Apuntes de álgebra lineal

Jose S. Cánovas Peña

2 de febrero de 2008

Índice General

Capítulo 1

Aprendemos a hacer cuentas

Sumario. Definición de matriz. Operaciones con matrices. Tipos de matrices. Operaciones elementales. Matrices elementales. Rango de una matriz. Matrices inversas. Definición de determinante. Propiedades básicas. Cálculo de la matriz inversa mediante determiantes. Definición de sistema de ecuaciones lineales. Forma matricial del sistema. Teorema de Rouché–Frobenius. Método de Gauss. Método de Cramer.

1.1 Matrices. Primeras definiciones

Sea $\mathbb{K}$ el cuerpo de los números reales o números complejos (que a veces llamaremos escalares) y sean m, n números naturales. Una *matriz* $n \times m$ es una aplicación $\mathbf{A} : \{1, 2, ..., n\} \times \{1, 2, ..., m\} \to \mathbb{K}$, es decir, dados $(i, j) \in \{1, 2, ..., n\} \times \{1, 2, ..., m\}$, $\mathbf{A}(i, j) = a_{ij}$ será un número real o complejo. Como el dominio de la matriz (aplicación) $\mathbf{A}$ es finito, es más usual escribir una matriz de la siguiente forma

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

Se dirá entonces que la matriz $\mathbf{A}$ tiene n filas y m columnas. Denotaremos por $\mathcal{M}_{n\times m}(\mathbb{K})$ al conjunto de matrices de tamaño $n \times m$ con coeficientes en $\mathbb{K}$. En caso de que $n = m$, es decir, tenemos el mismo número de filas que de columnas, diremos que la matriz $\mathbf{A}$ es cuadrada. Por ejemplo, la matriz

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

es cuadrada al tener 3 filas y columnas. Como notación escribiremos las matrices de la forma $\mathbf{A} = (a_{ij})_{i=1,2,...,n}^{j=1,2,...,m}$ o simplemente $\mathbf{A} = (a_{ij})$ si no hace falta especificar el número de filas y columnas.

La *fila* i–ésima de la matriz $\mathbf{A}$ será denotada por $\mathbf{A}_i = (a_{i1}, a_{i2}, ...a_{im})$, $1 \leq i \leq n$. Nótese que cada fila de la matriz $\mathbf{A}$ es a su vez una matriz de una fila y m columnas. A su vez, la *columna*

j–ésima de la matriz, $1 \le j \le m$, la representaremos por

$$\mathbf{A}^j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \dots \\ a_{nj} \end{pmatrix},$$

y será por tanto una matriz de n filas y una columna. Por ejemplo, en la matriz anterior la fila $\mathbf{A}_2 = (3, 2, 1)$ mientras que la columna es

$$\mathbf{A}^3 = \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}.$$

Dada la matriz $\mathbf{A}$ los elementos de la forma a_{ii}, $1 \le i \le \min\{n, m\}$ forman la *diagonal principal* de la matriz. Una matriz cuadrada que sólo tome valores no nulos en los elementos de la diagonal principal se dirá *diagonal*. Por ejemplo, las matrices

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & \pi \end{pmatrix},$$

serán diagonales. Como vemos las matrices diagonales cumplen que sus elementos $a_{ij} = 0$ si $i \neq j$. Si somos menos exigentes y sólo pedimos que o bien $a_{ij} = 0$ si $i < j$ o bien $a_{ij} = 0$ si $i > j$ tendremos la matrices (no necesariamente cuadradas) *triangulares*. Por ejemplo las matrices

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \end{pmatrix}, \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 6 & 8 & 0 \\ 1 & 1 & 11 \end{pmatrix}, \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 0 \end{pmatrix}, \begin{pmatrix} 2 & 0 & 1 \\ 0 & 2 & 7 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

son triangulares.

1.2 Operaciones con matrices

1.2.1 Suma de matrices

Dadas las matrices $\mathbf{A}, \mathbf{B} \in \mathcal{M}_{n\times m}(\mathbb{K})$, $m, n \in \mathbb{N}$, se define la *suma* de ambas matrices como

$$\mathbf{A} + \mathbf{B} = (a_{ij}) + (b_{ij}) = (a_{ij} + b_{ij}).$$

Por ejemplo

$$\begin{pmatrix} 0 & 1 & -2 \\ 4 & 3 & 4 \\ 9 & -4 & -4 \\ 1 & 1 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 1 & -5 \\ 1 & 5 & 8 \\ 2 & -1 & -1 \\ 1 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & -7 \\ 5 & 8 & 12 \\ 11 & -5 & -5 \\ 2 & 2 & 3 \end{pmatrix}.$$

Démonos cuenta que no es posible sumar matrices de distinto tamaño, es decir, con distinto número de filas o columnas. Por ejemplo no es posible calcular

$$\begin{pmatrix} 0 & 1 & -2 \\ 4 & 3 & 4 \\ 9 & -4 & -4 \\ 1 & 1 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 1 & 1 \\ 8 & 2 & 2 \\ 0 & 1 & 1 \end{pmatrix}.$$

La suma de matrices tiene las siguientes propiedades que se infieren directamente de las propiedades del cuerpo $\mathbb{K}$ y son las siguientes:

1. Propiedad asociativa. Dadas $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C} \in \mathcal{M}_{n\times m}(\mathbb{K})$ se verifica que $(\mathbf{A}+\mathbf{B})+\mathbf{C} = \mathbf{A}+(\mathbf{B}+\mathbf{C})$. Para demostrar esta propiedad consideramos

$$\begin{aligned}(\mathbf{A}+\mathbf{B})+\mathbf{C} &= ((a_{ij})+(b_{ij}))+(c_{ij}) = (a_{ij}+b_{ij})+(c_{ij}) = ((a_{ij}+b_{ij})+c_{ij}) \\ &= (a_{ij}+(b_{ij}+c_{ij})) = (a_{ij})+(b_{ij}+c_{ij}) = (a_{ij})+((b_{ij})+(c_{ij})) \\ &= \mathbf{A}+(\mathbf{B}+\mathbf{C}),\end{aligned}$$

   dado que la suma en el cuerpo $\mathbb{K}$ es asociativa.

2. Propiedad conmutativa. Dadas $\mathbf{A}$, $\mathbf{B} \in \mathcal{M}_{n\times m}(\mathbb{K})$ se verifica que $\mathbf{A}+\mathbf{B} = \mathbf{B}+\mathbf{A}$. Para demostrar esta propiedad consideramos

$$\mathbf{A}+\mathbf{B} = (a_{ij})+(b_{ij}) = (a_{ij}+b_{ij}) = (b_{ij}+a_{ij}) = (b_{ij})+(a_{ij}) = \mathbf{B}+\mathbf{A},$$

   dado que la suma en el cuerpo $\mathbb{K}$ es asociativa.

3. Elemento neutro. Se trata de la matriz $\mathbf{0} \in \mathcal{M}_{n\times m}(\mathbb{K})$, que es aquella que tiene cero en todas sus componentes. Es claro entonces que dada $\mathbf{A} \in \mathbf{M}_{n\times m}(\mathbb{K})$ se verifica que $\mathbf{A}+\mathbf{0} = \mathbf{0}+\mathbf{A} = \mathbf{A}$.

4. Elemento simétrico. Dado $\mathbf{A} \in \mathcal{M}_{n\times m}(\mathbb{K})$ existe un elemento $-\mathbf{A}$ de manera que $\mathbf{A}+(-\mathbf{A}) = (-\mathbf{A})+\mathbf{A} = \mathbf{0}$. Dada la matriz $\mathbf{A} = (a_{ij})$ se tiene que $-\mathbf{A} = (-a_{ij})$. Entonces es claro que

$$\mathbf{A}+(-\mathbf{A}) = (a_{ij})+(-a_{ij}) = (0) = \mathbf{0}.$$

Por verificarse estas cuatro propiedades, se dice que el par formado por el conjunto de matrices con la operación suma $(\mathcal{M}_{n\times m}(\mathbb{K}),+)$ es un grupo conmutativo.

1.2.2 Producto de matrices

Dadas las matrices $\mathbf{A} \in \mathcal{M}_{n\times m}(\mathbb{K})$ y $\mathbf{B} \in \mathcal{M}_{m\times k}(\mathbb{K})$, $m,n,k \in \mathbb{N}$, se define el *producto* $\mathbf{A}\cdot\mathbf{B} = (\sum_{k=1}^m a_{ik}b_{kj})$. Es decir, para poder multiplicar dos matrices en número de columnas de la primera ha de coincidir con el número de filas de la segunda. Por ejemplo

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

Las propiedades que cumple el producto de matrices son las siguientes.

1. Propiedad asociativa. Dadas $\mathbf{A} \in \mathcal{M}_{n\times m}(\mathbb{K})$, $\mathbf{B} \in \mathcal{M}_{m\times k}(\mathbb{K})$, $\mathbf{C} \in \mathcal{M}_{k\times l}(\mathbb{K})$ se verifica que $(\mathbf{A}\cdot\mathbf{B})\cdot\mathbf{C} = \mathbf{A}\cdot(\mathbf{B}\cdot\mathbf{C})$. Para ver la demostración consideramos

$$\begin{aligned}(\mathbf{A}\cdot\mathbf{B})\cdot\mathbf{C} &= ((a_{ij})\cdot(b_{ij}))\cdot(c_{ij}) = (\sum_{r=1}^{m} a_{ir}b_{rj})\cdot(c_{ij}) \\ &= (\sum_{s=1}^{k}\sum_{r=1}^{m} a_{ir}b_{rs}c_{sj}) = (\sum_{r=1}^{m} a_{ir}(\sum_{s=1}^{k} b_{rs}c_{sj})) \\ &= (a_{ij})\cdot(\sum_{s=1}^{k} b_{rs}c_{sj}) = \mathbf{A}\cdot(\mathbf{B}\cdot\mathbf{C}).\end{aligned}$$

2. En general no cabe plantearse si se cumple la propiedad conmutativa ya que como vemos en el ejemplo anterior, no se puede hacer el producto en orden inverso porque el número de columnas de la segunda matriz no coincide con el número de filas de la primera matriz. Ahora bien, en caso de poder realizarse el producto, por ejemplo si las matrices son cuadradas la propiedad conmutativa no se verifica como pone de manifiesto el siguiente ejemplo:

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

mientras que

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

3. Existe un elemento neutro, llamado *identidad* y que es la matriz diagonal que tiene 1 en cada elemento de la diagonal principal. Si llamamos $\mathbf{I}_n \in \mathcal{M}_{n\times n}(\mathbb{K})$ a la matriz identidad y $\mathbf{A} \in \mathcal{M}_{n\times m}(\mathbb{K})$, se verifica que $\mathbf{I}_n\cdot\mathbf{A} = \mathbf{A}\cdot\mathbf{I}_m = \mathbf{A}$. Si la matriz $\mathbf{A}$ es cuadrada, entonces $\mathbf{I}_n\cdot\mathbf{A} = \mathbf{A}\cdot\mathbf{I}_n = \mathbf{A}$.

4. No toda matriz cuadrada $\mathbf{A}$ no nula tiene matriz inversa, es decir, otra matriz $\mathbf{A}^{-1}$ tal que $\mathbf{A}\cdot\mathbf{A}^{-1} = \mathbf{A}^{-1}\cdot\mathbf{A} = \mathbf{I}_n$. En caso de existir diremos que la matriz $\mathbf{A}$ es *invertible* y que $\mathbf{A}^{-1}$ es su *matriz inversa*. Por ejemplo la matriz $\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$ no tiene inversa ya que si esta existiera tendría que verificarse que

$$\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}\cdot\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ a & b \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

y tendríamos que $a = 1$ y $a = 0$, lo cual es absurdo. Por ejemplo la matriz $\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ es invertible, siendo su inversa $\begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$. Veremos posteriormente cómo caracterizar las matrices invertibles y cómo se obtiene su inversa en caso de serlo.

5. Propiedad ditributiva del producto respecto de la suma de matrices. Dadas $\mathbf{A} \in \mathcal{M}_{n\times m}(\mathbb{K})$ $\mathbf{B}, \mathbf{C} \in \mathcal{M}_{m\times k}(\mathbb{K})$, se verifica que $\mathbf{A}\cdot(\mathbf{B}+\mathbf{C}) = \mathbf{A}\cdot\mathbf{B} + \mathbf{A}\cdot\mathbf{C}$. Para demostrar esta identidad

consideramos

$$\begin{aligned}\mathbf{A}\cdot(\mathbf{B}+\mathbf{C}) &= (a_{ij})\cdot(b_{ij}+c_{ij}) = (\sum_{r=1}^{m} a_{ir}(b_{rj}+c_{rj}))\\ &= (\sum_{r=1}^{m} a_{ir}b_{rj} + \sum_{r=1}^{m} a_{ir}c_{rj}) = (a_{ij})\cdot(b_{ij}) + (a_{ij}) + (c_{ij})\\ &= \mathbf{A}\cdot\mathbf{B}+\mathbf{A}\cdot\mathbf{C}.\end{aligned}$$

1.2.3 Multiplicación de una matriz por un escalar

Sean $\alpha \in \mathbb{K}$ y $\mathbf{A} \in \mathcal{M}_{n\times m}(\mathbb{K})$, $n, m \in \mathbb{N}$. Se define la *multiplicación* de α por $\mathbf{A} = (a_{ij})$ como la matriz $\alpha \cdot \mathbf{A} = (\alpha a_{ij})$. Por ejemplo

$$2\cdot\begin{pmatrix}1 & 2\\ 2 & 3\\ 3 & 4\end{pmatrix} = \begin{pmatrix}2 & 4\\ 4 & 6\\ 6 & 8\end{pmatrix}.$$

Veamos que propiedades tiene este producto.

1. Propiedad distributiva del producto respecto de la suma de matrices. Sean $\alpha \in \mathbb{K}$ y $\mathbf{A}, \mathbf{B} \in \mathcal{M}_{n\times m}(\mathbb{K})$, entonces se verifica que $\alpha\cdot(\mathbf{A}+\mathbf{B}) = \alpha\cdot\mathbf{A}+\alpha\cdot\mathbf{B}$. Para probar esta igualdad consideramos

$$\begin{aligned}\alpha\cdot(\mathbf{A}+\mathbf{B}) &= \alpha\cdot((a_{ij})+(b_{ij})) = \alpha\cdot(a_{ij}+b_{ij}) = (\alpha(a_{ij}+b_{ij}))\\ &= (\alpha a_{ij}+\alpha b_{ij}) = (\alpha a_{ij}) + (\beta a_{ij}) = \alpha(a_{ij}) + \alpha(b_{ij})\\ &= \alpha\cdot\mathbf{A}+\alpha\cdot\mathbf{B}.\end{aligned}$$

2. Propiedad distributiva del producto respecto de la suma de escalares. Sean $\alpha, \beta \in \mathbb{K}$ y $\mathbf{A} \in \mathcal{M}_{n\times m}(\mathbb{K})$, entonces se cumple que $(\alpha+\beta)\cdot\mathbf{A} = \alpha\cdot\mathbf{A}+\beta\cdot\mathbf{A}$. Para demostrar la igualdad tomamos

$$\begin{aligned}(\alpha+\beta)\cdot\mathbf{A} &= (\alpha+\beta)\cdot(a_{ij}) = ((\alpha+\beta)a_{ij}) = (\alpha a_{ij}+\beta a_{ij})\\ &= (\alpha a_{ij}) + (\beta a_{ij}) = \alpha\cdot(a_{ij}) + \beta\cdot(a_{ij}) = \alpha\cdot\mathbf{A}+\beta\cdot\mathbf{B}.\end{aligned}$$

3. Propiedad pseudoasociativa. Sean $\alpha, \beta \in \mathbb{K}$ y $\mathbf{A} \in \mathcal{M}_{n\times m}(\mathbb{K})$, entonces se cumple que $(\alpha\beta)\cdot\mathbf{A} = \alpha\cdot(\beta\cdot\mathbf{A})$. Para demostrar la igualdad sea

$$\begin{aligned}(\alpha\beta)\cdot\mathbf{A} &= (\alpha\beta)\cdot(a_{ij}) = ((\alpha\beta)a_{ij}) = (\alpha(\beta a_{ij}))\\ &= \alpha\cdot(\beta a_{ij}) = \alpha\cdot(\beta\cdot(a_{ij})) = \alpha\cdot(\beta\cdot\mathbf{A}).\end{aligned}$$

4. Para toda matriz $\mathbf{A} \in \mathcal{M}_{n\times m}(\mathbb{K})$ se verifica que $1\cdot\mathbf{A} = \mathbf{A}$. Para demostrarlo consideramos

$$1\cdot\mathbf{A} = 1\cdot(a_{ij}) = (1a_{ij}) = (a_{ij}) = \mathbf{A}.$$

Estas propiedades junto con las propiedades de la suma de matrices hace que la terna $(\mathcal{M}_{n\times m}(\mathbb{K}), +, \cdot)$ tenga estructura de *espacio vectorial*, como veremos con mayor detalle en el tema dedicado a los espacios vectoriales.

1.3 Matriz traspuesta

Dada una matriz $\mathbf{A} = (a_{ij})_{i=1,2,\ldots,n}^{j=1,2,\ldots,m} \in \mathcal{M}_{n\times m}(\mathbb{K})$, se define su *matriz traspuesta* de $\mathbf{A}$ como una matriz denotada $\mathbf{A}^t \in \mathcal{M}_{m\times n}(\mathbb{K})$ y que se contruye intercambiando las filas por las columnas de $\mathbf{A}$, esto es, $\mathbf{A}^t = (a_{ij}^t)_{i=1,2,\ldots,m}^{j=1,2,\ldots,n}$ cumple $a_{ij}^t = a_{ji}$. Por ejemplo la traspuesta de

$$\begin{pmatrix} 2 & 3 \\ 1 & 1 \\ 2 & 1 \end{pmatrix}$$

será la matriz

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

Veamos cúales son las propiedades de la trasposición de matrices.

Proposition 1 *Sean $\alpha \in \mathbb{K}$, $\mathbf{A}, \mathbf{B} \in \mathcal{M}_{m\times n}(\mathbb{K})$ y $\mathbf{C} \in \mathcal{M}_{n\times l}(\mathbb{K})$. Entonces*

(a) $(\mathbf{A}^t)^t = \mathbf{A}$.

(b) $(\alpha \cdot \mathbf{A})^t = \alpha \cdot \mathbf{A}^t$.

(c) $(\mathbf{A} + \mathbf{B})^t = \mathbf{A}^t + \mathbf{B}^t$.

(d) $(\mathbf{A} \cdot \mathbf{C})^t = \mathbf{C}^t \cdot \mathbf{A}^t$.

(e) Si $\mathbf{A} \in \mathcal{M}_{n\times n}(\mathbb{K})$ es una matriz invertible, entonces $(\mathbf{A}^{-1})^t = (\mathbf{A}^t)^{-1}$.

Demostración. La propiedad (a) es inmediata. Para probar (b) calculamos

$$\alpha \cdot \mathbf{A}^t = \alpha \cdot (a_{ij}^t) = \alpha \cdot (a_{ji}) = (\alpha a_{ji}) = (\alpha \cdot \mathbf{A})^t.$$

Para demostrar (c) calculamos

$$\mathbf{A}^t + \mathbf{B}^t = (a_{ij}^t) + (b_{ij}^t) = (a_{ji}) + (b_{ji}) = (a_{ji} + b_{ji}) = ((a_{ij} + b_{ij})^t) = (\mathbf{A} + \mathbf{B})^t.$$

Para la demostración de (d)

$$\mathbf{C}^t \cdot \mathbf{A}^t = (c_{ij}^t) \cdot (a_{ij}^t) = \left(\sum_{k=1}^n c_{ik}^t a_{kj}^t \right) = \left(\sum_{k=1}^n a_{jk} c_{ki} \right) = (\mathbf{A} \cdot \mathbf{C})^t.$$

Para demostar la última propiedad, sea $\mathbf{A}^{-1}$ la matriz inversa de $\mathbf{A}$. Entonces $\mathbf{A} \cdot \mathbf{A}^{-1} = \mathbf{I}_n$. Si aplicamos la propiedad anterior tenemos que

$$(\mathbf{A} \cdot \mathbf{A}^{-1})^t = (\mathbf{A}^{-1})^t \cdot \mathbf{A}^t = \mathbf{I}_n^t = \mathbf{I}_n,$$

dado que $\mathbf{I}_n$ es una matriz diagonal y su inversa es ella misma. Ahora bien, tenemos que $(\mathbf{A}^{-1})^t \cdot \mathbf{A}^t = \mathbf{I}_n$, de donde deducimos que $\mathbf{A}^t$ es invertible y su matriz inversa $(\mathbf{A}^t)^{-1} = (\mathbf{A}^{-1})^t$. ■

La matriz traspuesta se utiliza para definir dos tipos especiales de matrices. Una matriz cuadrada $\mathbf{A}$ se dice *simétrica* si $\mathbf{A}^t = \mathbf{A}$ y se dice *antisimétrica* si $\mathbf{A}^t = -\mathbf{A}$. Si una matriz $\mathbf{A}$ es simétrica se verifica que

$$\mathbf{A} = (a_{ij}) = (a_{ij}^t) = \mathbf{A}^t,$$

por lo que

$$a_{ij} = a_{ij}^t = a_{ji}.$$

Así, ejemplos de matrices simétricas son

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 3 \\ 0 & 3 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 & -1 \\ 1 & -1 & 4 & 0 \\ 0 & 4 & 4 & 3 \\ -1 & 0 & 3 & 0 \end{pmatrix}.$$

Si por el contrario la matriz $\mathbf{A}$ es antisimétrica, entonces tenemos que

$$\mathbf{A} = (a_{ij}) = -(a_{ij}^t) = -\mathbf{A}^t,$$

por lo que

$$a_{ij} = -a_{ij}^t = -a_{ji},$$

y si $i = j$, entonces $2a_{ii} = 0$, por lo que $a_{ii} = 0$. Son ejemplos de matrices antisimétricas

$$\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & -3 \\ 0 & 3 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 & -1 \\ -1 & 0 & 4 & 0 \\ 0 & -4 & 0 & 3 \\ 1 & 0 & -3 & 0 \end{pmatrix}.$$

1.4 Rango de una matriz. Sistemas de ecuaciones lineales

Sea $\mathbf{A} \in \mathcal{M}_{n\times m}(\mathbb{K})$ una matriz. Recordemos que las n filas de la matriz las denotamos por $\mathbf{A}_1, \mathbf{A}_2, ..., \mathbf{A}_n$ y las m columnas por $\mathbf{A}^1, \mathbf{A}^2, ..., \mathbf{A}^m$. Las filas $\mathbf{A}_1, \mathbf{A}_2, ..., \mathbf{A}_n$ (resp. columnas $\mathbf{A}^1, \mathbf{A}^2, ..., \mathbf{A}^m$) son linealmente independientes si dada la expresión $\alpha_1 \cdot \mathbf{A}_1 + \alpha_2 \cdot \mathbf{A}_2 + ... + \alpha_n \cdot \mathbf{A}_n = \mathbf{0}$ (resp. $\alpha_1 \cdot \mathbf{A}^1 + \alpha_2 \cdot \mathbf{A}^2 + ... + \alpha_m \cdot \mathbf{A}^m = \mathbf{0}$) se deduce que $\alpha_i = 0$ para todo $1 \leq i \leq n$ (resp. $\alpha_i = 0$ para todo $1 \leq i \leq m$). Se llama *rango* de la matriz $\mathbf{A}$, denotado $r(\mathbf{A})$, al número máximo de filas o columnas linealmente independientes que hay en $\mathbf{A}$.

Consideremos la matriz

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

Calculemos su rango tomando por ejemplo las dos filas $\mathbf{A}_1 = (1, 1)$ y $\mathbf{A}_2 = (0, 1)$. Planteamos entonces la expresión

$$\alpha_1 \cdot \mathbf{A}_1 + \alpha_2 \cdot \mathbf{A}_2 = \alpha_1 \cdot (1, 1) + \alpha_2 \cdot (0, 1) = (\alpha_1, \alpha_1 + \alpha_2) = (0, 0),$$

de donde obtenemos el examen

$$\begin{cases} \alpha_1 = 0, \\ \alpha_1 + \alpha_2 = 0, \end{cases}$$

de donde obviamente $\alpha_1 = \alpha_2 = 0$.

Remark 1 *En principio habría que distinguir el rango por filas como el número máximo de filas linealmente independientes y el rango por columnas definido intercambiando filas por columnas. Puede probarse (ver por ejemplo [**?**]) que el rango por filas y por columnas coinciden. De esta manera cobra sentido la definición establecida anteriormente.*

En general, dada una matriz $\mathbf{A} = (a_{ij}) \in \mathcal{M}_{n\times m}(\mathbb{K})$, calcular su rango supone, en el caso de obtener el rango utilizando las columnas de la matriz, resolver expresiones de la forma

$$\alpha_1 \cdot \mathbf{A}^1 + ... + \alpha_m \cdot \mathbf{A}^m = \mathbf{0},$$

o equivalentemente

$$\mathbf{A} \cdot \begin{pmatrix} \alpha_1 \\ ... \\ \alpha_m \end{pmatrix} = \begin{pmatrix} 0 \\ ... \\ 0 \end{pmatrix},$$

que en forma extendida se escribe como

$$\begin{cases} a_{11}\alpha_1 + ... + a_{m1}\alpha_m = 0, \\ .... \\ a_{1n}\alpha_1 + ... + a_{mn}\alpha_m = 0. \end{cases}$$

Esto nos da pie para introducir el concepto de *sistema de ecuaciones lineales* como una expresión de la forma

$$\begin{cases} a_{11}\, x_1 + a_{12}\, x_2 + \cdots + a_{1m}\, x_m = b_1 \\ a_{21}\, x_1 + a_{22}\, x_2 + \cdots + a_{2m}\, x_m = b_2 \\ .... \\ a_{n1}\, x_1 + a_{n2}\, x_2 + \cdots + a_{nm}\, x_m = b_n \end{cases}$$

donde los escalares $a_{ij} \in \mathbb{K}$ $(1 \leq i \leq n,\ 1 \leq j \leq m)$ reciben el nombre de *coeficientes del sistema*, b_i $(b_i \in \mathbb{K})$ se denominan *términos independientes* y x_j $(1 \leq j \leq m)$ son las *incógnitas* del sistema. Si se considera la matriz formada por los coeficientes

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

llamada *matriz del sistema* y los vectores columna

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \quad \text{y} \quad \mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \in \mathbb{K}^n,$$

denominados *vector incógnita* y *vector de términos independientes* respectivamente, el sistema anterior puede reescribirse como:

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b},$$

expresión denominada *forma matricial del sistema*. Obsérvese cómo la notación matricial permite escribir los sistemas de ecuaciones lineales de forma manejable y compacta.

Definition 1 *Una solución del sistema anterior es una m-tupla de elementos* $x_1^*, \ldots, x_m^* \in \mathbb{K}$ *de forma que al sustituir cada incógnita por el* x_j^* *correspondiente se verifican todas las ecuaciones. Usando la notación matricial, una solución será un vector* $\mathbf{x}^* \in \mathbb{K}^m$ *tal que* $\mathbf{A} \cdot \mathbf{x}^* = \mathbf{b}$.

Por ejemplo consideremos el sistema

$$\left\{ \begin{array}{l} x_1 + x_2 + x_3 = 3, \\ 2x_1 + x_2 = 3, \\ x_1 - x_2 + x_3 = 1, \end{array} \right.$$

o de forma matricial

$$\begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \\ 1 \end{pmatrix}.$$

Veamos que la matriz $(1,1,1)^t$, es decir $x_1 = x_2 = x_3 = 1$, es solución del sistema. Para ello multiplicamos

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

1.4.1 Tipos de sistemas

Según su estructura se distinguen dos tipos de sistemas de ecuaciones lineales.

- Los *sistemas homogéneos* son aquellos en que los términos independientes son todos nulos, es decir, $\mathbf{b} = \mathbf{0}$, manteniendo la notación anterior. Son los sistemas que hemos de resolver para calcular el rango de una matriz.

- Cuando $\mathbf{b} \neq \mathbf{0}$, es decir, alguno de los términos independientes es no nulo, el sistema se dice que es *no homogéneo*.

Una característica importante de los sistemas homogéneos es que el vector nulo es siempre solución de los mismos, algo que no ocurre en el caso no homogéneo donde no siempre los sistemas tienen solución. Por ejemplo, el sistema

$$\left\{ \begin{array}{l} x_1 + x_2 = 1, \\ 2x_1 + 2x_2 = 0, \end{array} \right.$$

que no puede tener solución porque entonces obtendríamos que $0 = 1$, lo cual es absurdo.

La existencia o no de solución, induce otra clasificación en la clase de los sistemas de ecuaciones lineales.

- *Sistemas incompatibles*. Un sistema se dice que es incompatible si no admite solución.

- *Sistemas compatibles*. Un sistema es compatible si tiene solución. Se distinguen a su vez dos tipos de sistemas compatibles atendiendo al número de soluciones.

  - *Sistemas compatibles determinados*. Son aquellos que tienen una única solución.
  - *Sistemas compatibles indeterminados*. Son los que tienen más de una solución.

Veremos a continuación cómo caracterizar los distintos tipos de sistemas y cómo obtener la solución de los mismos en caso de existir.

1.4.2 El teorema de Rouché-Frobenius

Dado un sistema de ecuaciones lineales, el primer paso consiste en determinar su carácter, es decir, ver si es compatible y de qué tipo o incompatible. Este proceso es lo que se conoce como *discutir* el sistema. Sea entonces el sistema $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$, con $\mathbf{A} \in \mathcal{M}_{n\times m}(\mathbb{K})$ y definamos la matriz $\mathbf{A}|\mathbf{b} = (\mathbf{A}\ |\mathbf{b}) \in \mathcal{M}_{n\times(m+1)}(\mathbb{K})$, es decir, la matriz que se obtiene añadiendo $\mathbf{b}$ como columna a la matriz $\mathbf{A}$. Recordemos que $\mathbf{A}^1, \ldots, \mathbf{A}^m \in \mathbb{K}^n$ son las columnas de la matriz $\mathbf{A}$. Usando esta notación el sistema puede reescribirse como:

$$x_1 \cdot \mathbf{A}^1 + x_2 \cdot \mathbf{A}^2 + \cdots + x_m \cdot \mathbf{A}^m = \mathbf{b},$$

de donde se deduce que la existencia de solución es equivalente a que el vector $\mathbf{b}$ pueda obtenerse como combinación lineal de $\mathbf{A}^1, \ldots, \mathbf{A}^m$. Entonces se tiene que:

- El sistema es incompatible, es decir, no existe solución alguna del sistema si y sólo si $r(\mathbf{A}) < r(\mathbf{A}|\mathbf{b})$.
- El sistema es compatible si y sólo si $r(\mathbf{A}) = r(\mathbf{A}|\mathbf{b})$.

La anterior discusión del carácter del sistema en términos de los rangos de su matriz y de su matriz ampliada es la tesis del conocido Teorema de Rouché-Frobenius.

Theorem 2 (Rouché-Frobenius) *Sea $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$, con $\mathbf{A} \in \mathcal{M}_{n\times m}(\mathbb{K})$ y $\mathbf{b} \in \mathcal{M}_{n\times 1}(\mathbb{K})$. Se verifica que el sistema es compatible si y sólo si $r(\mathbf{A}) = r(\mathbf{A}|\mathbf{b})$.*

Consideremos el siguiente ejemplo:

$$\begin{cases} x_1 + x_2 + x_3 = 1, \\ 2x_2 - 2x_3 = 0, \\ x_3 = 1. \end{cases}$$

De forma matricial el sistema se escribe como

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & -2 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

Como vemos, la matriz $\mathbf{A}$ del sistema es triangular. En este caso es fácil obtener la solución del sistema ya que, como es fácil ver, $x_3 = 1$, una vez obtenida x_3 obtenemos de la segunda ecuación $x_2 = x_3 = 1$, y finalmente, una vez obtenidos x_2 y x_3 obtenemos $x_1 = 1 - x_2 - x_3 = -1$. Obtenemos entonces resolviendo el sistema "de abajo a arriba", teniendo por solución $(-1, 1, 1)$.

Por otra parte, es fácil ver que el rango de la matriz

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

es tres ya que vemos fácilmente que el sistema homogéneo

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & -2 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

tiene por única solución $\alpha_1 = \alpha_2 = \alpha_3 = 0$, por lo que las columnas $\mathbf{A}^1$, $\mathbf{A}^2$ y $\mathbf{A}^3$ son linealmente independientes. El rango de la matriz

$$\mathbf{B} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & -2 \\ 0 & 0 & 0 \end{pmatrix}$$

es dos ya que una fila de ceros es siempre linealmente dependiente con cualesquiera otras filas. Nótese que siempre es posible escribir

$$0 \cdot (1,1,1) + 0 \cdot (0,2,-2) + \alpha \cdot (0,0,0) = (0,0,0)$$

para cualquier $\alpha \in \mathbb{R}$. Así, podemos "eliminar" la fila de ceros y calcular el rango de la matriz

$$\mathbf{C} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & -2 \end{pmatrix},$$

que tendrá como máximo rango dos al tener únicamente dos filas. Como

$$\alpha_1 \cdot \mathbf{A}^1 + \alpha_2 \cdot \mathbf{A}^2 = \alpha_1 \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \alpha_2 \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

tiene por solución $\alpha_1 = \alpha_2 = 0$ ambas columnas son linealmente independientes y por tanto $r(\mathbf{C}) = r(\mathbf{B}) = 2$.

Vamos a ver en el siguiente apartado cómo obtener la solución de un sistema y el rango de matrices cuando éstas no son triangulares. Para ello necesitamos introducir una manera de manipular sistemas o matrices por medio de unas operaciones que llamaremos elementales.

1.4.3 Resolución de sistemas. Método de Gauss

La idea de este método es encontrar un sistema equivalente al original que sea sencillo de resolver. Para ello se efectúan transformaciones elementales sobre las filas del sistema de tal manera que ambos sistemas, el original y el transformado, tengan las mismas soluciones. Estas operaciones que llamaremos elementales son las siguientes:

1. Intercambiar dos filas del sistema.
2. Multiplicar una fila del sistema por un escalar $\alpha \in \mathbb{K}$ no nulo.
3. Suma a una fila del sistema otra fila multiplicada por un escalar $\alpha \in \mathbb{K}$.

En general, si denotamos por $\mathcal{O}$ una de estas operaciones elementales, escribiremos

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b} \underset{\mathcal{O}}{\rightarrow} \mathbf{A}(\mathcal{O}) \cdot \mathbf{x} = \mathbf{b}(\mathcal{O}),$$

donde $\mathbf{A}(\mathcal{O})$ y $\mathbf{b}(\mathcal{O})$ son las nuevas matrices del sistema transformado. Por ejemplo, dado el sistema

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 2 & 3 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 6 \end{pmatrix},$$

realizamos las siguientes operaciones elementales. Si intercambiamos la segunda y tercera fila, operación que denotaremos por $F_2 \times F_3$, escribiremos

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 2 & 3 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 6 \end{pmatrix} \underset{F_2\times F_3}{\rightarrow} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & -1 & 0 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 6 \\ 0 \end{pmatrix},$$

y en este caso, como

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

tendremos que

$$\mathbf{A}(\mathcal{O}) = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & -1 & 0 \end{pmatrix} \quad \text{y} \quad \mathbf{b}(\mathcal{O}) = \begin{pmatrix} 1 \\ 6 \\ 0 \end{pmatrix}.$$

Si al sistema original le hacemos la operación $2F_1$, esto es, multiplicamos la segunda fila por 2 obtenemos

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 2 & 3 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 6 \end{pmatrix} \underset{2F_2}{\rightarrow} \begin{pmatrix} 1 & 1 & 1 \\ 2 & -2 & 0 \\ 1 & 2 & 3 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 6 \end{pmatrix},$$

y ahora

$$\mathbf{A}(2F_2) = \begin{pmatrix} 1 & 1 & 1 \\ 2 & -2 & 0 \\ 1 & 2 & 3 \end{pmatrix}.$$

Finalmente, si le sumamos a la tercera fila la primera multimplicada por tres escribimos

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 2 & 3 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 6 \end{pmatrix} \underset{F_3+3F_1}{\rightarrow} \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 4 & 5 & 6 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 9 \end{pmatrix},$$

y

$$\mathbf{A}(F_3 + 3F_1) = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 4 & 5 & 6 \end{pmatrix}.$$

Démonos cuenta que toda operación elemental se puede revertir, esto es, existe otra operación elemental que nos devuelve al sistema original. Es fácil ver que si realizamos la operación intercambiar las filas i y j, $F_i \times F_j$, entonces realizando la misma operación volvemos al sistema original, esto es,

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b} \underset{F_i\times F_j}{\rightarrow} \mathbf{A}(F_i \times F_j) \cdot \mathbf{x} = \mathbf{b}(F_i \times F_j) \underset{F_i\times F_j}{\rightarrow} \mathbf{A} \cdot \mathbf{x} = \mathbf{b}.$$

Si la operación es multiplicar la fila i por $\alpha \in \mathbb{K} \setminus \{0\}$, entonces es fácil ver que

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b} \underset{\alpha F_i}{\rightarrow} \mathbf{A}(\alpha F_i) \cdot \mathbf{x} = \mathbf{b}(\alpha F_i) \underset{\frac{1}{\alpha}F_i}{\rightarrow} \mathbf{A} \cdot \mathbf{x} = \mathbf{b}.$$

Finalmente, si la operación es sumar a la fila i la fila j multiplicada por α, entonces

$$\mathbf{A}\cdot\mathbf{x}=\mathbf{b}\underset{F_i+\alpha F_j}{\rightarrow}\mathbf{A}(F_i+\alpha F_j)\cdot\mathbf{x}=\mathbf{b}(F_i+\alpha F_j)\underset{F_i-\alpha F_j}{\rightarrow}\mathbf{A}\cdot\mathbf{x}=\mathbf{b}.$$

Esta propiedad nos resultará útil para probar el siguiente resultado, que como veremos es clave.

Proposition 3 *Sea* $\mathbf{A}\cdot\mathbf{x}=\mathbf{b}$ *un sistema dado en forma matricial con* $\mathbf{A}\in\mathcal{M}_{n\times m}(\mathbb{K})$ *y* $\mathbf{b}\in\mathcal{M}_{n\times 1}(\mathbb{K})$. *Realicemos la operación elemental* $\mathcal{O}$ *sobre el sistema. Entonces los sistemas* $\mathbf{A}\cdot\mathbf{x}=\mathbf{b}$ *y* $\mathbf{A}(\mathcal{O})\cdot\mathbf{x}=\mathbf{b}(\mathcal{O})$ *tienen las mismas soluciones.*

Demostración. Para fijar ideas sea $\mathbf{A}=(a_{ij})$ y $\mathbf{b}=(b_i)$. Sea $\mathbf{x}_0=(x_1^0,...,x_m^0)^t$ una solución arbitraria del sistema $\mathbf{A}\cdot\mathbf{x}=\mathbf{b}$ y veamos que también es solución de $\mathbf{A}(\mathcal{O})\cdot\mathbf{x}=\mathbf{b}(\mathcal{O})$. Distingamos para ello los siguientes tres casos. Si $\mathcal{O}=F_i\times F_j$, entonces es obvio darse cuenta que $\mathbf{x}_0$ también es solución. Si multiplicamos la fila i por $\alpha\in\mathbb{K}\setminus\{0\}$, entonces

$$\alpha a_{i1}x_1^0+...+\alpha a_{im}x_m^0=\alpha(a_{i1}x_1^0+...+a_{im}x_m^0)=\alpha b_i.$$

Como las demás filas del sistema no han variado también se satisfacen las ecuaciones y por tanto $\mathbf{x}_0$ es solución de $\mathbf{A}(\alpha F_i)\cdot\mathbf{x}=\mathbf{b}(\alpha F_i)$. Finalmente, si la operación es $F_i+\alpha F_j$, tenemos que

$$(a_{i1}+\alpha a_{j1})x_1^0+...+(a_{im}+\alpha a_{jm})x_m^0=a_{i1}x_1^0+...+a_{im}x_m^0+\alpha(a_{j1}x_1^0+...+a_{jm}x_m^0)=b_i+\alpha b_j.$$

Como de nuevo las demás filas del sistema no han variado también se satisfacen las ecuaciones y por tanto $\mathbf{x}_0$ es solución de $\mathbf{A}(F_i+\alpha F_j)\cdot\mathbf{x}=\mathbf{b}(F_i+\alpha F_j)$.

Hemos probado entonces que toda $\mathbf{x}_0$ solución de $\mathbf{A}\cdot\mathbf{x}=\mathbf{b}$, también es solución de $\mathbf{A}(\mathcal{O})\cdot\mathbf{x}=\mathbf{b}(\mathcal{O})$. Veamos ahora el recíproco. Para ello, sea $\mathcal{O}'$ la operación elemental tal que

$$\mathbf{A}\cdot\mathbf{x}=\mathbf{b}\underset{\mathcal{O}}{\rightarrow}\mathbf{A}(\mathcal{O})\cdot\mathbf{x}=\mathbf{b}(\mathcal{O})\underset{\mathcal{O}'}{\rightarrow}\mathbf{A}\cdot\mathbf{x}=\mathbf{b},$$

y entonces, vemos que igualmente y por la misma razón toda solución de $\mathbf{A}(\mathcal{O})\cdot\mathbf{x}=\mathbf{b}(\mathcal{O})$ lo es de $\mathbf{A}\cdot\mathbf{x}=\mathbf{b}$. ■

Veamos entonces como aprovechar esta propiedad para obtener las soluciones del sistema. Para fijar ideas, sea $\mathbf{A}\cdot\mathbf{x}=\mathbf{b}$ un sistema dado en forma matricial con $\mathbf{A}\in\mathcal{M}_{n\times m}(\mathbb{K})$. El *método de Gauss* consta de las siguientes etapas:

- En primer lugar se realizan transformaciones elementales por filas sobre la matriz ampliada del sistema $\mathbf{A}|\mathbf{b}$ hasta obtener una matriz equivalente de la forma

$$\left(\begin{array}{c|c}\mathbf{B}&\mathbf{c}\\\hline\mathbf{0}&\mathbf{d}\end{array}\right),$$

  con $\mathbf{B}$ una matriz lo más sencilla posible (triangular en la mayoría de los casos). El sistema original tiene las mismas soluciones que el sistema

$$\left(\begin{array}{c}\mathbf{B}\\\hline\mathbf{0}\end{array}\right)\cdot\mathbf{x}=\left(\begin{array}{c}\mathbf{c}\\\hline\mathbf{d}\end{array}\right),$$

  de donde tenemos que:

- Si $\mathbf{d} \neq \mathbf{0}$ el sistema es incompatible.
- Si $\mathbf{d} = \mathbf{0}$, entonces el sistema es compatible y equivalente a $\mathbf{B} \cdot \mathbf{x} = \mathbf{c}$, que es un sistema cuyas soluciones son fáciles de calcular.

Pongamos un ejemplo. Consideremos el sistema

$$\left\{ \begin{array}{l} 2x + 2y - 3z = 2, \\ -x + 5y - 4z = 4, \\ x + 7y - 7z = 7. \end{array} \right.$$

Escribimos la matriz asociada ampliada

$$\mathbf{A}|\mathbf{b} = \begin{pmatrix} 2 & 2 & -3 & 2 \\ -1 & 5 & -4 & 4 \\ 1 & 7 & -7 & 7 \end{pmatrix}$$

y hacemos operaciones elementales fila

$$\begin{pmatrix} 2 & 2 & -3 & 2 \\ -1 & 5 & -4 & 4 \\ 1 & 7 & -7 & 7 \end{pmatrix} \xrightarrow[F_1 \times F_3]{} \begin{pmatrix} 1 & 7 & -7 & 7 \\ -1 & 5 & -4 & 4 \\ 2 & 2 & -3 & 2 \end{pmatrix} \xrightarrow[F_2 + F_1]{} \begin{pmatrix} 1 & 7 & -7 & 7 \\ 0 & 12 & -11 & 11 \\ 2 & 2 & -3 & 2 \end{pmatrix} \xrightarrow[F_3 - 2F_1]{}$$

$$\xrightarrow[F_3 - 2F_1]{} \begin{pmatrix} 1 & 7 & -7 & 7 \\ 0 & 12 & -11 & 11 \\ 0 & -12 & 11 & 12 \end{pmatrix} \xrightarrow[F_3 + F_2]{} \begin{pmatrix} 1 & 7 & -7 & 7 \\ 0 & 12 & -11 & 11 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

por lo que $r(\mathbf{A}) = 2$ mientras que $r(\mathbf{A}|\mathbf{b}) = 3$ y el sistema es incompatible.

Consideremos ahora el sistema

$$\left\{ \begin{array}{l} x + 2y - 3z = 4, \\ 2x + 3y - 6z = 1 \\ -x - y + z = -2. \end{array} \right.$$

Su matriz ampliada es

$$\mathbf{A}|\mathbf{b} = \begin{pmatrix} 1 & 2 & -3 & 4 \\ 2 & 3 & -6 & 1 \\ -1 & -1 & 1 & -2 \end{pmatrix}$$

y haciendo operaciones elementales fila tenemos

$$\begin{pmatrix} 1 & 2 & -3 & 4 \\ 2 & 3 & -6 & 1 \\ -1 & -1 & 1 & -2 \end{pmatrix} \xrightarrow[F_2 - 2F_1]{} \begin{pmatrix} 1 & 2 & -3 & 4 \\ 0 & -1 & 0 & -7 \\ -1 & -1 & 1 & -2 \end{pmatrix} \xrightarrow[F_3 + F_1]{} \begin{pmatrix} 1 & 2 & -3 & 4 \\ 0 & -1 & 0 & -7 \\ 0 & 1 & -2 & 2 \end{pmatrix} \xrightarrow[F_3 + F_2]{}$$

$$\xrightarrow[F_3 + F_2]{} \begin{pmatrix} 1 & 2 & -3 & 4 \\ 0 & -1 & 0 & -7 \\ 0 & 0 & -2 & -5 \end{pmatrix},$$

que como vemos es una matriz triangular que da lugar al sistema

$$\left\{ \begin{array}{r} x + 2y - 3z = 4, \\ -y = -7 \\ -2z = -5, \end{array} \right.$$

de donde fácilmente obtenemos la solución

$$(x, y, z) = (-5/2, 7, 5/2).$$

Veamos otro ejemplo con el sistema

$$\begin{cases} x - 2y + 3z = 2, \\ 2x + 5y + 6z = 0. \end{cases}$$

Escribimos la matriz ampliada

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

y hacemos operaciones elementales fila

$$\begin{pmatrix} 1 & -2 & 3 & 2 \\ 2 & 5 & 6 & 0 \end{pmatrix} \underset{F_2-2F_1}{\longrightarrow} \begin{pmatrix} 1 & -2 & 3 & 2 \\ 0 & 9 & 0 & -4 \end{pmatrix},$$

de donde el sistema original tendrá las mismas soluciones que el sistema

$$\begin{cases} x - 2y + 3z = 2, \\ 9y = -4, \end{cases}$$

que fácilmente vemos que tiene por soluciones

$$(x, y, z) = (26/9 - 3t, -4/9, t), \ t \in \mathbb{R}.$$

Como el sistema original tiene las mismas soluciones que el sistema

$$\left(\frac{\mathbf{B}}{\mathbf{0}}\right) \cdot \mathbf{x} = \left(\frac{\mathbf{c}}{\mathbf{d}}\right),$$

con $\mathbf{B}$ una matriz triangular. Si el sistema es compatible, entonces $\mathbf{d} = \mathbf{0}$ y si el número total de incógnitas es m, entonces sólo podemos despejar $r(\mathbf{B})$ incónitas. El resto de incógnitas hemos de renunciar a calcularlas y asignarles un valor real arbitrario (parámetro). Entonces la solución del sistema dependerá de $m - r(\mathbf{B})$ parámetros. En el ejemplo anterior teníamos 3 incógnitas mientras que el rango de la matriz era dos, y así la solución dependía de un parámetro.

1.5 Operaciones elementales en matrices

Para calcular el rango de una matriz de forma práctica necesitamos unas herramientas que se conocen con el nombre de *operaciones elementales* fila y columna, y que son totalmente análogas a las operaciones elementales que hemos estudiado al resolver sistemas de ecuaciones lineales. Así, dada una matriz $\mathbf{A} \in \mathcal{M}_{n\times m}(\mathbb{K})$, estas operaciones son tres:

1. Intercambiar dos filas (columnas) de la matriz $\mathbf{A}$.

2. Multiplicar una fila (columna) por un escalar $\alpha \in \mathbb{K}$ no nulo.

3. Suma a una fila (columna) otra fila (columna) multiplicada por un escalar $\alpha \in \mathbb{K}$.

Para evitar equívocos que se pueden producir al utilizar indistintamente operaciones elementales fila y columna, a partir de ahora vamos a trabajar únicamente con operaciones elementales fila. Fijemos un poco de notación para entendernos. Consideremos una matriz $\mathbf{A}$ y realizemos sobre ella una operación elemental fila $\mathcal{O}$. Escribiremos entonces

$$\mathbf{A} \underset{\mathcal{O}}{\longrightarrow} \mathbf{A}(\mathcal{O}),$$

donde $\mathbf{A}(\mathcal{O})$ es la amtriz resultante de realizar en $\mathbf{A}$ la operación $\mathcal{O}$. Dada por ejemplo

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

si intercambiamos la fila 2 por la fila 3 escribiremos

$$\begin{pmatrix} 1 & 0 & 2 & -1 \\ 2 & 1 & 0 & 4 \\ -1 & 2 & 0 & 1 \end{pmatrix} \underset{F_2 \times F_3}{\longrightarrow} \begin{pmatrix} 1 & 0 & 2 & -1 \\ -1 & 2 & 0 & 1 \\ 2 & 1 & 0 & 4 \end{pmatrix} = \mathbf{A}(F_2 \times F_3).$$

Si muliplicamos la primera fila de $\mathbf{A}$ por 2 escribimos

$$\begin{pmatrix} 1 & 0 & 2 & -1 \\ 2 & 1 & 0 & 4 \\ -1 & 2 & 0 & 1 \end{pmatrix} \underset{2F_1}{\longrightarrow} \begin{pmatrix} 2 & 0 & 4 & -2 \\ 2 & 1 & 0 & 4 \\ -1 & 2 & 0 & 1 \end{pmatrix} = \mathbf{A}(2F_1).$$

Si sumamos a la primera fila de $\mathbf{A}$ la segunda multiplicada por 2 escribimos

$$\begin{pmatrix} 1 & 0 & 2 & -1 \\ 2 & 1 & 0 & 4 \\ -1 & 2 & 0 & 1 \end{pmatrix} \underset{F_1+2F_2}{\longrightarrow} \begin{pmatrix} 5 & 2 & 2 & 7 \\ 2 & 1 & 0 & 4 \\ -1 & 2 & 0 & 1 \end{pmatrix} = \mathbf{A}(F_1 + 2F_2).$$

Podemos concatenar operaciones elementales filas

$$\begin{pmatrix} 1 & 0 & 2 & -1 \\ 2 & 1 & 0 & 4 \\ -1 & 2 & 0 & 1 \end{pmatrix} \underset{F_3+F_1}{\longrightarrow} \begin{pmatrix} 1 & 0 & 2 & -1 \\ 2 & 1 & 0 & 4 \\ 0 & 2 & 2 & 0 \end{pmatrix} \underset{2F_2}{\longrightarrow} \begin{pmatrix} 1 & 0 & 2 & -1 \\ 4 & 2 & 0 & 8 \\ 0 & 2 & 2 & 0 \end{pmatrix} \underset{F_3 \times F_2}{\longrightarrow} \begin{pmatrix} 1 & 0 & 2 & -1 \\ 0 & 2 & 2 & 0 \\ 4 & 2 & 0 & 8 \end{pmatrix}.$$

Idéntica notación tenemos para operaciones columna cambiando F por C, aunque no usaremos estas.en general. La utilidad de las operaciones elementales para calcular el rango de una matriz se concreta en el siguiente resultado.

Proposition 4 *Sea $\mathbf{A} \in \mathcal{M}_{n\times m}(\mathbb{K})$ y sea $\mathcal{O}$ una operación elemental. Entonces $r(\mathbf{A}) = r(\mathbf{A}(\mathcal{O}))$.*

Demostración. Sean $k = r(\mathbf{A})$ e $i_1, ..., i_k \in \{1, ..., m\}$ tal que $\mathbf{A}_{i_1}, ..., \mathbf{A}_{i_k}$ son columnas linealmente independientes. Si $\mathcal{O} = F_i \times F_j$, entonces las columnas de la matriz no cambian, solo el orden, y es claro que el número de columnas linealmente independientes es idéntico, con lo que

$r(\mathbf{A}) = r(\mathbf{A}(F_i \times F_j))$. Supongamos que $\mathcal{O} = \alpha F_i$, $\alpha \in \mathbb{K} \setminus \{0\}$. Si $i \notin \{i_1, ..., i_k\}$, entonces las columnas $\mathbf{A}_{i_1}, ..., \mathbf{A}_{i_k}$ no han cambiado y $r(\mathbf{A}) \leq r(\mathbf{A}(\alpha F_i))$. Si $i \in \{i_1, ..., i_k\}$, por ejemplo $i = i_1$, entonces de la expresión $\alpha_1 \alpha \cdot \mathbf{A}_{i_1} + ... + \alpha_k \cdot \mathbf{A}_{i_k} = \mathbf{0}$, obtenemos por la independencia lineal de $\mathbf{A}_{i_1}, ..., \mathbf{A}_{i_k}$ que $\alpha_i = 0$ si $i = 2, ..., k$ y $\alpha\alpha_1 = 0$, de donde $\alpha_1 = 0$ ya que $\alpha \neq 0$. De nuevo $r(\mathbf{A}) \leq r(\mathbf{A}(\alpha F_i))$. Entonces $r(\mathbf{A}) \leq r(\mathbf{A}(\alpha F_i)) \leq r(\mathbf{A}(\alpha F_i)(\frac{1}{\alpha} F_i)) \leq r(\mathbf{A})$, por lo que $r(\mathbf{A}) = r(\mathbf{A}(\alpha F_i))$. Finalmente, supongamos que $\mathcal{O} = F_i + \alpha F_j$, $\alpha \in \mathbb{K}$. Si $i \notin \{i_1, ..., i_k\}$, entonces las columnas $\mathbf{A}_{i_1}, ..., \mathbf{A}_{i_k}$ no han cambiado y $r(\mathbf{A}) \leq r(\mathbf{A}(F_i + \alpha F_j))$. Si $i, j \in \{i_1, ..., i_k\}$, por ejemplo $i = i_1$ y $j = i_k$, entonces de la expresión

$$\alpha_1 \cdot (\alpha \cdot \mathbf{A}_j + \mathbf{A}_{i_1}) + ... + \alpha_k \cdot \mathbf{A}_{i_k} = \alpha_1 \cdot \mathbf{A}_{i_1} + ... + (\alpha\alpha_1 + \alpha_k) \cdot \mathbf{A}_{i_k} = \mathbf{0},$$

obtenemos por la independencia lineal de $\mathbf{A}_{i_1}, ..., \mathbf{A}_{i_k}$ que $\alpha_i = 0$ si $i = 1, ..., k-1$ y $\alpha\alpha_1 + \alpha_k = 0$, de donde $\alpha_k = 0$ ya que $\alpha_1 = 0$. De nuevo $r(\mathbf{A}) \leq r(\mathbf{A}(F_i + \alpha F_j))$. Si $i \in \{i_1, ..., i_k\}$, por ejemplo $i = i_1$, y distinguimos dos casos: si $\mathbf{A}_j, \mathbf{A}_{i_2}, ..., \mathbf{A}_{i_k}$ son linealmente independientes entonces $r(\mathbf{A}) \leq r(\mathbf{A}(F_i + \alpha F_j))$. En caso contrario existen $\beta_i \in \mathbb{K}$, $1 \leq i \leq k$, con $\beta_1 \neq 0$ de manera que

$$\beta_1 \cdot \mathbf{A}_j + \beta_2 \cdot \mathbf{A}_{i_2} + ... + \beta_k \cdot \mathbf{A}_{i_k} = \mathbf{0},$$

de donde

$$\mathbf{A}_j = \gamma_2 \cdot \mathbf{A}_{i_2} + ... + \gamma_k \cdot \mathbf{A}_{i_k}$$

con $\gamma_i = -\beta_i/\beta_1$, $2 \leq i \leq k$. Escribamos

$$\alpha_1 \cdot (\alpha \cdot \mathbf{A}_j + \mathbf{A}_{i_1}) + ... + \alpha_k \cdot \mathbf{A}_{i_k} = \alpha_1 \cdot \mathbf{A}_{i_1} + ... + (\alpha\alpha_1\gamma_k + \alpha_k) \cdot \mathbf{A}_{i_k} = \mathbf{0}$$

y como $\mathbf{A}_{i_1}, ..., \mathbf{A}_{i_k}$ son linealmente independientes $\alpha_1 = 0$ y así $\alpha\alpha_1\gamma_i + \alpha_i = \alpha_i = 0$, $2 \leq i \leq k$, por la independencia lineal de $\mathbf{A}_{i_2}, ..., \mathbf{A}_{i_k}$. Así $r(\mathbf{A}) \leq r(\mathbf{A}(F_i + \alpha F_j))$. Como

$$r(\mathbf{A}) \leq r(\mathbf{A}(F_i + \alpha F_j)) \leq r(\mathbf{A}(F_i + \alpha F_j)(F_i - \alpha F_j)) \leq r(\mathbf{A}),$$

obtenemos que $r(\mathbf{A}) \leq r(\mathbf{A}(F_i + \alpha F_j))$. ■

Las operaciones elementales conservan entonces el rango al transformar la matriz. Un método para calcular el rango de una matriz se basa en hacer operaciones elementales fila en una matriz hasta obtener una matriz triangular. Por ejemplo vamos a calcular el rango de la matriz del ejemplo anterior. Para ello hacemos operaciones elementales fila buscando una matriz triangular

$$\begin{pmatrix} 1 & 0 & 2 & -1 \\ 2 & 1 & 0 & 4 \\ -1 & 2 & 0 & 1 \end{pmatrix} \xrightarrow[F_2 - 2F_1]{} \begin{pmatrix} 1 & 0 & 2 & -1 \\ 0 & 1 & -2 & 5 \\ -1 & 2 & 0 & 1 \end{pmatrix} \xrightarrow[F_3 + F_1]{} \begin{pmatrix} 1 & 0 & 2 & -1 \\ 0 & 1 & -2 & 5 \\ 0 & 2 & 2 & 0 \end{pmatrix} \xrightarrow[F_3 - 2F_1]{} \begin{pmatrix} 1 & 0 & 2 & -1 \\ 0 & 1 & -2 & 5 \\ 0 & 0 & 6 & -10 \end{pmatrix},$$

y la última matriz es triangular y es fácil ver que sus tres primeras son linealmente independientes ya que el sistema

$$\alpha_1 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \alpha_2 \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \alpha_3 \cdot \begin{pmatrix} 2 \\ -2 \\ 6 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

tiene como única solución $\alpha_1 = \alpha_2 = \alpha_3 = 0$, por lo que su rango es tres.

1.6 Cálculo de matrices inversas

Las operaciones elementales también pueden emplearse para verificar si una matriz cuadrada es invertible y para calcular su inversa en caso de que lo sea. Para ello hemos de introducir las *matrices elementales*. Para fijar ideas, sea $\mathbf{I}_n$ la matriz identidad. Se llama matriz elemental a aquella que se obtiene al efectuar una operación elemental a $\mathbf{I}_n$. Por ejemplo, las matrices

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

son matrices elementales. Entonces se verifica la siguiente propiedad que es la clave que permite utilizar las operaciones elementales para calcular matrices inversas.

Proposition 5 *Sean* $\mathbf{A} \in \mathcal{M}_{n\times n}(\mathbb{K})$ *y* $\mathcal{O}$ *una operación elemental fila. Entonces* $\mathbf{A}(\mathcal{O}) = \mathbf{I}_n(\mathcal{O}) \cdot \mathbf{A}$.

Demostración. Supongamos en primer lugar que $\mathcal{O} = F_k \times F_l$. Entonces si $\mathbf{A}(F_k \times F_l) = (a^*_{ij})$, entonces $a^*_{kj} = a_{lj}$, $a^*_{lj} = a_{kj}$ y $a^*_{ij} = a_{ij}$ si $i \notin \{k, l\}$. Si denotamos por $\delta_{ij} = 1$ si $i = j$ y $\delta_{ij} = 0$ si $i \neq j$, entonces tenemos que $\mathbf{I}_n(F_k \times F_l) = (\delta^*_{ij})$, donde $\delta^*_{kj} = \delta_{lj}$, $\delta^*_{lj} = \delta_{kj}$ y $\delta^*_{ij} = \delta_{ij}$ si $i \notin \{k, l\}$. Entonces

$$\mathbf{I}_n(F_k \times F_l) \cdot \mathbf{A} = \left(\sum_{s=1}^{n} \delta^*_{is} a_{sj} \right) = \left\{ \begin{array}{c} a_{kj} \text{ si } i = l, \\ a_{lj} \text{ si } i = k, \\ a_{ij} \text{ si } i \neq k, l, \end{array} \right\} = (a^*_{ij}) = \mathbf{A}(F_k \times F_l).$$

Si $\mathcal{O} = \alpha F_k$, $\alpha \neq 0$, entonces $\mathbf{A}(\alpha F_k) = (a^*_{ij})$, con $a^*_{kj} = \alpha a_{kj}$ y $a^*_{ij} = a_{ij}$ si $i \neq k$. Similarmente $\mathbf{I}_n(\alpha F_k) = (\delta^*_{ij})$, donde $\delta^*_{kj} = \alpha \delta_{kj}$ y $\delta^*_{ij} = \delta_{ij}$ si $i \neq k$. Entonces

$$\mathbf{I}_n(\alpha F_k) \cdot \mathbf{A} = \left(\sum_{s=1}^{n} \delta^*_{is} a_{sj} \right) = \left\{ \begin{array}{c} \alpha a_{kj} \text{ si } i = k, \\ a_{ij} \text{ si } i \neq k, \end{array} \right\} = (a^*_{ij}) = \mathbf{A}(\alpha F_k).$$

Finalmente, supongamos que $\mathcal{O} = F_k + \alpha F_l$ y sean de nuevo $\mathbf{A}(F_k + \alpha F_l) = (a^*_{ij})$, con $a^*_{kj} = a_{kj} + \alpha a_{lj}$ y $a^*_{ij} = a_{ij}$ si $i \neq k$, e $\mathbf{I}_n(F_k + \alpha F_l) = (\delta^*_{ij})$, donde $\delta^*_{kj} = \delta_{kj} + \alpha \delta_{lj}$ y $\delta^*_{ij} = \delta_{ij}$ si $i \neq k$. Entonces

$$\mathbf{I}_n(F_k + \alpha F_l) \cdot \mathbf{A} = \left(\sum_{s=1}^{n} \delta^*_{is} a_{sj} \right) = \left\{ \begin{array}{c} a_{kj} + \alpha a_{lj} \text{ si } i = k, \\ a_{ij} \text{ si } i \neq k, \end{array} \right\} = (a^*_{ij}) = \mathbf{A}(F_k + \alpha F_l),$$

lo que concluye la prueba. ■

Para ejemplificar el cómo podemos aprovechar el resultado anterior para caracterizar matrices invertibles y obtener a la vez la matriz inversa, tomemos

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

y calculamos su inversa mediante operaciones elementales. Para ello realizamos operaciones elementales fila en la matriz buscando conseguir la matriz identidad $\mathbf{I}_3$.

$$
\begin{aligned}
\mathbf{A} &= \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \xrightarrow[F_2-F_1]{} \begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & 0 \\ 1 & 1 & 0 \end{pmatrix} \xrightarrow[F_3-F_1]{} \begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \\
&: \xrightarrow[F_1+F_3]{} \begin{pmatrix} 1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \xrightarrow[F_1+F_2]{} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \\
&: \xrightarrow[-1\cdot F_2]{} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \xrightarrow[-1\cdot F_3]{} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \mathbf{I}_3.
\end{aligned}
$$

Ahora bien, por la propiedad anterior, tenemos que

$$\mathbf{I}_3 = \mathbf{I}_3(-1\cdot F_3)\cdot\mathbf{I}_3(-1\cdot F_2)\cdot\mathbf{I}_3(F_1+F_2)\cdot\mathbf{I}_3(F_1+F_3)\cdot\mathbf{I}_3(F_3-F_1)\cdot\mathbf{I}_3(F_2-F_1)\cdot\mathbf{A},$$

de donde deducimos que la matriz $\mathbf{A}$ es invertible y

$$\mathbf{A}^{-1} = \mathbf{I}_3(-1\cdot F_3)\cdot\mathbf{I}_3(-1\cdot F_2)\cdot\mathbf{I}_3(F_1+F_2)\cdot\mathbf{I}_3(F_1+F_3)\cdot\mathbf{I}_3(F_3-F_1)\cdot\mathbf{I}_3(F_2-F_1).$$

Pero en vez de multiplicar todas estas matrices elementales, nos damos cuenta de que

$$\mathbf{A}^{-1} = \mathbf{I}_3(-1\cdot F_3)\cdot\mathbf{I}_3(-1\cdot F_2)\cdot\mathbf{I}_3(F_1+F_2)\cdot\mathbf{I}_3(F_1+F_3)\cdot\mathbf{I}_3(F_3-F_1)\cdot\mathbf{I}_3(F_2-F_1)\cdot\mathbf{I}_3,$$

y de nuevo por la propiedad anterior, la inversa la obtendremos haciendo las mismas operaciones elementales que hicimos en $\mathbf{A}$ pero ahora las hacemos sobre la identidad y obtenemos

$$
\begin{aligned}
\mathbf{I}_3 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow[F_2-F_1]{} \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow[F_3-F_1]{} \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \\
&: \xrightarrow[F_1+F_3]{} \begin{pmatrix} 0 & 0 & 1 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \xrightarrow[F_1+F_2]{} \begin{pmatrix} -1 & 1 & 1 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \\
&: \xrightarrow[-1\cdot F_2]{} \begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \xrightarrow[-1\cdot F_3]{} \begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix} = \mathbf{A}^{-1}.
\end{aligned}
$$

Para ahorrar tiempo, se suelen escribir juntas la matriz $\mathbf{A}$ y la identidad y realizar las operaciones fila un única vez sobre la matriz formada por $\mathbf{A}$ y la identidad, como en el siguiente ejemplo. Sea

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

Tomamos entonces la matriz

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

y realizamos las operaciones elementales

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{array}\right) \underset{F_2-F_1}{\longrightarrow} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{array}\right)$$

$$\underset{F_3-F_1}{\longrightarrow} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 1 & 0 \\ 0 & 1 & 1 & -1 & 0 & 1 \end{array}\right) \underset{F_3-F_2}{\longrightarrow} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & 0 & -1 & 1 \end{array}\right),$$

de donde la matriz inversa

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

1.7 Determinantes de matrices cuadradas. Definición

Dada una matriz cuadrada $\mathbf{A} \in \mathcal{M}_{n\times n}(\mathbb{K})$, se llama *determinante* de $\mathbf{A}$ a un elemento del cuerpo $\mathbb{K}$, denotado por $|\mathbf{A}|$ o $\det \mathbf{A}$, asociado a la matriz mediante la siguiente fórmula de recurrencia:

- Si $\mathbf{A} = (a_{11}) \in \mathcal{M}_{1\times 1}(\mathbb{K})$, entonces $|\mathbf{A}| = a_{11}$.
- Si $\mathbf{A} = (a_{ij}) \in \mathcal{M}_{n\times n}(\mathbb{K})$, entonces suponiendo conocidos los determinantes de matrices de orden $n-1$ se define:

  $$|\mathbf{A}| = \sum_{j=1}^{n} a_{1j}(-1)^{1+j}|\mathbf{\Delta}_{1j}|$$

  siendo $\mathbf{\Delta}_{1j}$ la matriz cuadrada de orden $n-1$ resultante de eliminar la primera fila y la j-ésima columna de $\mathbf{A}$.

De esta definición se deducen de forma inmediata las fórmula usuales para calcular determinantes de matrices de orden dos y tres. Así:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

En el caso de matrices cuadradas de orden tres:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = (a_{11}a_{22}a_{33} + a_{12}\, a_{31}a_{23} + a_{13}a_{21}a_{32}) - (a_{11}a_{23}a_{32} + a_{12}a_{21}a_{33} + a_{13}a_{31}a_{22}),$$

relación conocida como *regla de Sarrus*.

Así

$$\begin{vmatrix} 1 & 0 \\ 1 & 2 \end{vmatrix} = 2,$$

y

$$\begin{vmatrix} 1 & 0 & 0 \\ 1 & 3 & 0 \\ 0 & 1 & 1 \end{vmatrix} = 3.$$

En general, si $\mathbf{A} \in \mathcal{M}_{n\times n}(\mathbb{K})$ es diagonal entonces $|\mathbf{A}| = a_{11}a_{22}...a_{nn}$. Si $n = 1$, la fórmula es trivialmente cierta. Si suponemos cierta la relación para matrices de $\mathcal{M}_{(n-1)\times(n-1)}(\mathbb{K})$, entonces

$$|\mathbf{A}| = \sum_{j=1}^{n} a_{1j}(-1)^{1+j}|\mathbf{\Delta}_{1j}| = a_{11}|\mathbf{\Delta}_{11}| = a_{11}a_{22}...a_{nn},$$

dado que $\mathbf{\Delta}_{11} \in \mathcal{M}_{(n-1)\times(n-1)}(\mathbb{K})$ es la matriz triangular que tiene $a_{22}, ..., a_{nn}$ como coeficientes en la diagonal principal.

Remark 2 *La definición de determinante de una matriz cuadrada $\mathbf{A} = (a_{ij}) \in \mathcal{M}_{n\times n}(\mathbb{K})$ que hemos dado aquí no es la que suele darse en los libros de matemáticas como [**?**]. Esta definición más usual de determinante se basa en la noción de permutación, que es una aplicación biyectiva $\sigma : \{1, ..., n\} \to \{1, ..., n\}$. Si $\mathcal{S}_n$ es el conjunto de todas las permutaciones definidas sobre $\{1, ..., n\}$, entonces el determinante de la matriz $\mathbf{A}$ es*

$$|\mathbf{A}| = \sum_{\sigma\in\mathcal{S}_n} s(\sigma)a_{1\sigma(1)}...a_{n\sigma(n)},$$

donde $s(\sigma)$ es 1 o -1 y se conoce como signo de la permutación. Por ejemplo, si $n = 2$, entonces sólo hay dos permutaciones σ_1 y σ_2 que vienen dadas por $\sigma_1(1) = 1$ (y por lo tanto $\sigma_1(2) = 2$) y $\sigma_2(1) = 2$ (y $\sigma_2(2) = 1$). Los signos son $s(\sigma_1) = 1$ y $s(\sigma_2) = -1$ y entonces

$$\begin{aligned} \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} &= s(\sigma_1)a_{1\sigma_1(1)}a_{2\sigma_1(2)} + s(\sigma_2)a_{1\sigma_2(1)}a_{2\sigma_2(2)} \\ &= a_{11}a_{22} - a_{12}a_{21}, \end{aligned}$$

que es la definición que hemos dado para el determiante de matrices de orden o tamaño dos. A partir de esta definición más rigurosa se pueden probar todas las propiedades que daremos en la siguiente sección. Aquellas que no probemos pueden probarse a partir de esta definición, pero cuya demostración no es sencilla con la defínción inductiva que hemos adoptado, que a su vez, tiene la ventaja de entrar más directamente en el cálculo práctico de los determinantes.

1.7.1 Propiedades

Sean $\mathbf{A}, \mathbf{B} \in \mathcal{M}_{n\times n}(\mathbb{K})$ se verifican las siguientes propiedades:

D1. $|\mathbf{A}| = \sum_{j=1}^{n} a_{ij}\,(-1)^{i+j}|\mathbf{\Delta}_{ij}|$, para cada $1 \leq i \leq n$, y $|\mathbf{A}| = \sum_{i=1}^{n} a_{ij}(-1)^{i+j}|\mathbf{\Delta}_{ij}|$, para cada $1 \leq j \leq n$, donde $\mathbf{\Delta}_{ij}$ es la matriz resultante de eliminar la fila i y la columna j de la matriz original y se llama *menor* de orden i, j de $\mathbf{A}$.

D2. $|\mathbf{A} \cdot \mathbf{B}| = |\mathbf{A}| \cdot |\mathbf{B}|$.

D3. $|\mathbf{A}| = |\mathbf{A}^t|$.

Demostración. Es consecuencia de la propiedad D1. ■

D4. Si cambiamos dos filas o columnas de orden el determinante cambia el signo, es decir, $|\mathbf{A}(F_i \times F_j)| = |\mathbf{A}(C_i \times C_j)| = -|\mathbf{A}|$.

Demostración. Supongamos es primer lugar que $j = i + 1$. Si $\mathbf{\Delta}_{ij}$ es el menor de orden i, j de $\mathbf{A}$ y $\mathbf{\Phi}_{(i+1)j}$ es el menor de orden $i + 1, j$ de $\mathbf{A}(F_i \times F_{i+1})$, entonces es fácil darse cuenta de que $\mathbf{\Delta}_{ij} = \mathbf{\Phi}_{(i+1)j}$. Entonces, por D1 y dado que las filas i de $\mathbf{A}$ e $i + 1$ de $\mathbf{A}(F_i \times F_{i+1})$ son iguales tenemos que

$$|\mathbf{A}| = \sum_{j=1}^{n} a_{ij}(-1)^{i+j}|\mathbf{\Delta}_{ij}| = -\sum_{j=1}^{n} a_{ij}(-1)^{i+1+j}|\mathbf{\Phi}_{(i+1)j}| = -|\mathbf{A}(F_i \times F_{i+1})|.$$

Para probar la propiedad general démonos cuenta que podemos obtener $\mathbf{A}(F_i \times F_j)$ intercambiando $2(j - i) - 1$ filas contiguas en $2(j - i) - 1$ operaciones fila elementales, por lo que

$$|\mathbf{A}(F_i \times F_j)| = (-1)^{2(j-i)-1}|\mathbf{A}| = -|\mathbf{A}|.$$

Finalmente, por D3

$$|\mathbf{A}(C_i \times C_j)| = |\mathbf{A}(C_i \times C_j)^t| = |\mathbf{A}^t(F_i \times F_j)| = -|\mathbf{A}^t| = -|\mathbf{A}|,$$

que concluye la demostración. ■

D5. Si $\mathbf{A}$ tienes dos filas o columnas iguales, entonces $|\mathbf{A}| = 0$.

Demostración. Si $\mathbf{A}_i = \mathbf{A}_j$, entonces $\mathbf{A} = \mathbf{A}(F_i \times F_j)$ y por D4 se tendría

$$|\mathbf{A}| = -|\mathbf{A}(F_i \times F_j)| = -|\mathbf{A}|,$$

de donde $2|\mathbf{A}| = 0$ y por tanto $|\mathbf{A}| = 0$. La prueba en caso de dos columnas iguales es idéntica. ■

D6. Si sumamos a una fila o columna de $\mathbf{A} \in \mathbf{M}_{n\times n}(\mathbb{K})$ otra fila o columna multiplicada por un escalar el determinante no varía, esto es $|\mathbf{A}| = |\mathbf{A}(F_i + \alpha F_j)| = |\mathbf{A}(C_i + \alpha C_j)|$.

Demostración. Si $\mathbf{\Delta}_{ij}$ es el menor de orden i, j de $\mathbf{A}$ y $\mathbf{\Phi}_{ij}$ es el menor de orden i, j de $\mathbf{A}(F_i + \alpha F_j)$, entonces es fácil darse cuenta de que $\mathbf{\Delta}_{ij} = \mathbf{\Phi}_{ij}$ dado que la única fila distinta en ambas matrices es la i que es eliminada al obtener el menor. Entonces, utilizando D1 calculamos

$$\begin{aligned}
|\mathbf{A}(F_i + \alpha F_j)| &= \sum_{k=1}^{n}(a_{ik} + \alpha a_{jk})(-1)^{i+k}|\mathbf{\Phi}_{ik}| \\
&= \sum_{k=1}^{n} a_{ik}(-1)^{i+k}|\mathbf{\Delta}_{ik}| + \alpha\sum_{k=1}^{n} a_{jk}(-1)^{j+k}|\mathbf{\Delta}_{jk}| = |\mathbf{A}|,
\end{aligned}$$

dado que $\sum_{k=1}^{n} a_{jk}(-1)^{j+k}|\mathbf{\Delta}_{jk}| = 0$ es el determinante de una matriz que tiene las filas i y j iguales. Por D3 se prueba que

$$|\mathbf{A}(C_i + \alpha C_j)| = |\mathbf{A}(C_i + \alpha C_j)^t| = |\mathbf{A}^t(F_i + \alpha F_j)| = |\mathbf{A}^t| = |\mathbf{A}|,$$

con lo que la propiedad queda probada. ■

D7. Si se multiplica una fila o columna de una matriz $\mathbf{A}$ por un escalar $\alpha \neq 0$, entonces el determinante de $\mathbf{A}$ queda multiplicado por α. Es decir, $|\mathbf{A}(\alpha F_i)| = |\mathbf{A}(\alpha C_i)| = \alpha \cdot |\mathbf{A}|$.

Demostración. Si $\mathbf{\Delta}_{ij}$ es el menor de orden i, j de $\mathbf{A}$ y $\mathbf{\Phi}_{ij}$ es el menor de orden i, j de $\mathbf{A}(\alpha F_i)$, entonces es fácil darse cuenta de que $\mathbf{\Delta}_{ij} = \mathbf{\Phi}_{ij}$ dado que la única fila distinta en ambas matrices es la i que es eliminada al obtener el menor. Entonces, utilizando D1 calculamos

$$|\mathbf{A}(\alpha F_i)| = \sum_{j=1}^{n} \alpha a_{ij}(-1)^{i+j}|\mathbf{\Phi}_{ij}| = \alpha \sum_{j=1}^{n} a_{ij}(-1)^{i+j}|\mathbf{\Delta}_{ij}| = \alpha|\mathbf{A}|.$$

Por D3 se prueba que

$$|\mathbf{A}(\alpha C_i)| = |\mathbf{A}(\alpha C_i)^t| = |\mathbf{A}^t(\alpha F_i)| = \alpha \cdot |\mathbf{A}^t| = \alpha \cdot |\mathbf{A}|,$$

con lo que la propiedad queda probada. ■

D8. Si una matriz $\mathbf{A} \in \mathcal{M}_{n\times n}(\mathbb{K})$ tiene una fila o columna que es combinación lineal de las otras, entonces $|\mathbf{A}| = 0$. En particular, si una fila o columna de $\mathbf{A}$ es nula entonces su determinante es nulo.

Demostración. Supongamos por ejemplo que existen $i, i_1, ..., i_k \in \{1, ..., n\}$ tal que $\mathbf{A}_i = \alpha_1 \cdot \mathbf{A}_{i_1} + ... + \alpha_k \cdot \mathbf{A}_{i_k}$, con $\alpha_{i_j} \in \mathbb{K}$, $1 \leq j \leq k$. Entonces la matriz $\mathbf{A}(F_i - \alpha_1 F_{i_1})...(F_i - \alpha_k F_{i_k})$ tiene la fila i nula. Si $\mathbf{\Delta}_{ij}$ es el menor de orden i, j de dicha matriz, por D6 y D1 tenemos

$$|\mathbf{A}| = |\mathbf{A}(F_i - \alpha_1 F_{i_1})...(F_i - \alpha_k F_{i_k})| = \sum_{j=1}^{n} 0(-1)^{i+j}|\mathbf{\Delta}_{ij}| = 0,$$

con lo que concluye la prueba. ■

Una consecuencia de esta propiedad D8 que es de utilidad para calcular rangos de matrices es que si $|\mathbf{A}| \neq 0$, entonces $r(\mathbf{A})$ coincide con el número de filas de $\mathbf{A}$. Así, si por ejemplo

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 & 2 \\ 9 & 2 & 2 & 2 \end{pmatrix},$$

dado que

$$\begin{vmatrix} 1 & 0 \\ 9 & 2 \end{vmatrix} = 2 \neq 0,$$

se tiene que el rango de $\mathbf{A}$ es dos.

D9. Se verifica

$$\begin{vmatrix} a_{11} & ... & a_{1n} \\ ... & ... & ... \\ a_{i1} + a'_{i1} & ... & a_{n1} + a'_{n1} \\ ... & ... & ... \\ a_{n1} & ... & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & ... & a_{1n} \\ ... & ... & ... \\ a_{i1} & ... & a_{n1} \\ ... & ... & ... \\ a_{n1} & ... & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & ... & a_{1n} \\ ... & ... & ... \\ a'_{i1} & ... & a'_{n1} \\ ... & ... & ... \\ a_{n1} & ... & a_{nn} \end{vmatrix}.$$

Demostración. Sean

$$\mathbf{A}=\begin{pmatrix} a_{11} & ... & a_{1n} \\ ... & ... & ... \\ a_{i1}+a'_{i1} & ... & a_{n1}+a'_{n1} \\ ... & ... & ... \\ a_{n1} & ... & a_{nn} \end{pmatrix},$$

$$\mathbf{B}=\begin{pmatrix} a_{11} & ... & a_{1n} \\ ... & ... & ... \\ a_{i1} & ... & a_{n1} \\ ... & ... & ... \\ a_{n1} & ... & a_{nn} \end{pmatrix}$$

y

$$\mathbf{C}=\begin{pmatrix} a_{11} & ... & a_{1n} \\ ... & ... & ... \\ a'_{i1} & ... & a'_{n1} \\ ... & ... & ... \\ a_{n1} & ... & a_{nn} \end{pmatrix}.$$

Entonces los menores de orden i,j de $\mathbf{A}$, $\mathbf{B}$ y $\mathbf{C}$ son iguales y por tanto por D1

$$\begin{aligned} |\mathbf{A}| &= \sum_{j=1}^{n}(a_{ij}+a'_{ij})(-1)^{i+j}|\mathbf{\Delta}_{ij}| \\ &= \sum_{j=1}^{n} a_{ij}(-1)^{i+j}|\mathbf{\Delta}_{ij}| + \sum_{j=1}^{n} a'_{ij}(-1)^{i+j}|\mathbf{\Delta}_{ij}| = |\mathbf{B}|+|\mathbf{C}|, \end{aligned}$$

con lo que concluye la demostración. ■

Todas estas propiedades son de utilidad tanto desde el punto de cista teórico como a la hora de calcular determinantes de matrices grandes. Por ejemplo, calculemos

$$\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \end{vmatrix} \underset{F_2-F_1}{=} \begin{vmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \end{vmatrix} \underset{F_3-2F_1}{=} \begin{vmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 1 & 2 & 3 \\ 3 & 4 & 5 & 6 \end{vmatrix} = 0,$$

al tener la última matriz dos filas iguales.

1.7.2 Cálculo de la matriz inversa usando determinantes.

Dada $\mathbf{A}\in\mathcal{M}_{n\times n}(\mathbb{K})$ se define su *matriz adjunta* como la matriz cuadrada de orden n cuyos elementos son los adjuntos de $\mathbf{A}$. Es decir, $\bar{\mathbf{A}}\in\mathcal{M}_{n\times n}(\mathbb{K})$ es la matriz adjunta de $\mathbf{A}$ si $\bar{\mathbf{A}}=((-1)^{i+j}|\mathbf{\Delta}_{ij}|)$. Si $A=(a_{ij})$, de la definición de producto de matrices se sigue que

$$\mathbf{A}\cdot\bar{\mathbf{A}}^t=\left(\sum_{k=1}^{n} a_{ik}(-1)^{k+j}|\mathbf{\Delta}_{jk}|\right)$$

Usando ahora las propiedades de los determinantes:

- Si $i = j$:

$$\sum_{k=1}^{n} a_{ik}(-1)^{k+i}|\mathbf{\Delta}_{ik}| = |\mathbf{A}|.$$

- Si $i \neq j$:

$$\sum_{k=1}^{n} a_{ik}(-1)^{k+j}|\mathbf{\Delta}_{jk}| = \begin{array}{r} \\ \\ \\ \\ \text{fila } j \\ \\ \\ \end{array}\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = 0,$$

de donde se deduce que $\mathbf{A} \cdot \bar{\mathbf{A}}^t = |\mathbf{A}| \cdot \mathbf{I}_n$. Si $\mathbf{A}$ es invertible su determinante es no nulo, lo que permite obtener la fórmula:

$$\mathbf{A}^{-1} = |\mathbf{A}|^{-1} \cdot \bar{\mathbf{A}}^t$$

para el cálculo de la matriz inversa. Por ejemplo, si

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

entonces

$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|} \cdot \begin{pmatrix} 2 & 0 & 0 \\ -4 & 2 & 0 \\ -1 & -1 & 1 \end{pmatrix}^t = \frac{1}{2} \cdot \begin{pmatrix} 2 & -4 & -1 \\ 0 & 2 & -1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -2 & -\frac{1}{2} \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & \frac{1}{2} \end{pmatrix}.$$

Nótese que una matriz cuadrada con determinante no nulo es invertible por lo que su rango coincidirá con el número de filas de la matriz. Este hecho puede ayudar a calcular el rango de matrices. Por ejemplo, la matriz

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

verifica que

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

por lo que su rango es dos.

1.7.3 Resolución de sistemas de ecuaciones. Regla de Cramer

Veamos cómo los sistemas de ecuaciones lineales pueden resolverse mediante el uso de detrminantes. Se trata de un método aplicable cuando el sistema tiene la misma cantidad de ecuaciones que de incógnitas (este tipo de sistemas se llaman *cuadrados*). Bajo estas condiciones el sistema es compatible determinado si y sólo si el determinante de su matriz es no nulo. Sea un sistema $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ con

$\mathbf{A} \in \mathcal{M}_{n\times n}(\mathbb{K})$ y tal que $|\mathbf{A}| \neq 0$. Evidentemente la matriz $\mathbf{A}$ es invertible por lo que la solución única del sistema será $\mathbf{x} = \mathbf{A}^{-1}\cdot\mathbf{b}$ y usando la fórmula para la matriz inversa mediante determinantes se llega a que la solución es de la forma:

$$\mathbf{x} = |\mathbf{A}|^{-1}\cdot\bar{\mathbf{A}}^t\cdot\mathbf{b} = \frac{1}{|\mathbf{A}|}\cdot\begin{pmatrix} \sum_{k=1}^n(-1)^{k+1}|\mathbf{\Delta}_{k1}|b_k \\ \sum_{k=1}^n(-1)^{k+2}|\mathbf{\Delta}_{k2}|b_k \\ \cdots\cdots\cdots\cdots \\ \sum_{k=1}^n(-1)^{k+n}|\mathbf{\Delta}_{kn}|b_k \end{pmatrix},$$

de donde es inmediata la *regla de Cramer*.

Theorem 6 (Regla de Cramer) *Dado el sistema*

$$\left.\begin{array}{c} a_{11}\,x_1+\cdots+a_{1n}\,x_n=b_1 \\ a_{21}\,x_1+\cdots+a_{2n}\,x_n=b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}\,x_1+\cdots+a_{nn}\,x_n=b_n \end{array}\right\}$$

con

$$|\mathbf{A}| = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} \neq 0,$$

se tiene que es compatible determinado y la solución única viene dada por las fórmulas:

$$x_j = |\mathbf{A}|^{-1}\cdot\begin{vmatrix} a_{11} & \cdots & a_{1j-1} & b_1 & a_{1j+1} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2j-1} & b_2 & a_{2j+1} & \cdots & a_{2n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj-1} & b_n & a_{nj+1} & \cdots & a_{nn} \end{vmatrix}, \quad 1 \leq j \leq n.$$

La regla de Cramer puede adaptarse para resolver sistemas de ecuaciones lineales no necesariamente cuadrados y de manera que el determinante de la matriz sea nulo en caso de sistemas no cuadrados. Así dado $\mathbf{A}\cdot\mathbf{x} = \mathbf{b}$ un sistema con n ecuaciones y m incógnitas ($n < m$). Suponiendo que $r(\mathbf{A}) = n$, el sistema es compatible indeterminado como consecuencia del Teorema de Rouché-Frobenius. Suponiendo además que las n primeras columnas de $\mathbf{A}$ son linealmente independientes, se tiene que asignando parámetros a las variables asociadas a las últimas $m-n$ columnas $x_j = \mu_j,\ n+1 \leq j \leq m$, el valor de las restantes variables viene dado por el sistema:

$$\mathbf{B}\cdot\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \mathbf{b} - \sum_{j=n+1}^{m}\mu_j\cdot\mathbf{A}^j,$$

con $B = (\mathbf{A}^1, ..., \mathbf{A}^n) \in \mathcal{M}_{n\times n}(\mathbb{K})$. Evidentemente el sistema anterior es cuadrado y $|\mathbf{B}| \neq 0$ ya que $r(\mathbf{B}) = n$, por lo que puede resolverse usando la regla de Cramer. Por ejemplo, el sistema

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

verifica que

$$\begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix} = -2, \quad y \quad \begin{vmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 2 & 0 & 1 \end{vmatrix} = 0,$$

mientras que

$$\begin{vmatrix} 1 & 1 & 3 \\ 1 & -1 & 0 \\ 2 & 0 & 3 \end{vmatrix} = 0,$$

por lo que los rangos de las matrices del sistema es dos y el sistema es compatible. Despejamos la variable z, que no vamos a poder calcular

$$\begin{cases} x + y = 3 - z, \\ x - y = 0, \end{cases}$$

y entonces

$$\begin{aligned} x &= \frac{\begin{vmatrix} 3-z & 1 \\ 0 & -1 \end{vmatrix}}{\begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix}} = \frac{z-3}{2}, \\ y &= \frac{\begin{vmatrix} 1 & 3-z \\ 1 & 0 \end{vmatrix}}{\begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix}} = \frac{z-3}{2}, \end{aligned}$$

por lo que la solución del sistema es

$$\begin{cases} x = -\frac{3}{2} + \frac{\lambda}{2}, \\ y = -\frac{3}{2} + \frac{\lambda}{2}, \quad \lambda \in \mathbb{R}. \\ z = \lambda, \end{cases}$$

1.8 Ejercicios

1. Dadas las siguientes matrices realizar, si es posible, las operaciones que se indican:

$$\mathbf{A} = \begin{pmatrix} 1 & -1 & 2 \end{pmatrix} \quad \mathbf{B} = \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix} \quad \mathbf{C} = \begin{pmatrix} 2+i & -1 & 1 \\ 0 & 1 & 2 \\ 1 & 2+2i & 0 \end{pmatrix}$$

$$\mathbf{D} = \begin{pmatrix} 2 & 1 \\ 1 & 3 \\ 0 & -1 \end{pmatrix} \quad \mathbf{E} = \begin{pmatrix} i & 1 \\ 2 & 1-i \end{pmatrix}$$

(a) $(2\mathbf{A} + 3\mathbf{B}^t) \cdot \mathbf{C}$ (b) $\mathbf{C}^t \cdot \mathbf{A}^t - (\mathbf{A} \cdot \mathbf{C})^t$ (c) $\mathbf{D} \cdot \mathbf{C} + \mathbf{E}$ (d) $\mathbf{D} \cdot (\mathbf{E} + \mathbf{C})$

(e) $(\mathbf{C} - 2\mathbf{I}_3)^t \cdot (\mathbf{B} + \mathbf{A}^t)$ (f) $\mathbf{B} \cdot \mathbf{A} \cdot \mathbf{C} \cdot \mathbf{D}$ (g) $\mathbf{E} \cdot \mathbf{D}^t \cdot \mathbf{C}$ (h) $\mathbf{B}^t \cdot \mathbf{A}^t \cdot \mathbf{C} \cdot \mathbf{D}$

2. Calcular el rango de las siguientes matrices

$$\mathbf{A}=\begin{pmatrix}1&1&2&3\\2&1&1&2\\2&3&-1&2\end{pmatrix}\quad \mathbf{B}=\begin{pmatrix}1&0&1\\1&1&1\\2&1&2\end{pmatrix}\quad \mathbf{C}=\begin{pmatrix}2&-1&3&3\\1&1&3&3\end{pmatrix}$$

3. Determinar en función del parámetro real α el rango de las siguientes matrices

$$\mathbf{A}=\begin{pmatrix}\alpha&1&2&3\\0&\alpha&1&2\\0&0&\alpha&2\end{pmatrix}\quad \mathbf{B}=\begin{pmatrix}1&\alpha&1\\1&1&\alpha\\2&1&2\end{pmatrix}\quad \mathbf{C}=\begin{pmatrix}2&\alpha&3&3\\1&1&3&3\end{pmatrix}$$

4. Determinar si las siguientes matrices son invertibles y calcular en caso afirmativo su matriz inversa

$$\mathbf{A}=\begin{pmatrix}1&0&1&0\\0&1&0&1\\1&0&0&0\\0&0&0&1\end{pmatrix}\quad \mathbf{B}=\begin{pmatrix}1&1&1\\-1&0&-1\\2&1&2\end{pmatrix}\quad \mathbf{C}=\begin{pmatrix}1&0&1\\0&1&1\\1&1&0\end{pmatrix}$$

5. Calcular el rango de la siguiente matriz en función de los valores de a y b:

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

6. Sea la matriz $\mathbf{A}=\begin{pmatrix}0&0&0&0\\2&0&0&0\\2&2&0&0\\2&2&2&0\end{pmatrix}$, se pide:

   (a) Calcular las sucesivas potencias de $\mathbf{A}$.

   (b) Sea $\mathbf{B}=\mathbf{I}_4+\mathbf{A}$, expresar $\mathbf{B}^n$ en función de $\mathbf{I}_4, \mathbf{A}, \mathbf{A}^2$ y $\mathbf{A}^3$.

   (c) Demostrar que la inversa de $\mathbf{B}$ es $\mathbf{I}_4-\mathbf{A}+\mathbf{A}^2-\mathbf{A}^3$.

7. Hallar la potencia n–ésima de $\mathbf{A}=\begin{pmatrix}1&2&3\\0&1&2\\0&0&1\end{pmatrix}$ poniendo $\mathbf{A}=\mathbf{I}_3+\mathbf{B}$, siendo $\mathbf{B}$ una matriz a determinar.

8. Dada la matriz $\mathbf{A}=\begin{pmatrix}3&1\\5&2\end{pmatrix}$, se pide:

   (a) Hallar $3\mathbf{A}\cdot\mathbf{A}^t-2\mathbf{I}_2$.

   (b) Resolver la ecuación matricial $\mathbf{A}\cdot\mathbf{X}=\begin{pmatrix}2&0\\0&1\end{pmatrix}$ siendo $\mathbf{X}\in\mathbf{M}_{2\times2}(\mathbb{R})$.

9. Sea $\mathbf{A}$ una matriz cuadrada tal que $\mathbf{A}^2 = \mathbf{A}$. Si $\mathbf{B} = 2\mathbf{A} - \mathbf{I}_n$, demostrar que $\mathbf{B}^2$ es igual a $\mathbf{I}_n$.

10. Sea $\mathbf{A}$ una matriz cuadrada. Demostrar que las matrices $\mathbf{A} \cdot \mathbf{A}^t$ y $\mathbf{A}^t \cdot \mathbf{A}$ son siempre simétricas.

11. Demostrar que si una matriz cuadrada $\mathbf{A}$ verifica que $\mathbf{A}^2 - \mathbf{A} - \mathbf{I}_2 = 0$, entonces existe la inversa de $\mathbf{A}$. Calcularla.

12. Se considera la matriz con coeficientes reales $\mathbf{A} = \begin{pmatrix} a^2 & ab & ac \\ ab & b^2 & bc \\ ac & bc & c^2 \end{pmatrix}$. Demostrar que si $a^2 + b^2 + c^2 = 1$, entonces $\mathbf{A}^n = \mathbf{A}$ para todo entero positivo.

13. De las afirmaciones siguientes, demostrar las verdaderas y dar un contraejemplo para las falsas:

    (a) $(\mathbf{A} + \mathbf{B})^2 = \mathbf{A}^2 + 2\mathbf{A} \cdot \mathbf{B} + \mathbf{B}^2$.

    (b) $\mathbf{A}^2 - \mathbf{B}^2 = (\mathbf{A} - \mathbf{B}) \cdot (\mathbf{A} + \mathbf{B})$.

    (c) $\mathbf{A}^{m+1} - \mathbf{I}_n = (\mathbf{A} - \mathbf{I}_n)(\mathbf{I}_n + \mathbf{A} + \mathbf{A}^2 + ... + \mathbf{A}^m)$.

    (d) Si $\mathbf{P}$ es una matriz con determinante no nulo, entonces $(\mathbf{P} \cdot \mathbf{A} \cdot \mathbf{P}^{-1})^n = \mathbf{P} \cdot \mathbf{A}^n \mathbf{P}^{-1}$.

    (e) Si $\mathbf{A}$ es antisimétrica, entonces $\mathbf{A}^2$ es simétrica.

    (f) Si $\mathbf{A}$ es antisimétrica y $\mathbf{B}$ es simétrica, entonces $\mathbf{A} \cdot \mathbf{B}$ es antisimétrica si y sólo si $\mathbf{A} \cdot \mathbf{B} = \mathbf{B} \cdot \mathbf{A}$.

14. Hallar todas las soluciones de los siguientes sistemas:

(a) $\left\{ \begin{array}{l} -x + y - 2z - t = 0 \\ 2x - y + z - 2t = 2 \\ x + 2y - z + t = 3 \\ 3x + 4y - 3z - t = 1 \end{array} \right.$ (b) $\left\{ \begin{array}{l} 2x + 2y - 3z = 2 \\ -x + 5y - 4z = 4 \\ x + 7y - 7z = 7 \end{array} \right.$ (c) $\left\{ \begin{array}{l} x - y = -1 \\ -x + y = 1 \\ 2x - 2y = -2 \end{array} \right.$

(d) $\left\{ \begin{array}{l} 4x - y + 2z + t = 0 \\ 2x + 3y - z - 2t = 0 \\ 7y - 4z - 5t = 0 \\ 2x - 11y + 7z + 8t = 0 \end{array} \right.$ (e) $\left\{ \begin{array}{l} x - 2y + 3z = 0 \\ 2x + 5y + 6z = 0 \end{array} \right.$ (f) $\left\{ \begin{array}{l} x + 2y + 3z = 0 \\ 2x + 2y + 3z = 0 \\ 3x + 2y + z = 0 \end{array} \right.$

(g) $\left\{ \begin{array}{l} 2x + y + 4z = 0 \\ x - y + 2z = 4 \\ 2x + y - z = 14 \\ 3x + z = 18 \end{array} \right.$ (h) $\left\{ \begin{array}{l} x + 2y - 3z + 16t = 4 \\ y + 2z - 3t = 6 \\ -x - y + z + 9t = -2 \end{array} \right.$ (i) $\left\{ \begin{array}{l} x + 2y - 3z = 4 \\ 2x + 4y - 6z = 1 \\ -x - y + z = -2 \end{array} \right.$

15. Discutir y resolver según el valor de los parámetros que aparezcan:

(a) $\left\{ \begin{array}{l} \alpha x + y + 2z = 0 \\ x + 3y + z = 0 \\ 3x + 10y + 4z = 0 \end{array} \right.$ (b) $\left\{ \begin{array}{l} 3x - y + 2z = 1 \\ x + 4y + z = \beta \\ 2x - 5y + \alpha z = -2 \end{array} \right.$ (c) $\left\{ \begin{array}{l} 2y - z = \alpha \\ 3x - 2z = 11 \\ y + z = 6 \\ 2x + y - 4z = \alpha \end{array} \right.$

(d) $\left\{ \begin{array}{l} 2x - y - z = 3a \\ x - az = b \\ x - y + 2z = 7 \end{array} \right.$ (e) $\left\{ \begin{array}{l} 2\lambda x + \mu y + 2z = 1 \\ 2\lambda x + (2\mu - 1)y + 3z = 1 \\ 2\lambda x + \mu y + (\mu + 3)z = 2\mu - 1 \end{array} \right.$ (g) $\left\{ \begin{array}{l} \alpha x + \beta y + z = 1 \\ x + \alpha\beta y + z = \beta \\ x + \beta y + \alpha z = 1 \end{array} \right.$

16. Discutir la veracidad o falsedad de las siguientes afirmaciones:

    (a) Dado un sistema de m ecuaciones con n incógnitas, $\mathbf{A}\cdot\mathbf{x}=\mathbf{b}$, que admite solución única, entonces ésta es $\mathbf{x}=\mathbf{A}^{-1}\cdot\mathbf{b}$.

    (b) Si los sistemas $\mathbf{A}\cdot\mathbf{x}=\mathbf{b}_1$ y $\mathcal{A}\cdot\mathbf{x}=\mathbf{b}_2$ son compatibles, entonces lo es $\mathbf{A}\cdot\mathbf{x}=\mathbf{b}$ donde $\mathbf{b}=\mathbf{b}_1+\mathbf{b}_2$.

    (c) Un sistema con más ecuaciones que incógnitas es siempre incompatible.

    (d) Si un sistema de ecuaciones $\mathbf{A}\cdot\mathbf{x}=\mathbf{b}$ es compatible determinado, entonces $\mathbf{A}$ es una matriz cuadrada.

    (e) Si $\mathbf{A}\cdot\mathbf{x}=\mathbf{b}$ es un sistema incompatible con 5 ecuaciones y 4 incógnitas y el $r(\mathbf{A})=4$ entonces $r(\mathbf{A}|\mathbf{b})=5$.

17. Calcular las soluciones de los siguientes sistemas de ecuaciones tanto por el método de Gauss, como por el método de Kramer (por determinantes).

$$\text{(a)}\ \left\{\begin{array}{c} 8x+y+4z=9 \\ 5x-2y+4z=6 \\ x+y=1 \end{array}\right. \quad \text{(b)}\ \left\{\begin{array}{c} 6x-y+3z=6 \\ -6x+8y=-10 \\ 2x-5y-z=4 \end{array}\right. \quad \text{(c)}\ \left\{\begin{array}{c} x+y+z=1 \\ 3x-4y=5 \\ 7x-y-3z=8. \end{array}\right.$$

18. Discutir los siguientes sistemas de ecuaciones según los valores de a y b:

$$\text{(a)}\ \left\{\begin{array}{c} ax+2z=2 \\ 5x+2y=1 \\ x-2y+bz=3 \end{array}\right. \quad \text{(b)}\ \left\{\begin{array}{c} ax+by+z=1 \\ x+aby+z=b \\ x+by+az=1 \end{array}\right. \quad \text{(c)}\ \left\{\begin{array}{c} ax+y+z+t=1 \\ x+ay+z+t=b \\ x+y+az+t=b^2 \\ x+y+z+at=b^3 \end{array}\right.$$

19. Sea ω un número complejo raiz cúbica de la unidad. Discutir el sistema:

$$\left\{\begin{array}{c} x+y+z=a \\ x+\omega y+\omega^2 z=b \\ x+\omega^2 y+\omega z=c \end{array}\right.$$

    donde a,b,c son números reales.

20. Se tienen tres lingotes de oro de 100 gramos cuya composición es la siguiente

Lingote	Oro	Plata	Cobre
1	20	30	50
2	30	40	30
3	40	50	10

    ¿Que peso habrá que tomarse de cada uno de los tres lingotes para formar uno nuevo que contenga 60 gramos de oro, 50 gramos de plata y 45 gramos de cobre?

21. La suma de las tres cifras de un número es igual a 6. La cifra de las centenas es igual a la suma de las cifras de unidad y decena. Si se invierte el orden de las cifras, el número disminuye en 198 unidades. Calcular dicho número.

22. Una empresa hortofructícola tiene tres factorías diferentes en Castellón, Valencia y Alicante. En la época de la naranja cada una de estas factorías se dedica a envasar tres variedades diferentes de naranjas: navalate, navel y satsuma. La capacidad de envasado de la factoría de Castellón es de 4000Kg. de navalate, 3000Kg. de navel y 5000Kg de satsuma, todo ello por hora. La de Valencia es de 1000Kg. por hora de las tres variedades de naranjas. La de Alicante es de 2000Kg. de navalate, 4000Kg. de navel y 3000Kg. se satsuma, también por hora. ¿Cuántas horas se debe trabajar en cada factoría para satisfacer los dos siguientes pedidos?

    (a) 19000Kg. de navalate, 25000Kg. de navel y 25000Kg. de satsuma

    (b) 13000Kg. de navalate, 16000Kg. de navel y 16000Kg. de satsuma.

23. Una empresa tiene dos tipos de procesos productivos: torno y fresadora. Cada uno de estos procesos se utiliza para fabricar tres tipos de productos A, B y C. Se dispone de 120 horas semanales de torno y de 260 horas de fresadora, y las necesidades asociadas a cada proceso, por unidad de producto, son las siguientes:

Producto	Torno	Fresadora
A	0.1h	0.20h
B	0.25h	0.30h
C	-	0.40h

    Si el beneficio unitario que se obtiene con la venta se los productos A, B y C es de 3, 5 y 4 unidades monetarias, respectivamente. ¿Cómo debe de distribuirse la producción semanal para obtener un beneficio de 3800 u.m., si se utilizan todos los recursos disponibles?

24. Una empresa se dedica a la fabricación de cuatro tipos de jabón. Desde la compra de materias primas hasta la disposición para la distribución se realizan las siguientes fases: I) se mezclan los dos tipos de materias primas utilizadas, grasa vegetal y sosa cáustica; II) se introduce la mezcla obtenida en unos moldes preparados al efecto; III) los bloques obtenidos en la fase anterior se cortan y troquean, y IV) las pastillas así obtenidas se envasan en cajas de cartón de doscientas unidades.

    Los recursos necesarios para producir los cuatro tipos de jabones, por caja fabricada, vienen dados en la tabla siguiente:

Jabón	Sección Kg. Grasa	Mezclado Kg. Sosa	S. Moldeado Hora/Máquina	S. Troquelado Hora/Máquina
J_1	20	10	10	3
J_2	25	15	8	4
J_3	40	20	10	7
J_4	50	22	15	20

    Si se dispone durante una semana de 1970 Kg. de grasa vegetal, 970 Kg. de sosa cáustica, 601 hora/máquina en la sección de moldeado y de 504 horas/máquina en la sección de troquelado, ¿cuántas cajas de jabones de cada tipo se pueden producir, utilizando todos los recursos disponibles, en una semana?

25. Un agricultor produce maíz, trigo y cebada en las 12 hanegadas de tierra que posee. Cada Kg. de cereal plantado precisa de una cierta cantidad de dinero y de un determinado número de horas de trabajo semanales para su cultivo. En la tabla siguiente se especifica el capital necesario (en miles de pesetas), el trabajo preciso (en horas semanales) y el beneficio que produce (en miles de pesetas) cada uno de los cereales:

	Capital	Trabajo	Beneficio
Maíz	36	6	40
Trigo	24	6	30
Cebada	18	2	20

Calcular cuántos Kg. deberá de cultivar de cada tipo de cereal para obtener un beneficio de 400 mil pesetas si dispone de 360.000 pesetas y decide trabajar 48 horas semanales.

26. La ley de corriente de Kirchhoff dice que la suma algebraica de todas las corrientes que confluyen en un nudo de un circuito es nula. La ley de Ohm dice que la corriente a través de una resistencia entre dos nudos es el cociente entre la diferencia de voltaje entre cada nudo y el valor de la resistencia Dado el circuito de la figura, calcular las intensidades y los voltajes en cada nudo.

circui.eps

27. En un vecindario viven un fontanero, un electricista y un pintor. Deciden hacer reparaciones en sus tres casas, para lo cual cada uno de ellos va a trabajar diez jornadas en total. El fontanero trabajará tres días en su casa, dos días en la casa del electricista y cinco días en la casa del pintor. El electricista trabajará tres días en su propia casa, otros tres días en la casa del fontanero y cuatro días en la casa del pintor. Finalmente el pintor dedicará cinco días a la casa del fontanero y otros cinco días a la casa del electricista.

    Calcula cuál debería de ser el sueldo diario de cada uno para que en las diez jornadas ninguno de ellos pierda ni gane dinero, sabiendo de antemano que la suma de los tres sueldos es de 20000 ptas al día.

28. Una ciudad tiene tres industrias principales: una mina de carbón, una central eléctrica y un ferrocarril local. Para obtener 10 ptas de carbón, se utilizan 2 ptas de electricidad para hacer funcionar el equipamiento y 4 ptas para transportarlo a los almacenes. Para producir 10 ptas de electricidad la central eléctrica necesita 5 ptas de carbón de combustible, 1 pta de su propia electricidad y una peseta para dedicarla al transporte. Finalmente, para obtener 10 ptas en transporte se necesitan 5 ptas de carbón y 1 pta de electricidad. En una semana, la mina de carbón recibe un pedido valorado en 100000 ptas y la central eléctrica recibe otro pedido de 200000 ptas. El ferrocarril no satisface ninguna demanda externa. ¿Qué cantidad deben producir cada una de las industrias para satisfacer en esa semana tanto la demanda interna como la externa?

29. Encontrar el polinomio de grado 2 cuya gráfica que pasa por los puntos (1,4), (2,9) y (3,8).

30. Encuentra un polinomio $p(x)$ que verifique que $p(0) = 1$, $p(1) = 0$, $p'(0) = -1$ y $p''(1) = 1$.

31. Halla la ecuación de una circunferencia que pase por los puntos $(0, 1)$, $(-1, 1)$ y $(1, 0)$.

32. En una placa circular se ha establecido un mallado como el que se indica en el dibujo. Sabiendo las temperaturas en los puntos de la malla situados en el borde y que la temperatura en los demás puntos es igual a la media de la temperatura en los cuatro puntos adyacentes, calcula la temperatura en todos los puntos del mallado.

temp.eps

33. Calcular los siguientes determinantes:

$$\text{(a)} \begin{vmatrix} 1 & 3 & 0 \\ -1 & 2 & -4 \\ 1 & 1 & 2 \end{vmatrix} \quad \text{(b)} \begin{vmatrix} 5 & -1 & 7 \\ 6 & 4 & 3 \\ 3 & 2 & 1 \end{vmatrix} \quad \text{(c)} \begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} \quad \text{(d)} \begin{vmatrix} 3 & 5 & 7 & 2 \\ 2 & 4 & 1 & 1 \\ -2 & 0 & 0 & 0 \\ 1 & 1 & 3 & 4 \end{vmatrix}$$

$$\text{(e)} \begin{vmatrix} 1 & 1 & 1 & 6 \\ 2 & 4 & 1 & 6 \\ 4 & 1 & 2 & 9 \\ 2 & 4 & 2 & 7 \end{vmatrix} \quad \text{(f)} \begin{vmatrix} 1 & -2 & 3 & -4 \\ 2 & -1 & 4 & -3 \\ 2 & 3 & -4 & -5 \\ 3 & -4 & 5 & 6 \end{vmatrix} \quad \text{(g)} \begin{vmatrix} 1 & 2i & 3i \\ 4 & 5-i & 6 \\ 7i & 8 & 9 \end{vmatrix} \quad \text{(h)} \begin{vmatrix} 3-i & 5 & 7 & 2 \\ 2-6i & 4 & 1 & 1 \\ -2 & 0 & 0 & 0 \\ 1 & 1-i & 3 & 4 \end{vmatrix}$$

34. Dada una matriz cuadrada $\mathbf{A}$, ¿qué valores puede tomar $|\mathbf{A}|$ si:

   (a) $\mathbf{A}^2 = \mathbf{A}$?

   (b) $\mathbf{A} = \mathbf{A}^{-1}$?

35. Demostrar que si a, b, c son números reales se tiene que:

$$\begin{vmatrix} a-b-c & 2a & 2a \\ 2b & b-c-a & 2b \\ 2c & 2c & c-a-b \end{vmatrix} = (a+b+c)^3 .$$

36. Calcular los siguientes determinantes:

$$\text{(a)} \begin{vmatrix} x+a & b & c \\ a & x+b & c \\ a & b & x+c \end{vmatrix} \quad \text{(b)} \begin{vmatrix} x & a & b & c \\ a & x & 0 & 0 \\ b & 0 & x & 0 \\ c & 0 & 0 & x \end{vmatrix} \quad \text{(c)} \begin{vmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \\ 3 & 4 & 1 & 2 \\ 4 & 1 & 2 & 3 \end{vmatrix}$$

$$\text{(d)} \begin{vmatrix} a & 3 & 0 & 5 \\ 0 & b & 0 & 2 \\ 1 & 2 & c & 3 \\ 0 & 0 & 0 & d \end{vmatrix} \quad \text{(e)} \begin{vmatrix} a & b & 0 & 0 \\ 0 & a & b & 0 \\ 0 & 0 & a & b \\ b & 0 & 0 & a \end{vmatrix} \quad \text{(f)} \begin{vmatrix} a^2 & ab & ab & b^2 \\ ab & a^2 & b^2 & ab \\ ab & b^2 & a^2 & ab \\ b^2 & ab & ab & a^2 \end{vmatrix}$$

37. Calcular los siguientes determinantes de Vadermonde:

$$V_2 = \begin{vmatrix} 1 & 1 \\ a & b \end{vmatrix} \quad V_3 = \begin{vmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{vmatrix} \quad V_3 = \begin{vmatrix} 1 & 1 & 1 & 1 \\ a & b & c & d \\ a^2 & b^2 & c^2 & d^2 \\ a^3 & b^3 & c^3 & d^3 \end{vmatrix}$$

$$V_n = \begin{vmatrix} 1 & 1 & 1 & 1 & ... & 1 \\ a & b & c & d & ... & x \\ a^2 & b^2 & c^2 & d^2 & ... & x^2 \\ a^3 & b^3 & c^3 & d^3 & ... & x^3 \\ ... & ... & ... & ... & ... & ... \\ a^5 & b^5 & c^5 & d^5 & ... & x^5 \end{vmatrix}.$$

38. Resolver la ecuación $\Delta(x) = 0$, siendo $\Delta(X) = \begin{vmatrix} a & b & c \\ a & x & c \\ a & b & x \end{vmatrix}$, siendo a,b,c números reales.

39. Calcular los siguientes determinantes de orden n:

(a) $\begin{vmatrix} 1 & n & n & n & ... & n \\ n & 2 & n & n & ... & n \\ n & n & 3 & n & ... & n \\ n & n & n & 4 & ... & n \\ ... & ... & ... & ... & ... & ... \\ n & n & n & n & ... & n \end{vmatrix}$ (b) $\begin{vmatrix} 1 & 2 & 3 & 4 & ... & n \\ -1 & 0 & 3 & 4 & ... & n \\ -1 & -2 & 0 & 4 & ... & n \\ -1 & -2 & -3 & 0 & ... & n \\ ... & ... & ... & ... & ... & ... \\ -1 & -2 & -3 & -4 & ... & 0 \end{vmatrix}$

(c) $\begin{vmatrix} 1 & 1 & 1 & 1 & ... & 1 \\ 1 & 2+a & 1 & 1 & ... & 1 \\ 1 & 1 & 2+a & 1 & ... & 1 \\ 1 & 1 & 1 & 2+a & ... & 1 \\ ... & ... & ... & ... & ... & ... \\ 1 & 1 & 1 & 1 & ... & 2+a \end{vmatrix}$ (d) $\begin{vmatrix} 1 & n & n & n & ... & n \\ n & 2 & n & n & ... & n \\ n & n & 3 & n & ... & n \\ n & n & n & 4 & ... & n \\ ... & ... & ... & ... & ... & ... \\ n & n & n & n & ... & n \end{vmatrix}$

40. Demostrar que $\begin{vmatrix} 1 & \cos x & \cos 2x \\ \cos x & \cos 2x & \cos 3x \\ \cos 2x & \cos 3x & \cos 4x \end{vmatrix} = 0.$

41. De las afirmaciones siguientes, demostrar las verdaderas y dar un contraejemplo para las falsas:

   (a) Si $|\mathbf{A} \cdot \mathbf{B}| = 0$, entonces $|\mathbf{A}| = 0$ ó $|\mathbf{B}| = 0$.
   (b) $|\mathbf{A} + \mathbf{B}| = |\mathbf{A}| + |\mathbf{B}|$.
   (c) $|2\mathbf{A}| = 2\,|\mathbf{A}|$.

Capítulo 2

Espacio vectorial

Sumario. Axiomas de espacio vectorial. Combinación lineal. Dependencia e independencia lineal. Subespacios vectoriales. Operaciones con subespacios vectoriales. Sistema generador. Base. Espacio vectorial finitamente generado. Dimensión de un subespacio vetorial.

La línea seguida en este tema puede verse en [**?**], aunque algunas demostraciones han sido elaboradas por el autor para adaptarlas al nivel del alumno. Otras referencias de interés, de entre la gran cantidad de textos sobre álgebra lineal existentes, son [**?**].

2.1 Definiciones y propiedades básicas

Sea $\mathbb{K}$ el cuerpo de los números reales o complejos y sea $\mathcal{V}$ un conjunto en el que hay definidas una operación interna $+$ de manera que a cada $\mathbf{u}, \mathbf{v} \in \mathcal{V}$ le asocia un elemento $\mathbf{u}+\mathbf{v} \in \mathcal{V}$, y una operación externa $\cdot$ de manera que a cada $\alpha \in \mathbb{K}$ y cada $\mathbf{u} \in \mathcal{V}$ le asocia un elemento $\alpha \cdot \mathbf{u} \in \mathcal{V}$, cumpliendo las siguientes propiedades para todo $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w} \in \mathcal{V}$ y para todo α, $\beta \in \mathbb{K}$:

1. Propiedad asociativa para $+$: $(\mathbf{u}+\mathbf{v})+\mathbf{w} = \mathbf{u}+(\mathbf{v}+\mathbf{w})$.
2. Propiedad conmutativa para $+$: $\mathbf{u}+\mathbf{v} = \mathbf{v}+\mathbf{u}$.
3. Existencia de elemento neutro $\mathbf{0}$ para $+$: $\mathbf{0}+\mathbf{u} = \mathbf{u}$.
4. Para todo $\mathbf{u} \in \mathcal{V}$ existe elemento inverso o simétrico $-\mathbf{u}$ de manera que $\mathbf{u}+(-\mathbf{u}) = \mathbf{0}$.
5. Propiedad distributiva respecto de la suma en $\mathcal{V}$, es decir, $\alpha \cdot (\mathbf{u}+\mathbf{v}) = \alpha \cdot \mathbf{u} + \alpha \cdot \mathbf{v}$.
6. Propiedad distributiva respecto de la suma en $\mathbb{K}$, esto es, $(\alpha+\beta) \cdot \mathbf{u} = \alpha \cdot \mathbf{u} + \beta \cdot \mathbf{u}$.
7. Propiedad pseudoasociativa: $(\alpha\beta) \cdot \mathbf{u} = \alpha \cdot (\beta \cdot \mathbf{u})$.
8. $1 \cdot \mathbf{u} = \mathbf{u}$.

Entonces se dice que la terna $(\mathcal{V}, +, \cdot)$ tiene estructura de *espacio vectorial* sobre el cuerpo $\mathbb{K}$. Ejemplos de espacios vectoriales son los siguientes:

Example 1 Como ya vimos en el tema de matrices el conjunto de las matrices $\mathcal{M}_{n\times m}(\mathbb{K})$ con la suma de matrices y el producto por escalares tiene estructura de espacio vectorial ya que satisface las 8 propiedades anteriores. Cuando las matrices tengan una única fila, entonces escribiremos el conjunto como $\mathbb{K}^m$ (el caso que más trataremos será el de $\mathbb{R}^n$).

Example 2 Sea $\mathcal{P}_n[x]$ el conjunto de los polinomios de grado menor o igual que $n \in \mathbb{N}$ con coeficientes en $\mathbb{K}$. Con la suma usual de polinomios y el producto usual por escalares este conjunto es un espacio vectorial sobre $\mathbb{K}$.

Example 3 Sea $\mathcal{C}([a,b])$, $a,b \in \mathbb{R}$, el conjunto de funciones continuas definidas sobre $[a,b]$. Dadas $f,g \in \mathcal{C}([a,b])$ y $\alpha \in \mathbb{R}$ definimos $(f+g)(x) = f(x)+g(x)$ y $(\alpha\cdot f)(x) = \alpha f(x)$ para todo $x \in [a,b]$. Es fácil comprobar que $\mathcal{C}([a,b])$ con estas operaciones es un espacio vectorial sobre $\mathbb{R}$.

Veamos ahora algunas propiedades básicas que se derivan directamente de los 8 axiomas de espacio vectorial.

Proposition 7 *Sea $\mathcal{V}$ un espacio vectorial sobre $\mathbb{K}$. Entonces:*

(a) $0\cdot\mathbf{u} = \mathbf{0}$ *para todo* $\mathbf{u} \in \mathcal{V}$.

(b) $\alpha\cdot\mathbf{0} = \mathbf{0}$ *para todo* $\alpha \in \mathbb{K}$.

(c) $\alpha\cdot\mathbf{u} = \mathbf{0}$ *si y solo si* $\alpha = 0$ *o* $\mathbf{u} = \mathbf{0}$.

(d) $(-\alpha)\cdot\mathbf{u} = -(\alpha\cdot\mathbf{u}) = \alpha\cdot(-\mathbf{u})$ *para todo* $\alpha \in \mathbb{K}$ *y todo* $\mathbf{u} \in \mathcal{V}$.

(e) Si $\alpha\cdot\mathbf{u} = \alpha\cdot\mathbf{v}$ *y* $\alpha \neq 0$, *entonces* $\mathbf{u} = \mathbf{v}$.

(f) Si $\alpha\cdot\mathbf{u} = \beta\cdot\mathbf{u}$ *y* $\mathbf{u} \neq \mathbf{0}$, *entonces* $\alpha = \beta$.

(g) $(-\alpha)\cdot(-\mathbf{u}) = \alpha\cdot\mathbf{u}$ *para todo* $\alpha \in \mathbb{K}$ *y todo* $\mathbf{u} \in \mathcal{V}$.

Demostración. (a) $0+0=0$ y multiplicando por $\mathbf{u}$ tenemos $(0+0)\cdot\mathbf{u} = 0\cdot\mathbf{u}$ de donde

$$(0+0)\cdot\mathbf{u} = 0\cdot\mathbf{u} + 0\cdot\mathbf{u} = 0\cdot\mathbf{u}.$$

Sumando a ambos miembros el inverso de $0\cdot\mathbf{u}$ tenemos que $0\cdot\mathbf{u} = \mathbf{0}$.

(b) Ahora tenemos que $\mathbf{0}+\mathbf{0} = \mathbf{0}$. Multiplicando por α tenemos $\alpha\cdot(\mathbf{0}+\mathbf{0}) = \alpha\cdot\mathbf{0}$ de donde

$$\alpha\cdot(\mathbf{0}+\mathbf{0}) = \alpha\cdot\mathbf{0} + \alpha\cdot\mathbf{0} = \alpha\cdot\mathbf{0}.$$

Sumando en ambos miembros el inverso de $\alpha\cdot\mathbf{0}$ tenermos que$\alpha\cdot\mathbf{0} = \mathbf{0}$.

(c) Si $\alpha = 0$ o $\mathbf{u} = \mathbf{0}$, por los apartados anteriores tenemos que $\alpha\cdot\mathbf{u} = \mathbf{0}$. Supongamos ahora que $\alpha\cdot\mathbf{u} = \mathbf{0}$. Si $\alpha = 0$ ya hemos terminado así que supongamos que $\alpha \neq 0$. Entonces multiplicamos por el inverso de α

$$\alpha^{-1}\cdot(\alpha\cdot\mathbf{u}) = \alpha^{-1}\cdot\mathbf{0} = \mathbf{0}$$

y como

$$\alpha^{-1}\cdot(\alpha\cdot\mathbf{u}) = (\alpha^{-1}\alpha)\cdot\mathbf{u} = 1\cdot\mathbf{u} = \mathbf{u},$$

tenemos que $\mathbf{u} = \mathbf{0}$.

(d) Por un lado $\alpha + (-\alpha) = 0$ de donde multiplicando por $\mathbf{u}$ tenemos que

$$(\alpha + (-\alpha)) \cdot \mathbf{u} = 0 \cdot \mathbf{u} = \mathbf{0}.$$

Pero por otro lado

$$(\alpha + (-\alpha)) \cdot \mathbf{u} = \alpha \cdot \mathbf{u} + (-\alpha) \cdot \mathbf{u} = 0,$$

de donde tenemos que el inverso de $\alpha \cdot \mathbf{u}$ verifica que

$$-(\alpha \cdot \mathbf{u}) = (-\alpha) \cdot \mathbf{u}.$$

Por otro lado $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$. Multiplicando por α ambos miembros y procediendo como en el caso anterior tenemos que

$$-(\alpha \cdot \mathbf{u}) = \alpha \cdot (-\mathbf{u}).$$

(e) Si $\alpha \cdot \mathbf{u} = \alpha \cdot \mathbf{v}$ y $\alpha \neq 0$, entonces $\alpha \cdot (\mathbf{u} - \mathbf{v}) = \mathbf{0}$ y por el apartado (c) se tiene que $\mathbf{u} - \mathbf{v} = \mathbf{0}$, de donde $\mathbf{u} = \mathbf{v}$.

(f) Si $\alpha \cdot \mathbf{u} = \beta \cdot \mathbf{u}$ y $\mathbf{u} \neq \mathbf{0}$, entonces $(\alpha - \beta) \cdot \mathbf{u} = \mathbf{0}$ y por el apartado (c) se tiene que $\alpha - \beta = 0$, de donde $\alpha = \beta$.

(g) Consideramos $(-\alpha) \cdot (-\mathbf{u}) = -(\alpha \cdot (-\mathbf{u})) = -(-\alpha \cdot \mathbf{u}) = \alpha \cdot \mathbf{u}$. ■

2.2 Subespacios vectoriales

Definition 2 *Sea $\mathcal{V}$ un espacio vectorial sobre $\mathbb{K}$. Un subconjunto $\mathcal{W} \subseteq \mathcal{V}$ se dice un* subespacio vectorial *de $\mathcal{V}$ si para todo $\alpha, \beta \in \mathbb{K}$ y para todo $\mathbf{u}, \mathbf{v} \in \mathcal{W}$ se verifica que $\alpha \cdot \mathbf{u} + \beta \cdot \mathbf{v} \in \mathcal{W}$.*

Una primera consecuencia de la definición es que para todo $\alpha \in \mathbb{K}$ y todo $\mathbf{u}, \mathbf{v} \in \mathcal{W}$ se verifican que $\mathbf{u} + \mathbf{v} \in \mathcal{W}$ y $\alpha \cdot \mathbf{u} \in \mathcal{W}$. Como las operaciones siguen verificando los 8 axiomas de espacio vectorial, tenemos que $\mathcal{W}$ también es un espacio vectorial sobre $\mathbb{K}$, de donde inferimos que un subespacio vectorial es un espacio vectorial más pequeño dentro de uno más grande.

Example 4 Si $\mathcal{V}$ es espacio vectorial sobre $\mathbb{K}$, entonces $\{\mathbf{0}\}$ es un subespacio vectorial de $\mathcal{V}$. Mas aún, si $\mathcal{W}$ es un subespacio vectorial tenemos que para todo $\mathbf{v} \in \mathcal{W}$ se verifica que $\mathbf{0} = \mathbf{v} + (-\mathbf{v}) \in \mathcal{W}$.

Example 5 Supongamos que estamos en $\mathbb{R}^3$ y sea $\mathcal{W} = \{(x, y, z) \in \mathbb{R}^3 : y = z = 0\}$. Dados (x_1, y_1, z_1), $(x_2, y_2, z_2) \in \mathcal{W}$ y α, $\beta \in \mathbb{R}$, se verifica que

$$\alpha \cdot (x_1, y_1, z_1) + \beta \cdot (x_2, y_2, z_2) = (\alpha x_1 + \beta x_2, \alpha y_1 + \beta y_2, \alpha z_1 + \beta z_2),$$

de donde $\alpha y_1 + \beta y_2 = 0$ y $\alpha z_1 + \beta z_2 = 0$ por lo que $\alpha \cdot (x_1, y_1, z_1) + \beta \cdot (x_2, y_2, z_2) \in \mathcal{W}$ y así es un subespacio vectorial de $\mathbb{R}^3$.

Definition 3 *Dados $\mathbf{u}_1, ..., \mathbf{u}_n \in \mathcal{V}$ se define una* combinación lineal *de dichos vectores como una expresión de la forma*

$$\alpha_1 \cdot \mathbf{u}_1 + ... + \alpha_n \cdot \mathbf{u}_n$$

donde $\alpha_1, ..., \alpha_n \in \mathbb{K}$.

La noción de combinación lineal es central en la teoría de espacios vectoriales y aparecerá copiosamente durante el transcurso del tema. De hecho, para definir los subespacios vectoriales hemos utilizado una combinación lineal de dos elementos.

Definition 4 *Dado un subconjunto $\mathcal{S} \subset \mathcal{V}$ se define el* subespacio generado *por $\mathcal{S}$ como el conjunto de todas las combinaciones lineales finitas de elementos de $\mathcal{V}$. Lo denotamos por $\mathcal{L}(\mathcal{S})$, y podemos escribir que*

$$\mathcal{L}(\mathcal{S}) = \{\alpha_1 \cdot \mathbf{u}_1 + ...\alpha_n \cdot \mathbf{u}_n : \mathbf{u}_i \in \mathcal{S} \text{ y } \alpha_i \in \mathbb{K},\ i = 1, 2, ..., n\}.$$

Tenemos entonces la siguiente propiedad.

Proposition 8 *Sea $\mathcal{S} \subset \mathcal{V}$. Entonces el subespacio generado por $\mathcal{S}$, $\mathcal{L}(\mathcal{S})$ es un subespacio vectorial de $\mathcal{V}$.*

Demostración. Sean $\alpha, \beta \in \mathbb{K}$ y $\mathbf{u}, \mathbf{v} \in \mathcal{L}(\mathcal{S})$, y veamos que $\alpha \cdot \mathbf{u} + \beta \cdot \mathbf{v} \in \mathcal{L}(\mathcal{S})$. Para ello, sabemos que existen $\alpha_1, ..., \alpha_n, \beta_1, ..., \beta_m \in \mathbb{K}$ y $\mathbf{u}_1, ..., \mathbf{u}_n, \mathbf{v}_1, ..., \mathbf{v}_m \in \mathcal{S}$ tales que $\mathbf{u} = \alpha_1 \cdot \mathbf{u}_1 + ...\alpha_n \cdot \mathbf{u}_n$ y $\mathbf{v} = \beta_1 \cdot \mathbf{v}_1 + ... + \beta_m \cdot \mathbf{v}_m$. Entonces

$$\begin{aligned} \alpha \cdot \mathbf{u} + \beta \cdot \mathbf{v} &= \alpha \cdot (\alpha_1 \cdot \mathbf{u}_1 + ...\alpha_n \cdot \mathbf{u}_n) + \beta \cdot (\beta_1 \cdot \mathbf{v}_1 + ... + \beta_m \cdot \mathbf{v}_m) \\ &= (\alpha\alpha_1) \cdot \mathbf{u}_1 + ... + (\alpha\alpha_n) \cdot \mathbf{u}_n + (\beta\beta_1) \cdot \mathbf{v}_1 + ... + (\beta\beta_m) \cdot \mathbf{v}_m, \end{aligned}$$

de donde vemos que $\alpha \cdot \mathbf{u} + \beta \cdot \mathbf{v}$ es una combinación lineal finita de elementos de $\mathcal{S}$ y así $\alpha \cdot \mathbf{u} + \beta \cdot \mathbf{v}$ pertenece a $\mathcal{L}(\mathcal{S})$. ■

Example 6 Sea $\mathcal{S} = \{(1, 1, 1, 1), (0, 1, 1, 1)\} \subset \mathbb{R}^4$ y calculemos $\mathcal{L}(\mathcal{S})$. Para ello démonos cuenta que un vector $(x, y, z, t) \in \mathcal{L}(\mathcal{S})$ si ybsólo si existen $\alpha, \beta \in \mathbb{R}$ de manera que

$$(x, y, z, t) = \alpha \cdot (1, 1, 1, 1) + \beta \cdot (0, 1, 1, 1) = (\alpha, \alpha + \beta, \alpha + \beta, \alpha + \beta),$$

de donde obtenemos el sistema

$$\begin{cases} x = \alpha, \\ y = \alpha + \beta, \\ z = \alpha + \beta, \\ t = \alpha + \beta, \end{cases}$$

que al tener solución será compatible. Calculamos entonces los rangos de sus matrices asociadas

$$\left(\begin{array}{cc|c} 1 & 0 & x \\ 1 & 1 & y \\ 1 & 1 & z \\ 1 & 1 & t \end{array}\right) \to_{\substack{F_3 - F_2 \\ F_4 - F_2}} \left(\begin{array}{cc|c} 1 & 0 & x \\ 1 & 1 & y \\ 0 & 0 & z - y \\ 0 & 0 & t - y \end{array}\right),$$

y obtenemos que para que ambos sean iguales a dos debe verificarse que $z = y = t$. Entonces

$$\mathcal{L}(\mathcal{S}) = \{(x, y, z, t) \in \mathbb{R}^4 : y = z = t\}.$$

Veamos a continuación cómo se comporta la noción de subespacio vectorial con las operaciones entre conjuntos. Antes, definimos una nueva operación *suma* del siguiente modo. Sean $\mathcal{W}_1$ y $\mathcal{W}_2$ subespacios vectoriales de $\mathcal{V}$ y definimos la suma de ambos como

$$\mathcal{W}_1 + \mathcal{W}_2 = \{\mathbf{u} + \mathbf{v} : \mathbf{u} \in \mathcal{W}_1, \ \mathbf{v} \in \mathcal{W}_2\}.$$

Tenemos entonces el siguiente resultado.

Proposition 9 *Sean $\mathcal{W}_1$ y $\mathcal{W}_2$ subespacios vectoriales de $\mathcal{V}$. Entonces*

(a) $\mathcal{W}_1 \cap \mathcal{W}_2$ es un subespacio vectorial de $\mathcal{V}$.

(b) $\mathcal{W}_1 \cup \mathcal{W}_2$ no es un subespacio vectorial de $\mathcal{V}$ en general.

(c) $\mathcal{W}_1 + \mathcal{W}_2$ es un subespacio vectorial de $\mathcal{V}$.

Demostración. (a) Sean $\alpha, \beta \in \mathbb{K}$ y $\mathbf{u}, \mathbf{v} \in \mathcal{W}_1 \cap \mathcal{W}_2$ y veamos que $\alpha \cdot \mathbf{u} + \beta \cdot \mathbf{v} \in \mathcal{W}_1 \cap \mathcal{W}_2$. Como $\mathcal{W}_1$ es un subespacio vectorial se tiene que $\alpha \cdot \mathbf{u} + \beta \cdot \mathbf{v} \in \mathcal{W}_1$. Análogamente tenemos que $\alpha \cdot \mathbf{u} + \beta \cdot \mathbf{v} \in \mathcal{W}_2$. Así, $\alpha \cdot \mathbf{u} + \beta \cdot \mathbf{v} \in \mathcal{W}_1 \cap \mathcal{W}_2$.

(b) Consideremos el espacio vectorial $\mathbb{R}^2$ y los subespacios vectoriales $\mathcal{W}_1 = \{(x, y) \in \mathbb{R}^2 : x = 0\}$ y $\mathcal{W}_2 = \{(x, y) \in \mathbb{R}^2 : y = 0\}$. La unión de ambos subespacios es $\mathcal{W}_1 \cup \mathcal{W}_2 = \{(x, y) \in \mathbb{R}^2 : x = 0 \text{ ó } y = 0\}$. Entonces los vectores $(1, 0), (0, 1) \in \mathcal{W}_1 \cup \mathcal{W}_2$ y sin embargo $(1, 0) + (0, 1) = (1, 1) \notin \mathcal{W}_1 \cup \mathcal{W}_2$, por lo que no es un subespacio vectorial.

(c) Sean $\alpha, \beta \in \mathbb{K}$ y $\mathbf{u}, \mathbf{v} \in \mathcal{W}_1 + \mathcal{W}_2$ y veamos que $\alpha \cdot \mathbf{u} + \beta \cdot \mathbf{v} \in \mathcal{W}_1 + \mathcal{W}_2$. Dado que $\mathbf{u} \in \mathcal{W}_1 + \mathcal{W}_2$, existen vectores $\mathbf{u}_i \in \mathcal{W}_i$, $i = 1, 2$, tales que $\mathbf{u} = \mathbf{u}_1 + \mathbf{u}_2$. Similarmente y por la misma razón, existen vectores $\mathbf{v}_i \in \mathcal{W}_i$, $i = 1, 2$, tales que $\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2$. Como $\mathcal{W}_1$ es un subespacio vectorial se tiene que $\alpha \cdot \mathbf{u}_1 + \beta \cdot \mathbf{v}_1 \in \mathcal{W}_1$. Análogamente $\alpha \cdot \mathbf{u}_2 + \beta \cdot \mathbf{v}_2 \in \mathcal{W}_2$. Entonces

$$\alpha \cdot \mathbf{u} + \beta \cdot \mathbf{v} = \alpha \cdot (\mathbf{u}_1 + \mathbf{u}_2) + \beta \cdot (\mathbf{v}_1 + \mathbf{v}_2) = (\alpha \cdot \mathbf{u}_1 + \beta \cdot \mathbf{v}_1) + (\alpha \cdot \mathbf{u}_2 + \beta \cdot \mathbf{v}_2) \in \mathcal{W}_1 + \mathcal{W}_2,$$

por lo que $\mathcal{W}_1 + \mathcal{W}_2$ es un subespacio vectorial. ■

Se dirá que la suma de dos subespacios es directa si su intersección es el vector $\mathbf{0}$, esto es, $\mathcal{W}_1 \cap \mathcal{W}_2 = \{\mathbf{0}\}$. Escribiremos $\mathcal{W}_1 \oplus \mathcal{W}_2$ para indicar que los subespacios $\mathcal{W}_1$ y $\mathcal{W}_2$ estan en *suma directa*. La intersección y la suma de subespacios vectoriales son operaciones que permiten construir nuevos subespacios vectoriales más pequeños ($\mathcal{W}_1 \cap \mathcal{W}_2 \subseteq \mathcal{W}_i$, $i = 1, 2$) o más grandes ($\mathcal{W}_i \subseteq \mathcal{W}_1 + \mathcal{W}_2$, $i = 1, 2$). Estas operaciones serán de gran utilidad a la hora de estudiar la diagonalización de matrices cuadradas, como veremos posteriormente.

Example 7 Dado $\mathbb{R}^3$ y los subespacios $\mathcal{W}_1 = \{(x, y, z) \in \mathbb{R}^3 : x + y = 0, \ x + z = 0\}$ y $\mathcal{W}_2 = \{(x, y, z) \in \mathbb{R}^3 : x = 0, \ y = 0\}$, calculemos su suma e intersección. Para ello, démonos cuenta que $(x, y, z) \in \mathcal{W}_1 \cap \mathcal{W}_2$ si satisface a la vez las ecuaciones que definen ambos subespacios, esto es

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

de donde obtenemos resolviendo el sistema que $x = y = z = 0$, o lo que es lo mismo $\mathcal{W}_1 \cap \mathcal{W}_2 = \{(0,0,0)\}$. Por otro lado, las ecuaciones paramétricas de $\mathcal{W}_1$ y $\mathcal{W}_2$ son

$$\begin{cases} x = -\lambda, \\ y = \lambda, \quad \lambda \in \mathbb{R}, \\ z = \lambda, \end{cases}$$

y

$$\begin{cases} x = 0, \\ y = 0, \quad \lambda \in \mathbb{R}, \\ z = \lambda, \end{cases}$$

respectivamente. Entonces es fácil ver que $\mathcal{W}_1 = \mathcal{L}(-1,1,1)$ y $\mathcal{W}_2 = \mathcal{L}(0,0,1)$. Así $(x,y,z) \in \mathcal{W}_1 + \mathcal{W}_2$ si y sólo si existen $(x_i, y_i, z_i) \in \mathcal{W}_i$, $i = 1, 2$, donde

$$(x,y,z) = (x_1,y_1,z_1) + (x_2,y_2,z_2) = \alpha \cdot (-1,1,1) + \beta \cdot (0,0,1) = (-\alpha, \alpha, \alpha + \beta),$$

de donde tenemos el sistema

$$\begin{cases} x = -\alpha, \\ y = \alpha, \\ z = \alpha + \beta, \end{cases}$$

que es compatible. Al calcular los rangos de las matrices asociadas

$$\left(\begin{array}{cc|c} -1 & 0 & x \\ 1 & 0 & y \\ 1 & 1 & z \end{array} \right) \to_{\substack{F_2+F_1 \\ F_3+F_1}} \left(\begin{array}{cc|c} -1 & 0 & x \\ 0 & 0 & y+x \\ 0 & 1 & z+y \end{array} \right)$$

tenemos que ambos son iguales a dos si $y + x = 0$, por lo que

$$\mathcal{W}_1 + \mathcal{W}_2 = \{(x,y,z) \in \mathbb{R}^3 : x + y = 0\}.$$

2.3 Bases y dimensión de espacios vectoriales

Definition 5 *Dados* $\mathbf{u}_1, ..., \mathbf{u}_n \in \mathcal{V}$ *se dice que son* linealmente independientes *(abreviado LI) si dada la combinación lineal*

$$\alpha_1 \cdot \mathbf{u}_1 + ... + \alpha_n \cdot \mathbf{u}_n = \mathbf{0},$$

se verifica que $\alpha_1 = ... = \alpha_n = 0$. *En caso contrario se dirán* linealmente dependientes *(LD).*

Example 8 El vector $\mathbf{0}$ siempre es LD ya que para todo $\alpha \in \mathbb{K}$ se tiene que $\alpha \cdot \mathbf{0} = \mathbf{0}$.

Example 9 Dado $\mathbb{R}^2$ se verifica que $(1,0)$ y $(1,1)$ son LI. Para verificarlo construimos la combinación lineal

$$\alpha \cdot (1,0) + \beta \cdot (1,1) = (0,0),$$

de donde

$$(\alpha + \beta, \beta) = (0,0),$$

y obtenemos el sistema

$$\begin{cases} \alpha + \beta = 0, \\ \beta = 0, \end{cases}$$

que al resolverlo $\beta = \alpha = 0$.

Example 10 Dado $\mathcal{P}_2[x]$ se verifica que 1, x y x^2 son LI. Para comprobar ésto construimos la combinación lineal

$$\alpha + \beta x + \gamma x^2 = 0.$$

Dando los valores $x = 0$, $x = 1$ y $x = -1$ obtenemos el sistema

$$\begin{cases} \alpha = 0, \\ \alpha + \beta + \gamma = 0, \\ \alpha - \beta + \gamma = 0, \end{cases}$$

y al resolverlo tenemos las soluciones $\alpha = \beta = \gamma = 0$.

Tenemos entonces la siguiente propiedad.

Proposition 10 *Si $\mathbf{u}_1, ..., \mathbf{u}_n \in \mathcal{V}$ son LD entonces uno de ellos es el vector* $\mathbf{0}$ *o es combinación lineal de los restantes.*

Demostración. Si uno de ellos es $\mathbf{0}$, la tesis del resultado está probada. Supongamos entonces que son todos los vectores no nulos y que existe una combinación lineal $\alpha_1 \cdot \mathbf{u}_1 + ... + \alpha_n \cdot \mathbf{u}_n = \mathbf{0}$ tal que por ejemplo $\alpha_1 \neq 0$. Entonces podemos despejar $\mathbf{u}_1$ como

$$\mathbf{u}_1 = -\frac{\alpha_2}{\alpha_1} \cdot \mathbf{u}_2 - ... - \frac{\alpha_n}{\alpha_1} \cdot \mathbf{u}_n,$$

por lo que el vector $\mathbf{u}_1$ será combinación lineal de $\mathbf{u}_2, ..., \mathbf{u}_n$. ■

Definition 6 *Dados $\mathbf{u}_1, ..., \mathbf{u}_n \in \mathcal{V}$ se dice que* generan $\mathcal{V}$ *si $\mathcal{L}(\{\mathbf{u}_1, ..., \mathbf{u}_n\}) = \mathcal{V}$. En tal caso $\mathcal{V}$ se dirá un* espacio vectorial finitamente generado *y $\{\mathbf{u}_1, ..., \mathbf{u}_n\}$ un* conjunto generador *de $\mathcal{V}$.*

Example 11 El conjunto de vectores $\{(1, 1), (0, 1)\}$ generan el espacio vectorial $\mathbb{R}^2$. Para ello consideremos un vector arbitrario $(x, y) \in \mathbb{R}^2$ y veamos que existen $\alpha, \beta \in \mathbb{R}$ tales que $(x, y) = \alpha \cdot (1, 1) + \beta \cdot (0, 1)$. Desarrollamos

$$(x, y) = (\alpha, \alpha + \beta),$$

que da lugar al sistema

$$\begin{cases} x = \alpha, \\ y = \alpha + \beta, \end{cases}$$

de donde $\alpha = x$ y $\beta = y - x$ y así

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

y $\mathbb{R}^2 = \mathcal{L}(\{(1, 1), (0, 1)\})$.

Example 12 No todos los espacios vectoriales son finitamente generados. Por ejemplo, el conjunto de los polinomios con coeficientes reales

$$\mathcal{P}[x] = \{a_0 + a_1 x + ... + a_n x^n : n \in \mathbb{N}, a_i \in \mathbb{R}, 0 \leq i \leq n\}$$

es un subespacio vectorial sobre $\mathbb{R}$. Supongamos que existe polinomios $p_1(x), ..., p_k(x) \in \mathcal{P}[x]$ que generan dicho espacio vectorial. Sea m el grado del polinomio de más grado entre los polinomios $p_1(x), ..., p_k(x)$. Entonces nos es posible encontrar números reales $\alpha_1, ..., \alpha_k$ tales que

$$x^{m+1} = \alpha_1 p_1(x) + ... + \alpha_k p_k(x),$$

ya que en la igualdad anterior el polinomio de la izquierda de la igualdad es $m + 1$ y el de la derecha es a lo sumo m.

Veamos algunas propiedades de los conjuntos generadores.

Proposition 11 *Supongamos que $\mathcal{S} = \{\mathbf{u}_1, ..., \mathbf{u}_n\}$ generan $\mathcal{V}$. Entonces existe $\mathcal{B} \subseteq \mathcal{S}$ que también genera $\mathcal{V}$ y tal que sus elementos son LI.*

Demostración. Si los vectores de $\mathcal{S}$ son LI ya hemos terminado. Supongamos entonces que son LD y apliquemos la Proposición 10 y supongamos por ejemplo que existen $\alpha_2, ..., \alpha_n \in \mathbb{K}$ tales que

$$\mathbf{u}_1 = \alpha_2 \cdot \mathbf{u}_2 + ... + \alpha_n \cdot \mathbf{u}_n. \tag{2.1}$$

Entonces comprobemos que $\mathcal{L}(\{\mathbf{u}_1, ..., \mathbf{u}_n\} \setminus \{\mathbf{u}_2\}) = \mathcal{L}(\{\mathbf{u}_2, ..., \mathbf{u}_n\}) = \mathcal{V}$. Para ello, sea $\mathbf{v} \in \mathcal{V}$ arbitrario. Como $\mathcal{L}(\{\mathbf{u}_1, ..., \mathbf{u}_n\}) = \mathcal{V}$ existen $\beta_1, ..., \beta_n \in \mathbb{K}$ tales que $\mathbf{v} = \beta_1 \cdot \mathbf{u}_1 + ... + \beta_n \cdot \mathbf{u}_n$ y así sustituyendo $\mathbf{u}_1$ por la expresión (2.1) y simplificando tenemos

$$\begin{aligned} \mathbf{v} &= \beta_1(\alpha_2 \cdot \mathbf{u}_2 + ... + \alpha_n \cdot \mathbf{u}_n) + \beta_2 \cdot \mathbf{u}_{21} + ... + \beta_n \cdot \mathbf{u}_n \\ &= (\beta_1\alpha_2 + \beta_2) \cdot \mathbf{u}_2 + ... + (\beta_1\alpha_n + \beta_n) \cdot \mathbf{u}_n, \end{aligned}$$

por lo que $\mathbf{v}$ es combinación lineal de $\mathbf{u}_2, ..., \mathbf{u}_n$. De esta manera tenemos un método para eliminar vectores LD dentro de un conjunto generador. Como tenemos una cantidad finita de vectores, debe haber un momento en el cual los vectores que quedan al eliminar uno LD sean LI. ■

La siguiente propiedad establece relaciones entre el número de elementos en un conjunto generador de $\mathcal{V}$ y el número de vectores LI.

Proposition 12 *Supongamos que $\mathcal{V}$ está generado por n vectores. Entonces ningún conjunto LI de $\mathcal{V}$ tiene mas de n elementos.*

Demostración. Haremos la demostración por inducción en n.

Si $n = 1$ entonces $\mathcal{V} = \mathcal{L}(\{\mathbf{u}\})$ para algún $\mathbf{u} \in \mathcal{V}$. Entonces todo vector de $\mathcal{V}$ será proporcional a $\mathbf{u}$ y esto imposibilita que existan dos vectores linealmente independientes en $\mathcal{V}$.

Supongamos ahora que el resultado es cierto para espacios vectoriales generados por a lo sumo n elementos y probemos que es cierto para espacios generados por $n+1$ elementos. Para ello sean $\mathcal{V} = \mathcal{L}(\{\mathbf{v}_1, ..., \mathbf{v}_{n+1}\})$ y $\mathcal{S} = \{\mathbf{u}_1, ..., \mathbf{u}_m\}$ un conjunto de vectores linealmente independientes de $\mathcal{V}$, y veamos que $m \leq n+1$. Distinguimos dos casos. En primer lugar suponemos que $\mathcal{S} \subset \mathcal{L}(\{\mathbf{v}_1, ..., \mathbf{v}_n\})$ y entonces por la hipótesis inductiva $m \leq n < n+1$. Así, supongamos ahora que por ejemplo $\mathbf{u}_1 \notin \mathcal{L}(\{\mathbf{v}_1, ..., \mathbf{v}_n\})$. Entonces $\mathbf{u}_1 = \alpha_1 \cdot \mathbf{v}_1 + ... + \alpha_{n+1} \cdot \mathbf{v}_{n+1}$ y $\alpha_{n+1} \neq 0$, de donde $\mathbf{v}_{n+1} = \alpha_{n+1}^{-1}\alpha_1 \cdot \mathbf{v}_1 + ... + \alpha_{n+1}^{-1}\alpha_n \cdot \mathbf{v}_n + \alpha_{n+1}^{-1} \cdot \mathbf{u}_1$. Veamos entonces que $\mathcal{V} = \mathcal{L}(\{\mathbf{v}_1, ..., \mathbf{v}_n, \mathbf{u}_1\})$. Para ello sea $\mathbf{v} \in \mathcal{V}$ para el que existirán $\beta_i \in \mathbb{K}$, $i = 1, ..., n+1$, tales que

$$\begin{aligned} \mathbf{v} &= \beta_1 \cdot \mathbf{v}_1 + ... + \beta_n \cdot \mathbf{v}_n + \beta_{n+1} \cdot \mathbf{v}_{n+1} \\ &= \beta_1 \cdot \mathbf{v}_1 + ... + \beta_n \cdot \mathbf{v}_n + \beta_{n+1} \cdot (\alpha_{n+1}^{-1}\alpha_1 \cdot \mathbf{v}_1 + ... + \alpha_{n+1}^{-1}\alpha_n \cdot \mathbf{v}_n + \alpha_{n+1}^{-1} \cdot \mathbf{u}_1) \\ &= (\beta_1 + \beta_{n+1}\alpha_{n+1}^{-1}\alpha_1) \cdot \mathbf{v}_1 + ...(\beta_n + \beta_{n+1}\alpha_{n+1}^{-1}\alpha_n) \cdot \mathbf{v}_n + \beta_{n+1}\alpha_{n+1}^{-1} \cdot \mathbf{u}_1. \end{aligned}$$

Así $\mathcal{V} = \mathcal{L}(\{\mathbf{v}_1, ..., \mathbf{v}_n, \mathbf{u}_1\})$. Entonces cada $\mathbf{u}_i$, $i = 2, 3, ..., m$ puede ponerse en combinación lineal de $\mathbf{v}_1, ..., \mathbf{v}_n, \mathbf{u}_1$, esto es $\mathbf{u}_i = \sum_{j=1}^n \alpha_{ij} \cdot \mathbf{v}_j + \lambda_i \cdot \mathbf{u}_1$, $i = 2, ..., m$. Sean ahora $\mathbf{y}_i = \mathbf{u}_i - \lambda_i \cdot \mathbf{u}_1 \in \mathcal{L}(\{\mathbf{v}_1, ..., \mathbf{v}_n\})$ $i = 2, 3, ..., m$. Veamos que son LI. Para ello sea

$$\gamma_2 \cdot \mathbf{y}_2 + ... + \gamma_m \cdot \mathbf{y}_m = \mathbf{0}.$$

Pero entonces

$$\begin{aligned} \mathbf{0} &= \gamma_2 \cdot (\mathbf{u}_2 - \lambda_2 \cdot \mathbf{u}_1) + ... + \gamma_m \cdot (\mathbf{u}_m - \lambda_m \cdot \mathbf{u}_1) \\ &= \gamma_2 \cdot \mathbf{u}_2 + ... + \gamma_m \cdot \mathbf{u}_m - (\gamma_2 \lambda_2 + ... + \gamma_m \lambda_m) \cdot \mathbf{u}_1 \end{aligned}$$

y como $\mathbf{u}_1, ..., \mathbf{u}_m$ son LI se tiene que $\gamma_i = 0$, $i = 2, 3, ..., m$. De nuevo por la hipótesis inductiva $m - 1 \leq n$, de donde $m \leq n + 1$ como queríamos probar. ■

Definition 7 *Una* base *de $\mathcal{V}$ es un conjunto generador de $\mathcal{V}$ y LI.*

Example 13 El conjunto $\mathcal{B} = \{1, x, x^2\}$ es una base del conjunto $\mathcal{P}_2[x]$ de polinomios reales de grado menor o igual que 2. Ya vimos en el ejemplo 10 que era un conjunto LI. Por otra parte, es claro que es un conjunto generador dado que todo $p(x) \in \mathcal{P}_2[x]$ es de la forma $p(x) = a + bx + cx^2$, $a, b, c \in \mathbb{R}$.

A continuación probamos un resultado fundamental sobre las bases de espacios vectoriales finitamente generados que permite definir la noción de dimensión de estos espacios.

Theorem 13 *Si $\mathcal{V}$ es finitamente generado, entonces todas sus bases tienen el mismo número de elementos.*

Demostración. Sean $\mathcal{B}_1 = \{\mathbf{u}_1, ..., \mathbf{u}_n\}$ y $\mathcal{B}_2 = \{\mathbf{v}_1, ..., \mathbf{v}_m\}$ bases de $\mathcal{V}$ y veamos que $n = m$. Consideremos $\mathcal{B}_1$ como conjunto generador y $\mathcal{B}_2$ como conjunto linealmente independiente. Aplicamos la Proposición 12 para obtener que $n \geq m$. Si ahora consideramos $\mathcal{B}_1$ como linealmente independiente y $\mathcal{B}_2$ como conjunto generador y volvemos a aplicar la Proposición 12 y obtenemos $m \geq n$. Entonces $m = n$. ■

Definition 8 *Dado $\mathcal{V}$ un espacio vectorial finitamente generado. Se llama* dimensión *de $\mathcal{V}$,* $\dim \mathcal{V}$ *(o $\dim_{\mathbb{K}} \mathcal{V}$ si queremos enfatizar el cuerpo $\mathbb{K}$), al número de elementos en una base.*

Otra de las ventajas de trabajar con bases es la siguiente propiedad.

Proposition 14 *Sea $\mathcal{B} = \{\mathbf{u}_1, ..., \mathbf{u}_n\}$ una base de $\mathcal{V}$. Entonces todo vector $\mathbf{u} \in \mathcal{V}$ admite una única expresión como combinación lineal de los elementos de $\mathcal{B}$.*

Demostración. Supongamos dos combinaciones lineales

$$\mathbf{u} = \alpha_1 \cdot \mathbf{u}_1 + ... + \alpha_n \cdot \mathbf{u}_n = \beta_1 \cdot \mathbf{u}_1 + ... + \beta_n \cdot \mathbf{u}_n,$$

$\alpha_i, \beta_i \in \mathbb{K}$, $1 \leq i \leq n$. Entonces

$$(\alpha_1 - \beta_1) \cdot \mathbf{u}_1 + ... + (\alpha_n - \beta_n) \cdot \mathbf{u}_n = \mathbf{0}.$$

Como los vectores $\mathbf{u}_1, ..., \mathbf{u}_n$ son LI, se verifica que $\alpha_i - \beta_i = 0$, o lo que es lo mismo $\alpha_i = \beta_i$, $1 \leq i \leq n$. ■

Sean $\mathcal{B} = \{\mathbf{u}_1, ..., \mathbf{u}_n\}$ una base de $\mathcal{V}$ y $\mathbf{u} \in \mathcal{V}$. Entonces $\mathbf{u} = \alpha_1 \cdot \mathbf{u}_1 + ... + \alpha_n \cdot \mathbf{u}_n$ con $\alpha_1, ..., \alpha_n \in \mathbb{K}$ únicos. Si los escribimos como un vector de $\mathbb{K}^n$, diremos entonces que $(\alpha_1, ..., \alpha_n)_{\mathcal{B}}$ son las *coordenadas* de $\mathbf{u}$ en la base $\mathcal{B}$. Démonos cuenta que al cambiar la base cambian a su vez las coordenadas del vector. Por ejemplo, sean $\mathbb{R}^2$, las bases $\mathcal{B}_1 = \{(1,0),(0,1)\}$ y $\mathcal{B}_2 = \{(1,1),(1,-1)\}$ y el vector $(2,2)$. Este vector tiene coordenadas $(2,2)_{\mathcal{B}_1}$ en la base $\mathcal{B}_1$, mientras que son $(2,0)_{\mathcal{B}_2}$ en la base $\mathcal{B}_2$.

La base $\mathcal{B}_1$ es lo que conoceremos como base canónica, esto es, aquella que verifica que las coordenadas de un vector son las "naturales". Por ejemplo, todo vector $(x_1,...,x_n) \in \mathbb{K}^n$ tiene coordenadas $(x_1,...,x_n)_{\mathcal{C}}$ respecto de la base canónica $\mathcal{C} = \{(1,0,...,0),(0,1,0,...,0),...,(0,0,...,0,1)\}$. En el conjunto de los polinomios $\mathcal{P}_n[x]$ la base canónica es $\mathcal{C} = \{1,x,...,x^n\}$, de manera que cualquier polinomio $a_0 + a_1x + ... + a_nx^n \in \mathcal{P}_n[x]$ tiene por coordenadas en dicha base $(a_0,a_1,...,a_n)_{\mathcal{C}}$.

Proposition 15 *Sea $\mathcal{V}$ un espacio vectorial finitamente generado y $\{\mathbf{u}_1,...,\mathbf{u}_m\}$ un conjunto LI. Entonces existe una base $\mathcal{B}$ de $\mathcal{V}$ de manera que $\{\mathbf{u}_1,...,\mathbf{u}_m\} \subseteq \mathcal{B}$.*

Demostración. Si $\mathcal{L}(\{\mathbf{u}_1,...,\mathbf{u}_m\}) = \mathcal{V}$, ya hemos terminado. En caso contrario sea $\mathbf{v} \in \mathcal{V} \setminus \mathcal{L}(\{\mathbf{u}_1,...,\mathbf{u}_m\})$. Entonces toda combinación lineal

$$\alpha \cdot \mathbf{v} + \alpha_1 \cdot \mathbf{u}_1 + ... + \alpha_m \cdot \mathbf{u}_m = \mathbf{0},$$

debe verificar que $\alpha = 0$. Por la independencia lineal de $\mathbf{u}_1,...,\mathbf{u}_m$, se tiene además que $\alpha_i = 0$, $1 \leq i \leq m$. Así $\{\mathbf{v},\mathbf{u}_1,...,\mathbf{u}_m\}$ es un conjunto LI. De esta manera incrementamos en una unidad el conjunto LI. Como $\mathcal{V}$ es finitamente generado, en una cantidad finita de pasos logramos obtener un conjunto LI que genere $\mathcal{V}$. ■

A modo de resumen tenemos el siguiente resultado.

Theorem 16 *Sea $\mathcal{V}$ un espacio vectorial finitamente generado de dimensión n. Entonces*

(a) Todas las bases de $\mathcal{V}$ tienen n elementos.

(b) Todo conjunto LI de n elementos es una base de $\mathcal{V}$.

(c) Todo conjunto generador de $\mathcal{V}$ de n elementos es una base.

Demostración. El apartado (a) está probado en el Teorema 13. Para demostrar (b), sea $\{\mathbf{u}_1,...,\mathbf{u}_n\} \subset \mathcal{V}$ un conjunto LI. Si no fuera generador de $\mathcal{V}$, aplicando la Proposición 15, podríamos extenderlo a una base de $\mathcal{V}$ que necesariamente tendría mas de n elementos, lo cual contradice el Teorema 13. De igual modo y utilizando la Proposición 11 se prueba el apartado (c). ■

Para finalizar el tema, nótese que los subespacios vectoriales son a su vez espacios vectoriales que tendrán su dimensión. Sobre este particular tenemos la siguiente fórmula de las dimensiones de la suma e intersección de subespacios vectoriales.

Proposition 17 *Sean $\mathcal{W}_1$ y $\mathcal{W}_2$ subespacios vectoriales de $\mathcal{V}$ finitamente generados. Entonces*

$$\dim(\mathcal{W}_1 + \mathcal{W}_2) = \dim \mathcal{W}_1 + \dim \mathcal{W}_2 - \dim(\mathcal{W}_1 \cap \mathcal{W}_2).$$

En particular, si la suma es directa

$$\dim(\mathcal{W}_1 \oplus \mathcal{W}_2) = \dim \mathcal{W}_1 + \dim \mathcal{W}_2.$$

Demostración. Supongamos en primer lugar que $\mathcal{W}_1 \oplus \mathcal{W}_2$ y sean $\mathcal{B}_1 = \{\mathbf{u}_1, ..., \mathbf{u}_n\}$ y $\mathcal{B}_2 = \{\mathbf{v}_1, ..., \mathbf{v}_m\}$ bases de $\mathcal{W}_1$ y $\mathcal{W}_2$, respectivamente y comprobemos que $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ es una base de $\mathcal{W}_1 \oplus \mathcal{W}_2$. Veamos que $\mathcal{B}$ es un conjunto LI. Sea una combinación lineal

$$\alpha_1 \cdot \mathbf{u}_1 + ... + \alpha_n \cdot \mathbf{u}_n + \beta_1 \cdot \mathbf{v}_1 + ... + \beta_m \cdot \mathbf{v}_m = \mathbf{0}.$$

Entonces

$$\mathbf{v} = \alpha_1 \cdot \mathbf{u}_1 + ... + \alpha_n \cdot \mathbf{u}_n = -\beta_1 \cdot \mathbf{v}_1 - ... - \beta_m \cdot \mathbf{v}_m \in \mathcal{W}_1 \cap \mathcal{W}_2,$$

con lo que $\mathbf{v} = \mathbf{0}$ y

$$\begin{aligned} \alpha_1 \cdot \mathbf{u}_1 + ... + \alpha_n \cdot \mathbf{u}_n &= \mathbf{0}, \\ \beta_1 \cdot \mathbf{v}_1 + ... + \beta_m \cdot \mathbf{v}_m &= \mathbf{0}. \end{aligned}$$

Como $\mathcal{B}_1$ y $\mathcal{B}_2$ son LI, obtenemos que $\alpha_i = 0$ y $\beta_j = 0$ para $0 \leq i \leq n$ y $1 \leq j \leq m$. Veamos ahora que $\mathcal{L}(\mathcal{B})= \mathcal{W}_1 \oplus \mathcal{W}_2$. Para ello sea $\mathbf{v} \in \mathcal{W}_1 \oplus \mathcal{W}_2$ arbitrario y sea $\mathbf{v}_i \in \mathcal{W}_i$, $i = 1, 2$, de manera que $\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2$. Existen entonces $\alpha_i \in \mathbb{K}$ y $\beta_j \in \mathbb{K}$, $0 \leq i \leq n$ y $1 \leq j \leq m$, de manera que

$$\begin{aligned} \mathbf{v}_1 &= \alpha_1 \cdot \mathbf{u}_1 + ... + \alpha_n \cdot \mathbf{u}_n, \\ \mathbf{v}_2 &= \beta_1 \cdot \mathbf{v}_1 + ... + \beta_m \cdot \mathbf{v}_m. \end{aligned}$$

Entonces

$$\mathbf{v} = \alpha_1 \cdot \mathbf{u}_1 + ... + \alpha_n \cdot \mathbf{u}_n + \beta_1 \cdot \mathbf{v}_1 + ... + \beta_m \cdot \mathbf{v}_m \in \mathcal{L}(\mathcal{B}).$$

Así

$$\dim(\mathcal{W}_1 \oplus \mathcal{W}_2) = n + m = \dim \mathcal{W}_1 + \dim \mathcal{W}_2.$$

Supongamos ahora que $\mathcal{W}_1 \cap \mathcal{W}_2 \neq \{\mathbf{0}\}$ y sea $\{\mathbf{u}_1, ..., \mathbf{u}_k\}$ una base del mismo. Por la Proposición 15 extendemos $\{\mathbf{u}_1, ..., \mathbf{u}_k\}$ a sendas bases $\{\mathbf{u}_1, ..., \mathbf{u}_k, \mathbf{u}_{k+1}, ..., \mathbf{u}_n\}$ y $\{\mathbf{u}_1, ..., \mathbf{u}_k, \mathbf{v}_{k+1}, ..., \mathbf{v}_m\}$ de $\mathcal{W}_1$ y $\mathcal{W}_2$, respectivamente. Sea $\mathcal{V}_1 = \mathcal{L}(\{\mathbf{u}_{k+1}, ..., \mathbf{u}_n\})$. Entonces $\mathcal{V}_1 \cap \mathcal{W}_2 = \{\mathbf{0}\}$ y por tanto

$$\dim(\mathcal{V}_1 \oplus \mathcal{W}_2) = \dim \mathcal{V}_1 + \dim \mathcal{W}_2.$$

Por otra parte $\mathcal{V}_1 \oplus \mathcal{W}_2 = \mathcal{W}_1 + \mathcal{W}_2$ y entonces

$$\begin{aligned} \dim(\mathcal{W}_1 + \mathcal{W}_2) &= \dim(\mathcal{V}_1 \oplus \mathcal{W}_2) \\ &= \dim \mathcal{V}_1 + \dim \mathcal{W}_2 \\ &= \dim \mathcal{W}_1 + \dim \mathcal{W}_2 - \dim(\mathcal{W}_1 \cap \mathcal{W}_2), \end{aligned}$$

y la proposición queda demostrada. ■

Example 14 El resultado anterior tiene su aplicación a ejemplos como el siguiente. Sean

$$\mathcal{W}_1 = \{(x, y, z) \in \mathbb{R}^3 : x + y + z = 0\}$$

y

$$\mathcal{W}_2 = \{(x, y, z) \in \mathbb{R}^3 : x + y = 0,\ y + z = 0\}.$$

Sus dimensiones son 2 y 1, respectivamente. Por otra parte, $(x, y, z) \in \mathcal{W}_1 \cap \mathcal{W}_2$ si se satisfacen

$$\begin{cases} x + y + z = 0, \\ x + y = 0, \\ y + z = 0. \end{cases}$$

Resolvemos este sistema

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{array}\right) \rightarrow_{F_2 - F_1} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 \end{array}\right),$$

y obtenemos que $x = y = z = 0$, por lo que $\mathcal{W}_1 \cap \mathcal{W}_2 = \{(0, 0, 0)\}$. Entonces

$$\dim(\mathcal{W}_1 \oplus \mathcal{W}_2) = \dim \mathcal{W}_1 + \dim \mathcal{W}_2 = 2 + 1 = 3,$$

y dado que $\mathcal{W}_1 \oplus \mathcal{W}_2 \subseteq \mathbb{R}^3$, se verifica que $\mathcal{W}_1 \oplus \mathcal{W}_2 = \mathbb{R}^3$.

2.4 Ejercicios

1. Sea $\mathcal{M}_{2\times 2}(\mathbb{R})$ el espacio de las matrices de orden 2×2 sobre el cuerpo $\mathbb{R}$, estudiar si los siguientes conjuntos de matrices son subespacios vectoriales.

   (a) $\mathcal{M}_1 = \{\mathbf{A} \in \mathcal{M}_{2\times 2}(\mathbb{R}) : \mathbf{A} \text{ es simétrica}\}$

   (b) $\mathcal{M}_2 = \{\mathbf{A} \in \mathcal{M}_{2\times 2}(\mathbb{R}) : \mathbf{A}^2 = \mathbf{A}\}$

   (c) $\mathcal{M}_3 = \{\mathbf{A} \in \mathcal{M}_{2\times 2}(\mathbb{R}) : |\mathbf{A}| = 0\}$

   (d) $\mathcal{M}_4 = \{\mathbf{A} = \begin{pmatrix} 0 & 0 \\ 0 & a \end{pmatrix} \in \mathcal{M}_{2\times 2}(\mathbb{R}) : a \in \mathbb{R}\}$

2. En $\mathbb{R}^2$ definimos la operación interna $+$ dada por:

   $$(x, y) + (u, v) = (x + u, y + v)\ \forall (x, y), (u, v) \in \mathbb{R}^2,$$

   y la operación externa $* : \mathbb{R} \times \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dada por:

   $$\alpha * (x, y) = (\alpha x, y),\ \forall \alpha \in \mathbb{R}\ y\ \forall (x, y) \in \mathbb{R}^2.$$

   ¿Tiene la terna $(\mathbb{R}^2, +, *)$ estructura de espacio vectorial sobre $\mathbb{R}$?

3. Comprobar si los siguientes conjuntos de $\mathbb{R}^3$ son subespacios vectoriales:

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

   (b) $\mathcal{W} = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 + 2x_2 + x_3 = 1\}$.

   (c) $\mathcal{W} = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 = x_2 = 0\}$.

   (d) $\mathcal{W} = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 + x_2 = 0\ y\ x_3 - x_2 = 0\}$.

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

4. Decir si los siguientes vectores de $\mathbb{R}^4$ son linealmente independientes:

   (a) $\{(1,0,0,1), (0,1,1,0), (1,1,1,1)\}$.
   (b) $\{(1,0,0,1), (0,1,1,1), (1,1,1,1)\}$.
   (c) $\{(1,0,0,1), (0,1,1,0), (1,1,1,1), (2,0,0,0)\}$.

5. Extraer un conjunto linealmente independiente de los conjuntos del ejercicio 4.

6. ¿Alguno de los conjuntos del ejercicio 4 son una base de $\mathbb{R}^4$?.

7. Calcular el subespacio vectorial generado por los conjuntos del ejercicio 4.

8. Dados los subespacios de $\mathbb{R}^3$, $\mathcal{W}_1 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 = x_2 = 0\}$ y

$$\mathcal{W}_2 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 + x_2 + x_3 = 0\},$$

   calcular:

   (a) Un conjunto generador linealmente independiente de $\mathcal{W}_1$ y $\mathcal{W}_2$.
   (b) Calcular $\mathcal{W}_1 + \mathcal{W}_2$ y $\mathcal{W}_1 \cap \mathcal{W}_2$.
   (c) ¿Es la suma de $\mathcal{W}_1$ y $\mathcal{W}_2$ directa?.
   (d) Calcular las dimensiones de $\mathcal{W}_1$, $\mathcal{W}_2$, $\mathcal{W}_1 \cap \mathcal{W}_2$ y $\mathcal{W}_1 + \mathcal{W}_2$.

9. Sea $(\mathcal{V}, +, \cdot)$ un espacio vectorial sobre un cuerpo $\mathbb{K}$, y sea $\mathcal{B} = \{\mathbf{v}_1, ..., \mathbf{v}_n\}$ una base de $\mathcal{V}$. Demostrar que el conjunto de vectores $\mathcal{B}' = \{\mathbf{u}_1, ..., \mathbf{u}_n\}$ dado por $\mathbf{u}_1 = \mathbf{v}_1$, $\mathbf{u}_2 = \mathbf{v}_1 + \mathbf{v}_2, ...,$ $\mathbf{u}_n = \mathbf{v}_1 + \mathbf{v}_2 + ... + \mathbf{v}_n$ es también una base de $\mathcal{V}$.

10. Sea $\mathcal{P}_4[x]$ el espacio vectorial de los polinomios de grado menor o igual que n. Demostrar que los conjuntos $\mathcal{B} = \{1, x, x^2, x^3, x^4\}$ y $\mathcal{B}' = \left\{(1+x)^4, x(1+x)^3, x^2(1+x)^2, x^3(1+x), x^4\right\}$ son bases de $\mathcal{P}_4[x]$. Expresar los elementos de $\mathcal{B}'$ como combinación lineal de $\mathcal{B}$.

11. Se consideran en $\mathbb{R}^4$ los subespacios vectoriales $\mathcal{W}_1$ y $\mathcal{W}_2$ generados por los subconjuntos $\mathcal{S}_1 = \{(1,1,1,1), (1,-1,1,-1)\}$ y $\mathcal{S}_2 = \{(1,1,0,1), (1,2,-1,2), (3,5,-2,5)\}$ respectivamente. Encontrar:

    (a) $\dim(\mathcal{W}_1 + \mathcal{W}_2)$.
    (b) $\dim(\mathcal{W}_1 \cap \mathcal{W}_2)$.
    (c) Ecuaciones de $\mathcal{W}_1 + \mathcal{W}_2$.
    (d) Ecuaciones de $\mathcal{W}_1 \cap \mathcal{W}_2$.

12. Idéntica cuestión para los subespacios de $\mathbb{R}^3$

$$\mathcal{U} = \{(x, y, z) \in \mathbb{R}^3 : x = \lambda + \mu, y = \lambda - 2\mu, z = -\mu\}$$

    y

$$\mathcal{V} = \{(x, y, z) \in \mathbb{R}^3 : x = 0, y = 3z\}.$$

13. Indicar cuál es la dimensión de la intersección de los subespacios de $\mathbb{R}^3$ definidos como $\mathcal{U} = L[(1,1,\alpha),(\alpha,1,1)]$ y $\mathcal{V} = L[(-1,\alpha,-1),(1,1,1)]$ según los valores de α.

14. En $\mathbb{R}^3$ se consideran los subespacios vectoriales $\mathcal{A} = \{(x,y,z) \in \mathbb{R}^3 : y = 0\}$ y $\mathcal{B} = \langle(1,1,1),(1,a,3)\rangle$, donde a es un parámetro real.

    (a) Calcula la dimensión de $\mathcal{A}$, $\mathcal{B}$, $\mathcal{A}+\mathcal{B}$ y $\mathcal{A}\cap\mathcal{B}$ en función de a.

    (b) Si $a = 1$, averiguar si existe algún valor de b para el cual el vector $(b,2,1)$ pertenece al subespacio $\mathcal{A}+\mathcal{B}$.

15. Halla una base de $\mathbb{R}^4$ que contenga a los vectores $(1,0,1,1)$ y $(2,0,2,1)$.

16. Calcular las ecuaciones de los siguientes subespacios vectoriales:

$$\text{(a) } \langle(1,1,1)\rangle \quad \text{(b) } \langle(1,-1,0),(1,0,0\rangle \quad \text{(c) } \langle(1,1,1),(0,0,3)\rangle \quad \text{(d) } \langle(1,1),(1,0)\rangle$$

17. ¿Cuál es la dimensión de $\mathbb{C}$ considerado como espacio vectorial sobre $\mathbb{R}$? ¿Y si lo consideramos como un espacio vectorial sobre $\mathbb{C}$?

18. Hallar una base y la dimensión del subespacio de $\mathbb{R}^4$:

$$\{(x,y,z,t) \in \mathbb{R}^4 : x-y-t=0,\ x+y+z+t=0\}.$$

19. En el espacio vectorial $\mathbb{R}^3$ se consideran los subespacios vectoriales $\mathcal{A} = \{(x,y,z) \in \mathbb{R}^3 : y = 0\}$ y $\mathcal{B} = \langle(1,1,1),(2,2,2)\rangle$. Se pide:

    (a) Hallar una base y la dimensión de $\mathcal{A}$, $\mathcal{B}$, $\mathcal{A}+\mathcal{B}$ y $\mathcal{A}\cap\mathcal{B}$.

    (b) ¿Pertenece el vector $(3,2,1)$ al subespacio $\mathcal{A}+\mathcal{B}$?

20. De las siguientes afirmaciones, demostrar las que sean ciertas y dar un contraejemplo para las falsas:

    (a) Si $\{\mathbf{u},\mathbf{v}\}$ son linealmente independientes, entonces $\{\mathbf{u}-\mathbf{v},\mathbf{u}+\mathbf{v}\}$ también lo son.

    (b) Todo conjunto de vectores que no contenga al vector nulo es linealmente independiente.

    (c) Sean $\mathcal{V}_1$ y $\mathcal{V}_2$ subespacios vectoriales del mismo espacio vectorial. Si $\dim(\mathcal{V}_1) = \dim(\mathcal{V}_2)$, entonces $\mathcal{V}_1 = \mathcal{V}_2$.

    (d) El vector $(1,0,0)$ tiene por coordenadas $(1,1,-1)$ en la base $\mathcal{B} = \{(1,1,1),(1,2,1),(1,3,2)\}$.

    (e) Sean $\mathcal{S} = \{(1,1,2),(1,-2,-1),(3,-1,2)\}$ y $\mathcal{S}' = \{(1,0,0)\}$. Entonces:

        i. $L(\mathcal{S}) + L(\mathcal{S}') = L(\mathcal{S}\cup\mathcal{S}')$.

        ii. $L(\mathcal{S}) \cup L(\mathcal{S}') = L(\mathcal{S}\cup\mathcal{S}')$.

        iii. $L(\mathcal{S}) \cap L(\mathcal{S}') = L(\mathcal{S}\cap\mathcal{S}')$.

        iv. $\dim(L(\mathcal{S})) = 3$ y $\dim(L(\mathcal{S}')) = 1$.

    (f) Si $\{\mathbf{u},\mathbf{v}\}$ es un conjunto de vectores linealmente independientes tal que $\{\mathbf{u},\mathbf{w}\}$ y $\{\mathbf{v},\mathbf{w}\}$ son linealmente independientes, entonces $\{\mathbf{u},\mathbf{v},\mathbf{w}\}$ también son linealmente independientes.

21. Dados los subespacios de $\mathbb{R}^3$, $\mathcal{S} = \{(x, y, z) \in \mathbb{R}^3 : x = y = 0\}$ y $\mathcal{T} = \{(x, y, z) \in \mathbb{R}^3 : x + y + z = 0\}$ calcular:

    (a) Una base y la dimensión de ambos.

    (b) $\mathcal{S} + \mathcal{T}$ y $\mathcal{S} \cap \mathcal{T}$, dando las bases de dichos subespacios.

    (c) ¿La suma $\mathcal{S} + \mathcal{T}$ es directa?

22. Se consideran en $\mathbb{R}^4$ los subespacios vectoriales generados por

$$\mathcal{S}_1 = \{(1, 1, 1, 1), (1, -1, -1, 1)\} \text{ y } \mathcal{S}_2 = \{(1, 1, 0, 1), (1, 2, -1, 2), (3, 5, -2, 5)\}.$$

    Calcular:

    (a) La base y la dimensión de $L(\mathcal{S}_1)$ y $L(\mathcal{S}_2)$.

    (b) Calcular $L(\mathcal{S}_1) \cap L(\mathcal{S}_2)$ y $L(\mathcal{S}_1) + L(\mathcal{S}_2)$, dando bases de dichos subespacios.

    (c) ¿Pertenece el vector $(4, 0, -2, 1)$ a $L(\mathcal{S}_1) \cap L(\mathcal{S}_2)$? En caso afirmativo dar sus coordenadas respecto de la base considerada.

    (d) ¿Pertenece el vector $(4, 0, -2, 1)$ a $L(\mathcal{S}_1) + L(\mathcal{S}_2)$? En caso afirmativo dar sus coordenadas respecto de la base considerada.

23. Determinar $a \in \mathbb{R}$ para que los vectores $(a, 1, 1)$, $(1, a, 1)$ y $(1, 1, a)$ sean linealmente independientes. Determinar los tres subespacios vectoriales generados por dos de los 3 vectores anteriores en función de a y calcular la intersección y la suma de dichos subespacios vectoriales dos a dos. ¿Son las sumas directas?

24. Sea $\mathcal{C}(0, 1])$ el espacio vectorial de las funciones continuas $f : [0, 1] \to \mathbb{R}$ y sea $\mathcal{W}_1 = \{ax + bx^3 : a, b \in \mathbb{R}\}$ y $\mathcal{W}_2 = \{a + bx + cx^2 : a, b, c \in \mathbb{R}\}$.

    (a) ¿Son $\mathcal{W}_1$ y $\mathcal{W}_2$ subespacios vectoriales de $\mathcal{C}(0, 1])$? Razona la respuesta.

    (b) Determinar $\mathcal{W}_1 + \mathcal{W}_2$ y $\mathcal{W}_1 \cap \mathcal{W}_2$. ¿Es la suma directa? Determinar las dimensiones de ambos subespacios.

25. Sea $\mathcal{V}$ un espacio vectorial sobre $\mathbb{R}$ y sean $\mathcal{S}_1 = \{\mathbf{u}_1, ..., \mathbf{u}_n\}$ y $\mathcal{S}_2 = \{\mathbf{v}_1, ..., \mathbf{v}_m\}$. Razonar la validez o falsedad de las siguientes afirmaciones:

    (a) Si $\mathcal{S}_1$ y $\mathcal{S}_2$ son linealmente independientes, entonces $L(\mathcal{S}_1) + L(\mathcal{S}_2)$ es una suma directa.

    (b) Si $L(\mathcal{S}_1) \oplus L(\mathcal{S}_2)$, entonces $\mathcal{S}_1 \cup \mathcal{S}_2$ es un conjunto linealmente independiente.

    (c) $L(\mathcal{S}_1) + L(\mathcal{S}_2) = L(\mathcal{S}_1 \cup \mathcal{S}_2)$.

Capítulo 3

Aplicaciones lineales

Sumario. Definición de aplicación lineal. Núcleo e imagen. Fórmula de las dimensiones. Matriz de una aplicación lineal respecto a una base. Propiedades. Matrices de cambio de base.

Una gran cantidad de modelos de la ingeniería siguen las siguientes leyes. Podemos imaginarnos un aparato que cada vez que se le introduce un determinado estímulo devuelve ese estímulo modificado como una señal de salida. Dicho estímulo puede ser por ejemplo una diferencia de potencial de potencial en un circuito eléctrico que a su vez devolverá una cierta intensidad de corriente. Si denotamos por $V(t)$ el voltaje e $i(t)$ la intensidad, el circuito funciona de la manera siguiente: si $i_1(t)$ e $i_2(t)$ son las respuestas a los voltajes $V_1(t)$ y $V_2(t)$, entonces $i_1(t) + i_2(t)$ es la respuesta al voltaje $V_1(t) + V_2(t)$. Además, si el voltaje es amplificado multiplicando por un escalar α, entonces $\alpha i(t)$ es la respuesta a $\alpha V(t)$. Esta es la forma de funcionar de un circuito LRC y la de numerosos ingenios de la ingeniería. Como veremos a continuación, estos aparatos funcionan como una aplicación lineal.

3.1 Definiciones y propiedades básicas

Sean $\mathcal{V}$ y $\mathcal{V}'$ dos espacios vectoriales sobre el cuerpo $\mathbb{K}$. Una aplicación $\mathbf{f} : \mathcal{V} \to \mathcal{V}'$ se dice que es una *aplicación lineal* si verifica las siguientes propiedades:

1. $\mathbf{f}(\mathbf{u} + \mathbf{v}) = \mathbf{f}(\mathbf{u}) + \mathbf{f}(\mathbf{v})$, para todo $\mathbf{u}, \mathbf{v} \in \mathcal{V}$.

2. $\mathbf{f}(\alpha \cdot \mathbf{u}) = \alpha \cdot \mathbf{f}(\mathbf{u})$, para todo $\mathbf{u} \in \mathcal{V}$ y todo $\alpha \in \mathbb{K}$.

Equivalentemente la aplicación $\mathbf{f}$ será lineal si y solamente si para cada $\mathbf{u}, \mathbf{v} \in \mathcal{V}$ y $\alpha, \beta \in \mathbb{K}$ se verifica la igualdad

$$\mathbf{f}(\alpha \cdot \mathbf{u} + \beta \cdot \mathbf{v}) = \alpha \cdot \mathbf{f}(\mathbf{u}) + \beta \cdot \mathbf{f}(\mathbf{v}).$$

Example 15 Dado un espacio vectorial $\mathcal{V}$, la aplicación identidad $\mathbf{i} : \mathcal{V} \to \mathcal{V}$ dada por $\mathbf{i}(\mathbf{v}) = \mathbf{v}$ para todo $\mathbf{v} \in \mathcal{V}$ es trivialmente una aplicación lineal.

Example 16 Sea la aplicación $\mathbf{f} : \mathbb{R}^2 \to \mathbb{R}^3$ dada por $\mathbf{f}(x, y) = (x,\, x + y,\, x - 2y)$. Veamos que se trata de una aplicación lineal. Para ello sean $\alpha, \beta \in \mathbb{R}$ y $(x_1, y_1), (x_2, y_2) \in \mathbb{R}^2$ y calculemos

$$\begin{aligned}
\mathbf{f}(\alpha \cdot (x_1, y_1) + \beta \cdot (x_2, y_2)) &= \mathbf{f}(\alpha x_1 + \beta x_2, \alpha y_1 + \beta y_2) \\
&= (\alpha x_1 + \beta x_2, \alpha x_1 + \beta x_2 + \alpha y_1 + \beta y_2, \alpha x_1 + \beta x_2 - 2\alpha y_1 - 2\beta y_2) \\
&= \alpha \cdot (x_1, x_1 + y_1, x_1 - 2y_1) + \beta \cdot (x_2, x_2 + y_2, x_2 - 2y_2) \\
&= \alpha \cdot \mathbf{f}(x_1, y_1) + \beta \cdot \mathbf{f}(x_2, y_2),
\end{aligned}$$

por lo que $\mathbf{f}$ es lineal.

Example 17 Consideremos la aplicación $\mathbf{g} : \mathbb{R}^2 \to \mathbb{R}^3$ dada por $\mathbf{g}(x_1,\, x_2) = (x_1^2,\, x_1 + x_2,\, x_1 - 2x_2)$. En este caso no se trata de una aplicación lineal, ya que tomando el vector $(1, 1) \in \mathbb{R}^2$ y el escalar $-1 \in \mathbb{R}$ se tiene que

$$\mathbf{g}((-1) \cdot (1, 1)) = (1, -2, 1) \neq (-1, -2, 1) = (-1) \cdot \mathbf{g}(1, 1).$$

Example 18 Sea la aplicación $\mathbf{f} : \mathbb{R}^2 \to \mathcal{M}_{2\times 2}(\mathbb{R})$ dada por

$$\mathbf{f}(x_1,\, x_2) = \begin{pmatrix} x_1 & 0 \\ 1 & x_2 \end{pmatrix}.$$

Si tomamos dos vectores cualesquiera $(x_1, x_2),\ (y_1, y_2) \in \mathbb{R}^2$ es evidente que

$$\mathbf{f}((x_1, x_2) + (y_1, y_2)) = \begin{pmatrix} x_1 + y_1 & 0 \\ 1 & x_2 + y_2 \end{pmatrix} \neq \begin{pmatrix} x_1 & 0 \\ 1 & x_2 \end{pmatrix} + \begin{pmatrix} y_1 & 0 \\ 1 & y_2 \end{pmatrix} = \mathbf{f}(x_1, x_2) + \mathbf{f}(y_1, y_2),$$

y por tanto $\mathbf{f}$ no es lineal.

Dependiendo de sus características se distinguen diferentes clases de aplicaciones lineales. Así dada $\mathbf{f} : \mathcal{V} \to \mathcal{V}'$ una aplicación lineal se dice que es un:

Monomorfismo si es inyectiva, es decir, si $\mathbf{f}(\mathbf{u}) = \mathbf{f}(\mathbf{v})$ implica que $\mathbf{u} = \mathbf{v}$.

Epimorfismo si es suprayectiva o equivalentemente, si para cada $\mathbf{v} \in \mathcal{V}'$ se tiene $\mathbf{u} \in \mathcal{V}$ tal que $\mathbf{f}(\mathbf{u}) = \mathbf{v}$.

Isomorfismo si es biyectiva, o lo que es lo mismo, si es inyectiva y suprayectiva a la vez.

Endomorfismo si $\mathcal{V} = \mathcal{V}'$.

Automorfismo si $\mathcal{V} = \mathcal{V}'$ y $\mathbf{f}$ es biyectiva.

La siguiente propiedad es útil para determinar si una aplicación no es lineal.

Proposition 18 *Si $\mathbf{0} \in \mathcal{V}$ es el vector nulo y $\mathbf{f} : \mathcal{V} \to \mathcal{V}'$ es una aplicación lineal, se tiene que $\mathbf{f}(\mathbf{0}) = \mathbf{0}$.*

Demostración. Nótese que $\mathbf{0} + \mathbf{0} = \mathbf{0}$ y entonces $\mathbf{f}(\mathbf{0} + \mathbf{0}) = \mathbf{f}(\mathbf{0})$. Como $\mathbf{f}$ es lineal se verifica que $\mathbf{f}(\mathbf{0} + \mathbf{0}) = \mathbf{f}(\mathbf{0}) + \mathbf{f}(\mathbf{0}) = \mathbf{f}(\mathbf{0})$, de donde $\mathbf{f}(\mathbf{0}) = \mathbf{0}$. ■

Esta propiedad permite identificar fácilmente aplicaciones que no son lineales (como la del ejemplo 18), aunque las aplicaciones que mandan el vector nulo al vector nulo no son necesariamente lineales, como el caso del ejemplo 17.

3.2 Subespacios vectoriales asociados a una aplicación lineal

3.2.1 Imagen de una aplicación lineal.

Las aplicaciones lineales son aquellas que conservan los subespacios vectoriales, como el siguiente resultado demuestra.

Proposition 19 *Sea* $\mathbf{f} : \mathcal{V} \to \mathcal{V}'$ *una aplicación lineal y sea* $\mathcal{W} \subseteq \mathcal{V}$ *un subespacio vectorial. Entonces* $\mathbf{f}(\mathcal{W}) = \{\mathbf{f}(\mathbf{u}) : \mathbf{u} \in \mathcal{W}\}$ *es un subespacio vectorial de* $\mathcal{V}'$.

Demostración. Sean $\alpha, \beta \in \mathbb{K}$ y $\mathbf{u}, \mathbf{v} \in \mathbf{f}(\mathcal{W})$ y veamos que $\alpha \cdot \mathbf{u} + \beta \cdot \mathbf{v} \in \mathbf{f}(\mathcal{W})$. Para ello hemos de tener en cuenta que como $\mathbf{u}, \mathbf{v} \in \mathbf{f}(\mathcal{W})$, entonces existen vectores $\mathbf{u}_1, \mathbf{v}_1 \in \mathcal{W}$ tales que $\mathbf{f}(\mathbf{u}_1) = \mathbf{u}$ y $\mathbf{f}(\mathbf{v}_1) = \mathbf{v}$. Entonces, por la linealidad de $\mathbf{f}$ tenemos

$$\alpha \cdot \mathbf{u} + \beta \cdot \mathbf{v} = \alpha \cdot \mathbf{f}(\mathbf{u}_1) + \beta \cdot \mathbf{f}(\mathbf{v}_1) = \mathbf{f}(\alpha \cdot \mathbf{u}_1 + \beta \cdot \mathbf{v}_1),$$

de donde obtenemos que $\alpha \cdot \mathbf{u} + \beta \cdot \mathbf{v} \in \mathbf{f}(\mathcal{W})$. ■

Definition 9 *Dada una aplicación lineal* $\mathbf{f} : \mathcal{V} \to \mathcal{V}'$ *se llama* imagen de $\mathbf{f}$, *y se denota por* $\mathrm{Im}(\mathbf{f})$, *al conjunto* $\mathbf{f}(\mathcal{V})$ *que como hemos visto en la proposición anterior es un subespacio vectorial de* $\mathcal{V}'$. *Además se tiene que* $\mathbf{f}$ *es un epimorfismo si y solamente si* $\mathrm{Im}(\mathbf{f}) = \mathcal{V}'$.

Example 19 Calculamos la imagen de la aplicación lineal $\mathbf{f} : \mathbb{R}^3 \to \mathbb{R}^3$ definida por $\mathbf{f}(x, y, z) = (x + y, x - z, 2x + y - z)$. Nótese que $(x, y, z) \in \mathrm{Im}\,\mathbf{f}$ si existe $(\alpha, \beta, \gamma) \in \mathbb{R}^3$ de manera que

$$\begin{aligned}(x, y, z) &= \mathbf{f}(\alpha, \beta, \gamma)\\ &= (\alpha + \beta, \alpha - \gamma, 2\alpha + \beta - \gamma),\end{aligned}$$

de donde obtenemos el sistema compatible

$$\begin{cases} \alpha + \beta = x, \\ \alpha - \gamma = y, \\ 2\alpha + \beta - \gamma = z. \end{cases}$$

Calculando el rango de ambas matrices del sistema

$$\left(\begin{array}{ccc|c} 1 & 1 & 0 & x \\ 1 & 0 & -1 & y \\ 2 & 1 & -1 & z \end{array}\right) \to_{\substack{F_2 - F_1 \\ F_3 - 2F_1}} \left(\begin{array}{ccc|c} 1 & 1 & 0 & x \\ 0 & -1 & -1 & y - x \\ 0 & -1 & -1 & z - 2x \end{array}\right) \to_{F_3 - F_2} \left(\begin{array}{ccc|c} 1 & 1 & 0 & x \\ 0 & -1 & -1 & y - x \\ 0 & 0 & 0 & z - x - y \end{array}\right)$$

obtenemos que para que el rango de ambas sea dos debe verificarse que $z - x - y = 0$, y así

$$\mathrm{Im}\,\mathbf{f} = \{(x, y, z) \in \mathbb{R}^3 : z - x - y = 0\}.$$

Veamos una manera alternativa de calcular la imagen mediante la siguiente propiedad.

Proposition 20 *Sea* $\mathbf{f} : \mathcal{V} \to \mathcal{V}'$ *una aplicación lineal y* $\mathcal{B} = \{\mathbf{u}_1, ..., \mathbf{u}_n\}$ *una base de* $\mathcal{V}$. *Entonces* $\mathbf{f}(\mathcal{B})$ *es un sistema generador de la imagen de* $\mathbf{f}$.

Demostración. Sea $\mathbf{v} \in \operatorname{Im} \mathbf{f}$. Entonces existe $\mathbf{u} \in \mathcal{V}$ de manera que $\mathbf{v} = \mathbf{f}(\mathbf{u})$. Como $\mathcal{B}$ es una base de $\mathcal{V}$, existen $\alpha_1, ..., \alpha_n \in \mathbb{K}$ de manera que $\mathbf{u} = \alpha_1 \cdot \mathbf{u}_1 + ... + \alpha_n \cdot \mathbf{u}_n$. Entonces

$$\begin{aligned} \mathbf{v} &= \mathbf{f}(\mathbf{u}) = \mathbf{f}(\alpha_1 \cdot \mathbf{u}_1 + ... + \alpha_n \cdot \mathbf{u}_n) \\ &= \alpha_1 \cdot \mathbf{f}(\mathbf{u}_1) + ... + \alpha_n \cdot \mathbf{f}(\mathbf{u}_n), \end{aligned}$$

por lo que $\mathbf{f}(\mathcal{B})$ genera la imagen. ■

Calculemos de nuevo la imagen del ejemplo 19. Sea $\mathcal{C} = \{(1,0,0),(0,1,0),(0,0,1)\}$ la base canónica de $\mathbb{R}^3$ y calculamos $\mathbf{f}(\mathcal{C}) = \{(1,1,2),(1,0,1),(0,-1,-1)\}$. Entonces

$$\operatorname{Im} \mathbf{f} = \mathcal{L}(\{(1,1,2),(1,0,1),(0,-1,-1)\}),$$

esto es, un vector $(x,y,z) \in \operatorname{Im} \mathbf{f}$ si y sólo si existen $\alpha, \beta, \gamma \in \mathbb{R}$ tales que

$$\begin{aligned} (x,y,z) &= \alpha \cdot (1,1,2) + \beta \cdot (1,0,1) + \gamma \cdot (0,-1,-1) \\ &= (\alpha + \beta, \alpha - \gamma, 2\alpha + \beta - \gamma), \end{aligned}$$

de donde obtenemos el mismo sistema compatible del ejemplo 19, y de ahí obtendremos su ecuación.

3.2.2 Núcleo de una aplicación lineal.

Sea $\mathbf{f} : \mathcal{V} \to \mathcal{V}'$ una aplicación lineal. Se llama *núcleo de* $\mathbf{f}$, y se denota $\operatorname{Ker}(\mathbf{f})$, a los vectores cuya imagen por $\mathbf{f}$ es $\mathbf{0}$, es decir,

$$\operatorname{Ker}(\mathbf{f}) = \{\mathbf{u} \in \mathcal{V} : \mathbf{f}(\mathbf{u}) = \mathbf{0}\}.$$

En primer lugar, comprobemos que el núcleo es un subespacio vectorial de $\mathcal{V}$. Para ello sean $\alpha, \beta \in \mathbb{K}$ y $\mathbf{u}, \mathbf{v} \in \operatorname{Ker}(\mathbf{f})$ y comprobemos que $\alpha \cdot \mathbf{u} + \beta \cdot \mathbf{v} \in \operatorname{Ker}(\mathbf{f})$ comprobando que su imagen es nula.

$$\mathbf{f}(\alpha \cdot \mathbf{u} + \beta \cdot \mathbf{v}) = \alpha \cdot \mathbf{f}(\mathbf{u}) + \beta \cdot \mathbf{f}(\mathbf{v}) = \alpha \cdot \mathbf{0} + \beta \cdot \mathbf{0} = \mathbf{0}.$$

El núcleo permite caracterizar a las aplicaciones lineales inyectivas o monomorfismos.

Proposition 21 *Sea* $\mathbf{f} : \mathcal{V} \to \mathcal{V}'$ *una aplicación lineal. Entonces* $\mathbf{f}$ *es inyectiva si y solo si* $\operatorname{Ker}(\mathbf{f}) = \{\mathbf{0}\}$.

Demostración. Si $\mathbf{f}$ es inyectiva, es evidente que $\operatorname{Ker}(\mathbf{f}) = \{\mathbf{0}\}$ porque si $u \in \operatorname{Ker}(\mathbf{f})$ se tiene que $\mathbf{f}(\mathbf{u}) = \mathbf{0} = \mathbf{f}(\mathbf{0})$ y por la inyectividad de $\mathbf{f}$ se tiene que $\mathbf{u} = \mathbf{0}$. Recíprocamente, si $\operatorname{Ker}(\mathbf{f}) = \{\mathbf{0}\}$ y $\mathbf{f}(\mathbf{u}) = \mathbf{f}(\mathbf{v})$, por linealidad $\mathbf{0} = \mathbf{f}(\mathbf{u}) - \mathbf{f}(\mathbf{v}) = \mathbf{f}(\mathbf{u} - \mathbf{v})$ lo cual implica que $\mathbf{u} - \mathbf{v} \in \operatorname{Ker}(\mathbf{f}) = \{\mathbf{0}\}$, de donde $\mathbf{u} = \mathbf{v}$. ■

Las dimensiones del núcleo y la imagen de una aplicación lineal están ligadas por la siguiente relación.

Theorem 22 *Sea* $\mathbf{f} : \mathcal{V} \to \mathcal{V}'$ *una aplicación lineal con* $\mathcal{V}$ *de dimensión finita. Se verifica la igualdad:*

$$\dim \mathcal{V} = \dim \operatorname{Ker}(\mathbf{f}) + \dim \operatorname{Im}(\mathbf{f}).$$

Demostración. Sea $\mathcal{B}_{\mathrm{Ker}(\mathbf{f})} = \{\mathbf{u}_1, ..., \mathbf{u}_n\}$ una base de $\mathrm{Ker}(\mathbf{f})$ y la completamos a una base $\mathcal{B} = \{\mathbf{u}_1, ..., \mathbf{u}_n, \mathbf{v}_1, ..., \mathbf{v}_m\}$ de $\mathcal{V}$. El resultado estará probado si demostramos que $\mathcal{B}' = \{\mathbf{f}(\mathbf{v}_1), ..., \mathbf{f}(\mathbf{v}_m)\}$ es una base de $\mathrm{Im}(\mathbf{f})$.

Veamos es primer lugar que $\mathrm{Im}(\mathbf{f}) = \mathcal{L}(\mathcal{B}')$. Para ello sea $\mathbf{u} \in \mathrm{Im}(\mathbf{f})$ y entonces debe existir $\mathbf{v} \in \mathcal{V}$ de manera que $\mathbf{f}(\mathbf{v}) = \mathbf{u}$. Existen entonces escalares $\alpha_1, ..., \alpha_n, \beta_1, ..., \beta_m \in \mathbb{K}$ de manera que $\mathbf{v} = \alpha_1 \cdot \mathbf{u}_1 + ... + \alpha_n \cdot \mathbf{u}_n + \beta_1 \cdot \mathbf{v}_1 + ... + \beta_m \cdot \mathbf{v}_m$. Así

$$\begin{aligned} \mathbf{u} &= \mathbf{f}(\mathbf{v}) = \mathbf{f}(\alpha_1 \cdot \mathbf{u}_1 + ... + \alpha_n \cdot \mathbf{u}_n + \beta_1 \cdot \mathbf{v}_1 + ... + \beta_m \cdot \mathbf{v}_m) \\ &= \alpha_1 \cdot \mathbf{f}(\mathbf{u}_1) + ... + \alpha_n \cdot \mathbf{f}(\mathbf{u}_n) + \beta_1 \cdot \mathbf{f}(\mathbf{v}_1) + ... + \beta_m \cdot \mathbf{f}(\mathbf{v}_m) \\ &= \beta_1 \cdot \mathbf{f}(\mathbf{v}_1) + ... + \beta_m \cdot \mathbf{f}(\mathbf{v}_m), \end{aligned}$$

dado que $\mathbf{f}(\mathbf{u}_i) = \mathbf{0}$ por ser todo $\mathbf{u}_i \in \mathrm{Ker}(\mathbf{f})$, $1 \leq i \leq n$. De esta manera tenemos que

$$\mathbf{u} = \beta_1 \cdot \mathbf{f}(\mathbf{v}_1) + ... + \beta_m \cdot \mathbf{f}(\mathbf{v}_m)$$

y así $\mathrm{Im}(\mathbf{f}) = \mathcal{L}(\mathcal{B}')$.

Probemos ahora para terminar que los vectores de $\mathcal{B}'$ son LI. Para ello consideramos una combinación lineal igualada a cero

$$\beta_1 \cdot \mathbf{f}(\mathbf{v}_1) + ... + \beta_m \cdot \mathbf{f}(\mathbf{v}_m) = \mathbf{0}.$$

Como $\mathbf{f}$ es lineal, la expresión anterior podemos agruparla en

$$\mathbf{f}(\beta_1 \cdot \mathbf{v}_1 + ... + \beta_m \cdot \mathbf{v}_m) = \mathbf{0},$$

de donde tenemos que $\beta_1 \cdot \mathbf{v}_1 + ... + \beta_m \cdot \mathbf{v}_m \in \mathrm{Ker}(\mathbf{f})$. Entonces existirán $\alpha_1, ..., \alpha_n \in \mathbb{K}$ tales que

$$\beta_1 \cdot \mathbf{v}_1 + ... + \beta_m \cdot \mathbf{v}_m = \alpha_1 \cdot \mathbf{u}_1 + ... + \alpha_n \cdot \mathbf{u}_n,$$

de donde

$$\beta_1 \cdot \mathbf{v}_1 + ... + \beta_m \cdot \mathbf{v}_m - \alpha_1 \cdot \mathbf{u}_1 - ... - \alpha_n \cdot \mathbf{u}_n = \mathbf{0},$$

y así $\beta_1 = ... = \beta_m = \alpha_1 = ... = \alpha_n = 0$ por ser $\mathcal{B}$ una base de $\mathcal{V}$. Entonces ya hemos probado que los vectores de $\mathcal{B}'$ son LI y la prueba concluye. ■

Example 20 El resultado anterior resulta de utilidad como muestra el siguiente ejemplo. Sea $\mathbf{f} : \mathbb{R}^3 \to \mathbb{R}^3$ dada por

$$\mathbf{f}(x, y, z) = (x - y, x - z, y - z),$$

y calculemos su núcleo e imagen. Como sabemos $(x, y, z) \in \mathrm{Ker}(\mathbf{f})$ si y sólo si

$$(0, 0, 0) = \mathbf{f}(x, y, z) = (x - y, x - z, -y - z),$$

de donde obtenemos el sistema

$$\begin{cases} x - y = 0, \\ x - z = 0, \\ -y - z = 0, \end{cases}$$

y al resolverlo

$$\left(\begin{array}{ccc|c} 1 & -1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 \end{array}\right) \to_{F_2 - F_1} \left(\begin{array}{ccc|c} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & -1 & 0 \end{array}\right) \to_{F_3 + F_2} \left(\begin{array}{ccc|c} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & -2 & 0 \end{array}\right),$$

obtenemos que $x = y = z = 0$, por lo que $\mathrm{Ker}(\mathbf{f}) = \{(0,0,0)\}$ y su dimensión es cero. Entonces

$$\dim \mathrm{Im}\, \mathbf{f} = \dim \mathbb{R}^3 - \dim \mathrm{Ker}(\mathbf{f}) = 3,$$

por lo que $\mathrm{Im}\, \mathbf{f}$ es un subespacio vectorial de $\mathbb{R}^3$ de dimensión tres, y por tanto debe verificarse que $\mathrm{Im}\, \mathbf{f} = \mathbb{R}^3$.

3.3 Matriz asociada a una aplicación lineal.

Sea $\mathbf{f} : \mathcal{V} \to \mathcal{V}'$ una aplicación lineal y sean $\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ y $\mathcal{B}' = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ bases de $\mathcal{V}$ y $\mathcal{V}'$, respectivamente. Dado $\mathbf{u} \in \mathcal{V}$ un vector cualquiera, existen escalares $\alpha_1, \ldots, \alpha_m \in \mathbb{K}$ (sus coordenadas en la base $\mathcal{B}$) de forma que $\mathbf{u} = \alpha_1 \cdot \mathbf{u}_1 + \cdots + \alpha_m \mathbf{u}_m$, y por linealidad de $\mathbf{f}$ tenemos que

$$\mathbf{f}(\mathbf{u}) = \alpha_1 \cdot \mathbf{f}(\mathbf{u}_1) + \cdots + \alpha_m \cdot \mathbf{f}(\mathbf{u}_m). \tag{3.1}$$

De esta igualdad se desprende que, conociendo las imágenes por una aplicación lineal de los vectores de una base del espacio inicial, es posible calcular la imagen por la aplicación lineal de cualquier vector. Por otra parte los vectores $\mathbf{f}(\mathbf{u}_1), ..., \mathbf{f}(\mathbf{u}_m)$ están en $\mathcal{V}'$ por lo que pueden escribirse como combinación lineal de los vectores de $\mathcal{B}'$. Sean entonces $\lambda_{1j}, \ldots, \lambda_{nj} \in \mathbb{K}$ las coordenadas de $\mathbf{f}(\mathbf{u}_j)$ en la base $\mathcal{B}'$, esto es,

$$\mathbf{f}(\mathbf{u}_j) = \lambda_{1j} \cdot \mathbf{v}_1 + \cdots + \lambda_{nj} \cdot \mathbf{v}_n, \quad 1 \leq j \leq m. \tag{3.2}$$

Combinando las ecuaciones (3.1) y (3.2) se obtiene la igualdad:

$$\mathbf{f}(\mathbf{u}) = \sum_{j=1}^{m} \alpha_j \cdot \mathbf{f}(\mathbf{u}_j) = \sum_{j=1}^{m} \alpha_j \left(\sum_{i=1}^{n} \lambda_{ij} \cdot \mathbf{v}_i \right) = \sum_{i=1}^{n} \left(\sum_{j=1}^{m} \lambda_{ij} \alpha_j \right) \cdot \mathbf{v}_i.$$

Sean $\beta_1, \ldots, \beta_n \in \mathbb{K}$ las coordenadas del vector $\mathbf{f}(\mathbf{u})$ en la base $\mathcal{B}'$. Como consecuencia de la unicidad de las coordenadas de un vector en una base y de la relación anterior, se tiene que para cada $1 \leq i \leq n$:

$$\beta_i = \sum_{j=1}^{m} \lambda_{ij} \alpha_j.$$

Introduciendo la matriz,

$$\mathbf{M}_{\mathcal{B}'\mathcal{B}}(\mathbf{f}) = \begin{pmatrix} \lambda_{11} & \lambda_{12} & \cdots & \lambda_{1m} \\ \lambda_{21} & \lambda_{22} & \cdots & \lambda_{2m} \\ \vdots & \vdots & & \vdots \\ \lambda_{n1} & \lambda_{n2} & \cdots & \lambda_{nm} \end{pmatrix} \in \mathcal{M}_{n \times m}(\mathbb{K}),$$

cuyas columnas corresponden a las coordenadas en la base $\mathcal{B}'$ de las imágenes por $\mathbf{f}$ de los vectores de la base $\mathcal{B}$, la relación anterior puede escribirse en la forma:

$$\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} = \mathbf{M}_{\mathcal{B}'\mathcal{B}}(\mathbf{f}) \cdot \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix}.$$

La matriz $\mathbf{M}_{\mathcal{B}'\mathcal{B}}(\mathbf{f})$ se denomina *matriz de la aplicación* $\mathbf{f}$ *en las bases* $\mathcal{B}$ *y* $\mathcal{B}'$ y permite obtener las coordenadas en la base $\mathcal{B}'$ de la imagen por $\mathbf{f}$ de cualquier vector a partir de sus coordenadas en la base $\mathcal{B}$, como indica la relación anterior.

Example 21 *Sea* $\mathbf{f}:\mathbb{R}^2\to\mathbb{R}^3$ *la aplicación lineal del Ejemplo 16 y sean* $\mathcal{C}_2$ *y* $\mathcal{C}_3$ *las bases canónicas de* $\mathbb{R}^2$ *y* $\mathbb{R}^3$*, respectivamente. Se tiene que*

$$\mathbf{f}(1,0)=(1,1,1),\qquad \mathbf{f}(0,1)=(0,1,-2),$$

de donde:

$$\mathbf{M}_{\mathcal{C}_3\mathcal{C}_2}(\mathbf{f})=\begin{pmatrix}1&0\\1&1\\1&-2\end{pmatrix}.$$

Si en $\mathbb{R}^2$ *se considera ahora la base* $\mathcal{B}=\{(1,1),\,(1,-1)\}$*, se tiene que*

$$\mathbf{f}(1,1)=(1,2,-1),\qquad \mathbf{f}(1,-1)=(1,0,3),$$

y la matriz de $\mathbf{f}$ *en estas bases es de la forma:*

$$\mathbf{M}_{\mathcal{C}_3\mathcal{B}}(\mathbf{f})=\begin{pmatrix}1&1\\2&0\\-1&3\end{pmatrix}.$$

Finalmente, si sobre $\mathbb{R}^3$ *se considera la base* $\mathcal{B}'=\{(1,0,0),\,(1,1,0),\,(1,1,1)\}$ *se tiene que* $(-1,3,-1)$ *son las coordenadas de* $\mathbf{f}(1,1)$ *en* $\mathcal{B}'$ *y* $(1,-3,3)$ *las de* $\mathbf{f}(1,-1)$ *y*

$$\mathbf{M}_{\mathcal{B}'\mathcal{B}}(\mathbf{f})=\begin{pmatrix}-1&1\\3&-3\\-1&3\end{pmatrix}.$$

3.3.1 Matriz de la suma de aplicaciones lineales y del producto por escalares.

Dadas dos aplicaciones lineales $\mathbf{f},\mathbf{g}:\mathcal{V}\to\mathcal{V}'$ se define su *suma* como la aplicación $\mathbf{f}+\mathbf{g}:\mathcal{V}\to\mathcal{V}'$ dada por $(\mathbf{f}+\mathbf{g})(\mathbf{u})=\mathbf{f}(\mathbf{u})+\mathbf{g}(\mathbf{u})$ para todo $\mathbf{u}\in\mathcal{V}$. Tenemos entonces el siguiente resultado.

Proposition 23 *Dadas dos aplicaciones lineales* $\mathbf{f},\mathbf{g}:\mathcal{V}\to\mathcal{V}'$ *se verifica:*

(a) La aplicación $\mathbf{f}+\mathbf{g}$ *es lineal.*

(b) Dadas las bases $\mathcal{B}=\{\mathbf{u}_1,...,\mathbf{u}_m\}$ *y* $\mathcal{B}'=\{\mathbf{v}_1,...,\mathbf{v}_n\}$ *de* $\mathcal{V}$ *y* $\mathcal{V}'$*, respectivamente, se tiene que*

$$\mathbf{M}_{\mathcal{B}'\mathcal{B}}(\mathbf{f}+\mathbf{g})=\mathbf{M}_{\mathcal{B}'\mathcal{B}}(\mathbf{f})+\mathbf{M}_{\mathcal{B}'\mathcal{B}}(\mathbf{g}).$$

Demostración. (a) Sean $\alpha,\beta\in\mathbb{K}$ y $\mathbf{u},\mathbf{v}\in\mathcal{V}$. Entonces, por la linealidad de $\mathbf{f}$ y $\mathbf{g}$

$$\begin{aligned}(\mathbf{f}+\mathbf{g})(\alpha\cdot\mathbf{u}+\beta\cdot\mathbf{v}) &= \mathbf{f}(\alpha\cdot\mathbf{u}+\beta\cdot\mathbf{v})+\mathbf{g}(\alpha\cdot\mathbf{u}+\beta\cdot\mathbf{v})\\ &= \alpha\cdot\mathbf{f}(\mathbf{u})+\beta\cdot\mathbf{f}(\mathbf{v})+\alpha\cdot\mathbf{g}(\mathbf{u})+\beta\cdot\mathbf{g}(\mathbf{v})\\ &= \alpha\cdot(\mathbf{f}(\mathbf{u})+\mathbf{g}(\mathbf{u}))+\beta\cdot(\mathbf{f}(\mathbf{v})+\mathbf{g}(\mathbf{v}))\\ &= \alpha\cdot(\mathbf{f}+\mathbf{g})(\mathbf{u})+\beta\cdot(\mathbf{f}+\mathbf{g})(\mathbf{v}),\end{aligned}$$

por lo que $\mathbf{f}+\mathbf{g}$ es lineal.

(b) Supongamos que $\mathbf{M}_{\mathcal{B}'\mathcal{B}}(\mathbf{f}+\mathbf{g}) = (\lambda_{ij})$, $\mathbf{M}_{\mathcal{B}'\mathcal{B}}(\mathbf{f}) = (\alpha_{ij})$ y $\mathbf{M}_{\mathcal{B}'\mathcal{B}}(\mathbf{g}) = (\beta_{ij})$. Entonces para $1 \leq j \leq m$

$$(\mathbf{f}+\mathbf{g})(\mathbf{u}_j) = \sum_{i=1}^{n} \lambda_{ij} \cdot \mathbf{v}_i,$$

y por otro lado

$$(\mathbf{f}+\mathbf{g})(\mathbf{u}_j) = \mathbf{f}(\mathbf{u}_j) + \mathbf{g}(\mathbf{u}_j) = \sum_{i=1}^{n} \alpha_{ij} \cdot \mathbf{v}_i + \sum_{i=1}^{n} \beta_{ij} \cdot \mathbf{v}_i = \sum_{i=1}^{n} (\alpha_{ij}+\beta_{ij}) \cdot \mathbf{v}_i,$$

por lo que $\lambda_{ij} = \alpha_{ij} + \beta_{ij}$, $1 \leq i \leq n$ y $1 \leq j \leq m$, y la fórmula es cierta. ■

Dado un escalar $\alpha \in \mathbb{K}$ y la aplicación lineal $\mathbf{f} : \mathcal{V} \to \mathcal{V}'$ se define la aplicación producto como $\alpha \cdot \mathbf{f} : \mathcal{V} \to \mathcal{V}'$ dada por $(\alpha \cdot \mathbf{f})(\mathbf{u}) = \alpha \cdot \mathbf{f}(\mathbf{u})$ para todo $\mathbf{u} \in \mathcal{V}$. Tenemos entonces el siguiente resultado.

Proposition 24 *Dado un escalar $\lambda \in \mathbb{K}$ y la aplicación lineal $\mathbf{f} : \mathcal{V} \to \mathcal{V}'$ se verifica:*

(a) La aplicación $\lambda \cdot \mathbf{f}$ es lineal.

(b) Dadas las bases $\mathcal{B} = \{\mathbf{u}_1, ..., \mathbf{u}_m\}$ y $\mathcal{B}' = \{\mathbf{v}_1, ..., \mathbf{v}_n\}$ de $\mathcal{V}$ y $\mathcal{V}'$, respectivamente, se tiene que

$$\mathbf{M}_{\mathcal{B}'\mathcal{B}}(\lambda \cdot \mathbf{f}) = \lambda \cdot \mathbf{M}_{\mathcal{B}'\mathcal{B}}(\mathbf{f}).$$

Demostración. (a) Sean $\alpha, \beta \in \mathbb{K}$ y $\mathbf{u}, \mathbf{v} \in \mathcal{V}$. Entonces, por la linealidad de $\mathbf{f}$

$$\begin{aligned}
(\lambda \cdot \mathbf{f})(\alpha \cdot \mathbf{u} + \beta \cdot \mathbf{v}) &= \lambda \cdot \mathbf{f}(\alpha \cdot \mathbf{u} + \beta \cdot \mathbf{v}) \\
&= \lambda \cdot (\alpha \cdot \mathbf{f}(\mathbf{u}) + \beta \cdot \mathbf{f}(\mathbf{v})) \\
&= \alpha \cdot (\lambda \cdot \mathbf{f}(\mathbf{u})) + \beta \cdot (\lambda \cdot \mathbf{f}(\mathbf{v})) \\
&= \alpha \cdot (\lambda \cdot \mathbf{f})(\mathbf{u}) + \beta \cdot (\lambda \cdot \mathbf{f})(\mathbf{v}),
\end{aligned}$$

por lo que $\lambda \cdot \mathbf{f}$ es lineal.

(b) Supongamos que $\mathbf{M}_{\mathcal{B}'\mathcal{B}}(\lambda \cdot \mathbf{f} + \mathbf{g}) = (\alpha_{ij})$ y $\mathbf{M}_{\mathcal{B}'\mathcal{B}}(\mathbf{f}) = (\beta_{ij})$. Entonces para $1 \leq j \leq m$

$$(\lambda \cdot \mathbf{f})(\mathbf{u}_j) = \sum_{i=1}^{n} \alpha_{ij} \cdot \mathbf{v}_i,$$

y por otro lado

$$(\lambda \cdot \mathbf{f} + \mathbf{g})(\mathbf{u}_j) = \lambda \cdot \mathbf{f}(\mathbf{u}_j) = \lambda \cdot \sum_{i=1}^{n} \beta_{ij} \cdot \mathbf{v}_i = \sum_{i=1}^{n} (\lambda \beta_{ij}) \cdot \mathbf{v}_i,$$

por lo que $\alpha_{ij} = \lambda \beta_{ij}$, $1 \leq i \leq n$ y $1 \leq j \leq m$, y la fórmula es cierta. ■

Example 22 Sean $\mathbf{f}, \mathbf{g} : \mathbb{R}^3 \to \mathbb{R}^3$ las aplicaciones lineales dadas por

$$\mathbf{f}(x,y,z) = (y+z, z+x, x+y),$$

y

$$\mathbf{g}(x,y,z) = (x, x+y, x+y+z).$$

Si denotamos por $\mathcal{C} = \{(1,0,0),(0,1,0),(0,0,1)\}$ la base canónica de $\mathbb{R}^3$, entonces

$$\mathbf{M}_{\mathcal{C}\mathcal{C}}(\mathbf{f}) = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

y

$$\mathbf{M}_{\mathcal{C}\mathcal{C}}(\mathbf{g}) = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}.$$

Se tiene entonces que

$$\begin{aligned} \mathbf{M}_{\mathcal{C}\mathcal{C}}(\mathbf{f} + 2 \cdot \mathbf{g}) &= \mathbf{M}_{\mathcal{C}\mathcal{C}}(\mathbf{f}) + 2 \cdot \mathbf{M}_{\mathcal{C}\mathcal{C}}(\mathbf{g}) \\ &= \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} + 2 \cdot \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 1 \\ 3 & 2 & 1 \\ 3 & 3 & 2 \end{pmatrix}. \end{aligned}$$

3.3.2 Matriz de la composición de aplicaciones lineales.

Sean $\mathcal{V}$, $\mathcal{V}'$ y $\mathcal{V}''$ tres espacios vectoriales con bases $\mathcal{B}$, $\mathcal{B}'$ y $\mathcal{B}''$, respectivamente. Sean las aplicaciones lineales $\mathbf{f} : \mathcal{V} \to \mathcal{V}'$ y $\mathbf{g} : \mathcal{V}' \to \mathcal{V}''$. Es sabido que la aplicación composición $\mathbf{g} \circ \mathbf{f} : \mathcal{V} \to \mathcal{V}''$ se define por $(\mathbf{g} \circ \mathbf{f})(\mathbf{u}) = \mathbf{g}(\mathbf{f}(\mathbf{u}))$ para todo $\mathbf{u} \in \mathcal{V}$. Se verifica entonces el siguiente resultado:

Proposition 25 *Sean las aplicaciones lineales* $\mathbf{f} : \mathcal{V} \to \mathcal{V}'$ *y* $\mathbf{g} : \mathcal{V}' \to \mathcal{V}''$ *definidas arriba. Entonces:*

(a) La aplicación $\mathbf{g} \circ \mathbf{f}$ *es lineal.*

(b) Si $\mathcal{B} = \{\mathbf{u}_1, ..., \mathbf{u}_m\}$ *y* $\mathcal{B}' = \{\mathbf{v}_1, ..., \mathbf{v}_n\}$ *y* $\mathcal{B}'' = \{\mathbf{w}_1, ..., \mathbf{w}_l\}$ *son bases de* $\mathcal{V}$, $\mathcal{V}'$ *y* $\mathcal{V}''$, *respectivamente, entonces*

$$\mathbf{M}_{\mathcal{B}''\mathcal{B}}(\mathbf{g} \circ \mathbf{f}) = \mathbf{M}_{\mathcal{B}''\mathcal{B}'}(\mathbf{g}) \cdot \mathbf{M}_{\mathcal{B}'\mathcal{B}}(\mathbf{f}).$$

Demostración. (a) Sean $\alpha, \beta \in \mathbb{K}$ y $\mathbf{u}, \mathbf{v} \in \mathcal{V}$. Entonces, por la linealidad de $\mathbf{f}$ y $\mathbf{g}$

$$\begin{aligned} (\mathbf{g} \circ \mathbf{f})(\alpha \cdot \mathbf{u} + \beta \cdot \mathbf{v}) &= \mathbf{g}(\mathbf{f}(\alpha \cdot \mathbf{u} + \beta \cdot \mathbf{v})) \\ &= \mathbf{g}(\alpha \cdot \mathbf{f}(\mathbf{u}) + \beta \cdot \mathbf{f}(\mathbf{v})) \\ &= \alpha \cdot (\mathbf{g}(\mathbf{f}(\mathbf{u}))) + \beta \cdot (\mathbf{g}(\mathbf{f}(\mathbf{v}))) \\ &= \alpha \cdot (\mathbf{g} \circ \mathbf{f})(\mathbf{u}) + \beta \cdot (\mathbf{g} \circ \mathbf{f})(\mathbf{v}), \end{aligned}$$

por lo que $\mathbf{g} \circ \mathbf{f}$ es lineal.

(b) Supongamos que $\mathbf{M}_{\mathcal{B}'\mathcal{B}}(\lambda \cdot \mathbf{f} + \mathbf{g}) = (\lambda_{ij})$, $\mathbf{M}_{\mathcal{B}'\mathcal{B}}(\mathbf{f}) = (\alpha_{ij})$ y $\mathbf{M}_{\mathcal{B}'\mathcal{B}}(\mathbf{g}) = (\beta_{ij})$. Entonces para $1 \leq j \leq m$

$$(\mathbf{g} \circ \mathbf{f})(\mathbf{u}_j) = \sum_{i=1}^{l} \lambda_{ij} \cdot \mathbf{w}_i,$$

y por otro lado

$$\begin{aligned}(\mathbf{g}\circ\mathbf{f})(\mathbf{u}_j) &= \mathbf{g}(\mathbf{f}(\mathbf{u}_j)) = \mathbf{g}\left(\sum_{k=1}^{n}\alpha_{kj}\cdot\mathbf{v}_k\right) = \sum_{k=1}^{n}\alpha_{kj}\cdot\mathbf{g}(\mathbf{v}_k) \\ &= \sum_{k=1}^{n}\alpha_{kj}\cdot\sum_{i=1}^{l}\beta_{ik}\cdot\mathbf{w}_i = \sum_{i=1}^{l}\left(\sum_{k=1}^{n}\beta_{ik}\alpha_{kj}\right)\mathbf{w}_i,\end{aligned}$$

por lo que $\lambda_{ij} = \sum_{k=1}^{n}\beta_{ik}\alpha_{kj}$, $1 \leq i \leq l$ y $1 \leq j \leq m$, y la fórmula es cierta. ■

Example 23 Sean $\mathbf{f} : \mathbb{R}^2 \to \mathbb{R}^3$ y $\mathbf{g} : \mathbb{R}^3 \to \mathbb{R}^3$ aplicaciones lineales dadas por

$$\mathbf{f}(x, y) = (x - y, x + y, 0),$$

y

$$\mathbf{g}(x, y, z) = (x, x - y, x + y - z).$$

Si denotamos por $\mathcal{C}_2 = \{(1, 0), (0, 1\}$ y $\mathcal{C}_3 = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ las bases canónicas de $\mathbb{R}^2$ y $\mathbb{R}^3$, respectivamente, obtenemos que

$$\mathbf{M}_{\mathcal{C}_3\mathcal{C}_2}(\mathbf{f}) = \begin{pmatrix} 1 & -1 \\ 1 & 1 \\ 0 & 0 \end{pmatrix}$$

y

$$\mathbf{M}_{\mathcal{C}_3\mathcal{C}_3}(\mathbf{g}) = \begin{pmatrix} 1 & 0 & 0 \\ 1 & -1 & 0 \\ 1 & 1 & -1 \end{pmatrix}.$$

Entonces

$$\begin{aligned}\mathbf{M}_{\mathcal{C}_3\mathcal{C}_2}(\mathbf{g}\circ\mathbf{f}) &= \mathbf{M}_{\mathcal{C}_3\mathcal{C}_3}(\mathbf{g})\cdot\mathbf{M}_{\mathcal{C}_3\mathcal{C}_2}(\mathbf{f}) \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 1 & -1 & 0 \\ 1 & 1 & -1 \end{pmatrix}\cdot\begin{pmatrix} 1 & -1 \\ 1 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 0 & -2 \\ 2 & 0 \end{pmatrix}.\end{aligned}$$

3.3.3 Matriz asociada a las aplicaciones lineales inversas.

Sean $\mathcal{V}$ y $\mathcal{V}'$ dos espacios vectoriales y sea $\mathbf{f} : \mathcal{V} \to \mathcal{V}'$ un isomorfismo, esto es, una aplicación lineal biyectiva. Por el hecho de ser biyectiva existe su aplicación inversa, esto es, una aplicación $\mathbf{f}^{-1} : \mathcal{V}' \to \mathcal{V}$ de manera que se verifica que:

$$\mathbf{f}\circ\mathbf{f}^{-1} = \mathbf{i}, \qquad \mathbf{f}^{-1}\circ\mathbf{f} = \mathbf{i}.$$

En estas condiciones, se verifica el siguiente resultado.

Proposition 26 *Sean $\mathcal{V}$ y $\mathcal{V}'$ dos espacios vectoriales y sea $\mathbf{f} : \mathcal{V} \to \mathcal{V}'$ un isomorfismo. Entonces:*

(a) La aplicación $\mathbf{f}^{-1} : \mathcal{V}' \to \mathcal{V}$ es lineal.

(b) Si $\mathcal{B}$ es una base de $\mathcal{V}$ y $\mathcal{B}'$ es una base de $\mathcal{V}'$, entonces

$$\mathbf{M}_{\mathcal{B}\mathcal{B}'}(\mathbf{f}^{-1}) = (\mathbf{M}_{\mathcal{B}'\mathcal{B}}(\mathbf{f}))^{-1}.$$

Demostración. (a) Sean $\alpha, \beta \in \mathbb{K}$ y $\mathbf{u}, \mathbf{v} \in \mathcal{V}'$ y comprobemos que $\mathbf{f}^{-1}(\alpha \cdot \mathbf{u} + \beta \cdot \mathbf{v}) = \alpha \cdot \mathbf{f}^{-1}(\mathbf{u}) + \beta \cdot \mathbf{f}^{-1}(\mathbf{v})$. Para ello, démonos cuenta que

$$\mathbf{f}(\mathbf{f}^{-1}(\alpha \cdot \mathbf{u} + \beta \cdot \mathbf{v})) = \alpha \cdot \mathbf{u} + \beta \cdot \mathbf{v},$$

por un lado, y por otro, debido a la linealidad de $\mathbf{f}$,

$$\mathbf{f}(\alpha \cdot \mathbf{f}^{-1}(\mathbf{u}) + \beta \cdot \mathbf{f}^{-1}(\mathbf{v})) = \alpha \cdot \mathbf{f}(\mathbf{f}^{-1}(\mathbf{u})) + \beta \cdot \mathbf{f}(\mathbf{f}^{-1}(v)) = \alpha \cdot \mathbf{u} + \beta \cdot \mathbf{v}.$$

Dado que $\mathbf{f}$ es biyectiva, se tiene la igualdad pedida.

(b) Sea $n = \dim \mathcal{V}$ y $m = \dim \mathcal{V}'$. Veamos en primer lugar que $n = m$. Para ello, usamos que por la biyectividad de $\mathbf{f}$ se verifica que $\operatorname{Im} \mathbf{f} = \mathcal{V}'$ y $\operatorname{Ker}(\mathbf{f}) = \{\mathbf{0}\}$ en virtud de la Proposición 21. Por el Teorema 22 se tiene que

$$n = \dim \mathcal{V} = \dim \operatorname{Ker}(\mathbf{f}) + \dim \operatorname{Im} \mathbf{f} = 0 + \dim \mathcal{V}' = m.$$

Por otro lado, es fácil verificar que

$$\mathbf{M}_{\mathcal{B}\mathcal{B}}(\mathbf{i}) = \mathbf{M}_{\mathcal{B}'\mathcal{B}'}(\mathbf{i}) = \mathbf{I}_n.$$

De la Proposición 25 obtenemos que

$$\mathbf{I}_n = \mathbf{M}_{\mathcal{B}\mathcal{B}}(\mathbf{i}) = \mathbf{M}_{\mathcal{B}\mathcal{B}'}(\mathbf{f}^{-1}) \cdot \mathbf{M}_{\mathcal{B}'\mathcal{B}}(\mathbf{f})$$

e

$$\mathbf{I}_n = \mathbf{M}_{\mathcal{B}'\mathcal{B}'}(\mathbf{i}) = \mathbf{M}_{\mathcal{B}'\mathcal{B}}(\mathbf{f}) \cdot \mathbf{M}_{\mathcal{B}\mathcal{B}'}(\mathbf{f}^{-1}),$$

por lo que $\mathbf{M}_{\mathcal{B}'\mathcal{B}}(\mathbf{f})$ es invertible y $\mathbf{M}_{\mathcal{B}\mathcal{B}'}(\mathbf{f}^{-1}) = (\mathbf{M}_{\mathcal{B}'\mathcal{B}}(\mathbf{f}))^{-1}$. ■

Example 24 Calculemos la aplicación lineal inversa de $\mathbf{f} : \mathbb{R}^2 \to \mathbb{R}^2$ dada por

$$\mathbf{f}(x, y) = (x + y, x - 2y).$$

Sea $\mathcal{C} = \{(1, 0), (0, 1)\}$ la base canónica de $\mathbb{R}^2$ y calculamos

$$\mathbf{M}_{\mathcal{C}\mathcal{C}}(\mathbf{f}) = \begin{pmatrix} 1 & 1 \\ 1 & -2 \end{pmatrix}.$$

Veamos que dicha matriz es invertible y calculemos su inversa

$$\left(\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ 1 & -2 & 0 & 1 \end{array}\right) \to_{F_2 - F_1} \left(\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ 0 & -3 & -1 & 1 \end{array}\right) \to_{-\frac{1}{3}F_2} \left(\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ 0 & 1 & \frac{1}{3} & -\frac{1}{3} \end{array}\right)$$
$$\to_{F_2 - F_1} \left(\begin{array}{cc|cc} 1 & 0 & \frac{2}{3} & \frac{1}{3} \\ 0 & 1 & \frac{1}{3} & -\frac{1}{3} \end{array}\right),$$

de donde

$$\mathbf{M}_{\mathcal{C}\mathcal{C}}(\mathbf{f}^{-1}) = (\mathbf{M}_{\mathcal{C}\mathcal{C}}(\mathbf{f}))^{-1} = \begin{pmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & -\frac{1}{3} \end{pmatrix}.$$

Entonces

$$\begin{aligned} \mathbf{f}^{-1}(x, y) &= \left(\mathbf{M}_{\mathcal{C}\mathcal{C}}(\mathbf{f}^{-1}) \cdot \begin{pmatrix} x \\ y \end{pmatrix}\right)^t \\ &= \left(\begin{pmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & -\frac{1}{3} \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix}\right)^t = \left(\frac{2}{3}x + \frac{1}{3}y, \frac{1}{3}x - \frac{1}{3}y\right). \end{aligned}$$

3.3.4 Matrices de cambio de base.

Sean $\mathcal{V}$ un espacio vectorial de dimensión n y $\mathcal{B}$ una base cualquiera de $\mathcal{V}$. Es evidente que la matriz de la aplicación identidad $\mathbf{i}$ respecto de la base $\mathcal{B}$ es la identidad, esto es, $\mathbf{M}_{\mathcal{B}\mathcal{B}}(\mathbf{i}) = \mathbf{I}_n$. Sin embargo esto no ocurre cuando se consideran bases distintas, es decir, supongamos que $\mathcal{B}'$ es otra base de $\mathcal{V}$. Entonces la matriz $\mathbf{M}_{\mathcal{B}'\mathcal{B}}(\mathbf{i})$ y $\mathbf{M}_{\mathcal{B}\mathcal{B}'}(\mathbf{i})$ reciben el nombre de *matrices de cambio de base*. Veamos con un ejemplo cómo calcularlas.

Example 25 Sean las bases $\mathcal{B} = \{(1,1), (1,-1)\}$ y $\mathcal{B}' = \{(1,0), (2,1)\}$ de $\mathbb{R}^2$. Entonces:

$$\mathbf{i}(1,1) = (1,1) = (-1)\,(1,0) + 1\,(2,1), \quad \mathbf{i}(1,-1) = (1,-1) = 3\,(1,0) + (-1)\,(2,1),$$

de donde

$$\mathbf{M}_{\mathcal{B}'\mathcal{B}}(\mathbf{i}) = \begin{pmatrix} -1 & 3 \\ 1 & -1 \end{pmatrix}.$$

Dado $\mathbf{u} \in \mathcal{V}$ un vector cualquiera, si $(\alpha_1, \ldots, \alpha_n)_{\mathcal{B}}$ y $(\beta_1, \ldots, \beta_n)_{\mathcal{B}'}$ son sus coordenadas respecto de las bases $\mathcal{B}$ y $\mathcal{B}'$, respectivamente, se tiene la relación:

$$\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}_{\mathcal{B}'} = \mathbf{M}_{\mathcal{B}'\mathcal{B}}(\mathbf{i}) \cdot \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}_{\mathcal{B}}.$$

Por otra parte, usando la Proposición 25 y la igualdad $\mathbf{i} \circ \mathbf{i} = \mathbf{i}$, tenemos que

$$\mathbf{I}_n = \mathbf{M}_{\mathcal{B}\mathcal{B}}(\mathbf{i}) = \mathbf{M}_{\mathcal{B}\mathcal{B}}(\mathbf{i} \circ \mathbf{i}) = \mathbf{M}_{\mathcal{B}\mathcal{B}'}(\mathbf{i}) \cdot \mathbf{M}_{\mathcal{B}'\mathcal{B}}(\mathbf{i}),$$

y

$$\mathbf{I}_n = \mathbf{M}_{\mathcal{B}'\mathcal{B}'}(\mathbf{i}) = \mathbf{M}_{\mathcal{B}'\mathcal{B}'}(\mathbf{i} \circ \mathbf{i}) = \mathbf{M}_{\mathcal{B}'\mathcal{B}}(\mathbf{i}) \cdot \mathbf{M}_{\mathcal{B}\mathcal{B}'}(\mathbf{i}),$$

de donde se desprende que las matrices de cambio de base son invertibles y

$$(\mathbf{M}_{\mathcal{B}\mathcal{B}'}(\mathbf{i}))^{-1} = \mathbf{M}_{\mathcal{B}'\mathcal{B}}(\mathbf{i}).$$

Example 26 Calculamos la matriz $\mathbf{M}_{\mathcal{B}\mathcal{B}'}(\mathbf{i})$ en el ejemplo 25. Para ello basta calcular la inversa de la matriz obtenida en dicho ejemplo

$$\left(\begin{array}{cc|cc} -1 & 3 & 1 & 0 \\ 1 & -1 & 0 & 1 \end{array}\right) \to_{F_2+F_1} \left(\begin{array}{cc|cc} -1 & 3 & 1 & 0 \\ 0 & 2 & 1 & 1 \end{array}\right) \to_{\substack{(-1)F_1 \\ \frac{1}{2}F_2}} \left(\begin{array}{cc|cc} 1 & -3 & -1 & 0 \\ 0 & 1 & \frac{1}{2} & \frac{1}{2} \end{array}\right)$$
$$\to_{F_1+3F_2} \left(\begin{array}{cc|cc} 1 & 0 & \frac{1}{2} & \frac{3}{2} \\ 0 & 1 & \frac{1}{2} & \frac{1}{2} \end{array}\right),$$

por lo que

$$\mathbf{M}_{\mathcal{B}\mathcal{B}'}(\mathbf{i}) = \begin{pmatrix} \frac{1}{2} & \frac{3}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

3.3.5 Matrices asociadas a una aplicación lineal en bases diferentes.

Sean $\mathcal{V}$ y $\mathcal{V}'$ dos espacios vectoriales y consideremos bases disntintas $\mathcal{B}_1$ y $\mathcal{B}_2$ de V y $\mathcal{B}'_1$ y $\mathcal{B}'_2$ de $\mathcal{V}'$. Dada la aplicación lineal $\mathbf{f} : \mathcal{V} \to \mathcal{V}'$, sabemos que las matrices de $\mathbf{f}$ respecto de bases distintas son distintas, esto es, las matrices $\mathbf{M}_{\mathcal{B}'_1\mathcal{B}_1}(\mathbf{f})$ y $\mathbf{M}_{\mathcal{B}'_2\mathcal{B}_2}(\mathbf{f})$ no son iguales. Ahora bien, existe una relación entre ambas que viene dada por las matrices de cambio de base. Consideremos el siguiente esquema:

$$\begin{array}{ccccc} & & \mathbf{f} & & \\ & \mathcal{V}_{\mathcal{B}_1} & \longrightarrow & \mathcal{V}'_{\mathcal{B}'_1} & \\ \mathbf{i} & \downarrow & & \uparrow & \mathbf{i} \\ & \mathcal{V}_{\mathcal{B}_2} & \longrightarrow & \mathcal{V}'_{\mathcal{B}'_2} & \\ & & \mathbf{f} & & \end{array}$$

donde por $\mathcal{V}_{\mathcal{B}_1}$ queremos indicar la base que tomamos en $\mathcal{V}$ para hacer la matriz de la aplicación lineal. Entonces, teniendo en cuenta que $\mathbf{i} \circ \mathbf{f} \circ \mathbf{i} = \mathbf{f}$, se sigue que

$$\mathbf{M}_{\mathcal{B}'_1\mathcal{B}_1}(\mathbf{f}) = \mathbf{M}_{\mathcal{B}'_1\mathcal{B}_1}(\mathbf{i} \circ \mathbf{f} \circ \mathbf{i}) = \mathbf{M}_{\mathcal{B}'_1\mathcal{B}'_2}(\mathbf{i}) \cdot \mathbf{M}_{\mathcal{B}'_2\mathcal{B}_2}(\mathbf{f}) \cdot \mathbf{M}_{\mathcal{B}_2\mathcal{B}_1}(\mathbf{i}),$$

es decir, podemos pasar de una matriz a otra mediante las matrices de cambio de base.

Example 27 Calculamos una aplicación lineal $\mathbf{f} : \mathbb{R}^3 \to \mathbb{R}^3$ cuyo núcleo es $\mathrm{Ker}(\mathbf{f}) = \{(x, y, z) \in \mathbb{R}^3 : x + y + z = 0\}$ y $\mathbf{f}(1, 1, 1) = (2, 2, 2)$. Para ello obtenemos una base del núcleo a partir de sus ecuaciones paramétricas

$$\begin{cases} x = -\lambda - \mu, & \\ y = \lambda, & \lambda, \mu \in \mathbb{R}, \\ z = \mu, & \end{cases}$$

por lo que una base del núcleo es $\mathcal{B}_{\mathrm{Ker}(\mathbf{f})} = \{(-1, 1, 0), (-1, 0, 1)\}$. Veamos a continuación que $\mathcal{B} = \{(-1, 1, 0), (-1, 0, 1), (1, 1, 1)\}$ es una base de $\mathbb{R}^3$ calculando el rango de la matriz

$$\begin{pmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix} \to_{\substack{F_2 - F_1 \\ F_3 + F_1}} \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 2 & 1 \end{pmatrix} \to_{F_3 + 2F_2} \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 3 \end{pmatrix},$$

que como vemos es 3. Dado que

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

y si denotamos por $\mathcal{C} = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ la base canónica de $\mathbb{R}^3$, se tiene que

$$\mathbf{M}_{\mathcal{CB}}(\mathbf{f}) = \begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & 2 \\ 0 & 0 & 2 \end{pmatrix}.$$

Entonces

$$\mathbf{M}_{\mathcal{CC}}(\mathbf{f}) = \mathbf{M}_{\mathcal{CB}}(\mathbf{f}) \cdot \mathbf{M}_{\mathcal{BC}}(\mathbf{i}).$$

Como

$$\mathbf{M}_{\mathcal{BC}}(\mathbf{i}) = (\mathbf{M}_{\mathcal{BC}}(\mathbf{i}))^{-1} = \begin{pmatrix} -1 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}^{-1},$$

calculamos la inversa

$$\left(\begin{array}{ccc|ccc} -1 & -1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{array}\right) \longrightarrow_{F_2+F_1} \left(\begin{array}{ccc|ccc} -1 & -1 & 1 & 1 & 0 & 0 \\ 0 & -1 & 2 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{array}\right) \longrightarrow_{F_3+F_2} \left(\begin{array}{ccc|ccc} -1 & -1 & 1 & 1 & 0 & 0 \\ 0 & -1 & 2 & 1 & 1 & 0 \\ 0 & 0 & 3 & 1 & 1 & 1 \end{array}\right)$$

$$\longrightarrow_{\substack{(-1)F_1 \\ (-1)F_2 \\ \frac{1}{3}F_3}} \left(\begin{array}{ccc|ccc} 1 & 1 & -1 & -1 & 0 & 0 \\ 0 & 1 & -2 & -1 & -1 & 0 \\ 0 & 0 & 1 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{array}\right) \longrightarrow_{\substack{F_1+F_3 \\ F_2+2F_3}} \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & -\frac{2}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & 1 & 0 & -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \\ 0 & 0 & 1 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{array}\right)$$

$$\longrightarrow_{F_1-F_2} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ 0 & 1 & 0 & -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \\ 0 & 0 & 1 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{array}\right),$$

de donde

$$\mathbf{M}_{\mathcal{BC}}(\mathbf{i}) = \begin{pmatrix} -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

y

$$\mathbf{M}_{\mathcal{CC}}(\mathbf{f}) = \begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & 2 \\ 0 & 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix} = \begin{pmatrix} \frac{2}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & \frac{2}{3} \end{pmatrix}.$$

Entonces la aplicación lineal es

$$\begin{aligned} \mathbf{f}(x,y,z) &= \left(\mathbf{M}_{\mathcal{CC}}(\mathbf{f}) \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix}\right)^t \\ &= \left(\begin{pmatrix} \frac{2}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & \frac{2}{3} \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix}\right)^t \\ &= \frac{2}{3}(x+y+z, x+y+z, x+y+z). \end{aligned}$$

3.4 Ejercicios

1. Determinar cuáles de las siguientes aplicaciones son lineales:

   (a) $\mathbf{f} : \mathbb{R}^2 \to \mathbb{R}^2$ dada por $\mathbf{f}(x,y) = (x-y, 2x-y^2)$.

   (b) $\mathbf{f} : \mathbb{R}^4 \to \mathbb{R}^2$ dada por $\mathbf{f}(x,y,z,u) = (x-y,\ u+z,\ z,\ 2x-y)$.

   (c) $\mathbf{f} : \mathbb{R}^2 \to \mathbb{R}^3$ dada por $\mathbf{f}(x,y) = (x-y,\ x-2y,\ 3y)$.

   (d) $\mathbf{f} : \mathbb{R}^3 \to \mathbb{R}^3$ dada por $\mathbf{f}(x,y,z) = (x-z+y,\ 2x-y-3z,\ z+y)$.

2. Sea $\mathcal{V}$ un espacio vectorial sobre un cuerpo $\mathbb{K}$ y sean $\mathbf{f}$ y $\mathbf{g}$ dos aplicaciones lineales de $\mathcal{V}$ en $\mathcal{V}$ de manera que para una base $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2\}$ se tiene que:

$$\mathbf{f}(\mathbf{u}_1) = \mathbf{u}_2 \quad \mathbf{f}(\mathbf{u}_2) = \mathbf{u}_1 \qquad \mathbf{g}(\mathbf{u}_1) = -\mathbf{u}_1 \quad \mathbf{g}(\mathbf{u}_2) = \mathbf{u}_2$$

   Demuestrar que las aplicaciones $\mathbf{f} \circ \mathbf{g}$ y $\mathbf{g} \circ \mathbf{f}$ son distintas.

3. Sea $\mathcal{V}$ un espacio vectorial sobre un cuerpo $\mathbb{K}$ y sea $\mathbf{f} : \mathcal{V} \to \mathcal{V}$ una aplicacion lineal. Se define el **conjunto invariante** de $\mathbf{f}$, denotado $\mathrm{Inv}(\mathbf{f})$, como el conjunto de vectores los vectores v que permanecen invariantes por la aplicación, es decir,

$$\mathrm{Inv}(\mathbf{f}) = \{\mathbf{v} \in \mathcal{V} : \mathbf{f}(\mathbf{v}) = \mathbf{v}\}\,.$$

   Demuestra que el conjunto invariante de una aplicación lineal es un subespacio vectorial de $\mathcal{V}$.

4. Dada una aplicación lineal $\mathbf{f} : \mathcal{V} \to \mathcal{V}$ demostrar que $\mathbf{f}^2 = \mathbf{0}$ (es decir, $\mathbf{f} \circ \mathbf{f}$ es la aplicación nula) si y sólo si $\mathrm{Im}(\mathbf{f}) \subseteq \mathrm{Ker}(\mathbf{f})$.

5. Sea $\mathcal{B}_1 = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\}$ una base del espacio vectorial $\mathcal{V}$ y sea:

$$\mathcal{B}_2 = \{\mathbf{u}_1,\ \mathbf{u}_1 + \mathbf{u}_2,\ \mathbf{u}_1 + \mathbf{u}_2 + \mathbf{u}_3,\ \mathbf{u}_1 + \mathbf{u}_2 + \mathbf{u}_3 + \mathbf{u}_4\}.$$

   (a) Demostrar que $\mathcal{B}_2$ es una base de $\mathcal{V}$.
   (b) Encontrar la matriz de cambio de base.
   (c) Hallar las coordenadas respecto de $\mathcal{B}_1$ de un vector cuyas coordenadas con respecto a la base $\mathcal{B}_2$ son $(1, -1, 0, 1)$.

6. Sea $\mathcal{P}_4[x]$ el espacio vectorial de los polinomios con coeficientes reales de grado menor o igual que cuatro. Dadas las siguientes bases:

$$\mathcal{B}_1 = \{x,\ x^2 + 1,\ 2x^4 + x^3,\ x^3 - x^2 + x,\ x^2 + x\}$$
$$\mathcal{B}_2 = \{2x^4 + 1,\ x^3 - 1,\ x^3 + 2x,\ x^2,\ x^3 - x^2\}$$

   (a) Halla la matriz de cambio de base de $\mathcal{B}_1$ a $\mathcal{B}_2$.
   (b) Halla las coordenadas respecto de dichas bases del polinomio $\mathbf{p}(x) = x^4 + x^3 + x^2 + x + 1$.

7. Consideremos la aplicación $\mathbf{D} : \mathcal{P}_4[x] \longrightarrow \mathcal{P}_3[x]$ de forma que si $\mathbf{p}(x) = a_4x^4 + a_3x^3 + a_2x^2 + a_1x + a_0 \in \mathcal{P}_4[x]$, se tiene que:

$$\mathbf{D}[\mathbf{p}(x)] = 4a_4x^3 + 3a_3x^2 + 2a_2x + a_1$$

   (a) Probar que $\mathbf{D}$ es una aplicación lineal.
   (b) Encontrar la matriz de $\mathbf{D}$ asociada a las bases canónicas de $\mathcal{P}_4[x]$ y $\mathcal{P}_3[x]$.

(c) Si sobre $\mathcal{P}_4[x]$ se considera la base

$$\mathcal{B} = \{(1+x)^4,\ (1+x)^3x,\ (1+x)^2x^2,\ (1+x)x^3,\ x^4\}$$

obtener la matriz de $\mathbf{D}$ en esta nueva base.

8. Dado $\mathbb{R}^4$ y el subespacio vectorial $\mathcal{W}$ cuyas ecuaciones respecto de la base $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\}$ son:

$$x_1 + x_2 - x_3 = 0 \qquad x_1 + x_2 + x_3 + x_4 = 0$$

Se elige una nueva base $\mathcal{B}' = \{\mathbf{u}_1 - \mathbf{u}_2,\ \mathbf{u}_2 - \mathbf{u}_3,\ \mathbf{u}_3 - \mathbf{u}_4,\ \mathbf{u}_4\}$.

(a) Calcular las ecuaciones de $\mathcal{W}$ respecto de esta nueva base.

(b) Obtener las coordenadas de $\mathbf{v} = \mathbf{u}_1 - \mathbf{u}_2$ respecto de $\mathcal{B}'$. ¿Pertenece este vector a $\mathcal{W}$?

9. En $\mathbb{R}^3$ se considera la base $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ y la aplicación lineal $\mathbf{f} : \mathbb{R}^3 \to \mathbb{R}^3$ definida respecto a esta base por la relación:

$$\mathbf{f}(x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3) = (x_2 + x_3)\mathbf{e}_1 + (x_1 + x_3)\mathbf{e}_2 + (x_2 - x_1)\mathbf{e}_3$$

(a) Calcular la matriz de $\mathbf{f}$ respecto a la base $\mathcal{B}$.

(b) Encontrar los vectores invariantes de $\mathbf{f}$ (ver ejercicio 3.).

(c) Calcular el núcleo y la imagen de $\mathbf{f}$.

(d) Determinar una base de $\mathrm{Ker}(\mathbf{f})$ y la ampliar a una base de $\mathbb{R}^3$.

(e) Hallar la matriz de $\mathbf{f}$ respecto a esta nueva base.

10. Sea $\mathbf{g} : \mathbb{R}^3 \to \mathbb{R}^3$ una aplicación lineal dada por:

$$\mathbf{g}(-1,1,3) = (6,-4,16) \qquad \mathbf{g}(-2,1,1) = (-2,-5,1) \qquad \mathbf{g}(3,2,-1) = (1,14,-12)$$

Hallar la matriz de $\mathbf{g}$ respecto a la base canónica de $\mathbb{R}^3$, las ecuaciones del núcleo y la imagen y una base de ambos subespacios.

11. Sea $\mathbf{f} : \mathbb{R}^4 \to \mathbb{R}^3$ una aplicación lineal definida por:

$$\mathbf{f}(1,1,1,1) = (0,0,1) \qquad \mathbf{f}(1,0,1,0) = (1,1,-1)$$
$$\mathbf{f}(1,1,1,0) = (0,0,-1) \qquad \mathbf{f}(-1,-2,0,0) = (1,1,1)$$

(a) Calcular la matriz de $\mathbf{f}$ respecto a las bases canónicas.

(b) Calcular la dimensión y ecuaciones de $\mathrm{Ker}(\mathbf{f})$ e $\mathrm{Im}(\mathbf{f})$.

(c) Obtener la matriz de $\mathbf{f}$ respecto de la base canónica de $\mathbb{R}^4$ y la base de $\mathbb{R}^3$,

$$\mathcal{B} = \{(1,1,1),\ (1,0,1),\ (1,1,0)\}$$

así como las ecuaciones de la imagen de $\mathbf{f}$ en esta última base.

(d) Encontrar la matriz de $\mathbf{f}$ respecto de las bases

$$\mathcal{B}_1 = \{(1,1,1,0),\ (1,1,0,0),\ (1,0,0,0),\ (1,1,1,1)\}$$

de $\mathbb{R}^4$ y

$$\mathcal{B}_2 = \{(1,1,1),\ (0,1,1),\ (0,0,1)\}$$

de $\mathbb{R}^3$ así como las ecuaciones del núcleo y la imagen de $\mathbf{f}$ en estas bases.

12. Sea $\mathbf{f} : \mathbb{R}^4 \to \mathbb{R}^2$ la aplicación cuya matriz asociada en las bases canónicas es:

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

(a) Calcular la expresión analítica de $\mathbf{f}$ en las bases canónicas.

(b) Obtener el núcleo, la imagen y el espacio invariante de $\mathbf{f}$.

(c) Calcular la matriz de $\mathbf{f}$ respecto a las bases

$$\mathcal{B} = \{(1,1,1,0),\ (1,1,0,0),\ (1,0,0,0),\ (1,1,1,1)\}$$

y

$$\mathcal{B}' = \{(1,1),\ (0,2)\}.$$

13. Sean $\mathbf{g}$ y $\mathbf{f}$ las aplicaciones lineales de los ejercicios 10 y 11. Calcula la matriz de la aplicación compuesta $\mathbf{g} \circ \mathbf{f}$ respecto de las bases canónicas de $\mathbb{R}^4$ y $\mathbb{R}^3$. Calcula además el rango, el núcleo, la imagen y el espacio invariante de dicha aplicación.

14. Dada $\mathbf{f} : \mathbb{R}^3 \to \mathbb{R}^3$ por $\mathbf{f}(x,y,z) = (x+y, y, 0)$, se pide:

(a) Demostrar que $\mathbf{f}$ es lineal.

(b) Hallar la dimensión de los subespacios $\mathrm{Ker}(\mathbf{f})$ e $\mathrm{Im}(\mathbf{f})$ así como bases de los mismos.

(c) Representa gráficamente los dos subespacios anteriores.

15. Calcula la expresión analítica de una aplicación lineal $\mathbf{f} : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ sabiendo que

$$\mathrm{Ker}(\mathbf{f}) = \left\{(x,y,z) \in \mathbb{R}^3 : x+y+z=0,\ x-y+2z=0\right\}$$

y $\mathbf{f}(1,0,0) = (-1,2,0)$, $\mathbf{f}(0,1,0) = (1,1,0)$.

16. Contesta de forma razonada si son verdaderas o falsas las siguientes afirmaciones,

(a) No hay aplicaciones inyectivas de $\mathbb{R}^3$ en $\mathbb{R}^2$.

(b) Las relaciones $\mathbf{f}(1,1,1) = (1,2)$, $\mathbf{f}(1,-1,1) = (-1,-2)$ y $\mathbf{f}(0,0,2) = (3,6)$ definen una aplicación lineal sobreyectiva de $\mathbb{R}^3$ sobre $\mathbb{R}^2$.

(c) Todas las aplicaciones lineales de $\mathbb{R}^4$ en $\mathbb{R}$ son sobreyectivas.

(d) Existe una aplicación lineal de $\mathbb{R}^3$ en $\mathbb{R}^5$ tal que su núcleo es la recta $\mathcal{U} = L[(1,0,1)]$ y su imagen es la recta $\mathcal{V} = L[(1,-1,1,2,4)]$.

(e) Si $\mathbf{f} : \mathcal{V} \to \mathcal{V}$ es una aplicación lineal y $\mathbf{v} \in \mathcal{V}$ de forma que $\mathbf{f}(\mathbf{v}) = \mathbf{f}(-\mathbf{v})$, entonces $\mathbf{v} = \mathbf{0}$.

(f) Si $\mathbf{f} : \mathcal{V} \to \mathcal{V}$ es una aplicación lineal tal que $\dim \operatorname{Ker}(\mathbf{f}) > 0$, entonces $\mathbf{f}^{-1}$ es una aplicación.

(g) Sea una aplicación lineal $\mathbf{f} : \mathcal{V} \to \mathcal{W}$. Si $\mathbf{f}$ es suprayectiva, entonces $\dim(\mathcal{V}) \geq \dim(\mathcal{W})$.

(h) Si $\mathbf{f} : \mathcal{V} \to \mathcal{V}$, entonces $\mathcal{V} = \operatorname{Ker}(\mathbf{f}) \oplus \operatorname{Im}(\mathbf{f})$.

(i) Las matrices

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

representan la misma aplicación lineal respecto de bases diferentes.

17. Halla los valores de parámetro real a para los cuales $\mathbb{R}^3$ es suma directa del núcleo y la imagen de $\mathbf{f} : \mathbb{R}^3 \to \mathbb{R}^3$ definida por la matriz (respecto de las base usual de $\mathbb{R}^{\nvDash}$)

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

18. Calcula el valor de a para el cual la intersección del espacio $\mathcal{U} = L[(a, 1, 1), (0, 1, -1)]$ y el núcleo de la aplicación $\mathbf{f} : \mathbb{R}^3 \to \mathbb{R}^2$ dada por $\mathbf{f}(x, y, z) = (x - y - z, y + az)$ es distinta del subespacio $\{\mathbf{0}\}$.

19. Halla para qué valor de a el vector $(1, 1, a)$ está en la imagen de la aplicación lineal $\mathbf{f} : \mathbb{R}^2 \to \mathbb{R}^3$ definida mediante las relaciones $\mathbf{f}(1, 2) = (1, -1, 1)$ y $\mathbf{f}(2, 1) = (0, 1, 1)$.

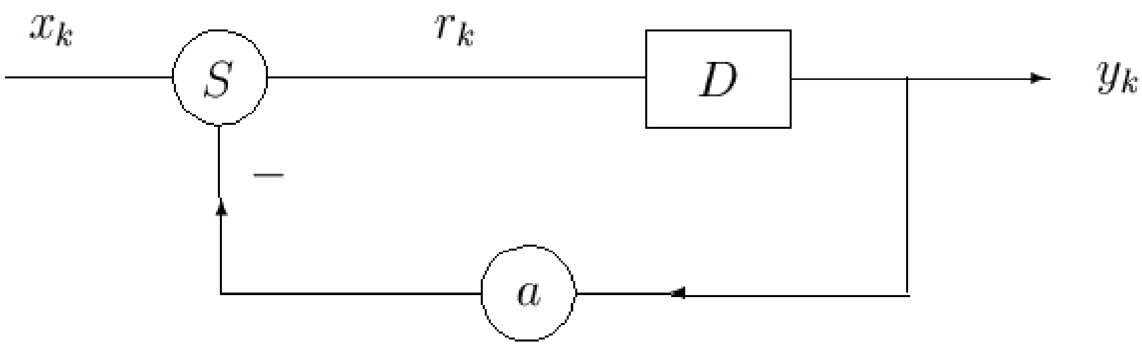
20. Calcula la dimensión del núcleo y la imagen de la aplicación $\mathbf{f}(x, y, z) = (ax - y - z, bx - z)$ en función de a y b.

Capítulo 4

Diagonalización de matrices cuadradas

Sumario. Valores y vectores propios. Polinomio característico. Bases de vectores propios. Caracterizción de matrices diagonalizables. Aplicaciones al cálculo de la potencia de una matriz. Aplicaciones.

Sea $\mathbf{A} \in \mathcal{M}_{n\times n}(\mathbb{R})$, $n \in \mathbb{N}$, una matriz cuadrada. A veces es necesario obtener la potencia $\mathbf{A}^k$ para $k \in \mathbb{N}$. Veámoslo con el siguiente ejemplo que proviene de la electrónica (ver [**?**]). Para fijar ideas, consideremos el siguiente ejemplo.



Este dispositivo está formado por dos elementos. El primero de ellos, marcado con una S, es un elemento que suma o resta datos, que a su vez vendrán modulados por números reales. El denotado por una D es un aparato que produce un retardo de una unidad temporal en la sucesión. La figura representa el tipo más sencillo de retroalimentación de una señal. Los datos de entrada vienen dados por la sucesión x_k y los de salida por

$$y_{k+1} = r_k. \tag{4.1}$$

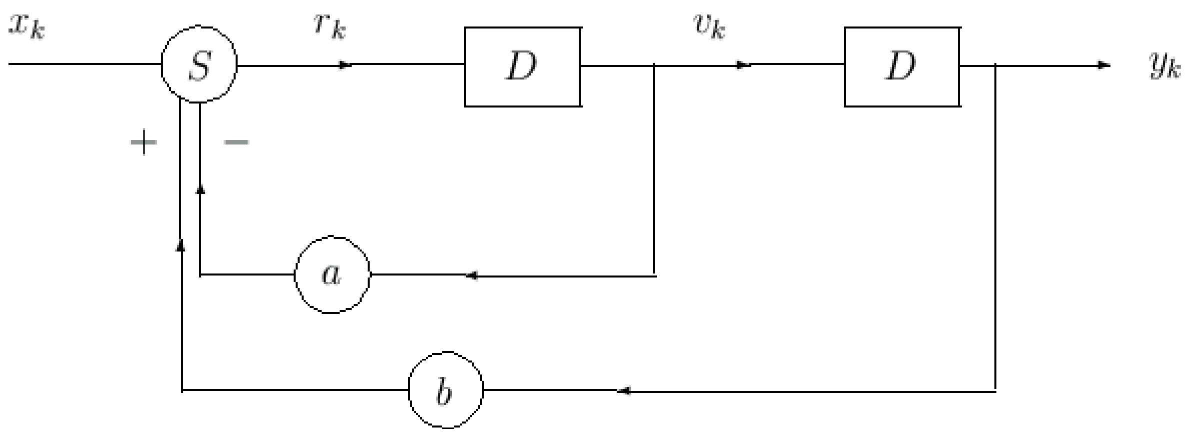
En el proceso, los datos intermedios r_k vienen dados por la expresión

$$r_k = x_k - a y_k, \tag{4.2}$$

donde a es un número real. Combinando (4.1) y (4.2) obtenemos la ecuación en diferencias de orden uno

$$y_{k+1} + a y_k = x_k.$$

Si complicamos el dispositivo, como se muestra en la figura,



se obtiene una ecuación de orden dos. Aquí

$$y_{k+1} = v_k,$$

$$v_{k+1} = r_k,$$

$$r_k = x_k + by_k - av_k,$$

de donde se obtiene la ecuación

$$y_{k+2} + ay_{k+1} - by_k = x_k.$$

Por ejemplo supongamos la ecuación

$$\begin{cases} y_{k+2} + y_{k+1} - 2y_k = 0; \\ y_0 = 0,\ y_1 = 1. \end{cases}$$

Introduciendo la variable $z_k = y_{k+1}$ la ecuación anterior se reescribe como

$$\begin{pmatrix} y_{k+1} \\ z_{k+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 2 & -1 \end{pmatrix} \cdot \begin{pmatrix} y_k \\ z_k \end{pmatrix}.$$

Entonces

$$\begin{pmatrix} y_{k+1} \\ z_{k+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 2 & -1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 2 & -1 \end{pmatrix} \cdot \begin{pmatrix} y_{k-1} \\ z_{k-1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 2 & -1 \end{pmatrix}^2 \cdot \begin{pmatrix} y_{k-1} \\ z_{k-1} \end{pmatrix},$$

e inductivamente

$$\begin{pmatrix} y_{k+1} \\ z_{k+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 2 & -1 \end{pmatrix}^{k+1} \cdot \begin{pmatrix} y_0 \\ z_0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 2 & -1 \end{pmatrix}^{k+1} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

dado que $z_0 = y_1 = 1$. Por lo tanto, para obtener la solución y_k hemos de obtener la potencia de una matriz. Para ello, vamos a intentar escribir la matriz en cuestión de forma “parecida” a una matriz diagonal ya que si $\mathbf{A} = (a_{ij})$ es diagonal, entonces $\mathbf{A}^k = (a_{ij}^k)$. Más precisamente, la matriz

$\mathbf{A}$ será *diagonalizable* si existe una matriz invertible $\mathbf{P}$ y una matriz diagonal $\mathbf{D}$ de manera que $\mathbf{A} = \mathbf{P} \cdot \mathbf{D} \cdot \mathbf{P}^{-1}$. Entonces

$$\mathbf{A}^k = \mathbf{P} \cdot \mathbf{D}^k \cdot \mathbf{P}^{-1}$$

para todo $k \geq 0$.

Vamos a ver en este tema cuándo la matriz cuadrada $\mathbf{A}$ es diagonalizable. Para ello, hemos de darnos cuenta en primer lugar que podemos construir una aplicación lineal $\mathbf{f_A} : \mathbb{R}^n \to \mathbb{R}^n$ de manera que para cada $(x_1, ..., x_n) \in \mathbb{R}^n$

$$\mathbf{f_A}(x_1, ..., x_n) = \left(\mathbf{A} \cdot \begin{pmatrix} x_1 \\ x_2 \\ ... \\ x_n \end{pmatrix} \right)^t .$$

Es inmediato ver que si $\mathcal{C}$ es la base canónica de $\mathbb{R}^n$, entonces $\mathbf{M}_{\mathcal{CC}}(\mathbf{f_A}) = \mathbf{A}$. Hablaremos entonces del núcleo de la matriz $\mathbf{A}$ como del núcleo de la aplicación lineal asociada $\mathbf{f_A}$, esto es, $\operatorname{Ker}(\mathbf{A}) = \operatorname{Ker}(\mathbf{f_A})$.

4.1 Valores y vectores propios de una matriz.

Empecemos estudiando en primer lugar qué propiedades tienen las matrices diagonales. Sea $\mathbf{D} = (d_{ij}) \in \mathcal{M}_{n\times n}(\mathbb{R})$ diagonal y sea como siempre $\mathcal{C} = \{(1,0,0,...,0),(0,1,0,...,0),...,(0,...,0,1)\}$ la base canónica de $\mathbb{R}^n$. Entonces para todo $1 \leq i \leq n$,

$$\mathbf{f_D}(0, ...0, \overset{i}{1}, 0, ..., 0) = \left(\mathbf{D} \cdot \begin{pmatrix} 0 \\ ... \\ 1 \\ ... \\ 0 \end{pmatrix} \right)^t = d_{ii} \cdot (0, ...0, \overset{i}{1}, 0, ..., 0).$$

Además, la matriz de $\mathbf{f_D}$ respecto a la base canónica es diagonal. Esta propiedad motiva la siguiente definición.

Definition 10 *Sea $\mathbf{A} \in M_{n\times n}(\mathbb{R})$. Se dice que $\lambda \in \mathbb{R}$ es un* valor propio *de la matriz $\mathbf{A}$ si existe un vector no nulo $\mathbf{x} \in \mathbb{R}^n$ de forma que*

$$\mathbf{A} \cdot \mathbf{x}^t = \lambda \cdot \mathbf{x}^t.$$

Si $\lambda \in \mathbb{R}$ es un valor propio de $\mathbf{A}$ y $\mathbf{x} \in \mathbb{R}^n$ es un vector verificando la relación anterior, se dice entonces que $\mathbf{x}$ es un vector propio *de $\mathbf{A}$ asociado al valor propio λ.*

Démonos cuenta que si

$$\mathbf{A} \cdot \mathbf{x}^t = \lambda \cdot \mathbf{x}^t$$

entonces

$$\mathbf{A} \cdot \mathbf{x}^t - \lambda \cdot \mathbf{I}_n \cdot \mathbf{x}^t = \mathbf{0}$$

o equivalentemente

$$(\mathbf{A} - \lambda \cdot \mathbf{I}_n) \cdot \mathbf{x}^t = \mathbf{0},$$

de donde se deduce que el conjunto de todos los vectores propios asociados a λ es el subespacio vectorial $\operatorname{Ker}(\mathbf{A} - \lambda \cdot \mathbf{I}_n)$ y recibe el nombre de *subespacio propio* asociado a λ. Veamos una importante propiedad de los subespacios propios.

Proposition 27 *Sean $\lambda_i \in \mathbb{R}$, $1 \leq i \leq k$, valores propios distintos de una matriz cuadrada* $\mathbf{A}$*. Entonces for $i \geq 2$,*

$$\begin{aligned} \{\mathbf{0}\} &= \mathrm{Ker}(\mathbf{A} - \lambda_i \cdot \mathbf{I}_n) \cap (\mathrm{Ker}(\mathbf{A} - \lambda_1 \cdot \mathbf{I}_n) + ... + \mathrm{Ker}(\mathbf{A} - \lambda_{i-1} \cdot \mathbf{I}_n)) \\ &= \mathrm{Ker}(\mathbf{A} - \lambda_i \cdot \mathbf{I}_n) \cap (\mathrm{Ker}(\mathbf{A} - \lambda_1 \cdot \mathbf{I}_n) \oplus ... \oplus \mathrm{Ker}(\mathbf{A} - \lambda_{i-1} \cdot \mathbf{I}_n)). \end{aligned}$$

En particular, $\mathrm{Ker}(\mathbf{A} - \lambda_1 \cdot \mathbf{I}_n) \oplus ... \oplus \mathrm{Ker}(\mathbf{A} - \lambda_i \cdot \mathbf{I}_n)$*.*

Demostración. Si $i = 2$, y $\mathbf{x} \in \mathrm{Ker}(\mathbf{A} - \lambda_1 \cdot \mathbf{I}_n) \cap \mathrm{Ker}(\mathbf{A} - \lambda_2 \cdot \mathbf{I}_n)$, se tiene que

$$\mathbf{A} \cdot \mathbf{x}^t = \lambda_1 \cdot \mathbf{x}^t = \lambda_2 \cdot \mathbf{x}^t,$$

y como $\lambda_1 \neq \lambda_2$, se deduce que $\mathbf{x} = \mathbf{0}$, con lo que $\mathrm{Ker}(\mathbf{A} - \lambda_1 \cdot \mathbf{I}_n) \cap \mathrm{Ker}(\mathbf{A} - \lambda_2 \cdot \mathbf{I}_n) = \{\mathbf{0}\}$, por lo que la suma es directa, esto es $\mathrm{Ker}(\mathbf{A} - \lambda_1 \cdot \mathbf{I}_n) \oplus \mathrm{Ker}(\mathbf{A} - \lambda_2 \cdot \mathbf{I}_n)$.

Suponiendo ahora que $\mathrm{Ker}(\mathbf{A} - \lambda_1 \cdot \mathbf{I}_n) \oplus ... \oplus \mathrm{Ker}(\mathbf{A} - \lambda_{i-1} \cdot \mathbf{I}_n)$, veamos que su intersección con $\mathrm{Ker}(\mathbf{A} - \lambda_i \cdot \mathbf{I}_n)$ es el vector nulo. Para ello, sea $\mathbf{x} \in \mathrm{Ker}(\mathbf{A} - \lambda_i \cdot \mathbf{I}_n) \cap (\mathrm{Ker}(\mathbf{A} - \lambda_1 \cdot \mathbf{I}_n) \oplus ... \oplus \mathrm{Ker}(\mathbf{A} - \lambda_{i-1} \cdot \mathbf{I}_n))$. Entonces $\mathbf{x} = \mathbf{x}_1 + ... + \mathbf{x}_{i-1}$, $\mathbf{x}_j \in \mathrm{Ker}(\mathbf{A} - \lambda_j \cdot \mathbf{I}_n)$, $j = 1, ..., i-1$. Así

$$\mathbf{A} \cdot \mathbf{x}^t = \lambda_i \cdot \mathbf{x}^t = \lambda_i \cdot (\mathbf{x}_1^t + ... + \mathbf{x}_{i-1}^t) = \lambda_i \cdot \mathbf{x}_1^t + ... + \lambda_i \cdot \mathbf{x}_{i-1}^t,$$

y por otro lado

$$\mathbf{A} \cdot \mathbf{x}^t = \mathbf{A} \cdot (\mathbf{x}_1^t + ... + \mathbf{x}_{i-1}^t) = \mathbf{A} \cdot \mathbf{x}_1^t + ... + \mathbf{A} \cdot \mathbf{x}_{i-1}^t = \lambda_1 \cdot \mathbf{x}_1^t + ... + \lambda_{i-1} \cdot \mathbf{x}_{i-1}^t.$$

Igualando las expresiones anteriores y simplificando

$$(\lambda_i - \lambda_1) \cdot \mathbf{x}_1^t + ... + (\lambda_i - \lambda_{i-1}) \cdot \mathbf{x}_{i-1}^t = \mathbf{0},$$

y como los vectores deben ser LI en caso de ser no nulos y λ_1,...,λ_i son distintos, debe verificarse que $\mathbf{x}_j = \mathbf{0}$ si $j = 1, ..., i-1$, y por tanto $\mathbf{x} = \mathbf{0}$. ■

El hecho de que la suma sea directa hace que si $\mathcal{B}_{\lambda_i}$ son bases de $\mathrm{Ker}(\mathbf{A} - \lambda_i \cdot \mathbf{I}_n)$, $1 \leq i \leq k$, entonces $\cup_{i=1}^k \mathcal{B}_{\lambda_i}$ es una base de $\mathrm{Ker}(\mathbf{A} - \lambda_1 \cdot \mathbf{I}_n) \oplus ... \oplus \mathrm{Ker}(\mathbf{A} - \lambda_k \cdot \mathbf{I}_n)$. Recordemos entonces que

$$\dim(\mathrm{Ker}(\mathbf{A} - \lambda_1 \cdot \mathbf{I}_n) \oplus ... \oplus \mathrm{Ker}(\mathbf{A} - \lambda_k \cdot \mathbf{I}_n)) = \sum_{i=1}^k \dim \mathrm{Ker}(\mathbf{A} - \lambda_i \cdot \mathbf{I}_n).$$

Tenemos entonces el siguiente resultado que es una primera caracterización de las matrices diagonalizables.

Theorem 28 *La matriz* $\mathbf{A}$ *es diagonalizable si y solo si $\lambda_1, ..., \lambda_k \in \mathbb{R}$ son todos los valores propios de la matriz* $\mathbf{A}$ *y $\sum_{i=1}^k \dim \mathrm{Ker}(\mathbf{A} - \lambda_i \cdot \mathbf{I}_n) = n$. Más aún, existe una base de vectores propios $\mathcal{B}$ de manera que:*

(a) La matriz asociada a $\mathbf{f_A}$ *en la base $\mathcal{B}$ es diagonal y de la forma*

$$\mathbf{M}_{\mathcal{B}\mathcal{B}}(\mathbf{f_A}) = \mathbf{D} = \begin{pmatrix} \lambda_1 & ... & 0 & ... & 0 & ... & 0 \\ ... & ... & ... & ... & ... & ... & ... \\ 0 & ... & \lambda_1 & ... & 0 & ... & 0 \\ ... & ... & ... & ... & ... & ... & ... \\ 0 & ... & 0 & ... & \lambda_k & ... & 0 \\ ... & ... & ... & ... & ... & ... & ... \\ 0 & ... & 0 & ... & 0 & ... & \lambda_k \end{pmatrix}$$

(b) La matriz $\mathbf{P}$ *es la matriz de cambio de base* $\mathbf{M}_{\mathcal{BC}}(\mathbf{i})$.

Demostración. Basta usar la Proposición 27 y las matrices asociadas a la aplicación lineal $\mathbf{f_A}$. ■

Hasta ahora no hemos puesto ningún ejemplo. La razón para esto es la siguiente: no sabemos cómo obtener los valores propios de una matriz cuadrada. Vamos a solucionar este problema en la siguiente sección.

4.2 El polinomio característico

El Teorema 28 nos dice que si tenemos una base de vectores propio la matriz $\mathbf{A}$ es diagonalizable. Sin embargo, no nos dice nada de cómo obtener algo clave en el resultado anterior como son los valores propios de la matriz. Una vez conocidos éstos podemos calcular los subespacios propios y sus dimensiones para posteriormente obtener bases de los mismos. Vamos a ver a continuación cómo calcular de una forma sencilla dichos valores propios.

Dada la matriz cuadrada $\mathbf{A} \in \mathcal{M}_{n\times n}(\mathbb{R})$, para que $\lambda \in \mathbb{R}$ sea un valor propio debe existir $\mathbf{x} \in \mathbb{R}^n$ no nulo de forma que:

$$\mathbf{A} \cdot \mathbf{x}^t = \lambda \cdot \mathbf{x}^t \quad \Leftrightarrow \quad (\mathbf{A} - \lambda \cdot \mathbf{I}_n) \cdot \mathbf{x}^t = \mathbf{0}.$$

Relación que en la terminología de los sistemas de ecuaciones lineales indica que el sistema homogéneo asociado a la matriz $\mathbf{A} - \lambda \cdot \mathbf{I}_n$ tiene una solución diferente de la nula y por ello es compatible indeterminado. Esto es equivalente, por el Teorema de Rouché-Frobenius, a que el rango de la matriz $\mathbf{A} - \lambda \cdot \mathbf{I}_n$ sea estrictamente menor que n, lo que a su vez equivale a que su determinante sea nulo. En resumen tenemos la siguiente propiedad:

Proposition 29 $\lambda \in \mathbb{R}$ *es un valor propio de* $\mathbf{A}$ *si y sólo si* $|\mathbf{A} - \lambda \cdot \mathbf{I}_n| = 0$.

Este resultado motiva la definición siguiente.

Definition 11 *Dada* $\mathbf{A} \in \mathcal{M}_{n\times n}(\mathbb{R})$*, se llama* polinomio característico *de* $\mathbf{A}$ *al polinomio de grado* n *y coeficientes reales definido por*

$$p(x) = |\mathbf{A} - x \cdot \mathbf{I}_n|.$$

El conjunto de las raíces del polinomio característico de $\mathbf{A}$, que se denota por $\sigma(\mathbf{A})$, se denomina *espectro* de la matriz $\mathbf{A}$. Evidentemente $\lambda \in \mathbb{R}$ será un valor propio de $\mathbf{A}$ si y solamente si $\lambda \in \sigma(\mathbf{A})$. Por otra parte, el Teorema Fundamental del Algebra asegura que $\sigma(\mathbf{A}) \neq \emptyset$, aunque puede que el espectro contenga números complejos, e incluso que ninguno de sus elementos sea un número real como ocurre con la matriz

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

Su polinomio característico es

$$p(x) = \begin{vmatrix} -x & -1 \\ 1 & -x \end{vmatrix} = x^2 + 1 = 0,$$

cuyo espectro es $\sigma(\mathbf{A}) = \{i, -i\}$, y por tanto dicha matriz no es diagonalizable en $\mathbb{R}$.

Dado un valor propio real $\lambda \in \sigma(\mathbf{A})$, sea $\mathcal{B}_\lambda$ una base del subespacio propio $\operatorname{Ker}(\mathbf{A} - \lambda \cdot \mathbf{I}_n)$ y extendámosla a una base $\mathcal{B}$ de $\mathbb{R}^n$. Entonces la matriz asociada

$$\mathbf{M}_{\mathcal{B}\mathcal{B}}(\mathbf{f_A}) = \begin{pmatrix} \lambda \cdot \mathbf{I}_k & \mathbf{M}_1 \\ \mathbf{0} & \mathbf{M}_2 \end{pmatrix},$$

donde $k = \dim \operatorname{Ker}(\mathbf{A} - \lambda \cdot \mathbf{I}_n)$, $\mathbf{0}$ es la matriz nula de tamaño $(n-k) \times k$ y $\mathbf{M}_1$ y $\mathbf{M}_2$ son matrices reales de tamaños $k \times (n-k)$ y $(n-k) \times (n-k)$, respectivamente. Entonces, si denotamos por $\mathcal{C}$ la base canónica de $\mathbb{R}^n$, el polinomio característico de $\mathbf{A}$ es

$$\begin{aligned} p(x) &= |\mathbf{A} - x \cdot \mathbf{I}_n| = |\mathbf{M}_{\mathcal{C}\mathcal{C}}(\mathbf{f_A}) - x \cdot \mathbf{I}_n| \\ &= |\mathbf{M}_{\mathcal{C}\mathcal{B}}(\mathbf{i}) \cdot \mathbf{M}_{\mathcal{B}\mathcal{B}}(\mathbf{f_A}) \cdot \mathbf{M}_{\mathcal{B}\mathcal{C}}(\mathbf{i}) - x \cdot \mathbf{M}_{\mathcal{C}\mathcal{B}}(\mathbf{i}) \cdot \mathbf{M}_{\mathcal{B}\mathcal{C}}(\mathbf{i})| \\ &= |\mathbf{M}_{\mathcal{C}\mathcal{B}}(\mathbf{i}) \cdot (\mathbf{M}_{\mathcal{B}\mathcal{B}}(\mathbf{f_A}) \cdot \mathbf{M}_{\mathcal{B}\mathcal{C}}(\mathbf{i}) - x \cdot \mathbf{I}_n) \cdot (\mathbf{M}_{\mathcal{C}\mathcal{B}}(\mathbf{i}))^{-1}| \\ &= |\mathbf{M}_{\mathcal{C}\mathcal{B}}(\mathbf{i})||\mathbf{M}_{\mathcal{B}\mathcal{B}}(\mathbf{f_A}) - x \cdot \mathbf{I}_n||(\mathbf{M}_{\mathcal{C}\mathcal{B}}(\mathbf{i}))^{-1}| \\ &= |\mathbf{M}_{\mathcal{B}\mathcal{B}}(\mathbf{f_A}) - x \cdot \mathbf{I}_n| = (x - \lambda)^k q(x), \end{aligned}$$

donde $q(x) = |\mathbf{M}_2 - x \cdot \mathbf{I}_{n-k}|$. Denotemos por $m(\lambda)$ la multiplicidad de λ, es decir, el polinomio característico se escribe como $p(x) = (x - \lambda)^{m(\lambda)} h(x)$, donde $h(x)$ es un polinomio tal que $h(\lambda) \neq 0$. Entonces se obtiene que $1 \leq k = \dim \operatorname{Ker}(\mathbf{A} - \lambda \cdot \mathbf{I}_n) \leq m(\lambda)$. Además si $\sigma(\mathbf{A}) = \{\lambda_1, \ldots, \lambda_r\}$:

$$\begin{aligned} r &\leq \dim\left(\operatorname{Ker}(\mathbf{A} - \lambda_1 \cdot \mathbf{I}_n) \oplus \cdots \oplus \operatorname{Ker}(\mathbf{A} - \lambda_r \cdot \mathbf{I}_n)\right) \\ &= \dim \operatorname{Ker}(\mathbf{A} - \lambda_1 \cdot \mathbf{I}_n) + \cdots + \dim \operatorname{Ker}(\mathbf{A} - \lambda_r \cdot \mathbf{I}_n) \leq n. \end{aligned}$$

Si $\mathbf{A} \in \mathcal{M}_{n \times n}(\mathbb{R})$ es diagonalizable, entonces $\sigma(\mathbf{A}) \subset \mathbb{R}$. El recíproco de la afirmación anterior es falso como nos muestra el ejemplo siguiente. Consideremos la matriz

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

cuyo polinomio característico es $p(\lambda) = (\lambda - 1)^2$. Por tanto $\sigma(A) = \{1\} \subset \mathbb{R}$, pero A no es diagonalizable ya que

$$\operatorname{Ker}(\mathbf{A} - \mathbf{I}_2) = \{(x, y) \in \mathbb{R}^2 : y = 0\},$$

que tiene dimensión 1. Entonces no podemos encontrar una base de vectores propios y por el Teorema 28 la matriz no es diagonalizable.

Las matrices diagonalizables quedan totalmente caracterizadas por el resultado siguiente que aumenta la información dada por el Teorema 28 y cuya demostración hemos obtenido anteriormente.

Theorem 30 *Sea $\mathbf{A} \in \mathcal{M}_{n \times n}(\mathbb{R})$. La matriz $\mathbf{A}$ es diagonalizable si y solamente si todas las raíces del polinomio característico son reales y además la dimensión del subespacio de los vectores propios asociados a cada valor propio coincide con la multiplicidad de dicho valor propio, es decir,*

$$\dim \operatorname{Ker}(\mathbf{A} - \lambda \cdot \mathbf{I}_n) = m(\lambda), \quad \forall\, \lambda \in \sigma(\mathbf{A}).$$

Sea la matriz

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

Su polinomio característico es $p(x) = |\mathbf{A} - x \cdot \mathbf{I}_3| = (3-x)(x^2-4x+3)$, y sus valores propios $\lambda_1 = 1$, $\lambda_2 = 3$ con multiplicidades $m(\lambda_1) = 1$ y $m(\lambda_2) = 2$. Calculamos a continuación los subespacios de vectores propios:

$$\begin{aligned} \mathrm{Ker}(\mathbf{A} - \mathbf{I}_n) &= \left\{(x,y,z) \in \mathbb{R}^3 : x - z = 0,\ y = 0\right\}, \\ \mathrm{Ker}(\mathbf{A} - 3 \cdot \mathbf{I}_n) &= \left\{(x,y,z) \in \mathbb{R}^3 : x - 2z = 0\right\}, \end{aligned}$$

de donde se deduce que $\dim \mathrm{Ker}(\mathbf{A} - \mathbf{I}_n) = 1$ y $\dim \mathrm{Ker}(\mathbf{A} - 3 \cdot \mathbf{I}_n) = 2$. La matriz es por tanto diagonalizable con matriz diagonal

$$\mathbf{D} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

Por otra parte obtenemos que $\{(1,0,1)\}$ y $\{(2,0,1),\ (0,1,0)\}$ son bases de $\mathrm{Ker}(\mathbf{A}-\mathbf{I}_n)$ y $\mathrm{Ker}(\mathbf{A}-3\cdot\mathbf{I}_n)$ respectivamente por lo que

$$\mathbf{P} = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

será la matriz de paso. Calculamos su inversa

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{array}\right) \quad \rightarrow_{F_3-F_1} \left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & -1 & 0 & -1 & 0 & 1 \end{array}\right) \rightarrow_{F_2\times F_3} \left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & -1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \end{array}\right)$$

$$\rightarrow_{F_1+2F_2} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & 0 & 2 \\ 0 & -1 & 0 & -1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \end{array}\right) \rightarrow_{(-1)F_2} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & 0 & 2 \\ 0 & 1 & 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 & 1 & 0 \end{array}\right),$$

de donde

$$\mathbf{P}^{-1} = \begin{pmatrix} -1 & 0 & 2 \\ 1 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$$

y así

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

Cabe hacer notar que las matrices de paso no son únicas: se tendrán diferentes matrices para cada familia de bases de los subespacios de vectores propios.

4.3 Aplicaciones

4.3.1 Circuitos digitales

Resolvamos ahora la ecuación en diferencias

$$\begin{cases} y_{k+2} + y_{k+1} - 2y_k = 0; \\ y_0 = 0,\ y_1 = 1, \end{cases}$$

introducida al comienzo del tema y que, con la variable $z_k = y_{k+1}$ se reescribe como

$$\begin{pmatrix} y_{k+1} \\ z_{k+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 2 & -1 \end{pmatrix} \cdot \begin{pmatrix} y_k \\ z_k \end{pmatrix},$$

por lo que

$$\begin{pmatrix} y_{k+1} \\ z_{k+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 2 & -1 \end{pmatrix}^{k+1} \cdot \begin{pmatrix} y_0 \\ z_0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 2 & -1 \end{pmatrix}^{k+1} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Calculemos los valores propios de la matriz

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

calculando previamente los valores propios

$$p(x) = \begin{vmatrix} -x & 1 \\ 2 & -1-x \end{vmatrix} = x^2 + x - 2 = 0,$$

de donde

$$x = \frac{-1 \pm \sqrt{1+8}}{2} = \frac{1 \pm 3}{2},$$

por lo que los valores propios son -1 y 2. Los subespacios propios son

$$\operatorname{Ker}(\mathbf{A} + \mathbf{I}_2) = \{(x, y) \in \mathbb{R}^2 : x = -y\},$$

$$\operatorname{Ker}(\mathbf{A} - 2 \cdot \mathbf{I}_2) = \{(x, y) \in \mathbb{R}^2 : 2x = y\},$$

y es fácil ver que $\mathcal{B} = \{(1, -1), (1, 2)\}$ es una base de vectores propios. La matriz diagonal es

$$\mathbf{D} = \begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix}$$

y las matrices de cambio de base son

$$\mathbf{P} = \begin{pmatrix} 1 & 1 \\ -1 & 2 \end{pmatrix}$$

y su inversa que calculamos a continuación.

$$\left(\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ -1 & 2 & 0 & 1 \end{array}\right) \to_{F_2+F_1} \left(\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ 0 & 3 & 1 & 1 \end{array}\right) \to_{\frac{1}{3}F_2} \left(\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ 0 & 1 & \frac{1}{3} & \frac{1}{3} \end{array}\right)$$
$$\to_{F_1-F_2} \left(\begin{array}{cc|cc} 1 & 0 & \frac{2}{3} & -\frac{1}{3} \\ 0 & 1 & \frac{1}{3} & \frac{1}{3} \end{array}\right),$$

con lo que

$$\mathbf{P}^{-1} = \begin{pmatrix} \frac{2}{3} & -\frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} \end{pmatrix}.$$

Así

$$\begin{aligned}
\begin{pmatrix} y_{k+1} \\ z_{k+1} \end{pmatrix} &= \begin{pmatrix} 0 & 1 \\ 2 & -1 \end{pmatrix}^{k+1} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix}^{k+1} \cdot \begin{pmatrix} \frac{2}{3} & -\frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} (-1)^{k+1} & 0 \\ 0 & 2^{k+1} \end{pmatrix} \cdot \begin{pmatrix} -\frac{1}{3} \\ \frac{1}{3} \end{pmatrix} \\
&= \frac{1}{3} \cdot \begin{pmatrix} (-1)^{k+2} + 2^{k+1} \\ (-1)^{k+2} + 2^{k+2} \end{pmatrix},
\end{aligned}$$

por lo que $y_{k+1} = \frac{1}{3}((-1)^{k+2} + 2^{k+1})$ y así la sucesión que buscamos es

$$y_k = \frac{(-1)^{k+1} + 2^k}{3}.$$

4.3.2 Procesos de Markov

Estudiamos la evolución de sistemas aislados (sin interacción con el exterior), cuya composición total no varía y de forma que su estado en un momento dado depende linealmente del estado en el momento inmediatamente anterior. Se trata de los llamados *procesos de Markov*. El ejemplo siguiente es ilustrativo de una situación de este tipo.

Supongamos que en una cierta ciudad existen dos compañías eléctricas X e Y encargadas del suministro. Cada año uno de cada diez usuarios de X decide cambiarse a Y y viceversa, dos de cada diez abonados a Y se pasan a B. Si denotamos por x_0 e y_0 la cantidad de clientes que X e Y respectivamente tienen este año y suponemos que la ciudad no crece, al año que viene se tendrá:

$$\begin{matrix} x_1 = 0.9\,x_0 + 0.2\,y_0 \\ y_1 = 0.1\,x_0 + 0.8\,y_0 \end{matrix} \quad \Leftrightarrow \quad \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix},$$

siendo x_1 e y_1 los clientes de X e Y al cabo de un año. Cuando pasen k años se tendrá (siempre suponiendo que el número de habitantes de la ciudad se mantiene constante):

$$\left\{ \begin{matrix} x_k = 0.9\,x_{k-1} + 0.2\,y_{k-1} \\ y_k = 0.1\,x_{k-1} + 0.8\,y_{k-1} \end{matrix} \right. \quad \Leftrightarrow \quad \begin{pmatrix} x_k \\ y_k \end{pmatrix} = \begin{pmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{pmatrix} \begin{pmatrix} x_{k-1} \\ y_{k-1} \end{pmatrix} = \begin{pmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{pmatrix}^k \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}.$$

Calculando los valores propios de la matriz

$$\mathbf{A} = \begin{pmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{pmatrix}$$

y estudiando sus subespacios de vectores propios se obtiene que es diagonalizable con lo que la expresión anterior puede reescribirse como

$$\begin{aligned}
\begin{pmatrix} x_k \\ y_k \end{pmatrix} &= \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0.7^k \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix}^{-1} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} \\
&= \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0.7^k \end{pmatrix} \begin{pmatrix} 1/3 & 1/3 \\ 1/3 & -2/3 \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}.
\end{aligned}$$

Esta fórmula ofrece una forma sencilla de calcular los clientes que tendrá cada una de las compañías al cabo de k años. Además cuando k sea grande 0.7^k se hará cada vez más pequeño por lo que cuando pase mucho tiempo ($k \to \infty$) se tendrá la relación

$$\begin{aligned} \begin{pmatrix} x_\infty \\ y_\infty \end{pmatrix} &= \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0.7^\infty \end{pmatrix} \begin{pmatrix} 1/3 & 1/3 \\ 1/3 & -2/3 \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1/3 & 1/3 \\ 1/3 & -2/3 \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} \frac{2}{3}(x_0 + y_0) \\ \frac{1}{3}(x_0 + y_0) \end{pmatrix}. \end{aligned}$$

Así la situación hacia la cual tiende el proceso y en la que se estabiliza es que $2/3$ de los habitantes de la ciudad contraten el suministro con la compañía X y $1/3$ lo haga con la Y.

4.4 Ejercicios

1. Hallar la forma diagonal, cuando sea posible, de las siguientes matrices, así como la matriz del cambio de base:

$$\text{(a)} \begin{pmatrix} 5 & 4 & 3 \\ -1 & 0 & -3 \\ 1 & -2 & 1 \end{pmatrix} \quad \text{(b)} \begin{pmatrix} 5 & 7 & 5 \\ -6 & -5 & -3 \\ 4 & 1 & 0 \end{pmatrix} \quad \text{(c)} \begin{pmatrix} -9 & 1 & 1 \\ -18 & 0 & 3 \\ -21 & 4 & 0 \end{pmatrix}$$

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

2. Sea $\mathbf{A}$ una matriz cuadrada. Demostrar que $\mathbf{A}$ y $\mathbf{A}^t$ tienen el mismo polinomio característico y, por tanto, los mismos valores propios. ¿Tendrían los mismos vectores propios? Comprobarlo para la matriz $\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$.

3. Sea $\mathbf{A}$ una matriz cuadrada de orden n, invertible. Demostrar que si λ es un valor propio de $\mathbf{A}$, entonces λ^{-1} es un valor propio de $\mathbf{A}^{-1}$.

4. Sea $\mathbf{A}$ una matriz 3×3 con valores propios 1 doble y 2 simple. Si los subespacios de vectores propios son $\{(x, y, z) : x = y = z\}$ y $\{(x, y, z) : x = -y\}$, respectivamente, se pide:

   (a) ¿Es $\mathbf{A}$ diagonalizable?

   (b) Calcular el rango y el determinante de $\mathbf{A}$.

   (c) Calcular la matriz $\mathbf{A}$.

   (d) Dado $\mathbf{b} \in \mathbb{R}^3$, ¿cómo es el sistema de ecuaciones $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$?

5. Sea $\mathbf{f} : \mathbb{R}^3 \to \mathbb{R}^3$ una aplicación lineal y $\mathbf{A}$ la matriz asociada a $\mathbf{f}$ en las bases canónicas de $\mathbb{R}^3$. Discutir la veracidad o falsedad de las siguientes afirmaciones:

   (a) Si $\mathbf{f}$ es biyectiva, entonces $\lambda = 0$ es un valor propio de $\mathbf{A}$.

(b) Si $\mathbf{f}$ no es sobre, entonces $\lambda = 0$ es un valor propio de $\mathbf{A}$.

6. Estudiar para qué valores de los parámetros las siguientes matrices son diagonalizables:

$$\text{(a)} \begin{pmatrix} \alpha & \beta & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{(b)} \begin{pmatrix} 5 & 0 & 3 \\ 0 & -1 & 0 \\ 3 & 0 & \alpha \end{pmatrix}.$$

7. Si $\mathbf{A}$ es diagonalizable, entonces existen matrices diagonal $\mathbf{D}$ e invertibe $\mathbf{P}$ tales que $\mathbf{A} = \mathbf{P} \cdot \mathbf{D} \cdot \mathbf{P}^{-1}$. Observa que si n es un número natural, entonces $\mathbf{A}^n = \mathbf{P} \cdot \mathbf{D}^n \cdot \mathbf{P}^{-1}$. Aplica este método para calcular la potencia n-ésima de la matriz

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

8. Sea $\mathbf{f}(x, y, z) = (4x-y,\ -6x+6y+3z,\ 3y)$ un endomorfismo de $\mathbb{R}^3$ en $\mathbb{R}^3$. Dados los subespacios $\mathcal{S} =< (1,1,1) >$ y $\mathcal{T} =< (1,1,0) >$, calcular $\mathbf{f}(\mathcal{S})$ y $\mathbf{f}(\mathcal{T})$. Comentar los resultados obtenidos en términos de valores y vectores propios de $\mathbf{f}$.

9. Sea $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^n$ una aplicación lineal con matriz asociada respecto a la base canónica $\mathbf{A}$ y $\mathbf{u}, \mathbf{v}$ dos vectores de $\mathbb{R}^n$ asociados a los valores propios distintos $\lambda, \mu \in \mathbb{R}$. Decidir cuales de las siguientes afirmaciones son ciertas.

   (a) El vector propio $\mathbf{u}$ tiene un único valor propio asociado.

   (b) Para todo $\alpha \in \mathbb{R}$, el vector $\alpha\mathbf{u}$ es un vector propio del valor propio λ.

   (c) Todo vector del núcleo de $\mathbf{f}$ es un vector propio.

   (d) El vector $\mathbf{w} = \mathbf{u} + \mathbf{v}$ es un vector propio de $\mathbf{f}$.

   (e) Si λ es un valor propio de $\mathbf{f}$, entonces λ^n es un valor propio de $\mathbf{f}^n$, donde $\mathbf{f}^n = \mathbf{f} \circ \mathbf{f} \circ .. \circ \mathbf{f}$ (n veces).

   (f) Una matriz tiene el valor propio 0 si y solo si su determinante es nulo.

   (g) Una matriz diagonalizable es invertible.

10. Sea $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^n$ una aplicación lineal de manera que $\mathbf{f} \circ \mathbf{f} = \mathbf{f}$. Calcular los valores propios de $\mathbf{f}$.

11. Calcular para las siguientes matrices los valores y vectores propios. Determinar si las matrices son o no diagonalizables y calcular la matriz diagonal en caso de serlo. Calcular además la base

respecto de la cual la matriz es diagonal y dar la matriz de cambio de base:

(a) $\begin{pmatrix} 1 & 1 \\ 3 & -1 \end{pmatrix}$ (b) $\begin{pmatrix} -26 & -15 \\ 50 & 29 \end{pmatrix}$ (c) $\begin{pmatrix} -2 & -1 \\ 4 & 2 \end{pmatrix}$

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

(g) $\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$ (h) $\begin{pmatrix} 1 & -2 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$ (i) $\begin{pmatrix} 1 & -1 & -1 \\ 1 & -1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$

(j) $\begin{pmatrix} 1 & 5 & 7 \\ 0 & 4 & 3 \\ 0 & 0 & 1 \end{pmatrix}$ (k) $\begin{pmatrix} 0 & 4 & 2 \\ -3 & 8 & 3 \\ 4 & -8 & -2 \end{pmatrix}$ (l) $\begin{pmatrix} 5 & 1 & -2 & 4 \\ 0 & 5 & 2 & 2 \\ 0 & 0 & 5 & 3 \\ 0 & 0 & 0 & 4 \end{pmatrix}$

(m) $\begin{pmatrix} 1 & 1 & 2 \\ 1 & 2 & 1 \\ 0 & 1 & 3 \end{pmatrix}$ (n) $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ (ñ) $\begin{pmatrix} 2 & 1 & -1 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$

12. Dada la sucesión de números reales definida por inducción como $x_{n+2} = x_{n+1} + x_n$, $x_1 = 1$, $x_2 = 2$, calcular x_n y el límite de la misma cuando $n \to \infty$.

13. En un cierto pais existen dos compañías eléctricas, luces y sombras S. A. y luz a gogó S. A. Cada año uno de cada diez consumidores se cambia de compañía. Si la población del pais no crece ni disminuye e inicialmente hay 10 millones de abonados a la primera compañía y 15 millones a la segunda, predecir la evolución del mercado a largo plazo.

14. En una ciudad existen tres supermercados $\mathcal{A}$, $\mathcal{B}$ y $\mathcal{C}$ que acaparan la totalidad de la población a la hora de comprar. Se ha observado que los compradores van cambiando año a año de supermercado según la siguiente ley: De los que compran en $\mathcal{A}$ un año vuelven al siguiente sólamente la mitad, mientras que la otra mitad pasa a comprar en $\mathcal{B}$. De los que compran en $\mathcal{B}$ la mitad permanece en $\mathcal{B}$, mientras que el resto compran el año siguiente en $\mathcal{C}$. Finalmente una cuarta parte de los compradores de $\mathcal{C}$ se cambian a $\mathcal{B}$, quedándose el resto en $\mathcal{C}$.

   (a) Si en un año determinado la mitad de la población compra en $\mathcal{A}$ y la otra mitad en $\mathcal{B}$, determinar cuál será la distribución el año siguiente.

   (b) Determinar cuál es el comportamiento a largo plazo de los compradores de la ciudad.

Capítulo 5

Espacio vectorial eucídeo

Sumario. Producto escalar. Norma. Ortogonalidad. Base ortonormal: teorema de Gramm–Schmidt. Diagonalización de matrices cuadradas simétricas. Proyección ortogonal. Teorema de la mejor aproximación. Aplicaciones lineales de significado geométrico.

5.1 Producto escalar

Sea $\mathcal{V}$ un espacio vectorial sobre el cuerpo $\mathbb{R}$ de los números reales. Una aplicación $\langle\cdot,\cdot\rangle : \mathcal{V}\times\mathcal{V} \to \mathbb{R}$, que a cada par de vectores $\mathbf{u},\mathbf{v} \in \mathcal{V}$ les asocia un número real $\langle\mathbf{u},\mathbf{v}\rangle \in \mathbb{R}$, se dice que es un *producto escalar* en $\mathcal{V}$ si verifica las siguientes propiedades:

(E1) Linealidad respecto de la primera coordenada, esto es, para cada $\alpha,\beta \in \mathbb{R}$ y cada $\mathbf{u},\mathbf{v},\mathbf{w} \in \mathcal{V}$, $\langle\alpha\mathbf{u}+\beta\mathbf{v},\mathbf{w}\rangle = \alpha\langle\mathbf{u},\mathbf{w}\rangle + \beta\langle\mathbf{v},\mathbf{w}\rangle$.

(E2) Conmutatividad. $\langle\mathbf{u},\mathbf{v}\rangle = \langle\mathbf{v},\mathbf{u}\rangle$, para cada $\mathbf{u},\mathbf{v} \in \mathcal{V}$.

(E3) Definida positiva. $\langle\mathbf{u},\mathbf{u}\rangle \geq 0$, para todo $\mathbf{u} \in\mathcal{V}$. Además $\langle\mathbf{u},\mathbf{u}\rangle = 0$ si y sólo si $\mathbf{u} = \mathbf{0}$.

Remark 3 *En ocasiones, especialmente en libros de Física, el producto escalar de dos vectores* $\mathbf{u},\mathbf{v} \in \mathcal{V}$ *se escribe como* $\mathbf{u}\cdot\mathbf{v}$*.*

El par $(\mathcal{V},\langle\cdot,\cdot\rangle)$ se llama *espacio vectorial euclídeo*. Como veremos a continuación, el producto escalar nos va a permitir medir distancias y algunas magnitudes asociadas a ésta.

Example 28 En el espacio vectorial $\mathbb{R}^n$, la aplicación que a cada par de vectores $(x_1,...,x_n)$ y $(y_1,...,y_n)$ les asocia el número real

$$\langle(x_1,...,x_n),(y_1,...,y_n)\rangle = (x_1,...,x_n)\cdot\begin{pmatrix} y_1 \\ ... \\ y_n \end{pmatrix} = x_1\,y_1 + \cdots + x_n\,y_n$$

es un ejemplo de producto escalar como se comprueba fácilmente. Este es el producto escalar usual de $\mathbb{R}^n$.

Example 29 Sea ahora $\mathcal{C}([a,b])$ el espacio vectorial de las funciones continuas definidas sobre el intervalo compacto $[a,b]$, $a,b \in \mathbb{R}$. Dadas $f,g\in \mathcal{C}([a,b])$ se define la aplicación

$$\langle f,g\rangle := \int_a^b f(x)g(x)dx.$$

De nuevo se comprueban fácilmente las propiedades (E1)–(E2) y que dado que $f(x)^2 \geq$, entonces $\langle f,f\rangle = \int_a^b f(x)^2dx \geq 0$. Es obvio que si $f(x)=0$, entonces $\langle f,f\rangle = 0$. Para comprobar el recíproco, supongamos que $\langle f,f\rangle = 0$ y que existe $x_0 \in [a,b]$ tal que $f(x_0) \neq 0$. Entonces $f(x_0)^2 > 0$, y por la continuidad de $f(x)^2$, debe existir un subintervalo $(c,d) \subset [a,b]$ tal que $f(x)^2 > 0$ para todo $x \in (c,d)$. Como $\langle f,f\rangle$ es el área determinada por $f(x)^2$ y el eje X, al ser $f(x)^2 > 0$ para todo $x \in (c,d)$, se tiene que dicha área es estrictamente positiva y por tanto $\langle f,f\rangle = 0$ es incompatible con la existencia de un $x_0 \in [a,b]$ tal que $f(x_0) \neq 0$, por lo que necesariamente $f(x)=0$ para todo $x \in [a,b]$.

De los axiomas de producto escalar se deducen fácilmente las siguientes propiedades.

Proposition 31 *Sea* $(\mathcal{V},\langle\cdot,\cdot\rangle)$ *un espacio vectorial euclídeo. Entonces*

(a) $\langle \mathbf{u},\mathbf{0}\rangle = 0$ *para todo* $\mathbf{u} \in \mathcal{V}$.

(b) Para cada $\alpha,\beta \in \mathbb{R}$ *y cada* $\mathbf{u},\mathbf{v},\mathbf{w} \in \mathcal{V}$, $\langle \mathbf{w}, \alpha\cdot\mathbf{u}+\beta\cdot\mathbf{v}\rangle = \alpha\langle\mathbf{w},\mathbf{u}\rangle + \beta\langle\mathbf{w},\mathbf{v}\rangle$.

Demostración. En primer lugar demostramos (a). Para ello basta tener en cuenta que

$$\langle \mathbf{u},\mathbf{0}\rangle = \langle \mathbf{u},\mathbf{0}+\mathbf{0}\rangle = \langle \mathbf{u},\mathbf{0}\rangle + \langle \mathbf{u},\mathbf{0}\rangle$$

de donde obtenemos que $\langle \mathbf{u},\mathbf{0}\rangle = 0$.

Para demostrar (b) consideramos

$$\langle \mathbf{w}, \alpha\cdot\mathbf{u}+\beta\cdot\mathbf{v}\rangle = \langle \alpha\cdot\mathbf{u}+\beta\cdot\mathbf{v}, \mathbf{w}\rangle = \alpha\langle\mathbf{u},\mathbf{w}\rangle + \beta\langle\mathbf{v},\mathbf{w}\rangle = \alpha\langle\mathbf{w},\mathbf{u}\rangle + \beta\langle\mathbf{w},\mathbf{v}\rangle,$$

mediante el uso de la conmutatividad y linealidad respecto de la primera componente del producto escalar. ■

Definition 12 *Sea* $(\mathcal{V},\langle\cdot,\cdot\rangle)$ *un espacio vectorial euclídeo. Se llama* norma *de* $\mathbf{u} \in \mathcal{V}$ *asociada al producto escalar a*

$$\|u\| = +\sqrt{\langle u,u\rangle}.$$

De los axiomas de producto escalar son inmediatas las siguientes propiedades de la norma asociada:

Proposition 32 *Sea* $(\mathcal{V},\langle\cdot,\cdot\rangle)$ *un espacio vectorial euclídeo. Entonces:*

(a) $\|\mathbf{u}\| \geq 0$ *para todo* $\mathbf{u} \in \mathcal{V}$. *Además* $\|\mathbf{u}\| = 0$ *si y sólo si* $\mathbf{u} = \mathbf{0}$.

(b) $\|\alpha\cdot\mathbf{u}\| = |\alpha|\|\mathbf{u}\|$, *para cada* $\alpha \in \mathbb{R}$ *y* $\mathbf{u} \in \mathcal{V}$.

(c) ***Desigualdad de Cauchy-Schwarz.*** *Dados* $\mathbf{u}, \mathbf{v} \in \mathcal{V}$ *se tiene la desigualdad*

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \, \|\mathbf{v}\|$$

que es estricta si $\mathbf{u}$ *y* $\mathbf{v}$ *son linealmente independientes.*

(d) ***Desigualdad de Minkowski o triangular.*** *Dados* $\mathbf{u}, \mathbf{v} \in \mathcal{V}$ *se verifica la relación:*

$$\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|.$$

(e) ***Regla del paralelogramo.*** $\|\mathbf{u} + \mathbf{v}\|^2 + \|\mathbf{u} - \mathbf{v}\|^2 = 2(\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2)$.

Demostración. La demostración de (a) es inmediata a partir de las definiciones. Para probar (b) hacemos el cálculo

$$\|\alpha \cdot \mathbf{u}\| = +\sqrt{\langle \alpha \cdot \mathbf{u}, \alpha \cdot \mathbf{u} \rangle} = +\sqrt{\alpha^2 \langle \mathbf{u}, \mathbf{u} \rangle} = |\alpha| \|\mathbf{u}\|.$$

Para probar (c) consideramos

$$0 \leq \langle \mathbf{u} + \lambda \cdot \mathbf{v}, \mathbf{u} + \lambda \cdot \mathbf{v} \rangle = \|\mathbf{u}\|^2 + 2\,\lambda \langle \mathbf{u}, \mathbf{v} \rangle + \lambda^2 \|\mathbf{v}\|^2$$

para cada $\lambda \in \mathbb{R}$, lo cual implica que el discriminante del anterior polinomio de grado dos en λ debe ser negativo, es decir,

$$4\langle \mathbf{u}, \mathbf{v} \rangle^2 - 4\,\|\mathbf{u}\|^2 \, \|\mathbf{v}\|^2 \leq 0 \quad \Leftrightarrow \quad |\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \, \|\mathbf{v}\|.$$

Notar que si $\mathbf{u}$ y $\mathbf{v}$ son linealmente independientes todas las desigualdades anteriores son estrictas.

Para probar (d) consideramos

$$\begin{aligned} \|\mathbf{u} + \mathbf{v}\|^2 &= \langle \mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{u} \rangle + 2\langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{v}, \mathbf{v} \rangle \\ &\leq \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2\|\mathbf{u}\| \, \|\mathbf{v}\| = (\|\mathbf{u}\| + \|\mathbf{v}\|)^2, \end{aligned}$$

de donde se deduce la propiedad.

Por último, para probar (e) desarrollamos

$$\begin{aligned} \|\mathbf{u} + \mathbf{v}\|^2 + \|\mathbf{u} - \mathbf{v}\|^2 &= \langle \mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v} \rangle + \langle \mathbf{u} - \mathbf{v}, \mathbf{u} - \mathbf{v} \rangle \\ &= \langle \mathbf{u}, \mathbf{u} \rangle + 2\langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{v}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{u} \rangle - 2\langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{v}, \mathbf{v} \rangle \\ &= 2(\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2). \end{aligned}$$

■

Como consecuencia de la desigualdad de Cauchy-Schwarz, dados dos vectores $\mathbf{u}, \mathbf{v}$ en el espacio euclídeo $\mathcal{V}$ se tiene que:

$$-1 \leq \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\| \, \|\mathbf{v}\|} \leq 1.$$

De aquí se define el *ángulo* formado por los vectores $\mathbf{u}$ y $\mathbf{v}$ como el número real θ de forma que:

$$\cos\theta = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\| \, \|\mathbf{v}\|}.$$

Así el producto escalar verifica la fórmula

$$\langle \mathbf{u}, \mathbf{v} \rangle = \|\mathbf{u}\| \, \|\mathbf{v}\| \cos\theta.$$

Remark 4 *En caso de $\mathbb{R}^2$, dados los vectores (x_1, y_1) y (x_2, y_2) y α el ángulo físico que forman, es decir, la longitud del arco de circunferencia de radio uno determinado por las semirectas $r_i = \{\lambda \cdot (x_i, y_i) : \lambda \geq 0\}$, $i = 1, 2$. Es fácil ver que existen $\alpha_i \in [0, 2\pi)$, $i = 1, 2$, de manera que*

$$(x_i, y_i) = ||(x_i, y_i)|| \cdot (\cos \alpha_i, \sin \alpha_i),\ i = 1, 2.$$

Entonces el ángulo generado por ambos vectores es $\alpha_1 - \alpha_2$. Así,

$$\begin{aligned} \frac{\langle (x_1, y_1), (x_2, y_2) \rangle}{||(x_1, y_1)||\, ||(x_2, y_2)||} &= \langle (\cos \alpha_1, \sin \alpha_1), (\cos \alpha_2, \sin \alpha_2) \rangle \\ &= \cos \alpha_1 \cos \alpha_2 + \sin \alpha_1 \sin \alpha_2 \\ &= \cos(\alpha_1 - \alpha_2), \end{aligned}$$

por lo que el ángulo formado por dos vectores coincide con la definición geométrica en el caso plano.

Example 30 Por ejemplo, el ángulo α generado por los vectores $(1, 1)$ y $(1, 0)$ se calcula mediante la expresión

$$\cos \alpha = \frac{\langle (1, 1), (1, 0) \rangle}{||(1, 1)||\, ||(1, 0)||} = \frac{1}{\sqrt{2}} = \frac{\sqrt{2}}{2},$$

de donde

$$\alpha = \arccos \frac{\sqrt{2}}{2} = \frac{\pi}{4}.$$

Un vector $\mathbf{u} \in \mathcal{V}$ se dice *normal* o *unitario* si $||\mathbf{u}|| = 1$. La siguiente proposición nos dice cómo obtener un vector normal de una forma sencilla.

Proposition 33 *Sea $(\mathcal{V}, \langle \cdot, \cdot \rangle)$ un espacio vectorial euclídeo y sea $\mathbf{u} \in \mathcal{V}$. Entonces $\frac{1}{||\mathbf{u}||} \cdot \mathbf{u}$ es normal.*

Demostración. Basta tener en cuenta que

$$\left\langle \frac{1}{||\mathbf{u}||} \cdot \mathbf{u}, \frac{1}{||\mathbf{u}||} \cdot \mathbf{u} \right\rangle = \frac{1}{||\mathbf{u}||^2} \langle \mathbf{u}, \mathbf{u} \rangle = 1,$$

con lo que concluye la demostración. ■

Por ejemplo, dado el vector $(1, 1) \in \mathbb{R}^2$, está claro que $||(1, 1)|| = \sqrt{2}$, por lo que no es normal. Entonces el vector

$$\frac{1}{\sqrt{2}}(1, 1) = \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2} \right)$$

es normal o unitario.

5.2 Ortogonalidad

Definition 13 *Dos vectores $\mathbf{u}, \mathbf{v} \in \mathcal{V}$ se dice que son* ortogonales *si su producto escalar es nulo, esto es, $\langle \mathbf{u}, \mathbf{v} \rangle = 0$. Un conjunto de vectores $\mathcal{S} = \{u_1, \ldots, u_r\} \subseteq \mathcal{V} \setminus \{\mathbf{0}\}$ se dice que es un* sistema ortogonal *si cada vector es ortogonal a todos los demás, es decir, si $\langle \mathbf{u}_i, \mathbf{u}_j \rangle = 0$ para cada $i \neq j$.*

Los vectores $(1,0)$ y $(0,1)$ son claramente ortogonales. Otro ejemplo menos claro es el siguiente: dados los polinomios 1 y x, definidos en $[-1,1]$, se verifica que con el producto escalar definido mediante la integral del producto de ambas funciones en dicho intervalo, se verifica

$$\langle 1,x\rangle = \int_{-1}^{1} x dx = 0.$$

Tenemos entonces el siguiente resultado que nos viene a decir que todo sistema ortogonal está compuesto por vectores linealmente independientes.

Proposition 34 *Sea $(\mathcal{V},\langle\cdot,\cdot\rangle)$ un espacio vectorial euclídeo y sea $\mathcal{S}=\{\mathbf{u}_1,...,\mathbf{u}_r\}$ un sistema ortogonal. Entonces los vectores de $\mathcal{S}$ son linealmente independientes.*

Demmostración. Planteamos la siguiente combinación lineal:

$$\alpha_1\cdot\mathbf{u}_1+\cdots+\alpha_r\cdot\mathbf{u}_r=\mathbf{0},$$

para cada $1\le j\le r$ se tiene que:

$$0=\langle\alpha_1\cdot\mathbf{u}_1+\cdots+\alpha_r\cdot\mathbf{u}_r,u_j\rangle=\sum_{i=1}^{r}\alpha_i\langle\mathbf{u}_i,\mathbf{u}_j\rangle=\alpha_j\left\|\mathbf{u}_j\right\|^2$$

y, dado que $\mathbf{u}_j\neq\mathbf{0}$, se deduce que $\alpha_j=0$. ■

Uno de los resultados clásicos, y nunca mejor dicho, de las familias de vectores ortogonales es el conocido Teorema de Pitágoras.

Theorem 35 ***(Pitágoras)*** *Sea $(\mathcal{V},\langle\cdot,\cdot\rangle)$ un espacio vectorial euclídeo. Dados $\mathbf{u},\mathbf{v}\in\mathcal{V}$ dos vectores ortogonales se verifica que*

$$\|\mathbf{u}+\mathbf{v}\|^2=\|\mathbf{u}\|^2+\|\mathbf{v}\|^2.$$

Desmostración. Consideremos

$$\|\mathbf{u}+\mathbf{v}\|^2=\langle\mathbf{u}+\mathbf{v},\mathbf{u}+\mathbf{v}\rangle=\langle\mathbf{u},\mathbf{u}\rangle+2\langle\mathbf{u},\mathbf{v}\rangle+\langle\mathbf{v},\mathbf{v}\rangle=\|\mathbf{u}\|^2+\|\mathbf{v}\|^2,$$

como queríamos probar. ■

5.2.1 Método de ortonormalización de Gram-Schmidt

Definition 14 *Una base $\mathcal{B}=\{u_1,\ldots,u_n\}$ se dice que es* ortogonal *si es un sistema ortogonal ($\langle u_i,u_j\rangle=0$, si $i\neq j$). Si además todos los vectores son normales ($\|u_i\|=1,\ i=1,...,n$), la base $\mathcal{B}$ se dirá* ortonormal.

Las bases canónicas de los espacios $\mathbb{R}^n$ son ejemplos de bases ortonormales (respecto de producto escalar canónico). ¿Cuál es la ventaja de disponer de este tipo de bases? Sea $\mathcal{B}=\{\mathbf{u}_1,\ldots,\mathbf{u}_n\}$ una base ortonormal de $\mathcal{V}$ y sea $\mathbf{u}\in\mathcal{V}$ un vector cualquiera con coordenadas $\alpha_1,\ldots,\alpha_n$ respecto de la base $\mathcal{B}$. Entonces

$$\langle\mathbf{u},\mathbf{u}_j\rangle=\sum_{i=1}^{n}\alpha_i\left\langle\mathbf{u}_i,\mathbf{u}_j\right\rangle=\alpha_j$$

para cada $1\le j\le n$. Relación que simplifica el cálculo de las coordenadas de los vectores en bases ortonormales.

El objetivo del método de Gram-Schmidt es obtener a partir de una base $\mathcal{B}=\{\mathbf{u}_1,\ldots,\mathbf{u}_n\}$ una base ortonormal $\mathcal{B}'=\{\mathbf{w}_1,\ldots,\mathbf{w}_n\}$. Para ello se procede en las siguientes etapas:

- Si $\mathcal{B}$ es una base ortogonal, entonces definiendo $\mathbf{w}_i = \|\mathbf{u}_i\|^{-1}\mathbf{u}_i,\ 1 \leq i \leq n$, se tiene que $\mathcal{B}'$ es una base ortonormal.

- Si $\mathcal{B}$ no es una base ortogonal el proceso es más laborioso. En primer lugar se define el vector $\mathbf{w}_1 = \|\mathbf{u}_1\|^{-1}\mathbf{u}_1$. A continuación se calcula el vector

$$\mathbf{v}_2 = \mathbf{u}_2 - \langle \mathbf{u}_2, \mathbf{w}_1 \rangle \cdot \mathbf{w}_1,$$

que es ortogonal a $\mathbf{w}_1$ dado que:

$$\langle \mathbf{v}_2, \mathbf{w}_1 \rangle = \langle \mathbf{u}_2, \mathbf{w}_1 \rangle - \langle \mathbf{u}_2, \mathbf{w}_1 \rangle \langle \mathbf{w}_1, \mathbf{w}_1 \rangle = 0,$$

y se define $\mathbf{w}_2 = \|\mathbf{v}_2\|^{-1}\mathbf{v}_2$. Recursivamente, para cada $2 \leq j \leq n$ se calcula el vector

$$\mathbf{v}_j = \mathbf{u}_j - \sum_{i=1}^{j-1} \langle \mathbf{u}_j, \mathbf{w}_i \rangle \cdot \mathbf{w}_i,$$

que es ortogonal a $\mathbf{w}_1, \ldots, \mathbf{w}_{j-1}$. Es fácil darse cuenta que cada nuevo construido $\mathbf{v}_j \neq \mathbf{0}$ ya que en otro caso $\mathbf{u}_j$ sería dependiente de $\mathbf{w}_1, ..., \mathbf{w}_{j-1}$, que a su vez son combinación lineal de $\mathbf{u}_1, ..., \mathbf{u}_{j-1}$, y esto no es posible porque $\mathcal{B}$ es una base y por tanto los vectores son LI. A continuación se toma $\mathbf{w}_j = \|\mathbf{v}_j\|^{-1}\mathbf{v}_j$. Con este proceso construimos la base ortonormal $\mathcal{B}' = \{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$.

Example 31 Sea la familia de vectores $\mathbf{u}_1 = (1,1,0)$, $\mathbf{u}_2 = (1,0,1)$, $\mathbf{u}_3 = (0,1,1)$ que constituyen una base de $\mathbb{R}^3$. Usaremos el método de Gram-Schmidt para obtener a partir de $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ una base ortonormal de $\mathbb{R}^3$ para el producto canónico. Para ello:

- Dado que $\|\mathbf{u}_1\| = \sqrt{2}$, se toma el vector $\mathbf{w}_1 = \frac{1}{\sqrt{2}}(1,1,0)$.

- Se calcula $\mathbf{v}_2 = \mathbf{u}_2 - \langle \mathbf{u}_2, \mathbf{w}_1 \rangle \mathbf{w}_1 = (\frac{1}{2}, \frac{-1}{2}, 1)$, y haciéndolo unitario tenemos $\mathbf{w}_2 = \frac{1}{\sqrt{6}}(1,-1,2)$.

- Finalmente se calcula

$$\mathbf{v}_3 = \mathbf{u}_3 - \langle \mathbf{u}_3, \mathbf{w}_1 \rangle \mathbf{w}_1 - \langle \mathbf{u}_3, \mathbf{w}_2 \rangle \mathbf{w}_2 = \frac{2}{3}(-1,1,1)$$

de donde $\mathbf{w}_3 = \frac{\sqrt{3}}{3}(-1,1,1)$.

5.2.2 Aplicación a la diagonalización ortogonal

En $\mathbb{R}^n$ con el producto escalar usual consideramos una base ortonormal $\mathcal{B} = \{\mathbf{u}_1, ..., \mathbf{u}_n\}$ y denotamos por $\mathcal{C}$ la base canónica de $\mathbb{R}^n$ que también es ortonormal. Si denotamos por $\mathbf{P} = \mathbf{M}_{\mathcal{C}\mathcal{B}}(\mathbf{i})$ la matriz de cambio de base, es inmediato darse cuenta que cada elemento del producto $\mathbf{P}^t \cdot \mathbf{P} = (a_{ij})$ viene dado por

$$a_{ij} = \langle \mathbf{u}_i, \mathbf{u}_j \rangle,\ i,j = 1, ..., n.$$

Dado que la base $\mathcal{B}$ es ortonormal se verifica que $\mathbf{P}^t \cdot \mathbf{P} = \mathbf{I}_n$, por lo que $\mathbf{P}^{-1} = \mathbf{P}^t$.

Sea ahora una matriz simétrica $\mathbf{A} \in \mathcal{M}_{n\times n}(\mathbb{R})$. Se verifica que dicha matriz es siempre diagonalizable y además la base que la diagonaliza puede construirse ortonormal, ya que si $\mathbf{u} \in \mathrm{Ker}(\mathbf{A} - \lambda \cdot \mathbf{I}_n)$

y $\mathbf{v} \in \mathrm{Ker}(\mathbf{A} - \mu \cdot \mathbf{I}_n)$ y $\lambda \neq \mu$, entonces $\langle \mathbf{v}, \mathbf{u} \rangle = 0$. Así, en el caso de matrices simétrica podemos escribir que

$$\mathbf{A} = \mathbf{P}^t \cdot \mathbf{D} \cdot \mathbf{P},$$

donde $\mathbf{D}$ es la correspondiente matriz diagonal.

Remark 5 *La demostración de este resultado sale fuera de los objetivos de este trabajo y puede verse en [**?**]. La prueba necesita una definición más amplia de la noción de producto escalar para que éste pueda tomar valores complejos. Para esto la condición (E2) de la definición de producto escalar debe cambiarse.*

Ilustremos esta aplicación con un ejemplo. Sea la matriz

$$\mathbf{A} = \begin{pmatrix} 5 & 4 & 2 \\ 4 & 5 & 2 \\ 2 & 2 & 2 \end{pmatrix},$$

que tiene como valores propios $\lambda_1 = 1$ con multiplicidad dos y $\lambda_2 = 10$. Los subespacios propios son $\mathrm{Ker}(\mathbf{A} - \mathbf{I}_3) = \{(x, y, z) \in \mathbb{R}^3 : 2x + 2y + z = 0\}$ y $\mathrm{Ker}(\mathbf{A} - 10 \cdot \mathbf{I}_3) = \{(x, y, z) \in \mathbb{R}^3 : -5x + 4y + 2z = 0$ y $4x - 5y + 2z = 0\}$, de donde obtenemos las bases $\mathcal{B}_1 = \{(1, -1, 0), (1, 0, -2)\}$ y $\mathcal{B}_{10} = \{(2, 2, 1)\}$, de donde $\mathcal{B} = \{(1, -1, 0), (1, 0, -2), (2, 2, 1)\}$ es una base de vectores propios. Aplicamos el método de Gram-Schmidt para encontrar una base ortonormal $\mathcal{B}' = \{(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}, 0), (\frac{\sqrt{2}}{6}, \frac{\sqrt{2}}{6}, -\frac{2\sqrt{2}}{3}), (\frac{2}{3}, \frac{2}{3}, \frac{1}{3})\}$ y así

$$\mathbf{A} = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{6} & \frac{2}{3} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{6} & \frac{2}{3} \\ 0 & -\frac{2\sqrt{2}}{3} & \frac{1}{3} \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 10 \end{pmatrix} \cdot \begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0 \\ \frac{\sqrt{2}}{6} & \frac{\sqrt{2}}{6} & -\frac{2\sqrt{2}}{3} \\ \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \end{pmatrix}.$$

5.3 Subespacios ortogonales

Sea $(\mathcal{V}, \langle \cdot, \cdot \rangle)$ un espacio vectorial euclídeo y consideremos un subespacio vectorial $\mathcal{W} \subset \mathcal{V}$. Se define el *subespacio ortogonal* a $\mathcal{W}$ como el subespacio

$$\mathcal{W}^\perp = \{\mathbf{v} \in \mathcal{V} : \langle \mathbf{v}, \mathbf{u} \rangle = 0 \ \forall \mathbf{u} \in \mathcal{W}\}.$$

Veamos cuáles son las propiedades de este nuevo subespacio.

Proposition 36 *$\mathcal{W}^\perp$ es un subespacio vectorial de $\mathcal{V}$.*

Demostración. Sean $\alpha, \beta \in \mathbb{R}$ y $\mathbf{u}, \mathbf{v} \in \mathcal{W}^\perp$ y veamos que su combinación lineal $\alpha \cdot \mathbf{u} + \beta \cdot \mathbf{v}$ también está en $\mathcal{W}^\perp$. Para ello tomamos $\mathbf{w} \in \mathcal{W}$ arbitrario y calculamos

$$\langle \alpha \cdot \mathbf{u} + \beta \cdot \mathbf{v}, \mathbf{w} \rangle = \alpha \langle \mathbf{u}, \mathbf{w} \rangle + \beta \langle \mathbf{v}, \mathbf{w} \rangle = 0,$$

por lo que dicha combinación lineal estará efectivamente en $\mathcal{W}^\perp$. ■

Proposition 37 *Sea $\mathcal{B} = \{\mathbf{v}_1, ..., \mathbf{v}_n\}$ una base de $\mathcal{W}$. El vector $\mathbf{u} \in \mathcal{W}^\perp$ si y sólo si $\langle \mathbf{u}, \mathbf{v}_i \rangle = 0$ para todo $i = 1, ..., n$.*

Demostración. Es claro que si $\mathbf{u} \in \mathcal{W}^{\perp}$ entonces $\langle \mathbf{u}, \mathbf{v}_i \rangle = 0$ para todo $i = 1, ..., n$. Supongamos ahora que $\langle \mathbf{u}, \mathbf{v}_i \rangle = 0$ para todo $i = 1, ..., n$ y veamos que $\mathbf{u} \in \mathcal{W}^{\perp}$. Para ello sea $\mathbf{v} \in \mathcal{W}$ arbitrario y pongámoslo como combinación lineal de los vectores de la base $\mathcal{B}$, esto es $\mathbf{v} = \alpha_1 \cdot \mathbf{v}_1 + ... + \alpha_n \cdot \mathbf{v}_n$, $\alpha_i \in \mathbb{R}$, $1 \leq i \leq n$. Calculemos

$$\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{u}, \alpha_1 \cdot \mathbf{v}_1 + ... + \alpha_n \cdot \mathbf{v}_n \rangle = \alpha_1 \langle \mathbf{u}, \mathbf{v}_1 \rangle + ... + \alpha_n \langle \mathbf{u}, \mathbf{v}_n \rangle = 0,$$

de donde vemos que $\mathbf{u} \in \mathcal{W}^{\perp}$. ■

El resultado anterior es bastante útil a la hora de hacer un cálculo práctico del subespacio ortogonal a uno dado, como vemos en el siguiente ejemplo.

Example 32 Sea $\mathbb{R}^3$ con el producto escalar usual y sea $\mathcal{W} = \{(x, y, z) \in \mathbb{R}^3 : x + y + z = 0\}$. Vamos a calcular $\mathcal{W}^{\perp}$. Para ello nos damos en primer lugar cuenta de que $\mathcal{B} = \{(1, -1, 0), (1, 0, -1)\}$ es una base de $\mathcal{W}$. Aplicamos la propiedad anterior para deducir que un vector $(x, y, z) \in \mathbb{R}^3$ está en $\mathcal{W}^{\perp}$ si se verifica que

$$\begin{aligned} \langle (x, y, z), (1, -1, 0) \rangle &= x - y = 0, \\ \langle (x, y, z), (1, 0, 1) \rangle &= x - z = 0, \end{aligned}$$

de donde $\mathcal{W}^{\perp} = \{(x, y, z) \in \mathbb{R}^3 : x = y = z\}$.

La siguiente propiedad nos dice que la suma de $\mathcal{W}$ con su ortogonal $\mathcal{W}^{\perp}$ es directa, esto es, $\mathcal{W} \oplus \mathcal{W}^{\perp}$.

Proposition 38 $\mathcal{W} \cap \mathcal{W}^{\perp} = \{\mathbf{0}\}$.

Demostración. Si $\mathbf{u} \in \mathcal{W} \cap \mathcal{W}^{\perp}$ entonces $\langle \mathbf{u}, \mathbf{u} \rangle = 0$, de donde $\mathbf{u} = \mathbf{0}$. ■

Aunque la suma sea directa, en general no siempre se verifica que $\mathcal{W} \oplus \mathcal{W}^{\perp} = \mathcal{V}$. El siguiente resultado nos garantiza que esto ocurre si la dimensión de $\mathcal{W}$ es finita.

Theorem 39 *Si $\mathcal{W}$ es de dimensión finita, entonces $\mathcal{W} \oplus \mathcal{W}^{\perp} = \mathcal{V}$.*

Demostración. Sea $\mathcal{B} = \{\mathbf{u}_1, ..., \mathbf{u}_n\}$ una base ortonormal de $\mathcal{W}$ y sea $\mathbf{v} \in \mathcal{V}$. Definimos $\mathbf{u} = \langle \mathbf{v}, \mathbf{u}_1 \rangle \cdot \mathbf{u}_1 + ... + \langle \mathbf{v}, \mathbf{u}_n \rangle \cdot \mathbf{u}_n \in \mathcal{W}$ y veamos que $\mathbf{v} - \mathbf{u} \in \mathcal{W}^{\perp}$. Para ello basta calcular para $i = 1, ..., n$

$$\begin{aligned} \langle \mathbf{v} - \mathbf{u}, \mathbf{u}_i \rangle &= \langle \mathbf{v} - \langle \mathbf{v}, \mathbf{u}_1 \rangle \cdot \mathbf{u}_1 + ... + \langle \mathbf{v}, \mathbf{u}_n \rangle \cdot \mathbf{u}_n, \mathbf{u}_i \rangle \\ &= \langle \mathbf{v}, \mathbf{u}_i \rangle - \langle \mathbf{v}, \mathbf{u}_1 \rangle \langle \mathbf{u}_1, \mathbf{u}_i \rangle - ... - \langle \mathbf{v}, \mathbf{u}_n \rangle \langle \mathbf{u}_n, \mathbf{u}_i \rangle \\ &= \langle \mathbf{v}, \mathbf{u}_i \rangle - \langle \mathbf{v}, \mathbf{u}_i \rangle = 0. \end{aligned}$$

Entonces $\mathbf{v} - \mathbf{u} \in \mathcal{W}^{\perp}$ y $\mathbf{v} = \mathbf{u} + (\mathbf{v} - \mathbf{u}) \in \mathcal{W} \oplus \mathcal{W}^{\perp}$. ■

El vector $\mathbf{u} = \langle \mathbf{v}, \mathbf{u}_1 \rangle \cdot \mathbf{u}_1 + ... + \langle \mathbf{v}, \mathbf{u}_n \rangle \cdot \mathbf{u}_n$ se conoce con el nombre de proyección ortogonal de $\mathbf{v}$ sobre $\mathcal{W}$ y la norma $||\mathbf{v} - \mathbf{u}||$ es la distancia de $\mathbf{v}$ a $\mathcal{W}$. La proyección ortogonal sobre $\mathcal{W}$ no depende de la base ortonormal $\mathcal{B} = \{\mathbf{u}_1, ..., \mathbf{u}_n\}$ de la Proposición 39. Si $\mathcal{B}' = \{\mathbf{v}_1, ..., \mathbf{v}_n\}$ es otra base ortonormal, entonces dado que $\mathbf{u} \in \mathcal{W}$ se verifica que

$$\begin{aligned}
\begin{pmatrix} \langle \mathbf{u}, \mathbf{v}_1 \rangle \\ ... \\ \langle \mathbf{u}, \mathbf{v}_n \rangle \end{pmatrix}_{\mathcal{B}'} &= \mathbf{M}_{\mathcal{B}'\mathcal{B}}(\mathbf{i}) \cdot \begin{pmatrix} \langle \mathbf{u}, \mathbf{u}_1 \rangle \\ ... \\ \langle \mathbf{u}, \mathbf{u}_n \rangle \end{pmatrix}_{\mathcal{B}} = \begin{pmatrix} \langle \mathbf{u}_1, \mathbf{v}_1 \rangle & ... & \langle \mathbf{u}_n, \mathbf{v}_1 \rangle \\ ... & ... & ... \\ \langle \mathbf{u}_1, \mathbf{v}_n \rangle & ... & \langle \mathbf{u}_n, \mathbf{v}_n \rangle \end{pmatrix} \cdot \begin{pmatrix} \langle \mathbf{v}, \mathbf{u}_1 \rangle \\ ... \\ \langle \mathbf{v}, \mathbf{u}_n \rangle \end{pmatrix} \\
&= \begin{pmatrix} \langle \mathbf{u}_1, \mathbf{v}_1 \rangle \langle \mathbf{v}, \mathbf{u}_1 \rangle + ... + \langle \mathbf{u}_n, \mathbf{v}_1 \rangle \langle \mathbf{v}, \mathbf{u}_n \rangle \\ ... \\ \langle \mathbf{u}_1, \mathbf{v}_n \rangle \langle \mathbf{v}, \mathbf{u}_1 \rangle + ... + \langle \mathbf{u}_n, \mathbf{v}_n \rangle \langle \mathbf{v}, \mathbf{u}_n \rangle \end{pmatrix} \\
&= \begin{pmatrix} \langle \mathbf{v}, \langle \mathbf{u}_1, \mathbf{v}_1 \rangle \cdot \mathbf{u}_1 ... + \langle \mathbf{u}_n, \mathbf{v}_1 \rangle \cdot \mathbf{u}_n \rangle \\ ... \\ \langle \mathbf{v}, \langle \mathbf{u}_1, \mathbf{v}_n \rangle \cdot \mathbf{u}_1 ... + \langle \mathbf{u}_n, \mathbf{v}_n \rangle \cdot \mathbf{u}_n \rangle \end{pmatrix} = \begin{pmatrix} \langle \mathbf{v}, \mathbf{v}_1 \rangle \\ ... \\ \langle \mathbf{v}, \mathbf{v}_n \rangle \end{pmatrix},
\end{aligned}$$

por lo que $\mathbf{u} = \langle \mathbf{v}, \mathbf{v}_1 \rangle \cdot \mathbf{v}_1 + ... + \langle \mathbf{v}, \mathbf{v}_n \rangle \cdot \mathbf{v}_n$. Veamos en un ejemplo cómo calcular la proyección ortogonal de un vector sobre un subespacio vectorial.

Example 33 Sea el espacio euclídeo $\mathbb{R}^3$ equipado con el producto escalar canónico y sea el subespacio vectorial

$$\mathcal{W} = \left\{ (x, y, z) \in \mathbb{R}^3 : x - y = 0, \quad x - z = 0 \right\}.$$

Dado el vector $(1, 1, 0)$ vamos a calcular su proyección ortogonal sobre $\mathcal{W}$. Evidentemente $\{(1, 1, 1)\}$ es una base de $\mathcal{W}$, por lo que $\{(\sqrt{3}/3, \sqrt{3}/3, \sqrt{3}/3)\}$ es una base ortonormal. Entonces la proyección ortogonal de $(1, 1, 0)$ sobre $\mathcal{W}$ viene dada por

$$\langle (1, 1, 0), (\sqrt{3}/3, \sqrt{3}/3, \sqrt{3}/3) \rangle (\sqrt{3}/3, \sqrt{3}/3, \sqrt{3}/3) = (2/3, 2/3, 2/3).$$

Example 34 En el ejemplo anterior, vamos a calcular la aplicación proyección ortogonal sobre $\mathcal{W}$. Para ello sea $(x, y, z) \in \mathbb{R}^3$ y definimos

$$\begin{aligned}
\mathbf{f}(x, y, z) &= \langle (x, y, z), ((\sqrt{3}/3, \sqrt{3}/3, \sqrt{3}/3)) \rangle (\sqrt{3}/3, \sqrt{3}/3, \sqrt{3}/3) \\
&= \frac{1}{3}(x + y + z, x + y + z, x + y + z).
\end{aligned}$$

Finalizamos el tema con este importante resultado que nos viene a decir que la mejor aproximación por los vectores de un subespacio vectorial de dimensión finita es la proyección ortogonal.

Theorem 40 *Sea $\mathcal{W}$ de dimensión finita. Sean $\mathbf{v} \in \mathcal{V}$ y $\mathbf{u}$ la proyección ortogonal de $\mathbf{v}$ sobre $\mathcal{W}$. Entonces*

$$||\mathbf{v} - \mathbf{u}|| = \min\{||\mathbf{v} - \mathbf{w}|| : \mathbf{w} \in \mathcal{W}\}.$$

Demostración. Sea $\mathbf{w} \in \mathcal{W}$ y calculamos

$$||\mathbf{v} - \mathbf{w}||^2 = ||(\mathbf{v} - \mathbf{u}) + (\mathbf{u} - \mathbf{w})||^2.$$

Ahora bien $\mathbf{v} - \mathbf{u} \in \mathcal{W}^\perp$ y $\mathbf{u} - \mathbf{w} \in \mathcal{W}$ por lo que $\langle \mathbf{v} - \mathbf{u}, \mathbf{u} - \mathbf{w} \rangle = 0$ y aplicando el Teorema de Pitágoras se verifica que

$$\begin{aligned}
||\mathbf{v} - \mathbf{w}||^2 &= ||(\mathbf{v} - \mathbf{u}) + (\mathbf{u} - \mathbf{w})||^2 \\
&= ||\mathbf{v} - \mathbf{u}||^2 + ||\mathbf{u} - \mathbf{w}||^2 \geq ||\mathbf{v} - \mathbf{u}||^2.
\end{aligned}$$

Además $||\mathbf{v} - \mathbf{w}|| > ||\mathbf{v} - \mathbf{u}||$ si $\mathbf{u} \neq \mathbf{w}$. ■

Example 35 Se considera el espacio euclídeo $\mathcal{V}$ de las funciones reales continuas definidas sobre $[1,2]$, con el producto escalar

$$\langle f,g\rangle := \int_1^2 f(x)g(x)dx.$$

Vamos a obtener la mejor aproximación de la función $\log x$ como un polinomio de grado menor o igual que dos. Para ello, en primer lugar obtenemos una base ortonormal del subespacio de los polinomios de grado menor o igual que dos a partir de la base canónica $\{1,x,x^2\}$. En primer lugar obtenemos una base ortogonal $\mathcal{O}=\{\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3\}$, donde $\mathbf{v}_1=1$ y

$$\mathbf{v}_2 = x - \frac{\langle x,1\rangle}{\langle 1,1\rangle}1 = x-\frac{3}{2},$$

$$\begin{aligned}\mathbf{v}_3 &= x^2 - \frac{\langle x^2,1\rangle}{\langle 1,1\rangle}1 - \frac{\langle x, x-\frac{3}{2}\rangle}{\langle x-\frac{3}{2}, x-\frac{3}{2}\rangle}\left(x-\frac{3}{2}\right)\\ &= x^2-\frac{7}{3}-\left(x-\frac{3}{2}\right) = x^2+x-\frac{5}{6}.\end{aligned}$$

Obtenemos ahora la base ortonormal $\mathcal{N}=\{\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_3\}$, donde

$$\begin{aligned}\mathbf{u}_1 &= \frac{1}{||\mathbf{v}_1||}\mathbf{v}_1 = 1,\\ \mathbf{u}_2 &= \frac{1}{||\mathbf{v}_2||}\mathbf{v}_2 = \sqrt{12}\left(x-\frac{3}{2}\right),\\ \mathbf{u}_3 &= \frac{1}{||\mathbf{v}_3||}\mathbf{v}_3 = \sqrt{\frac{180}{1861}}\left(x^2+x-\frac{5}{6}\right).\end{aligned}$$

El polinomio de grado menor o igual que dos que mejor aproxima $\log x$ será la proyección ortogonal de $\log x$ sobre el subespacio $\mathcal{W}$ dado por

$$\begin{aligned}\mathbf{p}(\log x) &= \langle \log x,\mathbf{u}_1\rangle\,\mathbf{u}_1 + \langle \log x,\mathbf{u}_2\rangle\,\mathbf{u}_2 + \langle \log x,\mathbf{u}_3\rangle\,\mathbf{u}_3\\ &= \langle \log x,1\rangle\,1 + \left\langle \log x, \sqrt{12}\left(x-\frac{3}{2}\right)\right\rangle\sqrt{12}\left(x-\frac{3}{2}\right)\\ &\quad + \left\langle \log x, \sqrt{\frac{180}{1861}}\left(x^2+x-\frac{5}{6}\right)\right\rangle\sqrt{\frac{180}{1861}}\left(x^2+x-\frac{5}{6}\right)\\ &= \langle \log x,1\rangle + 12\left\langle \log x, x-\frac{3}{2}\right\rangle\left(x-\frac{3}{2}\right) + \frac{180}{1861}\left\langle \log x, x^2+x-\frac{5}{6}\right\rangle\left(x^2+x-\frac{5}{6}\right)\\ &= 2\log 2 - 1 + 3(3-4\log 2)\left(x-\frac{3}{2}\right) + 5\frac{108\log 2-5}{1861}\left(x^2+x-\frac{5}{6}\right)\\ &= \frac{36770}{1861}\log 2 - \frac{80641}{5583} + \frac{1}{1861}(16624-21792\log 2)x + \frac{1}{1861}(540\log 2 - 125)x^2.\end{aligned}$$

El error cometido en la aproximación es

$$||\log x - \mathbf{p}(\log x)|| = \left(\int_1^2 (\log x - \mathbf{p}(\log x))^2 dx\right)^{1/2} \simeq 26.2219,$$

lo cual indica que la aproximación es mala ya que el error medio obtenido es bastante alta.

5.4 Endomorfismos con significado geométrico

En esta sección se estudian algunos tipos de aplicaciones lineales de evidente significado geométrico. Aunque algunos de los conceptos que se introducen no necesitan explícitamente la existencia de una norma o distancia, estas clases de endomorfismos toman una mayor relevancia en el marco de los espacios euclídeos.

5.4.1 Homotecias

Sea $(\mathcal{V}, \langle\cdot,\cdot\rangle)$ un espacio vectorial euclídeo. Se dice que una aplicación lineal $\mathbf{f} : \mathcal{V} \to \mathcal{V}$ es una *homotecia* de razón $\alpha \in \mathbb{R}$ si $\mathbf{f}(\mathbf{u}) = \alpha\,\mathbf{u}$ para todo $\mathbf{u} \in \mathcal{V}$. Es evidente que respecto de cualquier base $\mathcal{B}$ de $\mathcal{V}$ la matriz asociada a una homotecia es de la forma

$$\mathbf{M}_{\mathcal{B}\mathcal{B}}(\mathbf{f}) = \begin{pmatrix} \alpha & 0 & \cdots & 0 \\ 0 & \alpha & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \alpha \end{pmatrix} = \alpha\,\mathbf{I}_n$$

siendo $\dim \mathcal{V} = n$.

El efecto de una homotecia sobre cualquier subconjunto de $\mathcal{V}$ (o figura en el lenguaje del dibujo técnico) es el de contraerlo (si $|\alpha| < 1$) o expandirlo (si $|\alpha| > 1$), manteniendo su orientación (si $\alpha > 0$) o invirtiéndola (si $\alpha < 0$). Esto es, la imagen de cualquier conjunto contenido en $\mathcal{V}$ es el mismo conjunto aumentado de tamaño o empequeñecido y con la misma o diferente orientación.

5.4.2 Proyecciones

Sea $(\mathcal{V}, \langle\cdot,\cdot\rangle)$ un espacio vectorial euclídeo y sean $\mathcal{W}_1$ y $\mathcal{W}_2$ dos subespacios vectoriales de $\mathcal{V}$ de forma que $\mathcal{W}_1 \oplus \mathcal{W}_2 = \mathcal{V}$. Se dice que un endomorfismo $\mathbf{f} : \mathcal{V} \to \mathcal{V}$ es una *proyección* de base $\mathcal{W}_1$ y dirección $\mathcal{W}_2$ si se verifica que $\mathbf{f}(\mathbf{u}) = \mathbf{0}$ para cada $\mathbf{u} \in \mathcal{W}_2$ (es decir, $\mathrm{Ker}(\mathbf{f}) = \mathcal{W}_2$) y $\mathbf{f}(\mathbf{u}) = \mathbf{u}$ para todo $\mathbf{u} \in \mathcal{W}_1$. Si $\mathcal{W}_1^\perp = \mathcal{W}_2$ la proyección es la ortogonal, estudiada en el apartado anterior. Geométricamente, una proyección lleva una figura u objeto de $\mathcal{V}$ a una figura diferente contenida en la base de la proyección. Las proyecciones son bastante usadas en la asignatura de dibujo técnico de las carreras de Ingeniería.

5.4.3 Simetrías

El concepto de simetría se define de manera análoga al de proyección. Sea $(\mathcal{V}, \langle\cdot,\cdot\rangle)$ un espacio vectorial euclídeo y sean $\mathcal{W}_1$ y $\mathcal{W}_2$ dos subespacios vectoriales, de forma que $\mathcal{W}_1 \oplus \mathcal{W}_2 = \mathcal{V}$. Se dice que una aplicación lineal $\mathbf{f} : \mathcal{V} \to \mathcal{V}$ es una *simetría* de base $\mathcal{W}_1$ y dirección $\mathcal{W}_2$ si se verifica que $\mathbf{f}(\mathbf{u}) = -\mathbf{u}$ para todo $\mathbf{u} \in \mathcal{W}_2$ y $\mathbf{f}(\mathbf{u}) = \mathbf{u}$ para todo $\mathbf{u} \in \mathcal{W}_1$. Si $\mathcal{W}_1^\perp = \mathcal{W}_2$, la simetría se denomina *ortogonal* y únicamente hará falta indicar cuál es su base o su dirección. Es inmediato comprobar que la matriz asociada a una simetría es siempre diagonalizable y sus valores propios son -1 y 1. Un ejemplo clásico de simetría en $\mathbb{R}^2$ es la imagen reflejada en un espejo de una cierta figura. Cualquier objeto de $\mathcal{V}$ pasa a ser, mediante una simetría, otro objeto de $\mathcal{V}$ del mismo tamaño, pero en una posición diferente, marcada por la dirección y base de la simetría.

5.4.4 Rotaciones en el plano

El último tipo que transformaciones que vamos a estudiar son las rotaciones en un plano. Dado el espacio euclídeo $\mathbb{R}^2$ con el producto escalar usual, se dice que $\mathbf{f} : \mathbb{R}^2 \to \mathbb{R}^2$ es una *rotación* de ángulo θ si su matriz asociada respecto de alguna base $\mathcal{B}$ es de la forma

$$M_{\mathcal{B}\mathcal{B}}(\mathbf{f}) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Geométricamente $\mathbf{f}$ produce un giro de ángulo θ de cualquier vector de $\mathbb{R}^2$. Por ejemplo, dado un reloj, el pasar de una hora a otra posterior se hace mediante un giro de las manecillas.

5.5 Ejercicios

1. Sea el espacio vectorial $\mathbb{R}^3$ con el producto escalar estándar. Calcular:

   (a) El producto escalar de los vectores $(1, 3, -1)$ y $(1, -1, 1)$ así como el ángulo que forman.

   (b) Calcular el valor de α para que el vector $(\alpha, 1, 0)$ sea normal o unitario.

   (c) Calcular el valor de α para que los vectores $(\alpha, 1, 0)$ y $(\alpha, -1, -1)$ sean ortogonales.

2. Sea $\mathbb{R}^3$ el espacio vectorial de dimensión 3 dotado del producto escalar usual de $\mathbb{R}^3$. Sea la base canónica $B = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$, se pide:

   (a) Calcular α para que el ángulo formado por los vectores $\mathbf{u} = \mathbf{e}_1 + \alpha\mathbf{e}_2 + \mathbf{e}_3$ y $\mathbf{v} = \mathbf{e}_1 + 2\mathbf{e}_2$ sea de $\dfrac{\pi}{3}$ radianes.

   (b) Calcular α para que el módulo del vector $\mathbf{u} = \mathbf{e}_1 + \alpha\mathbf{e}_2 + \mathbf{e}_3$ sea 49.

   (c) Calcular todos los vectores que están a una distancia euclídea igual a 3 del vector $\mathbf{u} = 2\mathbf{e}_1 - \mathbf{e}_2$.

   (d) Calcular α para que los vectores $\mathbf{u} = \alpha\mathbf{e}_1 + (\alpha - 1)\mathbf{e}_2 + \alpha\mathbf{e}_3$ y $\mathbf{v} = 2\alpha\mathbf{e}_1 + \alpha\mathbf{e}_2 - 3\mathbf{e}_3$ sean ortogonales.

3. Demostrar

   (a) *Teorema de Pitágoras.* Si $\mathbf{u}$ y $\mathbf{v}$ son dos vectores ortogonales del espacio euclídeo $\mathbb{R}^n$. Entonces
   $$\| \mathbf{u} + \mathbf{v} \|^2 = \| \mathbf{u} \|^2 + \| \mathbf{v} \|^2 .$$

   (b) *Ley del Paralelogramo.* Si $\mathbf{u}$ y $\mathbf{v}$ son dos vectores cualesquiera del espacio euclídeo $\mathbb{R}^n$, entonces
   $$\| \mathbf{u} + \mathbf{v} \|^2 + \| \mathbf{u} - \mathbf{v} \|^2 = 2 \| \mathbf{u} \|^2 + 2 \| \mathbf{v} \|^2 .$$

4. Probar que si $\{\mathbf{u}, \mathbf{v}\}$ son vectores ortogonales de un espacio vectorial euclídeo, entonces forman un sistema linealmente independiente. ¿Es cierto el recíproco?

5. Obtener en el espacio euclídeo $(\mathbb{R}^3, \langle\rangle)$ donde $\langle\rangle$ es el producto escalar usual de $\mathbb{R}^3$, una base ortonormal aplicando el *Método de Gram-Schmidt* a las bases

   (a) $\mathcal{B} = \{(1,1,0),\ (1,0,1),\ (0,1,1)\}$.

   (b) $\mathcal{B} = \{(1,1,1), (0,1,1), (0,0,1)\}$.

   (c) $\mathcal{B} = \{(2,1,0), (-1,1,1), (1,0,3)\}$.

6. Consideremos $\mathbb{R}^4$ con el producto escalar usual. Se pide hallar una base ortonormal de los siguientes subespacios vectoriales:

   (a) $\mathcal{W} = \{(x,y,z,t) \in \mathbb{R}^4 : x+y+z+t = 0\}$.

   (b) $\mathcal{W} = \{(x,y,z,t) \in \mathbb{R}^4 : x+y = 0;\ z+t = 0\}$.

   (c) $\mathcal{W} = \{(x,y,z,t) \in \mathbb{R}^4 : x = 0;\ z = 0\}$.

   (d) $\mathcal{W} = \{(x,y,z,t) \in \mathbb{R}^4 : x = 0\}$.

7. Obtener la forma diagonal de las siguientes matrices simétricas, así como las matrices de cambio de base

$$\text{(a)} \begin{pmatrix} 5 & 4 & 2 \\ 4 & 5 & 2 \\ 2 & 2 & 2 \end{pmatrix} \quad \text{(b)} \begin{pmatrix} 1 & 2 & 5 \\ 2 & -2 & -2 \\ 5 & -2 & 1 \end{pmatrix} \quad \text{(c)} \begin{pmatrix} 1 & 5 \\ 5 & 1 \end{pmatrix}$$

$$\text{(d)} \begin{pmatrix} 2 & -4 & 0 \\ -4 & 0 & 0 \\ 0 & 0 & 6 \end{pmatrix} \quad \text{(e)} \begin{pmatrix} 2 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 2 \end{pmatrix}$$

8. Sea $\mathcal{P}_2[x]$ el conjuto de polinomios de grado menor o igual que 2 con coeficientes reales. Dados $\mathbf{p}(x), \mathbf{q}(x) \in \mathcal{P}_2[x]$, se define

$$< \mathbf{p}(x), \mathbf{q}(x) >= \int_{-1}^{1} x^4 \mathbf{p}(x)\mathbf{q}(x)dx.$$

   Probar que $\langle *, * \rangle$ es un producto escalar. Dada la base $\mathcal{B} = \{1, x, x^2\}$, obtener a partir de ella un base ortonormal de $\mathcal{P}_2[x]$ (usando el producto escalar anterior).

9. Sea $\mathcal{P}_3[x]$ el conjunto de los polinomios con coeficientes reales de grado a lo sumo tres. Definimos para todo $\mathbf{p}(x), \mathbf{q}(x) \in \mathcal{P}_3[x]$,

$$\langle \mathbf{p}(x), \mathbf{q}(x) \rangle := \int_0^1 \mathbf{p}(x)\mathbf{q}(x)dx.$$

   (a) Comprobar que se trata de un producto escalar.

   (b) Obtener una base ortonormal a partir de la base $\mathcal{B} = \{1, x, x^2, x^3\}$.

   (c) Calcular el valor del parámetro a para que los polinomios $ax^3 + x^2 + 1$ y $x + 1$ sean ortogonales.

(d) Calcular el valor de a para que $ax^2 + 1$ sea normal o unitario.

10. Dado el plano real $\mathbb{R}^2$ y el producto escalar usual calcular la *proyección ortogonal del vector* $(1,1)$ sobre los subespacios generados por los siguientes vectores:

$$\text{(a)}\ \left(\frac{1}{3}, \frac{2}{5}\right). \quad \text{(b)}\ (-1,-1)\,. \quad \text{(c)}\ (1,0)\,.$$

11. Sea el espacio vectorial $\mathbb{R}^3$ sobre el que tenemos definido el producto escalar usual. Dados los siguientes subespacios $\mathcal{S}$ de $\mathbb{R}^3$, calcular $\mathcal{S}^\perp$:

    (a) $\mathcal{S} = \{(x,y,z) \in \mathbb{R}^3 : x + y + z = 0\}\,.$

    (b) $\mathcal{S} = \{(x,y,z) \in \mathbb{R}^3 : x + y + z = 0 \text{ y } x - y = 0\}\,.$

    (c) $\mathcal{S} = \{(x,y,z) \in \mathbb{R}^3 : x - y + 2z = 0 \text{ y } x - z = 0\}\,.$

12. Calcular las ecuaciones de las siguientes aplicaciones lineales $\mathbf{f} : \mathbb{R}^3 \to \mathbb{R}^3$:

    (a) La proyección ortogonal de base $\{(x,y,z) : x = y = z\}$.

    (b) La proyección ortogonal con base $\{(x,y,z) : x = y\}$.

    (c) La proyección ortogonal cuya base es el subespacio ortogonal a $\{(x,y,z) : x = y = z\}$.

13. Obtener el núcleo, la imagen y los subespacios propios de las aplicaciones lineales del ejercicio anterior. ¿Qué conclusiones pueden obtenerse?

14. Hallar el polinomio de segundo grado que mejor aproxima la función $f(x) = \sqrt[3]{x}$ en el intervalo $[-1,1]$ con la norma asociada al producto escalar de las funciones reales continuas definidas en $[-1,1]$ dado por

$$\langle \mathbf{f}, \mathbf{g} \rangle = \int_{-1}^{1} \mathbf{f}(x)\mathbf{g}(x)dx.$$

15. Se considera el espacio euclídeo $\mathcal{V}$ de las funciones reales continuas definidas sobre $[1,2]$, con el producto escalar $\langle \mathbf{f}, \mathbf{g} \rangle := \int_1^2 \mathbf{f}(x)\mathbf{g}(x)dx$.

    (a) Hallar el ángulo entre $\mathbf{f}(x) = 1$ y $\mathbf{g}(x) = x$.

    (b) ¿Para qué valores de a son ortogonales los vectores $x - a$ y $x + a$?

    (c) Sea $\mathcal{W}$ el subespacio de los polinomios reales de grado menor o igual que 2. Ortonormalizar la base de dicho subespacio $\{1, x, x^2\}$.

    (d) ¿Cuál es el polinomio de grado menor o igual que 2 que mejor aproxima la función $\mathbf{f}(x) = \log x$. (**Ayuda**: tener en cuenta que $\int_1^2 x^n \log x dx = \frac{2^{n+1}}{n+1}(\log 2 - \frac{1}{n+1}) - \frac{1}{(n+1)^2}$ para todo $n \geq 0$.)

16. Calcula la proyección ortogonal del vector $\mathbf{u} = (1,1,-1)$ sobre $\mathcal{W} = \{(x,y,z) \in \mathbb{R}^3 : x - y - 2z = 0\}$.

17. Encuentra la expresión de la proyección ortogonal sobre la recta de $\mathbb{R}^3$ generada por el vector $(0,1,2)$.

18. Halla la distancia entre el vector $(1,0,2)$ y el plano $x-y-z=0$.

19. Calcula la distancia entre el punto $(2,2,2)$ y el plano $x-y-z=1$.

20. Dados un espacio vectorial eucídeo y un subespacio vectorial del mismo $\mathcal{W}$, se define la simetría ortogonal de base $\mathcal{W}$ como la aplicación lineal $\mathbf{f}:\mathcal{V}\to\mathcal{V}$ such that $\mathbf{f}(\mathbf{u})=\mathbf{u}$ for $\mathbf{u}\in\mathcal{W}$ and $\mathbf{f}(\mathbf{u})=-\mathbf{u}$ for $\mathbf{u}\in\mathcal{W}^{\perp}$. El subespacio $\mathcal{W}^{\perp}$ se dirá dirección de la simetría. Se pide obtener las expresiones analíticas de las siguientes simetrías ortogonales:

    (a) $\mathbf{f}:\mathbb{R}^3\to\mathbb{R}^3$ con base $\mathcal{W}=\{(x,y,z):x+y-z=0\}$.
    (b) $\mathbf{f}:\mathbb{R}^4\to\mathbb{R}^4$ con base $\mathcal{W}=\{(x,y,z,t):x+y-z=0,\ x=t\}$.
    (c) $\mathbf{f}:\mathbb{R}^3\to\mathbb{R}^3$ con dirección $\mathcal{W}=\{(x,y,z):-x+y+2z=0\}$.
    (d) $\mathbf{f}:\mathbb{R}^4\to\mathbb{R}^4$ con base $\mathcal{W}=\{(x,y,z):x+y-z=0,\ x-y+t=0,\ 2x-z+t\}$.

21. Dado un espacio vectorial eucídeo se define la homotecia de razón $\alpha\in\mathbb{R}$ como la aplicación lineal $\mathbf{f}:\mathcal{V}\to\mathcal{V}$ such that $\mathbf{f}(\mathbf{u})=\alpha\cdot\mathbf{u}$ for $\mathbf{u}\in\mathcal{V}$. Se pide obtener las expresiones analíticas de las siguientes homotecias ortogonales:

    (a) $\mathbf{f}:\mathbb{R}^3\to\mathbb{R}^3$ con razón $1/2$.
    (b) $\mathbf{f}:\mathbb{R}^4\to\mathbb{R}^4$ con razón -1.
    (c) $\mathbf{f}:\mathbb{R}^3\to\mathbb{R}^3$ con razón 2.

22. Obtener los valores propios y determinar si son diagonalizables las matrices respecto de las bases canónicas de las aplicaciones lineales de los ejercicios 20 y 21.

23. Obtener $\mathbf{f}\circ\mathbf{g}\circ\mathbf{h}$ con $\mathbf{f},\mathbf{g},\mathbf{h}:\mathbb{R}^3\to\mathbb{R}^3$ aplicaciones lineales, donde $\mathbf{f}$ es la proyección ortogonal de base $\mathcal{W}=\{(x,y,z):x+y+z=0\}$, $\mathbf{g}$ es la homotecia de razón -1 y $\mathbf{h}$ es la simetría ortogonal de dirección $\mathcal{S}=\{(x,y,z):z=y=z\}$.

Bibliografía

[IzTo] J. Izquierdo y J. R. Torregrosa, *Algebra y ecuaciones diferenciales,* Servicio de publicaciones, Universidad Politéctica de Valencia, 1991.

[Jef] A. Jeffrey, *Linear algebra and ordinary differential equations,* CRC Press, 1993.

[ToJo] J. R. Torregrosa y C. Jordan, *Algebra lineal y sus aplicaciones,* Schaum McGraw–Hill, 1987.