ÁLGEBRA LINEAL

Y APLICACIONES

CENTRO UNIVERSITARIO DE LA DEFENSA
2009
ZARAGOZA

Jorge Martín Morales
M$^{\underline{\text{a}}}$ Victoria Sebastián Guerrero
Raquel Villacampa Gutiérrez

2.ª edición revisada y actualizada, 2023

Edita: Centro Universitario de la Defensa
Ctra. de Huesca s/n. 50090 Zaragoza
http://cud.unizar.es

Impresión: Edelvives Talleres Gráficos

Impreso en España
Printed in Spain

Depósito Legal: Z 526-2012
ISBN: 978-84-938411-6-4

A todos aquellos que nos guiaron

en el saber matemático

PRÓLOGO

La familia *Textos Docentes*, que vio la luz el curso pasado, se incrementa con un nuevo miembro, ÁLGEBRA LINEAL Y APLICACIONES, cuyo contenido se corresponde en líneas generales con la asignatura Matemáticas II del Grado en Ingeniería de Organización Industrial que se imparte en este Centro Universitario de la Defensa. Sin embargo, el texto va más allá de la asignatura, por lo que estamos seguros de que será de gran utilidad a estudiantes de otros Grados científicos o técnicos, así como al lector interesado en ampliar sus conocimientos.

Pocas disciplinas matemáticas hay que tengan tanta aplicación como el Álgebra Lineal. Desde herramienta para otras materias como Ecuaciones Diferenciales, a áreas como la Logística o la animación de películas con ordenador.

La elaboración de un texto docente por varios autores siempre es compleja, sobre todo cuando todos ellos tienen responsabilidad sobre la totalidad del texto y no sobre unos capítulos concretos, pues cada profesor tiene su particular visión de cómo presentarlo, en qué hacer más énfasis, qué orden llevar, etc. Como es de esperar, y he sido testigo presencial, esta obra no ha sido ajena a estas sanas pugnas, pero la experiencia docente, su excelente formación matemática y su objetivo de publicar un texto que sirviese de ayuda a sus alumnos han hecho posible que tengamos entre las manos un excelente texto, con rigor, cuidadosamente escrito, con numerosos problemas planteados y resueltos.

En suma, con gran satisfacción damos la bienvenida a este nuevo título, al mismo tiempo que agradezco personalmente a los autores, los doctores Jorge Martín Morales, María Victoria Sebastián Guerrero y Raquel Villacampa Gutiérrez el esfuerzo, compromiso y tiempo que han dedicado para que podamos disfrutar de esta obra.

Antonio Elipe
Catedrático Director

ÍNDICE GENERAL

INTRODUCCIÓN

Este libro nace con el objetivo de guiar a los estudiantes en su recorrido por el mundo del Álgebra Lineal, herramienta fundamental para muchas de las asignaturas que componen cualquier Grado científico o técnico.

La asignatura de Matemáticas II que se imparte en el Grado en Ingeniería de Organización Industrial exige a los alumnos un alto grado de abstracción mental así como una elevada capacidad de razonamiento. Por este motivo, los profesores que impartimos esta asignatura en cursos anteriores en el Centro Universitario de la Defensa creímos conveniente la elaboración de un texto que ayudara a los alumnos al desarrollo de las competencias anteriores.

Aunque no hemos dejado de lado el rigor propio de las Matemáticas, el texto está diseñado para fomentar una lectura que facilite la comprensión de los conceptos explicados, con numerosas aclaraciones y ejemplos resueltos. Hemos prestado además especial atención a aplicaciones del Álgebra Lineal, presentando situaciones o problemas de la Ciencia y la Ingeniería, y también alguno de la Defensa, que requieren de esta disciplina para su resolución. Como consecuencia de la aplicación del Álgebra Lineal a otras ramas, surge la necesidad de utilizar algún tipo de software matemático que nos ayude con las operaciones a realizar en la resolución de problemas. Con este propósito, se incluyen en el texto nociones básicas relacionadas con el manipulador simbólico *wxMaxima*.

El temario que presentamos es estándar para un primer curso de estudios de Ingeniería. El primer bloque, correspondiente a los Capítulos 1–4, introduce las herramientas propias del Álgebra Lineal: matrices, sistemas de ecuaciones y determinantes. En el Capítulo 5 se introduce el concepto fundamental sobre el que gira toda la asignatura: el espacio vectorial, y en el Capítulo 6 el concepto que abre las puertas a la geometría: el producto escalar. En el último bloque, Capítulos 7–9, se relacionan espacios vectoriales entre sí a través de aplicaciones lineales.

Los capítulos siguen todos un mismo esquema interno:

- Comenzamos con una introducción, en la que se hace un breve recorrido por los matemáticos más relevantes que han tenido relación con los temas tratados en el capítulo; igualmente, se presentan algunas ideas ya conocidas que pueden ayudar a contextualizar los conceptos.

- El grueso del capítulo se centra en el desarrollo tanto de los contenidos teóricos como de ejemplos resueltos, que muestran los procedimientos que los alumnos deben dominar para superar la asignatura.
- En el apartado de aplicaciones se citan con bastante detalle situaciones concretas en las que los contenidos del capítulo son de especial importancia.
- Otra parte del capítulo va dedicada a la enumeración y breve explicación de comandos del manipulador simbólico *wxMaxima* que realizan los mismos procedimientos, o similares, que los estudiados en el tema. Para una explicación más detallada de los mismos se recomienda utilizar el libro *Álgebra lineal con wxMaxima*, [6].
- El capítulo finaliza con una colección de problemas que los alumnos deben resolver. Junto a ellos, hemos incluido las soluciones de los mismos, para promover el trabajo autónomo de los estudiantes.

Nuestro deseo es que el texto sea de utilidad tanto para los alumnos que cursen la asignatura como para aquellos que ya la han superado, sirviendo siempre como libro de consulta. Esperamos al menos llegar a despertar interés por unas grandes desconocidas: el Álgebra Lineal en particular y las Matemáticas en general.

CAPÍTULO 1

Matrices I

En la lectura diaria de la prensa aparecen datos ordenados frecuentemente en forma de tabla. Este modo claro y sencillo de introducir la información facilita su comprensión. En Matemáticas esta ordenación corresponde al trabajo con matrices. Una matriz es un rectángulo de datos donde lo importante no son solo estos, sino la posición que ocupan.

El concepto de matriz es muy antiguo, ya en la literatura china hacia el 650 a.C. aparecen alusiones a un cuadrado 3×3 y hacia el 200 a.C. se usan matrices para resolver sistemas de ecuaciones.

Históricamente la aparición del concepto de matriz se asocia al matemático inglés J.J. Silvester hacia 1848, aunque fue Hamilton en 1853 quien desarrolló la teoría sobre matrices. En 1858, en su trabajo *A Memoir on the Theory of Matrices*, Cayley introdujo la notación matricial como forma abreviada de escribir un sistema de ecuaciones lineales.

En el presente capítulo se introduce el concepto de matriz, las definiciones y aritmética necesarias para utilizar la que va a ser la herramienta fundamental en este curso de Álgebra Lineal.

Las matrices se utilizan para resolver sistemas de ecuaciones, en cálculo numérico, para representar aplicaciones lineales, resolver ecuaciones diferenciales y en derivadas parciales. Además aparecen en disciplinas tan distintas como Física, Estadística, Geometría, Economía, Informática, etc. En la actualidad las matrices son parte esencial de cualquier lenguaje de programación, ya que la manera natural de suministrar los datos al ordenador es mediante tablas organizadas en filas y columnas.

1.1. DEFINICIONES

Definición 1.1. Una **matriz** de n filas y m columnas es una ordenación rectangular de la forma

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{im} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nm} \end{pmatrix}.$$

La matriz anterior se representa abreviadamente como $A = (a_{ij})$, $1 \leq i \leq n$, $1 \leq j \leq m$, y se dice que A es una matriz de **orden** o **dimensión** $n \times m$. El elemento a_{ij} se denomina **término** (o **componente**) ij y es el elemento de A que está en la fila i-ésima y columna j-ésima.

Observación. Las componentes de una matriz pueden ser de naturaleza muy diversa: números reales, complejos, polinomios o incluso otras matrices.

Notación. Se denota por $\mathrm{Mat}_{n\times m}(\mathbb{R})$ el conjunto de todas las matrices $A = (a_{ij})$ de dimensión $n \times m$ donde $a_{ij} \in \mathbb{R}$.

Definición 1.2. Sea $A = (a_{ij}) \in \mathrm{Mat}_{n\times m}(\mathbb{R})$. Atendiendo a su dimensión se distinguen algunos casos especiales:

- Se dice que A es una **matriz cuadrada** si $n = m$. En tal caso los elementos a_{ii}, $1 \leq i \leq n$, forman la **diagonal principal** de A.

  El conjunto de matrices cuadradas de dimensión $n \times n$ donde los elementos a_{ij} son números reales se denota por $\mathrm{Mat}_n(\mathbb{R})$.

- Se dice que A es una **matriz** (o **vector**) **fila** si su dimensión es $1 \times m$:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \end{pmatrix}.$$

- Se dice que A es una matriz **matriz** (o **vector**) **columna** si su dimensión es $n \times 1$:

$$A = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix}.$$

✓ **Ejemplo 1.** Estas son algunas matrices reales de distintos órdenes:

$$\underbrace{\begin{pmatrix} 1 & -4 & 5 & 2 \\ 0 & 4 & -1 & 2 \\ 3 & 0 & 2 & -1 \end{pmatrix}}_{\text{orden: } 3\times 4}, \quad \underbrace{\begin{pmatrix} \cos\frac{\pi}{12} & -3 \\ 0 & \ln 2 \end{pmatrix}}_{2\times 2}, \quad \underbrace{\begin{pmatrix} -1 & \pi & 3 \end{pmatrix}}_{1\times 3}.$$

$\diamond$

Definición 1.3. Dos matrices $A = (a_{ij})$ y $B = (b_{ij})$ son **iguales** si tienen la misma dimensión y los elementos que ocupan la misma posición son iguales, es decir, $a_{ij} = b_{ij}$, $\forall i, j$.

✓ **Ejemplo 2.** Las matrices

$$A = \begin{pmatrix} 1 & a & 3 \\ 2 & 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 0 & b \end{pmatrix}$$

son iguales si y solo si $a = 2$ y $b = 0$. $\diamond$

1.1.1. TIPOS ESPECIALES DE MATRICES

A continuación se introducen algunos tipos especiales de matrices:

- **Matriz nula** de dimensión $n \times m$ es aquella cuyos términos son todos iguales a 0, ($a_{ij} = 0$, $\forall i, j$). La matriz nula se denota por $O_{n\times m}$ o simplemente $\mathbf{0}$:

$$O_{n\times m} = \begin{pmatrix} 0 & \cdots & \overset{m}{\cdots} & 0 \\ \vdots\, n & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 \end{pmatrix} \in \mathrm{Mat}_{n\times m}(\mathbb{R}).$$

- **Matriz identidad** de orden n es una matriz cuadrada que cumple $a_{ii} = 1$, $\forall i = 1, \ldots, n$ y $a_{ij} = 0$, $\forall i \neq j$:

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \in \mathrm{Mat}_n(\mathbb{R}).$$

- $A \in \mathrm{Mat}_n(\mathbb{R})$ se dice **matriz diagonal** si $a_{ij} = 0$, $\forall i \neq j$:

$$\mathrm{diag}(a_{11}, a_{22}, \ldots, a_{nn}) = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}.$$

  Observar que la matriz identidad es una matriz diagonal cuyos elementos a_{ii} valen todos 1, es decir, $I_n = \mathrm{diag}(1, \ldots, 1)$.

- $A \in \mathrm{Mat}_{n\times m}(\mathbb{R})$ se dice **matriz triangular superior** si $a_{ij} = 0$, $\forall i > j$. Por ejemplo:

$$\begin{pmatrix} a_{11} & \cdots & \cdots & a_{1n} & \cdots & a_{1m} \\ 0 & \ddots & \ddots & \vdots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nn} & \cdots & a_{nm} \end{pmatrix}, \quad \begin{pmatrix} a_{11} & \cdots & \cdots & a_{1n} \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nn} \\ 0 & \cdots & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 \end{pmatrix}.$$

- $A \in \mathrm{Mat}_{n\times m}(\mathbb{R})$ se dice **matriz triangular inferior** si $a_{ij} = 0$, $\forall i < j$. Por ejemplo:

$$\begin{pmatrix} a_{11} & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & \cdots & a_{nn} & 0 & \cdots & 0 \end{pmatrix}, \quad \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ a_{m1} & \cdots & \cdots & a_{mm} \\ \vdots & \ddots & \ddots & \vdots \\ a_{n1} & \cdots & \cdots & a_{nm} \end{pmatrix}.$$

- $A \in \mathrm{Mat}_{n}(\mathbb{R})$ se dice **matriz simétrica** si $a_{ij} = a_{ji}$, $\forall i, j$. Las matrices simétricas son de la forma:

$$A = \begin{pmatrix} a_{11} & \boxed{\boxed{a_{12}}} & \cdots & \boxed{a_{1n}} \\ \boxed{\boxed{a_{12}}} & a_{22} & \cdots & \boxed{a_{2n}} \\ \vdots & \vdots & \ddots & \vdots \\ \boxed{a_{1n}} & \boxed{a_{2n}} & \cdots & a_{nn} \end{pmatrix}.$$

- $A \in \mathrm{Mat}_{n}(\mathbb{R})$ se dice **matriz antisimétrica** si $a_{ij} = -a_{ji}$, $\forall i, j$. En consecuencia todos los elementos de la diagonal son cero:

$$A = \begin{pmatrix} 0 & \boxed{\boxed{a_{12}}} & \cdots & \boxed{a_{1n}} \\ \boxed{\boxed{-a_{12}}} & 0 & \cdots & \boxed{a_{2n}} \\ \vdots & \vdots & \ddots & \vdots \\ \boxed{-a_{1n}} & \boxed{-a_{2n}} & \cdots & 0 \end{pmatrix}.$$

Observación. La matriz I_n es triangular superior, triangular inferior, diagonal y simétrica. La matriz nula $O_{n\times m}$ es triangular superior y triangular inferior; si además es cuadrada también es diagonal, simétrica y antisimétrica.

1.2. OPERACIONES CON MATRICES

Se trata en este apartado la aritmética de matrices, estudiando las operaciones que se pueden realizar así como algunas de las propiedades que se verifican.

1.2.1. SUMA DE MATRICES

Dos matrices se pueden sumar si tienen la misma dimensión, siendo el resultado otra matriz de la misma dimensión. La suma de matrices es una operación interna dentro del conjunto $\mathrm{Mat}_{n\times m}(\mathbb{R})$:

$$\begin{array}{rccc} (+): & \mathrm{Mat}_{n\times m}(\mathbb{R}) \times \mathrm{Mat}_{n\times m}(\mathbb{R}) & \longrightarrow & \mathrm{Mat}_{n\times m}(\mathbb{R}) \\ & (A,B) & \mapsto & A+B. \end{array}$$

Si $A=(a_{ij})$, $B=(b_{ij}) \in \mathrm{Mat}_{n\times m}(\mathbb{R})$ entonces la **matriz suma** se define como $A+B=(a_{ij}+b_{ij}) \in \mathrm{Mat}_{n\times m}(\mathbb{R})$. Es decir, la suma de matrices se realiza elemento a elemento como se indica a continuación:

$$\underbrace{\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}}_{A} + \underbrace{\begin{pmatrix} b_{11} & \cdots & b_{1m} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{nm} \end{pmatrix}}_{B} = \underbrace{\begin{pmatrix} a_{11}+b_{11} & \cdots & a_{1m}+b_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1}+b_{n1} & \cdots & a_{nm}+b_{nm} \end{pmatrix}}_{A+B}.$$

✓ **Ejemplo 3.**

$$\begin{pmatrix} 1 & 2 & -4 \\ 0 & -4 & -3/2 \end{pmatrix} + \begin{pmatrix} 4 & -5 & 0 \\ 1 & \sqrt{3} & 5 \end{pmatrix} = \begin{pmatrix} 5 & -3 & -4 \\ 1 & -4+\sqrt{3} & 7/2 \end{pmatrix}.$$

◇

La suma de matrices verifica las siguientes propiedades:

Propiedades. Sean $A,B,C \in \mathrm{Mat}_{n\times m}(\mathbb{R})$.

1. Asociativa: $(A+B)+C = A+(B+C)$.
2. Conmutativa: $A+B=B+A$.
3. Elemento neutro: la matriz nula $O_{n\times m} \in \mathrm{Mat}_{n\times m}(\mathbb{R})$ es el elemento neutro de la suma, es decir, verifica $A+O_{n\times m} = O_{n\times m}+A = A$.
4. Elemento opuesto: dada una matriz $A=(a_{ij}) \in \mathrm{Mat}_{n\times m}(\mathbb{R})$, su matriz opuesta es $-A=(-a_{ij})$ y verifica $A+(-A)=(-A)+A=O_{n\times m}$.

Observación. Debido a las propiedades anteriores se dice que $(\mathrm{Mat}_{n\times m}(\mathbb{R}),+)$ es un **grupo abeliano**. Otros ejemplos de grupos abelianos, es decir, conjuntos de elementos dotados de una operación interna que verifican las cuatro propiedades anteriores, son:

- $(\mathbb{Z},+)$, $(\mathbb{R},+)$, $(\mathbb{C},+)$.
- $(\mathbb{R}[x],+)$, donde $\mathbb{R}[x]$ denota el conjunto de los polinomios en la variable x con coeficientes reales.

Como conjunto con una operación interna que no tiene estructura de grupo abeliano se puede mencionar:

- $(\mathbb{N},+)$ no es un grupo abeliano pues el opuesto de un número natural no es otro número natural, sino un número entero negativo.

Resta de matrices: Como consecuencia de las propiedades anteriores se puede definir la resta de dos matrices como la suma de la primera con la opuesta de la segunda, es decir, si $A,B\in\mathrm{Mat}_{n\times m}(\mathbb{R})$, entonces $\boxed{A-B=A+(-B)}\in\mathrm{Mat}_{n\times m}(\mathbb{R})$.

✓ **Ejemplo 4.**

$$\begin{pmatrix}1 & 2 & -4\\ 0 & -4 & -3/2\end{pmatrix}-\begin{pmatrix}4 & -5 & 0\\ 1 & \sqrt{3} & 5\end{pmatrix}=\begin{pmatrix}-3 & 7 & -4\\ -1 & -4-\sqrt{3} & -13/2\end{pmatrix}.$$

◇

1.2.2. PRODUCTO POR UN NÚMERO REAL

Toda matriz se puede multiplicar por un número real cualquiera obteniéndose como resultado otra matriz de la misma dimensión que la inicial. El producto de un número real por una matriz es una operación externa:

$$\begin{array}{rccc}(\,\cdot\,): & \mathbb{R}\times\mathrm{Mat}_{n\times m}(\mathbb{R}) & \longrightarrow & \mathrm{Mat}_{n\times m}(\mathbb{R})\\ & (\lambda,A) & \mapsto & \lambda A.\end{array}$$

Si $\lambda\in\mathbb{R}$ y $A=(a_{ij})\in\mathrm{Mat}_{n\times m}(\mathbb{R})$ entonces $\lambda A=(\lambda a_{ij})\in\mathrm{Mat}_{n\times m}(\mathbb{R})$. Es decir, la multiplicación de un número real por una matriz se realiza multiplicando el número real en cuestión por cada elemento de la matriz:

$$\lambda\underbrace{\begin{pmatrix}a_{11} & \cdots & a_{1m}\\ \vdots & \ddots & \vdots\\ a_{n1} & \cdots & a_{nm}\end{pmatrix}}_{A}=\underbrace{\begin{pmatrix}\lambda a_{11} & \cdots & \lambda a_{1m}\\ \vdots & \ddots & \vdots\\ \lambda a_{n1} & \cdots & \lambda a_{nm}\end{pmatrix}}_{\lambda A}.$$

✓ **Ejemplo 5.**

$$5\cdot\begin{pmatrix}-1 & 4 & 3\\ 2 & \sqrt{2} & 0\end{pmatrix}=\begin{pmatrix}-5 & 20 & 15\\ 10 & 5\sqrt{2} & 0\end{pmatrix}.$$

◇

Propiedades. Sean $\lambda, \mu \in \mathbb{R}$ y $A, B \in \mathrm{Mat}_{n\times m}(\mathbb{R})$.

1. $(\lambda + \mu)A = \lambda A + \mu A$.
2. $\lambda(A + B) = \lambda A + \lambda B$.
3. $\lambda(\mu A) = (\lambda\mu)A$.
4. $1A = A$.

Casos particulares:

- $\lambda O_{n\times m} = O_{n\times m}, \quad \forall \lambda \in \mathbb{R}$.
- $0A = O_{n\times m}, \quad \forall A \in \mathrm{Mat}_{n\times m}(\mathbb{R})$.
- $\lambda I_n = \mathrm{diag}(\lambda, \ldots, \lambda)$.

1.2.3. PRODUCTO DE MATRICES

Dos matrices se pueden multiplicar si el número de columnas de la primera es igual al número de filas de la segunda. El resultado es otra matriz que tiene tantas filas como la primera y tantas columnas como la segunda:

$$\begin{array}{rccc} (\,\cdot\,): & \mathrm{Mat}_{n\times m}(\mathbb{R}) \times \mathrm{Mat}_{m\times p}(\mathbb{R}) & \longrightarrow & \mathrm{Mat}_{n\times p}(\mathbb{R}) \\ & (A, B) & \mapsto & AB. \end{array}$$

Si $A = (a_{ik}) \in \mathrm{Mat}_{n\times m}(\mathbb{R})$ y $B = (b_{kj}) \in \mathrm{Mat}_{m\times p}(\mathbb{R})$, entonces la **matriz producto** se define como $AB = (c_{ij}) \in \mathrm{Mat}_{n\times p}(\mathbb{R})$, donde $1 \leq i \leq n, \;\; 1 \leq j \leq p$, y el elemento c_{ij} tiene la siguiente expresión:

$$\boxed{c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{im}b_{mj} = \sum_{k=1}^{m} a_{ik}b_{kj}}$$

Es decir, el elemento c_{ij} se obtiene multiplicando la fila i-ésima de la matriz A por la columna j-ésima de la matriz B, tal y como se muestra a continuación:

$$\underbrace{\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ \multicolumn{3}{c}{\boxed{a_{i1} \quad \cdots \quad a_{im}}} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} b_{11} & \cdots & \boxed{b_{1j}} & \cdots & b_{1p} \\ \vdots & \ddots & \boxed{\vdots} & \ddots & \vdots \\ b_{m1} & \cdots & \boxed{b_{mj}} & \cdots & b_{mp} \end{pmatrix}}_{B} = \underbrace{\begin{pmatrix} c_{11} & \cdots & c_{1j} & \cdots & c_{1p} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ c_{i1} & \cdots & \boxed{c_{ij}} & \cdots & c_{ip} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ c_{n1} & \cdots & c_{nj} & \cdots & c_{np} \end{pmatrix}}_{AB}.$$

Para calcular la dimensión de la nueva matriz producto se puede seguir el siguiente esquema:

$$\underbrace{\underset{n\times\boxed{m}}{A}\quad\cdot\quad\underset{\boxed{m}\times p}{B}}_{n\times p}\quad=\quad\underset{n\times p}{AB}$$

✓ **Ejemplo 6.** Calcular el producto de las matrices:

$$A=\begin{pmatrix}3&-1\\5&4\\1&2\end{pmatrix},\quad B=\begin{pmatrix}0&5&-4&6\\2&-5&1&-3\end{pmatrix}.$$

Solución: La matriz A es de orden 3×2 y la matriz B es de orden 2×4. Siguiendo el esquema anterior,

$$\underset{3\times\boxed{2}}{A}\quad\cdot\quad\underset{\boxed{2}\times 4}{B}\quad=\quad\underset{3\times 4}{AB},$$

tiene sentido calcular AB y es de orden 3×4.

$$AB=\begin{pmatrix}3&-1\\5&4\\1&2\end{pmatrix}\begin{pmatrix}0&5&-4&6\\2&-5&1&-3\end{pmatrix}$$

$$=\begin{pmatrix}3\cdot 0+(-1)\cdot 2&3\cdot 5+(-1)\cdot(-5)&3\cdot(-4)+(-1)\cdot 1&3\cdot 6+(-1)\cdot(-3)\\5\cdot 0+4\cdot 2&5\cdot 5+4\cdot(-5)&5\cdot(-4)+4\cdot 1&5\cdot 6+4\cdot(-3)\\1\cdot 0+2\cdot 2&1\cdot 5+2\cdot(-5)&1\cdot(-4)+2\cdot 1&1\cdot 6+2\cdot(-3)\end{pmatrix}$$

$$=\begin{pmatrix}-2&20&-13&21\\8&5&-16&18\\4&-5&-2&0\end{pmatrix}.$$

◇

Observación. Dadas dos matrices cualesquiera A y B, no siempre se pueden multiplicar. Se muestran algunos casos concretos:

- Sean $A\in\mathrm{Mat}_{2\times 3}(\mathbb{R})$ y $B\in\mathrm{Mat}_{1\times 4}(\mathbb{R})$. No tiene sentido hacer AB ni BA.
- Sean $A\in\mathrm{Mat}_{2\times 3}(\mathbb{R})$ y $B\in\mathrm{Mat}_{3\times 4}(\mathbb{R})$. Se puede realizar $AB\in\mathrm{Mat}_{2\times 4}(\mathbb{R})$ pero no tiene sentido hacer BA.
- Sean $A\in\mathrm{Mat}_{2\times 3}(\mathbb{R})$ y $B\in\mathrm{Mat}_{3\times 2}(\mathbb{R})$. Se puede realizar tanto $AB\in\mathrm{Mat}_{2}(\mathbb{R})$ como $BA\in\mathrm{Mat}_{3}(\mathbb{R})$.

El producto de matrices verifica las siguientes propiedades:

Propiedades. Sean A, B y C tres matrices de dimensiones adecuadas para que las operaciones que se indican se puedan realizar. Sea $\lambda \in \mathbb{R}$ un número real. Se verifica:

1. Asociativa: $(AB)C = A(BC)$.
2. Distributiva respecto de la suma (izquierda y derecha):
$$A(B+C) = AB + AC,$$
$$(A+B)C = AC + BC.$$
3. Producto y producto por número real: $\lambda(AB) = (\lambda A)B = A(\lambda B)$.
4. Sea $A \in \mathrm{Mat}_{n\times m}(\mathbb{R})$ y sean I_n, I_m las matrices identidad de orden n y m respectivamente. Se verifica:
    - $A\,I_m = A$. Se dice que I_m es el **elemento neutro a derecha** para las matrices $\mathrm{Mat}_{n\times m}(\mathbb{R})$.
    - $I_n A = A$. Se dice que I_n es el **elemento neutro a izquierda** para las matrices $\mathrm{Mat}_{n\times m}(\mathbb{R})$.
    - Si $n = m$ existe **elemento neutro**, la matriz identidad I_n. En este caso, el elemento neutro a derecha coincide con el elemento neutro a izquierda.

Observaciones. El producto de matrices presenta algunas diferencias al compararlo con el producto de números reales.

1. Existen divisores de cero, es decir, puede ocurrir $AB = \mathbf{0}$ siendo $A, B \neq \mathbf{0}$. Por tanto de la igualdad $AB = AC$ no se puede concluir $B = C$.

   ✓ **Ejemplo 7.**
$$\underbrace{\begin{pmatrix} 3 & 2 \\ 6 & 4 \end{pmatrix}}_{\neq O_{2\times 2}} \underbrace{\begin{pmatrix} 2 & -4 \\ -3 & 6 \end{pmatrix}}_{\neq O_{2\times 2}} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$
   ◇

2. El producto de matrices no cumple la propiedad conmutativa, es decir, en general, suponiendo que tenga sentido hacer tanto AB como BA,
$$\boxed{AB \neq BA}$$
   Se muestran algunos casos:
    - Sean $A \in \mathrm{Mat}_{2\times 3}(\mathbb{R})$ y $B \in \mathrm{Mat}_{3\times 2}(\mathbb{R})$, entonces se puede realizar tanto AB como BA. Sin embargo, AB y BA no pueden ser matrices iguales porque tienen distinta dimensión: $AB \in \mathrm{Mat}_2(\mathbb{R})$ mientras que $BA \in \mathrm{Mat}_3(\mathbb{R})$.

- Ni siquiera en el caso de matrices cuadradas se tiene la propiedad conmutativa:

  ✓ **Ejemplo 8.** Sean las matrices $A, B \in \mathrm{Mat}_2(\mathbb{R})$:

  $$A = \begin{pmatrix} 3 & -1 \\ 2 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} -2 & 5 \\ 4 & 1 \end{pmatrix}.$$

  Entonces $AB \neq BA$ pues:

  $$AB = \begin{pmatrix} -10 & 14 \\ 0 & 11 \end{pmatrix}, \quad BA = \begin{pmatrix} 4 & 7 \\ 14 & -3 \end{pmatrix}.$$

  ◇

Por las observaciones anteriores tiene sentido hacer la siguiente definición:

Definición 1.4. Dadas dos matrices cuadradas $A, B \in \mathrm{Mat}_n(\mathbb{R})$, se dice que **conmutan** si cumplen la propiedad conmutativa, es decir, si

$$AB = BA.$$

Se define el **conmutador** de A y B como $[A, B] = AB - BA$. Observar que A y B conmutan si y solo si su conmutador es cero:

$$[A, B] = 0 \Longleftrightarrow AB = BA.$$

✓ **Ejemplo 9.** Las siguientes matrices conmutan:

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} -1 & 3 \\ 3 & -1 \end{pmatrix} \quad \Longrightarrow \quad AB = BA = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}.$$

◇

Observaciones. Como consecuencia de las propiedades anteriores se puede concluir:

- Las matrices I_n y $O_{n\times n}$ conmutan con todas las matrices $A \in \mathrm{Mat}_n(\mathbb{R})$.
- Toda matriz cuadrada $A \in \mathrm{Mat}_n(\mathbb{R})$ conmuta consigo misma.

Nota. En el conjunto de matrices cuadradas $\mathrm{Mat}_n(\mathbb{R})$, el producto de matrices es una operación interna. Esta operación junto con la suma dotan a $\mathrm{Mat}_n(\mathbb{R})$ de estructura de **anillo** (no conmutativo si $n \geq 2$).

1.2.4. POTENCIAS DE MATRICES

Debido a la propiedad asociativa del producto de matrices tiene sentido definir las potencias de una matriz cuadrada $A \in \mathrm{Mat}_n(\mathbb{R})$.

Definición 1.5. Sea $A \in \mathrm{Mat}_n(\mathbb{R})$ y k un número entero positivo. La **potencia k-ésima** de A, A^k, es el producto de la matriz A por sí misma k-veces:

$$A^k = \underbrace{AA \dots A}_{k}.$$

✓ **Ejemplo 10.** Calcular las potencias A^2, A^3, A^4 de la siguiente matriz y deducir la potencia k-ésima:

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

Solución:

$$A^2 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}, \qquad A^3 = AA^2 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix},$$

$$A^4 = AA^3 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 4 & 1 \end{pmatrix}, \ \dots \ , \ A^k = \begin{pmatrix} 1 & 0 \\ k & 1 \end{pmatrix}.$$

La expresión de A^k se deduce fácilmente y se puede demostrar que es correcta por inducción.

◇

Observaciones.

- Al igual que sucede en el caso de los números reales, se utiliza el convenio $A^0 = I_n$, siempre que A no sea la matriz nula.
- $A^1 = A$.
- $(I_n)^k = I_n, \ \forall k \in \mathbb{Z}, \ k \geq 0.$
- $(O_{n\times n})^k = O_{n\times n}, \ \forall k \in \mathbb{Z}, \ k > 0.$

Proposición 1.6. Sea $A \in \mathrm{Mat}_n(\mathbb{R})$ una matriz diagonal, $A = \mathrm{diag}(a_{11}, \dots, a_{nn})$. La potencia k-ésima de A es también una matriz diagonal cuya expresión es:

$$A^k = \mathrm{diag}\ (a_{11}^k, \dots, a_{nn}^k) = \begin{pmatrix} a_{11}^k & 0 & \cdots & 0 \\ 0 & a_{22}^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn}^k \end{pmatrix}, \quad \forall k \in \mathbb{Z}, \, k > 0.$$

1.3. MATRIZ INVERSA

Definición 1.7. Sea $A \in \text{Mat}_n(\mathbb{R})$ una matriz cuadrada. Se dice que A es **invertible** o **regular** si existe una matriz $B \in \text{Mat}_n(\mathbb{R})$ tal que

$$\boxed{AB = BA = I_n}$$

La matriz B se dice **inversa** de A y se denota por A^{-1}. Las matrices no invertibles se dicen **singulares**.

Observación. La matriz A^{-1} anterior si existe es única, por tanto, normalmente se habla de **la** matriz inversa de A.

✓ **Ejemplo 11.** Sean las matrices

$$A = \begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & -5 \\ -1 & 3 \end{pmatrix}.$$

Entonces:

$$AB = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad BA = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Luego B es la inversa de A y se escribe $B = A^{-1}$.

◇

Propiedades. Sean $A, B \in \text{Mat}_n(\mathbb{R})$ dos matrices regulares. Sea $\lambda \in \mathbb{R}$, $\lambda \neq 0$. Se verifican las siguientes propiedades:

1. $(A^{-1})^{-1} = A$.
2. $(\lambda A)^{-1} = \frac{1}{\lambda} A^{-1}$.
3. $(AB)^{-1} = B^{-1}A^{-1}$. Si se tienen varias matrices regulares $A_1, \ldots, A_k \in \text{Mat}_n(\mathbb{R})$ entonces $(A_1 \ldots A_k)^{-1} = A_k^{-1} \ldots A_1^{-1}$.

Observación. A diferencia de las otras operaciones, la inversa no se comporta bien con respecto a la suma. En el Capítulo 2 se presenta cómo calcular la inversa de una matriz y se comprueba que en general $(A + B)^{-1} \neq A^{-1} + B^{-1}$.

Proposición 1.8. Sea $A \in \text{Mat}_n(\mathbb{R})$ una matriz diagonal, $A = \text{diag}(a_{11}, \ldots, a_{nn})$ con $a_{ii} \neq 0$. La inversa de A es también una matriz diagonal cuya expresión es:

$$A^{-1} = \text{diag}\left(\frac{1}{a_{11}}, \ldots, \frac{1}{a_{nn}}\right) = \begin{pmatrix} \frac{1}{a_{11}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{1}{a_{nn}} \end{pmatrix}.$$

1.4. MATRIZ TRASPUESTA

Dada una matriz A, se puede construir otra matriz intercambiando filas por columnas.

Definición 1.9. Sea $A \in \mathrm{Mat}_{n\times m}(\mathbb{R})$ una matriz. Se denomina **matriz traspuesta** de A a la matriz $B \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ tal que $b_{ij} = a_{ji}$, $\forall i, j$. Se denota por $B = A^t$.

✓ **Ejemplo 12.**

$$A = \begin{pmatrix} 2 & 1 & 3 & 4 \\ 1 & -1 & 2 & -3 \\ 3 & 0 & 5 & 1 \end{pmatrix} \quad \Longrightarrow \quad A^t = \begin{pmatrix} 2 & 1 & 3 \\ 1 & -1 & 0 \\ 3 & 2 & 5 \\ 4 & -3 & 1 \end{pmatrix}.$$

◇

Se han definido anteriormente las matrices simétricas y antisimétricas. Estas se pueden caracterizar en función de la matriz traspuesta.

Proposición 1.10. Sea $A \in \mathrm{Mat}_n(\mathbb{R})$, entonces:

- A es simétrica si y solo si $A^t = A$.
- A es antisimétrica si y solo si $A^t = -A$.

El siguiente resultado recoge el comportamiento de la trasposición con respecto a la aritmética matricial.

Propiedades. Sean A y B dos matrices de dimensiones adecuadas para que se puedan realizar las operaciones que se indican. Sea $\lambda \in \mathbb{R}$ un número real. Se verifican las siguientes propiedades:

1. $(A^t)^t = A$.
2. $(A+B)^t = A^t + B^t$.
3. $(\lambda A)^t = \lambda A^t$.
4. $(AB)^t = B^t A^t$. Si se tienen varias matrices $A_1, \ldots, A_k$ de dimensiones adecuadas para que se pueda realizar el producto, entonces $(A_1 \ldots A_k)^t = A_k^t \ldots A_1^t$.
5. Si $A \in \mathrm{Mat}_n(\mathbb{R})$ es regular: $(A^t)^{-1} = (A^{-1})^t$.

1.5. TRAZA DE UNA MATRIZ

Definición 1.11. Sea $A = (a_{ij}) \in \mathrm{Mat}_n(\mathbb{R})$ una matriz cuadrada. Se define la **traza** como la siguiente aplicación:

$$\begin{array}{rcl} \mathrm{tr} : \mathrm{Mat}_n(\mathbb{R}) & \longrightarrow & \mathbb{R}, \\ A & \mapsto & \mathrm{tr}\, A = a_{11} + a_{22} + \cdots + a_{nn}. \end{array}$$

Es decir, $\mathrm{tr}\, A$ es la suma de los elementos de la diagonal principal de A.

✓ **Ejemplo 13.**

$$\mathrm{tr} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = 1 + 5 + 9 = 15.$$

◇

Propiedades. Sean $A, B \in \mathrm{Mat}_n(\mathbb{R})$ dos matrices cuadradas de orden n.

1. $\mathrm{tr}\, A^t = \mathrm{tr}\, A$.
2. $\mathrm{tr}(A + B) = \mathrm{tr}\, A + \mathrm{tr}\, B$.
3. $\mathrm{tr}(\lambda A) = \lambda\, \mathrm{tr}\, A,\ \forall \lambda \in \mathbb{R}$.
4. Sean ahora $A \in \mathrm{Mat}_{n\times m}(\mathbb{R})$ y $B \in \mathrm{Mat}_{m\times n}(\mathbb{R})$, entonces $\mathrm{tr}(AB) = \mathrm{tr}(BA)$.

Observación. La traza no se comporta bien con respecto al producto. En general se tiene que $\boxed{\mathrm{tr}(AB) \neq (\mathrm{tr}\, A)(\mathrm{tr}\, B)}$

1.6. MATRICES POR BLOQUES

A veces es útil subdividir una matriz en matrices más pequeñas o bloques, por ejemplo, de la siguiente forma:

$$\left(\begin{array}{cccc|c} 1 & -1 & 3 & 4 & 0 \\ 2 & 0 & 6 & 1 & -2 \\ \hline 0 & 1 & 3 & -2 & 0 \end{array}\right), \quad \left(\begin{array}{c|c|c} 4 & 0 & 1 \\ -1 & 1 & -1 \\ 0 & 2 & 1 \\ 3 & -3 & 0 \\ -2 & 0 & 1 \end{array}\right), \quad \left(\begin{array}{c|cc|c} 3 & 0 & 0 & 0 \\ \hline 0 & 17 & -7 & 0 \\ 0 & 15 & 9 & 0 \\ \hline 0 & 0 & 0 & 2 \end{array}\right).$$

Se puede considerar que una **matriz por bloques** es aquella cuyos elementos son otras matrices. Observar que la subdivisión de una matriz en bloques no es única. Un caso especial consiste en considerar las filas o columnas de la matriz como bloques (ver la segunda matriz anterior). En el caso de matrices cuadradas otra descomposición importante es la conocida como **diagonal por bloques**, cuyos elementos diagonales son matrices cuadradas (ver tercera matriz anterior).

A la hora de operar con este tipo de matrices se utiliza la aritmética matricial explicada anteriormente, considerando que ahora los bloques juegan el papel de elementos, por tanto hay que prestar especial atención a las dimensiones de dichos bloques.

1.6.1. SUMA

Para sumar matrices $A, B \in \mathrm{Mat}_{n\times m}(\mathbb{R})$ por bloques se debe realizar la misma partición en cada una de ellas.

✓ **Ejemplo 14.** Sumar las matrices A y B utilizando la descomposición en bloques que se indica:

$$A = \left(\begin{array}{cc|c} 1 & 2 & 0 \\ 3 & 4 & 0 \\ \hline 0 & 0 & 5 \end{array}\right), \qquad B = \left(\begin{array}{cc|c} 6 & 7 & 0 \\ 8 & 9 & 0 \\ \hline 0 & 0 & 10 \end{array}\right).$$

Solución:

$$A + B = \left(\begin{array}{cc|c} 1 & 2 & 0 \\ 3 & 4 & 0 \\ \hline 0 & 0 & 5 \end{array}\right) + \left(\begin{array}{cc|c} 6 & 7 & 0 \\ 8 & 9 & 0 \\ \hline 0 & 0 & 10 \end{array}\right) = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} + \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} =$$

$$= \begin{pmatrix} A_{11} + B_{11} & A_{12} + B_{12} \\ A_{21} + B_{21} & A_{22} + B_{22} \end{pmatrix}.$$

- $A_{11} + B_{11} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} + \begin{pmatrix} 6 & 7 \\ 8 & 9 \end{pmatrix} = \begin{pmatrix} 7 & 9 \\ 11 & 13 \end{pmatrix}.$
- $A_{12} + B_{12} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$
- $A_{21} + B_{21} = \begin{pmatrix} 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \end{pmatrix}.$
- $A_{22} + B_{22} = (\,5\,) + (\,10\,) = (\,15\,).$

Así, la matriz buscada es:

$$A + B = \left(\begin{array}{cc|c} 7 & 9 & 0 \\ 11 & 13 & 0 \\ \hline 0 & 0 & 15 \end{array}\right).$$

Se observa que la matriz resultante es diagonal por bloques con la misma subdivisión que se tenía al principio en A y B.

◇

1.6.2. PRODUCTO

Para multiplicar matrices particionadas en bloques se debe tener en cuenta que los tamaños de los bloques permitan dicha operación. Entonces se aplican las reglas de multiplicación matricial utilizando los bloques a modo de elementos. A continuación se muestra cómo proceder con un ejemplo.

✓ **Ejemplo 15.** Multiplicar las matrices A y B utilizando la descomposición en bloques que se indica:

$$A = \left(\begin{array}{cccc|c} 1 & -1 & 3 & 4 & 0 \\ 2 & 0 & 6 & 1 & -2 \\ \hline 0 & 1 & 3 & -2 & 0 \end{array}\right), \qquad B = \left(\begin{array}{c|cc} 4 & 0 & 1 \\ -1 & 1 & -1 \\ 0 & 2 & 1 \\ 3 & -3 & 0 \\ \hline -2 & 0 & 1 \end{array}\right).$$

Solución:

$$\left(\begin{array}{cccc|c} 1 & -1 & 3 & 4 & 0 \\ 2 & 0 & 6 & 1 & -2 \\ \hline 0 & 1 & 3 & -2 & 0 \end{array}\right) \left(\begin{array}{c|cc} 4 & 0 & 1 \\ -1 & 1 & -1 \\ 0 & 2 & 1 \\ 3 & -3 & 0 \\ \hline -2 & 0 & 1 \end{array}\right) = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} =$$

$$= \begin{pmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{pmatrix}.$$

- $A_{11}B_{11} + A_{12}B_{21} = \begin{pmatrix} 1 & -1 & 3 & 4 \\ 2 & 0 & 6 & 1 \end{pmatrix} \begin{pmatrix} 4 \\ -1 \\ 0 \\ 3 \end{pmatrix} + \begin{pmatrix} 0 \\ -2 \end{pmatrix} (-2) = \begin{pmatrix} 17 \\ 15 \end{pmatrix}.$

- $A_{11}B_{12} + A_{12}B_{22} = \begin{pmatrix} 1 & -1 & 3 & 4 \\ 2 & 0 & 6 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & -1 \\ 2 & 1 \\ -3 & 0 \end{pmatrix} + \begin{pmatrix} 0 \\ -2 \end{pmatrix} \begin{pmatrix} 0 & 1 \end{pmatrix} =$

  $= \begin{pmatrix} -7 & 5 \\ 9 & 6 \end{pmatrix}.$

- $A_{21}B_{11} + A_{22}B_{21} = \begin{pmatrix} 0 & 1 & 3 & -2 \end{pmatrix} \begin{pmatrix} 4 \\ -1 \\ 0 \\ 3 \end{pmatrix} + (\,0\,)(-2) = (-7).$

- $A_{21}B_{12}+A_{22}B_{22}=\begin{pmatrix}0 & 1 & 3 & -2\end{pmatrix}\begin{pmatrix}0 & 1\\ 1 & -1\\ 2 & 1\\ -3 & 0\end{pmatrix}+(\,0\,)\begin{pmatrix}0 & 1\end{pmatrix}=\begin{pmatrix}13 & 2\end{pmatrix}.$

Así, la matriz buscada es:

$$AB=\left(\begin{array}{c|cc}17 & -7 & 5\\ 15 & 9 & 6\\ \hline -7 & 13 & 2\end{array}\right).$$

◇

Observación. Para multiplicar matrices diagonales por bloques las matrices deben tener la misma partición. En particular, para calcular la potencia de una matriz de este tipo basta hallar la potencia de cada bloque diagonal:

$$A=\left(\begin{array}{c|c|c}A_{11} & \cdots & \mathbf{0}\\ \hline \vdots & \ddots & \vdots\\ \hline \mathbf{0} & \cdots & A_{nn}\end{array}\right)\quad\Longrightarrow\quad A^k=\left(\begin{array}{c|c|c}A_{11}^k & \cdots & \mathbf{0}\\ \hline \vdots & \ddots & \vdots\\ \hline \mathbf{0} & \cdots & A_{nn}^k\end{array}\right).$$

1.6.3. INVERSA

En general, no es sencillo dar fórmulas para calcular la inversa de una matriz regular por bloques. Sin embargo, cuando esta es diagonal (por bloques) y cada bloque de la diagonal es una matriz regular, la situación es la siguiente:

$$A=\left(\begin{array}{c|c|c}A_{11} & \cdots & \mathbf{0}\\ \hline \vdots & \ddots & \vdots\\ \hline \mathbf{0} & \cdots & A_{nn}\end{array}\right)\quad\Longrightarrow\quad A^{-1}=\left(\begin{array}{c|c|c}A_{11}^{-1} & \cdots & \mathbf{0}\\ \hline \vdots & \ddots & \vdots\\ \hline \mathbf{0} & \cdots & A_{nn}^{-1}\end{array}\right).$$

✓ **Ejemplo 16.** Calcular la inversa de A usando la descomposición en bloques que se indica:

$$A=\left(\begin{array}{cc|c}1 & 2 & 0\\ 3 & 4 & 0\\ \hline 0 & 0 & 5\end{array}\right).$$

Solución:

$$A=\left(\begin{array}{cc|c}1 & 2 & 0\\ 3 & 4 & 0\\ \hline 0 & 0 & 5\end{array}\right)=\begin{pmatrix}A_{11} & \mathbf{0}\\ \mathbf{0} & A_{22}\end{pmatrix}\quad\Longrightarrow\quad A^{-1}=\begin{pmatrix}A_{11}^{-1} & \mathbf{0}\\ \mathbf{0} & A_{22}^{-1}\end{pmatrix}.$$

$$A_{11}^{-1}=\begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix}^{-1}=\begin{pmatrix}-2 & 1\\ 3/2 & -1/2\end{pmatrix};\quad A_{22}^{-1}=\left(5\right)^{-1}=\left(1/5\right).$$

Por tanto la matriz buscada es:

$$A^{-1} = \left(\begin{array}{cc|c} -2 & 1 & 0 \\ 3/2 & -1/2 & 0 \\ \hline 0 & 0 & 1/5 \end{array}\right).$$

◇

1.6.4. TRASPOSICIÓN

Para trasponer una matriz por bloques, aparte de intercambiar filas de bloques por columnas, hay que trasponer cada uno de los bloques que forman la matriz:

$$A = \left(\begin{array}{c|c|c} A_{11} & \cdots & A_{1m} \\ \hline \vdots & \ddots & \vdots \\ \hline A_{n1} & \cdots & A_{nm} \end{array}\right) \quad \Longrightarrow \quad A^t = \left(\begin{array}{c|c|c} A_{11}^t & \cdots & A_{n1}^t \\ \hline \vdots & \ddots & \vdots \\ \hline A_{1m}^t & \cdots & A_{nm}^t \end{array}\right).$$

✓ **Ejemplo 17.** Calcular la matriz traspuesta de A usando la descomposición en bloques indicada:

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

Solución:

$$A = \left(\begin{array}{cccc|c} 1 & -1 & 3 & 4 & 0 \\ 2 & 0 & 6 & 1 & -2 \\ \hline 0 & 1 & 3 & -2 & 0 \end{array}\right) = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \quad \Longrightarrow \quad A^t = \begin{pmatrix} A_{11}^t & A_{21}^t \\ A_{12}^t & A_{22}^t \end{pmatrix}.$$

- $A_{11}^t = \begin{pmatrix} 1 & -1 & 3 & 4 \\ 2 & 0 & 6 & 1 \end{pmatrix}^t = \begin{pmatrix} 1 & 2 \\ -1 & 0 \\ 3 & 6 \\ 4 & 1 \end{pmatrix}; \quad A_{21}^t = \begin{pmatrix} 0 & 1 & 3 & -2 \end{pmatrix}^t = \begin{pmatrix} 0 \\ 1 \\ 3 \\ -2 \end{pmatrix}.$
- $A_{12}^t = \begin{pmatrix} 0 \\ -2 \end{pmatrix}^t = \begin{pmatrix} 0 & -2 \end{pmatrix}; \quad A_{22}^t = \begin{pmatrix} 0 \end{pmatrix}^t = \begin{pmatrix} 0 \end{pmatrix}.$

Luego la matriz pedida es:

$$A^t = \left(\begin{array}{cc|c} 1 & 2 & 0 \\ -1 & 0 & 1 \\ 3 & 6 & 3 \\ 4 & 1 & -2 \\ \hline 0 & -2 & 0 \end{array}\right).$$

◇

1.7. APLICACIONES: SISTEMA BRAILLE

El braille es un sistema de lectura y escritura táctil pensado para personas ciegas. Fue creado por el francés Louis Braille en el siglo XIX. La nota curiosa es que este sistema fue diseñado a partir de otro creado en 1821 por Charles Barbier, oficial del ejército napoleónico, para que los soldados se comunicaran en silencio en plena oscuridad. Sin embargo, este método era tan complicado que fue rechazado por la milicia, hasta que Braille (quien quedó ciego a la edad de tres años por accidente) realizó varias mejoras y publicó el primer libro con su novedoso sistema en 1829.

Es un sistema basado en la combinación lógica de seis puntos en relieve sobre un espacio o celdilla (“cajetín”). Los seis puntos dispuestos en dos columnas forman el llamado *signo generador*. A cada punto se le asigna un número según un orden determinado dando lugar a una matriz 3 × 2 de la siguiente forma:

$$\boxed{\begin{array}{cc}\cdot & \cdot\\ \cdot & \cdot\\ \cdot & \cdot\end{array}} \Longrightarrow \begin{pmatrix}1 & 4\\ 2 & 5\\ 3 & 6\end{pmatrix}.$$

Con la presencia y ausencia de estos puntos, organizados en grupos o series, se generan un total de 64 ($= 2^6$) combinaciones diferentes, suficientes para crear las letras del alfabeto, los signos de puntuación y otros símbolos básicos, incluyendo la que no tiene ningún punto, que se utiliza como espacio en blanco.

El tamaño y distribución de los 6 puntos que forman el signo generador, no es un capricho sino el fruto de la experiencia de Louis Braille. Las terminaciones nerviosas de la yema del dedo están capacitadas para captar este tamaño en particular.

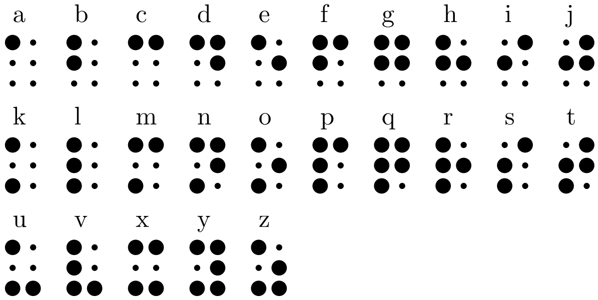


Figura 1.1: Alfabeto braille básico.

Se observa que los símbolos correspondientes a la primera fila ocupan solo los cuatro puntos superiores del signo generador. Los que corresponden a la segunda fila son iguales a los de la primera, pero se le agrega el punto inferior izquierdo (salvo la ñ que es propia del idioma español), y en los de la tercera se agregan los dos inferiores. Los puntos negros pequeños son los puntos del signo generador que no están en relieve, solo se han dibujado para una mejor comprensión de cada símbolo.

Para representar algunos signos especiales es preciso utilizar más de un carácter braille, pues las 64 combinaciones resultan insuficientes. Además, este sistema permite la notación musical, de forma que los músicos ciegos pueden leer las partituras en braille. Por ejemplo, para representar las letras mayúsculas se antepone el carácter braille formado por los puntos 4 y 6:

$$\begin{array}{cc} \cdot & \bullet \\ \cdot & \cdot \\ \cdot & \bullet \end{array}$$

Los números se hacen con el prefijo formado por los puntos 3, 4, 5 y 6 antes de las diez primeras letras. De este modo se indica que es un número en vez de una letra:

$$\begin{array}{cc} \cdot & \bullet \\ \cdot & \bullet \\ \bullet & \bullet \end{array}$$

Es importante destacar que el braille no es un idioma, sino un código. Por lo tanto, las particularidades y la sintaxis serán las mismas que para los caracteres propios de cada idioma.

En un ciego el área táctil no abarca más allá de la superficie de contacto del dedo lector y el papel, por lo que la velocidad lectora de un buen lector ciego es siempre inferior a la de un buen lector vidente. Un universitario ciego con buen nivel lector difícilmente alcanza las 180 palabras por minuto, mientras que un vidente en idénticas condiciones puede superar casi con seguridad el doble de esa velocidad. Si a un vidente se le disminuyera su área perceptiva a un solo carácter, las velocidades entre ciegos y videntes se asemejarían. Para que un ciego adquiera una buena velocidad lectora (alrededor de 150 palabras por minuto) es necesario combinar una correcta técnica de lectura bimanual con las estrategias de contextualización.

1.8. COMANDOS DE *wxMaxima*

A lo largo del texto se incluye una enumeración y breve explicación de algunos comandos básicos del manipulador simbólico *wxMaxima* relacionados con la temática presentada en cada capítulo. De esta manera se pretende que los alumnos aprendan a manejar este programa al mismo tiempo que asimilan la materia en cuestión.

El programa *wxMaxima* es gratuito y de libre acceso. Está distribuido bajo la licencia “GNU General Public License” y se puede descargar directamente desde la página web `wxmaxima.sourceforge.net/`.

Definición y manipulación de matrices

- Introducir una lista L:

  ```
  -->     L:[a1,a2,...,ar];
  ```

  Los elementos “`ai's`” pueden ser números u objetos de otra naturaleza.

- Introducir una matriz A:

  ```
  -->     A:matrix([a11,a12,...,a1m],...,[an1,an2,...,anm]);
  ```

  Observar que cada lista representa una fila de la matriz. Este comando se puede obtener directamente desde el menú “`\Álgebra\Introducir Matriz`”.

- Dimensión de la matriz A:

  ```
  -->     matrix_size(A);
  ```

  Devuelve una lista de la forma $[n, m]$, donde n es el número de filas y m es el número de columnas de la matriz A.

- Acceder al elemento (i, j) de la matriz A:

  ```
  -->     A[i,j];
  ```

- Acceder a la fila i-ésima de la matriz A:

  ```
  -->     row(A,i);
  ```

- Acceder a la columna j-ésima de la matriz A:

  ```
  -->     col(A,j);
  ```

- Obtener una submatriz de la matriz A:

```
-->        submatrix(i1,...,ip,A,j1,...,jq);
```

  El resultado es una submatriz de A en la que se han eliminado las filas “`i1,...,ip`” y las columnas “`j1,...,jq`”. Si, por ejemplo, no se quiere eliminar ninguna columna, la sintaxis queda así: “`submatrix(i1,...,ip,A)`”.

- Añadir filas a la matriz A:

```
-->        addrow(A,fila1,...,filak);
```

  Cada fila “`fila1,...,filak`” se introduce como una lista que se añade al final de la matriz A.

- Añadir columnas a una matriz:

```
-->        addcol(A,columna1,...,columnak);
```

  Cada columna “`columna1,...,columnak`” se introduce como una lista que se añade a la derecha de la matriz A.

Matrices especiales

- Matriz nula de dimensión $n \times m$:

```
-->        zeromatrix(n,m);
```

- Matriz identidad de dimensión n:

```
-->        ident(n);
```

- Matriz diagonal con elementos “`d1,...,dn`” en la diagonal:

```
-->        diag_matrix(d1,...,dn);
```

Operaciones con matrices

- Sumar dos matrices A y B de la misma dimensión:

```
-->        A+B;
```

- Restar dos matrices A y B de la misma dimensión:

```
-->     A-B;
```

- Producto de un número real “`a`” por una matriz A:

```
-->     a*A;
```

- Producto de dos matrices A y B de dimensiones adecuadas para que se pueda realizar el productor AB:

```
-->     A.B;
```

- Potencia de una matriz:

```
-->     A^^k;
```

Observaciones. (¡¡ CUIDADO !!)

- El producto de matrices se realiza con un punto “`.`” en lugar de con el símbolo “`*`”. Este último multiplica las matrices elemento a elemento.
- Las potencias de la matriz A se efectúan con el comando “`A^^k`” poniendo un doble circunflejo. Con uno solo se calculan las potencias de cada elemento de la matriz.

Otras operaciones

- Calcular la inversa de una matriz A cuadrada y regular:

```
-->     invert(A);
```

  Comando al que también se accede desde “`\Álgebra\Invertir Matriz`”.

- Matriz traspuesta de A:

```
-->     transpose(A);
```

  Comando al que también se accede desde “`\Álgebra\Trasponer Matriz`”.

- Traza de una matriz de A:

```
-->     load(nchrpl);
        mattrace(A);
```

  Observar que para que este comando funcione hay que cargar antes el paquete “`nchrpl`”.

Matrices por bloques

En caso de trabajar con matrices con bloques, es necesario ejecutar previamente los siguientes comandos:

```
-->     matrix_element_mult: ".";
        matrix_element_transpose: transpose;
```

- Introducir una matriz A por bloques:

```
-->     A:matrix([A11,A12,...,A1m],...,[An1,An2,...,Anm]);
```

  donde cada “`Aij`” es un bloque que se ha definido previamente con el comando “`matrix`”.

- Comprobar si la matriz A es una matriz por bloques o no:

```
-->     blockmatrixp(A);
```

- Deshacer el formato de bloques de una matriz A:

```
-->     mat_unblocker(A);
```

1.9. EJERCICIOS PROPUESTOS

Operaciones con Matrices

1. Calcular, siempre que sea posible, AB, BA, AC, CA, BC y CB, siendo:

$$A = \begin{pmatrix} 1 & -2 \end{pmatrix}, \qquad B = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \qquad C = \begin{pmatrix} 1 & -2 & 0 \end{pmatrix}.$$

2. Se consideran las siguientes matrices:

$$A = \begin{pmatrix} 1 & -4 & 2 \\ -1 & 4 & -2 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 2 \\ -1 & 3 \\ 5 & -2 \end{pmatrix}, \qquad C = \begin{pmatrix} 2 & 2 \\ 1 & -1 \\ 1 & -3 \end{pmatrix}.$$

Calcular $B + C$, AB, BA, CA, AC, $A(B + C)$, $AB + AC$, $A(2B - 3C)$.

3. Calcular $AB - BA$ y decir si las matrices A y B conmutan:

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

4. Dadas las matrices

$$A = \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix}, \qquad B = \begin{pmatrix} -2 & -1 \\ 3 & 1 \end{pmatrix},$$

calcular A^t, B^t, $(AB)^t$, A^tB^t y B^tA^t. Comprobar que $(AB)^t = B^tA^t$ pero que $(AB)^t \neq A^tB^t$.

5. Sean A, $B \in \mathrm{Mat}_2(\mathbb{R})$ dos matrices simétricas. Demostrar que AB no es necesariamente simétrica. Comprobarlo con el ejemplo:

$$A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} -2 & 1 \\ 1 & 1 \end{pmatrix}.$$

6. Sean A, $B \in \mathrm{Mat}_2(\mathbb{R})$ dos matrices simétricas. Probar que AB es simétrica si y solo si $AB = BA$.

7. Sea $A \in \mathrm{Mat}_n(\mathbb{R})$ una matriz cuadrada:

a) Demostrar que si A es simétrica, entonces A^2 es simétrica.

b) Demostrar que si A es simétrica, entonces A^p es simétrica para todo número natural p.

c) Buscar un contraejemplo para demostrar que el recíproco de a) no es cierto.

8. Sea $A \in \mathrm{Mat}_{n\times m}(\mathbb{R})$ una matriz cualquiera. Demostrar que AA^t y A^tA son matrices simétricas.

9. Sea $A \in \mathrm{Mat}_2(\mathbb{R})$. ¿Qué forma ha de tener la matriz A para que se verifique que $A^tA = AA^t$?

10. Encontrar todas las matrices reales $A \in \mathrm{Mat}_2(\mathbb{R})$ que verifican $AA^t = \mathbf{0}$. ¿Se puede generalizar el resultado anterior para matrices cuadradas de cualquier dimensión?

11. Obtener las matrices reales que conmutan con

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

12. Si $A = \begin{pmatrix} 2 & -1 \\ -2 & 3 \end{pmatrix}$ y $B = \begin{pmatrix} 7 & 6 \\ 9 & 8 \end{pmatrix}$, hallar las matrices C y D tales que $AC = B$ y $DA = B$.

13. Sean $A, B \in \mathrm{Mat}_n(\mathbb{R})$ dos matrices cuadradas:

a) Comprobar que las identidades algebraicas

$$(A+B)^2 = A^2 + 2AB + B^2 \quad \text{y} \quad (A+B)(A-B) = A^2 - B^2$$

no son ciertas para las matrices $A = \begin{pmatrix} 1 & -1 \\ 0 & 2 \end{pmatrix}$ y $B = \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix}$.

b) Modificar el segundo miembro de esas identidades para obtener fórmulas válidas para todas las matrices cuadradas A y B.

c) ¿Para qué matrices son válidas las identidades establecidas en a) ?

14. Sean las matrices

$$A_1 = \begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 3 & 1 \\ 0 & 2 \\ -1 & 4 \end{pmatrix}, \quad B_1 = \begin{pmatrix} 1 & 4 \end{pmatrix}, \quad B_2 = \begin{pmatrix} -1 & 2 \\ 0 & 1 \end{pmatrix}.$$

Evaluar cada una de las expresiones siguientes trabajando por bloques:

$$\begin{pmatrix} A_1 & A_2 \end{pmatrix}^t, \qquad \begin{pmatrix} A_1 & A_2 \end{pmatrix} \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}, \qquad \begin{pmatrix} B_2 \\ B_1 \end{pmatrix} A_2^t.$$

Comprobar los resultados obtenidos sin usar descomposición por bloques.

15. Demostrar que toda matriz cuadrada se puede escribir como suma de una matriz simétrica y otra antisimétrica. Aplicar este resultado para descomponer la matriz:

$$\begin{pmatrix} 1 & -1 & 3 \\ 3 & 0 & 7 \\ 5 & -3 & -1 \end{pmatrix}.$$

16. a) Calcular el siguiente producto de matrices:

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

b) Expresar la ecuación de la circunferencia $x^2+y^2+Ax+By+C = 0$ como producto de tres matrices igualadas a cero.

Potencias de Matrices

17. Comprobar que $A = \begin{pmatrix} 1 & 2 \\ 3 & 5 \end{pmatrix}$ satisface la ecuación $X^2 - 6X - I_2 = \mathbf{0}$.

18. Encontrar las potencias A^2, A^3, B^2, B^3, C^2, C^3, siendo:

$$A = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad C = AB.$$

19. Calcular A^2, A^3, A^4 y deducir la potencia A^n, donde

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

20. Sea $A = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$. Comprobar que $A^2 = 2A - I_2$ y calcular A^{100}.

21. Sea $A = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}$. Comprobar que $A^2 = \begin{pmatrix} \cos 2\theta & -\operatorname{sen} 2\theta \\ \operatorname{sen} 2\theta & \cos 2\theta \end{pmatrix}$, calcular A^3 y deducir A^n.

22. Calcular la potencia n-ésima de la matriz $A = \begin{pmatrix} \alpha & 1 \\ 0 & \alpha \end{pmatrix}$, donde $\alpha \in \mathbb{R}$.

23. Calcular las potencias de A, sabiendo que A es una matriz por bloques de la forma $A = \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix}$, donde

$$A_1 = (1), \qquad A_2 = \begin{pmatrix} 0 & 0 \end{pmatrix}, \qquad A_3 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad A_4 = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

24. Calcular A^2 y A^3 usando una descomposición en bloques adecuada:

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

25. Una matriz cuadrada $A \in \mathrm{Mat}_n(\mathbb{R})$ se dice **idempotente** si $A^2 = A$. ¿Qué ocurre con el resto de potencias de A?

26. Encontrar los valores de $a, b, c \in \mathbb{R}$ para los cuales la matriz

$$A = \begin{pmatrix} a & 0 \\ b & c \end{pmatrix}$$

es idempotente.

27. Determinar las condiciones que debe verificar una matriz cuadrada de orden 2 para que sea idempotente.

28. Sea $A \in \mathrm{Mat}_3(\mathbb{R})$ una matriz con todos los elementos en y debajo de la diagonal principal nulos. Probar que $A^3 = \mathbf{0}$.

29. Una matriz cuadrada $A \in \mathrm{Mat}_n(\mathbb{R})$ se dice **nilpotente** si existe un entero positivo k tal que $A^k = \mathbf{0}$. Más aún, A se dice **nilpotente de índice** p si p es el menor entero positivo que cumple la propiedad. Comprobar que la siguiente matriz es nilpotente de índice 3:

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

30. Comprobar que la matriz

$$A = \begin{pmatrix} 2 & 1 \\ -4 & -2 \end{pmatrix}$$

verifica la relación $A^2 = \mathbf{0}$. Encontrar las condiciones que deben verificar los coeficientes de una matriz no nula $A \in \mathrm{Mat}_2(\mathbb{R})$ para que sea nilpotente de índice 2, es decir, $A^2 = \mathbf{0}$.

1.10. SOLUCIONES A LOS EJERCICIOS PROPUESTOS

Operaciones con Matrices

1. Las siguientes operaciones NO tienen sentido porque las dimensiones de las matrices no son adecuadas: AC, CA, CB. Para las restantes:

$$AB = (\,0\,), \qquad BA = \begin{pmatrix} 2 & -4 \\ 1 & -2 \end{pmatrix}, \qquad BC = \begin{pmatrix} 2 & -4 & 0 \\ 1 & -2 & 0 \end{pmatrix}.$$

2. Todas las operaciones indicadas se pueden realizar y sus resultados son:

$$B+C = \begin{pmatrix} 3 & 4 \\ 0 & 2 \\ 6 & -5 \end{pmatrix}, \qquad AB = \begin{pmatrix} 15 & -14 \\ -15 & 14 \end{pmatrix}, \qquad BA = \begin{pmatrix} -1 & 4 & -2 \\ -4 & 16 & -8 \\ 7 & -28 & 14 \end{pmatrix},$$

$$CA = \begin{pmatrix} 0 & 0 & 0 \\ 2 & -8 & 4 \\ 4 & -16 & 8 \end{pmatrix}, \qquad AC = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \qquad A(B+C) = \begin{pmatrix} 15 & -14 \\ -15 & 14 \end{pmatrix},$$

$$AB + AC = \begin{pmatrix} 15 & -14 \\ -15 & 14 \end{pmatrix}, \qquad A(2B - 3C) = \begin{pmatrix} 30 & -28 \\ -30 & 28 \end{pmatrix}.$$

3. Las matrices no conmutan. Los productos AB y BA y el conmutador son:

$$AB = \begin{pmatrix} -2 & 9 & 3 \\ 6 & 8 & 4 \\ -1 & 11 & 4 \end{pmatrix}, \quad BA = \begin{pmatrix} 7 & 11 & 13 \\ 0 & -6 & -4 \\ 6 & 6 & 9 \end{pmatrix}, \quad AB - BA = \begin{pmatrix} -9 & -2 & -10 \\ 6 & 14 & 8 \\ -7 & 5 & -5 \end{pmatrix}.$$

4. Se realizan los cálculos oportunos y por observación se comprueba que $(AB)^t = B^t A^t$ pero $(AB)^t \neq A^t B^t$.

$$A^t = \begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix}, \qquad B^t = \begin{pmatrix} -2 & 3 \\ -1 & 1 \end{pmatrix}, \qquad (AB)^t = \begin{pmatrix} 4 & 11 \\ 1 & 4 \end{pmatrix},$$

$$A^t B^t = \begin{pmatrix} -1 & 2 \\ -7 & 9 \end{pmatrix}, \quad B^t A^t = \begin{pmatrix} 4 & 11 \\ 1 & 4 \end{pmatrix}.$$

5. La matriz $AB = \begin{pmatrix} -3 & 0 \\ 3 & 0 \end{pmatrix}$ no es simétrica.

6. Se trata de una cuestión teórica. Aplicar la Proposición 1.10 a la matriz $C = AB$ y usar las propiedades de la trasposición.

7. Los apartados a) y b) son cuestiones teóricas. A continuación se da una breve indicación de su resolución:

a) Aplicar la caracterización, A simétrica si y solo si $A^t = A$, a la matriz $B = A^2$ y usar las propiedades de la trasposición.

b) Análogamente, deducir que A^3, A^4 y A^5 verifican la misma propiedad, concluyendo por inducción que también es cierta para A^p siendo p un número natural cualquiera.

c) Todas las matrices de la forma $A = \begin{pmatrix} a & b \\ c & -a \end{pmatrix}$ con $b \neq c$.

8. Aplicar la Proposición 1.10 a las matrices $B = A^tA$ y $C = AA^t$. Usar también las propiedades de la trasposición.

9. Se tienen dos posibilidades para la matriz A:

$$A = \begin{pmatrix} a & b \\ b & d \end{pmatrix}, \quad A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}.$$

10. La única matriz cuadrada que verifica el enunciado, en cualquier dimensión, es la matriz nula.

11. Las matrices que conmutan con A son de la forma $B = \begin{pmatrix} a & 0 & 0 \\ b & a & 0 \\ c & b & a \end{pmatrix}$, con $a, b, c \in \mathbb{R}$.

12. Las matrices son $C = \begin{pmatrix} 15/2 & 13/2 \\ 8 & 7 \end{pmatrix}$ y $D = \begin{pmatrix} 33/4 & 19/4 \\ 43/4 & 25/4 \end{pmatrix}$.

13. a) Se realizan las operaciones pertinentes y se observa que no se cumplen las igualdades propuestas:

$$(A+B)^2 = \begin{pmatrix} 3 & -6 \\ 6 & 15 \end{pmatrix}, \qquad A^2 + 2AB + B^2 = \begin{pmatrix} 2 & -7 \\ 7 & 16 \end{pmatrix},$$

$$(A+B)(A-B) = \begin{pmatrix} 1 & -2 \\ -4 & -1 \end{pmatrix}, \qquad A^2 - B^2 = \begin{pmatrix} 0 & -3 \\ -3 & 0 \end{pmatrix}.$$

b) $(A+B)^2 = A^2 + AB + BA + B^2, \quad (A+B)(A-B) = A^2 - AB + BA - B^2.$

c) Las identidades son válidas para matrices que conmutan, es decir, $AB = BA$.

14. Cada una de las siguientes operaciones hay que realizarlas de dos maneras distintas, directamente y usando la descomposición en bloques:

$$\begin{pmatrix} A_1 & A_2 \end{pmatrix}^t = \begin{pmatrix} A_1^t \\ A_2^t \end{pmatrix} = \left(\begin{array}{ccc} 1 & -2 & 2 \\ \hline 3 & 0 & -1 \\ 1 & 2 & 4 \end{array}\right),$$

$$\begin{pmatrix} A_1 & A_2 \end{pmatrix} \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = A_1B_1 + A_2B_2 = \begin{pmatrix} -2 & 11 \\ -2 & -6 \\ 3 & 10 \end{pmatrix},$$

$$\begin{pmatrix} B_2 \\ B_1 \end{pmatrix} A_2^t = \begin{pmatrix} B_2A_2^t \\ B_1A_2^t \end{pmatrix} = \left(\begin{array}{ccc} -1 & 4 & 9 \\ 1 & 2 & 4 \\ \hline 7 & 8 & 15 \end{array}\right).$$

15. Sean $B = \frac{1}{2}(A + A^t)$ y $C = \frac{1}{2}(A - A^t)$. Se comprueba que $A = B + C$ y que B es simétrica y C antisimétrica. La descomposición para la matriz A es:

$$A = \begin{pmatrix} 1 & 1 & 4 \\ 1 & 0 & 2 \\ 4 & 2 & -1 \end{pmatrix} + \begin{pmatrix} 0 & -2 & -1 \\ 2 & 0 & 5 \\ 1 & -5 & 0 \end{pmatrix}.$$

16. a) $x^2 + y^2 - 2x + 4y + 3$.

b) $\begin{pmatrix} x & y & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & A/2 \\ 0 & 1 & B/2 \\ A/2 & B/2 & C \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = 0.$

Potencias de Matrices

17. Se comprueba que:

$$A^2 - 6A - I_2 = \underbrace{\begin{pmatrix} 7 & 12 \\ 18 & 31 \end{pmatrix}}_{A^2} - \underbrace{\begin{pmatrix} 6 & 12 \\ 18 & 30 \end{pmatrix}}_{6A} - \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}}_{I_2} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

18. $A^2 = A^3 = A$; $B^2 = I_2$, $B^3 = B$; $C = \begin{pmatrix} 1/2 & -1/2 \\ 1/2 & -1/2 \end{pmatrix}$, $C^2 = C^3 = \mathbf{0}$.

19.

$$A^2 = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}, \ A^3 = \begin{pmatrix} 1 & 3 & 6 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix}, \ A^4 = \begin{pmatrix} 1 & 4 & 10 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{pmatrix}, \ A^n = \begin{pmatrix} 1 & n & \frac{n(n+1)}{2} \\ 0 & 1 & n \\ 0 & 0 & 1 \end{pmatrix}.$$

Para calcular el término 13 de A^n utilizar progresiones aritméticas.

20. $A^{100} = 100A - 99I_2 = \begin{pmatrix} 1 & 0 \\ -100 & 1 \end{pmatrix}$.

21. Utilizando las fórmulas de $\operatorname{sen}(\alpha+\beta)$ y $\cos(\alpha+\beta)$ se deduce que $A^n = \begin{pmatrix} \cos n\theta & -\operatorname{sen} n\theta \\ \operatorname{sen} n\theta & \cos n\theta \end{pmatrix}$.

22. $A^n = \begin{pmatrix} \alpha^n & n\alpha^{n-1} \\ 0 & \alpha^n \end{pmatrix}$.

23. $A^n = \begin{pmatrix} A_1 & A_2 \\ A_3 & 2^{n-1}A_4 \end{pmatrix} = \left(\begin{array}{c|cc} 1 & 0 & 0 \\ \hline 1 & 2^{n-1} & -2^{n-1} \\ 1 & -2^{n-1} & 2^{n-1} \end{array}\right)$.

24. Se elige la descomposición en bloques que se indica y se calcula A^2 y A^3:

$$A = \left(\begin{array}{c|c} B & \mathbf{0} \\ \hline \mathbf{0} & C \end{array}\right) = \left(\begin{array}{cc|cc} 1 & 2 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 3 \end{array}\right) \Longrightarrow$$

$$A^2 = \left(\begin{array}{c|c} B^2 & \mathbf{0} \\ \hline \mathbf{0} & C^2 \end{array}\right) = \left(\begin{array}{cc|cc} 1 & 4 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 9 \end{array}\right), \ A^3 = \left(\begin{array}{c|c} B^3 & \mathbf{0} \\ \hline \mathbf{0} & C^3 \end{array}\right) = \left(\begin{array}{cc|cc} 1 & 6 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 27 \end{array}\right).$$

25. $A^k = A$, $\forall k \in \mathbb{N}$.

26. Posibilidades para la matriz $A = \begin{pmatrix} a & 0 \\ b & c \end{pmatrix}$:

$$A = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad A = \begin{pmatrix} 0 & 0 \\ b & 1 \end{pmatrix}, \quad A = \begin{pmatrix} 1 & 0 \\ b & 0 \end{pmatrix}.$$

27. Posibilidades para una matriz genérica $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$:

$$A = I_2, \quad A = \mathbf{0}, \quad A = \begin{pmatrix} a & b \\ c & 1-a \end{pmatrix} \text{ cumpliendo } \ a^2 + bc = a.$$

28. Se considera una matriz $A \in \mathrm{Mat}_3(\mathbb{R})$ verificando la propiedad del enunciado y se realizan los cálculos oportunos:

$$A = \begin{pmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{pmatrix} \quad \Longrightarrow \quad A^2 = \begin{pmatrix} 0 & 0 & a\,c \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad A^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

29.

$$A = \begin{pmatrix} 0 & 2 & -1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \quad \Longrightarrow \quad A^2 = \begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad A^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

30. Para una matriz no nula de orden 2, $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$:

$$A^2 = \mathbf{0} \iff A = \begin{pmatrix} a & b \\ c & -a \end{pmatrix} \text{ con } \ a^2 + bc = 0, \text{ siendo } a, b, c \neq 0.$$

CAPÍTULO 2

Matrices II. Operaciones Elementales

El objetivo de este capítulo es transformar una matriz $A \in \mathrm{Mat}_{n\times m}(\mathbb{R})$ en otra matriz más simple, con forma escalonada, que conserve ciertas propiedades de la matriz inicial. Las operaciones realizadas para dicha transformación solo pueden ser de tres tipos determinados y se denominan operaciones elementales. Estas se recogen a su vez en unas matrices denominadas matrices elementales. El proceso plantea las herramientas necesarias para la resolución de sistemas de ecuaciones que se estudian en el próximo capítulo.

El origen del algoritmo de las operaciones elementales y el concepto de matriz escalonada y su evolución hasta el llamado método de eliminación gaussiana se remonta a los Libros VII y VIII del Zhui Zhang Suan Shu (o El Arte de la Matemática en Nueve Libros, 152 a.C.), obra de la matemática oriental escrita por el insigne científico y hombre de estado Chuan Tsanom. Se trata de una obra clásica que se ha ido modificando a lo largo del tiempo, hasta lograr convertirse en el texto fundamental en el cual se basaron los científicos matemáticos en sus investigaciones entre los siglos VII y X.

No fue hasta el nacimiento del Álgebra Lineal, en un renovado intento de los matemáticos por encontrar métodos generales para la resolución de sistemas de ecuaciones lineales con varias incógnitas, cuando los conceptos anteriores vuelven a tomar importancia. Fue en 1805 cuando el matemático alemán K.F. Gauss formalizó el método que actualmente se conoce como método de eliminación gaussiana para encontrar todas las soluciones de un sistema de m ecuaciones lineales con n incógnitas.

2.1. MATRICES ESCALONADAS

Las matrices que se introducen en esta sección son básicas para el desarrollo de los siguientes capítulos del texto.

Definición 2.1. Una matriz $A \in \mathrm{Mat}_{n\times m}(\mathbb{R})$ se dice **matriz escalonada por filas** si verifica:

- Si hay filas enteras nulas, estas se sitúan en la parte inferior de la matriz.
- El primer elemento no nulo de cada fila, llamado **pivote**, es 1.
- El pivote de cada fila no nula está a la derecha del pivote de la fila anterior.
- Los elementos que aparecen por debajo del pivote en su misma columna son todos nulos.

✓ **Ejemplo 1.** Se muestra a continuación un ejemplo de matriz escalonada por filas, donde $*$ denota un valor real cualquiera. Los 1's representan los pivotes y se puede observar que no todos ocupan las posiciones diagonales de la matriz y además debajo de ellos todos los elementos son 0:

$$\begin{pmatrix}
\boxed{\mathbf{1}} & * & * & * & * & * & * & * & * \\
0 & 0 & \boxed{\mathbf{1}} & * & * & * & * & * & * \\
0 & 0 & 0 & \boxed{\mathbf{1}} & * & * & * & * & * \\
0 & 0 & 0 & 0 & 0 & \boxed{\mathbf{1}} & * & * & * \\
0 & 0 & 0 & 0 & 0 & 0 & \boxed{\mathbf{1}} & * & * \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}.$$

◇

Definición 2.2. Una matriz $A \in \mathrm{Mat}_{n\times m}(\mathbb{R})$ se dice **matriz escalonada reducida por filas** si es escalonada y además los elementos que aparecen en la misma columna que el pivote, por encima y por debajo, son todos cero.

✓ **Ejemplo 2.** La siguiente matriz 3×4 es escalonada reducida por filas:

$$\begin{pmatrix}
\mathbf{1} & 0 & 0 & 0 & 5 \\
0 & 0 & \mathbf{1} & 0 & 3 \\
0 & 0 & 0 & \mathbf{1} & 4 \\
0 & 0 & 0 & 0 & 0
\end{pmatrix}.$$

◇

2.2. OPERACIONES Y MATRICES ELEMENTALES

El objetivo de este capítulo es transformar una matriz cualquiera en una escalonada o escalonada reducida por filas. Para ello se plantean tres tipos de transformaciones que se deben realizar a la matriz. Se denominan *operaciones elementales.*

Definición 2.3. Dada $A \in \mathrm{Mat}_{n\times m}(\mathbb{R})$, para llevarla a una matriz escalonada o escalonada reducida por filas, se pueden realizar tres tipos de transformaciones denominadas **operaciones elementales**. Estas operaciones son:

- Tipo I: Intercambiar dos filas.
- Tipo II: Multiplicar una fila por un número real no nulo.
- Tipo III: Sumar a una fila un múltiplo de otra.

✓ **Ejemplo 3.**

$$\begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix} \xrightarrow[1^a fila \leftrightarrow 2^a fila]{\text{Tipo I}} \begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix} \xrightarrow[2^a fila - 2\cdot 1^a fila]{\text{Tipo III}} \begin{pmatrix} 1 & 3 \\ 0 & -5 \end{pmatrix} \xrightarrow[(-\frac{1}{5})\cdot 2^a fila]{\text{Tipo II}} \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}.$$

◇

A continuación se definen tres tipos de matrices que recogen los efectos de aplicar las operaciones elementales a la matriz $A \in \mathrm{Mat}_{n\times m}(\mathbb{R})$, y se denominan *matrices elementales.*

Definición 2.4. Una matriz de orden $n \times n$ se dice **elemental** si se puede obtener a partir de la matriz identidad I_n efectuando una única operación elemental en sus filas.

Matrices elementales de tipo I

Se denota P_{ij} a la matriz que se obtiene intercambiando las filas i, j de la matriz identidad.

$$P_{ij} = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 0 & \cdots & \boxed{1} & & \\ & & \vdots & \ddots & \vdots & & \\ & & \boxed{1} & \cdots & 0 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \begin{array}{l} \\ \\ \longrightarrow \text{posición } ij \\ \\ \\ \\ \\ \end{array}$$

$$\text{posición } ji \longleftarrow \boxed{1}$$

Matrices elementales de tipo II

Se denota $P_i(\alpha)$ a la matriz que se obtiene al multiplicar los elementos de la fila i-ésima de la matriz identidad por $\alpha \in \mathbb{R}, \alpha \neq 0$.

$$P_i(\alpha) = \operatorname{diag}(1,\dots,1,\overset{i)}{\alpha},1,\dots,1) = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & \boxed{\alpha} & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix} \longrightarrow \text{posición } ii$$

Matrices elementales de tipo III

Se denota $P_{ij}(\alpha)$ a la matriz que se obtiene al sumar a la fila i-ésima de la matriz identidad la j-ésima multiplicada por $\alpha \in \mathbb{R}$, siendo $i \neq j$.

$$\text{posición } ij \longleftarrow P_{ij}(\alpha) = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & \boxed{\alpha} & & \ddots & \\ & & & & 1 \end{pmatrix}$$

Nota. Observar que $P_{ij} = P_{ji}$, pero $P_{ij}(\alpha) \neq P_{ji}(\alpha)$.

A continuación se muestra cómo se transforma una matriz A al multiplicarla por una matriz elemental.

Proposición 2.5. Sea $A \in \mathrm{Mat}_{n\times m}(\mathbb{R})$ una matriz cualquiera. Sean P_{ij}, $P_i(\alpha)$, $P_{ij}(\alpha) \in \mathrm{Mat}_n(\mathbb{R})$ las matrices elementales que se acaban de definir, entonces:

- $P_{ij}A$ es la matriz que resulta al intercambiar las filas i, j de A, es decir, es el resultado de aplicar a A una operación elemental de tipo I.
- $P_i(\alpha)A$ es la matriz que resulta al multiplicar la fila i de A por α, siendo $\alpha \in \mathbb{R}$, $\alpha \neq 0$, es decir, es el resultado de aplicar a A una operación elemental de tipo II.
- $P_{ij}(\alpha)A$ es la matriz que resulta al sumar a la fila i de A la fila j multiplicada por α, $i \neq j$, es decir, es el resultado de aplicar a A una operación elemental de tipo III.

✓ **Ejemplo 4** (tipo I)**.** Sea $A \in \mathrm{Mat}_3(\mathbb{R})$. Se permutan las filas 1 y 3:

$$A = \begin{pmatrix} 0 & 3 & 5 \\ 2 & 0 & 4 \\ 1 & 1 & 3 \end{pmatrix} \xrightarrow[1^a fila \leftrightarrow 3^a fila]{} \begin{pmatrix} 1 & 1 & 3 \\ 2 & 0 & 4 \\ 0 & 3 & 5 \end{pmatrix}.$$

Lo mismo se puede lograr multiplicando la matriz elemental P_{13} por A:

$$P_{13}A = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 3 & 5 \\ 2 & 0 & 4 \\ 1 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 3 \\ 2 & 0 & 4 \\ 0 & 3 & 5 \end{pmatrix}.$$

◇

✓ **Ejemplo 5** (tipo II)**.** Sea $A \in \mathrm{Mat}_3(\mathbb{R})$. Se multiplica por 2 la tercera fila:

$$A = \begin{pmatrix} 0 & 3 & 5 \\ 2 & 0 & 4 \\ 1 & 1 & 3 \end{pmatrix} \xrightarrow[2\cdot 3^a fila]{} \begin{pmatrix} 0 & 3 & 5 \\ 2 & 0 & 4 \\ 2 & 2 & 6 \end{pmatrix}.$$

El mismo efecto se obtiene multiplicando la matriz elemental $P_3(2)$ por A:

$$P_3(2)A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 0 & 3 & 5 \\ 2 & 0 & 4 \\ 1 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 0 & 3 & 5 \\ 2 & 0 & 4 \\ 2 & 2 & 6 \end{pmatrix}.$$

◇

✓ **Ejemplo 6** (tipo III)**.** Sea $A \in \mathrm{Mat}_3(\mathbb{R})$. Se suma a la tercera fila la segunda multiplicada por $-\frac{1}{2}$:

$$A = \begin{pmatrix} 0 & 3 & 5 \\ 2 & 0 & 4 \\ 1 & 1 & 3 \end{pmatrix} \xrightarrow[3^a fila + (-\frac{1}{2})\cdot 2^a fila]{} \begin{pmatrix} 0 & 3 & 5 \\ 2 & 0 & 4 \\ 0 & 1 & 1 \end{pmatrix}.$$

Se llega a la misma matriz multiplicando la matriz elemental $P_{32}(-\frac{1}{2})$ por A como se muestra continuación:

$$P_{32}\left(-\tfrac{1}{2}\right)A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1/2 & 1 \end{pmatrix} \begin{pmatrix} 0 & 3 & 5 \\ 2 & 0 & 4 \\ 1 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 0 & 3 & 5 \\ 2 & 0 & 4 \\ 0 & 1 & 1 \end{pmatrix}.$$

◇

El siguiente resultado muestra cómo calcular la inversa de una matriz elemental P. Observar que esta inversa es a su vez una matriz elemental que deshace la transformación que realiza P.

Proposición 2.6. Las matrices elementales son regulares. Sus inversas son también matrices elementales y vienen dadas por:

- $(P_{ij})^{-1} = P_{ij}, \quad \forall i \neq j.$
- $(P_i(\alpha))^{-1} = P_i(\alpha^{-1}) = P_i\left(\frac{1}{\alpha}\right), \quad \forall \alpha \in \mathbb{R},\ \alpha \neq 0.$
- $(P_{ij}(\alpha))^{-1} = P_{ij}(-\alpha), \quad \forall \alpha \in \mathbb{R}.$

Observación. Las transformaciones elementales introducidas en la Definición 2.3 se podrían haber realizado por columnas, obteniendo así las operaciones elementales (de tipo I, II y III) por columnas. Más concretamente, si $A \in \mathrm{Mat}_{n\times m}(\mathbb{R})$ y P_{ij}, $P_i(\alpha)$, $P_{ij}(\alpha) \in \mathrm{Mat}_m(\mathbb{R})$ denotan las matrices elementales (de orden m), entonces:

- $A\,(P_{ij})^t$ es la matriz que resulta de intercambiar las columnas i, j de A.
- $A\,(P_i(\alpha))^t$ es la matriz que resulta de multiplicar la columna i de A por α, siendo $\alpha \in \mathbb{R}$, $\alpha \neq 0$.
- $A\,(P_{ij}(\alpha))^t$ es la matriz que resulta de sumar a la columna i de A la columna j multiplicada por α.

Notar que en este caso hay que colocar las traspuestas de las matrices elementales a la derecha de la matriz A. Este hecho es de utilidad en el Capítulo 9. Por otra parte, las matrices elementales de tipo I y II son simétricas, mientras que las de tipo III no. Más concretamente, se verifica:

$$(P_{ij})^t = P_{ij}, \qquad (P_i(\alpha))^t = P_i(\alpha), \qquad (P_{ij}(\alpha))^t = P_{ji}(\alpha).$$

✓ **Ejemplo 7.** Para sumar a la tercera columna la segunda multiplicada por $-\frac{1}{2}$ en la matriz del Ejemplo 4, hay que hacer la siguiente multiplicación:

$$A\,(P_{32}(-\tfrac{1}{2}))^t = \begin{pmatrix} 0 & 3 & 5 \\ 2 & 0 & 4 \\ 1 & 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1/2 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 3 & 7/2 \\ 2 & 0 & 4 \\ 1 & 1 & 5/2 \end{pmatrix}.$$

◇

2.3. MÉTODO DE ELIMINACIÓN GAUSSIANA

Para llevar una matriz A a una forma escalonada se realizan las operaciones elementales que se acaban de definir en un cierto orden de modo que los pivotes van apareciendo progresivamente, empezando por la primera fila. Este proceso se conoce como **método de eliminación gaussiana**, y sus etapas las recoge el algoritmo siguiente.

Algoritmo de Gauss

Dada una matriz A se aplican los siguientes pasos para llevarla a una escalonada:

1. Si la matriz A es la matriz nula, parar.
2. En otro caso, buscar en la primera columna no nula por la izquierda de A algún elemento no nulo λ (pivote) e intercambiar la primera fila con la que contiene al pivote (operación elemental de tipo I).
3. Multiplicar ahora la primera fila por $1/\lambda$ haciendo que el pivote valga 1 (operación elemental de tipo II).
4. Añadir a cada fila distinta de la primera un múltiplo escalar de la primera para obtener 0's debajo del pivote (operación elemental de tipo III). Como se hacen ceros por debajo del pivote, las matrices elementales que se utilizan son siempre $P_{ij}(\alpha)$ con $\boxed{i > j}$.
5. Repetir los pasos 1-4 con las restantes filas de A, a partir de la columna siguiente.

Aplicando este algoritmo a cualquier matriz se obtiene una matriz escalonada:

$$\begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nm} \end{pmatrix} \longrightarrow \begin{pmatrix} \boxed{1} & \cdots & * & \cdots & * \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & \boxed{1} & \cdots & * \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & * \end{pmatrix}$$

En la práctica el algoritmo permite variaciones, por ejemplo intercambiar los puntos 3 y 4 o trabajar con pivotes no necesariamente de valor 1.

✓ **Ejemplo 8.** Aplicar el algoritmo de Gauss a la matriz A para llevarla a una matriz escalonada por filas:

$$A = \begin{pmatrix} 0 & 3 & 5 \\ 2 & 0 & 4 \\ 1 & 1 & 3 \end{pmatrix}.$$

Solución:

$$A = \begin{pmatrix} 0 & 3 & 5 \\ 2 & 0 & 4 \\ 1 & 1 & 3 \end{pmatrix} \xrightarrow[1^a fila \leftrightarrow 3^a fila]{P_{13}} \begin{pmatrix} 1 & 1 & 3 \\ 2 & 0 & 4 \\ 0 & 3 & 5 \end{pmatrix} \xrightarrow[2^a fila - 2\cdot 1^a fila]{P_{21}(-2)} \begin{pmatrix} 1 & 1 & 3 \\ 0 & -2 & -2 \\ 0 & 3 & 5 \end{pmatrix}$$

$$\xrightarrow[(-\frac{1}{2})\cdot 2^a fila]{P_2(-\frac{1}{2})} \begin{pmatrix} 1 & 1 & 3 \\ 0 & 1 & 1 \\ 0 & 3 & 5 \end{pmatrix} \xrightarrow[3^a fila - 3\cdot 2^a fila]{P_{32}(-3)} \begin{pmatrix} 1 & 1 & 3 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix} \xrightarrow[(\frac{1}{2})\cdot 3^a fila]{P_3(\frac{1}{2})} \begin{pmatrix} \mathbf{1} & 1 & 3 \\ 0 & \mathbf{1} & 1 \\ 0 & 0 & \mathbf{1} \end{pmatrix}.$$

◇

Método de Gauss-Jordan

Tras aplicar el método de Gauss a la matriz A, se puede continuar realizando operaciones elementales a la matriz escalonada obtenida para llevarla a su forma escalonada reducida. Este proceso, conocido como **método de Gauss-Jordan**, consiste en aplicar operaciones elementales de tipo III con $i < j$, para obtener ceros por encima de los pivotes.

$$\begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nm} \end{pmatrix} \xrightarrow{\text{Método de Gauss}}$$

$$\underbrace{\begin{pmatrix} \boxed{1} & \cdots & * & \cdots & * \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \boxed{1} & \cdots & * \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \cdots & * \end{pmatrix}}_{\text{0's por debajo de los pivotes}} \xrightarrow[\text{Gauss-Jordan}]{\text{Método de}} \underbrace{\begin{pmatrix} \boxed{1} & \cdots & \mathbf{0} & \cdots & * \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & \boxed{1} & \cdots & * \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & * \end{pmatrix}}_{\text{0's por encima y debajo de los pivotes}}$$

✓ **Ejemplo 9.** Aplicar el algoritmo de Gauss-Jordan a la matriz A para llevarla a una matriz escalonada reducida por filas:

$$A = \begin{pmatrix} 0 & 3 & 5 \\ 2 & 0 & 4 \\ 1 & 1 & 3 \end{pmatrix}.$$

Solución:

$$A = \begin{pmatrix} 0 & 3 & 5 \\ 2 & 0 & 4 \\ 1 & 1 & 3 \end{pmatrix} \xrightarrow[\text{Ejemplo 8}]{\text{Operaciones elementales}} \begin{pmatrix} 1 & 1 & 3 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow[2^a fila - 3^a fila]{P_{23}(-1)} \begin{pmatrix} 1 & 1 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\xrightarrow[1^a fila - 3 \cdot 3^a fila]{P_{13}(-3)} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow[1^a fila - 2^a fila]{P_{12}(-1)} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

◇

Todo el proceso de eliminación gaussiana se puede resumir en un producto matricial, donde intervienen la matriz inicial, la matriz escalonada final y las matrices elementales que recogen las operaciones elementales realizadas.

Proposición 2.7. Toda matriz $A \in \mathrm{Mat}_{n\times m}(\mathbb{R})$ se puede transformar mediante operaciones elementales por filas en una matriz escalonada por filas (o escalonada reducida por filas).

Como consecuencia de esta proposición se puede enunciar el siguiente corolario.

Corolario 2.8. Sean $A, B \in \mathrm{Mat}_{n\times m}(\mathbb{R})$. Si B resulta de A tras aplicar una secuencia de operaciones elementales por filas, entonces existe una matriz regular $Q \in \mathrm{Mat}_n(\mathbb{R})$ tal que

$$B = QA.$$

Demostración. Realizando a la matriz A una secuencia de operaciones elementales por filas se obtiene la matriz B. Cada una de estas operaciones elementales se corresponde con una matriz elemental Q_i como se indica en el siguiente esquema:

$$A \xrightarrow[Q_1]{} A_1 \xrightarrow[Q_2]{} A_2 \rightarrow \ldots \xrightarrow[Q_r]{} A_r = B.$$

Por tanto,

$$\left.\begin{aligned} Q_1A &= A_1 \\ Q_2A_1 &= A_2 \end{aligned}\right\} \Longrightarrow Q_2Q_1A = A_2.$$

Repitiendo el mismo razonamiento se llega a

$$Q_r \ldots Q_2\, Q_1\, A = A_r = B.$$

Basta tomar $Q = Q_r \ldots Q_2Q_1 \in \mathrm{Mat}_n(\mathbb{R})$, que es regular por ser producto de matrices elementales que son regulares. ☐

Nota. Para hallar Q, en lugar de realizar el producto de todas las matrices,

$$Q = Q_r \ldots Q_2\, Q_1 = Q_r \ldots Q_2\, Q_1 I_n,$$

basta hacer a la matriz identidad las mismas operaciones elementales (y en el mismo orden) que se realizaron a la matriz A.

$$A \xrightarrow[\text{op. elem. filas}]{} B \qquad Q_r \ldots Q_1A = B,$$

$$I_n \xrightarrow[\text{mismas op. elem. filas}]{} Q \qquad Q_r \ldots Q_1I_n = Q.$$

Los dos procesos se puede realizar de manera simultánea:

$$(A \mid I_n) \xrightarrow[\text{op. elem. filas}]{} (B \mid Q).$$

✓ **Ejemplo 10.** Utilizar las operaciones elementales realizadas a la matriz A del Ejemplo 8 para hallar una matriz regular Q tal que $B = QA$, siendo B la matriz escalonada por filas obtenida en dicho ejemplo. Expresar Q como producto de matrices elementales.

Solución: Siguiendo la demostración del Corolario 2.8, se tiene que:

$$B = P_3(\tfrac{1}{2})\, P_{32}(-3)\, P_2(-\tfrac{1}{2})\, P_{21}(-2)\, P_{13}\, A = QA,$$

y así la expresión de Q es:

$$\begin{aligned} Q &= P_3(\tfrac{1}{2})\, P_{32}(-3)\, P_2(-\tfrac{1}{2})\, P_{21}(-2)\, P_{13} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1/2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & -1/2 & 1 \\ 1/2 & 3/4 & -3/2 \end{pmatrix}. \end{aligned}$$

El proceso se puede realizar en una sola vez:

$$(A \mid I_3) = \left(\begin{array}{ccc|ccc} 0 & 3 & 5 & 1 & 0 & 0 \\ 2 & 0 & 4 & 0 & 1 & 0 \\ 1 & 1 & 3 & 0 & 0 & 1 \end{array}\right) \xrightarrow[P_{13}]{} \left(\begin{array}{ccc|ccc} 1 & 1 & 3 & 0 & 0 & 1 \\ 2 & 0 & 4 & 0 & 1 & 0 \\ 0 & 3 & 5 & 1 & 0 & 0 \end{array}\right) \xrightarrow[P_{21}(-2)]{}$$

$$\longrightarrow \left(\begin{array}{ccc|ccc} 1 & 1 & 3 & 0 & 0 & 1 \\ 0 & -2 & -2 & 0 & 1 & -2 \\ 0 & 3 & 5 & 1 & 0 & 0 \end{array}\right) \xrightarrow[P_2(-\frac{1}{2})]{} \left(\begin{array}{ccc|ccc} 1 & 1 & 3 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & -1/2 & 1 \\ 0 & 3 & 5 & 1 & 0 & 0 \end{array}\right) \xrightarrow[P_{32}(-3)]{}$$

$$\longrightarrow \left(\begin{array}{ccc|ccc} 1 & 1 & 3 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & -1/2 & 1 \\ 0 & 0 & 2 & 1 & 3/2 & -3 \end{array}\right) \xrightarrow[P_3(\frac{1}{2})]{} \underbrace{\left(\begin{array}{ccc|ccc} 1 & 1 & 3 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & -1/2 & 1 \\ 0 & 0 & 1 & 1/2 & 3/4 & -3/2 \end{array}\right)}_{(B \mid Q)}.$$

De aquí se tienen las matrices buscadas:

$$B = \begin{pmatrix} 1 & 1 & 3 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \qquad Q = \begin{pmatrix} 0 & 0 & 1 \\ 0 & -1/2 & 1 \\ 1/2 & 3/4 & -3/2 \end{pmatrix}.$$

Se puede comprobar que efectivamente $B = QA$.

◇

2.4. FACTORIZACIÓN LU

Aplicando el Corolario 2.8 se puede obtener una factorización de la matriz $A \in \mathrm{Mat}_{n\times m}(\mathbb{R})$ como:

$$B = QA \Longrightarrow A = Q^{-1}B.$$

El caso más utilizado de esta factorización es aquel en el que B es una matriz escalonada por filas (con pivotes 1's) y Q^{-1} triangular inferior, denominándose *factorización* LU de la matriz A.

Para conseguir este propósito es necesario limitar el tipo de operaciones elementales en el algoritmo de Gauss. Únicamente se permite realizar operaciones de dos tipos:

- Tipo II: Multiplicación de una fila por un escalar no nulo, codificada a través de la matriz elemental $P_i(\alpha)$, con $\alpha \neq 0$.
- Tipo III: Adición a una fila i de un múltiplo escalar de una fila anterior j, codificada a través de la matriz elemental $P_{ij}(\alpha)$, con $\boxed{i > j}$

Observación. En el proceso anterior no se permite realizar intercambio de filas (tipo I) ni operaciones elementales de tipo III con $i < j$.

Las matrices $P_i(\alpha)$, $\alpha \neq 0$, y $P_{ij}(\alpha)$, $i > j$, son triangulares inferiores, así como sus inversas $P_i(\frac{1}{\alpha})$ y $P_{ij}(-\alpha)$. Si se denotan por $Q_1, \ldots, Q_s$ las matrices de los tipos anteriores que recogen las operaciones elementales realizadas para llegar de A hasta U, se tiene:

$$Q_s \ldots Q_2 Q_1 A = U.$$

Despejando A:

$$A = (Q_s \ldots Q_2 Q_1)^{-1} U = \underbrace{(Q_1^{-1} Q_2^{-1} \ldots Q_s^{-1})}_{L} U = LU.$$

La matriz A se puede expresar como el producto de dos matrices donde L es triangular inferior (por ser producto de triangulares inferiores) y regular, y U es triangular superior (o escalonada).

Se puede enunciar la siguiente proposición:

Proposición 2.9. Sea $A \in \mathrm{Mat}_{n\times m}(\mathbb{R})$. Si A puede ser reducida a una forma escalonada mediante el algoritmo de Gauss **sin intercambio de filas**, entonces existe una matriz $L \in \mathrm{Mat}_n(\mathbb{R})$ triangular inferior y regular, y una matriz $U \in \mathrm{Mat}_{n\times m}(\mathbb{R})$ escalonada (y por tanto triangular superior) tales que:

$$A = LU.$$

Definición 2.10. La descomposición de A como producto de las matrices L y U se denomina **factorización LU** de la matriz A.

Observaciones.

- Sea $A = LU$, con $A \in \mathrm{Mat}_{n\times m}(\mathbb{R})$. Entonces, las matrices A y U son de la misma dimensión y tienen el mismo rango (ver Definición 2.12). La matriz L es regular de tamaño $n \times n$.
- Dada una matriz A, las matrices L y U obtenidas en la Proposición 2.9 no son únicas, es decir, la factorización LU de una matriz A no es única.
- Regla nemotécnica:

$$L \longleftrightarrow \text{Lower triangular matrix.}$$

$$U \longleftrightarrow \text{Upper triangular matrix.}$$

✓ **Ejemplo 11.** Hallar una factorización LU de la matriz

$$A = \begin{pmatrix} 0 & 1 & -3 & -1 & 2 \\ 0 & -2 & 6 & 6 & 4 \\ 0 & -1 & 3 & 7 & 10 \end{pmatrix}.$$

Solución:

$$A = \begin{pmatrix} 0 & 1 & -3 & -1 & 2 \\ 0 & -2 & 6 & 6 & 4 \\ 0 & -1 & 3 & 7 & 10 \end{pmatrix} \xrightarrow[\substack{P_{21}(2)\\ P_{31}(1)}]{} \begin{pmatrix} 0 & 1 & -3 & -1 & 2 \\ 0 & 0 & 0 & 4 & 8 \\ 0 & 0 & 0 & 6 & 12 \end{pmatrix} \xrightarrow[P_2(1/4)]{}$$

$$\begin{pmatrix} 0 & 1 & -3 & -1 & 2 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 6 & 12 \end{pmatrix} \xrightarrow[P_{32}(-6)]{(*)} \begin{pmatrix} 0 & 1 & -3 & -1 & 2 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = U.$$

Recogiendo las matrices elementales utilizadas:

$$U = P_{32}(-6)\, P_2\,(1/4)\, P_{31}(1)\, P_{21}(2) A,$$

y despejando la matriz A:

$$\begin{aligned} A &= (P_{32}(-6)\, P_2\,(1/4)\, P_{31}(1)\, P_{21}(2))^{-1}\, U \\ &= \underbrace{P_{21}(-2) P_{31}(-1) P_2(4) P_{32}(6)}_{L}\, U. \end{aligned}$$

Para hallar L se pueden multiplicar las matrices elementales anteriores o aplicar las operaciones correspondientes a la matriz identidad.

$$L = P_{21}(-2) P_{31}(-1) P_2(4) P_{32}(6) I_3.$$

$$\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix} \xrightarrow[P_{32}(6)]{} \begin{pmatrix}1&0&0\\0&1&0\\0&6&1\end{pmatrix} \xrightarrow[P_2(4)]{} \begin{pmatrix}1&0&0\\0&4&0\\0&6&1\end{pmatrix} \xrightarrow[P_{31}(-1)]{}$$

$$\begin{pmatrix}1&0&0\\0&4&0\\-1&6&1\end{pmatrix} \xrightarrow[P_{21}(-2)]{} \begin{pmatrix}1&0&0\\-2&4&0\\-1&6&1\end{pmatrix} = L.$$

La descomposición LU obtenida es la siguiente:

$$A = \underbrace{\begin{pmatrix}1&0&0\\-2&4&0\\-1&6&1\end{pmatrix}}_{L}\underbrace{\begin{pmatrix}0&1&-3&-1&2\\0&0&0&1&2\\0&0&0&0&0\end{pmatrix}}_{U}.$$

◇

$\rightsquigarrow$ **Ejercicio.** Comprobar que si en lugar de la operación elemental $(*)$ se hace primero $P_3(1/6)$ y luego $P_{32}(-1)$, se obtiene otra descomposición LU de A distinta a la anterior. Concretamente:

$$A = \underbrace{\begin{pmatrix}1&0&0\\-2&4&0\\-1&6&6\end{pmatrix}}_{L}\underbrace{\begin{pmatrix}0&1&-3&-1&2\\0&0&0&1&2\\0&0&0&0&0\end{pmatrix}}_{U}.$$

Nota. Si en el proceso de llevar A a una matriz escalonada se necesita hacer intercambio de filas, se obtiene la llamada factorización $PA = LU$. En primer lugar se realizan a la matriz A todas las permutaciones de filas necesarias en el proceso, obteniéndose así una matriz B que se puede factorizar como $B = LU$. Las operaciones elementales de tipo I realizadas a la matriz A se recogen en una matriz P llamada *matriz de permutación* que verifica $B = PA$, de modo que

$$PA = LU.$$

✓ **Ejemplo 12.** Hallar una factorización LU de la matriz

$$A = \begin{pmatrix}0&1&-2\\0&0&4\\-1&2&1\end{pmatrix}.$$

Solución: En primer lugar se observa que es necesario hacer intercambios de filas para poder escalonar la matriz. Esto quiere decir que la factorización a obtener es del tipo $PA = LU$:

$$\underbrace{\begin{pmatrix}0&1&-2\\0&0&4\\-1&2&1\end{pmatrix}}_{A} \xrightarrow[P_{13}]{} \begin{pmatrix}-1&2&1\\0&0&4\\0&1&-2\end{pmatrix} \xrightarrow[P_{23}]{} \begin{pmatrix}-1&2&1\\0&1&-2\\0&0&4\end{pmatrix} = B \implies B = \underbrace{P_{23}P_{13}}_{P}A.$$

La matriz B se puede escalonar sin intercambios de filas:

$$B = \begin{pmatrix} -1 & 2 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 4 \end{pmatrix} \xrightarrow[\substack{P_1(-1) \\ P_3(1/4)}]{} \begin{pmatrix} 1 & -2 & -1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix} = U \Longrightarrow U = P_3(1/4)P_1(-1)B.$$

Se tiene que:

$$U = P_3(1/4)P_1(-1)B = P_3(1/4)P_1(-1)PA,$$

de donde

$$PA = \underbrace{[P_3(1/4)P_1(-1)]^{-1}}_{L} U,$$

$$P = P_{23}P_{13} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad L = P_1(-1)P_3(4) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix}.$$

La descomposición encontrada es la siguiente:

$$\underbrace{\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}}_{P} A = \underbrace{\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix}}_{L} \underbrace{\begin{pmatrix} 1 & -2 & -1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}}_{U}.$$

◇

2.5. RANGO DE UNA MATRIZ

Como se ha visto en la Proposición 2.7 toda matriz se puede transformar mediante operaciones elementales por filas en una matriz escalonada o escalonada reducida que proporciona información importante de la matriz inicial.

Proposición 2.11. Se considera una matriz $A \in \mathrm{Mat}_{n\times m}(\mathbb{R})$, y se transforma mediante una secuencia de operaciones elementales por filas en una matriz escalonada o escalonada reducida por filas $B \in \mathrm{Mat}_{n\times m}(\mathbb{R})$. El número de filas no nulas de la matriz B depende solo de la matriz original A y no de las operaciones elementales realizadas.

Definición 2.12. Dada una matriz $A \in \mathrm{Mat}_{n\times m}(\mathbb{R})$, se llama **rango** de A, y se denota por $\mathrm{rg}\, A$, al número de pivotes de una matriz escalonada, $B \in \mathrm{Mat}_{n\times m}(\mathbb{R})$, (o escalonada reducida) obtenida tras aplicar operaciones elementales por filas a la matriz A.

Nota. Observar que si $A \in \mathrm{Mat}_{n\times m}(\mathbb{R})$ es una matriz escalonada por filas, entonces $\operatorname{rg} A$ es directamente el número de filas no nulas de A. Así por ejemplo:

- $\operatorname{rg} I_n = n$.
- $\operatorname{rg} O_{n\times m} = 0$.

✓ **Ejemplo 13.** Calcular el rango de la matriz A:

$$A = \begin{pmatrix} 0 & 3 & 5 \\ 2 & 0 & 4 \\ 1 & 1 & 3 \end{pmatrix}.$$

Solución: Utilizando las operaciones del Ejemplo 8,

$$A = \begin{pmatrix} 0 & 3 & 5 \\ 2 & 0 & 4 \\ 1 & 1 & 3 \end{pmatrix} \xrightarrow[\text{op. elem. filas}]{} \begin{pmatrix} 1 & 1 & 3 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} = B.$$

Se observa que el número de pivotes de B es 3. Así, $\operatorname{rg} A = \operatorname{rg} B = 3$.

◇

✓ **Ejemplo 14.** Calcular el rango de la matriz A dependiente del parámetro a:

$$A = \begin{pmatrix} a & 2 \\ 1 & 4 \end{pmatrix}.$$

Solución: Se comienza permutando las filas 1 y 2 de la matriz A para asegurar que en la posición $(1,1)$ aparece un elemento no nulo que puede actuar como pivote.

$$A = \begin{pmatrix} a & 2 \\ 1 & 4 \end{pmatrix} \xrightarrow[P_{12}]{} \begin{pmatrix} 1 & 4 \\ a & 2 \end{pmatrix} \xrightarrow[P_{21}(-a)]{} \begin{pmatrix} 1 & 4 \\ 0 & 2-4a \end{pmatrix}.$$

Se observa que si $a = 1/2$, $\operatorname{rg} A = 1$, mientras que si $a \neq 1/2$, $\operatorname{rg} A = 2$. En este último caso, el elemento $2-4a$ se considera un pivote por ser un valor no nulo.

◇

Propiedades. Sean $A, B \in \mathrm{Mat}_{n\times m}(\mathbb{R})$, entonces se verifica:

1. $\operatorname{rg}(A+B) \leq \operatorname{rg} A + \operatorname{rg} B$.
2. $\operatorname{rg}(kA) = \operatorname{rg} A, \quad \forall\, k \in \mathbb{R} - \{0\}$.
3. $\operatorname{rg} A = \operatorname{rg} A^t$.
4. $\operatorname{rg}(PA) = \operatorname{rg} A, \ \forall P \in \mathrm{Mat}_n(\mathbb{R})$ regular.
5. $\operatorname{rg} A \leq \min\{n, m\}$.

2.6. CÁLCULO DE LA INVERSA DE UNA MATRIZ

Se pueden aprovechar las operaciones y matrices elementales para calcular la inversa de una matriz. Se enuncia el siguiente teorema que indica cuándo una matriz cuadrada tiene inversa.

Teorema 2.13. Sea $A \in \operatorname{Mat}_n(\mathbb{R})$ una matriz cuadrada. Las siguientes condiciones son equivalentes:

1. A es regular, es decir, $\exists A^{-1}$.
2. $\operatorname{rg} A = n$.
3. A se puede llevar por operaciones elementales a la identidad (que es su forma escalonada reducida).
4. A es producto de matrices elementales.
5. $|A| \neq 0$.

En el caso de que la matriz $A \in \operatorname{Mat}_n(\mathbb{R})$ cumpla una de las condiciones del teorema anterior, para calcular A^{-1} basta aplicar el Corolario 2.8 con $B = I_n$, es decir, llevando la matriz A hasta la identidad mediante operaciones elementales. Así,

$$QA = I_n$$

y Q es la inversa de A, $Q = A^{-1}$.

Por tanto se puede seguir el siguiente esquema para hallar la inversa de A:

$$\boxed{(A \mid I_n) \xrightarrow[\text{op. elem. filas}]{} (I_n \mid A^{-1})}$$

Con este método, si se recogen las operaciones elementales realizadas en sus correspondientes matrices elementales, se puede además factorizar A y A^{-1} como producto de matrices elementales:

$$\begin{aligned} Q_r \dots Q_2 Q_1 A &= I_n, \\ A^{-1} &= Q_r \dots Q_2 Q_1, \\ A &= Q_1^{-1} Q_2^{-1} \dots Q_r^{-1}. \end{aligned}$$

✓ **Ejemplo 15.** Dada la matriz

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

calcular A^{-1} y escribir A y A^{-1} como producto de matrices elementales.

Solución: Se hacen operaciones elementales por filas hasta transformar la matriz A en la identidad siguiendo el algoritmo de Gauss-Jordan:

$$(A \mid I_2) = \left(\begin{array}{cc|cc} 3 & -1 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{array}\right) \xrightarrow[P_{12}]{} \left(\begin{array}{cc|cc} 1 & 2 & 0 & 1 \\ 3 & -1 & 1 & 0 \end{array}\right) \xrightarrow[P_{21}(-3)]{} \left(\begin{array}{cc|cc} 1 & 2 & 0 & 1 \\ 0 & -7 & 1 & -3 \end{array}\right)$$

$$\xrightarrow[P_2(-\frac{1}{7})]{} \left(\begin{array}{cc|cc} 1 & 2 & 0 & 1 \\ 0 & 1 & -1/7 & 3/7 \end{array}\right) \xrightarrow[P_{12}(-2)]{} \left(\begin{array}{cc|cc} 1 & 0 & 2/7 & 1/7 \\ 0 & 1 & -1/7 & 3/7 \end{array}\right) = (\, I_2 \mid A^{-1}\,).$$

Luego la inversa de A es:

$$A^{-1} = \begin{pmatrix} 2/7 & 1/7 \\ -1/7 & 3/7 \end{pmatrix}.$$

Aplicando el Corolario 2.8,

$$P_{12}(-2)\, P_2\left(-\tfrac{1}{7}\right) P_{21}(-3)\, P_{12}\, A = I_2,$$

de donde se puede expresar A^{-1} y A como producto de matrices elementales

$$A^{-1} = P_{12}(-2)\, P_2\left(-\tfrac{1}{7}\right) P_{21}(-3)\, P_{12},$$

$$A = P_{12}\, P_{21}(3)\, P_2(-7)\, P_{12}(2).$$

◇

Se presenta un caso en el que no se dan las condiciones del Teorema 2.13 y por tanto no se puede calcular la inversa de la matriz dada.

✓ **Ejemplo 16.** En caso de que sea posible, escribir A como producto de matrices elementales:

$$A = \begin{pmatrix} 4 & 2 \\ 2 & 1 \end{pmatrix}.$$

Solución: Se realizan operaciones elementales a la matriz $(A \mid I_2)$ hasta obtener una forma escalonada por filas:

$$(A \mid I_2) = \left(\begin{array}{cc|cc} 4 & 2 & 1 & 0 \\ 2 & 1 & 0 & 1 \end{array}\right) \xrightarrow[P_{12}]{} \left(\begin{array}{cc|cc} 2 & 1 & 0 & 1 \\ 4 & 2 & 1 & 0 \end{array}\right) \xrightarrow[P_{21}(-2)]{} \left(\begin{array}{cc|cc} 2 & 1 & 0 & 1 \\ 0 & 0 & 1 & -2 \end{array}\right)$$

$$\xrightarrow[P_1(\frac{1}{2})]{} \left(\begin{array}{cc|cc} 1 & 1/2 & 0 & 1/2 \\ 0 & 0 & 1 & -2 \end{array}\right).$$

Se observa que $\operatorname{rg} A \neq 2$, y por tanto A no tiene inversa. Es decir, la matriz A no se puede llevar a la identidad mediante operaciones elementales y por tanto no se puede expresar como producto de matrices elementales.

◇

2.7. APLICACIONES: CRIPTOGRAFÍA

Una aplicación interesante de las matrices es la codificación de mensajes, de manera que una persona ajena no sea capaz de descifrar un mensaje que ha sido cifrado. Este proceso es conocido como **criptografía** y se realiza en varios pasos.

En primer lugar, se transforma el mensaje escrito en una lista numérica asignando a cada letra el número que ocupa en el alfabeto. Al espacio en blanco se le asigna el cero. Por simplicidad no se distingue entre mayúsculas y minúsculas, se ignoran las tildes y signos de puntuación, y la letra ñ se convierte en la letra n. De esta manera solo hacen falta los números del 0 al 26 como se muestra a continuación:

Letra	Número
␣	0
A	1
B	2
C	3
D	4
E	5
F	6

Letra	Número
G	7
H	8
I	9
J	10
K	11
L	12
M	13

Letra	Número
N	14
O	15
P	16
Q	17
R	18
S	19
T	20

Letra	Número
U	21
V	22
W	23
X	24
Y	25
Z	26

Por ejemplo, el texto “ACADEMIA GENERAL MILITAR”, después de transformar cada letra en número, se convierte en la lista:

$$L = \Big[\, \underbrace{1, 3, 1, 4, 5, 13, 9, 1}_{\text{ACADEMIA}},\ 0,\ \underbrace{7, 5, 14, 5, 18, 1, 12}_{\text{GENERAL}},\ 0,\ \underbrace{13, 9, 12, 9, 20, 1, 18}_{\text{MILITAR}} \,\Big].$$

En segundo lugar, hay que considerar una matriz regular C llamada **matriz de codificación** o **matriz de Hill**, por ejemplo

$$C = \begin{pmatrix} 2 & -1 & 1 \\ 1 & 3 & 1 \\ 1 & 2 & 1 \end{pmatrix},$$

y convertir la lista L anterior en una matriz A por bloques, donde cada uno de ellos es una matriz columna de longitud igual al orden de C. Esto es simplemente para que las dimensiones de las matrices C y A sean adecuadas para poder realizar el producto CA.

En nuestro ejemplo,

$$A = \left(\begin{array}{c|c|c|c|c|c|c|c} 1 & 4 & 9 & 7 & 5 & 12 & 9 & 20 \\ 3 & 5 & 1 & 5 & 18 & 0 & 12 & 1 \\ 1 & 13 & 0 & 14 & 1 & 13 & 9 & 18 \end{array}\right).$$

A continuación, se realiza la siguiente multiplicación de matrices:

$$CA = \left(\begin{array}{c|c|c|c|c|c|c|c} 0 & 16 & 17 & 23 & -7 & 37 & 15 & 57 \\ 11 & 32 & 12 & 36 & 60 & 25 & 54 & 41 \\ 8 & 27 & 11 & 31 & 42 & 25 & 42 & 40 \end{array}\right) = B.$$

Finalmente, se convierte la matriz por bloques columna B en una lista y se transmite el mensaje de la siguiente forma:

$$\Big[\ 0, 11, 8,\ \ 16, 32, 27,\ \ 17, 12, 11,\ \ 23, 36, 31,\ \ -7, 60, 42,\ \ 37, 25, 25,\ \ 15, 54, 42,\ \ 57, 41, 40\ \Big].$$

Para quienes desconozcan la matriz C de codificación, es difícil descifrar el mensaje anterior. Sin embargo, para un receptor que conozca C, descifrar el mensaje es simple, basta con repetir el mismo procedimiento descrito anteriormente con la matriz inversa C^{-1}, ya que

$$C^{-1}B = \underbrace{C^{-1}C}_{I_3}A = A,$$

recuperando así el mensaje original.

Siguiendo el esquema del cuadro de la página 48, la matriz inversa de C se puede calcular como:

$$\big(C \mid I_3\big) = \left(\begin{array}{ccc|ccc} 2 & -1 & 1 & 1 & 0 & 0 \\ 1 & 3 & 1 & 0 & 1 & 0 \\ 1 & 2 & 1 & 0 & 0 & 1 \end{array}\right) \xrightarrow{P_{13}} \left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 0 & 0 & 1 \\ 1 & 3 & 1 & 0 & 1 & 0 \\ 2 & -1 & 1 & 1 & 0 & 0 \end{array}\right)$$

$$\xrightarrow[P_{31}(-2)]{P_{21}(-1)} \left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & -1 \\ 0 & -5 & -1 & 1 & 0 & -2 \end{array}\right) \xrightarrow{P_{32}(5)} \left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & -1 \\ 0 & 0 & -1 & 1 & 5 & -7 \end{array}\right)$$

$$\xrightarrow{P_{3}(-1)} \left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & -1 & -5 & 7 \end{array}\right) \xrightarrow{P_{13}(-1)} \left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 5 & -6 \\ 0 & 1 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & -1 & -5 & 7 \end{array}\right)$$

$$\xrightarrow{P_{12}(-2)} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 3 & -4 \\ 0 & 1 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & -1 & -5 & 7 \end{array}\right) = \big(I_3 \mid C^{-1}\big).$$

Multiplicando C^{-1} por la matriz B se recupera el mensaje original escrito en forma de matriz A por bloques:

$$C^{-1}B = \left(\begin{array}{c|c|c|c|c|c|c|c} 1 & 4 & 9 & 7 & 5 & 12 & 9 & 20 \\ 3 & 5 & 1 & 5 & 18 & 0 & 12 & 1 \\ 1 & 13 & 0 & 14 & 1 & 13 & 9 & 18 \end{array}\right) = A.$$

Para terminar el proceso de descodificación se escribe en forma de lista y se sustituye cada número por la correspondiente letra del alfabeto:

$$L = \Big[\ \underbrace{1, 3, 1, 4, 5, 13, 9, 1}_{\text{ACADEMIA}},\ 0,\ \underbrace{7, 5, 14, 5, 18, 1, 12}_{\text{GENERAL}},\ 0,\ \underbrace{13, 9, 12, 9, 20, 1, 18}_{\text{MILITAR}}\ \Big].$$

Así, el mensaje original es “ACADEMIA GENERAL MILITAR” como ya se sabía.

Observación. A veces la longitud de la lista inicial no es divisible por el número de columnas de C. En este caso se añaden espacios en blanco, o equivalentemente, el número 0 hasta completar la matriz. Por ejemplo, el texto "INGENIEROS" da lugar a la siguiente matriz (observar que hay que añadir dos ceros en la matriz):

$$\begin{pmatrix} 9 & 5 & 5 & 19 \\ 14 & 14 & 18 & 0 \\ 7 & 9 & 15 & 0 \end{pmatrix}.$$

$\rightsquigarrow$ **Ejercicio.** Usar la inversa de la matriz de codificación

$$C = \begin{pmatrix} 1 & 2 & -5 & 0 \\ 3 & 4 & 1 & 2 \\ 0 & -1 & 2 & 3 \\ 4 & 7 & 0 & 1 \end{pmatrix}$$

para descifrar el siguiente mensaje:

$$29, 79, 12, 112, -32, 152, 70, 176, -85, 39, 33, 35,$$
$$31, 83, 3, 119, 8, 162, 53, 209, -95, 227, 128, 260.$$

A continuación vuelve a cifrar el mensaje obtenido usando esta vez la matriz C de tamaño 3×3 de la página anterior. Observar que el mensaje cifrado es diferente.

2.8. COMANDOS DE *wxMaxima*

Tras introducir una matriz A se pueden realizar las siguientes operaciones:

- Transformar una matriz en una escalonada por filas (con pivotes 1) mediante operaciones elementales:

```
-->      echelon(A);
```

- Transformar una matriz en triangular superior (similar a la escalonada) pero sin pivotes 1:

```
-->      triangularize(A);
```

- Calcular el rango de una matriz:

```
-->      rank(A);
```

- Hallar la inversa de una matriz A cuadrada y regular:

```
-->      invert(A);
```

  Comando al que también se accede desde “`\Álgebra\Invertir Matriz`”.

Los comandos anteriores pueden no funcionar correctamente en el caso de que la matriz A contenga parámetros.

Operaciones elementales

- Operación elemental de tipo I:

```
-->      rowswap(A,i,j);
```

  Este comando devuelve la matriz que se obtiene a partir de A intercambiando la fila i por la fila j. Si $A = I_n$ entonces el resultado de “`rowswap(ident(n),i,j)`” es la matriz elemental P_{ij} de orden n.

- Operación elemental de tipo III:

```
-->      rowop(A,i,j,-a);
```

  Este comando devuelve la matriz que resulta de sumar a la fila i de A la fila j multiplicada por “`a`”. Si $A = I_n$ entonces “`rowop(ident(n),i,j,-a)`” es la matriz elemental $P_{ij}(a)$ de orden n. Atención al signo menos que hay que añadir al parámetro "`a`" para que el comando coincida con la operación elemental de tipo III definida en el texto.

- Operación elemental de tipo II: *wxMaxima* no dispone de un comando específico para realizar la operación elemental de este tipo, pero se puede definir utilizando la operación elemental de tipo III. Multiplicar la fila i-ésima de A por a es equivalente a sumar a dicha fila i ella misma $a-1$ veces. Esto se consigue con el comando "`rowop`":

  ```
  -->     rowop(A,i,i,1-a);
  ```

  Si $A = I_n$ entonces "`rowop(ident(n),i,i,1-a)`" es la matriz elemental $P_i(a)$ de orden n.

Factorización LU

Para obtener la factorización LU de una matriz A cuadrada y regular los comandos son:

```
-->     A:matriz(fila1,...,filan)$
        M:lu_factor(A)$
        get_lu_factors(M);
```

El código solo funciona para matrices A cuadradas y devuelve una lista de tres matrices $[P_1, L, U]$ donde P_1 es una matriz de permutación, L es triangular inferior con 1's en la diagonal y U es triangular superior de modo que $A = P_1 LU$.

Observar que esta no es exactamente la descomposición $PA = LU$ que se ha definido en el texto. Presenta dos diferencias: por una parte, $P = P_1^{-1}$ y en el capítulo se ha definido la matriz U con pivotes 1's.

2.9. EJERCICIOS PROPUESTOS

Método de eliminación gaussiana

1. Escalonar las siguientes matrices aplicando el algoritmo de Gauss.

$$A = \begin{pmatrix} 1 & 1 & 2 \\ 3 & -1 & 1 \\ -1 & 3 & 4 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 & -1 & 3 \\ -1 & 4 & 5 & -2 \\ 1 & 6 & 3 & 4 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 2 & 3 \\ -1 & 1 & 1 \\ 3 & 3 & 5 \\ 0 & 3 & 4 \end{pmatrix}.$$

2. Para cada una de las siguientes matrices encontrar una matriz regular Q tal que $QA = B$, siendo B una matriz escalonada, y expresar Q como producto de matrices elementales:

$$\text{a)} \begin{pmatrix} 1 & 1 & -2 \\ -2 & 1 & 0 \end{pmatrix} \qquad \text{b)} \begin{pmatrix} 1 & 2 & -1 & 0 \\ 3 & 1 & 1 & 2 \\ 1 & -3 & 3 & 2 \end{pmatrix} \qquad \text{c)} \begin{pmatrix} 2 & 1 & -1 & 0 \\ 3 & -1 & 2 & 1 \\ 1 & -2 & 3 & 1 \end{pmatrix}$$

3. Encontrar una factorización LU de las siguientes matrices:

$$A_1 = \begin{pmatrix} 2 & 6 & -2 & 0 & 2 \\ 3 & 9 & -3 & 3 & 1 \\ -1 & -3 & 1 & -3 & 1 \end{pmatrix}, \qquad A_2 = \begin{pmatrix} 2 & 4 & 2 \\ 1 & -1 & 3 \\ -1 & 7 & -7 \end{pmatrix}.$$

Rango de Matrices

4. Determinar el rango de cada una de las matrices del Ejercicio 1.

5. Calcular p y q para que el rango de la siguiente matriz sea igual a 2:

$$\begin{pmatrix} 2 & 3 & 4 & 5 \\ 6 & 2 & -1 & 4 \\ -4 & -13 & p & q \end{pmatrix}.$$

6. Discutir el rango de las siguientes matrices según los valores de a y b:

$$A = \begin{pmatrix} 2a & b & 1 \\ 2 & ab & 1 \\ 2 & b & a \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 2 & 0 \\ a & 8 & 3 \\ 0 & b & 5 \end{pmatrix}.$$

Inversas de Matrices

7. Calcular la inversa de cada una de las siguientes matrices:

$$\text{a)} \begin{pmatrix} 4 & 1 \\ 3 & 2 \end{pmatrix} \quad \text{b)} \begin{pmatrix} 1 & -2 & 1 \\ -2 & 5 & -4 \\ 1 & -4 & 6 \end{pmatrix} \quad \text{c)} \begin{pmatrix} -1 & 4 & 5 & 2 \\ 0 & 0 & 0 & -1 \\ 1 & -2 & -2 & 0 \\ 0 & -1 & -1 & 0 \end{pmatrix} \quad \text{d)} \begin{pmatrix} 1 & 2 & 0 & 0 & 0 \\ 0 & 1 & 3 & 0 & 0 \\ 0 & 0 & 1 & 5 & 0 \\ 0 & 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

8. Calcular la inversa de las siguientes matrices:

$$A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 2 & 0 \\ 0 & 3 & 0 & 0 \\ 4 & 0 & 0 & 0 \end{pmatrix}.$$

9. Sea $A = \operatorname{diag}(a_{11}, \ldots, a_{nn}) \in \operatorname{Mat}_n(\mathbb{R})$ una matriz diagonal con $a_{ii} \neq 0$, $\forall i = 1, \ldots, n$. Calcular la inversa de A.

10. Calcular la inversa de las siguientes matrices triangulares:

$$\text{a)} \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad \text{b)} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad \text{c)} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

11. Calcular, en caso de que exista, la inversa de las siguientes matrices, expresando A y A^{-1} como producto de matrices elementales:

$$\text{a)} \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \qquad \text{b)} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} \qquad \text{c)} \begin{pmatrix} 2 & 3 & 4 \\ 2 & 1 & 1 \\ -1 & 1 & 2 \end{pmatrix}$$

$$\text{d)} \begin{pmatrix} 1 & 2 & 2 \\ 2 & -1 & 1 \\ 1 & -3 & -1 \end{pmatrix} \qquad \text{e)} \begin{pmatrix} 1 & 0 & -1 \\ 3 & 4 & -2 \\ 3 & 5 & -2 \end{pmatrix}$$

12. Se consideran las siguientes matrices:

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Comprobar que $(A+B)^{-1} \neq A^{-1} + B^{-1}$.

13. Encontrar dos matrices $A, B \in \operatorname{Mat}_3(\mathbb{R})$ singulares, tales que $A + B$ sea regular.

14. Expresar la condición que deben cumplir los elementos de la matriz $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ para que sea regular. En tal caso hallar la inversa de A.

15. Hallar la inversa de

$$\begin{pmatrix} a & b & 0 & 0 \\ c & d & 0 & 0 \\ 0 & 0 & a & b \\ 0 & 0 & c & d \end{pmatrix}.$$

16. En cada caso encontrar los valores de a y b para que las siguientes matrices sean regulares:

$$A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & a+1 & -1 \\ 0 & 1 & a-3 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & a & b \\ 1 & a^2 & b^2 \\ 1 & a^3 & b^3 \end{pmatrix}.$$

2.10. SOLUCIONES A LOS EJERCICIOS PROPUESTOS

Método de eliminación gaussiana

1. Unas posibles matrices escalonadas son:

$$A_{\text{esc}} = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 5/4 \\ 0 & 0 & 1 \end{pmatrix}, \quad B_{\text{esc}} = \begin{pmatrix} 1 & 1 & -1 & 3 \\ 0 & 1 & 4/5 & 1/5 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad C_{\text{esc}} = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 4/3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

2. La solución de este ejercicio NO es única. Aquí se propone una solución posible.

a) $Q = P_2(\frac{1}{3})P_{21}(2) = \begin{pmatrix} 1 & 0 \\ 2/3 & 1/3 \end{pmatrix}$,

$B = QA = \begin{pmatrix} 1 & 1 & -2 \\ 0 & 1 & -4/3 \end{pmatrix}$.

b) $Q = P_2(-\frac{1}{5})P_{32}(-1)P_{31}(-1)P_{21}(-3) = \begin{pmatrix} 1 & 0 & 0 \\ 3/5 & -1/5 & 0 \\ 2 & -1 & 1 \end{pmatrix}$,

$B = QA = \begin{pmatrix} 1 & 2 & -1 & 0 \\ 0 & 1 & -4/5 & -2/5 \\ 0 & 0 & 0 & 0 \end{pmatrix}$.

c) $Q = P_2(\frac{1}{5})P_{32}(-1)P_{31}(-2)P_{21}(-3)P_{13} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1/5 & -3/5 \\ 1 & -1 & 1 \end{pmatrix}$,

$B = QA = \begin{pmatrix} 1 & -2 & 3 & 1 \\ 0 & 1 & -7/5 & -2/5 \\ 0 & 0 & 0 & 0 \end{pmatrix}$.

3. Recordar que la factorización LU no es única. Una posibilidad sería:

$$A_1 = \begin{pmatrix} 2 & 0 & 0 \\ 3 & 3 & 0 \\ -1 & -3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 3 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & -2/3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

$$A_2 = \begin{pmatrix} 2 & 0 & 0 \\ 1 & -3 & 0 \\ -1 & 9 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & -2/3 \\ 0 & 0 & 0 \end{pmatrix}.$$

Rango de Matrices

4. $\operatorname{rg} A = 3, \qquad \operatorname{rg} B = 2, \qquad \operatorname{rg} C = 2.$

5. $\operatorname{rg} A = 2 \iff p = q = -21.$

6. $$\operatorname{rg} A = \begin{cases} 1 & \text{si} \quad a = 1, \quad \forall b \in \mathbb{R}. \\ 2 & \text{si} \quad \begin{cases} a = -2, \quad \forall b \in \mathbb{R}. \\ a \neq -2, \ a \neq 1 \quad \text{y} \quad b = 0. \end{cases} \\ 3 & \text{si} \quad a \neq -2, \ a \neq 1 \quad \text{y} \quad b \neq 0. \end{cases}$$

$$\operatorname{rg} B = \begin{cases} 2 & \text{si} \quad 10a + 3b - 40 = 0. \\ 3 & \text{si} \quad 10a + 3b - 40 \neq 0. \end{cases}$$

Inversas de Matrices

7.

a) $\begin{pmatrix} 2/5 & -1/5 \\ -3/5 & 4/5 \end{pmatrix}$ b) $\begin{pmatrix} 14 & 8 & 3 \\ 8 & 5 & 2 \\ 3 & 2 & 1 \end{pmatrix}$

c) $\begin{pmatrix} 0 & 0 & 1 & -2 \\ -1 & -2 & -1 & -3 \\ 1 & 2 & 1 & 2 \\ 0 & -1 & 0 & 0 \end{pmatrix}$ d) $\begin{pmatrix} 1 & -2 & 6 & -30 & 210 \\ 0 & 1 & -3 & 15 & -105 \\ 0 & 0 & 1 & -5 & 35 \\ 0 & 0 & 0 & 1 & -7 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$

8.

$$A^{-1} = \begin{pmatrix} 1/2 & 0 & 0 \\ 0 & 1/3 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad B^{-1} = \begin{pmatrix} 0 & 0 & 0 & 1/4 \\ 0 & 0 & 1/3 & 0 \\ 0 & 1/2 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

9. Realizar las operaciones elementales $P_i(\frac{1}{a_{ii}})$, $\forall i = 1, \ldots, n$. Observar que el resultado coincide con la Proposición 1.8.

10.

a) $\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ b) $\begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ c) $\begin{pmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 1 \end{pmatrix}$

11. Atención: Las operaciones elementales necesarias para realizar la inversa NO son únicas, aunque la inversa sí lo es. Aquí se muestra una solución posible.

a) $\begin{pmatrix} 1/2 & 1/2 & -1/2 \\ -1/2 & 1/2 & 1/2 \\ 1/2 & -1/2 & 1/2 \end{pmatrix}$

$A^{-1} = P_{23}(1)P_{13}(-1)P_3(\frac{1}{2})P_{32}(-1)P_{21}(-1)$,

$A = P_{21}(1)P_{32}(1)P_3(2)P_{13}(1)P_{23}(-1)$.

b) $\nexists A^{-1}$.

c) $\begin{pmatrix} -1 & 2 & 1 \\ 5 & -8 & -6 \\ -3 & 5 & 4 \end{pmatrix}$

$A^{-1} = P_3(-3)P_1(-1)P_{23}(5)P_{13}(1)P_{12}(-1)P_{32}(-5)P_2(\frac{1}{3})P_{31}(2)P_{21}(2)P_{13}$,
$A = P_{13}P_{21}(-2)P_{31}(-2)P_2(3)P_{32}(5)P_{12}(1)P_{13}(-1)P_{23}(-5)P_1(-1)P_3(-\frac{1}{3})$.

d) $\nexists A^{-1}$.

e) $\begin{pmatrix} -2 & 5 & -4 \\ 0 & -1 & 1 \\ -3 & 5 & -4 \end{pmatrix}$

$A^{-1} = P_3(-\frac{1}{4})P_{13}(-4)P_{23}(1)P_{32}(-5)P_2(\frac{1}{4})P_{31}(-3)P_{21}(-3)$,
$A = P_{21}(3)P_{31}(3)P_2(4)P_{32}(5)P_{23}(-1)P_{13}(4)P_3(-4)$.

12. Después de realizar los cálculos se observa que las matrices son diferentes:

$$(A+B)^{-1} = \begin{pmatrix} 2/3 & -1/3 \\ -1/3 & 2/3 \end{pmatrix}, \qquad A^{-1} + B^{-1} = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}.$$

13. La respuesta no es única. Una posible solución es la siguiente:

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

14. Condición para que A sea regular: $ad - bc \neq 0$.

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \implies A^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

15. Utilizando una descomposición en bloques adecuada y los resultados del ejercicio anterior:

$$A^{-1} = \frac{1}{ad-bc}\left(\begin{array}{cc|cc} d & -b & 0 & 0 \\ -c & a & 0 & 0 \\ \hline 0 & 0 & d & -b \\ 0 & 0 & -c & a \end{array}\right) \qquad \text{si} \qquad ad - bc \neq 0.$$

16. A es regular $\iff a \neq 1 \pm \sqrt{3}$.

B es regular $\iff ab(a-1)(b-1)(a-b) \neq 0$.

CAPÍTULO 3

Sistemas de Ecuaciones Lineales

Los sistemas de ecuaciones lineales aparecen desde los primeros testimonios matemáticos escritos de los que se tiene noticia. El primer uso demostrado del método de eliminación gaussiana aparece en el siglo III a.C. en China. No obstante, el estudio sistemático de los sistemas de ecuaciones lineales se debe a Leibniz y Cramer, a lo largo del siglo XVIII. Por otra parte, Von Neumann fue el primero en estudiar de qué forma contribuyen los errores de redondeo individual al error total y en 1948, el matemático inglés Alan Turing desarrolló la factorización LU. El método de Gauss en su formulación moderna fue presentado por Isaac Newton y recibió el apelativo de *método de eliminación gaussiana* a partir de la II Guerra Mundial.

El planteamiento y resolución de sistemas de ecuaciones lineales aparece en numerosas situaciones, tanto cotidianas como de la Ciencia y la Ingeniería, por lo que se considera una herramienta esencial para la resolución de problemas. Como aplicaciones se pueden citar los circuitos eléctricos, el ajuste de reacciones químicas, los modelos económicos de input-output o la modelización lineal de fenómenos físicos y económicos.

Este capítulo se centra en aplicar los métodos de eliminación gaussiana, aprendidos para matrices, a la resolución de sistemas de ecuaciones lineales. Además de los métodos anteriores, que se consideran directos, se proponen otros de tipo iterativo que calculan soluciones aproximadas de los sistemas de ecuaciones. Para estos últimos métodos es de gran importancia la utilización de un software matemático.

3.1. DEFINICIÓN DE UN SISTEMA DE ECUACIONES

La expresión general de un sistema de n ecuaciones lineales con m incógnitas es la siguiente:

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

donde a_{ij}, $b_i \in \mathbb{R}$ son datos conocidos mientras que $x_1, \ldots, x_m \in \mathbb{R}$ son las incógnitas a determinar. El sistema anterior se puede expresar matricialmente como $A\mathbf{x} = \mathbf{b}$:

$$\underbrace{\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}}_{\mathbf{x}} = \underbrace{\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}}_{\mathbf{b}},$$

donde $A \in \mathrm{Mat}_{n\times m}(\mathbb{R})$ se denomina **matriz de coeficientes**, $\mathbf{x} \in \mathrm{Mat}_{m\times 1}(\mathbb{R})$ es la **matriz** o **vector de incógnitas** y $\mathbf{b} \in \mathrm{Mat}_{n\times 1}(\mathbb{R})$ la **matriz** o **vector de términos independientes**.

En lo que sigue, es de gran utilidad considerar una nueva matriz que englobe a la matriz de coeficientes A y al vector de términos independientes $\mathbf{b}$ como una nueva columna.

Definición 3.1. Dado un sistema de n ecuaciones lineales con m incógnitas en forma matricial $A\mathbf{x} = \mathbf{b}$, se define la **matriz ampliada** del sistema, denotada por $(A\,|\,\mathbf{b}) \in \mathrm{Mat}_{n\times(m+1)}(\mathbb{R})$, como la matriz que surge al añadir a la matriz A la columna $\mathbf{b}$, correspondiente al vector de términos independientes.

$$(A\,|\,\mathbf{b}) = \left(\begin{array}{ccc|c} a_{11} & \cdots & a_{1m} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nm} & b_n \end{array}\right).$$

✓ **Ejemplo 1.** Expresar el siguiente sistema de ecuaciones lineales en notación matricial y construir la matriz ampliada asociada:

$$\begin{cases} x - y = 3 \\ 2x + z = -1. \end{cases}$$

Solución: En primer lugar se identifica el vector de incógnitas y el vector de términos independientes:

$$\mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathrm{Mat}_{3\times 1}(\mathbb{R}), \quad \mathbf{b} = \begin{pmatrix} 3 \\ -1 \end{pmatrix} \in \mathrm{Mat}_{2\times 1}(\mathbb{R}).$$

La matriz A de coeficientes del sistema debe ser de dimensión 2×3 para que sea compatible con las dimensiones de $\mathbf{x}$ y $\mathbf{b}$:

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

La expresión matricial del sistema es:

$$\begin{pmatrix} 1 & -1 & 0 \\ 2 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \end{pmatrix}.$$

La matriz ampliada es de dimensión 2×4:

$$(A \,|\, \mathbf{b}) = \left(\begin{array}{ccc|c} 1 & -1 & 0 & 3 \\ 2 & 0 & 1 & -1 \end{array} \right).$$

◇

✓ **Ejemplo 2.** Construir el sistema de ecuaciones lineales que tiene como matriz ampliada:

$$(A \,|\, \mathbf{b}) = \left(\begin{array}{ccc|c} 0 & 2 & -1 & 1 \\ -1 & 4 & -2 & 0 \\ 2 & -3 & 0 & -1 \end{array} \right).$$

Solución: La matriz $(A \,|\, \mathbf{b})$ tiene dimensión 3×4 por lo que el número de incógnitas es $4 - 1 = 3$. Llamando x, y, z a las incógnitas, la expresión matricial del sistema es:

$$\begin{pmatrix} 0 & 2 & -1 \\ -1 & 4 & -2 \\ 2 & -3 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}.$$

Realizando el producto de las matrices A y $\mathbf{x}$ se puede construir el sistema lineal:

$$\begin{cases} 2y - z = 1 \\ -x + 4y - 2z = 0 \\ 2x - 3y = -1. \end{cases}$$

◇

Un tipo especial de sistemas de ecuaciones lineales es el siguiente:

Definición 3.2. Un sistema de ecuaciones lineales se dice **homogéneo** si todos los términos independientes son nulos:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1m}x_m = 0 \\ \quad\vdots \qquad\qquad\qquad\qquad\quad \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nm}x_m = 0. \end{cases}$$

En formato matricial, $\mathbf{b} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$, por lo que el sistema se expresa $A\mathbf{x} = \mathbf{0}$.

Dado un sistema de ecuaciones, el problema que se plantea es encontrar los valores desconocidos de las variables x_i que satisfacen todas las ecuaciones del mismo.

Definición 3.3. Dado un sistema de ecuaciones lineales, se dice que una asignación de valores a las incógnitas,

$$x_1 = k_1, \ldots, x_n = k_n,$$

es una **solución del sistema** si hace que se verifiquen todas las ecuaciones del sistema:

$$\begin{cases} a_{11}k_1 + a_{12}k_2 + \cdots + a_{1m}k_m = b_1 \\ \quad\vdots \qquad\qquad\qquad\qquad\quad \vdots \\ a_{n1}k_1 + a_{n2}k_2 + \cdots + a_{nm}k_m = b_n. \end{cases}$$

Definición 3.4. Dado un sistema de ecuaciones, se denomina:

- **Solución general** al conjunto de todas las soluciones del sistema.
- **Solución particular** a una solución concreta del mismo.

Definición 3.5. Dos sistemas se dicen **equivalentes** si tienen la misma solución general.

✓ **Ejemplo 3.**

- El sistema

$$\begin{cases} x + y = 3 \\ 2x + 2y = 6 \end{cases}$$

tiene como solución general el conjunto $\{(\lambda, 3-\lambda)\ |\lambda \in \mathbb{R}\}$, mientras que $x=1$, $y=2$ es una solución particular del mismo. Observar que la solución particular pertenece al conjunto que define la solución general tomando $\lambda = 1$.

- El sistema anterior y el siguiente son equivalentes:

$$\left\{\begin{array}{l} -x-y=-3 \\ 2x+2y=6 \\ -3x-3y=-9. \end{array}\right.$$

◇

3.2. DISCUSIÓN DE UN SISTEMA DE ECUACIONES

Dependiendo del número de soluciones, los sistemas de ecuaciones lineales se clasifican del siguiente modo:

Definición 3.6. Un sistema de ecuaciones lineales se dice:

1. **Sistema compatible**, (SC), si tiene solución:
    - **Sistema compatible determinado**, (SCD), si tiene solución única.
    - **Sistema compatible indeterminado**, (SCI), si tiene infinitas soluciones.
2. **Sistema incompatible**, (SI), si no tiene solución.

En ocasiones, antes de resolver un sistema de ecuaciones resulta conveniente clasificarlo según la definición anterior. Si se trabaja en forma matricial, $A\mathbf{x} = \mathbf{b}$, la relación entre el rango de la matriz de coeficientes A y el rango de la matriz ampliada $(A\,|\,\mathbf{b})$ permite clasificar el sistema lineal, como muestra el siguiente teorema:

Teorema 3.7 (Teorema de Rouché-Frobenius)**.** Dado un sistema de n ecuaciones lineales con m incógnitas, $A\mathbf{x} = \mathbf{b}$, se verifica:

1. Si $\operatorname{rg} A = \operatorname{rg}(A\,|\,\mathbf{b})$, entonces el sistema es compatible.
    - Si $\operatorname{rg} A = \operatorname{rg}(A\,|\,\mathbf{b}) = m$ (número de incógnitas), el sistema es compatible determinado.
    - Si $\operatorname{rg} A = \operatorname{rg}(A\,|\,\mathbf{b}) < m$ (número de incógnitas), el sistema es compatible indeterminado.
2. Si $\operatorname{rg} A \neq \operatorname{rg}(A\,|\,\mathbf{b})$, entonces el sistema es incompatible.

Observaciones.

- Puesto que $A \in \mathrm{Mat}_{n\times m}(\mathbb{R})$ y $(A\,|\,\mathbf{b}) \in \mathrm{Mat}_{n\times(m+1)}(\mathbb{R})$, siempre se verifica
$$\operatorname{rg} A \leq \operatorname{rg}(A\,|\,\mathbf{b}).$$
- En el caso particular de los sistemas homogéneos ($\mathbf{b} = \mathbf{0}$) se tiene siempre que $\operatorname{rg} A = \operatorname{rg}(A\,|\,\mathbf{b})$, es decir, son siempre compatibles, siendo una de sus soluciones la **solución trivial**: $(x_1, \ldots, x_m) = (0, \ldots, 0)$. Más aún, si un sistema homogéneo es compatible determinado, su única solución es la trivial.

✓ **Ejemplo 4.** Clasificar el siguiente sistema dependiendo de los valores del parámetro a:

$$\begin{cases} x + 2y + z = 3 \\ x + 3y - z = 4 \\ x + 2y + (a^2 - 8)z = a. \end{cases}$$

Solución: Para clasificar el sistema se utiliza el teorema de Rouché-Frobenius. En primer lugar, se construye la matriz ampliada del sistema:

$$(A\,|\,\mathbf{b}) = \left(\begin{array}{ccc|c} 1 & 2 & 1 & 3 \\ 1 & 3 & -1 & 4 \\ 1 & 2 & a^2-8 & a \end{array}\right).$$

Para estudiar el rango de A y el rango de $(A\,|\,\mathbf{b})$ se lleva la matriz ampliada a una forma escalonada:

$$(A\,|\,\mathbf{b}) = \left(\begin{array}{ccc|c} 1 & 2 & 1 & 3 \\ 1 & 3 & -1 & 4 \\ 1 & 2 & a^2-8 & a \end{array}\right) \xrightarrow[\substack{P_{21}(-1)\\ P_{31}(-1)}]{} \left(\begin{array}{ccc|c} \mathbf{1} & 2 & 1 & 3 \\ 0 & \mathbf{1} & -2 & 1 \\ 0 & 0 & a^2-9 & a-3 \end{array}\right).$$

La existencia de dos pivotes garantiza que $\operatorname{rg} A \geq 2$, y por tanto $\operatorname{rg}(A\,|\,\mathbf{b}) \geq 2$. El $\operatorname{rg} A$ es 3 dependiendo de que se anule o no el elemento $a_{33} = a^2 - 9$, mientras que el rango de la matriz ampliada $(A\,|\,\mathbf{b})$ depende de la nulidad simultánea de los elementos $a_{33} = a^2 - 9$ y $a_{34} = a - 3$.

$$\begin{cases} a^2 - 9 = 0 & \Longleftrightarrow a = \pm 3 \\ a - 3 = 0 & \Longleftrightarrow a = 3. \end{cases}$$

Discusión del sistema:

$$\begin{cases} a \neq \pm 3, & \operatorname{rg} A = 3 = \operatorname{rg}(A\,|\,\mathbf{b}) = \text{número de incógnitas} & \Longrightarrow \boxed{\text{SCD}} \\ a = 3, & \operatorname{rg} A = 2 = \operatorname{rg}(A\,|\,\mathbf{b}) < \text{número de incógnitas} = 3 & \Longrightarrow \boxed{\text{SCI}} \\ a = -3, & \operatorname{rg} A = 2,\ \operatorname{rg}(A\,|\,\mathbf{b}) = 3 & \Longrightarrow \boxed{\text{SI}} \end{cases}$$

◇

Observación. Cuando se habla de un sistema de n ecuaciones con m incógnitas se tiende a pensar que $n \leq m$. Sin embargo, en numerosas ocasiones ocurre que $n > m$. En este caso, el sistema se dice **sobredeterminado**. En general, los sistemas sobredeterminados son incompatibles pero pueden modelizar problemas reales para los que es necesario conocer alguna solución.

Para la resolución de este tipo de sistemas se procede del siguiente modo:

a) Si el sistema es compatible, se resuelve con alguno de los métodos que se presentan en este capítulo.

b) Si el sistema es incompatible, en ocasiones es útil encontrar una buena aproximación a lo que se consideraría la solución del sistema, que tiene un gran componente geométrico como se observa en los siguientes ejemplos. Entre los procedimientos que se pueden utilizar para aproximar la solución de estos sistemas destaca el método de los mínimos cuadrados, para el que se puede utilizar la factorización QR de la matriz A asociada al sistema, como se ve en el Capítulo 6.

✓ **Ejemplo 5.** Determinar, en caso de que exista, la ecuación de la recta que pasa por los puntos $P_1(1,1)$, $P_2(2,0)$, $P_3(-1,3)$ y $P_4(\frac{1}{2},\frac{3}{2})$.

Solución: Observando que los cuatro puntos anteriores no están alineados verticalmente, la ecuación genérica de una recta r que pueda pasar por ellos toma la forma $y = mx + n$. Imponiendo que los puntos anteriores pertenezcan a dicha recta se obtiene el siguiente sistema:

$$\left.\begin{array}{lcl} P_1 \in r & \Longleftrightarrow & 1 = m + n \\ P_2 \in r & \Longleftrightarrow & 0 = 2m + n \\ P_3 \in r & \Longleftrightarrow & 3 = -m + n \\ P_4 \in r & \Longleftrightarrow & \frac{3}{2} = \frac{1}{2}m + n. \end{array}\right\}$$

Este sistema es sobredeterminado puesto que tiene 4 ecuaciones y solo 2 incógnitas. Para clasificarlo según el teorema de Roché-Frobenius, se procede en primer lugar a escalonar la matriz ampliada asociada al sistema:

$$\left(\begin{array}{cc|c} 1 & 1 & 1 \\ 2 & 1 & 0 \\ -1 & 1 & 3 \\ 1/2 & 1 & 3/2 \end{array}\right) \xrightarrow[\substack{P_{21}(-2) \\ P_{31}(1) \\ P_{41}(-1/2)}]{} \left(\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & 2 & 4 \\ 0 & 1/2 & 1 \end{array}\right) \xrightarrow[\substack{P_{32}(2) \\ P_{42}(1/2)}]{} \left(\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right).$$

Como $\operatorname{rg} A = 2 = \operatorname{rg}(A \,|\, \mathbf{b}) =$ número de incógnitas, el sistema es compatible determinado. La solución del mismo es $m = -1$ y $n = 2$, lo que da lugar a la recta buscada, $y = -x + 2$.

◇

✓ **Ejemplo 6.** Determinar, en caso de que exista, la ecuación de la recta que pasa por los puntos $P_1(1,1)$, $P_2(2,0)$, $P_3(-1,3)$ y $P_4(\frac{1}{2},0)$.

Solución: De manera análoga al ejemplo anterior, se plantea el siguiente sistema:

$$\begin{cases} 1 = m + n \\ 0 = 2m + n \\ 3 = -m + n \\ 0 = \frac{1}{2}m + n. \end{cases}$$

En este caso:

$$\left(\begin{array}{cc|c} 1 & 1 & 1 \\ 2 & 1 & 0 \\ -1 & 1 & 3 \\ 1/2 & 1 & 0 \end{array}\right) \xrightarrow[\substack{P_{21}(-2) \\ P_{31}(1) \\ P_{41}(-1/2)}]{} \left(\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & 2 & 4 \\ 0 & 1/2 & -1/2 \end{array}\right) \xrightarrow[\substack{P_{32}(2) \\ P_{42}(1/2)}]{} \left(\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & -3/2 \end{array}\right).$$

Puesto que $\operatorname{rg} A = 2$ y $\operatorname{rg}(A \,|\, \mathbf{b}) = 3$, se trata de un sistema incompatible, lo que quiere decir que no existe ninguna recta que pase por los cuatro puntos del enunciado. La pregunta que se plantea, y que se resuelve en el Capítulo 6, es cuál es la recta que, sin pasar por los cuatro puntos, mejor se ajusta a ellos.

◇

3.3. MÉTODOS DIRECTOS DE RESOLUCIÓN DE SISTEMAS DE ECUACIONES

El objetivo de esta sección es presentar algunos métodos matriciales que permiten obtener la solución exacta de un sistema de ecuaciones lineales, siempre, claro está, que el sistema sea compatible.

Como caso particular, suponer un sistema $A\mathbf{x} = \mathbf{b}$ donde $A \in \operatorname{Mat}_n(\mathbb{R})$ es una matriz regular. Por tanto, $\operatorname{rg} A = n = \operatorname{rg}(A \,|\, \mathbf{b})$ = número de incógnitas y el teorema de Rouché-Frobenius asegura que el sistema es compatible determinado, es decir, posee una única solución. Esta solución se puede obtener sin más que despejar $\mathbf{x}$ de la ecuación matricial $A\mathbf{x} = \mathbf{b}$:

$$\boxed{A\mathbf{x} = \mathbf{b} \xrightarrow{A \text{ regular}} \mathbf{x} = A^{-1}\mathbf{b}}$$

El procedimiento anterior tiene como limitación que solo es válido para sistemas de ecuaciones cuya matriz de coeficientes sea regular, lo que en particular implica que debe ser cuadrada.

Los métodos que se presentan a continuación, método de Gauss, Gauss-Jordan y LU son válidos para cualquier tipo de sistema.

3.3.1. MÉTODO DE GAUSS

Todo sistema de ecuaciones lineales se puede reducir a uno escalonado equivalente, es decir, a partir de un sistema dado se puede construir otro con las mismas soluciones y de resolución (casi) inmediata aplicando el método de sustitución. El procedimiento usado para este fin es el **método de eliminación gaussiana**, explicado en la Sección 2.3, aplicado a sistemas de ecuaciones.

Dado el sistema de ecuaciones lineales

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

se comienza construyendo la matriz ampliada del sistema. Aplicando el algoritmo de Gauss a dicha matriz se obtiene una nueva matriz escalonada:

$$\left(\begin{array}{ccccc|c} a_{11} & \cdots & a_{1j} & \cdots & a_{1m} & b_1 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} & b_i \\ \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nm} & b_n \end{array}\right) \longrightarrow \left(\begin{array}{ccccc|c} \boxed{1} & \cdots & * & \cdots & * & * \\ \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & \boxed{1} & \cdots & * & * \\ \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & \cdots & * & * \end{array}\right).$$

Se reconstruye el sistema con la matriz escalonada, obteniéndose un sistema equivalente al inicial y cuya solución, si la hay, se deduce de manera inmediata.

Para ilustrar el método, se presenta el siguiente ejemplo:

✓ **Ejemplo 7.** Considerar el sistema

$$\left\{\begin{array}{rcrcrcrcr} & & x_2 & + & 5x_3 & - & 3x_4 & = & -4 \\ x_1 & - & 2x_2 & & & + & 2x_4 & = & 5 \\ 2x_1 & + & 3x_2 & - & x_3 & + & x_4 & = & 0 \\ -3x_1 & + & x_2 & + & 8x_3 & - & 2x_4 & = & -6. \end{array}\right.$$

Un método ordenado y sistemático para hallar la solución del sistema es el de reducción, consistente en hacer operaciones con las ecuaciones del sistema para conseguir un nuevo sistema donde cada ecuación tiene un menor número de incógnitas que la anterior. De este modo se obtiene un sistema equivalente más sencillo.

El mismo resultado se puede obtener si se trabaja matricialmente y se aplica el método de eliminación gaussiana a la matriz ampliada del sistema:

$$(A \mid \mathbf{b}) = \left(\begin{array}{cccc|c} 0 & 1 & 5 & -3 & -4 \\ 1 & -2 & 0 & 2 & 5 \\ 2 & 3 & -1 & 1 & 0 \\ -3 & 1 & 8 & -2 & -6 \end{array}\right).$$

$$\left(\begin{array}{cccc|c} 0 & 1 & 5 & -3 & -4 \\ 1 & -2 & 0 & 2 & 5 \\ 2 & 3 & -1 & 1 & 0 \\ -3 & 1 & 8 & -2 & -6 \end{array}\right) \xrightarrow[P_{12}]{} \left(\begin{array}{cccc|c} \boxed{1} & -2 & 0 & 2 & 5 \\ 0 & 1 & 5 & -3 & -4 \\ 2 & 3 & -1 & 1 & 0 \\ -3 & 1 & 8 & -2 & -6 \end{array}\right) \xrightarrow[\substack{P_{21}(-2) \\ P_{41}(3)}]{}$$

$$\left(\begin{array}{cccc|c} \boxed{1} & -2 & 0 & 2 & 5 \\ 0 & \boxed{1} & 5 & -3 & -4 \\ 0 & 7 & -1 & -3 & -10 \\ 0 & -5 & 8 & 4 & 9 \end{array}\right) \xrightarrow[\substack{P_{32}(-7) \\ P_{42}(5)}]{} \left(\begin{array}{cccc|c} \boxed{1} & -2 & 0 & 2 & 5 \\ 0 & \boxed{1} & 5 & -3 & -4 \\ 0 & 0 & \boxed{-36} & 18 & 18 \\ 0 & 0 & 33 & -11 & -11 \end{array}\right) \xrightarrow[P_3(\frac{-1}{36})]{}$$

$$\left(\begin{array}{cccc|c} \boxed{1} & -2 & 0 & 2 & 5 \\ 0 & \boxed{1} & 5 & -3 & -4 \\ 0 & 0 & \boxed{1} & -1/2 & -1/2 \\ 0 & 0 & 33 & -11 & -11 \end{array}\right) \xrightarrow[P_{43}(-33)]{} \left(\begin{array}{cccc|c} \boxed{1} & -2 & 0 & 2 & 5 \\ 0 & \boxed{1} & 5 & -3 & -4 \\ 0 & 0 & \boxed{1} & -1/2 & -1/2 \\ 0 & 0 & 0 & \boxed{11/2} & 11/2 \end{array}\right) \xrightarrow[P_4(\frac{2}{11})]{}$$

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

Con la matriz escalonada obtenida a partir de la matriz ampliada $(A\,|\,\mathbf{b})$ se reconstruye el sistema:

$$\left\{\begin{array}{ccccccccc} x_1 & - & 2x_2 & & & + & 2x_4 & = & 5 \\ & & x_2 & + & 5x_3 & - & 3x_4 & = & -4 \\ & & & & x_3 & - & \frac{1}{2}x_4 & = & -\frac{1}{2} \\ & & & & & & x_4 & = & 1. \end{array}\right.$$

La solución se obtiene por *sustitución regresiva*. El valor de x_4 se sustituye en la tercera ecuación y se despeja x_3 y así sucesivamente hasta obtener la solución del sistema:

$$x_4 = 1, \quad x_3 = 0, \quad x_2 = -1, \quad x_1 = 1.$$

◇

3.3.2. MÉTODO DE GAUSS-JORDAN

Si tras aplicar el método de Gauss se continúa haciendo operaciones elementales a la matriz escalonada, obtenida a partir de la matriz ampliada $(A\,|\,\mathbf{b})$, para llevarla a su forma escalonada reducida, se llega a un sistema tan sencillo que las soluciones son inmediatas. Concretamente, si el sistema es compatible determinado y la matriz A es cuadrada, la forma escalonada reducida de la matriz de coeficientes A es la identidad y la solución del sistema aparece en la columna de los términos independientes. Este método se conoce como **método de Gauss-Jordan**.

$$\left(\begin{array}{ccccc|c} a_{11} & \cdots & a_{1j} & \cdots & a_{1m} & b_1 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} & b_i \\ \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nm} & b_n \end{array}\right) \xrightarrow{\text{Método de Gauss}}$$

$$\underbrace{\left(\begin{array}{ccccc|c} \boxed{1} & \cdots & * & \cdots & * & * \\ \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \cdots & \boxed{1} & \cdots & * & * \\ \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \cdots & * & * \end{array}\right)}_{\text{0's por debajo de los pivotes}} \xrightarrow[\text{Gauss-Jordan}]{\text{Método de}} \underbrace{\left(\begin{array}{ccccc|c} \boxed{1} & \cdots & \mathbf{0} & \cdots & * & * \\ \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & \boxed{1} & \cdots & * & * \\ \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & \cdots & * & * \end{array}\right)}_{\text{0's por encima y debajo de los pivotes}}$$

✓ **Ejemplo 8.** Resolver por el método de Gauss-Jordan el sistema del Ejemplo 7:

$$\left\{\begin{array}{rcrcrcrcr} & & x_2 & + & 5x_3 & - & 3x_4 & = & -4 \\ x_1 & - & 2x_2 & & & + & 2x_4 & = & 5 \\ 2x_1 & + & 3x_2 & - & x_3 & + & x_4 & = & 0 \\ -3x_1 & + & x_2 & + & 8x_3 & - & 2x_4 & = & -6. \end{array}\right.$$

Solución: Se parte de la última matriz obtenida en el Ejemplo 7 y se realizan operaciones elementales para hacer 0's por encima de los pivotes.

$$\left(\begin{array}{cccc|c} \boxed{1} & -2 & 0 & 2 & 5 \\ 0 & \boxed{1} & 5 & -3 & -4 \\ 0 & 0 & \boxed{1} & -1/2 & -1/2 \\ 0 & 0 & 0 & \boxed{1} & 1 \end{array}\right) \xrightarrow[\substack{P_{14}(-2) \\ P_{24}(3) \\ P_{34}(1/2)}]{} \left(\begin{array}{cccc|c} \boxed{1} & -2 & 0 & 0 & 3 \\ 0 & \boxed{1} & 5 & 0 & -1 \\ 0 & 0 & \boxed{1} & 0 & 0 \\ 0 & 0 & 0 & \boxed{1} & 1 \end{array}\right) \xrightarrow[P_{23}(-5)]{}$$

$$\left(\begin{array}{cccc|c} \boxed{1} & -2 & 0 & 0 & 3 \\ 0 & \boxed{1} & 0 & 0 & -1 \\ 0 & 0 & \boxed{1} & 0 & 0 \\ 0 & 0 & 0 & \boxed{1} & 1 \end{array}\right) \xrightarrow[P_{12}(2)]{} \left(\begin{array}{cccc|c} \boxed{1} & 0 & 0 & 0 & 1 \\ 0 & \boxed{1} & 0 & 0 & -1 \\ 0 & 0 & \boxed{1} & 0 & 0 \\ 0 & 0 & 0 & \boxed{1} & 1 \end{array}\right).$$

En este caso la última columna muestra directamente la solución del sistema:

$$x_1 = 1, \quad x_2 = -1, \quad x_3 = 0, \quad x_4 = 1.$$

◇

Observación. El método de Gauss-Jordan también es útil para obtener soluciones de sistemas compatibles indeterminados. Para ello, se recomienda tomar como parámetros las variables para las que no existe pivote en la correspondiente matriz escalonada.

✓ **Ejemplo 9.** Resolver el siguiente sistema compatible indeterminado utilizando el método de Gauss-Jordan:

$$\begin{cases} x+y+z=1 \\ x-y+2z=0. \end{cases}$$

Solución: Escalonando la matriz ampliada del sistema:

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 1 & -1 & 2 & 0 \end{array}\right) \xrightarrow[P_{21}(-1)]{} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & -2 & 1 & -1 \end{array}\right) \xrightarrow[P_2(-1/2)]{}$$

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & -1/2 & 1/2 \end{array}\right) \xrightarrow[P_{12}(-1)]{} \left(\begin{array}{ccc|c} \boxed{1} & 0 & 3/2 & 1/2 \\ 0 & \boxed{1} & -1/2 & 1/2 \end{array}\right).$$

Se observa que los pivotes aparecen en la primera y segunda columna, correspondientes a las variables x e y. Tomando como parámetro $z=\lambda$, se obtiene de manera inmediata la solución general del sistema:

$$(x,y,z)=\left(\tfrac{1-3\lambda}{2},\ \tfrac{1+\lambda}{2},\ \lambda\right), \qquad \lambda\in\mathbb{R}.$$

◇

3.3.3. FACTORIZACIÓN LU

Considerar el sistema de ecuaciones lineales $A\mathbf{x}=\mathbf{b}$, donde $A\in \mathrm{Mat}_{n\times m}(\mathbb{R})$. Si se tiene la factorización LU de A, es decir, $A=LU$, el sistema anterior se puede resolver utilizando el siguiente esquema:

$$A\mathbf{x}=\mathbf{b}\Longleftrightarrow (LU)\mathbf{x}=\mathbf{b}\Longleftrightarrow L\underbrace{(U\mathbf{x})}_{\mathbf{y}}=\mathbf{b}.$$

Haciendo un cambio de variable, llamando $\mathbf{y}=U\mathbf{x}$, resolver el sistema original $A\mathbf{x}=\mathbf{b}$ equivale a resolver los dos sistemas siguientes:

$$\boxed{A\mathbf{x}=\mathbf{b}\Longleftrightarrow \begin{cases} L\mathbf{y}=\mathbf{b} & (1) \\ U\mathbf{x}=\mathbf{y} & (2) \end{cases}}$$

- Sistema (1): $L\mathbf{y}=\mathbf{b}$. La matriz L y el vector $\mathbf{b}$ son conocidos, por lo que tiene sentido plantear el sistema $L\mathbf{y}=\mathbf{b}$. Por ser L regular, este sistema es compatible determinado por lo que se obtiene una única solución, llamada $\mathbf{y}$.
- Sistema (2): $U\mathbf{x}=\mathbf{y}$. La matriz U y el vector $\mathbf{y}$ son conocidos, por lo que al resolver este sistema se obtiene como solución $\mathbf{x}$, que es a su vez la solución del sistema original $A\mathbf{x}=\mathbf{b}$.

Observación. La ventaja de utilizar la factorización LU para resolver sistemas de ecuaciones consiste en que se sustituye el sistema original por dos cuya resolución es prácticamente inmediata ya que las matrices de coeficientes están escalonadas.

✓ **Ejemplo 10.** Resolver el sistema $A\mathbf{x} = \mathbf{b}$ utilizando una factorización LU de la matriz A.

$$A = \begin{pmatrix} 2 & 0 & 0 & 2 \\ 0 & 0 & -1 & -2 \\ 1 & 0 & 1 & 6 \end{pmatrix}, \qquad \mathbf{b} = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}.$$

Solución: Se puede comprobar que una factorización LU de A viene dada por:

$$L = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 1 & 1 & 3 \end{pmatrix}, \qquad U = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Utilizando el procedimiento anterior:

- Se resuelve el sistema (1), $L\mathbf{y} = \mathbf{b}$:

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 1 & 1 & 3 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} \Longrightarrow \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1/2 \\ 1 \\ 1/6 \end{pmatrix}.$$

- A continuación se resuelve el sistema (2), $U\mathbf{x} = \mathbf{y}$, obteniéndose la solución del sistema original $A\mathbf{x} = \mathbf{b}$:

$$\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1/2 \\ 1 \\ 1/6 \end{pmatrix} \Longrightarrow \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1/3 \\ \lambda \\ 2/3 \\ 1/6 \end{pmatrix}.$$

  Observar que al reconstruir el sistema la variable x_2 no aparece en ninguna ecuación, por lo que se considera un parámetro libre.

◇

La factorización LU es apropiada cuando se pretende resolver varios sistemas de ecuaciones que tienen la misma matriz de coeficientes:

$$A\mathbf{x}_1 = \mathbf{b}_1, \quad A\mathbf{x}_2 = \mathbf{b}_2, \quad \ldots, \quad A\mathbf{x}_s = \mathbf{b}_s.$$

En este caso es conveniente hallar la factorización LU de A y utilizarla con cada vector de términos independientes $\mathbf{b}_i$. El número de operaciones realizadas de este modo es menor que el de aplicar s veces el método de Gauss.

La resolución simultánea de sistemas con la misma matriz de coeficientes se utiliza, por ejemplo, para resolver sistemas del tipo $A^k\mathbf{x} = \mathbf{b}$ con $A \in \mathrm{Mat}_n(\mathbb{R})$, cuando no se quiere calcular la potencia k-ésima de A.

✓ **Ejemplo 11.** Para resolver el sistema $A^2\mathbf{x} = \mathbf{b}$, sin calcular A^2, se utiliza el siguiente esquema:

$$A^2\mathbf{x} = \mathbf{b} \Longleftrightarrow AA\mathbf{x} = \mathbf{b} \Longleftrightarrow A\underbrace{(A\mathbf{x})}_{\mathbf{y}} = \mathbf{b}.$$

Es decir, el sistema original es equivalente a resolver los dos sistemas siguientes:

$$\boxed{A^2\mathbf{x} = \mathbf{b} \Longleftrightarrow \begin{cases} A\mathbf{y} = \mathbf{b} & \text{(I)} \\ A\mathbf{x} = \mathbf{y} & \text{(II)} \end{cases}}$$

Se observa que los sistemas (I) y (II) tienen la misma matriz de coeficientes. Para su resolución se aplica la factorización LU de la matriz A siguiendo el esquema propuesto en el recuadro de la página 72, de modo que se obtienen cuatro sistemas de ecuaciones ya escalonados. La resolución de los mismos debe hacerse de manera ordenada ya que la solución de cada uno se utiliza como término independiente del sistema siguiente:

$$\text{(I)} \quad A\mathbf{y} = \mathbf{b} \Longleftrightarrow L\underbrace{U\mathbf{y}}_{\mathbf{z}} = \mathbf{b} \Longleftrightarrow \begin{cases} L\mathbf{z} = \mathbf{b}, \\ U\mathbf{y} = \mathbf{z}. \end{cases}$$

$$\text{(II)} \quad A\mathbf{x} = \mathbf{y} \Longleftrightarrow L\underbrace{U\mathbf{x}}_{\mathbf{t}} = \mathbf{y} \Longleftrightarrow \begin{cases} L\mathbf{t} = \mathbf{y}, \\ U\mathbf{x} = \mathbf{t}. \end{cases}$$

◇

3.4. ERRORES DE REDONDEO

Si un ordenador resuelve mediante el método de Gauss un sistema de 100 ecuaciones con 100 incógnitas, realiza aproximadamente $\frac{1}{3}10^6$ operaciones, cada una de ellas con sus posibles errores de redondeo. Esto puede afectar seriamente a las soluciones, haciendo que el resultado que devuelva el ordenador difiera mucho de la auténtica solución del sistema.

Imaginar, por simplicidad, que un ordenador trabaja con solo 3 cifras significativas, es decir, aproxima del siguiente modo:

$$0.345 + 0.00123 \cong 0.346; \qquad 0.345 + 0.00161 \cong 0.347.$$

Considerar el siguiente sistema,

$$\begin{cases} 0.0001x + y & = 1 \\ \quad\quad\;\; x + y & = 2, \end{cases}$$

cuya solución exacta es:

$$x = \frac{10000}{9999} \approx 1.000100010001, \quad y = \frac{9998}{9999} \approx 0.9998999899989999.$$

Si en el método de Gauss se utiliza como pivote en el primer paso el elemento 0.0001, el proceso se desarrolla del siguiente modo:

$$\left(\begin{array}{cc|c} 0.0001 & 1 & 1 \\ 1 & 1 & 2 \end{array}\right) \xrightarrow[P_1(10000)]{} \left(\begin{array}{cc|c} 1 & 10000 & 10000 \\ 1 & 1 & 2 \end{array}\right) \xrightarrow[P_{21}(-1)]{}$$

$$\left(\begin{array}{cc|c} 1 & 10000 & 10000 \\ 0 & -9999 & -9998 \end{array}\right) \xrightarrow[P_2(-1/9999)]{} \left(\begin{array}{cc|c} 1 & 10000 & 10000 \\ 0 & 1 & 9998/9999 \end{array}\right).$$

Aplicando el redondeo antes indicado

$$y = \frac{9998}{9999} = 0.99989999 \approx 1,$$

y despejando x de la ecuación

$$0.00001x + y = 1$$

se obtiene $(x, y) = (0, 1)$ que dista mucho de ser una solución del sistema original.

Sin embargo, si se utiliza como pivote el 1:

$$\left(\begin{array}{cc|c} 0.0001 & 1 & 1 \\ 1 & 1 & 2 \end{array}\right) \xrightarrow[P_{12}]{} \left(\begin{array}{cc|c} 1 & 1 & 2 \\ 0.0001 & 1 & 1 \end{array}\right) \xrightarrow[P_{21}(-0.0001)]{}$$

$$\left(\begin{array}{cc|c} 1 & 1 & 2 \\ 0 & 0.9999 & 0.9998 \end{array}\right) \xrightarrow[P_2(1/9999)]{} \left(\begin{array}{cc|c} 1 & 1 & 2 \\ 0 & 1 & 0.9998/0.9999 \end{array}\right).$$

De nuevo, el redondeo da lugar a:

$$y = \frac{0.9998}{0.9999} = 0.99989999 \approx 1.$$

Despejando de la ecuación $x + y = 2$ se obtiene la solución $x = y = 1$, que en este caso es el redondeo de la solución exacta.

Interesa, en general, escoger en cada paso el pivote posible más grande en valor absoluto (estrategia del **pivote parcial**). Los buenos programas de ordenador incorporan esta y otras estrategias para resolver sistemas de ecuaciones lineales.

3.5. MÉTODOS ITERATIVOS DE RESOLUCIÓN DE SISTEMAS DE ECUACIONES

En ocasiones no resulta posible o no es conveniente resolver un sistema de ecuaciones lineales por los métodos directos hasta ahora estudiados, debido al tamaño de la matriz, al coste computacional o a los errores de redondeo que se pueden generar. En esos casos se recurre a métodos iterativos que construyen una sucesión de vectores cuyo límite es la solución del sistema de ecuaciones lineales.

3.5.1. NORMAS VECTORIALES Y MATRICIALES

Antes de la construcción de los métodos iterativos es necesario introducir los siguientes conceptos.

Definición 3.8. Una **norma vectorial** en $\mathbb{R}^n$ es una aplicación,

$$\begin{array}{rccc} \|\cdot\|: & \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ & \mathbf{x} & \mapsto & \|\mathbf{x}\|, \end{array}$$

que verifica las siguientes propiedades:

- $\|\mathbf{x}\| > 0, \quad \forall \mathbf{x} \in \mathbb{R}^n,\ \mathbf{x} \neq \mathbf{0}; \qquad \|\mathbf{x}\| = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}.$
- $\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|, \quad \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n.$
- $\|\lambda\mathbf{x}\| = |\lambda|\|\mathbf{x}\|, \quad \forall \mathbf{x} \in \mathbb{R}^n,\ \forall \lambda \in \mathbb{R}.$

✓ **Ejemplo 12.** Sea $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$. Las normas vectoriales más utilizadas en $\mathbb{R}^n$ son las siguientes:

- *Norma* 1 *o norma taxi*:

$$\|\mathbf{x}\|_1 = |x_1| + |x_2| + \ldots + |x_n| = \sum_{i=1}^{n} |x_i|.$$

- *Norma* 2 *o norma euclídea*:

$$\|\mathbf{x}\|_2 = \left(|x_1|^2 + |x_2|^2 + \ldots + |x_n|^2\right)^{1/2} = \sqrt{\sum_{i=1}^{n} |x_i|^2}.$$

- *Norma* ∞ *o norma del máximo*:

$$\|\mathbf{x}\|_\infty = \text{máx}\{|x_1|, \ldots, |x_n|\} = \max_{1\leq i\leq n}\{|x_i|\}.$$

- *Norma p, ($p \geq 2$) o norma general*:

$$\|\mathbf{x}\|_p = (|x_1|^p + |x_2|^p + \ldots + |x_n|^p)^{1/p} = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}.$$

Para el vector $\mathbf{x} = (1, 0, -1) \in \mathbb{R}^3$ se tiene que:

$$\|\mathbf{x}\|_1 = 2, \quad \|\mathbf{x}\|_2 = \sqrt{2}, \quad \|\mathbf{x}\|_\infty = 1, \quad \|\mathbf{x}\|_p = 2^{1/p}.$$

◇

Observación. La norma de un vector proporciona una medida de la distancia entre dicho vector y el vector nulo. En el caso de dos vectores, $\|\mathbf{x} - \mathbf{y}\|$ representa la distancia entre ambos vectores.

Análogamente se puede definir el concepto de norma matricial para matrices cuadradas.

Definición 3.9. Una **norma matricial** para matrices $n \times n$ es una aplicación,

$$\begin{array}{rccc} \|\cdot\|: & \mathrm{Mat}_n(\mathbb{R}) & \longrightarrow & \mathbb{R} \\ & A & \mapsto & \|A\|, \end{array}$$

que verifica las siguientes propiedades:

- $\|A\| > 0, \quad \forall A \in \mathrm{Mat}_n(\mathbb{R}),\ A \neq \mathbf{0}; \quad \|A\| = 0 \Leftrightarrow A = \mathbf{0}.$
- $\|A + B\| \leq \|A\| + \|B\|, \quad \forall A, B \in \mathrm{Mat}_n(\mathbb{R}).$
- $\|\lambda A\| = |\lambda| \|A\|, \quad \forall A \in \mathrm{Mat}_n(\mathbb{R}),\ \forall \lambda \in \mathbb{R}.$

Existe relación entre normas vectoriales y matriciales.

Definición 3.10. Dada una norma matricial definida en $\mathrm{Mat}_n(\mathbb{R})$, $\|\cdot\|_M$, y una norma vectorial $\|\cdot\|_V$ de $\mathbb{R}^n$, se dice que ambas son normas **compatibles** si se verifica:

$$\|A\mathbf{x}\|_V \leq \|A\|_M \, \|\mathbf{x}\|_V, \qquad \forall A \in \mathrm{Mat}_n(\mathbb{R}),\ \forall \mathbf{x} \in \mathbb{R}^n.$$

Además se puede construir una norma matricial a partir de una norma vectorial como indica el siguiente teorema.

Teorema 3.11. Dada una norma vectorial en $\mathbb{R}^n$, $\|\cdot\|_V$, y una matriz $A \in \mathrm{Mat}_n(\mathbb{R})$, entonces la aplicación:

$$\|A\|_M = \max_{\|\mathbf{x}\|_V = 1} \|A\mathbf{x}\|_V$$

es una norma matricial, que es compatible con la norma vectorial que la induce.

Definición 3.12. La norma matricial anterior recibe el nombre de **norma matricial inducida** por la norma vectorial $\|\cdot\|_V$.

✓ **Ejemplo 13.** Se definen explícitamente la normas matriciales inducidas por las normas vectoriales usuales:

- Norma matricial inducida por la norma 1 vectorial:

$$\|A\|_1 = \max_{1\leq j\leq n} \left\{ \sum_{i=1}^{n} |a_{ij}| \right\}.$$

- Norma matricial inducida por la norma 2 vectorial:

$$\|A\|_2 = \big(\rho(A^t A)\big)^{1/2}.$$

  También recibe el nombre de *norma espectral*. Dada una matriz $B \in \mathrm{Mat}_n(\mathbb{R})$ con valores propios $\lambda_1, \ldots, \lambda_s$, se define el *radio espectral* de B como:

$$\rho(B) = \max_{1\leq i\leq s} |\lambda_i|.$$

  En el Capítulo 8 se trata el tema de valores propios.

- Norma matricial inducida por la norma ∞ vectorial:

$$\|A\|_\infty = \max_{1\leq i\leq n} \left\{ \sum_{j=1}^{n} |a_{ij}| \right\}.$$

◇

Observación. Se verifica además que para cualquier norma matricial y cualquier matriz $A \in \mathrm{Mat}_n(\mathbb{R})$:

$$\rho(A) \leq \|A\|,$$

es decir, el valor de cualquier norma matricial sobre A está acotado inferiormente por el radio espectral de la matriz A.

✓ **Ejemplo 14.** Dada la matriz $A = \begin{pmatrix} -1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & -8 & 9 \end{pmatrix}$, calcular $\|A\|_1$ y $\|A\|_\infty$.

Solución: Para el cálculo de $\|A\|_1$ se debe sumar el valor absoluto de los términos de cada columna:

- $j = 1: \quad |a_{11}| + |a_{21}| + |a_{31}| = 1 + 4 + 7 = 12.$
- $j = 2: \quad |a_{12}| + |a_{22}| + |a_{32}| = 2 + 5 + 8 = 15.$
- $j = 3: \quad |a_{13}| + |a_{23}| + |a_{33}| = 3 + 6 + 9 = 18.$

Ahora se halla el máximo de los valores obtenidos $\|A\|_1 = \max\{12, 15, 18\} = 18$.

Del mismo modo, para el cálculo de $\|A\|_\infty$ se debe sumar el valor absoluto de los términos de cada fila, obteniéndose 6, 15 y 24 respectivamente. En este caso, $\|A\|_\infty = \max\{6, 15, 24\} = 24$.

◇

Nota. En lo que resta de capítulo se consideran normas matriciales inducidas por normas vectoriales.

3.5.2. NOCIONES BÁSICAS SOBRE MÉTODOS ITERATIVOS

Se presentan a continuación las nociones básicas sobre métodos iterativos aplicados a la resolución de un sistema de ecuaciones lineales.

En lo que sigue, se considera un sistema compatible determinado de n ecuaciones con n incógnitas:

$$\boxed{A\mathbf{x} = \mathbf{b}, \quad A \in \operatorname{Mat}_n(\mathbb{R}), \quad \operatorname{rg} A = n}$$

Definición 3.13. Un **método iterativo** para encontrar la solución del sistema $A\mathbf{x} = \mathbf{b}$ consiste en la construcción de una sucesión de vectores de n componentes cada uno

$$\{\mathbf{x}^{(0)}, \mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(m)}, \ldots\}, \qquad \mathbf{x}^{(m)} = \begin{pmatrix} x_1^{(m)} \\ \vdots \\ x_n^{(m)} \end{pmatrix}, \qquad m \geq 0,$$

con la finalidad de que si $m \to \infty$, la sucesión anterior converge, componente a componente, a la solución del sistema $A\mathbf{x} = \mathbf{b}$, es decir:

$$\lim_{m\to\infty} \mathbf{x}^{(m)} = \mathbf{x} = A^{-1}\mathbf{b}.$$

El vector $\mathbf{x}^{(0)}$ de partida para iniciar un método iterativo es arbitrario y se denomina **vector inicial**. A partir de este vector se construye la sucesión recurrente de vectores $\{\mathbf{x}^{(0)}, \mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(m)}, \ldots\}$ considerando una función $F : \mathbb{R}^n \to \mathbb{R}^n$ de modo que

$$\mathbf{x}^{(m+1)} = F(\mathbf{x}^{(m)}), \qquad m = 0, 1, 2, \ldots$$

Observación. Aun fijada la función F se pueden construir diferentes sucesiones de vectores $\{\mathbf{x}^{(0)}, \mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(m)}, \ldots\}$ pues hay libertad para elegir el vector inicial $\mathbf{x}^{(0)}$. Es deseable que sea cual sea la sucesión de vectores construida, su límite sea la solución del sistema $A\mathbf{x} = \mathbf{b}$, lo que da pie a la siguiente definición.

Definición 3.14. Se dice que un método iterativo para el sistema $A\mathbf{x} = \mathbf{b}$ es **convergente** si para cualquier elección del vector inicial $\mathbf{x}^{(0)} \in \mathbb{R}^n$ se verifica

$$\lim_{m\to\infty} \mathbf{x}^{(m)} = \mathbf{x} = A^{-1}\mathbf{b}.$$

En el caso de métodos convergentes, el vector obtenido en cada iteración, $\mathbf{x}^{(m)}$, con m suficientemente grande, se considera una **solución aproximada** del sistema.

Un tema importante en el estudio de los métodos iterativos es controlar el error cometido al calcular una aproximación de la solución.

Definición 3.15. Sea el sistema de ecuaciones $A\mathbf{x} = \mathbf{b}$, y $\mathbf{x}^{(m)}$ la aproximación obtenida tras calcular m iteraciones de un método iterativo y $\|\cdot\|$ una norma vectorial cualquiera. Se define:

- El **error absoluto** como la distancia entre la iteración m-ésima y la solución del sistema
$$\|\mathbf{x}^{(m)} - \mathbf{x}\|.$$
- El **error de paso** como la distancia entre dos iteraciones sucesivas
$$\|\mathbf{x}^{(m)} - \mathbf{x}^{(m-1)}\|.$$

De aquí en adelante se consideran los métodos iterativos más sencillos, en los que F es una función de *tipo lineal*, $F(\mathbf{x}^{(m)}) = B\,\mathbf{x}^{(m)} + \mathbf{c}$, $m \geq 0$, donde $B \in \mathrm{Mat}_n(\mathbb{R})$ y $\mathbf{c} \in \mathbb{R}^n$. Así,

$$\boxed{\mathbf{x}^{(m+1)} = B\,\mathbf{x}^{(m)} + \mathbf{c}, \quad m \geq 0}$$

Al utilizar un método iterativo, en general, no es posible calcular el error absoluto puesto que la solución exacta del problema es desconocida. No obstante, en los métodos de tipo lineal se puede obtener una acotación del mismo utilizando el error de paso como indica el siguiente teorema.

Teorema 3.16. Dado el sistema de ecuaciones $A\mathbf{x} = \mathbf{b}$, se considera para su resolución un método iterativo lineal convergente, $\mathbf{x}^{(m+1)} = B\mathbf{x}^{(m)} + \mathbf{c}$, $m \geq 0$. El error absoluto cometido en la aproximación proporcionada por el método está acotado por:

$$\|\mathbf{x}^{(m)} - \mathbf{x}\| \leq \frac{\|B\|}{1 - \|B\|}\|\mathbf{x}^{(m)} - \mathbf{x}^{(m-1)}\|,$$

donde la norma matricial utilizada viene inducida por la norma vectorial.

Dado un método iterativo, para establecer un criterio de parada del mismo, se utilizan dos valores predefinidos por el usuario: un número natural M denominado *número máximo de iteraciones* y un valor numérico pequeño que recibe el nombre de *tolerancia*. El método se detiene si

- el error de paso tras m iteraciones, con $m < M$, es menor que la tolerancia predefinida, o bien
- se alcanza el número máximo de iteraciones permitidas.

En el primero de los casos se consigue una solución aproximada "suficientemente" buena, es decir, tan cercana como marca la tolerancia a la solución exacta del sistema.

En la práctica resulta sencillo comparar la tolerancia elegida con el error de paso y para relacionarla con el error absoluto se utiliza el teorema anterior.

3.5.3. CONSTRUCCIÓN DE MÉTODOS ITERATIVOS CONVERGENTES

El objetivo de esta sección es construir métodos iterativos convergentes del tipo $\mathbf{x}^{(m+1)} = B\,\mathbf{x}^{(m)} + \mathbf{c}$, $m \geq 0$. La elección de la matriz B y del vector $\mathbf{c}$ determina el método y sus propiedades son esenciales para la convergencia del mismo.

Definición 3.17. Sea $\mathbf{x}^{(m+1)} = B\,\mathbf{x}^{(m)} + \mathbf{c}$, $m \geq 0$, un método iterativo para hallar la solución del sistema $A\mathbf{x} = \mathbf{b}$. El método anterior se dice **consistente** si el sistema $\mathbf{x} = B\mathbf{x} + \mathbf{c}$ es equivalente a $A\mathbf{x} = \mathbf{b}$.

Para construir métodos convergentes se parte de métodos consistentes a los que es necesario imponer una condición adicional a la matriz de iteración.

Teorema 3.18. Un método iterativo $\mathbf{x}^{(m+1)} = B\,\mathbf{x}^{(m)} + \mathbf{c}$, $m \geq 0$, es convergente si y solo si es consistente y $\rho(B) < 1$.

Observación. La condición $\rho(B) < 1$ es equivalente a la existencia de alguna norma matricial de modo que $\|B\| < 1$.

A continuación se presenta un modo de construir métodos iterativos consistentes.

Teorema 3.19. El método iterativo $\mathbf{x}^{(m+1)} = B\,\mathbf{x}^{(m)} + \mathbf{c}$, $m \geq 0$, para el problema $A\mathbf{x} = \mathbf{b}$ es consistente si y solo si $\mathbf{c} = (I_n - B)A^{-1}\mathbf{b}$.

Para asegurar que $\mathbf{c} = (I_n - B)A^{-1}\mathbf{b}$ se puede considerar una descomposición de la matriz de coeficientes A del sistema como:

$$A = M - N,$$

siendo M una matriz regular, de modo que el sistema $A\mathbf{x} = \mathbf{b}$ se puede escribir como

$$M\mathbf{x} = N\mathbf{x} + \mathbf{b}, \quad \text{o bien,} \quad \mathbf{x} = M^{-1}N\mathbf{x} + M^{-1}\mathbf{b}.$$

Comparando esta última expresión con el método iterativo $\mathbf{x}^{(m+1)} = B\mathbf{x}^{(m)} + \mathbf{c}$, tiene sentido estudiar

$$\mathbf{x}^{(m+1)} = M^{-1}N\mathbf{x}^{(m)} + M^{-1}\mathbf{b},$$

es decir, considerar en el método

$$\boxed{B = M^{-1}N, \quad \mathbf{c} = M^{-1}\mathbf{b}}$$

Nota. Es fácil comprobar que el vector $\mathbf{c} = M^{-1}\mathbf{b}$ cumple la expresión del Teorema 3.19, es decir, que el método construido es consistente.

Una idea sencilla para construir métodos consistentes de la forma

$$\boxed{\mathbf{x}^{(m+1)} = \underbrace{M^{-1}N}_{B}\mathbf{x}^{(m)} + \underbrace{M^{-1}\mathbf{b}}_{\mathbf{c}}} \qquad (*)$$

es considerar como matriz M una matriz diagonal o triangular **sin ceros en la diagonal** (para asegurar su regularidad). Concretamente, se puede descomponer fácilmente la matriz A como

$$\boxed{A = D - L - U}$$

donde D es una matriz diagonal, L es una matriz triangular inferior y U es triangular superior:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \Longrightarrow$$

$$A = \underbrace{\begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}}_{D} - \underbrace{\begin{pmatrix} 0 & 0 & \cdots & 0 \\ -a_{21} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & 0 \end{pmatrix}}_{L} - \underbrace{\begin{pmatrix} 0 & -a_{12} & \cdots & -a_{1n} \\ 0 & 0 & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}}_{U}.$$

Ahora, dependiendo de la elección de las matrices intermedias M y N usando D, L y U se definen dos de los métodos más utilizados. Para asegurar la convergencia se aplica el Teorema 3.18.

3.5.4. MÉTODO DE JACOBI

El método de Jacobi para encontrar una solución aproximada del sistema $A\mathbf{x} = \mathbf{b}$ es un método de la forma $(*)$ donde $A = D - L - U$ y

$$\boxed{M = D, \qquad N = L + U}$$

Ahora, es necesario calcular $B = M^{-1}N$ y $\mathbf{c} = M^{-1}\mathbf{b}$ para la construcción del método iterativo.

Observación. En el método de Jacobi, $M = \operatorname{diag}(a_{11}, \ldots, a_{nn})$. Si algún a_{ii} fuese cero, M no sería regular por lo que antes de empezar el método sería necesario intercambiar las filas correspondientes de la matriz $(A \,|\, \mathbf{b})$ para obtener una matriz de coeficientes sin ningún cero en la diagonal. El cálculo de la matriz inversa M^{-1} es inmediato: $M^{-1} = \operatorname{diag}(a_{11}^{-1}, \ldots, a_{nn}^{-1})$.

El siguiente ejemplo ilustra cómo aplicar el método:

✓ **Ejemplo 15.** Obtener una aproximación a la solución del siguiente sistema de ecuaciones aplicando 5 iteraciones del método de Jacobi. Tomar como vector inicial $\mathbf{x}^{(0)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$:

$$\begin{cases} 7x_1 - x_2 = 6 \\ x_1 - 5x_2 = -4. \end{cases}$$

Solución: Fácilmente se puede comprobar que la solución exacta del sistema es $x_1 = 1$, $x_2 = 1$. La expresión matricial del sistema de ecuaciones es:

$$\underbrace{\begin{pmatrix} 7 & -1 \\ 1 & -5 \end{pmatrix}}_{A} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \underbrace{\begin{pmatrix} 6 \\ -4 \end{pmatrix}}_{\mathbf{b}}.$$

Se descompone la matriz A como $A = D - L - U$, siendo:

$$D = \begin{pmatrix} 7 & 0 \\ 0 & -5 \end{pmatrix}, \quad L = \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix}, \quad U = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Las matrices M y N del método de Jacobi son:

$$M = D = \begin{pmatrix} 7 & 0 \\ 0 & -5 \end{pmatrix}, \quad N = L + U = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Se observa que la matriz M es regular por ser diagonal y no tener ceros en la digonal. Con las matrices anteriores se construye la matriz de iteración B:

$$B = M^{-1}N = \begin{pmatrix} 1/7 & 0 \\ 0 & -1/5 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1/7 \\ 1/5 & 0 \end{pmatrix}.$$

Se comprueba que $\|B\|_\infty = \text{máx}\{1/7, 1/5\} = 1/5 < 1$, por lo que se asegura la convergencia del método. Se calcula a continuación el vector de iteración del método:

$$\mathbf{c} = M^{-1}\mathbf{b} = \begin{pmatrix} 1/7 & 0 \\ 0 & -1/5 \end{pmatrix} \begin{pmatrix} 6 \\ -4 \end{pmatrix} = \begin{pmatrix} 6/7 \\ 4/5 \end{pmatrix}.$$

La forma $(*)$ de expresar el método iterativo es:

$$\mathbf{x}^{(m+1)} = \begin{pmatrix} 0 & 1/7 \\ 1/5 & 0 \end{pmatrix} \mathbf{x}^{(m)} + \begin{pmatrix} 6/7 \\ 4/5 \end{pmatrix}.$$

Tomando como vector inicial $\mathbf{x}^{(0)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, las 5 primeras iteraciones del método son:

$$\mathbf{x}^{(1)} = B\mathbf{x}^{(0)} + \mathbf{c} = \begin{pmatrix} 0 & 1/7 \\ 1/5 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 6/7 \\ 4/5 \end{pmatrix} = \begin{pmatrix} 6/7 \\ 4/5 \end{pmatrix};$$

$$\mathbf{x}^{(2)} = B\mathbf{x}^{(1)} + \mathbf{c} = \begin{pmatrix} 0 & 1/7 \\ 1/5 & 0 \end{pmatrix} \begin{pmatrix} 6/7 \\ 4/5 \end{pmatrix} + \begin{pmatrix} 6/7 \\ 4/5 \end{pmatrix} = \begin{pmatrix} 34/35 \\ 34/35 \end{pmatrix};$$

$$\mathbf{x}^{(3)} = B\mathbf{x}^{(2)} + \mathbf{c} = \begin{pmatrix} 0 & 1/7 \\ 1/5 & 0 \end{pmatrix} \begin{pmatrix} 34/35 \\ 34/35 \end{pmatrix} + \begin{pmatrix} 6/7 \\ 4/5 \end{pmatrix} = \begin{pmatrix} 244/245 \\ 174/175 \end{pmatrix};$$

$$\mathbf{x}^{(4)} = B\mathbf{x}^{(3)} + \mathbf{c} = \begin{pmatrix} 0 & 1/7 \\ 1/5 & 0 \end{pmatrix} \begin{pmatrix} 244/245 \\ 174/175 \end{pmatrix} + \begin{pmatrix} 6/7 \\ 4/5 \end{pmatrix} = \begin{pmatrix} 1224/1225 \\ 1224/1225 \end{pmatrix};$$

$$\mathbf{x}^{(5)} = B\mathbf{x}^{(4)} + \mathbf{c} = \begin{pmatrix} 0 & 1/7 \\ 1/5 & 0 \end{pmatrix} \begin{pmatrix} 1224/1225 \\ 1224/1225 \end{pmatrix} + \begin{pmatrix} 6/7 \\ 4/5 \end{pmatrix} = \begin{pmatrix} 8574/8575 \\ 6124/6125 \end{pmatrix} \approx \begin{pmatrix} 0.9998 \\ 0.9998 \end{pmatrix}.$$

◇

3.5.5. MÉTODO DE GAUSS-SEIDEL

El método de Gauss-Seidel para encontrar una solución aproximada del sistema $A\mathbf{x} = \mathbf{b}$ es un método de la forma $(*)$ donde $A = D - L - U$ y

$$\boxed{M = D - L, \qquad N = U}$$

Hay que calcular $B = M^{-1}N$ y $\mathbf{c} = M^{-1}\mathbf{b}$ para la construcción del método iterativo.

Desde el punto de vista computacional, si el tamaño de la matriz es grande, este método tiene menos coste que el de Jacobi y requiere menos iteraciones.

Nota. Una variante del método de Gauss-Seidel consiste en tomar $M = D - U$ y $N = L$.

Observación. En este método la matriz M es triangular inferior. Para asegurar su regularidad, ningún elemento de su diagonal debe ser cero. Igual que en el método de Jacobi, si algún a_{ii} fuese cero, antes de empezar el método sería necesario intercambiar las filas correspondientes de la matriz $(A \,|\, \mathbf{b})$ para obtener una matriz de coeficientes sin ningún cero en la diagonal.

✓ **Ejemplo 16.** Dado el sistema de ecuaciones $\begin{cases} 7x_1 - x_2 = 6 \\ x_1 - 5x_2 = -4. \end{cases}$

- Obtener una aproximación a su solución aplicando 3 iteraciones del método de Gauss-Seidel y tomando como vector inicial el vector nulo.
- Calcular el error de paso entre las dos últimas iteraciones calculadas, utilizando $\|\cdot\|_\infty$. ¿Son suficientes tres iteraciones si se considera una tolerancia $\mathrm{tol} = 10^{-3}$?
- Obtener una cota del error absoluto utilizando el Teorema 3.16 con $m = 3$.

Solución: Se considera la descomposición anterior $A = D - L - U$, siendo

$$D = \begin{pmatrix} 7 & 0 \\ 0 & -5 \end{pmatrix}, \quad L = \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix}, \quad U = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Las matrices M (que es regular por no tener ceros en la diagonal) y N del método de Gauss-Seidel son:

$$M = D - L = \begin{pmatrix} 7 & 0 \\ 1 & -5 \end{pmatrix}, \qquad N = U = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

La matriz y el vector de iteración son:

$$B = M^{-1}N = \begin{pmatrix} 0 & 1/7 \\ 0 & 1/35 \end{pmatrix}, \quad \mathbf{c} = M^{-1}\mathbf{b} = \begin{pmatrix} 6/7 \\ 34/35 \end{pmatrix}.$$

Se puede asegurar la convergencia del método ya que se verifica $\|B\|_\infty = {}^{1}\!/\!{}_{7} < 1$. La expresión $(*)$ del método es por tanto:

$$\mathbf{x}^{(m+1)} = \begin{pmatrix} 0 & 1/7 \\ 0 & 1/35 \end{pmatrix} \mathbf{x}^{(m)} + \begin{pmatrix} 6/7 \\ 34/35 \end{pmatrix}.$$

Tomando como vector inicial $\mathbf{x}^{(0)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, las 3 primeras iteraciones del método son:

$$\mathbf{x}^{(1)} = \begin{pmatrix} 0 & 1/7 \\ 0 & 1/35 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 6/7 \\ 34/35 \end{pmatrix} = \begin{pmatrix} 6/7 \\ 34/35 \end{pmatrix};$$

$$\mathbf{x}^{(2)} = \begin{pmatrix} 0 & 1/7 \\ 0 & 1/35 \end{pmatrix} \begin{pmatrix} 6/7 \\ 34/35 \end{pmatrix} + \begin{pmatrix} 6/7 \\ 34/35 \end{pmatrix} = \begin{pmatrix} 244/245 \\ 1224/1225 \end{pmatrix};$$

$$\mathbf{x}^{(3)} = \begin{pmatrix} 0 & 1/7 \\ 0 & 1/35 \end{pmatrix} \begin{pmatrix} 244/245 \\ 1224/1225 \end{pmatrix} + \begin{pmatrix} 6/7 \\ 34/35 \end{pmatrix} = \begin{pmatrix} 8574/8575 \\ 42874/42875 \end{pmatrix} \approx \begin{pmatrix} 0.99988 \\ 0.99998 \end{pmatrix}.$$

Se calcula el error de paso con los resultados obtenidos en las dos últimas iteraciones:

$$\mathbf{x}^{(3)} - \mathbf{x}^{(2)} = \begin{pmatrix} 34/8575 \\ 34/42875 \end{pmatrix} \Longrightarrow \|\mathbf{x}^{(3)} - \mathbf{x}^{(2)}\|_\infty = \frac{34}{8575} \cong 3.96 \cdot 10^{-3}.$$

Se observa que el error obtenido es mayor que la tolerancia dada, por lo que tres iteraciones no son suficientes. Para conseguir un error menor que la tolerancia fijada es necesario realizar alguna iteración más.

Una cota del error absoluto viene dada por:

$$\|\mathbf{x}^{(3)} - \mathbf{x}\|_\infty \leq \frac{\|B\|_\infty}{1 - \|B\|_\infty} \|\mathbf{x}^{(3)} - \mathbf{x}^{(2)}\|_\infty = \frac{1/7}{6/7} 3.96 \cdot 10^{-3} = 6.608 \cdot 10^{-4}$$

◇

3.5.6. CONVERGENCIA DE LOS MÉTODOS DE JACOBI Y GAUSS-SEIDEL

Los métodos de Jacobi y Gauss-Seidel son métodos consistentes por construcción pero no siempre son convergentes. La condición que garantiza su convergencia, $\rho(B) < 1$, de momento no se utiliza por no estar definido el concepto de valor propio (se introduce en el Capítulo 8). La condición equivalente de existencia de una norma matricial tal que $\|B\| < 1$, en ocasiones puede no ser fácil de calcular. Se presenta a continuación una condición más sencilla de verificar que garantiza la convergencia de los métodos anteriores.

Definición 3.20. Una matriz $A \in \mathrm{Mat}_n(\mathbb{R})$ es **estrictamente diagonal dominante** (por filas) si el valor absoluto de cada elemento de la diagonal principal es mayor que la suma de los valores absolutos de los demás elementos de la fila:

$$|a_{ii}| > |a_{i1}| + \ldots + |a_{i,i-1}| + |a_{i,i+1}| + \ldots + |a_{in}|, \quad \forall i = 1, \ldots, n.$$

Observación. Si la matriz A no es estrictamente diagonal dominante, en ocasiones se puede obtener una matriz con esa propiedad intercambiando el orden de sus filas.

✓ **Ejemplo 17.** Estudiar si A_1 y A_2 son estrictamente diagonal dominantes por filas.

$$A_1 = \begin{pmatrix} -2 & 1 & 0 \\ -1 & 3 & -1 \\ 0 & 0 & 1 \end{pmatrix}, \qquad A_2 = \begin{pmatrix} 1 & 1 & 3 \\ -1 & 3 & -1 \\ 2 & 0 & 1 \end{pmatrix}.$$

Solución: La matriz A_1 si que lo es, ya que:

- Fila 1: $|a_{11}| = 2 > 1 = 1 + 0 = |a_{12}| + |a_{13}|$.
- Fila 2: $|a_{22}| = 3 > 2 = 1 + 1 = |a_{21}| + |a_{23}|$.
- Fila 3: $|a_{33}| = 1 > 0 = 0 + 0 = |a_{31}| + |a_{32}|$.

La matriz A_2 no lo es debido a que falla la condición en las filas 1 y 3 pero se puede conseguir una que sí lo sea intercambiando las filas 2 y 3.

◇

El hecho de que una matriz sea estrictamente diagonal dominante proporciona una condición suficiente pero no necesaria de convergencia para los métodos.

Teorema 3.21. Sea el sistema de ecuaciones $A\mathbf{x} = \mathbf{b}$ con $A \in \mathrm{Mat}_n(\mathbb{R})$ una matriz regular. Si A es estrictamente diagonal dominante (por filas), los métodos de Jacobi y de Gauss-Seidel son ambos convergentes.

Observación. El hecho de que A sea estrictamente diagonal dominante asegura que la matriz de iteración B en los métodos de Jacobi y Gauss-Seidel verifica $\|B\|_\infty < 1$, y por tanto la convergencia del método. En caso contrario no se puede afirmar nada sobre la convergencia.

A la vista de los resultados anteriores se proponen unas etapas a seguir para estudiar la convergencia de los métodos de Jacobi y de Gauss-Seidel:

1. Estudiar si la matriz A es estrictamente diagonal dominante. En caso afirmativo se tiene la convergencia del método, en otro caso,
2. comprobar $\|B\|_1 < 1$ o $\|B\|_\infty < 1$. Si fallan las condiciones anteriores buscar otra norma matricial que verifique $\|B\| < 1$ o recurrir a la condición,
3. $\rho(B) < 1$.

✓ **Ejemplo 18.** Estudiar la convergencia del método de Jacobi para el sistema:

$$\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Se observa que la matriz de coeficientes A no es estrictamente diagonal dominante, incluso si se permuta el orden de las ecuaciones del sistema la nueva matriz de coeficientes tampoco lo es.

La matriz del método iterativo de Jacobi es $B = \begin{pmatrix} 0 & -2 \\ 0 & 0 \end{pmatrix}$. De manera inmediata se obtiene que $\|B\|_1 = 2$ y $\|B\|_\infty = 2$, por lo que ninguna de las dos primeras condiciones de la observación anterior se verifican. Sin embargo, la matriz B verifica $\rho(B) = 0 < 1$ (de momento no se sabe realizar este cálculo) por lo que el método es convergente. El hecho de que $\rho(B) = 0 < 1$ es equivalente a la existencia de una norma matricial (no necesariamente usual) para la cual $\|B\| < 1$.

◇

A continuación, se presenta un cuadro resumen de cómo aplicar los métodos de Jacobi y Gauss-Siedel.

	Jacobi	**Gauss-Seidel**
Expresión matricial del sistema	$A\mathbf{x} = \mathbf{b}$	
Descomposición de A	$A = D - L - U$	
Paso intermedio	$M = D$ $N = L + U$	$M = D - L$ $N = U$
Matriz de iteración	$B = M^{-1}N$	
Condiciones necesarias de convergencia	Etapas 1, 2, 3	
Vector de iteración	$\mathbf{c} = M^{-1}\mathbf{b}$	
Método	$\mathbf{x}^{(m+1)} = B\mathbf{x}^{(m)} + \mathbf{c}$	
Vector inicial	Elegir $\mathbf{x}^{(0)}$	
Iteraciones	$\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(r)}$	
Criterios de parada	Error $\leq$ tol Máximo número de iteraciones	

3.6. APLICACIONES

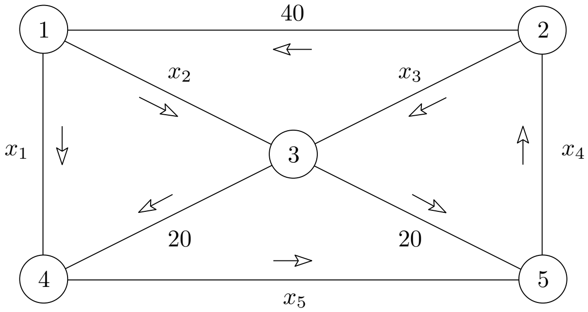
Los sistemas de ecuaciones aparecen en multitud de situaciones tanto cotidianas como de la Ciencia y la Ingeniería. Al resolver problemas utilizando sistemas de ecuaciones es muy importante la interpretación de las soluciones obtenidas ya que en ocasiones hay que desechar algunas de ellas. Pensar, por ejemplo, en un problema en el que hay que calcular el número de tropas destinadas en una misión. El resultado del problema ha de ser un número entero positivo, por lo que se deberían desechar las soluciones que fueran números negativos o decimales.

A continuación se presentan dos aplicaciones del uso de sistemas de ecuaciones lineales, una en un contexto físico y otra en un contexto geométrico.

3.6.1. ANÁLISIS DE REDES

Muchos de los problemas en Física e Ingeniería que se pueden resolver a través de sistemas de ecuaciones lineales tienen que ver con el análisis de redes. Este tipo de situaciones se basan siempre en la misma premisa: el flujo que entra en un nodo o unión debe ser igual al que sale. De esta manera, cada nodo da lugar a una ecuación lineal.

✓ **Ejemplo 19.** Calcular el flujo que circula por cada arista de la siguiente red:



Solución: Para plantear el problema hay que igualar el flujo que entra y sale en cada nodo. Así, el sistema de ecuaciones lineales asociado y su expresión matricial es:

$$\left\{\begin{array}{ll} \text{Nodo 1}: & 40 = x_1 + x_2 \\ \text{Nodo 2}: & x_4 = 40 + x_3 \\ \text{Nodo 3}: & x_2 + x_3 = 40 \\ \text{Nodo 4}: & 20 + x_1 = x_5 \\ \text{Nodo 5}: & 20 + x_5 = x_4 \end{array}\right. \Longrightarrow \left(\begin{array}{ccccc|c} 1 & 1 & 0 & 0 & 0 & 40 \\ 0 & 0 & 1 & -1 & 0 & -40 \\ 0 & 1 & 1 & 0 & 0 & 40 \\ 1 & 0 & 0 & 0 & -1 & -20 \\ 0 & 0 & 0 & 1 & -1 & 20 \end{array}\right).$$

Se resuelve el ejercicio utilizando el método de Gauss-Jordan:

$$\left(\begin{array}{ccccc|c} 1 & 1 & 0 & 0 & 0 & 40 \\ 0 & 0 & 1 & -1 & 0 & -40 \\ 0 & 1 & 1 & 0 & 0 & 40 \\ 1 & 0 & 0 & 0 & -1 & -20 \\ 0 & 0 & 0 & 1 & -1 & 20 \end{array}\right) \xrightarrow[P_{41}(-1)]{} \left(\begin{array}{ccccc|c} 1 & 1 & 0 & 0 & 0 & 40 \\ 0 & 0 & 1 & -1 & 0 & -40 \\ 0 & 1 & 1 & 0 & 0 & 40 \\ 0 & -1 & 0 & 0 & -1 & -60 \\ 0 & 0 & 0 & 1 & -1 & 20 \end{array}\right) \xrightarrow[P_{23}]{}$$

$$\left(\begin{array}{ccccc|c} 1 & 1 & 0 & 0 & 0 & 40 \\ 0 & 1 & 1 & 0 & 0 & 40 \\ 0 & 0 & 1 & -1 & 0 & -40 \\ 0 & -1 & 0 & 0 & -1 & -60 \\ 0 & 0 & 0 & 1 & -1 & 20 \end{array}\right) \xrightarrow[P_{42}(1)]{} \left(\begin{array}{ccccc|c} 1 & 1 & 0 & 0 & 0 & 40 \\ 0 & 1 & 1 & 0 & 0 & 40 \\ 0 & 0 & 1 & -1 & 0 & -40 \\ 0 & 0 & 1 & 0 & -1 & -20 \\ 0 & 0 & 0 & 1 & -1 & 20 \end{array}\right) \xrightarrow[P_{43}(-1)]{}$$

$$\left(\begin{array}{ccccc|c} 1 & 1 & 0 & 0 & 0 & 40 \\ 0 & 1 & 1 & 0 & 0 & 40 \\ 0 & 0 & 1 & -1 & 0 & -40 \\ 0 & 0 & 0 & 1 & -1 & 20 \\ 0 & 0 & 0 & 1 & -1 & 20 \end{array}\right) \xrightarrow[P_{54}(-1)]{} \left(\begin{array}{ccccc|c} 1 & 1 & 0 & 0 & 0 & 40 \\ 0 & 1 & 1 & 0 & 0 & 40 \\ 0 & 0 & 1 & -1 & 0 & -40 \\ 0 & 0 & 0 & 1 & -1 & 20 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \xrightarrow[P_{34}(1)]{}$$

$$\left(\begin{array}{ccccc|c} 1 & 1 & 0 & 0 & 0 & 40 \\ 0 & 1 & 1 & 0 & 0 & 40 \\ 0 & 0 & 1 & 0 & -1 & -20 \\ 0 & 0 & 0 & 1 & -1 & 20 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \xrightarrow[P_{23}(-1)]{} \left(\begin{array}{ccccc|c} 1 & 1 & 0 & 0 & 0 & 40 \\ 0 & 1 & 0 & 0 & 1 & 60 \\ 0 & 0 & 1 & 0 & -1 & -20 \\ 0 & 0 & 0 & 1 & -1 & 20 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \xrightarrow[P_{12}(-1)]{}$$

$$\left(\begin{array}{ccccc|c} 1 & 0 & 0 & 0 & -1 & -20 \\ 0 & 1 & 0 & 0 & 1 & 60 \\ 0 & 0 & 1 & 0 & -1 & -20 \\ 0 & 0 & 0 & 1 & -1 & 20 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Se observa que se trata de un sistema compatible indeterminado, por lo que la red tiene infinitas soluciones. Tomando como parámetro $x_5 = \lambda$, los valores buscados son:

$$x_1 = -20 + \lambda, \quad x_2 = 60 - \lambda, \quad x_3 = -20 + \lambda, \quad x_4 = 20 + \lambda, \quad x_5 = \lambda.$$

Como los valores del flujo han de ser todos positivos, el valor de λ debe estar acotado: $20 < \lambda < 60$.

◇

3.6.2. APLICACIONES GEOMÉTRICAS

Los sistemas de ecuaciones lineales así como su clasificación según el número de soluciones tienen una clara interpretación geométrica. A continuación se presenta lo que ocurre en el plano y en el espacio.

Sistemas de ecuaciones lineales en el plano

La ecuación general de una recta en el plano es una ecuación lineal en dos variables:

$$r : ax + by + c = 0.$$

Los sistemas de dos ecuaciones lineales con dos incógnitas estudian la posición relativa de dos rectas en el plano. Considerar el sistema formado por dos rectas:

$$\begin{cases} r_1 : a_1x + b_1y + c_1 = 0 \\ r_2 : a_2x + b_2y + c_2 = 0. \end{cases}$$

La matriz ampliada asociada al sistema es:

$$(A\,|\,\mathbf{b}) = \left(\begin{array}{cc|c} a_1 & b_1 & -c_1 \\ a_2 & b_2 & -c_2 \end{array} \right).$$

Mediante el teorema de Rouché-Frobenius se puede clasificar el sistema anterior, siendo su interpretación geométrica la siguiente:

Posición relativa de 2 rectas:

$\mathrm{rg}(A\,\vert\,\mathbf{b})$	$\mathrm{rg}\,A$	Incógnitas	Clasificación	Interpretación geométrica
2	2	2	SCD	rectas que se cortan en un punto
2	1	2	SI	rectas paralelas
1	1	2	SCI	rectas coincidentes

Para estudiar la posición relativa de 3 rectas r_1, r_2 y r_3 la situación es similar:

Posición relativa de 3 rectas:

$\mathrm{rg}(A\,\vert\,\mathbf{b})$	$\mathrm{rg}\,A$	Incógnitas	Clasif.	Interpretación geométrica
3	2	2	SI	dos rectas paralelas y una transversal, o bien, las rectas se cortan dos a dos
3	1	2	SI	rectas paralelas
2	2	2	SCD	las rectas se cortan en un punto
2	1	2	SI	dos rectas coincidentes y la tercera paralela
1	1	2	SCI	tres rectas coincidentes

Sistemas de ecuaciones lineales en el plano

La ecuación general de un plano en el espacio es una ecuación lineal en 3 variables:

$$\pi : ax + by + cz + d = 0.$$

De manera análoga al estudio anterior de la posición relativa de dos y tres rectas en el plano, se puede estudiar la posición relativa de dos y tres planos en el espacio, siendo esta como sigue:

Posición relativa de 2 planos:

$\mathrm{rg}(A \mid \mathbf{b})$	$\mathrm{rg}\, A$	Incógnitas	Clasificación	Interpretación geométrica
2	2	3	SCI	planos que se cortan en una recta
2	1	3	SI	planos paralelos
1	1	3	SCI	planos coincidentes

Posición relativa de 3 planos:

$\mathrm{rg}(A \mid \mathbf{b})$	$\mathrm{rg}\, A$	Incógnitas	Clasif.	Interpretación geométrica
3	3	3	SCD	los planos se cortan en un punto
3	2	3	SI	dos planos paralelos y uno transversal, o bien, los planos se cortan dos a dos
3	1	3	SI	planos paralelos
2	2	3	SCI	haz de planos (tienen una recta en común)
2	1	3	SI	dos planos coincidentes y el tercero paralelo
1	1	3	SCI	tres planos coincidentes

Como se acaba de ver al estudiar la posición de dos planos, se puede interpretar las rectas en el espacio como la intersección de dos planos:

$$r = \pi_1 \cap \pi_2 = \cap \begin{cases} \pi_1 : a_1x + b_1y + c_1z + d_1 = 0 \\ \pi_2 : a_2x + b_2y + c_2z + d_2 = 0, \end{cases} \quad \text{donde } \mathrm{rg} \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix} = 2.$$

Para estudiar la posición relativa de un plano π y una recta $r = \pi_1 \cap \pi_2$ se considera un sistema de tres ecuaciones con tres incógnitas y para estudiar la posición relativa de dos rectas se considera un sistema de cuatro ecuaciones con tres incógnitas. En ambos casos hay que suponer que el rango de la matriz de coeficientes es mayor o igual que 2:

Posición relativa de una recta y un plano:

$\text{rg}(A \mid \mathbf{b})$	$\text{rg}\, A$	Incógnitas	Clasificación	Interpretación geométrica
3	3	3	SCD	la recta corta al plano en un punto
3	2	3	SI	la recta y el plano son paralelos
2	2	3	SCI	la recta está contenida en el plano

Posición relativa de 2 rectas:

$\text{rg}(A \mid \mathbf{b})$	$\text{rg}\, A$	Incógnitas	Clasificación	Interpretación geométrica
4	3	3	SI	las rectas se cruzan
3	3	3	SCD	las rectas se cortan en un punto
3	2	3	SI	las rectas son paralelas
2	2	3	SI	las rectas son coincidentes

Observación. En este contexto geométrico, el hecho de que una ecuación sea homogénea se interpreta como que la recta o el plano correspondiente pasa por el origen de coordenadas.

3.7. COMANDOS DE *wxMaxima*

- Resolver directamente un sistema de ecuaciones lineales:

  ```
  -->     linsolve([ec1, ..., ecn], [var1, ..., varm]);
  ```

  donde `eci` hace referencia a la ecuación i-ésima del sistema y `varj` a la variable j-ésima.

  Directamente desde el menú: “`\Ecuaciones\Resolver sistema lineal`”.

- Eliminar variables de un sistema de ecuaciones:

  ```
  -->     eliminate([ec1, ..., ecn], [var1, ..., vars]);
  ```

  donde las variables que se eliminan son las que se han escrito en el comando.

  Directamente desde el menú: “`\Ecuaciones\Eliminar variable`”.

- Extraer la matriz de coeficientes de un sistema de ecuaciones:

  ```
  -->     coefmatrix([ec1, ..., ecn], [var1, ..., varm]);
  ```

- Construir la matriz ampliada de un sistema de ecuaciones $A\mathbf{x} = \mathbf{b}$:

  ```
  -->     addcol(A,b);
  ```

  donde **b** puede estar introducido como una lista.

- El comando

  ```
  -->     augcoefmatrix([ec1, ..., ecn], [var1, ..., varm]);
  ```

  construye directamente la matriz ampliada de un sistema de ecuaciones lineales aunque no lo hace como se ha definido en este texto sino que cambia el signo del vector de términos independientes. No obstante, esta modificación no es relevante para el estudio del rango de la matriz.

- Para escalonar matrices se pueden utilizar los comandos vistos en el Capítulo 2 “`echelon`” y “`triangularize`”.

- Resolución del sistema lineal $A\mathbf{x} = \mathbf{b}$ mediante matrices, siendo $A \in \mathrm{Mat}_n(\mathbb{R})$ y regular.

  ```
  -->     A:matrix(fila1,...,filam)$
          b:matrix(columna)$
          linsolve_by_lu(A,b);
  ```

- Resolución de sistemas lineales homogéneos, $A\mathbf{x} = \mathbf{0}$:

```
-->      nullspace(A);
```

- Normas matriciales: Si $A \in \mathrm{Mat}_{n\times m}(\mathbb{R})$, se pueden calcular $||A||_1$ y $||A||_\infty$.

```
-->      mat_norm(A,1);

-->      mat_norm(A,inf);
```

3.8. EJERCICIOS PROPUESTOS

Algoritmo de Gauss y Teorema de Rouché-Frobenius

1. Usar el método de Gauss o de Gauss-Jordan para reducir los siguientes sistemas a una forma escalonada y encontrar, si existen, las soluciones de los mismos:

$$\text{a)}\begin{cases}2x-3y=3\\4x-5y+z=7\\2x-y-3z=5\end{cases}\quad \text{b)}\begin{cases}y-z=0\\x-y=0\\x-z=0\end{cases}\quad \text{c)}\begin{cases}y-z=1\\x-y=1\\x-z=1\end{cases}\quad \text{d)}\begin{cases}x-2y+z+2u=-2\\2x+3y-z-5u=9\\4x-y+z-u=5\\5x-3y+2z+u=3\end{cases}$$

2. Resolver según los valores del parámetro m los siguientes sistemas homogéneos:

$$\text{a)}\begin{cases}2x-my+4z=0\\x+y+7z=0\\mx-y+13z=0\end{cases}\qquad \text{b)}\begin{cases}x+y-z=0\\my-z=0\\x+y+mz=0\end{cases}$$

3. Discutir el siguiente sistema según los valores de m y resolverlo en los casos que sea posible:

$$\begin{cases}mx+y+z=1\\x+my+z=m\\x+y+mz=m^2\end{cases}$$

4. En cada uno de los siguientes sistemas, encontrar (si es posible) condiciones sobre a, b y c de modo que el sistema no tenga soluciones, tenga solo una o posea infinitas soluciones:

$$\text{a)}\begin{cases}3x+y-z=a\\x-y+2z=b\\5x+3y-4z=c\end{cases}\qquad \text{b)}\begin{cases}2x+y-z=a\\2y+3z=b\\x-z=c\end{cases}\qquad \text{c)}\begin{cases}x+ay=0\\y+bz=0\\z+cx=0\end{cases}$$

Resolución de sistemas por LU

5. En cada uno de los siguientes casos, resolver el sistema $A\mathbf{x}=\mathbf{b}$ utilizando la factorización LU de la matriz A:

$$\text{a)}\ A=\begin{pmatrix}-2&2&-4&2\\1&-2&1&3\\-1&1&0&0\\0&1&1&-2\end{pmatrix},\quad \mathbf{b}=\begin{pmatrix}1\\-1\\2\\0\end{pmatrix}.$$

$$\text{b)}\ A=\begin{pmatrix}1&-2&2&2\\-1&2&0&0\\3&-3&1&4\end{pmatrix},\quad \mathbf{b}=\begin{pmatrix}-6\\4\\5\end{pmatrix}.$$

6. En cada uno de los siguientes casos, resolver el sistema $A^2\mathbf{x}=\mathbf{b}$ utilizando la factorización LU de la matriz A:

a) $A = \begin{pmatrix} 1 & 2 & -1 \\ -2 & -5 & 3 \\ -1 & -3 & 0 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} -2 \\ 4 \\ 10 \end{pmatrix}.$

b) $A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & -1 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 2 \\ 3 \\ 2 \\ 1 \end{pmatrix}.$

Problemas

7. Las puntuaciones de 3 jugadores en un torneo se han perdido. La información que se tiene es la suma de las puntuaciones de los jugadores 1 y 2, la suma de los jugadores 2 y 3 y la suma de los jugadores 1 y 3.

a) Probar que de esta información se pueden deducir las puntuaciones de cada jugador.

b) ¿Es cierto lo mismo si tuviéramos 4 jugadores y conociéramos la suma de 1 y 2, de 2 y 3, de 3 y 4 y de 1 y 4?

8. Un Caballero Cadete tiene un presupuesto de 400 euros para gastar durante los próximos 12 fines de semana. El gasto que le supone ir un fin de semana a casa es de 50 euros, mientras que si sale por Zaragoza son 40 euros y si se queda en la AGM solo 10 euros. Calcular todas las posibilidades para los citados fines de semana sabiendo que al final de los doce ha agotado todo su presupuesto.

9. Un curso de inglés de verano cuesta 700 euros para Comandantes, 200 euros para Capitanes y 50 euros para Caballeros y Damas Cadetes. Se sabe que han adjudicado 150 plazas para la AGM de Zaragoza y que se ha pagado un total de 10.000 euros. ¿Cuántos Comandantes, Capitanes y Caballeros y Damas Cadetes de la AGM asistirán al curso? (Ayuda: por motivos estrictamente humanitarios no vale partir a una persona.)

Métodos iterativos

10. a) Partiendo del vector inicial $\mathbf{x}^{(0)} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, aplicar cuatro iteraciones del método de Jacobi, calculando en cada iteración el error de paso con la norma infinito, para encontrar una solución aproximada del siguiente sistema:

$$\begin{cases} 5x + 2y = 1 \\ x - 4y = 0. \end{cases}$$

b) Haciendo uso del Teorema 3.16, aplicar tantas iteraciones como sean necesarias del método de Gauss-Seidel para obtener una aproximación con un error absoluto en norma infinito menor que una tolerancia de 0.001.

c) Sabiendo que la solución exacta del sistema anterior es $(x, y) = ({}^{2}/_{11}, {}^{1}/_{22})$, comprobar que efectivamente se cumple el Teorema 3.16, para la norma infinito, tras la última iteración del apartado anterior.

11. Hallar un valor aproximado de la solución del sistema

$$\begin{cases} 9x - 2y = 5 \\ -2x + 4y - z = 1 \\ -y + z = {}^{-5}\!/_{6} \end{cases}$$

tomando $\mathbf{x}^{(0)} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ como vector inicial y realizando tantas iteraciones como sean necesarias del método de Gauss-Seidel para obtener un error de paso menor que una tolerancia de 10^{-3}. Si la solución exacta del sistema es $x = {}^{2}\!/_{3}$, $y = {}^{1}\!/_{2}$, $z = {}^{-1}\!/_{3}$, obtener el error absoluto, usando la norma ∞, tras la última iteración del apartado anterior.

12. Aplicar cinco iteraciones del método de Jacobi al siguiente sistema:

$$\begin{cases} 3x + 12y - z = -2 \\ 11x - 4y + 3z = -3 \\ -3x - 2y - 12z = -2. \end{cases}$$

Calcular el error de paso desde la segunda iteración. A la vista de los resultados obtenidos ¿qué se puede decir sobre la convergencia del método? ¿Se puede manipular el sistema antes de aplicar las iteraciones del método de Jacobi para asegurar la convergencia?

13. Considerar el siguiente sistema de ecuaciones lineales dependiente del parámetro $a \in \mathbb{R}$:

$$\begin{cases} 2x + y = 5 \\ 3x + ay + z = 5 \\ x + y + 3z = 5. \end{cases}$$

Responder a las siguientes cuestiones:

a) ¿Para qué valores del parámetro a la matriz de coeficientes del sistema es estrictamente diagonal dominante por filas?

b) Para el valor $a = 5$, ¿se puede afirmar algo acerca de la convergencia de los métodos de Jacobi y Gauss-Seidel? ¿Y para el valor $a = -5$?

c) Considerar $a = 3$. ¿Se puede afirmar la convergencia de alguno de los dos métodos?

(Ayuda: Utilizar las dos primeras etapas de la observación de la página 86.)

3.9. SOLUCIONES A LOS EJERCICIOS PROPUESTOS

Algoritmo de Gauss y Teorema de Rouché-Frobenius

1. a) SCD. Solución: $(x, y, z) = (3, 1, 0)$.

b) SCI. Solución: $(x, y, z) = (\lambda, \lambda, \lambda), \quad \lambda \in \mathbb{R}$.

c) SI.

d) SCI. Solución: $(x, y, z, u) = \left(\frac{12+4\mu-\lambda}{7},\, \frac{13+3\lambda+9\mu}{7},\, \lambda,\, \mu\right), \quad \lambda,\, \mu \in \mathbb{R}$.

2. a) $\begin{cases} m = -12/7 : & \text{SCI.} & \text{Solución: } (x,y,z) = (28\lambda, -35\lambda, \lambda), & \lambda \in \mathbb{R}. \\ m = 3 : & \text{SCI.} & \text{Solución: } (x,y,z) = (-5\lambda, -2\lambda, \lambda), & \lambda \in \mathbb{R}. \\ m \neq -12/7,\, m \neq 3 : & \text{SCD.} & \text{Solución trivial: } (x,y,z) = (0,0,0). \end{cases}$

b) $\begin{cases} m = 0 : & \text{SCI.} & \text{Solución: } (x,y,z) = (\lambda, -\lambda, 0), & \lambda \in \mathbb{R}. \\ m = -1 : & \text{SCI.} & \text{Solución: } (x,y,z) = (2\lambda, -\lambda, \lambda), & \lambda \in \mathbb{R}. \\ m \neq 0,\, m \neq -1 : & \text{SCD.} & \text{Solución trivial: } (x,y,z) = (0,0,0). \end{cases}$

3. $\begin{cases} m = -2 : & \text{SI.} \\ m = 1 : & \text{SCI.} & \text{Solución: } (x,y,z) = (1-\lambda-\mu, \lambda, \mu), \quad \lambda,\, \mu \in \mathbb{R}. \\ m \neq -2,\, m \neq 1 : & \text{SCD.} & \text{Solución: } (x,y,z) = \left(\frac{-m-1}{m+2},\, \frac{1}{m+2},\, \frac{(m+1)^2}{m+2}\right). \end{cases}$

4. a) $\begin{cases} -2a+b+c = 0, & \text{Sistema Compatible Indeterminado} \\ -2a+b+c \neq 0, & \text{Sistema Incompatible} \end{cases}$

b) Sistema Compatible Determinado $\forall a,\, b,\, c$.

c) $\begin{cases} 1+abc = 0, & \text{Sistema Compatible Indeterminado} \\ 1+abc \neq 0, & \text{Sistema Compatible Determinado} \end{cases}$

Resolución de sistemas por LU

5. a) $A = \underbrace{\begin{pmatrix} -2 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ -1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 2 \end{pmatrix}}_{L} \underbrace{\begin{pmatrix} 1 & -1 & 2 & -1 \\ 0 & 1 & 1 & -4 \\ 0 & 0 & 1 & -1/2 \\ 0 & 0 & 0 & 1 \end{pmatrix}}_{U}$.

$$A\mathbf{x} = \mathbf{b} \Longleftrightarrow \begin{cases} (1)\, L\mathbf{y} = \mathbf{b} \rightsquigarrow \text{obtener } \mathbf{y} \\ (2)\, U\mathbf{x} = \mathbf{y} \rightsquigarrow \text{obtener } \mathbf{x} \end{cases} \qquad \mathbf{y} = \begin{pmatrix} -1/2 \\ 1/2 \\ 3/4 \\ -1/4 \end{pmatrix},\ \mathbf{x} = \begin{pmatrix} -25/8 \\ -9/8 \\ 5/8 \\ -1/4 \end{pmatrix}.$$

b) $$\underbrace{\begin{pmatrix}1&0&0\\0&0&1\\0&1&0\end{pmatrix}}_{P} A = \underbrace{\begin{pmatrix}1&0&0\\3&3&0\\-1&0&2\end{pmatrix}}_{L}\underbrace{\begin{pmatrix}1&-2&2&2\\0&1&-5/3&-2/3\\0&0&1&1\end{pmatrix}}_{U}.$$

$$A\mathbf{x}=\mathbf{b} \Longleftrightarrow \begin{cases}(1)\, L\mathbf{y}=P\mathbf{b} \rightsquigarrow \text{obtener } \mathbf{y}\\ (2)\, U\mathbf{x}=\mathbf{y} \rightsquigarrow \text{obtener } \mathbf{x}\end{cases} \qquad \mathbf{y}=\begin{pmatrix}-6\\23/3\\-1\end{pmatrix},\ \mathbf{x}=\begin{pmatrix}8-2\lambda\\6-\lambda\\-1-\lambda\\\lambda\end{pmatrix},\ \lambda\in\mathbb{R}.$$

6. a) $$A=\underbrace{\begin{pmatrix}1&0&0\\-2&-1&0\\-1&-1&-2\end{pmatrix}}_{L}\underbrace{\begin{pmatrix}1&2&-1\\0&1&-1\\0&0&1\end{pmatrix}}_{U},\quad A^2\mathbf{x}=\mathbf{b} \Longleftrightarrow \begin{cases}(1)\, L\mathbf{z}=\mathbf{b} \rightsquigarrow \text{obtener } \mathbf{z}\\ (2)\, U\mathbf{y}=\mathbf{z} \rightsquigarrow \text{obtener } \mathbf{y}\\ (3)\, L\mathbf{t}=\mathbf{y} \rightsquigarrow \text{obtener } \mathbf{t}\\ (4)\, U\mathbf{x}=\mathbf{t} \rightsquigarrow \text{obtener } \mathbf{x}.\end{cases}$$

$$\mathbf{z}=\begin{pmatrix}-2\\0\\-4\end{pmatrix},\quad \mathbf{y}=\begin{pmatrix}2\\-4\\-4\end{pmatrix},\quad \mathbf{t}=\begin{pmatrix}2\\0\\1\end{pmatrix},\quad \mathbf{x}=\begin{pmatrix}1\\1\\1\end{pmatrix}.$$

b) $$\underbrace{\begin{pmatrix}1&0&0&0\\0&0&1&0\\0&1&0&0\\0&0&0&1\end{pmatrix}}_{P} A = \underbrace{\begin{pmatrix}1&0&0&0\\0&1&0&0\\1&0&1&0\\0&0&1&1\end{pmatrix}}_{L}\underbrace{\begin{pmatrix}1&1&0&0\\0&1&1&-1\\0&0&1&-1\\0&0&0&1\end{pmatrix}}_{U}$$

$$A^2\mathbf{x}=\mathbf{b} \Longleftrightarrow \begin{cases}(1)\, L\mathbf{z}=P\mathbf{b} \rightsquigarrow \mathbf{z}\\ (2)\, U\mathbf{y}=\mathbf{z} \rightsquigarrow \mathbf{y}\\ (3)\, L\mathbf{t}=P\mathbf{y} \rightsquigarrow \mathbf{t}\\ (4)\, U\mathbf{x}=\mathbf{t} \rightsquigarrow \mathbf{x}.\end{cases} \qquad \mathbf{z}=\begin{pmatrix}2\\2\\1\\0\end{pmatrix}, \mathbf{y}=\begin{pmatrix}1\\1\\1\\0\end{pmatrix}, \mathbf{t}=\begin{pmatrix}1\\1\\0\\0\end{pmatrix}, \mathbf{x}=\begin{pmatrix}0\\1\\0\\0\end{pmatrix}.$$

Problemas

7. Ayuda: Utilizar el teorema de Rouché-Frobenius.

8. Si x representa el número de fines de semana que el Caballero Cadete se va a casa, y los que sale por Zaragoza y z los que se queda en la AGM, las únicas posibilidades son: $(x,y,z)=(1,8,3)$, $(4,4,4)$, $(7,0,5)$.

9. Asisten al curso 140 Caballeros y Damas Cadetes, 8 Capitanes y 2 Comandantes.

Métodos iterativos

10. a) La expresión del método es: $\mathbf{x}^{(m+1)}=\begin{pmatrix}0&-2/5\\1/4&0\end{pmatrix}\mathbf{x}^{(m)}+\begin{pmatrix}1/5\\0\end{pmatrix}$. Aplicándolo cuatro veces se obtiene la siguiente sucesión de vectores:

$$\mathbf{x}^{(1)}=\begin{pmatrix}-3/5\\1/4\end{pmatrix},\quad \mathbf{x}^{(2)}=\begin{pmatrix}1/10\\-3/20\end{pmatrix},\quad \mathbf{x}^{(3)}=\begin{pmatrix}13/50\\1/40\end{pmatrix},\quad \mathbf{x}^{(4)}=\begin{pmatrix}19/100\\13/200\end{pmatrix}.$$

$$\|\mathbf{x}^{(1)}-\mathbf{x}^{(0)}\|_\infty=\frac{7}{4},\ \|\mathbf{x}^{(2)}-\mathbf{x}^{(1)}\|_\infty=\frac{7}{10},\ \|\mathbf{x}^{(3)}-\mathbf{x}^{(2)}\|_\infty=\frac{7}{40},\ \|\mathbf{x}^{(4)}-\mathbf{x}^{(3)}\|_\infty=\frac{7}{100}.$$

b) Las matrices del método son $B = \begin{pmatrix} 0 & -2/5 \\ 0 & -1/10 \end{pmatrix}$ y $\mathbf{c} = \begin{pmatrix} 1/5 \\ 1/20 \end{pmatrix}$. Para $m = 5$ se alcanza la condición

$$\frac{\|B\|_\infty}{1 - \|B\|_\infty} \|\mathbf{x}^{(m)} - \mathbf{x}^{(m-1)}\|_\infty < 0.001.$$

Por tanto el error absoluto en la quinta iteración es menor que la tolerancia dada.

c) Se tiene que $\mathbf{x}^{(4)} = \begin{pmatrix} 913/5000 \\ 913/20000 \end{pmatrix}$, $\mathbf{x}^{(5)} = \begin{pmatrix} 9087/50000 \\ 9087/200000 \end{pmatrix}$,

$$\|\mathbf{x}^{(5)} - \mathbf{x}\|_\infty = 7.81 \cdot 10^{-5}, \quad \frac{\|B\|_\infty}{1 - \|B\|_\infty} = \frac{2}{3}, \quad \|\mathbf{x}^{(5)} - \mathbf{x}^{(4)}\|_\infty = 8.6 \cdot 10^{-4}.$$

De aquí se comprueba que se cumple el Teorema 3.16 para este caso particular.

11. El método de Gauss-Seidel se expresa como:

$$\mathbf{x}^{(m+1)} = \begin{pmatrix} 0 & 2/9 & 0 \\ 0 & 1/9 & 1/4 \\ 0 & 1/9 & 1/4 \end{pmatrix} \mathbf{x}^{(m)} + \begin{pmatrix} 5/9 \\ 19/36 \\ -11/36 \end{pmatrix}.$$

Tras la quinta iteración se obtiene un error de paso $\|\mathbf{x}^{(5)} - \mathbf{x}^{(4)}\|_\infty = 8.357 \cdot 10^{-4} < 10^{-3}$.
La solución aproximada es $\mathbf{x}^{(5)} = \begin{pmatrix} 5041045/7558272 \\ 30261649/60466176 \\ -20126831/60466176 \end{pmatrix} \approx \begin{pmatrix} 0.66696 \\ 0.50047 \\ -0.33286 \end{pmatrix}$.
El error absoluto tras la quinta iteración es $\|\mathbf{x}^{(5)} - \mathbf{x}\|_\infty = 4.72347 \cdot 10^{-4} < 10^{-3}$.

12. El método de Jacobi no es convergente, ya que el error de paso aumenta tras cada iteración:

$$\|\mathbf{x}^{(2)} - \mathbf{x}^{(1)}\|_\infty \approx 2.945, \quad \|\mathbf{x}^{(3)} - \mathbf{x}^{(2)}\|_\infty \approx 8.066,$$
$$\|\mathbf{x}^{(4)} - \mathbf{x}^{(3)}\|_\infty \approx 32.604, \quad \|\mathbf{x}^{(5)} - \mathbf{x}^{(4)}\|_\infty \approx 89.386.$$

Si se intercambian las ecuaciones 1 y 2 del sistema inicial, se puede asegurar la convergencia del método de Jacobi ya que la nueva matriz de coeficientes obtenida es estrictamente diagonal dominante (por filas).

13. a) La matriz A es estrictamente diagonal dominante por filas si y solo si $|a| > 4$.

b) Para los valores $a = \pm 5$, la matriz A es estrictamente diagonal dominante por filas, por lo que se puede afirmar la convergencia de ambos métodos.

c) Para el valor $a = 3$, la matriz A no es estrictamente diagonal dominante. Las matrices B de ambos métodos son:

$$B_J = \begin{pmatrix} 0 & -1/2 & 0 \\ -1 & 0 & -1/3 \\ -1/3 & -1/3 & 0 \end{pmatrix}, \quad B_{GS} = \begin{pmatrix} 0 & -1/2 & 0 \\ 0 & 1/2 & -1/3 \\ 0 & 0 & 1/9 \end{pmatrix}.$$

Se tiene que $\|B_J\|_1 = \|B_J\|_\infty = 4/3 > 1$ y por tanto no se puede afirmar nada sobre la convergencia del método de Jacobi. Por otra parte, puesto que $\|B_{GS}\|_1 = 1$ y $\|B_{GS}\|_\infty = 5/6 < 1$, se puede afirmar que el método de Gauss-Seidel converge.

CAPÍTULO 4

Determinantes

A toda matriz cuadrada real se le puede asociar un número real denominado determinante. La idea de determinante apareció en Occidente antes que el concepto de matriz, en el siglo XVI, ligada al problema de la resolución de sistemas de ecuaciones. En la actualidad no se suelen usar para dicho propósito, ya que el método de eliminación gaussiana propuesto en capítulos anteriores es más adecuado.

El cálculo de un determinante no solo se asocia al estudio y resolución de sistemas de ecuaciones, sino que permite establecer el rango de una matriz, calcular su inversa, resolver problemas de tipo geométrico (fórmulas sencillas para el cálculo de áreas y volúmenes) o calcular los valores propios de una matriz como se muestra en el Capítulo 8. Son además muchas las aplicaciones de los determinantes en Ingeniería, Física, Economía y otras ciencias.

En 1683 el matemático japonés Takakasu ya utilizó la idea de determinante, aunque el mérito se le suele atribuir a Leibniz, que en 1693 realizando un proceso de eliminación para resolver un sistema, obtuvo una expresión que coincide con lo que hoy conocemos como determinante de la matriz de coeficientes del sistema, pero no publicó este trabajo.

McLaurin, Cramer, Bézout, Vandermonde, Laplace, Jacobi, Lagrange y Gauss son otros de los matemáticos que realizaron aportaciones a la teoría de los determinantes.

El nombre de determinante y el significado actual que conocemos se debe a Cauchy, que fue el matemático que realmente propuso la teoría general, dando demostraciones rigurosas y completas. Se encargó de realizar una síntesis de los conocimientos anteriores sobre el tema y publicó en 1812 la fórmula y demostración del determinante de un producto junto con el enunciado y demostración del desarrollo por adjuntos de un determinante.

4.1. DEFINICIÓN DE DETERMINANTE

Existen varias teorías que permiten introducir el concepto de determinante de una matriz cuadrada. Se ha elegido un método sencillo, en el que la definición se presenta de forma inductiva, es decir, usando determinantes de orden $n-1$ para definir el determinante de orden n.

Definición 4.1. La aplicación **determinante**, que a cada matriz cuadrada $A \in \mathrm{Mat}_n(\mathbb{R})$ le hace corresponder un número real

$$\begin{array}{rcl} \det : \mathrm{Mat}_n(\mathbb{R}) & \longrightarrow & \mathbb{R} \\ A & \longmapsto & \det A = |A|, \end{array}$$

se define de modo inductivo mediante:

- Si $n = 1$, $\quad A = (a_{11}) \implies \boxed{|A| = a_{11}}$
- Si $n > 1$, supuesto definido el determinante para matrices cuadradas de orden $n-1$, se define el determinante de $A \in \mathrm{Mat}_n(\mathbb{R})$ como:s

  $$\boxed{|A| = a_{11}A_{11} + a_{21}A_{21} + \cdots + a_{n1}A_{n1}}$$

  donde a_{i1}, $i = 1, \dots, n$, son los elementos de la primera columna de A y A_{i1}, $i = 1, \dots, n$, sus **adjuntos**, (ver siguiente definición).

Definición 4.2. Sea $A = (a_{ij}) \in \mathrm{Mat}_n(\mathbb{R})$. Si se suprime la fila i y la columna j de A, resulta una matriz cuadrada de orden $n-1$ cuyo determinante se llama **menor complementario** del elemento a_{ij} y se representa por M_{ij}. Se denomina **adjunto** del elemento a_{ij}, y se denota mediante A_{ij}, al número

$$A_{ij} = (-1)^{i+j} M_{ij} = (-1)^{i+j} \left| \begin{array}{ccc|ccc} a_{11} & \dots & a_{1,j-1} & a_{1,j+1} & \dots & a_{1n} \\ \vdots & & & & & \vdots \\ a_{i-1,1} & & & & & a_{i-1,n} \\ \hline a_{i+1,1} & & & & & a_{i+1,n} \\ \vdots & & & & & \vdots \\ a_{n1} & \dots & \dots & \dots & \dots & a_{nn} \end{array} \right| .$$

Cuando se usan las definiciones anteriores para evaluar un determinante, se dice que se efectúa un *desarrollo por adjuntos*. Aunque en la Definición 4.1 se utiliza el desarrollo por los elementos de la primera columna, se muestra a lo largo del capítulo que se puede calcular el determinante desarrollando por cualquier fila o columna.

Observación. El signo $(-1)^{i+j}$ que acompaña a los menores forma una imitación de un tablero de ajedrez sobre la matriz, con el signo + sobre la diagonal principal.

$$\begin{pmatrix} + & - & + & \cdots & \cdots \\ - & + & - & \cdots & \cdots \\ + & - & + & \cdots & \cdots \\ \vdots & & & \ddots & \vdots \\ & & & & + \end{pmatrix}$$

✓ **Ejemplo 1.** Dada la matriz cuadrada $A \in \mathrm{Mat}_2(\mathbb{R})$

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

el menor complementario del elemento a_{11} es el determinante $M_{11} = a_{22} = 1$ y el adjunto correspondiente a dicho elemento es

$$A_{11} = (-1)^{1+1} M_{11} = 1.$$

Del mismo modo, se puede calcular por ejemplo, el menor complementario del elemento a_{21} como $M_{21} = a_{12} = 2$ y el adjunto

$$A_{21} = (-1)^{2+1} M_{21} = -2.$$

◇

Observación. Se proponen a continuación fórmulas explícitas para el cálculo del determinante de una matriz de orden 2 o 3, ya que el cálculo del determinante de una matriz de cualquier orden se reduce a uno de estos tipos.

- $n = 2, \quad A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \Longrightarrow \boxed{|A| = a_{11} \underbrace{A_{11}}_{a_{22}} + a_{21} \underbrace{A_{21}}_{-a_{12}} = a_{11}\, a_{22} - a_{12}\, a_{21}}$

- $n = 3, \quad A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \Longrightarrow$

$$\boxed{\begin{gathered} |A| = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21} \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} = \\ a_{11}\, a_{22}\, a_{33} - a_{11}\, a_{23}\, a_{32} - a_{21}\, a_{12}\, a_{33} + a_{21}\, a_{13}\, a_{32} + a_{31}\, a_{12}\, a_{23} - a_{31}\, a_{13}\, a_{22} = \\ a_{11}\, a_{22}\, a_{33} + a_{21}\, a_{32}\, a_{13} + a_{12}\, a_{23}\, a_{31} - a_{13}\, a_{22}\, a_{31} - a_{11}\, a_{23}\, a_{32} - a_{21}\, a_{12}\, a_{33} \end{gathered}}$$

Esta última fórmula desarrollada es conocida como **regla de Sarrus**. Para recordarla se suelen usar los diagramas siguientes:

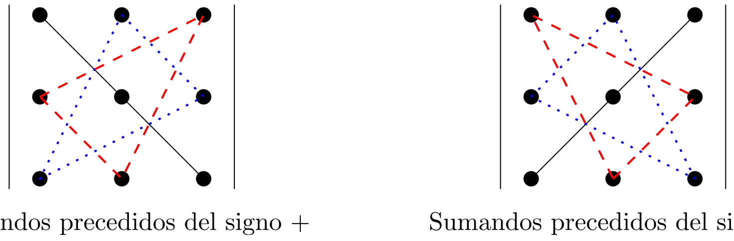

Sumandos precedidos del signo $+$ Sumandos precedidos del signo $-$

Figura 4.1: Regla de Sarrus.

✓ **Ejemplo 2.** Calcular el determinante de las siguientes matrices:

$$A = \begin{pmatrix} 3 & 2 \\ 4 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}.$$

Solución:

- Para calcular el determinante de la matriz $A \in \mathrm{Mat}_2(\mathbb{R})$ basta usar la fórmula dada en la observación anterior:

$$|A| = \begin{vmatrix} 3 & 2 \\ 4 & 1 \end{vmatrix} = 3 \cdot 1 - 2 \cdot 4 = -5.$$

- Para resolver el determinante de la matriz $B \in \mathrm{Mat}_3(\mathbb{R})$ se puede aplicar directamente la regla de Sarrus:

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} = 1 \cdot 5 \cdot 9 + 2 \cdot 6 \cdot 7 + 4 \cdot 8 \cdot 3 - 7 \cdot 5 \cdot 3 - 8 \cdot 6 \cdot 1 - 4 \cdot 2 \cdot 9 = 0.$$

◇

⇝ **Ejercicio.** Escribir una fórmula para calcular el determinante de una matriz de orden 4 en función de determinantes de orden 3.

Observación. Para calcular el determinante de una matriz de orden 5 siguiendo la definición, se deben calcular 5 determinantes de orden 4. Para cada uno de ellos hay que calcular 4 determinantes 3×3, sumando un total de 20 determinantes de orden 3, que a su vez requieren realizar 3 determinantes 2×2 cada uno. Es decir, para calcular el determinante de la matriz inicial 5×5 es necesario realizar 60 determinantes de orden 2.

El cálculo de un determinante utilizando la definición resulta muy largo sobre todo cuando se trata de matrices de orden elevado. El estudio del determinante de algunas matrices sencillas así como las propiedades generales de los determinantes pueden ayudar en la solución del problema anterior.

4.1.1. DETERMINANTE DE ALGUNAS MATRICES ESPECIALES

De la definición de determinante propuesta en la sección anterior, se deducen fácilmente los determinantes de las siguientes matrices:

- El determinante de la matriz identidad de orden n, I_n, es igual a 1,

$$|I_n| = \begin{vmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{vmatrix} = 1.$$

- El determinante de la matriz nula de orden n es cero,

$$\begin{vmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{vmatrix} = 0.$$

- El determinante de una matriz diagonal de orden n es igual al producto de sus elementos diagonales,

$$\begin{vmatrix} a_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & a_{nn} \end{vmatrix} = a_{11} \, \dots \, a_{nn}.$$

- El determinante de una matriz cuadrada triangular (superior o inferior) de orden n es igual al producto de sus elementos diagonales,

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & a_{22} & a_{23} & \dots & a_{2n} \\ 0 & 0 & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & a_{nn} \end{vmatrix} = a_{11}\, a_{22}\, a_{33} \, \dots \, a_{nn}.$$

✓ **Ejemplo 3.** El siguiente determinante se calcula usando la definición y se comprueba que se obtiene el resultado de la propiedad anterior.

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & 5 & 6 & 7 \\ 0 & 0 & 8 & 9 \\ 0 & 0 & 0 & 10 \end{vmatrix} = 1 \begin{vmatrix} 5 & 6 & 7 \\ 0 & 8 & 9 \\ 0 & 0 & 10 \end{vmatrix} = 1 \cdot 5 \begin{vmatrix} 8 & 9 \\ 0 & 10 \end{vmatrix} = 1 \cdot 5 \cdot 8 \cdot 10 = 400.$$

◇

4.2. PROPIEDADES DE LOS DETERMINANTES

Se presentan a continuación una serie de propiedades de los determinantes que son útiles para el cálculo de estos.

Propiedades.

- El determinante de una matriz cuadrada A coincide con el de su traspuesta,

$$|A| = \begin{vmatrix} a_{11} & \cdots & a_{1i} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{i1} & \cdots & a_{ii} & \cdots & a_{in} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{ni} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & \cdots & a_{i1} & \cdots & a_{n1} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{1i} & \cdots & a_{ii} & \cdots & a_{ni} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{1n} & \cdots & a_{in} & \cdots & a_{nn} \end{vmatrix} = |A^t|.$$

✓ **Ejemplo 4.** Dada la matriz $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{R})$, se calculan los determinantes de A y A^t, comprobando que se verifica la propiedad anterior.

$$|A| = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc, \qquad |A^t| = \begin{vmatrix} a & c \\ b & d \end{vmatrix} = ad - cb.$$

◇

- La suma de los determinantes de dos matrices A_1 y $A_2 \in \mathrm{Mat}_n(\mathbb{R})$ que se diferencian solo en los elementos de una fila cualquiera, por ejemplo la i-ésima, es igual al determinante de una matriz cuyos elementos de la fila i-ésima son la suma de los elementos correspondientes de las filas i-ésimas de A_1 y A_2 y el resto de filas se mantienen invariantes.

$$\underbrace{\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a'_{i1} & \cdots & a'_{in} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}}_{|A_1|} + \underbrace{\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a''_{i1} & \cdots & a''_{in} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}}_{|A_2|} = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a'_{i1} + a''_{i1} & \cdots & a'_{in} + a''_{in} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}.$$

✓ **Ejemplo 5.** Se puede comprobar fácilmente que

$$\begin{vmatrix} 1 & 2 & 3 \\ 2 & 0 & 2 \\ 4 & 5 & 6 \end{vmatrix} + \begin{vmatrix} 1 & 2 & 3 \\ 1 & 0 & 1 \\ 4 & 5 & 6 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 2+1 & 0+0 & 2+1 \\ 4 & 5 & 6 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 3 & 0 & 3 \\ 4 & 5 & 6 \end{vmatrix}.$$

Calculando por separado los determinantes se tiene:

$$\begin{vmatrix} 1 & 2 & 3 \\ 2 & 0 & 2 \\ 4 & 5 & 6 \end{vmatrix} = 12; \quad \begin{vmatrix} 1 & 2 & 3 \\ 1 & 0 & 1 \\ 4 & 5 & 6 \end{vmatrix} = 6; \quad \begin{vmatrix} 1 & 2 & 3 \\ 3 & 0 & 3 \\ 4 & 5 & 6 \end{vmatrix} = 18.$$

Observar que se cumple la propiedad.

◇

En general, esta propiedad se suele utilizar en el sentido inverso para calcular el determinante de matrices cuyos elementos son parámetros.

✓ **Ejemplo 6.**

$$\begin{vmatrix} p+x & q+y \\ a+x & b+y \end{vmatrix} = \begin{vmatrix} p & q \\ a+x & b+y \end{vmatrix} + \begin{vmatrix} x & y \\ a+x & b+y \end{vmatrix} =$$

$$\begin{vmatrix} p & q \\ a & b \end{vmatrix} + \begin{vmatrix} p & q \\ x & y \end{vmatrix} + \begin{vmatrix} x & y \\ a & b \end{vmatrix} + \begin{vmatrix} x & y \\ x & y \end{vmatrix} = pb - qa + py - qx + xb - ya.$$

◇

- Si en una matriz $A \in \mathrm{Mat}_n(\mathbb{R})$ se permutan dos filas entre sí (por ejemplo las filas i, j), el determinante de la matriz $B \in \mathrm{Mat}_n(\mathbb{R})$ que resulta es el determinante de A cambiado de signo:

$$|B| = \begin{vmatrix} a_{11} & \cdots & \cdots & a_{1n} \\ \vdots & & & \vdots \\ \boxed{a_{j1} \quad \cdots \quad \cdots \quad a_{jn}} \\ \vdots & & & \vdots \\ \boxed{a_{i1} \quad \cdots \quad \cdots \quad a_{in}} \\ \vdots & & & \vdots \\ a_{n1} & \cdots & \cdots & a_{nn} \end{vmatrix} = - \begin{vmatrix} a_{11} & \cdots & \cdots & a_{1n} \\ \vdots & & & \vdots \\ \boxed{a_{i1} \quad \cdots \quad \cdots \quad a_{in}} \\ \vdots & & & \vdots \\ \boxed{a_{j1} \quad \cdots \quad \cdots \quad a_{jn}} \\ \vdots & & & \vdots \\ a_{n1} & \cdots & \cdots & a_{nn} \end{vmatrix} = -|A|.$$

Recordando la nomenclatura usada en capítulos anteriores para las operaciones elementales, $B = P_{ij}A$ es la matriz que resulta tras aplicar a A una operación elemental de tipo I. Es decir, $|B| = |P_{ij}A| = -|A|$.

✓ **Ejemplo 7.** Dada la matriz $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{R})$, si se permuta la primera fila con la segunda se obtiene la matriz $B \in \mathrm{Mat}_2(\mathbb{R})$,

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \xrightarrow[P_{12}]{} \begin{pmatrix} 3 & 4 \\ 1 & 2 \end{pmatrix} = B.$$

Calculando el determinante de A y B se observa que se verifica la propiedad anterior.

$$|A| = \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = -2; \qquad |B| = \begin{vmatrix} 3 & 4 \\ 1 & 2 \end{vmatrix} = 2 = -|A|.$$

De otro modo $|A| = \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} \overset{P_{12}}{=} - \begin{vmatrix} 3 & 4 \\ 1 & 2 \end{vmatrix} = -2.$

◇

- Si una matriz $A \in \mathrm{Mat}_n(\mathbb{R})$ tiene dos filas iguales, su determinante es cero:

$$|A| = \begin{vmatrix} a_{11} & \cdots & \cdots & a_{1n} \\ \vdots & & & \vdots \\ \boxed{a_{i1} & \cdots & \cdots & a_{in}} \\ \vdots & & & \vdots \\ \boxed{a_{i1} & \cdots & \cdots & a_{in}} \\ \vdots & & & \vdots \\ a_{n1} & \cdots & \cdots & a_{nn} \end{vmatrix} = 0.$$

Esta propiedad es una consecuencia inmediata de la anterior.

✓ **Ejemplo 8.** Dada la matriz $A = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{R})$, su determinante es

$$|A| = \begin{vmatrix} 1 & 2 \\ 1 & 2 \end{vmatrix} = 1 \cdot 2 - 1 \cdot 2 = 0.$$

◇

- Si se multiplica una fila de la matriz $A \in \mathrm{Mat}_n(\mathbb{R})$ por un número real $\lambda \neq 0$, el determinante de la matriz $B \in \mathrm{Mat}_n(\mathbb{R})$ resultante es igual al determinante de A multiplicado por λ:

$$|B| = \begin{vmatrix} a_{11} & \cdots & \cdots & a_{1n} \\ \vdots & & & \vdots \\ \lambda\, a_{i1} & \cdots & \cdots & \lambda\, a_{in} \\ \vdots & & & \vdots \\ a_{n1} & \cdots & \cdots & a_{nn} \end{vmatrix} = \lambda \begin{vmatrix} a_{11} & \cdots & \cdots & a_{1n} \\ \vdots & & & \vdots \\ a_{i1} & \cdots & \cdots & a_{in} \\ \vdots & & & \vdots \\ a_{n1} & \cdots & \cdots & a_{nn} \end{vmatrix} = \lambda |A|.$$

Volviendo a la notación de las operaciones elementales, $B = P_i(\lambda)A$ y $|B| = |P_i(\lambda)A| = \lambda|A|$, o equivalentemente, $|A| = \frac{1}{\lambda}|B|$.

✓ **Ejemplo 9.** Dada la matriz $A = \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{R})$, si se multiplica la primera fila de A por dos se obtiene la matriz $B \in \mathrm{Mat}_2(\mathbb{R})$

$$A = \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix} \xrightarrow[P_1(2)]{} \begin{pmatrix} 2 & 4 \\ -1 & 1 \end{pmatrix} = B.$$

Hallando el determinante de las matrices A y B, se comprueba que se verifica la propiedad anterior.

$$|A| = \begin{vmatrix} 1 & 2 \\ -1 & 1 \end{vmatrix} = 3; \qquad |B| = \begin{vmatrix} 2 & 4 \\ -1 & 1 \end{vmatrix} = 6 = 2 \cdot |A|.$$

De otro modo $|A| = \begin{vmatrix} 1 & 2 \\ -1 & 1 \end{vmatrix} \overset{P_1(2)}{=} \frac{1}{2} \begin{vmatrix} 2 & 4 \\ -1 & 1 \end{vmatrix} = \frac{1}{2} 6 = 3.$

◇

- Si todos los elementos de una fila de la matriz $A \in \mathrm{Mat}_n(\mathbb{R})$ son nulos, entonces su determinante es cero:

$$|A| = \begin{vmatrix} a_{11} & \cdots & \cdots & a_{1n} \\ \vdots & & & \vdots \\ 0 & \cdots & \cdots & 0 \\ \vdots & & & \vdots \\ a_{n1} & \cdots & \cdots & a_{nn} \end{vmatrix} = 0.$$

✓ **Ejemplo 10.** Dada la matriz $A = \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{R})$, su determinante es

$$|A| = \begin{vmatrix} 1 & 2 \\ 0 & 0 \end{vmatrix} = 1 \cdot 0 - 0 \cdot 2 = 0.$$

◇

- Si en la matriz $A \in \mathrm{Mat}_n(\mathbb{R})$, la fila i se sustituye por la fila i más la fila j multiplicada por un número real λ, el determinante de la matriz $B \in \mathrm{Mat}_n(\mathbb{R})$ resultante es igual al determinante de A:

$$|A| = \begin{vmatrix} a_{11} & \cdots & \cdots & a_{1n} \\ \vdots & & & \vdots \\ a_{j1} & \cdots & \cdots & a_{jn} \\ \vdots & & & \vdots \\ a_{i1} & \cdots & \cdots & a_{in} \\ \vdots & & & \vdots \\ a_{n1} & \cdots & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & \cdots & \cdots & a_{1n} \\ \vdots & & & \vdots \\ a_{j1} & \cdots & \cdots & a_{jn} \\ \vdots & & & \vdots \\ a_{i1} + \lambda a_{j1} & \cdots & \cdots & a_{in} + \lambda a_{jn} \\ \vdots & & & \vdots \\ a_{n1} & \cdots & \cdots & a_{nn} \end{vmatrix} = |B|.$$

En términos de operaciones elementales $B = P_{ij}(\lambda)A$, y $|B| = |P_{ij}(\lambda)A| = |A|$.

✓ **Ejemplo 11.** Dada la matriz $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{R})$, si se suma a la segunda fila de A la primera multiplicada por -3, se obtiene la matriz $B \in \mathrm{Mat}_2(\mathbb{R})$

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \xrightarrow[P_{21}(-3)]{} \begin{pmatrix} 1 & 2 \\ 0 & -2 \end{pmatrix} = B.$$

Hallando el determinante de las matrices A y B, se puede comprobar que se verifica la propiedad anterior

$$|A| = \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = 4 - 6 = -2; \qquad |B| = \begin{vmatrix} 1 & 2 \\ 0 & -2 \end{vmatrix} = -2 = |A|.$$

De otro modo $|A| = \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} \overset{P_{21}(-3)}{=} \begin{vmatrix} 1 & 2 \\ 0 & -2 \end{vmatrix} = -2.$ ◇

- Si en la matriz $A \in \mathrm{Mat}_n(\mathbb{R})$ la fila i-ésima es combinación lineal de otras filas, su determinante es nulo:

$$|A| = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ \sum\limits_{k \neq i} \lambda_k a_{k1} & \overset{\text{fila } i}{\cdots} & \sum\limits_{k \neq i} \lambda_k a_{kn} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = 0.$$

Esta propiedad es consecuencia inmediata de las anteriores. De aquí se deduce que si la matriz A tiene dos filas proporcionales, su determinante es nulo.

✓ **Ejemplo 12.** Dada la matriz $A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{R})$, si se calcula su determinante se obtiene

$$|A| = \begin{vmatrix} 1 & 2 \\ 2 & 4 \end{vmatrix} = 1 \cdot 4 - 2 \cdot 2 = 0.$$

Si se observa directamente que sus filas son proporcionales, aplicando la propiedad anterior se obtiene 0 sin necesidad de calcularlo.

◇

Observación. Teniendo en cuenta la primera propiedad, $|A^t| = |A|$, se pueden enunciar las propiedades anteriores de modo análogo por columnas.

4.3. OPERACIONES DE MATRICES Y DETERMINANTES

El comportamiento del determinante con respecto a las operaciones habituales de matrices: suma, producto por un número real y producto de matrices, es el siguiente:

1. **Suma de matrices**: El determinante no se comporta bien con respecto a la suma de matrices, es decir, en general dadas $A, B \in \mathrm{Mat}_n(\mathbb{R})$:

$$\boxed{|A + B| \neq |A| + |B|}$$

2. **Producto por un número real**: Si se considera la matriz $A \in \mathrm{Mat}_n(\mathbb{R})$, y el número real $\lambda \in \mathbb{R}$, entonces

$$\boxed{|\lambda A| = \lambda^n |A|}$$

3. **Producto de matrices**: Dadas las matrices cuadradas $A, B \in \mathrm{Mat}_n(\mathbb{R})$, entonces:

$$\boxed{|AB| = |A|\,|B|}$$

   Es decir, el determinante de un producto de matrices cuadradas es igual al producto de sus determinantes.

Como consecuencia de la propiedad 3 anterior y usando algunos resultados obtenidos en la Sección 4.2 se puede deducir el valor del determinante de cada uno de los tipos de matrices elementales.

Proposición 4.3. Los determinantes de las matrices elementales de tipo I, II y III son, respectivamente:

I. $|P_{ij}| = -1$.

II. $|P_i(\lambda)| = \lambda$.

III. $|P_{ij}(\lambda)| = 1$.

Demostración. Sea A una matriz regular, $|A| \neq 0$, entonces:

I. $|P_{ij}A| = |P_{ij}||A| = -|A| \Rightarrow |P_{ij}| = -1$.

II. $|P_i(\lambda)A| = |P_i(\lambda)||A| = \lambda|A| \Rightarrow |P_i(\lambda)| = \lambda$.

III. $|P_{ij}(\lambda)A| = |P_{ij}(\lambda)||A| = |A| \Rightarrow |P_{ij}(\lambda)| = 1$.

□

✓ **Ejemplo 13.** Dadas las matrices

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

comprobar que se verifica $|AB| = |A|\,|B|$, mientras que $|A+B| \neq |A| + |B|$.

Solución: En efecto, se tiene que:

- $|AB| = \left| \begin{pmatrix} 1 & 2 \\ -2 & 3 \end{pmatrix} \cdot \begin{pmatrix} 0 & -3 \\ 2 & 2 \end{pmatrix} \right| = \begin{vmatrix} 4 & 1 \\ 6 & 12 \end{vmatrix} = 42.$
- $|A|\,|B| = \begin{vmatrix} 1 & 2 \\ -2 & 3 \end{vmatrix} \cdot \begin{vmatrix} 0 & -3 \\ 2 & 2 \end{vmatrix} = 7 \cdot 6 = 42.$

Por tanto $|AB| = |A|\,|B|$. Sin embargo:

- $|A+B| = \left|\begin{pmatrix} 1 & 2 \\ -2 & 3 \end{pmatrix} + \begin{pmatrix} 0 & -3 \\ 2 & 2 \end{pmatrix}\right| = \begin{vmatrix} 1 & -1 \\ 0 & 5 \end{vmatrix} = 5.$
- $|A| + |B| = \begin{vmatrix} 1 & 2 \\ -2 & 3 \end{vmatrix} + \begin{vmatrix} 0 & -3 \\ 2 & 2 \end{vmatrix} = 7 + 6 = 13.$

Se observa que $|A+B| = 5 \neq |A| + |B| = 13$.

◇

Para las matrices por bloques se verifica el siguiente resultado.

Proposición 4.4. Sean tres matrices $A \in \text{Mat}_p(\mathbb{R})$, $B \in \text{Mat}_{p\times q}(\mathbb{R})$ y $C \in \text{Mat}_q(\mathbb{R})$ que dan lugar a una matriz por bloques del tipo

$$M = \left(\begin{array}{c|c} A & B \\ \hline \mathbf{0} & C \end{array}\right) \in \text{Mat}_n(\mathbb{R}),$$

con $n = p + q$. Entonces, $|M| = |A|\,|C|$.

✓ **Ejemplo 14.** Calcular el determinante de la siguiente matriz

$$M = \begin{pmatrix} 1 & 2 & 0 & 1 & 0 \\ -1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 4 & 5 & 6 \end{pmatrix}.$$

Solución: La matriz M se puede dividir en bloques $M = \left(\begin{array}{c|c} A & B \\ \hline \mathbf{0} & C \end{array}\right)$ siendo

$$A = \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 0 & 1 \\ 4 & 5 & 6 \end{pmatrix}.$$

Aplicando la Proposición 4.4 se sabe que

$$|M| = |A|\,|C| = \begin{vmatrix} 1 & 2 \\ -1 & 1 \end{vmatrix} \begin{vmatrix} 1 & 2 & 3 \\ 1 & 0 & 1 \\ 4 & 5 & 6 \end{vmatrix} = 3 \cdot 6 = 18.$$

Los determinantes de A y C se conocen de los Ejemplos 3 y 4.

◇

Notas.

- La Proposición 4.4 es válida si los bloques son de la forma
$$M = \left(\begin{array}{c|c} A & \mathbf{0} \\ \hline B & C \end{array}\right) \in \mathrm{Mat}_n(\mathbb{R}),$$
con $A \in \mathrm{Mat}_p(\mathbb{R})$, $B \in \mathrm{Mat}_{q\times p}(\mathbb{R})$, $C \in \mathrm{Mat}_q(\mathbb{R})$ y $n = p + q$. En este caso también se verifica $|M| = |A|\,|C|$.
- Como consecuencia de la Proposición 4.4 se tiene que si M es diagonal por bloques, su determinante es el producto de los determinantes de los bloques diagonales. Es decir, si
$$M = \left(\begin{array}{c|c} A & \mathbf{0} \\ \hline \mathbf{0} & C \end{array}\right) \in \mathrm{Mat}_n(\mathbb{R}),$$
con $A \in \mathrm{Mat}_p(\mathbb{R})$, $C \in \mathrm{Mat}_q(\mathbb{R})$ y $n = p+q$, el determinante es $|M| = |A|\,|C|$.

4.4. CÁLCULO DE DETERMINANTES

Dada una matriz $A \in \mathrm{Mat}_n(\mathbb{R})$, la Definición 4.1 de determinante dada al inicio del capítulo propone, para calcularlo, desarrollar por los elementos de la primera columna. Esta definición puede generalizarse desarrollando por los elementos de cualquier fila o columna de la matriz A.

Teorema 4.5. Sea $A = (a_{ij}) \in \mathrm{Mat}_n(\mathbb{R})$ una matriz cuadrada de orden n, entonces:

$$|A| = a_{i1}A_{i1} + a_{i2}A_{i2} + \ldots + a_{in}A_{in} = \sum_{k=1}^{n} a_{ik}A_{ik} \qquad \textit{desarrollo por la fila i-ésima}$$

o

$$|A| = a_{1j}A_{1j} + a_{2j}A_{2j} + \ldots + a_{nj}A_{nj} = \sum_{k=1}^{n} a_{kj}A_{kj} \quad \textit{desarrollo por la columna j-ésima.}$$

✓ **Ejemplo 15.** Calcular el determinante:

$$\begin{vmatrix} 1 & 2 & 3 \\ 1 & 0 & 1 \\ 2 & 3 & 4 \end{vmatrix}.$$

Solución: Se desarrolla por la segunda fila por tener el mayor número de ceros:

$$\begin{vmatrix} 1 & 2 & 3 \\ 1 & 0 & 1 \\ 2 & 3 & 4 \end{vmatrix} = 1\,(-1)^3 \begin{vmatrix} 2 & 3 \\ 3 & 4 \end{vmatrix} + 1\,(-1)^5 \begin{vmatrix} 1 & 2 \\ 2 & 3 \end{vmatrix} = 1 + 1 = 2.$$

◇

El cálculo del determinante de una matriz de orden grande puede ser bastante engorroso si se sigue la definición. El uso del Teorema 4.5 y de las propiedades de la Sección 4.2 permite simplificar los cálculos.

Se proponen dos formas de abordar el cálculo de determinantes, aunque estas no son las únicas maneras de obtenerlo.

- Hacer operaciones elementales a la matriz hasta obtener una matriz triangular, cuyo determinante es inmediato multiplicando los elementos de la diagonal principal. Hay que recordar como afectan las operaciones elementales al determinante.

✓ **Ejemplo 16.** Calcular el determinante de la matriz $A = \begin{pmatrix} 3 & 3 & -2 \\ 1 & 2 & 0 \\ -1 & 4 & 1 \end{pmatrix}$.

Solución:

$$|A| = \begin{vmatrix} 3 & 3 & -2 \\ 1 & 2 & 0 \\ -1 & 4 & 1 \end{vmatrix} \overset{P_{12}}{=} - \begin{vmatrix} 1 & 2 & 0 \\ 3 & 3 & -2 \\ -1 & 4 & 1 \end{vmatrix} \underset{P_{31}(1)}{\overset{P_{21}(-3)}{=}} - \begin{vmatrix} 1 & 2 & 0 \\ 0 & -3 & -2 \\ 0 & 6 & 1 \end{vmatrix} \overset{P_2(-\frac{1}{3})}{=}$$

$$-(-3) \begin{vmatrix} 1 & 2 & 0 \\ 0 & 1 & 2/3 \\ 0 & 6 & 1 \end{vmatrix} \overset{P_{32}(-6)}{=} 3 \begin{vmatrix} 1 & 2 & 0 \\ 0 & 1 & 2/3 \\ 0 & 0 & -3 \end{vmatrix} = 3 \cdot [1 \cdot 1 \cdot (-3)] = -9.$$

◇

- Combinar las operaciones elementales con las propiedades de la Sección 4.2, desarrollando el determinante por la fila o columna con el mayor número de ceros.

✓ **Ejemplo 17.** Calcular el determinante de la matriz $A = \begin{pmatrix} 3 & 3 & -2 \\ 1 & 2 & 0 \\ -1 & 4 & 1 \end{pmatrix}$.

Solución: En este caso se elige la tercera columna que ya tiene un cero. Se hace una operación elemental para convertir el elemento a_{13} en cero y se desarrolla por la tercera columna que tiene un único elemento no nulo. Se obtiene así un determinante de orden 2 que se resuelve directamente mediante la definición.

$$|A| = \begin{vmatrix} 3 & 3 & -2 \\ 1 & 2 & 0 \\ -1 & 4 & 1 \end{vmatrix} \overset{P_{13}(2)}{=} \begin{vmatrix} 1 & 11 & 0 \\ 1 & 2 & 0 \\ -1 & 4 & 1 \end{vmatrix} = (-1)^{3+3} \begin{vmatrix} 1 & 11 \\ 1 & 2 \end{vmatrix} = 1 \cdot 2 - 1 \cdot 11 = -9.$$

◇

4.5. APLICACIONES

Los determinantes son de utilidad en la resolución de algunos problemas que se han planteado en capítulos anteriores. Por ejemplo, con ellos se puede dar una condición necesaria y suficiente para la regularidad de una matriz, calcular la inversa o resolver sistemas de ecuaciones cuya matriz de coeficientes sea invertible.

4.5.1. RANGO Y EXISTENCIA DE LA INVERSA DE UNA MATRIZ

En los capítulos anteriores se ha definido el rango de una matriz A como el número de filas no nulas que tiene una cualquiera de sus formas escalonadas. Para matrices cuadradas de orden n, recordando algunos de los apartados del Teorema 2.13, el rango de una matriz cuadrada se usaba para caracterizar la regularidad de la matriz. Ahora se puede hacer también con el determinante:

$$\boxed{\forall A \in \operatorname{Mat}_n(\mathbb{R}), \quad \operatorname{rg} A = n \Longleftrightarrow |A| \neq 0 \Longleftrightarrow \exists A^{-1} \in \operatorname{Mat}_n(\mathbb{R})}$$

✓ **Ejemplo 18.** Se considera la matriz A del Ejemplo 17, ¿tiene A rango máximo? ¿Es A regular?

Solución: Por los cálculos realizado en la solución del Ejemplo 17 se sabe que $|A| = -9 \neq 0$, por tanto $\operatorname{rg} A = 3$ y A es regular, es decir, $\exists A^{-1}$.

◇

Observación. Si $|A| \neq 0$ entonces existe A^{-1} y se verifica $AA^{-1} = I_n$. De aquí se deduce, usando la propiedad multiplicativa del determinante, que $|A||A^{-1}| = 1$. Así,

$$\boxed{|A^{-1}| = \frac{1}{|A|}}$$

4.5.2. CÁLCULO DE LA MATRIZ INVERSA

El determinante permite además dar una forma explícita para calcular la inversa de una matriz, en el caso de que sea regular. No se aconseja el uso de este método para el cálculo de la inversa, sino el método de eliminación gaussiana propuesto en capítulos anteriores.

Teorema 4.6. Sea $A \in \operatorname{Mat}_n(\mathbb{R})$ una matriz cuadrada regular, entonces:

$$A^{-1} = \frac{1}{|A|}(\operatorname{adj} A)^t$$

donde $\operatorname{adj} A$ es la matriz adjunta de A, matriz formada por los adjuntos de los elementos de A (Definición 4.2),

$$\operatorname{adj} A = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \dots & A_{nn} \end{pmatrix}.$$

✓ **Ejemplo 19.** Calcular la inversa de la matriz A, aplicando el método de los adjuntos:

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

Solución: Se procede de la siguiente forma:

- Calcular el determinante de A:

  $$|A| = \begin{vmatrix} 1 & 2 & 3 \\ -1 & 1 & -2 \\ 1 & 0 & 2 \end{vmatrix} = -1.$$

  Como $|A| \neq 0$, entonces $\exists A^{-1}$.

- Calcular la matriz $\operatorname{adj} A$:

  $$A_{11} = (-1)^{1+1} \begin{vmatrix} 1 & -2 \\ 0 & 2 \end{vmatrix} = 2; \quad A_{12} = (-1)^{1+2} \begin{vmatrix} -1 & -2 \\ 1 & 2 \end{vmatrix} = 0; \quad \dots$$

  En este caso:

  $$\operatorname{adj} A = \begin{pmatrix} 2 & 0 & -1 \\ -4 & -1 & 2 \\ -7 & -1 & 3 \end{pmatrix}.$$

- Por último, hallar la traspuesta de la adjunta y la multiplicarla por $\frac{1}{|A|}$, para obtener la matriz inversa:

  $$A^{-1} = \frac{1}{|A|} (\operatorname{adj} A)^t = \begin{pmatrix} -2 & 4 & 7 \\ 0 & 1 & 1 \\ 1 & -2 & -3 \end{pmatrix}.$$

◇

⇝ **Ejercicio.** Considerar

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Suponiendo que A es regular, calcular A^{-1} usando el método explicado en esta sección. Comprobar que el resultado obtenido verifica $AA^{-1} = A^{-1}A = I_2$.

4.5.3. REGLA DE CRAMER

El determinante proporciona una regla, conocida como **regla de Cramer**, para la resolución de sistemas de n ecuaciones con n incógnitas cuya matriz de coeficientes sea regular. Se presenta este método aunque no se aconseja su uso, ya que el número de operaciones que hay que realizar cuando $n \geq 3$ es muy grande. El método de Gauss se plantea como método óptimo de resolución de sistemas.

Teorema 4.7. Sea $A \in \text{Mat}_n(\mathbb{R})$ una matriz cuadrada regular de orden n. El sistema de ecuaciones lineales $A\mathbf{x} = \mathbf{b}$ es compatible determinado $\forall \mathbf{b}$, y las componentes x_i, $i = 1, \ldots, n$, del vector solución $\mathbf{x}$ se calculan mediante la fórmula

$$x_i = \frac{|C_i|}{|A|}, \quad i = 1, 2, \ldots, n,$$

donde C_i es la matriz que resulta de sustituir en la matriz A la columna i por el vector $\mathbf{b}$,

$$C_i = \begin{pmatrix} a_{11} & \cdots & a_{1,i-1} & b_1 & a_{1,i+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i1} & \cdots & a_{i,i-1} & b_i & a_{i,i+1} & \cdots & a_{in} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,i-1} & b_n & a_{n,i+1} & \cdots & a_{nn} \end{pmatrix}.$$

4.5.4. APLICACIONES GEOMÉTRICAS

Además de las aplicaciones anteriores, los determinantes tienen muchas aplicaciones en geometría. Se presentan a continuación algunas de ellas.

Aplicaciones geométricas en $\mathbb{R}^2$

- Cálculo del área de un triángulo de vértices (x_1, y_1), (x_2, y_2), (x_3, y_3):

  $$A = \frac{1}{2} \begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}.$$

  El área del triángulo es el valor absoluto de la cantidad A calculada.

- Los puntos (x_1, y_1), (x_2, y_2), (x_3, y_3) están en la misma recta o son colineales si y solo si

  $$\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} = 0.$$

- Cálculo de la ecuación de la recta que pasa por dos puntos distintos (x_1, y_1), (x_2, y_2):
$$\begin{vmatrix} x & y & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{vmatrix} = 0.$$

Aplicaciones geométricas en $\mathbb{R}^3$

- Cálculo del volumen de un tetraedro cuyos vértices son los puntos (x_1, y_1, z_1), (x_2, y_2, z_2), (x_3, y_3, z_3), (x_4, y_4, z_4):
$$V = \frac{1}{6}\begin{vmatrix} x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \\ x_4 & y_4 & z_4 & 1 \end{vmatrix}.$$
El volumen del tetraedro es el valor absoluto de la cantidad V calculada.

- Los puntos (x_1, y_1, z_1), (x_2, y_2, z_2), (x_3, y_3, z_3), (x_4, y_4, z_4) son coplanarios si y solo si
$$\begin{vmatrix} x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \\ x_4 & y_4 & z_4 & 1 \end{vmatrix} = 0.$$

- Cálculo de la ecuación del plano que pasa por los puntos (x_1, y_1, z_1), (x_2, y_2, z_2), (x_3, y_3, z_3):
$$\begin{vmatrix} x & y & z & 1 \\ x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \end{vmatrix} = 0.$$

- Cálculo del producto vectorial de los vectores $\mathbf{u} = (x_1, y_1, z_1)$, $\mathbf{v} = (x_2, y_2, z_2)$:
$$\mathbf{u} \times \mathbf{v} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \end{vmatrix} = \begin{vmatrix} y_1 & z_1 \\ y_2 & z_2 \end{vmatrix}\vec{i} - \begin{vmatrix} x_1 & z_1 \\ x_2 & z_2 \end{vmatrix}\vec{j} + \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix}\vec{k},$$
es decir, el vector $\mathbf{u} \times \mathbf{v}$ se expresa como:
$$\mathbf{u} \times \mathbf{v} = \left(\begin{vmatrix} y_1 & z_1 \\ y_2 & z_2 \end{vmatrix}, - \begin{vmatrix} x_1 & z_1 \\ x_2 & z_2 \end{vmatrix}, \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} \right)$$
y es ortogonal a los vectores $\mathbf{u}$ y $\mathbf{v}$.

4.6. COMANDOS DE *wxMaxima*

Tras introducir una matriz cuadrada A, se pueden realizar las siguientes operaciones:

- Calcular el determinante de la matriz A:

  ```
  -->     determinant(A);
  ```

  Comando al que también se puede acceder desde "`\Álgebra\Determinante`".

- Calcular la traspuesta de la matriz adjunta de la matriz A dada:

  ```
  -->     adjoint(A);
  ```

  Si se quiere obtener el adjunto del elemento a_{ij} de la matriz A, es necesario extraer el elemento que ocupa la posición (j, i) en la matriz que proporciona el comando "`adjoint`".

- Calcular la matriz que resulta de eliminar la fila i y la columna j de la matriz A:

  ```
  -->     minor(A,i,j);
  ```

  Si se quiere obtener el menor complementario M_{ij} del elemento a_{ij} de la matriz A, es necesario calcular el determinante de la matriz que proporciona el comando "`minor`".

- Construir una matriz de Vandermonde de tamaño n, de parámetros $x_1, \ldots, x_n$, cuya i-ésima fila es $[1, x_i, x_i^2, ..., x_i^{(n-1)}]$:

  ```
  -->     vandermonde_matrix([x1,...,xn]);
  ```

- Pueden usarse además otros comandos definidos en capítulos anteriores como "`submatrix`", "`invert`".

4.7. EJERCICIOS PROPUESTOS

1. Sean $A, B \in \mathrm{Mat}_2(\mathbb{R})$, demostrar la propiedad del producto $|AB| = |A||B|$.

2. Expresar la condición que deben cumplir las matrices $A, B \in \mathrm{Mat}_2(\mathbb{R})$ para que se verifique $|A + B| = |A| + |B|$.

3. Calcular el determinante de la matriz

$$A = \begin{pmatrix} 2 & 3 & 1 \\ -2 & 5 & 3 \\ 4 & 1 & 2 \end{pmatrix}$$

de tres formas distintas:

- Desarrollando por una fila o una columna.
- Usando la regla de Sarrus.
- Haciendo operaciones elementales.

¿Qué método ha resultado más sencillo?

4. Calcular los siguientes determinantes:

$$\text{a)} \begin{vmatrix} -1 & 3 & 1 \\ 2 & 5 & 3 \\ 1 & -2 & 1 \end{vmatrix} \quad \text{b)} \begin{vmatrix} -1 & -1 & 1 & 0 \\ 2 & 1 & 1 & 3 \\ 0 & 1 & 1 & 2 \\ 1 & 3 & -1 & 2 \end{vmatrix} \quad \text{c)} \begin{vmatrix} 2 & 3 & 1 & 1 \\ 0 & 2 & -1 & 3 \\ 0 & 5 & 1 & 1 \\ 1 & 1 & 2 & 5 \end{vmatrix} \quad \text{d)} \begin{vmatrix} 1 & 2 & 1 & 2 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 2 \\ 1 & 2 & 2 & 1 & 1 \end{vmatrix}.$$

5. Demostrar que:

$$\begin{vmatrix} p+x & q+y & r+z \\ a+x & b+y & c+z \\ a+p & b+q & c+r \end{vmatrix} = 2 \begin{vmatrix} a & b & c \\ p & q & r \\ x & y & z \end{vmatrix}.$$

6. Probar que:

$$\text{a)} \begin{vmatrix} 1 & x & x^2 & x^3 \\ p & 1 & x & x^2 \\ 0 & q & 1 & x \\ 0 & 0 & r & 1 \end{vmatrix} = (1 - px)(1 - qx)(1 - rx).$$

$$\text{b)} \begin{vmatrix} x & -1 & 0 & 0 \\ 0 & x & -1 & 0 \\ 0 & 0 & x & -1 \\ p & q & r & x+s \end{vmatrix} = p + qx + rx^2 + sx^3 + x^4.$$

7. Probar que:

$$\begin{vmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{vmatrix} = (b-a)(c-a)(c-b).$$

Este determinante se denomina determinante de **Vandermonde**. Generalizar la expresión anterior a una matriz de tamaño $n \times n$.

8. Sabiendo que $\begin{vmatrix} a & b & c \\ p & q & r \\ x & y & z \end{vmatrix} = 2$, calcular sin desarrollar $\begin{vmatrix} a & x & p \\ b & y & q \\ c & z & r \end{vmatrix}$.

9. Calcular los siguientes determinantes:

$$\text{a)} \begin{vmatrix} x & y & 1 \\ y & 1 & x \\ 1 & x & y \end{vmatrix}, \qquad \text{b)} \begin{vmatrix} x^2 & y^2 & z^2 \\ (y+z)^2 & (x+z)^2 & (x+y)^2 \\ yz & xz & xy \end{vmatrix}.$$

10. Resolver las siguientes ecuaciones:

$$\text{a)} \begin{vmatrix} 1 & 1 & 1 \\ 1 & x & -1 \\ 1 & x^2 & 1 \end{vmatrix} = 0, \qquad \text{b)} \begin{vmatrix} 2x-1 & 3x & x-2 \\ 2x+1 & x & 2x+1 \\ 4x-1 & 3x & 3x-2 \end{vmatrix} = 0.$$

11. Dada la matriz

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

resolver la ecuación $|A - \lambda I_3| = 0$.

4.8. SOLUCIONES A LOS EJERCICIOS PROPUESTOS

1. Considerar $A = (a_{ij})$, $B = (b_{ij})$ con $i, j = 1, 2$, y realizar las operaciones.

2. Sean $A = (a_{ij})$, $B = (b_{ij})$ con $i, j = 1, 2$. La condición que deben cumplir es:

$$a_{11}b_{22} + b_{11}a_{22} - a_{21}b_{12} - b_{21}a_{12} = 0.$$

3. $|A| = 40$.

4. a) $|A| = -17$, b) $|A| = 2$, c) $|A| = 106$, d)$|A| = 2$.

5. Se puede calcular el determinante utilizando las propiedades adecuadas o bien realizando operaciones elementales y aplicando propiedades.

6. Realizar operaciones elementales y aplicar propiedades de los determinantes.

7. Para una matriz $n \times n$:

$$\begin{vmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} \end{vmatrix} = \prod_{1 \leq j < i \leq n} (a_i - a_j).$$

8. $\begin{vmatrix} a & x & p \\ b & y & q \\ c & z & r \end{vmatrix} = -2.$

9. a) $\begin{vmatrix} x & y & 1 \\ y & 1 & x \\ 1 & x & y \end{vmatrix} = -(1 + x + y)\left[x^2 + y^2 - xy - x - y + 1\right].$

b) $\begin{vmatrix} x^2 & y^2 & z^2 \\ (y+z)^2 & (x+z)^2 & (x+y)^2 \\ yz & xz & xy \end{vmatrix} =$

$$(x^2 + y^2 + z^2)(x + y + z)(x - y)(x - z)(y - z).$$

10. a) $\begin{vmatrix} 1 & 1 & 1 \\ 1 & x & -1 \\ 1 & x^2 & 1 \end{vmatrix} = 0 \Longleftrightarrow x = \pm 1.$

b) $\begin{vmatrix} 2x-1 & 3x & x-2 \\ 2x+1 & x & 2x+1 \\ 4x-1 & 3x & 3x-2 \end{vmatrix} = 0 \Longleftrightarrow x = 0,\ x = -1.$

11. $\lambda = -1$, $\lambda = 1$, $\lambda = 3$.

CAPÍTULO 5

Espacios Vectoriales Reales

El concepto de vector como segmento orientado se remonta a Aristóteles, aunque tiene su máxima aplicación en Física a partir del siglo XVII con el desarrollo de la Geometría Analítica, cuyas bases fueron sentadas por Descartes y Fermat.

Los avances producidos además con el desarrollo de la teoría de matrices y sistemas de ecuaciones permitieron el paso del plano y el espacio a espacios de dimensión mayor. Destacan los aportes de Gauss, Hamilton (que fue quien inventó el nombre de vector), Cayley y Grassman que manejan las propiedades algebraicas del espacio n-dimensional.

La formalización de la teoría de vectores en el plano y en el espacio se realizó en el siglo XIX. En 1888 Peano introdujo la definición axiomática del espacio vectorial.

Un desarrollo de estos espacios vectoriales se dio ya en pleno siglo XX con la construcción de los espacios de funciones por Lebesgue, formalizados posteriormente por Banach y Hilbert.

Los espacios vectoriales son aplicables al estudio de otras partes de las Matemáticas, así como a la Física o la Ingeniería. Estos espacios son necesarios para la construcción de series de Fourier, la resolución de ecuaciones diferenciales y en derivadas parciales y para tratar algunos problemas geométricos.

Se estudian en este capítulo distintos conjuntos dotados de ciertas operaciones que verifican las mismas propiedades. Así, todos tienen una misma estructura que se denomina espacio vectorial, y a sus elementos se les llama vectores.

El capítulo se centra en el estudio de espacios vectoriales reales de dimensión finita, se comienza por el conocido $\mathbb{R}^2$, se amplia el estudio a $\mathbb{R}^n$ y de ahí se generaliza a un espacio real cualquiera V.

5.1. EL ESPACIO VECTORIAL $\mathbb{R}^n$

Cuando se habla de un vector lo primero que viene a la mente es un segmento orientado con punto de aplicación, módulo, dirección y sentido en el plano o en el espacio tal y como se usa en Física y que tanto se utiliza para representar magnitudes llamadas vectoriales, como la fuerza o la velocidad. En esta sección se pretende extender y generalizar el concepto de vector al caso de $\mathbb{R}^n$, siendo n cualquier valor natural.

Se comienza recordando las propiedades de los vectores del plano $\mathbb{R}^2$, ya que su representación geométrica ayuda a comprender el concepto general de vector.

5.1.1. VECTORES EN EL PLANO $\mathbb{R}^2$

Se considera un vector geométrico en el plano $\mathbb{R}^2$, representado por un segmento orientado que comienza en un punto inicial llamado *origen* o *punto de aplicación*, y tiene una longitud denominada *módulo* que puede representar una cierta magnitud. El segmento se encuentra sobre una recta que indica su *dirección* y dentro de la cual tiene un *sentido*.

En general se suele trabajar con vectores determinados por su longitud y dirección independientes de su origen, es decir, *vectores libres*.

Cuando se considera un sistema de referencia cartesiano y como origen de todos los vectores el punto $(0, 0)$, si el punto donde termina el vector es el (x_1, x_2), este par ordenado se usa para caracterizar el vector, denotado para evitar confusiones por $\mathbf{x} = (x_1, x_2)$. Los valores x_1, x_2 se denominan *componentes del vector*.

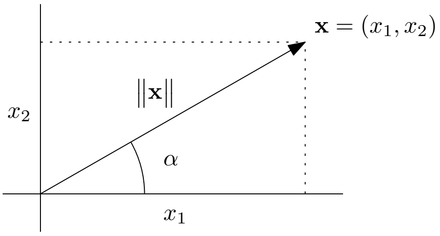


Figura 5.1: Vectores en $\mathbb{R}^2$

Observación. Los vectores se representan mediante letras minúsculas en negrita, $\mathbf{x}$, o con una flecha encima, $\vec{x}$.

Definición 5.1. Dado el vector $\mathbf{x} \in \mathbb{R}^2$ con $\mathbf{x} = (x_1, x_2)$, se define la **norma** o longitud de $\mathbf{x}$ como

$$\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2}.$$

Así una vez definida la norma, si se observa la Figura 5.1 y se considera el ángulo α que forma la dirección positiva del eje OX con el vector, se verifica:

$$x_1 = \|\mathbf{x}\| \cos\alpha, \qquad x_2 = \|\mathbf{x}\| \operatorname{sen}\alpha.$$

Definición 5.2. Dos vectores $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$, con $\mathbf{x} = (x_1, x_2)$ e $\mathbf{y} = (y_1, y_2)$ son **iguales** si y solo si $x_1 = y_1$, $x_2 = y_2$.

Se presentan a continuación dos de las operaciones que se pueden definir en este conjunto de vectores así como las propiedades que estas verifican.

SUMA DE VECTORES DE $\mathbb{R}^2$

Definición 5.3. La **suma vectorial** en $\mathbb{R}^2$ es una operación interna

$$\begin{array}{rcl} + : \mathbb{R}^2 \times \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ (\mathbf{x}, \mathbf{y}) & \mapsto & \mathbf{x} + \mathbf{y} \end{array}$$

que a cada par de vectores $\mathbf{x}, \mathbf{y}$ le asocia otro vector dado por la suma de sus componentes:

$$\mathbf{x} + \mathbf{y} = (x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 + y_2).$$

Geométricamente, la suma de dos vectores $\mathbf{x}, \mathbf{y}$ con origen común y distinta dirección es la diagonal del paralelogramo formado con lados $\mathbf{x}$ e $\mathbf{y}$.

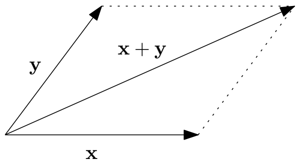


Figura 5.2: Regla del paralelogramo

La suma de vectores geométricos de $\mathbb{R}^2$ verifica las siguientes propiedades:

Propiedades.

- Asociativa: $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$, $\forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^2$.
- Conmutativa: $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$, $\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^2$.
- Elemento neutro: existe el vector nulo, denotado por $\mathbf{0} = (0, 0) \in \mathbb{R}^2$ tal que $\mathbf{x} + \mathbf{0} = \mathbf{0} + \mathbf{x} = \mathbf{x}$, $\forall \mathbf{x} \in \mathbb{R}^2$.

- Elemento opuesto: $\forall \mathbf{x} = (x_1, x_2) \in \mathbb{R}^2$ existe $-\mathbf{x} = (-x_1, -x_2) \in \mathbb{R}^2$ tal que $\mathbf{x}+(-\mathbf{x}) = (-\mathbf{x})+\mathbf{x} = \mathbf{0}$. (Observar que $-\mathbf{x}$ tiene el mismo módulo y dirección que $\mathbf{x}$ pero sentido opuesto).

Las propiedades anteriores se pueden demostrar geométricamente con ayuda de las representaciones gráficas o bien usando las componentes de los vectores.

Debido a las cuatro propiedades anteriores, el conjunto $\mathbb{R}^2$ dotado de la operación suma que se acaba de definir, $(\mathbb{R}^2, +)$, se dice que tiene estructura de **grupo abeliano**.

A lo largo del capítulo se sigue profundizando en algunas estructuras algebraicas importantes.

PRODUCTO DE UN VECTOR DE $\mathbb{R}^2$ POR UN ESCALAR

Definición 5.4. El **producto por un escalar** es una operación externa

$$\begin{array}{rcl} \cdot : \mathbb{R} \times \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ (\lambda, \mathbf{x}) & \mapsto & \lambda\mathbf{x} \end{array}$$

que a cada par formado por un escalar $\lambda \in \mathbb{R}$ y un vector $\mathbf{x} \in \mathbb{R}^2$ le asocia otro vector dado por:

$$\lambda\mathbf{x} = \lambda(x_1, x_2) = (\lambda x_1, \lambda x_2).$$

Geométricamente, el producto por un escalar representa la dilatación o contracción del vector $\mathbf{x}$ (dependiendo de si $|\lambda| > 1$ o $|\lambda| < 1$). El sentido del vector se conserva o no según sea el signo de λ ($\lambda > 0$ conserva el sentido).

El conjunto $\mathbb{R}^2$ con las dos operaciones definidas verifica las siguientes propiedades:

Propiedades.

- $\lambda(\mathbf{x} + \mathbf{y}) = \lambda\mathbf{x} + \lambda\mathbf{y}$, $\forall\lambda \in \mathbb{R}$, $\forall\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$.
- $(\lambda + \mu)\mathbf{x} = \lambda\mathbf{x} + \mu\mathbf{x}$, $\forall\lambda, \mu \in \mathbb{R}$, $\forall\mathbf{x} \in \mathbb{R}^2$.
- $\lambda(\mu\mathbf{x}) = (\lambda\mu)\mathbf{x}$, $\forall\lambda, \mu \in \mathbb{R}$, $\forall\mathbf{x} \in \mathbb{R}^2$.
- $1\mathbf{x} = \mathbf{x}$, $\forall\mathbf{x} \in \mathbb{R}^2$.

El conjunto $\mathbb{R}^2$ dotado de las dos operaciones anteriores, una interna y otra externa, $(\mathbb{R}^2, +, \cdot)$ verificando las 8 propiedades de la suma y el producto por un escalar tiene estructura de *espacio vectorial*, como se ve en la Sección 5.2.

5.1.2. VECTORES EN $\mathbb{R}^n$

Una vez estudiado el conjunto de vectores geométricos del plano $\mathbb{R}^2$ con sus operaciones y propiedades, se generaliza el estudio al conjunto $\mathbb{R}^n$, que se define del siguiente modo:

Definición 5.5. Se denota mediante $\mathbb{R}^n$ al conjunto de todas las n-tuplas de números reales, es decir:

$$\boxed{\mathbb{R}^n = \{(x_1, \ldots, x_n) \mid x_i \in \mathbb{R},\, i = 1, \ldots, n\}}$$

Los elementos del conjunto se denominan **vectores**. Dado un vector del espacio $\mathbb{R}^n$, $\mathbf{x} = (x_1, \ldots, x_n)$, el número real x_i se denomina **i-ésima componente** de $\mathbf{x}$.

Los elementos del conjunto $\mathbb{R}^n$ se escriben con letras minúsculas en negrita, por ejemplo, $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$, $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$, o bien con una flecha, $\vec{x}$, y sus componentes mediante la misma letra sin negrita y con el subíndice que indica la posición que ocupa. En ocasiones para valores pequeños de n se puede usar una notación sin subíndices, por ejemplo, en $\mathbb{R}^3$ un vector genérico se suele denotar (x, y, z).

✓ **Ejemplo 1.** El elemento $\mathbf{x} = (-1, \sqrt{2}, 5/8)$ es un vector de $\mathbb{R}^3$; el número real $\sqrt{2}$ es la segunda componente de $\mathbf{x}$, $x_2 = \sqrt{2}$.

◇

Se presentan algunos ejemplos concretos de espacios $\mathbb{R}^n$ ya conocidos que se pueden identificar y representar gráficamente de una manera sencilla:

- $\mathbb{R}^1 = \mathbb{R}$ es el conjunto de los números reales.
- $\mathbb{R}^2$ es el conjunto de pares ordenados de números reales, denominado plano real, revisado en el apartado anterior.
- $\mathbb{R}^3$ es el conjunto de ternas ordenadas de números reales, denominado espacio tridimensional.

Geométricamente, el espacio $\mathbb{R}^n$ se puede pensar como el construido a partir de n ejes perpendiculares entre sí. De este modo la coordenada i-ésima indica el desplazamiento en el eje i-ésimo.

Definición 5.6. Dos vectores $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, son **iguales** si y solo si sus componentes correspondientes son iguales.

$$\mathbf{x} = \mathbf{y} \iff x_i = y_i,\ \forall i = 1, \ldots, n.$$

Observación. Los vectores de $\mathbb{R}^n$ se pueden representar de diferentes formas, que se utilizan a lo largo de texto:

- Como una n-tupla.

$$\mathbf{x} = (x_1, \dots, x_n).$$

- Como una matriz columna de orden $n \times 1$, o vector columna, con coeficientes reales.

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \operatorname{Mat}_{n\times 1}(\mathbb{R}).$$

- Como una matriz fila de orden $1 \times n$, o vector fila, con coeficientes reales.

$$\mathbf{x}^t = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} \in \operatorname{Mat}_{1\times n}(\mathbb{R}).$$

Al igual que en $\mathbb{R}^2$ se pueden definir en $\mathbb{R}^n$ las operaciones suma y producto por un escalar. Estas operaciones se denominan *operaciones usuales* o *habituales.*

Definición 5.7. Sean $\mathbf{x} = (x_1, \dots, x_n)$, $\mathbf{y} = (y_1, \dots, y_n) \in \mathbb{R}^n$, y $\lambda \in \mathbb{R}$. Se definen la operación interna **suma**

$$\begin{array}{rcl} + : \mathbb{R}^n \times \mathbb{R}^n & \longrightarrow & \mathbb{R}^n \\ (\mathbf{x}, \mathbf{y}) & \mapsto & \mathbf{x} + \mathbf{y} = (x_1 + y_1, \dots, x_n + y_n), \end{array}$$

y la operación externa **producto por un escalar**

$$\begin{array}{rcl} \cdot : \mathbb{R} \times \mathbb{R}^n & \longrightarrow & \mathbb{R}^n \\ (\lambda, \mathbf{x}) & \mapsto & \lambda\mathbf{x} = (\lambda x_1, \dots, \lambda x_n). \end{array}$$

Se puede observar que el conjunto $\mathbb{R}^n$ con las operaciones que se acaban de definir conserva las mismas propiedades que en el caso $\mathbb{R}^2$.

Propiedades. Se comprueba que para todo $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z} \in \mathbb{R}^n$ la suma verifica:

1. Asociativa: $\mathbf{x} + (\mathbf{y} + \mathbf{z}) = (\mathbf{x} + \mathbf{y}) + \mathbf{z}$.
2. Conmutativa: $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$.
3. Elemento neutro: existe $\mathbf{0} = (0, \dots, 0) \in \mathbb{R}^n$ tal que $\mathbf{x} + \mathbf{0} = \mathbf{0} + \mathbf{x} = \mathbf{x}$.
4. Elemento opuesto: $\forall \mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ existe $-\mathbf{x} = (-x_1, \dots, -x_n) \in \mathbb{R}^n$ tal que $\mathbf{x} + (-\mathbf{x}) = (-\mathbf{x}) + \mathbf{x} = \mathbf{0}$.

Es decir, $(\mathbb{R}^n, +)$ es un grupo abeliano.

Además dados $\lambda, \mu \in \mathbb{R}$, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, se cumplen las siguientes propiedades con respecto a la operación externa:

5. $\lambda(\mathbf{x} + \mathbf{y}) = \lambda\mathbf{x} + \lambda\mathbf{y}$.
6. $(\lambda + \mu)\mathbf{x} = \lambda\mathbf{x} + \mu\mathbf{x}$.
7. $\lambda(\mu\mathbf{x}) = (\lambda\mu)\mathbf{x}$.
8. $1\mathbf{x} = \mathbf{x}$.

El conjunto $(\mathbb{R}^n, +, \cdot)$ con las operaciones habituales que verifican las 8 propiedades anteriores, se dice que tiene **estructura de espacio vectorial** sobre $\mathbb{R}$, o que es un espacio vectorial real.

De las propiedades 1-8 anteriores se deducen las siguientes:

Propiedades. Dados $\lambda, \mu \in \mathbb{R}$, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, se verifica:

- $0\,\mathbf{x} = \mathbf{0}$ y $\lambda\mathbf{0} = \mathbf{0}$.
- $\lambda\mathbf{x} = \mathbf{0} \implies \lambda = 0$ o $\mathbf{x} = \mathbf{0}$.
- $\lambda\mathbf{x} = \mu\mathbf{x}$ con $\mathbf{x} \neq \mathbf{0} \implies \lambda = \mu$.
- $\lambda\mathbf{x} = \lambda\mathbf{y}$ con $\lambda \neq 0 \implies \mathbf{x} = \mathbf{y}$.
- $(-1)\mathbf{x} = -\mathbf{x}$.
- $(-\lambda)\mathbf{x} = \lambda(-\mathbf{x}) = -(\lambda\mathbf{x})$.

Además de las operaciones usuales estudiadas hasta el momento, se pueden definir en $\mathbb{R}^n$ otras operaciones tipo "suma" y tipo "producto por un escalar". Si con las operaciones definidas se siguen verificando las 8 propiedades descritas, $\mathbb{R}^n$ sigue conservando la estructura algebraica anterior, como se ve en la sección siguiente.

5.2. ESPACIOS VECTORIALES REALES

Se generaliza a continuación la estructura algebraica de $(\mathbb{R}^n, +, \cdot)$. Se define en general un espacio vectorial como un conjunto V en el que se consideran dos operaciones que verifican las mismas 8 propiedades que $\mathbb{R}^n$ con las operaciones habituales. Con esto se amplia el concepto que se tenía hasta el momento de vector no solo a los vectores de $\mathbb{R}^n$, sino a cualquier elemento que pertenezca al conjunto V. Se observa que el caso anterior, $\mathbb{R}^n$ con sus operaciones habituales, es solo un caso particular de espacio vectorial real.

Definición 5.8. Un **espacio vectorial real** es un conjunto no vacío V, cuyos elementos se llaman **vectores**, dotado de dos operaciones, una interna que se denomina *suma*:

$$\begin{array}{rcl} + : V \times V & \longrightarrow & V \\ (\mathbf{u}, \mathbf{v}) & \mapsto & \mathbf{u} + \mathbf{v}, \end{array}$$

y una externa, denominada *producto por un escalar*:

$$\begin{array}{rcl} \cdot : \mathbb{R} \times V & \longrightarrow & V \\ (\lambda, \mathbf{v}) & \mapsto & \lambda \mathbf{v}, \end{array}$$

verificando las siguientes propiedades:

- Con respecto a la operación interna, $(V, +)$ es un grupo abeliano, es decir cumple:
    1. Asociativa: $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$, $\forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in V$.
    2. Conmutativa: $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$, $\forall \mathbf{u}, \mathbf{v} \in V$.
    3. Elemento neutro: $\exists \mathbf{e} \in V$ tal que $\mathbf{v} + \mathbf{e} = \mathbf{e} + \mathbf{v} = \mathbf{v}$, $\forall \mathbf{v} \in V$.
    4. Elemento opuesto: $\forall \mathbf{v} \in V$, $\exists \mathbf{v}' \in V$ tal que $\mathbf{v} + \mathbf{v}' = \mathbf{v}' + \mathbf{v} = \mathbf{e}$.
- Con respecto a la operación externa:
    5. $\lambda(\mathbf{u} + \mathbf{v}) = \lambda\mathbf{u} + \lambda\mathbf{v}$, $\forall \lambda \in \mathbb{R}$, $\forall \mathbf{u}, \mathbf{v} \in V$.
    6. $(\lambda + \mu)\mathbf{v} = \lambda\mathbf{v} + \mu\mathbf{v}$, $\forall \lambda, \mu \in \mathbb{R}$, $\forall \mathbf{v} \in V$.
    7. $(\lambda\mu)\mathbf{v} = \lambda(\mu\mathbf{v})$, $\forall \lambda, \mu \in \mathbb{R}$, $\forall \mathbf{v} \in V$.
    8. $1\mathbf{v} = \mathbf{v}$, $\forall \mathbf{v} \in V$.

El espacio vectorial se denota mediante $(V, +, \cdot)$.

Dado un espacio vectorial $(V, +, \cdot)$ se pueden demostrar las siguientes propiedades a partir de la definición anterior:

Propiedades.

- $0\,\mathbf{v} = \mathbf{e}$ y $\lambda\mathbf{e} = \mathbf{e}$.
- $\lambda\mathbf{v} = \mathbf{e} \implies \lambda = 0$ o $\mathbf{v} = \mathbf{e}$.
- $\lambda\mathbf{v} = \mu\mathbf{v}$ con $\mathbf{v} \neq \mathbf{e} \implies \lambda = \mu$.
- $\lambda\mathbf{u} = \lambda\mathbf{v}$ con $\lambda \neq 0 \implies \mathbf{u} = \mathbf{v}$.
- $(-1)\mathbf{v} = \mathbf{v}'$.
- $(-\lambda)\mathbf{v} = \lambda\mathbf{v}' = (\lambda\mathbf{v})'$.

Notación. Normalmente, si no hay confusión, el elemento neutro respecto de la suma se denota por $\mathbf{0}_V$ y se llama *elemento nulo*, y el opuesto de $\mathbf{v} \in V$ se denota por $-\mathbf{v}$.

Observaciones.

- A partir de ahora un vector ya no es solo un elemento de $\mathbb{R}^n$, sino cualquier elemento perteneciente a un espacio vectorial V.
- Es importante destacar el hecho de que la estructura de espacio vectorial no la posee el conjunto, sino este unido a dos operaciones. Es decir, un mismo conjunto puede ser o no espacio vectorial dependiendo de las operaciones definidas sobre él. Otro hecho a tener en cuenta es que la operación interna se denota por $+$, lo cual no quiere decir que sea la suma habitual del conjunto (ver Ejercicio 2 de la sección de Ejercicios Propuestos).
- Para demostrar si una terna $(V, +, \cdot)$ tiene estructura de espacio vectorial hay que comprobar primero que la operación suma es interna, la operación producto por un escalar es externa y luego que se verifican las 8 propiedades de la Definición 5.8. Para probar que $(V, +, \cdot)$ no tiene estructura de espacio vectorial basta encontrar una propiedad que no se verifique, dando para ello un contraejemplo.

✓ **Ejemplo 2.** Se considera $\mathbb{R}^3$ con la suma habitual y un nuevo producto por un escalar definido mediante $\lambda(x, y, z) = (\lambda x, y, \lambda z)$. Comprobar que $(\mathbb{R}^3, +, \cdot)$ no es un espacio vectorial real.

Solución: Por tratarse de la suma habitual, $(\mathbb{R}^3, +)$ verifica las propiedades 1-4 y por tanto es un grupo abeliano. Se puede comprobar que no se verifica la propiedad 6 calculando:

- $(\lambda + \mu)(x, y, z) = ((\lambda + \mu)x, y, (\lambda + \mu)z)$,
- $\lambda(x, y, z) + \mu(x, y, z) = (\lambda x, y, \lambda z) + (\mu x, y, \mu z) = ((\lambda + \mu)x, 2y, (\lambda + \mu)z)$.

Se observa que

$$(\lambda + \mu)(x, y, z) \neq \lambda(x, y, z) + \mu(x, y, z),$$

por tanto, se concluye que $\mathbb{R}^3$ con estas operaciones no es un espacio vectorial real.

◇

Observación. Se pueden definir espacios vectoriales no reales, donde el conjunto de escalares que actúa en la operación externa es un cuerpo $\mathbb{K}$, por ejemplo el conjunto de números complejos $\mathbb{C}$. En este caso se dice que el conjunto V es un $\mathbb{K}$-espacio vectorial o un espacio vectorial sobre $\mathbb{K}$. El texto se centra en el estudio de espacios vectoriales reales, es decir, sobre $\mathbb{R}$. Por comodidad en ocasiones se omite la palabra "real" al hacer referencia a dichos espacios.

5.2.1. EJEMPLOS DE ESPACIOS VECTORIALES

Se presentan a continuación algunos conjuntos ya conocidos que tienen estructura de espacio vectorial.

1. El conjunto $\mathbb{R}^n$ con las operaciones usuales descritas anteriormente es un espacio vectorial real.
2. El conjunto de las matrices de orden $n \times m$ sobre $\mathbb{R}$, $\text{Mat}_{n\times m}(\mathbb{R})$, con la suma y producto por un escalar, estudiados en el Capítulo 1, tiene estructura de espacio vectorial real.
3. El conjunto de los polinomios de grado menor o igual que n con coeficientes reales en una indeterminada, denotado por $\mathbb{R}_n[x]$, con las operaciones habituales entre polinomios es un espacio vectorial real, donde

   $$\mathbb{R}_n[x] = \{a_0 + a_1x + \cdots + a_nx^n \mid a_i \in \mathbb{R},\, i = 0, \ldots, n\}.$$

   Se recuerda que la suma habitual entre polinomios se realiza sumando los coeficientes de los monomios del mismo grado.
4. También el conjunto de polinomios de cualquier grado con coeficientes en $\mathbb{R}$ en una indeterminada, denotado por $\mathbb{R}[x]$, es un espacio vectorial real.
5. Dados dos espacios vectoriales reales V y W, el producto cartesiano $V \times W$ también es un espacio vectorial real.
6. El conjunto de n-tuplas complejas

   $$\mathbb{C}^n = \{(x_1, \ldots, x_n) \mid x_i \in \mathbb{C}, i = 1, \ldots, n\}$$

   con las operaciones habituales es un espacio vectorial, que es real cuando el conjunto que actúa en el producto es $\mathbb{R}$ y complejo si es $\mathbb{C}$.
7. El conjunto $\mathcal{C}([a, b])$ de las funciones reales continuas en el intervalo $[a, b]$ es un espacio vectorial real con las operaciones usuales entre funciones, donde

   $$\mathcal{C}([a, b]) = \{f : [a, b] \to \mathbb{R} \mid f \text{ es continua en } [a, b]\}.$$

5.3. SUBESPACIOS VECTORIALES

Dado el espacio vectorial real $(V, +, \cdot)$ se buscan subconjuntos de V que a su vez tengan estructura de espacio vectorial con las operaciones “heredadas” de V. En general, en los problemas lo interesante no es trabajar con todo el espacio, sino con una parte de este que conserve las propiedades del espacio total.

Definición 5.9. Sea el espacio vectorial real $(V, +, \cdot)$ y $S \subseteq V$ un subconjunto no vacío. Se dice que S es un **subespacio vectorial** de V si y solo si $(S, +, \cdot)$ tiene estructura de espacio vectorial.

Para comprobar si un subconjunto de V es o no subespacio vectorial basta aplicar el siguiente teorema de caracterización de subespacios.

Teorema 5.10. Sea $(V, +, \cdot)$ un espacio vectorial real y sea $S \subseteq V$ un subconjunto *no vacío*. El subconjunto S es un subespacio vectorial de V si y solo si verifica:

1. $\forall \mathbf{u}, \mathbf{v} \in S \implies \mathbf{u} + \mathbf{v} \in S,$
2. $\forall \lambda \in \mathbb{R},\ \forall \mathbf{u} \in S \implies \lambda \mathbf{u} \in S,$

o equivalentemente

$$\forall \mathbf{u}, \mathbf{v} \in S,\ \forall \lambda, \mu \in \mathbb{R} \implies \lambda \mathbf{u} + \mu \mathbf{v} \in S.$$

Observaciones.

- Para demostrar que S es un subespacio vectorial de V basta comprobar que es no vacío, que la operación "+" es interna en el conjunto S y la operación "·" es externa, ya que las 8 propiedades se cumplen automáticamente por ser $(V, +, \cdot)$ un espacio vectorial.
- Todo subespacio debe contener al elemento neutro de la suma del espacio vectorial. La falta de dicho elemento en un subconjunto de un espacio vectorial basta para justificar que no es subespacio vectorial.

✓ **Ejemplo 3.** Comprobar que el conjunto

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

es un subespacio vectorial de $\mathbb{R}^3$ con las operaciones habituales.

Solución: Se verifica primero que $S \neq \emptyset$. En efecto, el elemento $(0, 0, 0) \in S$ por cumplir que $0 + 3 \cdot 0 - 2 \cdot 0 = 0$. Se comprueban a continuación las dos condiciones del Teorema 5.10 anterior:

1. Sean $\mathbf{x} = (x_1, x_2, x_3)$, $\mathbf{y} = (y_1, y_2, y_3)$ dos elementos de S. De la definición de S se sigue que

   $$\begin{cases} x_1 + 3x_2 - 2x_3 &= 0, \\ y_1 + 3y_2 - 2y_3 &= 0. \end{cases}$$

   Se calcula el vector suma $\mathbf{x}+\mathbf{y} = (x_1+y_1, x_2+y_2, x_3+y_3) \in \mathbb{R}^3$ y se comprueba que verifica la condición para pertenecer a S:

   $$(x_1 + y_1) + 3(x_2 + y_2) - 2(x_3 + y_3) = \underbrace{(x_1 + 3x_2 - 2x_3)}_{=0} + \underbrace{(y_1 + 3y_2 - 2y_3)}_{=0} = 0.$$

   De este modo $\mathbf{x} + \mathbf{y} \in S$.

2. Para la segunda condición se considera $\lambda \in \mathbb{R}$, $\mathbf{x} = (x_1, x_2, x_3) \in S$. De nuevo de la definición de S, $x_1 + 3x_2 - 2x_3 = 0$. Se calcula $\lambda\mathbf{x} = (\lambda x_1, \lambda x_2, \lambda x_3) \in \mathbb{R}^3$ y se comprueba que cumple la condición de pertenencia a S:

$$(\lambda x_1) + 3(\lambda x_2) - 2(\lambda x_3) = \lambda \underbrace{(x_1 + 3x_2 - 2x_3)}_{=0} = \lambda 0 = 0.$$

De todo lo anterior se deduce que S es un subespacio vectorial de $\mathbb{R}^3$.

◇

✓ **Ejemplo 4.** Comprobar que $T = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1 + 3x_2 - 2x_3 = 5\}$ no es un subespacio vectorial de $\mathbb{R}^3$.

Solución: El conjunto T no contiene al elemento neutro de la suma $(0, 0, 0)$, por tanto no puede ser un subespacio. Otra forma de probar que T no es subespacio vectorial consiste en refutar una de las dos condiciones del Teorema 5.10 dando un contraejemplo. Por ejemplo, tomando $\lambda = 2 \in \mathbb{R}$ y $\mathbf{x} = (-1, 2, 0) \in T$, se observa que $2\mathbf{x} = (-2, 4, 0)$ no está en T puesto que $(-2) + 3 \cdot (4) - 2 \cdot (0) = 10 \neq 5$.

◇

✓ **Ejemplo 5.** Demostrar que el conjunto $S = \{p(x) \in \mathbb{R}_2[x] \,|\, p(0) = 0\}$ es un subespacio vectorial del espacio vectorial de los polinomios reales de grado menor o igual que dos, $\mathbb{R}_2[x]$.

Solución: Se puede realizar el ejercicio de dos modos distintos, según como se considere definido el conjunto S:

- Tomando S con la expresión del enunciado, $S = \{p(x) \in \mathbb{R}_2[x] \,|\, p(0) = 0\}$, se observa que $S \neq \emptyset$, ya que el polinomio nulo $p = 0$ pertenece a S porque verifica $p(0) = 0$. Se comprueban ahora las dos condiciones del Teorema 5.10:

    1. Dados dos elementos $p(x)$, $q(x) \in S$, se comprueba que el polinomio suma $s(x) = (p + q)(x)$ pertenece a S:

    $$s(0) = (p + q)(0) = p(0) + q(0) = 0 + 0 = 0,$$

    por tanto $s(x) = (p + q)(x) \in S$.

    2. Dados $p(x) \in S$ y $\lambda \in \mathbb{R}$, se comprueba que el polinomio $r(x) = (\lambda p)(x)$ pertenece a S:

    $$r(x) = (\lambda p)(0) = \lambda(p(0)) = \lambda 0 = 0,$$

    por tanto $r(x) = (\lambda p)(x) \in S$.

    Se concluye que S es un subespacio vectorial de $\mathbb{R}_2[x]$.

- Desarrollando la expresión del conjunto, se considera un polinomio genérico de $\mathbb{R}_2[x]$, que es de la forma $p(x) = ax^2 + bx + c$, y se impone la condición de pertenencia a S:

  $$p(x) \in S \Longleftrightarrow p(0) = 0 \Longleftrightarrow a \cdot 0 + b \cdot 0 + c = 0 \Longleftrightarrow c = 0,$$

  por lo que S se puede expresar como: $S = \{ax^2 + bx \,|\, a, b \in \mathbb{R}\}$, es decir, S es el conjunto de polinomios de grado menor o igual que dos con término independiente nulo.

  1. Dados dos elementos $ax^2 + bx$, $a'x^2 + b'x \in S$, se calcula el polinomio suma $(ax^2 + bx) + (a'x^2 + b'x) = (a + a')x^2 + (b + b')x$ y se comprueba que verifica la condición de pertenecer a S, ya que es un polinomio de $\mathbb{R}_2[x]$ con término independiente nulo.
  2. Dados $ax^2 + bx \in S$ y $\lambda \in \mathbb{R}$, se calcula el polinomio $\lambda(ax^2 + bx) = \lambda ax^2 + \lambda bx$ y se comprueba que verifica la condición de pertenencia a S, por ser un polinomio de $\mathbb{R}_2[x]$ de término independiente nulo.

  Se puede concluir por tanto que S es un subespacio vectorial de $\mathbb{R}_2[x]$.

◇

5.3.1. FORMAS DE PRESENTAR UN SUBESPACIO

Según lo visto en el último ejemplo, un subespacio vectorial se puede representar de maneras distintas, dependiendo de cómo se exprese la condición de pertenencia al conjunto, entre las que cabe destacar:

- Condición dada mediante **ecuaciones implícitas**. Por ejemplo:

  $$S = \{ \underbrace{(x, y, z)}_{\text{vector genérico}} \in \mathbb{R}^3 \mid \underbrace{x + 3y - 2z = 0}_{\text{condiciones}} \}.$$

- Condición metida dentro de las componentes del vector, también denominada **forma paramétrica**. Por ejemplo el subespacio S del caso anterior se puede escribir como:

  $$S = \{(-3y + 2z, y, z) \in \mathbb{R}^3 \mid y, z \in \mathbb{R}\}$$

- El subespacio se puede dar mediante un conjunto de vectores que lo representan, que se denomina **sistema generador** y que se estudia en la próxima sección. Siguiendo con el ejemplo anterior:

  $$S = \langle (-3, 1, 0), (2, 0, 1) \rangle.$$

El uso de una representación u otra depende de cada ejercicio, y el paso de una a otra se puede hacer de manera sencilla.

✓ **Ejemplo 6.** Considerar el siguiente subespacio vectorial expresado mediante ecuaciones implícitas:

$$S = \left\{A \in \mathrm{Mat}_2(\mathbb{R}) \mid A = A^t\right\}.$$

Para obtener su expresión en forma paramétrica hay que incorporar la condición $A = A^t$ en la propia matriz A. Si

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{R}),$$

entonces $A = A^t$ si y solo si la matriz A es simétrica, es decir, si y solo si $b = c$. De este modo, la expresión en forma paramétrica del subespacio S es:

$$S = \left\{\begin{pmatrix} a & b \\ b & d \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{R}) \,\middle|\, a, b, d \in \mathbb{R}\right\}.$$

◇

✓ **Ejemplo 7.** Considerar el siguiente subespacio vectorial expresado en forma paramétrica:

$$S = \left\{\begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{R}) \,\middle|\, \lambda \in \mathbb{R}\right\}.$$

Para obtener su expresión mediante ecuaciones implícitas hay que encontrar la o las ecuaciones que verifican las matrices de S. En este caso, si $A \in S$ se tiene que:

$$A = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} = \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \lambda I_2.$$

De este modo, la expresión mediante ecuaciones implícitas del subespacio S es:

$$S = \{A \in \mathrm{Mat}_2(\mathbb{R}) \mid A - \lambda I_2 = \mathbf{0},\ \lambda \in \mathbb{R}\}.$$

◇

✓ **Ejemplo 8.** Sea S el subespacio vectorial dado de forma paramétrica mediante $S = \{(-3\lambda + 2\beta, \lambda, \beta) \in \mathbb{R}^3 \mid \lambda, \beta \in \mathbb{R}\}$. Para calcular su expresión en ecuaciones implícitas es necesario resolver el siguiente sistema de ecuaciones, de variables λ y β:

$$\begin{cases} x = -3\lambda + 2\beta \\ y = \lambda \\ z = \beta. \end{cases}$$

El sistema anterior es un sistema sobredeterminado por tener mayor número de ecuaciones que de incógnitas. Se resuelve despejando las variables λ y β de las dos últimas ecuaciones del sistema en función de x, y, z. Sustituyendo el resultado obtenido en la primera ecuación se obtienen las ecuaciones paramétricas del subespacio.

En este caso concreto, se obtiene directamente que $\lambda = y$ y $\beta = z$, y sustituyendo en la primera ecuación:

$$x = -3\lambda + 2\beta \quad \Longrightarrow \quad x = -3y + 2z \quad \Longrightarrow \quad x + 3y - 2z = 0.$$

De este modo, la expresión del subespacio S en ecuaciones implícitas es:

$$S = \{(x, y, z) \in \mathbb{R}^3 \mid \; x + 3y - 2z = 0\}.$$

◇

5.3.2. ALGUNOS SUBESPACIOS VECTORIALES IMPORTANTES

Se presentan a continuación algunos ejemplos de subespacios vectoriales reales que se manejan en el futuro y que tienen una especial importancia.

1. En el espacio vectorial $\mathbb{R}^3$ todas las rectas que pasan por el origen de coordenadas son subespacios vectoriales (que se pueden identificar con $\mathbb{R}$). Del mismo modo, todos los planos que pasan por el origen son subespacios vectoriales (que se pueden identificar con $\mathbb{R}^2$).

2. Sea V un espacio vectorial real. Se considera el conjunto S formado por el elemento nulo de V, $S = \{\mathbf{0}_V\}$. S es un subespacio vectorial de V.

3. Sea V un espacio vectorial real. Tomando $S = V$, S es un subespacio vectorial de V.

4. El conjunto de soluciones de un sistema homogéneo de ecuaciones lineales $A\mathbf{x} = \mathbf{0}$ es un subespacio vectorial. Más concretamente, sea $A = (a_{ij}) \in \mathrm{Mat}_{n\times m}(\mathbb{R})$ la matriz de coeficientes del sistema

   $$\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

   El conjunto $S = \{\mathbf{x} \in \mathbb{R}^m \mid A\mathbf{x} = \mathbf{0}\}$ es un subespacio vectorial de $\mathbb{R}^m$.

   En este capítulo se aprende a expresar este subespacio vectorial de diversas formas y en secciones posteriores se ve que el rango de la matriz A juega un papel importante.

5. El conjunto $\mathbb{R}_n[x]$ es un subespacio vectorial del espacio $\mathbb{R}[x]$, $\forall n \geq 0$. El conjunto $\mathbb{R}_n[x]$ es un subespacio vectorial del espacio $\mathbb{R}_m[x]$, $\forall n \leq m$.

6. Sean $S_1, \ldots, S_r$ subespacios vectoriales del espacio vectorial real V, entonces la **suma** $S_1 + \cdots + S_r$ y la **intersección** $S_1 \cap \cdots \cap S_r$ son subespacios vectoriales de V. Sin embargo la **unión** de subespacios vectoriales $S_1 \cup \cdots \cup S_r$, no es, en general, un subespacio vectorial. Estos conjuntos se estudian en una sección posterior.

5.4. GENERADORES Y DEPENDENCIA LINEAL

Se introducen a continuación una serie de conceptos básicos para el estudio de los espacios vectoriales.

Definición 5.11. Sea V un espacio vectorial real y $\{\mathbf{v}_1, \dots, \mathbf{v}_r\}$ una familia de vectores de V. Un vector $\mathbf{v} \in V$ se dice **combinación lineal** de los vectores de dicha familia si se puede expresar de la forma:

$$\mathbf{v} = \lambda_1 \mathbf{v}_1 + \cdots + \lambda_r \mathbf{v}_r, \quad \text{donde } \lambda_1, \dots, \lambda_r \in \mathbb{R}.$$

Observaciones.

- El vector $\mathbf{0}$ es combinación lineal de cualquier familia de vectores.
- Para determinar si un vector es combinación lineal de otros es necesario resolver un sistema de ecuaciones lineales, cuyas incógnitas son λ_i.

✓ **Ejemplo 9.** En el espacio vectorial $\mathbb{R}^3$, el vector $\mathbf{v} = (5, -18, -6)$ es combinación lineal de $\mathbf{v}_1 = (1, 2, 3)$ y $\mathbf{v}_2 = (0, 4, 3)$, ya que se puede escribir como

$$\underbrace{(5, -18, -6)}_{\mathbf{v}} = \underbrace{5}_{\lambda_1} \underbrace{(1, 2, 3)}_{\mathbf{v}_1} + \underbrace{(-7)}_{\lambda_2} \underbrace{(0, 4, 3)}_{\mathbf{v}_2}.$$

◇

✓ **Ejemplo 10.** En el espacio vectorial $\mathbb{R}_2[x]$, el polinomio $p(x) = 2x^2 + x + 3$ es combinación lineal de $p_1(x) = x^2 + x$ y $p_2(x) = x - 3$, ya que se puede escribir como

$$2x^2 + x + 3 = 2(x^2 + x) - (x - 3).$$

◇

Definición 5.12. Se considera una familia $\{\mathbf{v}_1, \dots, \mathbf{v}_r\}$ de vectores del espacio vectorial real V. El conjunto de todas las combinaciones lineales de $\mathbf{v}_1, \dots, \mathbf{v}_r$ es un subespacio vectorial de V, denominado **clausura lineal** de la familia, y denotado por $L(\mathbf{v}_1, \dots, \mathbf{v}_r)$ o $\langle \mathbf{v}_1, \dots, \mathbf{v}_r \rangle$:

$$L(\mathbf{v}_1, \dots, \mathbf{v}_r) = \langle \mathbf{v}_1, \dots, \mathbf{v}_r \rangle = \big\{\lambda_1 \mathbf{v}_1 + \cdots + \lambda_r \mathbf{v}_r \mid \lambda_i \in \mathbb{R},\ i = 1, \dots, r\big\}.$$

Observación. La clausura lineal es el menor subespacio vectorial que contiene a la familia de vectores $\{\mathbf{v}_1, \dots, \mathbf{v}_r\}$, de ahí el nombre de clausura.

✓ **Ejemplo 11.** La clausura lineal de los vectores $(1, 0)$ y $(0, 1)$ es:

$$\langle (1, 0), (0, 1) \rangle = \{\lambda_1(1, 0) + \lambda_2(0, 1) \mid \lambda_i \in \mathbb{R}\} = \{(\lambda_1, \lambda_2) \mid \lambda_i \in \mathbb{R}\} = \mathbb{R}^2.$$

◇

Definición 5.13. Sea S un subespacio vectorial del espacio vectorial real V. Se considera una familia de vectores $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ de S. Se dice que la familia anterior de vectores **genera** el subespacio S o es un **sistema generador** de S si:

$$S = \langle \mathbf{v}_1, \ldots, \mathbf{v}_r \rangle,$$

es decir, todo elemento de S se puede escribir como combinación lineal de los vectores $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$:

$$\forall \mathbf{s} \in S,\ \exists \lambda_1, \ldots, \lambda_r \in \mathbb{R} \quad \text{tales que} \quad \mathbf{s} = \lambda_1 \mathbf{v}_1 + \cdots + \lambda_r \mathbf{v}_r.$$

✓ **Ejemplo 12.** El conjunto $\{(1,0,0), (0,1,0), (0,0,1)\}$ genera todo $\mathbb{R}^3$, ya que cualquier elemento $(x_1, x_2, x_3) \in \mathbb{R}^3$ se puede escribir como:

$$(x_1, x_2, x_3) = x_1(1,0,0) + x_2(0,1,0) + x_3(0,0,1).$$

◇

✓ **Ejemplo 13.** El conjunto de matrices $\left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$ es un sistema generador de $\mathrm{Mat}_2(\mathbb{R})$, ya que cualquier matriz $A \in \mathrm{Mat}_2(\mathbb{R})$ se puede escribir como:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = a \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + b \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + c \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + d \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

◇

✓ **Ejemplo 14.** El conjunto $\{x^2, x\}$ es un sistema generador del subespacio S del Ejemplo 5 dado por $S = \{p(x) \in \mathbb{R}_2[x] \,|\, p(0) = 0\} = \{ax^2 + bx \,|\, a, b \in \mathbb{R}\}$, ya que cualquier polinomio $p(x) \in S$ se puede escribir como:

$$p(x) = ax^2 + bx = a\,(x^2) + b\,(x).$$

◇

Definición 5.14. Sea V un espacio vectorial real. Se dice que los vectores $\mathbf{v}_1, \ldots, \mathbf{v}_r \in V$ son **linealmente independientes** (o **libres**) si la única manera de expresar el vector nulo como combinación lineal de $\mathbf{v}_1, \ldots, \mathbf{v}_r$ es la trivial, es decir,

$$\boxed{\lambda_1 \mathbf{v}_1 + \cdots + \lambda_r \mathbf{v}_r = \mathbf{0} \implies \lambda_1 = \cdots = \lambda_r = 0}$$

Los vectores $\mathbf{v}_1, \ldots, \mathbf{v}_r \in V$ son **linealmente dependientes** (o **ligados**) si $\boxed{\text{no}}$ son linealmente independientes, es decir, existen $\lambda_1, \ldots, \lambda_r \in \mathbb{R}$ *no todos nulos* tales que

$$\lambda_1 \mathbf{v}_1 + \cdots + \lambda_r \mathbf{v}_r = \mathbf{0}. \tag{5.1}$$

Las siguientes propiedades ayudan a establecer la dependencia o independencia lineal de una familia de vectores.

Propiedades. Dado un espacio vectorial real V se tiene:

1. Cualquier familia de vectores que contenga al vector nulo es ligada.
2. El conjunto formado por un único vector $\{\mathbf{v}\}$ es libre si y solo si $\mathbf{v} \neq \mathbf{0}$.
3. Dos vectores $\mathbf{u}$ y $\mathbf{v}$ de V son linealmente dependientes si y solo si uno de ellos es múltiplo escalar del otro, es decir, si y solo si son proporcionales.
4. La familia $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ de vectores de V es ligada si y solo si existe un vector $\mathbf{v}_i$ que depende linealmente del resto.
5. Si el vector $\mathbf{w} \in V$ depende linealmente de la familia de vectores de V, $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$, entonces $\langle \mathbf{w}, \mathbf{v}_1, \ldots, \mathbf{v}_r \rangle = \langle \mathbf{v}_1, \ldots, \mathbf{v}_r \rangle$.

Observación. El estudio de la dependencia lineal de una familia de vectores se traduce en la discusión de un sistema de ecuaciones lineales homogéneo. Si el sistema es compatible determinado, entonces la familia de vectores es libre; si por el contrario el sistema es compatible indeterminado, la familia de vectores es ligada y la resolución del sistema proporciona combinaciones lineales explícitas entre los vectores.

Caso particular $V = \mathbb{R}^n$

En el caso de que el espacio vectorial sea $V = \mathbb{R}^n$, dada una familia de vectores $\mathcal{F} = \{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ se pueden estudiar las relaciones en dicha familia mediante el uso de matrices aplicando el método de eliminación gaussiana.

- Para decidir si un vector $\mathbf{v} \in \mathbb{R}^n$ pertenece a la clausura lineal de una familia libre $\mathcal{F}$ se hace uso del siguiente resultado:

$$\boxed{\mathbf{v} \in \langle \underbrace{\mathbf{v}_1, \ldots, \mathbf{v}_r}_{\text{familia libre}} \rangle \Longleftrightarrow \operatorname{rg} \begin{pmatrix} \mathbf{v}_1 \\ \hline \vdots \\ \hline \mathbf{v}_r \\ \hline \mathbf{v} \end{pmatrix} = \operatorname{rg} \begin{pmatrix} \mathbf{v}_1 \\ \hline \vdots \\ \hline \mathbf{v}_r \end{pmatrix}}$$

- Para estudiar la dependencia o independencia lineal de la familia $\mathcal{F}$ se utiliza:

$$\boxed{\mathbf{v}_1, \ldots, \mathbf{v}_r \in \mathbb{R}^n \text{ son libres} \Longleftrightarrow \operatorname{rg}(\mathbf{v_1}|\cdots|\mathbf{v}_r) = r \Longleftrightarrow \operatorname{rg} \begin{pmatrix} \mathbf{v}_1 \\ \hline \vdots \\ \hline \mathbf{v}_r \end{pmatrix} = r}$$

  El valor $k = \operatorname{rg}(\mathbf{v_1}|\cdots|\mathbf{v}_r) \leq r$ indica el número máximo de vectores linealmente independientes de la familia $\{\mathbf{v_1}, \cdots, \mathbf{v}_r\}$. Del resultado anterior se deduce que toda familia escalonada de vectores de $\mathbb{R}^n$ es libre.

Si se consideran los vectores de $\mathbb{R}^n$, $\mathbf{v}_1 = (v_{11}, \dots, v_{1n}), \dots,\ \mathbf{v}_r = (v_{r1}, \dots, v_{rn})$, la expresión (5.1) se puede desarrollar como:

$$\lambda_1(v_{11}, \dots, v_{1n}) + \cdots + \lambda_r(v_{r1}, \dots, v_{rn}) = (0, \dots, 0),$$

dando lugar al siguiente sistema de ecuaciones, igualando componente a componente:

$$\left\{\begin{array}{l} \lambda_1 v_{11} + \cdots + \lambda_r v_{r1} = 0 \\ \quad\vdots \\ \lambda_1 v_{1n} + \cdots + \lambda_r v_{rn} = 0 \end{array}\right. \Longrightarrow \underbrace{\begin{pmatrix} v_{11} & \cdots & v_{r1} \\ \vdots & \ddots & \vdots \\ v_{1n} & \cdots & v_{rn} \end{pmatrix}}_{A} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_r \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

$$\underbrace{\begin{matrix} v_{11} \\ \vdots \\ v_{1n}\end{matrix}}_{\mathbf{v}_1} \quad \cdots \quad \underbrace{\begin{matrix} v_{r1} \\ \vdots \\ v_{rn}\end{matrix}}_{\mathbf{v}_r}$$

El teorema de Rouché-Frobenius asegura que la única solución del sistema homogéneo anterior es la trivial $(\lambda_1, \dots, \lambda_r) = (0, \dots, 0)$ si y solo si el rango de la matriz de coeficientes A es r (número de vectores), en cuyo caso los vectores $\mathbf{v}_1, \dots, \mathbf{v}_r$ son linealmente independientes. De este modo el problema se reduce a estudiar el carácter del sistema de ecuaciones (compatible determinado o indeterminado) que depende del rango de la matriz de coeficientes A (que coincide con el rango de A^t).

✓ **Ejemplo 15.** En el espacio vectorial $\mathbb{R}^3$, se consideran los vectores $\mathbf{v}_1 = (1, 1, 2)$, $\mathbf{v}_2 = (-1, 0, -1)$ y $\mathbf{v} = (2, -1, a)$.

- Estudiar si el vector $\mathbf{v}$ pertenece a la clausura lineal de $\mathcal{F} = \{\mathbf{v}_1, \mathbf{v}_2\}$, dependiendo de los valores del parámetro $a \in \mathbb{R}$.
- Para los casos en que $\mathbf{v} \in \langle \mathbf{v}_1, \mathbf{v}_2 \rangle$, encontrar una dependencia lineal explícita entre los vectores.

Solución: En primer lugar se observa que los vectores de $\mathcal{F}$ son linealmente independientes por no ser proporcionales, por lo que $\operatorname{rg}\begin{pmatrix} \mathbf{v}_1 \\ \hline \mathbf{v}_2 \end{pmatrix} = 2$.

- Es necesario estudiar el rango de la matriz formada por los tres vectores:

$$\begin{pmatrix} \mathbf{v}_1 \\ \hline \mathbf{v}_2 \\ \hline \mathbf{v} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 2 \\ -1 & 0 & -1 \\ 2 & -1 & a \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 0 & -3 & a-4 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & a-1 \end{pmatrix}.$$

  Se sigue que el rango de la matriz anterior es 2 si y solo si $a = 1$, lo que indica que $\mathbf{v} \in \langle \mathbf{v}_1, \mathbf{v}_2 \rangle$ si y solo si $a = 1$.

- Para el valor $a = 1$, se busca una dependencia lineal del tipo

$$\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 = \mathbf{v} \quad \Longrightarrow \quad \lambda_1(1, 1, 2) + \lambda_2(-1, 0, -1) = (2, -1, 1),$$

para lo cual es necesario resolver el siguiente sistema, de variables λ_1 y λ_2:

$$\left\{\begin{array}{rcr} \lambda_1-\lambda_2 & = & 2\\ \lambda_1 & = & -1\\ 2\lambda_1-\lambda_2 & = & 1\end{array}\right. \Longrightarrow \begin{pmatrix}\underbrace{\begin{matrix}1\\1\\2\end{matrix}}_{\mathbf{v}_1} & \underbrace{\begin{matrix}-1\\0\\-1\end{matrix}}_{\mathbf{v}_2}\end{pmatrix}\begin{pmatrix}\lambda_1\\ \lambda_2\end{pmatrix}=\underbrace{\begin{pmatrix}2\\-1\\1\end{pmatrix}}_{\mathbf{v}}.$$

Calculando una forma escalonada, el sistema es equivalente a este otro:

$$\begin{pmatrix}1 & -1\\ 0 & 1\\ 0 & 0\end{pmatrix}\begin{pmatrix}\lambda_1\\ \lambda_2\end{pmatrix}=\begin{pmatrix}2\\-3\\0\end{pmatrix} \Longrightarrow \left\{\begin{array}{rcr} \lambda_1-\lambda_2 & = & 2\\ \lambda_2 & = & -3.\end{array}\right.$$

La única solución es $\lambda_1 = -1$, $\lambda_2 = -3$, y por tanto $\mathbf{v} = -\mathbf{v}_1 - 3\mathbf{v}_2$. Se podría haber iniciado el ejercicio buscando una relación más general del tipo $\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \lambda_3\mathbf{v} = \mathbf{0}$.

◇

5.5. BASES Y DIMENSIÓN

El concepto de base de un espacio vectorial V es uno de los más importantes de este capítulo. Se trata de encontrar conjuntos de vectores que generen el espacio V y sean linealmente independientes. Es decir, se buscan familias de vectores que conteniendo el menor número de elementos posible permitan reconstruir todo el espacio.

Recordar que un subespacio vectorial es un espacio vectorial que está dentro de otro, de modo que todo lo que en este apartado se diga para espacios vectoriales es válido también para subespacios.

Definición 5.15. Sea V un espacio vectorial real y $\mathcal{B} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ una familia de vectores de V. Se dice que $\mathcal{B}$ es una **base** de V si:

- Los vectores $\mathbf{u}_1, \dots, \mathbf{u}_n$ son linealmente independientes.
- El conjunto $\mathcal{B}$ es un sistema generador de V, es decir, $V = \langle \mathbf{u}_1, \dots, \mathbf{u}_n \rangle$.

Observación. Si el espacio vectorial V es el espacio nulo, $V = \{\mathbf{0}\}$, entonces la base es el conjunto vacío, $\mathcal{B} = \emptyset$.

Notación. Sea un espacio vectorial $V \neq \{\mathbf{0}\}$. Dados los vectores $\mathbf{u}_1, \dots, \mathbf{u}_n \in V$:

- Para expresar que dichos vectores forman una base de V, estos se escriben entre *llaves* de la siguiente manera:

$$\text{Base de } V = \mathcal{B}_V = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}.$$

- Para expresar que el espacio V está generado por los vectores $\mathbf{u}_i$, estos se escriben entre *ángulos*:
$$V = \langle \mathbf{u}_1, \ldots, \mathbf{u}_n \rangle.$$

Observar que la diferencia reside en que la base es un conjunto finito formado únicamente por los n vectores $\mathbf{u}_i$, mientras que en el segundo caso están además todas sus combinaciones lineales, obteniéndose el espacio vectorial V, que tiene infinitos elementos.

✓ **Ejemplo 16.** Los espacios vectoriales con los que habitualmente se trabaja $(\mathbb{R}^n, \mathrm{Mat}_n(\mathbb{R}), \mathbb{R}_n[x])$ tienen una base destacada, que es la que se considera salvo que se indique lo contrario. Esta base se denomina **base canónica**.

- El conjunto $\{(1,0,0),(0,1,0),(0,0,1)\}$ es la base canónica de $\mathbb{R}^3$. De modo análogo se considera la base canónica de $\mathbb{R}^n$, $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$, siendo
$$\mathbf{e}_i = (0, \ldots, 0, \overset{i)}{1}, 0, \ldots, 0).$$
- La base canónica del espacio de matrices cuadradas 2 × 2 reales, $\mathrm{Mat}_2(\mathbb{R})$ es:
$$\left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}.$$
En general para el espacio $\mathrm{Mat}_{n\times m}(\mathbb{R})$ la base canónica es $\{E_{ij}, i = 1, \ldots, n,\ j = 1, \ldots, m\}$, donde
$$\begin{array}{r} \\ \\ \\ \text{posición } ij \longleftarrow \\ \\ \end{array} E_{ij} = \begin{pmatrix} 0 & & & & \\ & \ddots & & & \\ & & 0 & & \\ & \boxed{1} & & \ddots & \\ & & & & 0 \end{pmatrix}.$$
- La base canónica del espacio $\mathbb{R}_n[x]$ es $\{1, x, x^2, \ldots, x^n\}$ o bien cambiando el orden $\{x^n, \ldots, x^2, x, 1\}$.

◇

✓ **Ejemplo 17.** Una base del espacio $\mathbb{R}[x]$ es $\{1, x, x^2, \ldots, x^n, \ldots\}$. En este caso la base tiene infinitos elementos. En este texto no se consideran espacios de este tipo.

◇

Uno de los problemas más habituales del Álgebra Lineal consiste en hallar una base de un subespacio vectorial, estudiando la independencia lineal de un sistema generador del mismo. En el caso de que el espacio esté presentado mediante ecuaciones implícitas, para hallar un sistema generador es necesario resolver el sistema de ecuaciones lineales que definen el subespacio. El siguiente ejemplo muestra cómo proceder en este último caso:

✓ **Ejemplo 18.** Encontrar una base del subespacio de $\mathbb{R}^4$:

$$S = \left\{ (x_1, x_2, x_3, x_4) \left| \begin{array}{c} x_1 + x_2 - 4x_3 + 2x_4 = 0 \\ -x_1 + x_2 + 2x_3 - 4x_4 = 0 \end{array} \right. \right\}.$$

Solución: Para resolver el sistema de ecuaciones se extrae la matriz de coeficientes y se aplica el método de Gauss-Jordan:

$$\begin{pmatrix} 1 & 1 & -4 & 2 \\ -1 & 1 & 2 & -4 \end{pmatrix} \longrightarrow \cdots \longrightarrow \begin{pmatrix} \boxed{1} & 0 & -3 & 3 \\ 0 & \boxed{1} & -1 & -1 \end{pmatrix}.$$

El rango de la matriz es $2 < 4$ (= número de incógnitas) y por tanto el sistema es compatible indeterminado con dos parámetros libres. Se reconstruye el sistema de ecuaciones y se resuelve tomando x_3 y x_4 como parámetros libres:

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

Se sustituyen las componentes x_1 y x_2 del vector genérico de S, obteniéndose la expresión paramétrica del subespacio:

$$S = \{(3x_3 - 3x_4, x_3 + x_4, x_3, x_4) \mid x_3, x_4 \in \mathbb{R}\}.$$

Sacando los parámetros x_3 y x_4 factor común, se obtiene la expresión de S como clausura lineal de dos vectores:

$$S = \{x_3(3,1,1,0) + x_4(-3,1,0,1) \mid x_3, x_4 \in \mathbb{R}\} = \langle \underbrace{(3,1,1,0)}_{\mathbf{s}_1}, \underbrace{(-3,1,0,1)}_{\mathbf{s}_2} \rangle.$$

Los vectores $\{\mathbf{s}_1, \mathbf{s}_2\}$ son un sistema generador de S y además son linealmente independientes por no ser proporcionales, por tanto, se puede concluir que forman una base de S:

$$\mathcal{B}_S = \{(3,1,1,0), (-3,1,0,1)\}.$$

◇

✓ **Ejemplo 19.** Es fácil comprobar que la familia $\mathcal{F} = \{(-3,1,0),(2,0,1)\}$ es una base del subespacio vectorial $S = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1 + 3x_2 - 2x_3 = 0\}$ del Ejemplo 3. Efectivamente, en primer lugar se observa que los elementos de la familia pertenecen al subespacio S. Aprovechando los cálculos ya realizados, $S = \{(-3\alpha + 2\beta, \alpha, \beta) | \alpha, \beta \in \mathbb{R}\}$. Un vector genérico se puede escribir en combinación lineal de $(-3,1,0)$ y $(2,0,1)$ como:

$$(-3\alpha + 2\beta, \alpha, \beta) = (-3\alpha, \alpha, 0) + (2\beta, 0, \beta) = \alpha(-3,1,0) + \beta(2,0,1).$$

La familia $\mathcal{F}$ es por tanto generadora de S y además, por ser libre, es base.

◇

5.5.1. TEOREMAS DE CARACTERIZACIÓN DE BASES. DIMENSIÓN

Se presentan a continuación una serie de resultados teóricos que facilitan la obtención de bases.

Teorema 5.16. Sea V un espacio vectorial real, y sea $\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ una familia de vectores de V. La familia $\mathcal{B}$ es una base de V si y solo si todo vector $\mathbf{v} \in V$ se puede escribir de **modo único** como combinación lineal de $\mathbf{u}_1, \ldots, \mathbf{u}_n$.

Teorema 5.17. Sea $V \neq \{\mathbf{0}\}$ un espacio vectorial real generado por un número finito de elementos. Todo sistema generador de V contiene una base de V.

La idea de la demostración es sencilla. Se parte de un sistema generador finito G de V. Si G es además un conjunto linealmente independiente, entonces G es una base y se ha terminado. En caso contrario hay que eliminar vectores del sistema generador hasta obtener un conjunto linealmente independiente, que será una base del espacio vectorial dado.

✓ **Ejemplo 20.** En $\mathbb{R}^3$ se considera la siguiente familia de vectores:

$$G = \{(1,1,2), (-1,0,-1), (2,-1,1), (0,1,1), (2,-1,2)\}.$$

Encontrar una base del espacio S generado por los vectores de G.

Solución: La familia es generadora por definición, por tanto, para hallar una base basta eliminar los vectores linealmente dependientes. Aplicando el método de eliminación gaussiana a la matriz cuyas filas son los vectores dados:

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

se observa que las filas 3 y 4 de la matriz anterior se han convertido en $\mathbf{0}$ tras dicho proceso, lo que quiere decir que el tercer y cuarto vector de G son combinación lineal de los restantes, de modo que se pueden eliminar del sistema generador. Así:

$$S = \langle (1,1,2), (-1,0,-1), (2,-1,2) \rangle.$$

Si se quiere expresar la base hay que escribir los vectores entre llaves,

$$\mathcal{B}_S = \{(1,1,2), (-1,0,-1), (2,-1,2)\}.$$

Se podrían tomar también los vectores del paso final de la eliminación gaussiana. De este modo, $S = \langle (1,1,2), (0,1,1), (0,0,1) \rangle$, dando lugar a la base escalonada $\{(1,1,2), (0,1,1), (0,0,1)\}$.

◇

Un resultado fundamental del Álgebra Lineal es el siguiente:

Teorema 5.18. Sea $V \neq \{\mathbf{0}\}$ un espacio vectorial real finitamente generado. Entonces todas las bases de V tienen el mismo número de elementos.

Debido al teorema anterior se puede enunciar el siguiente concepto.

Definición 5.19. La **dimensión** de un espacio vectorial real V finitamente generado, $\dim V$, es el número de elementos que tiene una cualquiera de sus bases.

✓ **Ejemplo 21.**

- La dimensión de $\mathbb{R}^3$ es 3, ya que la base $\{(1,0,0),(0,1,0),(0,0,1)\}$ de $\mathbb{R}^3$ tiene tres elementos.
- El subespacio vectorial S del Ejemplo 18 tiene dimensión 2 pues el conjunto $\{(3,1,1,0),(-3,1,0,1)\}$ es una base de ese subespacio.

◇

Se indican a continuación las dimensiones de algunos espacios vectoriales con los que se trabaja habitualmente.

✓ **Ejemplo 22.** Al considerar las bases canónicas se deduce que:

- $\dim \mathbb{R}^n = n$.
- $\dim(\mathrm{Mat}_{n\times m}(\mathbb{R})) = n \cdot m$.
- $\dim \mathbb{R}_n[x] = n+1$.

◇

En adelante el texto se centra en espacios vectoriales reales de dimensión finita.

Proposición 5.20. Sea V un espacio vectorial real de dimensión finita, $\dim V = n$. Si S es un subespacio vectorial de V, entonces $0 \leq \dim S \leq n$. Además:

- $\dim S = n \iff S = V$.
- $\dim S = 0 \iff S = \{\mathbf{0}\}$.

Observación. En general, si $S = \{\mathbf{x} \in \mathbb{R}^{\mathbf{n}} \mid A\mathbf{x} = \mathbf{0}\}$, entonces, dim $S = n - \operatorname{rg} A$.

La dimensión de un espacio vectorial impone restricciones sobre el tamaño que pueden tener los conjuntos linealmente independientes y los sistemas generadores de ese espacio.

Proposición 5.21. Sea V un espacio vectorial real de dimensión finita $\dim V = n$. Sea $\mathcal{F} = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ una familia de vectores de V. Se tiene:

1. Si $\mathcal{F}$ es un sistema generador de V, entonces $k \geq \dim V = n$. Es decir, todo sistema generador de V tiene al menos n vectores.

2. Si $\mathcal{F}$ es linealmente independiente, entonces $k \leq \dim V = n$. Es decir, toda familia de vectores linealmente independientes de V tiene como mucho n vectores.
3. Si $\mathcal{F}$ es un sistema generador y $k = \dim V = n$, entonces $\mathcal{F}$ es una base de V.
4. Si $\mathcal{F}$ es linealmente independiente y $k = \dim V = n$, entonces $\mathcal{F}$ es una base de V.

✓ **Ejemplo 23.** Se presentan algunos ejemplos que facilitan la comprensión de los resultados anteriores.

- El conjunto $\mathcal{F} = \{(1,0,1),(0,0,2)\}$ no puede generar $\mathbb{R}^3$ ya que $\dim \mathbb{R}^3 = 3$ y para generar un espacio de dimensión 3 se necesitan como mínimo 3 vectores.
- En $\mathbb{R}^2$ el conjunto $\mathcal{F} = \{(1,0),(0,2),(1,4)\}$ es linealmente dependiente, puesto que el número de elementos de $\mathcal{F}$ supera la dimensión del espacio total que es 2.
- El conjunto $\mathcal{F} = \{(1,5),(2,4)\}$ es una base de $\mathbb{R}^2$ ya que son vectores linealmente independientes por no ser proporcionales y el número de elementos de $\mathcal{F}$ coincide con $\dim \mathbb{R}^2 = 2$.

◇

Las familias linealmente independientes se pueden completar hasta obtener bases del espacio que las contiene, como indica el siguiente teorema.

Teorema 5.22 (Teorema de la base incompleta)**.** Sea $V \neq \{\mathbf{0}\}$ un espacio vectorial real de dimensión finita n, $\dim V = n$. Toda familia linealmente independiente de vectores de V se puede completar hasta obtener una base de V.

Dado un conjunto de vectores de V linealmente independiente que no lleguen a generar todo V, la idea es ir añadiendo vectores de V a los dados, manteniendo el carácter libre de la familia, hasta conseguir tantos vectores linealmente independientes como $\dim V$. Observar que todos ellos generan V por la Proposición 5.21. El resultado anterior se puede aplicar también a subespacios vectoriales, con la precaución de que los vectores que se añaden han de pertenecer al subespacio.

✓ **Ejemplo 24.** En el espacio vectorial $\mathbb{R}^5$ se considera el subespacio

$$S = \{(x_1, x_2, 0, x_4, x_5) \mid x_1, x_2, x_4, x_5 \in \mathbb{R}\}.$$

Completar la familia $\{(1,1,0,3,4),(1,2,0,2,4)\} \subseteq S$, hasta una base de S.

Solución: El subespacio S tiene dimensión 4, ya que una base de S es el conjunto:

$$\mathcal{B}_S = \{(1,0,0,0,0),(0,1,0,0,0),(0,0,0,1,0),(0,0,0,0,1)\}.$$

Los vectores $\{(1,1,0,3,4),(1,2,0,2,4)\}$ son linealmente independientes por no ser proporcionales. Si se añaden los vectores $\{(0,0,0,1,0),(0,0,0,0,1)\}$ a los anteriores, se puede comprobar fácilmente que la familia $\mathcal{B}$ es una base de S,

$$\mathcal{B} = \{(1,1,0,3,4),(1,2,0,2,4),(0,0,0,1,0),(0,0,0,0,1)\}.$$

- Son linealmente independientes pues la matriz que forman tiene rango 4:

$$\begin{pmatrix} 1 & 1 & 0 & 3 & 4 \\ 1 & 2 & 0 & 2 & 4 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow{P_{21}(-1)} \begin{pmatrix} 1 & 1 & 0 & 3 & 4 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

- Por el apartado 4 de la Proposición 5.21 se sabe que 4 vectores linealmente independientes en un espacio de dimensión 4 son base. Con esto se evita demostrar que la familia es generadora. No obstante, para probar que $\mathcal{B}$ es un sistema generador de S habría que expresar un vector cualquiera de S como combinación lineal de los vectores de la familia, resolviendo el siguiente sistema de ecuaciones, de incógnitas α, β, γ y λ:

  $$(x_1, x_2, 0, x_4, x_5) = \alpha(1,1,0,3,4)+\beta(1,2,0,2,4)+\gamma(0,0,0,1,0)+\lambda(0,0,0,0,1),$$

  cuya solución es: $\alpha = 2x_1-x_2$, $\beta = x_2-x_1$, $\gamma = -4x_1+x_2+x_4$, $\lambda = -4x_1+x_5$.

◇

5.6. OPERACIONES CON SUBESPACIOS VECTORIALES

Sean $S_1, \dots, S_r$ subespacios vectoriales de un espacio vectorial real V de dimensión finita. Se dedica este apartado al estudio de nuevos subespacios que surgen al operar con los dados.

5.6.1. INTERSECCIÓN DE SUBESPACIOS

Sean $S_1, \dots, S_r$ subespacios vectoriales de V, entonces el subespacio **intersección**, denotado por $S_1 \cap \cdots \cap S_r$, se define como:

$$S_1 \cap \cdots \cap S_r = \{\mathbf{v} \in V \mid \mathbf{v} \in S_i, \ \forall i = 1, \dots, r\}.$$

Proposición 5.23. Dados los subespacios vectoriales de V, $S_1, \dots, S_r$, el subespacio intersección $S_1 \cap \cdots \cap S_r$ es un subespacio vectorial de cada subespacio S_i, $i = 1, \dots, r$, por tanto,

$$\dim(S_1 \cap \cdots \cap S_r) \leq \dim S_i, \quad \forall i = 1, \dots, r.$$

En la práctica, si se conocen los subespacios $S_1, \dots, S_r$, y se quiere calcular el subespacio $S_1 \cap \cdots \cap S_r$, se deben buscar vectores que satisfagan las condiciones de pertenencia a cada uno de los subespacios $S_1, \dots, S_r$.

Dependiendo de la forma en que vengan dados los subespacios se actúa de un modo u otro. Por ejemplo, suponiendo que los subespacios están dados mediante sus ecuaciones implícitas, entonces el conjunto formado por las ecuaciones de $S_1, \dots, S_r$

da lugar a unas ecuaciones implícitas de $S_1 \cap \cdots \cap S_r$. Esto se debe a que un vector está en $S_1 \cap \cdots \cap S_r$, es decir, está al mismo tiempo en cada S_i para $i = 1, \ldots, r$, si y solo si verifica a la vez las ecuaciones implícitas de S_1, las de S_2, …, hasta las de S_r.

✓ **Ejemplo 25.** Se consideran los subespacios vectoriales de $\mathbb{R}^4$ dados mediante sus ecuaciones implícitas

$$S_1 = \left\{ (x_1, x_2, x_3, x_4) \left| \begin{array}{c} x_1 - x_3 = 0 \\ x_2 + x_4 = 0 \end{array} \right. \right\}, \quad S_2 = \{(x_1, x_2, x_3, x_4) |\ x_1 + x_4 = 0\}.$$

El subespacio intersección está dado por:

$$S_1 \cap S_2 = \left\{ (x_1, x_2, x_3, x_4) \left| \begin{array}{c} x_1 - x_3 = 0 \\ x_2 + x_4 = 0 \\ x_1 + x_4 = 0 \end{array} \right. \right\}.$$

◇

En caso de que se disponga de una base de cada uno de los subespacios, para calcular un vector perteneciente al subespacio intersección se construyen vectores genéricos como combinación lineal de cada una de las bases y se igualan las expresiones dos a dos. Se resuelven los sistemas de ecuaciones obtenidos como se muestra en el siguiente ejemplo.

✓ **Ejemplo 26.** Hallar una base del subespacio intersección $S_1 \cap S_2$ siendo:

$$S_1 = \langle (1, 3, 2, -8), (0, 1, -1, 3) \rangle, \quad S_2 = \langle (1, 5, 0, -2) \rangle.$$

Solución: Es inmediato comprobar los vectores que generan S_1 son linealmente independientes, por lo que $\dim S_1 = 2$ y $\dim S_2 = 1$. Por la Proposición 5.23 se deduce que $\dim(S_1 \cap S_2) \leq 1$.

Se considera un vector genérico de la intersección, $\mathbf{v} \in S_1 \cap S_2$. Por una parte $\mathbf{v} \in S_1$, es decir, existen $\alpha, \beta \in \mathbb{R}$ tales que

$$\mathbf{v} = \alpha(1, 3, 2, -8) + \beta(0, 1, -1, 3).$$

Por otro lado, $\mathbf{v} \in S_2$, por lo que existe $\lambda \in \mathbb{R}$ tal que

$$\mathbf{v} = \lambda(1, 5, 0, -2).$$

Igualando las dos expresiones de $\mathbf{v}$

$$\alpha(1, 3, 2, -8) + \beta(0, 1, -1, 3) = \lambda(1, 5, 0, -2),$$

se obtiene el sistema:

$$\left\{ \begin{array}{lcl} \alpha & = & \lambda \\ 3\alpha + \beta & = & 5\lambda \\ 2\alpha - \beta & = & 0 \\ -8\alpha + 3\beta & = & -2\lambda \end{array} \right.$$

cuya solución es $\alpha = \lambda$, $\beta = 2\lambda$. Así, el vector $\mathbf{v} \in S_1 \cap S_2$ tiene la forma

$$\mathbf{v} = \lambda(1,3,2,-8) + 2\lambda(0,1,-1,3) = (\lambda, 5\lambda, 0, -2\lambda) = \lambda(1,5,0,-2),$$

y se puede concluir que $S_1 \cap S_2 = \langle(1,5,0,-2)\rangle = S_2$, y por tanto, $\dim(S_1 \cap S_2) = 1$.

◇

5.6.2. SUMA DE SUBESPACIOS

Sean $S_1, \dots, S_r$ subespacios vectoriales del espacio vectorial real V, entonces el subespacio **suma**, denotado por $S_1 + \cdots + S_r$, se define como:

$$S_1 + \cdots + S_r = \{\mathbf{s}_1 + \cdots + \mathbf{s}_r \mid \mathbf{s}_i \in S_i, \ \forall i = 1, \dots, r\}.$$

Proposición 5.24. Dados los subespacios vectoriales de V, $S_1, \dots, S_r$, el subespacio suma $S_1 + \cdots + S_r$ es un subespacio vectorial que contiene a todos los S_i, $i = 1, \dots, r$, por tanto,

$$\dim S_i \leq \dim(S_1 + \cdots + S_r), \quad \forall i = 1, \dots, r.$$

En la práctica para hallar el subespacio suma $S_1 + \cdots + S_r$, basta tomar una base de cada subespacio S_i, dado por la familia $\mathcal{B}_i = \{\mathbf{v}_{i1}, \dots, \mathbf{v}_{is_i}\}$, y al juntar todas las bases se obtiene un sistema generador de $S_1 + \cdots + S_r$. Es decir,

$$S_1 + \cdots + S_2 = \langle\underbrace{\mathbf{v}_{11}, \dots, \mathbf{v}_{1s_1}}_{\mathcal{B}_1}, \dots, \underbrace{\mathbf{v}_{r1}, \dots, \mathbf{v}_{rs_r}}_{\mathcal{B}_r}\rangle.$$

A partir del sistema generador se consigue una base sin más que eliminar los vectores linealmente dependientes.

✓ **Ejemplo 27.** Hallar una base del subespacio suma $S_1 + S_2$ siendo S_1 y S_2 los subespacios del Ejemplo 26.

Solución: Un sistema generador del subespacio suma viene dado al juntar las bases de S_1 y S_2:

$$S_1 + S_2 = \langle(1,3,2,-8), (0,1,-1,3), (1,5,0,-2)\rangle.$$

Para dar una base de la suma, es necesario estudiar la dependencia lineal de los vectores anteriores:

$$\begin{pmatrix} 1 & 3 & 2 & -8 \\ 0 & 1 & -1 & 3 \\ 1 & 5 & 0 & -2 \end{pmatrix} \xrightarrow{P_{31}(-1)} \begin{pmatrix} 1 & 3 & 2 & -8 \\ 0 & 1 & -1 & 3 \\ 0 & 2 & -3 & 6 \end{pmatrix} \xrightarrow{P_{32}(-2)} \begin{pmatrix} 1 & 3 & 2 & -8 \\ 0 & 1 & -1 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Se sigue que $\dim(S_1 + S_2) = 2$ y $\mathcal{B}_{S_1+S_2} = \{(1,3,2,-8), (0,1,-1,3)\}$.

◇

Se presenta un resultado importante del Álgebra Lineal que relaciona las dimensiones de dos subespacios vectoriales cualesquiera, con las del subespacio suma y el subespacio intersección.

Teorema 5.25 (Fórmula de la dimensión)**.** Sean S y T dos subespacios vectoriales de un espacio vectorial real V de dimensión finita. Entonces se cumple:

$$\boxed{\dim(S+T) = \dim S + \dim T - \dim(S \cap T)}$$

Este teorema, válido únicamente para dos subespacios, permite conocer la dimensión del subespacio suma (o intersección) sin necesidad de hallarlo, conociendo las dimensiones de cada subespacio y la del subespacio intersección (o suma).

✓ **Ejemplo 28.** Utilizando los datos del Ejemplo 26 calcular la dimensión del subespacio suma $S_1 + S_2$ sin hallarlo.

Solución: Se sabe que $\dim S_1 = 2$, $\dim S_2 = 1$, $\dim(S_1 \cap S_2) = 1$. Aplicando la fórmula de la dimensión, se deduce la dimensión del subespacio suma:

$$\dim(S_1 + S_2) = \dim S_1 + \dim S_2 - \dim(S_1 \cap S_2) = 2 + 1 - 1 = 2.$$

◇

5.6.3. SUMA DIRECTA

Definición 5.26. Sea V un espacio vectorial real de dimensión finita y sean $S_1, \ldots, S_r$ subespacios de V. La suma $S_1 + \cdots + S_r$ se dice **suma directa**, y se escribe $S_1 \oplus \cdots \oplus S_r$, si cada vector $\mathbf{v} \in S_1 + \cdots + S_r$ se escribe *de manera única* como suma de elementos de cada subespacio

$$\mathbf{v} = \mathbf{s}_1 + \cdots + \mathbf{s}_r,$$

con $\mathbf{s}_i \in S_i$, $i = 1, \ldots, r$.

✓ **Ejemplo 29.** Se consideran los subespacios: $S_1 = \{(0, x_2, x_3) \mid x_2, x_3 \in \mathbb{R}\}$ y $S_2 = \{(x_1, 0, x_3) \mid x_1, x_3 \in \mathbb{R}\}$ de $\mathbb{R}^3$. Comprobar que $S_1 + S_2$ no es suma directa.

Solución: Se comprueba fácilmente que el subespacio suma viene dado por:

$$S_1 + S_2 = \{(x_1, x_2, x_3) \mid x_1, x_2, x_3 \in \mathbb{R}\} = \mathbb{R}^3.$$

El vector $\mathbf{v} = (1, 1, 0) \in S_1 + S_2$, y se puede expresar de varios modos como suma de un vector de S_1 y otro de S_2:

$$\mathbf{v} = (1,1,0) = \underbrace{(0,1,0)}_{\in S_1} + \underbrace{(1,0,0)}_{\in S_2},$$
$$\mathbf{v} = (1,1,0) = \underbrace{(0,1,1)}_{\in S_1} + \underbrace{(1,0,-1)}_{\in S_2},$$

por tanto, la suma $S_1 + S_2$ **no** es directa.

◇

La comprobación de que una suma es directa aplicando la definición anterior no siempre es fácil. En la práctica las propiedades siguientes pueden ayudar a realizar dicha comprobación.

Propiedades. Sean $S_1, \dots, S_r$ subespacios vectoriales del espacio vectorial V.

- La suma $S_1 + \cdots + S_r$ es directa si y solo si el único modo de escribir

  $$\mathbf{0} = \mathbf{s}_1 + \cdots + \mathbf{s}_r \quad \text{con} \quad \mathbf{s}_i \in S_i,\ i = 1, \dots, r$$

  es con todos los sumandos nulos, es decir, $\mathbf{s}_1 = \cdots = \mathbf{s}_r = \mathbf{0}$.
- La suma de dos subespacios S_1, S_2 de V es directa si y solo si $S_1 \cap S_2 = \{\mathbf{0}\}$.
- La suma $S_1 + \cdots + S_r,\ r > 2$, es directa si y solo si
  - $L = S_1 + \cdots + S_{r-1}$ es directa.
  - $L + S_r$ es directa, es decir $L \cap S_r = \{\mathbf{0}\}$.

✓ **Ejemplo 30.** Es fácil ver que los subespacios del Ejemplo 29 verifican

$$S_1 \cap S_2 = \{(0, 0, x_3) \mid x_3 \in \mathbb{R}\}.$$

Como $S_1 \cap S_2 \neq \{\mathbf{0}\}$, la suma de ambos no es directa.

Definición 5.27. Sea V un espacio vectorial real de dimensión finita. Sean S y T dos subespacios de V, verificando $S \oplus T = V$, entonces se dice que S y T son subespacios vectoriales **suplementarios**.

✓ **Ejemplo 31.** Se consideran en $\mathbb{R}^3$ los subespacios vectoriales

$$S_1 = \{(0, x_2, x_3) \mid x_2, x_3 \in \mathbb{R}\},\ S_2 = \{(x_1, 0, x_3) \mid x_1, x_3 \in \mathbb{R}\},\ S_3 = \{(x_1, 0, 0) \mid x_1 \in \mathbb{R}\}.$$

Se comprueba fácilmente que:

- $S_1 \cap S_2 = \{(0, 0, x_3) \mid x_3 \in \mathbb{R}\}$.
- $S_1 \cap S_3 = \{(0, 0, 0)\}$.
- $S_2 \cap S_3 = \{(x_1, 0, 0) \mid x_1 \in \mathbb{R}\}$.

Luego la suma $S_1 + S_3$ es directa, mientras que las sumas $S_1 + S_2$ y $S_2 + S_3$ no lo son. Además $S_1 \oplus S_3 = \mathbb{R}^3$ y por tanto S_1 y S_3 son suplementarios.

◇

5.6.4. UNIÓN DE SUBESPACIOS

Sean $S_1, \ldots, S_r$ subespacios vectoriales del espacio vectorial real V, entonces se define la **unión**, denotada por $S_1 \cup \cdots \cup S_r$, como:

$$S_1 \cup \cdots \cup S_r = \{\mathbf{s} \mid \mathbf{s} \in S_i, \text{ para algún } i\}.$$

Observación. La unión de subespacios vectoriales $S_1 \cup \cdots \cup S_r$ no es en general un subespacio vectorial, como se ve en el siguiente ejemplo:

✓ **Ejemplo 32.** Se considera el espacio vectorial $\mathbb{R}^2$ con las operaciones habituales y los subespacios de $\mathbb{R}^2$:

$$S_1 = \{(x, 0) \mid x \in \mathbb{R}\}, \quad S_2 = \{(0, y) \mid y \in \mathbb{R}\}.$$

Se observa que S_1 coincide con el eje OX y S_2 con el eje OY. El conjunto unión

$$S_1 \cup S_2 = \{(x, 0), (0, y) \mid x, y \in \mathbb{R}\}$$

viene dado por los ejes coordenados. $S_1 \cup S_2$ no es un subespacio vectorial ya que si se consideran dos elementos de dicho conjunto, por ejemplo, $\mathbf{s}_1 = (1, 0)$ y $\mathbf{s}_2 = (0, 1)$, el elemento suma $\mathbf{s}_1 + \mathbf{s}_2 = (1, 1)$ no pertenece al conjunto unión, porque no es un punto de ninguno de los dos ejes.

◇

5.7. COORDENADAS RESPECTO DE UNA BASE

Sea V un espacio vectorial real de dimensión finita n y $\mathcal{B}$ una base de V. Como consecuencia del Teorema 5.16 todo elemento de V se escribe de *manera única* como combinación lineal de los elementos de $\mathcal{B}$, lo que da pie a la siguiente definición.

Definición 5.28. Sea V un espacio vectorial real de dimensión finita, $\dim V = n$, y $\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ una base de V. Dado un vector $\mathbf{v} \in V$ se llaman **coordenadas** de $\mathbf{v}$ en la base $\mathcal{B}$ a la n-tupla $\mathbf{v}_{\mathcal{B}} = (\lambda_1, \ldots, \lambda_n) \in \mathbb{R}^n$ que verifica

$$\mathbf{v} = \lambda_1 \mathbf{u}_1 + \ldots + \lambda_n \mathbf{u}_n. \tag{5.2}$$

Observación. En el caso particular de trabajar con bases canónicas, las coordenadas de un vector son inmediatas de obtener:

- En $\mathbb{R}^n$, las coordenadas coinciden con el vector.
- En $\mathbb{R}_n[x]$, las coordenadas son los coeficientes del polinomio.
- En $\mathrm{Mat}_{n\times m}(\mathbb{R})$, las coordenadas son los elementos de la matriz.

Fijada una base, las coordenadas de un vector son únicas, sin embargo si se consideran bases distintas las coordenadas cambian. Posteriormente se verá la relación entre las coordenadas de un vector en distintas bases.

✓ **Ejemplo 33.** Hallar las coordenadas del vector $\mathbf{v} = (1, 2, 3) \in \mathbb{R}^3$ respecto de las bases $\mathcal{B}_1 = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ y $\mathcal{B}_2 = \{(1, 1, 1), (0, 1, 0), (0, 0, 2)\}$. Hallar el vector $\mathbf{w}$ de $\mathbb{R}^3$ cuyas coordenadas respecto de la base $\mathcal{B}_2$ son $\mathbf{w}_{\mathcal{B}_2} = (-1, 0, 1)$.

Solución: Por ser $\mathcal{B}_1$ la base canónica, las coordenadas de $\mathbf{v}$ en $\mathcal{B}_1$ son $\mathbf{v}_{\mathcal{B}_1} = (1, 2, 3)$. Para hallar las coordenadas de $\mathbf{v}$ en $\mathcal{B}_2$ hay que resolver el sistema

$$(1, 2, 3) = \alpha(1, 1, 1) + \beta(0, 1, 0) + \lambda(0, 0, 2), \quad \Longrightarrow \quad \begin{cases} \alpha = 1 \\ \alpha + \beta = 2 \\ \alpha + 2\lambda = 3 \end{cases} \quad \Longrightarrow \quad \begin{cases} \alpha = 1 \\ \beta = 1 \\ \lambda = 1. \end{cases}$$

La solución proporciona las coordenadas de $\mathbf{v}$ en la base $\mathcal{B}_2$, $\mathbf{v}_{\mathcal{B}_2} = (1, 1, 1)$. Para construir el vector $\mathbf{w}$ basta multiplicar cada coordenada por el vector correspondiente de la base.

$$\mathbf{w} = -1\,(1, 1, 1) + 0\,(0, 1, 0) + 1\,(0, 0, 2) = (-1, -1, 1).$$

◇

✓ **Ejemplo 34.** Hallar las coordenadas del polinomio $p(x) = 1 + 2x + 3x^2 \in \mathbb{R}_2[x]$ respecto de las bases $\mathcal{B}_1 = \{1, x, x^2\}$ y $\mathcal{B}_2 = \{1 + x + x^2, x, 2x^2\}$.

Solución: Por ser $\mathcal{B}_1$ la base canónica de $\mathbb{R}_2[x]$, las coordenadas de $p(x)$ en $\mathcal{B}_1$ son $p(x)_{\mathcal{B}_1} = (1, 2, 3)$. Para hallar las coordenadas de $p(x)$ en $\mathcal{B}_2$ hay que resolver:

$$1+2x+3x^2 = \alpha(1+x+x^2)+\beta(x)+\lambda(2x^2), \quad \Longrightarrow \quad \begin{cases} \alpha = 1 \\ \alpha + \beta = 2 \\ \alpha + 2\lambda = 3 \end{cases} \quad \Longrightarrow \quad \begin{cases} \alpha = 1 \\ \beta = 1 \\ \lambda = 1. \end{cases}$$

La solución proporciona las coordenadas de $p(x)$ en la base $\mathcal{B}_2$, $p(x)_{\mathcal{B}_2} = (1, 1, 1)$.

◇

✓ **Ejemplo 35.** Sea el subespacio $S = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1 + 3x_2 - 2x_3 = 0\}$ de $\mathbb{R}^3$, y $\mathcal{B}_S = \{(-3, 1, 0), (2, 0, 1)\}$ una base de S. Calcular las coordenadas del vector $\mathbf{v} = (1, -5, -7) \in S$.

Solución: Para hallar las coordenadas de $\mathbf{v}$ respecto de la base $\mathcal{B}_S$ basta encontrar $\lambda_1, \lambda_2 \in \mathbb{R}$ tales que

$$\lambda_1(-3, 1, 0) + \lambda_2(2, 0, 1) = (1, -5, -7) \quad \Longrightarrow \quad \begin{cases} -3\lambda_1 + 2\lambda_2 = 1 \\ \lambda_1 = -5 \\ \lambda_2 = -7. \end{cases}$$

De donde se sigue que $\mathbf{v}_{\mathcal{B}} = (-5, -7) \in \mathbb{R}^2$. Se observa que las coordenadas son una 2-tupla, ya que dim $S = 2$.

◇

La siguiente definición permite relacionar cualquier espacio vectorial de dimensión n con el espacio $\mathbb{R}^n$.

Definición 5.29. Considerar un espacio vectorial real de dimensión finita n $(V, +, \cdot)$ y $\mathbb{R}^n$ con sus operaciones habituales. Dada una base $\mathcal{B}$ de V, la aplicación

$$\begin{array}{rcl} \mathcal{C}_{\mathcal{B}} : V & \longrightarrow & \mathbb{R}^n \\ \mathbf{v} & \mapsto & \mathcal{C}_{\mathcal{B}}(\mathbf{v}) = \mathbf{v}_{\mathcal{B}} = (\lambda_1, \dots, \lambda_n) \end{array}$$

que a cada vector $\mathbf{v}$ de V le asocia sus coordenadas en la base $\mathcal{B}$, se llama **aplicación coordenada** relativa a la base $\mathcal{B}$.

La aplicación coordenada posee las siguientes propiedades:

Propiedades.

1. $\mathcal{C}_{\mathcal{B}}$ es biyectiva.
2. $\forall \mathbf{u}, \mathbf{v} \in V \implies \mathcal{C}_{\mathcal{B}}(\mathbf{u} + \mathbf{v}) = \mathcal{C}_{\mathcal{B}}(\mathbf{u}) + \mathcal{C}_{\mathcal{B}}(\mathbf{v})$.
3. $\forall \lambda \in \mathbb{R},\ \mathbf{v} \in V \implies \mathcal{C}_{\mathcal{B}}(\lambda \mathbf{v}) = \lambda\, \mathcal{C}_{\mathcal{B}}(\mathbf{v})$.

Como consecuencia de las propiedades anteriores se verifica que:

- Las coordenadas del vector nulo es la n-tupla nula sea cual sea la base elegida, $\mathbf{0}_{\mathcal{B}} = (0, \dots, 0) \in \mathbb{R}^n$.
- La aplicación coordenada $\mathcal{C}_{\mathcal{B}}$ conserva las combinaciones lineales, el carácter ligado o libre de una familia de vectores y por tanto las dimensiones de los subespacios. En particular dada una familia de vectores de V:

$$\{\mathbf{v}_1, \dots, \mathbf{v}_r\} \text{ es libre} \iff \{(\mathbf{v}_1)_{\mathcal{B}}, \dots, (\mathbf{v}_r)_{\mathcal{B}}\} \text{ es libre.}$$

A partir de ahora siempre que se tenga un espacio vectorial V de dimensión finita n, se puede identificar con $(\mathbb{R}^n, +, \cdot)$ con sus operaciones habituales, trabajando en coordenadas.

✓ **Ejemplo 36.** Probar, utilizando la aplicación coordenada, que el conjunto de vectores $\mathcal{C} = \{1, 1 + x, 1 + x^2\}$ es una base del espacio vectorial $\mathbb{R}_2[x]$.

Solución: Como $\dim \mathbb{R}_2[x] = 3$, aplicando la Proposición 5.21 basta comprobar que los vectores de $\mathcal{C}$ son linealmente independientes. Considerando la base de $\mathbb{R}_2[x]$, $\mathcal{B} = \{1, x, x^2\}$, las coordenadas de los vectores de $\mathcal{C}$ son:

$$1_{\mathcal{B}} = (1, 0, 0), \quad (1 + x)_{\mathcal{B}} = (1, 1, 0), \quad (1 + x^2)_{\mathcal{B}} = (1, 0, 1).$$

Aplicando el método de Gauss a la matriz que tiene por filas las coordenadas anteriores se observa que los vectores son independientes, y por tanto los polinomios de $\mathcal{C}$ son una base de $\mathbb{R}_2[x]$.

◇

5.8. CAMBIO DE COORDENADAS

Como ya se ha dicho anteriormente las coordenadas de un vector no son únicas, puesto que dependen de la base escogida. Se presenta a continuación la relación de las coordenadas de un vector en distintas bases.

Sea V un espacio vectorial real de dimensión finita n. Se consideran las bases $\mathcal{B}_1 = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ y $\mathcal{B}_2 = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ de V. Entonces, dado un vector $\mathbf{v} \in V$ existen $\lambda_i, \beta_i \in \mathbb{R}$, $i = 1, \dots, n$, tales que:

$$\mathbf{v}_{\mathcal{B}_1} = (\lambda_1, \dots, \lambda_n) \quad \text{y} \quad \mathbf{v}_{\mathcal{B}_2} = (\beta_1, \dots, \beta_n),$$

es decir,

$$\mathbf{v} = \lambda_1\mathbf{u}_1 + \lambda_2\mathbf{u}_2 + \ldots + \lambda_n\mathbf{u}_n = \begin{pmatrix} \lambda_1 & \cdots & \lambda_n \end{pmatrix} \left(\begin{array}{c} \mathbf{u}_1 \\ \hline \vdots \\ \hline \mathbf{u}_n \end{array}\right),$$

$$\mathbf{v} = \beta_1\mathbf{v}_1 + \beta_2\mathbf{v}_2 + \ldots + \beta_n\mathbf{v}_n = \begin{pmatrix} \beta_1 & \cdots & \beta_n \end{pmatrix} \left(\begin{array}{c} \mathbf{v}_1 \\ \hline \vdots \\ \hline \mathbf{v}_n \end{array}\right).$$

Se pretende estudiar la relación que existe entre las coordenadas del vector $\mathbf{v}$ en ambas bases. Para ello, se empieza relacionando las dos bases, por ejemplo expresando los vectores de la base $\mathcal{B}_2$ como combinación lineal de los vectores de la base $\mathcal{B}_1$.

$$\begin{cases} \mathbf{v}_1 = p_{11}\mathbf{u}_1 + \cdots + p_{n1}\mathbf{u}_n, \\ \mathbf{v}_2 = p_{12}\mathbf{u}_1 + \cdots + p_{n2}\mathbf{u}_n, \\ \vdots \\ \mathbf{v}_n = p_{1n}\mathbf{u}_1 + \cdots + p_{nn}\mathbf{u}_n. \end{cases}$$

Es decir, $\mathbf{v}_j = p_{1j}\mathbf{u}_1 + \cdots + p_{nj}\mathbf{u}_n$, $\forall j = 1, \dots, n$, o equivalentemente trabajando en coordenadas, $(\mathbf{v}_j)_{\mathcal{B}_1} = (p_{1j}, \dots, p_{nj}) \in \mathbb{R}^n$, $j = 1, \dots, n$, que en forma matricial se puede expresar como:

$$\underbrace{\left(\begin{array}{c} \mathbf{v}_1 \\ \hline \vdots \\ \hline \mathbf{v}_n \end{array}\right)}_{2^{\text{a}}\text{ base}} = \underbrace{\begin{pmatrix} p_{11} & \dots & p_{n1} \\ \vdots & \ddots & \vdots \\ p_{1n} & \dots & p_{nn} \end{pmatrix}}_{Q} \underbrace{\left(\begin{array}{c} \mathbf{u}_1 \\ \hline \vdots \\ \hline \mathbf{u}_n \end{array}\right)}_{1^{\text{a}}\text{ base}}.$$

Nota. Observar que en las matrices anteriores $\mathbf{v}_j$ y $\mathbf{u}_i$ son vectores y p_{ij} coordenadas. Se está utilizando notación de matrices por bloques, donde cada vector es un bloque fila.

La matriz Q anterior relaciona bases, pero el objetivo es relacionar coordenadas. Se sabe que

$$\mathbf{v} = \begin{pmatrix}\lambda_1 & \cdots & \lambda_n\end{pmatrix}\begin{pmatrix}\mathbf{u}_1\\ \hline \vdots \\ \hline \mathbf{u}_n\end{pmatrix} = \begin{pmatrix}\beta_1 & \cdots & \beta_n\end{pmatrix}\begin{pmatrix}\mathbf{v}_1\\ \hline \vdots \\ \hline \mathbf{v}_n\end{pmatrix} = \begin{pmatrix}\beta_1 & \cdots & \beta_n\end{pmatrix} Q \begin{pmatrix}\mathbf{u}_1\\ \hline \vdots \\ \hline \mathbf{u}_n\end{pmatrix}.$$

Como las coordenadas de un vector respecto de una base, en este caso $\mathcal{B}_1$, son únicas se tiene que

$$\begin{pmatrix}\lambda_1 & \cdots & \lambda_n\end{pmatrix} = \begin{pmatrix}\beta_1 & \cdots & \beta_n\end{pmatrix} Q$$

o trasponiendo la expresión

$$\begin{pmatrix}\lambda_1 & \cdots & \lambda_n\end{pmatrix}^t = \left(\begin{pmatrix}\beta_1 & \cdots & \beta_n\end{pmatrix} Q\right)^t = Q^t \begin{pmatrix}\beta_1 & \cdots & \beta_n\end{pmatrix}^t$$

y llamando $P = Q^t$

$$\begin{pmatrix}\lambda_1\\ \vdots \\ \lambda_n\end{pmatrix} = P \begin{pmatrix}\beta_1\\ \vdots \\ \beta_n\end{pmatrix}.$$

Escribiendo los valores de la matriz P se observa que las columnas de dicha matriz son las coordenadas de los vectores de la base $\mathcal{B}_2$ con respecto a la base $\mathcal{B}_1$:

$$\underbrace{\begin{pmatrix}\lambda_1\\ \vdots \\ \lambda_n\end{pmatrix}}_{\mathbf{v}_{\mathcal{B}_1}} = \underbrace{\left(\begin{array}{c|c|c} \overbrace{p_{11}}^{(\mathbf{v}_1)_{\mathcal{B}_1}} & \cdots & \overbrace{p_{1n}}^{(\mathbf{v}_n)_{\mathcal{B}_1}} \\ \vdots & \ddots & \vdots \\ p_{n1} & \cdots & p_{nn}\end{array}\right)}_{P} \underbrace{\begin{pmatrix}\beta_1\\ \vdots \\ \beta_n\end{pmatrix}}_{\mathbf{v}_{\mathcal{B}_2}} \implies \boxed{\mathbf{v}_{\mathcal{B}_1} = P \cdot \mathbf{v}_{\mathcal{B}_2}}$$

La matriz P se llama **matriz de cambio de coordenadas** y se puede denotar por $P_{\mathcal{B}_2,\mathcal{B}_1}$, donde se hace referencia a que la base $\mathcal{B}_2$ se ha expresado en función de $\mathcal{B}_1$. Se trata de una matriz regular, cuya inversa es $P^{-1}_{\mathcal{B}_2,\mathcal{B}_1} = P_{\mathcal{B}_1,\mathcal{B}_2}$, es decir, es la matriz de cambio de coordenadas tomando las bases en orden cambiado, y además verifica

$$P^{-1} \cdot \mathbf{v}_{\mathcal{B}_1} = \mathbf{v}_{\mathcal{B}_2}.$$

La relación entre $\mathbf{v}_{\mathcal{B}_1}$ y $\mathbf{v}_{\mathcal{B}_2}$ se hace mediante P o P^{-1}, calculando en cada ejercicio la matriz que resulte más sencilla, o bien resolviendo un sistema de ecuaciones.

En páginas anteriores se propuso el siguiente ejercicio que ahora se va a resolver usando la matriz del cambio de coordenadas.

✓ **Ejemplo 37.** Hallar las coordenadas del vector $\mathbf{v} = (1, 2, 3) \in \mathbb{R}^3$ respecto de las bases $\mathcal{B}_1 = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ y $\mathcal{B}_2 = \{(1, 1, 1), (0, 1, 0), (0, 0, 2)\}$.

Solución: Es fácil observar que las coordenadas de $\mathbf{v}$ en la base $\mathcal{B}_1$ coinciden con sus componentes por ser la base canónica, es decir $\mathbf{v}_{\mathcal{B}_1} = (1, 2, 3)$, y por lo mismo es más sencillo poner la segunda base en función de la primera, de modo que la matriz del cambio de coordenadas se construye sin más que escribir los vectores de $\mathcal{B}_2$ en columnas:

$$P_{\mathcal{B}_2,\mathcal{B}_1} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 2 \end{pmatrix}.$$

Para hallar $\mathbf{v}_{\mathcal{B}_2}$ se puede resolver un sistema de ecuaciones

$$\mathbf{v}_{\mathcal{B}_1} = P \cdot \mathbf{v}_{\mathcal{B}_2},$$

o bien calcular una inversa:

$$\mathbf{v}_{\mathcal{B}_2} = P^{-1} \cdot \mathbf{v}_{\mathcal{B}_1},$$

obteniéndose $\mathbf{v}_{\mathcal{B}_2} = (1,1,1)$.

◇

Observación. Cuando el espacio vectorial $V = \mathbb{R}^n$ se puede seguir el siguiente esquema para calcular la matriz de cambio de coordenadas, que corresponde a la resolución simultánea de n sistemas de ecuaciones con n incógnitas:

$$\boxed{(\,\mathcal{B}_1 \,|\, \mathcal{B}_2\,) \xrightarrow{\text{Op. elementales}} (\,I_n \,|\, P_{\mathcal{B}_2,\mathcal{B}_1}\,)}$$

Esto es válido en general cuando V es un espacio vectorial cualquiera. Basta entonces considerar las coordenadas de los vectores de las bases $\mathcal{B}_1$ y $\mathcal{B}_2$ en la base canónica de V.

✓ **Ejemplo 38.** Dadas las bases de $\mathbb{R}^3$

$$\mathcal{B}_1 = \{(1,1,0),(1,0,1),(0,1,1)\} \quad \text{y} \quad \mathcal{B}_2 = \{(1,-1,0),(1,0,-1),(1,0,0)\},$$

calcular la matriz de cambio $P_{\mathcal{B}_2,\mathcal{B}_1}$.

Solución: Aplicando la observación anterior se resuelven varios sistemas de ecuaciones simultáneamente usando el método de Gauss:

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

Se obtienen los vectores de la base $\mathcal{B}_2$ en función de los de la base $\mathcal{B}_1$, de modo que la matriz de cambio de coordenadas es

$$P_{\mathcal{B}_2,\mathcal{B}_1} = \begin{pmatrix} 0 & 1 & 1/2 \\ 1 & 0 & 1/2 \\ -1 & -1 & -1/2 \end{pmatrix}.$$

◇

5.9. APLICACIONES: ESPACIOS DE FILAS Y COLUMNAS DE UNA MATRIZ

Se proponen en esta sección algunos ejemplos de subespacios vectoriales que son de utilidad en el desarrollo del texto. Se considera una matriz $A \in \mathrm{Mat}_{n\times m}(\mathbb{R})$:

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

Recordando la descomposición en bloques, la matriz A se puede separar en filas

$$A = \left(\begin{array}{cccc} a_{11} & a_{12} & \cdots & a_{1m} \\ \hline a_{21} & a_{22} & \cdots & a_{2m} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline a_{n1} & a_{n2} & \cdots & a_{nm} \end{array}\right) = \left(\begin{array}{c} \mathbf{l}_1 \\ \hline \mathbf{l}_2 \\ \hline \vdots \\ \hline \mathbf{l}_n \end{array}\right)$$

donde cada fila $\mathbf{l}_i = \begin{pmatrix} a_{i1} & a_{i2} & \cdots & a_{im} \end{pmatrix} \in \mathbb{R}^m$. O bien en columnas:

$$A = \left(\begin{array}{c|c|c|c} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{array}\right) = \left(\mathbf{c}_1|\mathbf{c}_2|\cdots|\mathbf{c}_m\right)$$

donde cada columna $\mathbf{c}_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix} \in \mathbb{R}^n$.

Definición 5.30. Sea la matriz $A \in \mathrm{Mat}_{n\times m}(\mathbb{R})$. Las filas de A generan un subespacio vectorial de $\mathbb{R}^m$ denominado **espacio de filas de** A

$$\text{e. filas } A = \langle \mathbf{l}_1, \mathbf{l}_2, \ldots, \mathbf{l}_n \rangle.$$

Las columnas de A generan un subespacio vectorial de $\mathbb{R}^n$ denominado **espacio de columnas de** A

$$\text{e. col. } A = \langle \mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_m \rangle.$$

Proposición 5.31. Dada la matriz $A \in \mathrm{Mat}_{n\times m}(\mathbb{R})$, si tras aplicar una secuencia de operaciones elementales por filas se obtiene la matriz $B \in \mathrm{Mat}_{n\times m}(\mathbb{R})$, entonces los espacios generados por las filas de A y B son iguales, es decir,

$$\text{e. filas } A = \text{e. filas } B.$$

Proposición 5.32. Dada la matriz $A \in \mathrm{Mat}_{n\times m}(\mathbb{R})$, la dimensión del espacio de filas de la matriz A es el rango de dicha matriz A. La dimensión del espacio de columnas de la matriz A es también el rango de dicha matriz A.

Observar que los espacios generados por las filas y las columnas de una matriz A tienen la misma dimensión, aunque son subespacios de espacios vectoriales distintos (el de filas es subespacio de $\mathbb{R}^m$ y el de columnas de $\mathbb{R}^n$).

Observar también que tras aplicar operaciones elementales por filas solo se mantiene el espacio de filas y no el de columnas, por tanto, para dar una base del espacio de filas basta tomar las filas no nulas que se obtienen tras aplicar el método de Gauss a la matriz A. Este procedimiento es útil también para encontrar sistemas generadores y bases escalonadas. Si lo que se quiere dar es una base del espacio de columnas se toman tantas columnas linealmente independientes de la matriz inicial A como indique $\operatorname{rg} A$.

✓ **Ejemplo 39.** Calcular una base y la dimensión del espacio de filas y del espacio de columnas de

$$A = \begin{pmatrix} 0 & 2 & -6 & -2 & 4 \\ 0 & -1 & 3 & 3 & 2 \\ 0 & -1 & 3 & 7 & 10 \end{pmatrix}.$$

Solución: Se aplica el método de eliminación gaussiana para hallar $\operatorname{rg} A$:

$$A = \begin{pmatrix} 0 & 2 & -6 & -2 & 4 \\ 0 & -1 & 3 & 3 & 2 \\ 0 & -1 & 3 & 7 & 10 \end{pmatrix} \xrightarrow[P_1(\frac{1}{2})]{} \begin{pmatrix} 0 & 1 & -3 & -1 & 2 \\ 0 & -1 & 3 & 3 & 2 \\ 0 & -1 & 3 & 7 & 10 \end{pmatrix} \xrightarrow[\substack{P_{21}(1)\\ P_{31}(1)}]{}$$

$$\begin{pmatrix} 0 & 1 & -3 & -1 & 2 \\ 0 & 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 6 & 12 \end{pmatrix} \xrightarrow[P_2(\frac{1}{2})]{} \begin{pmatrix} 0 & 1 & -3 & -1 & 2 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 6 & 12 \end{pmatrix} \xrightarrow[P_{32}(-6)]{}$$

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

Como $\operatorname{rg} A = 2$, la dimensión de ambos espacios es 2. Una base del espacio de filas de A puede ser:

$$\mathcal{B}_{\text{filas}} = \{(0,1,-3,-1,2),(0,0,0,1,2)\}$$

o bien las correspondientes filas iniciales de A,

$$\mathcal{B}_{\text{filas}} = \{(0,2,-6,-2,4),(0,-1,3,3,2)\}.$$

La base del espacio de columnas de A la forman dos columnas linealmente independientes de A, por ejemplo:

$$\mathcal{B}_{\text{columnas}} = \left\{ \begin{pmatrix} 2 \\ -1 \\ -1 \end{pmatrix}, \begin{pmatrix} -2 \\ 3 \\ 7 \end{pmatrix} \right\}.$$

◇

5.10. COMANDOS DE *wxMaxima*

- Introducir un vector de $\mathbb{R}^n$ como una lista de n elementos:

```
-->     v:[v1,v2,....,vn];
```

- También se puede introducir un vector como una matriz fila o columna:

```
-->     v:matrix([v1,v2,...,vn]);
-->     v:matrix([v1],[v2],...,[vn]);
```

- Calcular el número de componentes de un vector $\mathbf{v}$:

```
-->     length(v);
```

- La suma de vectores se realiza con el comando "`+`".
- Para multiplicar un escalar por un vector se utiliza el comando "`*`".
- Calcular una base de un subespacio dado mediante ecuaciones implícitas, $S = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \mathbf{0}\}$:

```
-->     nullspace(A);
```

- Calcular una base del espacio de columnas de A:

```
-->     columnspace(A);
```

- Calcular una base del espacio de filas de A:

```
-->     columnspace(transpose(A));
```

- Para estudiar la dependencia o independencia lineal de una familia de vectores, hallar bases, dimensiones, comprobar si la suma de subespacios es directa, etc., se puede hacer uso de algunos comandos citados en otros capítulos: "`matrix`", "`submatrix`", "`addrow`", "`rank`", "`echelon`", "`triangularize`", y aquellos que se han aplicado anteriormente a la resolución de sistemas de ecuaciones.

5.11. EJERCICIOS PROPUESTOS

Espacios vectoriales

1. Comprobar que $\mathbb{R}^3$ con las operaciones habituales es un espacio vectorial real.

2. Comprobar que $\mathbb{R}^2$ con las siguientes operaciones es un espacio vectorial real:

$$(x, y) + (x', y') = (x + x', y + y' + 1), \qquad \lambda(x, y) = (\lambda\, x, \lambda\, y + \lambda - 1).$$

Comprobar que se cumplen las propiedades de la página 130.

3. Considerar $\mathbb{R}^3$ con la suma habitual. ¿Con cuáles de los siguientes productos por un escalar es $(\mathbb{R}^3, +, \cdot)$ un espacio vectorial real?

a) $\lambda(x, y, z) = (\lambda\, x, y, \lambda\, z)$.
b) $\lambda(x, y, z) = (0, 0, 0)$.
c) $\lambda(x, y, z) = (2\lambda\, x, 2\lambda\, y, 2\lambda\, z)$.

4. Demostrar que el conjunto de polinomios de grado menor o igual que 2 con coeficientes reales $\mathbb{R}_2[x]$ es un espacio vectorial sobre $\mathbb{R}$ con las operaciones habituales entre polinomios, mientras que el conjunto de polinomios de grado igual a 2 no es un espacio vectorial.

5. Determinar cuáles de los siguientes conjuntos, con las operaciones habituales entre matrices, son espacios vectoriales reales:

a) El conjunto de matrices cuadradas no regulares.
b) El conjunto de matrices cuadradas diagonales.

Subespacios vectoriales

6. Considerar el espacio vectorial $\mathbb{R}^2$ con las operaciones habituales. Determinar si los siguientes subconjuntos son subespacios vectoriales:

a) $S = \{(x, kx) \,|\, x \in \mathbb{R},\, k \in \mathbb{R}\}$.
b) $S = \{(x, y) \,|\, x - y = 0\}$.
c) $S = \{(x, y) \,|\, x = y - 3\}$.
d) $S = \{(x, y) \,|\, x < y\}$.
e) El conjunto de puntos de una recta que no pasa por el origen de coordenadas.

7. Considerar el espacio vectorial $\mathbb{R}^3$ con las operaciones habituales. Determinar si los siguientes subconjuntos son subespacios vectoriales:

a) $S = \{(x, y, 1) \,|\, x, y \in \mathbb{R}\}$.
b) $S = \{(x, x, y) \,|\, x, y \in \mathbb{R}\}$.
c) $S = \{(x, y, z) \,|\, x + 2y - z = 0\}$.
d) $S = \{(x, y, z) \,|\, x^2 + y^2 = z^2\}$.

8. Considerar el espacio vectorial $\mathbb{R}^4$ con las operaciones habituales. Determinar si los siguientes subconjuntos son subespacios vectoriales:

a) $S = \{(x, y, z, t) \,|\, y = z\}$.
b) $S = \{(x, y, z, t) \,|\, z = x - t\}$.
c) $S = \{(x, y, z, t) \,|\, 2x + y = 1\}$.

9. Considerar el espacio vectorial $\mathbb{R}^n$. Determinar si los siguientes subconjuntos son subespacios vectoriales:

a) $S = \{(x_1, \dots, x_n) \,|\, x_1 + \cdots + x_n = 0\}$.
b) $S = \{(x_1, \dots, x_n) \,|\, x_1 + \cdots + x_n = 1\}$.

10. Sea $V = \mathrm{Mat}_n(\mathbb{R})$ el espacio vectorial de las matrices reales de orden n. Determinar si los siguientes subconjuntos son subespacios vectoriales de V:

a) $S = \{A \in \mathrm{Mat}_n(\mathbb{R}) | A \,\text{es simétrica}\}$.
b) $S = \{A \in \mathrm{Mat}_n(\mathbb{R}) | A \,\text{es antisimétrica}\}$.
c) $S = \{A \in \mathrm{Mat}_n(\mathbb{R}) | \,\text{si}\, A = (a_{ij}),\, a_{ij} = 0 \,\forall i \neq j\}$.

11. Determinar si los siguientes conjuntos son subespacios vectoriales del espacio $\mathbb{R}_2[x]$:

a) $S = \{p(x) \in \mathbb{R}_2[x] \,|\, p(0) = 0\}$.
b) $S = \{p(x) \in \mathbb{R}_2[x] \,|\, p(x) = p(-x)\}$.
c) $S = \{p(x) \in \mathbb{R}_2[x] \,|\, p(2) = 1\}$.

Dependencia e independencia lineal. Generadores

12. Estudiar la dependencia o independencia lineal de las siguientes familias de vectores de $\mathbb{R}^n$. En caso de dependencia, encontrar una relación entre los vectores.

a) $\{(1, -1, 0),\, (3, 2, -1),\, (3, 5, -2)\} \subset \mathbb{R}^3$.
b) $\{(1, 1, 1),\, (1, -1, 1),\, (0, 1, 0)\} \subset \mathbb{R}^3$.
c) $\{(1, -1, 1, -1),\, (2, 0, 1, 0),\, (0, -2, 1, -2)\} \subset \mathbb{R}^4$.

13. Considerar el espacio vectorial $\mathbb{R}_2[x]$. Determinar cuáles de las siguientes familias de vectores son linealmente independientes:

a) $\{x^2 + 1,\, x + 1,\, x\}$.
b) $\{x^2 - x + 3,\, 2x^2 + x + 5,\, x^2 + 5x + 1\}$.

14. Considerar el espacio vectorial $\mathrm{Mat}_2(\mathbb{R})$. Estudiar la dependencia o independencia lineal de la siguiente familia:

$$\left\{ \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \right\}.$$

15. ¿Existe algún valor de γ para el cual sean linealmente dependientes los vectores de $\mathbb{R}^4$: $\mathbf{a} = (\gamma, -1, 0, 1)$, $\mathbf{b} = (0, \gamma, -1, 1)$ y $\mathbf{c} = (1, 0, -1, 2)$? En caso afirmativo encontrar la relación de dependencia entre los vectores $\mathbf{a}, \mathbf{b}$ y $\mathbf{c}$.

16. Determinar el valor de $\lambda \in \mathbb{R}$ para que $\mathbf{a} = (3, 1, -4, 6)$, $\mathbf{b} = (1, 1, 4, 4)$, $\mathbf{c} = (1, 0, 4, \lambda)$ sean vectores linealmente dependientes en $\mathbb{R}^4$.

17. En cada uno de los siguientes casos determinar si $\mathbf{w}_1$ y $\mathbf{w}_2$ pertenecen o no a la clausura lineal de $\{\mathbf{u}, \mathbf{v}\}$.

a) En $\mathbb{R}^3$: $\mathbf{w}_1 = (1, -1, 2)$, $\mathbf{w}_2 = (3, 1, -3)$, $\mathbf{u} = (1, 1, 1)$ y $\mathbf{v} = (0, 1, 3)$.

b) En $\mathbb{R}_2[x]$: $\mathbf{w}_1 = 3x^2 - 2x - 1$, $\mathbf{w}_2 = x$, $\mathbf{u} = x^2 + 1$ y $\mathbf{v} = x + 2$.

c) En $\mathrm{Mat}_2(\mathbb{R})$: $\mathbf{w}_1 = \begin{pmatrix} 1 & 3 \\ -1 & 1 \end{pmatrix}$, $\mathbf{w}_2 = \begin{pmatrix} 1 & -4 \\ 5 & 3 \end{pmatrix}$, $\mathbf{u} = \begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix}$ y $\mathbf{v} = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}$.

18. Determinar γ para que los vectores

$$\mathbf{a} = (\gamma, 8, 4), \quad \mathbf{b} = (-1, 2, 0), \quad \mathbf{c} = (0, 1, 2)$$

satisfagan la condición: $\langle \mathbf{a}, \mathbf{b}, \mathbf{c} \rangle = \langle \mathbf{b}, \mathbf{c} \rangle$.

19. Hallar γ y γ' para que el vector $\mathbf{a} = (1, 4, \gamma, \gamma')$ pertenezca al subespacio vectorial de $\mathbb{R}^4$ generado por los vectores $\mathbf{u} = (1, 2, -1, 2)$ y $\mathbf{v} = (0, 1, 2, 1)$.

20. Hallar γ y γ' para que los vectores $\mathbf{a} = (-1, 5, 4)$ y $\mathbf{b} = (\gamma, -2, -2)$ generen el mismo subespacio vectorial de $\mathbb{R}^3$ que los vectores $\mathbf{c} = (\gamma', 3, 2)$ y $\mathbf{d} = (5, 1, 0)$.

21. En el espacio vectorial $\mathbb{R}^3$ se consideran los siguientes conjuntos de vectores:

$$A = \{(1, 5, 1), (2, 1, 0)\}, \quad C = \{(1, 5, 1), (2, 1, 0), (1, 0, 0), (0, 0, 1)\}.$$

Probar que A es un conjunto linealmente independiente y que C es un sistema de generadores de $\mathbb{R}^3$.

22. Sea V un espacio vectorial real y consideremos las siguientes familias de vectores de V:

$$\mathcal{B}_1 = \{\mathbf{a}_1, \ldots, \mathbf{a}_n\}, \qquad \mathcal{B}_2 = \{\mathbf{a}_1, \mathbf{a}_1 + \mathbf{a}_2, \ldots, \mathbf{a}_1 + \mathbf{a}_2 + \cdots + \mathbf{a}_n\}.$$

Demostrar que:

a) Si $\mathcal{B}_1$ es una familia libre, entonces $\mathcal{B}_2$ es libre.

b) Si $\mathcal{B}_1$ es un sistema generador de V, entonces $\mathcal{B}_2$ es un sistema generador de V.

Bases y dimensión

23. Encontrar una base y la dimensión de cada uno de los siguientes subespacios:

a) $S = \{(a, b, c, d) \,|\, a + b = c - d\}$.

b) $T = \{(a, b, c) \,|\, 2a - b + c = 0,\ a + 2b - c = 0,\ a - 3b + 2c = 0\}$.

c) $U = \left\{ \begin{pmatrix} a & b \\ -a & b \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{R}) | \, a, b \in \mathbb{R} \right\}$.

d) $W = \{p(x) \in \mathbb{R}_2[x] \,|\, p(x) = p(-x)\}$.

24. Averiguar si los siguientes conjuntos son bases de $\mathbb{R}_3[x]$:

a) $\{1, 1+x, 1+x+x^2, 1+x+x^2+x^3\}$.

b) $\{1-x+x^2-x^3, 1+x+x^2+x^3, 2+3x+4x^2+5x^3, 1+x^2\}$.

25. En el espacio vectorial $\mathbb{R}^4$ se consideran las siguientes familias de vectores:

a) $\mathcal{F}_1 = \{(1,-1,2,1), (2,-1,2,1), (0,1,-1,-1)\}$.

b) $\mathcal{F}_2 = \{(2,-2,3,1), (-1,4,-6,-2), (1,2,-3,-1)\}$.

Completar cada familia F_i hasta obtener una base de $\mathbb{R}^4$.

26. En el espacio vectorial $\mathbb{R}^4$ se consideran los vectores

$$\mathbf{a} = (3,2,\gamma,5), \quad \mathbf{b} = (2,-3,5,\gamma), \quad \mathbf{c} = (0,13,\gamma',7).$$

Hallar γ y γ' para que la dimensión del subespacio generado por dichos vectores sea 2. Hallar una base de dicho subespacio.

27. En el espacio vectorial $\mathbb{R}^3$ se consideran los siguientes conjuntos de vectores:

$$A = \{(1,5,1), (2,1,0)\}, \quad C = \{(1,5,1), (2,1,0), (1,0,0), (0,0,1)\}.$$

Encontrar una base que contenga al conjunto A y esté contenida en el conjunto C.

28. Determinar el subconjunto S de $\mathbb{R}^3$ caracterizado por la siguiente condición:

$$(x,y,z) \in S \iff \text{ la matriz } \begin{pmatrix} x & y \\ z & x \end{pmatrix} \text{ conmuta con } \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Demostrar que S es un subespacio vectorial de $\mathbb{R}^3$, hallar una base y la dimensión de S y completarla hasta obtener una base de $\mathbb{R}^3$.

29. En el espacio vectorial $\mathbb{R}^3$ se consideran los vectores:

$$\mathbf{a} = (\gamma,1,-2), \quad \mathbf{b} = (1,\gamma,2), \quad \mathbf{c} = (2\gamma,1,0).$$

a) ¿Para qué valores de γ el conjunto $\{\mathbf{a},\mathbf{b},\mathbf{c}\}$ es base de $\mathbb{R}^3$?

b) ¿Existe algún valor de γ para el cual $\dim(\langle \mathbf{a},\mathbf{b},\mathbf{c}\rangle) = 1$?

c) Estudiar, en función de γ, los distintos valores que puede tener $\dim(\langle \mathbf{a},\mathbf{b}\rangle)$.

30. Considerar el espacio vectorial $\mathbb{R}^4$. Sea S el subespacio generado por la solución del sistema homogéneo:

$$\begin{cases} 2x_1 - x_2 - 3x_3 = 0 \\ x_1 - 2x_2 + 6x_3 - 6x_4 = 0. \end{cases}$$

a) Hallar la dimensión y una base $\mathcal{B}_S$ del subespacio S.

b) Completar $\mathcal{B}_S$ hasta una base de $\mathbb{R}^4$.

31. Considerar el espacio vectorial de las matrices reales de orden n, $\mathrm{Mat}_n(\mathbb{R})$. Determinar la dimensión de cada uno de los siguientes subespacios vectoriales:

a) El subespacio de las matrices diagonales.

b) El subespacio de las matrices triangulares superiores.

Coordenadas respecto de una base y cambio de coordenadas

32. Consideramos las siguientes bases del espacio vectorial $\mathbb{R}^2$:

$$\mathcal{B}_1 = \{(1,-2),\, (3,-4)\} \quad \text{y} \quad \mathcal{B}_2 = \{(1,3),\, (3,8)\}.$$

a) Encontrar las coordenadas de un vector arbitrario (a,b) relativas a la base $\mathcal{B}_1$.

b) Hallar una matriz de cambio de coordenadas entre las bases $\mathcal{B}_1$ y $\mathcal{B}_2$.

c) Encontrar las coordenadas de un vector arbitrario (a,b) relativas a la base $\mathcal{B}_2$.

33. Sea $\mathcal{B} = \{(1,0,1,3),\, (0,1,-1,0),\, (1,0,1,2),\, (-1,0,0,1)\}$. Probar que $\mathcal{B}$ es una base de $\mathbb{R}^4$. Calcular las coordenadas de los siguientes vectores respecto de la base $\mathcal{B}$:

$$\mathbf{u} = (1,0,1,0), \qquad \mathbf{v} = (1,1,1,1), \qquad \mathbf{w} = (1,-1,2,1).$$

34. Sean V un espacio vectorial real de dimensión 3 y $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ una base de V. Se consideran los vectores de V cuyas coordenadas respecto de $\mathcal{B}$ son:

a) $\mathbf{a}_1 = (2,3,5)$, $\mathbf{a}_2 = (3,7,8)$, $\mathbf{a}_3 = (1,-6,1)$, $\mathbf{b} = (7,-2,m)$.

b) $\mathbf{a}_1 = (4,4,3)$, $\mathbf{a}_2 = (7,2,1)$, $\mathbf{a}_3 = (4,1,6)$, $\mathbf{b} = (5,9,m)$.

c) $\mathbf{a}_1 = (3,4,2)$, $\mathbf{a}_2 = (6,8,7)$, $\mathbf{b} = (9,12,m)$.

En cada caso, hallar m para que el vector $\mathbf{b}$ sea combinación lineal de los vectores $\mathbf{a}_i$.

35. Dado el subespacio vectorial de $\mathbb{R}^3$, $S = \{(x_1, x_2, 0) \,|\, x_1, x_2 \in \mathbb{R}\}$, se consideran las bases $\mathcal{B}_1 = \{(1,0,0),\, (1,1,0)\}$ y $\mathcal{B}_2 = \{(2,3,0),\, (1,4,0)\}$ de S. Encontrar la matriz del cambio de coordenadas entre dichas bases. Dado el vector $\mathbf{v} = (1,2,0) \in S$, hallar sus coordenadas en dichas bases usando la matriz anterior.

36. Se consideran dos bases del espacio vectorial $\mathbb{R}^3$: $\mathcal{B} = \{(1,1,1),\, (1,1,0),\, (1,0,0)\}$, y $\mathcal{C}$ la base canónica. Se pide:

a) Hallar la matriz de cambio de coordenadas.

b) Calcular las coordenadas del vector $\mathbf{v} = (2,0,-3)$ en dichas bases utilizando la matriz anterior.

37. Se consideran dos bases del espacio vectorial $\mathbb{R}^3$:

$$\mathcal{B}_1 = \{(1,1,1),\, (0,1,1),\, (0,0,1)\}, \qquad \mathcal{B}_2 = \{(1,2,5),\, (0,1,0),\, (0,0,1)\}.$$

Hallar la matriz del cambio de coordenadas. Dado el vector $\mathbf{v} = (1,1,1)$, calcular las coordenadas de $\mathbf{v}$ en dichas bases utilizando la matriz anterior.

38. En el espacio vectorial $\mathrm{Mat}_2(\mathbb{R})$ se consideran las bases:

$$\mathcal{B}_1 = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\},$$

$$\mathcal{B}_2 = \left\{ \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix} \right\}.$$

Hallar las coordenadas del vector $\mathbf{v} = \begin{pmatrix} 1 & 2 \\ -1 & 0 \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{R})$ con respecto a las bases anteriores usando la matriz de cambio de coordenadas.

39. Hallar las coordenadas del vector $\mathbf{v} \in \mathbb{R}_2[x]$ con respecto a la base $\mathcal{B} = \{x^2,\, x+1,\, 3\}$, sabiendo que sus coordenadas en base canónica $\mathcal{C} = \{x^2,\, x,\, 1\}$ son $(2, 1, -1)$.

40. Sea V un espacio vectorial real con $\dim V = 3$ y $\mathcal{B} = \{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ una base de V. Se consideran los subespacios vectoriales S y T generados por los vectores que se indican:

$$\begin{aligned} S &= \langle \mathbf{u} + 2\mathbf{v} + 3\mathbf{w}, 3\mathbf{u} + 6\mathbf{v} + 7\mathbf{w} \rangle, \\ T &= \langle 2\mathbf{u} - 3\mathbf{v} + \mathbf{w}, 3\mathbf{u} - \mathbf{v} + 5\mathbf{w}, \mathbf{u} + 2\mathbf{v} + 4\mathbf{w} \rangle. \end{aligned}$$

Hallar la dimensión de cada uno de ellos, una base contenida en el sistema de generadores dado y expresar los restantes vectores que generan los subespacios como combinación lineal de los vectores de la base obtenida.

Operaciones con subespacios vectoriales

41. Se considera el espacio vectorial $\mathbb{R}^4$. En cada caso se consideran los subespacios vectoriales S y T generados por:

$$\text{a)} \left\{ \begin{array}{l} S = \langle (1,2,0,1), (1,1,1,0) \rangle \\ T = \langle (1,0,1,0), (1,3,0,0) \rangle \end{array} \right.$$

$$\text{b)} \left\{ \begin{array}{l} S = \langle (1,2,1,-2), (2,3,1,0), (1,2,2,-3) \rangle \\ T = \langle (1,1,1,1), (1,0,1,-1), (1,3,0,-4) \rangle \end{array} \right.$$

Hallar una base y la dimensión de $S + T$ y $S \cap T$.

42. Considerar en el espacio vectorial $\mathbb{R}^4$ los subespacios S y T:

$$S = \langle (1,-1,2,1), (0,1,-1,3), (2,0,1,-1) \rangle$$

$$T = \left\{ (x_1, x_2, x_3, x_4) \left| \begin{array}{r} 2x_1 - x_2 - 3x_3 = 0 \\ x_1 - 2x_2 + 6x_3 - 6x_4 = 0 \end{array} \right. \right\}$$

Calcular una base y la dimensión de S, T, $S + T$ y $S \cap T$.

43. En el espacio vectorial $\mathbb{R}^3$ se consideran los subespacios siguientes:

$$S = \{(x,y,z) \in \mathbb{R}^3 \mid y - z = 0\}, \quad T = \{(x,0,z) \in \mathbb{R}^3 \mid x + z = 0\}.$$

Probar que $\mathbb{R}^3 = S \oplus T$.

44. En $\mathbb{R}^4$ se consideran los siguientes subespacios vectoriales:

$$S = \{(x,y,z,t) \in \mathbb{R}^4 \mid x - y = z - t = 0\},$$

$$T = \{(x,y,z,t) \in \mathbb{R}^4 \mid x + y + z = x + 2y + t = 0\}.$$

a) ¿Es $S + T$ directa?

b) ¿Son S y T suplementarios?

45. Se consideran en $\mathbb{R}^4$ los subespacios vectoriales definidos por:

$$S = \left\{ (x_1, x_2, x_3, x_4) \left| \begin{array}{l} x_1 - x_3 = 0 \\ x_2 + x_4 = 0 \end{array} \right. \right\}, \qquad T = \langle (0,0,1,0) \rangle.$$

Calcular una base y la dimensión de S, T, $S + T$ y $S \cap T$. ¿Son S y T suplementarios?

46. En el espacio vectorial $\mathbb{R}_3[x]$ se consideran los subespacios:

$$S = \langle 1 + x^2, 1 - x^2 \rangle, \quad T = \langle 1 + 3x + 5x^2, -1 + 2x^2, 3x + 7x^2 \rangle.$$

Calcular una base y la dimensión de los subespacios $S + T$ y $S \cap T$. Con los resultados obtenidos razonar si S y T son suplementarios. ¿Se puede completar la base de $S + T$ hasta una base de $\mathbb{R}_3[x]$?

47. Considerar los siguientes subespacios vectoriales de $V = \mathrm{Mat}_2(\mathbb{R})$:

$$S = \left\{ \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{R}) \,\middle|\, a, b \in \mathbb{R} \right\}, \quad T = \left\{ A \in \mathrm{Mat}_2(\mathbb{R}) \mid A + A^t = \mathbf{0} \right\}.$$

Hallar una base y la dimensión de los subespacios S, T, $S + T$ y $S \cap T$. ¿Es la suma de los subespacios S y T directa? ¿Son S y T subespacios suplementarios? Si U es un subespacio de $\mathrm{Mat}_2(\mathbb{R})$ suplementario a S, ¿cuál es su dimensión?

48. Sea $V = \mathrm{Mat}_n(\mathbb{R})$. Sea S el subconjunto de V de las matrices simétricas y T el de las matrices antisimétricas. Se pide:

a) Obtener una base y la dimensión de S y T.

b) Demostrar que $V = S \oplus T$.

5.12. SOLUCIONES A LOS EJERCICIOS PROPUESTOS

Espacios vectoriales

1. Hay que comprobar que se cumplen las propiedades 1-8 de la Definición 5.8.

2. Hay que comprobar que se cumplen las propiedades 1-8. Merece la pena resaltar que el elemento neutro de la suma es $(0, -1)$ y el elemento opuesto de un elemento genérico (x, y) es $(-x, -y-2)$.

3. Ninguno de los 3 casos confiere a $\mathbb{R}^3$ con la suma habitual la estructura de espacio vectorial.

4. En la primera parte basta comprobar que se cumplen las propiedades 1-8. En la segunda parte es suficiente dar un contraejemplo para comprobar que la suma no es interna, es decir, encontrar dos elementos del conjunto cuya suma no esté en él.

5. a) El conjunto de matrices cuadradas no regulares no es un espacio vectorial real.

b) El conjunto de matrices cuadradas diagonales es un espacio vectorial sobre $\mathbb{R}$.

Subespacios vectoriales

6. Solamente son subespacios los conjuntos de los apartados a) y b).

7. Solamente son subespacios los conjuntos de los apartados b) y c).

8. Solamente son subespacios los conjuntos de los apartados a) y b).

9. Solamente es subespacio el conjunto del apartado a).

10. Los tres conjuntos son subespacios vectoriales de $\mathrm{Mat}_n(\mathbb{R})$.

11. Solamente son subespacios de $\mathbb{R}_2[x]$ los conjuntos de los apartados a) y b).

Dependencia e independencia lineal. Generadores

12. a) La familia es linealmente independiente.

b) La familia es ligada y una relación de dependencia es: $\mathbf{v}_3 = \frac{1}{2}\mathbf{v}_1 - \frac{1}{2}\mathbf{v}_2$.

c) La familia es ligada y una relación de dependencia es: $\mathbf{v}_3 = 2\mathbf{v}_1 - \mathbf{v}_2$.

13. Los vectores de la familia a) son linealmente independientes, mientras que los de la familia b) son linealmente dependientes.

14. La familia es linealmente dependiente.

15. La familia es linealmente dependiente si $\gamma = 1$. La relación de dependencia es $\mathbf{c} = \mathbf{a}+\mathbf{b}$.

16. Los vectores $\mathbf{a}, \mathbf{b}$ y $\mathbf{c}$ son linealmente independientes $\forall \lambda \in \mathbb{R}$.

17. a) El vector $\mathbf{w}_1 \notin \langle \mathbf{u}, \mathbf{v} \rangle$ mientras que $\mathbf{w}_2 \in \langle \mathbf{u}, \mathbf{v} \rangle$, siendo $\mathbf{w}_2 = 3\mathbf{u} - 2\mathbf{v}$ la relación de dependencia.

b) El vector $\mathbf{w}_1 \in \langle \mathbf{u}, \mathbf{v} \rangle$, y verifica $\mathbf{w}_1 = 3\mathbf{u} - 2\mathbf{v}$. El vector $\mathbf{w}_2 \notin \langle \mathbf{u}, \mathbf{v} \rangle$.

c) El vector $\mathbf{w}_1 \notin \langle \mathbf{u}, \mathbf{v} \rangle$ pero $\mathbf{w}_2 \in \langle \mathbf{u}, \mathbf{v} \rangle$, siendo la relación de dependencia $\mathbf{w}_2 = 3\mathbf{u} - \mathbf{v}$.

18. $\gamma = -3$. En ese caso, la relación de dependencia es $\mathbf{a} = 3\mathbf{b} + 2\mathbf{c}$.

19. El vector $\mathbf{a} \in \langle \mathbf{u}, \mathbf{v} \rangle$ si y solo si $\gamma = 3, \gamma' = 2$.

20. Tras imponer que $\mathbf{a}, \mathbf{b} \in \langle \mathbf{c}, \mathbf{d} \rangle$ y $\mathbf{c}, \mathbf{d} \in \langle \mathbf{a}, \mathbf{b} \rangle$ se obtiene: $\gamma = 3, \gamma' = 4$.

21. Basta aplicar la Proposición 5.21.

22. a) Se resuelve el sistema homogéneo que se plantea al igualar a cero una combinación lineal de los vectores de $\mathcal{B}_2$, imponiendo que los vectores de $\mathcal{B}_1$ son linealmente independientes.

b) Denominando $\mathbf{b}_i$ a los elementos de $\mathcal{B}_2$, basta cambiar cada vector $\mathbf{b}_i$, con $i = 2, \ldots, n$ por $\mathbf{b}_i - \mathbf{b}_{i-1}$.

Bases y dimensión

23. a) $\mathcal{B}_S = \{(1,0,0,-1), (0,1,0,-1), (0,0,1,1)\}$, $\dim S = 3$.

b) $\mathcal{B}_T = \{(-1,3,5)\}$, $\dim T = 1$.

c) $\mathcal{B}_U = \left\{ \begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \right\}$, $\dim U = 2$.

d) $\mathcal{B}_W = \{1, x^2\}$, $\dim W = 2$.

24. a) La familia es base ya que son 4 vectores linealmente independientes del espacio $\mathbb{R}_3[x]$ que tiene dimensión 4.

b) La familia no es base por ser linealmente dependiente.

25. Para completar la familia $\mathcal{F}_1$ basta añadir el vector $\{(0,0,0,1)\}$. Para completar la familia $\mathcal{F}_2$ se añaden los vectores $\{(0,0,1,0), (0,0,0,1)\}$.

26. $\gamma = 1, \gamma' = -13$. Para estos valores, los vectores $\mathbf{a}$ y $\mathbf{b}$ constituyen una base del subespacio y el vector $\mathbf{c}$ se puede expresar como $\mathbf{c} = 2\mathbf{a} - 3\mathbf{b}$.

27. Por ejemplo, $\mathcal{B} = \{(1,5,1), (2,1,0), (1,0,0)\}$.

28. $S = \{(x,y,y) \mid x, y \in \mathbb{R}\}$. $\mathcal{B}_S = \{(1,0,0), (0,1,1)\}$, $\dim S = 2$. Base de $\mathbb{R}^3$ completando la anterior: $\mathcal{B}_{\mathbb{R}^3} = \{(1,0,0), (0,1,1), (0,0,1)\}$.

29. a) Si $\gamma \neq -1$, $\gamma \neq \frac{1}{2}$, entonces los vectores dados son base.

b) No existe ningún valor de γ que verifique la condición.

c) Si $\gamma = -1$, $\dim(\langle \mathbf{a}, \mathbf{b} \rangle) = 1$; en otro caso, la dimensión es 2.

30. a) $\dim S = 2$, y una base de S puede ser $\mathcal{B}_S^1 = \{(4,5,1,0), (-2,-4,0,1)\}$ o bien $\mathcal{B}_S^2 = \{(2,4,0,-1), (0,-3,1,2)\}$.

b) Para completar la base $\mathcal{B}_S^1$ hasta una base del espacio basta con añadir los dos primeros vectores de la base canónica de $\mathbb{R}^4$, en el caso de $\mathcal{B}_S^2$ se pueden añadir los dos últimos vectores de la base canónica de $\mathbb{R}^4$.

31. a) La dimensión es n.

b) La dimensión es $\frac{n(n+1)}{2}$.

Coordenadas respecto de una base y cambio de coordenadas

32. a) $(a,b)_{\mathcal{B}_1} = (-2a - \frac{3}{2}b,\ a + \frac{1}{2}b)$.

b) $P_{\mathcal{B}_1,\mathcal{B}_2} = \begin{pmatrix} -14 & -36 \\ 5 & 13 \end{pmatrix}, \quad P_{\mathcal{B}_2,\mathcal{B}_1} = \begin{pmatrix} -13/2 & -18 \\ 5/2 & 7 \end{pmatrix}$.

c) $(a,b)_{\mathcal{B}_2} = (-8a + 3b,\ 3a - b)$.

33. $\mathbf{u}_{\mathcal{B}} = (-2,0,3,0), \quad \mathbf{v}_{\mathcal{B}} = (-4,1,6,1), \quad \mathbf{w}_{\mathcal{B}} = (-1,-1,2,0)$.

34. En el apartado a) $m = 15$. Los apartados b) y c) se verifican $\forall m \in \mathbb{R}$.

35. $P_{\mathcal{B}_2,\mathcal{B}_1} = \begin{pmatrix} -1 & -3 \\ 3 & 4 \end{pmatrix}, \quad P_{\mathcal{B}_1,\mathcal{B}_2} = \begin{pmatrix} 4/5 & 3/5 \\ -3/5 & -1/5 \end{pmatrix}, \quad \mathbf{v}_{\mathcal{B}_2} = \begin{pmatrix} 2/5 \\ 1/5 \end{pmatrix}, \quad \mathbf{v}_{\mathcal{B}_1} = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$.

36. $P_{\mathcal{B},\mathcal{C}} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad P_{\mathcal{C},\mathcal{B}} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 0 \end{pmatrix}, \quad \mathbf{v}_{\mathcal{C}} = \begin{pmatrix} 2 \\ 0 \\ -3 \end{pmatrix}, \quad \mathbf{v}_{\mathcal{B}} = \begin{pmatrix} -3 \\ 3 \\ 2 \end{pmatrix}$.

37. $P_{\mathcal{B}_1,\mathcal{B}_2} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -4 & 1 & 1 \end{pmatrix}, \quad P_{\mathcal{B}_2,\mathcal{B}_1} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 3 & -1 & 1 \end{pmatrix}, \quad \mathbf{v}_{\mathcal{B}_1} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{v}_{\mathcal{B}_2} = \begin{pmatrix} 1 \\ -1 \\ -4 \end{pmatrix}$.

38. Las matrices de cambio de coordenadas son

$$P_{\mathcal{B}_2,\mathcal{B}_1} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix} \qquad P_{\mathcal{B}_1,\mathcal{B}_2} = \begin{pmatrix} 1/2 & 1/2 & -1/2 & 1/2 \\ 1/2 & -1/2 & 1/2 & -1/2 \\ -1/2 & 1/2 & 1/2 & 1/2 \\ -1/2 & 1/2 & 1/2 & -1/2 \end{pmatrix},$$

y las coordenadas son: $\mathbf{v}_{\mathcal{B}_1} = (1,2,-1,0)$, $\mathbf{v}_{\mathcal{B}_2} = (2,-1,0,0)$.

39. $P_{\mathcal{B},\mathcal{C}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 3 \end{pmatrix}, \quad P_{\mathcal{C},\mathcal{B}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1/3 & 1/3 \end{pmatrix}, \quad \mathbf{v}_{\mathcal{B}} = \begin{pmatrix} 2 \\ 1 \\ -2/3 \end{pmatrix}$.

40. El sistema generador de S constituye una base, dim $S = 2$. El subespacio T tiene dimensión 2. Una base de T está formada por el primer y el tercer vector del sistema generador, siendo el segundo vector suma de ellos.

Operaciones con subespacios vectoriales

41. a) $\mathcal{B}_{S+T} = \{(1,2,0,1),(1,1,1,0),(1,0,1,0),(1,3,0,0)\} \Longrightarrow \dim(S+T) = 4$.
Otra posibilidad es: $\mathcal{B}_{S+T} = \{(1,2,0,1),(0,1,-1,1),(0,0,1,-1),(0,0,0,1)\}$.
$\dim(S \cap T) = 0$, por tanto, la base de $S \cap T$ es el conjunto vacío.

b) $\mathcal{B}_{S+T} = \{(1,2,1,-2),(2,3,1,0),(1,2,2,-3),(1,0,1,-1)\} \Longrightarrow \dim(S+T) = 4$.
Otra posibilidad es: $\mathcal{B}_{S+T} = \{(1,2,1,-2),(0,1,1,-4),(0,0,1,-1),(0,0,0,1)\}$.
$\mathcal{B}_{S\cap T} = \{(1,1,1,1),(1,3,0,-4)\} \Longrightarrow \dim(S \cap T) = 2$.
Otra posibilidad es: $\mathcal{B}_{S\cap T} = \{(1,1,1,1),(0,-2,1,5)\}$.

42. $\mathcal{B}_S = \{(1,-1,2,1),(0,1,-1,3),(0,0,1,9)\} \Longrightarrow \dim S = 3$.
$\mathcal{B}_T = \{(4,5,1,0),(0,-3,1,2)\} \Longrightarrow \dim T = 2$.
$\mathcal{B}_{S+T} = \{(1,-1,2,1),(0,1,-1,3),(0,0,1,9),(0,0,0,1)\} \Longrightarrow \dim\ (S+T) = 4$.
$\mathcal{B}_{S\cap T} = \{(58,-1,39,49)\} \Longrightarrow \dim(S \cap T) = 1$.

43. $\mathbb{R}^3 = S \oplus T$, ya que $\dim S = 2$, $\dim T = 1$, $\dim(S+T) = 3 = \dim \mathbb{R}^3$, $\dim(S \cap T) = 0$.

44. La suma $S+T$ es directa y los subespacios son suplementarios.

45. $\mathcal{B}_S = \{(1,0,1,0),(0,-1,0,1)\} \Longrightarrow \dim S = 2$.
$\mathcal{B}_T = \{(0,0,1,0)\} \Longrightarrow \dim T = 1$.
$S \cap T = \{\mathbf{0}\} \Longrightarrow \dim(S \cap T) = 0$.
$\mathcal{B}_{S+T} = \{(1,0,1,0),(0,-1,0,1),(0,0,1,0)\} \Longrightarrow \dim(S+T) = 3$.
La suma de los subespacios es directa pero no son suplementarios.

46. $\mathcal{B}_{S+T} = \{1+x^2, -2x^2, 3x+4x^2\} \Longrightarrow \dim(S+T) = 3$.
$\mathcal{B}_{S\cap T} = \{1-2x^2\} \Longrightarrow \dim(S \cap T) = 1$.
Los subespacios no son suplementarios. La base de $S+T$ se puede completar hasta una base de $\mathbb{R}_3[x]$ añadiendo, por ejemplo, el polinomio x^3.

47. $\mathcal{B}_S = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\} \Longrightarrow \dim S = 2$. $\mathcal{B}_T = \left\{ \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right\} \Longrightarrow \dim T = 1$.

$\mathcal{B}_{S+T} = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right\} \Longrightarrow \dim(S+T) = 3$.

$S \cap T = \{\mathbf{0}\} \Longrightarrow \dim(S \cap T) = 0$.

La suma de los subespacios es directa pero no son suplementarios. La dimensión de un subespacio suplementario a S es 2.

48. a) La dimensión del espacio de matrices simétricas es $\dim S = \frac{n(n+1)}{2}$. En el caso de las matrices antisimétricas es $\dim T = \frac{n(n-1)}{2}$.

b) Basta comprobar que $S \cap T = \left\{ \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \right\}$ y aplicar la fórmula de la dimensión.

CAPÍTULO 6

Espacios Euclídeos

En el Capítulo 5 se ha generalizado la estructura del espacio $\mathbb{R}^2$, de vectores libres del plano, hasta llegar a los espacios vectoriales reales. En ese proceso de generalización hay nociones como la de longitud (distancia) o ángulo entre vectores que no se han tenido en cuenta.

En este capítulo se abordan estas cuestiones, generalizando para ello la noción de producto escalar de $\mathbb{R}^2$. Se trata de una nueva operación que a cada par de vectores de un espacio vectorial cualquiera V les asocia un *escalar* (en el caso estudiado en el texto un número real), y que verifica ciertas propiedades.

Si un espacio vectorial se dota de un producto escalar, se pueden trasladar las nociones de longitud y ángulo entre vectores al terreno abstracto.

La denominación de *espacio euclídeo* para denotar un espacio vectorial en el que se ha definido un producto escalar se debe a que este espacio satisface los axiomas de Euclides de la geometría. Se generalizan con estos espacios las ya conocidas construcciones del plano real euclídeo o el espacio tridimensional de la geometría euclídea.

El término euclídeo se utiliza para distinguir estos espacios de los espacios curvos de las geometrías no euclidianas y del espacio de la teoría de la relatividad de Einstein.

6.1. PRODUCTO ESCALAR USUAL DE $\mathbb{R}^n$

Además de las operaciones citadas en el Capítulo 5, en el espacio $\mathbb{R}^n$ se puede definir otra operación entre vectores denominada *producto escalar*. Se comienza con el espacio $\mathbb{R}^2$.

Definición 6.1. Se llama **producto escalar usual** en $\mathbb{R}^2$ a la operación

$$\begin{array}{rcl} (\,\cdot\,) : \mathbb{R}^2 \times \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (\mathbf{x}, \mathbf{y}) & \mapsto & \mathbf{x} \cdot \mathbf{y} \end{array}$$

que a cada par de vectores $\mathbf{x}$, $\mathbf{y} \in \mathbb{R}^2$ le asocia un número real dado por

$$\mathbf{x} \cdot \mathbf{y} = \begin{cases} 0 & \text{si } \mathbf{x} = \mathbf{0} \text{ o } \mathbf{y} = \mathbf{0} \\ \|\mathbf{x}\| \, \|\mathbf{y}\| \cos\alpha & \text{en otro caso,} \end{cases}$$

siendo α el ángulo que forman los vectores.

Usando coordenadas cartesianas el producto escalar anterior se expresa como:

$$\mathbf{x} \cdot \mathbf{y} = (x_1, x_2) \cdot (y_1, y_2) = x_1 y_1 + x_2 y_2 \in \mathbb{R}.$$

Este producto escalar se puede generalizar a $\mathbb{R}^n$ del siguiente modo.

Definición 6.2. Dados dos vectores $\mathbf{x}$, $\mathbf{y} \in \mathbb{R}^n$, se define el **producto escalar usual** como:

$$\mathbf{x} \cdot \mathbf{y} = (x_1, \ldots, x_n) \cdot (y_1, \ldots, y_n) = x_1 y_1 + \ldots + x_n y_n \in \mathbb{R}.$$

La expresión matricial del producto escalar es la siguiente:

$$\mathbf{x}^t \, \mathbf{y} = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = x_1 y_1 + \ldots + x_n y_n \in \mathbb{R}.$$

Propiedades. El espacio vectorial $\mathbb{R}^n$ con esta nueva operación verifica:

1. $\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}, \ \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.
2. $(\lambda \mathbf{x} + \mu \mathbf{y}) \cdot \mathbf{z} = \lambda\, (\mathbf{x} \cdot \mathbf{z}) + \mu\, (\mathbf{y} \cdot \mathbf{z}), \ \forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n, \ \forall \lambda, \mu \in \mathbb{R}$.
3. $\mathbf{x} \cdot \mathbf{x} = \|\mathbf{x}\|^2 \geq 0, \ \forall \mathbf{x} \in \mathbb{R}^n$. Además, $\mathbf{x} \cdot \mathbf{x} = 0 \iff \mathbf{x} = \mathbf{0}$.

Debido a estas propiedades se dice que $\mathbb{R}^n$ con el producto escalar usual es un *espacio euclídeo*.

6.2. PRODUCTO ESCALAR EN UN ESPACIO VECTORIAL

A continuación se generaliza la estructura euclídea de $\mathbb{R}^n$ estudiando espacios vectoriales con las mismas tres propiedades anteriores.

Definición 6.3. Sea V un espacio vectorial real. Un **producto escalar** en V es una aplicación $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{R}$ que verifica las siguientes propiedades:

1. *Simétrica:* $\langle \mathbf{u}, \mathbf{v}\rangle = \langle \mathbf{v}, \mathbf{u}\rangle, \quad \forall \mathbf{u}, \mathbf{v} \in V.$
2. *Lineal (primera componente):*
   - $\langle \mathbf{u}+\mathbf{v}, \mathbf{w}\rangle = \langle \mathbf{u}, \mathbf{w}\rangle + \langle \mathbf{v}, \mathbf{w}\rangle, \quad \forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in V.$
   - $\langle \lambda\mathbf{u}, \mathbf{v}\rangle = \lambda\langle \mathbf{u}, \mathbf{v}\rangle, \quad \forall \mathbf{u}, \mathbf{v} \in V, \quad \forall \lambda \in \mathbb{R}.$
3. *Definida positiva:*
   - $\langle \mathbf{u}, \mathbf{u}\rangle \geq 0, \quad \forall \mathbf{u} \in V.$
   - $\langle \mathbf{u}, \mathbf{u}\rangle = 0 \Longleftrightarrow \mathbf{u} = \mathbf{0}.$

Se dice que V es un **espacio vectorial euclídeo**, es decir, es un espacio vectorial real dotado de un producto escalar, y se representa por $(V, \langle\cdot,\cdot\rangle)$.

La segunda propiedad se puede resumir en una sola:

$$\langle \lambda\mathbf{u}+\mu\mathbf{v}, \mathbf{w}\rangle = \lambda\langle \mathbf{u}, \mathbf{w}\rangle + \mu\langle \mathbf{v}, \mathbf{w}\rangle, \quad \forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in V, \quad \forall \lambda, \mu \in \mathbb{R}.$$

Debido a la simetría, el producto escalar también es lineal en la segunda componente, es decir, verifica:

$$\langle \mathbf{u}, \lambda\mathbf{v}+\mu\mathbf{w}\rangle = \lambda\langle \mathbf{u}, \mathbf{v}\rangle + \mu\langle \mathbf{u}, \mathbf{w}\rangle, \quad \forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in V, \quad \forall \lambda, \mu \in \mathbb{R}.$$

Así, el producto escalar es *bilineal*, es decir, es lineal en ambas componentes.

Observación. Como consecuencia inmediata de la bilinealidad todo producto escalar verifica:

$$\langle \mathbf{0}, \mathbf{v}\rangle = 0, \quad \langle \mathbf{v}, \mathbf{0}\rangle = 0, \quad \forall \mathbf{v} \in V.$$

Sin embargo, puede ocurrir que $\langle \mathbf{u}, \mathbf{v}\rangle = 0$ con $\mathbf{u} \neq \mathbf{0}$ y $\mathbf{v} \neq \mathbf{0}$. Por ejemplo en $\mathbb{R}^2$ con el producto escalar usual tomar $\mathbf{u} = (1, 0)$ y $\mathbf{v} = (0, 1)$.

✓ **Ejemplo 1.** Además de $\mathbb{R}^n$ con el producto escalar usual, existen otros ejemplos de espacios vectoriales euclídeos:

- Si $V = \mathrm{Mat}_n(\mathbb{R})$ es el espacio vectorial de las matrices reales cuadradas de orden n, la aplicación dada por

  $$\langle A, B\rangle = \mathrm{tr}(AB^t)$$

  es un producto escalar. Por tanto, $(\mathrm{Mat}_n(\mathbb{R}), \langle\cdot,\cdot\rangle)$ es un espacio euclídeo.

- Si $V = \mathcal{C}([0,1])$ es el espacio vectorial de las funciones reales continuas definidas en el intervalo cerrado $[0,1]$, la aplicación definida por
$$\langle f, g\rangle = \int_0^1 f(t)g(t)\,dt$$
es un producto escalar. Así, $(\mathcal{C}([0,1]), \langle\cdot,\cdot\rangle)$ es un espacio euclídeo.

◇

6.3. NORMAS, DISTANCIAS Y ÁNGULOS

El producto escalar es la herramienta que permite medir en un espacio vectorial. En esta sección se introducen los conceptos de norma, distancia y ángulo en los espacios euclídeos.

Definición 6.4. Sea V un espacio vectorial euclídeo. Se define la **norma** (o **longitud**) de un vector $\mathbf{u} \in V$, y se denota por $\|\mathbf{u}\|$, como el número real positivo:
$$\|\mathbf{u}\| = +\sqrt{\langle \mathbf{u}, \mathbf{u}\rangle}.$$
Queda definida la **aplicación norma** como:
$$\begin{array}{rccc} \|\cdot\|: & V & \longrightarrow & \mathbb{R}^+ \\ & \mathbf{u} & \mapsto & \|\mathbf{u}\|. \end{array}$$

Observar que la tercera propiedad del producto escalar dice que $\langle \mathbf{u}, \mathbf{u}\rangle \geq 0$, y por tanto siempre tiene sentido calcular su raíz cuadrada, siendo esta un número real positivo.

A lo largo del capítulo los vectores con norma 1 juegan un papel importante.

Definición 6.5. Un vector $\mathbf{u}$ de un espacio vectorial euclídeo V se dice **unitario** si tiene norma 1, es decir, $\|\mathbf{u}\| = 1$.

Observación. Dado un vector $\mathbf{u} \neq \mathbf{0}$, se puede construir otro vector $\mathbf{v}$ proporcional a $\mathbf{u}$ de norma 1, denominado *vector normalizado*. El proceso recibe el nombre de *normalización*.
$$\mathbf{v} = \frac{\mathbf{u}}{\|\mathbf{u}\|}.$$

✓ **Ejemplo 2.**

- Sea $\mathbb{R}^3$ con el producto escalar usual. El vector $\mathbf{x} = (1,2,3)$ tiene norma $\sqrt{14}$ y por tanto su vector normalizado es $\frac{1}{\sqrt{14}}\mathbf{x} = \left(\frac{1}{\sqrt{14}}, \frac{2}{\sqrt{14}}, \frac{3}{\sqrt{14}}\right)$.

- Sea el espacio vectorial $\mathrm{Mat}_2(\mathbb{R})$ con el producto escalar dado por la traza definido en el Ejemplo 1. La matriz $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ tiene norma $\sqrt{30}$, por lo que su matriz normalizada es:
$$\frac{1}{\sqrt{30}} A = \begin{pmatrix} 1/\sqrt{30} & 2/\sqrt{30} \\ 3/\sqrt{30} & 4/\sqrt{30} \end{pmatrix}.$$

◇

La aplicación norma verifica las siguientes propiedades:

Propiedades.

1. $\|\mathbf{u}\| \geq 0, \ \forall \mathbf{u} \in V$. Más aún, $\|\mathbf{u}\| = 0 \iff \mathbf{u} = \mathbf{0}$.
2. $\|\lambda \mathbf{u}\| = |\lambda| \, \|\mathbf{u}\|, \quad \forall \lambda \in \mathbb{R}, \quad \forall \mathbf{u} \in V$.
3. *Desigualdad triangular:* $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|, \quad \forall \mathbf{u}, \mathbf{v} \in V$.

Como consecuencia de las propiedades anteriores se deduce el siguiente resultado.

Teorema 6.6 (Desigualdad de Cauchy-Schwarz)**.** Sea V un espacio vectorial euclídeo y se consideran $\mathbf{u}, \mathbf{v} \in V$. Entonces,
$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \, \|\mathbf{v}\|.$$
Además, $|\langle \mathbf{u}, \mathbf{v} \rangle| = \|\mathbf{u}\| \, \|\mathbf{v}\| \iff \{\mathbf{u}, \mathbf{v}\}$ es una familia ligada.

Observación. La desigualdad de Cauchy-Schwarz se puede escribir como:
$$-1 \leq \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\| \, \|\mathbf{v}\|} \leq 1, \qquad (\mathbf{u}, \mathbf{v} \neq \mathbf{0})$$
lo que da pie a la siguiente definición.

Definición 6.7. Sean $\mathbf{u}, \mathbf{v}$ dos vectores no nulos de V. El **ángulo** $\alpha \in [-\pi, \pi]$ que forman $\mathbf{u}$ y $\mathbf{v}$ es aquel que cumple
$$\cos \alpha = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\| \, \|\mathbf{v}\|}.$$

Por último, se presenta cómo medir la distancia entre dos vectores $\mathbf{u}$ y $\mathbf{v}$.

Definición 6.8. Sea $(V, \langle\cdot,\cdot\rangle)$ un espacio vectorial euclídeo y sean $\mathbf{u}, \mathbf{v} \in V$. Se define la **distancia** del vector $\mathbf{u}$ al vector $\mathbf{v}$ como:

$$d(\mathbf{u},\mathbf{v}) = \|\mathbf{u}-\mathbf{v}\|.$$

La **aplicación distancia** se define como:

$$\begin{array}{rccl} d: & V \times V & \longrightarrow & \mathbb{R}^+ \\ & (\mathbf{u},\mathbf{v}) & \mapsto & d(\mathbf{u},\mathbf{v}) = \|\mathbf{u}-\mathbf{v}\|. \end{array}$$

La aplicación distancia verifica las siguientes propiedades:

Propiedades. Para todo $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ se tiene:

1. $d(\mathbf{u},\mathbf{v}) \geq 0$. Además, $d(\mathbf{u},\mathbf{v}) = 0 \iff \mathbf{u} = \mathbf{v}$.
2. $d(\mathbf{u},\mathbf{v}) = d(\mathbf{v},\mathbf{u})$.
3. *Desigualdad triangular:* $d(\mathbf{u},\mathbf{v}) \leq d(\mathbf{u},\mathbf{w}) + d(\mathbf{w},\mathbf{v})$.

✓ **Ejemplo 3.** Se considera $\mathbb{R}^n$ con el producto escalar usual. Sean $\mathbf{x} = (x_1, \ldots, x_n)$, $\mathbf{y} = (y_1, \ldots, y_n) \in \mathbb{R}^n$, entonces

$$\|\mathbf{x}\| = \sqrt{\langle \mathbf{x},\mathbf{x}\rangle} = \sqrt{\big\langle (x_1,\ldots,x_n),(x_1,\ldots,x_n)\big\rangle} = \sqrt{x_1^2 + \ldots + x_n^2}\,,$$

$$d(\mathbf{x},\mathbf{y}) = \|\mathbf{x}-\mathbf{y}\| = \sqrt{(x_1-y_1)^2 + \ldots + (x_n-y_n)^2}\,.$$

Observar que en este caso la norma coincide con la norma euclídea (denotada por $\|\cdot\|_2$) definida en el Capítulo 3.

◇

6.4. BASES ORTONORMALES

Una noción importante en la geometría del espacio es la de ortogonalidad. En el espacio $\mathbb{R}^2$ usual dos vectores son ortogonales o perpendiculares si forman un ángulo $\alpha = \frac{\pi}{2}$, o equivalentemente, si $\cos\alpha = 0$. Esto a su vez es equivalente a que el producto escalar de estos dos vectores sea cero. Extendiendo esta situación se tiene lo siguiente.

Definición 6.9. Sea $(V, \langle\cdot,\cdot\rangle)$ un espacio vectorial euclídeo. Dos vectores $\mathbf{u}$ y $\mathbf{v}$ son **ortogonales** si $\langle \mathbf{u},\mathbf{v}\rangle = 0$.

Definición 6.10. Sea V un espacio vectorial euclídeo. Una **familia** de vectores $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ se dice **ortogonal** si

$$\langle \mathbf{u}_i, \mathbf{u}_j \rangle = 0, \quad \forall i \neq j.$$

La familia $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ se dice **ortonormal** si es ortogonal y además todos los vectores $\mathbf{u}_i$'s son unitarios.

Observaciones.

- El vector $\mathbf{0}_V$ es ortogonal a todo vector $\mathbf{u} \in V$.
- El vector $\mathbf{0}_V$ no puede pertenecer a ninguna familia ortonormal ya que no se puede normalizar por tener norma 0, pero sí a una ortogonal.

✓ **Ejemplo 4.**

- La base canónica de $\mathbb{R}^n$, $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$, es ortonormal con respecto al producto escalar usual:

$$\langle \mathbf{e}_i, \mathbf{e}_j \rangle = 0, \quad \forall i \neq j; \qquad \langle \mathbf{e}_i, \mathbf{e}_i \rangle = 1, \quad \forall i = 1, \ldots, n.$$

- La base canónica de $\mathrm{Mat}_n(\mathbb{R})$, $\{E_{11}, \ldots, E_{nn}\}$, es ortonormal con respecto al producto escalar dado por la traza definido en el Ejemplo 1.

◇

El siguiente resultado es útil para el cálculo de bases de un subespacio.

Proposición 6.11. Toda familia ortogonal de vectores no nulos (en particular, toda familia ortonormal) es linealmente independiente.

El concepto de ortogonalidad tiene su análogo para subespacios.

Definición 6.12. Sea V un espacio vectorial euclídeo. Dos subespacios S_1 y S_2 de V se dicen **ortogonales entre sí** cuando verifican:

$$\langle \mathbf{u}, \mathbf{v} \rangle = 0, \quad \forall \mathbf{u} \in S_1, \quad \forall \mathbf{v} \in S_2.$$

Dicho de otro modo, todo vector de S_1 es ortogonal a todo vector de S_2.

Proposición 6.13. Sean $S_1 = \langle \mathbf{u}_1, \ldots, \mathbf{u}_r \rangle$ y $S_2 = \langle \mathbf{v}_1, \ldots, \mathbf{v}_s \rangle$ dos subespacios vectoriales de un espacio euclídeo V. Entonces S_1 y S_2 son ortogonales si y solo si

$$\langle \mathbf{u}_i, \mathbf{v}_j \rangle = 0, \quad \forall i = 1, \ldots, r, \quad \forall j = 1, \ldots, s.$$

Es decir, basta comprobar la propiedad de ortogonalidad para los generadores de los subespacios.

Dado un subespacio S de un espacio vectorial V, el conjunto de todos los vectores de V ortogonales a S tiene estructura de subespacio vectorial.

Definición 6.14. Sea V un espacio vectorial euclídeo. Dado S un subespacio vectorial de V se define el **subespacio ortogonal** a S, denotado por $S^{\perp}$, como:

$$S^{\perp} = \{\mathbf{v} \in V \mid \langle \mathbf{u}, \mathbf{v} \rangle = 0, \quad \forall \mathbf{u} \in S\}.$$

El subespacio anterior es ortogonal a S y verifica:

Proposición 6.15. Sea V un espacio vectorial euclídeo y S un subespacio vectorial de V. Entonces,

$$\boxed{V = S \oplus S^{\perp}}$$

Por tanto, $\dim V = \dim S + \dim S^{\perp}$ y en particular,

$$\dim S^{\perp} = \dim V - \dim S.$$

El siguiente ejemplo proporciona un método para hallar una base de $S^{\perp}$.

✓ **Ejemplo 5.** Dado el subespacio $S = \langle (1,2,3) \rangle$ de $\mathbb{R}^3$ con el producto escalar usual, calcular una base de $S^{\perp}$ y comprobar que se cumple la fórmula de la dimensión anterior.

Solución: Siguiendo la definición, $S^{\perp} = \{(x,y,z) \in \mathbb{R}^3 \mid \langle (x,y,z), (1,2,3) \rangle = 0\}$. Teniendo en cuenta que

$$\langle (x,y,z), (1,2,3) \rangle = 0 \iff x + 2y + 3z = 0 \iff x = -2y - 3z,$$

el subespacio ortogonal a S se puede escribir como:

$$\begin{aligned} S^{\perp} &= \{(x,y,z) \in \mathbb{R}^3 \mid x = -2y - 3z\} \\ &= \{(-2y - 3z, y, z) \mid y, z \in \mathbb{R}\} \\ &= \{y(-2,1,0) + z(-3,0,1) \mid y, z \in \mathbb{R}\} \\ &= \langle (-2,1,0), (-3,0,1) \rangle. \end{aligned}$$

Los dos vectores anteriores son sistema generador y linealmente independientes, por tanto forman una base de $S^{\perp}$. Luego $\dim S^{\perp} = 2$ y la fórmula de la dimensión se cumple pues $\dim \mathbb{R}^3 - \dim S = 3 - 1 = 2 = \dim S^{\perp}$.

◇

Observación. Se recomienda comprobar que el resultado obtenido es correcto, realizando los productos escalares correspondientes:

$$\langle (-2,1,0), (1,2,3) \rangle = (-2) \cdot 1 + 1 \cdot 2 + 0 \cdot 3 = 0,$$

$$\langle (-3,0,1), (1,2,3) \rangle = (-3) \cdot 1 + 0 \cdot 2 + 1 \cdot 3 = 0.$$

6.4.1. MÉTODO DE GRAM-SCHMIDT

Sea V un espacio vectorial euclídeo y $S = \langle \mathbf{a}_1, \ldots, \mathbf{a}_r \rangle$ un subespacio de V. El problema que se plantea a continuación es encontrar a partir de los vectores $\mathbf{a}_i$'s una familia de vectores *ortonormales* que sigan generando S, es decir, $S = \langle \mathbf{q}_1, \ldots, \mathbf{q}_s \rangle$. En particular, y utilizando la Proposición 6.11, se tiene que la familia ortonormal construida $\{\mathbf{q}_1, \ldots, \mathbf{q}_s\}$ es una base del subespacio S.

El método para encontrar una familia ortonormal $\{\mathbf{q}_1, \ldots, \mathbf{q}_s\}$ a partir de una familia cualquiera $\{\mathbf{a}_1, \ldots, \mathbf{a}_r\}$ se articula en tres pasos:

- En el paso 1, llamado *método de Gram-Schmidt*, partiendo de $\{\mathbf{a}_1, \ldots, \mathbf{a}_r\}$, se construye una familia ortogonal $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$, sistema generador del subespacio S. Observar que se obtiene el mismo número de vectores.
- En el paso 2, se eliminan los vectores nulos resultantes del proceso anterior, obteniéndose una base ortogonal de S, $\{\mathbf{v}_1, \ldots, \mathbf{v}_s\}$.
- En el paso 3, se normaliza esta base $\{\mathbf{v}_1, \ldots, \mathbf{v}_s\}$ dividiendo cada vector $\mathbf{v}_j$ por su norma, obteniendo de este modo la familia ortonormal buscada $\{\mathbf{q}_1, \ldots, \mathbf{q}_s\}$, y base ortonormal del subespacio S.

Con más detenimiento, cada etapa se realiza del siguiente modo:

$\boxed{\text{Paso 1}}$

Es conocido como el **método de Gram-Schmidt**. Se parte de la familia de vectores $\{\mathbf{a}_1, \ldots, \mathbf{a}_r\}$ que genera el subespacio S. La construcción de los vectores $\mathbf{v}_j$ se realiza de manera recursiva a partir de la familia anterior, es decir, es necesario conocer $\{\mathbf{v}_1, \ldots, \mathbf{v}_{j-1}\}$ para poder construir $\mathbf{v}_j$. Concretamente:

$$\begin{aligned}
\mathbf{v}_1 &= \mathbf{a}_1, \\
\mathbf{v}_2 &= \mathbf{a}_2 - \alpha_{12}\mathbf{v}_1\,, \\
\mathbf{v}_3 &= \mathbf{a}_3 - \alpha_{13}\mathbf{v}_1 - \alpha_{23}\mathbf{v}_2\,, \\
&\vdots \\
\mathbf{v}_j &= \mathbf{a}_j - \alpha_{1j}\mathbf{v}_1 - \alpha_{2j}\mathbf{v}_2 - \ldots - \alpha_{j-1,j}\mathbf{v}_{j-1} = \mathbf{a}_j - \sum_{i=1}^{j-1} \alpha_{ij}\mathbf{v}_i, \\
&\vdots
\end{aligned}$$

donde los coeficientes α_{ij} vienen dados por:

$$\begin{cases}
\alpha_{ij} = \dfrac{\langle \mathbf{v}_i, \mathbf{a}_j \rangle}{\|\mathbf{v}_i\|^2} = \dfrac{\langle \mathbf{v}_i, \mathbf{a}_j \rangle}{\langle \mathbf{v}_i, \mathbf{v}_i \rangle} & \text{si} \;\; \mathbf{v}_i \neq \mathbf{0}, \\[2ex]
\alpha_{ij} = 0 & \text{si} \;\; \mathbf{v}_i = \mathbf{0}.
\end{cases}$$

Observar que si $\mathbf{a}_j$ depende linealmente de los vectores anteriores $\{\mathbf{a}_1, \dots, \mathbf{a}_{j-1}\}$, entonces el vector construido $\mathbf{v}_j$ es nulo: $\mathbf{v}_j = \mathbf{0}$.

$\boxed{\text{Paso 2}}$

En la familia $\{\mathbf{v}_1, \dots, \mathbf{v}_r\}$ obtenida en el paso 1 se eliminan los vectores nulos, de modo que se obtiene la familia $\{\mathbf{v}_1, \dots, \mathbf{v}_s\}$ ortogonal que ya es base de S aplicando la Proposición 6.11.

$\boxed{\text{Paso 3}}$

Partiendo de la base ortogonal $\{\mathbf{v}_1, \dots, \mathbf{v}_s\}$ se construye la base $\{\mathbf{q}_1, \dots, \mathbf{q}_s\}$ sin más que normalizar cada vector $\mathbf{v}_j$ para $j = 1, \dots, s$:

$$\mathbf{q}_j = \frac{\mathbf{v}_j}{\|\mathbf{v}_j\|}.$$

Se muestra un ejemplo de aplicación del método.

✓ **Ejemplo 6.** Utilizar el método de Gram-Schmidt para construir una base ortonormal del subespacio S de $\mathbb{R}^3$ dado por:

$$S = \langle \underbrace{(1,1,1)}_{\mathbf{a}_1}, \underbrace{(0,1,1)}_{\mathbf{a}_2}, \underbrace{(0,0,1)}_{\mathbf{a}_3}, \underbrace{(1,2,3)}_{\mathbf{a}_4} \rangle.$$

Solución: Siguiendo el paso 1 anterior:

- $\mathbf{v}_1 = \mathbf{a}_1 = (1,1,1)$.

- $\mathbf{v}_2 = \mathbf{a}_2 - \alpha_{12}\mathbf{v}_1$, donde $\alpha_{12} = \dfrac{\langle \mathbf{v}_1, \mathbf{a}_2 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle}$.

$$\alpha_{12} = \frac{\langle (1,1,1), (0,1,1) \rangle}{\langle (1,1,1), (1,1,1) \rangle} = \frac{0+1+1}{1+1+1} = \frac{2}{3}.$$

$$\mathbf{v}_2 = \mathbf{a}_2 - \frac{2}{3}\mathbf{v}_1 = (0,1,1) - \frac{2}{3}(1,1,1) = \Big(-\frac{2}{3}, \frac{1}{3}, \frac{1}{3} \Big).$$

Es conveniente comprobar que efectivamente $\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = 0$.

- $\mathbf{v}_3 = \mathbf{a}_3 - \alpha_{13}\mathbf{v}_1 - \alpha_{23}\mathbf{v}_2$, donde $\alpha_{13} = \dfrac{\langle \mathbf{v}_1, \mathbf{a}_3 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle}$ y $\alpha_{23} = \dfrac{\langle \mathbf{v}_2, \mathbf{a}_3 \rangle}{\langle \mathbf{v}_2, \mathbf{v}_2 \rangle}$.

$$\alpha_{13} = \frac{\langle (1,1,1), (0,0,1) \rangle}{\langle (1,1,1), (1,1,1) \rangle} = \frac{0+0+1}{1+1+1} = \frac{1}{3}.$$

$$\alpha_{23} = \frac{\langle (-\frac{2}{3}, \frac{1}{3}, \frac{1}{3}), (0,0,1) \rangle}{\langle (-\frac{2}{3}, \frac{1}{3}, \frac{1}{3}), (-\frac{2}{3}, \frac{1}{3}, \frac{1}{3}) \rangle} = \frac{0+0+\frac{1}{3}}{\frac{4}{9}+\frac{1}{9}+\frac{1}{9}} = \frac{1/3}{6/9} = \frac{1}{2}.$$

$$\mathbf{v}_3 = (0,0,1) - \frac{1}{3}(1,1,1) - \frac{1}{2}\Big(-\frac{2}{3}, \frac{1}{3}, \frac{1}{3} \Big) = \Big(0, -\frac{1}{2}, \frac{1}{2} \Big).$$

De nuevo resulta interesante comprobar que $\langle \mathbf{v}_1, \mathbf{v}_3 \rangle = 0$ y $\langle \mathbf{v}_2, \mathbf{v}_3 \rangle = 0$.

- $\mathbf{v}_4 = \mathbf{a}_4 - \alpha_{14}\mathbf{v}_1 - \alpha_{24}\mathbf{v}_2 - \alpha_{34}\mathbf{v}_3$, donde $\alpha_{i4} = \dfrac{\langle \mathbf{v}_i, \mathbf{a}_4\rangle}{\langle \mathbf{v}_i, \mathbf{v}_i\rangle}$, $i = 1, 2, 3$.

$$\alpha_{14} = \frac{\langle \mathbf{v}_1, \mathbf{a}_4\rangle}{\langle \mathbf{v}_1, \mathbf{v}_1\rangle} = \frac{\langle (1,1,1),(1,2,3)\rangle}{\langle (1,1,1),(1,1,1)\rangle} = \frac{1+2+3}{1+1+1} = 2.$$

$$\alpha_{24} = \frac{\langle \mathbf{v}_2, \mathbf{a}_4\rangle}{\langle \mathbf{v}_2, \mathbf{v}_2\rangle} = \frac{\langle (-\frac{2}{3},\frac{1}{3},\frac{1}{3}),(1,2,3)\rangle}{\langle (-\frac{2}{3},\frac{1}{3},\frac{1}{3}),(-\frac{2}{3},\frac{1}{3},\frac{1}{3})\rangle} = \frac{-\frac{2}{3}+\frac{2}{3}+1}{\frac{4}{9}+\frac{1}{9}+\frac{1}{9}} = \frac{1}{2/3} = \frac{3}{2}.$$

$$\alpha_{34} = \frac{\langle \mathbf{v}_3, \mathbf{a}_4\rangle}{\langle \mathbf{v}_3, \mathbf{v}_3\rangle} = \frac{\langle (0,-\frac{1}{2},\frac{1}{2}),(1,2,3)\rangle}{\langle (0,-\frac{1}{2},\frac{1}{2}),(0,-\frac{1}{2},\frac{1}{2})\rangle} = \frac{0-1+\frac{3}{2}}{0+\frac{1}{4}+\frac{1}{4}} = \frac{1/2}{1/2} = 1.$$

$$\mathbf{v}_4 = (1,2,3) - 2(1,1,1) - \frac{3}{2}\Big(-\frac{2}{3},\frac{1}{3},\frac{1}{3}\Big) - 1\Big(0,-\frac{1}{2},\frac{1}{2}\Big) = (0,0,0).$$

Esto indica que $\mathbf{a}_4$ es combinación lineal de $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$.

Se obtiene la siguiente familia ortogonal:

$$\Big\{ \underbrace{(1,1,1)}_{\mathbf{v}_1}, \underbrace{\Big(-\frac{2}{3},\frac{1}{3},\frac{1}{3}\Big)}_{\mathbf{v}_2}, \underbrace{\Big(0,-\frac{1}{2},\frac{1}{2}\Big)}_{\mathbf{v}_3}, \underbrace{(0,0,0)}_{\mathbf{v}_4} \Big\}.$$

En el paso 2 se elimina el cuarto vector $\mathbf{v}_4$, obteniendo así una base de vectores ortogonales $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. En el paso paso 3, utilizando que

$$\|\mathbf{v}_1\| = \sqrt{3}, \quad \|\mathbf{v}_2\| = \sqrt{\frac{2}{3}}, \quad \|\mathbf{v}_3\| = \sqrt{\frac{1}{2}},$$

se obtiene la siguiente base ortonormal de S:

$$\Big\{ \underbrace{\Big(\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}}\Big)}_{\mathbf{q}_1}, \underbrace{\Big(-\frac{2}{\sqrt{6}},\frac{1}{\sqrt{6}},\frac{1}{\sqrt{6}}\Big)}_{\mathbf{q}_2}, \underbrace{\Big(0,-\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}\Big)}_{\mathbf{q}_3} \Big\}.$$

En este caso, como S es un subespacio de $\mathbb{R}^3$ de dimensión 3, se tiene que $S = \mathbb{R}^3$.

◇

Proposición 6.16. A partir del método anterior se deducen las siguientes afirmaciones:

- En un espacio euclídeo, todo subespacio finitamente generado tiene una base ortonormal.
- En un espacio euclídeo de dimensión finita, toda familia ortonormal se puede completar hasta una base ortonormal.

6.5. FACTORIZACIÓN QR

El proceso anterior permite realizar una descomposición de una matriz A cualquiera en producto de dos matrices, $A = QR$, donde Q es una matriz de columnas ortonormales y R es una matriz triangular superior. Para obtener esta descomposición se procede de la siguiente forma:

1. Partiendo de una matriz $A \in \mathrm{Mat}_{n\times r}(\mathbb{R})$, se aplica el método de Gram-Schmidt a las columnas de la matriz:

$$A = (\, \mathbf{a}_1 \,|\, \mathbf{a}_2 \,|\, \cdots \,|\, \mathbf{a}_r \,) \in \mathrm{Mat}_{n\times r}(\mathbb{R}).$$

2. Con los datos obtenidos en el paso anterior, se realiza una descomposición intermedia $A = Q_0R_0$, donde:

   - Q_0 es la matriz cuyas columnas son los vectores ortogonales $\{\mathbf{v}_1, \dots, \mathbf{v}_r\}$ obtenidos al finalizar el paso 1 del método de Gram-Schmidt:

$$Q_0 = (\, \mathbf{v}_1 \,|\, \mathbf{v}_2 \,|\, \cdots \,|\, \mathbf{v}_r \,) \in \mathrm{Mat}_{n\times r}(\mathbb{R}), \quad \mathrm{rg}\, Q_0 = \mathrm{rg}\, A.$$

   - R_0 es la matriz triangular superior con 1's en la diagonal, y por tanto regular, formada por los coeficientes α_{ij} del método de Gram-Schmidt, de modo que cada uno de ellos se coloca en la matriz en el lugar correspondiente a sus índices:

$$R_0 = \begin{pmatrix} 1 & \alpha_{12} & \cdots & \alpha_{1r} \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \alpha_{r-1,r} \\ 0 & \cdots & 0 & 1 \end{pmatrix} \in \mathrm{Mat}_r(\mathbb{R}).$$

3. En la matriz Q_0 se eliminan las columnas nulas y en la matriz R_0 se eliminan las filas asociadas a las columnas suprimidas en Q_0. Se obtienen de esta manera dos nuevas matrices, Q_0' y R_0', que siguen cumpliendo $A = Q_0'R_0'$ y además verifican:

$$\begin{aligned} Q_0' &\in \mathrm{Mat}_{n\times(\mathrm{rg}\, A)}(\mathbb{R}), & \mathrm{rg}\, Q_0' &= \mathrm{rg}\, A; \\ R_0' &\in \mathrm{Mat}_{(\mathrm{rg}\, A)\times r}(\mathbb{R}), & \mathrm{rg}\, R_0' &= \mathrm{rg}\, A. \end{aligned}$$

4. Se divide la columna k-ésima de la matriz Q_0' por su norma y se multiplica la fila k-ésima de la matriz R_0' por la norma anterior. Este proceso se repite para todas las columnas de Q_0' y todas las filas de R_0'. De esta manera se obtienen las matrices Q y R que cumplen $A = QR$. Observar que

$$Q = (\, \mathbf{q}_1 \,|\, \mathbf{q}_2 \,|\, \cdots \,|\, \mathbf{q}_s \,),$$

es decir, las columnas de la matriz Q son los vectores de la base ortonormal obtenida tras aplicar el método de la Sección 6.4.1 a las columnas de la matriz A.

Teorema 6.17 (Descomposición QR)**.** Sea la matriz $A \in \mathrm{Mat}_{n\times r}(\mathbb{R})$. Entonces existe una matriz $Q \in \mathrm{Mat}_{n\times(\mathrm{rg}\,A)}(\mathbb{R})$ cuyas columnas son ortonormales (con $\mathrm{rg}\,Q = \mathrm{rg}\,A$) y una matriz triangular superior $R \in \mathrm{Mat}_{(\mathrm{rg}\,A)\times r}(\mathbb{R})$ (con $\mathrm{rg}\,R = \mathrm{rg}\,A$) tales que

$$A = QR.$$

✓ **Ejemplo 7.** Hallar una descomposición QR de la matriz A cuyas columnas son los vectores del Ejemplo 6:

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

Solución: La construcción de las matrices Q_0 y R_0 es la siguiente:

$$Q_0 = \big(\, \mathbf{v}_1 \,|\, \mathbf{v}_2 \,|\, \mathbf{v}_3 \,|\, \mathbf{v}_4 \,\big) = \begin{pmatrix} 1 & -2/3 & 0 & 0 \\ 1 & 1/3 & -1/2 & 0 \\ 1 & 1/3 & 1/2 & 0 \end{pmatrix}.$$

$$R_0 = \begin{pmatrix} 1 & \alpha_{12} & \alpha_{13} & \alpha_{14} \\ 0 & 1 & \alpha_{23} & \alpha_{24} \\ 0 & 0 & 1 & \alpha_{34} \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2/3 & 1/3 & 2 \\ 0 & 1 & 1/2 & 3/2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Eliminando la cuarta columna de Q_0 y, por tanto, la cuarta fila de R_0 se obtienen las matrices Q_0', R_0':

$$Q_0 \longmapsto \underbrace{\begin{pmatrix} 1 & -2/3 & 0 \\ 1 & 1/3 & -1/2 \\ 1 & 1/3 & 1/2 \end{pmatrix}}_{Q_0'}, \qquad R_0 \longmapsto \underbrace{\begin{pmatrix} 1 & 2/3 & 1/3 & 2 \\ 0 & 1 & 1/2 & 3/2 \\ 0 & 0 & 1 & 1 \end{pmatrix}}_{R_0'}.$$

Lo último que falta para conseguir la factorización QR buscada es dividir (y multiplicar) por la norma de los vectores $\mathbf{v}_i$:

- Como $\|\mathbf{v}_1\| = \sqrt{3}$, se divide la primera columna de Q_0' entre $\sqrt{3}$ y se multiplica la primera fila de R_0' por $\sqrt{3}$.
- Como $\|\mathbf{v}_2\| = \sqrt{\frac{2}{3}}$, se divide la segunda columna de Q_0' entre $\sqrt{\frac{2}{3}}$ y se multiplica la segunda fila de R_0' por $\sqrt{\frac{2}{3}}$.
- Como $\|\mathbf{v}_3\| = \sqrt{\frac{1}{2}}$, se divide la tercera columna de Q_0' entre $\sqrt{\frac{1}{2}}$ y se multiplica la tercera fila de R_0' por $\sqrt{\frac{1}{2}}$.

Se obtiene la siguiente factorización:

$$A = QR = \begin{pmatrix} 1/\sqrt{3} & -2/\sqrt{6} & 0 \\ 1/\sqrt{3} & 1/\sqrt{6} & -1/\sqrt{2} \\ 1/\sqrt{3} & 1/\sqrt{6} & 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} \sqrt{3} & 2/\sqrt{3} & 1/\sqrt{3} & 2\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{6} & 3/\sqrt{6} \\ 0 & 0 & 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}.$$

◇

Observación. Las columnas de la matriz Q son ortonormales por lo que se cumple que

$$Q^t Q = I.$$

Notar que las filas de Q no tienen por qué ser ortonormales, es decir, en general no se cumple $Q\,Q^t = I$.

Las matrices que verifican las dos igualdades anteriores se definen a continuación.

6.5.1. MATRICES ORTOGONALES

Definición 6.18. Una matriz cuadrada $P \in \mathrm{Mat}_n(\mathbb{R})$ se dice **ortogonal** si:

$$PP^t = P^t P = I_n,$$

o equivalentemente,

$$P^{-1} = P^t.$$

Nota. Las matrices ortogonales cumplen las siguientes propiedades:

- Sus filas y columnas constituyen bases ortonormales de $\mathbb{R}^n$.
- Su determinante vale 1 o -1.

✓ **Ejemplo 8.** Comprobar que la matriz $A = \begin{pmatrix} \cos\alpha & \operatorname{sen}\alpha \\ -\operatorname{sen}\alpha & \cos\alpha \end{pmatrix}$ es ortogonal.

Solución: Hay que verificar que $A^t A = AA^t = I_2$:

$$\begin{aligned} A^t A &= \begin{pmatrix} \cos\alpha & -\operatorname{sen}\alpha \\ \operatorname{sen}\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} \cos\alpha & \operatorname{sen}\alpha \\ -\operatorname{sen}\alpha & \cos\alpha \end{pmatrix} \\ &= \begin{pmatrix} \cos^2\alpha + \operatorname{sen}^2\alpha & \cos\alpha\,\operatorname{sen}\alpha - \operatorname{sen}\alpha\,\cos\alpha \\ \operatorname{sen}\alpha\,\cos\alpha - \cos\alpha\,\operatorname{sen}\alpha & \operatorname{sen}^2\alpha + \cos^2\alpha \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned}$$

Análogamente se comprueba que $AA^t = I_2$.

◇

Observación. Si la matriz A es regular y se realiza la descomposición $A = QR$, entonces la matriz Q es una matriz ortogonal.

Si en el cambio de coordenadas visto en la Sección 5.8 se consideran las dos bases de V ortonormales, la matriz que se obtiene es del tipo anterior, como indica el siguiente resultado.

Proposición 6.19. Sea V un espacio vectorial euclídeo y sean $\mathcal{B}_1 = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$, $\mathcal{B}_2 = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ dos bases ortonormales de V. La matriz P de cambio de coordenadas asociada a dichas bases es una matriz ortogonal.

6.5.2. RESOLUCIÓN DE SISTEMAS MEDIANTE FACTORIZACIÓN QR

Considerar un sistema de ecuaciones lineales $A\mathbf{x} = \mathbf{b}$, donde $A \in \mathrm{Mat}_{n\times m}(\mathbb{R})$, $\mathbf{x} \in \mathbb{R}^m$ y $\mathbf{b} \in \mathrm{Mat}_{n\times 1}(\mathbb{R})$. El sistema anterior puede ser incompatible. En tal caso la estructura de espacio euclídeo de $\mathbb{R}^m$ permite encontrar un vector $\mathbf{x} \in \mathbb{R}^m$ que verifique que $\|A\mathbf{x} - \mathbf{b}\|_2$ sea mínima.

Observar que en el caso de que el sistema $A\mathbf{x} = \mathbf{b}$ fuese compatible, un vector $\mathbf{x} \in \mathbb{R}^m$ que cumple lo anterior es una solución del propio sistema, siendo entonces $\|A\mathbf{x} - \mathbf{b}\|_2 = 0$.

Es posible calcular este vector apoyándose en el siguiente resultado teórico:

$$\boxed{\mathbf{x} \in \mathbb{R}^m \text{ verifica } \|A\mathbf{x} - \mathbf{b}\|_2 \text{ es mínima} \Longleftrightarrow \mathbf{x} \text{ es solución de } A^tA\mathbf{x} = A^t\mathbf{b}}$$

El sistema $A^tA\mathbf{x} = A^t\mathbf{b}$ se conoce como *sistema de ecuaciones normales o de Gauss* (o sistema normal asociado al sistema $A\mathbf{x} = \mathbf{b}$). La resolución del sistema anterior puede ser tediosa puesto que, en general, se trata de un sistema de dimensiones grandes cuyos valores son elevados. Se puede hacer uso de la factorización $A = QR$ para resolver dicho sistema:

$$A^tA\mathbf{x} = A^t\mathbf{b} \iff (QR)^t(QR)\mathbf{x} = (QR)^t\mathbf{b} \iff$$

$$R^tQ^tQR\mathbf{x} = R^tQ^t\mathbf{b} \iff R^tI_rR\mathbf{x} = R^tQ^t\mathbf{b} \overset{(*)}{\iff} R\mathbf{x} = Q^t\mathbf{b}.$$

La equivalencia $(*)$ proviene del hecho de que R posee inversa a derecha ya que $\mathrm{rg}\, R$ es igual al número de filas de R.

$$\boxed{\mathbf{x} \in \mathbb{R}^m \text{ verifica } \|A\mathbf{x} - \mathbf{b}\|_2 \text{ es mínima} \Longleftrightarrow \mathbf{x} \text{ es solución de } R\mathbf{x} = Q^t\mathbf{b}}$$

✓ **Ejemplo 9.** Dado el sistema $A\mathbf{x} = \mathbf{b}$ incompatible, resolver el sistema normal asociado mediante la factorización QR:

$$\begin{pmatrix} -1 & 1/2 \\ 0 & 1 \\ 1 & 2 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 0 \\ 1 \\ 4 \end{pmatrix}.$$

Solución: El sistema normal asociado $A^tA\mathbf{x} = A^t\mathbf{b}$ es:

$$\begin{pmatrix} 2 & 3/2 \\ 3/2 & 21/4 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 4 \\ 9 \end{pmatrix}. \tag{6.1}$$

Para resolver el sistema anterior se comienza calculando la factorización QR de la matriz A. Se aplica el método de Gram-Schmidt a las columnas de la matriz $A = (\,\mathbf{a}_1 \,|\, \mathbf{a}_2\,)$, siendo $\mathbf{a}_1 = (-1, 0, 1)$ y $\mathbf{a}_2 = (\frac{1}{2}, 1, 2)$.

- $\mathbf{v}_1 = \mathbf{a}_1 = (-1, 0, 1)$.
- $\mathbf{v}_2 = \mathbf{a}_2 - \alpha_{12}\mathbf{v}_1$, donde $\alpha_{12} = \dfrac{\langle \mathbf{v}_1, \mathbf{a}_2\rangle}{\langle \mathbf{v}_1, \mathbf{v}_1\rangle}$.

$$\alpha_{12} = \frac{\langle \mathbf{v}_1, \mathbf{a}_2\rangle}{\langle \mathbf{v}_1, \mathbf{v}_1\rangle} = \frac{\langle(-1,0,1),(\frac{1}{2},1,2)\rangle}{\langle(-1,0,1),(-1,0,1)\rangle} = \frac{-\dfrac{1}{2}+0+2}{1+0+1} = \frac{3}{4}.$$

$$\mathbf{v}_2 = \mathbf{a}_2 - \frac{3}{4}\mathbf{v}_1 = \Big(\frac{1}{2}, 1, 2\Big) - \frac{3}{4}\Big(-1, 0, 1\Big) = \Big(\frac{5}{4}, 1, \frac{5}{4}\Big).$$

Por tanto, la factorización intermedia $A = Q_0R_0$ que se obtiene es la siguiente:

$$A = \underbrace{\begin{pmatrix} -1 & 5/4 \\ 0 & 1 \\ 1 & 5/4 \end{pmatrix}}_{Q_0} \underbrace{\begin{pmatrix} 1 & 3/4 \\ 0 & 1 \end{pmatrix}}_{R_0}.$$

Como no hay que quitar ninguna columna de Q_0, para conseguir la factorización QR solo hay que dividir y multiplicar por la norma de los vectores $\mathbf{v}_i$.

- Como $\|\mathbf{v}_1\| = \sqrt{2}$, se divide la primera columna de Q_0 entre $\sqrt{2}$ y se multiplica la primera fila de R_0 por $\sqrt{2}$.
- Como $\|\mathbf{v}_2\| = \sqrt{\frac{33}{8}}$, se divide la segunda columna de Q_0 entre $\sqrt{\frac{33}{8}}$ y se multiplica la segunda fila de R_0 por $\sqrt{\frac{33}{8}}$.

Así la factorización QR de A es:

$$A = \underbrace{\begin{pmatrix} \frac{-1}{\sqrt{2}} & \frac{5}{4}\sqrt{\frac{8}{33}} \\ 0 & \sqrt{\frac{8}{33}} \\ \frac{1}{\sqrt{2}} & \frac{5}{4}\sqrt{\frac{8}{33}} \end{pmatrix}}_{Q} \underbrace{\begin{pmatrix} \sqrt{2} & \frac{3}{4}\sqrt{2} \\ 0 & \sqrt{\frac{33}{8}} \end{pmatrix}}_{R}.$$

El sistema normal es equivalente al sistema $R\mathbf{x} = Q^t\mathbf{y}$:

$$\underbrace{\begin{pmatrix} \sqrt{2} & \frac{3}{4}\sqrt{2} \\ 0 & \sqrt{\frac{33}{8}} \end{pmatrix}}_{R} \underbrace{\begin{pmatrix} \alpha \\ \beta \end{pmatrix}}_{\mathbf{x}} = \underbrace{\begin{pmatrix} \frac{-1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{5}{4}\sqrt{\frac{8}{33}} & \sqrt{\frac{8}{33}} & \frac{5}{4}\sqrt{\frac{8}{33}} \end{pmatrix}}_{Q^t} \underbrace{\begin{pmatrix} 0 \\ 1 \\ 4 \end{pmatrix}}_{\mathbf{y}},$$

$$\begin{pmatrix} \sqrt{2} & \frac{3}{4}\sqrt{2} \\ 0 & \sqrt{\frac{33}{8}} \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} \frac{4}{\sqrt{2}} \\ 6\sqrt{\frac{8}{33}} \end{pmatrix}.$$

Este sistema está en forma escalonada y su resolución es muy sencilla:

- $\sqrt{\dfrac{33}{8}}\,\beta = 6\sqrt{\dfrac{8}{33}} \implies \beta = \dfrac{16}{11}$.
- $\sqrt{2}\,\alpha + \dfrac{3}{4}\sqrt{2}\,\dfrac{16}{11} = \dfrac{4}{\sqrt{2}} \implies \alpha = \dfrac{10}{11}$.

◇

Observación. Debido a que el sistema normal (6.1) del ejemplo anterior no presenta gran complejidad, se podría haber resuelto directamente utilizando el método de Gauss. Sin embargo, es importante señalar que en general el coste computacional usando la factorización QR es menor.

6.6. APLICACIONES: MÍNIMOS CUADRADOS

Uno de los problemas que se plantean habitualmente en muchas ciencias experimentales es el ajuste de un conjunto de datos por una curva.

Dado un conjunto de puntos $D = \{(x_i, y_i)\}_{i=1}^{n}$ de $\mathbb{R}^2$, se pretende encontrar una curva $\mathcal{C}$ de un tipo determinado que pase por todos ellos. En caso de que esto no sea posible se busca una curva que se ajuste al máximo a estos puntos, minimizando el error cuadrático medio. El procedimiento anterior se conoce como *método de mínimos cuadrados.*

Las curvas en el plano se pueden describir de manera general como un subconjunto de puntos de $\mathbb{R}^2$ que cumplen una ecuación implícita:

$$\mathcal{C}_f = \{(x, y) \in \mathbb{R}^2 \mid f(x, y) = 0\}.$$

En ocasiones es posible disponer de la curva anterior en forma explícita, es decir, $y = g(x)$. En el caso particular de que la curva sea una recta se tiene la *recta de regresión.*

Se va a suponer que la función que describe la curva $\mathcal{C}_f$ contiene una serie de parámetros desconocidos $\alpha_1, \ldots, \alpha_m \in \mathbb{R}$ que se deben determinar y además la dependencia respecto de ellos es lineal.

Un procedimiento para determinar los parámetros es resolver el sistema de ecuaciones lineales, $A\mathbf{x} = \mathbf{b}$, que se obtiene al imponer la condición $f(x_i, y_i) = 0$ para todos los $i = 1, \ldots, n$. Si el sistema anterior es compatible, existe una curva del tipo dado que pasa por todos los puntos de D. Por el contrario, si es incompatible, se resuelve el sistema normal asociado, $A^tA\mathbf{x} = A^t\mathbf{b}$, cuya solución proporciona los valores de los parámetros $\alpha_1, \ldots, \alpha_m$ que dan lugar a una curva $\mathcal{C}$ que es la que más se ajusta a los puntos de D por mínimos cuadrados.

✓ **Ejemplo 10.** Se consideran los siguientes pares de datos de $\mathbb{R}^2$,

$$\{(-1,0),(0,1),(1,4)\}.$$

Encontrar la función de la forma $y = \alpha x + \beta\, 2^x$ que mejor ajusta por mínimos cuadrados los datos anteriores.

Solución: Para encontrar una función $y = \alpha x + \beta 2^x$ que pase por los datos, hay que calcular α y β. Observar que esta función no es lineal en la variable x pero sí que es lineal en los parámetros α y β, que se obtienen con el siguiente sistema:

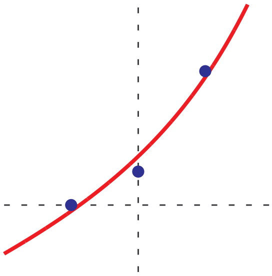
$$\left\{\begin{array}{rcc} 0 & = & -\alpha + \frac{1}{2}\beta \\ 1 & = & \beta \\ 4 & = & \alpha + 2\beta \end{array}\right. \iff \underbrace{\begin{pmatrix} -1 & 1/2 \\ 0 & 1 \\ 1 & 2 \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} \alpha \\ \beta \end{pmatrix}}_{\mathbf{x}} = \underbrace{\begin{pmatrix} 0 \\ 1 \\ 4 \end{pmatrix}}_{\mathbf{b}}.$$

Se puede comprobar que este sistema es incompatible, de modo que no hay una curva del tipo anterior que pase por los tres puntos. Se busca una aproximación por mínimos cuadrados resolviendo el sistema normal asociado $A^t A\mathbf{x} = A^t\mathbf{b}$,

$$\begin{pmatrix} 2 & 3/2 \\ 3/2 & 21/4 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 4 \\ 9 \end{pmatrix}.$$

Este sistema normal se ha resuelto en el Ejemplo 9, obteniéndose $\alpha = \frac{10}{11}$ y $\beta = \frac{16}{11}$. Por tanto, la función pedida es $y = \frac{10}{11}\, x + \frac{16}{11}\, 2^x$.

En el siguiente dibujo se muestra la representación gráfica de la función obtenida junto con los puntos de partida:



Como se observa la función no pasa por ninguno de los tres puntos pero se aproxima a ellos.

◇

6.7. COMANDOS DE *wxMaxima*

- Calcular el producto escalar usual de los vectores `u,v` de $\mathbb{R}^n$:

```
-->     load(eigen);
        innerproduct(u,v);
```

- Para calcular la norma de un vector `v` de $\mathbb{R}^n$ con el producto escalar usual, es necesario recurrir a la definición:

```
-->     load(eigen);
        sqrt(innerproduct(v,v));
```

- Normalizar un vector `v` de $\mathbb{R}^n$ con el producto escalar usual:

```
-->     load(eigen);
        unitvector(v);
```

- Calcular una base del subespacio $S^\perp$ en $\mathbb{R}^n$ con el producto escalar usual, donde $S = \langle \mathbf{v}_1, \ldots, \mathbf{v}_r \rangle$. Los vectores $\mathbf{v}_i$ hay que introducirlos como vectores columna. Para ello es útil usar el comando “`columnvector`” del paquete “`eigen`”.

```
-->     load(eigen);
        orthogonal_complement(v1,...,vr);
```

- Calcular una base ortogonal a partir de la familia de vectores linealmente independientes $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ de $\mathbb{R}^n$ con el producto escalar usual, aplicando el método de Gram-Schmidt.

```
-->     load(eigen);
        gramschmidt([v1,v2,...,vr]);
```

  También es posible ejecutar “`gramschmidt(A)`” siendo A una matriz cuyas filas son los vectores $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$. En ocasiones se puede utilizar el comando “`ratsimp`” para mejorar la expresión de la salida obtenida.

- Calcular una base ortogonal en un espacio vectorial euclídeo cualquiera, indicando previamente el producto escalar que se va a usar. Este tiene que estar definido como una función de dos variables.

```
-->     load(eigen);
        pe(u,v):= ...;
        gramschmidt([v1,v2,...,vr],pe);
```

- Dado un conjunto de datos $D = \{(x_i, y_i)\}_{i=1}^{n}$ de $\mathbb{R}^2$, para determinar la curva $\mathcal{C}$ de variables x e y, dependiendo linealmente de m parámetros $a_1, \ldots, a_m$, que mejor ajusta dichos datos por el método de mínimos cuadrados, se construye una matriz `D` cuyas filas son los datos del conjunto D, y a continuación se aplica el comando:

```
-->     load(lsquares);
        D:matrix([x1,y1], ..., [xn,yn]);
        lsquares_estimates(D, [x,y], C, [a1,...,am]);
```

6.8. EJERCICIOS PROPUESTOS

Producto escalar

1. Considerar la aplicación de $\mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ dada por:

$$\langle (x_1, x_2, x_3), (y_1, y_2, y_3) \rangle = x_1y_1 - 2x_1y_2 - 2x_2y_1 + 6x_2y_2 + x_2y_3 + x_3y_2 + x_3y_3.$$

Comprobar que $(\mathbb{R}^3, \langle \cdot, \cdot \rangle)$ es un espacio vectorial euclídeo.

2. Normalizar el vector $\mathbf{u} = (1, 2, -1) \in \mathbb{R}^3$ con respecto al producto escalar usual de $\mathbb{R}^3$ y con respecto al producto escalar definido en el Ejercicio 1.

3. Calcular el ángulo que forman los vectores $\mathbf{u} = (0, 3, -1)$ y $\mathbf{v} = (-2, 5, 1)$ en $\mathbb{R}^3$ con el producto escalar usual y con el producto escalar definido en el Ejercicio 1.

4. Determinar los valores de k para los que la aplicación:

$$\langle (x_1, x_2), (y_1, y_2) \rangle = x_1y_1 - 3x_1y_2 - 3x_2y_1 + k\, x_2y_2$$

es un producto escalar sobre $\mathbb{R}^2$.

5. Sea $V = \mathbb{R}_1[x]$ el espacio vectorial de los polinomios de grado menor o igual que 1. Demostrar que la aplicación $V \times V \to \mathbb{R}$ dada por:

$$\langle p, q \rangle = p(0)\, q(0) + p(1)\, q(1),$$

donde $p, q \in \mathbb{R}_1[x]$, es un producto escalar.

6. Sean $\mathbf{u}$ y $\mathbf{v}$ dos vectores de un espacio vectorial euclídeo V. Demostrar que si $\mathbf{u}$ es ortogonal a $\mathbf{v}$, entonces cualquier múltiplo escalar de $\mathbf{u}$ también es ortogonal a $\mathbf{v}$.

7. Se considera $\mathbb{R}^3$ con el producto escalar usual. Hallar un vector unitario ortogonal a $\mathbf{v}_1 = (1, 1, 2)$ y $\mathbf{v}_2 = (0, 1, 3)$.

Bases ortonormales

8. Considerar $(\mathbb{R}^3, \langle \cdot, \cdot \rangle)$ con el producto escalar usual. Obtener una base ortonormal de $\mathbb{R}^3$ a partir de $\{(1, 0, -1), (1, 1, 0), (1, 1, 1)\}$.

9. En cada caso, utilizar el método de Gram-Schmidt para encontrar una base ortogonal y una ortonornal del subespacio S de V que se indica:

a) $V = \mathbb{R}^3$, $S = \langle (1, 1, 1), (0, 1, 1) \rangle$.

b) $V = \mathbb{R}^4$, $S = \langle (1, 0, 1, 0), (1, 1, 1, 0), (1, 1, 0, 0) \rangle$.

10. Utilizar el método de Gram-Schmidt para convertir la base $\{1, x, x^2\}$ de $\mathbb{R}_2[x]$ en una base ortogonal usando el producto escalar que se indica:

a) $\langle p, q \rangle = p(0)q(0) + p(1)q(1) + p(2)q(2)$.

b) $\langle p, q \rangle = \int_0^2 p(x)q(x)\, dx$.

11. Obtener una base ortonormal de $\mathbb{R}^4$ para el producto escalar estándar cuyos primeros vectores sean $\mathbf{q}_1 = \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right)$ y $\mathbf{q}_2 = \left(\frac{1}{6}, \frac{1}{6}, \frac{3}{6}, \frac{-5}{6}\right)$. Repetir el ejercicio imponiendo que $\mathbf{q}_3$ tenga su primera componente nula.

12. Se considera la aplicación $f : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ dada por:

$$f((x_1, x_2, x_3), (y_1, y_2, y_3)) = \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} 3 & -1 & -1 \\ -1 & 5 & 0 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}.$$

a) Comprobar que se trata de un producto escalar.

b) Obtener una base ortonormal de $\mathbb{R}^3$ con respecto a este producto escalar.

Matrices y subespacios ortogonales

13. Hallar una matriz ortogonal A cuya primera columna sea $\mathbf{v}_1 = (\frac{1}{3}, \frac{2}{3}, \frac{2}{3})$. ¿Es única?

14. Encontrar la forma general de las matrices ortogonales de orden 2.

15. Demostrar que si una matriz ortogonal es triangular, entonces es diagonal.

16. Se considera el subespacio vectorial de $\mathbb{R}^3$, $S = \langle(1, -2, 3), (-1, 1, 1)\rangle$:

a) Hallar el subespacio $S^\perp$.

b) Escribir el vector $(1, 5, 7)$ como suma de uno de S y otro de $S^\perp$.

17. Se considera el subespacio vectorial de $\mathbb{R}^4$, $S = \langle(1, 1, 1, 1), (1, 0, 0, 1), (1, 1, -1, 3)\rangle$:

a) Hallar una base ortonormal de S.

b) Hallar una base de $S^\perp$.

Factorización QR y mínimos cuadrados

18. Encontrar la factorización QR de la matriz A en los siguientes casos:

$$a)\ A = \begin{pmatrix} 1 & -1 \\ -1 & 0 \end{pmatrix} \qquad b)\ A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

19. Resolver los siguinetes sistemas $A\mathbf{x} = \mathbf{b}$ utilizando la factorización QR de la matriz A:

a) $A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 2 \\ 0 & 2 & 5 \end{pmatrix} \quad \mathbf{b} = \begin{pmatrix} 3 \\ 3 \\ -1 \end{pmatrix}.$

b) $A = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 0 & 0 \\ 1 & 0 & -2 \\ 1 & 1 & 1 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 0 \\ 1 \\ 3 \\ -2 \end{pmatrix}.$

20. Encontrar la función de la forma $y = \alpha x^2 + \beta \operatorname{sen}(\frac{\pi x}{2})$ que mejor aproxima por mínimos cuadrados el conjunto de datos $\{(-1, 0), (0, 0), (1, 2), (2, 3)\}$. Usar la factorización QR para resolver las ecuaciones normales de Gauss.

6.9. SOLUCIONES A LOS EJERCICIOS PROPUESTOS

Producto escalar

1. Basta comprobar que la aplicación $\langle\cdot,\cdot\rangle$ verifica las tres propiedades del producto escalar. Para ver que es definida positiva se puede usar que:

$$\langle (x_1,x_2,x_3),(x_1,x_2,x_3)\rangle = (x_1-2x_2)^2+(x_2+x_3)^2+x_2^2.$$

2.
- Con respecto al producto escalar usual: $\frac{\mathbf{v}}{\|\mathbf{v}\|} = \left(\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, \frac{-1}{\sqrt{6}}\right)$.
- Con respecto al producto escalar del Ejercicio 1: $\frac{\mathbf{v}}{\|\mathbf{v}\|} = \left(\frac{1}{\sqrt{14}}, \frac{2}{\sqrt{14}}, \frac{-1}{\sqrt{14}}\right)$.

3.
- Con respecto al producto escalar usual: $\cos\alpha = \frac{14}{10\sqrt{3}}$.
- Con respecto al producto escalar del Ejercicio 1: $\cos\alpha = \frac{99}{7\sqrt{205}}$.

4. La aplicación dada es bilineal y simétrica para cualquier valor de $k\in\mathbb{R}$. Sin embargo, es definida positiva solo cuando $k>9$. Para justificar esto último se puede usar que:

$$\langle (x_1,x_2),(x_1,x_2)\rangle = (x_1-3x_2)^2+(k-9)x_2^2.$$

5. Cuestión teórica. Se trata de comprobar que la aplicación $\langle\cdot,\cdot\rangle$ verifica las tres propiedades del producto escalar.

6. Para resolver esta cuestión teórica basta usar que $\langle\lambda\mathbf{u},\mathbf{v}\rangle = \lambda\langle\mathbf{u},\mathbf{v}\rangle$, $\forall\lambda\in\mathbb{R}$, y la definición de vectores ortogonales.

7. Hay dos posibles soluciones: $\left(\frac{1}{\sqrt{11}}, \frac{-3}{\sqrt{11}}, \frac{1}{\sqrt{11}}\right)$ y su opuesto $\left(\frac{-1}{\sqrt{11}}, \frac{3}{\sqrt{11}}, \frac{-1}{\sqrt{11}}\right)$.

Bases ortonormales

8. Base ortonormal: $\mathbf{q}_1 = \left(\frac{1}{\sqrt{2}}, 0, \frac{-1}{\sqrt{2}}\right)$, $\mathbf{q}_2 = \left(\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right)$, $\mathbf{q}_3 = \left(\frac{1}{\sqrt{3}}, \frac{-1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)$.

9. a)
- Base ortogonal: $\mathbf{v}_1 = (1,1,1)$, $\mathbf{v}_2 = \left(\frac{-2}{3}, \frac{1}{3}, \frac{1}{3}\right)$.
- Base ortonormal: $\mathbf{q}_1 = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)$, $\mathbf{q}_2 = \left(\frac{-2}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right)$.

b)
- Base ortogonal: $\mathbf{v}_1 = (1,0,1,0)$, $\mathbf{v}_2 = (0,1,0,0)$, $\mathbf{v}_3 = \left(\frac{1}{2}, 0, \frac{-1}{2}, 0\right)$.
- Base ortonormal: $\mathbf{q}_1 = (\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}, 0)$, $\mathbf{q}_2 = (0,1,0,0)$, $\mathbf{q}_3 = \left(\frac{1}{\sqrt{2}}, 0, \frac{-1}{\sqrt{2}}, 0\right)$.

10. a) Base ortogonal: $\mathbf{v}_1 = 1$, $\mathbf{v}_2 = -1+x$, $\mathbf{v}_3 = \frac{1}{3} - 2x + x^2$.

b) Base ortogonal: $\mathbf{v}_1 = 1$, $\mathbf{v}_2 = -1+x$, $\mathbf{v}_3 = \frac{2}{3} - 2x + x^2$.

11. Una posibilidad es aplicar el método de Gram-Schmidt a $\{\mathbf{q}_1, \mathbf{q}_2, (0,1,0,0), (0,0,0,1)\}$. En este caso se obtiene: $\mathbf{q}_3 = \left(\frac{-5}{3\sqrt{26}}, \frac{13}{3\sqrt{26}}, \frac{-2}{\sqrt{26}}, \frac{-2}{3\sqrt{26}}\right)$, $\mathbf{q}_4 = \left(\frac{-4}{\sqrt{26}}, 0, \frac{3}{\sqrt{26}}, \frac{1}{\sqrt{26}}\right)$. La solución de la segunda parte es única salvo el signo:

$$\mathbf{q}_3 = \left(0, \frac{-4}{\sqrt{26}}, \frac{3}{\sqrt{26}}, \frac{1}{\sqrt{26}}\right),\ \mathbf{q}_4 = \left(\frac{-13}{3\sqrt{26}}, \frac{5}{3\sqrt{26}}, \frac{2}{\sqrt{26}}, \frac{2}{3\sqrt{26}}\right).$$

12. a) Basta comprobar las tres propiedades del producto escalar. Para ver que la aplicación es definida positiva observar que:

$$\langle (x_1,x_2,x_3),(x_1,x_2,x_3)\rangle = (x_1-x_2)^2+(x_1-x_3)^2+x_1^2+4x_2^2.$$

b) Se parte de la base canónica y se obtiene la siguiente base ortonormal:

$$\mathbf{q}_1=\left(\tfrac{1}{\sqrt{3}},0,0\right),\quad \mathbf{q}_2=\left(\tfrac{1}{\sqrt{42}},\tfrac{3}{\sqrt{42}},0\right),\quad \mathbf{q}_3=\left(\tfrac{5}{3\sqrt{14}},\tfrac{1}{3\sqrt{14}},\tfrac{14}{3\sqrt{14}}\right).$$

Matrices y subespacios ortogonales

13. La solución no es única. Por ejemplo, podemos tomar:

$$A=\begin{pmatrix} 1/3 & 0 & -4/\sqrt{18} \\ 2/3 & 1/\sqrt{2} & 1/\sqrt{18} \\ 2/3 & -1/\sqrt{2} & 1/\sqrt{18} \end{pmatrix}.$$

Esta matriz se ha calculado aplicando el método de Gram-Schmidt a la familia de vectores $\{(\frac{1}{3},\frac{2}{3},\frac{2}{3}),(0,1,-1),(0,1,0)\}$ y colocando los vectores obtenidos por columnas.

14. La forma general de las matrices ortogonales de orden 2 es:

$$A=\begin{pmatrix} \cos\alpha & \operatorname{sen}\alpha \\ -\operatorname{sen}\alpha & \cos\alpha \end{pmatrix}.$$

Para llegar a esta expresión hay que usar que todo par de números reales (a,b) verificando $a^2+b^2=1$ se puede escribir de la forma $a=\cos\alpha$ y $b=\operatorname{sen}\alpha$ para cierto $\alpha\in\mathbb{R}$.

15. Cuestión teórica. Se escribe una matriz triangular (cuadrada) A de dimensión n y se impone que verifique $AA^t=I_n$. Se concluye que $a_{ij}=0$, $\forall i\neq j$.

16. a) $S^\perp=\langle(5,4,1)\rangle$.

b) $(1,5,7)=s+t$ donde $s=\frac{6}{7}(1,-2,3)+\frac{11}{3}(-1,1,1)\in S,\ t=\frac{16}{21}(5,4,1)\in S^\perp$.

17. a) Base ortonormal de S:

$$\mathbf{q}_1=\left(\tfrac{1}{2},\tfrac{1}{2},\tfrac{1}{2},\tfrac{1}{2}\right),\quad \mathbf{q}_2=\left(\tfrac{1}{2},\tfrac{-1}{2},\tfrac{-1}{2},\tfrac{1}{2}\right),\quad \mathbf{q}_3=\left(\tfrac{-1}{2},\tfrac{1}{2},\tfrac{-1}{2},\tfrac{1}{2}\right).$$

b) $S^\perp=\langle(1,1,-1,-1)\rangle$.

Factorización QR y mínimos cuadrados

18. a) $A=\begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ -1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix}\begin{pmatrix} \sqrt{2} & -1/\sqrt{2} \\ 0 & 1/\sqrt{2} \end{pmatrix}.$

b) $A=\begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{6} & 1/\sqrt{2} \\ 1/\sqrt{3} & 1/\sqrt{6} & -1/\sqrt{2} \\ 1/\sqrt{3} & -2/\sqrt{6} & 0 \end{pmatrix}\begin{pmatrix} \sqrt{3} & 2/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{6} & 1/\sqrt{6} \\ 0 & 0 & 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}.$

19. a) $A = \begin{pmatrix} 1/\sqrt{2} & -\sqrt{2}/6 & 2/3 \\ 1/\sqrt{2} & \sqrt{2}/6 & -2/3 \\ 0 & 2\sqrt{2}/3 & 1/3 \end{pmatrix} \begin{pmatrix} \sqrt{2} & 3/\sqrt{2} & \sqrt{2} \\ 0 & 3/\sqrt{2} & 11\sqrt{2}/3 \\ 0 & 0 & 1/3 \end{pmatrix}$. Solución del sistema: $(1, 2, -1)$.

b) $A = \begin{pmatrix} 1/2 & 1/2 & -1/2 \\ 1/2 & -1/2 & 1/2 \\ 1/2 & -1/2 & -1/2 \\ 1/2 & 1/2 & 1/2 \end{pmatrix} \begin{pmatrix} 2 & 1 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix}$. Solución del sistema: $(1, -2, -1)$.

20. La matriz A junto con su descomposición QR y el vector $\mathbf{b}$ del sistema inicial son:

$$\underbrace{\begin{pmatrix} 1 & -1 \\ 1 & 1 \\ 4 & 0 \end{pmatrix}}_{A} = \underbrace{\begin{pmatrix} \sqrt{2}/6 & -\sqrt{2}/2 \\ \sqrt{2}/6 & \sqrt{2}/2 \\ 2\sqrt{2}/3 & 0 \end{pmatrix}}_{Q} \underbrace{\begin{pmatrix} 3\sqrt{2} & 0 \\ 0 & \sqrt{2} \end{pmatrix}}_{R}, \qquad \mathbf{b} = \begin{pmatrix} 0 \\ 2 \\ 3 \end{pmatrix}.$$

El sistema $R\mathbf{x} = Q^t\mathbf{b}$ que hay que resolver es:

$$\begin{pmatrix} 3\sqrt{2} & 0 \\ 0 & \sqrt{2} \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 14/3\sqrt{2} \\ 2/\sqrt{2} \end{pmatrix}$$

cuya solución es $\alpha = \frac{7}{9}$, $\beta = 1$.

CAPÍTULO 7

Aplicaciones Lineales

El concepto de aplicación lineal, tal y como ahora se conoce, se debe a Peano al axiomatizar la definición de espacio vectorial. No obstante, el nacimiento de las aplicaciones lineales hay que buscarlo en los intentos iniciales de Descartes, y posteriores de Euler y Lagrange, por “algebrizar” la geometría. Cayley fue quien hizo patente la relación entre las matrices y las aplicaciones lineales al escribir de forma matricial las ecuaciones de los diferentes tipos de transformaciones geométricas.

Hoy día las aplicaciones lineales son importantes tanto en Matemáticas como en otras ciencias. Además de modelizar las transformaciones geométricas del plano y del espacio (como simetrías, traslaciones y rotaciones), sirven para resolver problemas complicados que se aproximan mediante un proceso de linealización.

Las aplicaciones lineales se definen entre dos espacios vectoriales y la propiedad que las caracteriza es que preservan la estructura de espacio vectorial, es decir, preservan la suma de vectores y el producto por escalares. El ejemplo más simple de aplicación lineal corresponde a la ecuación de una recta en el plano de pendiente m que pasa por el origen de coordenadas:

$$\begin{array}{rccl} f: & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & x & \mapsto & f(x) = mx. \end{array}$$

Esta aplicación f cumple dos propiedades:

$$f(x_1 + x_2) = m(x_1 + x_2) = mx_1 + mx_2 = f(x_1) + f(x_2).$$

$$f(\lambda x) = m(\lambda x) = \lambda(mx) = \lambda f(x).$$

Las aplicaciones lineales entre espacios vectoriales generalizan este ejemplo.

7.1. DEFINICIONES Y EJEMPLOS

A lo largo de todo el capítulo se supone que V y W son espacios vectoriales reales con dimensión finita, $\dim V = m$ y $\dim W = n$.

Definición 7.1. Sean V y W dos espacios vectoriales reales. Una aplicación $f : V \to W$ se dice **lineal** (u **homomorfismo**) si verifica las dos condiciones siguientes:

1. $f(\mathbf{v}_1 + \mathbf{v}_2) = f(\mathbf{v}_1) + f(\mathbf{v}_2), \quad \forall \mathbf{v}_1, \mathbf{v}_2 \in V.$
2. $f(\lambda \mathbf{v}) = \lambda f(\mathbf{v}), \quad \forall \lambda \in \mathbb{R},\ \forall \mathbf{v} \in V.$

Definición 7.2. Una aplicación lineal de un espacio vectorial V en sí mismo, $f : V \to V$, se denomina **endomorfismo**.

Notas.

- Las aplicaciones lineales preservan la estructura de espacio vectorial: transforman sumas en sumas y son compatibles con el producto por escalares.
- Las dos condiciones que definen las aplicaciones lineales se pueden resumir en una sola:

$$\boxed{f(\lambda \mathbf{v}_1 + \mu \mathbf{v}_2) = \lambda f(\mathbf{v}_1) + \mu f(\mathbf{v}_2), \quad \forall \lambda, \mu \in \mathbb{R}, \quad \forall \mathbf{v}_1, \mathbf{v}_2 \in V}$$

Observación. En la definición de aplicación lineal hay que entender correctamente las operaciones que se utilizan en cada momento. Más detalladamente, si se denotan por $(V, +_V, \cdot_V)$ y $(W, +_W, \cdot_W)$ las ternas completas que definen la estructura de espacio vectorial en V y W respectivamente, la definición queda como sigue:

1. $f(\mathbf{v}_1 +_V \mathbf{v}_2) = f(\mathbf{v}_1) +_W f(\mathbf{v}_2), \quad \forall \mathbf{v}_1, \mathbf{v}_2 \in V.$
2. $f(\lambda \cdot_V \mathbf{v}) = \lambda \cdot_W f(\mathbf{v}), \quad \forall \lambda \in \mathbb{R},\ \forall \mathbf{v} \in V.$

Se presentan a continuación algunos ejemplos de cómo estudiar si una aplicación es lineal o no.

✓ **Ejemplo 1.** Probar que la siguiente aplicación es lineal

$$\begin{array}{rccc} f: & \mathbb{R}_2[x] & \longrightarrow & \mathbb{R}^2 \\ & a + bx + cx^2 & \mapsto & (a+b,\ a-b+c). \end{array}$$

Solución: Para ver que f es lineal, hay que comprobar las dos condiciones de la definición:

1. Sean $\mathbf{v}_1 = a + bx + cx^2$, $\mathbf{v}_2 = a' + b'x + c'x^2 \in \mathbb{R}_2[x]$. Aplicando f a ambos vectores se obtiene:

$$f(\mathbf{v}_1) = f\big(a + bx + cx^2\big) = (a+b, a-b+c),$$
$$f(\mathbf{v}_2) = f\big(a' + b'x + c'x^2\big) = (a'+b', a'-b'+c').$$

Sumando las dos imágenes:

$$\begin{aligned} f(\mathbf{v}_1) + f(\mathbf{v}_2) \quad &= (a+b, a-b+c) + (a'+b', a'-b'+c') \\ &= (a+a'+b+b',\ a+a'-b-b'+c+c'). \end{aligned}$$

Por otra parte, primero se suman los vectores:

$$\mathbf{v}_1 + \mathbf{v}_2 = (a+bx+cx^2) + (a'+b'x+c'x^2) = (a+a') + (b+b')x + (c+c')x^2,$$

y después se aplica f:

$$\begin{aligned} f(\mathbf{v}_1 + \mathbf{v}_2) \quad &= ((a+a') + (b+b'),\ (a+a') - (b+b') + (c+c')) \\ &= (a+a'+b+b',\ a+a'-b-b'+c+c'). \end{aligned}$$

Se observa que efectivamente $f(\mathbf{v}_1 + \mathbf{v}_2) = f(\mathbf{v}_1) + f(\mathbf{v}_2)$.

2. Sea $\mathbf{v} = a + bx + cx^2 \in \mathbb{R}_2[x]$ y $\lambda \in \mathbb{R}$.

$$\begin{aligned} f(\lambda\mathbf{v}) = f\big(\lambda a + \lambda bx + \lambda cx^2\big) &= (\lambda a + \lambda b, \lambda a - \lambda b + \lambda c) = \\ &= (\lambda\,(a+b),\ \lambda\,(a-b+c)) = \lambda(a+b, a-b+c) = \lambda f(\mathbf{v}). \end{aligned}$$

De las dos condiciones anteriores se deduce que f es una aplicación lineal.

◇

✓ **Ejemplo 2.** Comprobar que la siguiente aplicación no es lineal.

$$\begin{array}{cccc} g: & \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ & (x,y) & \mapsto & (x, y+1). \end{array}$$

Solución: Para demostrar que g no es lineal, basta encontrar un ejemplo concreto que no satisfaga alguna de las dos condiciones que caracterizan a las aplicaciones lineales. Por ejemplo, se puede encontrar un vector $\mathbf{v}$ y un escalar λ de modo que falle la condición 2 de la definición. En efecto, tomando $\mathbf{v} = (1,0)$ y $\lambda = 2$, se tiene que:

$$g(2\mathbf{v}) = g(2,0) = (2,1), \qquad 2g(\mathbf{v}) = 2g(1,0) = 2(1,1) = (2,2).$$

Como $g(2\mathbf{v}) \neq 2g(\mathbf{v})$, la segunda condición de la definición no se cumple y por tanto g no es lineal.

◇

A continuación se muestra la linealidad de algunas aplicaciones ya conocidas.

✓ **Ejemplo 3.**

1. Sea V un espacio vectorial real y $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ una base de V. La aplicación coordenada $\mathcal{C}_\mathcal{B}$ es lineal.

$$\begin{array}{rcl} \mathcal{C}_\mathcal{B} : V & \longrightarrow & \mathbb{R}^m \\ \mathbf{v} & \mapsto & \mathbf{v}_\mathcal{B} = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_m \end{pmatrix}, \quad \mathbf{v} = \lambda_1 \mathbf{v}_1 + \cdots + \lambda_m \mathbf{v}_m. \end{array}$$

2. Sean $A \in \mathrm{Mat}_{n\times m}(\mathbb{R})$ una matriz, $V = \mathbb{R}^m$ y $W = \mathbb{R}^n$. La siguiente aplicación es lineal:

$$\begin{array}{rccc} f: & \mathbb{R}^m & \longrightarrow & \mathbb{R}^n \\ & \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} & \mapsto & \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = A \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}. \end{array}$$

3. Sean V y W los espacios de funciones siguientes:

$$V = \{f : [a,b] \longrightarrow \mathbb{R} \mid f \text{ es derivable}\}, \quad W = \{f : [a,b] \longrightarrow \mathbb{R}\}.$$

La aplicación que a cada función $f \in V$ le asocia su derivada es lineal.

$$\begin{array}{rcl} \Phi : V & \longrightarrow & W \\ f & \mapsto & f' = \dfrac{df}{dx}. \end{array}$$

4. Sean V, W, Z tres espacios vectoriales reales. Considerar las aplicaciones lineales

$$f : V \longrightarrow W, \qquad g : W \longrightarrow Z.$$

La composición $g \circ f : V \to Z$ definida como

$$\begin{array}{rccccc} g \circ f : & V & \xrightarrow{f} & W & \xrightarrow{g} & Z \\ & \mathbf{v} & \mapsto & f(\mathbf{v}) & \mapsto & g(f(\mathbf{v})), \end{array}$$

es una aplicación lineal. Notar que primero actúa f y después g.

◇

⇝ **Ejercicio.** Demostrar la linealidad de los ejemplos anteriores.

Como consecuencia directa de la Definición 7.1 se tienen:

Propiedades. Sea $f : V \to W$ una aplicación lineal entre dos espacios vectoriales reales. Se verifican las siguientes propiedades:

1. $f(\mathbf{0}_V) = \mathbf{0}_W$.
2. $f(-\mathbf{v}) = -f(\mathbf{v}), \quad \forall \mathbf{v} \in V$.
3. $f(\mathbf{u} - \mathbf{v}) = f(\mathbf{u}) - f(\mathbf{v}), \quad \forall \mathbf{u}, \mathbf{v} \in V$.
4. $f(\lambda_1 \mathbf{v}_1 + \cdots + \lambda_k \mathbf{v}_k) = \lambda_1 f(\mathbf{v}_1) + \cdots + \lambda_k f(\mathbf{v}_k)$.

   En particular, si el conjunto de vectores $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} \subset V$ es linealmente dependiente, entonces $\{f(\mathbf{v}_1), \ldots, f(\mathbf{v}_k)\} \subset W$ también lo es.

   $$\boxed{\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} \text{ l.d.} \Longrightarrow \{f(\mathbf{v}_1), \ldots, f(\mathbf{v}_k)\} \text{ l.d.}}$$

   Equivalentemente, si el conjunto $\{f(\mathbf{v}_1), \ldots, f(\mathbf{v}_k)\} \subset W$ es linealmente independiente, entonces también lo es $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$.

   $$\boxed{\{f(\mathbf{v}_1), \ldots, f(\mathbf{v}_k)\} \text{ l.i.} \Longrightarrow \{\mathbf{v}_1, \ldots, \mathbf{v}_k\} \text{ l.i.}}$$

5. Si $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ es una base de V y $\{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ son m elementos arbitrarios de W, existe una única aplicación lineal $f : V \to W$ tal que

   $$f(\mathbf{v}_i) = \mathbf{w}_i, \quad \forall i = 1, \ldots, m.$$

   Es decir, para definir una aplicación lineal basta conocer la imagen de los elementos de una base de V.

$\rightsquigarrow$ **Ejercicio.** Demostrar las propiedades anteriores.

✓ **Ejemplo 4.** En el Ejemplo 2, para demostrar que g no es lineal basta observar que no se verifica la propiedad 1 anterior. Efectivamente, se tiene que, $g(0,0) = (0,1) \neq (0,0)$.

◇

Se considera el conjunto de todas las aplicaciones lineales entre dos espacios vectoriales reales:

Definición 7.3. Sean V y W dos espacios vectoriales reales.

- El conjunto de todas la aplicaciones lineales de V en W se denota por $Hom(V, W)$, es decir:

  $$Hom(V, W) = \{f : V \to W \mid f \text{ lineal}\}.$$

- En el caso $V = W$, el conjunto de todos los endomorfismos de V se denota por $End(V)$, es decir:

  $$End(V) = Hom(V, V) = \{f : V \to V \mid f \text{ lineal}\}.$$

Observación. Los conjuntos anteriores tienen estructura de espacio vectorial con las operaciones habituales entre aplicaciones lineales.

✓ **Ejemplo 5.**

1. El **homomorfismo nulo** viene dado por:

$$\begin{array}{rcl} 0_{Hom(V,W)} : V & \longrightarrow & W \\ \mathbf{v} & \mapsto & \mathbf{0}_{W.} \end{array}$$

2. El **endomorfismo identidad** es:

$$\begin{array}{rcl} Id_V : V & \longrightarrow & V \\ \mathbf{v} & \mapsto & \mathbf{v}. \end{array}$$

◇

7.1.1. SUBESPACIOS ASOCIADOS A UNA APLICACIÓN LINEAL

A continuación se estudia el comportamiento de las aplicaciones lineales de V a W al actuar sobre subespacios vectoriales de V y W.

Proposición 7.4. Sea $f : V \to W$ una aplicación lineal. Se tiene:

- Si U es un subespacio vectorial de V, entonces la *imagen* de U, $f(U) = \{f(\mathbf{u}) \,|\, \mathbf{u} \in U\}$, es un subespacio vectorial de W.
- Si Z es un subespacio vectorial de W, entonces la *preimagen* de Z, $f^{-1}(Z) = \{\mathbf{v} \in V \,|\, f(\mathbf{v}) \in Z\}$, es un subespacio vectorial de V.

Definición 7.5. Sea $f : V \to W$ una aplicación lineal. Sea $U \subset V$ un subespacio vectorial de V. Se define la **restricción de f a U** como la aplicación lineal:

$$\begin{array}{rccc} f|_U : & U & \longrightarrow & W \\ & \mathbf{u} & \mapsto & f(\mathbf{u}). \end{array}$$

✓ **Ejemplo 6.** Considerar la aplicación lineal dada por:

$$\begin{array}{rccc} f : & \mathrm{Mat}_2(\mathbb{R}) & \longrightarrow & \mathrm{Mat}_2(\mathbb{R}) \\ & \begin{pmatrix} a & b \\ c & d \end{pmatrix} & \mapsto & \begin{pmatrix} a+b & b+c \\ c+d & d+a \end{pmatrix}. \end{array}$$

Sea S el subespacio de las matrices simétricas de orden 2. Calcular $f(S)$, dar la expresión de $f|_S$ y determinar $f^{-1}(S)$.

Solución: Las matrices de S son de la forma $\begin{pmatrix} a & b \\ b & d \end{pmatrix}$, es decir, $c = b$. Así,

$$f(S) = \left\{ f\begin{pmatrix} a & b \\ b & d \end{pmatrix} \,\middle|\, a, b, d \in \mathbb{R} \right\} = \left\{ \begin{pmatrix} a+b & 2b \\ b+d & d+a \end{pmatrix} \,\middle|\, a, b, d \in \mathbb{R} \right\} =$$

$$\left\{ a\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + b\begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix} + d\begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} \,\middle|\, a, b, d \in \mathbb{R} \right\} = \left\langle \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} \right\rangle.$$

La restricción $f|_S$ viene dada por:

$$\begin{array}{cccc} f: & S & \longrightarrow & f(S) \\ & \begin{pmatrix} a & b \\ b & d \end{pmatrix} & \mapsto & \begin{pmatrix} a+b & 2b \\ b+d & d+a \end{pmatrix}. \end{array}$$

Para calcular $f^{-1}(S)$ se buscan matrices $A \in \mathrm{Mat}_2(\mathbb{R})$ de modo que $f(A)$ sea simétrica. Una matriz $\begin{pmatrix} a+b & b+c \\ c+d & d+a \end{pmatrix}$ es simétrica si y solo si $b+c = c+d$, es decir, $b = d$. Así,

$$f^{-1}(S) = \left\{ \begin{pmatrix} a & b \\ c & b \end{pmatrix} \,\middle|\, a, b, c \in \mathbb{R} \right\} = \left\langle \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \right\rangle.$$

◇

NÚCLEO E IMAGEN DE UNA APLICACIÓN LINEAL

Se presentan a continuación dos subespacios vectoriales que son importantes para el estudio de las aplicaciones lineales.

Definición 7.6. Sea $f : V \to W$ una aplicación lineal. Se definen los siguientes conjuntos:

1. **Núcleo** de f:
$$\operatorname{Ker} f = \{\mathbf{v} \in V \mid f(\mathbf{v}) = \mathbf{0}_W\} \subset V.$$
2. **Imagen** de f:
$$\operatorname{Im} f = \{f(\mathbf{v}) \,|\, \mathbf{v} \in V\} = \{\mathbf{w} \in W \mid \exists \mathbf{v} \in V,\ f(\mathbf{v}) = \mathbf{w}\} = f(V) \subset W.$$

Observaciones.

- Se tiene que $\operatorname{Ker} f = f^{-1}(\{\mathbf{0}_W\})$, es decir, $f|_{\operatorname{Ker} f}$ es la aplicación nula.
- $\operatorname{Im} f = f(V)$.

Proposición 7.7. Sea $f : V \to W$ una aplicación lineal. Entonces:

1. $\operatorname{Ker} f$ es un subespacio vectorial de V.
2. $\operatorname{Im} f$ es un subespacio vectorial de W.

Demostración. Aunque el resultado es inmediato a la vista de la Proposición 7.4, se demuestra este resultado de manera independiente. Recordar que un subconjunto de un espacio vectorial tiene estructura de subespacio vectorial si es no vacío y cerrado para la suma y el producto por escalares.

$\boxed{\operatorname{Ker} f \subset V}$

- El conjunto $\operatorname{Ker} f$ es no vacío ya que el elemento $\mathbf{0}_V \in \operatorname{Ker} f$: En efecto, $f(\mathbf{0}_V) = \mathbf{0}_W$ por ser f una aplicación lineal.
- Sean $\mathbf{v}_1$, $\mathbf{v}_2 \in \operatorname{Ker} f$ y $\alpha, \beta \in \mathbb{R}$. El elemento $\mathbf{v} = \alpha\mathbf{v}_1 + \beta\mathbf{v}_2$ pertenece también a $\operatorname{Ker} f$ ya que $f(\mathbf{v}) \ = \mathbf{0}_W$:

  $$f(\mathbf{v}) = f(\alpha\mathbf{v}_1 + \beta\mathbf{v}_2) \overset{(1)}{=} \alpha f(\mathbf{v}_1) + \beta f(\mathbf{v}_2) \overset{(2)}{=} \alpha\mathbf{0}_W + \beta\mathbf{0}_W = \mathbf{0}_W + \mathbf{0}_W = \mathbf{0}_W,$$

  donde en (1) se utiliza que f es una aplicación lineal y en (2) la hipótesis de que los vectores $\mathbf{v}_1$ y $\mathbf{v}_2$ pertenecen al núcleo de f.

Las propiedades anteriores indican que $\operatorname{Ker} f$ es un subespacio vectorial de V.

$\boxed{\operatorname{Im} f \subset W}$

- El conjunto $\operatorname{Im} f$ es no vacío ya que el elemento $\mathbf{0}_W \in \operatorname{Im} f$: En efecto, $\mathbf{0}_W = f(\mathbf{0}_V)$ por ser f una aplicación lineal.
- Sean $\mathbf{w}_1$, $\mathbf{w}_2 \in \operatorname{Im} f$ y $\alpha, \beta \in \mathbb{R}$. El elemento $\mathbf{w} = \alpha\mathbf{w}_1 + \beta\mathbf{w}_2$ pertenece también a $\operatorname{Im} f$ ya que $\exists \mathbf{v} \in V$ tal que $f(\mathbf{v}) \ = \mathbf{w}$:

  $$\mathbf{w}_1 \in \operatorname{Im} f \Longrightarrow \exists \mathbf{v}_1 \text{ cumpliendo } f(\mathbf{v}_1) = \mathbf{w}_1.$$

  $$\mathbf{w}_2 \in \operatorname{Im} f \Longrightarrow \exists \mathbf{v}_2 \text{ cumpliendo } f(\mathbf{v}_2) = \mathbf{w}_2.$$

  Sea $\mathbf{v} = \alpha\mathbf{v}_1 + \beta\mathbf{v}_2$. Entonces:

  $$f(\mathbf{v}) = f(\alpha\mathbf{v}_1 + \beta\mathbf{v}_2) \overset{(1)}{=} \alpha f(\mathbf{v}_1) + \beta f(\mathbf{v}_2) = \alpha\mathbf{w}_1 + \beta\mathbf{w}_2 = \mathbf{w},$$

  donde en (1) se utiliza que f es una aplicación lineal.

Por tanto, se concluye que $\operatorname{Im} f$ es un subespacio vectorial de W.

□

Observaciones.

- Si f es el homomorfismo nulo:

$$\operatorname{Ker} f = V \quad \operatorname{Im} f = \{\mathbf{0}_W\}.$$

- Para cualquier aplicación lineal f se tiene que:

$$\dim(\operatorname{Ker} f) \leq \dim V, \quad \dim(\operatorname{Im} f) \leq \dim W.$$

- Un sistema generador de $\operatorname{Im} f$ viene dado por la imagen de una base de V.

En el siguiente ejemplo se muestra cómo calcular explícitamente los subespacios núcleo e imagen de una aplicación lineal.

✓ **Ejemplo 7.** Dada la aplicación lineal

$$\begin{array}{cccc} f: & \mathbb{R}_2[x] & \longrightarrow & \mathbb{R}^2 \\ & a+bx+cx^2 & \mapsto & (a+b,\ a-b+c) \end{array}$$

calcular una base y dimensión del núcleo y de la imagen.
Solución:

$\boxed{\operatorname{Ker} f}$ Se expresa $\operatorname{Ker} f$ a través de un sistema generador:

$$\begin{aligned} \operatorname{Ker} f &= \{\mathbf{v} \in \mathbb{R}_2[x] \mid f(\mathbf{v}) = \mathbf{0}\} \\ &= \left\{a+bx+cx^2 \in \mathbb{R}_2[x] \mid f(a+bx+cx^2) = (0,0)\right\} \\ &= \left\{a+bx+cx^2 \in \mathbb{R}_2[x] \mid (a+b, a-b+c) = (0,0)\right\} \\ &= \left\{a+bx+cx^2 \in \mathbb{R}_2[x] \;\middle|\; \begin{array}{r} a+b=0 \\ a-b+c=0 \end{array}\right\} \\ &= \left\{a+bx+cx^2 \in \mathbb{R}_2[x] \;\middle|\; \begin{array}{c} a=-b \\ c=2b \end{array}\right\} \\ &= \left\{-b+bx+2bx^2 \in \mathbb{R}_2[x] \;\middle|\; b \in \mathbb{R}\right\} \\ &= \left\{b(-1+x+2x^2) \;\middle|\; b \in \mathbb{R}\right\} \\ &= \langle -1+x+2x^2 \rangle. \end{aligned}$$

Así, $\operatorname{Ker} f = \langle -1+x+2x^2 \rangle \subset \mathbb{R}_2[x]$. El subespacio núcleo está generado por un único polinomio no nulo, por lo que:

$$\mathcal{B}_{\operatorname{Ker} f} = \{-1+x+2x^2\}, \quad \dim(\operatorname{Ker} f) = 1.$$

$\boxed{\operatorname{Im} f}$ Se obtiene en primer lugar un sistema generador de $\operatorname{Im} f$:

$$\begin{aligned} \operatorname{Im} f &= \{(a+b,\ a-b+c) \in \mathbb{R}^2 \mid a,b,c \in \mathbb{R}\} \\ &= \{a(1,1)+b(1,-1)+c(0,1) \in \mathbb{R}^2 \mid a,b,c \in \mathbb{R}\} \\ &= \langle (1,1),(1,-1),(0,1) \rangle. \end{aligned}$$

Eliminando el segundo vector del sistema generador se obtiene una base:

$$\mathcal{B}_{\mathrm{Im}\, f} = \{(1,1),(0,1)\}, \quad \dim(\mathrm{Im}\, f) = 2.$$

Como la dimensión del subespacio imagen alcanza la dimensión del espacio de llegada, $\mathrm{Im}\, f = \mathbb{R}^2$.

◇

Se presenta a continuación cómo calcular el núcleo y la imagen en un caso particular importante de aplicaciones lineales.

Observación. Sea $f : \mathbb{R}^m \to \mathbb{R}^n$ una aplicación lineal dada por:

$$\begin{array}{rcl} f : \mathbb{R}^m & \longrightarrow & \mathbb{R}^n \\ \mathbf{x} & \mapsto & \mathbf{y} = A\mathbf{x}, \end{array}$$

donde $A \in \mathrm{Mat}_{n\times m}(\mathbb{R})$. Entonces el núcleo y la imagen de f se pueden calcular de manera sencilla:

- $\mathbf{x} \in \mathrm{Ker}\, f \Longleftrightarrow \mathbf{x}$ verifica $A\mathbf{x} = \mathbf{0}$.
- $\mathbf{b} \in \mathrm{Im}\, f \Longleftrightarrow$ el sistema $A\mathbf{x} = \mathbf{b}$ es compatible $\overset{(*)}{\Longleftrightarrow}$ $\mathbf{b}$ pertenece al espacio vectorial generado por las columnas de la matriz A (ver Definición 5.30).

  Demostración de $(*)$: El sistema $A\mathbf{x} = \mathbf{b}$ es compatible $\Longleftrightarrow \exists x_1, \ldots, x_m \in \mathbb{R}$ tales que

$$\left(\begin{array}{c|c|c} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{array}\right)\begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \Longleftrightarrow \left(\mathbf{c}_1 \mid \cdots \mid \mathbf{c}_m\right)\begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} = \mathbf{b} \Longleftrightarrow$$

$$\mathbf{c}_1 x_1 + \ldots + \mathbf{c}_m x_m = \mathbf{b} \Longleftrightarrow \mathbf{b} \in \langle \mathbf{c}_1, \, \ldots, \, \mathbf{c}_m \rangle \, ,$$

  es decir, $\mathbf{b}$ pertenece al espacio vectorial generado por las columnas de A.

La observación anterior tiene una gran aplicación a la hora de resolver problemas. De hecho, en las secciones siguientes se muestra cómo obtener una matriz asociada a una aplicación lineal entre espacios vectoriales genéricos V y W de dimensión finita.

✓ **Ejemplo 8.** Calcular el núcleo y la imagen de la siguiente aplicación lineal:

$$\begin{array}{rccc} f : & \mathbb{R}^3 & \longrightarrow & \mathbb{R}^2 \\ & \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} & \mapsto & \underbrace{\begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix}}_{A}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}. \end{array}$$

Solución:

$\boxed{\operatorname{Ker} f}$ Se busca la solución del sistema $A\mathbf{x} = \mathbf{0}$.

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Este sencillo sistema se puede resolver de manera directa o utilizando el método de Gauss. Utilizando esta última opción:

$$\left(\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 1 & -1 & 1 & 0 \end{array}\right) \xrightarrow[P_{21}(-1)]{} \left(\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 0 & -2 & 1 & 0 \end{array}\right).$$

Tomando como parámetro $x_2 = \lambda$ la solución que se obtiene es:

$$(x_1, x_2, x_3) = (-\lambda, \lambda, 2\lambda) \Longrightarrow \operatorname{Ker} f = \langle(-1, 1, 2)\rangle.$$

$\boxed{\operatorname{Im} f}$ Es el subespacio vectorial generado por las columnas de A.

$$\operatorname{Im} f = \left\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\rangle = \mathbb{R}^2.$$

◇

En general, para cualquier aplicación lineal f se cumple la siguiente relación:

Teorema 7.8 (Teorema de las dimensiones)**.** Sean V y W dos espacios vectoriales reales de dimensión finita y sea $f : V \to W$ una aplicación lineal entre ellos. Se verifica:

$$\boxed{\dim V = \dim(\operatorname{Ker} f) + \dim(\operatorname{Im} f)}$$

Este teorema es muy importante y tiene gran aplicación, al igual que ocurría con la fórmula de las dimensiones (Teorema 5.25). Observar que se cumple en los Ejemplos 7 y 8.

7.1.2. TIPOS DE APLICACIONES

Definición 7.9. Sea $f : V \to W$ una aplicación cualquiera no necesariamente lineal. Se dice que f es:

- **Inyectiva** si $\forall \mathbf{u}, \mathbf{v} \in V$ tales que $f(\mathbf{u}) = f(\mathbf{v})$, entonces $\mathbf{u} = \mathbf{v}$.
- **Suprayectiva** si $\forall \mathbf{w} \in W$ existe $\mathbf{v} \in V$ tal que $f(\mathbf{v}) = \mathbf{w}$.
- **Biyectiva** si es inyectiva y suprayectiva. En este caso $\exists f^{-1} : W \to V$ cumpliendo:

$$\begin{cases} f^{-1} \circ f = Id_V \\ f \circ f^{-1} = Id_{W.} \end{cases}$$

Notación. A lo largo de este texto se siguen las siguientes convenciones:

- **Homomorfismo:** Aplicación lineal.
- **Monomorfismo:** Aplicación lineal inyectiva.
- **Epimorfismo:** Aplicación lineal suprayectiva.
- **Isomorfismo:** Aplicación lineal biyectiva.
- **Endomorfismo:** Aplicación lineal de un espacio V en sí mismo, $f : V \to V$.
- **Automorfismo:** Endomorfismo biyectivo, es decir, aplicación lineal biyectiva de un espacio en sí mismo.

A continuación se caracterizan los tipos de aplicaciones lineales utilizando los subespacios núcleo e imagen.

Teorema 7.10. Sea $f : V \to W$ una aplicación lineal. Entonces:

1. f es inyectiva $\Longleftrightarrow \operatorname{Ker} f = \{\mathbf{0}_V\}$.
2. f es suprayectiva $\Longleftrightarrow \operatorname{Im} f = W$.

Demostración. La demostración de cada parte del teorema consta de dos implicaciones:

1. ($\Longrightarrow$) Sea f lineal e inyectiva y sea $\mathbf{v} \in \operatorname{Ker} f$. Se tiene que:
   - $\mathbf{v} \in \operatorname{Ker} f \Longrightarrow f(\mathbf{v}) = \mathbf{0}_W$.
   - Como f es lineal: $f(\mathbf{0}_V) = \mathbf{0}_W$.

   Por tanto, existen dos vectores de V, $\mathbf{v}$ y $\mathbf{0}_V$, que tienen la misma imagen. Por ser f inyectiva ocurre que $\mathbf{v} = \mathbf{0}_V$.

   ($\Longleftarrow$) Sea f lineal con $\operatorname{Ker} f = \{\mathbf{0}_V\}$ y sean $\mathbf{v}_1$, $\mathbf{v}_2 \in V$ tales que $f(\mathbf{v}_1) = f(\mathbf{v}_2)$. Veamos que $\mathbf{v}_1 = \mathbf{v}_2$:

   $$f(\mathbf{v}_1) = f(\mathbf{v}_2) \Longrightarrow \mathbf{0}_W = f(\mathbf{v}_1) - f(\mathbf{v}_2) = f(\mathbf{v}_1 - \mathbf{v}_2) \Longrightarrow \mathbf{v}_1 - \mathbf{v}_2 \in \operatorname{Ker} f.$$

   Como $\operatorname{Ker} f = \{\mathbf{0}_V\}$, entonces, $\mathbf{v}_1 - \mathbf{v}_2 = \mathbf{0}_V \Longrightarrow \mathbf{v}_1 = \mathbf{v}_2$, y por tanto f es inyectiva.

2. ($\Longrightarrow$) Sea f lineal y suprayectiva. Para ver que $\operatorname{Im} f = W$ basta ver que $W \subseteq \operatorname{Im} f$ puesto que siempre ocurre que $\operatorname{Im} f \subseteq W$.
   Sea $\mathbf{w} \in W$. Por ser f suprayectiva, existe $\mathbf{v} \in V$ de modo que $\mathbf{w} = f(\mathbf{v})$, es decir, $\mathbf{w} \in \operatorname{Im} f$.

   ($\Longleftarrow$) Sea f lineal con $\operatorname{Im} f = W$. Sea $\mathbf{w} \in W = \operatorname{Im} f = \{f(\mathbf{v}) \,|\, \mathbf{v} \in V\}$. Entonces, existe $\mathbf{v} \in V$ tal que $f(\mathbf{v}) = \mathbf{w}$, es decir, f es suprayectiva.

$\square$

Otra caracterización de las aplicaciones lineales inyectivas, suprayectivas y biyectivas viene dada por:

Teorema 7.11. Sea $f : V \to W$ una aplicación lineal entre dos espacios vectoriales reales. Entonces:

- f es inyectiva $\Longleftrightarrow$ f transforma familias linealmente independientes de V en familias linealmente independientes de W.
- f es suprayectiva $\Longleftrightarrow$ f transforma sistemas generadores de V en sistemas generadores de W.
- f es biyectiva $\Longleftrightarrow$ f transforma bases de V en bases de W.

Se termina esta sección con la noción de "igualdad" entre espacios vectoriales:

Definición 7.12. Dos espacios vectoriales V y W se dicen **isomorfos**, y se denota $V \cong W$, si existe una aplicación lineal biyectiva, o isomorfismo, entre ellos.

Teorema 7.13. Dos espacios vectoriales reales V y W de dimensión finita son isomorfos $\Longleftrightarrow$ $\dim V = \dim W$.

Proposición 7.14. Si V es un espacio vectorial real de dimensión n, entonces $V \cong \mathbb{R}^n$. Fijada una base $\mathcal{B}$ de V la aplicación coordenada $\mathcal{C}_{\mathcal{B}} : V \to \mathbb{R}^n$ es un isomorfismo.

7.2. MATRIZ COORDENADA DE UN HOMOMORFISMO

Sean V y W dos espacios vectoriales reales con $\dim V = m$ y $\dim W = n$ y sea $f : V \to W$ una aplicación lineal. El problema que se plantea es buscar una matriz $A \in \operatorname{Mat}_{n\times m}(\mathbb{R})$ que represente dicha aplicación; equivalentemente, se busca una aplicación lineal

$$\begin{array}{rccl} f_A : & \mathbb{R}^m & \longrightarrow & \mathbb{R}^n \\ & \mathbf{x} & \mapsto & f_A(\mathbf{x}) = A\mathbf{x}, \end{array}$$

de modo que a través de A se puedan estudiar propiedades de f.

Para ello lo primero que hay que hacer es fijar una base $\mathcal{B}_V$ del espacio de partida V y otra $\mathcal{B}_W$ del espacio de llegada W con el fin de relacionar las coordenadas de un vector genérico $\mathbf{v}$ con las coordenadas de su imagen $f(\mathbf{v})$. Esta relación se expresa mediante la matriz A. La construcción de dicha matriz es la siguiente:

Si las bases de V y W vienen dadas por:

$$\mathcal{B}_V = \{\mathbf{v}_1, \dots, \mathbf{v}_m\}, \quad \mathcal{B}_W = \{\mathbf{w}_1, \dots, \mathbf{w}_n\},$$

se puede expresar un vector genérico $\mathbf{v} \in V$ y su imagen $f(\mathbf{v}) \in W$ en términos de las bases anteriores, obteniendo de ese modo sus coordenadas:

$$\mathbf{v} = x_1\mathbf{v}_1 + \cdots + x_m\mathbf{v}_m \Longrightarrow \mathcal{C}_{\mathcal{B}_V}(\mathbf{v}) = (x_1, \ldots, x_m).$$

$$f(\mathbf{v}) = y_1\mathbf{w}_1 + \cdots + y_n\mathbf{w}_n \Longrightarrow \mathcal{C}_{\mathcal{B}_W}(f(\mathbf{v})) = (y_1, \ldots, y_n).$$

Para relacionar $\mathcal{C}_{\mathcal{B}_V}(\mathbf{v})$ y $\mathcal{C}_{\mathcal{B}_W}(f(\mathbf{v}))$, se utiliza la definición de aplicación lineal:

$$\mathbf{v} = x_1\mathbf{v}_1 + \cdots + x_m\mathbf{v}_m \Longrightarrow f(\mathbf{v}) = x_1 f(\mathbf{v}_1) + \cdots + x_m f(\mathbf{v}_m). \tag{7.1}$$

Los vectores $f(\mathbf{v}_i) \in W$, por lo que se pueden expresar en combinación lineal de la base $\mathcal{B}_W$:

$$f(\mathbf{v}_i) = c_{i1}\mathbf{w}_1 + \cdots + c_{in}\mathbf{w}_n \Longrightarrow \mathcal{C}_{\mathcal{B}_W}(f(\mathbf{v}_i)) = (c_{i1}, \ldots, c_{in}). \tag{7.2}$$

Sustituyendo esta expresión de $f(\mathbf{v}_i)$ en la ecuación (7.1) se obtiene:

$$\begin{aligned} f(\mathbf{v}) &= x_1 f(\mathbf{v}_1) + \cdots + x_m f(\mathbf{v}_m) \\ &= x_1\left(c_{11}\mathbf{w}_1 + \cdots + c_{1n}\mathbf{w}_n\right) + \cdots + x_m\left(c_{m1}\mathbf{w}_1 + \cdots + c_{mn}\mathbf{w}_n\right) \\ &= \left(x_1c_{11} + \cdots + x_mc_{m1}\right)\mathbf{w}_1 + \cdots + \left(x_1c_{1n} + \cdots + x_mc_{mn}\right)\mathbf{w}_n. \end{aligned}$$

Dado que $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ es la base considerada en W, de esta última expresión se obtiene que

$$\mathcal{C}_{\mathcal{B}_W}(f(\mathbf{v})) = \big((x_1c_{11} + \cdots + x_mc_{m1}), \ldots, (x_1c_{1n} + \cdots + x_mc_{mn})\big).$$

Según el Teorema 5.16, las coordenadas de un vector en una base son únicas, por lo que:

$$(y_1, \ldots, y_n) = \left((x_1c_{11} + \cdots + x_mc_{m1}), \ldots, (x_1c_{1n} + \cdots + x_mc_{mn})\right).$$

Equivalentemente:

$$\begin{cases} y_1 = x_1c_{11} + \cdots + x_mc_{m1}, \\ \vdots \\ y_n = x_1c_{1n} + \cdots + x_mc_{mn}, \end{cases}$$

o en forma matricial:

$$\underbrace{\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}}_{\mathbf{y}} = \underbrace{\begin{pmatrix} c_{11} & \cdots & c_{m1} \\ \vdots & \ddots & \vdots \\ c_{1n} & \cdots & c_{mn} \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}}_{\mathbf{x}}. \tag{7.3}$$

El proceso anterior se puede representar mediante el siguiente diagrama:

$$\begin{tikzcd}
\mathbf{v} \arrow[d, mapsto] \arrow[rrr, mapsto, bend left] & V \arrow[r, "f"] \arrow[d, "\mathcal{C}_{\mathcal{B}_V}"'] \arrow[rd, phantom, "\#"] & W \arrow[d, "\mathcal{C}_{\mathcal{B}_W}"] & f(\mathbf{v}) \arrow[d, mapsto] \\
\mathbf{x} \arrow[rrr, mapsto, bend right] & \mathbb{R}^m \arrow[r, "A"'] & \mathbb{R}^n & \mathbf{y} = A\mathbf{x}
\end{tikzcd}$$

Definición 7.15. Sea $f : V \to W$ una aplicación lineal. La matriz A anteriormente calculada se llama **matriz coordenada (por columnas)** de f en las bases $\mathcal{B}_V$ y $\mathcal{B}_W$.

Observación. Sea $f : V \to W$ una aplicación lineal y sea A la matriz coordenada asociada en bases $\mathcal{B}_V$ y $\mathcal{B}_W$. La construcción de A se puede realizar de modo directo sin más que observar que sus columnas son las coordenadas de los vectores $f(\mathbf{v}_i)$ en base $\mathcal{B}_W$ (ver ecuaciones (7.2) y (7.3)):

$$A = \left(\begin{array}{c|c|c} \mathcal{C}_{\mathcal{B}_W}(f(\mathbf{v}_1)) & \cdots & \mathcal{C}_{\mathcal{B}_W}(f(\mathbf{v}_m)) \end{array} \right)$$

Los pasos para la construcción de A son simplemente:

1. Fijar bases de V y W: $\mathcal{B}_V$ y $\mathcal{B}_W$.
2. Calcular la imagen de los vectores de la base $\mathcal{B}_V$: $f(\mathbf{v}_i)$.
3. Expresar $f(\mathbf{v}_i)$ como combinación lineal de la base $\mathcal{B}_W$.
4. Colocar las coordenadas obtenidas en las columnas correspondientes de una matriz.

Se concluye con la siguiente observación:

Fijadas las bases $\mathcal{B}_V$ y $\mathcal{B}_W$ existe una correspondencia biyectiva entre aplicaciones lineales y matrices.

$$f \in Hom(V, W) \quad \longleftrightarrow \quad A \in \mathrm{Mat}_{n\times m}(\mathbb{R})$$

donde $n = \dim W$ y $m = \dim V$.

Observación. La matriz coordenada de una aplicación no es única, ya que para construirla es necesario fijar bases de los espacios vectoriales considerados. En las siguientes secciones se ve como relacionar matrices coordenadas de una misma aplicación lineal en bases distintas.

✓ **Ejemplo 9.**

- Sea $f : V \to W$ el homomorfismo nulo. Entonces, la matriz coordenada en cualquier base es la matriz nula.
- Sea $f : V \to V$ el endomorfismo identidad. Entonces, fijada una base cualquiera de V, la matriz coordenada es la identidad.
- Sea $f : V \to V$ el endomorfismo identidad. Fijadas dos bases de V, $\mathcal{B}_1$ y $\mathcal{B}_2$, la matriz coordenada de f en esas bases es la matriz de cambio de coordenadas $P_{\mathcal{B}_2,\mathcal{B}_1}$ definida en el Capítulo 5.

◇

✓ **Ejemplo 10.** Dada la aplicación $f : \mathbb{R}_2[x] \to \mathbb{R}^2$ definida por

$$f(a + bx + cx^2) = (a + b, a - b + c),$$

hallar la matriz coordenada de f en las bases canónicas:

$$\begin{aligned}\mathcal{B}_1 &= \text{Base canónica de } \mathbb{R}_2[x] = \{1, x, x^2\},\\ \mathcal{C}_1 &= \text{Base canónica de } \mathbb{R}^2 = \{(1,0),(0,1)\}.\end{aligned}$$

Solución: Siguiendo el esquema, es necesario calcular la imagen por f de los elementos de la base escogida de $\mathbb{R}_2[x]$ y expresar las imágenes anteriores en combinación lineal de los elementos de la base elegida de $\mathbb{R}^2$:

$$\begin{array}{lcll} f(1) & = (1,1) & \Longrightarrow & \mathcal{C}_{\mathcal{B}_W}(f(1)) = (1,1),\\ f(x) & = (1,-1) & \Longrightarrow & \mathcal{C}_{\mathcal{B}_W}(f(x)) = (1,-1),\\ f(x^2) & = (0,1) & \Longrightarrow & \mathcal{C}_{\mathcal{B}_W}(f(x^2)) = (0,1). \end{array}$$

Observar que por trabajar con la base canónica en $W = \mathbb{R}^2$, $f(\mathbf{v})$ coincide con $\mathcal{C}_{\mathcal{B}_W}(f(\mathbf{v}))$. La matriz coordenada buscada es:

$$A = \left(\begin{array}{c|c|c} \mathcal{C}_{\mathcal{B}_W}(f(1)) & \mathcal{C}_{\mathcal{B}_W}(f(x)) & \mathcal{C}_{\mathcal{B}_W}(f(x^2)) \end{array} \right) = \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix}.$$

Esta matriz da lugar a la siguiente aplicación:

$$\begin{array}{cccc} f_A : & \mathbb{R}^3 & \longrightarrow & \mathbb{R}^2 \\ & \begin{pmatrix} a \\ b \\ c \end{pmatrix} & \mapsto & \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} a+b \\ a-b+c \end{pmatrix}. \end{array}$$

◇

⇝ **Ejercicio.** Comprobar que la matriz coordenada de la aplicación del Ejemplo 10 en bases $\mathcal{B}_2 = \{1 - x, 1 + x, x^2 - 3x + 2\}$ y $\mathcal{C}_2 = \{(0,-1),(2,3)\}$ viene dada por la matriz

$$B = \begin{pmatrix} -2 & 3 & -15/2 \\ 0 & 1 & -1/2 \end{pmatrix}.$$

7.2.1. MATRIZ COORDENADA DE LA COMPOSICIÓN

Sean V, W, Z tres espacios vectoriales reales de dimensión finita y se consideran las aplicaciones lineales $f : V \to W$ y $g : W \to Z$. Para determinar la matriz coordenada de la composición $g \circ f : V \to Z$ se fijan bases $\mathcal{B}_V$, $\mathcal{B}_W$ y $\mathcal{B}_Z$ de los espacios. Sea A la matriz coordenada de f en las bases $\mathcal{B}_V$ y $\mathcal{B}_W$ y sea B la matriz coordenada de g en las bases $\mathcal{B}_W$ y $\mathcal{B}_Z$. Considerar un vector $\mathbf{v} \in V$ cualquiera.

$$\begin{aligned} \mathbf{v} \in V \quad &\Longrightarrow \quad \mathcal{C}_{\mathcal{B}_V}(\mathbf{v}) = \mathbf{x}, \\ f(\mathbf{v}) \in W \quad &\Longrightarrow \quad \mathcal{C}_{\mathcal{B}_W}(f(\mathbf{v})) = \mathbf{y}, \\ g(f(\mathbf{v})) \in Z \quad &\Longrightarrow \quad \mathcal{C}_{\mathcal{B}_Z}(g(f(\mathbf{v})) = \mathbf{z}. \end{aligned}$$

Trabajando con coordenadas y matrices se tiene:

$$\mathbf{y} = A\mathbf{x}, \quad \mathbf{z} = B\mathbf{y} \quad \Longrightarrow \quad \mathbf{z} = BA\mathbf{x}.$$

Esto quiere decir que la matriz coordenada de $g \circ f$ es el producto de matrices coordenadas BA (en este orden). Nótese que primero actúa A y después actúa B.

$$\boxed{\begin{array}{l} f \longleftrightarrow A \\ g \longleftrightarrow B \end{array} \Longrightarrow g \circ f \longleftrightarrow BA}$$

7.2.2. USOS DE LA MATRIZ COORDENADA

La matriz coordenada se utiliza para estudiar propiedades de la aplicación f a la que representa. Concretamente, permite determinar fácilmente los subespacios $\operatorname{Ker} f$ e $\operatorname{Im} f$; decidir si f es inyectiva, suprayectiva o biyectiva y calcular explícitamente la aplicación f^{-1} en este último caso.

CÁLCULO DE LOS SUBESPACIOS $\operatorname{Ker} f$, $\operatorname{Im} f$

Si A es la matriz coordenada de f fijadas las bases $\mathcal{B}_V$ y $\mathcal{B}_W$, entonces se puede decir que:

$$\boxed{\begin{array}{l} \operatorname{Ker} f_A = \text{coordenadas en } \mathcal{B}_V \text{ de } \operatorname{Ker} f. \\ \operatorname{Im} f_A = \text{coordenadas en } \mathcal{B}_W \text{ de } \operatorname{Im} f. \end{array}}$$

Como la aplicación f_A viene dada por una matriz, los subespacios $\operatorname{Ker} f_A$ e $\operatorname{Im} f_A$ (también denotados como $\operatorname{Ker} A$ e $\operatorname{Im} A$) se calculan como en la observación de la página 208. Para obtener los subespacios $\operatorname{Ker} f$ e $\operatorname{Im} f$ se deben construir los vectores cuyas coordenadas vienen dadas en $\operatorname{Ker} f_A$ e $\operatorname{Im} f_A$.

✓ **Ejemplo 11.** Calcular los subespacios $\operatorname{Ker} f$ e $\operatorname{Im} f$ de la aplicación

$$\begin{array}{cccc} f: & \mathbb{R}_2[x] & \longrightarrow & \mathbb{R}^2 \\ & a+bx+cx^2 & \mapsto & (a+b,\ a-b+c) \end{array}$$

utilizando la matriz coordenada A calculada en el Ejemplo 10.

Solución: Para hallar el núcleo se calcula en primer lugar $\operatorname{Ker} f_A$:

$$\operatorname{Ker} f_A = \left\{ (a,b,c) \in \mathbb{R}^3 \;\middle|\; \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\} = \langle (-1,1,2) \rangle.$$

Observar que este subespacio se ha calculado previamente en el Ejemplo 8. Ahora, $\operatorname{Ker} f = \langle p(x) \rangle$ donde $p(x) \in \mathbb{R}_2[x]$ es un polinomio cuyas coordenadas en base $\{1, x, x^2\}$ son $(-1,1,2)$:

$$p(x) = -1 \cdot 1 + 1 \cdot x + 2 \cdot x^2 = -1 + x + 2x^2 \Longrightarrow \operatorname{Ker} f = \langle -1 + x + 2x^2 \rangle.$$

Se obtiene el mismo resultado que con el cálculo directo de $\operatorname{Ker} f$ hecho en el Ejemplo 7.

El cálculo de $\operatorname{Im} f$ es inmediato conociendo $\operatorname{Im} f_A$:

$$\operatorname{Im} f_A \overset{\text{Ej.8}}{=} \left\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\rangle.$$

Como el espacio de llegada es $\mathbb{R}^2$ y se ha considerado su base canónica, entonces los vectores coinciden con sus coordenadas, por lo que $\operatorname{Im} f = \langle (1,1),\ (0,1) \rangle$.

◇

ESTUDIO DEL TIPO DE APLICACIÓN

Sin necesidad de calcular los subespacios $\operatorname{Ker} f$ e $\operatorname{Im} f$ se puede decidir si f es inyectiva, suprayectiva o biyectiva, simplemente conociendo el rango de la matriz A:

$$\dim(\operatorname{Im} f) = \dim(\operatorname{Im} f_A) = \dim\{\text{espacio de columnas de } A\} = \operatorname{rg} A.$$

Despejando del teorema de las dimensiones (Teorema 7.8):

$$\dim(\operatorname{Ker} f) = \dim V - \dim(\operatorname{Im} f) = \dim V - \operatorname{rg} A.$$

En consecuencia se tiene el siguiente resultado.

Proposición 7.16. Sea $f : V \to W$ una aplicación lineal y sea A la matriz coordenada de f en bases $\mathcal{B}_V$ y $\mathcal{B}_W$.

1. f es inyectiva $\Longleftrightarrow \operatorname{rg} A = \dim V$.

2. f es suprayectiva $\Longleftrightarrow \operatorname{rg} A = \dim W$.

3. f es biyectiva $\Longleftrightarrow \dim V = \dim W = \operatorname{rg} A$.

✓ **Ejemplo 12.** Utilizando la proposición anterior estudiar las propiedades de la aplicación del Ejemplo 10.

Solución: Como $\dim V = 3,\ \dim W = 2$ y $\operatorname{rg} A = 2$, se tiene que f es una aplicación suprayectiva (o epimorfismo) y no inyectiva.

◇

APLICACIÓN INVERSA

Sea $f : V \to W$ un isomorfismo, es decir, una aplicación lineal biyectiva. Por ser f biyectiva, existe $f^{-1} : W \to V$ cumpliendo

$$\begin{cases} f^{-1} \circ f = Id_V \\ f \circ f^{-1} = Id_{W.} \end{cases}$$

Sea A la matriz coordenada de f en bases $\mathcal{B}_V$ y $\mathcal{B}_W$. El problema que se plantea es determinar la matriz coordenada B de f^{-1} en bases $\mathcal{B}_W$ y $\mathcal{B}_V$. Utilizando la expresión de la matriz coordenada de la composición se tiene que:

$$\begin{cases} f^{-1} \circ f = Id_V \Longrightarrow BA = I \\ f \circ f^{-1} = Id_W \Longrightarrow AB = I \end{cases} \quad \Longrightarrow B = A^{-1}.$$

Así pues:

$$\boxed{f \longleftrightarrow A \Longrightarrow f^{-1} \longleftrightarrow A^{-1}}$$

Una vez obtenida la matriz coordenada A^{-1}, para reconstruir la aplicación inversa $f^{-1} : W \to V$ se utilizan las bases $\mathcal{B}_W$ y $\mathcal{B}_V$ como se muestra en el siguiente ejemplo.

✓ **Ejemplo 13.** Dada la aplicación

$$\begin{array}{rccc} f: & \operatorname{Mat}_2(\mathbb{R}) & \longrightarrow & \mathbb{R}^4 \\ & \begin{pmatrix} a & b \\ c & d \end{pmatrix} & \mapsto & (a, a+b, c, c+d) \end{array}$$

probar que f es biyectiva y construir explícitamente f^{-1}.

Solución: Se comienza calculando la matriz coordenada de f en las bases canónicas:

Base canónica de $\operatorname{Mat}_2(\mathbb{R})$: $\mathcal{B}_1 = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}.$

Base canónica de $\mathbb{R}^4$: $\mathcal{C}_1 = \{(1,0,0,0),\ (0,1,0,0),\ (0,0,1,0),\ (0,0,0,1)\}.$

La matriz coordenada de f en las bases anteriores es:

$$\left\{\begin{array}{l} f\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix} = (1,1,0,0),\\ f\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix} = (0,1,0,0),\\ f\begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix} = (0,0,1,1),\\ f\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix} = (0,0,0,1), \end{array}\right. \Longrightarrow \quad A = \begin{pmatrix}1 & 0 & 0 & 0\\ 1 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 1 & 1\end{pmatrix}.$$

Como $\operatorname{rg} A = 4 = \dim V = \dim W$, por la Proposición 7.16 se tiene que f es biyectiva y por tanto existe f^{-1}, que es de la forma:

$$\begin{array}{rccc} f^{-1}: & \mathbb{R}^4 & \longrightarrow & \mathrm{Mat}_2(\mathbb{R})\\ & (x,y,z,t) & \mapsto & ? \end{array}$$

Al calcular A^{-1} se obtiene la matriz coordenada de f^{-1} en bases $\mathcal{C}_1$ y $\mathcal{B}_1$:

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

$$\begin{array}{rccc} f_{A^{-1}} = (f_A)^{-1}: & \mathbb{R}^4 & \longrightarrow & \mathbb{R}^4\\ & (x,y,z,t) & \mapsto & A^{-1}\begin{pmatrix}x\\ y\\ z\\ t\end{pmatrix} = \begin{pmatrix}x\\ -x+y\\ z\\ -z+t\end{pmatrix}. \end{array}$$

La aplicación anterior está dada en coordenadas, luego para reconstruir f^{-1} es necesario utilizar las bases $\mathcal{C}_1$ y $\mathcal{B}_1$:

$$\begin{array}{rccc} f^{-1}: & \mathbb{R}^4 & \longrightarrow & \mathrm{Mat}_2(\mathbb{R})\\ & (x,y,z,t) & \mapsto & \begin{pmatrix}x & -x+y\\ z & -z+t\end{pmatrix}. \end{array}$$

◇

✓ **Ejemplo 14.** Con los datos del ejemplo anterior comprobar que $f \circ f^{-1} = Id_{\mathbb{R}^4}$.

Solución:

$$(x,y,z,t) \overset{f^{-1}}{\longmapsto} \begin{pmatrix}x & -x+y\\ z & -z+t\end{pmatrix} \overset{f}{\longmapsto} (x, x-x+y, z, z-z+t) = (x,y,z,t).$$

◇

⇝ **Ejercicio.** Comprobar que $f^{-1} \circ f = Id_{\mathrm{Mat}_2(\mathbb{R})}$.

7.2.3. MATRIZ COORDENADA EN OTRAS BASES

La matriz coordenada de una aplicación lineal depende de las bases elegidas en los espacios V y W. Si se cambian las bases la matriz también cambia aunque no lo hacen sus propiedades básicas, como el rango. En esta sección se trata de encontrar una relación entre las matrices coordenadas de una aplicación lineal en distintas bases.

Sea $f : V \to W$ una aplicación lineal entre dos espacios vectoriales reales con dimensión finita $\dim V = m$ y $\dim W = n$. Sean:

- $\mathcal{B}_V^1 = \{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ una base de V, $\mathcal{B}_W^1 = \{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ una base de W y A la matriz coordenada de f en $\mathcal{B}_V^1$ y $\mathcal{B}_W^1$.
- $\mathcal{B}_V^2 = \{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ otra base de V, $\mathcal{B}_W^2 = \{\mathbf{z}_1, \ldots, \mathbf{z}_n\}$ otra base de W y B la matriz coordenada de f en estas bases.

A continuación se estudia la relación que existe entre las matrices A y B:

Sea $\mathbf{v} \in V$. Entonces:

$$\begin{aligned} \mathbf{v} \in V \quad &\Longrightarrow \quad \mathcal{C}_{\mathcal{B}_V^1}(\mathbf{v}) = \mathbf{x}, \qquad \mathcal{C}_{\mathcal{B}_V^2}(\mathbf{v}) = \widetilde{\mathbf{x}}. \\ f(\mathbf{v}) \in W \quad &\Longrightarrow \quad \mathcal{C}_{\mathcal{B}_W^1}(f(\mathbf{v})) = \mathbf{y}, \quad \mathcal{C}_{\mathcal{B}_W^2}(f(\mathbf{v})) = \widetilde{\mathbf{y}}. \end{aligned}$$

La relación entre las coordenadas utilizando las matrices es:

$$\mathbf{y} = A\mathbf{x}, \quad \widetilde{\mathbf{y}} = B\widetilde{\mathbf{x}}.$$

En V se han considerado dos bases distintas, $\mathcal{B}_V^1$ y $\mathcal{B}_V^2$. Las coordenadas de un vector $\mathbf{v}$ en ambas bases se relacionan a través de la matriz de cambio de coordenadas estudiada en el Capítulo 5. Concretamente:

$$\exists P \in \mathrm{Mat}_m(\mathbb{R}) \quad \text{regular} \quad \text{tal que} \quad \mathbf{x} = P\widetilde{\mathbf{x}}.$$

Aplicando el mismo argumento en W, se tiene:

$$\exists Q \in \mathrm{Mat}_n(\mathbb{R}) \quad \text{regular} \quad \text{tal que} \quad \mathbf{y} = Q\widetilde{\mathbf{y}}.$$

Entonces:

$$\mathbf{y} = A\mathbf{x} \quad \Longrightarrow \quad Q\widetilde{\mathbf{y}} = AP\widetilde{\mathbf{x}} \quad \Longrightarrow \quad \widetilde{\mathbf{y}} = Q^{-1}AP\widetilde{\mathbf{x}}.$$

Como $\widetilde{\mathbf{y}} = B\widetilde{\mathbf{x}}$, entonces

$$B\widetilde{\mathbf{x}} = Q^{-1}AP\widetilde{\mathbf{x}}, \quad \forall\widetilde{\mathbf{x}} \quad \Longrightarrow \quad \boxed{B = Q^{-1}AP}$$

El proceso anterior se puede resumir con el siguiente diagrama:

$$\begin{tikzcd}
V \arrow[rrr, dotted, "f"] \arrow[dr, "\mathcal{C}_{\mathcal{B}_V^1}"'] & & & W \arrow[dl, "\mathcal{C}_{\mathcal{B}_W^1}"] \\
& \mathbb{R}^m \arrow[r, "A"] & \mathbb{R}^n & \\
& \mathbb{R}^m \arrow[u, "P"] \arrow[r, "B"'] \arrow[ur, phantom, "\#"] & \mathbb{R}^n \arrow[u, "Q"'] & \\
V \arrow[rrr, dotted, "f"'] \arrow[ur, "\mathcal{C}_{\mathcal{B}_V^2}"] & & & W \arrow[ul, "\mathcal{C}_{\mathcal{B}_W^2}"']
\end{tikzcd}$$

En la práctica se simplifica el diagrama anterior y se trabaja directamente con coordenadas, es decir, se utiliza el diagrama conmutativo central.

$$\begin{tikzcd}
(\mathcal{B}_V^1) & \mathbb{R}^m \arrow[r, "A"] & \mathbb{R}^n & (\mathcal{B}_W^1) \\
(\mathcal{B}_V^2) & \mathbb{R}^m \arrow[u, "P"] \arrow[r, "B"'] \arrow[ur, phantom, "\#"] & \mathbb{R}^n \arrow[u, "Q"'] & (\mathcal{B}_W^2)
\end{tikzcd}$$

Notas.

- La matriz P es la matriz de cambio de coordenadas de la base $\mathcal{B}_V^2$ a $\mathcal{B}_V^1$, es decir, $P = P_{\mathcal{B}_2,\mathcal{B}_1}$, y eso se refleja en el diagrama anterior mediante el sentido de la flecha. Lo mismo sucede con la matriz Q.
- El diagrama anterior tiene dos lecturas posibles, dependiendo de cuál sea la matriz despejada: $B = Q^{-1}AP$ o bien $A = QBP^{-1}$.
- Lo más recomendable es utilizar como $\mathcal{B}_V^1$ la base canónica de V y como $\mathcal{B}_W^1$ la base canónica de W.

Definición 7.17. Dos matrices $A, B \in \mathrm{Mat}_{n\times m}(\mathbb{R})$ se dicen **equivalentes** si existen matrices $P \in \mathrm{Mat}_m(\mathbb{R})$ y $Q \in \mathrm{Mat}_n(\mathbb{R})$ regulares tales que $B = Q^{-1}AP$.

Proposición 7.18. Dos matrices son equivalentes si y solo si son matrices coordenadas asociadas a una misma aplicación lineal en bases distintas.

Proposición 7.19. Dos matrices son equivalentes si y solo si tienen la misma dimensión y el mismo rango.

✓ **Ejemplo 15.** Considerar la aplicación lineal

$$\begin{array}{cccc} f: & \mathbb{R}_2[x] & \longrightarrow & \mathbb{R}^2 \\ & a + bx + cx^2 & \mapsto & (a+b,\ a-b+c) \end{array}$$

Sea A la matriz coordenada de f en bases canónicas dada en el Ejemplo 10. Determinar la matriz coordenada B de f en bases $\mathcal{B}_2 = \{1-x, 1+x, x^2-3x+2\}$ y $\mathcal{C}_2 = \{(0,-1), (2,3)\}$ sin calcularla directamente, hallando las matrices P y Q de cambio de coordenadas y la relación de estas con A y B.

Solución: Según el Ejemplo 10, $A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix}$. El diagrama que resume la situación es

$$\begin{tikzcd}
(\mathcal{B}_1) & \mathbb{R}^3 \arrow[r, "A"] & \mathbb{R}^2 \arrow[d, dotted, bend left, "Q^{-1}"] & (\mathcal{C}_1) \\
(\mathcal{B}_2) & \mathbb{R}^3 \arrow[u, "P"] \arrow[r, "B"'] \arrow[ur, phantom, "\#"] & \mathbb{R}^2 \arrow[u, "Q"] & (\mathcal{C}_2)
\end{tikzcd}$$

De esta manera $B = Q^{-1}AP$. Las matrices de cambio de coordenadas se obtienen de manera directa por ser $\mathcal{B}_1$ y $\mathcal{C}_1$ las bases canónicas de los espacios correspondientes:

$$P = \begin{pmatrix} 1 & 1 & 2 \\ -1 & 1 & -3 \\ 0 & 0 & 1 \end{pmatrix}, \quad Q = \begin{pmatrix} 0 & 2 \\ -1 & 3 \end{pmatrix}, \quad Q^{-1} = \begin{pmatrix} 3/2 & -1 \\ 1/2 & 0 \end{pmatrix}.$$

Multiplicando las matrices se obtiene la matriz B buscada:

$$B = \underbrace{\begin{pmatrix} 3/2 & -1 \\ 1/2 & 0 \end{pmatrix}}_{Q^{-1}} \underbrace{\begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} 1 & 1 & 2 \\ -1 & 1 & -3 \\ 0 & 0 & 1 \end{pmatrix}}_{P} = \begin{pmatrix} -2 & 3 & -15/2 \\ 0 & 1 & -1/2 \end{pmatrix}.$$

Se observa que se obtiene la matriz calculada en el Ejercicio de la página 214.

◇

7.2.4. ENDOMORFISMOS: MATRICES SEMEJANTES

Sea $f : V \to V$ un endomorfismo. Para hallar una matriz coordenada de f se utiliza el proceso antes expuesto. Sin embargo los endomorfismo presentan dos peculiaridades:

- Las matrices coordenadas en cualquier base son cuadradas.
- Normalmente se utiliza la misma base para el espacio de partida y de llegada.

En lo que sigue, se busca la relación que existe entre dos matrices coordenadas de un endomorfismo f en bases distintas.

Sea $\mathcal{B}_1 = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ una base de V y A la matriz coordenada de f asociada. Sea $\mathcal{B}_2 = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ otra base de V y B la matriz coordenada:

$$\mathbf{y} = A\mathbf{x}, \quad \widetilde{\mathbf{y}} = B\widetilde{\mathbf{x}}.$$

Sea $P \in \mathrm{Mat}_n(\mathbb{R})$ la matriz regular de cambio de coordenadas entre $\mathcal{B}_2$ y $\mathcal{B}_1$:

$$\mathbf{x} = P\widetilde{\mathbf{x}}, \quad \mathbf{y} = P\widetilde{\mathbf{y}}.$$

De este modo se tiene:

$$\mathbf{y} = A\mathbf{x} \Longleftrightarrow P\widetilde{\mathbf{y}} = AP\widetilde{\mathbf{x}} \Longleftrightarrow \widetilde{\mathbf{y}} = P^{-1}AP\widetilde{\mathbf{x}} \Longleftrightarrow B\widetilde{\mathbf{x}} = P^{-1}AP\widetilde{\mathbf{x}} \quad \forall\widetilde{\mathbf{x}}$$

es decir,

$$\boxed{B = P^{-1}AP}$$

El siguiente diagrama representa el proceso anterior:

$$\begin{tikzcd}
(\mathcal{B}_1) & \mathbb{R}^n \arrow[r, "A"] & \mathbb{R}^n & (\mathcal{B}_1) \\
(\mathcal{B}_2) & \mathbb{R}^n \arrow[u, "P"] \arrow[r, "B"'] \arrow[ur, phantom, "\#"] & \mathbb{R}^n \arrow[u, "P"'] & (\mathcal{B}_2)
\end{tikzcd}$$

Nota. Equivalentemente se verifica $PBP^{-1} = A$.

Para matrices cuadradas se tiene la siguiente definición:

Definición 7.20. Dos matrices $A, B \in \mathrm{Mat}_n(\mathbb{R})$ se dicen **semejantes** si existe una matriz $P \in \mathrm{Mat}_n(\mathbb{R})$ regular tal que $B = P^{-1}AP$.

Proposición 7.21. Las matrices semejantes representan un mismo endomorfismo en bases distintas.

✓ **Ejemplo 16.** Sea $f : \mathbb{R}^3 \to \mathbb{R}^3$ el endomorfismo definido por:

$$\begin{array}{rccl} f: & \mathbb{R}^3 & \longrightarrow & \mathbb{R}^3 \\ & (a,b,c) & \mapsto & (2a-b, b+c, c-3a). \end{array}$$

Se considera $\mathcal{B}_1$ la base canónica de $\mathbb{R}^3$ y $\mathcal{B}_2 = \{(1,1,0),(1,0,1),(0,1,0)\}$. Encontrar la matriz coordenada A de f en base $\mathcal{B}_1$ y una matriz P regular de modo que $P^{-1}AP$ sea la matriz coordenada de f en base $\mathcal{B}_2$.

Solución: En primer lugar se plantea el diagrama asociado:

$$\begin{tikzcd}
(\mathcal{B}_1) & \mathbb{R}^3 \arrow[r, "A"] & \mathbb{R}^3 \arrow[d, dotted, bend left, "P^{-1}"] & (\mathcal{B}_1) \\
(\mathcal{B}_2) & \mathbb{R}^3 \arrow[u, "P"] \arrow[r, "B"'] \arrow[ur, phantom, "\#"] & \mathbb{R}^3 \arrow[u, "P"] & (\mathcal{B}_2)
\end{tikzcd}$$

Por ser $\mathcal{B}_1$ la base canónica, la construcción de P es inmediata: $P = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$.

La matriz P^{-1} viene dada por: $P^{-1} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 1 \\ -1 & 1 & 1 \end{pmatrix}$.

La matriz coordenada de f en base canónica $\mathcal{B}_1$ es A:

$$\left.\begin{array}{l} f(1,0,0) = (2,0,-3) \\ f(0,1,0) = (-1,1,0) \\ f(0,0,1) = (0,1,1) \end{array}\right\} \Longrightarrow A = \begin{pmatrix} 2 & -1 & 0 \\ 0 & 1 & 1 \\ -3 & 0 & 1 \end{pmatrix}.$$

Finalmente, la matriz coordenada de f en base $\mathcal{B}_2$ se calcula haciendo el siguiente producto:

$$P^{-1}AP = \begin{pmatrix} 4 & 4 & -1 \\ -3 & -2 & 0 \\ -3 & -3 & 2 \end{pmatrix}.$$

Se deja como ejercicio comprobar que el resultado obtenido es correcto.

◇

Observación. El problema que se plantea de aquí en adelante es encontrar una base de V en la que la matriz coordenada de un endomorfismo sea lo más sencilla posible. Equivalentemente, dada una matriz cuadrada, encontrar otra semejante a ella lo más sencilla posible. Este problema se trata con detalle en el capítulo siguiente.

7.3. APLICACIONES

A continuación se presenta una interpretación geométrica de ciertas aplicaciones lineales en el plano $\mathbb{R}^2$ y en el espacio $\mathbb{R}^3$.

7.3.1. TRANSFORMACIONES EN EL PLANO

Considerar aplicaciones lineales del siguiente tipo:

$$\begin{array}{cccc} f: & \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ & \begin{pmatrix} x \\ y \end{pmatrix} & \mapsto & \begin{pmatrix} x' \\ y' \end{pmatrix} = A \begin{pmatrix} x \\ y \end{pmatrix} \end{array}$$

Dependiendo de la matriz A se obtienen distintas transformaciones del plano:

1. **Reflexiones**: El resultado de aplicar una reflexión a un vector es la imagen especular del vector, donde el espejo es el eje de la reflexión.
   - Reflexión respecto al eje OX: $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.
   - Reflexión respecto al eje OY: $A = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$.

- Reflexión respecto de la recta $x = y$: $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.
- Reflexión respecto al origen de coordenadas: $A = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$.

2. **Giro de ángulo** θ: El resultado de aplicar un giro de ángulo θ a un vector es, precisamente, el giro de ángulo θ del vector en sentido antihorario.

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

Utilizando coordenadas polares se observa cómo el argumento del vector se incrementa en θ:

$$(x, y) = (r\cos\alpha,\, r\operatorname{sen}\alpha) \Longrightarrow$$

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix} \begin{pmatrix} r\cos\alpha \\ r\operatorname{sen}\alpha \end{pmatrix} = \begin{pmatrix} r\cos(\alpha+\theta) \\ r\operatorname{sen}(\alpha+\theta) \end{pmatrix}.$$

3. **Proyecciones**: El resultado de aplicar una proyección a un vector es equivalente a ver el vector en una dimensión menos.

   - Proyección en el eje OX: $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$.
   - Proyección en el eje OY: $A = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$.

4. **Homotecias**: El resultado de aplicar una homotecia a un vector es una dilatación o contracción del mismo, es decir, un vector de la misma dirección y sentido que el original pero de diferente módulo:

$$A = \begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix} \quad \rightsquigarrow \quad \begin{cases} \text{Si } 0 < k < 1, \text{ la homotecia es } \textit{contractiva.} \\ \text{Si } k > 1, \text{ la homotecia es } \textit{expansiva.} \end{cases}$$

Observación. Considerar la norma euclídea de $\mathbb{R}^n$, una aplicación $F : \mathbb{R}^n \to \mathbb{R}^n$ es una **isometría** si $\|F(\mathbf{v})\| = \|\mathbf{v}\|$. Las isometrías se pueden ver como aplicaciones que conservan las distancias. De las cuatro transformaciones lineales del plano que se han presentado, solamente las reflexiones y los giros son isometrías. Observar además que en esos casos se tiene que la matriz A es ortogonal, es decir $AA^t = I$.

7.3.2. TRANSFORMACIONES EN EL ESPACIO

Las transformaciones del plano tienen su análogo en el espacio:

1. **Reflexiones**:

   - Reflexión respecto al eje OX: $A = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

- Reflexión respecto al eje OY: $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.
- Reflexión respecto al eje OZ: $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$.
- Reflexión respecto al origen de coordenadas: $A = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$.

2. **Giro de ángulo θ**:
   - Alrededor del eje OX: $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\operatorname{sen}\theta \\ 0 & \operatorname{sen}\theta & \cos\theta \end{pmatrix}$.
   - Alrededor del eje OY: $A = \begin{pmatrix} \cos\theta & 0 & \operatorname{sen}\theta \\ 0 & 1 & 0 \\ -\operatorname{sen}\theta & 0 & \cos\theta \end{pmatrix}$.
   - Alrededor del eje OZ: $A = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta & 0 \\ \operatorname{sen}\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$.
3. **Proyecciones**:
   - Proyección en el plano $x = 0$: Transforma una pieza tridimensional en su vista en perfil:
   $$A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$
   - Proyección en el plano $y = 0$: Transforma una pieza tridimensional en su vista en alzado:
   $$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$
   - Proyección en el plano $z = 0$: Transforma una pieza tridimensional en su vista en planta:
   $$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

   Las transformaciones anteriores se realizan punto a punto.
4. **Homotecias**:
   $$A = \begin{pmatrix} k & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & k \end{pmatrix}.$$

7.4. COMANDOS DE *wxMaxima*

Dada una aplicación lineal $f : V \to W$, *wxMaxima* no dispone de comandos específicos para el cálculo de los subespacios $\operatorname{Ker} f$ e $\operatorname{Im} f$. Se pueden obtener sus coordenadas utilizando una matriz coordenada A de f en bases $\mathcal{B}_V$ y $\mathcal{B}_W$ y los siguientes comandos:

- Calcular la dimensión del núcleo de la matriz A, $\dim(\operatorname{Ker} A)$, o equivalentemente la dimensión del espacio de soluciones del sistema homogéneo $A\mathbf{x} = \mathbf{0}$:

```
-->     nullity(A);
```

- Calcular una base del núcleo de la matriz A, o equivalentemente una base del espacio de soluciones del sistema homogéneo $A\mathbf{x} = \mathbf{0}$:

```
-->     nullspace(A);
```

- Calcular una base del espacio de columnas de A, es decir, una base de $\operatorname{Im} A$:

```
-->     columnspace(A);
```

7.5. EJERCICIOS PROPUESTOS

Aplicaciones lineales y subespacios

1. Determinar cuáles de las siguientes aplicaciones son lineales:

a) $f : \mathbb{R}^2 \to \mathbb{R}^2; \quad f(x,y) = (x-1, x-y)$.

b) $f : \mathbb{R}^3 \to \mathbb{R}^2; \quad f(x,y,z) = (0, 2x-3y+5z)$.

c) $f : \mathrm{Mat}_n(\mathbb{R}) \to \mathrm{Mat}_n(\mathbb{R}); \quad f(A) = A^t + A$.

d) $f : \mathrm{Mat}_{n\times m}(\mathbb{R}) \to \mathbb{R}; \quad f(A) = \operatorname{rg} A$.

e) $f : \mathbb{R}_n[x] \to \mathbb{R}; \quad f(p(x)) = p(0)$.

f) $f : \mathbb{R}_3[x] \to \mathbb{R}_3[x]; \quad f(p(x)) = 2p(x) + p'(x)$.

g) $f : \mathbb{R} \to \mathbb{R}; \quad f(x) = x^3$.

h) $f : \mathrm{Mat}_2(\mathbb{R}) \to \mathbb{R}; \quad f\begin{pmatrix} a & b \\ c & d \end{pmatrix} = a^2 + b^2$.

2. Sea la aplicación lineal $f : \mathbb{R}_1[x] \to \mathbb{R}_1[x]$ dada por

$$f(-1+x) = -8+5x, \quad f(2+2x) = 4+9x.$$

Determinar $f(a+bx)$ y $f(10x-15)$.

3. Hallar una aplicación lineal $f : \mathbb{R}^5 \to \mathbb{R}^4$ tal que $\operatorname{Im} f = \{(x,0,x,x+y) \mid x,y \in \mathbb{R}\}$.

4. Sea $f : \mathbb{R}^4 \to \mathbb{R}^3$ la aplicación lineal dada por $f(x,y,z,t) = (x+2y+t, y+3z-t, 0)$. Dar una base de la imagen de los siguientes subespacios:

a) $S = \langle (1,0,1,0), (2,3,0,-1) \rangle$.

b) $T = \langle (0,0,3,2), (4,6,3,-1), (1,0,0,2) \rangle$.

5. Sea $f : \mathbb{R}^3 \to \mathbb{R}^4$ la aplicación lineal dada por $f(x,y,z) = (x+y+z, z, y, x)$. Hallar todos los vectores cuya imagen sea $\mathbf{w} = (4,1,3,0)$.

6. Sea $f : \mathbb{R}^3 \to \mathbb{R}^2$ una aplicación lineal dada por $f(\mathbf{x}) = \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \end{pmatrix} \mathbf{x}$. Hallar las dimensiones y una base de los siguientes subespacios:

a) $S = \{\mathbf{x} \in \mathbb{R}^3 \mid f(\mathbf{x}) = \mathbf{0}\}$.

b) $T = \{\mathbf{y} \in \mathbb{R}^2 \mid f(\mathbf{x}) = \mathbf{y} \ \text{ con } \ \mathbf{x} \in \mathbb{R}^3\}$.

7. Hallar en cada caso una base del núcleo y de la imagen:

a) $f : \mathbb{R}_2[x] \to \mathbb{R}^2; \quad f(p(x)) = (p(0), p(1))$.

b) $g : \mathrm{Mat}_2(\mathbb{R}) \to \mathrm{Mat}_2(\mathbb{R}); \quad g\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a+b & b+c \\ c+d & d+a \end{pmatrix}$.

8. Considerar las siguientes restricciones de las aplicaciones del ejercicio anterior.

a) $f|_U$, siendo $U = \mathbb{R}_1[x]$.

b) $g|_U$, siendo U el subespacio vectorial de las matrices simétricas de orden 2.

Hallar en cada caso una base del núcleo y de la imagen.

9. Se consideran las aplicaciones lineales $f : V \to W$ y $g : W \to Z$. Probar que $\operatorname{Ker} f \subseteq \operatorname{Ker}\,(g \circ f)$.

10. Sea una matriz $A \in \operatorname{Mat}_{5\times 7}(\mathbb{R})$ con $\operatorname{rg} A = 4$.

a) ¿Cuál es la dimensión del espacio de soluciones de $A\mathbf{x} = \mathbf{0}$?

b) ¿Es compatible $A\mathbf{x} = \mathbf{b}$ para todo $\mathbf{b}$ perteneciente a $\mathbb{R}^5$? ¿Por qué?

Matriz coordenada. Equivalencia y semejanza de matrices

11. Estudiar cuáles de las siguientes matrices son equivalentes:

a) $A = \begin{pmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 4 \end{pmatrix}$ y $B = \begin{pmatrix} 2 & 3 & 4 \\ 1 & 2 & 3 \end{pmatrix}$.

b) $A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \end{pmatrix}$ y $B = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 0 & -1 & 1 \end{pmatrix}$.

12. Se considera la aplicación

$$\begin{array}{rccc} f: & \operatorname{Mat}_2(\mathbb{R}) & \to & \operatorname{Mat}_2(\mathbb{R}) \\ & \begin{pmatrix} a & b \\ c & d \end{pmatrix} & \mapsto & \begin{pmatrix} -1 & 2 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix}. \end{array}$$

a) Demostrar que f es lineal.

b) Hallar la matriz coordenada A de f respecto de la base canónica de $\operatorname{Mat}_2(\mathbb{R})$.

c) Hallar el núcleo y la imagen de f así como sus dimensiones y bases.

d) Hallar la matriz coordenada B de f respecto de la base:

$$\mathcal{B} = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}.$$

e) Demostrar que las matrices A y B son semejantes.

13. Sea $f : \mathbb{R}^3 \to \mathbb{R}^3$ dada por

$$f\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \lambda x_1 + \mu x_2 + x_3 \\ x_1 + \lambda\mu x_2 + x_3 \\ x_1 + \mu x_2 + \lambda x_3 \end{pmatrix}.$$

a) Hallar la matriz coordenada A de f en base canónica.

b) Encontrar los valores de λ y μ para los que la imagen de f es $\mathbb{R}^3$. En esos casos determinar $\operatorname{Ker} f$.

c) Sea $\lambda = 1$. Encontrar una base de $\operatorname{Ker} f$.

d) Sea $\lambda = 1$ y $\mu = 0$. Encontrar la matriz coordenada B_1 de f respecto de la base $\mathcal{B}_1$, la matriz coordenada B_2 de f respecto de la base $\mathcal{B}_2$ y la matriz P que proporciona la relación de semejanza entre las dos matrices anteriores, siendo

$$\mathcal{B}_1 = \{(-1,0,1),(0,1,0),(4,1,2)\}, \quad \mathcal{B}_2 = \{(1,1,2),(1,1,0),(-1,1,-1)\}.$$

14. Sea $f : \mathbb{R}^3 \to \operatorname{Mat}_2(\mathbb{R})$ la aplicación lineal dada por

$$f(x_1,x_2,x_3) = \begin{pmatrix} x_1 - x_2 & x_2 - x_3 \\ x_3 - x_1 & x_1 - x_2 \end{pmatrix}.$$

a) Hallar la matriz coordenada de f en bases canónicas.

b) Hallar el núcleo y la imagen de f.

15. Sea el endomorfismo $f : \mathbb{R}^3 \to \mathbb{R}^3$ cuya matriz coordenada en base $\mathcal{B}$ es A:

$$\mathcal{B} = \{(1,-1,0),(0,1,-1),(-1,0,0)\}, \quad A = \begin{pmatrix} -2 & a & 1 \\ 1 & -2a & 1 \\ 1 & a & -2 \end{pmatrix}.$$

a) Hallar la matriz coordenada de f en base canónica.

b) Calcular los subespacios $\operatorname{Ker} f$ e $\operatorname{Im} f$ según los valores de a.

16. Dadas las bases $\mathcal{B}_1 = \{(1,4),(1,3)\}$ de $\mathbb{R}^2$ y $\mathcal{B}_2 = \{(2,0,1),(0,-1,0),(3,0,0)\}$ de $\mathbb{R}^3$, hallar la matriz coordenada B en bases $\mathcal{B}_1$ y $\mathcal{B}_2$ de una aplicación $f : \mathbb{R}^2 \to \mathbb{R}^3$ si su matriz en bases canónicas es

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

Comprobar que las matrices A y B son equivalentes.

17. Dado el endomorfismo $f : \mathbb{R}^4 \to \mathbb{R}^4$ definido por

$$f(\mathbf{e}_1) = 2\mathbf{e}_2, \quad f(\mathbf{e}_2) = \mathbf{e}_1 + a\,\mathbf{e}_2, \quad f(\mathbf{e}_3) = \mathbf{e}_1 + 2\mathbf{e}_2 - \mathbf{e}_3, \quad f(\mathbf{e}_4) = \mathbf{e}_1 - \mathbf{e}_2 + \mathbf{e}_3 - \mathbf{e}_4,$$

donde $\mathcal{B} = \{\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3,\mathbf{e}_4\}$ es una base de $\mathbb{R}^4$:

a) Calcular la matriz coordenada de f respecto a la base $\mathcal{B}$.

b) Indicar si el endomorfismo es inyectivo, suprayectivo o biyectivo dependiendo de los valores del parámetro a.

c) Para $a = 0$ encontrar la preimagen del vector $\mathbf{v}$ cuyas coordenadas en base $\mathcal{B}$ son $(1,-1,1,-1)$.

18. Sea f un endomorfismo de un espacio vectorial V de dimensión 3. Sean $\mathcal{B}_1 = \{\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_3\}$ y $\mathcal{B}_2 = \{\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3\}$ dos bases de V de modo que

$$f(\mathbf{u}_1) = 2\mathbf{e}_1 + a\,\mathbf{e}_2 + \mathbf{e}_3, \quad f(\mathbf{u}_2) = \mathbf{e}_1 + a\,\mathbf{e}_2 + 2\mathbf{e}_3, \quad f(\mathbf{u}_3) = 3\mathbf{e}_1 + 3\mathbf{e}_3.$$

$$\mathbf{u}_1 = \mathbf{e}_1 + \mathbf{e}_2, \quad \mathbf{u}_2 = \mathbf{e}_1 + \mathbf{e}_3, \quad \mathbf{u}_3 = \mathbf{e}_2 + \mathbf{e}_3.$$

a) Hallar la matriz coordenada A de f en base $\mathcal{B}_1$.

b) Hallar la matriz coordenada B de f en base $\mathcal{B}_2$.

c) Hallar una matriz P regular de modo que $B = P^{-1}AP$.
d) Hallar $\operatorname{Ker} f$ e $\operatorname{Im} f$ según los distintos valores de a.

19. Hallar la matriz coordenada asociada a la aplicación lineal $f : V \to W$ respecto de $\mathcal{B}_V = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ y $\mathcal{B}_W = \{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \mathbf{w}_4\}$ sabiendo que

$$f(\mathbf{v}_1+2\mathbf{v}_2-3\mathbf{v}_3) = \mathbf{w}_1-\mathbf{w}_3-\mathbf{w}_4, \quad f(2\mathbf{v}_1+\mathbf{v}_3) = 2\mathbf{w}_2-\mathbf{w}_4, \quad f(3\mathbf{v}_1+\mathbf{v}_2) = \mathbf{w}_2+2\mathbf{w}_3.$$

Tipos de aplicaciones

20. Sean V y W dos espacios vectoriales ambos con dimensión finita n y $f : V \to W$ una aplicación lineal. Probar que si f es suprayectiva, entonces f es biyectiva.

21. Demostrar que una aplicación lineal $f : \mathbb{R}^{n+1} \to \mathbb{R}^n$ no es inyectiva. Demostrar también que es suprayectiva si y solo si $\dim(\operatorname{Ker} f) = 1$.

22. Construir una matriz A que pueda corresponder a una aplicación lineal:

a) $f : \mathbb{R}^3 \to \mathbb{R}^4$ inyectiva.
b) $f : \mathbb{R}^4 \to \mathbb{R}^3$ suprayectiva.
c) entre dos espacios cualesquiera biyectiva.

23. Razonar las respuestas:

a) ¿Una aplicación lineal $f : \mathbb{R}^2 \to \mathbb{R}^4$ puede ser inyectiva?
b) ¿Una aplicación lineal $f : \mathbb{R}^3 \to \mathbb{R}^2$ puede ser inyectiva?
c) ¿Una aplicación lineal $f : \mathbb{R}^2 \to \mathbb{R}^4$ puede ser suprayectiva?
d) ¿Una aplicación lineal $f : \mathbb{R}^2 \to \mathbb{R}^4$ puede ser inyectiva sin ser suprayectiva?
e) ¿Una aplicación lineal $f : \mathbb{R}^6 \to \mathbb{R}^6$ puede ser inyectiva sin ser suprayectiva?

24. Sea f un endomorfismo de $\mathbb{R}^3$ con matriz coordenada respecto de la base canónica

$$A = \begin{pmatrix} 2 & a+2 & a \\ 3 & 3 & 1 \\ 1 & 2 & 1 \end{pmatrix}.$$

Discutir según los valores de $a \in \mathbb{R}$ la existencia de f^{-1} y hallar su matriz coordenada respecto de la base canónica en los casos que exista.

25. Se define la aplicación $f : \mathbb{R}_n[x] \to \mathbb{R}_n[x]$ mediante $f(p(x)) = mp(x) + p'(x)$.

a) Probar que f es lineal.
b) Hallar la matriz coordenada de f en base canónica.
c) Hallar $\operatorname{Ker} f$ e $\operatorname{Im} f$ según los valores de m. ¿Existe algún valor de m para el que f sea un isomorfismo?

26. Considerar el subespacio $E = \left\{ \begin{pmatrix} 0 & a \\ b & a+b \end{pmatrix} \mid a, b \in \mathbb{R} \right\} \subset \operatorname{Mat}_2(\mathbb{R})$ y el endomorfismo $f : E \to E$ dado por

$$f\begin{pmatrix} 0 & a \\ b & a+b \end{pmatrix} = \begin{pmatrix} 0 & 6a+3b \\ -2a-b & 4a+2b \end{pmatrix}.$$

a) Comprobar que $\left\{\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix}\right\}$ forma una base de E.

b) Hallar la matriz coordenada de f respecto a la base anterior.

c) Hallar $\operatorname{Ker} f$ e $\operatorname{Im} f$. ¿Es f inyectiva? ¿Y suprayectiva?

27. Considerar $f : \mathbb{R}^3 \to \mathbb{R}^3$ un endomorfismo que verifica:

$$f(0,1,1) = (0,a,2), \quad f(0,1,0) = (-5,0,-3), \quad f(1,-1,0) = (8,-2,6).$$

a) Hallar la expresión general de f: $f(x,y,z)$.

b) Hallar la matriz coordenada de f respecto a la base canónica.

c) Indicar si f es inyectiva, suprayectiva o biyectiva en función de los valores del parámetro a.

28. Encontrar la matriz coordenada de f en bases canónicas, comprobar que es biyectiva y determinar f^{-1}, siendo:

$$\begin{array}{rccc} f: & \mathbb{R}_2[x] & \longrightarrow & \mathbb{R}^3 \\ & a+bx+cx^2 & \mapsto & (a-c, b, 2a-c). \end{array}$$

7.6. SOLUCIONES A LOS EJERCICIOS PROPUESTOS

Aplicaciones lineales y subespacios

1. Son aplicaciones lineales los casos: b), c), e), f).

2. $f(a+bx) = (5a-3b) + \left(\frac{19b-a}{4}\right)x$. $f(10x-15) = -105 + \frac{205}{4}x$.

3. La manera más simple de definir f es: $f(x,y,z,t,w) = (x,0,x,x+y)$. Otra forma es dar la imagen de los elementos de una base de $\mathbb{R}^5$. Por ejemplo, podemos considerar:

$$f(\mathbf{e}_1) = f(\mathbf{e}_2) = f(\mathbf{e}_3) = (1,0,1,1), \quad f(\mathbf{e}_4) = f(\mathbf{e}_5) = (0,0,0,1).$$

4. a) $\mathcal{B}_{f(S)} = \{(1,3,0),(7,4,0)\}$. b) $\mathcal{B}_{f(T)} = \{(2,7,0),(3,-2,0)\}$.
En ambos casos también sirve la siguiente base: $\{(1,0,0),(0,1,0)\}$.

5. $f^{-1}(\mathbf{w}) = (0,3,1)$.

6. a) El subespacio S es precisamente $\operatorname{Ker} f$. $\mathcal{B}_S = \{(0,0,1)\}$, $\dim S = 1$.

b) El subespacio T es precisamente $\operatorname{Im} f$. $\mathcal{B}_T = \{(1,1),(1,-1)\}$ o, alternativamente, $\{(1,0),(0,1)\}$. $\dim T = 2$.

7. a) $\operatorname{Ker} f = \langle x^2 - x\rangle$. $\operatorname{Im} f = \langle (1,0),(0,1)\rangle$.

b) $\operatorname{Ker} g = \left\langle \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix} \right\rangle$. $\operatorname{Im} g = \left\langle \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \right\rangle$.

En ambos casos, las bases vienen dadas por los elementos de los sistemas generadores.

8. a) $\operatorname{Ker} f|_U = \{\mathbf{0}\} \Longrightarrow \mathcal{B}_{\operatorname{Ker} f|_U} = \emptyset$. $\mathcal{B}_{\operatorname{Im} f|_U} = \{(1,0),(0,1)\}$.

b) $\operatorname{Ker} g|_U = \{\mathbf{0}\} \Longrightarrow \mathcal{B}_{\operatorname{Ker} g|_U} = \emptyset$.

$$\mathcal{B}_{\operatorname{Im} g|_U} = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} \right\}.$$

9. Utilizar que $g(\mathbf{0}) = \mathbf{0}$ por ser g una aplicación lineal.

10. a) $\dim(\operatorname{Ker} A) = 3$.

b) El sistema no es compatible pues $\dim(\operatorname{Im} A) = \operatorname{rg} A = 4 \neq 5$.

Matriz coordenada. Equivalencia y semejanza de matrices

11. Para que dos matrices sean equivalentes han de tener la misma dimensión y el mismo rango. Esas dos propiedades solo se cumplen en el apartado b).

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

c) $\operatorname{Ker} f = \{\mathbf{0}\} \Longrightarrow \mathcal{B}_{\operatorname{Ker} f} = \emptyset$ y $\dim(\operatorname{Ker} f) = 0$.
$\operatorname{Im} f = \operatorname{Mat}_2(\mathbb{R})$. Como base de $\operatorname{Im} f$ se puede considerar, por ejemplo, la base canónica del espacio $\operatorname{Mat}_2(\mathbb{R}) \Longrightarrow \dim(\operatorname{Im} f) = 4$.

d) $B = \begin{pmatrix} -1 & 2 & 4 & 0 \\ 0 & \nicefrac{-5}{3} & \nicefrac{-4}{3} & \nicefrac{4}{3} \\ 0 & \nicefrac{4}{3} & \nicefrac{5}{3} & \nicefrac{-2}{3} \\ 0 & 0 & 0 & 1 \end{pmatrix}.$

e) Las matrices A y B verifican la relación $B = P^{-1}AP$, siendo $P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$

13. a) La matriz coordenada en base canónica es $A = \begin{pmatrix} \lambda & \mu & 1 \\ 1 & \lambda\mu & 1 \\ 1 & \mu & \lambda \end{pmatrix}.$

b) $\operatorname{Im} f = \mathbb{R}^3 \iff \operatorname{rg} A = 3 \iff \mu(\lambda-1)^2(\lambda+2) \neq 0$. En esos casos, $\operatorname{Ker} f = \{\mathbf{0}\}$.

c) $\mathcal{B}_{\operatorname{Ker} f} = \{(1,0,-1),(0,1,-\mu)\}$.

d) $B_1 = \begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & 4 \\ 0 & 0 & 2 \end{pmatrix}$. $B_2 = \begin{pmatrix} \nicefrac{3}{2} & \nicefrac{1}{2} & -1 \\ \nicefrac{3}{2} & \nicefrac{1}{2} & -1 \\ 0 & 0 & 0 \end{pmatrix}$. $P = P_{\mathcal{B}_2,\mathcal{B}_1} = \begin{pmatrix} 1 & \nicefrac{-1}{3} & \nicefrac{-1}{3} \\ \nicefrac{1}{2} & \nicefrac{5}{6} & \nicefrac{4}{3} \\ \nicefrac{1}{2} & \nicefrac{1}{6} & \nicefrac{-1}{3} \end{pmatrix}.$

14. a) $A = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix}.$

b) $\operatorname{Ker} f = \langle (1,1,1) \rangle$. $\operatorname{Im} f = \left\langle \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \right\rangle.$

15. a) $B = \begin{pmatrix} -3 & 0 & 0 \\ 0 & -3 & 3(a-1) \\ 1 & 2 & -2(a-1) \end{pmatrix}.$

b) $\operatorname{Ker} f = \langle (0, a-1, 1) \rangle$. $\operatorname{Im} f = \langle (-3,0,1),(0,-3,2) \rangle$.

16. La matriz $B = \begin{pmatrix} 1 & 1 \\ -15 & -12 \\ -4 & \nicefrac{-10}{3} \end{pmatrix}$ y se obtiene haciendo $B = Q^{-1}AP$, siendo $P = \begin{pmatrix} 1 & 1 \\ 4 & 3 \end{pmatrix}$ y $Q = \begin{pmatrix} 2 & 0 & 3 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$

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

b) $\operatorname{rg} A = 4 \quad \forall a \in \mathbb{R}$, por lo que f es biyectivo para todo valor de a.

c) $f^{-1}(1,-1,1,-1) = \mathbf{e}_4$.

18. a) $A = \begin{pmatrix} \nicefrac{(a+1)}{2} & \nicefrac{(a-1)}{2} & 0 \\ \nicefrac{(3-a)}{2} & \nicefrac{(3-a)}{2} & 3 \\ \nicefrac{(a-1)}{2} & \nicefrac{(a+1)}{2} & 0 \end{pmatrix}$. b) $B = \begin{pmatrix} 0 & 2 & 1 \\ a & 0 & 0 \\ 0 & 1 & 2 \end{pmatrix}$. c) $P = \begin{pmatrix} \nicefrac{1}{2} & \nicefrac{1}{2} & \nicefrac{-1}{2} \\ \nicefrac{1}{2} & \nicefrac{-1}{2} & \nicefrac{1}{2} \\ \nicefrac{-1}{2} & \nicefrac{1}{2} & \nicefrac{1}{2} \end{pmatrix}.$

d) En términos de $\mathcal{B}_2$: $\operatorname{Ker} f = \begin{cases} \langle \mathbf{e}_1 \rangle, & a = 0 \\ \mathbf{0}, & a \neq 0, \end{cases}$ $\qquad \operatorname{Im} f = \begin{cases} \langle \mathbf{e}_1, \mathbf{e}_3 \rangle, & a = 0 \\ V & a \neq 0. \end{cases}$

19. $A = \begin{pmatrix} 1 & -3 & -2 \\ 4 & -11 & -6 \\ -5 & 17 & 10 \\ -4 & 12 & 7 \end{pmatrix}.$

Tipos de aplicaciones

20. Ayuda: Utilizar el teorema de las dimensiones.

21. Ayuda: Utilizar el teorema de las dimensiones.

22. a) $A \in \mathrm{Mat}_{4\times 3}(\mathbb{R})$ y $\operatorname{rg} A = 3$. b) $A \in \mathrm{Mat}_{3\times 4}(\mathbb{R})$ y $\operatorname{rg} A = 3$. c) A debe ser cuadrada y regular.

23. Ayuda: Utilizar el teorema de las dimensiones. Los apartados a) y d) tienen respuesta afirmativa.

24. Existe $f^{-1} \iff a \neq 2$. En ese caso, $A^{-1} = \frac{1}{a-2}\begin{pmatrix} 1 & a-2 & 2(1-a) \\ -2 & 2-a & 3a-2 \\ 3 & a-2 & -3a \end{pmatrix}.$

25. b) $A = \begin{pmatrix} m & 1 & 0 & \cdots & \cdots \\ 0 & m & 2 & \cdots & \cdots \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & m & n \\ 0 & 0 & \cdots & 0 & m \end{pmatrix} \in \mathrm{Mat}_{n+1}(\mathbb{R}).$

c) A es una matriz triangular superior. Rango de $A = n+1 \iff m \neq 0$.

$$\begin{cases} m \neq 0 \Longrightarrow & f \text{ es isomorfismo.} \\ m = 0 \Longrightarrow & \begin{cases} \operatorname{Ker} f = \langle 1 \rangle \\ \operatorname{Im} f = \langle 1,\, x,\, \ldots, x^{n-1} \rangle. \end{cases} \end{cases}$$

26. b) $A = \begin{pmatrix} 2 & 6 \\ 1 & 3 \end{pmatrix}.$

c) $\operatorname{Ker} f = \left\langle \begin{pmatrix} 0 & 1 \\ -2 & -1 \end{pmatrix} \right\rangle$. $\operatorname{Im} f = \left\langle \begin{pmatrix} 0 & 3 \\ -1 & 2 \end{pmatrix} \right\rangle$. f no es inyectiva ni suprayectiva.

27. a) $f(x, y, z) = (3x - 5y + 5z, -2x + az, 3x - 3y + 5z)$.

b) $A = \begin{pmatrix} 3 & -5 & 5 \\ -2 & 0 & a \\ 3 & -3 & 5 \end{pmatrix}.$

c) $\operatorname{rg} A = 3 \iff a \neq \frac{-10}{3}$. Por tanto, si $a \neq \frac{-10}{3}$, entonces f es un isomorfismo. Si $a = \frac{-10}{3}$, f no es ni inyectiva ni suprayectiva.

28. $A = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 2 & 0 & -1 \end{pmatrix}$. Como el rango de A es 3, entonces f es un isomorfimso. La matriz coordenada de f^{-1} es $A^{-1} = \begin{pmatrix} -1 & 0 & 1 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix} \Longrightarrow$ por lo que

$$f^{-1}(d, e, f) = (-d + f) + ex + (-2d + f)x^2.$$

CAPÍTULO 8

Valores y Vectores Propios: Diagonalización y Jordan

Tras el estudio de las aplicaciones lineales en general, este capítulo se centra en un caso particular de estas, los endomorfismos. Se recuerda que un endomorfismo es una aplicación lineal de un espacio vectorial V en sí mismo. Si el espacio V tiene dimensión n, la matriz coordenada del endomorfismo en una determinada base de V es cuadrada, $A \in \mathrm{Mat}_n(\mathbb{R})$.

Dado un endomorfismo $f : V \to V$, un problema interesante es encontrar una base de V en la que la matriz coordenada de f adopte la forma "más sencilla posible". En particular se pretende encontrar, cuando sea posible, una base en la que la matriz coordenada sea una matriz diagonal. En caso de existir dicha base y la matriz diagonal correspondiente, el endomorfismo se dice diagonalizable. Para resolver el problema anterior es necesario introducir las definiciones de valor y vector propio. Estos conceptos surgieron en el siglo XVIII con el estudio de ecuaciones diferenciales desarrollado por Lagrange y Laplace. Cauchy también los utilizó, así como Sylvester en sus trabajos aplicados a física que generalizó Weierstrass.

Para aquellos endomorfismos f en los que no sea posible encontrar la base en la que la matriz coordenada sea diagonal, es decir endomorfismos no diagonalizables, se buscan bases de V en las que la matriz coordenada sea "casi diagonal". Utilizando la noción de matrices semejantes, Jordan en su *Traité des substitutions* (1870) resolvió el problema, introduciendo la denominada forma canónica de Jordan.

8.1. VALORES Y VECTORES PROPIOS

Se introducen a continuación algunos de los conceptos necesarios para abordar el problema de la diagonalización de un endomorfismo $f : V \to V$, es decir, encontrar una base de V en la que la matriz coordenada de f sea diagonal.

Para ello se estudian aquellos vectores no nulos de V tales que la imagen por f es proporcional al vector. Se dice que tales vectores son *vectores propios* de f y la constante de proporcionalidad se denomina *valor propio* de f.

Las definiciones se pueden plantear en términos del endomorfismo f, o paralelamente, para una matriz cuadrada A.

Definición 8.1.

- Sea V un espacio vectorial real y $f : V \to V$ un endomorfismo. Sea $\mathbf{v} \in V$ un vector no nulo ($\mathbf{v} \neq \mathbf{0}$) y $\lambda \in \mathbb{R}$. Se dice que $\mathbf{v}$ es un **vector propio** de f de **valor propio** λ si:
$$f(\mathbf{v}) = \lambda \mathbf{v}.$$
- Sea $A \in \mathrm{Mat}_n(\mathbb{R})$ una matriz cuadrada. Sea $\mathbf{x} \in \mathbb{R}^n$ un vector no nulo ($\mathbf{x} \neq \mathbf{0}$) y $\lambda \in \mathbb{R}$. Se dice que $\mathbf{x}$ es un **vector propio** de A de **valor propio** λ si:
$$A\mathbf{x} = \lambda \mathbf{x}.$$

Hay que observar que los conceptos de valor y vector propio van ligados, es decir, para que exista uno debe existir el otro.

✓ **Ejemplo 1.** Dado el endomorfismo

$$\begin{array}{cccc} f : & \mathbb{R}^3 & \longrightarrow & \mathbb{R}^3 \\ & (x, y, z) & \mapsto & (x - y - z,\ -x + y - z,\ -x - y + z) \end{array}$$

comprobar que el vector $\mathbf{v} = (1, 1, 1)$ es vector propio de f de valor propio $\lambda = -1$.

Solución:

$$f(\mathbf{v}) = f(1, 1, 1) = (-1, -1, -1) = -1\,(1, 1, 1) = -1\,\mathbf{v}.$$

◇

Los apartados 1 y 2 de la Definición 8.1 se relacionan del siguiente modo:

Sea V un espacio vectorial real de dimensión finita n, $f : V \to V$ un endomorfismo y $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ una base de V. Sea $A \in \mathrm{Mat}_n(\mathbb{R})$ la matriz coordenada de f en la base $\mathcal{B}$. Sea $\mathbf{v} \in V$ un vector no nulo cuyas coordenadas en $\mathcal{B}$ son

$$\mathcal{C}_{\mathcal{B}}(\mathbf{v}) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \mathbf{x}.$$

Entonces:

> $\mathbf{v}$ es un vector propio de f de valor propio λ si y solo si $\mathbf{x}$ es un vector propio de A de valor propio λ.

Nota. En la literatura pueden encontrarse otras nomenclaturas para los conceptos de valor y vector propio, por ejemplo *autovalor* para los valores propios y *autovector* para los vectores propios.

El cálculo de los valores y vectores propios directamente con la definición puede no ser del todo efectivo, sobre todo si se pretende utilizar f. Se propone el siguiente teorema que proporciona un resultado para el cálculo práctico de los valores propios. En la sección siguiente se introduce el cálculo de los vectores propios.

Teorema 8.2. Sea V un espacio vectorial real de dimensión finita n y $f : V \to V$ un endomorfismo. Sea $A \in \mathrm{Mat}_n(\mathbb{R})$ la matriz coordenada de f en una cierta base $\mathcal{B}$ de V. Las siguientes afirmaciones son equivalentes:

1. λ es un valor propio de A.
2. El sistema de ecuaciones $(A-\lambda I_n)\mathbf{x} = \mathbf{0}$ tiene soluciones distintas de la trivial.
3. $|A - \lambda I_n| = 0$.

Demostración. Por su propia definición $\lambda \in \mathbb{R}$ es un valor propio de la matriz $A \in \mathrm{Mat}_n(\mathbb{R}) \Leftrightarrow$ existe un vector $\mathbf{x} \in \mathbb{R}^n$ no nulo tal que $A\mathbf{x} = \lambda\mathbf{x} \Leftrightarrow$ el sistema homogéneo $(A - \lambda I_n)\mathbf{x} = \mathbf{0}$ tiene solución no trivial. Esto ocurre si y solo si $\mathrm{rg}(A - \lambda I_n) < n \Leftrightarrow |A - \lambda I_n| = 0$.

□

Definición 8.3. Sea $A \in \mathrm{Mat}_n(\mathbb{R})$ una matriz cuadrada. El polinomio de grado n que resulta de calcular $p(t) = |tI_n - A|$, se llama **polinomio característico** de A:

$$p(t) = |tI_n - A| = t^n - (\mathrm{tr}\, A)t^{n-1} + \cdots + (-1)^n|A|.$$

Observación. Tal como se ha definido el polinomio característico es mónico, es decir, su coeficiente director es $+1$. En otros textos puede aparecer definido como $p(t) = |A - tI_n|$.

Corolario 8.4. Los valores propios de una matriz cuadrada $A \in \mathrm{Mat}_n(\mathbb{R})$ son las raíces *reales* de su polinomio característico.

Observación. El polinomio característico de una matriz $A \in \mathrm{Mat}_n(\mathbb{R})$, cuadrada con coeficientes reales, puede tener todas sus raíces reales (en este caso coinciden con los valores propios) o puede tener raíces no reales, como muestran los siguientes ejemplos:

- La matriz $A = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{R})$, cuadrada real, tiene polinomio característico $p(t) = t^2 - 1$, cuyas raíces son -1 y 1. Es decir, los valores propios de A son $\lambda_1 = -1$ y $\lambda_2 = 1$.

- La matriz $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{R})$, cuadrada real, tiene polinomio característico $p(t) = t^2 + 1$, cuyas raíces son complejas i, $-i$. Por tanto, según la definición A no tiene valores propios.

Observación. De aquí en adelante se consideran matrices $A \in \mathrm{Mat}_n(\mathbb{R})$ en las que todas las raíces del polinomio característico son reales.

Definición 8.5. Sea V un espacio vectorial real de dimensión finita n, $f : V \to V$ un endomorfismo y $\mathcal{B}$ una base cualquiera de V. Sea $A \in \mathrm{Mat}_n(\mathbb{R})$ la matriz coordenada de f en la base $\mathcal{B}$. Se llama **polinomio característico** de f al polinomio característico de A.

Observación. Esta definición no depende de la base escogida, ya que si se considera otra base $\mathcal{B}_1$ en la que la matriz coordenada de f sea B, entonces, como se ha visto en el Capítulo 7, existe una matriz P regular tal que $B = P^{-1}AP$ y,

$$\begin{aligned} |tI_n - B| &= |tI_n - P^{-1}AP| = |P^{-1}tI_nP - P^{-1}AP| \\ &= |P^{-1}(tI_n - A)P| = |P^{-1}|\,|tI_n - A|\,|P| \\ &= |tI_n - A|. \end{aligned}$$

Consecuencias.

- Matrices semejantes tienen el mismo polinomio característico y por tanto los mismos valores propios.

- λ es valor propio de $f \iff \lambda \in \mathbb{R}$ y es raíz de su polinomio característico.

Definición 8.6. Sea V un espacio vectorial real, $f : V \to V$ un endomorfismo y $A \in \mathrm{Mat}_n(\mathbb{R})$ la matriz coordenada de f en una base $\mathcal{B}$ de V. Sea $\lambda \in \mathbb{R}$ uno de los valores propios de f, o equivalentemente, de la matriz A. Se define la **multiplicidad** de λ como el número de veces que aparece como raíz del polinomio característico de f.

Según esto el polinomio característico se puede escribir como:

$$p(t) = (t - \lambda_1)^{m_1} \cdots (t - \lambda_r)^{m_r},$$

donde $\lambda_1, \dots, \lambda_r \in \mathbb{R}$ son los valores propios (distintos) y $m_1, \dots, m_r$ son sus multiplicidades respectivas.

Así, para calcular los valores propios de un endomorfismo $f : V \to V$ se propone seguir los siguientes pasos:

- Hallar la matriz coordenada A de f en una cierta base de V (por simplicidad se recomienda la base canónica).
- Hallar el polinomio característico de A.
- Hallar las raíces reales de dicho polinomio. Es decir, resolver $|\lambda I_n - A| = 0$, o equivalentemente, $|A - \lambda I_n| = 0$.

✓ **Ejemplo 2.** Dado el endomorfismo:

$$\begin{array}{rccc} f: & \mathbb{R}^3 & \longrightarrow & \mathbb{R}^3 \\ & (x,y,z) & \mapsto & (x-y-z,\ -x+y-z,\ -x-y+z), \end{array}$$

calcular los valores propios de f y sus multiplicidades asociadas.

Solución: Se calcula la matriz coordenada de f en la base canónica de $\mathbb{R}^3$:

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

Para determinar los valores propios de f se utiliza la matriz A, calculando las raíces de su polinomio característico (con sus correspondientes multiplicidades):

$$|A - \lambda I_3| = \begin{vmatrix} 1-\lambda & -1 & -1 \\ -1 & 1-\lambda & -1 \\ -1 & -1 & 1-\lambda \end{vmatrix} = \begin{vmatrix} -1-\lambda & -1-\lambda & -1-\lambda \\ -1 & 1-\lambda & -1 \\ -1 & -1 & 1-\lambda \end{vmatrix} =$$

$$-(1+\lambda) \begin{vmatrix} 1 & 1 & 1 \\ -1 & 1-\lambda & -1 \\ -1 & -1 & 1-\lambda \end{vmatrix} = -(1+\lambda) \begin{vmatrix} 1 & 1 & 1 \\ 0 & 2-\lambda & 0 \\ 0 & 0 & 2-\lambda \end{vmatrix} = -(\lambda+1)(\lambda-2)^2.$$

Así el polinomio característico y los valores propios con sus respectivas multiplicidades son:

$$p(t) = (t+1)(t-2)^2 \quad \Longrightarrow \quad \begin{cases} \lambda_1 = -1, & m_1 = 1 \\ \lambda_2 = 2, & m_2 = 2. \end{cases}$$

◇

8.2. SUBESPACIOS FUNDAMENTALES

Dado un endomorfismo f sobre el espacio vectorial real V, el conjunto de los vectores propios asociados al mismo valor propio, junto con el vector nulo, constituyen un subespacio vectorial de V, como se demuestra a continuación. Si se considera un valor propio λ y un vector propio asociado $\mathbf{v} \in V$, entonces todo múltiplo escalar de $\mathbf{v}$ distinto de cero también es un vector propio de f, es decir, dado $\alpha \in \mathbb{R}$:

$$f(\alpha\mathbf{v}) = \alpha f(\mathbf{v}) = \alpha(\lambda\mathbf{v}) = \lambda(\alpha\mathbf{v}).$$

Además si se consideran dos vectores propios $\mathbf{v}_1, \mathbf{v}_2 \in V$ asociados al mismo valor propio λ, entonces su suma también es un vector propio correspondiente al mismo valor propio:

$$f(\mathbf{v}_1 + \mathbf{v}_2) = f(\mathbf{v}_1) + f(\mathbf{v}_2) = \lambda\mathbf{v}_1 + \lambda\mathbf{v}_2 = \lambda(\mathbf{v}_1 + \mathbf{v}_2).$$

Es decir, se puede dar la siguiente definición:

Definición 8.7. Sea V un espacio vectorial real, $f : V \to V$ un endomorfismo y $\lambda \in \mathbb{R}$ un valor propio de f. El subespacio

$$S(f, \lambda) = \{\mathbf{v} \in V \mid f(\mathbf{v}) = \lambda\mathbf{v}\}$$

se dice **subespacio fundamental** de valor propio λ para f.

Observaciones.

- El subespacio anterior $S(f, \lambda) \neq \{\mathbf{0}\}$, es decir $\dim S(f, \lambda) \geq 1$, por la propia definición de valor y vector propio, ya que si λ es un valor propio de f, entonces existe $\mathbf{v} \neq \mathbf{0}$ cumpliendo $f(\mathbf{v}) = \lambda\mathbf{v}$, es decir, existe $\mathbf{v} \in S(f, \lambda)$.
- $S(f, \lambda)$ está formado por los vectores propios de f de valor propio λ y por el vector nulo de V, ya que si este no estuviese no podría ser subespacio.

De manera análoga si se trabaja con la matriz cuadrada A se puede dar la siguiente definición:

Definición 8.8. Sea $A \in \mathrm{Mat}_n(\mathbb{R})$ una matriz cuadrada y $\lambda \in \mathbb{R}$ un valor propio de A. El subespacio

$$S(A, \lambda) = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \lambda\mathbf{x}\}$$

se dice **subespacio fundamental** de valor propio λ para A.

Nota. El subespacio fundamental $S(A, \lambda)$ está formado por los vectores propios de A de valor propio λ y el vector nulo. Además este subespacio no puede ser nulo pues en ese caso no existiría el valor propio λ.

Observaciones.

- Si A es la matriz coordenada de $f : V \to V$ en la base $\mathcal{B}$ del espacio vectorial real V, entonces $S(A, \lambda)$ está formado por las n-tuplas coordenadas de los vectores de $S(f, \lambda)$ en la base $\mathcal{B}$.
- En la práctica suele resultar más sencillo calcular $S(A, \lambda)$, para lo cual se resuelve $A\mathbf{x} = \lambda\mathbf{x}$ o equivalentemente el sistema homogéneo compatible indeterminado $(A - \lambda I_n)\mathbf{x} = \mathbf{0}$. Es decir,

$$\boxed{S(A, \lambda) = \operatorname{Ker}(A - \lambda I_n)}$$

✓ **Ejemplo 3.** Dado el endomorfismo $f : \mathbb{R}^3 \to \mathbb{R}^3$ definido por

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix}}_{A} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix},$$

calcular los subespacios fundamentales. (Observar que se trata del mismo endomorfismo del Ejemplo 2).

Solución: Los valores propios ya han sido calculados previamente

$$\begin{cases} \lambda_1 = -1, & m_1 = 1 \\ \lambda_2 = 2, & m_2 = 2. \end{cases}$$

En este caso, por tratarse de un endomorfismo de $\mathbb{R}^3$ y la matriz A ser matriz coordenada de f en base canónica, el subespacio $S(f, \lambda)$ coincide con el subespacio $S(A, \lambda)$.

1. $\boxed{S(A, -1) = \operatorname{Ker}(A + I_3)}$

   Es decir, se debe encontrar la solución general del sistema $(A + I_3)\mathbf{x} = \mathbf{0}$. Para ello se usa el método de Gauss como se hace habitualmente.

$$\underbrace{\left(\begin{array}{ccc|c} 2 & -1 & -1 & 0 \\ -1 & 2 & -1 & 0 \\ -1 & -1 & 2 & 0 \end{array}\right)}_{(\,A+I_3\ |\ \mathbf{0}\,)} \xrightarrow{P_{12}} \left(\begin{array}{ccc|c} -1 & 2 & -1 & 0 \\ 2 & -1 & -1 & 0 \\ -1 & -1 & 2 & 0 \end{array}\right) \xrightarrow[P_{31}(-1)]{P_{21}(2)}$$

$$\left(\begin{array}{ccc|c} -1 & 2 & -1 & 0 \\ 0 & 3 & -3 & 0 \\ 0 & -3 & 3 & 0 \end{array}\right) \xrightarrow{P_{32}(1)} \left(\begin{array}{ccc|c} -1 & 2 & -1 & 0 \\ 0 & 3 & -3 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) \xrightarrow{P_2(1/3)}$$

$$\left(\begin{array}{ccc|c} -1 & 2 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) \xrightarrow{P_{12}(-2)} \left(\begin{array}{ccc|c} -1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

Las ecuaciones implícitas y paramétricas del sistema anterior son:

$$\left\{\begin{array}{r} -x+z=0 \\ y-z=0 \end{array}\right. \quad \Longrightarrow \quad \left\{\begin{array}{l} x=\alpha \\ y=\alpha \\ z=\alpha \end{array}\right. \quad \alpha \in \mathbb{R},$$

de donde se deduce que $S(A,-1)=\langle(1,1,1)\rangle$. Una base de este subespacio es $\{(1,1,1)\}$ y por tanto tiene dimensión 1.

Para asegurar que los cálculos están bien, se puede comprobar, siguiendo la definición, que $\mathbf{v}_1=(1,1,1)$ es un vector propio de A de valor propio $\lambda_1=-1$.

$$\underbrace{\begin{pmatrix} 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}}_{\mathbf{v}_1} = \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix} = \underbrace{-1}_{\lambda_1} \underbrace{\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}}_{\mathbf{v}_1}.$$

2. $\boxed{S(A,2)=\mathrm{Ker}(A-2I_3)}$

   En este caso se trata de resolver el sistema homogéneo $(A-2I_3)\mathbf{x}=\mathbf{0}$,

$$\underbrace{\left(\begin{array}{ccc|c} -1 & -1 & -1 & 0 \\ -1 & -1 & -1 & 0 \\ -1 & -1 & -1 & 0 \end{array}\right)}_{(\,A-2I_3\ \mid\ \mathbf{0}\,)} \xrightarrow[P_{31}(-1)]{P_{21}(-1)} \left(\begin{array}{ccc|c} -1 & -1 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) \xrightarrow{P_1(-1)} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right),$$

   cuyas soluciones son:

$$x+y+z=0 \quad \Longrightarrow \quad \left\{\begin{array}{l} x=\alpha \\ y=\beta \\ z=-\alpha-\beta \end{array}\right. \quad \alpha,\,\beta \in \mathbb{R}.$$

   Es decir, el subespacio fundamental es:

$$S(A,2)=\{(\alpha,\beta,-\alpha-\beta) \mid \alpha,\beta \in \mathbb{R}\}=\langle(1,0,-1),(0,1,-1)\rangle.$$

   De los cálculos anteriores se deduce que la dimensión de $S(A,2)$ es 2 y una base puede ser $\{(1,0,-1),(0,1,-1)\}$.

◇

Una vez entendidos los conceptos de valor y vector propio y subespacio fundamental, en el siguiente resultado se muestran algunas propiedades de estos subespacios.

Teorema 8.9. Sea V un espacio vectorial real de dimensión finita n. Sea $f: V \to V$ un endomorfismo y sean $\lambda_1,\ldots,\lambda_r \in \mathbb{R}$ los distintos valores propios de f con multiplicidades $m_1,\ldots,m_r$, respectivamente. Se considera $A \in \mathrm{Mat}_n(\mathbb{R})$ la matriz coordenada de f en una base $\mathcal{B}$ de V, entonces:

1. La suma de los subespacios fundamentales, $S(f,\lambda_i)$, $i = 1,\ldots,r$, es directa,
$$S(f,\lambda_1) \oplus \cdots \oplus S(f,\lambda_r).$$
2. La dimensión de los subespacios fundamentales se puede calcular como:
$$\dim S(f,\lambda_i) = n - \operatorname{rg}(A - \lambda_i I_n), \quad \forall i = 1,\ldots,r.$$
3. La dimensión de los subespacios fundamentales se acota como sigue:
$$1 \leq \dim S(f,\lambda_i) \leq m_i, \quad \forall i = 1,\ldots,r.$$

Observaciones.

- Se verifica: $\displaystyle\sum_{i=1}^{r} \dim S(f,\lambda_i) \leq n.$
- Si λ es un valor propio de multiplicidad 1, entonces $\dim S(f,\lambda) = 1$.

Nota. En el Ejemplo 3, $\dim S(A,-1) = 1 = m_1$ y $\dim S(A,2) = 2 = m_2$. En particular, como la suma de los subespacios fundamentales es directa y el espacio vectorial $V = \mathbb{R}^3$ tiene dimensión tres, se verifica además la igualdad:
$$\mathbb{R}^3 = S(A,-1) \oplus S(A,2).$$

El hecho de que la suma de los subespacios fundamentales sea el espacio total no tiene por qué ocurrir siempre (ver Teorema 8.11).

$\rightsquigarrow$ **Ejercicio.** Comprobar que los subesapcios fundamentales del Ejemplo 3 verifican $S(A,-1) \cap S(A,2) = \{\mathbf{0}\}$.

8.3. DIAGONALIZACIÓN

Dado un espacio vectorial real V de dimensión finita, $\dim V = n$, un endomorfismo $f : V \to V$ y una matriz coordenada de f, $A \in \operatorname{Mat}_n(\mathbb{R})$, el objetivo de este capítulo es encontrar, si existe, una base $\mathcal{D}$ de V en la que la matriz coordenada de f sea diagonal D, o equivalentemente encontrar una matriz D diagonal semejante a A. En caso de existir D, el endomorfismo f, y en consecuencia la matriz A, se dice *diagonalizable.*

Se comienza con un ejemplo sencillo antes de resolver el problema de manera general.

✓ **Ejemplo 4.** Se considera el endomorfismo $f : \mathbb{R}^2 \to \mathbb{R}^2$ dado por:

$$f(x, y) = (-x + 2y, y).$$

Se puede comprobar que la matriz coordenada de f en la base canónica de $\mathbb{R}^2$ es

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

mientras que la matriz coordenada en la base $\mathcal{D} = \{(1, 1), (1, 0)\}$ es diagonal:

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

Si se calculan los valores propios de A se obtienen $\lambda_1 = 1$ y $\lambda_2 = -1$. Calculando los subespacios fundamentales se observa que $S(f, 1) = \langle(1, 1)\rangle$ y $S(f, -1) = \langle(1, 0)\rangle$. Es decir, la matriz D tiene en su diagonal los valores propios de A y la base $\mathcal{D}$ está formada por vectores propios.

◇

Definición 8.10.

- Sea V un espacio vectorial real de dimensión finita n y $f : V \to V$ un endomorfismo. Se dice que f es **diagonalizable** si V posee una base en la que la matriz coordenada de f es diagonal.
- Sea $A \in \mathrm{Mat}_n(\mathbb{R})$, se dice que A es **diagonalizable** si es semejante a una matriz diagonal, es decir, existe una matriz P regular tal que $P^{-1}AP = D$, siendo $D \in \mathrm{Mat}_n(\mathbb{R})$ una matriz diagonal.

Observación. Los dos conceptos anteriores se encuentran relacionados, como se muestra a continuación. Sea A la matriz coordenada del endomorfismo $f : V \to V$ en una base $\mathcal{B}$ de V. Entonces f es diagonalizable si y solo si existe una base $\mathcal{D}$ de V con matriz coordenada diagonal D, es decir, A y D son matrices coordenadas de un mismo endomorfismo en bases distintas. Entonces, existe una matriz regular P, de cambio de coordenadas tal que $P^{-1}AP = D$, equivalentemente, A es diagonalizable. Conclusión:

$$\boxed{f \text{ es diagonalizable} \quad \Longleftrightarrow \quad A \text{ es diagonalizable}}$$

El proceso de diagonalización se puede resumir en el siguiente diagrama conmutativo:

$$\begin{tikzcd}
(\mathcal{B}) & \mathbb{R}^n \arrow[r, "A"] & \mathbb{R}^n & (\mathcal{B}) \\
(\mathcal{D}) & \mathbb{R}^n \arrow[u, "P"] \arrow[r, "D"'] \arrow[ur, phantom, "\#"] & \mathbb{R}^n \arrow[u, "P"'] & (\mathcal{D})
\end{tikzcd}$$

El problema que se plantea es encontrar condiciones bajo las que un endomorfismo es diagonalizable. El siguiente resultado proporciona la solución.

Teorema 8.11. Sea V un espacio vectorial real con $\dim V = n$ y $f : V \to V$ un endomorfismo cuyo polinomio característico tiene todas sus raíces reales. Si $\lambda_1, \ldots, \lambda_r \in \mathbb{R}$ son los distintos valores propios de f con multiplicidades respectivas $m_1, \ldots, m_r$, entonces son equivalentes:

1. f es diagonalizable.
2. $\dim S(f, \lambda_i) = m_i, \quad \forall i = 1, \ldots, r.$
3. $V = S(f, \lambda_1) \oplus \cdots \oplus S(f, \lambda_r)$.
4. Existe una base de V formada por vectores propios de f.

Corolario 8.12. Sea $f : V \to V$ un endomorfismo. Si el polinomio característico de f es $p(t) = (t - \lambda_1) \cdots (t - \lambda_n)$, con todos los λ_i's $\in \mathbb{R}$ distintos, entonces f es diagonalizable. Es decir, si todos los valores propios de f tienen multiplicidad 1, entonces f es diagonalizable.

Observación. Teniendo en cuenta la identificación que hay entre endomorfismos y matrices cuadradas se puede enunciar el análogo del teorema y del corolario anterior para matrices cuadradas $A \in \mathrm{Mat}_n(\mathbb{R})$.

Corolario 8.13. Sea $A \in \mathrm{Mat}_n(\mathbb{R})$ una matriz cuadrada de orden n. Si A tiene n valores propios distintos (y reales), entonces A es diagonalizable.

El Teorema 8.11 no solo indica cuando el endomorfismo es diagonalizable, sino cómo construir la base en la que la matriz coordenada es diagonal D, y por tanto P. Se propone un esquema de trabajo:

1. Dado el endomorfismo f calcular la matriz coordenada A (se aconseja en base canónica) y hallar los valores propios de A y sus multiplicidades.
2. Calcular la dimensión de los subespacios fundamentales y comprobar si cada una coincide con la multiplicidad del correspondiente valor propio. En caso afirmativo seguir, en caso negativo parar y decir "A (o el endomorfismo f) no es diagonalizable", ver Sección 8.5.
3. Si A es diagonalizable, construir la matriz diagonal D que tiene como entradas en la diagonal los valores propios de A repetidos según su multiplicidad.
4. Calcular bases de los subespacios fundamentales $S(A, \lambda_i)$ y a partir de ellas obtener bases de los subespacios $S(f, \lambda_i)$ juntándolas (en el orden adecuado a como se han colocado los valores propios en la diagonal) para obtener una base de vectores propios de f en la que la matriz coordenada es D.

5. Hallar la matriz P de cambio de coordenadas, entre la base de vectores propios calculada y la base inicial, tal que $P^{-1}AP = D$.

✓ **Ejemplo 5.** Sea A la matriz coordenada de un endomorfismo $f : \mathbb{R}^3 \to \mathbb{R}^3$ en la base canónica:

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

¿Es f diagonalizable? En caso afirmativo, hallar la matriz coordenada diagonal D y una base $\mathcal{D}$ en la cual D es la matriz coordenada del endomorfismo. Hallar también una matriz P regular tal que $P^{-1}AP = D$.

Solución: Observando que se trata del endomorfismo de los ejemplos anteriores, se aprovechan los cálculos ya realizados.

1. En primer lugar se hallan los valores propios de A y sus multiplicidades,

$$\begin{cases} \lambda_1 = -1, & m_1 = 1 \\ \lambda_2 = 2, & m_2 = 2. \end{cases}$$

2. Se calcula a continuación la dimensión de cada subespacio fundamental, comprobando si esta dimensión coincide con la multiplicidad del valor propio correspondiente.

   - $\dim S(A, -1) = 1$ ya que la multiplicidad de $\lambda_1 = -1$, m_1, es igual a 1.
   - $\dim S(A, 2) = 3 - \mathrm{rg}(A - 2I) = 3 - 1 = 2 = m_2$.

   Tras comprobar la igualdad de las dimensiones con las multiplicidades respectivas, se deduce que A es diagonalizable, y por tanto lo es f.

3. Se construye la matriz diagonal D, que tiene en la diagonal los valores propios de A:

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

4. Para hallar la base $\mathcal{D}$ de vectores propios en la cual D es la matriz coordenada del endomorfismo, se calculan bases de cada uno de los subespacios fundamentales.

   - $\mathcal{B}_{S(A,-1)} = \{(1, 1, 1)\}$.
   - $\mathcal{B}_{S(A,2)} = \{(1, 0 - 1), (0, 1, -1)\}$.

   Por trabajar en $\mathbb{R}^3$ con base canónica se tiene que $S(A, \lambda_i) = S(f, \lambda_i)$. Juntando las bases de los subespacios fundamentales (en el orden indicado), se obtiene por tanto una base de vectores propios:

$$\mathcal{D} = \{(1, 1, 1), (1, 0, -1), (0, 1, -1)\}.$$

5. El diagrama conmutativo que representa el cambio de coordenadas es:

$$\begin{tikzcd}
(\mathcal{C}) & \mathbb{R}^3 \arrow[r, "A"] & \mathbb{R}^3 & (\mathcal{C}) \\
(\mathcal{D}) & \mathbb{R}^3 \arrow[u, "P"] \arrow[r, "D"'] \arrow[ur, phantom, "\#"] & \mathbb{R}^3 \arrow[u, "P"'] & (\mathcal{D})
\end{tikzcd}$$

donde $\mathcal{C}$ es la base canónica de $\mathbb{R}^3$. La matriz P regular tal que $P^{-1}AP = D$ es la matriz del cambio de coordenadas entre las bases $\mathcal{D}$ y $\mathcal{C}$. En este ejemplo concreto por ser la base inicial la canónica, para hallar P basta colocar los vectores de la nueva base $\mathcal{D}$ en columnas,

$$P = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & -1 \end{pmatrix} \quad \Longrightarrow \quad P^{-1} = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 2/3 & -1/3 & -1/3 \\ -1/3 & 2/3 & -1/3 \end{pmatrix}.$$

Para cerciorarse de que el resultado es correcto se comprueba que $P^{-1}AP = D$.

◇

Observación. Notar que la matriz P del cambio de coordenadas no es única, ya que las bases de cada subespacio fundamental $S(A, \lambda_i)$ no lo son. Sin embargo la matriz D, si existe, es única salvo orden de los valores propios.

✓ **Ejemplo 6.** Sea A la matriz coordenada de un endomorfismo $f : \mathbb{R}_2[x] \to \mathbb{R}_2[x]$ en la base canónica $\mathcal{B} = \{1, x, x^2\}$:

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

¿Es f diagonalizable? En caso afirmativo, hallar la matriz coordenada diagonal D y una base $\mathcal{D}$ en la cual D es la matriz coordenada del endomorfismo. Hallar también una matriz P regular tal que $P^{-1}AP = D$.

Solución: Observar que la matriz A coincide con la del Ejemplo 5 por lo tanto son válidos todos los cálculos anteriores. En particular A y f son diagonalizables. La única diferencia con dicho ejemplo es la construcción de la base $\mathcal{D}$ de vectores propios de f pertenecientes a $\mathbb{R}_2[x]$.

Para hallar la base $\mathcal{D}$ de vectores propios en la cual D es la matriz coordenada del endomorfismo f, se reconstruyen los subespacios fundamentales $S(f, \lambda_i)$ a partir de los $S(A, \lambda_i)$.

- $\mathcal{B}_{S(A,-1)} = \{(1,1,1)\} \quad \rightsquigarrow \quad \mathcal{B}_{S(f,-1)} = \{1 + x + x^2\}$.
- $\mathcal{B}_{S(A,2)} = \{(1, 0-1), (0,1,-1)\} \quad \rightsquigarrow \quad \mathcal{B}_{S(f,2)} = \{1 - x^2, x - x^2\}$.

Juntando las bases de los subespacios fundamentales (en el orden adecuado), se obtiene por tanto una base de vectores propios:

$$\mathcal{D} = \{1 + x + x^2, 1 - x^2, x - x^2\}.$$

◇

8.4. DIAGONALIZACIÓN ORTOGONAL

En este apartado se trata de resolver el mismo problema de la diagonalización propuesto anteriormente pero añadiendo alguna condición de ortogonalidad sobre la nueva base.

Sea V un espacio vectorial euclídeo de dimensión n, es decir, un espacio vectorial real en el que se ha definido un producto escalar. Dado un endomorfismo $f : V \to V$, el objetivo ahora es, cuando sea posible, encontrar una base ortonormal de V respecto de la cual la matriz coordenada de f sea diagonal. En caso de existir dicha matriz y la base ortonormal de V, el endomorfismo f se dice ortogonalmente diagonalizable.

Definición 8.14. Sea V un espacio vectorial euclídeo de dimensión finita n y sea $f : V \to V$ un endomorfismo. Se dice que f es **ortogonalmente diagonalizable** si existe una base ortonormal de V en la que la matriz coordenada de f es diagonal.

Definición 8.15. Sea $A \in \mathrm{Mat}_n(\mathbb{R})$, se dice que A es **ortogonalmente diagonalizable** si existe una matriz $P \in \mathrm{Mat}_n(\mathbb{R})$ ortogonal ($P^{-1} = P^t$) tal que $P^{-1}AP = D$ (o equivalentemente $P^tAP = D$), siendo $D \in \mathrm{Mat}_n(\mathbb{R})$ una matriz diagonal.

Observación. Los conceptos anteriores se relacionan como en el caso de la diagonalización. Considerando A la matriz coordenada del endomorfismo $f : V \to V$ en una base *ortonormal* de V,

$$\boxed{f \text{ es ortogonalmente diagonalizable} \iff A \text{ es ortogonalmente diagonalizable}}$$

Para poder proponer un teorema que indique cuándo un endomorfismo f, o una matriz $A \in \mathrm{Mat}_n(\mathbb{R})$, es ortogonalmente diagonalizable se necesitan algunos conceptos previos.

Definición 8.16. Sea $(V, \langle\cdot,\cdot\rangle)$ un espacio vectorial euclídeo. Un endomorfismo $f : V \to V$ se dice **simétrico** si verifica

$$\langle f(\mathbf{u}), \mathbf{v}\rangle = \langle \mathbf{u}, f(\mathbf{v})\rangle \qquad \forall \mathbf{u}, \mathbf{v} \in V.$$

✓ **Ejemplo 7.** Sea $f : \mathbb{R}^2 \to \mathbb{R}^2$ un endomorfismo dado por:

$$f(x, y) = (x + y, x - y).$$

Se considera el producto escalar usual de $\mathbb{R}^2$. Se comprueba que:

- $\langle f(x, y), (x', y')\rangle = \langle (x + y, x - y), (x', y')\rangle = xx' + yx' + xy' - yy'.$

- $\langle (x,y), f(x',y')\rangle = \langle (x,y), (x'+y', x'-y')\rangle = xx' + xy' + yx' - yy'$.

Por tanto $\langle f(x,y), (x',y')\rangle = \langle (x,y), f(x',y')\rangle$, y f es un endomorfismo simétrico.

◇

Comprobar si un endomorfismo es o no simétrico con la definición anterior en general no resulta cómodo. La siguiente proposición facilita dicha comprobación.

Proposición 8.17. Sea V un espacio vectorial euclídeo, $f : V \to V$ un endomorfismo y A la matriz coordenada de f en una base ortonormal de V. Entonces f es un endomorfismo simétrico si y solo si A es una matriz simétrica.

Los valores propios de los endomorfismos simétricos verifican:

Proposición 8.18. Sea V un espacio vectorial euclídeo y $f : V \to V$ un endomorfismo simétrico. Entonces todas las raíces del polinomio característico de f son reales. Dicho de otro modo, todas las raíces del polinomio característico son valores propios de f.

El resultado que se propone a continuación proporciona la respuesta a cuándo un endomorfismo es ortogonalmente diagonalizable.

Teorema 8.19. Sea V un espacio vectorial euclídeo, $f : V \to V$ un endomorfismo y $\lambda_1, \dots, \lambda_r \in \mathbb{R}$ las distintas raíces del polinomio característico de f con multiplicidades respectivas $m_1, \dots, m_r$. Entonces son equivalentes:

1. f es ortogonalmente diagonalizable.
2. f es un endomorfismo simétrico.
3. $V = S(f,\lambda_1) \oplus \cdots \oplus S(f,\lambda_r)$, siendo los subespacios fundamentales $S(f,\lambda_i)$'s ortogonales entre sí.
4. V posee una base ortonormal formada por vectores propios de f.

Puede enunciarse un teorema análogo al anterior redactado en términos de matrices $A \in \mathrm{Mat}_n(\mathbb{R})$.

Teorema 8.20. Sea $A \in \mathrm{Mat}_n(\mathbb{R})$ una matriz cuadrada real, $\lambda_1, \dots, \lambda_r \in \mathbb{R}$ las distintas raíces del polinomio característico de A con multiplicidades respectivas $m_1, \dots, m_r$. Entonces son equivalentes:

1. A es ortogonalmente diagonalizable.
2. A es una matriz simétrica.
3. $\mathbb{R}^n = S(A,\lambda_1) \oplus \cdots \oplus S(A,\lambda_r)$, siendo los subespacios fundamentales $S(A,\lambda_i)$'s ortogonales entre sí.
4. $\mathbb{R}^n$ posee una base ortonormal formada por vectores propios de A.

Observación. Recordar que la matriz del cambio de coordenadas entre dos bases ortonormales es una matriz ortogonal, es decir $PP^t = P^tP = I_n$, (ver Proposición 6.19).

Para resolver un problema de diagonalización ortogonal se comienza estudiando si el endomorfismo f es simétrico. Para ello se calcula una matriz coordenada de f en una base ortonormal y se comprueba si es simétrica. Para obtener una base ortonormal en la que la matriz coordenada de f sea diagonal basta aplicar el método de Gram-Schmidt a una base de cada subespacio fundamental $S(A, \lambda_i)$ normalizando los resultados obtenidos y reconstruyendo los subesapcios $S(f, \lambda_i)$. Juntando las bases ortonormales que se acaban de construir se obtiene la base buscada.

Con el siguiente ejemplo se propone un esquema de trabajo con los pasos que se pueden seguir para resolver este tipo de problemas.

✓ **Ejemplo 8.** Sea el espacio euclídeo $\mathbb{R}^3$ y sea A la matriz coordenada de un endomorfismo $f : \mathbb{R}^3 \to \mathbb{R}^3$, en la base canónica:

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

¿Es f ortogonalmente diagonalizable? En caso afirmativo, hallar la matriz coordenada diagonal D y una base ortonormal $\mathcal{D}$ en la cual D es la matriz coordenada del endomorfismo. Hallar también una matriz P ortogonal tal que $P^tAP = D$.

Solución: Observando que se trata del endomorfismo de los ejemplos anteriores, se aprovechan los cálculos ya realizados.

1. En primer lugar se comprueba que A es simétrica y además está dada en la base canónica que es ortonormal con respecto al producto escalar usual, por tanto A y f son ortogonalmente diagonalizables.

2. Se calculan a continuación los valores propios de A y sus multiplicidades,

$$\begin{cases} \lambda_1 = -1, & m_1 = 1 \\ \lambda_2 = 2, & m_2 = 2. \end{cases}$$

   Se deduce que A es semejante a la matriz

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

3. Para hallar la base ortonormal $\mathcal{D}$ de vectores propios de f en la cual D es la matriz coordenada del endomorfismo, se calculan bases de cada uno de los subespacios fundamentales.

- $\mathcal{B}_{S(A,-1)} = \{(1,1,1)\}$.
- $\mathcal{B}_{S(A,2)} = \{(1,0,-1),(0,1,-1)\}$.

Se aplica el método de Gram-Schmidt y la normalización a cada una de ellas.

- Base ortonormal de $S(A,-1) = \{(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})\}$.
- Base ortonormal de $S(A,2) = \{(\frac{1}{\sqrt{2}}, 0, \frac{-1}{\sqrt{2}}), (\frac{-1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, \frac{-1}{\sqrt{6}})\}$.

Por trabajar en $\mathbb{R}^3$ con base canónica se tiene que $S(A,\lambda_i) = S(f,\lambda_i)$. Juntando las bases ortonormales de los subespacios fundamentales (en el orden indicado), se obtiene una base ortonormal de vectores propios:

$$\mathcal{D} = \left\{ \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right), \left(\frac{1}{\sqrt{2}}, 0, \frac{-1}{\sqrt{2}}\right), \left(\frac{-1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, \frac{-1}{\sqrt{6}}\right) \right\}.$$

4. El diagrama conmutativo que representa el cambio de coordenadas es:

$$\begin{tikzcd}
(\mathcal{C}) & \mathbb{R}^3 \arrow[r, "A"] & \mathbb{R}^3 & (\mathcal{C}) \\
(\mathcal{D}) & \mathbb{R}^3 \arrow[u, "P"] \arrow[r, "D"'] \arrow[ur, phantom, "\#"] & \mathbb{R}^3 \arrow[u, "P"'] & (\mathcal{D})
\end{tikzcd}$$

La matriz P regular tal que $P^tAP = D$ es la matriz del cambio de coordenadas entre las bases $\mathcal{D}$ de vectores propios y $\mathcal{C}$ la base canónica de $\mathbb{R}^3$, como se aprecia en el diagrama anterior. En este ejemplo concreto por ser la base inicial la canónica, para hallar P basta colocar los vectores de la nueva base $\mathcal{D}$ en columnas.

$$P = \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{3} & 0 & 2/\sqrt{6} \\ 1/\sqrt{3} & -1/\sqrt{2} & -1/\sqrt{6} \end{pmatrix}.$$

5. Para tener la certeza de que el resultado es correcto basta comprobar que $P^tAP = D$.

◇

$\rightsquigarrow$ **Ejercicio.** Comprobar que $S(A,-1) = S(A,2)^{\perp}$.

8.5. FORMA CANÓNICA DE JORDAN

El objetivo de este apartado es encontrar, para aquellos endomorfismos no diagonalizables, una base en la que la matriz coordenada sea “lo más sencilla posible”, es decir, lo más parecida a una matriz diagonal. Más concretamente una matriz diagonal por bloques donde en la diagonal aparecen los valores propios repetidos según su multiplicidad.

A lo largo de toda la sección se consideran endomorfismos $f : V \to V$ donde V es un espacio vectorial real de dimensión finita. Como además se está considerando el caso en el que todas las raíces del polinomio característico de f son reales, siempre existe una matriz coordenada de f de la forma que se presenta a continuación.

Se comienza proponiendo un ejemplo sencillo antes de plantear el problema general.

✓ **Ejemplo 9.** Se considera el endomorfismo $f : \mathbb{R}^2 \to \mathbb{R}^2$ dado por

$$f(x, y) = (2y, -2x + 4y).$$

La matriz coordenada en base canónica es

$$A = \begin{pmatrix} 0 & 2 \\ -2 & 4 \end{pmatrix}.$$

Se calculan los valores propios de A, $|A - \lambda I_2| = 0$, y se obtiene $\lambda = 2$ con multiplicidad $m = 2$. Para ver si A es diagonalizable se debe comparar el valor de m con la dimensión del subespacio fundamental $S(A, 2)$.

Recordando que $\dim S(A, 2) = 2 - \text{rg}(A - 2I)$,

$$A - 2I = \begin{pmatrix} -2 & 2 \\ -2 & 2 \end{pmatrix} \quad \Rightarrow \quad \text{rg}(A - 2I) = 1,$$

de modo que

$$\dim S(A, 2) = 2 - \text{rg}(A - 2I) = 2 - 1 = 1 \neq m.$$

Se observa que la dimensión del subespacio fundamental no coincide con la multiplicidad, es decir, A (o f) no es diagonalizable. Se plantea por tanto hallar una matriz lo más parecida a una matriz diagonal que sea semejante a la matriz A. En este caso se puede adelantar que la matriz buscada es de la forma

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

La base $\mathcal{D} = \{\mathbf{v}_1, \mathbf{v}_2\}$ en la que la matriz coordenada de f es J, debe verificar:

$$f(\mathbf{v}_1) = 2\mathbf{v}_1; \qquad f(\mathbf{v}_2) = \mathbf{v}_1 + 2\mathbf{v}_2.$$

De lo anterior se observa que el vector $\mathbf{v}_1$ es un vector propio de valor propio 2, es decir, $\mathbf{v}_1 \in S(A, 2)$.

$$S(A, 2) = \text{Ker}(A - 2I) = \text{Ker} \begin{pmatrix} -2 & 2 \\ -2 & 2 \end{pmatrix} = \{(x, y) \in \mathbb{R}^2 \mid x = y\} = \{(x, x) \mid x \in \mathbb{R}\}.$$

Tomando, por ejemplo, el vector $\mathbf{v}_1 = (1, 1)$, se observa que $\mathbf{v}_2$ verifica $f(\mathbf{v}_2) - 2\mathbf{v}_2 = \mathbf{v}_1$, es decir, es solución del sistema $(A - 2I)\mathbf{v}_2 = \mathbf{v}_1$:

$$\begin{pmatrix} -2 & 2 \\ -2 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Del conjunto de soluciones $\{(x, x + {}^1\!/_2) \mid x \in \mathbb{R}\}$ se elige por ejemplo

$$\mathbf{v}_2 = (-1, -{}^1\!/_2).$$

En la base $\{\mathbf{v}_1, \mathbf{v}_2\} = \{(1,1), (-1, -{}^1\!/_2)\}$ la matriz coordenada de f es J. La base anterior se denomina *base de cadenas*. La matriz P de cambio es

$$P = \begin{pmatrix} 1 & -1 \\ 1 & -{}^1\!/_2 \end{pmatrix},$$

que verifica $P^{-1}AP = J$ (comprobarlo como ejercicio), siendo J la matriz que se denomina *forma canónica de Jordan* o *matriz de Jordan.*

◇

En el ejemplo anterior se ha dado por conocida la matriz de Jordan J, y a partir de ella se ha calculado la base en la que J es matriz coordenada. El objetivo ahora es, dado un endomorfismo f o equivalentemente una matriz $A \in \mathrm{Mat}_n(\mathbb{R})$, determinar la existencia de su forma canónica de Jordan J asociada, calcularla y encontrar la nueva base en la cual J es matriz coordenada.

8.5.1. MATRIZ DE JORDAN

Se comienza por determinar la forma que deben tener la matrices "más sencillas posibles" que ocupan el lugar de las matrices diagonales.

Definición 8.21. Un **bloque de Jordan** de orden n es una matriz cuadrada, triangular superior de tamaño $n \times n$, cuyos elementos diagonales son iguales y los elementos que están en la línea por encima de la diagonal valen 1, es decir, $a_{ii} = a$, $\forall i = 1, \ldots, n$, y $a_{i,i+1} = 1$, $\forall i = 1, \ldots, n-1$.

✓ **Ejemplo 10.** Los bloques de Jordan de orden 1, 2 y 3 son, respectivamente, de la siguiente forma:

$$J_1 = (a); \quad J_2 = \begin{pmatrix} a & 1 \\ 0 & a \end{pmatrix}; \quad J_3 = \begin{pmatrix} a & 1 & 0 \\ 0 & a & 1 \\ 0 & 0 & a \end{pmatrix}, \qquad a \in \mathbb{R}.$$

Definición 8.22. Una **matriz de Jordan** J es una matriz diagonal por bloques, donde cada bloque en la diagonal es un bloque de Jordan.

✓ **Ejemplo 11.** La siguiente matriz es una matriz de Jordan:

$$J = \left(\begin{array}{c|c|ccc|cc} 3 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 4 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 \end{array}\right)$$

formada por 2 bloques de orden 1, un bloque de orden 3 y un bloque de orden 2.

Definición 8.23. Sea $f : V \to V$ un endomorfismo y $A \in \mathrm{Mat}_n(\mathbb{R})$ la matriz coordenada de f en una cierta base $\mathcal{B}$ del espacio vectorial real V de dimensión finita n. Si existe una matriz de Jordan J que sea semejante con A, se dice que J es su **forma canónica de Jordan**.

Notas.

- Observar que una matriz diagonal es una matriz de Jordan cuyos bloques son todos de orden 1. Así, las formas canónicas de Jordan de los endomorfismos diagonalizables coinciden con las matrices diagonales estudiadas en las secciones anteriores.
- Para cualquier endomorfismo, entre sus matrices coordenadas no puede haber dos matrices de Jordan distintas, salvo las obtenidas reordenando los bloques, es decir, la forma canónica de Jordan es única.

8.5.2. SUBESPACIOS FUNDAMENTALES GENERALIZADOS

Dado un endomorfismo f, o una matriz cuadrada $A \in \mathrm{Mat}_n(\mathbb{R})$, no diagonalizable, para abordar el problema de determinar su forma de Jordan asociada, así como la nueva base es necesario introducir los subespacios vectoriales que se definen a continuación.

Definición 8.24. Sea V un espacio vectorial real de dimensión finita n. Sea $f : V \to V$ un endomorfismo, $A \in \mathrm{Mat}_n(\mathbb{R})$ la matriz coordenada de f en una cierta base $\mathcal{B}$ de V y $\lambda \in \mathbb{R}$ un valor propio de f. Los **subespacios fundamentales generalizados** asociados a dicho valor propio λ vienen dados por:

$$S^j(f,\lambda) = \{\mathbf{v} \in V \mid (f - \lambda Id)^j(\mathbf{v}) = \mathbf{0}\} = \mathrm{Ker}[(f - \lambda Id)^j],$$

que son los formados por los vectores cuyas coordenadas respecto de $\mathcal{B}$ pertenecen a:

$$S^j(A,\lambda) = \{\mathbf{x} \in \mathbb{R}^n \mid (A - \lambda I_n)^j \mathbf{x} = \mathbf{0}\} = \mathrm{Ker}[(A - \lambda I_n)^j].$$

Los subespacios que se acaban de definir verifican las siguientes propiedades.

Propiedades.

- $S^1(f,\lambda)$ es el subespacio fundamental $S(f,\lambda)$, dado en Definición 8.7.
- Cada subespacio fundamental generalizado está contenido en el siguiente, formando la siguiente cadena:

  $$S^1(f,\lambda) \subseteq S^2(f,\lambda) \subseteq \ldots \subseteq S^{k-1}(f,\lambda) \subseteq S^k(f,\lambda) \subseteq \ldots$$

  ya que si $\mathbf{v} \in S^j(f,\lambda)$, $(f-\lambda Id)^j(\mathbf{v}) = \mathbf{0}$, y por tanto

  $$(f-\lambda Id)^{j+1}(\mathbf{v}) = (f-\lambda Id)\left[(f-\lambda Id)^j(\mathbf{v})\right] = (f-\lambda Id)(\mathbf{0}) = \mathbf{0},$$

  de modo que $\mathbf{v} \in S^{j+1}(f,\lambda)$.
- De la propiedad anterior se deduce que si un vector pertenece a $S^{j+1}(f,\lambda)$, su imagen por $f-\lambda Id$ pertenece a $S^j(f,\lambda)$.
- Como V es un espacio de dimensión finita, la cadena anterior debe estabilizarse, es decir, a partir de un índice todos los subespacios coinciden. Si k es el primer valor que verifica $S^k(f,\lambda) = S^{k+1}(f,\lambda)$, entonces $S^k(f,\lambda) = S^j(f,\lambda)\ \forall j \geq k$, es decir, la cadena se estabiliza definitivamente la primera vez que se para.
- Si la multiplicidad del valor propio λ es m, entonces:

  $$1 \leq \dim S^j(f,\lambda) \leq m.$$

Definición 8.25. Sea V un espacio vectorial real de dimensión finita, $f : V \to V$ un endomorfismo y $\lambda \in \mathbb{R}$ un valor propio de f. Sean $S^j(f,\lambda)$ los subespacios fundamentales generalizados asociados a λ. Si k es el primer valor que verifica $S^k(f,\lambda) = S^{k+1}(f,\lambda)$, entonces el subespacio $S^k(f,\lambda)$ se denomina **subespacio máximo** asociado al valor propio λ, y se denota por $S^*(f,\lambda)$.

Nota. De las propiedades anteriores es claro deducir que el subespacio máximo contiene a todos los demás subespacios $S^j(f,\lambda)$.

$$S^1(f,\lambda) \subseteq S^2(f,\lambda) \subseteq \ldots \subseteq S^{k-1}(f,\lambda) \subseteq S^*(f,\lambda).$$

Los subespacios máximos verifican:

Proposición 8.26. Sea V un espacio vectorial real de dimensión finita, $f : V \to V$ un endomorfismo y $\lambda \in \mathbb{R}$ uno de sus valores propios. El subespacio $S^*(f,\lambda)$ es invariante por f, es decir, si $\mathbf{v} \in S^*(f,\lambda)$ entonces $f(\mathbf{v}) \in S^*(f,\lambda)$.

Proposición 8.27. Sea V un espacio vectorial real de dimensión finita y $f : V \to V$ un endomorfismo. Sea λ un valor propio de f de multiplicidad m. Se verifica:

$$\dim S^*(f, \lambda) = m.$$

Conociendo que el subespacio máximo tiene dimensión igual a la multiplicidad del valor propio, los cálculos de dicho subespacio se agilizan, ya que se indica donde se estabiliza la cadena de subespacios fundamentales generalizados.

8.5.3. BASE DEL SUBESPACIO $S^*(f, \lambda)$

Definidos los subespacios máximos de un endomorfismo $f : V \to V$, se propone un método para encontrar una base de ellos.

Proposición 8.28. Sea V un espacio vectorial real de dimensión finita, $f : V \to V$ un endomorfismo y $\lambda \in \mathbb{R}$ uno de sus valores propios. Existe una base de $S^*(f, \lambda)$ de modo que la matriz coordenada de la restricción $f_{|S^*(f,\lambda)} : S^*(f, \lambda) \to S^*(f, \lambda)$ respecto de dicha base es una matriz de Jordan.

Pueden darse demostraciones de la proposición anterior que proporcionan un método práctico para obtener la base buscada. Siguiendo la tónica general del texto de evitar las demostraciones largas, se propone directamente un esquema para el cálculo de la base de $S^*(f, \lambda)$.

Dado el endomorfismo $f : V \to V$ se considera uno de sus valores propios, $\lambda \in \mathbb{R}$, con multiplicidad m.

1. Se calcula la cadena de subespacios fundamentales generalizados, dando una base y sus dimensiones, hasta llegar al subespacio máximo $S^*(f, \lambda)$.

2. Se construye un esquema siguiendo los siguientes criterios:

   - Se escribe la cadena de subespacios fundamentales generalizados hasta llegar al subespacio máximo y sobre cada uno de ellos su dimensión.

   $$\begin{array}{ccccccccc} \boxed{d_1} & & \boxed{d_2} & & \boxed{d_3} & & & & \boxed{m} \\ S^1(f,\lambda) & \subset & S^2(f,\lambda) & \subset & S^3(f,\lambda) & \subset & \dots & \subset & S^*(f,\lambda) \end{array}$$

   - Debajo de cada subespacio $S^j(f, \lambda)$ se dibujan tantos puntos como indique la diferencia de las siguientes dimensiones:

   $$\dim S^j(f, \lambda) - \dim S^{j-1}(f, \lambda).$$

   En particular, debajo del subespacio $S^1(f, \lambda)$ deben aparecer tantos puntos como indique su dimensión.

- Se unen los puntos de cada fila mediante flechas de derecha a izquierda hasta llegar a $S^1(f,\lambda)$.

  Como ejemplo, se supone un endomorfismo $f : \mathbb{R}^5 \to \mathbb{R}^5$, con valor propio λ de multiplicidad 5, y tal que $S^*(f,\lambda) = S^3(f,\lambda)$. Si las dimensiones de los subespacios fundamentales generalizados son: $\dim S^1(f,\lambda) = 2$, $\dim S^2(f,\lambda) = 4$ y $\dim S^3(f,\lambda) = 5$, el esquema correspondiente es de la forma:

$$\begin{array}{ccccc} \boxed{2} & & \boxed{4} & & \boxed{5} \\ S^1(f,\lambda) & \subset & S^2(f,\lambda) & \subset & S^3(f,\lambda) \\ \bullet & \leftarrow & \bullet & \leftarrow & \bullet \\ \bullet & \leftarrow & \bullet & & \end{array}$$

- Observar que el total de puntos dibujados en el esquema coincide con m, la multiplicidad del valor propio correspondiente.

3. La matriz de Jordan de la restricción de f al subespacio máximo se construye considerando lo siguiente:

   - Cada una de las filas del esquema da lugar a una cadena que corresponde en la forma canónica de Jordan con un bloque de Jordan.
   - El número de bloques de Jordan asociados al valor propio λ viene dado por la dimensión del subespacio fundamental $S^1(f,\lambda) = S(f,\lambda)$, que coincide con el número de filas del esquema.
   - El tamaño de cada bloque de Jordan viene determinado por el número de vectores que aparecen en la fila correspondiente.

     Así en el ejemplo anterior la forma canónica de Jordan asociada a f es:

$$J = \left(\begin{array}{ccc|cc} \lambda & 1 & 0 & 0 & 0 \\ 0 & \lambda & 1 & 0 & 0 \\ 0 & 0 & \lambda & 0 & 0 \\ \hline 0 & 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & 0 & \lambda \end{array}\right).$$

4. Para construir la base en la que la matriz coordenada es J hay que sustituir los puntos del esquema construido en el apartado 2, por vectores adecuados del subespacio correspondiente.

   - Una vez que se tiene el primer vector a la derecha de cada fila, los siguientes se obtienen calculando, por cada flecha, la imagen del vector del que se parte por $f - \lambda Id$.

$$\begin{array}{ccccc} \boxed{2} & & \boxed{4} & & \boxed{5} \\ S^1(f,\lambda) & \subset & S^2(f,\lambda) & \subset & S^3(f,\lambda) \\ \mathbf{v}_1 & \leftarrow & \mathbf{v}_2 & \leftarrow & \mathbf{v}_3 \\ \mathbf{w}_1 & \leftarrow & \mathbf{w}_2 & & \end{array}$$

Cada una de las filas constituye una cadena de vectores. Es decir, una vez elegido $\mathbf{v}_3$ los vectores de la primera fila del esquema se calculan:

$$\mathbf{v}_2 = (f - \lambda Id)(\mathbf{v}_3), \quad \mathbf{v}_1 = (f - \lambda Id)(\mathbf{v}_2).$$

- Los extremos de cada fila, correspondientes a $S^j(f,\lambda)$ se obtienen tomando los (posibles) vectores ya construidos de $S^j(f,\lambda)$ y los de una base cualquiera de $S^{j-1}(f,\lambda)$ (si $j>1$) hasta completar una base de $S^j(f,\lambda)$.

✓ **Ejemplo 12.** Se considera el endomorfismo $f : \mathbb{R}^5 \to \mathbb{R}^5$ cuya matriz coordenada en base canónica es:

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

que tiene un único valor propio $\lambda = 1$ con multiplicidad $m = 5$.

1. Se comienza calculando los subespacios fundamentales generalizados y sus dimensiones correspondientes.

   - $S^1(A,1) = S(A,1) = \mathrm{Ker}(A - I_5) = \mathrm{Ker}\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ -2 & 0 & 0 & 2 & 3 \\ -2 & 0 & 0 & 1 & 3 \\ 3 & 0 & 0 & 0 & -4 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} =$

     $\{(0,x_2,x_3,0,0) \,|\, x_2,x_3 \in \mathbb{R}\} = \langle (0,1,0,0,0),(0,0,1,0,0)\rangle$.

     $\dim S(A,1) = 2$.

   - $S^2(A,1) = \mathrm{Ker}(A - I_5)^2 = \mathrm{Ker}\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 6 & 0 & 0 & 0 & -8 \\ 3 & 0 & 0 & 0 & -4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} =$

     $\{\left(x_1,x_2,x_3,x_4,\frac{3}{4}x_1\right) \,|\, x_1,x_2,x_3,x_4 \in \mathbb{R}\} =$
     $\langle (4,0,0,0,3),(0,1,0,0,0)(0,0,1,0,0),(0,0,0,1,0)\rangle$.

     $\dim S^2(A,1) = 4$.

   - $S^3(A,1) = \mathrm{Ker}(A - I_5)^3 = \mathrm{Ker}\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = \mathbb{R}^5 =$

     $\langle (1,0,0,0,0),(0,1,0,0,0)(0,0,1,0,0),(0,0,0,1,0),(0,0,0,0,1)\rangle$.

$\dim S^3(A,1) = 5$. Como la dimensión del subespacio $S^3(A,1)$ coincide con la multiplicidad del valor propio, ya se ha llegado al subespacio máximo, $S^3(A,1) = S^*(A,1)$.

2. Se construye el esquema con los criterios establecidos anteriormente.

   Como $\dim S(A,1) = 2$, el esquema tiene dos filas y por tanto la matriz de Jordan dos bloques.

$$\begin{array}{ccccc} \boxed{2} & & \boxed{4} & & \boxed{5} \\ S^1(A,1) & \subset & S^2(A,1) & \subset & S^3(A,1) \end{array}$$

   - Como $\dim S^3(A,1) - \dim S^2(A,1) = 5 - 4 = 1$, debajo de $S^3(A,1)$ solo hay un punto.
   - Calculando $\dim S^2(A,1) - \dim S^1(A,1) = 4-2 = 2$, se observa que debajo de $S^2(A,1)$ hay que poner 2 puntos.
   - Todas las cadenas deben terminar en $S^1(A,1)$, luego a este espacio le corresponden 2 puntos.

   Así se obtiene este esquema final:

$$\begin{array}{ccccc} \boxed{2} & & \boxed{4} & & \boxed{5} \\ S^1(A,1) & \subset & S^2(A,1) & \subset & S^3(A,1) \\ \bullet & \leftarrow & \bullet & \leftarrow & \bullet \\ \bullet & \leftarrow & \bullet & & \end{array}$$

3. Se construye la matriz de Jordan. Del esquema anterior se deduce que hay dos cadenas en la base, una de longitud 3 y otra de longitud 2 que corresponden a dos cajas de Jordan, y la forma canónica de Jordan es:

$$J = \left(\begin{array}{ccc|cc} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right).$$

4. Se construye la base sustituyendo los puntos del esquema anterior por los vectores correspondientes.

$$\begin{array}{ccccc} \boxed{2} & & \boxed{4} & & \boxed{5} \\ S^1(A,1) & \subset & S^2(A,1) & \subset & S^3(A,1) \\ \mathbf{v}_1 & \leftarrow & \mathbf{v}_2 & \leftarrow & \mathbf{v}_3 \\ \mathbf{w}_1 & \leftarrow & \mathbf{w}_2 & & \end{array}$$

   Recordando las bases de los subespacios fundamentales generalizados:

- $\mathcal{B}_{S^1(A,1)} = \{(0,1,0,0,0),(0,0,1,0,0)\}$.
- $\mathcal{B}_{S^2(A,1)} = \{(4,0,0,0,3),(0,1,0,0,0),(0,0,1,0,0),(0,0,0,1,0)\}$.
- $\mathcal{B}_{S^3(A,1)} = \{(1,0,0,0,0),(0,1,0,0,0),(0,0,1,0,0),(0,0,0,1,0),$
  $(0,0,0,0,1)\}$.

El vector $\mathbf{v}_3$ se elige considerando que debe estar en $S^3(A,1)$ pero no en $S^2(A,1)$, observando las bases anteriores basta tomar $\mathbf{v}_3 = (1,0,0,0,0)$.

Para calcular $\mathbf{v}_2$ hay que realizar $\mathbf{v}_2 = (A - I_5)\mathbf{v}_3$, de lo que se deduce que $\mathbf{v}_2 = (0,-2,-2,3,0)$. Para calcular $\mathbf{v}_1$ hay que realizar $\mathbf{v}_1 = (A - I_5)\mathbf{v}_2$, de lo que se deduce que $\mathbf{v}_1 = (0,6,3,0,0)$.

De este modo se ha completado la primera cadena $\{\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3\}$.

Análogamente para construir la segunda cadena se elige un $\mathbf{w}_2$ de $S^2(A,1)$ con la condición $\mathbf{w}_2 \notin S^1(A,1) + \langle \mathbf{v}_2 \rangle$. De este modo $\mathbf{w}_2$, $\mathbf{v}_2$ y una base cualquiera de $S^1(A,1)$ dan lugar a una base de $S^2(A,1)$. En este caso se puede tomar $\mathbf{w}_2 = (4,0,0,0,3)$.

Ahora se calcula $\mathbf{w}_1 = (A - I_5)\,\mathbf{w}_2 = (0,1,1,0,0)$, y se tiene la base completa $\{\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3,\mathbf{w}_1,\mathbf{w}_2\}$ de $S^3(A,1)$.

◇

8.5.4. EXISTENCIA Y CÁLCULO DE LA FORMA DE JORDAN

En este apartado solo queda recopilar todo lo hecho hasta el momento para calcular la forma canónica de Jordan de una matriz cualquiera $A \in \mathrm{Mat}_n(\mathbb{R})$. Dado un espacio vectorial real V de dimensión finita y un endomorfismo $f : V \to V$, se comienza mostrando un teorema de existencia de dicha forma canónica y para construir la base se propone una descomposición del espacio V, de manera análoga a como se hizo en el caso de la diagonalización con los subespacios fundamentales.

Teorema 8.29. Sea V un espacio vectorial real con $\dim V = n$ y $f : V \to V$ un endomorfismo con polinomio característico $p(t) = (t-\lambda_1)^{m_1} \dots (t-\lambda_r)^{m_r}$, cuyas raíces son todas reales. Entonces se verifica:

- $\displaystyle\sum_{i=1}^{r} m_i = n$.
- $V = S^*(f,\lambda_1) \oplus \cdots \oplus S^*(f,\lambda_r)$.
- Existe una matriz de Jordan J asociada al endomorfismo f.

Corolario 8.30. Si $A \in \mathrm{Mat}_n(\mathbb{R})$ es una matriz cuadrada real y las raíces de su polinomio característico son todas reales, entonces A es semejante a una matriz de Jordan J. (Observar que por ser semejantes A y J tienen el mismo rango).

Con los resultados anteriores ya se está en disposición de encontrar una base en la que la matriz coordenada sea la forma canónica de Jordan. Basta encontrar bases de cada $S^*(f, \lambda_i)$ y juntarlas.

Sea V un espacio vectorial de dimensión finita y sea $f : V \to V$ un endomorfismo con polinomio característico $(t-\lambda_1)^{m_1} \dots (t-\lambda_r)^{m_r}$, siendo $\lambda_i \in \mathbb{R}$, $\forall i = 1, \dots, r$. Sea A la matriz coordenada de f en una cierta base $\mathcal{B}$.

Se propone el siguiente esquema para hallar la forma canónica de Jordan y la base de cadenas correspondiente.

1. Se calculan los valores propios de A, con sus multiplicidades correspondientes.

2. Para cada valor propio λ_i se calcula la cadena de subespacios fundamentales generalizados hasta llegar al subespacio máximo $S^*(A, \lambda_i)$. Se aplica el método expuesto en la Sección 8.5.3 para hallar una base de cadenas de cada subespacio $S^*(A, \lambda_i)$. A partir de las bases anteriores se reconstruyen las bases de los subespacios $S^*(f, \lambda_i)$.

3. Juntando las bases de cada $S^*(f, \lambda_i)$ se obtiene una base del espacio V en la que la matriz coordenada de f es la forma canónica de Jordan de f.

4. La matriz de Jordan se construye con los bloques de Jordan que corresponden a cada cadena, según se ha indicado en la Sección 8.5.3.

5. Para hallar la matriz regular P tal que $P^{-1}AP = J$, basta hallar el cambio de coordenadas entre la nueva base de cadenas obtenida y la base inicial $\mathcal{B}$. En el caso en el que la base inicial sea la base canónica, P se forma poniendo por columnas los vectores de la base de cadenas de A debidamente ordenados.

✓ **Ejemplo 13.** Sea el endomorfismo $f : \mathbb{R}^5 \to \mathbb{R}^5$ dado por:

$$f(x_1, x_2, x_3, x_4, x_5) = (x_1 + 2x_2 + 3x_3, x_2 + 2x_3, x_3 + 2x_4, 2x_4 + x_5, x_5).$$

Calcular la forma canónica de Jordan J del endomorfismo. Hallar una base en la que la matriz coordenada de f sea J. Comprobar que existe una matriz regular P tal que $P^{-1}AP = J$, siendo A la matriz coordenada de f en base canónica.

Solución: La matriz coordenada de f en base canónica es:

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

Aplicando el método anterior:

1. Se calculan los valores propios de A con sus multiplicidades:

   $$\begin{cases} \lambda_1 = 1, & m_1 = 4 \\ \lambda_2 = 2, & m_2 = 1. \end{cases}$$

   Por tanto

   $$\mathbb{R}^5 = S^*(A,1) \oplus S^*(A,2).$$

2. Se calculan los subespacios fundamentales generalizados asociados a cada valor propio. Se aplica el esquema propuesto anteriormente para hallar una base de cada uno de los subespacios máximos.

   $\boxed{\lambda_1 = 1}$

   - $S^1(A,1) = S(A,1) = \operatorname{Ker}(A - I_5) = \{(x_1, 0, 0, 0, 0) \mid x_1 \in \mathbb{R}\}$;
     $\dim S(A,1) = 1$.
   - $S^2(A,1) = \operatorname{Ker}(A - I_5)^2 = \{(x_1, x_2, 0, 0, 0) \mid x_1, x_2 \in \mathbb{R}\}$;
     $\dim S^2(A,1) = 2$.
   - $S^3(A,1) = \operatorname{Ker}(A - I_5)^3 = \{(x_1, x_2, x_3, 0, 0) \mid x_1, x_2, x_3 \in \mathbb{R}\}$;
     $\dim S^3(A,1) = 3$.
   - $S^4(A,1) = \operatorname{Ker}(A - I_5)^4 = \{(x_1, x_2, x_3, x_4, -x_4) \mid x_1, x_2, x_3, x_4 \in \mathbb{R}\}$;
     $\dim S^4(A,1) = 4 = m_1$, por tanto, $S^4(A,1) = S^*(A,1)$ es el subespacio máximo y se para la cadena de subespacios.

   Se plantea el esquema

   $$\begin{array}{ccccccc} \boxed{1} & & \boxed{2} & & \boxed{3} & & \boxed{4} \\ S(A,1) & \subset & S^2(A,1) & \subset & S^3(A,1) & \subset & S^4(A,1) \\ \bullet\mathbf{v_1} & \leftarrow & \bullet\mathbf{v}_2 & \leftarrow & \bullet\mathbf{v}_3 & \leftarrow & \bullet\mathbf{v}_4 \end{array}$$

   Para construir la base se elige un vector $\mathbf{v}_4 \in S^4(A,1)$, pero $\mathbf{v}_4 \notin S^3(A,1)$,

   $$\mathbf{v}_4 = (0, 0, 0, 1, -1).$$

   El resto de vectores se obtienen de forma inmediata:

   $$\begin{aligned} \mathbf{v}_3 &= (A - I_5)\mathbf{v}_4 = (0, 0, 2, 0, 0), \\ \mathbf{v}_2 &= (A - I_5)\mathbf{v}_3 = (6, 4, 0, 0, 0), \\ \mathbf{v}_1 &= (A - I_5)\mathbf{v}_2 = (8, 0, 0, 0, 0). \end{aligned}$$

   En este caso $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ es una base de $S^*(A,1)$ formada por una cadena de longitud 4.

$\boxed{\lambda_2 = 2}$

- $S^1(A,2) = \operatorname{Ker}(A - 2I_5) = \{(14x_4, 4x_4, 2x_4, x_4, 0) \mid x_4 \in \mathbb{R}\}$.
  Como $\dim S^1(A,2) = 1 = m_2$, entonces $S^1(A,2) = S^*(A,2)$, por tanto, basta elegir un vector de este espacio,

$$\mathbf{w}_1 = (14, 4, 2, 1, 0).$$

  $\{\mathbf{w}_1\}$ es una base de $S^*(A,2)$ formada por una cadena de longitud 1.

3. Se juntan las bases de los subespacios $S^*(A,1)$ y $S^*(A,2)$, de modo que

$$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4, \mathbf{w}_1\}$$

   es una base de cadenas de $\mathbb{R}^5$. En esta base la matriz coordenada del endomorfismo es J.

4. Se construye la forma canónica de Jordan J y la matriz regular P tal que $P^{-1}AP = J$ como sigue:

$$J = \left(\begin{array}{cccc|c} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 0 & 2 \end{array}\right), \qquad P = \begin{pmatrix} 8 & 6 & 0 & 0 & 14 \\ 0 & 4 & 0 & 0 & 4 \\ 0 & 0 & 2 & 0 & 2 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & -1 & 0 \end{pmatrix}.$$

◇

8.6. APLICACIONES

8.6.1. MODELOS DE CRECIMIENTO

Las matrices de transición se utilizan para modelizar el crecimiento de una población, y permiten calcular el vector de estado en un momento conociendo dicho vector en el periodo anterior, es decir,

$$\mathbf{v}_{i+1} = A\mathbf{v}_i.$$

Cuando se usan modelos de crecimiento estables es necesario que el vector de estado en el momento $(i+1)$-ésimo sea múltiplo escalar del i-ésimo vector, es decir, $\mathbf{v}_{i+1} = \lambda\mathbf{v}_i$. Aplicando esto a la expresión que modeliza el crecimiento

$$A\mathbf{v}_i = \mathbf{v}_{i+1} = \lambda\mathbf{v}_i,$$

se observa que en este caso los vectores de estado son vectores propios de A, siendo la constante de proporcionalidad λ el valor propio asociado.

8.6.2. CÁLCULO DE POTENCIAS DE UNA MATRIZ

En el Capítulo 1 se propuso un método inductivo para el cálculo de la potencia n-ésima de una matriz $A \in \mathrm{Mat}_n(\mathbb{R})$: A partir del cálculo de las potencias sucesivas A^2, A^3, A^4, A^5 ... se plantean $n \times n$ sucesiones con los elementos de dichas matrices que ocupan la misma posición $a_{ij}(n)$, calculando su término general. Para expresar A^n basta colocar los términos generales calculados en la posición que corresponde.

En ocasiones, el método anterior no resulta útil, sobre todo si el tamaño de A es grande.

Si A es una matriz diagonalizable, el cálculo de A^n se puede simplificar. Basta hallar la matriz diagonal $D \in \mathrm{Mat}_n(\mathbb{R})$ y la matriz regular P tales que $D = P^{-1}AP$ o bien $A = PDP^{-1}$, y aplicar la Proposición 1.6 para el cálculo de la potencia n-ésima de una matriz diagonal.

$$\begin{aligned}
&A = PDP^{-1}\\
&A^2 = AA = (PDP^{-1})(PDP^{-1}) = PDI_nDP^{-1} = PD^2P^{-1}\\
&A^3 = A^2A = (PD^2P^{-1})(PDP^{-1}) = PD^2I_nDP^{-1} = PD^3P^{-1}\\
&\vdots\\
&A^n = A^{n-1}A = (PD^{n-1}P^{-1})(PDP^{-1}) = PD^{n-1}I_nDP^{-1} = PD^nP^{-1}.
\end{aligned}$$

En caso de que $A \in \mathrm{Mat}_n(\mathbb{R})$ no sea diagonalizable se puede seguir un proceso similar calculando su forma canónica de Jordan, $J = P^{-1}AP$, de modo que

$$\boxed{A^n = PJ^nP^{-1}}$$

En este caso el cálculo de J^n no es tan sencillo como el de D^n, pero se puede considerar J como una matriz diagonal por bloques y calcular las potencias de los bloques de Jordan.

Estos métodos se suelen aplicar para la resolución de ecuaciones y sistemas de ecuaciones diferenciales, problemas de Markov, cálculo de términos generales de sucesiones recurrentes, etc.

8.6.3. SISTEMAS DE ECUACIONES DIFERENCIALES

La diagonalización y la forma canónica de Jordan pueden ayudar a resolver sistemas de ecuaciones diferenciales como se ve en el ejemplo que sigue.

✓ **Ejemplo 14.** Encontrar tres funciones derivables f, g, h verificando

$$\begin{cases} f'(t) = f(t) - g(t) - h(t) \\ g'(t) = -f(t) + g(t) - h(t) \\ h'(t) = -f(t) - g(t) + h(t) \end{cases}$$

y $f(0) = -1$, $g(0) = 1$, $h(0) = 3$.

Solución: Se expresa el sistema de ecuaciones diferenciales en forma matricial.

$$\underbrace{\begin{pmatrix} 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} f(t) \\ g(t) \\ h(t) \end{pmatrix}}_{\mathbf{v}} = \underbrace{\begin{pmatrix} f'(t) \\ g'(t) \\ h'(t) \end{pmatrix}}_{\dot{\mathbf{v}}}.$$

La matriz A se diagonalizó en el Ejemplo 5, obteniéndose como matriz diagonal y matriz de cambio de coordenadas

$$D = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \qquad P = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & -1 \end{pmatrix},$$

tales que $D = P^{-1}AP$.

Si se aplica el cambio de coordenadas anterior, $\mathbf{v} = P\widetilde{\mathbf{v}}$, al sistema de ecuaciones diferenciales $A\mathbf{v} = \dot{\mathbf{v}}$, se obtiene

$$AP\widetilde{\mathbf{v}} = P\dot{\widetilde{\mathbf{v}}}, \qquad P^{-1}AP\widetilde{\mathbf{v}} = \dot{\widetilde{\mathbf{v}}}, \qquad D\widetilde{\mathbf{v}} = \dot{\widetilde{\mathbf{v}}}.$$

Este último sistema es de resolución inmediata

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} \widetilde{f}(t) \\ \widetilde{g}(t) \\ \widetilde{h}(t) \end{pmatrix} = \begin{pmatrix} \widetilde{f}'(t) \\ \widetilde{g}'(t) \\ \widetilde{h}'(t) \end{pmatrix},$$

$$\begin{cases} \widetilde{f}(t) = C_1 e^{-t} \\ \widetilde{g}(t) = C_2 e^{2t} \\ \widetilde{h}(t) = C_3 e^{2t}. \end{cases}$$

Deshaciendo el cambio para volver a las funciones de partida,

$$\mathbf{v} = P\widetilde{\mathbf{v}} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & -1 \end{pmatrix} \begin{pmatrix} C_1 e^{-t} \\ C_2 e^{2t} \\ C_3 e^{2t} \end{pmatrix},$$

se obtiene la solución:

$$\begin{cases} f(t) = C_1 e^{-t} + C_2 e^{2t} \\ g(t) = C_1 e^{-t} + C_3 e^{2t} \\ h(t) = C_1 e^{-t} - C_2 e^{2t} - C_3 e^{2t}, \end{cases}$$

que imponiendo las condiciones iniciales proporciona la solución del problema, siendo $C_1 = 1$, $C_2 = -2$, $C_3 = 0$:

$$\begin{cases} f(t) = e^{-t} - 2e^{2t} \\ g(t) = e^{-t} \\ h(t) = e^{-t} - 2e^{2t}. \end{cases}$$

◇

8.6.4. EXPONENCIAL DE UNA MATRIZ

En ocasiones resulta necesario utilizar la función exponencial de una matriz cuadrada $A \in \mathrm{Mat}_n(\mathbb{R})$, que viene definida por la siguiente serie de potencias:

$$\exp(A) = e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!}.$$

Para entender la definición anterior es necesario conocer los conceptos de límite de una sucesión de matrices y convergencia de una serie de matrices que se presentan a continuación.

Definición 8.31. Sea $\{A_k\}_{k=0}^{\infty}$ una sucesión de matrices, $A_k = (a_{ij})_k \in \mathrm{Mat}_n(\mathbb{R})$, y sea $B = (b_{ij}) \in \mathrm{Mat}_n(\mathbb{R})$. Se dice que

$$\lim_{k\to\infty} A_k = B \quad \Leftrightarrow \quad \lim_{k\to\infty} [a_{ij}]_k = b_{ij}, \quad \forall i, j = 1, \ldots, n.$$

Definición 8.32. Se dice que la serie de matrices $\sum_{k=0}^{\infty} A^k$ es convergente y su suma es igual a la matriz B si:

$$\lim_{m\to\infty} \sum_{k=0}^{m} A^k = B.$$

Calcular la exponencial de una matriz aplicando la definición y usando series no siempre resulta cómodo. Si se aplican propiedades de la exponencial de matrices y algunos de los métodos presentados en el capítulo, la resolución se simplifica.

Si $D = \mathrm{diag}(d_{11}, \ldots, d_{nn}) \in \mathrm{Mat}_n(\mathbb{R})$ es una matriz diagonal, e^D es también una matriz diagonal que se calcula tomando las exponenciales de cada uno de los elementos de la diagonal principal de D:

$$e^D = \begin{pmatrix} e^{d_{11}} & 0 & \cdots & 0 \\ 0 & e^{d_{22}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{d_{nn}} \end{pmatrix}.$$

Además dadas dos matrices $X, Q \in \mathrm{Mat}_n(\mathbb{R})$ con Q regular se verifica:

$$e^{QXQ^{-1}} = Qe^X Q^{-1}.$$

Si se quiere calcular e^A siendo A una matriz cualquiera, basta diagonalizarla (si se puede) hallando la matriz regular P de cambio de coordenadas tal que $D = P^{-1}AP$, o equivalentemente, $A = PAP^{-1}$. Hallando la exponencial de la matriz diagonal y realizando el producto por la matriz P se obtiene:

$$e^A = Pe^D P^{-1}.$$

En caso de que la matriz A no sea diagonalizable se calcula su forma de Jordan J tal que $J = P^{-1}AP$, y procediendo de un modo análogo al anterior se obtiene $e^A = Pe^J P^{-1}$. Para calcular J^n basta considerar que la exponencial de un bloque de Jordan es:

$$J_i = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ 0 & 0 & \lambda & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda \end{pmatrix} \quad \Rightarrow \quad e^{J_i} = \begin{pmatrix} e^\lambda & e^\lambda/1! & e^\lambda/2! & \cdots & e^\lambda/(n-1)! \\ 0 & e^\lambda & e^\lambda/1! & \cdots & e^\lambda/(n-2)! \\ 0 & 0 & e^\lambda & \cdots & e^\lambda/(n-3)! \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & e^\lambda \end{pmatrix}.$$

Como J es una diagonal por bloques, J^n es inmediato aplicando lo anterior.

El cálculo de la exponencial de una matriz se utiliza para resolver sistemas de ecuaciones diferenciales. Dado un sistema del tipo

$$\begin{cases} \dot{\mathbf{v}}(t) = A\mathbf{v}(t) + f(t) \\ \mathbf{v}(t_0) = \mathbf{v}_0, \end{cases}$$

la solución general se obtiene como:

$$\mathbf{v}(t) = e^{A(t-t_0)}\mathbf{v}_0 + \int_{t_0}^{t} e^{A(t-s)} f(s)\, ds.$$

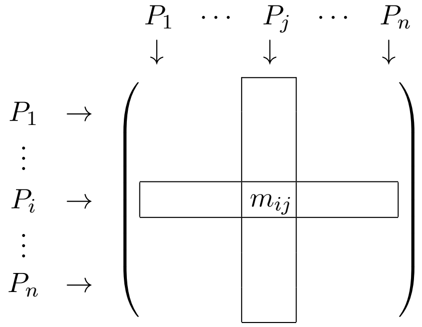
El sistema del Ejemplo 14 se podría haber resuelto con la fórmula anterior.

8.6.5. BÚSQUEDA EN INTERNET

La aparición del buscador Google en 1998 supuso una revolución en el universo de la información. El diseño fue realizado por dos jóvenes estudiantes de doctorado de la Universidad de Stanford, Sergei Brin (graduado en Informática) y Lawrance Page (graduado en Matemáticas). En el momento en el que fue diseñado Google existían otros buscadores, pero el objetivo que se propusieron fue crear un buscador en el que al menos una de las diez primeras páginas que se muestran tras la búsqueda contuviese información útil para el que realiza la consulta.

Para lograr este objetivo necesitaron un *criterio de ordenación de la información*, etiquetando cada una de las páginas de la red con un símbolo P_j, y asignando a cada P_j un número x_j que representase la *importancia* de dicha información. Los valores de x_j podían variar entre 0 y 1.

Con estos valores la red se describe mediante un *grafo* dirigido, cuya matriz de adyacencia es $M \in \mathrm{Mat}_n(\mathbb{R})$. El elemento m_{ij} indica si hay o no enlace desde la página P_j a la P_i. En caso afirmativo vale 1 y en caso negativo 0:



Se define además el vector $\mathbf{x}$ de importancias de modo que cada elemento x_j indica la importancia de cada página P_j. El criterio seguido para establecer la importancia se basa en dos principios:

- páginas muy citadas,
- páginas poco citadas, pero desde sitios muy importantes.

Con este criterio la decisión de la importancia x_j de cada página P_j es proporcional a la suma de las importancias de las páginas que enlazan con P_j. Matricialmente esto se traduce en resolver el sistema

$$\mathbf{x} = KM\mathbf{x}$$

o equivalentemente se busca $\mathbf{x}$ tal que:

$$M\mathbf{x} = \lambda\mathbf{x},$$

es decir, el problema se ha convertido en una cuestión de cálculo de valores y vectores propios.

En principio la solución al problema no es única, ya que como se ha visto el conjunto de vectores propios asociados a un mismo valor propio (junto con el vector nulo) es un subespacio vectorial denotado por $S(f, \lambda)$.

Ahora dentro de este subespacio se buscan vectores con todas las entradas no negativas $x_j \geq 0, \forall j$, denotado por $\mathbf{x} \geq 0$. Pero se trata de llegar a una solución única que proporcione la primera salida de Google tras la búsqueda.

Considerando que el modelo que describe el problema es probabilístico y dinámico se puede construir una nueva matriz $M' \in \mathrm{Mat}_n(\mathbb{R})$ de modo que cada elemento $m'_{ij} = m_{ij}/N_j$, siendo N_j el número de enlaces desde la página P_j (es decir, la suma de las entradas de cada columna de M). De este modo la matriz M' tiene todos sus elementos no negativos (valores entre 0 y 1) y además la suma de los registros de cada columna es 1. Es decir, la matriz M' es una matriz de Markov, que se denomina *matriz de transición*, donde cada m'_{ij} es la probabilidad de pasar del estado P_j al estado P_i. Los registros de las sucesivas potencias de la matriz son las probabilidades de pasar de P_j a P_i tras varios instantes de tiempo.

El teorema de Perron-Frobenius puede resolver el problema anterior, ya que asegura que para matrices no negativas e irreducibles existe un único vector propio asociado al valor propio λ con elementos todos positivos y cuya suma es 1.

El algoritmo con el que Google ordena los resultados se denomina PageRank.

8.6.6. APLICACIONES GEOMÉTRICAS

El cálculo de valores y vectores propios tiene aplicaciones en geometría analítica en el estudio de cónicas y cuádricas. Este apartado se aborda en el Capítulo 9 con el estudio de las formas cuadráticas.

8.7. COMANDOS DE *wxMaxima*

- Calcular los valores propios de una matriz A:

  ```
  -->      eigenvalues(A);
  ```

  Este comando devuelve dos listas, la primera con los valores propios de la matriz y la segunda con las multiplicidades respectivas. Se puede acceder desde el menú “`\Álgebra\Valores propios`”.

- Calcular los vectores propios de una matriz A:

  ```
  -->      eigenvectors(A);
  ```

  Este comando devuelve dos listas, la primera con los valores propios y sus multiplicidades, y la segunda con las listas que componen los correspondientes vectores propios, (desde el menú “`\Álgebra\Vectores propios`”).

- Estudiar si una matriz A es diagonalizable o no

  ```
  -->     eigenvectors(A)$
          nondiagonalizable;
  ```

  Devuelve “`true`” en caso de que la matriz cuadrada A sea no diagonalizable y “`false`” en caso de que sí lo sea.

- Calcular el polinomio característico de la matriz A con variable t:

  ```
  -->      charpoly(A,t);
  ```

  Directamente desde el menú “`\Álgebra\Polinomio característico`”. En ocasiones es aconsejable utilizar órdenes como “`expand`” y/o “`factor`” para obtener una expresión más manejable del polinomio.

- Calcular una matriz cuyo polinomio característico sea $p(t)$:

  ```
  -->      load(linearalgebra);
           polytocompanion(p,t);
  ```

- Para hallar los subespacios fundamentales se puede hacer uso de algunos comandos antes citados: “`matrix`”,“`determinant`”,“`solve`”, “`nullspace`”.

- Calcular un bloque de Jordan de tamaño n asociado al valor propio a:

```
-->     load(diag);
        JF(a,n);
```

- Calcular la forma canónica de Jordan de la matriz $A \in \mathrm{Mat}_n(\mathbb{R})$:

```
-->     load(diag);
        dispJordan(jordan(A));
```

- Calcular la matriz P de cambio de coordenadas que relaciona la matriz $A \in \mathrm{Mat}_n(\mathbb{R})$ con su forma canónica de Jordan J, $P^{-1}AP = J$:

```
-->     load(diag);
        P:ModeMatrix(A,jordan(A));
```

8.8. EJERCICIOS PROPUESTOS

1. Hallar los valores propios y los subespacios fundamentales de las matrices

$$\text{a) } A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix} \quad \text{b) } A = \begin{pmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{pmatrix} \quad \text{c) } A = \begin{pmatrix} -1 & -1 & -1 \\ 2 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

2. ¿Son diagonalizables las matrices A del ejercicio anterior? En caso afirmativo encontrar una matriz P regular tal que $P^{-1}AP$ sea diagonal.

3. Comprobar si las siguientes matrices son diagonalizables. En caso afirmativo encontrar P regular tal que $P^{-1}AP$ sea diagonal.

$$\text{a) } A = \begin{pmatrix} 1 & 4 \\ 3 & 2 \end{pmatrix} \text{ b) } A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix} \text{ c) } A = \begin{pmatrix} 3 & 1 & 6 \\ 2 & 1 & 0 \\ -1 & 0 & -3 \end{pmatrix} \text{ d) } A = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix}$$

4. Sean $A = \begin{pmatrix} 6 & 2 \\ 4 & -1 \end{pmatrix}$ y $B = \begin{pmatrix} 1 & 0 \\ 4 & -14 \end{pmatrix}$. Comprobar que A y B son diagonalizables pero la matriz producto AB no lo es.

5. Comprobar si los siguientes endomorfismos son diagonalizables. En caso afirmativo encontrar una matriz diagonal D y una matriz P regular tal que $P^{-1}AP = D$.

a) $f : \mathbb{R}^2 \to \mathbb{R}^2$, $f(a, b) = (2a - 4b, -a + 2b)$

b) $f : \mathbb{R}_2[x] \to \mathbb{R}_2[x]$, $f(a + bx + cx^2) = (3a + 6c) - 3b\,x + (5a + 2c)x^2$.

6. ¿Son diagonalizables los siguientes endomorfismos? En caso afirmativo encontrar una base de vectores propios para la aplicación lineal f y hallar una matriz P regular tal que $P^{-1}AP$ sea diagonal siendo A una matriz coordenada de f.

a) $f : \mathbb{R}_2[x] \to \mathbb{R}_2[x]$, dado por

$$f(a + bx + cx^2) = (a + 3b + 7c) - (3a + 5b + 7c)x + (3a + 3b + 5c)x^2.$$

b) $f : \mathbb{R}^4 \to \mathbb{R}^4$, dado por $f(a, b, c, d) = (2a + b + 4c, -b, 3a + 3c, 5d)$.

c) $f : \mathrm{Mat}_2(\mathbb{R}) \to \mathrm{Mat}_2(\mathbb{R})$, dado por $f\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 2a + 2c & 4b + 6d \\ -2a - 3c & -3b - 5d \end{pmatrix}$.

d) $f : \mathbb{R}_2[x] \to \mathbb{R}_2[x]$, dado por $f(a + bx + cx^2) = (3c - 4a) + 2bx + (2a + b + c)x^2$.

7. Sea $f : V \to V$ un endomorfismo idempotente, es decir, $f^2 = f$. Probar que los únicos valores propios de f son 0 y 1.

8. Sea $f : \mathbb{R}_2[x] \to \mathbb{R}_2[x]$, el endomorfismo dado por

$$f(a + bx + cx^2) = (a + 2c) + (ka + 6b)x + (5a + 4c)x^2.$$

Encontrar el valor o valores de k para los que f es diagonalizable. Elegir un valor de k para el que f sea diagonalizable y en ese caso encontrar una base de $\mathbb{R}_2[x]$ formada por vectores propios de f.

9. Sea $A \in \mathrm{Mat}_n(\mathbb{R})$, demostrar que si $\lambda = 0$ es valor propio de A entonces A es una matriz no regular.

10. Sea $A \in \mathrm{Mat}_n(\mathbb{R})$, demostrar que si λ es valor propio de A entonces λ^p es valor propio de A^p, $\forall p \in \mathbb{N}$.

11. Se considera la matriz

$$A = \begin{pmatrix} 0 & a & 0 \\ a & 0 & b \\ 0 & b & 0 \end{pmatrix}$$

donde al menos uno de los números reales a o b son no nulos. Hallar el polinomio característico de A, los valores propios de A y encontrar, si es posible, una matriz regular P tal que $P^{-1}AP$ sea diagonal.

12. Probar que la única matriz $n \times n$ diagonalizable con un solo valor propio λ es $A = \lambda I_n$.

13. Calcular A^{10} y A^{11}, siendo $A = \begin{pmatrix} -2 & 7 \\ 1 & 2 \end{pmatrix}$.

14. Sea $x_0, x_1, \ldots, x_n, \ldots$ una sucesión de números reales determinada por las condiciones $x_0 = x_1 = 1$ y $x_{n+2} = 6x_n + x_{n+1}$ para todo $n \geq 0$. Se define el vector $\mathbf{v}_n = \begin{pmatrix} x_n \\ x_{n+1} \end{pmatrix}$ para todo $n \geq 0$ y se considera la matriz $A = \begin{pmatrix} 0 & 1 \\ 6 & 1 \end{pmatrix}$.

- a) Comprobar que $\mathbf{v}_{n+1} = A\mathbf{v}_n$, para todo n.
- b) Deducir que $\mathbf{v}_n = A^n \mathbf{v}_0$, para todo n.
- c) Comprobar que A es diagonalizable y encontrar P regular tal que $P^{-1}AP = D$, siendo $D \in \mathrm{Mat}_2(\mathbb{R})$ una matriz diagonal.
- d) Calcular la potencia n-ésima de A usando la matriz D del apartado anterior.
- e) Encontrar una fórmula para x_n, $n \geq 0$, dependiente de n pero no de los valores anteriores.

15. Se considera el endomorfismo $f : \mathbb{R}^3 \to \mathbb{R}^3$ dado por:

$$f(x, y, z) = (3x, 2y + 2z, 2y + 5z).$$

¿Es un endomorfismo ortogonalmente diagonalizable? En caso afirmativo encontrar una matriz ortogonal P tal que P^tAP sea diagonal.

16. Se consideran las siguientes matrices

$$\text{a) } A = \begin{pmatrix} 3 & 0 & 7 \\ 0 & 5 & 0 \\ 7 & 0 & 3 \end{pmatrix} \qquad \text{b) } A = \begin{pmatrix} 5 & 3 & 0 & 0 \\ 3 & 5 & 0 & 0 \\ 0 & 0 & 7 & 1 \\ 0 & 0 & 1 & 7 \end{pmatrix} \qquad \text{c) } A = \begin{pmatrix} 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix}$$

Encontrar para cada una de ellas, en caso de que exista, una matriz ortogonal P tal que P^tAP sea diagonal.

17. Calcular la forma canónica de Jordan de las siguientes matrices, hallando una matriz P regular tal que $P^{-1}AP = J$, siendo:

$$\text{a) } A = \begin{pmatrix} 0 & 3 & 1 \\ 2 & -1 & -1 \\ -2 & -1 & -1 \end{pmatrix} \qquad \text{b) } A = \begin{pmatrix} 4 & 0 & 2 \\ 0 & 4 & 3 \\ 0 & 0 & 4 \end{pmatrix}$$

$$\text{c) } A = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & -2 & -1 & -2 \\ 1 & -1 & 0 & -1 \\ 1 & -2 & -1 & -2 \end{pmatrix} \qquad \text{d) } A = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ -1 & 3 & 0 & 0 & 0 \\ -1 & 1 & 2 & 0 & 0 \\ -1 & 1 & 0 & 1 & 1 \\ -1 & 1 & 0 & -1 & 3 \end{pmatrix}$$

8.9. SOLUCIONES A LOS EJERCICIOS PROPUESTOS

1. a)
- Polinomio característico: $p(t) = (t+1)(t-1)(t-3)$.
- Subespacios fundamentales:

$$S(A,-1) = \langle (1,0,-1) \rangle, \quad S(A,1) = \langle (0,1,0) \rangle, \quad S(A,3) = \langle (1,0,1) \rangle.$$

b)
- Polinomio característico: $p(t) = (t-2)^2(t-5)$.
- Subespacios fundamentales:

$$S(A,2) = \langle (1,0,-1),(0,1,-1) \rangle, \quad S(A,5) = \langle (1,1,1) \rangle.$$

c)
- Polinomio característico: $p(t) = t(t-1)^2$.
- Subespacios fundamentales:

$$S(A,0) = \langle (1,-1,0) \rangle, \quad S(A,1) = \langle (1,-1,-1) \rangle.$$

2. a) La matriz A es diagonalizable. Una posible matriz $P = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$ siendo $D = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$.

b) La matriz A es diagonalizable. Una posible matriz $P = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ -1 & -1 & 1 \end{pmatrix}$ siendo $D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 5 \end{pmatrix}$.

c) La matriz A no es diagonalizable, por tanto, no existe P regular tal que $P^{-1}AP$ sea diagonal.

3. a) A sí es diagonalizable.
- Polinomio característico: $p(t) = (t+2)(t-5)$.
- Subespacios fundamentales:

$$S(A,-2) = \langle (4,-3) \rangle, \quad S(A,5) = \langle (1,1) \rangle.$$

- $D = \begin{pmatrix} -2 & 0 \\ 0 & 5 \end{pmatrix}, \quad P = \begin{pmatrix} 4 & 1 \\ -3 & 1 \end{pmatrix}.$

b) A sí es diagonalizable.
- Polinomio característico: $p(t) = (t-1)^2(t-2)$.
- Subespacios fundamentales:

$$S(A,1) = \langle (1,0,-1),(0,1,-1) \rangle, \quad S(A,2) = \langle (0,1,0) \rangle.$$

- $D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ -1 & -1 & 0 \end{pmatrix}.$

c) A no es diagonalizable.
  - Polinomio característico: $p(t) = (t+1)^2(t-3)$.
  - Subespacio fundamental: $S(A,-1) = \langle(2,-2,-1)\rangle$, $\dim S(A,-1) = 1 \neq m_1 = 2$.

d) A no es diagonalizable. No todas las raíces del polinomio característico $p(t) = (t-1)(t^2+1)$ son reales.

4.
- Polinomio característico de A: $p_A(t) = (t+2)(t-7)$. Como A tiene dos valores propios de multiplicidad 1, entonces es diagonalizable.
- Polinomio característico de B: $p_B(t) = (t+14)(t-1)$. Como B tiene dos valores propios de multiplicidad 1, entonces es diagonalizable.
- Polinomio característico de AB: $p_{AB}(t) = (t-14)^2$. Subespacio fundamental, $S(AB,14) = \langle(1,0)\rangle$. Como $\dim S(AB,14) = 1 \neq m = 2$, la matriz AB no es diagonalizable.

5. a) f es diagonalizable. Tomando la matriz $A = \begin{pmatrix} 2 & -4 \\ -1 & 2 \end{pmatrix}$ coordenada de f en base canónica, $D = \begin{pmatrix} 4 & 0 \\ 0 & 0 \end{pmatrix}$, $P = \begin{pmatrix} -2 & 2 \\ 1 & 1 \end{pmatrix}$.

b) f es diagonalizable. Tomando la matriz $A = \begin{pmatrix} 3 & 0 & 6 \\ 0 & -3 & 0 \\ 5 & 0 & 2 \end{pmatrix}$ coordenada de f en base canónica, $D = \begin{pmatrix} -3 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 8 \end{pmatrix}$, $P = \begin{pmatrix} 1 & 0 & 6 \\ 0 & 1 & 0 \\ -1 & 0 & 5 \end{pmatrix}$.

6. a) f es diagonalizable.
  - Matriz coordenada de f en base canónica $\{1, x, x^2\}$: $A = \begin{pmatrix} 1 & 3 & 7 \\ -3 & -5 & -7 \\ 3 & 3 & 5 \end{pmatrix}$.
  - Polinomio característico: $p(t) = (t+2)^2(t-5)$.
  - Base de vectores propios: $\mathcal{D} = \{-7+3x^2, -1+x, 1-x+x^2\}$.
  - $D = \begin{pmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 5 \end{pmatrix}$, $P = \begin{pmatrix} -7 & -1 & 1 \\ 0 & 1 & -1 \\ 3 & 0 & 1 \end{pmatrix}$.

b) f no es diagonalizable.
  - Matriz coordenada de f en base canónica: $\begin{pmatrix} 2 & 1 & 4 & 0 \\ 0 & -1 & 0 & 0 \\ 3 & 0 & 3 & 0 \\ 0 & 0 & 0 & 5 \end{pmatrix}$.
  - Polinomio característico: $p(t) = (t+1)^2(t-5)(t-6)$.
  - $\dim S(f,-1) = 1 \neq m_1 = 2$.

c) f es diagonalizable.
  - Matriz coordenada de f en la base canónica de $\mathrm{Mat}_2(\mathbb{R})$: $\begin{pmatrix} 2 & 0 & 2 & 0 \\ 0 & 4 & 0 & 6 \\ -2 & 0 & -3 & 0 \\ 0 & -3 & 0 & -5 \end{pmatrix}$.

- Polinomio característico: $p(t) = (t+2)^2(t-1)^2$.
- Una base de vectores propios es:

$$\mathcal{D} = \left\{ \begin{pmatrix} 1 & 0 \\ -2 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} -2 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -2 \\ 0 & 1 \end{pmatrix} \right\}.$$

- $D = \begin{pmatrix} -2 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad P = \begin{pmatrix} 1 & 0 & -2 & 0 \\ 0 & 1 & 0 & -2 \\ -2 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix}.$

d) f no es diagonalizable.

- Matriz coordenada de f en base canónica $\{1, x, x^2\}$: $\begin{pmatrix} -4 & 0 & 3 \\ 0 & 2 & 0 \\ 2 & 1 & 1 \end{pmatrix}$.
- Polinomio característico: $p(t) = (t+5)(t-2)^2$.
- $\dim S(f,2) = 1 \neq m_2 = 2$.

7. Aplicar la definición de valor y vector propio y el hecho de que f es lineal.

8.
- Matriz coordenada de f en la base canónica $\{1, x, x^2\}$: $\begin{pmatrix} 1 & 0 & 2 \\ k & 6 & 0 \\ 5 & 0 & 4 \end{pmatrix}$.
- Polinomio característico: $p(t) = (t+1)(t-6)^2$.
- f es diagonalizable $\Longleftrightarrow k = 0$. Para $k = 0$, una base de vectores propios es:

$$\mathcal{D} = \{1 - x^2, 2 + 5x^2, x\}.$$

9. Utilizar que si λ es valor propio, entonces $|A - \lambda I| = 0$.

10. Utilizar inducción.

11.
- Polinomio característico:

$$p(t) = t(t^2 - (a^2 + b^2)) = t(t + \sqrt{a^2 + b^2})(t - \sqrt{a^2 + b^2}).$$

- Valores propios: $\lambda_1 = -\sqrt{a^2 + b^2}, \quad \lambda_2 = 0, \quad \lambda_3 = \sqrt{a^2 + b^2}$.
- Subespacios fundamentales:

$$S(A, \lambda_1) = \langle (a, -\sqrt{a^2 + b^2}, b) \rangle, \quad S(A, 0) = \langle (b, 0, -a) \rangle,$$

$$S(A, \lambda_3) = \langle (a, \sqrt{a^2 + b^2}, b) \rangle.$$

- $D = \begin{pmatrix} -\sqrt{a^2 + b^2} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \sqrt{a^2 + b^2} \end{pmatrix}, \quad P = \begin{pmatrix} a & b & a \\ -\sqrt{a^2 + b^2} & 0 & \sqrt{a^2 + b^2} \\ b & -a & b \end{pmatrix}.$

12. Ayuda: A diagonalizable $\Longleftrightarrow A = P^{-1}DP$ con P regular y D diagonal.

13. $A^{10} = \begin{pmatrix} 11^5 & 0 \\ 0 & 11^5 \end{pmatrix}$, $A^{11} = \begin{pmatrix} -2 \cdot 11^5 & 7 \cdot 11^5 \\ 11^5 & 2 \cdot 11^5 \end{pmatrix}$, $(11^5 = 161051)$.

14. a) Comprobación.

b) Utilizar inducción.

c) $D = \begin{pmatrix} -2 & 0 \\ 0 & 3 \end{pmatrix}$, $P = \begin{pmatrix} 1 & 1 \\ -2 & 3 \end{pmatrix}$.

d) $A^n = \frac{1}{5}\begin{pmatrix} 2\cdot 3^n + 3\cdot(-2)^n & 3^n - (-2)^n \\ 2\cdot 3^{n+1} + 3\cdot(-2)^{n+1} & 3^{n+1} - (-2)^{n+1} \end{pmatrix}$.

e) $x_n = \frac{1}{5}\left[3^{n+1} + (-1)^n \cdot 2^{n+1}\right]$, $n \geq 0$.

15.

- f es ortogonalmente diagonalizable, ya que la matriz $A = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 2 \\ 0 & 2 & 5 \end{pmatrix}$ coordenada de f en base canónica es simétrica.
- $D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 6 \end{pmatrix}$ y $P = \begin{pmatrix} 0 & 1 & 0 \\ -2/\sqrt{5} & 0 & 1/\sqrt{5} \\ 1/\sqrt{5} & 0 & 2/\sqrt{5} \end{pmatrix}$.

16. a) A simétrica, por tanto, ortogonalmente diagonalizable.

$$D = \begin{pmatrix} -4 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 10 \end{pmatrix}, \qquad P = \begin{pmatrix} 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 1 & 0 \\ -1/\sqrt{2} & 0 & 1/\sqrt{2} \end{pmatrix}.$$

b) A simétrica, por tanto, ortogonalmente diagonalizable.

$$D = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 \\ 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 8 \end{pmatrix}, \qquad P = \begin{pmatrix} 1/\sqrt{2} & 0 & 1/\sqrt{2} & 0 \\ -1/\sqrt{2} & 0 & 1/\sqrt{2} & 0 \\ 0 & 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & -1/\sqrt{2} & 0 & 1/\sqrt{2} \end{pmatrix}.$$

c) A simétrica, por tanto, ortogonalmente diagonalizable.

$$D = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \qquad P = \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{3} & 0 & 2/\sqrt{6} \\ 1/\sqrt{3} & -1/\sqrt{2} & -1/\sqrt{6} \end{pmatrix}.$$

17. a) $J = \begin{pmatrix} -2 & 1 & 0 \\ 0 & -2 & 1 \\ 0 & 0 & -2 \end{pmatrix}, \qquad P = \begin{pmatrix} -1 & 0 & 1 \\ -1 & -1 & 0 \\ -1 & 0 & 0 \end{pmatrix}.$

b) $J = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 4 & 1 \\ 0 & 0 & 4 \end{pmatrix}, \qquad P = \begin{pmatrix} 0 & 2 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$

c) $J = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -5 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad P = \begin{pmatrix} 1 & -1 & -1 & 0 \\ 0 & 2 & -1 & -1 \\ 1 & 1 & 1 & 0 \\ 0 & 2 & 0 & 1 \end{pmatrix}.$

d) $J = \begin{pmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}, \qquad P = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{pmatrix}.$

CAPÍTULO 9

Formas Bilineales y Cuadráticas

En el Capítulo 7 se han introducido las aplicaciones lineales generalizando las funciones lineales de $\mathbb{R}$ en $\mathbb{R}$, o rectas. Para presentar el concepto de forma cuadrática se puede tomar como modelo geométrico las cónicas en el plano $\mathbb{R}^2$ o las cuádricas en el espacio $\mathbb{R}^3$, ambas expresadas como polinomios de grado dos en dos y tres variables respectivamente. Concretamente, la ecuación general de una cónica en el plano viene dada por

$$\underbrace{a_{11}x^2 + 2a_{12}xy + a_{22}y^2}_{\text{parte cuadrática}} + \underbrace{2a_1x + 2a_2y}_{\text{parte lineal}} + a_0 = 0.$$

La parte cuadrática se puede representar matricialmente como $\mathbf{x}^t A\mathbf{x}$, siendo A una matriz simétrica:

$$\begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Las matrices simétricas han sido muy estudiadas por grandes matemáticos. Cabe mencionar a Cauchy quien probó que los valores propios de una matriz simétrica son reales y, si la matriz es de orden 2×2, determinan la longitud de los ejes de la cónica asociada a esta matriz simétrica.

Por otra parte, si se considera la matriz A como la matriz identidad, entonces la expresión $\mathbf{x}^t A\mathbf{x}$ se reduce a $x^2 + y^2$ o equivalentemente,

$$\mathbf{x}^t\mathbf{x} = \langle \mathbf{x}, \mathbf{x} \rangle,$$

donde $\langle \cdot, \cdot \rangle$ es el producto escalar usual de $\mathbb{R}^2$. Desde este punto de vista, las formas cuadráticas pueden ser vistas como una generalización del producto escalar introducido en el Capítulo 6.

9.1. FORMAS BILINEALES

A lo largo de este capítulo se trabaja con un tipo especial de aplicaciones llamadas *formas*, que son aplicaciones de un espacio vectorial en su cuerpo de escalares $\mathbb{K}$. Puesto que los espacios vectoriales con los que se trabaja en este texto son reales, las formas que se consideran son aplicaciones de un espacio vectorial en $\mathbb{R}$.

Definición 9.1. Sea V un espacio vectorial real. Una **forma bilineal sobre** V es una aplicación

$$\begin{array}{rcl} F : V \times V & \longrightarrow & \mathbb{R} \\ (\mathbf{u}, \mathbf{v}) & \mapsto & F(\mathbf{u}, \mathbf{v}) \end{array}$$

que cumple las siguientes propiedades:

- $F(\mathbf{u}_1 + \mathbf{u}_2, \mathbf{v}) = F(\mathbf{u}_1, \mathbf{v}) + F(\mathbf{u}_2, \mathbf{v}), \quad \forall \mathbf{u}_1, \mathbf{u}_2, \mathbf{v} \in V.$
- $F(\mathbf{u}, \mathbf{v}_1 + \mathbf{v}_2) = F(\mathbf{u}, \mathbf{v}_1) + F(\mathbf{u}, \mathbf{v}_2), \quad \forall \mathbf{u}, \mathbf{v}_1, \mathbf{v}_2 \in V.$
- $F(\lambda \mathbf{u}, \mathbf{v}) = \lambda F(\mathbf{u}, \mathbf{v}), \quad \forall \mathbf{u}, \mathbf{v} \in V, \quad \forall \lambda \in \mathbb{R}.$
- $F(\mathbf{u}, \lambda \mathbf{v}) = \lambda F(\mathbf{u}, \mathbf{v}), \quad \forall \mathbf{u}, \mathbf{v} \in V, \quad \forall \lambda \in \mathbb{R}.$

Observación. Las propiedades anteriores se pueden resumir diciendo que una forma bilineal es lineal en cada componente.

Directamente de la definición anterior se tienen las siguientes propiedades:

Propiedades. Sea $F : V \times V \to \mathbb{R}$ una forma bilineal. Entonces:

- $F(\mathbf{u}, \mathbf{0}) = 0$, $\forall \mathbf{u} \in V$.
- $F(\mathbf{0}, \mathbf{v}) = 0$, $\forall \mathbf{v} \in V$.
- $F(-\mathbf{u}, \mathbf{v}) = -F(\mathbf{u}, \mathbf{v})$, $\forall \mathbf{u}, \mathbf{v} \in V$.
- $F(\mathbf{u}, -\mathbf{v}) = -F(\mathbf{u}, \mathbf{v})$, $\forall \mathbf{u}, \mathbf{v} \in V$.

Definición 9.2. Sea F una forma bilineal sobre un espacio vectorial V.

- F es **simétrica** si $F(\mathbf{u}, \mathbf{v}) = F(\mathbf{v}, \mathbf{u}), \quad \forall \mathbf{u}, \mathbf{v} \in V.$
- F es **antisimétrica** si $F(\mathbf{u}, \mathbf{v}) = -F(\mathbf{v}, \mathbf{u}), \quad \forall \mathbf{u}, \mathbf{v} \in V.$

$\rightsquigarrow$ **Ejercicio.** Sea F una forma bilineal antisimétrica sobre un espacio vectorial V. Probar que $F(\mathbf{u}, \mathbf{u}) = 0$, $\forall \mathbf{u} \in V$.

✓ **Ejemplo 1.** Sea F la forma bilineal definida por:

$$\begin{array}{rcl} F:\mathbb{R}^2\times\mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ ((x_1,x_2),(y_1,y_2)) & \mapsto & x_1y_2+k\,x_2y_1, \end{array}$$

donde $k\in\mathbb{R}$. Estudiar los valores de k para los que F es una forma bilineal simétrica o antisimétrica.

Solución: Para ver si F es simétrica o antisimétrica hay que relacionar las expresiones de $F(\mathbf{u},\mathbf{v})$ y $F(\mathbf{v},\mathbf{u})$. Si $\mathbf{u}=(x_1,x_2)$ y $\mathbf{v}=(y_1,y_2)$:

$$F(\mathbf{u},\mathbf{v})=x_1y_2+k\,x_2y_1,\quad F(\mathbf{v},\mathbf{u})=y_1x_2+k\,y_2x_1.$$

- F es simétrica $\Longleftrightarrow x_1y_2+k\,x_2y_1=k\,x_1y_2+x_2y_1,\ \forall x_1,x_2,y_1,y_2\in\mathbb{R}\Longleftrightarrow k=1$.
- F es antisimétrica $\Longleftrightarrow x_1y_2+k\,x_2y_1=-k\,x_1y_2-x_2y_1,\ \forall x_1,x_2,y_1,y_2\in\mathbb{R}\Longleftrightarrow k=-1$.
- Si $k\neq\pm1$, entonces F es una forma bilineal que no es ni simétrica ni antisimétrica.

◇

✓ **Ejemplo 2.** Sea $(V,\langle\cdot,\cdot\rangle)$ un espacio vectorial euclídeo. Según la Definición 6.3 el producto escalar $\langle\cdot,\cdot\rangle$ es una forma bilineal simétrica. Además, es definida positiva, es decir, cumple que $\langle\mathbf{u},\mathbf{u}\rangle\geq 0$, $\forall\mathbf{u}\in V$, obteniéndose únicamente la igualdad para el vector $\mathbf{u}=\mathbf{0}$.

◇

9.1.1. EXPRESIÓN COORDENADA DE UNA FORMA BILINEAL

Las formas bilineales sobre espacios de dimensión finita se pueden representar mediante matrices, de modo análogo a lo que sucede con las aplicaciones lineales.

Definición 9.3. Sea V un espacio vectorial real con $\dim V=n$, y $F:V\times V\to\mathbb{R}$ una forma bilineal. Fijada una base $\mathcal{B}=\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ de V, la **matriz coordenada** de la forma bilineal F respecto de la base $\mathcal{B}$ es aquella cuyas entradas son $a_{ij}=F(\mathbf{v}_i,\mathbf{v}_j)$; es decir:

$$A=\begin{pmatrix} F(\mathbf{v}_1,\mathbf{v}_1) & F(\mathbf{v}_1,\mathbf{v}_2) & \cdots & F(\mathbf{v}_1,\mathbf{v}_n) \\ F(\mathbf{v}_2,\mathbf{v}_1) & F(\mathbf{v}_2,\mathbf{v}_2) & \cdots & F(\mathbf{v}_2,\mathbf{v}_n) \\ \vdots & \vdots & \ddots & \vdots \\ F(\mathbf{v}_n,\mathbf{v}_1) & F(\mathbf{v}_n,\mathbf{v}_2) & \cdots & F(\mathbf{v}_n,\mathbf{v}_n) \end{pmatrix}\in\mathrm{Mat}_n(\mathbb{R}).$$

La matriz coordenada A se puede utilizar para calcular la imagen por F de cualquier par de vectores $\mathbf{u},\mathbf{v}\in V$: si las coordenadas del vector $\mathbf{u}$ en la base $\mathcal{B}$ son

$\mathcal{C}_{\mathcal{B}}(\mathbf{u}) = \mathbf{x} = (x_1, \ldots, x_n)$ y las del vector $\mathbf{v}$ son $\mathcal{C}_{\mathcal{B}}(\mathbf{v}) = \mathbf{y} = (y_1, \ldots, y_n)$, entonces:

$$F(\mathbf{u}, \mathbf{v}) = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} A \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \mathbf{x}^t A \mathbf{y} = \sum_{i,j} a_{ij} x_i y_j,$$

donde se está usando la notación matricial de los vectores de $\mathbb{R}^n$ introducida en la página 127. A esta manera de expresar la forma bilineal se le llama *expresión coordenada.*

✓ **Ejemplo 3.** Considerar la siguiente forma bilineal

$$\begin{array}{cccc} F: & \mathbb{R}^2 \times \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ & ((x_1, x_2), (y_1, y_2)) & \mapsto & 2x_1y_1 - x_2y_2, \end{array}$$

y los vectores $\mathbf{u} = (1, 1)$, $\mathbf{v} = (0, 2) \in \mathbb{R}^2$.

a) Hallar la matriz coordenada de f respecto de la base canónica $\mathcal{B}_1 = \{(1, 0), (0, 1)\}$ y utilizarla para calcular la imagen $F(\mathbf{u}, \mathbf{v})$.

b) Repetir el ejercicio utilizando la base $\mathcal{B}_2 = \{(1, 1), (-1, 1)\}$.

Solución:

a) La matriz coordenada respecto de la base canónica es $A = \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}$ puesto que

$$\begin{array}{ll} a_{11} = F((1, 0), (1, 0)) = 2, & a_{12} = F((1, 0), (0, 1)) = 0, \\ a_{21} = F((0, 1), (1, 0)) = 0, & a_{22} = F((0, 1), (0, 1)) = -1. \end{array}$$

Así pues, la expresión coordenada de F en la base canónica es

$$F((x_1, x_2), (y_1, y_2)) = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

En concreto, para los vectores $\mathbf{u}$ y $\mathbf{v}$, la imagen por F se calcula como

$$F(\mathbf{u}, \mathbf{v}) = \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 \\ 2 \end{pmatrix} = -2.$$

b) Para hallar la matriz coordenada en base $\mathcal{B}_2$ es necesario calcular:

$$\begin{array}{ll} b_{11} = F((1, 1), (1, 1)) = 1, & b_{12} = F((1, 1), (-1, 1)) = -3, \\ b_{21} = F((-1, 1), (1, 1)) = -3, & b_{22} = F((-1, 1), (-1, 1)) = 1, \end{array}$$

por lo que $B = \begin{pmatrix} 1 & -3 \\ -3 & 1 \end{pmatrix}$. Para calcular $F(\mathbf{u}, \mathbf{v})$ utilizando la matriz B se necesitan las coordenadas de los vectores $\mathbf{u}$ y $\mathbf{v}$ respecto de la base $\mathcal{B}_2$:

$$\mathcal{C}_{\mathcal{B}_2}(\mathbf{u}) = (1, 0), \quad \mathcal{C}_{\mathcal{B}_2}(\mathbf{v}) = (1, 1).$$

Así pues,

$$F(\mathbf{u},\mathbf{v}) = \begin{pmatrix}1 & 0\end{pmatrix}\begin{pmatrix}1 & -3\\ -3 & 1\end{pmatrix}\begin{pmatrix}1\\ 1\end{pmatrix} = -2.$$

◇

✓ **Ejemplo 4.** Sea $A = \begin{pmatrix}1 & 1\\ 1 & 3\end{pmatrix}$ la matriz coordenada de una forma bilineal F sobre un espacio vectorial V en base canónica. Dar la expresión coordenada de F si:

a) $V = \mathbb{R}^2$.

b) $V = \mathbb{R}_1[x]$.

Solución:

a) Sea $V = \mathbb{R}^2$. Considerar $\mathbf{u} = (x_1, x_2)$, $\mathbf{v} = (y_1, y_2) \in \mathbb{R}^2$. Sus coordenadas en base canónica son precisamente $\mathbf{x} = (x_1, x_2)$, $\mathbf{y} = (y_1, y_2)$. El siguiente producto,

$$\mathbf{x}^t A\mathbf{y} = \begin{pmatrix}x_1 & x_2\end{pmatrix}\begin{pmatrix}1 & 1\\ 1 & 3\end{pmatrix}\begin{pmatrix}y_1\\ y_2\end{pmatrix} = x_1y_1 + x_1y_2 + x_2y_1 + 3x_2y_2,$$

determina la expresión coordenada de F:

$$\begin{array}{rccl} F: & \mathbb{R}^2 \times \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ & ((x_1,x_2),(y_1,y_2)) & \mapsto & x_1y_1 + x_1y_2 + x_2y_1 + 3x_2y_2. \end{array}$$

b) Sea $V = \mathbb{R}_1[x]$. Considerar $\mathbf{u} = a_1 + a_2x$, $\mathbf{v} = b_1 + b_2x$ dos vectores de $\mathbb{R}_1[x]$. Sus coordenadas en base canónica son $\mathbf{x} = (a_1, a_2)$, $\mathbf{y} = (b_1, b_2)$. Ahora:

$$\mathbf{x}^t A\mathbf{y} = \begin{pmatrix}a_1 & a_2\end{pmatrix}\begin{pmatrix}1 & 1\\ 1 & 3\end{pmatrix}\begin{pmatrix}b_1\\ b_2\end{pmatrix} = a_1b_1 + a_1b_2 + a_2b_1 + 3a_2b_2.$$

La expresión coordenada de F viene dada por:

$$\begin{array}{rccl} F: & \mathbb{R}_1[x] \times \mathbb{R}_1[x] & \longrightarrow & \mathbb{R} \\ & (a_1 + a_2x, b_1 + b_2x) & \mapsto & a_1b_1 + a_1b_2 + a_2b_1 + 3a_2b_2. \end{array}$$

◇

✓ **Ejemplo 5.** Considerar el espacio euclídeo usual $(\mathbb{R}^n, \langle\cdot,\cdot\rangle)$. Como se ha visto en el Ejemplo 2, $\langle\cdot,\cdot\rangle$ es una forma bilineal sobre $\mathbb{R}^n$. La matriz coordenada en términos de la base canónica es la matriz identidad I_n, por lo que la expresión matricial resulta:

$$\langle(x_1,\ldots,x_n),(y_1,\ldots,y_n)\rangle = \begin{pmatrix}x_1 & \cdots & x_n\end{pmatrix}\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix} = \mathbf{x}^t\mathbf{y}.$$

◇

Observación. Sea $F : V \times V \to \mathbb{R}$ una forma bilineal sobre V. Si $\dim V = n$, entonces la matriz coordenada de F en cualquier base es una matriz cuadrada de dimensión $n \times n$.

Más aún, se tiene el siguiente resultado sobre matrices cuadradas:

Proposición 9.4. Sea V un espacio vectorial real de dimensión finita. Fijada una base $\mathcal{B}$ de V, toda forma bilineal $F : V \times V \to \mathbb{R}$ se puede representar mediante una matriz cuadrada. Recíprocamente, a toda matriz cuadrada le corresponde una única forma bilineal.

A través de la matriz coordenada de una forma bilineal F se pueden obtener propiedades de F. Concretamente:

Proposición 9.5. Sea V un espacio vectorial real de dimensión finita. Considerar una forma bilineal $F : V \times V \to \mathbb{R}$ y A su matriz coordenada en una base cualquiera de V. Entonces:

- F es simétrica $\iff$ A es una matriz simétrica.
- F es antisimétrica $\iff$ A es una matriz antisimétrica.

✓ **Ejemplo 6.** La forma bilineal del Ejemplo 3 es simétrica ya que su matriz coordenada (tanto en base $\mathcal{B}_1$ como en base $\mathcal{B}_2$) es simétrica.

◇

✓ **Ejemplo 7.** La matriz coordenada en base canónica de la forma bilineal del Ejemplo 1 es $A = \begin{pmatrix} 0 & 1 \\ k & 0 \end{pmatrix}$, de donde se deduce que F es simétrica si y solo si $k = 1$ y F es antisimétrica si y solo si $k = -1$.

◇

9.1.2. CAMBIO DE COORDENADAS EN FORMAS BILINEALES

Del mismo modo que ocurre con las aplicaciones lineales, es útil estudiar la relación entre las matrices coordenadas de una forma bilineal en bases distintas.

Teorema 9.6. Sea V un espacio vectorial real con $\dim V = n$ y $F : V \times V \to \mathbb{R}$ una forma bilineal sobre V. Sean $\mathcal{B}_1$ y $\mathcal{B}_2$ dos bases de V y sean A y B las matrices coordenadas de F respecto de las bases $\mathcal{B}_1$ y $\mathcal{B}_2$ respectivamente. Si P es la matriz de cambio de coordenadas de la base $\mathcal{B}_2$ a $\mathcal{B}_1$, entonces:

$$\boxed{B = P^t A P}$$

Demostración. Sean $\mathbf{u}$ y $\mathbf{v}$ dos vectores de V. Considerar sus coordenadas en las bases $\mathcal{B}_1$ y $\mathcal{B}_2$:

$$\mathcal{C}_{\mathcal{B}_1}(\mathbf{u}) = \mathbf{x}, \quad \mathcal{C}_{\mathcal{B}_2}(\mathbf{u}) = \widetilde{\mathbf{x}}.$$

$$\mathcal{C}_{\mathcal{B}_1}(\mathbf{v}) = \mathbf{y}, \quad \mathcal{C}_{\mathcal{B}_2}(\mathbf{v}) = \widetilde{\mathbf{y}}.$$

Sea P la matriz de cambio de coordenadas de la base $\mathcal{B}_2$ a $\mathcal{B}_1$. Entonces:

$$P\widetilde{\mathbf{x}} = \mathbf{x}, \quad P\widetilde{\mathbf{y}} = \mathbf{y}.$$

De este modo, la expresión coordenada de F es:

$$F(\mathbf{u}, \mathbf{v}) = \mathbf{x}^t A\mathbf{y} = (P\widetilde{\mathbf{x}})^t A(P\widetilde{\mathbf{y}}) = \widetilde{\mathbf{x}}^t (P^t AP)\widetilde{\mathbf{y}},$$

es decir, $P^t AP$ es la matriz coordenada de F en base $\mathcal{B}_2$. □

El teorema anterior indica que todas las matrices coordenadas de una misma forma bilineal en bases distintas están relacionadas por la expresión anterior.

Definición 9.7. Dos matrices $A, B \in \mathrm{Mat}_n(\mathbb{R})$ se dice que son **congruentes** si existe una matriz regular P tal que $B = P^t AP$.

Observación. No confundir matrices congruentes, $B = P^t AP$, con matrices semejantes, $B = P^{-1}AP$.

Las matrices congruentes se caracterizan por la siguiente propiedad:

Proposición 9.8. Dos matrices A y B son congruentes si y solo si ambas representan la misma forma bilineal en bases distintas.

✓ **Ejemplo 8.** Comprobar que las matrices A y B del Ejemplo 3 son congruentes.

Solución: La matriz P de cambio de coordenadas de $\mathcal{B}_2$ a $\mathcal{B}_1$ es $P = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$. Se tiene que:

$$P^t AP = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -3 \\ -3 & 1 \end{pmatrix} = B.$$

◇

Observación. Dos matrices congruentes $A, B \in \mathrm{Mat}_n(\mathbb{R})$ tienen el mismo rango ya que P y P^t son matrices regulares. Este hecho da sentido a la siguiente definición:

Definición 9.9. Sea $F : V \times V \to \mathbb{R}$ una forma bilineal sobre un espacio vectorial V de dimensión finita n:

- El **rango** de la forma bilineal F, denotado como $\operatorname{rg} F$, es el rango de una de sus matrices coordenadas.
- Una forma bilineal de rango máximo, es decir $\operatorname{rg} F = n$, se dice **regular**.
- Una forma bilineal se dice **singular** o **degenerada** si no es regular.

✓ **Ejemplo 9.** La forma bilineal del Ejemplo 3 es regular, ya que $\operatorname{rg} F = \operatorname{rg} A = 2$.

◇

9.1.3. DESCOMPOSICIÓN DE UNA FORMA BILINEAL

El siguiente resultado muestra que las formas bilineales simétricas y antisimétricas permiten construir cualquier forma bilineal.

Teorema 9.10. Sea V un espacio vectorial real de dimensión finita. Toda forma bilineal $F : V \times V \to \mathbb{R}$ se puede descomponer de manera única como suma de una forma bilineal simétrica y otra antisimétrica:

$$F = F_S + F_A.$$

Demostración. Dada la forma bilineal F se definen las siguientes formas bilineales:

$$F_S(\mathbf{u}, \mathbf{v}) = \frac{1}{2}\left(F(\mathbf{u}, \mathbf{v}) + F(\mathbf{v}, \mathbf{u})\right), \quad F_A(\mathbf{u}, \mathbf{v}) = \frac{1}{2}\left(F(\mathbf{u}, \mathbf{v}) - F(\mathbf{v}, \mathbf{u})\right).$$

Es inmediato comprobar que F_S es simétrica y F_A es antisimétrica. Además $F = F_S + F_A$.

Para demostrar la unicidad basta considerar dos posibles descomposiciones de la forma bilineal F: $F = F_S + F_A$ y $F = G_S + G_A$, donde F_S y G_S son formas simétricas mientras que F_A y G_A son antisimétricas. Igualando ambas expresiones:

$$F_S + F_A = G_S + G_A \Longleftrightarrow \underbrace{F_S - G_S}_{\text{simétrica}} = \underbrace{G_A - F_A}_{\text{antisimétrica}}.$$

Se obtiene que una forma simétrica es igual a una antisimétrica, por lo que debe ser la forma nula. Así: $F_S = G_S$ y $F_A = G_A$, con lo que se tiene la unicidad de la descomposición.

□

Corolario 9.11. Sea V un espacio vectorial real de dimensión finita. Sea F una forma bilineal sobre V. Se considera la descomposición $F = F_S + F_A$. Entonces:

- F es simétrica $\Longleftrightarrow F_A = 0$.
- F es antisimétrica $\Longleftrightarrow F_S = 0$.

✓ **Ejemplo 10.** Descomponer la forma bilineal dada en el Ejemplo 1, en su parte simétrica y antisimétrica.

Solución: La expresión coordenada de F en base canónica es $F(\mathbf{u}, \mathbf{v}) = x_1y_2 + kx_2y_1$. Aplicando el teorema anterior,

$$\begin{aligned} F_S(\mathbf{u}, \mathbf{v}) &= \tfrac{1}{2}\left(F(\mathbf{u}, \mathbf{v}) + F(\mathbf{v}, \mathbf{u})\right) = \tfrac{1}{2}(x_1y_2 + kx_2y_1 + x_2y_1 + kx_1y_2) \\ &= \tfrac{1+k}{2}\left(x_1y_2 + x_2y_1\right). \\ F_A(\mathbf{u}, \mathbf{v}) &= \tfrac{1}{2}\left(F(\mathbf{u}, \mathbf{v}) - F(\mathbf{v}, \mathbf{u})\right) = \tfrac{1}{2}(x_1y_2 + kx_2y_1 - x_2y_1 - kx_1y_2) \\ &= \tfrac{1-k}{2}\left(x_1y_2 - x_2y_1\right). \end{aligned}$$

Se observa que $F = F_S$ (F es simétrica) si y solo si $F_A = 0$, es decir, $k = 1$. Análogamente, $F = F_A$ (F es antisimétrica) si y solo si $F_S = 0$, es decir, $k = -1$.

◇

9.2. FORMAS CUADRÁTICAS

Esta sección se dedica al estudio de las formas bilineales simétricas, ya que estas generan un tipo especial de formas muy importante en geometría, denominadas *formas cuadráticas.*

Definición 9.12. Sea V un espacio vectorial real y $F : V \times V \to \mathbb{R}$ una forma bilineal sobre V. La **forma cuadrática** asociada a F es la aplicación

$$\begin{array}{rcl} Q : V & \longrightarrow & \mathbb{R} \\ \mathbf{v} & \mapsto & Q(\mathbf{v}) = F(\mathbf{v}, \mathbf{v}). \end{array}$$

⇝ **Ejercicio.** Sea F una forma bilineal sobre V y sea Q su forma cuadrática asociada. Comprobar:

$$Q(\mathbf{u} + \mathbf{v}) = Q(\mathbf{u}) + Q(\mathbf{v}) + F(\mathbf{u}, \mathbf{v}) + F(\mathbf{v}, \mathbf{u}).$$

✓ **Ejemplo 11.** Hallar la forma cuadrática asociada a la forma bilineal dada en el Ejemplo 1.

Solución: La expresión coordenada de F en base canónica es $F(\mathbf{u}, \mathbf{v}) = x_1y_2 + kx_2y_1$, por lo que la forma cuadrática es:

$$Q(\mathbf{v}) = F(\mathbf{v}, \mathbf{v}) = x_1x_2 + kx_2x_1 = (k + 1)x_1x_2.$$

◇

Directamente de la definición se sigue que:

Propiedades.

1. $Q(\mathbf{0}) = 0$.
2. $Q(\lambda\mathbf{v}) = \lambda^2 Q(\mathbf{v}), \quad \forall \mathbf{v} \in V, \quad \forall \lambda \in \mathbb{R}$. (Notar que una forma cuadrática no es una aplicación lineal).

Observaciones.

- Formas bilineales distintas pueden tener asociada la misma forma cuadrática.
- De todas las formas bilineales que tienen asociada la misma forma cuadrática Q, solo una es simétrica. Esta forma simétrica, denotada por F_p, recibe un nombre especial (ver siguiente definición).

Definición 9.13. Sea $Q : V \to \mathbb{R}$ una forma cuadrática sobre V. La forma bilineal simétrica asociada a Q se denomina **forma polar** asociada a Q. Su expresión es:

$$F_p(\mathbf{u}, \mathbf{v}) = \frac{Q(\mathbf{u} + \mathbf{v}) - Q(\mathbf{u}) - Q(\mathbf{v})}{2}.$$

✓ **Ejemplo 12.** Considerar las formas bilineales representadas (en la base canónica) por las matrices

$$A=\begin{pmatrix}0&2\\0&0\end{pmatrix},\qquad B=\begin{pmatrix}0&0\\2&0\end{pmatrix}.$$

Comprobar que ambas tienen la misma forma cuadrática asociada y calcular la forma polar de esta.

Solución: Las matrices A y B no son simétricas, por lo que las formas bilineales que definen, F_A y F_B, tampoco lo son. Las formas cuadráticas asociadas son, respectivamente:

$$Q_A(\mathbf{v})=F_A(\mathbf{v},\mathbf{v})=F_A((x_1,x_2),(x_1,x_2))=\begin{pmatrix}x_1&x_2\end{pmatrix}\begin{pmatrix}0&2\\0&0\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}=2x_1x_2.$$

$$Q_B(\mathbf{v})=F_B(\mathbf{v},\mathbf{v})=F_B((x_1,x_2),(x_1,x_2))=\begin{pmatrix}x_1&x_2\end{pmatrix}\begin{pmatrix}0&0\\2&0\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}=2x_1x_2.$$

En ambos casos se tiene la misma forma cuadrática:

$$\begin{array}{cccc}Q: & \mathbb{R}^2 & \longrightarrow & \mathbb{R}\\ & (x_1,x_2) & \mapsto & 2x_1x_2.\end{array}$$

La forma polar de Q viene dada por

$$F_p((x_1,x_2),(y_1,y_2))=\frac{2(x_1+y_1)(x_2+y_2)-2x_1x_2-2y_1y_2}{2}=x_1y_2+x_2y_1,$$

cuya matriz coordenada en base canónica es $C=\begin{pmatrix}0&1\\1&0\end{pmatrix}$. Observar que C es una matriz simétrica, como corresponde a una forma bilineal simétrica.

◇

Proposición 9.14. Sea F una forma bilineal, Q su forma cuadrática asociada y F_p la forma polar asociada a Q. Si F se descompone en su parte simétrica y antisimétrica, se tiene que

$$F_p=F_S,$$

es decir, la forma polar es la parte simétrica de la forma bilineal F. Más aún:

- F es simétrica $\Longleftrightarrow F_p=F$.
- F es antisimétrica $\Longleftrightarrow F_p=0$. (En este caso, $Q=0$.)

✓ **Ejemplo 13.** Considerar la forma cuadrática $Q(\mathbf{v})=(k+1)x_1x_2$ asociada a la forma bilineal del Ejemplo 1. Hallar su forma polar.

Solución: Sean $\mathbf{u}=(x_1,x_2)$ y $\mathbf{v}=(y_1,y_2)$ dos vectores de $\mathbb{R}^2$. Entonces:

$$\begin{aligned}F_p(\mathbf{u},\mathbf{v}) &= \tfrac{Q(\mathbf{u}+\mathbf{v})-Q(\mathbf{u})-Q(\mathbf{v})}{2}=\tfrac{1+k}{2}\left((x_1+y_1)(x_2+y_2)-x_1x_2-y_1y_2\right)\\ &= \tfrac{1+k}{2}\left(x_1y_2+x_2y_1\right).\end{aligned}$$

Se observa que, efectivamente, F_p coincide con F_S, calculada en el Ejemplo 10.

◇

9.2.1. EXPRESIÓN COORDENADA DE UNA FORMA CUADRÁTICA

Definición 9.15. Sea V un espacio vectorial real de dimensión finita n y sea $Q : V \to \mathbb{R}$ una forma cuadrática. Considerar una base $\mathcal{B}$ de V. La **matriz coordenada** de Q en la base $\mathcal{B}$ es la matriz coordenada A de la forma bilineal polar de Q, F_p.

La expresión matricial de la forma cuadrática Q es

$$Q(\mathbf{v}) = \mathbf{x}^t A\mathbf{x},$$

donde $A \in \mathrm{Mat}_n(\mathbb{R})$ es una matriz simétrica y $\mathcal{C}_{\mathcal{B}}(\mathbf{v}) = \mathbf{x} = (x_1, \ldots, x_n)$ son las coordenadas del vector $\mathbf{v}$ en la base $\mathcal{B}$:

$$Q(\mathbf{v}) = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{1n} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \Longrightarrow$$

$$\boxed{Q(\mathbf{v}) = \sum_{i=1}^{n} a_{ii}\, x_i^2 + \sum_{i<j} 2a_{ij}\, x_i x_j}$$

La expresión anterior es un polinomio homogéneo de grado 2 en las coordenadas x_i con coeficientes reales.

Por otra parte, si se conoce la expresión polinómica de la forma cuadrática como polinomio homogéneo de grado 2, debidamente ordenado, $Q(\mathbf{v}) = \sum_{i \leq j} c_{ij} x_i x_j$, la obtención de la matriz coordenada de Q, $A = (a_{ij})$, es inmediata:

- $a_{ii} = c_{ii}$, es decir, el coeficiente de x_i^2 en $Q(\mathbf{v})$.
- $a_{ij} = \frac{1}{2} c_{ij}$, es decir, $\frac{1}{2}$ del coeficiente de $x_i x_j$ en $Q(\mathbf{v})$, si $i \neq j$.

La expresión polinómica de Q recibe el nombre de *expresión coordenada.*

✓ **Ejemplo 14.** Dada la forma cuadrática Q, hallar su expresión matricial:

$$\begin{array}{rccl} Q: & \mathbb{R}^3 & \longrightarrow & \mathbb{R} \\ & (x_1, x_2, x_3) & \mapsto & x_1^2 + 3x_2^2 - 2x_3^2 + 2x_1x_2 + x_2x_3. \end{array}$$

Solución: La matriz $A = (a_{ij}) \in \mathrm{Mat}_3(\mathbb{R})$ se construye como sigue:
Los elementos de la diagonal se obtienen directamente:

- a_{11} es el coeficiente de $x_1^2 \Longrightarrow a_{11} = 1$.
- a_{22} es el coeficiente de $x_2^2 \Longrightarrow a_{22} = 3$.

- a_{33} es el coeficiente de $x_3^2 \Longrightarrow a_{33} = -2$.

Los elementos que no están en la diagonal:

- $a_{12} = a_{21}$ es el coeficiente de x_{12} dividido entre dos $\Longrightarrow a_{12} = a_{21} = 1$.
- $a_{13} = a_{31}$ es el coeficiente de x_{13} dividido entre dos $\Longrightarrow a_{12} = a_{21} = 0$.
- $a_{23} = a_{32}$ es el coeficiente de x_{23} dividido entre dos $\Longrightarrow a_{12} = a_{21} = 1/2$.

Así, la expresión matricial de Q es:

$$Q(x_1, x_2, x_3) = \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \underbrace{\begin{pmatrix} 1 & 1 & 0 \\ 1 & 3 & 1/2 \\ 0 & 1/2 & -2 \end{pmatrix}}_{A} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

◇

9.2.2. CAMBIO DE COORDENADAS EN FORMAS CUADRÁTICAS

La expresión coordenada de una forma cuadrática $Q : V \to \mathbb{R}$ depende de la base elegida en el espacio vectorial V. La relación entre las expresiones de Q al considerar dos bases distintas de V viene dada en el siguiente teorema.

Teorema 9.16. Sea $Q : V \to \mathbb{R}$ una forma cuadrática sobre un espacio vectorial V de dimensión finita n. Sean $\mathcal{B}_1$ y $\mathcal{B}_2$ dos bases de V y sean A y B las matrices coordenadas de Q respecto de las bases $\mathcal{B}_1$ y $\mathcal{B}_2$ respectivamente. Si P es la matriz de cambio de coordenadas de la base $\mathcal{B}_2$ a $\mathcal{B}_1$, entonces:

$$\boxed{B = P^t A P}$$

Demostración. Si las coordenadas de un vector $\mathbf{v} \in V$ en las bases $\mathcal{B}_1$ y $\mathcal{B}_2$ vienen dadas por:

$$\mathbf{x} = \mathcal{C}_{\mathcal{B}_1}(\mathbf{v}), \quad \widetilde{\mathbf{x}} = \mathcal{C}_{\mathcal{B}_2}(\mathbf{v}),$$

y su relación a través de la matriz P es $\mathbf{x} = P\widetilde{\mathbf{x}}$, se tiene que:

$$Q(\mathbf{v}) = \mathbf{x}^t A \mathbf{x} = (P\widetilde{\mathbf{x}})^t A (P\widetilde{\mathbf{x}}) = \widetilde{\mathbf{x}}^t (P^t A P)\widetilde{\mathbf{x}},$$

es decir, $P^t AP$ es la matriz coordenada de Q en base $\mathcal{B}_2$.

□

Observaciones.

- Como A es una matriz simétrica, también $P^t AP$ es simétrica.
- Dos matrices simétricas congruentes representan una misma forma cuadrática en bases distintas.

Definición 9.17. Sea V un espacio vectorial de dimensión finita y $Q : V \to \mathbb{R}$ una forma cuadrática. Se llama **rango** de Q al rango de una cualquiera de sus matrices coordenadas.

✓ **Ejemplo 15.** El rango de la forma cuadrática del Ejemplo 14 es 3, ya que su matriz coordenada $A \in \mathrm{Mat}_3(\mathbb{R})$ es regular, es decir, su rango es 3.

◇

9.3. DIAGONALIZACIÓN DE FORMAS CUADRÁTICAS

Dada una forma cuadrática $Q : V \to \mathbb{R}$ se pretende encontrar una expresión coordenada de Q lo más sencilla posible; equivalentemente, encontrar una base de V en la que la matriz coordenada de Q sea diagonal. Esto es siempre posible debido al siguiente resultado:

Teorema 9.18. Sea $A \in \mathrm{Mat}_n(\mathbb{R})$ una matriz simétrica. Entonces, A es congruente a una matriz diagonal.

Recordar que toda matriz simétrica es ortogonalmente diagonalizable y, por tanto, congruente a una matriz diagonal.

En la siguiente definición se presenta la expresión coordenada de una forma cuadrática cuya matriz coordenada es diagonal.

Definición 9.19. Sea $Q : V \to \mathbb{R}$ una forma cuadrática sobre un espacio vectorial real V de dimensión finita n. Se llama **expresión canónica** de la forma cuadrática a cualquier expresión del tipo:

$$Q(\mathbf{v}) = \sum_{i=1}^{n} k_i\, x_i^2,$$

siendo $(x_1, \ldots, x_n)$ las coordenadas de $\mathbf{v}$ en una determinada base $\mathcal{D}$.

El objetivo de esta sección es buscar expresiones canónicas de formas cuadráticas donde los k_i tomen el valor $1, -1, 0$.

Definición 9.20. Sea $\mathcal{D}$ una base de V en la que la expresión de la forma cuadrática Q es canónica. La base $\mathcal{D}$ recibe el nombre de **base conjugada**.

Observación. La matriz coordenada de Q en una base conjugada es diagonal pues en la expresión canónica no aparecen términos cruzados del tipo $x_i x_j$ con $i \neq j$.

Antes de diagonalizar una forma cuadrática Q, es conveniente clasificarla atendiendo a la siguiente definición:

Definición 9.21. Sea $Q : V \to \mathbb{R}$ una forma cuadrática.

- Q se dice **definida positiva** si $Q(\mathbf{v}) > 0$, $\forall \mathbf{v} \in V$, $\mathbf{v} \neq \mathbf{0}$.
- Q se dice **definida negativa** si $Q(\mathbf{v}) < 0$, $\forall \mathbf{v} \in V$, $\mathbf{v} \neq \mathbf{0}$.
- Q se dice **semidefinida positiva** si $Q(\mathbf{v}) \geq 0$, $\forall \mathbf{v} \in V$, $\mathbf{v} \neq \mathbf{0}$, (en particular, existe algún $\mathbf{v} \neq \mathbf{0}$ tal que $Q(\mathbf{v}) = 0$).
- Q se dice **semidefinida negativa** si $Q(\mathbf{v}) \leq 0$, $\forall \mathbf{v} \in V$, $\mathbf{v} \neq \mathbf{0}$, (en particular, existe algún $\mathbf{v} \neq \mathbf{0}$ tal que $Q(\mathbf{v}) = 0$).
- Q se dice **indefinida** en otro caso, es decir, $\exists\, \mathbf{u}, \mathbf{v} \in V$ tales que $Q(\mathbf{u}) > 0$ y $Q(\mathbf{v}) < 0$.

Usando la relación entre matrices simétricas y formas cuadráticas, tiene sentido definir los conceptos anteriores para el caso de matrices.

Definición 9.22. Una matriz $A \in \mathrm{Mat}_n(\mathbb{R})$ simétrica real se dice de uno de los tipos anteriores si la forma cuadrática que define $Q(\mathbf{v}) = \mathbf{x}^t A \mathbf{x}$ es de ese tipo.

✓ **Ejemplo 16.** Comprobar que la matriz $A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 3 & 1/2 \\ 0 & 1/2 & -2 \end{pmatrix}$ es indefinida.

Solución: Observar que la forma cuadrática que define la matriz A es la dada en el Ejemplo 14. Para comprobar que A es indefinida basta ver que lo es la forma cuadrática asociada Q, encontrando dos vectores $\mathbf{u} \neq \mathbf{0}$ y $\mathbf{v} \neq \mathbf{0}$ tales que $Q(\mathbf{u}) > 0$ y $Q(\mathbf{v}) < 0$.

Las comprobaciones anteriores se pueden realizar con la expresión analítica de la forma cuadrática o bien con la expresión matricial.

Por ejemplo, tomando $\mathbf{u} = (1, 0, 0)$ y $\mathbf{v} = (0, 0, 1)$, entonces:

$$Q(\mathbf{u}) = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 1 & 3 & 1/2 \\ 0 & 1/2 & -2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = 1 > 0,$$

$$Q(\mathbf{v}) = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 1 & 3 & 1/2 \\ 0 & 1/2 & -2 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = -2 < 0.$$

◇

En general no es sencillo clasificar matrices o formas cuadráticas usando las definiciones. El siguiente resultado es fundamental para resolver este problema.

Teorema 9.23 (Ley de inercia de Sylvester)**.** Sea Q una forma cuadrática sobre un espacio vectorial real V de dimensión finita n. Suponer que existen dos bases de V, $\mathcal{B}_1$ y $\mathcal{B}_2$, en las que la expresión de Q es canónica, es decir, $\mathcal{B}_1$ y $\mathcal{B}_2$ son bases conjugadas. Entonces, el número de términos que aparecen con coeficientes positivos en la expresión de Q, así como el número de términos que aparecen con coeficientes negativos, es el mismo en ambos casos.

El resultado anterior permite dar la siguiente definición:

Definición 9.24. El número de términos positivos que aparece en una expresión canónica cualquiera de Q se denomina **signatura** de la forma cuadrática.

Proposición 9.25. Sea V un espacio vectorial real de dimensión finita y $Q : V \to \mathbb{R}$ una forma cuadrática. Sea D una matriz coordenada diagonal de Q. Suponiendo que se ordenan los valores de la diagonal en positivos, negativos y ceros:

$$D = \operatorname{diag}(\underbrace{\alpha_1, \ldots, \alpha_s}_{>0}, \underbrace{\alpha_{s+1}, \ldots, \alpha_r}_{<0}, 0, \ldots, 0).$$

Entonces:

- $r = \operatorname{rg} Q =$ número de elementos no nulos en la diagonal.
- $s = \operatorname{sign} Q =$ número de elementos positivos en la diagonal.

Consecuencia. Del resultado anterior se deduce que:

$$\boxed{\operatorname{sign} Q \leq \operatorname{rg} Q}$$

El rango y la signatura de una forma cuadrática permiten su clasificación:

Teorema 9.26. Sea V un espacio vectorial real de dimensión finita y $Q : V \to \mathbb{R}$ una forma cuadrática sobre V. Entonces:

- Q es definida positiva si y solo si $\operatorname{rg} Q = \operatorname{sign} Q = \dim V$.
- Q es definida negativa si y solo si $\operatorname{rg} Q = \dim V$ y $\operatorname{sign} Q = 0$.
- Q es semidefinida positiva si y solo si $\operatorname{rg} Q = \operatorname{sign} Q < \dim V$.
- Q es semidefinida negativa si y solo si $\operatorname{rg} Q < \dim V$ y $\operatorname{sign} Q = 0$.
- Q es indefinida si $0 < \operatorname{sign} Q < \operatorname{rg} Q$.

En forma de tabla, la clasificación es la siguiente:

Signatura	**Rango**	**Clasificación**
$\operatorname{sign} Q = 0$	$\operatorname{rg} Q = \dim V$	Definida negativa
	$\operatorname{rg} Q < \dim V$	Semidefinida negativa
$0 < \operatorname{sign} Q < \operatorname{rg} Q$		Indefinida
$\operatorname{sign} Q = \operatorname{rg} Q$	$\operatorname{rg} Q < \dim V$	Semidefinida positiva
	$\operatorname{rg} Q = \dim V$	Definida positiva

✓ **Ejemplo 17.** Determinar el rango y la signatura de la matriz $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & k \end{pmatrix}$ en función de los valores del parámetro k y clasificar la forma cuadrática que define.

Solución: Aplicando la Proposición 9.25 se deduce que:

$$\begin{cases} k < 0 : & \operatorname{rg} A = 3,\ \operatorname{sign} A = 1, \\ k = 0 : & \operatorname{rg} A = 2,\ \operatorname{sign} A = 1, \\ k > 0 : & \operatorname{rg} A = 3,\ \operatorname{sign} A = 2. \end{cases}$$

Utilizando el Teorema 9.26 se tiene que para todo valor de k la matriz A, y por tanto la forma cuadrática asociada, es indefinida.

◇

A continuación se pretende encontrar una base conjugada de V, es decir, una base en la que la forma cuadrática se represente mediante una matriz diagonal o, equivalentemente, la expresión de Q sea canónica. Para ello, se utilizan tres procedimientos:

- Método matricial basado en la diagonalización ortogonal de matrices estudiada en el Capítulo 8.
- Método polinómico o de Lagrange basado en el proceso de completar cuadrados.
- Representación matricial del método de Lagrange utilizando la relación de congruencia entre matrices simétricas.

9.3.1. DIAGONALIZACIÓN POR VALORES PROPIOS

Teorema 9.27 (Teorema de los ejes principales)**.** Sea $Q : V \to \mathbb{R}$ una forma cuadrática sobre un espacio vectorial real de dimensión finita n y sea $A \in \mathrm{Mat}_n(\mathbb{R})$ su matriz coordenada en una cierta base $\mathcal{B}$. Entonces, existe una base $\mathcal{D}$ de V en la cual la matriz coordenada de Q es diagonal, con los valores propios de A, λ_i, en la diagonal (repetidos según su multiplicidad). La expresión coordenada de Q en base $\mathcal{D}$ es:

$$Q(\mathbf{v}) = \mathbf{y}^t D \mathbf{y} = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2.$$

Además:

- $\operatorname{rg} Q =$ número de valores propios no nulos (contados con multiplicidades).
- $\operatorname{sign} Q =$ número de valores propios positivos (contados con multiplicidades).

Demostración. Recordar en primer lugar el Teorema 8.20 según el cual toda matriz simétrica es ortogonalmente diagonalizable. Aplicando la diagonalización ortogonal a la matriz A, existe una base $\mathcal{D}$ de V ortonormal formada por vectores propios de modo que $P^{-1}AP = D$, donde P es la matriz de cambio de coordenadas entre la base de vectores propios y la base ortonormal inicial $\mathcal{B}$. Por ser P una matriz ortogonal, $P^{-1} = P^t$, por lo que $D = P^tAP$. Si la primera base es la base canónica de V, entonces las columnas de P son los vectores propios de A y reciben el nombre de **ejes principales** de la forma cuadrática. La matriz D es diagonal y está formada por los valores propios de la matriz A.

□

✓ **Ejemplo 18.** Considerar la siguiente forma cuadrática sobre $\mathbb{R}^3$:

$$Q(\mathbf{v}) = x_1^2 + x_2^2 + x_3^2 - 2x_1x_2 - 2x_1x_3 - 2x_2x_3,$$

expresada en términos de la base canónica. Aplicar el teorema anterior para encontrar la diagonalización de la forma Q, indicando además la matriz P y los ejes principales. Clasificar Q.

Solución: En términos de la base canónica de $\mathbb{R}^3$, la matriz A asociada a la forma cuadrática Q es

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

Los valores propios de A son $\lambda_1 = 2$, con multiplicidad 2 y $\lambda_2 = -1$ con multiplicidad 1. La diagonalización ortogonal de esta matriz aparece en el Ejemplo 8 del Capítulo 8 salvo el orden elegido:

$$\underbrace{\begin{pmatrix} 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ -1/\sqrt{6} & 2/\sqrt{6} & -1/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \end{pmatrix}}_{P^t} \underbrace{\begin{pmatrix} 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{3} \\ -1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \end{pmatrix}}_{P} = \underbrace{\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix}}_{D}.$$

En términos de la base

$$\mathcal{D} = \left\{ \left(\tfrac{1}{\sqrt{2}}, 0, \tfrac{1}{\sqrt{2}} \right), \left(\tfrac{-1}{\sqrt{6}}, \tfrac{2}{\sqrt{6}}, \tfrac{-1}{\sqrt{6}} \right), \left(\tfrac{1}{\sqrt{3}}, \tfrac{1}{\sqrt{3}}, \tfrac{1}{\sqrt{3}} \right) \right\},$$

se obtiene la siguiente expresión canónica de Q:

$$Q(\mathbf{v}) = 2y_1^2 + 2y_2^2 - y_3^2,$$

donde $\mathcal{C}_{\mathcal{D}}(\mathbf{v}) = (y_1, y_2, y_3)$. Además, los vectores de la base $\mathcal{D}$ constituyen los ejes principales. Por último, el rango de Q es 3 (pues ningún valor propio es nulo) y la signatura es 2 (porque hay dos valores propios positivos), por lo que la forma cuadrática Q es indefinida.

◇

Observación. El teorema indica la existencia de una base en la que la matriz coordenada es diagonal con valores propios en la diagonal. Sin embargo, pueden existir otras bases conjugadas en las que la matriz coordenada sea diagonal y no aparezcan los valores propios en ella, como se ve en el siguiente corolario.

Corolario 9.28. Sea $Q : V \to \mathbb{R}$ una forma cuadrática. Entonces existe una base de V en la cual la matriz coordenada de Q es diagonal con 1's, -1's y 0's:

$$\operatorname{diag}\,(\overbrace{\underbrace{1,\ldots,1}_{s}, -1,\ldots,-1}^{r}, 0,\ldots,0),$$

es decir, existe una base de V en la que la forma cuadrática se expresa de manera canónica como:

$$Q(\mathbf{v}) = y_1^2 + y_2^2 + \ldots + y_s^2 - y_{s+1}^2 - \ldots - y_r^2,$$

donde:

- $r = \operatorname{rg} Q =$ número de 1's y -1's en la diagonal.
- $s = \operatorname{sign} Q =$ número de 1's en la diagonal.

Demostración. Por el Teorema 9.27, dada A una matriz coordenada de Q en una base $\mathcal{B}$, existe una base conjugada $\mathcal{D}$ de V de modo que $P^tAP = D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. Sin pérdida de generalidad, se puede suponer que los valores propios aparecen en la diagonal ordenados del siguiente modo: $\lambda_1, \ldots, \lambda_s$ son positivos; $\lambda_{s+1}, \ldots, \lambda_r$ son negativos; $\lambda_{r+1} = \ldots = \lambda_n = 0$.

Considerar la matriz diagonal $U \in \operatorname{Mat}_n(\mathbb{R})$:

$$U = \operatorname{diag}\left(\frac{1}{\sqrt{|\lambda_1|}}, \ldots, \frac{1}{\sqrt{|\lambda_r|}}, 1, \ldots, 1 \right).$$

Si se multiplica U a derecha y a izquierda de D, se obtiene una matriz diagonal con 1's, -1's y 0's, concretamente:

$$UDU = \widetilde{D} = \text{diag}\,(\overbrace{\underbrace{1,\ldots,1}_{s}, -1,\ldots,-1}^{r}, 0,\ldots,0).$$

Ahora,

$$P^tAP = D \Longrightarrow U(P^tAP)U = UDU \Longrightarrow (PU)^tA(PU) = \widetilde{D}.$$

Llamando $R = PU$, se tiene que $R^tAR = \widetilde{D}$. Las columnas de la matriz R proporcionan la base de vectores en los que la matriz coordenada de Q es $\widetilde{D}$.

$\square$

✓ **Ejemplo 19.** Hallar una expresión canónica con coeficientes 1's, -1's y 0's de la forma cuadrática dada en el Ejemplo 18.

Solución: La matriz diagonal hallada en el ejemplo anterior es $D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix}$, por lo que la matriz U necesaria es: $U = \begin{pmatrix} 1/\sqrt{2} & 0 & 0 \\ 0 & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

Las matrices R y $\widetilde{D}$ vienen dadas por:

$$R = PU = \begin{pmatrix} 1/2 & -1/\sqrt{12} & 1/\sqrt{3} \\ 0 & 2/\sqrt{12} & 1/\sqrt{3} \\ -1/2 & -1/\sqrt{12} & 1/\sqrt{3} \end{pmatrix}, \quad \widetilde{D} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Por trabajar en el espacio vectorial $V = \mathbb{R}^3$ y ser la base $\mathcal{B}$ la base canónica, las columnas de la matriz R constituyen la base buscada. Así, en base

$$\widetilde{\mathcal{D}} = \left\{ \left(\tfrac{1}{2}, 0, \tfrac{-1}{2}\right), \left(\tfrac{-1}{\sqrt{12}}, \tfrac{2}{\sqrt{12}}, \tfrac{-1}{\sqrt{12}}\right), \left(\tfrac{1}{\sqrt{3}}, \tfrac{1}{\sqrt{3}}, \tfrac{1}{\sqrt{3}}\right) \right\}$$

la expresión coordenada de la forma cuadrática es

$$Q(\mathbf{v}) = y_1^2 + y_2^2 - y_3^2,$$

donde $\mathcal{C}_{\widetilde{\mathcal{D}}}(\mathbf{v}) = (y_1, y_2, y_3)$. De esta expresión se deduce inmediatamente que el rango de Q es 3 (aparecen y_1, y_2, y_3 en la expresión de Q) y la signatura es 2 (pues son dos variables las que llevan el signo positivo, y_1, y_2).

◇

9.3.2. DIAGONALIZACIÓN POR EL MÉTODO DE LAGRANGE

En la práctica no siempre resulta útil obtener una expresión canónica de una forma cuadrática realizando la diagonalización ortogonal de una de sus matrices coordenadas, ya que puede ser difícil el cálculo de los valores propios de la misma. Si se trabaja con la expresión coordenada de la forma cuadrática se puede utilizar el siguiente procedimiento.

Método de Lagrange

Considerar la expresión general de una forma cuadrática Q:

$$Q(\mathbf{v}) = \sum_{i \leq j} a_{ij} x_i x_j .$$

Para obtener una expresión canónica de la forma cuadrática, o equivalentemente, para encontrar una base conjugada de V se procede utilizando los cambios de variables que a continuación se detallan:

1. Si para algún índice i se tiene que $a_{ii} \neq 0$,es decir, si aparece alguna variable al cuadrado, x_i^2 , se completan cuadrados con todos los términos que contengan x_i para obtener

   $$Q(\mathbf{v}) = a_{ii} \left(x_i + \frac{1}{2} \sum_{j \neq i} \frac{a_{ij}}{a_{ii}} x_j \right)^2 + Q_1(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n),$$

   donde Q_1 es una nueva forma cuadrática con $n-1$ variables, con la que se inicia de nuevo el método. El cambio de variables que se utiliza es:

   $$\begin{cases} y_i = x_i + \frac{1}{2} \displaystyle\sum_{j \neq i} \frac{a_{ij}}{a_{ii}} x_j \\ y_j = x_j, \text{ para } j \neq i. \end{cases}$$

2. Si $a_{11} = \ldots = a_{nn} = 0$, es decir, si no aparece ninguna variable al cuadrado, se elige un coeficiente $a_{ij} \neq 0$. Haciendo el cambio de variables

   $$\begin{cases} x_i = y_i + y_j \\ x_j = y_i - y_j \\ x_k = y_k \text{ para } k \neq i, j, \end{cases}$$

   la expresión $Q(y_1, \ldots, y_n)$ contiene términos cuadráticos, por lo que se vuelve al caso anterior.

Hay que aplicar tantos cambios de los tipos anteriores como sea necesario y en el orden adecuado.

Para obtener la base conjugada de V es necesario concatenar todos los cambios de variables que se han realizado en el proceso anterior, obteniendo una relación matricial entre las variables iniciales y finales, de la cual se extrae la base buscada. En el ejemplo siguiente se explican los pasos a seguir con más detalle.

✓ **Ejemplo 20.** Diagonalizar según el método de Lagrange la forma cuadrática dada en el Ejemplo 18.

Solución: La expresión en base canónica de Q es la siguiente:

$$Q(\mathbf{v}) = x_1^2 + x_2^2 + x_3^2 - 2x_1x_2 - 2x_1x_3 - 2x_2x_3.$$

Como el coeficiente que acompaña a x_1^2 es no nulo, $a_{11} = 1 \neq 0$, se comienza el método por el paso 1, agrupando todos los términos que contienen x_1:

$$Q(\mathbf{v}) = (x_1^2 - 2x_1x_2 - 2x_1x_3) + x_2^2 + x_3^2 - 2x_2x_3.$$

A continuación, se completan cuadrados en la parte que contiene x_1:

$$x_1^2 - 2x_1x_2 - 2x_1x_3 = (x_1 - x_2 - x_3)^2 - (x_2^2 + x_3^2 + 2x_2x_3).$$

Sustituyendo la expresión anterior en $Q(\mathbf{v})$:

$$\begin{aligned} Q(\mathbf{v}) &= (x_1^2 - 2x_1x_2 - 2x_1x_3) + x_2^2 + x_3^2 - 2x_2x_3 \\ &= [(x_1 - x_2 - x_3)^2 - (x_2^2 + x_3^2 + 2x_2x_3)] + x_2^2 + x_3^2 - 2x_2x_3 \\ &= (x_1 - x_2 - x_3)^2 - 4x_2x_3. \end{aligned}$$

Con el cambio

$$\begin{cases} y_1 = x_1 - x_2 - x_3 \\ y_2 = x_2 \\ y_3 = x_3 \end{cases}$$

la nueva expresión coordenada de Q es

$$Q(\mathbf{v}) = y_1^2 - 4y_2y_3 = y_1^2 + Q_1(y_2, y_3).$$

La nueva forma cuadrática Q_1 se encuentra en el paso 2 del método de Lagrange por lo que se plantea el siguiente cambio, donde la variable y_1 simplemente se renombra:

$$\begin{cases} y_1 = u_1 \\ y_2 = u_2 + u_3 \\ y_3 = u_2 - u_3. \end{cases}$$

Se obtiene:

$$Q(\mathbf{v}) = u_1^2 - 4u_2^2 + 4u_3^2.$$

Por último si se reordenan y normalizan las variables mediante

$$\begin{cases} w_1 = u_1 \\ w_2 = 2u_3 \\ w_3 = 2u_2 \end{cases}$$

se concluye que

$$Q(\mathbf{v}) = w_1^2 + w_2^2 - w_3^2,$$

de modo que el rango de Q es 3 y la signatura 2, como ya se ha visto en ejemplos anteriores.

Finalmente, para obtener la matriz P cumpliendo $P^tAP = D$ es necesario expresar las primeras variables en función de las últimas, es decir $\mathbf{x} = P\mathbf{w}$:

$$\begin{cases} x_1 = y_1 + y_2 + y_3 = u_1 + 2u_2 = w_1 + w_3 \\ x_2 = y_2 = u_2 + u_3 = \frac{1}{2}(w_2 + w_3) \\ x_3 = y_3 = u_2 - u_3 = \frac{1}{2}(w_3 - w_2) \end{cases} \Longrightarrow \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1/2 & 1/2 \\ 0 & -1/2 & 1/2 \end{pmatrix}}_{P} \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix}.$$

Por trabajar en $\mathbb{R}^3$ con base canónica, la base que diagonaliza la forma cuadrática está formada por las columnas de P: $\{(1,0,0)\,(0,\frac{1}{2},\frac{-1}{2}),\ (1,\frac{1}{2},\frac{1}{2})\}$. La matriz coordenada de Q es

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

◇

9.3.3. DIAGONALIZACIÓN POR CONGRUENCIA

Este método está basado en las transformaciones elementales, aunque se puede interpretar como una expresión matricial del método de Lagrange. Dadas dos matrices congruentes A y B, existe una matriz regular P de modo que $B = P^tAP$. Por ser P^t regular, es producto de matrices elementales, $P^t = P_k \cdots P_1$. De esta manera:

$$B = P^tAP = (P_k \cdots P_1)A(P_k \cdots P_1)^t = (P_k \cdots P_1)A(P_1^t \cdots P_k^t).$$

Recordando el Capítulo 2, si P_i es una matriz elemental, P_i^t es la matriz que representa la misma transformación elemental por columnas. Por tanto, $P_iAP_i^t$ es la matriz que resulta de A tras aplicar la misma operación elemental por filas y por columnas.

Proposición 9.29. Dos matrices cuadradas A y B son congruentes si y solo si una puede obtenerse a partir de la otra haciendo transformaciones elementales, las mismas por filas que por columnas.

Sea $Q : V \to \mathbb{R}$ una forma cuadrática y sea A su matriz coordenada en la base canónica de V. Para encontrar una matriz diagonal congruente con A se sigue el siguiente esquema:

$$\left(\begin{array}{c|c} I_n & A \\ \hline & I_n \end{array}\right) \xrightarrow{\text{mismas operaciones por filas y columnas}} \left(\begin{array}{c|c} P^t & D \\ \hline & P \end{array}\right).$$

Observar que las operaciones por filas afectan a $(I_n \,|\, A)$, mientras que las operaciones por columnas afectan a $\left(\begin{array}{c} A \\ \hline I_n \end{array}\right)$. Las columnas de P proporcionan una base conjugada de V.

✓ **Ejemplo 21.** Diagonalizar por congruencias la matriz

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

Solución: Aplicando el esquema anterior, es necesario hacer operaciones elementales, las mismas por filas que por columnas a la matriz $\left(\begin{array}{c|c} I_3 & A \\ \hline & I_3 \end{array}\right)$ hasta transformar A en una matriz diagonal:

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -1 & -1 \\ 0 & 1 & 0 & -1 & 1 & -1 \\ 0 & 0 & 1 & -1 & -1 & 1 \\ \hline & & & 1 & 0 & 0 \\ & & & 0 & 1 & 0 \\ & & & 0 & 0 & 1 \end{array}\right) \xrightarrow[\substack{P_{21}(1) \\ P_{31}(1)}]{} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -1 & -1 \\ 1 & 1 & 0 & 0 & 0 & -2 \\ 1 & 0 & 1 & 0 & -2 & 0 \\ \hline & & & 1 & 0 & 0 \\ & & & 0 & 1 & 0 \\ & & & 0 & 0 & 1 \end{array}\right) \xrightarrow[\text{columnas}]{}$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & -2 \\ 1 & 0 & 1 & 0 & -2 & 0 \\ \hline & & & 1 & 1 & 1 \\ & & & 0 & 1 & 0 \\ & & & 0 & 0 & 1 \end{array}\right) \xrightarrow[P_{23}(1)]{} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 2 & 1 & 1 & 0 & -2 & -2 \\ 1 & 0 & 1 & 0 & -2 & 0 \\ \hline & & & 1 & 1 & 1 \\ & & & 0 & 1 & 0 \\ & & & 0 & 0 & 1 \end{array}\right) \xrightarrow[\text{columnas}]{}$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 2 & 1 & 1 & 0 & -4 & -2 \\ 1 & 0 & 1 & 0 & -2 & 0 \\ \hline & & & 1 & 2 & 1 \\ & & & 0 & 1 & 0 \\ & & & 0 & 1 & 1 \end{array}\right) \xrightarrow[P_{32}(-1/2)]{} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 2 & 1 & 1 & 0 & -4 & -2 \\ 0 & -1/2 & 1/2 & 0 & 0 & 1 \\ \hline & & & 1 & 2 & 1 \\ & & & 0 & 1 & 0 \\ & & & 0 & 1 & 1 \end{array}\right) \xrightarrow[\text{columnas}]{}$$

$$\left(\begin{array}{ccc|ccc} 1&0&0&1&0&0\\ 2&1&1&0&-4&0\\ 0&-1/2&1/2&0&0&1\\ \hline &&&1&2&0\\ &&&0&1&-1/2\\ &&&0&1&1/2 \end{array}\right) \xrightarrow[P_2(1/2)]{} \left(\begin{array}{ccc|ccc} 1&0&0&1&0&0\\ 1&1/2&1/2&0&-2&0\\ 0&-1/2&1/2&0&0&1\\ \hline &&&1&2&0\\ &&&0&1&-1/2\\ &&&0&1&1/2 \end{array}\right) \xrightarrow[\text{columnas}]{}$$

$$\left(\begin{array}{ccc|ccc} 1&0&0&1&0&0\\ 1&1/2&1/2&0&-1&0\\ 0&-1/2&1/2&0&0&1\\ \hline &&&1&1&0\\ &&&0&1/2&-1/2\\ &&&0&1/2&1/2 \end{array}\right) \xrightarrow[P_{23}]{} \left(\begin{array}{ccc|ccc} 1&0&0&1&0&0\\ 0&-1/2&1/2&0&0&1\\ 1&1/2&1/2&0&-1&0\\ \hline &&&1&1&0\\ &&&0&1/2&-1/2\\ &&&0&1/2&1/2 \end{array}\right) \xrightarrow[\text{columnas}]{}$$

$$\left(\begin{array}{ccc|ccc} 1&0&0&1&0&0\\ 0&-1/2&1/2&0&1&0\\ 1&1/2&1/2&0&0&-1\\ \hline &&&1&0&1\\ &&&0&-1/2&1/2\\ &&&0&1/2&1/2 \end{array}\right) = \left(\begin{array}{c|c} P^t & D\\ \hline & P \end{array}\right).$$

Las columnas de P proporcionan la base buscada. Así, en términos de la base $\{(1,0,0),\ (0,\frac{-1}{2},\frac{1}{2}),\ (1,\frac{1}{2},\frac{1}{2})\}$, la matriz coordenada de la forma canónica es diagonal: $D = \operatorname{diag}(1,1,-1) = \begin{pmatrix} 1&0&0\\ 0&1&0\\ 0&0&-1 \end{pmatrix}$.

◇

9.4. FORMAS BILINEALES ANTISIMÉTRICAS

En las secciones anteriores se han estudiado las formas bilineales simétricas a través de las formas cuadráticas. Ahora, se presenta brevemente la manera de expresar canónicamente las formas bilineales antisimétricas.

Teorema 9.30. Sea V un espacio vectorial real con $\dim V = n$ y sea $F : V \times V \to \mathbb{R}$ una forma bilineal antisimétrica. Entonces, el rango de f es par, $\operatorname{rg} F = 2k$. Además, existe una base de V en la que la matriz coordenada de F es diagonal por bloques con $n-2k$ bloques nulos de dimensión 1 y k bloques de la forma

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

Observación. El procedimiento para encontrar la base del teorema es similar a la diagonalización por congruencia vista para formas cuadráticas.

✓ **Ejemplo 22.** Considerar la siguiente forma bilineal antisimétrica sobre $\mathbb{R}^3$:

$$\begin{array}{rccl} F: & \mathbb{R}^3\times\mathbb{R}^3 & \longrightarrow & \mathbb{R} \\ & ((x_1,x_2,x_3),(y_1,y_2,y_3)) & \mapsto & (x_1y_2-x_2y_1)+2(x_1y_3-x_3y_1)+(x_2y_3-x_3y_2). \end{array}$$

Hallar una base de $\mathbb{R}^3$ de modo que la matriz coordenada de F sea

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

Solución: La matriz coordenada de F en base canónica es $A=\begin{pmatrix} 0 & 1 & 2 \\ -1 & 0 & 1 \\ -2 & -1 & 0 \end{pmatrix}$. Es inmediato comprobar que su rango es 2 por lo que, aplicando el teorema anterior, existe una base en la que la matriz coordenada es B. Se puede llegar a esta matriz realizando a la matriz A las mismas operaciones por filas y por columnas.

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 1 & 2 \\ 0 & 1 & 0 & -1 & 0 & 1 \\ 0 & 0 & 1 & -2 & -1 & 0 \\ \hline & & & 1 & 0 & 0 \\ & & & 0 & 1 & 0 \\ & & & 0 & 0 & 1 \end{array}\right) \xrightarrow[P_{32}(-2)]{} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 1 & 2 \\ 0 & 1 & 0 & -1 & 0 & 1 \\ 0 & -2 & 1 & 0 & -1 & -2 \\ \hline & & & 1 & 0 & 0 \\ & & & 0 & 1 & 0 \\ & & & 0 & 0 & 1 \end{array}\right) \xrightarrow[\text{columnas}]{}$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 & 0 & 1 \\ 0 & -2 & 1 & 0 & -1 & 0 \\ \hline & & & 1 & 0 & 0 \\ & & & 0 & 1 & -2 \\ & & & 0 & 0 & 1 \end{array}\right) \xrightarrow[P_{31}(1)]{} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 & 0 & 1 \\ 1 & -2 & 1 & 0 & 0 & 0 \\ \hline & & & 1 & 0 & 0 \\ & & & 0 & 1 & -2 \\ & & & 0 & 0 & 1 \end{array}\right) \xrightarrow[\text{columnas}]{}$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 & 0 \\ \hline & & & 1 & 0 & 1 \\ & & & 0 & 1 & -2 \\ & & & 0 & 0 & 1 \end{array}\right) = \left(\begin{array}{c|c} P^t & B \\ \hline & P \end{array}\right).$$

La base buscada la forman las columnas de P: $\{(1,0,0),\ (0,1,0),\ (1,-2,1)\}$.

◇

9.5. APLICACIONES

En esta sección se ve una aplicación de la diagonalización de formas cuadráticas definidas sobre $\mathbb{R}^2$ para la clasificación y estudio de las cónicas. Se puede plantear un estudio análogo para las cuádricas a través de diagonalización de formas cuadráticas sobre $\mathbb{R}^3$. Concretamente, se ve que toda ecuación polinómica de segundo grado en dos variables

$$a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + 2a_1x + 2a_2y + a_0 = 0,$$

donde alguno de los coeficientes a_{11}, a_{12}, a_{22} es distinto de cero, representa una cónica, entendiendo por cónica también los casos degenerados. De hecho, la ecuación anterior puede representar:

- Una elipse, una parábola o una hipérbola.
- Un par de rectas secantes, paralelas o coincidentes.
- Un punto o nada.

Además, se estudia bajo qué condiciones la cónica está girada respecto a los ejes coordenados y cómo determinar sus elementos característicos en ese caso.

En primer lugar se destacan dos partes diferenciadas en la ecuación de la cónica:

$$\underbrace{a_{11}x^2 + 2a_{12}xy + a_{22}y^2}_{\text{parte cuadrática}} + \underbrace{2a_1x + 2a_2y}_{\text{parte lineal}} + a_0 = 0.$$

El proceso para llevar la ecuación anterior a la ecuación reducida de la cónica tiene dos etapas:

1. Mediante un cambio de variables ortogonal hay que conseguir que la parte cuadrática de la ecuación no contenga términos cruzados, es decir, se necesita un cambio de variables $\mathbf{x} = P\widetilde{\mathbf{x}}$ con P una matriz ortogonal. Observar que si el coeficiente $a_{12} = 0$, se pasa directamente al siguiente punto.

2. Una vez obtenida la ecuación de segundo grado sin términos cruzados, basta completar cuadrados, o equivalentemente, hacer una traslación: $\mathbf{x}' = \widetilde{\mathbf{x}} - \mathbf{v}$.

Finalmente, para obtener los elementos característicos y la representación gráfica en el sistema de ejes original, es necesario deshacer los cambios de variables anteriores: $\mathbf{x} = P(\mathbf{x}' + \mathbf{v})$.

A continuación se analiza cada uno de los pasos con más detenimiento:

1. En términos matriciales, la ecuación de la cónica adopta la siguiente forma:

$$\begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + 2 \begin{pmatrix} a_1 & a_2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + a_0 = 0. \tag{9.1}$$

Realizando la diagonalización ortogonal de la matriz asociada a la parte cuadrática $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix}$: Si λ_1 y λ_2 son los valores propios de A, se eligen dos vectores propios ortonormales $\mathbf{v}_1$ y $\mathbf{v}_2$ de modo que el ángulo de $\mathbf{v}_1$ a $\mathbf{v}_2$ es $\frac{\pi}{2}$ en sentido positivo (contrario a las agujas del reloj). De este modo se obtiene:

$$P^t A P = D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \quad \text{donde} \quad P = (\mathbf{v}_1 \,|\, \mathbf{v}_2)\,.$$

Si se definen nuevas coordenadas $(\widetilde{x}, \widetilde{y})$ mediante

$$\begin{pmatrix} x \\ y \end{pmatrix} = P \begin{pmatrix} \widetilde{x} \\ \widetilde{y} \end{pmatrix},$$

la ecuación de la cónica (9.1), en términos de las nuevas variables, es:

$$\begin{pmatrix} \widetilde{x} & \widetilde{y} \end{pmatrix} P^t A P \begin{pmatrix} \widetilde{x} \\ \widetilde{y} \end{pmatrix} + 2 \begin{pmatrix} a_1 & a_2 \end{pmatrix} P \begin{pmatrix} \widetilde{x} \\ \widetilde{y} \end{pmatrix} + a_0 = 0,$$

$$\begin{pmatrix} \widetilde{x} & \widetilde{y} \end{pmatrix} D \begin{pmatrix} \widetilde{x} \\ \widetilde{y} \end{pmatrix} + 2 \begin{pmatrix} b_1 & b_2 \end{pmatrix} \begin{pmatrix} \widetilde{x} \\ \widetilde{y} \end{pmatrix} + a_0 = 0,$$

es decir, la ecuación de la cónica en las coordenadas $(\widetilde{x}, \widetilde{y})$ es:

$$\lambda_1 \widetilde{x}^2 + \lambda_2 \widetilde{y}^2 + 2b_1 \widetilde{x} + 2b_2 \widetilde{y} + a_0 = 0. \tag{9.2}$$

Los ejes de esta cónica son las rectas que pasan por el origen de coordenadas y tienen como vectores directores los vectores propios de A: $\mathbf{v}_1$ y $\mathbf{v}_2$. El nuevo sistema de ejes $O\widetilde{X}\widetilde{Y}$ se obtiene del sistema OXY haciendo un giro, con centro el origen de coordenadas, y ángulo el que determina $\mathbf{v}_1$ con el semieje OX^+.

2. Una vez obtenida la ecuación (9.2), hay que completar cuadrados mediante cambios del tipo $x' = \widetilde{x} - \alpha$ e $y' = \widetilde{y} - \beta$ para obtener una ecuación de uno de los tipos siguientes:

   - <u>Caso elíptico:</u> $\lambda_1 \lambda_2 > 0$ (valores propios no nulos del mismo signo):

     $$a^2 x'^2 + b^2 y'^2 = c.$$

     Las posibilidades son: $\begin{cases} c > 0, & \text{elipse} \\ c = 0, & \text{punto} \\ c < 0, & \text{nada.} \end{cases}$

   - <u>Caso hiperbólico:</u> $\lambda_1 \lambda_2 < 0$ (valores propios no nulos de distinto signo):

     $$a^2 x'^2 - b^2 y'^2 = c.$$

     Las posibilidades son: $\begin{cases} c \neq 0, & \text{hipérbola} \\ c = 0, & \text{dos rectas secantes.} \end{cases}$

- Caso parabólico: $\lambda_1\lambda_2 = 0$ (un valor propio nulo y el otro no). Si se supone que $\lambda_2 = 0$:
$$a^2x'^2 + by' = 0, \quad \text{o bien} \quad a^2x'^2 + c = 0.$$
$$\text{Las posibilidades son: } \begin{cases} b \neq 0, & \text{parábola} \\ c < 0, & \text{dos rectas paralelas} \\ c = 0, & \text{dos rectas coincidentes} \\ c > 0, & \text{nada.} \end{cases}$$

Finalmente, para obtener los elementos característicos de la cónica y su representación gráfica, basta obtenerlos en las coordenadas (x', y') y deshacer los cambios de variables hechos en el proceso.

✓ **Ejemplo 23.** Considerar la cónica
$$3x^2 + 3y^2 - 2xy + 2x - 4y + 1 = 0.$$

La matriz asociada a la parte cuadrática es $A = \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}$. Es inmediato comprobar que los valores propios son $\lambda_1 = 4$ y $\lambda_2 = 2$ (notar que como $\lambda_1\lambda_2 > 0$, el caso correspondiente es el elíptico). Como vectores propios ortonormales se pueden considerar $\mathbf{v}_1 = \frac{\sqrt{2}}{2}(1, -1)$ y $\mathbf{v}_2 = \frac{\sqrt{2}}{2}(1, 1)$. Además, el ángulo que forma $\mathbf{v}_1$ con el semieje positivo OX es $\frac{-\pi}{4}$.

Tras utilizar el primer cambio de coordenadas, la nueva ecuación de la cónica es:
$$\mathbf{x} = P\widetilde{\mathbf{x}} \Longrightarrow 4\widetilde{x}^2 + 2\widetilde{y}^2 + 3\sqrt{2}\widetilde{x} - \sqrt{2}\widetilde{y} + 1 = 0.$$

Para completar cuadrados en la ecuación anterior se necesita hacer las siguientes traslaciones:
$$x' = \widetilde{x} + \frac{3\sqrt{2}}{8}, \quad y' = \widetilde{y} - \frac{\sqrt{2}}{4}$$
de modo que la ecuación queda simplemente:
$$4x'^2 + 2y'^2 = \frac{3}{8}.$$

Como $c = \frac{3}{8} > 0$, la cónica es una elipse.

Por ejemplo, para calcular el centro de la elipse se haría lo siguiente: En coordenadas (x', y') el centro es el $(0, 0)$. Como $\mathbf{x} = P(\mathbf{x}' + \mathbf{v})$, entonces el centro de la elipse original es:
$$\begin{pmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ -\sqrt{2}/2 & \sqrt{2}/2 \end{pmatrix} \begin{pmatrix} -3\sqrt{2}/8 \\ \sqrt{2}/4 \end{pmatrix} = \begin{pmatrix} -1/8 \\ 5/8 \end{pmatrix}.$$

◇

9.6. COMANDOS DE *wxMaxima*

Para resolver los ejercicios de este capítulo no hay nuevos comandos de *xwMaxima*. Algunos que pueden ser de utilidad pero que ya se han presentado en otros capítulos son, por ejemplo: “`rank`”, “`eigenvalues`”, “`dispJordan`”, “`signum`”.

9.7. EJERCICIOS PROPUESTOS

Formas bilineales

1. Sea $F : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ la aplicación dada por $F((x_1, x_2), (y_1, y_2)) = x_1y_1 - 2x_1y_2$. Probar que F es bilineal pero no simétrica.

2. Sean los vectores $\mathbf{u} = (x_1, x_2, x_3)$ y $\mathbf{v} = (y_1, y_2, y_3)$ y sea F la forma bilineal sobre $\mathbb{R}^3$ dada por

$$F(\mathbf{u}, \mathbf{v}) = 3x_1y_1 - 2x_1y_2 + 5x_2y_1 + 7x_2y_2 - 8x_2y_3 + 4x_3y_2 - x_3y_3.$$

Obtener la expresión coordenada de F.

3. Sea F la forma bilineal sobre $\mathbb{R}^2$ definida por:

$$F((x_1, x_2), (y_1, y_2)) = 2x_1y_1 - 3x_1y_2 + x_2y_2.$$

a) Hallar la matriz coordenada A de F en la base $\mathcal{B} = \{(1, 0), (1, 1)\}$.

b) Hallar la matriz coordenada B de F en la base $\mathcal{C} = \{(2, 1), (1, -1)\}$.

c) Hallar la matriz de cambio de coordenadas P y verificar que $B = P^tAP$.

Formas cuadráticas

4. Para la forma bilineal F sobre $\mathbb{R}^3$ dada por

$$F((x_1, x_2, x_3), (y_1, y_2, y_3)) = x_1y_1 + x_2y_2 + x_3y_3 + x_1y_3,$$

describir su forma cuadrática asociada y la forma polar de esta.

5. Se considera la siguiente forma cuadrática sobre $\mathbb{R}^2$, $Q(\mathbf{v}) = 2x_1^2 - 3x_1x_2 + 4x_2^2$, donde (x_1, x_2) son las coordenadas del vector $\mathbf{v}$ en una base $\{\mathbf{e}_1, \mathbf{e}_2\}$ de $\mathbb{R}^2$. Calcular la expresión coordenada de Q:

a) en la base $\{\mathbf{e}_1, \mathbf{e}_2\}$,

b) en la base $\{\mathbf{v}_1, \mathbf{v}_2\}$, donde las coordenadas de $\mathbf{v}_1$ y $\mathbf{v}_2$ en términos de la base $\{\mathbf{e}_1, \mathbf{e}_2\}$ vienen dadas por $(1, -1)$ y $(1, 2)$ respectivamente.

6. Dada la forma bilineal

$$F((x_1, x_2, x_3), (y_1, y_2, y_3)) = x_1y_1 + x_1y_2 + x_1y_3 + x_2y_1 + x_2y_2 - x_3y_1 + 2x_3y_3,$$

hallar su forma cuadrática asociada Q, su forma polar, la matriz asociada a la forma polar y la signatura y el rango de Q.

7. Hallar una matriz regular P tal que P^tAP sea diagonal con 1's, -1's y 0's en la diagonal, siendo A:

$$\text{a) } A = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \\ 2 & 1 & 0 \end{pmatrix} \qquad \text{b) } A = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$$

8. Calcular, según los valores del parámetro $b \in \mathbb{R}$, el rango y la signatura de la matriz

$$A = \begin{pmatrix} 2-b & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 4 & b-2 \\ 0 & 0 & b-2 & 4 \end{pmatrix}.$$

9. Dada la forma cuadrática Q sobre $\mathbb{R}^3$

$$Q(x,y,z) = \alpha\, x^2 + \alpha\, y^2 + (\alpha - 1)z^2 + 2xy,$$

donde $\alpha \in \mathbb{R}$, hallar el rango y la signatura de Q para los distintos valores de α. Clasificar, en función de dichos valores, la forma cuadrática.

10. Considerar las siguientes formas cuadráticas en $\mathbb{R}^3$:

a) $Q(x,y,z) = 2x^2 + y^2 + z^2 + 2xy + 2xz$.

b) $Q(x,y,z) = xy + 2xz$.

c) $Q(x,y,z) = x^2 - z^2 - 2xy + xz$.

d) $Q(x,y,z) = 2x^2 + y^2 + 5z^2 - 2xy - 2yz + 6xz$.

e) $Q(x,y,z) = -x^2 - 2y^2 - z^2 + 2xy + 2yz$.

Escribirlas en forma matricial y clasificarlas.

11. Diagonalizar las siguientes formas cuadráticas reales, indicando su clasificación:

a) $Q(x_1,x_2,x_3) = 2x_1^2 + x_1x_2 + x_2x_3 + 2x_3^2$.

b) $Q(x_1,x_2,x_3) = x_1^2 - 3x_3^2 - 2x_1x_2 + 2x_1x_3 - 6x_2x_3$.

c) $Q(x_1,x_2,x_3) = x_1^2 - 2x_2x_3 + x_1x_3$.

d) $Q(x_1,x_2,x_3,x_4) = x_1x_2 + x_1x_3 + x_1x_4 + x_2x_3 + x_2x_4 + x_3x_4$.

12. Diagonalizar por congruencias las formas cuadráticas del ejercicio anterior, obteniendo una matriz diagonal con 1's, -1's y 0's. Comprobar que se obtienen los mismos resultados que en el ejercicio anterior.

Formas bilineales antisimétricas

13. Dada la matriz antisimétrica $A = \begin{pmatrix} 0 & 1 & 2 \\ -1 & 0 & 3 \\ -2 & -3 & 0 \end{pmatrix}$, hallar una matriz regular P de modo que $P^tAP = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.

14. Sea F la forma bilineal antisimétrica sobre $\mathbb{R}^4$ definida por:

$$F(\mathbf{x},\mathbf{y}) = (x_1y_2 - x_2y_1) + (x_1y_3 - x_3y_1) + 2(x_1y_4 - x_4y_1) + 3(x_2y_4 - x_4y_2) + (x_2y_3 - x_3y_2).$$

Hallar una base de $\mathbb{R}^4$ en la que la matriz coordenada de F sea diagonal por bloques.

9.8. SOLUCIONES A LOS EJERCICIOS PROPUESTOS

Formas bilineales

1. Para ver que no es simétrica tomar por ejemplo, $\mathbf{u} = (1, 0)$ y $\mathbf{v} = (0, 1)$.

2. La matriz coordenada en base canónica es $A = \begin{pmatrix} 3 & -2 & 0 \\ 5 & 7 & -8 \\ 0 & 4 & -1 \end{pmatrix}$.

3. $A = \begin{pmatrix} 2 & -1 \\ 2 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 3 & 9 \\ 0 & 6 \end{pmatrix}, \quad P = \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}$.

Formas cuadráticas

4. Forma cuadrática: $Q(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 + x_1x_3$.
Forma polar: $F_p((x_1, x_2, x_3), (y_1, y_2, y_3)) = x_1y_1 + x_2y_2 + x_3y_3 + \dfrac{x_1y_3}{2} + \dfrac{y_1x_3}{2}$.

5. a) Matriz coordenada: $\begin{pmatrix} 2 & {}^{-3}\!/_{2} \\ {}^{-3}\!/_{2} & 4 \end{pmatrix}$. b) Matriz coordenada: $\begin{pmatrix} 9 & {}^{-15}\!/_{2} \\ {}^{-15}\!/_{2} & 12 \end{pmatrix}$.

6. Forma cuadrática: $Q(x_1, x_2, x_3) = x_1^2 + x_2^2 + 2x_3^2 + 2x_1x_2$.

Forma polar: $F_p((x_1, x_2, x_3), (y_1, y_2, y_3)) = x_1y_1 + x_2y_2 + 2x_3y_3 + x_1y_2 + x_2y_1$.

Matriz asociada a la forma polar: $A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$.

$\operatorname{sign} Q = \operatorname{rg} Q = 2$. (Los valores propios de A son: 0, con multiplicidad 1, y 2 con multiplicidad 2).

7. Una posible solución es:

a) $P = \begin{pmatrix} 0 & 1 & 1 \\ 1 & -1 & -2 \\ 0 & 0 & 1 \end{pmatrix}$. b) $P = \begin{pmatrix} {}^{1}\!/_{\sqrt{2}} & {}^{1}\!/_{\sqrt{2}} & {}^{-\sqrt{2}}\!/_{\sqrt{3}} & 0 \\ {}^{1}\!/_{\sqrt{2}} & {}^{-1}\!/_{\sqrt{2}} & {}^{-\sqrt{2}}\!/_{\sqrt{3}} & 0 \\ 0 & 0 & {}^{1}\!/_{\sqrt{6}} & {}^{1}\!/_{\sqrt{2}} \\ 0 & 0 & {}^{1}\!/_{\sqrt{6}} & {}^{-1}\!/_{\sqrt{2}} \end{pmatrix}$.

8. Los valores propios son: $\lambda_1 = 2$, $\lambda_2 = 2 - b$, $\lambda_3 = 6 - b$, $\lambda_4 = 2 + b$. Observar que, dependiendo del valor de b puede haber valores propios repetidos.

	Rango	**Signatura**
$b < -2$	4	3
$b = -2$	3	3
$b \in (-2, 2)$	4	4
$b = 2$	3	3
$b \in (2, 6)$	4	3
$b = 6$	3	2
$b > 6$	4	2

9. Los valores propios son: $\lambda_1 = \alpha - 1$ (multiplicidad 2), $\lambda_2 = \alpha + 1$.

	Rango	**Signatura**	**Clasificación**
$\alpha < -1$	3	0	Definida negativa
$\alpha = -1$	2	0	Semidefinida negativa
$\alpha \in (-1, 1)$	3	1	Indefinida
$\alpha = 1$	1	1	Semidefinida positiva
$\alpha > 1$	3	3	Definida positiva

10. a) $A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$. Valores propios: $\lambda_1 = 0$, $\lambda_2 = 1$, $\lambda_3 = 3$.

$\operatorname{rg} Q = \operatorname{sign} Q = 2 < 3 = \dim V \Longrightarrow$ Semidefinida positiva.

b) $A = \begin{pmatrix} 0 & 1/2 & 1 \\ 1/2 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$. Valores propios: $\lambda_1 = 0$, $\lambda_2 = \frac{-\sqrt{5}}{2}$, $\lambda_3 = \frac{\sqrt{5}}{2}$.

$\operatorname{rg} Q = 2$, $\operatorname{sign} Q = 1 \Longrightarrow$ Indefinida.

c) $A = \begin{pmatrix} 1 & -1 & 1/2 \\ -1 & 0 & 0 \\ 1/2 & 0 & -1 \end{pmatrix}$. Valores propios: $\lambda_1 = -\frac{1}{2}$, $\lambda_2 = \frac{1+\sqrt{33}}{4}$, $\lambda_3 = \frac{1-\sqrt{33}}{4}$.

$\operatorname{rg} Q = 3$, $\operatorname{sign} Q = 1 \Longrightarrow$ Indefinida.

d) $A = \begin{pmatrix} 2 & -1 & 3 \\ -1 & 1 & -1 \\ 3 & -1 & 5 \end{pmatrix}$. Valores propios: $\lambda_1 = 0$, $\lambda_2 = 4 + \sqrt{10}$, $\lambda_3 = 4 - \sqrt{10}$.

$\operatorname{rg} Q = \operatorname{sign} Q = 2 < 3 = \dim V \Longrightarrow$ Semidefinida positiva.

e) $A = \begin{pmatrix} -1 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -1 \end{pmatrix}$. Valores propios: $\lambda_1 = 0$, $\lambda_2 = -1$, $\lambda_3 = -3$.

$\operatorname{rg} Q = 2$, $\operatorname{sign} Q = 0 \Longrightarrow$ Semidefinida negativa.

11. a) Base $\begin{cases} y_1 = \sqrt{2}(x_1 + \frac{1}{4}x_2) \\ y_2 = \sqrt{2}(x_3 + \frac{1}{4}x_2) \\ y_3 = \frac{1}{2}x_2 \end{cases} \Longrightarrow Q(y_1, y_2, y_3) = y_1^2 + y_2^2 - y_3^2.$

$\operatorname{rg} Q = 3$, $\operatorname{sign} Q = 2 \Longrightarrow$ Indefinida.

b) Base $\begin{cases} y_1 = x_1 - x_2 + x_3 \\ y_2 = x_2 + 2x_3 \\ y_3 = x_3 \end{cases} \Longrightarrow Q(y_1, y_2, y_3) = y_1^2 - y_2^2.$

$\operatorname{rg} Q = 2$, $\operatorname{sign} Q = 1 \Longrightarrow$ Indefinida.

c) Base $\begin{cases} y_1 = x_1 + \frac{1}{2}x_3 \\ y_2 = 2x_2 \\ y_3 = 2x_2 + \frac{1}{2}x_3 \end{cases} \Longrightarrow Q(y_1, y_2, y_3) = y_1^2 + y_2^2 - y_3^2.$

$\operatorname{rg} Q = 3$, $\operatorname{sign} Q = 2 \Longrightarrow$ Indefinida.

d) Base $\begin{cases} y_1 = \frac{1}{2}(x_1 + x_2) + x_3 + x_4 \\ y_2 = \frac{1}{2}(x_1 - x_2) \\ y_3 = \frac{\sqrt{3}}{2}(x_3 + x_4) \\ y_4 = \frac{1}{2}(x_3 - x_4) \end{cases} \Longrightarrow Q(y_1, y_2, y_3, y_4) = y_1^2 - y_2^2 - y_3^2 - y_4^2.$

$\operatorname{rg} Q = 4$, $\operatorname{sign} Q = 1 \Longrightarrow$ Indefinida.

Formas bilineales antisimétricas

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

14. $\mathcal{B}_{\mathbb{R}^4} = \{(1, 0, 0, 0),\ (0, 1, 0, 0),\ (1, 0, -2, 1),\ (0, -1, 3, -1)\}.$

BIBLIOGRAFÍA

[1] Juan de Burgos. *Álgebra Lineal y geometría cartesiana.* McGraw-Hill, Universidad Politécnica de Madrid, 2000.

[2] Eugenio Hernández. *Álgebra y geometría.* Addison Wesley Iberoamericana, Universidad Autónoma de Madrid, 1994.

[3] Roland E. Larson y Bruce H. Edwards. *Introduccion al Álgebra Lineal.* Limusa Noriega Editores, 2006.

[4] David C. Lay. *Álgebra Lineal y sus Aplicaciones.* Pearson Addison Wesley, University of Maryland, 2007.

[5] Luis Merino González y Evangelina Santos Aláez. *Álgebra Lineal con métodos elementales.* Paraninfo Thomson Learning, Granada, 2006.

[6] Antonio Otal, Maria Victoria Sebastián y Raquel Villacampa. *Álgebra Lineal con wxMaxima.* Centro Universitario de la Defensa, Textos Docentes 11, Zaragoza, 2013.