\ \hline & -1 \end{array} \right)$$

que tiene determinante  $-1$ . Por tanto los giros mantienen la orientación y las reflexiones la cambian.

## 7.4 Método de Gram-Schmidt

Vamos a ver ahora que todo espacio hermitico de dimensión finita tiene una base ortonormal. De hecho veremos como construir una a partir de una base arbitraria. Ya hemos visto cómo construir una base ortonormal a partir de una ortogonal, simplemente dividiendo cada vector de la base por su norma. Por tanto lo que ahora necesitamos es conseguir una ortogonal a partir de una base arbitraria  $\{v_1, \dots, v_n\}$ . Esto se llama el *Método de Gram-Schmidt* y es lo que hace la siguiente

**Proposición 7.4.1** *Sea  $v_1, \dots, v_n$  una lista linealmente independiente formado por elementos distintos de 0. Construimos otra lista  $w_1, \dots, w_n$  poniendo*

$$\begin{aligned} w_1 &= v_1, \\ w_{i+1} &= v_{i+1} - \sum_{k=1}^i \frac{v_{i+1} \cdot w_k}{\|w_k\|^2} w_k. \end{aligned}$$

*Entonces cada  $w_k \neq 0$  (lo que permite construir  $w_{i+1}$  a partir de  $v_1, \dots, v_i$ ),  $w_1, \dots, w_n$  es ortogonal y  $\langle v_1, \dots, v_n \rangle = \langle w_1, \dots, w_n \rangle$ .*

**Demostración.** Razonamos por inducción en  $i$ . Está claro que  $w_1 = v_1 \neq 0$  y  $\langle w_1 \rangle = \langle v_1 \rangle$  y por supuesto  $\{w_1\}$  es ortogonal.

Supongamos que ya hemos demostrado que  $w_1, \dots, w_i$  es ortogonal y está formado por elementos distintos de 0 y que  $\langle w_1, \dots, w_i \rangle = \langle v_1, \dots, v_n \rangle$ . Tenemos que ver que  $w_1, \dots, w_{i+1}$  es ortogonal, está formado por elementos distintos de 0 y  $\langle w_1, \dots, w_{i+1} \rangle = \langle v_1, \dots, v_{i+1} \rangle$ . De la definición de  $w_i$  tenemos que  $v_{i+1} = \sum_{k=1}^i \frac{v_{i+1} \cdot w_k}{\|w_k\|^2} w_k + w_{i+1} \in \langle w_1, \dots, w_i, w_{i+1} \rangle$ . Como también sabemos que  $\langle v_1, \dots, v_i \rangle = \langle w_1, \dots, w_i \rangle$ , de la Proposición 2.3.9.(1) deducimos que  $\langle v_1, \dots, v_{i+1} \rangle \subseteq \langle w_1, \dots, w_{i+1} \rangle$ . Como  $v_1, \dots, v_{i+1}$  son ortogonales y no nulos, de la Proposición 7.3.3, tenemos que son linealmente independientes lo que implica que la dimensión de  $\langle w_1, \dots, w_{i+1} \rangle$  es  $i+1$  y por tanto  $w_{i+1} \neq 0$  y  $\langle v_1, \dots, v_{i+1} \rangle = \langle w_1, \dots, w_{i+1} \rangle$ , por el Corolario 4.1.12.

Como ya sabemos que  $w_1, \dots, w_i$  es ortogonal para demostrar que  $w_1, \dots, w_{i+1}$  es ortogonal basta ver que  $w_{i+1} \neq 0$  y  $w_{i+1} \cdot w_j = 0$  para todo  $j = 1, \dots, i$ . Efectivamente

$$\begin{aligned} w_{i+1} \cdot w_j &= \left( v_{i+1} - \sum_{k=1}^i \frac{v_{i+1} \cdot w_k}{\|w_k\|^2} w_k \right) \cdot w_j = v_{i+1} \cdot w_j - \sum_{k=1}^i \frac{v_{i+1} \cdot w_k}{\|w_k\|^2} (w_k \cdot w_j) \\ &= v_{i+1} \cdot w_j - \frac{v_{i+1} \cdot w_j}{\|w_j\|^2} (w_{i+1} \cdot w_j) = v_{i+1} \cdot w_j - (v_{i+1} \cdot w_j) \frac{\|w_j\|^2}{\|w_j\|^2} = 0. \end{aligned}$$

■

**Observación 7.4.2** Si  $v_1, \dots, v_n$  es una lista como en la Proposición 7.4.1 y  $v_1, \dots, v_i$  es ortogonal entonces en la lista ortogonal  $w_1, \dots, w_n$  del Método de Gram-Schmidt (Proposición 7.4.1), tenemos  $w_j = v_j$  para  $j = 1, \dots, i$ .

**Corolario 7.4.3** Todo espacio hermítico de dimensión finita tiene una base ortonormal. Además toda conjunto ortogonal formado por elementos no nulos está contenido en una base ortogonal y todo conjunto ortonormal está contenido en una base ortonormal.

**Demostración.** Aplicando el Método de Gram-Schmidt a una base obtendremos una base ortogonal y dividiendo sus elementos por su norma obtendremos una base ortonormal.

Sea  $w_1, \dots, w_k$  un conjunto ortogonal formado por elementos distintos de 0. Por la Proposición 7.3.3,  $w_1, \dots, w_k$  es linealmente independiente y por el Teorema 4.1.4, podemos completar esta lista a una base  $w_1, \dots, w_k, v_{k+1}, \dots, v_n$  de  $V$ . Aplicando el Método de Gram-Schmidt obtenemos una base ortogonal de  $V$ , pero como  $w_1, \dots, w_k$  es ortogonal la base obtenida es de la forma  $w_1, \dots, w_k, w_{k+1}, \dots, w_n$ . Si  $w_1, \dots, w_k$  es ortonormal dividiendo los otros  $w_i$  por su norma tenemos una base ortonormal.

■

**Ejemplo 7.4.4** Consideremos el subespacio vectorial de  $\mathbb{R}^4$  formado por los vectores  $(x_1, x_2, x_3, x_4)$  que satisfacen la ecuación  $x_1 + 2x_2 + 3x_3 + 4x_4 = 0$ . Vamos a encontrar una base ortonormal de este espacio con respecto al producto escalar estándar. Para ello comenzamos con una base cualquiera, por ejemplo:

$$v_1 = (2, -1, 0, 0), \quad v_2 = (3, 0, -1, 0), \quad v_3 = (4, 0, 0, -1).$$

Ahora aplicamos el Método de Gram-Schmidt para conseguir una base ortogonal:

$$\begin{aligned} w_1 &= v_1 = (2, -1, 0, 0), \\ w_2 &= v_2 - \frac{v_2 \cdot w_1}{\|w_1\|^2} w_1 = (3, 0, -1, 0) - \frac{6}{5}(2, -1, 0, 0) = \left(\frac{3}{5}, \frac{6}{5}, -1, 0\right), \\ w_3 &= v_3 - \frac{v_3 \cdot w_1}{\|w_1\|^2} w_1 - \frac{v_3 \cdot w_2}{\|w_2\|^2} w_2 = (4, 0, 0, -1) - \frac{8}{5}(2, -1, 0, 0) - \frac{12/5}{14/5} \left(\frac{3}{5}, \frac{6}{5}, -1, 0\right) \\ &= \left(\frac{2}{7}, \frac{4}{7}, \frac{6}{7}, -1\right) \end{aligned}$$

Multiplicando el segundo por 5 y el tercero por 7 conseguimos librarnos de los denominadores sin perder la ortogonalidad:

$$u_1 = (2, -1, 0, 0), \quad u_2 = (3, 6, -5, 0), \quad u_3 = (2, 4, 6, -7).$$

Finalmente dividiendo cada uno por su norma obtenemos la base ortonormal:

$$t_1 = \frac{1}{\sqrt{5}}(2, -1, 0, 0), \quad t_2 = \frac{1}{\sqrt{70}}(3, 6, -5, 0), \quad t_3 = \frac{1}{\sqrt{105}}(2, 4, 6, -7).$$

## 7.5 Subespacio ortogonal

Sea  $V$  un espacio hermítico y sea  $A$  un subconjunto de  $V$ . El siguiente conjunto se llama *subespacio ortogonal* de  $A$  en  $V$ :

$$A^\perp = \{v \in V : v \cdot a = 0 \text{ para todo } a \in A\}.$$

**Proposición 7.5.1** *Si  $V$  es un espacio hermítico y  $A \subseteq V$  entonces  $A^\perp$  es un subespacio vectorial de  $V$ .*

**Demostración.** De (7.1), si  $a \in A$  entonces  $0 \cdot a = 0$ . Esto demuestra que  $0 \in A^\perp$ . Si  $v, w \in A^\perp$  y  $\alpha \in \mathbb{K}$  entonces

$$\begin{aligned} (v+w) \cdot a &= v \cdot a + w \cdot a = 0 + 0 = 0, \\ (\alpha v) \cdot a &= \alpha(v \cdot a) = \alpha 0 = 0. \end{aligned}$$

Usando la Proposición 2.3.5 se deduce que  $A^\perp$  es un subespacio vectorial de  $V$ . ■

**Proposición 7.5.2** *Sean  $A$  y  $B$  subconjuntos de un espacio hermítico. Entonces*

- (1) *Si  $A \subseteq B$  entonces  $B^\perp \subseteq A^\perp$ .*
- (2)  *$A \subseteq (A^\perp)^\perp$ .*
- (3)  *$A^\perp = \langle A \rangle^\perp$ .*

**Demostración.** (1) y (2) es un sencillo ejercicio.

(3) Como  $A \subseteq \langle A \rangle$ , de (1) se deduce que  $\langle A \rangle^\perp \subseteq A^\perp$ . Sea  $v \in A^\perp$  y sea  $w \in \langle A \rangle$ . Entonces  $w = \sum_{i=1}^n \alpha_i a_i$  para ciertos  $a_1, \dots, a_n \in A$  y  $\alpha_1, \dots, \alpha_n \in K$ . Luego

$$w \cdot v = \left( \sum_{i=1}^n \alpha_i a_i \right) \cdot v = \sum_{i=1}^n \alpha_i (a_i \cdot v) = 0$$

pues  $a_i \cdot v = 0$  para todo  $i$ . ■

**Proposición 7.5.3** *Si  $V$  es un espacio hermítico de dimensión finita y  $W$  es un subespacio vectorial de  $V$  entonces*

$$V = W \oplus W^\perp \quad \text{y} \quad (W^\perp)^\perp = W.$$

*En particular*

$$\dim V = \dim W + \dim W^\perp.$$

**Demostración.** Sea  $v \in W \cap W^\perp$ . Entonces  $0 = v \cdot v = \|v\|^2$  y por tanto  $v = 0$ , por la Proposición 7.2.1.1. Esto demuestra que  $W \cap W^\perp = 0$  con lo que  $W + W^\perp = W \oplus W^\perp$ . Para demostrar que  $V = W \oplus W^\perp$ , falta demostrar que esto es igual a  $V$  y para eso basta demostrar que contiene a  $V$ . En efecto, sea  $w_1, \dots, w_k$  una base de ortogonal de  $W$  y completémosla a una base ortogonal  $w_1, \dots, w_k, v_{k+1}, \dots, v_n$  de  $V$ , lo que es posible por el Corolario 7.4.3. Como esta base es ortogonal y aplicando la Proposición 7.5.2.(1) se tiene que  $w_{k+1}, \dots, w_n \in \{w_1, \dots, w_k\}^\perp = \langle w_1, \dots, w_k \rangle^\perp = W^\perp$ . Por tanto  $V = \langle w_1, \dots, w_n \rangle = \langle w_1, \dots, w_k \rangle + \langle w_{k+1}, \dots, w_n \rangle \subseteq W + W^\perp$ , como queríamos demostrar.

De la Proposición 7.5.2.(2) se tiene que  $W \subseteq (W^\perp)^\perp$ . Por tanto  $\dim W \leq \dim(W^\perp)^\perp$ , pero como  $V = W \oplus W^\perp = W^\perp \oplus (W^\perp)^\perp$ , de (4.3) deducimos que  $\dim W = \dim V - \dim W^\perp = \dim(W^\perp)^\perp$  y por tanto  $W = (W^\perp)^\perp$ . ■

Aplicando la Proposición 7.5.3 al subespacio vectorial  $W = A^\perp$  tenemos el siguiente:

**Corolario 7.5.4** *Si  $A$  es un subconjunto de un espacio hermítico entonces  $A^\perp = ((A^\perp)^\perp)^\perp$ .*

**Corolario 7.5.5** *Si  $W_1$  y  $W_2$  son subespacios vectoriales de un espacio hermítico de dimensión finita entonces*

$$(W_1 \cap W_2)^\perp = W_1^\perp + W_2^\perp \quad y \quad (W_1 + W_2)^\perp = W_1^\perp \cap W_2^\perp$$

**Demostración.** Aplicando la Proposición 7.5.2.(1) a las inclusiones  $W_1 \cap W_2 \subseteq W_i \subseteq W_1 + W_2$  tenemos  $(W_1 + W_2)^\perp \subseteq W_i^\perp \subseteq (W_1 \cap W_2)^\perp$  y por tanto

$$(W_1 + W_2)^\perp \subseteq W_1^\perp \cap W_2^\perp \quad y \quad W_1^\perp + W_2^\perp \subseteq (W_1 \cap W_2)^\perp.$$

La segunda inclusión aplicada a  $W_1^\perp$  y  $W_2^\perp$ , junto con la Proposición 7.5.3 nos daría

$$W_1 + W_2 = (W_1^\perp)^\perp + (W_2^\perp)^\perp \subseteq (W_1^\perp \cap W_2^\perp)^\perp$$

Aplicando las Proposiciones 7.5.2.(2) y 7.5.3 con las desigualdades anteriores tenemos que

$$W_1 + W_2 \subseteq (W_1^\perp \cap W_2^\perp)^\perp \subseteq ((W_1 + W_2)^\perp)^\perp = W_1 + W_2$$

Eso implica que los contenidos intermedios son igualdades, es decir

$$((W_1 + W_2)^\perp)^\perp = (W_1^\perp \cap W_2^\perp)^\perp \quad y \quad (W_1^\perp \cap W_2^\perp)^\perp = W_1 + W_2$$

Aplicando otra vez la Proposición 7.5.3 obtenemos que

$$(W_1 + W_2)^\perp = (((W_1 + W_2)^\perp)^\perp)^\perp = ((W_1^\perp \cap W_2^\perp)^\perp)^\perp = W_1^\perp \cap W_2^\perp$$

y

$$W_1^\perp \cap W_2^\perp = ((W_1^\perp \cap W_2^\perp)^\perp)^\perp = (W_1 + W_2)^\perp.$$

■

Sea  $V$  un espacio hermitico y sea  $W$  un subespacio vectorial de  $V$ . Por la Proposición 7.5.3 tenemos  $V = W \oplus W^\perp$ . Llamamos *proyección ortogonal* de  $V$  en  $W$  a la proyección de  $V$  en  $W$  en la dirección de  $W^\perp$ . La denotamos  $P_W$ . O sea

$$P_W = P_{W, W^\perp}.$$

Usando que  $W = (W^\perp)^\perp$  tenemos que

$$v = P_W(v) + P_{W^\perp}(v).$$

La siguiente proposición nos va a dar una fórmula para  $P_W$ :

**Proposición 7.5.6** *Sea  $W$  un subespacio vectorial de un espacio hermitico  $V$  de dimensión finita y sea  $v_1, \dots, v_k$  una base ortogonal de  $W$ . Entonces*

$$P_W(v) = \frac{v \cdot v_1}{\|v_1\|^2} v_1 + \dots + \frac{v \cdot v_k}{\|v_k\|^2} v_k.$$

*En particular si  $v_1, \dots, v_k$  es una base ortonormal de  $W$  entonces*

$$P_W(v) = (v \cdot v_1)v_1 + \dots + (v \cdot v_k)v_k.$$

**Demostración.** Por el Corolario 7.4.3,  $V$  tiene una base ortogonal  $w_1, \dots, w_k, w_{k+1}, \dots, w_n$  y según vimos en la demostración de la Proposición 7.5.3,  $w_{k+1}, \dots, w_n$  es una base de  $W^\perp$ . Aplicando la Proposición 7.3.5 tenemos la siguiente igualdad para todo  $v \in V$ :

$$v = \frac{v \cdot v_1}{\|v_1\|^2} v_1 + \dots + \frac{v \cdot v_n}{\|v_n\|^2} v_n$$

Como la primera parte de la suma  $\frac{v \cdot v_1}{\|v_1\|^2} v_1 + \dots + \frac{v \cdot v_k}{\|v_k\|^2} v_k$  está en  $W$  y la segunda  $\frac{v \cdot v_{k+1}}{\|v_{k+1}\|^2} v_{k+1} + \dots + \frac{v \cdot v_n}{\|v_n\|^2} v_n$  está en  $W^\perp$  deducimos que

$$P_W(v) = \frac{v \cdot v_1}{\|v_1\|^2} v_1 + \dots + \frac{v \cdot v_k}{\|v_k\|^2} v_k.$$

■

La siguiente proposición nos asegura que entre todos los elementos de  $W$ ,  $P_W(v)$  es que está más cerca de  $v$ .

**Proposición 7.5.7** *Sea  $W$  un subespacio vectorial de un espacio vectorial hermitico de dimensión finita  $V$ . Para todo  $v \in V$  y todo  $w \in W$  se tiene que*

$$d(v, w) \geq d(v, P_W(v))$$

*y la igualdad se verifica si y solo si  $w = P_W(v)$ .*

**Demostración.** Sea  $u = P_W(v)$  y  $z = P_{W^\perp}$ . Entonces  $v = u + z$ . Si  $w \in W$  entonces  $v - w = (u - w) + z$ ,  $u - w \in W$  y  $z \in W^\perp$ . Del Teorema de Pitágoras deducimos que

$$d(v, w)^2 = \|v - w\|^2 = \|u - w\|^2 + \|z\|^2 \geq \|z\|^2 = \|v - u\|^2 = d(v, u)^2$$

y por tanto  $d(v, w) \geq d(v, P_W(v))$ . Además la igualdad se verifica si y solo si  $\|u - w\| = 0$  si y solo si  $u = w$ . ■

La Proposición 7.5.7 nos permite definir la distancia de  $v$  a  $W$  como

$$d(v, W) = d(v, P_W(v)).$$

**Ejemplo 7.5.8** Vamos a obtener una fórmula para la proyección ortogonal en el plano  $\pi$  de  $\mathbb{R}^3$  dado por la ecuación

$$x + y - 2z = 0.$$

Comenzamos eligiendo una base de  $\pi$ :  $v_1 = (2, 0, 1)$ ,  $v_2 = (0, 2, 1)$ . A partir de ella construimos una base ortogonal aplicamos el Método de Gram-Schmidt:  $w_1 = v_1$  y

$$w_2 = v_2 - \frac{v_2 \cdot v_1}{\|v_1\|^2} v_1 = (0, 2, 1) - \frac{1}{5}(2, 0, 1) = \left(-\frac{2}{5}, 2, \frac{4}{5}\right).$$

Aplicando la fórmula de la Proposición 7.5.7 tenemos que

$$\begin{aligned} P_\pi(x, y, z) &= \frac{(x, y, z) \cdot (2, 0, 1)}{5} (2, 0, 1) + \frac{(x, y, z) \cdot \left(-\frac{2}{5}, 2, \frac{4}{5}\right)}{14/5} \left(-\frac{2}{5}, 2, \frac{4}{5}\right) \\ &= \frac{2x + z}{5} (2, 0, 1) + \frac{-2x + 10y + 4z}{14} \left(-\frac{2}{5}, 2, \frac{4}{5}\right) \\ &= \frac{1}{7} (6x - 2y + 2z, -2x + 10y + 4z, 2x + 4y - 3z). \end{aligned}$$

Por ejemplo, la proyección del punto  $(1, 0, -1)$  en  $\pi$  es

$$P_\pi(1, 0, -1) = \frac{1}{7} (4, -6, -1).$$

Por tanto

$$\begin{aligned} d((1, 0, -1), \pi) &= d\left((1, 0, 1), \frac{1}{7}(4, -6, -1)\right) = \sqrt{\left(1 - \frac{4}{7}\right)^2 + \left(\frac{6}{7}\right)^2 + \left(-1 - \frac{-1}{7}\right)^2} \\ &= \frac{1}{7} \sqrt{3^2 + 6^2 + 8^2} = \frac{\sqrt{109}}{7}. \end{aligned}$$

**Ejemplo 7.5.9** Supongamos que queremos calcular una fórmula para la proyección ortogonal en el subespacio vectorial  $W = \langle (0, 2, 0, -1), (1, 2, 4, -1), (0, 1, -1, 2) \rangle$  del espacio euclídeo  $\mathbb{R}^4$  con respecto al producto escalar estándar. En primer lugar calculamos una base ortogonal de

$W$ . Para ello aplicamos el Método de Gram-Schmidt a  $v_1 = (0, 2, 0, -1)$ ,  $v_2 = (1, 2, 0, -6)$ ,  $v_3 = (0, 1, -1, 2)$ :

$$\begin{aligned} w_1 &= v_1 = (0, 2, 0, -1) \\ w_2 &= v_2 - \frac{v_2 \cdot w_1}{\|w_1\|^2} w_1 = (1, 2, 0, -6) - \frac{10}{5}(0, 2, 0, -1) = (1, -2, 0, -4) \\ w_3 &= v_3 - \frac{v_3 \cdot w_1}{\|w_1\|^2} w_1 - \frac{v_3 \cdot w_2}{\|w_2\|^2} w_2 = (0, 1, -1, 2) - \frac{0}{5} w_1 - \frac{-10}{21}(1, -2, 0, -4) \\ &= \frac{1}{21}(10, 1, -21, 2) \end{aligned}$$

En realidad nos podemos librar de los denominadores multiplicando  $w_3$  por 21 lo que no cambiará el que siga siendo una base ortogonal de  $W$ . Por tanto podemos usar la siguiente base ortogonal de  $W$ :

$$u_1 = (0, 2, 0, -1), \quad u_2 = (1, -2, 0, -4) \quad u_3 = (10, 1, -21, 2).$$

Entonces

$$\begin{aligned} P_W \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} &= \frac{(x, y, z, w) \cdot (0, 2, 0, -1)}{5} \begin{pmatrix} 0 \\ 2 \\ 0 \\ -1 \end{pmatrix} + \frac{(x, y, z, w) \cdot (1, -2, 0, -4)}{21} \begin{pmatrix} 1 \\ -2 \\ 0 \\ -4 \end{pmatrix} + \\ &\quad \frac{(x, y, z, w) \cdot (10, 1, -21, 2)}{546} \begin{pmatrix} 10 \\ 1 \\ -21 \\ 2 \end{pmatrix} \\ &= \frac{1}{5 \cdot 546} (546(2y - w)) \begin{pmatrix} 0 \\ 2 \\ 0 \\ -1 \end{pmatrix} + 130(x - 2y - 4w) \begin{pmatrix} 1 \\ -2 \\ 0 \\ -4 \end{pmatrix} \\ &\quad + 5(10x + y - 21z + 2w) \begin{pmatrix} 10 \\ 1 \\ -21 \\ 2 \end{pmatrix} \\ &= \frac{1}{2730} \begin{pmatrix} 630x - 210y - 1050z - 420w \\ -210x + 2752y - 105z - 66w \\ -1050x - 105y + 2205z - 210w \\ -420x - 42y - 210z + 2646w \end{pmatrix}. \end{aligned}$$

**Ejemplo 7.5.10** Vamos ahora a considerar el espacio  $\mathcal{C}([-1, 1])$  para aproximar funciones continuas en el intervalo  $[-1, 1]$  a polinomios de un grado dado. Para fijar ideas vamos a usar polinomios de grado como mucho 2, es decir consideramos el espacio  $W$  de las funciones polinómicas de grado menor o igual que 2 y lo primero que hacemos es buscar una base

ortogonal de  $W$ . Para ello partimos de la base  $f_1 = 1, f_2 = X, f_3 = X^2$  y aplicamos el Método de Gram-Schmidt y comenzamos calculando

$$f_2 \cdot f_1 = \int_1^{-1} x dx = \left[ \frac{x^2}{2} \right]_{-1}^1 = 0.$$

Por tanto  $f_1$  y  $f_2$  son ortogonales y los dos primeros vectores de la base ortogonal son  $w_1 = 1$  y  $w_2$ . Para calcular el tercero  $w_3$  calculamos

$$\|f_1\|^2 = \int_{-1}^1 dx = 2, \quad f_3 \cdot f_1 = \int_{-1}^1 x^2 dx = \left[ \frac{x^3}{3} \right]_{-1}^1 = \frac{2}{3}, \quad f_3 \cdot f_2 = \int_{-1}^1 x^3 dx = \left[ \frac{x^4}{4} \right]_{-1}^1 = 0.$$

De donde deducimos que el tercer vector de la base ortogonal es

$$w_3 = X^2 - \frac{2/3}{2}1 = X^2 - \frac{1}{3}.$$

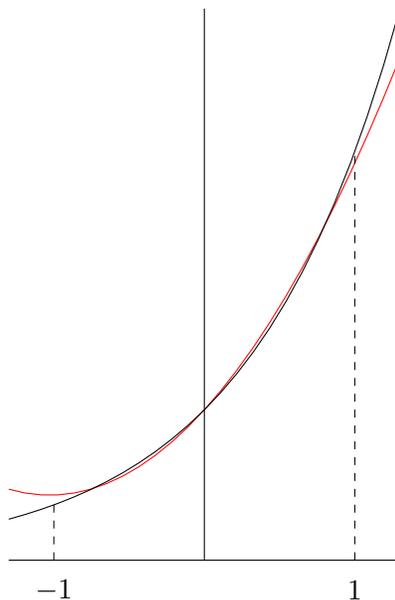


Figure 7.2: Comparación entre la gráfica de  $f(x) = e^x$  (negro) y su aproximación a un polinomio de grado 2,  $P_W(f)$  (rojo) en el intervalo  $[-1, 1]$ .

Ahora calculamos la proyección ortogonal de la función  $f(x) = e^x$ . Según la fórmula de la Proposición 7.5.7 necesitamos calcular

$$f \cdot f_1 = \int_{-1}^1 e^x dx = [e^x]_{-1}^1 = e - e^{-1} = \frac{e^2 - 1}{e}$$

$$f \cdot f_2 = \int_{-1}^1 x e^x dx = [x e^x]_{-1}^1 - \int_{-1}^1 e^x dx = e + e^{-1} - e + e^{-1} = 2e^{-1}$$

y  $f \cdot w_3$ , pero nos resulta más cómodo calcular primero

$$f \cdot f_3 = \int_{-1}^1 x^2 e^x dx = [x^2 e^x]_{-1}^1 - 2 \int_{-1}^1 x e^x dx = e - e^{-1} - 4e^{-1} = e - 5e^{-1} = \frac{e^2 - 5}{e}$$

de donde deducimos que

$$f \cdot w_3 = f \cdot f_3 - \frac{1}{3} f \cdot f_1 = \frac{e^2 - 5}{e} - \frac{e^2 - 1}{3e} = \frac{2e^2 - 14}{3e}.$$

Como ya hemos calculado  $\|f_1\|^2 = 2$  y obviamente  $\|f_2\|^2 = f_3 \cdot f_1 = \frac{2}{3}$  nos faltaría calcular

$$\|w_3\|^2 = \int_{-1}^1 \left(x^2 - \frac{1}{3}\right)^2 dx = \frac{1}{9} \int_{-1}^1 (9x^4 - 6x^2 + 1) dx = \frac{1}{9} \left[ \frac{9x^5}{5} - 2x^3 + x \right]_{-1}^1 = \frac{8}{45}.$$

Por tanto

$$P_W(f) = \frac{e^2 - 1}{2e} + \frac{3}{e}x + \frac{15(e^2 - 7)}{4e}x^2.$$

La Figura 7.2 muestra las gráficas de  $f(x) = e^x$  y  $P_W(x)$  en el intervalo  $[-1, 1]$ .

## 7.6 Aplicación: Ajuste por mínimos cuadrados

Supongamos que tenemos un sistema de ecuaciones lineales incompatible  $AX = b$ . Supongamos que  $A \in M_{m,n}(\mathbb{K})$ ,  $b \in \mathbb{K}^m$  y vamos a considerar en  $\mathbb{K}^n$  el producto hermitico del Ejemplo 7.1.2.(1), o sea

$$(x_1, \dots, x_n) \cdot (y_1, \dots, y_n) = x_1 \overline{y_1} + \dots + x_n \overline{y_n}.$$

El que el sistema sea incompatible quiere decir que no tiene ninguna solución o sea  $Ax \neq b$  para todo  $x \in \mathbb{K}^n$  pero queremos encontrar lo que más se parezca a una solución, o sea buscamos un  $x_0 \in \mathbb{K}^n$  de forma que la distancia entre  $Ax$  y  $b$  sea lo menor posible o sea de forma que  $\|Ax_0 - b\| \leq \|Ax - b\|$  para todo  $x \in \mathbb{K}^n$ . Nos referiremos a esto como la *solución aproximada* al sistema de ecuaciones lineales. Observa que los elementos de la forma  $Ax$  son precisamente los del espacio columna  $C(A)$  de la matriz. Por tanto queremos encontrar  $x_0$  de forma que  $Ax_0$  es el vector de  $C(A)$  más cercano a  $b$ . La Proposición 7.5.7 nos proporciona un método para calcular la solución aproximada pues el elemento de  $C(A)$  más próximo a  $b$  es precisamente  $P_{C(A)}(b)$ , o sea  $Ax_0 = P_{C(A)}(b)$ . Entonces  $b - Ax_0 = b - P_{C(A)}(b) \in C(A)^\perp$  lo que quiere decir que  $b - Ax_0$  es ortogonal a las columnas de  $A$  o lo que es lo mismo que  $A^T \cdot \overline{b - Ax_0} = 0$ , o sea  $\overline{A}^T Ax_0 = \overline{A}^T b$ . Aplicando la Proposición 7.5.7 obtenemos la siguiente

**Proposición 7.6.1** *Si  $AX = b$  es un sistema de ecuaciones lineales incompatible entonces el sistema  $\overline{A}^T AX = \overline{A}^T b$  es compatible y si  $x_0$  es una solución de este sistema entonces  $\|Ax_0 - b\| \leq \|Ax - b\|$  para todo  $x$ .*

En otras palabras las soluciones aproximadas de un sistema de ecuaciones lineales incompatible  $Ax = b$  son las soluciones del sistema de ecuaciones lineales  $\overline{A}^T AX = \overline{A}^T b$ .

**Ejemplo 7.6.2** Consideremos el siguiente sistema de ecuaciones lineales

$$\begin{array}{rccccrcrcl} x & + & (2+i)y & - & 3z & = & 1 \\ 2x & & & & + & z & = & 1 \\ & & 5y & + & 7z & = & 2 \\ 3x & - & 2y & + & z & = & 3 \end{array}$$

Vamos a usar Maxima para calcular el determinante de la matriz del sistema

```
B:matrix([1,2+%i,-3,1],[2,0,1,1],[0,5,7,2],[3,-2,1,3]);
ratsimp(determinant(B));
-23*%i-138
```

El determinante de la matriz del sistema es diferente de cero, lo que implica que la matriz del sistema tiene rango 4 y el de la de coeficientes necesariamente será 3. Por tanto el sistema es incompatible. Sin embargo podemos intentar calcular lo que llamaríamos la “mejor solución” que se trata de los valores para los que  $Ax$  está más cerca de  $b$ . Para calcular los valores de  $x, y, z$  que están más cercanos a la solución ponemos

$$A = \begin{pmatrix} 1 & 2+i & -3 \\ 2 & 0 & 1 \\ 0 & 5 & 7 \\ 3 & -2 & 1 \end{pmatrix} \quad y \quad b = \begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \end{pmatrix}$$

y calculamos  $\overline{A}^T A$  y  $\overline{A}^T b$  usando Maxima:

```
A:matrix([1,2+%i,-3],[2,0,1],[0,5,7],[3,-2,1]);
b:[1,1,2,3];
ratsimp(conjugate(transpose(A)).A);
conjugate(transpose(A)).b;
(A) matrix(
[1,%i+2,-3],
[2,0,1],
[0,5,7],
[3,-2,1]
)
(b) [1,1,2,3]
matrix(
[14,%i-4,2],
[-%i-4,34,3*%i+27],
[2,27-3*%i,60]
)
matrix(
```

[12],  
 [6\-%i],  
 [15]  
 )

O sea

$$\overline{A}^T A = \begin{pmatrix} 14 & i-4 & 2 \\ -i-4 & 34 & 27+3i \\ 2 & 27-3i & 60 \end{pmatrix} \quad \text{y} \quad \overline{A}^T b = \begin{pmatrix} 12 \\ 6-i \\ 15 \end{pmatrix}.$$

Resolviendo el sistema  $(A^T A)X = A^T b$

```
linsolve([14*x+(%i-4)*y+2*z=12, (-%i-4)*x+34*y+(3*%i+27)*z=6-%i,
2*x+(27-3*%i)*y+60*z=15], [x,y,z]);
[x=\-(465*%i\ -14637)/16628, y=\-(352*%i\ -1377)/8314, z=(470*%i+2465)/16628]
```

obtenemos que la solución aproximada viene dada por

$$x = \frac{14637 - 465i}{16628}, \quad y = \frac{1377 - 352i}{8314}, \quad z = \frac{2465 + 470i}{16628}.$$

Vamos a ver una aplicaciones de esto. Imaginemos que tenemos una serie de puntos del plano  $(x_1, y_1), \dots, (x_n, y_n)$  y queremos encontrar una recta que esté lo más cerca posible de estos puntos. Supongamos que la ecuación de esta recta es  $Y = a + bX$ . La situación ideal sería que todos los puntos estuvieran en esta recta es decir que se cumpliera:

$$\begin{aligned} a + bx_1 &= y_1 \\ a + bx_2 &= y_2 \\ &\vdots \\ a + bx_n &= y_n \end{aligned}$$

En tal caso  $(a, b)$  sería una solución del sistema de ecuaciones lineales  $AX = y$  donde

$$A = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix} \quad \text{e} \quad y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$

Pero si no todos los puntos están alineados lo más a lo que podemos aspira es a encontrar la solución más cercana, es decir los valores  $(a, b)$  para los que  $\left\| A \begin{pmatrix} a \\ b \end{pmatrix} - y \right\|$  es mínimo. Para ello aplicando la Proposición 7.6.1 tenemos que resolver el sistema

$$\overline{A}^T AX = \overline{A}^T y.$$

Observa que este sistema tiene la forma

$$\begin{pmatrix} n & \sum_{i=1}^n x_i \\ \sum_{i=1}^n \overline{x}_i & \sum_{i=1}^n |x_i|^2 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n \overline{x}_i y_i \end{pmatrix}.$$

Si consideramos los vectores  $x = (x_1, \dots, x_n)$ ,  $y = (y_1, \dots, y_n)$  y  $u = (1, 1, \dots, 1)$  y consideramos el producto hermitico de  $\mathbb{C}^n$  del Ejemplo 7.1.2.(1), o sea

$$(a_1, \dots, a_n) \cdot (b_1, \dots, b_n) = a_1 \bar{b}_1 + \dots + a_n \bar{b}_n$$

tenemos que el sistema toma la forma siguiente:

$$\begin{pmatrix} \|v\|^2 & u \cdot v \\ v \cdot u & \|u\|^2 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} y \cdot u \\ y \cdot x \end{pmatrix}.$$

El determinante de la matriz de coeficientes es

$$\|u\|^2 \|v\|^2 - |u \cdot v|^2.$$

De la Desigualdad de Schwarz (Proposición 7.2.1.(4)) deducimos que la matriz es invertible si y solo si  $u \in \langle v \rangle$  si y solo si  $x_1 = \dots = x_n$ . En caso contrario, es decir si no todos los  $x_i$  son iguales, tenemos que la solución del sistema es única con lo que aplicando la Regla de Cramer tenemos que

$$a = \frac{(\sum_{i=1}^n |x_i|^2) (\sum_{i=1}^n y_i) - (\sum_{i=1}^n \bar{x}_i) (\sum_{i=1}^n \bar{x}_i y_i)}{n \sum_{i=1}^n |x_i|^2 - |\sum_{i=1}^n \bar{x}_i|^2}$$

$$b = \frac{n \sum_{i=1}^n \bar{x}_i y_i - (\sum_{i=1}^n \bar{x}_i) (\sum_{i=1}^n y_i)}{n \sum_{i=1}^n |x_i|^2 - |\sum_{i=1}^n \bar{x}_i|^2}.$$

La recta así calculada se llama *recta de regresión de mínimos cuadrados* porque es la que minimiza  $\sum_{i=1}^n \|y_i - (ax_i + b)\|^2$ .

**Ejemplo 7.6.3** Supongamos que hemos realizado mediciones de dos variables  $x$  e  $y$  obteniendo los siguientes valores:

$x$	2	1.5	1.2	0.2	2	3	1	0.1
$y$	123	81	96.3	13	140	250	89	10

que aparecen representados en la Figura 7.3. Aunque está claro que estos puntos no están alineados, parece que se aproximan a los puntos de una recta. Para buscar la recta  $y = a + bx$  a la que estos puntos se aproximan más aplicamos las fórmulas anteriores. Para ello calculamos

$$\sum_{i=1}^8 x_i = 11, \quad \sum_{i=1}^8 y_i = 802,3, \quad \sum_{i=1}^8 x_i^2 = 21,74, \quad \sum_{i=1}^8 x_i y_i = 1605,66$$

y con ello tenemos que

$$a = \frac{21,74 \cdot 802,3 - 11 \cdot 1605,66}{8 \cdot 21,74 - 11^2} \approx -4,16 \quad \text{y} \quad b = \frac{8 \cdot 1605,66 - 11 \cdot 802,3}{8 \cdot 21,74 - 11^2} \approx 75,96.$$

La recta pintada en azul en la gráfica es la que tiene por ecuación  $y = 75,96x - 4.16$ .

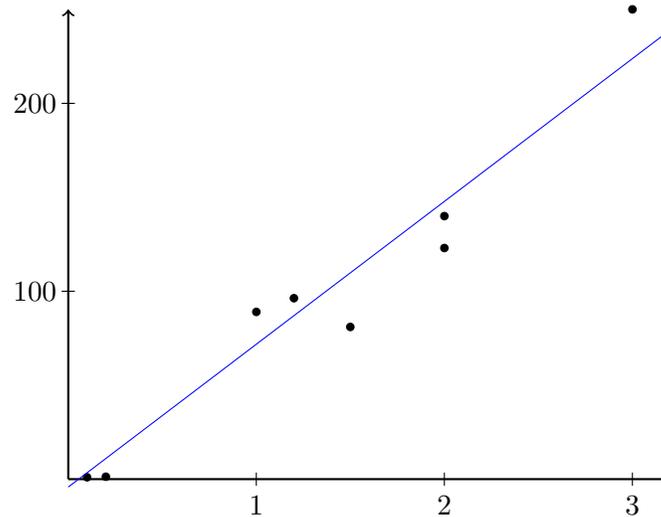


Figure 7.3:

En resumen podemos detectar si dos variables sobre las que tenemos datos experimentales están relacionados de forma lineal y la recta de mínimos cuadrados proporcionaría el modelo lineal que más se aproxima a los datos. Cuanto más cercanos a la recta estén los puntos representando las parejas de datos más confianza tendremos en que el modelo lineal es adecuado.

Supongamos ahora que los datos para dos variables de un experimento son los siguientes

$x$	0.1	2.1	1.3	0.3	1.2	1.3	1.8
$y$	2	20.3	5.6	1.65	4.45	5.65	13

Los hemos representado en la Figura 7.4 junto con la recta de regresión de mínimos cuadrados que en este caso no se ajusta tan bien a los puntos. De hecho parece como que una parábola podría aproximarse más a los puntos. Por tanto vamos a buscar una parábola  $y = a + bx + cx^2$ . Si todos los puntos pasaran por una parábola tendríamos que  $(a, b, c)$  sería la solución del siguiente sistema lineal de ecuaciones:

$$\begin{pmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$

Para buscar la parábola más cercana multiplicamos el sistema por la transpuesta conjugada de la matriz de coeficientes y obtenemos el sistema

$$\begin{pmatrix} n & \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 \\ \sum_{i=1}^n \bar{x}_i & \sum_{i=1}^n |x_i|^2 & \sum_{i=1}^n \bar{x}_i |x_i|^2 \\ \sum_{i=1}^n \bar{x}_i^2 & \sum_{i=1}^n \bar{x}_i |x_i|^2 & \sum_{i=1}^n |x_i|^4 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n \bar{x}_i y_i \\ \sum_{i=1}^n \bar{x}_i^2 y_i \end{pmatrix}$$

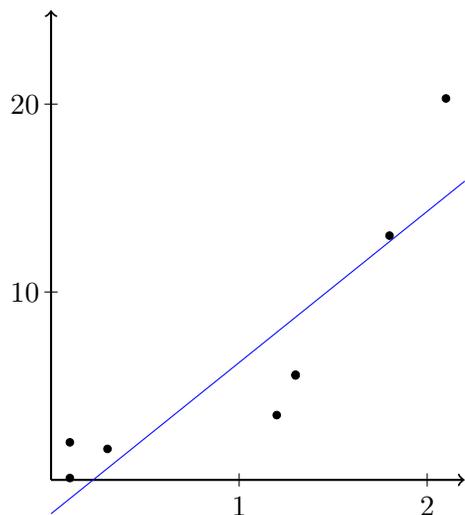


Figure 7.4:

En nuestro caso como todos los valores son reales esto toma la forma

$$\begin{pmatrix} n & \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 \\ \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i^3 \\ \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i^3 & \sum_{i=1}^n x_i^4 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n x_i y_i \\ \sum_{i=1}^n x_i^2 y_i \end{pmatrix}$$

Para los valores de la tabla anterior esto toma la forma

$$\begin{pmatrix} 7 & 8.1 & 12.57 \\ 8.1 & 12.57 & 21.243 \\ 12.57 & 21.243 & 37.7397 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 52,65 \\ 86,69 \\ 157,232 \end{pmatrix}$$

Resolviendo el sistema obtenemos

$$a \approx 3, \quad b \approx -8, \quad c \approx 7,66$$

La Figura 7.5 muestra los puntos y la parábola  $y = 3 - 8x + 7.66x^2$  y observamos que efectivamente se ajusta bastante bien a los datos.

O sea podemos estudiar si unos datos obtenidos experimentalmente se aproximan a un modelo polinomial  $y = P(x)$  pongamos de grado  $n$ . Para ello consideramos los  $n+1$  coeficientes de un polinomio genérico  $P = p_0 + p_1X + \dots + p_nX^n$  como variables y con los datos de los que disponemos  $\{(x_i, y_i) : i = 1, \dots, n\}$  planteamos un sistema de ecuaciones lineales

$$p_0 + p_1x_i + p_2x_i^2 + \dots + p_nx_i^n = y_i \quad (i = 1, 2, \dots, n).$$

y buscamos los valores de los  $p_i$  que más se aproximan a una solución del sistema.

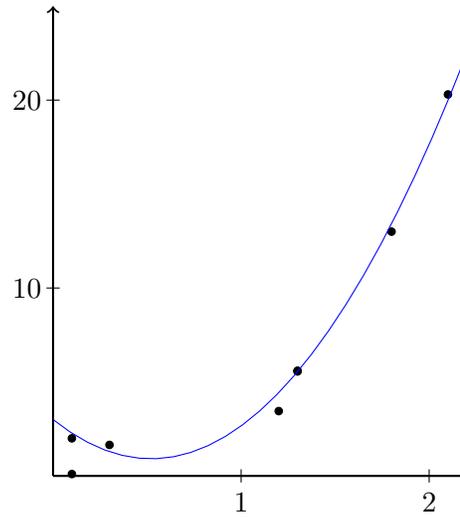


Figure 7.5:

## 7.7 Diagonalización ortogonal

Un endomorfismo de un espacio vectorial  $V$  de dimensión finita se dice que es *simétrico* si  $f_B$  es una matriz simétrica para alguna base  $B$  de  $V$ .

Sea  $A \in M_n(K)$ . Decimos que  $A$  es *diagonalizable ortogonalmente* si existe una matriz ortogonal  $U$  tal que  $U^{-1}AU$  es diagonal (o lo que es lo mismo  $U^T U = I$  y  $U^T A U$  es diagonal). En tal caso  $U^T A U$  es diagonal y por tanto simétrica. Por tanto

$$A^T = (U U^T A U U^T)^T = U (U^T A U)^T U^T = U U^T A U U^T = A$$

es decir  $A$  tiene que ser una matriz simétrica.

Un endomorfismo  $f$  de un espacio euclídeo  $V$  se dice que es *diagonalizable ortogonalmente* si  $f_B$  es una matriz diagonal para alguna base ortonormal  $B$  de  $V$ .

Si  $f$  es diagonalizable ortogonalmente, digamos con la base ortonormal  $B$  y  $B'$  es otra base ortonormal entonces la matriz  $U = B'_B$  es ortogonal, por la Proposición 7.3.10. Por tanto  $f_B = U^T f_{B'} U$  es diagonal y por tanto  $f_{B'}$  es una matriz simétrica con lo que  $f$  es un endomorfismo simétrico.

Vamos a usar el producto escalar estándar en  $\mathbb{C}^n$ , o sea  $v \cdot u = v \bar{u}^T$ . Por tanto, si  $A \in M_n(\mathbb{C})$  entonces

$$(vA) \cdot u = v A \bar{u}^T = v (\bar{u} A^T)^T = v \cdot u \bar{A}^T.$$

Hasta ahora cuando  $v$  era un autovector de una matriz  $A$  con valor propio  $\lambda$  habíamos escrito  $Av = \lambda v$  pero en realidad  $v$  era un vector columna con lo que si vemos  $v$  como vector fila esto significa que  $Av^T = \lambda v^T$ , o lo que es lo mismo  $vA^T = \lambda v$ . En el caso en que  $A$  sea una matriz simétrica esto es lo mismo que  $vA = \lambda v$ . Si además  $A$  tiene entradas reales entonces tenemos que  $(vA) \cdot u = v \cdot (uA)$ .

**Proposición 7.7.1** *Si  $A$  es una matriz simétrica real entonces el polinomio característico de  $A$  es completamente factorizable en  $\mathbb{R}[X]$ .*

**Demostración.** Por el Teorema Fundamental del Álgebra,  $\chi_A$  es completamente factorizable en  $\mathbb{C}[X]$ . Por tanto basta con demostrar que todos las raíces de  $\chi_A$ , o sea los valores propios de  $A$  son números reales. Sea  $\lambda$  un valor propio de  $A$  y  $v$  un vector propio distinto de cero con valor propio  $\lambda$ . Luego  $vA = \lambda v$ . Tomando conjugados y teniendo en cuenta que las entradas de  $A$  son números reales tenemos que  $\bar{v}A = \bar{\lambda}\bar{v}$ . Por tanto

$$\lambda\|v\|^2 = \lambda(v \cdot v) = (\lambda v) \cdot v = (vA) \cdot v = v \cdot (vA) = v \cdot (\lambda v) = \bar{\lambda}(v \cdot v) = \bar{\lambda}\|v\|^2.$$

Como  $\|v\| \neq 0$  deducimos que  $\lambda = \bar{\lambda}$  y por tanto  $\lambda \in \mathbb{R}$ . ■

**Teorema 7.7.2** *Si  $f$  es un endomorfismo de un espacio euclídeo de dimensión finita  $V$  tal que la matriz asociada a  $f$  en alguna base ortonormal  $B$  de  $V$  es simétrica entonces  $f$  es diagonalizable ortogonalmente.*

**Demostración.** Vamos a razonar por inducción en la dimensión  $n$  de  $V$ . El resultado está claro si  $n = 1$  con lo que supondremos que  $n > 1$  y que el resultado se verifica si la dimensión fuera menor que  $n$ . Como  $f_B$  es una matriz simétrica para alguna base, se deduce de la Proposición 7.7.1 que los autovalores de  $f$  son números reales y al menos tendrá uno. Por tanto  $f$  tiene al menos un autovector no nulo  $v$  que podemos suponer unitario. Sea  $W = v^\perp$  y sea  $B_1$  una base ortonormal de  $W$ . Entonces  $B = \{v\} \cup B_1$  es una base ortonormal de  $V$ . Como  $f_{B'}$  es simétrica para otra base ortonormal se tiene que  $U = B_{B'}$  es ortonormal y  $A = f_B = U^T f_{B'} U$  con lo que  $A$  es una matriz simétrica. Si  $w \in W$  entonces

$$v \cdot f(w) = v \cdot wA = (vA) \cdot w = \lambda v \cdot w = 0.$$

Esto demuestra que  $f(W) \subseteq W$  y por tanto  $f$  es un endomorfismo de  $W$ . Además  $f_B = \begin{pmatrix} \lambda & 0 \\ 0 & f_{B_1} \end{pmatrix}$  con lo que  $f_{B_1}$  es una matriz simétrica. Como  $\dim W = n - 1$ , podemos aplicar la hipótesis de inducción a la restricción  $g$  de  $f$  a  $W$  visto como un endomorfismo de  $W$ . Por tanto  $W$  tiene una base ortonormal  $v_2, \dots, v_n$  tal  $f_{B_1}$  es diagonal. Entonces  $C = \{v_1, v_2, \dots, v_n\}$  es una base ortonormal de  $V$  con  $f_C$  una matriz diagonal. ■

**Corolario 7.7.3** *Toda matriz simétrica con entradas reales es diagonalizable ortogonalmente.*

**Demostración.** Sea  $A$  una matriz satisfaciendo las hipótesis del corolario. Considérese el endomorfismo  $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$  dado por  $f(x) = Ax$ . Entonces  $A$  es la matriz asociada a  $f$  en la base canónica la cual es una base ortonormal de  $\mathbb{R}^n$  para el producto estándar. Por el Teorema 7.7.2 se  $\mathbb{R}^n$  tiene una base ortonormal  $B$  para la que  $f_B$  es diagonal. Entonces la matriz formada poniendo los vectores de  $B$  en columna es una matriz ortogonal  $U$  con  $U^T A U$  diagonal. ■

**Proposición 7.7.4** *Sea  $A$  una matriz simétrica con entradas reales. Dos autovectores con autovalores diferentes son ortogonales respecto al producto escalar estándar.*

**Demostración.** Supongamos que  $v_1$  y  $v_2$  son autovectores con autovalores  $\lambda_1$  y  $\lambda_2$  respectivamente. Entonces  $\lambda_1, \lambda_2 \in \mathbb{R}$  y

$$(\lambda_1 v_1) \cdot v_2 = (v_1 A) \cdot v_2 = v_1 \cdot (v_2 A) = v_1 \cdot (\lambda_2 v_2) = \lambda_2 (v_1 \cdot v_2).$$

Luego  $(\lambda_1 - \lambda_2)(v_1 \cdot v_2) = 0$  y por tanto si  $\lambda_1 \neq \lambda_2$  tenemos que  $v_1 \cdot v_2 = 0$ . ■

El proceso para diagonalizar ortogonalmente una matriz simétrica con entradas reales solo requiere un paso más al de la diagonalización habitual. Es decir simplemente calculamos los valores propios  $\lambda_1, \dots, \lambda_r$ , que serán todos reales por la Proposición 7.7.1. Para cada uno de ellos calculamos una base de  $N(A - \lambda_i I)$  y le aplicamos el Método de Gram-Schmidt para asegurarnos de que es una base ortonormal de  $N(A - \lambda_i I)$ . Poniendo todos los vectores obtenidos juntos tenemos una base ortonormal como consecuencia de la Proposición 7.7.4.

## 7.8 Producto vectorial

Lo que vamos a ver en esta sección solo vale para el espacio euclídeo  $\mathbb{R}^3$  con el producto escalar estándar. Se define el *producto vectorial* de dos elementos  $u$  y  $v$  de  $\mathbb{R}^3$  como el único vector  $u \wedge v$  que cumple las siguientes condiciones:

- (1)  $\|u \wedge v\| = \|u\| \|v\| |\widehat{\text{sen}(u, v)}|$ .
- (2)  $u \wedge v$  es ortogonal a  $u$  y  $v$ .
- (3) El determinante de la matriz cuyas filas con  $u$ ,  $v$  y  $u \wedge v$  es positivo.

Vamos a ver que efectivamente  $u \wedge v$  queda completamente determinado por las condiciones dadas. Si  $u, v$  es linealmente dependiente entonces o bien  $u = 0$  ó  $v = 0$  ó  $\widehat{u, v}$  es  $0$  ó  $\pi$ . Por tanto, de la condición (1) se deduce que  $u \wedge v = 0$ . En caso contrario,  $\langle u, v \rangle$  tiene dimensión 2 y por tanto  $\{u, v\}^\perp = \langle u, v \rangle^\perp$  tiene dimensión 1, por la Proposición 7.5.3. Sea  $w$  un generador de  $\langle u, v \rangle^\perp$ . Por la condición (2)  $u \wedge v = \alpha w$  para algún número real  $\alpha$  que será diferente de cero por la condición (3). Pero

$$\|\beta w\| = |\beta| \|w\|$$

para todo  $\beta \in \mathbb{R}$  con lo que hay exactamente dos números reales  $\beta$  que cumplen

$$\|\beta w\| = \|u\| \|v\| |\widehat{\text{sen}(u, v)}|$$

y esos dos números reales son opuestos uno del otro. O sea si  $\beta$  cumple la propiedad anterior entonces  $u \wedge v = \alpha w$  es  $\beta w$  ó  $-\beta w$ , o sea  $\alpha$  es  $\beta$  ó  $-\beta$ . Para decidir cuál de los dos es nos fijamos en que las matrices  $(u, v, \beta w)$  y  $(u, v, -\beta w)$  tienen determinantes opuestos con lo que  $\alpha = \beta$  si  $\det(u, v, \beta w) > 0$  y en caso contrario  $\alpha = -\beta$ .

**Proposición 7.8.1** *Si  $u, v, w \in \mathbb{R}^3$  entonces*

- (1)  $|u \wedge v|$  es el área del paralelogramo determinado por los vectores  $u$  y  $v$  (Figura de la izquierda en 7.6).
- (2)  $|(u \wedge v) \cdot w|$  es el volumen del paralelogramo determinado por los vectores  $u$ ,  $v$  y  $w$  (Figura de la derecha en 7.6).

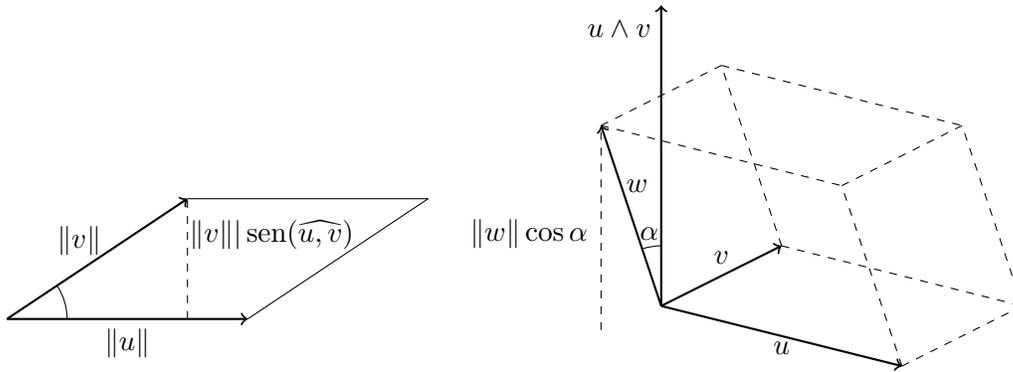


Figure 7.6:

La siguiente proposición nos va proporcionar una fórmula para calcular  $u \wedge v$ :

**Proposición 7.8.2** Si  $u = (u_1, u_2, u_3)$  y  $v = (v_1, v_2, v_3)$  son dos elementos de  $\mathbb{R}^3$  entonces

$$u \wedge v = \left( \begin{vmatrix} u_2 & u_3 \\ v_2 & v_3 \end{vmatrix}, \begin{vmatrix} u_3 & u_1 \\ v_3 & v_1 \end{vmatrix}, \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix} \right).$$

**Demostración.** Basta con comprobar que la fórmula dada define un producto que verifican las tres condiciones de la definición. Si  $u$  y  $v$  son linealmente dependientes entonces también lo son las filas de las tres matrices que aparecen en la fórmula y por tanto  $u \wedge v = 0$ . Supongamos que  $u$  y  $v$  no son linealmente dependientes. Supongamos que  $u$  y  $v$  son linealmente independientes. Eso implica que uno de los tres menores de la matriz formada por las coordenadas de  $u$  y  $w$  es distinto de cero y en particular  $u \wedge v \neq 0$ .

Si  $w = (w_1, w_2, w_3)$  es un tercer vector entonces desarrollando el siguiente determinante por la primera fila tenemos lo siguiente:

$$\begin{vmatrix} w_1 & w_2 & w_3 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} = w_1 \begin{vmatrix} u_2 & u_3 \\ v_2 & v_3 \end{vmatrix} + w_2 \begin{vmatrix} u_3 & u_1 \\ v_3 & v_1 \end{vmatrix} + w_3 \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix} = w \cdot (u \wedge v).$$

En el caso particular en que  $w = u$  ó  $w = v$  el determinante será cero y por tanto  $u \wedge v$  es ortogonal a  $u$  y a  $v$  lo que implica que pertenece a  $\langle u, v \rangle^\perp$  por la Proposición 7.5.2.(3). Por otro lado si  $w = u \wedge v$  entonces el determinante es precisamente

$$\begin{vmatrix} w_1 & w_2 & w_3 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} = w_1^2 + w_2^2 + w_3^2 > 0.$$

Esto demuestra que se verifica la propiedad (2).

Además, por un lado

$$\begin{aligned}\|u \wedge v\|^2 &= \begin{vmatrix} u_2 & u_3 \\ v_2 & v_3 \end{vmatrix}^2 + \begin{vmatrix} u_3 & u_1 \\ v_3 & v_1 \end{vmatrix}^2 + \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix}^2 \\ &= (u_2v_3 - u_3v_2)^2 + (u_1v_3 - u_3v_1)^2 + (u_1v_2 - u_2v_1)^2 \\ &= \sum_{i \neq j} u_i^2 v_j^2 - 2 \sum_{i < j} u_i u_j v_i v_j.\end{aligned}$$

y por otro

$$\begin{aligned}\|u\|^2 \|v\|^2 \operatorname{sen}^2(\widehat{u, v}) &= \|u\|^2 \|v\|^2 (1 - \cos^2(\widehat{u, v})) = \|u\|^2 \|v\|^2 - (u \cdot v)^2 \\ &= (u_1^2 + u_2^2 + u_3^2)(v_1^2 + v_2^2 + v_3^2) - (u_1v_1 + u_2v_2 + u_3v_3)^2 \\ &= \sum_{i=1}^3 u_i^2 v_i^2 + \sum_{i \neq j} u_i^2 v_j^2 - \left( \sum_{i=1}^3 u_i^2 v_i^2 + 2 \sum_{i < j} u_i u_j v_i v_j \right) \\ &= \sum_{i \neq j} u_i^2 v_j^2 - 2 \sum_{i < j} u_i u_j v_i v_j.\end{aligned}$$

Esto demuestra que se cumple (1). ■

De la Proposición 7.8.2 se deduce fácilmente el siguiente

**Corolario 7.8.3** Si  $\alpha \in \mathbb{R}$  y  $u, v, w \in \mathbb{R}^3$  entonces

- (1)  $v \wedge u = -u \wedge v$ .
- (2)  $(\alpha u) \wedge v = u \wedge \alpha v = \alpha(u \wedge v)$ .
- (3)  $u \wedge (v + w) = (u \wedge v) + (u \wedge w)$  y  $(u + v) \wedge w = (u \wedge w) + (v \wedge w)$ .
- (4)  $u \cdot (v \wedge w) = \det(u, v, w)$ .

## 7.9

**7.1** Decidir cuáles de los siguientes productos en  $\mathbb{R}^2$  son productos escalares donde  $x = (x_1, x_2)$  e  $y = (y_1, y_2)$ :

- (1)  $x \cdot y = x_1 y_1 + 2x_2 y_2 + 2x_1 y_2 + x_2 y_1$ .
- (2)  $x \cdot y = x_1 y_1 - x_2 y_2$ .
- (3)  $x \cdot y = x_1 y_1 + 3x_2 y_2 + 2x_1 y_2 + 2x_2 y_1$ .
- (4)  $x \cdot y = 2x_1 y_1 + x_2 y_2 + x_1 y_2 + x_2 y_1$ .

**7.2** Decidir para cuáles de las siguientes matrices el producto dado como en el Ejemplo 7.1.2.(2) es un producto escalar:

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

**7.3** ¿Cuáles de las siguientes matrices con hermíticas?

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ -1 & 1 & 2 \end{pmatrix} \quad \begin{pmatrix} 1 & 1 & 2 \\ 1 & 0 & 1 \\ 2 & 1 & 2 \end{pmatrix} \quad \begin{pmatrix} 0 & i & 1+i \\ -i & 0 & i-2 \\ 1-i & i+2 & 2 \end{pmatrix} \quad \begin{pmatrix} 0 & i & 1+i \\ -i & 0 & i-2 \\ 1-i & -i-2 & 2 \end{pmatrix}.$$

**7.4** ¿Para qué números reales  $\alpha$  la siguiente fórmula define un producto escalar en  $\mathbb{R}^2$ :  $(x_1, x_2) \cdot (y_1, y_2) = x_1y_1 - x_1y_2 - x_2y_1 + \alpha x_2y_2$ ?

**7.5** La traza de una matriz cuadrada  $A$  es la suma de los elementos de su diagonal y la denotamos como  $\text{tr}(A)$ . Demostrar que se cumplen las siguientes propiedades:

- (1)  $\text{tr}(A^T) = \text{tr}(A)$ .
- (2) Si  $A = (a_{ij})$  y  $B = (b_{ij})$  entonces la entrada  $i$ -ésima de la diagonal de  $AB$  es  $\sum_{j=1}^n a_{ij}b_{ji}$ .
- (3)  $\text{tr}(AB) = \text{tr}(BA)$ .
- (4) El siguiente es un producto escalar en  $M_n(\mathbb{R})$ :  $A \cdot B = \text{tr}(AB^T)$ .
- (5)  $\text{tr}(\bar{A}) = \overline{\text{tr}(A)}$ , donde si  $A = (a_{ij})$  entonces  $\bar{A} = (\bar{a}_{ij})$ .
- (6) El siguiente es un producto hermítico en  $M_n(\mathbb{C})$ :  $A \cdot B = \text{tr}(A\bar{B}^T)$ .

**7.6** Sean  $x_1, \dots, x_k \in \mathbb{K}$ . Demostrar que el siguiente es un producto hermítico en el espacio vectorial  $\mathbb{K}[X]$ :

$$P \cdot Q = \sum_{i=1}^k P(x_i)\overline{Q(x_i)}.$$

**7.7** Para aquellos productos definidos en el Problema 7.1 que sean productos escalares calcular las normas de  $(1, -2)$  y  $(1, -1)$  y la distancia y el ángulo entre ellos. Para dichos productos, obtener una fórmula para la norma de un vector en función de sus coordenadas y para la distancia y el coseno del ángulo entre vectores arbitrarios en función de sus coordenadas.

**7.8** Para las matrices del Problema 7.2 que definan un producto escalar dar una fórmula para la norma y coseno del ángulo de vectores arbitrarios en función de las coordenadas.

**7.9** Consideremos los siguientes elementos del espacio euclídeo  $\mathcal{C}([-1, 1])$  del Ejemplo 7.1.2.(3):

$$f(x) = x, \quad g(x) = x^2, \quad h(x) = e^x.$$

Calcular las normas de  $f$ ,  $g$  y  $h$  y las distancias y ángulos entre cada dos de ellos.

**7.10** Demostrar que se verifican las siguientes en un espacio hermítico:

- (1)  $\|u\| - \|v\| \leq \|u - v\| \leq \|u\| + \|v\|$ .
- (2)  $\|u + v\|^2 + \|u - v\|^2 = 2(\|u\|^2 + \|v\|^2)$  (Ley del paralelogramo).
- (3)  $\|u + v\|^2 - \|u - v\|^2 = 4\operatorname{Re}(u \cdot v)$ .
- (4)  $\|u + v\|^2 - \|u\|^2 - \|v\|^2 = 2\operatorname{Re}(u \cdot v)$ .

**7.11** Demostrar que si  $u$  y  $v$  son elementos de un espacio euclídeo entonces

$$(u + v) \cdot (u - v) = \|u\|^2 - \|v\|^2.$$

**7.12** Sean  $u, v$  elementos no nulos de un espacio euclídeo.

- (1) Demuestra que  $|u \cdot v| = \|u\| \|v\|$  si y solo si  $\{u, v\}$  es linealmente dependiente. Indicación: Revisa la demostración de la Desigualdad de Schwarz.
- (2) Demuestra que las siguientes condiciones son equivalentes:
  - (a)  $\widehat{u, v} = 0$ .
  - (b)  $\|u + v\| = \|u\| + \|v\|$ .
  - (c)  $v = \alpha u$  para algún número real positivo  $\alpha$ .

Indicación: Revisa la demostración de la Desigualdad Triangular.

- (3) Demuestra que las siguientes condiciones son equivalentes:
  - (a)  $\widehat{u, v} = \pi$ .
  - (b)  $\|u - v\| = \|u\| + \|v\|$ .
  - (c)  $v = \alpha u$  para algún número real negativo  $\alpha$ .

**7.13** Calcular un vector de  $\mathbb{R}^4$  que sea unitario y ortogonal a los siguientes vectores con respecto al producto escalar estándar:  $(1, 2, 3, 0)$ ,  $(2, 2, 1, 1)$ ,  $(-1, 3, 2, 1)$  Realiza el mismo problema cambiando para el producto escalar dado por  $x \cdot y = xAy^T$  para la matriz:

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 5 & 1 & 0 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & -1 & 6 \end{pmatrix}$$

**7.14** Consideremos el espacio euclídeo  $\mathcal{C}([0, 1])$  de las funciones continuas  $[0, 1] \rightarrow \mathbb{R}$ . Demostrar que si  $p = p_0 + p_1X + p_2X^2$  y  $q = q_0 + q_1X + q_2X^2$  son dos polinomios entonces

$$p \cdot q = (p_0, p_1, p_2) \begin{pmatrix} 1 & 1/2 & 1/3 \\ 1/2 & 1/3 & 1/4 \\ 1/3 & 1/4 & 1/5 \end{pmatrix} \begin{pmatrix} q_0 \\ q_1 \\ q_2 \end{pmatrix}.$$

Calcular un polinomio de grado 2 que sea unitario y ortogonal a  $1 + X$  y  $2 - X$ .

**7.15** Demostrar que las siguientes matrices son ortogonales en el espacio euclídeo  $M_2(\mathbb{R})$  del Problema 7.5

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

Demostrar que la matriz  $\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$  es ortogonal a  $A$  si y solo si  $a_{11} + a_{21} = 2a_{12}$ . ¿Qué condición tiene que cumplir una matriz para que sea ortogonal a  $B$ ?

**7.16** Sean  $A, B \in M_n(\mathbb{R})$  y supongamos que  $A$  es simétrica y  $B$  es antisimétrica, o sea  $B^T = -B$ . Demostrar que  $A$  y  $B$  son ortogonales como elementos del espacio euclídeo del Problema 7.5.

**7.17** Encontrar un polinomio  $f_3$  de grado 2 con coeficientes en  $\mathbb{R}$  que sea ortogonal a los polinomios  $f_1 = 1 + X$  y  $f_2 = 2 - X$  con respecto al producto escalar en  $([-1, 1])$  del Ejemplo 3. Encuentra una base ortonormal del subespacio vectorial generado por  $f_1, f_2$  y  $f_3$ .

**7.18** Encontrar la expresión de un producto escalar en  $\mathbb{R}^3$  en el que los vectores:  $(1, 0, 0)$ ,  $(-1, 1, 0)$ ,  $\frac{1}{2}(1, -1, 2)$  forman una base ortonormal.

**7.19** Consideremos el espacio unitario  $M_3(\mathbb{C})$  con el producto hermítico definido en el Problema 7.5. Sean

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 & 1 \\ 0 & i & 0 \\ -i & 0 & 0 \end{pmatrix}, \quad C = \begin{pmatrix} 0 & 0 & 1 \\ i & 0 & 0 \\ 0 & -i & 0 \end{pmatrix}.$$

Calcular las matrices de la forma  $A + \beta B + \gamma C$  ortogonales a  $B$  y  $C$ .

**7.20** Consideremos  $M_n(\mathbb{K})$  como espacio hermítico con el producto definido en el Problema 7.5. Consideremos la base canónica  $B$  de  $M_n(\mathbb{K})$ , es decir aquella formada por los vectores de la forma  $e_{ij}$  que tienen un 1 en la entrada  $(i, j)$  y 0 en las demás entradas. Demostrar

(1)  $e_{ij}^T = e_{ij}$  para todo  $1 \leq i, j \leq n$ .

(2) Si  $1 \leq i, j, k, l \leq n$  entonces  $e_{ij}e_{kl} = \begin{cases} e_{il}, & \text{si } i = k; \\ 0, & \text{si } i \neq k. \end{cases}$

(3)  $B$  es una base ortonormal.

(4) Si  $A = (a_{ij})$  y  $B = (b_{ij})$  entonces  $A \cdot B = \sum_{i=1}^n \sum_{j=1}^n a_{ij}b_{ij}$ .

**7.21** Calcular una base ortonormal del espacio nulo, el espacio fila y el espacio columna de las siguientes matrices con respecto al producto hermítico estándar.

$$\begin{pmatrix} -1 & 2 & 0 & 1 \\ 2 & 1 & 2 & 3 \\ 5 & 5 & 6 & 10 \\ -9 & 8 & -4 & -1 \end{pmatrix} \quad \begin{pmatrix} 1 & 2 & 3 & 1 \\ 0 & 1 & 2 & 3 \\ 1 & 1 & 1 & 1 \\ -3 & 1 & 5 & 6 \end{pmatrix} \quad \begin{pmatrix} i & 2 & 3 & 1 \\ 0 & 1 & 2 & 3 \\ 1 & 1+2i & 1 & 1 \\ i-2 & 1-4i & 3 & 2 \end{pmatrix}$$

**7.22** Calcular una base ortonormal del subespacio de  $\mathcal{C}([0, 1])$  formado por los polinomios de grado menor o igual que 2

**7.23** Calcular una base ortonormal del subespacio de  $\mathcal{C}\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$  generado por las funciones  $f(x) = \sin(x)$ ,  $g(x) = \sin(2x)$ ,  $h(x) = \sin(3x)$ .

**7.24** Calcular una base ortonormal del subespacio de  $M_3(\mathbb{C})$ , con respecto al producto hermitico definido en el Problema 7.5, generado por las matrices  $A, A^T, B$  y  $B^T$  donde

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

**7.25** Considérese el espacio euclideo  $\mathbb{R}^4$  con el producto escalar estándar. Sea  $W$  el subespacio de  $\mathbb{R}^4$  formado por los vectores que satisfacen las siguientes ecuaciones:

$$x_1 + x_3 + x_4 = x_1 - x_2 + 2x_3 = 0.$$

Calcular

- (1) Una base ortogonal de  $W$  y una base ortogonal de  $W^\perp$ .
- (2) Las proyecciones ortogonales de  $\mathbb{R}^4$  en  $W$  y  $W^\perp$ .

**7.26** Consideremos en  $\mathbb{R}^4$  el producto dado por

$$x \cdot y = x_1y_1 + x_1y_2 + x_2y_1 + 2x_2y_2 + x_3y_3 - x_3y_4 - x_4y_3 + 3x_4y_4.$$

- (1) Demostrar que es un producto escalar.
- (2) Calcular una base ortonormal de  $\mathbb{R}^4$  con respecto a este producto escalar.
- (3) Calcular una base ortonormal del subespacio  $W$  dado por las ecuaciones  $x_1 = x_2 = 0$  y la proyección ortogonal en dicho espacio.
- (4) Calcular una base ortonormal de  $W^\perp$  y la proyección ortogonal en dicho subespacio.
- (5) Calcular una base ortonormal del subespacio  $U$  dado por las ecuaciones  $x_1 = x_3 = 0$  y la proyección ortogonal en dicho espacio.
- (6) Calcular una base ortonormal de  $U^\perp$  y la proyección ortogonal en dicho subespacio.

**7.27** Sea  $V$  el espacio euclideo  $M_2(\mathbb{R})$  con el producto escalar dado en el Problema 7.5. Para cada uno de los siguientes subespacios de  $V$  calcular una base ortogonal del subespacio y de su complemento ortogonal y fórmulas para las proyecciones ortogonales en ambos:

$$\left\langle I, J = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right\rangle, \quad \left\langle \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix} \right\rangle, \quad \langle I \rangle.$$

**7.28** Calcular las soluciones aproximadas de los siguientes sistemas de ecuaciones lineales incompatibles del Ejemplo 3.4.18.

**7.29** Calcular las soluciones aproximadas de los siguientes sistemas de ecuaciones lineales:

$$\left. \begin{array}{l} x + y = 0 \\ x - y = 2 \\ x + 2y = 1 \end{array} \right\} \left. \begin{array}{l} x + y + z = 3 \\ x - y - z = 0 \\ 2x + 2y + z = -2 \\ 2y + 3z = 1 \\ x + 2y = 0 \end{array} \right\}$$

**7.30** Calcular la recta de regresión de mínimos cuadrados de la siguiente lista de datos:

$x$	0.1	1.1	0.2	-3.1	0.1	-2.1	-0.2	-0.85
$y$	-2.65	4.67	-0.85	-11.32	1.82	6.8	-3.64	-3.07

**7.31** Calcular la recta de regresión de mínimos cuadrados de la siguiente lista de datos y la parábola que más se ajusta a los datos

$x$	0.1	1.1	0.2	-3.1	0.1	-2.1	-0.2	-0.85
$y$	-0.56	4.56	1.11	-15.63	3.91	2.49	-1.28	-0.75

**7.32** Demostrar que si un sistema es compatible las soluciones aproximadas son las mismas que las soluciones.

**7.33** Diagonalizar ortogonalmente las siguientes matrices simétricas:

$$\begin{pmatrix} 3 & -1 & 0 \\ -1 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix} \quad \begin{pmatrix} 5 & -10 & 8 \\ -10 & 2 & 2 \\ 8 & 2 & 11 \end{pmatrix} \quad \begin{pmatrix} 0 & 2 & 2 \\ 2 & 0 & 2 \\ 2 & 2 & 0 \end{pmatrix}.$$

# Index

- ángulo
  - en un espacio euclídeo, 207
- anillo, 9
  - conmutativo, 9
- aplicación inversa, 35
- aplicación lineal, 29, 125
- argumento
  - de un número complejo, 10
- asociativa, 6
- autovalor
  - de un endomorfismo, 152
- autovector
  - de un endomorfismo, 152
  - de una matriz, 152
- base, 43, 66
  - canónica
    - de  $K[X]$ , 41
    - de  $K^n$ , 35
    - de  $M_{m,n}(K)$ , 67
- biyectiva, 97
- cero
  - de un anillo, 9
  - de un cuerpo, 7
- cociente, 14
- coeficiente, 12
  - principal de un polinomio, 12
- coeficientes, 36
- combinación lineal, 36
  - un vector es, 37
- compatible
  - sistema de ecuaciones lineales, 61
- compatible determinado
  - sistema de ecuaciones lineales, 61
- compatible indeterminado
  - sistema de ecuaciones lineales, 61
- complementarios
  - subespacios vectoriales, 111
- complemento, 111
- completamente factorizable
  - polinomio, 16
- conjugado
  - de un número complejo, 10
- conmutativa, 6
- cuadrada
  - matriz-, 26
- cuerpo, 7
- definida positiva
  - matriz, 203
- diagonaliza
  - base que, 151
  - matriz que, 151
- diagonalizable
  - endomorfismo, 151
  - matriz, 151
- diagonalización, 153
- diagonalizar
  - un endomorfismo, 151
  - una matriz, 151
- dimensión, 103
  - de una variedad, 117
  - finita, 103
- ecuaciones
  - implícitas, 116
  - paramétricas del plano, 116
- ecuación
  - paramétrica de la recta, 114
  - continua de la recta, 115
  - lineal

- homogénea, 107
  - vectorial de la recta, 114
  - vectorial del plano, 116
- ecuación lineal, 60
  - homogénea, 39
- Eliminación de Gauss
  - Método, 53
- endomorfismo
  - diagonalizable, 151
- endomorfismo
  - diagonalizable
    - ortogonalmente, 229
  - simétrico, 229
- equivalentes
  - sistemas de ecuaciones lineales, 61
- escalonada
  - matriz, 53
- espacio
  - columna, 61, 65
  - fila, 61, 65
  - nulo, 61, 65
- espacio euclídeo, 201
- espacio hermítico, 201
- espacio unitario, 201
- espacio vectorial, 25
- evaluación
  - de un polinomio, 13
- factorización
  - de un polinomio, 17
- forma de Jordan, 170
- Fórmula de Grassman, 110
- fuertemente escalonada
  - matriz, 53
- Gauss-Jordan
  - Método, 53
- grados de libertad
  - de un sistema de ecuaciones lineales, 65
- Gram-Schmidt
  - Metodo de, 215
- hermítica
  - matriz, 203
- incompatible
  - sistema de ecuaciones lineales, 61
- independientes
  - subespacios, 113
- inversa
  - de una aplicación, 35
- inverso, 7
- invertible, 9
  - matriz, 34
- inyectiva, 97
- Jordan
  - forma, 170
  - matriz, 170
  - matriz elemental, 169
- libre, 43
- ligado, 43
- lineal
  - aplicación, 29, 125
- linealmente
  - dependiente, 43
  - independiente, 43
- matrices elementales, 56
- matriz, 26
  - asociada
    - a un producto hermítico, 204
  - cuadrada, 26
  - de Jordan, 170
    - real, 188
  - de un sistema de ecuaciones lineales, 60
  - definida positiva, 203
  - diagonal, 32
  - diagonalizable, 151
    - ortogonalmente, 229
  - elemental de Jordan, 169
  - escalonada, 53
  - fuertemente escalonada, 53
  - hermítica, 203
  - invertible, 34
  - ortogonal, 213
  - simétrica, 57
- matriz de los coeficientes, 61
- matriz de paso, 156, 170

- menor complementario, 76
- Método de Gauss, 55
- Método de Gauss-Jordan, 55
- Método de los Contrarrecíprocos, 45
- módulo
  - de un número complejo, 10
- monomio, 12
- multiplicidad
  - de una raíz, 16
- neutro, 6
- norma, 206
- núcleo, 40
- números complejos, 6
- números enteros, 6
- números naturales, 6
- números racionales, 6
- números reales, 6
- operación binaria, 5
- operación elemental
  - en un sistema de ecuaciones lineales, 61
- opuesto, 7, 25
  - en un anillo, 9
- orientación, 214
- ortogonal
  - lista, 209
  - matriz, 213
  - subespacio, 217
- ortogonales, 209
- parte imaginaria, 10
- parte real, 10
- pivote
  - de una matriz escalonada, 53
- plano, 116
- polinomio, 11
- polinomio característico
  - de un endomorfismo, 154
  - de una matriz, 154
- producto
  - escalar
    - estándar, 202
  - hermítico
    - estándar, 202
- producto cartesiano, 5
- producto escalar, 201
- producto hermítico, 201
- producto vectorial, 231
- proyección
  - en una dirección, 112
  - ortogonal, 219
- raíz
  - de un polinomio, 13
- rango, 81
  - de una aplicación lineal, 139
  - de una matriz, 65
  - de una matriz escalonada, 53
- recta de regresión de mínimos cuadrados, 226
- Regla de Sarrus, 79
- resto, 14
- semejantes
  - matrices, 151
- simétrica
  - matriz, 57
- simétrico, 6
- sistema de ecuaciones lineales, 60
  - compatible, 61
  - determinado, 61
  - indeterminado, 61
  - incompatible, 61
- sistema generador, 37
- sistemas de ecuaciones lineales
  - equivalentes, 61
- solución aproximada, 223
- subespacio, 38
  - generado, 37
- subespacio ortogonal, 217
- subespacios vectoriales
  - complementarios, 111
- sucesión recurrente lineal, 164
- suma
  - de subespacios vectoriales, 108
- suprayectiva, 97
- Teorema
  - de la Dimensión, 103

Teorema de Pitágoras, 209

término

independiente de un polinomio, 12

transpuesta, 57

unitario, 206

uno

de un anillo, 9

de un cuerpo, 7

valor propio

de un endomorfismo, 152

variedad lineal, 117

vector

columna, 32

fila, 32

vector cero, 25

vector de coordenadas

en una base, 98

vector propio

de un endomorfismo, 152

de una matriz, 152

vectores, 25