ALGEBRA LINEAL. CURSO 2010-11

ALGEBRA LINEAL. CURSO 2010-11

Date: 24 de abril de 2011.

Índice

1. Espacios vectoriales, subespacios y espacios cociente

1.1. Grupos abelianos y Cuerpos.

1.1.1. **Grupos abelianos.** Las propiedades de la operación de suma de vectores de $\mathbb{R}^n$ se pueden expresar en forma axiomática, dando lugar al concepto de grupo abeliano o grupo conmutativo.

Definición 1.1.1. Un *grupo conmutativo* es un conjunto E junto con

1. un elemento distinguido (llamado *elemento neutro, cero* o *unidad*) $0 \in E$,
2. una aplicación (llamada *operación*, *suma*, o *producto*)
$$+ : E \times E \longrightarrow E, \quad (x, y) \mapsto x + y,$$
3. una aplicación (llamada *opuesto* o *inverso*)
$$-(?) : E \longrightarrow E, \quad x \mapsto -x,$$

que cumplen los 3 axiomas de grupo:

1. (*Axioma del neutro*) $x + 0 = x = 0 + x$, para todo $x \in E$,
2. (*Axioma de asociatividad*) $(x + y) + z = x + (y + z)$, para todo $x, y, z \in E$,
3. (*Axioma del opuesto o inverso*) $x + (-x) = 0 = (-x) + x$, para todo $x \in E$.

y el axioma de conmutatividad

(4) (*Axioma de conmutatividad*) $x + y = y + x$, para todo $x, y \in E$,

Para definir una estructura de grupo abeliano sobre un conjunto E es suficiente con definir la operación $+$ y probar la existencia del 0 y del opuesto $-(?)$, que, si existen, están determinados de modo único por la operación $+$, como se prueba en el siguiente ejercicio.

Ejercicio 1.1.2. Sea $(E, 0, +, -(?))$ un grupo conmutativo.

1. Supongamos que $e \in E$ cumple $e + x = x$, para todo $x \in E$, entonces $e = 0$.
2. Supongamos que $x + y = 0$, entonces $y = -x$ y $x = -y$.
3. Prueba que $-0 = 0$ y $-(-x) = x$.

Notación 1.1.3. Denotaremos por $\mathbb{N}$ el conjunto de los números naturales $\{0, 1, 2, \ldots\}$, por $\mathbb{Z}$ el conjunto de los números enteros $\{0, \pm 1, \pm 2, \cdots\}$, por $\mathbb{Q}$ el conjunto de los números racionales $\{\frac{p}{q}; p, q \in \mathbb{Z}, q \neq 0\}$, por $\mathbb{R}$ el conjunto de los números reales, y por $\mathbb{C}$ el conjunto de los números complejos, $\{a + bi; a, b \in \mathbb{R}^2\}$.

Ejemplo 1.1.4.
1. El conjunto $E = \{0\}$ tiene una única estructura de grupo conmutativo.
2. El conjunto $E = \{0, 1\}$ tiene una única estructura de grupo conmutativo, en la cual 0 es el elemento neutro. Esta estructura está definida por la suma $1 + 1 = 0$.
3. Los datos $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$ y $(\mathbb{C}, +)$ son ejemplos de grupos conmutativos. En este caso solemos decir que $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ y $\mathbb{C}$ son grupos conmutativos con la suma.
4. El conjunto $(\mathbb{N}, +)$ no define una estructura de grupo puesto que el opuesto de un número natural $n > 0$ no es un número natural (no existe en $\mathbb{N}$).
5. El dato $(\mathbb{R}, +)$ es un grupo conmutativo.
6. Sea $\mathbb{R}[x]$ el conjunto de polinomios en una indeterminada x con coeficientes en $\mathbb{R}$. Entonces $\mathbb{R}[x]$ con la suma es un grupo conmutativo.

7. Consideremos ahora en $E = \mathbb{Q} \setminus \{0\}$, el elemento distinguido 1, la multiplicación $\cdot$, y para cada $x \in E$, su inverso $\frac{1}{x}$. Entonces E con estos datos también es un grupo conmutativo, es decir,
   a) (neutro) $x \cdot 1 = x = 1 \cdot x$, para todo $x \in E$,
   b) (asociativa) $(x \cdot y) \cdot z = x \cdot (y \cdot z)$, para todo $x, y, z \in E$, y
   c) (opuesto) $x \cdot \frac{1}{x} = 1 = \frac{1}{x} \cdot x$, para todo $x \in E$.
   d) (conmutativa) $x \cdot y = y \cdot x$, para todo $x, y \in E$,
8. De la misma manera que en el ejemplo anterior, $(\mathbb{R} \setminus \{0\}, 1, \cdot, \frac{1}{?})$ y $(\mathbb{C} \setminus \{0\}, 1, \cdot, \frac{1}{?})$ son también grupos abelianos. Por el contrario $(\mathbb{Z} \setminus \{0\}, 1, \cdot)$ no soporta una estructutra de grupo porque el inverso $\frac{1}{n}$ de $n \neq \pm 1$ no es un número entero.

En un grupo conmutativo se puede sumar una familia finita de elementos:

Definición 1.1.5. Sean E un grupo abeliano, $n \geq 1$ un entero y $\{x_i\}_{1 \leq i \leq n}$ una familia finita de elementos de E. Se define su suma $x_1 + ... + x_n$, según la fórmula recurrente

$$x_1 + x_2 + ... + x_n := (x_1 + ... + x_{n-1}) + x_n.$$

Si $n = 0$, se conviene que la suma es 0.

Por la propiedad asociativa, esta suma no depende de la forma en que se dispongan los paréntesis. Puesto que, además, la suma es conmutativa, la suma de la familia tampoco depende del orden en que se dispongan los sumandos x_i.

Ejemplo 1.1.6. Consideremos los números enteros del 1 al 5. Su suma es por definición

$$1 + 2 + 3 + 4 + 5 = (1 + 2 + 3 + 4) + 5 = \cdots = ((((1 + 2) + 3) + 4) + 5),$$

pero obviamente coincide con cualquier otra reordenación de los paréntesis, por ejemplo

$$1 + 2 + 3 + 4 + 5 = (1 + 2) + (3 + (4 + 5)).$$

En este caso, puesto que además la operación es conmutativa, también podríamos cambiar el orden de los sumandos y el resultado sería el mismo.

1.1.2. **Cuerpos.** Un cuerpo es un conjunto en el que tienen sentido las cuatro operaciones de la aritmética, y éstas satisfacen las propiedades habituales que encontramos en $\mathbb{Q}$ o en $\mathbb{R}$.

Definición 1.1.7. Un *cuerpo* es un conjunto K, junto con

1. dos elementos distinguidos $0 \in K$, $1 \in K \setminus \{0\}$;
2. dos operaciones (suma y producto)
$$+ : K \times K \longrightarrow K,\ (x, y) \mapsto x + y,$$
$$\cdot : K \times K \longrightarrow K,\ (x, y) \mapsto x \cdot y,$$
3. las aplicaciones opuesto e inverso
$$-? : K \longrightarrow K,\ x \mapsto -x,$$
$$\frac{1}{?} : K \setminus \{0\} \longrightarrow K \setminus \{0\},\ x \mapsto \frac{1}{x},$$

tales que

1. $(K, 0, +, -?)$ es un grupo conmutativo, es decir,
   a) $x + 0 = x = 0 + x$, $\forall x \in K$,
   b) $(x + y) + z = x + (y + z)$, $\forall x, y, z \in K$,
   c) $X + (-x) = 0 = (-x) + x$, $\forall x \in K$,
   d) $x + y = y + x$, $\forall x, y \in K$,

2. $(K \setminus \{0\}, 1, \cdot, \frac{1}{?})$ es un grupo conmutativo, es decir
   a) $x \cdot 1 = x = 1 \cdot x$, $\forall x \in K \setminus \{0\}$,
   b) $(x \cdot y) \cdot z = x \cdot (y \cdot z)$, $\forall x, y, z \in K \setminus \{0\}$,
   c) $x \cdot \frac{1}{x} = 1 = \frac{1}{x} \cdot x$, $\forall x \in K \setminus \{0\}$,
   d) $x \cdot y = y \cdot x$, $\forall x, y \in K \setminus \{0\}$.
3. (*Axioma de distributividad por la izquierda*) $(x + y) \cdot z = x \cdot z + y \cdot z, \quad \forall\, x, y, z \in K$.
4. (*Axioma de distributividad por la derecha*) $x \cdot (y + z) = x \cdot y + x \cdot z, \quad \forall\, x, y, z \in K$.

Los elementos de un cuerpo se llaman *escalares*.

Al igual que en el caso de grupos, usualmente solo mencionaremos las operaciones de suma y de producto, dando por sobreentendido cuales son los elementos neutros $0, 1$ y las operaciones de opuesto e inverso.

Proposición 1.1.8. Sea K un cuerpo. Entonces

$$0 \cdot \lambda = 0 = \lambda \cdot 0, \quad (-1) \cdot \lambda = -\lambda = \lambda \cdot (-1),$$

para todo $\lambda \in K$, y

$$\frac{1}{-\lambda} = -\frac{1}{\lambda}$$

si $\lambda \in K \setminus \{0\}$.

Ejemplo 1.1.9.
1. El conjunto $K = \{0, 1\}$ es el cuerpo más pequeño que existe. La suma está caracterizada por $1 + 1 = 0$. Los axiomas de cuerpo determinan el resto de operaciones y se cumple $0 + 0 = 0, 0 + 1 = 1 + 0 = 1, 1 \cdot 0 = 0 = 0 \cdot 1, 1 \cdot 1 = 1,$.
2. Los conjuntos $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, y $\mathbb{Z}/(p)$, con $p \in \mathbb{Z}$ primo, con la suma y el producto, son ejemplos de cuerpos.
3. El conjunto $\mathbb{R}[x]$, con la suma y el producto, no es un cuerpo, porque el inverso de un polinomio no nulo en general no es un polinomio.
4. El conjunto de fracciones racionales
   $$\mathbb{R}(x) = \left\{ \frac{P(x)}{Q(x)};\ P(x), Q(x) \in \mathbb{R}[x],\ Q(x) \neq 0 \right\},$$
   con la suma y el producto, es un cuerpo.

Notación 1.1.10. En lo que sigue K denota un cuerpo arbitrario.

1.2. Espacios y subespacios.

1.2.1. **Espacio vectorial.** La noción de espacio vectorial sobre $\mathbb{R}$ resume, en forma abstracta, las propiedades básicas de los espacios $\mathbb{R}^n$ dotados de la suma y del producto por escalares, de modo que otros conjuntos, como los de matrices o de polinomios, también son espacios vectoriales. Las operaciones básicas son la suma de vectores, que da como resultado un vector, y el producto de un escalar por un vector, que da como resultado un vector.

Igualmente, la consideración de un cuerpo base arbitrario en lugar de $\mathbb{R}$ permite el estudio de $\mathbb{Q}^n$, $\mathbb{C}^n$, y los espacios de polinomios o de matrices correspondientes, como casos particulares de la teoría general de espacios vectoriales.

Así, por una parte extendemos los coeficientes de $\mathbb{R}$ a un cuerpo K, y las operaciones con vectores de $\mathbb{R}^n$ como espacio vectorial sobre $\mathbb{R}$ a las de un espacio vectorial E sobre K.

Definición 1.2.1. Sea K un cuerpo. Un *espacio vectorial* sobre K, o K-espacio vectorial, es un grupo conmutativo $(E, 0, +, -?)$, cuyos elementos se llaman *vectores*, y una aplicación $\cdot : K \times E \longrightarrow E$ llamada producto por escalares, verificando:

1. (*Axioma de unidad*) $1 \cdot u = u$, para todo $u \in E$.
2. (*Axioma de asociatividad*) $(\lambda \cdot \mu) \cdot u = \lambda \cdot (\mu \cdot u)$, $\forall\, \lambda, \mu \in K, u \in E$.
3. (*Axioma de distributividad por la izquierda*) $(\lambda+\mu)\cdot u = \lambda\cdot u+\mu\cdot u$, $\forall\, \lambda, \mu \in K, u \in E$.
4. (*Axioma de distributividad por la derecha*) $\lambda\cdot(u+v) = \lambda\cdot u+\lambda\cdot v$, $\forall\lambda \in K, u, v \in E$.

Siguiendo la linea establecida para los grupos y los cuerpos, para dotar a un conjunto de una estructura de espacio vectorial daremos únicamente las operaciones de suma y de producto por escalares.

A continuación proponemos la prueba de unas propiedades elementales a partir únicamente de los axiomas.

Proposición 1.2.2. Sea E un espacio vectorial sobre K.

1. Para todo $\lambda \in K$ y para todo $u \in E$,
$$\lambda \cdot \mathbf{0} = \mathbf{0}, \quad 0 \cdot u = \mathbf{0}, \quad (-1) \cdot u = -u.$$
2. Si $\lambda \in K$ y $u \in E$ son no nulos, entonces $\lambda \cdot u \neq \mathbf{0}$.

Ejemplo 1.2.3. Los ejemplos básicos de espacios vectoriales son los siguientes.

1. El conjunto $E = \{0\}$ es un espacio vectorial sobre cualquier cuerpo K.
2. El espacio vectorial de n coordenadas $\mathbb{R}^n$, $n \geq 1$, es un espacio vectorial sobre $\mathbb{R}$. Más generalmente, para todo cuerpo K, y todo natural $n \geq 1$, el conjunto
$$K^n = \{(x_1, \cdots, x_n);\ x_1, \cdots, x_n \in K\}$$
es un espacio vectorial sobre K, con las operaciones de suma y producto por escalares realizadas coordenada a coordenada:
$$(x_1, \cdots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \cdots, x_n + y_n), \quad \forall (x_1, \cdots, x_n), (y_1, \ldots, y_n) \in K^n,$$
$$\lambda \cdot (x_1, \cdots, x_n) = (\lambda \cdot x_1, \cdots, \lambda \cdot x_n), \quad \forall \lambda \in K, \forall (x_1, \cdots, x_n) \in K^n$$
Mencionemos que el vector $\mathbf{0}$ es $\mathbf{0} := (0, 0, \ldots, 0)$ y que el opuesto de $(x_1, \cdots, x_n)$ es $-(x_1, \cdots, x_n) := (-x_1, \cdots, -x_n)$.
3. El conjunto $\mathbb{R}(n, m)$ de matrices de tipo $n \times m$ con coeficientes en $\mathbb{R}$ es un espacio vectorial sobre $\mathbb{R}$. Una matriz $A = (a_i^j)_{ij} \in \mathbb{R}(n, m)$ tiene $n \times m$ coordenadas, las entradas a_i^j.
   En general el conjunto $K(n, m)$ de matrices de tipo $n\times m$ con coeficientes en un cuerpo K, junto con las operaciones correspondientes realizadas coordenada a coordenada, es un espacio vectorial sobre K. Así por ejemplo podemos hablar de matrices con entradas en $\mathbb{C}$, o con entradas en $\mathbb{R}(x)$, etc.
   Usualmente se identifica un vector $(x_1, \ldots, x_n) \in K^n$ con la matriz columna de sus componentes
$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$
Por esta razón en lo que sigue identificaremos K^n con el espacio de matrices de n filas y una columna $K(n, 1)$.

Otros ejemplos:

Ejemplo 1.2.4. El conjunto $\mathbb{R}[x]$ es un espacio vectorial sobre $\mathbb{R}$, en el cual las operaciones se realizan coordenada a coordenada. Las “coordenadas“ de un polinomio $\sum\limits_{0\leq i\leq n} a_i x^i$ son sus coeficientes $(a_0, a_1, \ldots, a_n)$.

Análogamente, el conjunto $K[x]$ de polinomios en una indeterminada x con coeficientes en K, es un espacio vectorial sobre K.

Ejemplo 1.2.5. Sea $\mathcal{F}(\mathbb{N}, \mathbb{R})$, o abreviadamente $\mathcal{F}(\mathbb{N})$, el conjunto de sucesiones de números reales. Las operaciones se realizan coordenada a coordenada. Las "coordenadas" de una sucesión $(a_n)_{n\in\mathbb{N}}$ son sus términos a_n, por tanto una sucesión tiene infinitas coordenadas posiblemente distintas. Con estas operaciones $\mathcal{F}(\mathbb{N})$ es un $\mathbb{R}$-espacio vectorial.

Ejemplo 1.2.6. Sea X un conjunto, por ejemplo un intervalo de $\mathbb{R}$. Denotamos por $\mathcal{F}(X, \mathbb{R})$, o simplemente por $\mathcal{F}(X)$ si se sobreentiende $\mathbb{R}$, el conjunto de funciones de X en $\mathbb{R}$. Dotamos a $\mathcal{F}(X)$ de las siguientes operaciones:

Si $f, g \in \mathcal{F}(X)$, $\lambda \in \mathbb{R}$, se definen la suma de f y g, $f + g$, el producto de λ por f, $\lambda \cdot f$, la función opuesta de f, $-f$, y la función cero, $\mathbf{0}$, por

$$\begin{aligned} \mathbf{0}(x) &:= 0, \\ (-f)(x) &:= -(f(x)), \\ (f+g)(x) &:= f(x) + g(x), \\ (\lambda \cdot f)(x) &:= \lambda \cdot (f(x)), \end{aligned}$$

para todo $x \in X$. Con estas operaciones, $\mathcal{F}(X)$ es un $\mathbb{R}$-espacio vectorial.

Este ejemplo generaliza el espacio vectorial $\mathbb{R}^n$ (tomando como X el conjunto $\{1, 2, \cdots, n\}$), los espacios de matrices $\mathbb{R}(n, m)$ (tomando como I el conjunto $\{1, 2, \cdots, n\} \times \{1, 2, \cdots, m\}$), y el espacio de sucesiones $\mathcal{F}(\mathbb{N})$ (tomando como X el conjunto $\mathbb{N}$).

Ejemplo 1.2.7. El ejemplo anterior es igualmente válido cambiando $\mathbb{R}$ por un cuerpo cualquiera K, por ejemplo, los cuerpos $\mathbb{Q}$, $\mathbb{C}$, $\mathbb{Z}/(p)$ o $\mathbb{R}(x)$.

En conclusión, podemos decir que un polinomio, una matriz, una sucesión o una función también son vectores. Por ejemplo: la sucesión $\{\frac{1}{n}\}_{n\in\mathbb{N}}$ es un vector del $\mathbb{R}$-espacio vectorial $\mathcal{F}(\mathbb{N})$; la función $f(x) = 3e^x \sin x$ es un vector del $\mathbb{R}$-espacio vectorial $\mathcal{F}(\mathbb{R})$.

Por otra parte también las matrices A de un tamaño dado, cuyos coeficientes son fracciones racionales, son vectores. Por ejemplo, la matriz de tamaño 2×2

$$\begin{pmatrix} \frac{3x+1}{x^2+1} & \frac{x^3-1}{x^4+2} \\ \frac{1}{x^5} & \frac{x^3}{x^2-1} \end{pmatrix},$$

es un vector del espacio vectorial $\mathbb{R}(x)(2, 2)$.

1.2.2. Subespacios vectoriales.

Sea E un espacio vectorial sobre K.

Ciertos subconjuntos de E son ellos mismos espacios vectoriales. Los ejemplos básicos son las rectas y planos de $\mathbb{R}^3$ que pasan por el origen de coordenadas. En general, tales subconjuntos son los subespacios vectoriales. Para que un subconjunto F de E sea un subespacio vectorial basta que las operaciones de suma y de producto por escalares se puedan realizar en F. Más formalmente:

Definición 1.2.8. Un *subespacio vectorial* de E es un subconjunto F de E tal que

(0) $0 \in F$,
(+) $u + v \in F$, para todo $u, v \in F$, y
($\cdot$) $\lambda \cdot u \in F$, para todo $\lambda \in K$, $u \in F$.

Proposición 1.2.9. Sea F un subespacio vectorial F de un espacio vectorial E. Entonces F con el 0, la suma y el producto por escalares es un espacio vectorial sobre K.

Ejercicio 1.2.10. Sea F un subconjunto no vacío de un espacio vectorial E. Prueba que si

$$x + \lambda y \in F, \qquad \text{para todo } x, y \in F,\ \lambda \in K,$$

entonces F es un subespacio vectorial de E.

Ejemplo 1.2.11.

1. Para cualquier espacio vectorial E, $\{0\}$ es un subespacio vectorial de E, llamado el subespacio *trivial* o nulo. También se denota por $\mathbf{0}$. Nótese que el conjunto vacío $\emptyset$ no es un subespacio vectorial.
2. Para cualquier espacio vectorial E, E es un subespacio vectorial de E, llamado el subespacio *total*.

Ejemplo 1.2.12. De las propiedades de las sucesiones convergentes deducimos que el subconjuto $\mathcal{C}(\mathbb{N})$ de $\mathcal{F}(\mathbb{N})$ formado por las sucesiones convergentes es un subespacio vectorial de $\mathcal{F}(\mathbb{N})$.

Ejemplo 1.2.13.
1. El conjunto $\mathcal{C}^0(I)$ de las funciones continuas, es un subespacio vectorial de $\mathcal{F}(I)$. Para cada $n \geq 1$, $\mathcal{C}^n(I)$ el subconjunto de las funciones n veces derivables, con derivada n-ésima continua, es un subespacio vectorial de $\mathcal{F}(I)$. Se tienen inclusiones
$$\cdots \subset \mathcal{C}^{n+1}(I) \subset \mathcal{C}^n(I) \subset \ldots \mathcal{C}^0(I) \subset \mathcal{F}(I).$$
2. Para todo $n \geq 0$ las funciones polinomiales definen un subespacio vectorial de En el espacio vectorial $\mathcal{C}^n(\mathbb{R})$, que se identifica con $\mathbb{R}[x]$.

1.3. Aplicaciones lineales. Las aplicaciones lineales serán estudiadas con más detalle en el capítulo siguiente. Ahora nos limitamos a una breve introducción que nos facilitará la comprobación de que ciertos subconjuntos son subespacios vectoriales.

1.3.1. **Aplicación lineal.**

Definición 1.3.1. Sean E, F dos espacios vectoriales sobre K. Una aplicación $f : E \longrightarrow F$ se llama *aplicación lineal* si

(0) $f(0) = 0$,
(+) $f(u + v) = f(u) + f(v)$, para todo $u, v \in E$,
($\cdot$) $f(\lambda \cdot u) = \lambda \cdot f(u)$, para todo $u \in E, \lambda \in K$.

Proposición 1.3.2. Sea $f : E \longrightarrow F$ una aplicación entre dos K-espacios vectoriales.Si

$$f(x + \lambda y) = f(x) + \lambda \cdot f(y), \quad \text{para todo } x, y \in E,\ \lambda \in K,$$

entonces f es una aplicación lineal.

Proposición 1.3.3.

1. La aplicación identidad, $Id : E \longrightarrow E$, $Id(u) = u$, $\forall u \in E$, es una aplicación lineal.
2. Sean $f : E \longrightarrow F$, $g : F \longrightarrow G$ dos aplicaciones lineales. Entonces la composición
$$g \circ f : E \longrightarrow G, \quad (g \circ f)(x) = g(f(x)),\ \forall x \in E,$$
es lineal.
3. La aplicación nula, $\mathbf{0} : E \longrightarrow F$, definida por $\mathbf{0}(u) := 0$, $\forall u \in E$, es una aplicación lineal.

4. La suma de dos aplicaciones $f, g : E \longrightarrow F$, es la aplicación $f + g : E \longrightarrow F$ definida por
$$(f + g)(x) := f(x) + g(x).$$
Si f y g son aplicaciones lineales, también lo es $f + g$.
5. Si $f : E \longrightarrow F$ es una aplicación, y $\lambda \in K$ es un escalar, el producto de λ por f es la aplicación $\lambda \cdot f : E \longrightarrow F$ definida por
$$(\lambda \cdot f)(x) := \lambda \cdot f(x).$$
Si f es una aplicación lineal, también lo es $\lambda \cdot f$.

Demostración. La primera parte es trivial. Veamos la segunda. Por 1.3.2, basta observar que
$$(g\circ f)(u+\lambda\cdot v) = g(f(u+\lambda\cdot v)) = g(f(u)+\lambda\cdot f(v)) = g(f(u))+\lambda\cdot g(f(v)) = (d\circ f)(u)+\lambda\cdot(g\circ f)(v),$$
para todo $u, v \in E$, $\lambda \in \mathbf{K}$.

La tercera parte es obvia. La cuarta y la quinta las dejamos a cargo del lector. □

Ejemplo 1.3.4. Damos a continuación unos ejemplos básicos de aplicaciones lineales.

1. Sea $A \in \mathbb{R}(n, m)$. La multiplicación de vectores columna por A,
$$L_A : \mathbb{R}^m \longrightarrow \mathbb{R}^n, \quad L_A(u) = A \cdot u,$$
es lineal.
   Por ejemplo, si A es la matriz $A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{pmatrix} \in \mathbb{R}(2, 4)$, entonces
$$L_A : \mathbb{R}^4 \longrightarrow \mathbb{R}^2, \quad (x_1, x_2, x_3, x_4) \mapsto A \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} x_1 + 2x_2 + 3x_3 + 4x_4 \\ 5x_1 + 6x_2 + 7x_3 + 8x_4 \end{pmatrix}.$$
2. Sea $A \in \mathbb{R}(n, m)$. La multiplicación de matrices por A por la izquierda,
$$L_A : \mathbb{R}(m, l) \longrightarrow \mathbb{R}(n, l), \quad L_A(B) = A \cdot B,$$
es lineal.
   Por ejemplo, si $A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{pmatrix} \in \mathbb{R}(2, 4)$, y $l = 2$, entonces
$$L_A : \mathbb{R}(4, 2) \longrightarrow \mathbb{R}(2, 2), \quad B = \begin{pmatrix} b_1^1 & b_2^1 \\ b_1^2 & b_2^2 \\ b_1^3 & b_2^3 \\ b_1^4 & b_2^4 \end{pmatrix} \mapsto A \cdot B = \begin{pmatrix} b_1^1 + 2b_1^2 + 3b_1^3 + 4b_1^4 \\ 5b_2^1 + 6b_2^2 + 7b_2^3 + 8b_2^4 \end{pmatrix}$$
   Analogamente la multiplicación de matrices por A por la derecha
$$R_A : \mathbb{R}(l, n) \longrightarrow \mathbb{R}(l, m), \quad R_A(B) = B \cdot A,$$
también es lineal.
   Por ejemplo, si $A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{pmatrix} \in \mathbb{R}(2, 4)$, y $l = 2$, entonces
$$R_A : \mathbb{R}(3, 2) \longrightarrow \mathbb{R}(3, 4),$$
$$B = \begin{pmatrix} b_1^1 & b_2^1 \\ b_1^2 & b_2^2 \\ b_1^3 & b_2^3 \end{pmatrix} \mapsto B \cdot A = \begin{pmatrix} b_1^1 + 5b_2^1 & 2b_1^1 + 6b_2^1 & 3b_1^1 + 7b_2^1 & 4b_1^1 + 8b_2^1 \\ b_1^2 + 5b_2^2 & 2b_1^2 + 6b_2^2 & 3b_1^2 + 7b_2^2 & 4b_1^2 + 8b_2^2 \\ b_1^3 + 5b_2^3 & 2b_1^3 + 6b_2^3 & 3b_1^3 + 7b_2^3 & 4b_1^3 + 8b_2^3 \end{pmatrix}$$

Veamos otros ejemplos.

Ejercicio 1.3.5.

1. Sea $a \in \mathbb{R}$. La evaluación de un polinomio en el punto a,
$$E_a : \mathbb{R}[x] \longrightarrow \mathbb{R}, \quad E_a(p(x)) = p(a),$$
es una aplicación lineal.
2. Sea $a(x) \in \mathbb{R}[x]$. La multiplicación de polinomios por $a(x)$,
$$L : \mathbb{R}[x] \to \mathbb{R}[x], \quad L(p(x)) = a(x) \cdot p(x)$$
es lineal.
3. La derivación de polinomios
$$D : \mathbb{R}[x] \longrightarrow \mathbb{R}[x], \quad D(p) = p'$$
es lineal.
4. Para todo $n \geq 1$, la derivada n-ésima de polinomios
$$D^n : \mathbb{R}[x] \to \mathbb{R}[x],$$
definida recurrentemente por $D^n = D \circ D^{n-1}$, es una aplicación lineal.
5. Sean $a, b \in \mathbb{R}$. La integración definida entre a y b de un polinomio
$$\int_a^b : \mathbb{R}[x] \longrightarrow \mathbb{R}, \quad p(x) \mapsto \int_a^b p(x)dx,$$
es una aplicación lineal.
6. Si $a \in \mathbb{R}$, la integración con extremo variable
$$\int_a^x : \mathbb{R}[x] \longrightarrow \mathbb{R}[x], \quad p(x) \mapsto \int_a^x p(x)dx,$$
es una aplicación lineal.

Los ejemplos anteriores para polinomios se pueden repetir para funciones definidas en un intervalo I de $\mathbb{R}$.

Ejercicio 1.3.6. Sea I un intervalo de $\mathbb{R}$.

1. Las aplicaciones siguientes son lineales:
$$\begin{aligned} E_a : \mathcal{F}(I) &\longrightarrow \mathbb{R}, \quad E_a(f) = f(a), \\ D : \mathcal{C}^1(I) &\longrightarrow \mathcal{C}^0(I), \quad D(f) = f', \\ D^n : \mathcal{C}^n(I) &\longrightarrow \mathcal{C}^0(I), \quad D^n(f) = D \circ D^{n-1}, \\ \int_a^b : \mathcal{C}^0(I) &\longrightarrow \mathbb{R}, \quad f \mapsto \int_a^b f(x)dx, \\ \int_a^x : \mathcal{C}^0(I) &\longrightarrow \mathcal{C}^1(I), \quad f \mapsto \int_a^x f(x)dx. \end{aligned}$$
2. Sean I un intervalo de $\mathbb{R}$, y $a(t) = a_0 + a_1 t + \cdots + a_n t^n \in \mathbb{R}[t]$ un polinomio de grado $\leq n$. Para todo $f \in \mathcal{C}^\infty(I)$, denotamos por $a(D)(f)$ la función
$$a(D)(f) = a_0 f + a_1 D(f) + \cdots + a_n D^n(f).$$
Entonces $a(D) : \mathcal{C}^\infty(I) \longrightarrow \mathcal{C}^\infty(I)$ es lineal.

Ejercicio 1.3.7. Sea $\mathcal{C}(\mathbb{N})$ el $\mathbb{R}$-espacio vectorial de las sucesiones de números reales convergentes, entonces
$$\lim_{n\to\infty} : \mathcal{C}(\mathbb{N}) \longrightarrow \mathbb{R}, \quad (a_n)_n \mapsto \lim_{n\to\infty} a_n,$$
es una aplicación lineal.

1.3.2. **Isomorfismos.**

En el caso de aplicaciones entre conjuntos un isomorfismo, es decir, una aplicación que tiene inversa, es simplemente una aplicación biyectiva. Veremos que esto también sucede para las aplicaciones lineales.

En primer lugar veamos que la aplicación identidad de un espacio vectorial es una aplicación lineal, y también lo es la composición de aplicaciones lineales.

Un isomorfismo es una aplicación lineal que tiene una inversa lineal, aunque veremos que la linealidad de la inversa es una condición redundante y basta que la inversa como aplicación entre conjuntos. Sin embargo queremos poner este hecho de manifiesto porque en otras situaciones la definición de isomorfismo sí requiere hipótesis para la inversa. Por ejemplo, la función $f : \mathbb{R} \longrightarrow \mathbb{R}$, $f(x) = x^3$, es diferenciable y biyectiva, pero su inversa $f^{-1}(y) = \sqrt[3]{y}$ no es diferenciable (en $y = 0$), por tanto f no es un isomorfismo (en el sentido diferenciable).

Definición 1.3.8. Una aplicación lineal $f : E \longrightarrow F$ se dice que es un *isomorfismo* si existe una aplicación lineal $g : F \longrightarrow E$ tal que $f \circ g = Id_F$ y $g \circ f = Id_E$.

Dos espacios vectoriales E y F se llaman *isomorfos* si existe un isomorfismo entre ellos. En tal caso escribiremos $E \cong F$.

Todo isomorfismo es obviamente una aplicación lineal biyectiva. El recíproco también es cierto.

Proposición 1.3.9. La inversa de una aplicación lineal biyectiva es lineal. En particular, una aplicación lineal es un isomorfismo si y solo si es biyectiva.

Demostración. Sea $f : E \longrightarrow F$ lineal y biyectiva, y sea g su inversa. Veamos que g es lineal. Por 1.3.2, basta ver que

$$g(u + \lambda \cdot v) = g(u) + \lambda \cdot g(v),$$

para todo $u, v \in F$, $\lambda \in \mathbf{K}$. En efecto, por la linealidad de f, y teniendo en cuenta que $f \circ g = Id$, se tiene:

$$f(g(u) + \lambda \cdot g(v)) = f(g(u)) + \lambda \cdot f(g(v)) = u + \lambda \cdot v = (f \circ g)(u + \lambda \cdot v) = f(g(u + \lambda \cdot v)).$$

Puesto que f es inyectiva, de aquí se sigue que $g(u + \lambda \cdot v) = g(u) + \lambda \cdot g(v)$. □

Ejercicio 1.3.10. Sea $f : \mathbb{R}[x]_{\leq n} \longrightarrow \mathbb{R}[x]_{\leq n}$ definida por

$$f(p(x)) = p(x + 1).$$

Prueba que f es un isomorfismo y halla su inverso.

1.3.3. **Núcleo e imagen de una aplicación lineal.**

Así como el conjunto de soluciones de un sistema homogéneo de ecuaciones lineales en n incógnitas es un subespacio vectorial de $\mathbb{R}^n$, en general el llamado "núcleo"de una aplicación lineal también es un subespacio vectorial.

Por otra parte, los subespacios vectoriales de $\mathbb{R}^n$ admiten una descripción en forma paramétrica. Este hecho se generaliza a los subconjuntos de imágenes de una aplicación lineal.

Definición 1.3.11. Si $f : E \longrightarrow F$ es una aplicación lineal, su *núcleo* y su *imagen* son, respectivamente,

$$\mathrm{Nuc}(f) = f^{-1}(0) = \{x \in E;\ f(x) = 0\,\}, \quad \mathrm{Im}(f) = f(E) = \{y \in F;\ \exists x \in E,\ f(x) = y\,\}.$$

El núcleo $\mathrm{Nuc}(f)$ se denota también por $\mathrm{Ker}\,(f)$.

Proposición 1.3.12. Sea $f : E \longrightarrow F$ una aplicación lineal entre dos K-espacios vectoriales. El núcleo Nuc f es un subespacio vectorial de E y la imagen Im f es un subespacio vectorial de F.

Observación 1.3.13. Si $f : E \longrightarrow F$ es una aplicación lineal, y $v \in F$ un vector no nulo. El subconjunto de E, $f^{-1}(v) = \{u \in E; f(u) = v\}$, no es un subespacio vectorial de E.

Ejemplo 1.3.14.

1. El ejemplo (1) de 1.3.4 nos da el ejemplo básico de subespacio vectorial de $\mathbb{R}^n$, el conjunto de soluciones de un sistema homogéneo de ecuaciones lineales en n incógnitas: Si A es una matriz (de tipo $n \times m$), el conjunto de soluciones del sistema homogéneo de m ecuaciones lineales en las incógnitas $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$, definido por la ecuación matricial $A \cdot \mathbf{x} = 0$, es un subespacio vectorial de $\mathbb{R}^n$. Por el contrario, si $\mathbf{b} = (b_1, \cdots, b_m) \in \mathbb{R}^m$ es no nulo, el conjunto de soluciones $\mathbf{x} \in \mathbb{R}^n$ de $A\mathbf{x} = \mathbf{b}$ no es un subespacio vectorial de $\mathbb{R}^n$.

   Por ejemplo, el conjunto de soluciones del sistema de ecuaciones lineales
$$\begin{cases} 3x_1 + 2x_2 - 3x_3 = 0 \\ 4x_1 - 2x_2 + x_3 + x_4 = 0 \end{cases}$$
   en las incógnitas $(x_1, x_2, x_3, x_4) \in \mathbb{R}^4$, es un subespacio vectorial de $\mathbb{R}^4$.

   Por el contrario, el conjunto de soluciones del sistema de ecuaciones "no homogeneo"
$$\begin{cases} 3x_1 + 2x_2 - 3x_3 = 1 \\ 4x_1 - 2x_2 + x_3 + x_4 = 0 \end{cases}$$
   no es un subespacio vectorial.
2. El conjunto de matrices $C \in \mathbb{R}(3, 4)$ de la forma
$$\begin{pmatrix} b_1^1 + 5b_2^1 & 2b_1^1 + 6b_2^1 & 3b_1^1 + 7b_2^1 & 4b_1^1 + 8b_2^1 \\ b_1^2 + 5b_2^2 & 2b_1^2 + 6b_2^2 & 3b_1^2 + 7b_2^2 & 4b_1^2 + 8b_2^2 \\ b_1^3 + 5b_2^3 & 2b_1^3 + 6b_2^3 & 3b_1^3 + 7b_2^3 & 4b_1^3 + 8b_2^3 \end{pmatrix}$$
   con $b_1^1, b_1^2, b_1^3, b_1^4, b_2^1, b_2^2, b_2^3, b_2^4 \in \mathbb{R}$, es la imagen de la aplicación lineal R_A del ejemplo (2) de 1.3.4, por tanto es un subespacio vectorial de $\mathbb{R}(3, 4)$.

Ejemplo 1.3.15. Sea $\mathcal{C}(\mathbb{N})$ el $\mathbb{R}$-espacio vectorial de las sucesiones de números reales convergentes. Puesto que la aplicación $\lim_{n\to\infty} : \mathcal{C}(\mathbb{N}) \longrightarrow \mathbb{R}$ es lineal, su núcleo $\mathcal{C}_0$, el subconjunto de las sucesiones convergentes a cero, es un subespacio vectorial de $\mathcal{C}(\mathbb{N})$.

Ejemplo 1.3.16.
1. Sea $[a, b]$ un intervalo cerrado de $\mathbb{R}$. El subconjunto de funciones $f \in \mathcal{C}^0([a, b])$ tales que $\int_a^b f(x)dx = 0$ es un subespacio vectorial de $\mathcal{C}^0([a, b])$.
2. El núcleo de la aplicación lineal $D^{n+1} : \mathcal{C}^\infty(\mathbb{R}) \longrightarrow \mathcal{C}^\infty(\mathbb{R})$ es $\mathbb{R}[x]_{\leq n}$, el subespacio vectorial de los polinomios de grado menor o igual que n.

Ejemplo 1.3.17. Sean I un intervalo de $\mathbb{R}$, y $E = \mathcal{C}^\infty(I)$. Sea $a(t) = a_0 + a_1 t + \cdots + a_n t^n$ un polinomio de grado $\leq n$. Entonces el núcleo de la aplicación lineal $a(D) : E \longrightarrow E$,
$$\mathcal{G} = \{f \in \mathcal{C}^\infty(I);\ a(D)(f) = \mathbf{0}\}$$
es un subespacio vectorial de E. En la teoría de ecuaciones diferenciales este hecho se conoce como principio de superposición de las soluciones de una ecuación diferencial homogénea.

Veamos algunos casos particulares:

1. Para el polinomio $a(t) = D^{n+1}$, se tiene $a(D) = D^{n+1}$, y el núcleo de $D^{n+1} : E \longrightarrow E$ es el conjunto de soluciones en $f \in E$ de la ecuación $D^{n+1}(f) = 0$, es decir, el subespacio vectorial $\mathbb{R}[x]_{\leq n}$ de E formado por las funciones polinomiales de grado $\leq n$.

2. Sea $\lambda \in \mathbb{R}$ no nulo. Para el polinomio $a(t) = (t - \lambda)^n$, se tiene $a(D) = (D - \lambda \cdot Id)^n$, y el núcleo de $(D - \lambda \cdot Id)^n : E \longrightarrow E$ es el conjunto de soluciones en f de la ecuación $(D - \lambda \cdot Id)^n(f) = 0$. Éste es el subespacio vectorial de E formado por las funciones de la forma $f(x) = p(x)e^{\lambda x}$, donde $p(x)$ es un polinomio de grado $< n$. (Prueba: si expresamos una solución en la forma $f = ge^{\lambda x}$, entonces, por inducción sobre n podemos comprobar que $(D - \lambda \cdot Id)^n(ge^{\lambda x}) = (D^n g)e^{\lambda x}$ y la ecuación $(D - \lambda \cdot Id)^n(f) = 0$ equivale a $D^n g = 0$, es decir, a $g = p(x)$, donde $p(x)$ es un polinomio de grado $< n$.)
3. Sea $\omega \in \mathbb{R}$ no nulo. Para el polinomio $a(x) = x^2 + \omega^2$, se tiene $a(D) = D^2 + Id \cdot \omega^2$, y el núcleo de $a(D)$ es el conjunto de soluciones en $f \in E$ de la ecuación
$$D^2(f) = -\omega^2 f$$
Éste es el subespacio vectorial de E formado por las funciones de la forma $A \cos \omega x + B \sin \omega x$, con $A, B \in \mathbb{R}$ constantes arbitrarias. (En efecto. Obviamente toda función de la forma $f(x) = A \sin \omega x + B \cos \omega x$ cumple la ecuación $D^f = -\omega^2 f$. Además observeos que si $f(x) = A \sin \omega x + B \cos \omega x$, se cumple
$$A = f(0), \quad B = \frac{D(f)(0)}{\omega}.$$
Ahora, si f es del núcleo de $D^2 + \omega \cdot Id$, restando a f la función $f(0) \cos \omega x + \frac{(Df)(0)}{\omega} \sin \omega x$, podemos suponer que $f(0) = (Df)(0) = 0$. Multiplicando la ecuación $D^2(f) + \omega^2 f = 0$ por Df se obtiene
$$D(f)D^2(f) + \omega^2 f D(f) = 0,$$
por tanto
$$D(D(f)^2 + \omega^2 f^2) = 0,$$
por lo cual $D(f)^2 + \omega^2 f^2$ es constante. Como $f(0) = (Df)(0) = 0$ resulta $f(x) = 0$ para todo $x \in \mathbb{R}$.)

Ejercicio 1.3.18. Dada una función $g \in \mathcal{C}^0(\mathbb{R})$ no nula, el conjunto de primitivas de f, es decir, de soluciones $f \in \mathcal{C}^1(\mathbb{R})$ de la ecuación $D(f) = g$, no es un subespacio vectorial de $\mathcal{C}^1(\mathbb{R})$.

1.4. Combinaciones lineales y dependencia lineal, generadores. En lo que sigue, K denota un cuerpo y E un espacio vectorial sobre K.

1.4.1. **Intersección de subespacios. Subespacio generado por un subconjunto.**

Sea X un conjunto. Dados dos subconjuntos A y B de X su intersección es el subconjunto de X
$$A \cap B = \{x \in X;\ x \in A, x \in B\ \}.$$
Más generalmente, podemos definir la intersección de una familia de subconjuntos de X. Recordemos que, si I es un conjunto (de índices), una familia de subconjuntos de X indexada por I es una aplicación de I en el conjunto de partes de X,
$$I \longrightarrow \mathfrak{P}(X), \quad i \mapsto A_i.$$
Una tal familia se denota por $\{A_i\}_{i \in I}$. La intersección de esta familia es:
$$\bigcap_{i \in I} A_i = \{x \in X;\ x \in A_i,\ \forall\, i \in I\ \}.$$

La operación de intersección de conjuntos es habitual en la consideración de problemas geométricos. Por ejemplo sabemos que una recta es una intersección de dos planos en $\mathbb{R}^3$.

Sea E un espacio vectorial sobre K. En general tenemos:

Proposición 1.4.1. La intersección de una familia de subespacios vectoriales de E es un subespacio vectorial de E.

Observación. El resultado anterior no es cierto para la reunión de subespacios vectoriales. Por ejemplo, en $\mathbb{R}^3$ la reunión conjuntista de dos rectas distintas nunca es un plano ni una recta, por tanto no es un subespacio vectorial.

Ejemplo 1.4.2. La intersección $\mathcal{C}^\infty(I) = \bigcap\limits_{n\in\mathbb{N}} \mathcal{C}^n(I)$ es un subespacio vectorial de $\mathcal{F}(I)$. Sus funciones se llaman infinitamente derivables. Esta intersección infinita no se puede reducir a una intersección finita.

Definición 1.4.3. Dado un subconjunto S de E, se llama *subespacio vectorial generado por* S al mínimo subespacio vectorial $\langle S\rangle$ que contiene a S. Es decir, $\langle S\rangle$ es un subespacio de E que contiene a S y, si F es un subespacio que contiene a S entonces F contiene a $\langle S\rangle$.

Se dice también que S es un *conjunto de generadores* del subespacio vectorial $\langle S\rangle$.

Veamos que este subespacio vectorial existe efectivamente.

Proposición 1.4.4. Sea S un subconjunto de un espacio vectorial E. Entonces $\langle S\rangle$ coincide con la intersección de todos los subespacios de E que contienen a S. En particular, existe y es único.

Demostración. En efecto, si G denota la intersección de todos los subespacios de E que contienen a S, entonces G es un subespacio vectorial que contiene a S y, además, si F es un subespacio que contiene a S entonces $G \subset F$, por tanto G es mínimo. □

Definición 1.4.5. Un espacio vectorial (o un subespacio) se dice que es *finitamente generado* si admite un conjunto finito de generadores.

Ejemplo 1.4.6.
1. Para $S = \emptyset$ se tiene $\langle\emptyset\rangle = \{0\}$.
2. Para todo subespacio vectorial F de E se tiene $\langle F\rangle = F$.
3. Si $v \in E$ entonces $\langle\{v\}\rangle = \{\lambda\cdot v; \lambda \in K\}$. En particuar, si $v = 0$ obtenemos $\langle\{0\}\rangle = \{0\}$.
4. Sean $u, v \in E$. Entonces $\langle\{u, v\}\rangle = \{\lambda \cdot u + \mu \cdot v; \lambda, \mu \in K\}$.

1.4.2. Combinación lineal.

En E, la combinación de las operaciones de suma de vectores y producto por escalares da lugar al concepto de combinación lineal.

Definición 1.4.7. Si $\lambda, \mu \in K, u, v \in E$, entonces el vector

$$\lambda \cdot u + \mu \cdot v \in E$$

se llama *combinación lineal* de u, v con coeficientes λ, μ.

Más generalmente, si S es un subconjunto de E, posiblemente infinito, una combinación lineal de elementos de S es un vector $x \in E$ tal que existen $r \in \mathbb{N}$, $u_i \in S$, $\lambda_i \in K$, $1 \leq i \leq r$, tales que

$$x = \lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_r u_r.$$

Ejemplo 1.4.8. Puesto que la suma de 0 elementos es el vector cero, el vector cero es una combinación lineal de S, para cualquier subconjunto S de E, incluso para $S = \emptyset$.

Ejemplo 1.4.9. En $\mathbb{R}[x]$ toda combinación lineal de los monomios $\{x^0, x^1, x^2, \dots, x^n, \dots\}$ es un polinomio (de grado $\leq n$).

Las sumas infinitas no son, en general, polinomios. Por ejemplo, $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$ no es un polinomio.

Ejemplo 1.4.10.

1. El vector $(1,3,5) \in \mathbb{R}^3$ es combinación lineal de $S = \{(4,10,12),(1,2,1)\}$, pues $(1,3,5) = \frac{1}{2}(4,10,12) - (1,2,1)$.
2. En $E = \mathcal{F}(\mathbb{R})$ la función constante 1 es combinación lineal de $S_1 = \{\sin^2 x, \cos^2 x\}$, pero no lo es de $S_2 = \{\sin x, \cos x\}$.

Definición 1.4.11. Dado un conjunto de vectores S de E, se dice que un vector $v \in E$ *depende linealmente de* S si v se puede expresar como combinación lineal de los elementos de S.

Ejemplo 1.4.12. Siguiendo con el ejemplo 1.4.10 podemos decir que

1. El vector $(1,3,5) \in \mathbb{R}^3$ depende linealmente de $S = \{(4,10,12),(1,2,1)\}$.
2. En $E = \mathcal{F}(\mathbb{R})$ la función constante 1 depende linealmente de $S_1 = \{\sin^2 x, \cos^2 x\}$, pero no de $S_2 = \{\sin x, \cos x\}$.

Como aplicación, damos otra manera de describir el subespacio vectorial generado por un subconjunto.

Proposición 1.4.13. Sea S un subconjunto de E. El subespacio vectorial $\langle S \rangle$ coincide con el conjunto de vectores que son combinaciones lineales de S, es decir, el conjunto de vectores que dependen linealmente de S.

Demostración. Si C_S denota el conjunto de combinaciones lineales de S, hemos de ver que C_S es un subespacio vectorial, que $S \subset C_S$ y que, para todo subespacio vectorial F tal que $S \subset F$ se cumple $C_S \subset F$.

En efecto, 0 es una combinación lineal de $\emptyset$ y por tanto de S. La suma de dos combinaciones lineales de S es una combinación lineal de S, y el producto de un escalar por una combinación lineal de S es una combinación lineal de S. Por tanto C_S es un subespacio vectorial de E.

Para cada $v \in S$, $v = 1 \cdot v$ expresa v como combinación lineal de S, por tanto $S \subset C_S$.

Finalmente, si F es un subespacio vectorial de E tal que $S \subset F$, entonces F contiene todo combinación lineal de S, por tanto $C_S \subset F$. □

En la siguiente proposición vemos una condición suficiente para que al quitar un elemento de un conjunto generador, el subconjunto restante siga siendo un conjunto generador.

Proposición 1.4.14. Sean E un K-espacio vectorial, S un subconjunto de E y $\mathbf{v} \in E$. Si $S \cup \{\mathbf{v}\}$ un conjunto de generadores de E y $\mathbf{v}$ es combinación lineal de S, entonces S es un conjunto de generadores de E.

Demostración. En efecto. Si $S \cup \{v\}$ es un conjunto de generadores, dado $x \in E$ existen vectores $u_1, \cdots, u_n$ de S y una familia de escalares $(\mu, \lambda_1, \cdots, \lambda_n)$ tales que

$$x = \mu v + \lambda_1 u_1 + \cdots + \lambda_n u_n.$$

Por otra parte, como v es combinación lineal de S, existen vectores $v_1, \cdots, v_r$ de S y una familia de escalares $(\phi_1, \cdots, \phi_r)$ tales que

$$v = \phi_1 v_1 + \cdots + \phi_r v_r,$$

por tanto

$$x = \mu(\phi_1 v_1 + \cdots + \phi_r v_r) + \lambda_1 u_1 + \cdots + \lambda_n u_n = \mu\phi_1 v_1 + \cdots + \mu\phi_r v_r + \lambda_1 u_1 + \cdots + \lambda_n u_n,$$

es una combinación lineal de S. □

Ejemplo 1.4.15. 1. En $\mathbb{R}^3$ el subespacio vectorial generado por $S = \{(1,1,0),(1,-1,0)\}$ es $F = \{(x,y,z); z = 0\}$.
2. En $\mathbb{R}[x]$ el subespacio vectorial generado por $S = \{x + x^2, x - x^2\}$ son los polinomios de grado ≤ 2 con término independiente igual a 0.

Ejemplo 1.4.16. Sea $E = \mathcal{C}^\infty(\mathbb{R})$.

1. Para todo $\lambda \in \mathbb{R}$, el núcleo de $(D-\lambda{\cdot}Id)^n$ está generado por $S = \{e^{\lambda x}, xe^{\lambda x}, x^2e^{\lambda x}, \ldots, x^{n-1}e^{\lambda x}\}$ (ver Ejercicio 1.3.17(2)).
2. Si $\omega \in \mathbb{R}$ $\omega > 0$. El núcleo de $D^2 + \omega^2$ está generado por $S = \{\cos \omega x, \sin \omega x\}$ (ver Ejercicio 1.3.17(3)).

1.4.3. Dependencia e independencia lineal.

Definición 1.4.17. Sea S un conjunto de vectores de E. Se dice que S es *linealmente independiente* si toda combinación lineal de vectores de S que tenga algún coeficiente no nulo es un vector no nulo.

Equivalentemente: S es linealmente independiente si la única combinación lineal de elementos de S que da como resultado el vector 0 es la que tiene todos los coeficientes nulos, es decir,

"Si $\lambda_1u_1 + \lambda_2u_2 + \cdots + \lambda_ru_r$ es una combinación lineal de r elementos diferentes de S tal que $\lambda_1u_1 + \lambda_2u_2 + \cdots + \lambda_ru_r = 0$, entonces $\lambda_i = 0, \forall i$".

S se llama *linealmente dependiente* si no es linealmente independiente.

Equivalentemente: S es linealmente dependiente si existe una familia de coeficientes no todos nulos $\{\lambda_1, \cdots, \lambda_r\}$ (no necesariamente diferentes), y r vectores diferentes $u_1, \cdots, u_r$ de S tales que la correspondiente combinación lineal $\lambda_1u_1 + \lambda_2u_2 + \cdots + \lambda_ru_r$ es el vector nulo.

Ejercicio 1.4.18. Sea E un espacio vectorial. Prueba que un subconjunto S de E formado por un solo elemento, $S = \{u\}$ es lineamente independiente si, y solo si, $u \neq 0$.

Ejemplo 1.4.19. 1. Si un subconjunto S de un espacio vectorial E contiene el vector 0, entonces S es linealmente dependiente. Por ejemplo, $S = \{(0,0),(1,0)\}$ es linealmente dependiente en $\mathbb{R}^2$.
2. El subconjunto $S_1 = \{(1,2,3),(4,5,6),(6,9,12)\}$ de $\mathbb{R}^3$ es linealmente dependiente, pues $(6,9,12) = (4,5,6) + 2\cdot(1,2,3)$.
3. El subconjunto $S_2 = \{(1,2,3),(5,7,9)\}$ es linealmente independiente.

Ejercicio 1.4.20. Sea $E = \mathcal{F}(\mathbb{R})$. Demuestra que

1. El subconjunto $S_1 = \{1, \sin x, \cos x\}$ de $\mathcal{F}(\mathbb{R})$ es linealmente independiente.
2. El subconjunto $S_2 = \{1, \cos^2 x, \sin^2 x\}$ de $\mathcal{F}(\mathbb{R})$ es linealmente dependiente.
3. Sea $\omega > 0$. El subconjunto $S_3 = \{\sin \omega x, \cos \omega x\}$ de $\mathcal{F}(\mathbb{R})$ es linealmente independiente.

Ejercicio 1.4.21. Sean $E = \mathcal{F}(\mathbb{R})$, $\alpha, \beta \in \mathbb{R}$, $n \in \mathbb{N}, n > 0$. Prueba que los siguientes subconjuntos de E son linealmente independientes

$$S_1 = \{e^{\alpha x}, e^{\beta x}\}, \alpha \neq \beta; \quad S_2 = \{\sin \alpha x, \cos \alpha x, \sin \beta x, \cos \beta x\}, \alpha, \beta \neq 0, \alpha \neq \beta;$$
$$S_3 = \{e^{\alpha x}, xe^{\alpha x}, \cdots, x^{n-1}e^{\alpha x}\}; \quad S_4 = \{\sin \alpha x, \cos \alpha x, x\sin \alpha x, x\cos \alpha x\}, \alpha \neq 0;$$
$$S_5 = \{e^{\alpha x}\cos \beta x, e^{\alpha x}\sin \beta x\}, \beta \neq 0.$$

En el siguiente lema vemos una condición para que al añadir a un conjunto linealmente independiente S un vector v el conjunto $S \cup \{v\}$ sea también linealmente independiente.

Lema 1.4.22. Sea S un subconjunto finito de E. Supongamos que S es linealmente independiente, y sea $v \in E, v \notin S$. Entonces $S \cup \{v\}$ es linealmente independiente si, y sólo si, v no es combinación lineal de S. Dicho de otro modo, $S \cup \{v\}$ es linealmente dependiente si, y sólo si, v es combinación lineal de S.

Demostración. Supongamos que S tiene r elementos $\{u_1, \cdots, u_r\}$. Probaremos la segunda aserción. Si $S \cup \{v\}$ es linealmente dependiente, existen escalares $\mu, \lambda_1, \ldots, \lambda_n$, no todos nulos, y vectores $u_1, u_2, \ldots, u_r \in S$ diferentes entre sí, tales que

$$\mu v + \lambda_1 u_1 + \cdots + \lambda_n u_n = 0.$$

Puesto que S es linealmente, $\mu \neq 0$. Por tanto podemos despejar v en la forma

$$v = \frac{-\lambda_1}{\mu} u_1 + \cdots + \frac{-\lambda_n}{\mu} u_n$$

y v es combinación lineal de S.

Recíprocamente, si v es combinación lineal de S, existen escalares $\lambda_1, \ldots, \lambda_r$ tales que

$$v = \lambda_1 u_1 + \cdots + \lambda_r u_r,$$

entonces se tiene

$$(-1)v + \lambda_1 u_1 + \cdots + \lambda_r u_r = 0$$

lo que demuestra que $S \cup \{v\}$ es linealmente dependiente, pues al menos el coeficiente de v es no nulo. □

Ejercicio 1.4.23. Sean E un K-espacio vectorial, S un subconjunto de E y $\mathbf{v} \in E$. Prueba que S es linealmente dependiente si, y sólo si, algun vector $\mathbf{v} \in S$ es combinación lineal de los restantes vectores de S, $S \setminus \{\mathbf{v}\}$.

Sea X un conjunto ordenado. Un elemento $x_0 \in X$ se llama *minimal* (resp. *maximal*) si ningún elemento de X es menor (resp. mayor) que x_0.

Teorema 1.4.24. Si X es un conjunto ordenado finito, entonces X tiene un elemento minimal (resp. maximal).

Demostración. Sea $x_0 \in X$. Si x es minimal ya hemos terminado. Si no lo es, existe $x_1 \in X$ tal que $x_1 < x_0$. Si x_1 es minimal ya hemos terminado, si no existe $x_2 \in X$ tal que $x_2 < x_1$. De esta manera obtenemos una cadena descendente

$$x_0 > x_1 > x_2 > \cdots,$$

Si la cadena termina en x_n, entonces x_n es minimal. En caso contrario la cadena es infinita, lo que no es posible pues X tiene solo un número finito de elementos. □

Proposición 1.4.25. Sean E un K-espacio vectorial y F un subespacio de E. Supongamos que F tiene un subconjunto finito S de generadores. Prueba que

1. Todo subconjunto minimal de S que genera F es linealmente independiente.
2. Todo conjunto maximal de S de vectores linealmente independiente genera F.

1.5. Bases, Teorema de Steitnitz.

En lo que sigue E denota un espacio vectorial sobre K.

1.5.1. **Bases.** Recordemos que el cardinal de un conjunto finito X es el número de elementos de X, y se denota por $\#X$.

Definición 1.5.1. Un subconjunto B (finito o infinito) de E se llama una *base* de E si B genera E y B es linealmente independiente.

Proposición 1.5.2. Sea B un subconjunto finito de E. Entonces B es una base de E si, y solo si, todo vector de E se expresa como combinación lineal de B de forma única. Más precisamente, si $B = \{e_1, e_2, \cdots, e_n\}$ tiene n elementos, B es una base si, y solo si, para todo vector $x \in E$ existe una única familia de n escalares $(\lambda_1, \lambda_2, \cdots, \lambda_n)$ tal que

$$x = \lambda_1 \cdot e_1 + \lambda_2 \cdot e_2 + \cdots + \lambda_n \cdot e_n.$$

Demostración. Supongamos que B es una base. Entonces B genera E, por tanto existe una familia de n escalares $(\lambda_1, \lambda_2, \cdots, \lambda_n)$ tal que

$$x = \lambda_1 \cdot e_1 + \lambda_2 \cdot e_2 + \cdots + \lambda_n \cdot e_n.$$

Veamos que esta familia es única. Si

$$x = \phi_1 \cdot e_1 + \phi_2 \cdot e_2 + \cdots + \phi_n \cdot e_n,$$

restando ambas expresiones resulta

$$0 = (\lambda_1 - \phi_1) \cdot e_1 + (\lambda_2 - \phi_2) \cdot e_2 + \cdots + (\lambda_n - \phi_n) \cdot e_n.$$

Puesto que B es linealmente independiente, obtenemos $\lambda_i - \phi_i = 0$, para todo i, lo que prueba la unicidad.

Recíprocamente, si se cumple la propiedad es obvio que B genera E. Además, si $(\lambda_1, \dots, \lambda_n)$ son escalares tales que

$$0 = \lambda_1 \cdot e_1 + \lambda_2 \cdot e_2 + \cdots + \lambda_n \cdot e_n$$

entonces de

$$0 = 0 \cdot e_1 + 0 \cdot e_2 + \cdots + 0 \cdot e_n$$

y de la unicidad, deducimos que $\lambda_i = 0$ para todo i. $\square$

Proposición 1.5.3. Sea E un espacio vectorial sobre K finitamente generado. Si T es un conjunto finito de generadores de E, entonces existe una base B de E tal que $B \subset T$. En particular E tiene una base finita.

Demostración. Puesto que el cardinal del conjunto de partes de T que son generadores de E es finito, existe un subconjunto minimal B de T que genera E, entonces B es una base de E. En efecto, basta ver que B es linealmente dependiente. Puesto que B es un conjunto minimal de generadores, por 1.4.25, B es linealmente independiente. $\square$

Observación 1.5.4. La propiedad anterior es válida también sin la hipótesis de que B sea finito. En este caso las coordenadas de un vector v en un base $B = \{e_i\}_{i\in I}$ es una familia de escalares $\{\lambda_i\}_{i\in I}$ tal que $\lambda_i = 0$ salvo para un subconjunto finito J de índices, y $v = \sum_{i\in J} \lambda_i \cdot e_i$.

Ejemplo 1.5.5.
1. En $\mathbb{R}^3$, sean $e_1 = (1,0,0)$, $e_2 = (0,1,0)$, $e_3 = (0,0,1)$. Entonces $B = \{e_1, e_2, e_3\}$ es una base de $\mathbb{R}^3$. De la misma manera se obtiene una base de $\mathbb{R}^n$, para todo $n \geq 1$.
2. En $\mathbb{R}(2,3)$ las matrices $E_{11} = \left(\begin{smallmatrix} 1&0&0\\0&0&0 \end{smallmatrix}\right)$, $E_{12} = \left(\begin{smallmatrix} 0&1&0\\0&0&0 \end{smallmatrix}\right)$, $E_{13} = \left(\begin{smallmatrix} 0&0&1\\0&0&0 \end{smallmatrix}\right)$, $E_{21} = \left(\begin{smallmatrix} 0&0&0\\1&0&0 \end{smallmatrix}\right)$, $E_{22} = \left(\begin{smallmatrix} 0&0&0\\0&1&0 \end{smallmatrix}\right)$, $E_{23} = \left(\begin{smallmatrix} 0&0&0\\0&0&1 \end{smallmatrix}\right)$, forman una base. Análogamente se obtiene una base para $\mathbb{R}(n,m)$, para todo $n, m \geq 1$.

Ejercicio 1.5.6.
1. En el espacio $\mathbb{R}[x]_{\leq 4}$ de polinomios de grado menor o igual que 4, los monomios $\{1, x, x^2, x^3, x^4\}$ forman una base.

2. Análogamente, en el espacio $\mathbb{R}[x]_{\leq n}$ de polinomios de grado $\leq n$, los monomios $\{1, x, \ldots, x^n\}$ forman una base.
3. En el espacio $\mathbb{R}[x]$ de polinomios, el conjunto de todos los monomios $\{x^i; i \geq 0\}$ es una base. Esta base tiene infinitos elementos.

Ejercicio 1.5.7. En el espacio vectorial $\mathcal{F}(\mathbb{N})$ de las sucesiones, sean

$$\begin{aligned}
e_0 &= (1, 0, \cdots, 0, \cdots), \\
e_1 &= (0, 1, 0, \cdots, 0, \cdots), \\
e_2 &= (0, 0, 1, 0, \cdots, 0, \cdots), \\
&\vdots \\
e_n &= (0, 0, \cdots, 0, 1, 0, \cdots), \text{ 1 en el lugar } n\text{-ésimo}, \\
&\vdots
\end{aligned}$$

Entonces el conjunto infinito $\{e_i; i \in \mathbb{N}\}$ es linealmente independiente, pero NO es una base de $\mathcal{F}(\mathbb{N})$. En particular es falso que la sucesión constante $a_n = 1, \forall\, n$, sea la suma de todas las sucesiones e_i, pues esa suma no sería finita, y por tanto no está definida. No conocemos una base explícita de $\mathcal{F}(\mathbb{N})$.

Ejercicio 1.5.8. En $\mathbb{R}^3$, con coordenadas (x, y, z),

1. Las soluciones de la ecuación $x + y - z = 0$ forman un subespacio vectorial que tiene por base $B = \{(1, 0, 1), (0, 1, 1)\}$.
2. Las soluciones del sistema $x + y - z = 0, y + z = 0, x + 2y = 0$ forman un subespacio vectorial que tiene por base $B = \{(2, -1, 1)\}$.
3. El conjunto $\{(2x + 2y, x + y, -3x - 3y); x, y \in \mathbb{R}\}$ es un subespacio vectorial que tiene por base $\{(2, 1, -3)\}$.
4. El conjunto $\{(2x + 2y, x + y, -3x + 3y); x, y \in \mathbb{R}\}$ es un subespacio vectorial que tiene por base $\{(2, 1, -3), (2, 1, 3)\}$.

Ejercicio 1.5.9. En el espacio de matrices $\mathbb{R}(2, 2)$, el conjunto de soluciones $X \in \mathbb{R}(2, 2)$ de la ecuación $\left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right) \cdot X = 0$, tiene por base $\{\left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right)\}$.

Ejercicio 1.5.10. En el espacio de polinomios $\mathbb{R}[x]$, los polinomios $p(x)$ que cumplen $p(0) + p'(1) = 0, D^4 p(x) = 0$ es un subespacio vectorial que tiene por base $\{-2x + x^2, -3x + x^3\}$.

Ejercicio 1.5.11. Sea $E = \mathcal{C}^\infty(\mathbb{R})$. Prueba que, en cada caso, B es una base de subespacio de E formado por las soluciones f de la ecuación $a(D)(f) = 0$, donde $a(t)$ es el polinomio indicado (ver Ejercicios 1.3.17, 1.4.16 y 1.4.21).

1. $a(t) = (t - \alpha)^n$, con $\alpha \in \mathbb{R}$, $n \in \mathbb{N}, n > 0$, $B = \{e^{\alpha x}, xe^{\alpha x}, \cdots, x^{n-1}e^{\alpha x}\}$.
2. $a(t) = t^2 + \omega^2$, con $\omega \in \mathbb{R}$ no nulo, $B = \{\cos \omega x, \sin \omega x\}$.

1.5.2. **El teorema de Steitnitz.** El siguiente lema es la primera etapa en la prueba de un resultado clave de la teoría de la dimensión de los espacios vectoriales de dimensión finita, el teorema de Steitnitz.

Lema 1.5.12. Sea $B = \{e_1, e_2, \cdots, e_n\}$ una base de E de cardinal n, $v \in E$ un vector no nulo. Sea

$$v = \lambda_1 e_1 + \lambda_2 e_2 + \cdots + \lambda_n e_n.$$

Si $\lambda_1 \neq 0$, entonces $B' = \{v, e_2, \cdots, e_n\}$ es una base de E.

Demostración. Hemos de ver que B' genera E y que B' linealmente independiente.

Veamos que B' genera E. Puesto que $\lambda_1 \neq 0$, podemos despejar e_1 como combinación lineal de B',

$$e_1 = \frac{\lambda_2}{-\lambda_1}e_2 + \cdots \frac{\lambda_n}{-\lambda_1}e_n + \frac{1}{\lambda_1}v.$$

Sea $x \in E$. Veamos que x es combinación lineal de $\{v, e_2, \ldots, e_n\}$. Puesto que B es un conjunto de generadores, existen escalares $\phi_1, \ldots, \phi_n$ tales que

$$x = \phi_1 e_1 + \phi_2 e_2 + \cdots + \phi_n e_n.$$

Sustituyendo e_1 por su expresión como combinación lineal de $\{v, e_2, \ldots, e_n\}$ obtenemos

$$\begin{aligned} x =& \phi_1 \left(\frac{\lambda_2}{-\lambda_1}e_2 + \cdots \frac{\lambda_n}{-\lambda_1}e_n + \frac{1}{\lambda_1}v \right) + \phi_2 e_2 + \cdots + \phi_n e_n \\ =& \left(\phi_1 \frac{\lambda_2}{-\lambda_1} + \phi_2 \right) e_2 + \cdots + \left(\phi_1 \frac{\lambda_n}{-\lambda_1} + \phi_n \right) e_n + \frac{\phi_1}{\lambda_1}v, \end{aligned}$$

lo que prueba que x es combinación lineal de $\{v, u_2, \ldots, u_n\}$. De aquí concluimos que $\{v, e_2, \ldots, e_n\}$ es un conjunto de generadores de E.

Veamos que B' es linealmente independiente. Supongamos que

$$0 = \phi_1 v + \phi_2 e_2 + \cdots + \phi_n e_n.$$

Sustituyendo v por su expresión como combinación lineal de $\{e_1, e_2, \ldots, e_n\}$, obtenemos

$$\begin{aligned} 0 =& \phi_1(\lambda_1 e_1 + \lambda_2 e_2 + \cdots + \lambda_n e_n) + \phi_2 e_2 + \cdots + \phi_n e_n \\ =& (\phi_1 \lambda_1)e_1 + (\phi_1 \lambda_2 + \phi_2)e_2 + \cdots + (\phi_1 \lambda_n + \phi_n)e_n. \end{aligned}$$

y puesto que $\{e_1, \ldots, e_n\}$ es linealmente independiente, todos los coeficientes han de ser nulos:

$$\phi_1\lambda_1 = 0,\ \phi_1\lambda_2 + \phi_2 = 0,\ \cdots, \phi_1\lambda_n + \phi_n = 0.$$

Puesto que $\lambda_1 \neq 0$, resulta $\phi_1 = 0$, y por tanto $\phi_i = 0$, para todo i. Esto prueba que B' es linealmente independiente. □

Ejemplo 1.5.13. En $\mathbb{R}^3$, consideremos la base canónica $B = \{e_1, e_2, e_3\}$, y $v = (3, 0, 2)$. Entonces $v = 3e_1 + 0e_2 + 2e_3$, por tanto v puede sustituir a e_1 o a e_3 en la base B, pero no puede sustituir a e_2.

A continuación damos el teorema más importante de este capítulo.

Teorema 1.5.14. (Teorema de Steitnitz para espacios con una base finita.) Sea E un espacio vectorial $E \neq \{0\}$. Si $B = \{e_1, e_2, \ldots, e_n\}$ es una base finita de E, de cardinal $n \geq 1$, para todo conjunto linealmente independiente S de E se cumple

1. S es finito, y $\#S \leq n$.
2. Existe un subconjunto $T \subset B$, de cardinal $n - r$, tal que $S \cup T$ es una base de E de cardinal n.

Demostración. Supongamos en primer lugar que S es finito, y razonamos por inducción sobre $r = \#S$. Si $r = 0$, entonces $0 \leq n$ y $T = B$. Si $r = 1$, $S = \{v_1\}$ tiene un solo elemento, y puesto que S es linealmente independiente, $v_1 \neq 0$, y por el lema 1.5.12 anterior se sigue el teorema.

Supongamos que el teorema es cierto para todo subconjunto con menos de r elementos, y que S tiene r elementos, $S = \{v_1, v_2, \ldots, v_{r-1}, v_r\}$. Sea $S' = \{v_1, v_2, \ldots, v_{r-1}\}$. Entonces S' es linealmente independiente, y por hipótesis de inducción, $r - 1 \leq n$ y podemos sustituir $r - 1$

elementos en B por los $r-1$ vectores de S'. Reordenando la base B podemos suponer que hemos sustituido los $r-1$ primeros elementos B, y que $B' = \{v_1, v_2, \dots, v_{r-1}, e_r, \dots, e_n\}$ es una base de E. Ahora expresamos v_r en esta base,

$$v_r = \lambda_1 v_1 + \cdots + \lambda_{r-1} v_{r-1} + \lambda_r e_r + \cdots + \lambda_n e_n$$

Puesto que S es linealmente independiente, los coeficientes λ_i con $i \geq r$ no pueden ser todos nulos, por tanto $n-(r-1) > 0$, es decir $r \leq n$. Reordenado los índices podemos suponer que $\lambda_r \neq 0$. Aplicando el lema 1.5.12 anterior, $B'' = \{v_1, v_2, \dots, v_r, e_{r+1}, \dots, e_n\}$ es una base de E. Lo cual prueba el teorema en este caso.

Si S es arbitrario, para cualquier subconjunto finito T de S, T tiene cardinal $\leq n$, por lo que acabamos de demostrar. Por lo tanto S es finito y de cardinal $\leq n$. □

Ejemplo 1.5.15. En $\mathbb{R}^3$ consideremos la base canónica $B = \{e_1, e_2, e_3\}$ y el conjunto de vectores $S = \{(1,2,3), (2,4,5)\}$. Entonces, S es linealmente independiente y $S \cup \{e_1\}$ es una base de $\mathbb{R}^3$. También $S \cup \{e_2\}$ es una base, pero $S \cup \{e_3\}$ no lo es.

Corolario 1.5.16. Sea E un espacio vectorial sobre K finitamente generado. Si S es un conjunto finito de vectores linealmente independiente de E, entonces existe una base B de E tal que $S \subset B$.

Corolario 1.5.17. Sea E un espacio vectorial sobre un cuerpo K. Si E tiene una base finita con n elementos, todo subconjunto de E con más de n vectores es linealmente dependiente.

El teorema de Steitnitz tiene una versión para espacios que pueden no tener una base finita.

Ejercicio 1.5.18. Si E es un espacio vectorial, y S es un subconjunto de E linealmente independiente, existe una base B de E tal que $S \subset B$. En particular, todo espacio vectorial tiene una base. (Indicación: Consideremos el conjunto τ de todos los subconjuntos T de E tales que T es linealmente independiente y $S \subset T$. Entonces τ, ordenado con la inclusión, es un conjunto inductivo y no vacío. En efecto, $S \in \tau$ y, si τ' es un subconjunto totalmente ordenado de τ, entonces su reunión $T = \bigcup \tau'$ es de τ, pues cualquier subconjunto finito F de T está en algún elemento T_i de τ', y por tanto F es linealmente independiente. Por el lema de Zorn, τ tiene un elemento maximal B. Es obvio que B es una base, y que $S \subset B$.)

1.6. Dimensión y coordenadas.

1.6.1. **Dimensión.** Gracias al teorema de Steitnitz podemos afirmar que todas las bases de un espacio vectorial finitamente generado E tienen el mismo número de elementos.

Corolario 1.6.1. Si E tiene una base finita, todas las demás bases de E son finitas y tienen el mismo número de elementos.

Demostración. Sea B_1 una base finita de E. Si B_2 es otra base de E, entonces B_2 es linealmente independiente. Aplicando el teorema de Steitnitz con la base $B = B_1$ y $S = B_2$, resulta que B_2 es finito y $\#B_2 \leq \#B_1$. Aplicando de nuevo el teorema de Steitnitz, con $B = B_2$ y $S = B_1$, resulta $\#B_1 \leq \#B_2$, por tanto $\#B_1 = \#B_2$. □

El resultado anterior permite dar la definición de dimensión de un espacio vectorial finitamente generado.

Definición 1.6.2. Si E tiene una base finita, se define su *dimensión* como el cardinal de una base cualquiera de E. El espacio vectorial $\{0\}$ se dice que tiene dimensión 0. Por convenio el conjunto vacío (con 0 elementos) es una base del espacio vectorial $\{0\}$. Si E tiene una base finita se dice que E es de *dimensión finita*.

Corolario 1.6.3. Supongamos que E es un espacio vectorial de dimensión finita.

1. Si F es un subespacio vectorial de E, entonces F tiene una base finita y $\dim F \leq \dim E$.
2. Toda base de F se puede completar a una base de E.
3. Si $\dim F = \dim E$, entonces $F = E$.

Demostración. Sea $n = \dim E$. Supondremos $n \geq 1$.

Veamos (1). Si $F = \{0\}$, entonces $\dim F = 0$. Si $F \neq \{0\}$, existe $v_1 \in F$, con $v_1 \neq 0$. Por tanto $S_1 = \{v_1\}$ es linealmente independiente. Si $\langle S_1 \rangle = F$, entonces S_1 es una base de F y $\dim F = 1 \leq n$. Si $\langle S_1 \rangle \neq F$, existe $v_2 \in F$ tal que $v_2 \notin \langle S_1 \rangle$. Por el lema 1.4.22, $S_2 = S_1 \cup \{v_2\}$ es linealmente independiente. Si $\langle S_2 \rangle = F$, entonces S_2 es una base de F. Repitiendo este proceso obtenemos una sucesión de subconjuntos S_i de F tales que

$$S_1 \subset S_2 \subset \cdots \subset S_r \subset S_{r+1} \subset \cdots$$

Cada S_i tiene i elementos, y es linealmente independiente. Por el teorema de Steitnitz, $i \leq n$, por tanto esta sucesión no se puede prolongar más allá de S_n. Supongamos que la sucesión se termina en S_r. Entonces $\langle S_r \rangle = F$, pues de lo contrario se podría construir S_{r+1}. Por tanto S_r genera F y es linealmente independiente, es decir S_r es una base de F. Además $r \leq n$.

(2) Sea B una base de E. Si S es una base de F, S es un conjunto linealmente independiente y por el teorema de Steitnitz podemos completar S con elementos de B hasta obtener una base B' de E tal que $S \subset B'$.

(3) Si $n = \dim F = \dim E$, y S es una base de F, S tiene n elementos. Completemos S a una base B' de E, tal que $S \subset B'$. Puesto que $\#B' = n$, se tiene $\#S = \#B'$ y por tanto $S = B'$. De aquí se deduce que S genera E. Como S genera F, resulta $E = F$. □

Corolario 1.6.4. Si E tiene dimensión n, y S es un subconjunto de E con n elementos, las siguientes afirmaciones son equivalentes

1. S genera E.
2. S es linealmente independiente.
3. S es una base de E.

Ejercicio 1.6.5. Sea E un espacio vectorial sobre un cuerpo K. Prueba que si, para todo entero $n > 0$ existen n vectores linealmente independientes, entonces E no tiene dimensión finita.

Ejercicio 1.6.6.
1. Prueba que $\mathbb{R}[x]$ no tiene dimensión finita.
2. Prueba que, para cualquier intervalo no degenerado I de $\mathbb{R}$, el espacio vectorial $\mathcal{C}^\infty(I, \mathbb{R})$ no tiene dimensión finita.

1.6.2. **Coordenadas.**

Sea E un espacio vectorial de dimensión finita $n \geq 1$.

Definición 1.6.7. Sea $\mathbf{e} = (e_1, e_2, \ldots, e_n)$ una base ordenada de E, y $u \in E$. La única familia de escalares $(\lambda^1, \lambda^2, \ldots, \lambda^n)$ tal que

$$u = \lambda^1 e_1 + \lambda^2 e_2 + \cdots + \lambda^n e_n,$$

se llama *las coordenadas de u en la base* $\mathbf{e}$. La matriz de una sola columna y n filas formada con estas coordenadas

$$M_{\mathbf{e}}(u) = \begin{pmatrix} \lambda^1 \\ \lambda^2 \\ \vdots \\ \lambda^n \end{pmatrix} \in \mathbf{K}(n,1),$$

se llama el *vector de coordenadas* de u en la base $\mathbf{e}$.

Las coordenadas de un vector $u \in E$ se pueden considerar como una etiqueta $(\lambda^1, \lambda^2, \dots, \lambda^n) \in K^n$ que asignamos a u.

Las propiedades básicas de las coordenadas de un vector con sus coordenadas en una base fijada son las siguientes:

1. Biyectividad. Cada vector tiene una sola n-pla de coordenadas, y dos vectores son iguales si y solo si tienen las mismas coordenadas.
2. Linealidad. Las coordenadas del vector 0 es la n-pla $(0, 0, , \dots, 0)$ (en cualquier base). Las coordenadas de la suma de dos vectores se obtienen sumando las coordenadas de cada uno de los vectores. Las coordenadas de un vector v que es producto de un escalar λ por un vector u, es decir $v = \lambda \cdot u$, se obtienen multiplicando por λ las coordenadas de u.

La biyectividad nos permite identificar un vector con sus coordenadas, y la linealidad nos garantiza que podemos operar con las coordenadas igual que lo haríamos con los vectores.

Si cambiamos de base, las coordenadas de un vector cambian. Es por esto que la identificación entre vectores y sus coordenadas está sujeta a limitaciones y no se puede aplicar de forma arbitraria.

Con esta identificación podemos trabajar en E de la misma manera que lo haríamos en K^n. En particular obtendremos criterios de dependencia lineal de vectores de E expresados en términinos de sus coordenadas.

Observación 1.6.8. Por definición se cumple:

$$u = \begin{pmatrix} e_1 & e_2 & \dots & e_n \end{pmatrix} \cdot \begin{pmatrix} \lambda^1 \\ \lambda^2 \\ \vdots \\ \lambda^n \end{pmatrix},$$

que abreviaremos en la forma $\quad u = \mathbf{e} \cdot M_{\mathbf{e}}(u)$.

El siguiente corolario es un ejemplo fundamental de isomorfismo, el que permite en cierto modo identificar, cuando se fija una base, un espacio vectorial de dimensión finita con el espacio numérico $\mathbf{K}^n$, estudiado en el curso de Matrius i Vectors.

Corolario 1.6.9. Sea $\mathbf{u} = (u_1, u_2, \dots, u_n)$ una base ordenada de E. Sea

$$L_{\mathbf{u}} : \mathbf{K}^n \longrightarrow E, \quad L_{\mathbf{u}}(X) = \mathbf{u} \cdot X = u_1 \cdot X^1 + \cdots + u_n \cdot X^n, \quad \forall X = \begin{pmatrix} X^1 \\ \vdots \\ X^n \end{pmatrix} \in \mathbf{K}^n,$$

y sea

$$M_{\mathbf{u}} : E \longrightarrow \mathbf{K}^n,$$

la aplicación de coordenadas asociada a la base $\mathbf{u}$. Sea $\{e_i\}_{1 \le i \le n}$ la base canónica de $\mathbf{K}^n$.

1. La aplicación $L_{\mathbf{u}} : \mathbf{K}^n \longrightarrow E$ es lineal, es decir,
$$\mathbf{u} \cdot 0 = 0, \quad \mathbf{u} \cdot (X + X') = \mathbf{u} \cdot X + \mathbf{u} \cdot X', \quad \mathbf{u} \cdot (a \cdot X) = a \cdot (\mathbf{u} \cdot X),$$
es biyectiva, y cumple $L_{\mathbf{u}}(e_i) = u_i$, para todo i.
2. La inversa de $L_{\mathbf{u}}$ es $M_{\mathbf{u}} : E \longrightarrow \mathbf{K}^n$. En particular $M_{\mathbf{u}}$ es lineal, es decir,
$$M_{\mathbf{u}}(0) = 0, \quad M_{\mathbf{u}}(x + y) = M_{\mathbf{u}}(x) + M_{\mathbf{u}}(y), \quad M_{\mathbf{u}}(a \cdot x) = a \cdot M_{\mathbf{u}}(x),$$
es biyectiva, y cumple $M_{\mathbf{u}}(u_i) = e_i$, para todo i.

Demostración. Puesto que
$$M_{\mathbf{u}}(\mathbf{u} \cdot X) = X, \quad \mathbf{u} \cdot M_{\mathbf{u}}(x) = x,$$
las aplicaciones $M_{\mathbf{u}}$ y $L_{\mathbf{u}}$ son inversas una de otra, y por tanto ambas son biyectivas. La linealidad de $L_{\mathbf{u}}$ se sigue de las propiedades del producto de matrices. La linealidad de $M_{\mathbf{u}}$ se sigue de 1.3.9. $\square$

Ejemplo 1.6.10. Sean $E = \mathcal{C}^\infty(\mathbb{R})$, $w > 0$ un número real, y
$$F = \{f \in E;\ f'' + w^2 f = 0,\ \}.$$
Sabemos (ver Ejercicio 1.5.11, (2)) que F es un subespacio vectorial de dimensión 2 de $\mathcal{C}^\infty(\mathbb{R})$, y que tiene una base formada por las funciones $e_1 = \cos wx, e_2 = \sin wx$.

Sean $\mathbf{e} = (e_1, e_2)$ y $f \in E$. ¿Cuáles son las coordenadas de f en la base $\mathbf{e}$?

Para responder a esta cuestión, supongamos que (c_1, c_2) son las coordenadas que buscamos, es decir
$$f(x) = c_1 \cos wx\ + c_2 \sin wx.$$
Entonces $f(0) = c_1$.

Por otra parte
$$f'(x) = -wc_1 \sin wx\ + wc_2 \cos wx\,,$$
por tanto $f'(0) = wc_2$. Deducimos que
$$c_1 = f(0), \quad c_2 = \frac{f'(0)}{w},$$
es decir,
$$M_{\mathbf{e}}(f) = \begin{pmatrix} f(0) \\ \frac{f'(0)}{w} \end{pmatrix}$$

Corolario 1.6.11. Sea $\mathbf{e} = (e_1, e_2, \ldots, e_n)$ una base ordenada de E. Dados $u_1, u_2, \ldots, u_r, v$ vectores de E, y $a_1, a_2, \ldots, a_r$ escalares, entonces
$$v = a_1 u_1 + a_2 u_2 + \cdots + a_r u_r \iff M_{\mathbf{e}}(v) = a_1 M_{\mathbf{e}}(u_1) + a_2 M_{\mathbf{e}}(u_2) + \cdots + a_r M_{\mathbf{e}}(u_r)$$

Definición 1.6.12. Sea $\mathbf{e} = (e_1, e_2, \ldots, e_n)$ una base ordenada de E. Si $(u_1, u_2, \ldots, u_m)$ es una familia finita de vectores de E, la matriz
$$M_{\mathbf{e}}(u_1, \ldots, u_m) = \begin{pmatrix} \lambda_1^1 & \lambda_2^1 & \cdots & \lambda_i^1 & \cdots & \lambda_m^1 \\ \lambda_1^2 & \lambda_2^2 & \cdots & \lambda_i^2 & \cdots & \lambda_m^2 \\ \vdots & & \vdots & & \vdots & \\ \lambda_1^n & \lambda_2^n & \cdots & \lambda_i^n & \cdots & \lambda_m^n \end{pmatrix} \in K(n, m),$$
cuya columna i-ésima es el vector de coordenadas de u_i en la base $\mathbf{e}$, se llama matriz de coordenadas de $(u_1, u_2, \ldots, u_m)$ en la base $\mathbf{e}$.

Observación 1.6.13. Se tiene la ecuación

$$\begin{pmatrix} u_1 & u_2 & \cdots & u_m \end{pmatrix} = \mathbf{e} \cdot M_{\mathbf{e}}(u_1, u_2, \ldots, u_m)$$

Si $U = \begin{pmatrix} u_1 & u_2 & \cdots & u_m \end{pmatrix}$, esta ecuación se puede abreviar en la forma

$$U = \mathbf{e} \cdot M_{\mathbf{e}}(U).$$

Corolario 1.6.14. Sean $(u_1, \ldots, u_r)$ una familia de r vectores de E, $\mathbf{e}$ una base ordenada de E, $v \in E$.

1. Criterio de dimensión: $\dim\langle\{u_1, \ldots, u_r\}\rangle = \text{rg } M_{\mathbf{e}}(u_1, \ldots, u_r)$.
2. Criterio de dependencia lineal en coordenadas (ecuación de un subespacio)

$$v \in \langle\{u_1, \ldots, u_r\}\rangle \iff \text{rg } M_{\mathbf{e}}(u_1, \ldots, u_r, v) = \text{rg } M_{\mathbf{e}}(u_1, \ldots, u_r).$$

3. Criterio para independencia lineal:

$$\{u_1, \ldots, u_r\} \text{ es linealmente independiente} \iff \text{rg } M_{\mathbf{e}}(u_1, \ldots, u_r) = r.$$

4. Criterio para generar: Si $\dim E = n$, entonces

$$\{u_1, \ldots, u_r\} \text{ genera } E \iff \text{rg } M_{\mathbf{e}}(u_1, \ldots, u_r) = n.$$

5. Criterio para base: Si $\dim E = n$, entonces

$$\{u_1, \ldots, u_n\} \text{ es base de } E \iff \text{rg } M_{\mathbf{e}}(u_1, \ldots, u_n) = n.$$

Ejemplo 1.6.15. Si $u_1 = (1, 0, 2), u_2 = (0, 1, 0), u_3 = (1, 1, 2)$, y $F = \langle\{u_1, u_2, u_3\}\rangle$, entonces

$$\text{rg } (u_1, u_2, u_3) = \text{rg } (u_1, u_2) = 2$$

y $v = (x, y, z) \in F$ si y solo si

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

es decir, $2x - z = 0$.

A continuación estudiamos las ecuaciones de las coordenadas de los vectores de un subespacio del cual se conoce una base, o más generalmente, un conjunto generador.

Proposición 1.6.16. (Ecuaciones de un subespacio.) Sean E un espacio vectorial de dimensión finita, $\mathbf{e} = (e_1, \ldots, e_n)$ una base de E, y F un subespacio de E de dimensión d generado por $S = \{u_1, \cdots, u_r\}$. Si $v \in E$ entonces

$$v \in F \iff \text{rg } M_{\mathbf{e}}(u_1, \cdots, u_r, v) = d.$$

Observación 1.6.17. En general la noción de coordenadas en el sentido definido en esta sección coincide el utilizado en 1.2.3 para los espacios de dimensión finita, y también para el espacio de los polinomios. Sin embargo, para una sucesión de números reales dijimos que sus coordenadas eran los términos de la sucesión, pero las coordenadas de un vector en una base siempre son un número finito de escalares. Así, una sucesión no se puede expresar como una combinación lineal de vectores de una base usando sus coordenadas, porque estas coordenadas son en general un número infinito de escalares. Algo similar sucede con la función exponencial y su desarrollo de Taylor, $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$. En este curso nos limitaremos a hablar de coordenadas en el caso de espacios vectoriales de dimensión finita.

1.7. Suma de subespacios.

En lo que sigue E denota un espacio vectorial sobre K.

En general la reunión de subespacios no es un subespacio. Sin embargo podemos introducir una operación de suma entre los subespacios. Esta operación no debe confundirse con la suma de vectores, aunque está directamente relacionada. Es una operación entre conjuntos y cumple algunas de las reglas de la suma de vectores, pero no todas.

1.7.1. Suma de subespacios.

Definición 1.7.1. Si F y G son subespacios vectoriales de E, se define su *suma* por

$$F + G := \{u + v; u \in F, v \in G\}.$$

En general, la suma de una familia finita de subespacios $\{F_i\}_{1\leq i\leq r}$ de E es

$$F_1 + \cdots + F_r := \left\{ \sum_{i=1}^{r} u_i \in E;\ u_1 \in F_1,\ \ldots, u_r \in F_r \right\}.$$

Esta suma también se denota por $\sum\limits_{1\leq i\leq r} F_i$.

Proposición 1.7.2. Sean F y G dos subespacios vectoriales de E. Si $F = \langle S\rangle$ y $G = \langle T\rangle$ entonces $F + G = \langle S \cup T\rangle$. En particular $F + G$ es un subespacio vectorial de E.

Sea $\{F_i\}_{1\leq i\leq r}$ una familia finita de subespacios de E. Si $F_i = \langle S_i\rangle$, para todo i, entonces

$$\sum_{1\leq i\leq r} F_i := \left\langle \bigcup_{1\leq i\leq r} S_i \right\rangle.$$

En particular $\sum\limits_{1\leq i\leq r} F_i$ es un subespacio vectorial de E.

Demostración. Se deja como ejercicio. □

Ejemplo 1.7.3. Sean $F = \langle\{e_1, e_2\}\rangle$, $G = \langle\{e_1+e_3, e_2+e_3\}\rangle$ subespacios de $\mathbb{R}^3$ Consideremos el subespacio suma es $F + G = \langle\{e_1, e_2, e_1 + e_3, e_2 + e_3\}\rangle$. Puesto que el rango de

$$M_{\mathbf{e}}(e_1, e_2, e_1 + e_3, e_2 + e_3) = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

es 3, el subespacio $F + G$ tiene dimensión 3, por tanto $F + G = \mathbb{R}^3$.

Ejemplo 1.7.4. Sean F un subespacio vectorial de E y $S = \{u_1, \cdots, u_r\}$ un conjunto de vectores de E. Entonces S es un conjunto generador de F si, y solo si, $F = \sum\limits_{1\leq i\leq r} \langle\{u_i\}\rangle$.

1.7.2. Suma directa de subespacios.

El concepto de vectores linealmente independientes tiene una traducción para subespacios, la suma directa.

Definición 1.7.5.

1. Sean F y G subespacios de E. La suma $F + G$ se llama *suma directa* si todo vector de la suma se expresa de forma única como suma de un vector de F y uno de G. Esto es equivalente a que si $0 = u + v$, con $u \in F$, $v \in G$, entonces $u = v = 0$. En tal caso la suma se denota por $F \oplus G$.

2. Sea $\{F_i\}_{i\in I}$ una familia (finita o infinita) de subespacios vectoriales diferentes de E. Su suma $\sum\limits_{i\in I} F_i$ se llama directa, y se denota por $\bigoplus\limits_{i\in I} F_i$, si todo vector de la suma se expresa de forma única como suma de vectores de los F_i. Esto equivale a la siguiente aserción: Si 0 se expresa como una suma de un número finito de vectores de los diferentes subespacios F_i, $0 = \sum_{j=1}^r u_j$, con $u_j \in F_j$, entonces $u_j = 0, \forall j$.

La siguiente proposición nos da un criterio para obtener sumas directas. Es el análogo del lema 1.4.22.

Proposición 1.7.6. Sean F, G subespacios vectoriales de E.

1. La suma $F + G$ es directa si y solo si $F \cap G = \{\mathbf{0}\}$.
2. Sean $F_1, \dots, F_r$ subespacios vectoriales de E. La suma $F_1 + \cdots + F_r$, es directa si, y solo si, $(F_1 \oplus \cdots \oplus F_i) \cap F_{i+1} = \{\mathbf{0}\}$, para todo $i = 1, \dots, r-1$.

Demostración. Se deja como ejercicio. □

Ejemplo 1.7.7.

1. $\mathbb{R}^3 = \langle\{e_1\}\rangle \oplus \langle\{e_2, e_3\}\rangle$,
2. Sea $\{u_1, \dots, u_r\}$ un conjunto de vectores de E. Entonces $\{u_1, \dots, u_r\}$ es linealmente independiente si y solo si la suma $\sum\limits_{1\le i\le r} \langle\{u_i\}\rangle$ es directa.
3. Si E es un espacio vectorial y $B = \{e_1, \dots, e_n\}$ un conjunto de n vectores de E. Entonces B es una base de E si y solo si $E = \langle\{e_1\}\rangle \oplus \cdots \oplus \langle\{e_n\}\rangle$.

Ejercicio 1.7.8. Si F es un subespacio vectorial de E, existe un subespacio G tal que $E = F \oplus G$. (Se dice que G es un *suplemento* de F.) (Indicación: Si E tiene una base finita, aplicando el teorema de Steitnitz podemos completar una base B_F de F a una base $B = B_F \sqcup B_1$ de E. Entonces, el subespacio G generado por B_1 es un suplemento de F. En el caso en que E sea de dimensión infinita, usamos 1.5.18 para ampliar una base de F a una base de E y concluimos de la misma manera.)

Ejemplo 1.7.9. Sea $E = \mathcal{C}^\infty(\mathbb{R})$. Sean $P(X)$ y $Q(X)$ dos polinomios primos entre sí. Consideremos las aplicaciones lineales $P(D), Q(D)$ y $P(D)\circ Q(D) = Q(D)\circ P(D)$ de E en E. Entonces los núcleos de $P(D)$ y $Q(D)$ forman suma directa y además

$$\operatorname{Ker} P(D) \oplus \operatorname{Ker} Q(D) = \operatorname{Ker} P(D)\circ Q(D).$$

En efecto, por la identidad de Bezout, existen polinomios $A(X)$ y $B(X)$ tales que

$$1 = A(X)\cdot P(X) + B(X)\cdot Q(X).$$

Por tanto, para todo $v \in E$ se cumple

$$v = A(D)P(D)v + B(D)Q(D)v.$$

Si $v \in \operatorname{Ker} P(D) \cap \operatorname{Ker} Q(D)$, entonces

$$v = A(D)P(D)v + B(D)Q(D)v = 0,$$

por tanto la suma $\operatorname{Ker} P(D) + \operatorname{Ker} Q(D)$ es directa. Por otra parte

$$\ker P(D) \subset \operatorname{Ker} Q(D)\circ P(D), \quad \ker Q(D) \subset \operatorname{Ker} P(D)\circ Q(D),$$

por tanto

$$\operatorname{Ker} P(D) \oplus \operatorname{Ker} Q(D) \subset \operatorname{Ker} P(D)\circ Q(D).$$

Para ver la inclusión inversa observemos que, si $v \in \operatorname{Ker} P(D)\circ Q(D)$, entonces $A(D)P(D)v \in \operatorname{Ker} Q(D)$ y $B(D)Q(D)v \in \operatorname{Ker} P(D)$. Por tanto $v \in \operatorname{Ker} P(D) \oplus Q(D)$.

Ejercicio 1.7.10. Sea $E = \mathcal{C}^\infty(\mathbb{R})$. Prueba que, en cada caso, B es una base del núcleo de $a(D)$, donde $a(t)$ es el polinomio indicado.

1. $a(t) = (t-\alpha)(t-\beta)$, con $\alpha, \beta \in \mathbb{R}$ distintos, $B = \{e^{\alpha x}, e^{\beta x}\}$.
2. $a(t) = (t-\alpha)^2(t-\beta)$, con $\alpha, \beta \in \mathbb{R}$ distintos, $B = \{e^{\alpha x}, xe^{\alpha x}, e^{\beta x}\}$.
3. $a(t) = (t-\alpha)^2(t-\beta)^2$, con $\alpha, \beta \in \mathbb{R}$ distintos, $B = \{e^{\alpha x}, xe^{\alpha x}, e^{\beta x}, xe^{\beta x}\}$.
4. $a(t) = (t+\alpha^2)(t+\beta^2)$, $\alpha, \beta \in \mathbb{R}$, $\alpha, \beta > 0$ y $\alpha \neq \beta$, $B = \{\sin \alpha x, \cos \alpha x, \sin \beta x, \cos \beta x\}$.

1.7.3. **Fórmula de Grassmann.**

La fórmula de Grassmann relaciona las dimensiones de dos subespacios, la de su suma y la de su intersección. Esta fórmula es una versión geométrica de la fórmula siguiente: Sea X un conjunto, si A y B son dos subconjuntos finitos de X, entonces $A \cup B$ y $A \cap B$ son finitos y

$$\#(A \cup B) = \#A + \#B - \#(A \cap B).$$

Teorema 1.7.11. (Fórmula de Grassmann) Si F, G son subespacios vectoriales de dimensión finita de un espacio vectorial E, entonces $F + G$ es de dimensión finita y

$$\dim(F+G) = \dim F + \dim G - \dim(F \cap G)$$

Demostración. Sea $B_0 = \{u_1, \cdots, u_r\}$ una base de $F \cap G$. Por el teorema de Steitnitz, ampliamos B_0 a una base $B_1 = \{u_1, \cdots, u_r, v_1, \cdots, v_s\}$ de F y a una base $B_2 = \{u_1, \cdots, u_r, w_1, \cdots, w_t\}$ de G. Entonces $B = B_1 \cup B_2$ es una base de $F + G$. En efecto, B genera $F + G$ por la proposición 1.7.2. Veamos que $B = \{u_1, \cdots, u_r, v_1, \cdots, v_s, w_1, \cdots, w_t\}$ es linealmente independiente. Sean λ_i, ϕ_j, ν_k, con $1 \leq i \leq r, 1 \leq j \leq s, 1 \leq k \leq t$ escalares tales que

$$\sum_i \lambda_i u_i + \sum_j \phi_j v_j + \sum_k \nu_k w_k = 0,$$

entonces

$$\sum_i \lambda_i u_i + \sum \phi_j v_j = -\sum_k \nu_k w_k \in F \cap G.$$

Por tanto existen $\alpha_i \in K$, con $1 \leq i \leq r$, tales que

$$-\sum_k \nu_k w_k = \sum_i \alpha_i u_i,$$

es decir

$$\sum_i \alpha_i u_i + \sum_k \nu_k w_k = 0.$$

Como B_2 es una base de G, B_2 es linealmente independiente, por tanto $\alpha_i, \nu_k = 0$ para todo i, k. En consecuencia,

$$\sum_i \lambda_i u_i + \sum \phi_j v_j = -\sum_k \nu_k w_k = 0.$$

Usando ahora que B_1 es linealmente independiente resulta $\lambda_i, \phi_j = 0$, para todo i, j. Por tanto B es linealmente independiente.

Por otra parte,

$$\#B = \#(B_1 \cup B_2) = \#B_1 + \#B_2 - \#(B_1 \cap B_2) = \#B_1 + \#B_2 - \#B_0 = r + s + t.$$

Por tanto

$$\dim(F+G) = r+s+t = (r+s)+(r+t)-r = \dim F + \dim G - \dim(F \cap G)$$

□

Ejemplo 1.7.12. Sean $F = \langle\{e_1, e_2\}\rangle$, $G = \langle\{e_1+e_3, e_2+e_3\}\rangle$ subespacios de $\mathbb{R}^3$. Hemos visto en 1,7,3 que $F + G = \mathbb{R}^3$. Usando la fórmula de Grassmann deducimos que $\dim F \cap G = 1$.

1.8. Espacio producto y espacio cociente.

1.8.1. Producto de dos espacios vectoriales.

Recordemos que el producto cartesiano de dos conjuntos E y F es el conjunto de pares ordenados

$$E \times F = \{(x, y);\ x \in E, y \in F\ \}.$$

Cuando E y F son $\mathbf{K}$-espacios vectoriales hay una manera natural de dotar a $E \times F$ de una estructura de $\mathbf{K}$-espacio vectorial.

Definimos la suma en $E \times F$ por

$$(x_1, y_1) + (x_2, y_2) := (x_1 + x_2, y_1 + y_2),\ \forall (x_1, y_1), (x_2, y_2) \in E \times F,$$

y el producto por escalares por

$$\lambda \cdot (x, y) := (\lambda \cdot x, \lambda \cdot y),\ \forall \lambda \in \mathbf{K}, (x, y) \in E \times F.$$

Proposición 1.8.1. Sean E y F dos espacios vectoriales sobre $\mathbf{K}$. Con la suma y el producto por escalares definidos más arriba, $E \times F$ es un $\mathbf{K}$-espacio vectorial, llamado espacio vectorial producto de E y F. (También se le llama espacio suma.) El elemento neutro es $(0, 0)$ y el opuesto de (x, y) es $(-x, -y)$

Ejemplo 1.8.2. Si $E = \mathbf{K}$, entonces $E \times E = \mathbf{K}^2$. En general, $\mathbf{K}^n = \mathbf{K}^{n-1} \times \mathbf{K}$, para todo $n \geq 2$.

Proposición 1.8.3. Sean E y F dos espacios vectoriales sobre $\mathbf{K}$. Las proyecciones

$$p : E \times F \longrightarrow E,\ (x, y) \mapsto p(x, y) = x; \quad q : E \times F \longrightarrow E,\ (x, y) \mapsto q(x, y) = y$$

son aplicaciones lineales. Además $\operatorname{Ker} p = \{0\} \times F,\ \operatorname{Ker} q = E \times \{0\}$.

Proposición 1.8.4. La aplicación

$$E \longrightarrow E \times F, \quad x \mapsto (x, 0)$$

es lineal e inyectiva, e identifica E con el subespacio $E \times \{0\}$ de $E \times F$. Análogamente, la aplicación

$$E \longrightarrow E \times F, \quad y \mapsto (0, y)$$

es lineal e inyectiva, e identifica F con el subespacio $\{0\} \times F$ de $E \times F$. Finalmente, $E \times F = (E \times \{0\}) \oplus (\{0\} \times F)$.

Proposición 1.8.5. Sean E y F dos espacios vectoriales sobre $\mathbf{K}$. Si E es de dimensión finita m, y F es de dimensión finita n, entonces $E \times F$ es de dimensión finita $m + n$. Más precisamente, si $\mathbf{u} = \{u_1, \cdots, u_m\}$ es una base de E y $\mathbf{v} = \{v_1, \cdots, v_n\}$ es una base de F entonces

$$\mathbf{u} \sqcup \mathbf{v} = \{(u_1, 0), \ldots, (u_m, 0), (0, v_1), \cdots, (0, v_n)\}$$

es una base de $E \times F$.

1.8.2. **Relaciones de equivalencia compatibles.** Sea $R \subset E \times E$ una relación en un conjunto E. Denotaremos la condición $(x, y) \in R$ por $x \sim y$. Recordemos que R es de equivalencia si

1. Es reflexiva, es decir $x \sim x$, para todo $x \in E$.
2. Es simétrica, es decir, dados $x, y \in E$, entonces $x \sim y$ si y solo si $y \sim x$.
3. Es transitiva, es decir, si $x \sim y$, $y \sim z$ entonces $x \sim z$.

Dada una relación de equivalencia $\sim$, la clase de equivalencia de un elemento $x \in E$ es $[x] = \{y \in E;\ x \sim y\}$. El conjunto de las clases de equivalencia se denota por $E/\sim$. Se tiene una aplicación exhaustiva

$$\pi : E \longrightarrow E/\sim, \quad x \mapsto [x],$$

llamada aplicación canónica, tal que $\pi^{-1}\{[x]\} = [x]$, es decir, $\pi(x) = \pi(y)$ si y solo si $x \sim y$.

Los problemas de clasificación son frecuentes dentro y fuera del ámbito de la matemáticas. Clasificar un "universo"de objetos (o de individuos) consiste en separarlos en "clases". Dos clases diferentes no tienen elementos en común, y cualquier objeto del universo está en una de las clases.

Detrás de todos los problemas de clasificación hay una relación de equivalencia. Veamos un ejemplo con el cual todos estais familiarizados. Se trata de una clasificación de los números reales que se utiliza cuando se quiere operar con un objetivo práctico, no formal. Por ejemplo, podemos aproximar un número considerando solo sus dos primeros decimales. Hablamos de la aproximación a la centésima. La relación de equivalencia asociada es la siguiente: dados dos números reales $x, y \in \mathbb{R}$, decimos que x e y están relacionados si tienen iguales la parte entera y los dos primeros decimales. Por ejemplo $\pi \sim\ 3,14$. En particular, la clase de 0 está formada por los números x tales que $|x| < 0,01$. Pero esta forma de proceder tiene algunos inconvenientes con los que os habréis encontrado alguna vez. La aproximación de una combinacion lineal de dos números x_1, x_2 no se obtiene en general como la correspondiente combinación lineal de sus aproximaciones. Por ejemplo, $\sqrt{3} \sim 1,73$, $\sqrt{2} \sim 1,41$, y $3 * \sqrt{3} + 2 * \sqrt{2} \sim 8,02$, pero $3 * 1,73 + 2 * 1,41 = 8,01$, es decir

$$3 * \sqrt{3} + 2 * \sqrt{2} \nsim 3 * 1,73 + 2 * 1,41.$$

Otro ejemplo más apropiado de relación de equivalencia, que se comporta bien respecto de las combinaciones lineales, lo habéis trabajado en el terreno de la aproximación de funciones. Cuando se aproxima una función suficientemente derivable $f(x)$ por su polinomio de Taylor en el entorno de $x_0 = 0$, hasta un cierto orden, pongamos hasta orden 1 para simplificar, se tiene

$$f(x) = f(0) + f'(0)x + R_1(x),$$

donde $\lim\limits_{x \to 0} \frac{R_1(x)}{x} = 0$. Entonces, "despreciando" el resto R_1, aproximamos $f(x)$ por

$$f(x) \backsimeq f(0) + f'(0)x,$$

si x está suficientemente próximo a 0. Esta manera de aproximar se basa también en una relación de equivalencia

$$\text{"}f \backsimeq g \text{ si y solo si } \lim_{x \to 0} \tfrac{f(x) - g(x)}{x} = 0.\text{"}$$

La clase de la función cero se denota por $o(x)$ (o pequeña de x). En este caso la relación de equivalencia se comporta bien respecto de las operaciones de suma y de producto, y por ello es lícito operar con las clases de equivalencia. A continuación vamos a ver como establecer esta propiedad en abstracto.

En lo que sigue E denota un espacio vectorial sobre K.

Definición 1.8.6. Una relación de equivalencia R en E (que denotaremos por $\sim$) se dice que es *compatible* con la estructura de espacio vectorial si

$$\begin{aligned} u_1 \sim v_1, u_2 \sim v_2 &\implies u_1 + u_2 \sim v_1 + v_2, \\ u \sim v, \lambda \in K &\implies \lambda u \sim \lambda v \end{aligned}$$

En la proposición siguiente estudiamos como son las relaciones de equivalencia compatibles y vemos que siempre están asociadas a un subespacio vectorial.

Proposición 1.8.7. Sea E un espacio vectorial sobre K.

1. Sea $\sim$ una relación de equivalencia en E, compatible con la estructura de espacio vectorial. Si
$$F = [0] = \{u;\ u \sim 0\ \},$$
entonces
a) F es un subespacio vectorial de E.
b) Para todo $u \in E$,
$$[u] = \{u + v; v \in F\} = u + F.$$
c) Se cumple
$$u \sim v \iff u - v \sim 0 \Leftrightarrow u - v \in F.$$
2. Recíprocamente, sea F un subespacio vectorial de E. La relación en E definida por
$$u \sim v \iff u - v \in F$$
cumple
a) Es una relación de equivalencia.
b) Es compatible con la estructura de espacio vectorial.
c) $[0] = F$.

Demostración. Se deja como ejercicio. □

Ejemplo 1.8.8. En $E = \mathbb{R}^3$ consideremos la relación $(x_1, x_2, x_3) \sim (y_1, y_2, y_3)$ si y solo si $x_1 = y_1,\ x_2 = y_2$. Es fácil ver que $\sim$ es una relación de equivalencia en E, que es compatible con la estructura de espacio vectorial. En este caso $F = \{(0, 0, x_3);\ x_3 \in \mathbb{R}\ \}$.

1.8.3. **Espacios cociente. Dimensión de un cociente.** El cociente de un espacio vectorial por una relación de equivalencia compatible con la estructura de espacio vectorial hereda la estructura de espacio vectorial.

Proposición 1.8.9. Dada una relación de equivalencia $\sim$ en E compatible con la estructura de espacio vectorial, el conjunto cociente $E/\sim$ es un espacio vectorial con las siguientes definiciones

1. Neutro: $0 := [0]$.
2. Suma: $[u] + [v] := [u + v]$.
3. Opuesto: $-[u] := [-u]$.
4. Producto por escalares: $\lambda[u] := [\lambda u]$.

Además, la proyección
$$\pi : E \longrightarrow E/\sim, \quad x \mapsto \pi(x) := [x]$$
es una aplicación lineal, y $\operatorname{Ker} \pi = F$.

Notación 1.8.10. El espacio vectorial $E/\sim$ se llama *espacio cociente* de E por la relación de equivalencia $\sim$, o cociente de E por F, y se denota también por E/F, donde $F = [0]$.

Demostración. Lo primero que se ha de comprobar es que las definiciones son consistentes, para lo cual es imprecindible que $\sim$ sea compatible. En efecto, para la suma hemos de comprobar que si $[u] = [u']$, $[v] = [v']$ entonces $[u + v] = [u' + v']$. Es decir, que $u \sim u'$, $v \sim v'$ implica $u + v \sim u' + v'$, lo cual es cierto pues $\sim$ es compatible con la suma. De la misma manera se comprueba que si $[u] = [u']$ entonces $[\lambda \cdot u] = [\lambda \cdot u']$. De aquí deducimos, para $\lambda = -1$, que si $[u] = [u']$, entonces $[-u] = [-u']$. Con lo cual las operaciones están bien definidas. La comprobación de las propiedades de espacio vectorial es un mero trámite que dejamos como ejercicio. □

En el espacio cociente E/F a veces es posible escoger representantes especiales de las clases de equivalencia. Así en el ejemplo 1.8.8, podemos tomar como representante especial de la clase $[(x_1, x_2, x_3)]$ el vector $(x_1, x_2, 0)$. De esta manera es más fácil visualizar el espacio cociente, y también sus operaciones. De hecho, el siguiente teorema demuestra que esto se puede hacer en general.

Teorema 1.8.11. Sea F un subespacio vectorial de E.

1. Si $E = F \oplus G$ entonces la aplicación $\pi_G : G \longrightarrow E/F, \quad x \mapsto [x]$ es lineal y biyectiva.
2. Si E es de dimensión finita, $\{u_1, \ldots, u_r\}$ es una base de F y $\{v_1, \ldots, v_s\}$ es una base de G, entonces $\{[v_1], \ldots, [v_s]\}$ es una base de E/F.
3. Si E es de dimensión finita, entonces E/F es de dimensión finita y
$$\dim E/F = \dim E - \dim F.$$

Demostración. La segunda y la tercera aserciones son consecuencias de la primera. Veamos pues la primera. La aplicación π_G es lineal, pues es la composición de la inclusión $G \longrightarrow E$ y de la proyección canónica $\pi : E \longrightarrow E/F$. Veamos que es inyectiva. Sean $x, y \in G$. Si $\pi(x) = \pi(y)$ entonces $x \sim y$, es decir $x - y \in F$. Como $x - y \in F \cap G = \{0\}$, resulta $x = y$, luego es inyectiva. Finalmente veamos que pi_G es exhaustiva. Dado $[x] \in E/F$, expresamos $x = u + v$, con $u \in F, v \in G$. Entonces $[x] = [v] = \pi_G(v)$, por tanto π_G es exhaustiva. $\square$

Ejemplo 1.8.12. En el ejemplo 1.8.8, $\mathbb{R}^3$ se expresa como suma directa de $G = \langle\{e_1, e_2\}\rangle$ y $F = \langle\{e_3\}$ por tanto las clases de equivalencia de e_1 y e_2 definen una base $\{[e_1], [e_2]\}$ de E/F. Por ejemplo, puesto que el vector $(1, 2, 3)$ se expresa en la forma $(1, 2, 3) = u_1 + 2u_2 + 3e_3$, su clase se expresa en la forma $[(1, 2, 3)] = [e_1] + 2[e_2]$.

2. Aplicaciones lineales

En lo que sigue **K** denota un cuerpo.

2.1. Matriz de una aplicación lineal.

2.1.1. Aplicaciones lineales.

Recordemos la definición de aplicación lineal dada en el capítulo 1.

Definición 2.1.1. Sean E, F dos espacios vectoriales sobre **K**. Una aplicación $f : E \longrightarrow F$ se llama *aplicación lineal* si

1. Para todo $u, v \in E$,
$$f(u+v) = f(u) + f(v).$$
2. Para todo $u \in E, \lambda \in K$.
$$f(\lambda u) = \lambda f(u).$$

El ejemplo básico de aplicación lineal es la multiplicación por una matriz.

Ejemplo 2.1.2. Aplicación lineal definida por una matriz. Identificamos $\mathbf{K}^n = \mathbf{K}(n, 1)$.

Sea $A \in \mathbf{K}(n, m)$, entonces
$$L_A : \mathbf{K}^m \longrightarrow \mathbf{K}^n, \quad L_A(X) = A \cdot X$$
es una aplicación lineal. En efecto, las propiedades
$$A \cdot (X + Y) = A \cdot X + A \cdot Y, \quad A \cdot (\lambda \cdot X) = \lambda \cdot (A \cdot X)$$
se siguen de las propiedades de las operaciones con matrices. Observemos que, si $\{e_i\}_{1 \leq i \leq n}$ es la base canónica de $\mathbf{K}^n$, y A_i denota la i-ésima columna de la matriz A, entonces
$$L_A(e_i) = A_i.$$

Veamos un ejemplo concreto.

Ejemplo 2.1.3. Sea
$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}.$$
La aplicación $L_A : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ está definida por
$$L_A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \cdot x + 2 \cdot y + 3 \cdot z \\ 4 \cdot x + 5 \cdot y + 6 \cdot z \end{pmatrix}.$$
En particular
$$L_A \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 4 \end{pmatrix} \; ; \; L_A \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 5 \end{pmatrix} \; ; \; L_A \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 6 \end{pmatrix}.$$

Ejercicio 2.1.4. Sea B una matriz $m \times n$ con coeficientes en **K**. Determina cuáles de las siguientes aplicaciones son lineales:

(i) $\mathrm{tr} : \mathbf{K}(n, n) \to \mathbf{K}$, tal que $\mathrm{tr}\,(A) :=$ traza de A.
(ii) $\det : \mathbf{K}(n, n) \to \mathbf{K}$, tal que $\det(A) :=$ determinante de A.
(iii) $R_B : \mathbf{K}(p, m) \to \mathbf{K}(p, n)$, tal que $R_B(A) := AB$, para todo $A \in \mathbf{K}(p, m)$.
(iv) $L_B : \mathbf{K}(n, p) \to \mathbf{K}(m, p)$, tal que $L_B(A) := BA$, para todo $A \in \mathbf{K}(n, p)$.

Puesto que las aplicaciones lineales preservan la suma y el producto por escalares también preservan las combinaciones lineales.

Proposición 2.1.5. Sea $f : E \longrightarrow F$ una aplicación lineal. Si $x = \lambda_1 \cdot u_1 + \cdots + \lambda_r \cdot u_r \in E$ es una combinación lineal, entonces

$$f(x) = \lambda_1 \cdot f(u_1) + \cdots + \lambda_r \cdot f(u_r).$$

En particular, $f(0) = 0$ y $f(-u) = -f(u)$.

Demostración. Es una consecuencia inmediata de las definiciones y se deja como ejercicio. □

Debido a la propiedad anterior, toda aplicación lineal queda determinada por su valor sobre una base. Dicho valor, por otra parte, puede prefijarse de forma arbitraria.

Proposición 2.1.6. Sea E y F dos espacios vectoriales. Supongamos que $\mathbf{u} = \{u_i\}_{1 \le i \le n}$ es una base de E.

1. Sean $f : E \longrightarrow F$ una aplicación lineal y $w_i := f(u_i)$, para todo i. Sean $x \in E$, y $(\lambda_1, \ldots, \lambda_n) \in \mathbf{K}^n$ las coordenadas de x en la base $\mathbf{u}$. Entonces
$$f(x) = \sum_i \lambda_i w_i.$$
2. Para toda familia $\{w_i\}_{1 \le i \le n}$ de vectores de F existe una única aplicación lineal $f : E \longrightarrow F$ tal que $f(u_i) = w_i$, para todo i. La aplicación lineal f está definida por
$$f\left(\sum_i \lambda_i u_i\right) := \sum_i \lambda_i w_i$$

Demostración. (1) se sigue de la proposición 2.1.5, teniendo en cuenta que $x = \sum_i \lambda_i u_i$, y $w_i = f(u_i)$. La unicidad de la parte (2) se sigue de la parte (1). Para comprobar la existencia de f basta ver que la fórmula del enunciado define efectivamente una aplicación lineal, lo cual se deja como ejercicio. □

Veamos como se manifiesta la propiedad anterior en un ejemplo concreto.

Ejemplo 2.1.7. Volviendo al ejemplo 2.1.3 anterior observamos $\mathbf{u} = (e_1, e_2, e_3)$ es una base de $\mathbf{K}^3$, y

$$w_1 = \begin{pmatrix} 1 \\ 4 \end{pmatrix},\ w_2 = \begin{pmatrix} 2 \\ 5 \end{pmatrix},\ w_3 = \begin{pmatrix} 3 \\ 6 \end{pmatrix},$$

son tres vectores (cualesquiera) de $\mathbf{K}^2$. La aplicación L_A satisface $L_A(e_i) = w_i$ y, siendo

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = x \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + z \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = xe_1 + ye_2 + ze_3,$$

L_A satisface

$$L_A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = x \begin{pmatrix} 1 \\ 4 \end{pmatrix} + y \begin{pmatrix} 2 \\ 5 \end{pmatrix} + z \begin{pmatrix} 3 \\ 6 \end{pmatrix}.$$

2.1.2. Matriz de una aplicación lineal.

La propiedad 2.1.1 da lugar a una propiedad fundamental de las aplicaciones lineales entre dos espacios vectoriales: dadas bases de los dos espacios vectoriales, toda aplicación lineal determina, y está determinada por, una matriz. La matriz es única fijada las bases, pero cambiará cuando cambiemos las bases, al igual que lo hacen las coordenadas de un vector en una base.

Definición 2.1.8. Sean E y F dos espacios vectoriales sobre $\mathbf{K}$ de dimensión finita, $f : E \longrightarrow F$ una aplicación lineal. Sean $\mathbf{u} = (u_1, \ldots, u_n)$ una base ordenada de E, y $\mathbf{v} = (v_1, \ldots, v_m)$ una base ordenada de F. La *matriz de f en las bases* $(\mathbf{u}, \mathbf{v})$ es la matriz de tipo $m \times n$ cuya columna i-ésima son las coordenadas de $f(u_i)$ en la base $\mathbf{v}$, es decir, la matriz

$$M_{\mathbf{v}}(f(u_1), \ldots, f(u_n)).$$

Denotaremos eventualmente esta matriz por $M_{\mathbf{v}}(f; \mathbf{u})$.

Ejemplo 2.1.9. Sea $E = \mathbb{R}[x]_{\leq 3}$, $F = \left\{ \begin{pmatrix} a & b \\ c & -a \end{pmatrix} ;\ a, b, c \in \mathbb{R} \right\}$ el espacio vectorial de matrices 2×2 de traza nula. Sea $f : E \longrightarrow F$ la aplicación lineal definida por

$$f(p(x)) = \begin{pmatrix} -p(1) & p'(0) \\ \int_0^1 p(x) & p(1) \end{pmatrix}$$

Consideremos las bases $\mathbf{u} = \{1, x, x^2, x^3\}$ de E y $\mathbf{v} = \{\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)\}$ de F. Entonces

$$\begin{aligned} f(1) &= \left(\begin{smallmatrix} -1 & 0 \\ 1 & 1 \end{smallmatrix}\right) = -v_1 + v_3 \\ f(x) &= \left(\begin{smallmatrix} -1 & 1 \\ \frac{1}{2} & 1 \end{smallmatrix}\right) = -v_1 + v_2 + \frac{1}{2} v_3 \\ f(x^2) &= \left(\begin{smallmatrix} -1 & 0 \\ \frac{1}{3} & 1 \end{smallmatrix}\right) = -v_1 + \frac{1}{3} v_3 \\ f(x^3) &= \left(\begin{smallmatrix} -1 & 0 \\ \frac{1}{4} & 1 \end{smallmatrix}\right) = -v_1 + \frac{1}{4} v_3 \end{aligned}$$

por tanto

$$M_{\mathbf{u}}(f; \mathbf{v}) = \begin{pmatrix} -1 & -1 & -1 & -1 \\ 0 & 1 & 0 & 0 \\ 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \end{pmatrix}$$

La matriz de la aplicación lineal permite el cálculo de las imágenes de los vectores utilizando coordenadas por la fórmula siguiente:

Proposición 2.1.10. Con las hipótesis de la definición anterior, para todo $x \in E$,

$$M_{\mathbf{v}}(f(x)) = M_{\mathbf{v}}(f; \mathbf{u}) \cdot M_{\mathbf{u}}(x).$$

Además $M_{\mathbf{v}}(f; \mathbf{u})$ es la única matriz que cumple esta propiedad.

Demostración. Si $x \in E$, y

$$M_{\mathbf{u}}(x) = \begin{pmatrix} X^1 \\ \vdots \\ X^n \end{pmatrix} \in \mathbf{K}^n$$

son sus coordenadas en la base $\mathbf{u}$, por la proposición 2.1.5 se tiene

$$f(x) = \sum_i X^i \cdot f(u_i).$$

Para calcular las coordenadas de $f(x)$ en la base $\mathbf{v}$ usamos la linealdad de $M_{\mathbf{v}}$ (ver Proposición 1.6.9)

$$M_{\mathbf{v}}(f(x)) = \sum_i M_{\mathbf{v}}(f(u_i)) \cdot X^i = \begin{pmatrix} M_{\mathbf{v}}(f(u_1)) & \dots & M_{\mathbf{v}}(f(u_n)) \end{pmatrix} \cdot \begin{pmatrix} X^1 \\ \vdots \\ X^n \end{pmatrix}$$

$$= M_{\mathbf{v}}(f(u_1), \dots, f(u_n)) \cdot \begin{pmatrix} X^1 \\ \vdots \\ X^n \end{pmatrix} = M_{\mathbf{v}}(f; \mathbf{u}) \cdot M_{\mathbf{u}}(x).$$

Supongamos ahora que existiera otra matriz A cumpliendo

$$M_{\mathbf{v}}(f(x)) = A \cdot M_{\mathbf{u}}(x),$$

para todo $x \in E$. Sustituyendo x por u_i resulta

$$A \cdot M_{\mathbf{u}}(u_i) = A \cdot e_i = \text{columna } i\text{-ésima de } A$$

y por otro lado

$$M_{\mathbf{v}}(f(u_i)) = \text{columna } i\text{-ésima de } M_{\mathbf{v}}(f; \mathbf{u})$$

por tanto A y $M_{\mathbf{v}}(f; \mathbf{u})$ tienen las mismas columnas.

$\square$

Observación 2.1.11. Si

$$f(u_i) = \sum_{j=1}^{m} a_i^j v_j = \begin{pmatrix} v_1 & \cdots & v_m \end{pmatrix} \cdot \begin{pmatrix} a_i^1 \\ \vdots \\ a_i^m \end{pmatrix}, \quad \forall\, i,\ 1 \le i \le n,$$

entonces la matriz de f en las bases $(\mathbf{u}, \mathbf{v})$ es la matriz

$$A = \begin{pmatrix} a_1^1 & \dots & a_i^1 & \dots & a_n^1 \\ & & \vdots & & \\ a_1^m & \dots & a_i^m & \dots & a_n^m \end{pmatrix}.$$

Si además $X = M_{\mathbf{u}}(x)$ y $Y = M_{\mathbf{v}}(f(x))$, entonces, por la proposición anterior, se tiene

$$Y = A \cdot X\ .$$

Resumiremos esta información en una forma abreviada expresada por el siguiente diagrama

$$(E, \mathbf{u}) \xrightarrow{(f,A)} (F, \mathbf{v})$$

$$(x, X) \longmapsto (f(x), Y = A \cdot X)\,.$$

Veamos este resultado en el ejemplo 2.1.9 desarrollado anteriormente

Ejemplo 2.1.12. En el ejemplo 2.1.9 anterior, la matriz A de f en las bases $\mathbf{u} = (\{1, x, x^2, x^3\}$, y $\mathbf{v} = \{\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)\}$) es

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

Sea $p = 1-3x+6x^2-5x^3 \in E$, entonces las coordenadas de $f(p)$ en la base $\{\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)\}$) son

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

por tanto

$$f(p) = 1 \cdot v_1 + (-3) \cdot v_2 + \frac{1}{4} v_3 \quad = \begin{pmatrix} 1 & -3 \\ \frac{1}{4} & -1 \end{pmatrix}$$

Ejercicio 2.1.13. 1. Sea $f : \mathbb{R}[x]_{\leq 2} \longrightarrow \mathbb{R}(2,2)$ la aplicación definida por

$$f(a+bx+cx^2) = \begin{pmatrix} b+c & a \\ b & c \end{pmatrix}.$$

Prueba que f es lineal y halla su matriz en las bases

$$\{1, x, x^2\}$$

de $\mathbb{R}[x]_{\leq 2}$ y

$$\{\left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right)\}$$

de $\mathbb{R}(2,2)$.

2. Sean $n \geq 0$, y $E = \mathbb{R}[x]_{\leq n}$.
   (i) Halla la matriz de la aplicación derivada $D : E \longrightarrow E$ en la base $\{1, x, x^2, \dots, x^n\}$.
   (ii) En el caso $n = 3$, halla las matrices de las potencias de D^i de D.
3. Sean $E = \mathbb{R}[x]_{\leq 1}$, $F = \mathbb{R}[x]_{\leq 2}$, y $f : E \longrightarrow F$ la aplicación lineal tal que

$$f(1) = 1 - x + 3x^2, \quad f(x) = 2 + x^2.$$

Sean

$$\mathbf{u} = \{u_1 = 1 + 2x, u_2 = 3 + 2x\},$$

y

$$\mathbf{v} = \{v_1 = 1 + 3x^2, v_2 = 2 + x + x^2, v_3 = 2 + 5x^2\}$$

bases de E y F respectivamente. Halla la matriz de f en las bases $\mathbf{u}$ y $\mathbf{v}$.

2.2. Matriz de la composición de aplicaciones lineales, cambios de base.

El resultado fundamental de esta lección es que la matriz de la composición de dos aplicaciones lineales es el producto de las matrices de los factores.

2.2.1. Matriz de la composición.

Ejemplo 2.2.1. 1. Sea $A \in \mathbf{K}(n, m)$, recordemos que

$$L_A : \mathbf{K}^m \longrightarrow \mathbf{K}^n, \quad L_A(X) = A \cdot X$$

es una aplicación lineal. Si $A \in \mathbf{K}(n, m)$ y $B \in \mathbf{K}(m, l)$, entonces la composición

$$L_B \circ L_A : \mathbf{K}^n \xrightarrow{\quad L_A \quad} \mathbf{K}^m \xrightarrow{\quad L_B \quad} \mathbf{K}^l$$

coincide con $L_{B \cdot A}$, por la propiedad asociativa del producto de matrices:

$$(L_B \circ L_A)(X) = L_B(L_A(X)) = L_B(A \cdot X) = B \cdot (A \cdot X) = (A \cdot B) \cdot X = L_{B \cdot A}(X),$$

para todo $X \in \mathbf{K}^n$.

2. Si $\mathbf{1} \in \mathbf{K}(n, n)$ es la matriz unidad, la aplicación lineal $L_{\mathbf{1}} : \mathbf{K}^n \longrightarrow \mathbf{K}^n$ es la identidad de $\mathbf{K}^n$.

3. Si $A \in \mathbf{K}(n,n)$ es una matriz cuadrada invertible entonces L_A es un isomorfismo, cuya inversa es $L_{A^{-1}} : \mathbf{K}^n \longrightarrow \mathbf{K}^n$. En efecto, $L_A \circ L_{A^{-1}} = L_{A \cdot A^{-1}} = L_{\mathbf{1}} = Id_{\mathbf{K}^n}$, por los apartados 1 y 2 anteriores. Análogamente $L_{A^{-1}} \circ L_A = Id_{\mathbf{K}^n}$.

Llegamos a la proposición que hemos adelantado al principio de la lección: "La matriz de una composición es el producto de matrices".

Proposición 2.2.2.

1. Sea E un espacio vectorial de dimensión finita n, $\mathbf{u}$ una base de E. La matriz de la aplicación identidad $Id_E : E \longrightarrow E$ en las bases $(\mathbf{u},\mathbf{u})$ es la matriz unidad $n \times n$.
2. Sean E, F, y G tres espacios vectoriales, $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ bases respectivas, y $f : E \longrightarrow F$, $g : F \longrightarrow G$ aplicaciones lineales. Se cumple
$$M_{\mathbf{w}}(g \circ f; \mathbf{u}) = M_{\mathbf{w}}(g; \mathbf{v}) \circ M_{\mathbf{v}}(f; \mathbf{u})$$
3. Sean E, y F espacios vectoriales de dimensión finita, $\mathbf{u}$ y $\mathbf{v}$ bases respectivas, y $f : E \longrightarrow F$ un isomorfismo lineal y $f^{-1} : F \longrightarrow E$ su isomorfismo inverso. Se cumple
$$M_{\mathbf{u}}(f^{-1}; \mathbf{v}) = M_{\mathbf{v}}(f; \mathbf{u})^{-1}$$

Demostración. (1) se sigue inmediatamente de las definiciones. Veamos (2). Sean $A = M_{\mathbf{v}}(f;\mathbf{u})$ y $B = M_{\mathbf{w}}(g;\mathbf{v})$. Sean $x \in E$, $X = M_{\mathbf{u}}(x)$, $Y = M_{\mathbf{v}}(f(x))$, y $Z = M_{\mathbf{w}}(g(f(x)))$. Por hipótesis
$$Y = A \cdot X, \quad Z = B \cdot Y,$$
por tanto
$$Z = B \cdot (A \cdot X) = (B \cdot A) \cdot X$$
de donde se deduce que $M_{\mathbf{w}}(g \circ f; \mathbf{u}) = B \cdot A$. La prueba de (3) se sigue de (1) y (2) y se deja como ejercicio. $\square$

La fórmula anterior se puede expresar mediante el triángulo conmutativo siguiente

$$\begin{tikzcd}
(E,\mathbf{u}) \arrow[rr, "{(f,A)}"] \arrow[dr, "{(g\circ f,B\cdot A)}"'] & & (F,\mathbf{v}) \arrow[dl, "{(g,B)}"] \\
& (G,\mathbf{w}) &
\end{tikzcd}$$

Veamos un ejemplo concreto.

Ejemplo 2.2.3. Sean $E = \mathbf{K}^3$, $F = \mathbf{K}^2$, $G = \mathbf{K}^2$, $f : \mathbf{K}^3 \longrightarrow \mathbf{K}^2$, $g : \mathbf{K}^2 \longrightarrow \mathbf{K}^2$ las aplicaciones lineales definidas por
$$f\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x - 2y + z \\ x - z \end{pmatrix} = \begin{pmatrix} 1 & -2 & 1 \\ 1 & 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix},$$
$$g\begin{pmatrix} t \\ s \end{pmatrix} = \begin{pmatrix} 2t - s \\ t + s \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} t \\ s \end{pmatrix}.$$
Las matrices A de f, B de g son ,
$$A = \begin{pmatrix} 1 & -2 & 1 \\ 1 & 0 & -1 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix};$$

de modo que, si $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, $Y = \begin{pmatrix} s \\ t \end{pmatrix}$, entonces

$$f(X) = A \cdot X, \quad g(Y) = B \cdot Y.$$

Por un cálculo directo de la composición obtenemos:

$$(g \circ f) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2(x-2y+z)-(x-z) \\ (x-2y+z)+(x-z) \end{pmatrix} = \begin{pmatrix} x-4y+3z \\ 2x-2y \end{pmatrix} = \begin{pmatrix} 1 & -4 & 3 \\ 2 & -2 & 0 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

La proposición anterior establece que podemos calcular el resultado anterior simplemente multiplicando las matrices B y A (en este orden) para obtener la matriz de $g \circ f$. En efecto, el producto $B \cdot A$ es

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

2.2.2. **Cambios de base.**

En este apartado estudiaremos la transformación de las coordenadas de un vector por un cambio de base.

Sean E un espacio vectorial de dimensión n, $\mathbf{u}$ y $\widehat{\mathbf{u}}$ dos bases de E. Para cada vector $x \in E$ tenemos dos vectores de coordenadas $X = M_{\mathbf{u}}(x) \in \mathbf{K}^n$ y $\widehat{X} = M_{\widehat{\mathbf{u}}}(x) \in \mathbf{K}^n$, tales que

$$\mathbf{u} \cdot X = x = \widehat{\mathbf{u}} \cdot \widehat{X}.$$

Problema 2.2.4. Hallar la relación entre X y $\widehat{X}$.

Para resolver este problema necesitamos la relación entre las bases $\mathbf{u}$ y $\widehat{\mathbf{u}}$, es decir la matriz de tipo $n \times n$, $M_{\mathbf{u}}(\widehat{u}_1, \dots, \widehat{u}_n)$, cuya columna i-ésima son las componentes de $\widehat{u}_i$ en la base $\mathbf{u}$.

Definición 2.2.5. Con las notaciones anteriores, llamaremos matriz de paso de la base $\widehat{\mathbf{u}}$ a la base $\mathbf{u}$, o simplemente matriz de cambio de base, a la matriz

$$M_{\mathbf{u}}(\widehat{\mathbf{u}}) := M_{\mathbf{u}}(\widehat{u}_1, \dots, \widehat{u}_n).$$

Observemos que $M_{\mathbf{u}}(\widehat{\mathbf{u}})$ es la matriz $M_{\mathbf{u}}(Id_E, \widehat{\mathbf{u}})$ de la aplicación identidad de E en las bases $(\widehat{\mathbf{u}}, \mathbf{u})$. De las definiciones y de la proposición 2.1.10 se obtiene la siguiente relación entre X y $\widehat{X}$.

Proposición 2.2.6. (Fórmula de cambio de base para las coordenadas.) Sean E un espacio vectorial de dimensión finita, $\mathbf{u}$ y $\widehat{\mathbf{u}}$ dos bases de E. Se cumple

$$M_{\mathbf{u}}(x) = M_{\mathbf{u}}(\widehat{\mathbf{u}}) \cdot M_{\widehat{\mathbf{u}}}(x).$$

En forma abreviada, si notamos $P = M_{\mathbf{u}}(\widehat{\mathbf{u}})$:

$$(E, \widehat{\mathbf{u}}) \xrightarrow{(Id_E, P)} (E, \mathbf{u}),$$

$$(x, \widehat{X}) \longmapsto \left(x, X = P \cdot \widehat{X}\right).$$

A continuación nos planteamos obtener la relación que hay entre las diferentes matrices de cambio de base cuando tenemos tres bases diferentes. La respuesta se suele conocer como la propiedad transitiva del cambio de base. Más precisamente, se cumplen las tres reglas básicas siguientes.

Proposición 2.2.7. Sean E un espacio vectorial de dimensión finita, $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ tres bases de E. Entonces

$$
\begin{aligned}
M_{\mathbf{u}}(\mathbf{u}) &= \mathbf{1}\\
M_{\mathbf{u}}(\mathbf{w}) &= M_{\mathbf{u}}(\mathbf{v}) \cdot M_{\mathbf{v}}(\mathbf{w})\\
M_{\mathbf{v}}(\mathbf{u}) &= M_{\mathbf{u}}(\mathbf{v})^{-1}
\end{aligned}
$$

Demostración. Se sigue de la Proposición 2.2.2. □

Ejercicio 2.2.8. Sean E un espacio vectorial y $\mathbf{u} = \{u_1, u_2, u_3\}$ una base de E. Sea $\widehat{\mathbf{u}} = \{\widehat{u}_1, \widehat{u}_2, \widehat{u}_3\}$ la base de E definida por $\widehat{u}_1 = u_1 + 3u_3$, $\widehat{u}_2 = 2u_1 + u_2 + u_3$, $\widehat{u}_3 = 2u_1 + 5u_3$.

Dado un vector arbitrario $x \in E$, sean $X = (x^1, x^2, x^3)$ las componentes de x en la base $\mathbf{u}$ e $\widehat{X} = (\widehat{x}^1, \widehat{x}^2, \widehat{x}^3)$ las componentes de x en la base $\widehat{\mathbf{u}}$.

1. Halla la relación entre X e $\widehat{X}$ y expresa el resultado mediante el cálculo matricial. Utiliza la fórmula que has obtenido para hallar las componentes de $\widehat{u}_1 + \widehat{u}_2 - 2\widehat{u}_3$ en la base $\mathbf{u}$.
2. Calcula las componentes de los vectores u_i en la base $\widehat{\mathbf{u}}$. Halla a continuación las componentes de $-u_1 + u_2 - 6u_3$ en la base $\widehat{\mathbf{u}}$. *Indicación:* $\begin{pmatrix} 1 & 2 & 2 \\ 0 & 1 & 0 \\ 3 & 1 & 5 \end{pmatrix} \cdot \begin{pmatrix} -5 & 8 & 2 \\ 0 & 1 & 0 \\ 3 & -5 & -1 \end{pmatrix} = \mathbf{1}$.
3. Sea $\mathbf{w}$ la base definida por $w_1 = -u_1 + 2u_3$, $w_2 = -u_1 + u_3$, $w_3 = 2u_1 + u_2$. Halla la matriz $M_{\widehat{\mathbf{u}}}(\mathbf{w})$ de paso de la base $\mathbf{w}$ a la base $\widehat{\mathbf{u}}$.

2.2.3. Efecto de un cambio de bases en la matriz de una aplicación lineal.

Veamos como cambia la matriz de una aplicación lineal cuando cambiamos las bases de los espacios.

Sean E y F dos espacios vectoriales de dimensión finita.

(i) Sean $\mathbf{u}$, $\widehat{\mathbf{u}}$ dos bases de E, $P = M_{\mathbf{u}}(\widehat{\mathbf{u}})$ la matriz de cambio de base.
(ii) Sean $\mathbf{v}$, $\widehat{\mathbf{v}}$ dos bases de F, $Q = M_{\mathbf{v}}(\widehat{\mathbf{v}})$ la matriz de cambio de base.
(iii) Sean $f : E \longrightarrow F$ una aplicación lineal,
  (a) A la matriz de f en las bases $(\mathbf{u}, \mathbf{v})$,
  ($\widehat{a}$) $\widehat{A}$ la matriz de f en las bases $(\widehat{\mathbf{u}}, \widehat{\mathbf{v}})$.

Problema 2.2.9. ¿Qué relación existe entre A y $\widehat{A}$?

En el siguiente cuadrado

$$
\begin{tikzcd}
(E, \widehat{\mathbf{u}}) \arrow[r, "{(f,\widehat{A})}"] \arrow[d, "{(Id_E,P)}"'] & (F, \widehat{\mathbf{v}}) \arrow[d, "{(Id_F,Q)}"] \\
(E, \mathbf{u}) \arrow[r, "{(f,A)}"] & (F, \mathbf{v})\ .
\end{tikzcd}
$$

se cumple que las dos composiciones de E a F coinciden, $Id_F \circ f = f \circ Id_E$ (decimos que el cuadrado es conmutativo), por tanto también coinciden los productos de matrices correspondientes:

Teorema 2.2.10. (Fórmula de cambio de base para la matriz de una aplicación lineal). Con las hipótesis anteriores,

$$A \cdot P = Q \cdot \widehat{A},$$

es decir

$$M_{\mathbf{v}}(f; \mathbf{u}) \cdot M_{\mathbf{u}}(\widehat{\mathbf{u}}) = M_{\mathbf{v}}(\widehat{\mathbf{v}}) \cdot M_{\widehat{\mathbf{v}}}(f; \widehat{\mathbf{u}}).$$

Ejemplo 2.2.11. Sea $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ la aplicación lineal que en las coordenadas canónicas está expresada por

$$f(x, y) = (2x - y, x + y, -y).$$

Sean $\mathbf{u}$ y $\mathbf{v}$ las bases canónicas de $\mathbb{R}^2$ y $\mathbb{R}^3$ respectivamente. Consideremos las nuevas bases $\widehat{\mathbf{u}} = (\widehat{u}_1, \widehat{u}_2)$ de $\mathbb{R}^2$ y $\widehat{\mathbf{v}} = (\widehat{v}_1,\ \widehat{v}_2,\ \widehat{v}_3)$ de $\mathbb{R}^3$ definidas por

$$\widehat{u}_1 = (2, 3),\ \widehat{u}_2 = (1, 2);\quad \widehat{v}_1 = (1, 0, 3),\ \widehat{v}_2 = (2, 0, 5),\ \widehat{v}_3 = (0, 1, 0)$$

Halla la matriz $\widehat{A}$ de f en las bases $(\widehat{\mathbf{u}}, \widehat{\mathbf{v}})$.

Solución: $\widehat{A} = Q^{-1} \cdot A \cdot P = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 3 & 5 & 0 \end{pmatrix}^{-1} \cdot \begin{pmatrix} 2 & -1 \\ 1 & 1 \\ 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix} = \begin{pmatrix} -11 & -4 \\ 6 & 4 \\ 5 & 3 \end{pmatrix}.$

2.3. Núcleo e imagen de una aplicación lineal.

En lo que sigue, E y F son dos espacios vectoriales sobre $\mathbf{K}$.

En esta lección repasaremos las definiciones de núcleo e imagen de una aplicación lineal dadas en el capítulo 1 y desarrollaremos sus principales propiedades.

En el curso de "Matrius i vectors" habeis visto que, dados $A = (a_{ij})_{ij} \in \mathbf{K}(m, n)$ y $b \in \mathbf{K}^m$, el conjunto de soluciones del sistema de ecuaciones lineales en la incógnita $(x_1, \ldots, x_n) \in \mathbf{K}^n$,

$$\begin{cases} a_{11}x_1 + \cdots + a_{n1}x_n & = b_1 \\ \qquad \cdots & \\ a_{1m}x_1 + \cdots + a_{nm}x_n & = b_m \end{cases}$$

tiene asociado un conjunto de soluciones

$$S_b = \{X \in \mathbf{K}^n; A \cdot X = b\}.$$

Para $b = 0$, S_0 es un subespacio vectorial de $\mathbf{K}^n$, que coincide con el núcleo de la aplicación lineal $L_A : \mathbf{K}^n \longrightarrow \mathbf{K}^m$.

Si $b \neq 0$, entonces S_b puede ser vacío (el sistema se llama incompatible), o no vacío (el sistema se llama compatible). Si el sistema es compatible y $x_0 \in S_b$ es una solución particular, la solución general se obtiene en la forma $S_b = x_0 + S_0$

El subespacio vectorial S_0 se dice que está expresado *por un sistema de ecuaciones*. En contraste con esta presentación, existe la expresión de un subespacio vectorial siguiente. Dada una familia $(u_1, \ldots, u_r)$ de vectores de E, podemos considerar el subespacio vectorial de E generado por $\{u_1, \cdots, u_r\}$,

$$F = \langle\{u_1, \cdots, u_r\}\rangle = \{\lambda_1 u_1 + \cdots + \lambda_r u_r \in E; \lambda_1, \ldots, \lambda_r \in \mathbf{K}\}.$$

Al expresar los vectores de F en la forma $\lambda_1 u_1 + \cdots + \lambda_r u_r$, donde los parámetros λ_i son arbitrarios, se dice que F está expresado en *forma parámetrica* (especialmente cuando la familia $(u_1, \ldots, u_r)$ es linealmente independiente). El subespacio F coincide con la imagen de la única aplicación lineal $f : \mathbf{K}^r \longrightarrow E$ tal que $f(e_i) = u_i$, para todo i.

2.3.1. Núcleo de una aplicación lineal.

En primer lugar se define el núcleo de una aplicación lineal, que es el conjunto de vectores que se aplican en el cero.

Definición 2.3.1. Sea $f : E \longrightarrow F$ una aplicación lineal. Se llama *núcleo* de f al subconjunto de E definido por

$$\mathrm{Ker}\,(f) = f^{-1}(0) = \{x \in E;\ f(x) = 0\ \}.$$

Proposición 2.3.2. Si $f : E \longrightarrow F$ es una aplicación lineal, $\operatorname{Ker} f$ es un subespacio vectorial de E.

Ejemplo 2.3.3. En el ejemplo 2.1.3, el núcleo de L_A es el subespacio de $\mathbb{R}^3$ definido por

$$\{(x, y, z) \in \mathbb{R}^3; (x + 2y + 3z, 4x + 5y + 6z) = (0, 0)\},$$

es decir, $\operatorname{Ker} L_A = \langle\{(1, -2, 1)\}\rangle$.

En la siguiente proposición mostramos, entre otras cosas, que el núcleo de una aplicación lineal $f : E \longrightarrow F$ controla la inyectividad de la aplicación lineal (1) y que la inyectividad de la aplicación f equivale a la conservación de la independencia lineal de vectores (3). En el caso en que E y F sean de dimensión finita y f inyectiva, en (4) damos la forma más simple de la matriz de f, en bases adecuadas de los espacios E y F.

Proposición 2.3.4. Sea $f : E \longrightarrow F$ una aplicación lineal.

1. f es inyectiva si y solo si $\operatorname{Ker}(f) = \{0\}$.
2. Si existe una base B de E tal que $f_{|B})$ es inyectiva y $f(B)$ es linealmente independiente, entonces f es inyectiva.
3. Si f es inyectiva, para todo $S \subset E$ linealmente independiente, $f(S)$ es linealmente independiente.
4. Si E y F son de dimensión finita y f es inyectiva, entonces $n = \dim E \leq \dim F$, y existen bases $\mathbf{u}$ de E y $\mathbf{v}$ de F tales que la matriz de f en estas bases es $\begin{pmatrix} \mathbf{1}_n \\ \mathbf{0} \end{pmatrix}$.

Demostración. (2) Con las hipótesis de (2), veamos que $\operatorname{Ker} f = 0$. Sea $x \in \operatorname{Ker} f$. Expresemos x en la base $B = \{u_i\}_{i\in I}$, $x = \sum_i \lambda_i u_i$. Por la linealidad de f,

$$0 = f(x) = \sum_i \lambda_i f(u_i),$$

pero como los vectores $f(u_i)$ son distintos y linealmente independientes, todos los λ_i se anulan. Por tanto $x = 0$, es decir, f es inyectiva, por el apartado (1).

(3) Supongamos f inyectiva y S linealmente independiente. Veamos que $f(S)$ es linealmente independiente. Si $\sum_i \lambda_i f(u_i)$ es una combinación lineal de elementos distintos de $f(S)$ tal que $\sum_i \lambda_i f(u_i) = 0$, entonces $f(\sum_i \lambda_i u_i) = \sum_i \lambda_i f(u_i) = 0$, por tanto $\sum_i \lambda_i u_i \in \operatorname{Ker} f = \{0\}$, pero como S es linealmente independiente, resulta que todos los λ_i son nulos.

(4) Supongamos que f es inyectiva, $\dim E = n$, y $B = \{u_i\}_{1\leq i\leq n}$ una base de E. Entonces $f(B)$ es un subconjunto de F de n elementos (pues f es inyectiva) y linealmente independiente (por el apartado (3)). Por el teorema de Steitnitz, $n \leq \dim F$, y además $f(B)$ se puede completar a una base $B' = \{v_j\}_{1\leq j\leq m}$ de F, de modo que $v_i = f(u_i)$ para $1 \leq i \leq n$. En las bases $(\mathbf{u}, \mathbf{v})$ la matriz de f es $\left(\begin{smallmatrix} \mathbf{1}_n \\ \mathbf{0} \end{smallmatrix}\right)$ □

Observación 2.3.5. Sea E un espacio vectorial de dimensión finita n. Entonces $0 \leq \dim \operatorname{Ker} f \leq n$, y $\dim \operatorname{Ker}$ es una medida de como f difiere de ser inyectiva. El valor mínimo de esta dimensión es 0 (f es inyectiva), y el valor máximo es $\dim E$ ($f = 0$).

Observación 2.3.6. El siguiente ejemplo muestra que, en el item 3 del teorema anterior, la hipótesis de que los vectores $f(u_i)$ sean diferentes es importante. Sea $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ definida por $f(x, y) = (x + y, 0, 0)$. La aplicación f no es inyectiva, pero, si $B = \{e_1, e_2\}$ denota la base canónica de $\mathbb{R}^2$, $f(B) = \{(1, 0, 0)\}$ es linealmente independiente.

2.3.2. Imagen de una aplicación lineal. Rango.

Definición 2.3.7. Sea $f : E \longrightarrow F$ una aplicación lineal. Se llama *imagen* de f al subconjunto de F definido por

$$\operatorname{Im} f = f(E) = \{y \in F;\ \exists x \in E,\ y = f(x)\ \}.$$

Proposición 2.3.8. Si $f : E \longrightarrow F$ una aplicación lineal, $\operatorname{Im} f$ es un subespacio vectorial de F.

Ejemplo 2.3.9. Sean $A = \left(\begin{smallmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{smallmatrix}\right)$ y $L_A : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ la aplicacion lineal $L_A(X) = A \cdot X$. Entonces la imagen de L_A está generada por las imágenes de los vectores de la base canónica de $\mathbb{R}^2$, es decir, las columnas de A, $\operatorname{Im} L_A = \langle\{(1, 3, 5), (2, 4, 6)\}\rangle$.

Definición 2.3.10. Sea $f : E \longrightarrow F$ una aplicación lineal. Si $\operatorname{Im} f$ tiene dimensión finita, su dimensión se llama el *rango* de f,

$$\operatorname{rg}(f) = \dim \operatorname{Im} f.$$

En la siguiente proposición, que es un resultado paralelo al obtenido para el núcleo (ver proposición 2.3.4), vemos que la imagen de una aplicación lineal $f : E \longrightarrow F$ controla si la aplicación f es exhaustiva, aunque este criterio solo tiene utilidad cuando se expresa en términos de dimensiones (1). También vemos que las aplicaciones exhaustivas se pueden caracterizar como aquellas que envían generadores de E a generadores de F (3). En el caso en que E y F sean de dimensión finita y f exhaustiva, en (4) damos la forma más simple de la matriz de f, en bases adecuadas en los espacios E y F.

Proposición 2.3.11. Sea $f : E \longrightarrow F$ una aplicación lineal.

1. La aplicación f es exhaustiva si y solo si $\operatorname{Im}(f) = F$. En particular, si F es de dimensión finita, f es exhaustiva si y solo si $\operatorname{rg}(f) = \dim F$.
2. Si existe una base B de E tal que $f(B)$ genera F, entonces f es exhaustiva.
3. Si f es exhaustiva, para todo subconjunto S que genera E, el subconjunto $f(S)$ genera F.
4. Si f es exhaustiva, entonces $\dim E \geq \dim F = m$, y existen bases $\mathbf{u}$ de E y $\mathbf{v}$ de F tales que la matriz de f en estas bases es $\begin{pmatrix} \mathbf{1}_m & \mathbf{0} \end{pmatrix}$.

Demostración. (1) Si $\operatorname{rg}(f) = \dim F$, entonces, por el Corolario 1.6.3(3) se tiene $\operatorname{Im} f = F$.

(4) Sea $\mathbf{v} = \{v_j\}_{1 \leq j \leq m}$ una base de F. Para cada j, sea $u_j \in E$ tal que $f(u_j) = v_j$. Sea $S = \{u_j; 1 \leq j \leq m\}$. Puesto que $f(S)$ es linealmente independiente también lo es S. En efecto, si existen escalares $\lambda_i \in \mathbf{K}$ tales que $\sum_{j=1}^{m} \lambda_j u_j = 0$, entonces $\sum_{j=1}^{m} \lambda_j f(u_j) = 0$, y como $\{v_j\}_{1 \leq j \leq m}$ es linealmente independiente, resulta $\lambda_j = 0$, para todo j, por tanto S es linealmente independiente.

Veamos que $\langle S \rangle \cap \operatorname{Ker} f = \{0\}$ En efecto, si $x \in \langle S \rangle \cap \operatorname{Ker} f$, entonces $f(x) = 0$ y existen escalares $\lambda_i \in \mathbf{K}$ tales que $x = \sum_{j=1}^{m} \lambda_j u_j$. Por tanto $\sum_{j=1}^{m} \lambda_j f(u_j) = f(x) = 0$, y como $\{v_j\}_{1 \leq j \leq m}$ es linealmente independiente, resulta $\lambda_j = 0$, para todo j, por tanto $x = 0$.

Veamos que $\langle S \rangle + \operatorname{Ker} f = E$. Si $x \in E$, entonces existen escalares $\lambda_i \in \mathbf{K}$ tales que $f(x) = \sum_{j=1}^{m} \lambda_j v_j$. Sea $u = \sum_{j=1}^{m} \lambda_j u_j \in \langle S \rangle$. Entonces $w := x - u \in \operatorname{Ker} f$ y $x = u + w \in \langle S \rangle + \operatorname{Ker} f$.

Sea $\{w_k\}_{1 \leq k \leq n-m}$ una base de $\operatorname{Ker} f$. Entonces $\mathbf{u} = (u_1, \ldots, u_m, w_1, \ldots, w_{n-m})$ es una base de E. En las bases $(\mathbf{u}, \mathbf{v})$ la matriz de f es $\begin{pmatrix} \mathbf{1}_m & \mathbf{0} \end{pmatrix}$

□

Observación 2.3.12. Si F es un espacio vectorial de dimensión finita y $f : E \longrightarrow F$ es una aplicación lineal, $\dim F - \operatorname{rg} f \geq 0$ puede considerarse una medida de como f difiere de ser exhaustiva. El valor mínimo es 0 (f exhaustiva) y el máximo es $\dim F$ ($f = 0$).

Corolario 2.3.13. Sea $f : E \longrightarrow F$ una aplicación lineal entre espacios vectoriales de dimensión finita, entonces

1. $\operatorname{rg}(f) \leq \text{máx}(\dim E, \dim F)$,
2. f es exhaustiva si y solo si $\operatorname{rg} f = \dim F$.
3. f es inyectiva si y solo si $\operatorname{rg} f = \dim E$.

Demostración. (1) Puesto que $\operatorname{Im} f$ es un subespacio de F, se cumple $\dim \operatorname{Im} f \leq \dim F$. Por otra parte, si B es una base de E con n elementos, $f(B)$ es un conjunto generador de $\operatorname{Im} f$ con $\leq n$ elementos, por tanto $\dim \operatorname{Im} f \leq n$.

(2) f es exhaustiva si y solo si $\operatorname{Im} f = F$, que equivale a $\operatorname{rg}(f) = \dim F$ por el Corolario 1.6.3(3).

(3) Si f es inyectiva, y B es una base de E con n elementos, entonces $f(B)$ es un conjunto de n vectores linealmente independientes de $\operatorname{Im} f$, que además son generadores y por tanto una base de $\operatorname{Im} f$, de donde resulta $\operatorname{rg} f = \dim E$. Recíprocamente, si $n = \dim E \leq \dim \operatorname{Im} f$, puesto que f transforma una base B de n elementos de E en un conjunto $f(B)$ de generadores de $\operatorname{Im} f$, y todo conjunto de n generadores de $\operatorname{Im} f$ es una base, resulta que $f(B)$ es linealmente independiente, y por tanto f es inyectiva. □

Proposición 2.3.14. Sea $f : E \longrightarrow F$ una aplicación lineal entre espacios vectoriales de dimensión finita. Si $\mathbf{u}$ y $\mathbf{v}$ son bases de E y F respectivamente, y A es la matriz de f en esta base, entonces

$$\operatorname{rg}(f) = \operatorname{rg}(A)$$

En particular, f es exhaustiva si y solo si $\operatorname{rg} A = \dim F$.

Demostración. Sea $\mathbf{u} = \{u_i\}_i$ una base de E. El rango de la matriz A coincide con el máximo número de columnas linealmente independientes, que, a su vez (ver 1.6.14), es la dimensión del subespacio $\langle \{f(u_i); i\} \rangle$ generado por las imágenes de los vectores de la base. □

2.4. Isomorfismos y fórmula de las dimensiones.

2.4.1. Isomorfismos y dimensión.

Recordemos que una aplicación lineal $f : E \longrightarrow F$ se llama isomorfismo si existe una aplicación lineal $g : F \longrightarrow E$ tal que $f \circ g = Id_F$ y $g \circ f = Id_E$. En tal caso E y F se llaman espacios vectoriales isomorfos, y esta relación se denota por $E \cong F$.

Puesto que una aplicación lineal es un isomorfismo si y solo si es biyectiva, tenemos la siguiente propiedad.

Proposición 2.4.1. Sea $f : E \longrightarrow F$ una aplicación lineal. Las siguientes propiedades son equivalentes:

1. f es un isomorfismo.
2. $\operatorname{Ker} f = \{0\}$ y $\operatorname{Im}(f) = F$.

Los isomorfismos de espacios vectoriales conservan todas las propiedades que hemos desarrollado en este contexto: combinaciones lineales, independencia lineal, generadores, subespacios, dimensión, bases, intersecciones y sumas de subespacios, sumas directas, etc. Un criterio suficiente para decidir si una aplicación lineal es un isomorfismo es estudiando como se transforma una base.

Proposición 2.4.2. Sea $f : E \longrightarrow F$ una aplicación lineal.

1. Si existe una base B de E tal que $f_{|B}$ es inyectiva y $f(B)$ es una base de F, entonces f es un isomorfismo.
2. f es un isomorfismo si y solo si para toda base B de E, $f(B)$ es una base de F.
3. Si E y F son de dimensión finita y f es un isomorfismo, entonces
$$\dim E = \dim F.$$

La dimensión es un invariante que clasifica las clases de isomorfismo de los espacios vectoriales de dimensión finita.

Corolario 2.4.3. Sean E y F dos espacios vectoriales isomorfos. Entonces E tiene dimensión finita si y solo si F tiene dimensión finita, y ambos tienen la misma dimensión.

Si E y F son espacios vectoriales de dimensión finita y de la misma dimensión, entonces E y F son isomorfos.

2.4.2. Teorema de isomorfismo y fórmula de las dimensiones.

En un sistema de ecuaciones lineales $AX = b$, si $b \neq 0$, entonces S_b puede ser vacío (el sistema se llama incompatible), o no vacío (el sistema se llama compatible). Si el sistema es compatible y $x_0 \in S_b$ es una solución particular, la solución general se obtiene en la forma $S_b = x_0 + S_0$

Para una aplicación lineal $f : E \longrightarrow F$, la relación de equivalencia definida por f, "$x \sim x'$ si y solo si $f(x) = f(x')$", coincide con la relación de equivalencia definida por el subespacio $\operatorname{Ker} f$. En particular, para $y \in F$, el conjunto de sus antiimágenes
$$f^{-1}(\{y\}) = \{x \in E;\ f(x) = y\ \}$$
es vacío o bien, para todo $x_0 \in f^{-1}(\{y\})$, se cumple
$$f^{-1}(\{y\}) = x_0 + \operatorname{Ker} f = \{x_0 + x;\ f(x) = 0\ \}.$$

Una forma alternativa y más estructural de expresar este resultado es el teorema de isomorfismo siguiente.

Teorema 2.4.4. (**Teorema de isomorfismo**). Sean E y F espacios vectoriales, y $f : E \longrightarrow F$ es una aplicación lineal. La aplicación
$$\widehat{f} : E/\operatorname{Ker} f \longrightarrow \operatorname{Im} f, \quad \widehat{f}([x]) = f(x), \forall x \in E,$$
está bien definida y es un isomorfismo.

Corolario 2.4.5. Fórmula de las dimensiones. Sean E y F espacios vectoriales, y $f : E \longrightarrow F$ una aplicación lineal. Supongamos que E es de dimensión finita, entonces
$$\dim E = \dim \operatorname{Ker} f + \dim \operatorname{Im} f.$$

La fórmula anterior es la que aparece en el teorema de Rouché-Frobenius sobre la resolución de un sistema homogéneo de ecuaciones lineales. En efecto, si $A \in \mathbf{K}(n, m)$ es una matriz, el conjunto de soluciones S_A de la ecuación
$$A \cdot X = 0$$

en la incógnita $X \in \mathbf{K}^n$ es el núcleo de $L_A : \mathbf{K}^n \longrightarrow \mathbf{K}^m$. Por otra parte, $\operatorname{rg} A = \dim \operatorname{Im} L_A$, por tanto la fórmula de las dimensión en este caso se expresa por

$$\dim S_A = n - \operatorname{rg} A.$$

Corolario 2.4.6. Sean E y F dos espacios vectoriales de dimensión finita y supongamos que ambos tienen la misma dimensión. Sea $f : E \longrightarrow F$ una aplicación lineal, entonces las siguientes afirmaciones son equivalentes

1. f es inyectiva.
2. f es un isomorfismo.
3. f es exhaustiva.

Demostración. Se sigue de 2.3.13 □

Ejercicio 2.4.7. Sea $f : E \longrightarrow F$ una aplicación lineal entre espacios vectoriales de la misma dimensión finita. Sean $\mathbf{u}$ y $\mathbf{v}$ son bases de E y F respectivamente, y A la matriz de f en esta base. Entonces f es un isomorfismo si y solo si la matriz A es inversible.

Ejercicio 2.4.8. Sean F, G dos subespacios vectoriales de un espacio vectorial E. Sea

$$\phi : F/(F \cap G) \longrightarrow (F + G)/G$$

la aplicación definida por

$$\phi([u]) = [[u]]$$

donde $[u]$ denota la clase de equivalencia de u en $F/(F \cap G)$ y $[[u]]$ denota la clase de equivalencia de u en $(F + G)/G$.

(i) Prueba que ϕ está bien definida.
(ii) Prueba que ϕ es un isomorfismo.
(iii) Deduce de (ii) la fórmula de Grassmann.

Ejercicio 2.4.9. Sean n, m enteros ≥ 1 y r un entero tal que $0 \leq r \leq n, m$. Sea $\pi : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ la aplicación lineal definida por $\pi(x_1, \cdots, x_n) = (x_1, \ldots, x_r, 0, \ldots, 0)$. Halla la matriz de π en las bases canónicas y determina el rango de f.

Ejercicio 2.4.10. Sean E y F dos espacios vectoriales de dimensión finita sobre $\mathbb{R}$, y $f : F \longrightarrow E$ una aplicación lineal de rango r. Prueba que existen bases $\mathcal{U}$ de E y $\mathcal{V}$ de F tales que la matriz de f en las bases $\mathcal{V}$ y $\mathcal{U}$ es

$$\begin{pmatrix} \mathbf{I}_r & 0 \\ 0 & 0 \end{pmatrix}$$

2.5. Algebra de endomorfismos.

2.5.1. **Espacio vectorial de aplicaciones lineales.** Sean E y F dos espacios vectoriales sobre un mismo cuerpo $\mathbf{K}$. Denotamos por $\mathcal{L}(E, F)$ el conjunto de las aplicaciones lineales de E en F. Nuestro objetivo es dotar a este conjunto de una estructura de espacio vectorial.

El conjunto de matrices de un tipo dado $\mathbf{K}(n, m)$ es un espacio vectorial, de dimensión $n \cdot m$. Además $\mathbf{K}(n, m)$ se identifica con $\mathcal{L}(\mathbf{K}^n, \mathbf{K}^m)$, mediante la aplicación biyectiva

$$\mathbf{K}(m, n) \longrightarrow \mathcal{L}(\mathbf{K}^n, \mathbf{K}^m), \quad A \mapsto L_A,$$

donde $L_A(X) = A \cdot X$, para todo $X \in \mathbf{K}^n = \mathbf{K}(n, 1)$. Esta aplicación satisface la propiedad de linealidad

$$L_{A+B} = L_A + L_B, \quad L_{\lambda \cdot A} = \lambda \cdot L_A,$$

interpretanto adecuadamente las operaciones $\mathcal{L}(K^n, \mathbf{K}^m)$, de modo que esta biyección es un isomorfismo de espacios vectoriales. También satisface la regla de la composición:

$$L_{\mathbf{a}} = Id; \quad L_{B\cdot A} = L_B \circ L_A,$$

si el producto $B \cdot A$ está definido.

Más generalmente, sean E y F dos espacios vectoriales sobre un mismo cuerpo $\mathbf{K}$. La suma $f + g$ de dos aplicaciones $f, g : E \longrightarrow F$, y el producto $\lambda \cdot f$ de f por un escalar $\lambda \in \mathbf{K}$ se definen por

$$(f + g)(x) := f(x) + g(x), \quad (\lambda \cdot f)(x) := \lambda \cdot (f(x)), \quad \forall\, x \in E.$$

La aplicación nula 0 y la opuesta de una aplicación f se definen respectivamente por

$$0(x) := 0, \quad (-f)(x) := -(f(x)), \quad \forall x \in E.$$

Con estas operaciones se tiene el siguiente resultado.

Proposición 2.5.1. Sean E y F dos espacios vectoriales sobre un mismo cuerpo $\mathbf{K}$.

1. Si $f, g : E \longrightarrow F$ son lineales, y $\lambda \in \mathbf{K}$, entonces las aplicaciones
$$f + g,\ \lambda \cdot f,\ 0,\ -f$$
son lineales
2. Con las operaciones anteriores, $\mathcal{L}(E, F)$ es un espacio vectorial sobre $\mathbf{K}$.
3. Supongamos que E y F son espacios vectoriales de dimensiones n y m respectivamente. Sean $\mathbf{u}$ una base de E y $\mathbf{v}$ una base de F. La aplicación
$$M : \mathcal{L}(E, F) \longrightarrow \mathbf{K}(m, n), \quad f \mapsto M_{\mathbf{v}}(f; \mathbf{u})$$
es lineal:
$$M(f + g) = M(f) + M(g),$$
$$M(\lambda \cdot f) = \lambda \cdot M(f)$$
En particular,
$$M(0) = 0, \quad M(-f) = -M(f).$$
4. La aplicación M del apartado anterior es un isomorfismo. En particular $\mathcal{L}(E, F)$ tiene dimensión finita y
$$\dim \mathcal{L}(E, F) = \dim E \cdot \dim F.$$

2.5.2. **Estructura multiplicativa. Endomorfismos y automorfismos.** Las operaciones de espacio vectorial en el conjunto de aplicaciones lineales son compatibles con la operación de composición, en el sentido siguiente:

Proposición 2.5.2. Sean E, F y G tres espacios vectoriales sobre $\mathbf{K}$. La aplicación de composición

$$\mathcal{L}(F, G) \times \mathcal{L}(E, F) \longrightarrow \mathcal{L}(E, G), \quad (g, f) \mapsto g \circ f,$$

es bilineal, es decir, satisface

$$g \circ (f_1 + f_2) = g \circ f_1 + g \circ f_2,$$
$$(g_1 + g_2) \circ f = g_1 \circ f + g_2 \circ f,$$
$$g \circ (\lambda \cdot f) = \lambda \cdot (g \circ f) = (\lambda \cdot g) \circ f$$

En el estudio de matrices habeis observado que las matrices cuadradas de un mismo tamaño, digamos $n \times n$, se pueden sumar, multiplicar por escalares y también multiplicar entre ellas. La estructura a que dan lugar es la de un anillo (no necesariamente conmutativo), o también de $\mathbf{K}$-álgebra.

El conjunto de todas las aplicaciones lineales sobre espacios vectoriales variables sobre un mismo cuerpo $\mathbf{K}$ es una estructura más compleja que se llama "una categoría $\mathbf{K}$-lineal" y que se estudia en cursos superiores. Sin embargo existe un caso particular especialmente interesante en que se pueden combinar la estructura de espacio vectorial con la de la composición. Se trata del caso en que se fija un único espacio vectorial E y se considera el espacio vectorial de aplicaciones lineales de E en sí mismo, $f : E \longrightarrow E$. Tales aplicaciones lineales se llaman *endomorfismos*. El conjunto de los endomorfismos de E se denota por $\operatorname{End}(E)$.

Dada una aplicación lineal $f : E \longrightarrow F$ entre dos espacios vectoriales de dimensión n, escogiendo bases en los espacios E y F asociamos a f una matriz cuadrada A. Por tanto podemos considerar los escalares $\det A$ o $\operatorname{tr} A$. Sin embargo estos escalares dependen de las bases escogidas. Puesto que f es un isomorfismo si y solo si A es inversible, lo cual a su vez equivale a que $\det A \neq 0$, observamos que la nulidad de $\det A$ no depende de las bases escogidas. Cuando $E = F$, es decir, cuando f es un *endomorfismo*, entonces podemos (y en general debemos) escoger la misma base en el espacio de salida y en el espacio de llegada. Haciendo esto obtendremos resultados aún más interesantes. Por ejemplo, calcular las potencias de f como las potencias de A, y también que el determinante o la traza de A es independiente de la base escogida.

Proposición 2.5.3. Sea E un espacio vectorial de dimensión n sobre $\mathbf{K}$. Sean $\mathbf{u}$, $\mathbf{v}$ dos bases cualesquiera de E, y $P = M_{\mathbf{u}}(\mathbf{v})$ la matriz de cambio de base, entonces, para todo $f \in \operatorname{End}(E)$,

$$M_{\mathbf{u}}(f;\mathbf{u}) = P \cdot M_{\mathbf{v}}(f;\mathbf{v}) \cdot P^{-1}.$$

En particular,

$$\det M_{\mathbf{u}}(f;\mathbf{u}) = \det M_{\mathbf{v}}(f;\mathbf{v})$$
$$\operatorname{tr} M_{\mathbf{u}}(f;\mathbf{u}) = \operatorname{tr} M_{\mathbf{v}}(f;\mathbf{v})$$

La proposición anterior garantiza la corrección de la siguiente definición.

Definición 2.5.4. Sea E un espacio vectorial de dimensión finita, y f un endomorfismo de E. Se definen el *determinante* de f y la *traza* de f respectivamente como los escalares

$$\det f = \det M_{\mathbf{u}}(f), \quad \operatorname{tr} f = \operatorname{tr} M_{\mathbf{u}}(f),$$

siendo $\mathbf{u}$ una base cualquiera de E.

Ejemplo 2.5.5. Sean $A = \left(\begin{smallmatrix} 2 & 3 \\ 4 & 5 \end{smallmatrix}\right)$, $P = \left(\begin{smallmatrix} 1 & 2 \\ 2 & 3 \end{smallmatrix}\right)$, entonces

$$B = P^{-1} \cdot A \cdot P = \left(\begin{smallmatrix} 4 & 7 \\ 2 & 3 \end{smallmatrix}\right)$$

Las matrices A y B tienen la misma traza y el mismo determinante.

Proposición 2.5.6. Sea E un espacio vectorial sobre $\mathbf{K}$.

1. El conjunto $\operatorname{End}(E)$, junto con las operaciones de espacio vectorial sobre $\mathbf{K}$ y la operación de composición

$$\operatorname{End}(E) \times \operatorname{End}(E) \longrightarrow \operatorname{End}(E), \quad (g, f) \mapsto g \circ f,$$

es una K-álgebra (no conmutativa), es decir, $\mathrm{End}\,(E)$ es un espacio vectorial sobre $\mathbf{K}$ tal que, para todo $g, f, g_1, g_2, f_1, f_2 \in \mathrm{End}\,(E), \lambda \in \mathbf{K}$, se cumple

$$\begin{aligned} g \circ (f_1 + f_2) &= g \circ f_1 + g \circ f_2, \\ (g_1 + g_2) \circ f &= g_1 \circ f + g_2 \circ f, \\ g \circ (\lambda \cdot f) &= \lambda \cdot (g \circ f) = (\lambda \cdot g) \circ f \end{aligned}$$

2. Si además suponemos que E tiene dimensión n y fijamos una base $\mathbf{u}$ en E, la aplicación $f \mapsto M_{\mathbf{u}}(f, \mathbf{u})$, que denotaremos simplemente por $M_{\mathbf{u}}$ para poner énfasis en que solo consideramos una base,
$$M_{\mathbf{u}} : \mathrm{End}\,(E) \longrightarrow \mathbf{K}(n, n),$$
es un isomorfismo de $\mathbf{K}$-álgebras, es decir, es un isomorfismo lineal que transforma la composición en producto:
$$M_{\mathbf{u}}(g \circ f) = M_{\mathbf{u}}(g) \cdot M_{\mathbf{u}}(f)$$
En particular, para todo entero $p \geq 0$, y para f inversible,
$$\begin{aligned} M_{\mathbf{u}}(f^p) &= M_{\mathbf{u}}(f)^p \\ M_{\mathbf{u}}(f^{-1}) &= M_{\mathbf{u}}(f)^{-1} \end{aligned}$$
3. Si E tiene dimensión finita, la aplicación $\mathrm{tr} : \mathrm{End}\,(E) \longrightarrow \mathbf{K}$ es lineal.

Definición 2.5.7. Los endomorfismos de E que además son isomorfismos se llaman *automorfismos* de E. El subconjunto de $\mathrm{End}\,(E)$ formado por los automorfismos de E se denota por $GL(E)$ (*grupo general lineal de* E), o $Aut(E)$ (*grupo de automorfismos de* E). En el caso $E = \mathbf{K}^n$, este conjunto se denota por $GL(n)$, y coincide con el conjunto de matrices cuadradas inversibles de orden n.

Proposición 2.5.8. Sea E un espacio vectorial.

1. $Aut(E)$ es un grupo (no conmutativo si $\dim E > 1$).
2. Si $\dim E = n$, y $\mathbf{u}$ es una base, entonces
$$M_{\mathbf{u}} : Aut(E) \longrightarrow GL(n)$$
es un isomorfismo de grupos.
3. Si $\dim E = n$, la aplicación $\det : GL(E) \longrightarrow \mathbf{K} \setminus \{0\}$ es un morfismo de grupos multiplicativos.

Ejemplo 2.5.9. Sea $t \in \mathbb{R}$. Una rotación del plano euclídeo de ángulo $t \in \mathbb{R}$ es el endomorfismo de $\mathbb{R}^2$ cuya matriz en la base canónica es $R_t = \left(\begin{smallmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{smallmatrix}\right)$. La matriz R_{-t} es su inversa, por tanto R_t es un automorfismo. Sus potencias son $(R_t)^n = R_{nt}$, para todo $n \in \mathbb{Z}$.

Ejemplo 2.5.10. Sea $t \in \mathbb{R}$. Una rotación del plano de Lorentz de ángulo $t \in \mathbb{R}$ es el endomorfismo de $\mathbb{R}^2$ cuya matriz en la base canónica es $\Omega_t = \left(\begin{smallmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{smallmatrix}\right)$. La matriz Ω_{-t} es su inversa, por tanto Ω_t es un automorfismo. Sus potencias son $(\Omega_t)^n = \Omega_{nt}$, para todo $n \in \mathbb{Z}$.

2.6. Espacio dual.

2.6. Espacio dual. El estudio del espacio dual de un espacio vectorial se enmarca en un contexto más amplio, que es el de las funciones sobre un conjunto. Si X es un conjunto, una función definida sobre X es una aplicación que toma valores en un conjunto numérico, por ejemplo, el conjunto de los números reales, $f : X \longrightarrow \mathbb{R}$. Dado un conjunto X las funciones permiten definir características de los elementos del conjunto.

En álgebra lineal, el universo X de objetos que queremos medir es siempre un espacio vectorial E. Una función lineal sobre E se llama forma lineal. La mayoría de las funciones no son formas.

Por ejemplo, la longitud de un vector

$$\mathbb{R}^3 \longrightarrow \mathbb{R}, \quad (x,y,z) \mapsto \sqrt{x^2+y^2+z^2}$$

es una función, pero no es lineal. Ejemplos de formas lineales sobre $\mathbb{R}^3$ son las funciones cooordenadas, por ejemplo, la primera coordenada

$$\mathbb{R}^3 \longrightarrow \mathbb{R}, \quad (x,y,z) \mapsto x.$$

Si E es un espacio vectorial sobre $\mathbf{K}$, una función lineal $\omega : E \longrightarrow \mathbf{K}$ se llama una *forma lineal* sobre E, o también *covector*. Las formas lineales son una de las primeras expresiones que hemos encontrado en álgebra lineal. En el estudio de un sistema de ecuaciones lineales, por ejemplo en 2 variables,

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

las expresiones de los términos de la izquierda $2x - y$, $x + 3y$ son formas lineales dobre $\mathbb{R}^2$. Cuando manipulamos las ecuaciones por el método de Gauss para resolver el sistema estamos considerando las ecuaciones como vectores. Es en el contexto de las formas lineales donde esta manipulación tiene un significado preciso.

Si E es un espacio vectorial la aplicación traza definida sobre el espacio vectorial de los endomorfismos de E,

$$\mathrm{tr} : \mathrm{End}(E) \longrightarrow \mathbf{K}$$

es una forma lineal de $\mathrm{End}(E)$ y nos da una información de un endomorfismo que no depende de ninguna base. La función determinante también es una función escalar sobre $\mathrm{End}(E)$, pero no es lineal, sino multilineal, que es un concepto de álgebra lineal que no trataremos en este curso mas que tangencialmente.

2.6.1. **Formas lineales y el espacio dual.** Sea E un espacio vectorial sobre $\mathbf{K}$.

Definición 2.6.1. Se llama *forma lineal* de E a toda aplicación lineal

$$\omega : E \longrightarrow \mathbf{K}.$$

El espacio vectorial $\mathcal{L}(E, \mathbf{K})$ de todas las formas lineales de E se llama el *espacio dual de E*, y se denota por E^*. El espacio dual de E^* se llama el espacio bidual de E, y se denota por E^{**}.

Notación. Si $\omega \in E^*$, $x \in E$, la evaluación $\omega(x)$ de ω en x se denota habitualmente por

$$\langle \omega, x \rangle = \omega(x).$$

Otras notaciones usadas son $\langle \omega | x \rangle$ o $\omega \cdot x$.

Observación 2.6.2. La linealidad de ω se expresa por la regla:

$$\langle \omega, \lambda \cdot x + \mu \cdot y \rangle = \lambda \cdot \langle \omega, x \rangle + \mu \cdot \langle \omega, y \rangle, \quad \forall\, x, y \in E, \forall\, \lambda, \mu \in \mathbf{K}.$$

Por otra parte, fijado $x \in E$, la aplicación de evaluación en x define una aplicación lineal

$$\langle \ \ , x \rangle : E^* \longrightarrow \mathbf{K}, \quad \omega \mapsto \langle \omega, x \rangle,$$

pues se cumple

$$\langle \lambda \cdot \omega + \mu \cdot \theta, x \rangle = \lambda \cdot \langle \omega, x \rangle + \mu \cdot \langle \theta, x \rangle, \quad \forall\, \omega, \theta \in E^*, \forall\, \lambda, \mu \in \mathbf{K},$$

por definición de las operaciones $\omega + \theta$ y $\lambda \cdot \omega$ en E^*. Por tanto, $\langle \ \ , x \rangle$ es un elemento de E^{**}.

Definición 2.6.3. La aplicación *bilineal*

$$\langle \ \ , \ \ \rangle : E^* \times E \longrightarrow \mathbf{K}, \quad (\omega, x) \mapsto \langle \omega, x \rangle,$$

se llama el *apareamiento* canónico.

Observación 2.6.4. Supongamos que $E = \mathbf{K}$. Puesto que en $\mathbf{K}$ tenemos una base canónica, la formada por $\{1\}$, su espacio dual $\mathbf{K}^* = \mathcal{L}(\mathbf{K}, \mathbf{K}) = \mathrm{End}(\mathbf{K})$ se identifica canónicamente con $\mathbf{K}$ con el siguiente isomorfismo:

$$\mathbf{K}^* \longrightarrow \mathbf{K}, \quad \omega \mapsto \langle \omega, 1 \rangle.$$

A partir de ahora usaremos esta identificación cuando sea conveniente.

Ejemplo 2.6.5.

1. Para cada i, $1 \leq i \leq n$, la función coordenada i-ésima de $\mathbb{R}^n$,
$$e_i^* : \mathbb{R}^n \longrightarrow \mathbb{R}, \quad \mathbf{x} = (x_1, \ldots, x_i, \ldots, x_n) \mapsto x_i$$
es una forma lineal. Su matriz en las bases canónicas es $\begin{pmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \end{pmatrix}$, donde 1 ocupa la columna i-ésima.
2. Toda matriz $\omega = \begin{pmatrix} \omega_1 & \cdots & \omega_n \end{pmatrix} \in \mathbf{K}(1, n)$ define una forma lineal de $\mathbf{K}^n$,
$$\mathbf{K}^n \longrightarrow \mathbf{K}, \quad X = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} \mapsto \omega \cdot X = \omega_1 \cdot X_1 + \cdots + \omega_n \cdot X_n.$$
Se tiene que ω es una combinación lineal de las funciones coordenadas, $\omega = \sum_i \omega_i e_i^*$.
3. Toda forma lineal ω de $\mathbf{K}^n$ está definida por una matriz de $\mathbf{K}(1, n)$. Así, si identificamos los vectores de $\mathbf{K}^n$ con "vectores columna", $\mathbf{K}^n = \mathbf{K}(n, 1)$, las formas lineales sobre $\mathbf{K}^n$ se identifican con las "matrices fila", $\mathbf{K}(1, n)$, es decir identificamos $\mathbf{K}(n, 1)^*$ con $\mathbf{K}(1, n)$.
   Observemos también que el núcleo de la forma lineal ω es precisamente el conjunto de soluciones de la ecuación lineal homogénea
$$\omega_1 \cdot X_1 + \cdots + \omega_n \cdot X_n = 0.$$
   Podemos interpretar las transformaciones elementales de las ecuaciones lineales tratadas en el curso de Matrius i vectors como combinaciones lineales de las correspondientes formas lineales.
4. Sea $I = [a, b]$ un intervalo de $\mathbb{R}$. Si $E = \mathcal{F}(I, \mathbb{R})$, para cada $x_0 \in I$, la aplicación de evaluación
$$e_{x_0} : E \longrightarrow K, \quad f \mapsto f(x_0),$$
es una forma lineal.
5. Si $E = \mathcal{C}^0([a, b], \mathbb{R})$, la aplicación integral
$$\mathcal{C}^0([a, b], \mathbb{R}) \longrightarrow \mathbb{R}, \quad f \mapsto \int_a^b f(x)dx,$$
es una forma lineal de E.
6. Si $E = \mathcal{C}^\infty([a, b], \mathbb{R})$, la aplicación derivada n-ésima en un punto $x_0 \in (a, b)$,
$$\mathcal{C}^0([a, b], \mathbb{R}) \longrightarrow \mathbb{R}, \quad f \mapsto D^n(f)(x_0)$$
es una forma lineal de E.

2.6.2. **Base dual.** La base dual $\mathbf{u}^*$ de una base $\mathbf{u}$ son las funciones coordenadas asociadas a la base $\mathbf{u}$, y es una de las piezas clave de la teoría de la dualidad.

Teorema 2.6.6. Si E es un espacio vectorial de dimensión finita, y $B = \{u_1, \ldots, u_n\}$ es una base de E, para cada i sea u_i^* la única forma lineal tal que

$$\langle u_i^*, u_j \rangle = \delta_{ij}.$$

Entonces se cumple:

1. Para todo $\omega \in E^*$, $\omega = \langle \omega, u_1 \rangle u_1^* + \cdots + \langle \omega, u_n \rangle u_n^*$, es decir $M_{\mathbf{u}^*}(\omega) = M(\omega; \mathbf{u})^t$.
2. $B^* = \{u_1^*, \ldots, u_n^*\}$ es una base de E^*, llamada la *base dual de la base B*.

3. Para todo $x \in E$, $x = \langle u_1^*, x\rangle u_1 + \cdots + \langle u_n^*, x\rangle u_n$, es decir u_i^* es la función "coordenada i-ésima respecto de la base $\mathbf{u}$".
4. Si $(f, A) : (E, \mathbf{u}) \longrightarrow (F, \mathbf{v})$, entonces $a_i^j = \langle v_j^*, f(u_i)\rangle$. (Aquí a_i^j denota el elemento que ocupa la columna i y la fila j en la matriz A.)
5. La aplicación
$$\Phi : E \longrightarrow E^{**}, \quad x \mapsto \langle \quad , x\rangle$$
es un isomorfismo lineal.

Demostración. Veamos (1). Sea $x = \sum_{i=1}^n x_i u_i$, entonces
$$\langle u_j^*, x\rangle = \left\langle u_j^*, \sum_{i=1}^n x_i u_i \right\rangle = \sum_{i=1}^n x_i \langle u_j^*, u_i\rangle = x_j.$$

Veamos (2). Para todo j se cumple
$$\left\langle \sum_{i=1}^n \langle \omega, u_i\rangle u_i^*, u_j \right\rangle = \sum_{i=1}^n \langle \omega, u_i\rangle \langle u_i^*, u_j\rangle = \langle \omega, u_j\rangle,$$
por tanto $\sum_{i=1}^n \langle \omega, u_i\rangle u_i^*$ y w coinciden sobre los elementos de la base. Por ser aplicaciones lineales, ambas son iguales.

Veamos (3). Por (2), $\{u_i^*\}_{1\le i\le n}$ genera E^*. Como $\dim E^* = \dim E$, $\{u_i^*\}_i$ es base de E^*.

Veamos (4). Por definición de A, a_i^j es la coordenada j-ésima de $f(u_i)$ en la base $\mathbf{v}$, es decir, por (1), $a_i^j = \langle v_j^*, f(u_i)\rangle$.

Veamos (5) La linealidad de Φ es la linealidad del apareamiento $\langle \omega, x\rangle$ en la variable x. Puesto que $\dim E = \dim E^{**}$, basta ver la inyectividad. Sea $x \in E$. Si $\langle \omega, x\rangle = 0$, para todo $\omega \in E^*$, entonces $\langle u_i^*, x\rangle = 0$ para todo i. Por tanto $x = 0$.

□

2.7. Aplicación dual y ortogonalidad.

2.7.1. **Aplicación traspuesta de una aplicación lineal.** La trasposición de matrices tiene su avatar en el contexto de las aplicaciones lineales.

Definición 2.7.1. Si $f : E \longrightarrow F$ es una aplicación lineal entre dos espacios vectoriales E y F, se define su aplicación *dual* o *transpuesta* $f^* : F^* \longrightarrow E^*$ por
$$f^*(\beta) = \beta \circ f, \quad \forall \beta \in F^*,$$
es decir,
$$\langle f^*(\beta), x\rangle = \langle \beta, f(x)\rangle, \quad \forall \beta \in F^*, x \in E$$

Proposición 2.7.2. Sean E y F espacios vectoriales.

1. Si $f : E \longrightarrow F$ es lineal, entonces $f^* : F^* \longrightarrow E^*$ es lineal.
2. La aplicación
$$\mathcal{L}(E, F) \longrightarrow \mathcal{L}(F^*, E^*), \quad f \mapsto f^*$$
es lineal, es decir
$$(\lambda f + \mu g)^* = \lambda f^* + \mu g^*, \quad \forall\, f, g \in \mathcal{L}(E, F), \forall\, \lambda, \mu \in \mathbf{K}.$$

Ejemplo 2.7.3. Sea $x \in E$, y sea $L_x : \mathbf{K} \longrightarrow E$ definida por $L_x(\lambda) = \lambda x$. Sea

$$(L_x)^* : E^* \longrightarrow \mathbf{K}^* \cong \mathbf{K}$$

su dual. Si $\omega \in F^*$, via el isomorfismo canónico $\mathbf{K}^* \to \mathbf{K}$, la forma lineal $L_x(\omega) \in K^*$ se identifica con $\langle L_x(\omega), 1\rangle = \langle \omega, x\rangle$, es decir $(L_x)^*$ se identifica con la aplicación de evaluación $\langle \quad, x\rangle$.

Ejercicio. Sea $A \in \mathbf{K}(m,n)$, y consideremos las aplicaciones lineales

$$L_A : \mathbf{K}(n,1) \longrightarrow \mathbf{K}(m,1), \quad R_A : \mathbf{K}(1,m) \longrightarrow \mathbf{K}(1,n)$$

definidas por

$$L_A(X) = A \cdot X,\ \forall X \in \mathbf{K}(n,1); \quad R_A(W) = W \cdot A,\ \forall W \in \mathbf{K}(1,m).$$

Prueba que, mediante el isomorfismo $\mathbf{K}(p,1)^* \cong \mathbf{K}(1,p)$, la aplicación $(L_A)^*$ se identifica con R_A.

Las propiedades multiplicativas del operador de trasposición son análogas a las del operador de trasposición de matrices:

Proposición 2.7.4.
1. $(Id_E)^* = Id_{E^*}$
2. $(g \circ f)^* = f^* \circ g^*$.
3. Si f es inversible, f^* es invertible y $(f^*)^{-1} = \left(f^{-1}\right)^*$.

Demostración. Veamos (2). Sean $f : E \longrightarrow F$, y $g : F \longrightarrow G$ aplicaciones lineales. Entonces $g \circ f : E \longrightarrow G$ cumple que, para todo $\omega \in G^*$,

$$(g \circ f)^*(\omega) = \omega \circ (g \circ f) = (\omega \circ g) \circ f = g^*(\omega) \circ f = f^*(g^*(\omega)) = (f^* \circ g^*)(\omega).$$

por tanto $(g \circ f)^* = f^* \circ g^*$. □

Proposición 2.7.5. Sean E y F dos espacios vectoriales de dimensión finita, y $\mathbf{u}$, $\mathbf{v}$ bases de E y F respectivamente. Sea $f : E \longrightarrow F$ una aplicación lineal y A la matriz de f en las bases $(\mathbf{u}, \mathbf{v})$, es decir,

$$(f, A) : (E, \mathbf{u}) \longrightarrow (F, \mathbf{v}).$$

Entonces

$$(f^*, A^t) : (F^*, \mathbf{v}^*) \longrightarrow (E^*, \mathbf{u}^*).$$

Demostración. Sea B la matriz de f^* en las bases $(\mathbf{v}^*, \mathbf{u}^*)$. Entonces $B_i^j = \langle f(v_i^*), u_j\rangle = \langle v_i^*, f(u_j)\rangle A_j^i$, por tanto $B = A^t$. □

Proposición 2.7.6. Sean $\mathbf{u}, \mathbf{v}$ son dos bases de E y sea $P = M_{\mathbf{u}}(\mathbf{v})$ la matriz de cambio de base, entonces

$$M_{\mathbf{v}^*}(\mathbf{u}^*) = P^t$$

Demostración. Puesto que $M_{\mathbf{u}}(\mathbf{v}) = M_{\mathbf{u}}(Id_E; \mathbf{v})$, resulta

$$M_{\mathbf{v}^*}(\mathbf{u}^*) = M_{\mathbf{v}^*}((Id_E)^*; \mathbf{u}^*) = M_{\mathbf{u}}(Id_E; \mathbf{v})^t = M_{\mathbf{u}}(\mathbf{v})^t = P^t$$

□

Ejercicio. Sea $\mathbf{e}$ la base canónica de $\mathbb{R}^3$, y $\mathbf{v}$ la base definida por

$$v_1 = (1,3,1),\ v_2 = (0,2,1),\ v_3 = (-1,2,2).$$

Halla la base dual $\mathbf{v}^*$ en función de la dual de la base canonica $\mathbf{e}^*$.

Ejercicio 2.7.7.

1. Si $f : F \longrightarrow E$ es una aplicación lineal inyectiva, entonces su aplicación traspuesta $f^* : E^* \longrightarrow F^*$, es una aplicación exhaustiva. Más precisamente, si G es un suplemento de F en E, toda forma lineal $w \in F^*$ se extiende a una forma lineal $\widetilde{w} \in E^*$, tal que $\widetilde{w}_{|G} = 0$.
2. Si $E = F \oplus G$, y $i_F : F \longrightarrow E$, $i_G : G \longrightarrow E$ denotan las aplicaciones de inclusión, entonces las aplicaciones i_F^*, i_G^* inducen un isomorfismo

$$i_F^* \times i_G^* : E^* \longrightarrow F^* \times G^*, \quad w \mapsto (i_F^*(w), i_G^*(w)).$$

2.7.2. **Ortogonalidad.**

Definición 2.7.8. Sea $A \subset E$ un subconjunto de E, se define su ortogonal por

$$A^\perp = \{w \in E^*;\ \langle \omega, x\rangle = 0, \forall x \in A\ \}.$$

Sea $\Omega \subset E^*$ un subconjunto de E^*, se define su ortogonal por

$$\Omega^\perp = \{x \in E;\ \langle \omega, x\rangle = 0, \forall \omega \in \Omega\ \}.$$

Proposición 2.7.9. Sea E un espacio vectorial sobre $\mathbf{K}$.

(1) $\{0\}^\perp = E^*$, $E^\perp = \{0\}$.
(2) Si $A \subset B \subset E$, entonces $A^\perp \supset B^\perp$.
(3) $A^\perp$ es un subespacio vectorial de E^*, para todo $A \subset E$.
(4) Si F es un subespacio de E, el núcleo de la aplicación exhaustiva $E^* \longrightarrow F^*$ es $F^\perp$ y $F^* \cong E^*/F^\perp$. En particular, si la dimensión de E es finita, $\dim F + \dim F^\perp = \dim E$.
(5) $(A^\perp)^\perp = \langle A\rangle$, para todo $A \subset E$.
(6) Si A, B son subespacios vectoriales de E, entonces $(A \cap B)^\perp = A^\perp + B^\perp$, $(A + B)^\perp = A^\perp \cap B^\perp$.
(7) Si $f : E \longrightarrow F$ es una aplicación lineal, entonces
   (a) $(\operatorname{Im} f)^\perp = \operatorname{Ker} f^*$
   (b) $\operatorname{Ker} f = (\operatorname{Im} f^*)^\perp$
   (c) $(\operatorname{Ker} f)^\perp = \operatorname{Im} f^*$,
   (d) $\operatorname{Im} f = (\operatorname{Ker} f^*)^\perp$

Demostración. Las propiedades (1) y (2) se deducen inmediatamente de las definiciones.

(3) Se sigue de la fórmula $A^\perp = \bigcap\limits_{x\in A} \operatorname{Ker}\langle\ , x\rangle$.

(4) Sea $w \in E^*$. Entonces $w \in E^*$ es cero en F^*, si y solo si $w \cdot F = 0$, es decir $w \in F^\perp$

(5) En primer lugar observemos que $A^\perp = \langle A\rangle^\perp$. En efecto, la inclusión $A^\perp \subset \langle A\rangle^\perp$ se sigue de (2). Veamos la otra inclusión. Si $x_1, \dots, x_r \in A$, $\omega \in A^\perp$, entonces $\langle \omega, \sum_i \lambda_i x_i\rangle = \sum_i \lambda_i \langle \omega, x_i\rangle = 0$, para toda familia de escalares $\lambda_1, \dots, \lambda_r$. Por tanto $\omega \in \langle A\rangle^\perp$. Ahora, para probar (5), podemos suponer que A es un subespacio vectorial. Sea $x \in A$, veamos que $x \in A^{\perp\perp}$. En efecto, para todo $\omega \in A^\perp$ se cumple $\langle \omega, x\rangle = 0$, por tanto$\langle A^\perp, x\rangle = 0$, y así $x \in A^{\perp\perp}$. Si E es de dimensión finita podemos concluir utilizando un cálculo de dimensiones, teniendo en cuenta la fórmula del apartado (4). Sino, veamos que $x \in A^{\perp\perp}$ implica $x \in A$. En efecto, en caso contrario, el subespacio $A \oplus \langle\{x\}\rangle$ tendría un suplemento B tal que $A \oplus \langle\{x\}\rangle \oplus B = E$ y prodríamos definir $w \in E^*$ de modo que $w(A) = 0$, $w(x) \neq 0$, $w(B) = 0$. Ahora bien, como $x \in A^{\perp\perp}$, $w(x) = 0$ para todo $w \in A^\perp$, lo cual es una contradicción.

(6) De (2) se sigue que $A^\perp \subset (A \cap B)^\perp$, $B^\perp \subset (A \cap B)^\perp$, por tanto $A^\perp + B^\perp \subset (A \cap B)^\perp$. Si E es de dimensión finita se puede concluir con la fórmula de Grassmann. En caso contrario,

podemos usar 1.7.8 y el ejercicio 2.7.7. En efecto, sea $I = A \cap B$, $S = A + B$. Entonces existe subespacios A_1, B_1, S_1 de E tales que

$$I \oplus A_1 = A, \quad I \oplus B_1 = B, \quad S \oplus S_1 = E.$$

Por tanto, $A \oplus B_1 \oplus S_1 = E$. Ahora, dada $w \in (A \cap B)^\perp$, la restricción de w a A tiene una extensión w_1 a E que vale cero sobre $B_1 \oplus S_1$. Entonces $w_1 \in B^\perp$, y $w_2 := w - w_1 \in A^\perp$. Por tanto $w = w_1 + w_2 \in A^\perp + B^\perp$.

De (2) se sigue que $(A+B)^\perp \subset A^\perp$, $(A+B)^\perp \subset B^\perp$, por tanto $(A+B)^\perp \subset A^\perp \cap B^\perp$. Por otra parte, si $w \in A^\perp \cap B^\perp$, se tiene $w \cdot A = 0 = w \cdot B$, por tanto $w \cdot (A+B) = 0$, es decir $w \in (A+B)^\perp$.

(7) Veamos (a). Sea $\theta \in F^*$. Entonces $\theta \in (\operatorname{Im} f)^\perp$ si y solo si $\langle \theta, f(x) \rangle = 0$, para todo $x \in E$, lo cual equivale a $\langle f^*(\theta), x \rangle = 0$, para todo $x \in E$, es decir a $f^*(\theta) = 0$.

Veamos (b). Sea $x \in E$. Entonces $x \in (\operatorname{Im} f^*)^\perp$ si y solo si $\langle f^*(\theta), x \rangle = 0$ para todo $\theta \in F^*$, es decir, $\langle \theta, f(x) \rangle = 0$ para todo $\theta \in F^*$, lo cual equivale a $f(x) = 0$, es decir, $x \in \operatorname{Ker} f$.

Veamos (c). En efecto, la inclusión $\operatorname{Im} f^* \subset (\operatorname{Ker} f)^\perp$ es sencilla, pues, si $\theta \in F^*$, se tiene $\langle f^*(\theta), x \rangle = \langle \theta, f(x) \rangle = 0$, para todo $x \in \operatorname{Ker} f$. Para ver la otra inclusión, sea $w \in (\operatorname{Ker} f)^\perp$. Queremos hallar $\theta \in F^*$ tal que $f^*(\theta) = w$. Para ello observamos que $\operatorname{Ker} w \subset \operatorname{Ker} f$, por tanto w induce una forma lineal $\widetilde{w}$ sobre el espacio cociente $E/\operatorname{Ker} f$. Componiendo con el inverso del isomorfismo $\widehat{f} : E/(\operatorname{Ker} f) \longrightarrow \operatorname{Im} f$ obtenemos una forma lineal sobre $\operatorname{Im} f$ que, por 2.7.7 se extiende a una forma lineal θ sobre F que cumple $f^*(\theta) = w$.

Veamos (d). Se tiene

$$\operatorname{Im} f = (\operatorname{Im} f)^{\perp\perp} = (\operatorname{Ker} f^*)^\perp.$$

□

Ejercicio 2.7.10. Sea E un espacio vectorial sobre $\mathbf{K}$.

(1) $\{0_{E^*}\}^\perp = E$ y $E^{*\perp} = \{0\}$.
(2) Si $A \subset B \subset E^*$, entonces $A^\perp \supset B^\perp$.
(3) $A^\perp$ es un subespacio vectorial de E, para todo $A \subset E^*$.
(4) Si la dimensión de E es finita, y A es un subespacio vectorial de E^*, $\dim A + \dim A^\perp = \dim E$.
(5) $A \subset (A^\perp)^\perp$, para todo subespacio A de E^*. Si E tiene dimensión finita, se cumple la igualdad.
(6) Si A, B son subespacios vectoriales de E^*, entonces $(A+B)^\perp = A^\perp \cap B^\perp$ y $(A \cap B)^\perp \supset A^\perp + B^\perp$. Si E es de dimensión finita, se cumple la igualdad.

Observación 2.7.11. Las propiedades $A^{\perp\perp} = A$ y $(A \cap B)^\perp = A^\perp + B^\perp$ son falsas en general, para A, B subespacios de E^*. Veamos contraejemplos.

Sea $E = \mathbb{R}[X]$, y $e_i = X^i$, entonces las formas lineales $\{e_i^*; i \in \mathbb{N}\}$ generan un subespacio vectorial H de E^* tal que $H \neq E^*$, pero $H^\perp = 0$. En particular, $E^* = H^{\perp\perp} \neq H$.

Sea $E = \mathbb{R}[X]$, $w \in E^*$ definida por $w \cdot e_i = 1$, para todo i, $w_p \in E^*$ definida por $w \cdot e_{2i} = 1$, para todo i. Sea $A = \langle \{e_i^*; i > 0\} \rangle + \langle w \rangle$, $B = \langle \{e_i^*; i > 0\} \rangle + \langle w_p \rangle$. Entonces $A^\perp + B^\perp = 0$, y $e_0 \in (A \cap B)^\perp \neq 0$, luego $(A \cap B)^\perp \neq A^\perp + B^\perp$.

Observación 2.7.12. Sea E un espacio vectorial de dimensión n. Si $\Omega = \{\omega_1, \ldots, \omega_r\}$ es un conjunto de formas lineales de E, el conjunto de las soluciones del sistema de ecuaciones lineales $\omega_i(x) = 0$, $1 \leq i \leq r$, es decir,

$$\Omega^\perp = \{x \in E;\ \omega_1(x) = 0, \ldots, \omega_r(x) = 0\,\},$$

es un subespacio vectorial de E tal que $\dim \Omega^\perp = n - \mathrm{rg}\Omega$.

Recíprocamente si A es un subconjunto de E, entonces $A^\perp$ es un subespacio vectorial de E^* tal que $\dim A^\perp = n - \mathrm{rg}A$. Si $\omega_1, \dots \omega_p$ es una base de $A^\perp$, entonces $\langle A \rangle$ se expresa como un conjunto de soluciones de un sistema de ecuaciones lineales en $x \in E$:

$$\langle A \rangle = \{x \in E; \omega_1(x) = 0, \dots, \omega_p(x) = 0\,\}.$$

Ejemplo 2.7.13. Sea F el subespacio vectorial de $\mathbb{R}^3$ generado por los vectores $u_1 = (1, 0, -1), u_2 = (1, 1, 0)$. Entonces $F^\perp$ es el conjunto de formas $(w_1, w_2, w_3) \in (\mathbb{R}^3)^*$ que son solución de

$$\begin{pmatrix} w_1 & w_2 & w_3 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ -1 & 0 \end{pmatrix} = 0$$

Resolviendo el sistema obtenemos que $w = (1, -1, 1)$ es una base de $F^\perp$, por tanto

$$F = \{(x, y, z) \in \mathbb{R}^3;\ x - y + z = 0\ \}.$$

Observemos que en este caso, si $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ -1 & 0 \end{pmatrix}$ se cumple $F = \operatorname{Im} A$, y el cálculo anterior corresponde a la fórmula

$$(\operatorname{Im} A)^\perp = \operatorname{Ker} A^t.$$

3. Diagonalización de endomorfismos

3.1. Introducción. Sea A una matriz cuadrada, por ejemplo de tipo 2×2, con coeficientes en un cuerpo $\mathbf{K}$, por ejemplo $\mathbb{R}$. Uno de los objetivos de este capítulo es abordar el cálculo de las potencias A^p de A.

Para una matriz diagonal,

$$D(\lambda_1, \lambda_2) = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix},$$

sus potencias también son matrices diagonales, y se cumple

$$D^p = D(\lambda_1^p, \lambda_2^p)$$

para todo entero $p \geq 0$.

Aunque una matriz A no sea diagonal es posible que se transforme en una matriz diagonal después de efectuar un cambio de base. Es decir, es posible que exista una matriz invertible U, de tipo 2×2, tal que el producto $D = U^{-1} \cdot A \cdot U$ sea una matriz diagonal $D = D(\lambda_1, \lambda_2)$. Se dice entonces que la matriz A es diagonalizable. En tal caso, de la ecuación $D = U^{-1} \cdot A \cdot U$ deducimos que

$$A = U \cdot D \cdot U^{-1}$$

y por tanto

$$A^p = U \cdot D^p \cdot U^{-1} = U \cdot D(\lambda_1^p, \lambda_2^p) \cdot U^{-1},$$

lo cual resuelve el problema del cálculo de las potencias de A, en el caso de que A sea diagonalizable.

Ejemplo 3.1.1. Por ejemplo, si $A = \left(\begin{smallmatrix} 11 & -8 \\ 16 & -13 \end{smallmatrix}\right)$, y $U = \left(\begin{smallmatrix} 1 & 1 \\ 1 & 2 \end{smallmatrix}\right)$, entonces mediante el cálculo explícito podemos comprobar que

$$A \cdot U = U \cdot D(3, -5),$$

es decir, $U^{-1} \cdot A \cdot U = D(3, -5)$ y A es diagonalizable. Por tanto

$$A = U \cdot D(3, -5) \cdot U^{-1}$$

y, puesto que

$$U^{-1} = \left(\begin{smallmatrix} 2 & -1 \\ -1 & 1 \end{smallmatrix}\right),$$

se tiene

$$A^p = U \cdot D(3^p, (-5)^p) \cdot U^{-1} = \left(\begin{smallmatrix} 2\cdot 3^p - (-5)^p & -3^p + (-5)^p \\ 2\cdot 3^p - 2\cdot(-5)^p & -3^p + 2\cdot(-5)^p \end{smallmatrix}\right)$$

Para entender como afectan los cambios de base a la matriz A es útil considerar el endomorfismo L_A de $\mathbf{K}^2$ asociado a A. En efecto, dada una matriz invertible U, sus columnas forman una base $\mathbf{u} = (u_1, u_2)$ de $\mathbf{K}^2$, para la cual U es la matriz de cambio de base, de la base canónica $\mathbf{e}$ a la base $\mathbf{u}$,

$$U = M_{\mathbf{e}}(\mathbf{u}),$$

y, gracias a la fórmula de cambio de base, la matriz $U^{-1} \cdot A \cdot U$ es la matriz del endomorfismo L_A en la base $\mathbf{u}$,

$$M_{\mathbf{u}}(L_A) = M_{\mathbf{e}}(\mathbf{u})^{-1} \cdot M_{\mathbf{e}}(L_A) \cdot M_{\mathbf{e}}(\mathbf{u}) = U^{-1} \cdot A \cdot U.$$

El argumento anterior conduce a dos preguntas naturales:

Pregunta 3.1.2. ¿Existe una base $\mathbf{u} = (u_1, u_2)$ de $\mathbf{K}^2$ tal que la matriz de L_A sea una matriz diagonal, $D = D(\lambda_1, \lambda_2)$, para ciertos escalares λ_1, λ_2?

Pregunta 3.1.3. En caso de que la respuesta a la pregunta anterior sea afirmativa, ¿cómo hallar la base $\mathbf{u}$?

La primera observación que conviene destacar es que hemos conseguido formular estas cuestiones en el contexto de los endomorfismos, y no meramente en el contexto de las matrices. La segunda es que, aunque una vez conocida la base $\mathbf{u}$, los escalares λ_1, λ_2 quedan determinados por el producto $U^{-1} \cdot A \cdot U$, veremos que el proceso de cálculo se realiza en el sentido inverso, es decir, en primer lugar se obtienen los posibles escalares λ_1, λ_2 y a continuación la base $\mathbf{u}$.

Para responder a la pregunta 3.1.3, observemos que, puesto que la matriz de L_A en la base $\mathbf{u}$ es la matriz diagonal $D(\lambda_1, \lambda_2)$, los vectores u_1, u_2, y los escalares λ_1, λ_2, han de satisfacer las ecuaciones

$$L_A(u_1) = \lambda_1 \cdot u_1, \quad L_A(u_2) = \lambda_2 \cdot u_2,$$

que son el punto de partida de nuestro estudio.

En el ejemplo 3.1.1, los vectores columna de la matriz U, $u_1 = (1, 1)$ y $u_2 = (1, 2)$, satisfacen, respectivamente,

$$\begin{pmatrix} 11 & -8 \\ 16 & -13 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 3 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 11 & -8 \\ 16 & -13 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} = -5 \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix},$$

es decir,

$$A \cdot u_1 = 3 \cdot u_1, \quad A \cdot u_2 = -5 \cdot u_2\,.$$

Observemos que, como consecuencia, para todo entero $p \geq 0$ se cumple

$$A^p \cdot u_1 = 3^p \cdot u_1, \quad A^p \cdot u_2 = (-5)^p \cdot u_2\,.$$

Como u_1, u_2 forman una base de $\mathbf{K}^2$, corroboramos así que la matriz A^p queda determinada por estas ecuaciones.

3.2. Valores y vectores propios. En lo que sigue E denota un espacio vectorial sobre $\mathbf{K}$ y f un endomorfismo de E. El endomorfismo identidad de E lo denotaremos por Id_E, $\mathbf{1}_E$ o simplemente $\mathbf{1}$. La matriz identidad se denota por I o $\mathbf{1}$. Si es preciso denotar su tamaño n escribiremos $\mathbf{1}_n$.

Definición 3.2.1. Sea $P(t) = a_0 + a_1 t + \cdots + a_m t^m \in \mathbf{K}[t]$ un polinomio con coeficientes en $\mathbf{K}$. Definimos el endomorfismo $P(f)$ de E por

$$P(f) = a_0 Id_E + a_1 f + \cdots + a_m f^m.$$

En la práctica el sumando $a_0 Id_E$ se escribirá abreviadamente a_0 si no hay peligro de confusión.

Proposición 3.2.2. 1. Si $P(t), Q(t) \in \mathbf{K}[t]$ son dos polinomios, entonces

$$(P + Q)(f) = P(f) + Q(f), \quad (P \cdot Q)(f) = P(f) \circ Q(f).$$

2. Los endomorfismos $P(f)$ y $Q(f)$ conmutan, es decir $P(f) \circ Q(f) = Q(f) \circ P(f)$.

Ejemplo 3.2.3. (1) Si $P(t) = t - \lambda$, entonces $P(f) = f - \lambda Id_E$.

(2) Si $P(t) = (t - 2)(t - 1) = t^2 - 3t + 2$, entonces

$$P(f) = f^2 - 3f + 2Id_E = (f - 2Id_E) \circ (f - Id_E) = (f - Id_E) \circ (f - 2Id_E).$$

Dados n escalares $\lambda_1, \ldots, \lambda_n$, denotaremos por $D(\lambda_1, \ldots, \lambda_n)$ la matriz diagonal

$$D(\lambda_1, \ldots, \lambda_n) = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

Definición 3.2.4. (Vectores y valores propios.) Un vector no nulo $v \in E$ se llama un *vector propio de* f si existe un escalar $\lambda \in \mathbf{K}$ tal que $f(v) = \lambda \cdot v$. Si v es un vector propio de f, el escalar λ es único y se llama *el valor propio de* v. También se dice que λ es un *valor propio de* f.

Observaciones 3.2.5.

1. La definición de vector propio y de valor propio de un endomorfismo no utiliza bases.
2. En la definición que hemos dado de vector propio se exige que el vector sea distinto de cero. El vector $0 \in E$ cumple la ecuación $f(0) = \lambda \cdot 0$ para todo $\lambda \in \mathbf{K}$, pero 0 no se suele considerar como vector propio.
3. Toda combinación lineal no nula de vectores propios del mismo valor propio λ es un vector propio de valor propio λ.
4. Los vectores no nulos del núcleo de f son vectores propios de valor propio $\lambda = 0$.
5. Un vector propio se llama también un autovector ("eigenvector" en inglés), y un valor propio se llama también un autovalor ("eigenvalue" en inglés). El prefijo "eigen" es una palabra alemana que significa "propio". El conjunto de los valores propios de un endomorfismo f se llama el "espectro" de f, especialmente en el caso de dimensión infinita, caso de gran interés en el Análisis Funcional y en la Mecánica Cuántica. El significado de esta nomenclatura es el siguiente. Espectro significa "lo que se ve"(por ejemplo, la palabra "espectador", el que mira, es un derivado). Por otra parte, los coeficientes de la matriz de f en una base dependen de la base, por tanto no tienen ningún significado si desconocemos la base en que está calculada la matriz. Sin embargo, cuando el endomorfismo es diagonalizable, veremos que su matriz diagonal es única, salvo el orden de los valores propios. Esto significa que el espectro no depende de la base escogida, es un invariante intrínseco del endomorfismo (igual que lo son la traza o el determinante), es decir, es lo único auténtico que vemos de f en su matriz. Este hecho contrasta fuertemente con las componentes de un vector, que, si el espacio vectorial no tiene una estructura adicional, no tienen ningún valor intrínseco. Sin embargo, si, por ejemplo, el espacio vectorial está dotado de un producto escalar, y las componentes de un vector se calculan en una base ortonormal, entonces su módulo se calcula a partir únicamente de las componentes y representa la longitud del vector, que tiene un significado intrínseco en este contexto.

Ejemplo 3.2.6.

1. Sea $\lambda \in \mathbf{K}$. Si $f = \lambda \cdot \mathbf{1}_E$, todo vector no nulo de E es un vector propio de f, de valor propio λ.
2. Si A es una matriz $n \times n$, los vectores propios de L_A (abreviadamente, de A) son los vectores $X \in \mathbf{K}^n = \mathbf{K}(n, 1)$ tales que existe $\lambda \in \mathbf{K}$ con $A \cdot X = \lambda \cdot X$.
3. Si $A = D(\lambda_1, \lambda_2) = \left(\begin{smallmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{smallmatrix}\right)$, entonces
$$A \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \lambda_1 \\ 0 \end{pmatrix} = \lambda_1 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad A \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ \lambda_2 \end{pmatrix} = \lambda_2 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$
por tanto $(1, 0)$ (o cualquier vector no nulo de la forma $(a, 0)$) es un vector propio de valor propio λ_1, y $(0, 1)$ (o cualquier vector no nulo de la forma $(0, b)$) es un vector propio de valor propio λ_2. Si $\lambda_1 \neq \lambda_2$, la matriz A no tiene más vectores propios.
4. Si $A = D(\lambda_1, \dots, \lambda_n)$, para cada $i = 1, \dots, n$, el vector $e_i \in \mathbf{K}^n$ de la base canónica es vector propio de valor propio λ_i.

Ejemplo 3.2.7. Continuemos con el ejemplo 3.1.1. El vector $u_1 = (1, 1) \in \mathbb{R}^2$ es un un vector propio de A valor propio $\lambda_1 = 3$, y $u_2 = (1, 2) \in \mathbb{R}^2$ es un vector propio de A de valor propio $\lambda_2 = -5$. Por tanto $\lambda_1 = 3$ y $\lambda_2 = -5$ son valores propios de A. Veremos más adelante que A no tiene más valores propios.

Ejemplo 3.2.8. Sea
$$A = \left(\begin{smallmatrix} 0 & 2 & 2 \\ 1 & 4 & 2 \\ -1 & -1 & 1 \end{smallmatrix}\right).$$
Los vectores
$$u_1 = (-2, -5, 1), \quad u_2 = (0, 1, -1), \quad u_3 = (-2, 0, 1)$$
satisfacen
$$Au_1 = 4u_1, \quad Au_2 = 2u_2, \quad Au_3 = -u_3.$$
Por tanto u_1 es un vector propio de valor propio 4, u_2 es un vector propio de valor propio 2 y u_3 es un vector propio de valor propio -1,

Ejemplo 3.2.9. Sea
$$B = \left(\begin{smallmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{smallmatrix}\right),$$
entonces
$$v_1 = (-1, 1, 0), \quad v_2 = (-1, 0, 1), \quad v_3 = (1, 1, 1)$$
satisfacen
$$Bv_1 = -v_1, \quad Bv_2 = -v_2, \quad Bv_3 = 2v_3.$$
Por tanto v_1 y v_2 son vectores propios de valor propio -1 y v_3 es un vector propio de valor propio 2,

A continuación vamos a responder a la pregunta siguiente.

Pregunta 3.2.10. ¿Cómo calcular los vectores propios de un valor propio λ dado?

Sea $\lambda \in \mathbf{K}$. La ecuación de los vectores propios de valor propio λ del endomorfismo f es la ecuación
$$f(v) = \lambda \cdot v$$
en la incógnita $v \in E \setminus \{0\}$. Esta ecuación es equivalente a la ecuación
$$f(v) - \lambda \cdot v = 0,$$
es decir, a la condición
$$v \in \operatorname{Ker}(f - \lambda \cdot \mathbf{1}_E).$$
Por tanto $\operatorname{Ker}(f - \lambda \cdot \mathbf{1}_E)$ contiene todos los vectores propios de valor propio λ, y además el vector 0. Esto conduce a la definición siguiente, que responde a la pregunta 3.2.10 anterior.

Definición 3.2.11. (Autoespacios.) El subespacio vectorial $\operatorname{Ker}(f - \lambda \cdot \mathbf{1}_E)$ se llama *autoespacio de f correspondiente a λ* o el λ-*autoespacio* de f. Si E es de dimensión finita, la dimensión de $\operatorname{Ker}(f - \lambda \cdot \mathbf{1}_E)$ se llama *multiplicidad geométrica de λ como valor propio de f*, y la denotaremos usualmente por d_λ.

Observaciones 3.2.12.
1. La palabra inglesa para "autoespacio" es "eigenspace".
2. Para $\lambda \in \mathbf{K}$ arbitrario, el λ-autoespacio de f es nulo, salvo que λ sea un valor propio de f.
3. Si A es una matriz $n \times n$, U una matriz invertible y $B = U^{-1}AU$, entonces la dimensión del λ-autoespacio de A coincide con la del λ-autoespacio de B.

Ejemplo 3.2.13. Si $A = \left(\begin{smallmatrix} 11 & -8 \\ 16 & -13 \end{smallmatrix}\right)$ (ver ejemplo 3.1.1) los únicos autoespacios no nulos de A son los correspondientes a $\lambda_1 = 3$, $\lambda_2 = -5$, que son, respectivamente,
$$\operatorname{Ker}(A - 3 \cdot \mathbf{I}) = \langle (1, 1) \rangle, \quad \operatorname{Ker}(A + 5 \cdot \mathbf{I}) = \langle (1, 2) \rangle.$$

Ejemplo 3.2.14. Continuando con la matriz B del ejemplo 3.2.9, el subespacio propio $\mathrm{Ker}\,(B-(-1)I)$ tiene dimensión 2 y el subespacio propio $\mathrm{Ker}\,(B-2I)$ tiene dimensión 1.

Si $V=\begin{pmatrix} -1 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$, $V^{-1}BV=D=\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$. Los únicos autoespacios no nulos de D son

$$\mathrm{Ker}\,(D-(-1)\cdot \mathbf{I})=\langle (1,0,0),(0,1,0)\rangle,\quad \mathrm{Ker}\,(D-5\cdot \mathbf{I})=\langle (0,0,1)\rangle.$$

Por tanto los autoespacios de B y D no son iguales, pero tienen la misma dimensión.

El resultado que exponemos a continación es un punto clave de la teoría de la diagonalización.

Teorema 3.2.15. Vectores propios de valores propios diferentes son linealmente independientes. Sea f un endomorfismo de un espacio vectorial E (no necesariamente de dimensión finita), $\lambda_1,\dots,\lambda_r$, r escalares distintos, entonces

(1) La suma de los autoespacios

$$\mathrm{Ker}\,(f-\lambda_1 Id_E)+\cdots+\mathrm{Ker}\,(A-\lambda_r Id_E)$$

es directa (alguno de estos subespacios puede ser nulo). En particular, si E es de dimensión finita n, y $d_i=\dim \mathrm{Ker}\,(f-\lambda_i Id_E)$, entonces $\sum\limits_{i=1}^{r} d_i\le n$.

(2) Si $v_1,\cdots,v_r$ son r vectores propios de valores propios $\lambda_1,\dots,\lambda_r$ respectivamente, entonces $\{v_1,\cdots,v_r\}$ es linealmente independiente.

(3) Si $P(t)=\prod\limits_{i=1}^{r}(t-\lambda_i)$, entonces $\mathrm{Ker}\,P(f)=\bigoplus_i \mathrm{Ker}\,(A-\lambda_i Id_E)$.

Demostración. Veamos (1) y (3) por inducción sobre r. Si $r=1$ es trivial. Supongamos el resultado cierto para $r-1$. Para $r\ge 2$, los polinomios $Q_1=t-\lambda_1$, $Q_2=\prod\limits_{i=2}^{r}(t-\lambda_i)$ son primos entre sí, por tanto satisfacen una identidad de Bézout del tipo

$$1=C_1(t)Q_1(t)+C_2(t)Q_2(t),$$

donde $C_1,C_2\in \mathbf{K}[t]$. Además $P(t)=Q_1(t)Q_2(t)$. Por la hipótesis de inducción podemos suponer que

$$\mathrm{Ker}\,Q_2(f)=\mathrm{Ker}\,(f-\lambda_2 Id_E)\oplus\cdots\oplus \mathrm{Ker}\,(f-\lambda_r Id_E),$$

por tanto basta que la suma $\mathrm{Ker}\,Q_2(f)+\mathrm{Ker}\,Q_1(f)$ es directa y que $\mathrm{Ker}\,P(f)=\mathrm{Ker}\,Q_2(f)+\mathrm{Ker}\,Q_1(f)$. Sea $v\in \mathrm{Ker}\,Q_2(f)\cap \mathrm{Ker}\,Q_1(f)$. Entonces, por la identidad de Bézout

$$v=C_1(f)Q_1(f)(v)+C_2(f)Q_2(f)(v)=0,$$

y la suma es directa. Además, si $P(f)(v)=0$ se cumple

$$v=C_1(f)Q_1(f)(v)+C_2(f)Q_2(f)(v)\in \mathrm{Ker}\,Q_2(f)+\mathrm{Ker}\,Q_1(f).$$

por tanto $\mathrm{Ker}\,P(f)\subset \mathrm{Ker}\,Q_2(f)+\mathrm{Ker}\,Q_1(f)$. La otra inclusión es trivial, pues $P(f)(v)=Q_1(f)\circ Q_2(f)(v)=Q_2(f)\circ Q_1(f)(v)$.

(2) se sigue inmediatamente de (1), pues si $0=v_1+\cdots+v_r$, donde $v_i\in \mathrm{Ker}\,(f-\lambda_i I)$, $\forall i$. Los vectores v_i son nulos o bien vectores propios de valores propios diferentes. Por tanto, $v_i=0$, para todo i. $\square$

Corolario 3.2.16. Una matriz de tipo $n\times n$ tiene, a lo sumo, n valores propios distintos.

3.3. Polinomio característico. En esta sección E denota un espacio vectorial sobre $\mathbf{K}$ de dimensión finita n, y f un endomorfismo de E.

Hemos visto que para hallar los vectores propios de valor propio $\lambda \in \mathbf{K}$ basta calcular el autoespacio $\operatorname{Ker}(f - \lambda \cdot \mathbf{1}_E)$. El problema es que necesitamos saber a priori el valor de λ. Por ello, la pregunta que hemos de responder previamente es la siguiente.

Pregunta 3.3.1. ¿Cómo calcular los valores propios?.

Esto se consigue, al menos de forma teórica, con el polinomio característico.

Proposición 3.3.2. (Caracterización de los valores propios de f.) Sea f un endomorfismo de un espacio vectorial de dimensión finita n. Dado un escalar $\lambda \in \mathbf{K}$ son equivalentes:

(1) El escalar λ es un valor propio de f.
(2) El endomorfismo $f - \lambda \cdot \mathbf{1}_E$ no es inyectivo.
(3) $\det(f - \lambda \cdot \mathbf{1}_E) = 0$.

Demostración. Sea $\lambda \in \mathbf{K}$, entonces $f - \lambda \cdot \mathbf{1}_E$ es un endomorfismo de E. La condición de que λ sea un valor propio de f es equivalente a que $\operatorname{Ker}(f - \lambda \cdot \mathbf{1}_E)$ sea no nulo, es decir, que $f - \lambda \cdot \mathbf{1}_E$ no sea inyectiva. Si E tiene dimensión finita, entonces, la no inyectividad de $f - \lambda \cdot \mathbf{1}_E$ es equivalente a que su determinante sea nulo. $\square$

Si E tiene dimensión finita, el determinante $\det(f - \lambda \cdot Id_E)$ es una función de λ,

$$\mathbf{K} \longrightarrow \mathbf{K}, \quad \lambda \mapsto \det(f - \lambda \cdot Id_E),$$

y la forma de detectar los valores propios de f es hallando los ceros de esta función. A continuación vamos a analizar esta función. Para ello usaremos el determinante de una matriz con coeficientes en un anillo conmutativo.

Observación 3.3.3. Un anillo conmutativo es un conjunto R, con dos elementos distinguidos $0, 1 \in R$, dos operaciones binarias, la suma $+$ y la multiplicación $\cdot$, y la aplicación opuesto $-?$. Estos datos $(R, 0, 1, +, \cdot, -?)$ han de satisfacer los axiomas de cuerpo, salvo el correspondiente al inverso para la multiplicación.

El determinante de una matriz cuadrada A con coeficientes en un anillo conmutativo R se obtiene por las mismas fórmulas que en el caso de los determinantes de las matrices con coeficientes en $\mathbb{R}$, pues en esas fórmulas solo se utilizan la suma, la resta y el producto de elementos de $\mathbb{R}$, y nunca la división. El resultado del determinante de una matriz A con coeficientes en un anillo R es pues un elemento de R. Este determinante satisface propiedades análogas a las del determinante con coeficientes en el cuerpo de los números reales $\mathbb{R}$ que hemos estudiado en el curso de Matrius i vectors, y que no detallaremos.

El anillo que usaremos será $R = \mathbf{K}[t]$, el anillo de polinomios con coeficientes en $\mathbf{K}$ en una indeterminada t.

Proposición 3.3.4. (Polinomio característico de una matriz.) Sea A una matriz de tipo $n \times n$. Para toda matriz invertible U de tipo $n \times n$ se cumple

$$\det(U^{-1}AU - t\mathbf{1}) = \det(A - t\mathbf{1}).$$

Demostración. En primer lugar se cumple que

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

y por las propiedades multiplicativas del determinante,

$$\det(U^{-1} \cdot A \cdot U - t\mathbf{1}) = (\det U^{-1}) \det(A - t\mathbf{1}) \det U = (\det U)^{-1} \det U \det(A - t\mathbf{1}) = \det(A - t\mathbf{1}).$$

□

Teorema 3.3.5. (Polinomio característico.) Supongamos que E tiene dimensión finita n. Sea $\mathbf{u}$ una base de E y A la matriz de f en la base $\mathbf{u}$. Entonces

1. El determinante
$$\chi_f(t) = \det(A - t \cdot I).$$
no depende de la base $\mathbf{u}$ escogida.
2. $\chi_f(t)$ es un polinomio de grado n en la indeterminada t.
3. El coeficiente de grado p de $\chi_f(t)$ es $(-1)^p c_p$, donde c_p es la suma de todos los menores de orden $n-p$ cuya diagonal está sobre la diagonal principal de $A - t \cdot \mathbf{1}$. En particular
$$\chi_f(t) = (-1)^n t^n + (-1)^{n-1} \mathrm{tr}\,(f) t^{n-1} + \cdots + \det(f) t^0.$$
4. Las raíces de $\chi_f(t)$ son exactamente los valores propios de f.

Definición 3.3.6. Si f es un endomorfismo de un espacio vectorial E de dimensión finita sobre $\mathbf{K}$, el polinomio $\chi_f(t) \in \mathbf{K}[t]$ se llama el *polinomio característico de f*.

Demostración. Prueba de la proposición 3.3.5. Veamos (1). Si $\mathbf{v}$ es otra base de E y B denota la matriz de f en la base $\mathbf{v}$, entonces $B = U^{-1}AU$, para una cierta matriz invertible U, por tanto, por el lema anterior, $\det(B - tI) = \det(A - tI)$.

Veamos (2). El determinante de $A - t \cdot \mathbf{1}$ es
$$\begin{vmatrix} a_1^1 - t & a_2^1 & \cdots & a_n^1 \\ a_1^2 & a_2^2 - t & \cdots & a_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_1^n & a_2^n & \cdots & a_n^n - t \end{vmatrix}.$$

Este determinante se puede expresar como una suma cuyos términos son todos del mismo tipo: aparte del signo, cada uno de ellos es un producto de n factores escogidos entre los elementos de la matriz, de forma que no haya dos de la misma fila o columna. En particular, el producto de los elementos de la diagonal $(a_1^1 - t) \cdot \cdots \cdot (a_n^n - t)$ es uno de estos sumandos, que al desarrollar en potencias de t es un polinomio en t de grado n y coeficiente dominante $(-1)^n$. El resto de los sumandos son también polinomios en t de grado $< n$. Por tanto $\det(A - t \cdot \mathbf{1}_E)$ es un polinomio en t de grado n.

Veamos (3). Daremos solo una idea de la prueba. Para determinar el coeficiente de grado p en t, observamos que este coeficiente es una suma de ciertos productos formados cada uno de ellos por n factores, p de los cuales son de la forma $(-t)$ y corresponden a términos que están en la diagonal de la matriz. Para cada elección de p lugares de la diagonal, con índices $i_1, \ldots, i_p$, para los términos $(-t)$, los $n-p$ factores restantes son de la forma a_k^j, con $j, k \neq i_1, \ldots, i_p$, y corresponden a las $n-p$ filas y columnas de los índices diferentes de $i_1, \ldots, i_p$. Cada una de estas elecciones da lugar a un menor de orden $n-p$ de A. □

Ejemplo 3.3.7. Supongamos que E es de dimensión finita n y que tiene una base $\mathbf{u}$ de vectores propios de f, con valores propios $\lambda_1, \ldots, \lambda_n$, entonces la matriz de f en la base $\mathbf{u}$ es la matriz diagonal $D(\lambda_1, \ldots, \lambda_n)$, por tanto
$$\chi_f(t) = \det\left(D(\lambda_1, \ldots, \lambda_n) - t\mathbf{1}\right) = \det D(\lambda_1 - t, \ldots, \lambda_n - t) = (\lambda_1 - t) \cdot \cdots \cdot (\lambda_n - t).$$

Ejemplo 3.3.8. Si A es una matriz de tipo 2×2, entonces
$$\chi_A(t) = \begin{vmatrix} a_1^1 - t & a_2^1 \\ a_1^2 & a_2^2 - t \end{vmatrix} = t^2 - (a_1^1 + a_2^2)t + (a_1^1 \cdot a_2^2 - a_2^1 \cdot a_1^2) = t^2 - (\mathrm{tr}\,A) \cdot t + \det A.$$

Continuando el ejemplo 3.1.1, si $A = \left(\begin{smallmatrix} 11 & -8 \\ 16 & -13 \end{smallmatrix}\right)$,

$$\chi_A(t) = \begin{vmatrix} 11-t & -8 \\ 16 & -13-t \end{vmatrix} = t^2 - (11-13)t + (11 \cdot (-13) - 16 \cdot (-8)) = t^2 + 2t - 15.$$

Como hemos visto anteriormente, si $U = \left(\begin{smallmatrix} 1 & 1 \\ 1 & 2 \end{smallmatrix}\right)$, $D = \left(\begin{smallmatrix} 3 & 0 \\ 0 & -5 \end{smallmatrix}\right)$ entonces se cumple $U^{-1}AU = D$. Por tanto $\operatorname{tr} A = \operatorname{tr} D = 3 - 5 = -2$, $\det A = \det D = 3 \cdot (-5) = -15$.

Ejemplo 3.3.9. Si A es una matriz de tipo 3×3, entonces

$$\begin{aligned}\chi_A(t) &= \begin{vmatrix} a_1^1 - t & a_2^1 & a_3^1 \\ a_1^2 & a_2^2 - t & a_3^2 \\ a_1^3 & a_2^3 & a_3^3 - t \end{vmatrix} \\ &= -t^3 + (a_1^1 + a_2^2 + a_3^3)t^2 - \left(\begin{vmatrix} a_1^1 & a_2^1 \\ a_1^2 & a_2^2 \end{vmatrix} + \begin{vmatrix} a_1^1 & a_3^1 \\ a_1^3 & a_3^3 \end{vmatrix} + \begin{vmatrix} a_2^2 & a_3^2 \\ a_2^3 & a_3^3 \end{vmatrix} \right) t + \det A\end{aligned}$$

Ejemplo 3.3.10. Si

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

donde $a \in \mathbb{R}$ es un parámetro, entonces $\chi_A(t) = (2-t)^2(3-t)$ (en este caso, el polinomio característico no depende del parámetro a). Los valores propios son $\lambda_1 = 2$, con multiplicidad $m_1 = 2$, y $\lambda_2 = 3$, con multiplicidad $m_2 = 1$.

Ejemplo 3.3.11. Si

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

entonces $\chi_A(t) = (t^2+1)(3-t)$. Si consideramos A como matriz sobre $\mathbb{R}$, el único valor propio es $\lambda_1 = 3$, con multiplicidad $m_1 = 1$. Si consideramos A como matriz sobre $\mathbb{C}$, entonces A tiene tres valores propios de multiplicidad 1, que son $\{i, -i, 3\}$.

Definición 3.3.12. (Multiplicidad algebraica de un valor propio.) Supongamos que E tiene dimensión finita. Sea λ un valor propio de f. Se llama *multiplicidad algebraica de λ como valor propio de f* a la multiplicidad de λ como raíz de $\chi_f(t)$, y se le denota por m_λ.

Corolario 3.3.13. (La suma de las multiplicidades es menor o igual que la dimensión.) Sea E un espacio vectorial de dimensión finita n sobre $\mathbf{K}$. Si f es un endomorfismo de E, f tiene a lo sumo n valores propios distintos, contados con sus multiplicidades algebraicas. Es decir, si $\lambda_1, \dots, \lambda_r$ son los valores propios distintos de f y $m_1, \dots, m_r$ sus multiplicidades algebraicas, entonces $m_1 + \cdots + m_r \leq n$.

Proposición 3.3.14. Sean λ un valor propio de f, $d_\lambda = \dim \operatorname{Ker}(f - \lambda I)$ su multiplicidad geométrica y m_λ su multiplicidad algebraica. Entonces

$$d_\lambda \leq m_\lambda.$$

Demostración. Sea $\{u_1, \dots, u_d\}$ una base de $\operatorname{Ker}(f - \lambda I)$. Completamos $\{u_1, \dots, u_d\}$ a una base $\mathbf{u} = \{u_1, \dots, u_d, u_{d+1}, \dots, u_n\}$ de E. Entonces la matriz de f en la base $\mathbf{u}$ es de la forma

$$\begin{pmatrix} \lambda \mathbf{1}_d & C \\ 0 & B \end{pmatrix},$$

para matrices C, B convenientes. Por tanto

$$\chi_f(t) = (\lambda - t)^d \cdot \chi_B(t),$$

de donde deducimos que $d \leq m$. □

Observación 3.3.15. Si E tiene dimensión finita, f solo puede tener un número finito de valores propios. Pero en caso contrario, f puede tener cualquier número de valores propios. Por ejemplo, si $E = \mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$, y consideramos el endomorfismo de derivación, $D(h) = h'$, para toda $h \in E$, entonces, para todo $\lambda \in \mathbb{R}$, la función $e^{\lambda t}$ satisface $D(e^{\lambda t}) = \lambda \cdot e^{\lambda t}$, por tanto $e^{\lambda t}$ es un vector propio de valor propio λ del operador D, es decir $\operatorname{Ker}(D - \lambda \cdot \mathbf{1}_E)$ es no nulo para todo $\lambda \in \mathbb{R}$.

3.4. Diagonalización.

En lo que sigue E denota un espacio vectorial sobre $\mathbf{K}$ y f un endomorfismo de E.

Definición 3.4.1. (Endomorfismo diagonalizable.) Se dice que f es un *endomorfismo diagonalizable* si existe una base de E formada por vectores propios de f.

Una consecuencia inmediata de 3.2.15 es la siguiente caracterización de los endomorfismos diagonalizables.

Teorema 3.4.2. Sea f un endomorfismo de un espacio vectorial E (no necesariamente de dimensión finita). Sean $\{\lambda_i\}_{i \in I}$ los valores propios distintos de f.

Entonces f es diagonalizable. si y solo si $E = \bigoplus_{i \in I} \operatorname{Ker}(f - \lambda_i Id_E)$.

Además, en tal caso, si B_i es una base de $\operatorname{Ker}(f - \lambda_i \cdot \mathbf{1}_E)$, para todo i, entonces $B = \bigcup_{i \in I} B_i$ es una base de vectores propios de f.

Proposición 3.4.3. Supongamos que E es de dimensión finita. Entonces f es diagonalizable si y solo existe una base de E tal que la matriz de f en dicha base es una matriz diagonal.

Definición 3.4.4. (Matriz diagonalizable.) Se dice que una matriz cuadrada A de tipo $n \times n$ es una *matriz diagonalizable* si el endomorfismo L_A de $\mathbf{K}^n$ asociado a A es diagonalizable, es decir, si existe una matriz invertible U de tipo $n \times n$ tal que $U^{-1} \cdot A \cdot U$ es una matriz diagonal.

Observación 3.4.5. Si A es una matriz diagonalizable, la matriz U no es en general única, pero los valores propios de la forma diagonal están determinados por el polinomio característico de A y son únicos.

Ejemplo 3.4.6. Continuando con la matriz A el ejemplo 3.2.8, la matriz $U = \begin{pmatrix} -2 & 0 & -2 \\ -5 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix}$ satisface $AU = U \begin{pmatrix} 4 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix}$, por tanto

$$U^{-1}AU = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix} = D(4, 2, -1),$$

y A es diagonalizable.

Ejemplo 3.4.7. Continuando con la matriz B del ejemplo 3.2.9,

entonces la matriz $V = \begin{pmatrix} -1 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$ satisface $BV = V \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$, por tanto

$$V^{-1}BV = D(-1, -1, 2),$$

y B es diagonalizable.

La matriz $W = \begin{pmatrix} -2 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & -1 & 1 \end{pmatrix}$ también satisface $W^{-1}BW = D(-1, -1, 2)$.

Teorema 3.4.8. (Primera condición necesaria de diagonalizabilidad: Descomposición del polinomio característico.) Si E es un espacio vectorial de dimensión finita n, y f es un endomorfismo diagonalizable de E, entonces su polinomio característico descompone completamente en factores lineales. Dicho de otro modo, la suma de las multiplicidades algebraicas de sus valores propios es n, $\sum_i m_i = n$.

Ejemplo 3.4.9. A continuación damos un ejemplo en que no se cumple la condición anterior. Sea $C = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$, entonces $\chi_C(t) = t^2 + 1$, que es un polinomio sin raíces reales. Por tanto C no diagonaliza sobre $\mathbb{R}$.

Observemos sin embargo que C sí que diagonaliza sobre $\mathbb{C}$, pues la matriz inversible compleja $W = \left(\begin{smallmatrix} 1 & i \\ 1 & -i \end{smallmatrix}\right)$ satisface $CW = W \left(\begin{smallmatrix} i & 0 \\ 0 & -i \end{smallmatrix}\right)$.

Ejemplo 3.4.10. La condición del teorema no es, sin embargo, suficiente, como muestra el siguiente ejemplo. Sea $A = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$. Entonces $\chi_A(t) = t^2$ descompone en factores lineales, pero A no es diagonalizable. En efecto, puesto que el único valor propio es $\lambda = 0$, si A fuera diagonalizable, su forma diagonal D sería 0, por tanto $A = UDU^{-1} = 0$.

Sobre el cuerpo de los números complejos, un polinomio de grado n siempre tiene n raíces, y si el polinomio es "genérico", no tiene raíces repetidas. Las matrices que provienen de observaciones experimentales no suelen tener valores propios múltiples debido a que han de ser estables bajo el efecto de pequeños errores de observación o simplemente de las variaciones sufridas por las aproximaciones de las mediciones. A estas matrices se les aplica el siguiente resultado.

Teorema 3.4.11. (Una condición suficiente de diagonalizabilidad: Espectro máximo) Supongamos que E es de dimensión finita n. Si el polinomio característico $\chi_f(t)$ tiene n raíces distintas (las máximas posibles), entonces f es diagonalizable.

Demostración. Sean $\{\lambda_i\}_{1\le i\le n}$ los valores propios distintos de f. Por 3.2.15 la suma de autoespacios

$$\operatorname{Ker}(f - \lambda_1 \cdot \mathbf{1}_E) + \cdots + \operatorname{Ker}(f - \lambda_n \cdot \mathbf{1}_E)$$

es directa, y $\dim \operatorname{Ker}(f - \lambda_i \cdot \mathbf{1}_E) \geq 1$, pues cada λ_i es un valor propio de f. De aquí se obtiene

$$\dim \Big(\operatorname{Ker}(f - \lambda_1 \cdot \mathbf{1}_E) \oplus \cdots \oplus \operatorname{Ker}(f - \lambda_n \cdot \mathbf{1}_E)\Big) \geq n.$$

Por tanto

$$\operatorname{Ker}(f - \lambda_1 \cdot \mathbf{1}_E) \oplus \cdots \oplus \operatorname{Ker}(f - \lambda_n \cdot \mathbf{1}_E) = E,$$

y f es diagonalizable, por el teorema 3.4.2. □

Ejemplo 3.4.12. Sea

$$A = \begin{pmatrix} 1 & 0 & 0 & 2 \\ 0 & 3 & 0 & 0 \\ 4 & 5 & 6 & 7 \\ 0 & 8 & 0 & 9 \end{pmatrix}.$$

Entonces $\chi_A(t) = (1-t)(3-t)(6-t)(9-t)$. Por tanto A tiene cuatro valores propios diferentes $1, 3, 6, 9$. Usando el resultado anterior deducimos que A es diagonalizable, y su forma diagonal es $D(1, 3, 6, 9)$. Un cálculo de vectores propios correspondientes a estos valores propios conduce a una base de vectores propios $u_1 = (5, 0, -4, 0)$, $u_2 = (12, -9, -29, 12)$, $u_3 = (0, 0, 1, 0)$, $u_4 = (3, 0, 32, 12)$. Por tanto, si

$$U = \begin{pmatrix} 5 & 12 & 0 & 3 \\ 0 & -9 & 0 & 9 \\ -4 & -29 & 1 & 32 \\ 0 & 12 & 0 & 12 \end{pmatrix}, \quad D = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 9 \end{pmatrix},$$

entonces $A = U \cdot D \cdot U^{-1}$.

Teorema 3.4.13. (**Primer caracterización de diagonalizabilidad: Criterio de las multiplicidades geométricas.**) Sea f un endomorfismo de un espacio vectorial E de dimensión finita n. Sean $\{\lambda_i\}_{1\leq i\leq r}$ los valores propios distintos de f, y $d_i = \dim \operatorname{Ker}(f - \lambda_i Id_E)$ la multiplicidad geométrica de λ_i, y m_i la multiplicidad algebraica de λ_i, para todo i. Entonces las siguientes condiciones son equivalentes:

(1) f es diagonalizable.
(2) $d_1 + \cdots + d_r = n$
(3) $d_i = m_i$, para todo i, y $\sum_i m_i = n$.

Demostración. En efecto, puesto que $d_i \leq m_i$ para todo i, por 3.3.14, y $\sum_i m_i \leq n$ por 3.3.13, las condiciones (2) y (3) del enunciado son equivalentes.

Veamos la equivalencia entre (2) y (1). Si f es diagonalizable, existe una base $B = \{u_1, \ldots, u_n\}$ de E formada por vectores propios de f. Esta base B se puede ordenar agrupando los vectores propios del mismo valor propio. Así podemos expresar B como reunión de partes disjuntas $B = B_1 \cup \cdots \cup B_r$, donde B_i está formado por k_i vectores propios de valor propio λ_i. La matriz de f en la base B es diagonal y su polinomio característico es

$$(\lambda_1 - t)^{k_1} \cdot \ \ldots \ \cdot (\lambda_r - t)^{k_r},$$

y coincide con el polinomio característico de f, por tanto $k_i = m_i$ para todo i. Por otra parte, $\operatorname{Ker}(f - \lambda_i I)$ contiene los m_i vectores linealmente independientes de B_i, por tanto $m_i \leq d_i$. Puesto que $\sum_i m_i = n$, y $\sum_i d_i \leq n$, resulta $d_i = m_i$ para todo i, y $\sum_i d_i = n$.

Recíprocamente, si $d_1 + \cdots + d_r = n$, escogiendo una base de cada autoespacio y tomando su reunión obtenemos n vectores propios linealmente independientes. Por tanto forman una base de vectores propios y f es diagonalizable.

□

Ejemplo 3.4.14. Continuando el ejemplo 3.3.11 observamos que A no es diagonalizable considerada como matriz sobre $\mathbb{R}$, pues falla la primera condición del teorema anterior. Pero A sí que es diagonalizable considerada como matriz sobre $\mathbb{C}$.

Ejemplo 3.4.15. Continuando el ejemplo 3.3.10, usando el criterio anterior obtenemos que A es diagonalizable si $a = 0$, y no lo es si $a \neq 0$, pues falla la segunda condición del teorema anterior.

Ejemplo 3.4.16. Sea

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

una matriz dependiente de un parámetro $p \in \mathbb{R}$. Se tiene

$$\chi_C(t) = (-1 - t)^2(-2 - t).$$

Por tanto los valores propios son $\lambda_1 = 1, \lambda_2 = -2$, con multiplicidades algebraicas $m_1 = 2$, $m_2 = 1$, para todo valor de p.

Por otra parte, por un cálculo explícito, la multiplicidad geométrica del primer valor propio es $d_1 = 1$, si $p = -1$, y $d_1 = 2$, si $p \neq -1$. La del segundo valor propio no hace falta calcularla, pues es necesariamente 1.

Del teorema anterior deducimos que C no diagonaliza si $p = -1$, y diagonaliza para $p \neq -1$.

Ejemplo 3.4.17. Veamos de nuevo que la matriz $\left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ no es diagonalizable (compárese con la prueba de 3.4.10). En efecto, el polinomio característico es t^2. El único valor propio es $\lambda = 0$, con multiplicidad algebraica 2 y multiplicidad geométrica $1 \neq 2$.

A continuación vamos a estudiar un segundo criterio de diagonalizabilidad que, además, tiene un considerable interés en las aplicaciones.

Proposición 3.4.18. Sean f es un endomorfismo diagonalizable con un número finito de valores propios $\lambda_1, \ldots, \lambda_r$, y $P(t) = \prod_{i=1}^{r} (t - \lambda_i)$. Entonces

$$P(f) := (f - \lambda_1 Id_E) \cdot (f - \lambda_2 Id_E) \cdot \ \ldots \ \cdot (f - \lambda_r Id_E) = 0.$$

Demostración. En efecto los endomorfismos $f - \lambda_i Id_E$ conmutan 2 a 2, por tanto, para cada vector propio u de f, si λ_i es su valor propio, $P(f) \cdot u$ se puede calcular aplicando el factor $f - \lambda_i Id_E$ en primer lugar, por tanto $P(f) \cdot u = 0$. Puesto que E tiene una base formada por vectores propios, $P(f) = 0$. □

Una consecuencia del resultado anterior es un nuevo método para calcular vectores propios, que indicamos en el siguiente ejercicio.

Ejercicio 3.4.19. Sean f un endomorfismo de un espacio vectorial E, y $P(t)$, $Q(t)$ dos polinomios primos entre sí tales que $(P \cdot Q)(f) = 0$. Entonces $\operatorname{Im} P(f) = \operatorname{Ker} P(f)$.

Como aplicación, si A es una matriz diagonalizable, con valores propios $\lambda_1, \cdots, \lambda_r$, entonces

$$\operatorname{Ker}(A - \lambda_1 \mathbf{1}) = \operatorname{Im}(A - \lambda_2 \mathbf{1}) \cdot \ \cdots \ \cdot (A - \lambda_r \mathbf{1}).$$

Para obtener el recíproco de la proposición 3.4.18veamos primero un resultado que es consecuencia del teorema clave 3.2.15. Este resultado es de gran utilidad en las aplicaciones a la resolución de ecuaciones en diferencias, y de ecuaciones diferenciales.

Corolario 3.4.20. (Segunda caracterización de diagonalizabilidad: Criterio del polinomio anulador) Si f es un endomorfismo de un espacio vectorial E (no necesariamente de dimensión finita). Entonces f es diagonalizable con un número finito de valores propios si, y solo si, existe un polinomio $P(t)$ de grado r con r raíces simples tal que $P(f) = 0$.

Demostración. Si f es diagonalizable, el resultado de sigue de 3.4.18. Recíprocamente, si el polinomio $P(t)$ existe, por 3.2.15(3) y 3.4.2 se obtiene que f es diagonalizable. □

3.5. Un procedimiento para la diagonalización de una matriz. Sean $\mathbf{K}$ un cuerpo y A una matriz de tipo $n \times n$ con coeficientes en $\mathbf{K}$.

Resumiendo los resultados de las secciones anteriores, tenemos el siguiente procedimiento en 5 pasos para decidir si una matriz es diagonalizable y, en caso afirmativo, hallar su forma diagonal, y una base de vectores propios.

Paso 1. Hallar el polinomio característico, $\chi_A(t) = \det(A - tI)$.

Paso 2. Hallar las diferentes raíces $\lambda_1, \ldots, \lambda_r$ de $\chi_A(t)$ en $\mathbf{K}$, y determinar sus multiplicidades $m_1, \ldots, m_r$ como raíces de $\chi_A(t)$. Comprobar que $m_1 + \cdots + m_r = n$ (es decir, $\chi_A(t)$ se descompone en producto de factores de grado 1). Si no es así, la matriz no es diagonalizable (sobre $\mathbf{K}$).

Paso 3. Para cada λ_i, tal que $m_i > 1$, comprobar si $rg(A - \lambda_i I) \leq n - m_i$. En tal caso se tendrá necesariamente $rg(A - \lambda_i I) = n - m_i$, y la matriz será diagonalizable. En caso contrario, la matriz no es diagonalizable sobre $\mathbf{K}$.

Paso 4. Para cada i, hallar una base $U_i = (u_{1,i}, \ldots, u_{d_i,i})$ de $\operatorname{Ker}(A - \lambda_i I)$.

Paso 5. Puesto que $\sum_i d_i = \sum_i m_i = n$ y la suma $\sum_i \operatorname{Ker}(A - \lambda_i I)$ es directa, reuniendo las bases U_i obtendremos un conjunto U linealmente independiente de n vectores (sin necesidad de comprobarlo!). Por tanto U es una base de vectores propios, cuya matriz

$$U = \begin{pmatrix} u_{1,1} & u_{2,1} & \cdots & u_{d_1,1} & u_{1,2} & \ldots & u_{1,r} & \cdots & u_{d_r,r} \end{pmatrix}$$

es un cambio de base que diagonaliza a la matriz A, es decir,

$$U^{-1}AU = \lambda_1 \cdot I_{d_1} \oplus \lambda_2 \cdot I_{d_2} \oplus \cdots \oplus \lambda_r \cdot I_{d_r}$$

donde I_d denota la matriz unidad de tamaño $d \times d$.

Ejemplo 3.5.1. Sea $\mathbf{K} = \mathbb{R}$, $n = 5$ y A la matriz

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

donde $p \in \mathbb{R}$ es un parámetro. Estudiamos la diagonalizabilidad de A en función de p.

Paso 1. Usando la regla de Laplace obtenemos

$$\chi_A(t) = \begin{vmatrix} -1-t & -2 & 0 & 0 & 3 \\ 0 & 1-t & 0 & 0 & 2 \\ 1 & 1 & -t & 1 & 1 \\ 1 & p & 1 & -t & 0 \\ 0 & 0 & 0 & 0 & 2-t \end{vmatrix} = (2-t) \cdot \begin{vmatrix} -1-t & -2 & 0 & 0 \\ 0 & 1-t & 0 & 0 \\ 1 & 1 & -t & 1 \\ 1 & p & 1 & -t \end{vmatrix}$$

$$= (2-t) \cdot \begin{vmatrix} -1-t & -2 \\ 0 & 1-t \end{vmatrix} \cdot \begin{vmatrix} -t & 1 \\ 1 & -t \end{vmatrix} = (2-t)(t^2-1)(t^2-1)$$

y por tanto $\chi_A(t) = (t-2)(t+1)^2(t-1)^2$.

Paso 2. Las raíces de $\chi_A(t)$ son $\lambda_1 = 1, \lambda_2 = -1, \lambda_3 = 2$, con multipilicidades $m_1 = m_2 = 2, m_3 = 1$. Por tanto $m_1 + m_2 + m_3 = 5$ y el polinomio $\chi_A(t)$ se descompone en producto de factores de grado 1. Por tanto, para todo p, A cumple la primera condición necesaria de diagonalizabilidad.

Paso 3. Para $\lambda_1 = 1$ estudiamos el rango de $A - I$ y lo comparamos con $5 - m_1 = 3$. Aplicando Gauss por filas se obtiene

$$A - I = \begin{pmatrix} -2 & -2 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 2 \\ 1 & 1 & -1 & 1 & 1 \\ 1 & p & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \sim_f \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & p-1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

por tanto el rango de $A - I$ es 3 si, solo si, $p = 1$.

Analogamente, para $\lambda_2 = -1$ estudiamos el rango de $A + I$ y lo comparamos con $5 - m_2 = 3$. Aplicando Gauss por filas se obtiene

$$A + I = \begin{pmatrix} 0 & -2 & 0 & 0 & 3 \\ 0 & 2 & 0 & 0 & 2 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & p & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix} \sim_f \begin{pmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

por tanto su rango es $5 - m_2 = 3$ para todo p.

Para λ_3 no es necesario satisfacer ninguna comprobación, pues $m_3 = 1$.

En resumen, se tiene $d_i = m_i$, para todo i, si y solo si $p = 1$. Por tanto A diagonaliza si y solo si $p = 1$. Para $p = 1$ la forma diagonal de A es

$$D = D(1, 1, -1, -1, 2).$$

Paso 4. Cálculo de una base de vectores propios de A si $p = 1$. Para $\lambda_1 = 1$ obtenemos, a partir de los cálculos anteriores, la siguiente base de $\operatorname{Ker}(A - I)$: $u_{1,1} = (-1, 1, 0, 0, 0), u_{2,1} = (0, 0, 1, 1, 0)$.

Para $\lambda_2 = -1$ obtenemos, a partir de los cálculos anteriores, la siguiente base de $\operatorname{Ker}(A + I)$: $u_{1,2} = (-1, 0, 1, 0, 0), u_{2,2} = (-1, 0, 0, 1, 0)$.

Finalmente, para $\lambda_3 = 2$ hemos de hallar una base de $\operatorname{Ker}(A - 2I)$. Usando Gauss por filas:

$$A - 2I = \begin{pmatrix} -3 & -2 & 0 & 0 & 3 \\ 0 & -1 & 0 & 0 & 2 \\ 1 & 1 & -2 & 1 & 1 \\ 1 & 1 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \sim_f \begin{pmatrix} 1 & 0 & 0 & 0 & \frac{1}{3} \\ 0 & 1 & 0 & 0 & -2 \\ 0 & 0 & 1 & 0 & -\frac{7}{3} \\ 0 & 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

por tanto una base de $\operatorname{Ker} A - 2I$ es $u_{1,3} = (-1, 6, 7, 6, 3)$.

Paso 5. Finalmente la matriz U formada por la base de vectores propios es

$$U = \begin{pmatrix} -1 & 0 & -1 & -1 & -1 \\ 1 & 0 & 0 & 0 & 6 \\ 0 & 1 & 1 & 0 & 7 \\ 0 & 1 & 0 & 1 & 6 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix}.$$

Para comprobar que está bien resuelto comprobamos, mediante un cálculo explícito, la siguiente identidad:

$$A \cdot U = U \cdot D.$$

3.6. Aplicaciones.

3.6.1. **Potencias de matrices.** A continuación vamos a ver como se calculan las potencias de una matriz diagonalizable.

Proposición 3.6.1. Sea A una matriz diagonalizable, U una matriz invertible tal que $U^{-1} \cdot A \cdot U = D(\lambda_1, \ldots, \lambda_n)$ es una matriz diagonal, entonces, para todo entero $p \geq 0$,

$$A^p = U \cdot D(\lambda_1^p, \ldots, \lambda_n^p) \cdot U^{-1}.$$

Más generalmente estamos interesados en el cálculo de cualquier combinación lineal de potencias de una matriz cuadrada A.

Teorema 3.6.2. Si A es diagonalizable, y U es un cambio de base U tal que $U^{-1} \cdot A \cdot U = D = D(\lambda_1, \ldots, \lambda_n)$. entonces

$$P(A) = U \cdot D(P(\lambda_1), \ldots, P(\lambda_n)) \cdot U^{-1},$$

para todo $P(t) \in \mathbf{K}[t]$.

Demostración. Si $D = D(\lambda_1, \ldots, \lambda_n)$ es diagonal, entonces

$$P(D) = D(P(\lambda_1), \ldots, P(\lambda_n)).$$

Por tanto

$$P(A) = U \cdot P(D) \cdot U^{-1} = U \cdot D(P(\lambda_1), \ldots, P(\lambda_n)) \cdot U^{-1},$$

□

Ejemplo 3.6.3. Continuando con la matriz $A = \begin{pmatrix} 11 & -8 \\ 16 & -13 \end{pmatrix}$ del ejemplo 3.1.1, sabemos que $U = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$ satisface $A = U \cdot D(3, -5) \cdot U^{-1}$. Entonces

1. Si $P(t) = t^6 - t^3 + 1$, entonces

$$A^6 - A^3 + \mathbf{1} = U^{-1} \cdot D(P(3), P(-5)) \cdot U = U \cdot D(703, 15.751) \cdot U^{-1} = \begin{pmatrix} -14{,}345 & 15{,}048 \\ 20{,}093 & 20{,}799 \end{pmatrix}.$$

2. Buscamos ahora una matriz B tal que $B^2 = A$, es decir, raíces cuadradas de A. Si trabajamos sobre $\mathbb{C}$, podemos hallar soluciones de esta ecuación que sean diagonalizables de la forma $B = U \cdot D(\phi_1, \phi_2) \cdot U^{-1}$. En efecto, la ecuación $B^2 = A$ es equivalente a $D((\phi_1)^2, (\phi_2)^2) = D(3, -5)$, cuyas soluciones son $\phi_1 = \pm\sqrt{3}, \phi_2 = \pm\sqrt{5}i$, por tanto obtenemos 4 matrices, que se pueden escribir en la forma
$$B = U \cdot D(\pm\sqrt{3}, \pm\sqrt{5}i) \cdot U^{-1}$$
3. Consideremos la ecuación matricial en la incógnita B,
$$B^2 + 4B + A = 0.$$
Las soluciones diagonalizables en la base U, $B = U \cdot D(\phi_1, \phi_2) \cdot U^{-1}$ satisfacen
$$(\phi_1)^2 + 4\phi_1 + 3 = 0, \quad (\phi_2)^2 + 4\phi_2 - 5 = 0$$
por tanto $\phi_1 \in \{-1, -3\}$, $\phi_2 \in \{1, -5\}$, y combinando estos valores obtenemos 4 soluciones para B, que son matrices sobre $\mathbb{R}$.

Ejemplo 3.6.4. Continuemos con el ejemplo 3.2.8. Sea $A = \begin{pmatrix} 0 & 2 & 2 \\ 1 & 4 & 2 \\ -1 & -1 & 1 \end{pmatrix}$. La matriz $U = \begin{pmatrix} -2 & 0 & -2 \\ -5 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix}$ satisface
$$U^{-1}AU = D(4, 2, -1).$$
La inversa de U es
$$U^{-1} = \frac{-1}{10} \begin{pmatrix} 1 & 2 & 2 \\ 5 & 0 & 10 \\ 4 & -2 & -2 \end{pmatrix}$$
Las potencias de A son
$$A^p = U \cdot D(4^p, 2^p, (-1)^p) \cdot U^{-1} = \frac{-1}{10} \begin{pmatrix} -2\cdot4^p-8\cdot(-1)^p & -4\cdot4^p+4\cdot(-1)^p & -4\cdot4^p+4\cdot(-1)^p \\ -5\cdot4^p+5\cdot2^p & -10\cdot4^p & -10\cdot4^p+10\cdot2^p \\ 4^p-5\cdot2^p+4\cdot(-1)^p & 2\cdot4^p-2\cdot(-1)^p & 2\cdot4^p-10\cdot2^p-2\cdot(-1)^p \end{pmatrix}$$
Si $P(t) = t^2 - t + 1$, entonces
$$P(A) = U \cdot D(P(4), P(2), P(-1)) \cdot U^{-1} = U \cdot D(13, 3, 3) \cdot U^{-1} = \begin{pmatrix} 5 & 4 & 4 \\ 5 & 13 & 10 \\ -1 & -2 & 1 \end{pmatrix}.$$

3.6.2. **Resolución de ecuaciones diferenciales y ecuaciones en diferencias.** Supongamos que E es un espacio vectorial, no necesariamente de dimensión finita, y f es un endomorfismo de E. Las aplicaciones que siguen se basan en el teorema 3.2.15.

Ejemplo 3.6.5. Consideremos en $E = \mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$ el endomorfismo de derivación D, definido por $D(f) = f'$, para todo $f \in E$.

Sean a, b escalares tal que el polinomio $P(t) = t^2 + at + b$ tiene dos raíces diferentes λ_1, λ_2. Sea F el subespacio de E de las soluciones de la ecuación
$$D^2(f) + aD(f) + bf = 0, \quad f \in E.$$
Supongamos que $P(t) = t^2 + at + b = (t - \lambda_1)(t - \lambda_2)$, con $\lambda_1 \neq \lambda_2$. Entonces
$$F = \operatorname{Ker} P(D) = \operatorname{Ker}(D - \lambda_1 Id_E) \oplus \operatorname{Ker}(D - \lambda_2 Id_E),$$
por el teorema 3.2.15. Puesto que las soluciones de $D(f) = \lambda f$ son las funciones de la forma $c \cdot e^{\lambda \cdot x}$, con $c \in \mathbb{R}$, toda función $f \in F$ es de la forma
$$f(x) = c_1 e^{\lambda_1 x} + c_2 e^{\lambda_2 x}.$$
Los escalares c_1, c_2 se suelen determinar a partir de los valores $f(0), f'(0)$, llamados condiciones iniciales.

Ejemplo 3.6.6. Por ejemplo, la solución de la ecuación

$$f'' - 3f' + 2f = 0, \quad f \in E,$$

con las condiciones iniciales $f(0) = 0, f'(0) = 1$ es la función

$$f(t) = c_1 e^x + c_2 e^{2x},$$

donde c_1, c_2 satisfacen

$$c_1 + c_2 = 0, \quad c_1 + 2c_2 = 1,$$

por tanto

$$f(t) = -e^x + e^{2x},$$

Ejemplo 3.6.7. Estudiamos a continuación la solución de un sistema de ecuaciones diferenciales lineales homogéneas, cuya matriz sea diagonalizable.

Sea $E = \mathcal{C}^\infty(\mathbb{R}, \mathbb{R})^n$. Un vector de E es una n-pla de funciones $\overrightarrow{f} = (f_1, \cdots, f_n)$, con $f_i \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$, para todo i. La derivación define una aplicación lineal $D : E \longrightarrow E$, para la cual $D(f_1, \cdots, f_n) = (D(f_1), \cdots, D(f_n))$.

Sea A una matriz de tipo $n \times n$ con coeficientes en $\mathbb{R}$. Entonces A define una aplicación $\mathbb{R}$-lineal

$$L_A : E \longrightarrow E, \quad \overrightarrow{f} \mapsto A \cdot \overrightarrow{f},$$

donde $\overrightarrow{f}$ se identifica a una matriz de una sola columna. Dado $\overrightarrow{f}_0 \in \mathbb{R}^n$, consideremos la ecuación diferencial con condición inicial

$$D(\overrightarrow{f}) = A \cdot \overrightarrow{f}, \quad \overrightarrow{f}(0) = \overrightarrow{f}_0$$

Si A es una matriz diagonal $D(\lambda_1, \cdots, \lambda_n)$, esta ecuación es equivalente a n ecuaciones escalares:

$$D(f_i) = \lambda_i \cdot f_i, \quad f_i(0) = f_{0,i}, \quad \forall i.$$

La solución es $f_i(x) = f_{0,i} \cdot e^{\lambda_i \cdot x}$, para todo i. Por tanto,

$$\overrightarrow{f} = Diag(e^{\lambda_1 \cdot x}, \cdots, e^{\lambda_n \cdot x}) \cdot \overrightarrow{f}_0.$$

Si A es diagonalizable, existe una matriz invertible U, cuyos coeficientes son constantes reales, tal que

$$A = U^{-1} \cdot Diag(\lambda_1, \cdots, \lambda_n) \cdot U.$$

Efectuando el cambio de coordenadas $\overrightarrow{g} = U \cdot \overrightarrow{f}$ en E, se cumple

$$D(\overrightarrow{g}) = U \cdot D(\overrightarrow{f}), \quad U \cdot A \cdot \overrightarrow{f} = Diag(e^{\lambda_1 \cdot x}, \cdots, e^{\lambda_n \cdot x}) \cdot U \cdot \overrightarrow{f} = Diag(e^{\lambda_1 \cdot x}, \cdots, e^{\lambda_n \cdot x}) \cdot \overrightarrow{g},$$

por tanto la ecuación inicial se transforma en la ecuación

$$D(\overrightarrow{g}) = Diag(\lambda_1, \cdots, \lambda_n) \cdot \overrightarrow{g}, \quad \overrightarrow{g}(0) = U \cdot \overrightarrow{f}_0,$$

cuya solución es $\overrightarrow{g} = Diag(e^{\lambda_1 \cdot x}, \cdots, e^{\lambda_n \cdot x}) \cdot \overrightarrow{g}_0$, por tanto $\overrightarrow{f} = U^{-1} \cdot \overrightarrow{g}$ es

$$\overrightarrow{f} = U^{-1} \cdot Diag(e^{\lambda_1 \cdot x}, \cdots, e^{\lambda_n \cdot x}) \cdot U \cdot \overrightarrow{f}_0.$$

Ejemplo 3.6.8. Consideremos de nuevo la matriz del ejemplo 3.1.1, $A = \left(\begin{smallmatrix} 11 & -8 \\ 16 & -13 \end{smallmatrix}\right)$, con $U = \left(\begin{smallmatrix} 1 & 1 \\ 1 & 2 \end{smallmatrix}\right)$, tal que

$$A \cdot U = U \cdot Diag(3, -5), \quad U^{-1} = \left(\begin{smallmatrix} 2 & -1 \\ -1 & 1 \end{smallmatrix}\right).$$

El sistema de ecuaciones diferenciales asociado es

$$\begin{cases} D(f_1) = 11f_1 - 8f_2 \\ D(f_2) = 16f_1 - 13f_2. \end{cases}$$

Su solución es

$$\begin{pmatrix} f_1 \\ f_2 \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} e^{3x} & 0 \\ 0 & e^{-5x} \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} f_1(0) \\ f_2(0) \end{pmatrix} = \begin{pmatrix} 2e^{3x} - e^{-5x} & -e^{3x} + e^{-5x} \\ 2e^{3x} - 2e^{-5x} & -e^{3x} + 2e^{-5x} \end{pmatrix} \cdot \begin{pmatrix} f_1(0) \\ f_2(0) \end{pmatrix}.$$

Ejemplo 3.6.9. Estudiamos a continuación las sucesiones recurrentes con ley de recurrencia lineal. Consideremos en el espacio de sucesiones $E = \mathbb{R}^{\mathbb{N}}$, el endomorfismo de traslación T, definido por $T(s)_n = s_{n+1}$, para toda sucesión $s = (s_n)_n \in E$.

Sean a, b escalares tal que el polinomio $P(t) = t^2 + at + b$ tiene dos raíces diferentes λ_1, λ_2. Sea F el subespacio de sucesiones s tales que

$$T^2(s) + aT(s) + bs = 0.$$

La ecuación anterior se escribe también en la forma

$$s_{n+2} = -as_{n+1} - bs_n,$$

que es una ley de recurrencia lineal. Puesto que $P(t) = (t - \lambda_1)(t - \lambda_2)$, por el teorema 3.2.15, se cumple

$$F = \operatorname{Ker} P(T) = \operatorname{Ker}(T - \lambda_1 Id_E) \oplus \operatorname{Ker}(T - \lambda_2 Id_e).$$

La ecuación $T(s) = \lambda s$ tiene como soluciones las progresiones geométricas de razón λ. Por tanto las sucesiones de F tienen la forma $s_n = c_1\lambda_1^n + c_2\lambda_2^n$, donde las constantes c_1, c_2 se determinan a partir de los valores iniciales s_0, s_1 de la sucesión.

Por ejemplo, estudiemos la ecuación

$$s_{n+2} = s_{n+1} + s_n.$$

Sus soluciones son conocidas como sucesiones de Fibonacci.

Las raíces de $t^2 - t - 1$ son el número aureo, $\phi = \frac{1+\sqrt{5}}{2}$ y su conjugado, $\psi = \frac{1-\sqrt{5}}{2}$. Por tanto toda sucesión de Fibonacci es de la forma

$$s_n = c_1\phi^n + c_2\psi^n.$$

para ciertos valores c_1, c_2 . Por ejemplo, si imponemos las condiciones iniciales $s_0 = 0, s_1 = 1$, resulta

$$c_1 = \frac{1}{\sqrt{5}}, \quad c_2 = -\frac{1}{\sqrt{5}},$$

por tanto el número natural s_n se expresa en la forma (fórmula de Binet)

$$s_n = \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n - \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^n.$$

3.7. Matrices reales que diagonalizan sobre los complejos.

Sea A una matriz de tipo $n \times n$ con coeficientes reales. Entonces A es también una matriz $n \times n$ con coeficientes complejos, pero denotaremos por $A^{\mathbb{C}}$ la matriz A como elemento de $\mathbb{C}(n, n)$ para evitar confusiones.

Supongamos que $n = 2$. Si $A^{\mathbb{C}}$ diagonaliza, pero sus valores propios $\lambda_1, \lambda_2 \in \mathbb{C}$ no son reales, entonces $\lambda_2 = \overline{\lambda_1}$. Si $u = (z_1, z_2) \in \mathbb{C}^2$ es un vector propio de valor propio $\lambda := \lambda_1$, entonces $\overline{w} = (\overline{z_1}, \overline{z_2})$ es un vector propio de valor propio $\overline{\lambda}$. Puesto que existen $\alpha, \beta \in \mathbb{R}$ tales que $\lambda = \alpha + i\beta$ y $\beta \neq 0$, resulta $\lambda \neq \overline{\lambda}$, y, por el teorema 3.2.15, los vectores w y $\overline{w}$ son linealmente independientes. Sea $W = \left(w \ \overline{w} \right)$. La matriz del endomorfismo $L_{A^{\mathbb{C}}}$ en la base $(w, \overline{w})$ es

$$W^{-1}A^{\mathbb{C}}W = \left(\begin{smallmatrix} \lambda & 0 \\ 0 & \lambda \end{smallmatrix}\right).$$

Vamos a ver como obtener una forma simplificada de la matriz del endomorfismo real L_A mediante un cambio de base real. Para ello sean

$$u = \frac{1}{2}(w + \overline{w}) \in \mathbb{R}^2, \quad v = \frac{1}{2i}(w - \overline{w}) \in \mathbb{R}^2,$$

las partes real e imaginaria, respectivamente, de w. Entonces $w = u + iv$, $\overline{w} = u - iv$, por tanto u, v también es una $\mathbb{C}$-base de $\mathbb{C}^2$. Pero además $u, v \in \mathbb{R}^2$, luego u, v es una base de $\mathbb{R}^2$ como $\mathbb{R}$-espacio vectorial.

Sean $U = (\,u\ v\,)$ y $C = M_{(u\,v)}(w\,\overline{w}) = \left(\begin{smallmatrix} 1 & 1 \\ i & -i \end{smallmatrix}\right)$, entonces

$$C^{-1} = M_{(w\,\overline{w})}(u\,v) = \frac{1}{2}\left(\begin{smallmatrix} 1 & -i \\ 1 & i \end{smallmatrix}\right), \quad W = U \cdot C, \quad U = W \cdot C^{-1}.$$

Un cálculo explícito que no detallaremos conduce a que la matriz de L_A en la base U es

$$U^{-1}AU = CW^{-1}AWC^{-1} = CD(\lambda, \overline{\lambda})C^{-1} = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}.$$

Ejemplo 3.7.1. Consideremos la matriz $A = \left(\begin{smallmatrix} 4 & 1 \\ -5 & 0 \end{smallmatrix}\right)$. Se tiene $\chi_A(t) = t^2 - 4t + 5$ cuyas raíces son $\lambda = 2 + i$ y $\overline{\lambda} = 2 - i$. Un vector propio correspondiente a λ es $w = (1, i - 2)$. Entonces $u = \mathcal{R}(w) = (1, -2)$, $v = \mathcal{I}(w) = (0, 1)$ forman una $\mathbb{R}$-base de $\mathbb{R}^2$ tal que la matriz del endomorfismo L_A en dicha base es $D_\lambda = \left(\begin{smallmatrix} 2 & 1 \\ -1 & 2 \end{smallmatrix}\right)$.

Definición 3.7.2. Para cada $\lambda = \alpha + \beta i \in \mathbb{C}$, denotamos $D_\lambda = (\alpha)$, si $\beta = 0$, y $D_\lambda = \left(\begin{smallmatrix} \alpha & \beta \\ -\beta & \alpha \end{smallmatrix}\right)$ si $\beta \neq 0$.

Si $A_1, A_2, \ldots, A_r$ son r matrices cuadradas denotaremos por $A_1 \oplus A_2 \oplus \cdots \oplus A_r$ la matriz diagonal por bloques definida por

$$\begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_r \end{pmatrix}.$$

Ejercicio 3.7.3. Si

$$A = D_{\lambda_1} \oplus \cdots \oplus D_{\lambda_r},$$

entonces

$$A^n = D_{\lambda_1^n} \oplus \cdots \oplus D_{\lambda_r^n},$$

para todo $n \in \mathbb{N}$.

Para generalizar las observaciones anteriores a una matriz de tamaño arbitrario, consideremos el $\mathbb{C}$-espacio vectorial numérico n-dimensional $\mathbb{C}^n$. Éste es también un $\mathbb{R}$-espacio vectorial, de dimensión $2n$. Identificaremos $\mathbb{R}^n$ con el $\mathbb{R}$-subespacio vectorial de $\mathbb{C}^n$ formado por los vectores $(z_1, \ldots, z_n) \in \mathbb{C}^n$, con $z_j \in \mathbb{R}$ para todo j.

Definimos la conjugación

$$c : \mathbb{C}^n \longrightarrow \mathbb{C}^n, \quad c(z_1, \ldots, z_n) = (\overline{z_1}, \ldots, \overline{z_n})$$

que es un $\mathbb{R}$-endomorfismo de $\mathbb{C}^n$ tal que $\operatorname{Ker}(c - Id) = \mathbb{R}^n$. Si $w \in \mathbb{C}^n$, denotaremos $\overline{w} := c(w)$

La parte real y la parte imaginaria de un vector de $\mathbb{C}^n$ definen aplicaciones $\mathbb{R}$-lineales

$$\mathcal{R} : \mathbb{C}^n \longrightarrow \mathbb{R}^n, \quad \mathcal{R}\,(z_1, \ldots, z_n) = (\mathcal{R}\,z_1, \ldots, \mathcal{R}\,z_n)$$
$$\mathcal{I} : \mathbb{C}^n \longrightarrow \mathbb{R}^n, \quad \mathcal{I}\,(z_1, \ldots, z_n) = (\mathcal{I}\,z_1, \ldots, \mathcal{I}\,z_n)$$

que satisfacen

$$\mathcal{R}\,(w) = \frac{1}{2}(w + \overline{w}), \quad \mathcal{I}\,(w) = \frac{1}{2i}(w - \overline{w}).$$

Proposición 3.7.4. Sean W un subespacio complejo de $\mathbb{C}^n$ de dimensión p, $\{w_1, \dots, w_p\}$ una $\mathbb{C}$-base de W, y notemos $u_j = \mathcal{R}w_j$, $v_j = \mathcal{I}w_j$, para $1 \leq j \leq p$.

Si $W \cap c(W) = 0$ entonces $\{u_1, \dots, u_p, v_1, \dots, v_p\}$ es una $\mathbb{R}$-base de $(W \oplus \overline{W}) \cap \mathbb{R}^n$.

Demostración. Para cada j, sea $T_j = \mathbb{C}w_j + \mathbb{C}\overline{w_j}$. Puesto que $W \cap \overline{W} = 0$, los vectores $w_j, \overline{w_j}$ son linealmente independientes y son una $\mathbb{C}$-base de T_j. Los vectores u_j, v_j también son una $\mathbb{C}$-base de T_j, por tanto son linealmente independientes sobre $\mathbb{C}$, y sobre $\mathbb{R}$. Puesto que, por 3.2.15, la suma $\sum_j T_j$ también es directa, los vectores $\{u_1, \dots, u_p, v_1, \dots, v_p\}$ son linealmente independientes sobre $\mathbb{R}$ y generan un $\mathbb{R}$-subespacio V de $\mathbb{R}^n$.

Obviamente $V \subset (W \oplus \overline{W}) \cap \mathbb{R}^n$. Veamos la otra inclusión. Sea $x \in (W \oplus \overline{W}) \cap \mathbb{R}^n$. Puesto que $w_1, \dots, w_p, \overline{w_1}, \dots, \overline{w_p}$ es una $\mathbb{C}$-base de $W \oplus \overline{W}$, existen escalares $\alpha_j, \beta_j \in \mathbb{C}$ tales que

$$x = \alpha_1 w_1 + \cdots + \alpha_p w_p + \beta_1 \overline{w_1} + \cdots + \beta_p \overline{w_p}$$

La ecuación $x = \overline{x}$ es equivalente a $\beta_j = \overline{\alpha_j}$, para todo j. Por tanto

$$x = \alpha_1 w_1 + \cdots + \alpha_p w_p + \overline{\alpha_1 w_1} + \cdots + \overline{\alpha_p w_p} = 2\sum_j \mathcal{R}(\alpha_j) u_j - 2\sum_j \mathcal{I}(\alpha_j) v_j,$$

lo que prueba que $x \in V$. $\square$

Proposición 3.7.5. Sea A una matriz real de tipo $n \times n$. Para cada valor propio no real $\lambda = \alpha + i\beta$ de $A^{\mathbb{C}}$, sea $q_\lambda(t) = t^2 - 2\alpha + |\lambda|^2$.

Sea $w_1, \dots, w_p$ una $\mathbb{C}$-base del $\mathbb{C}$-subespacio $\operatorname{Ker}(A^{\mathbb{C}} - \lambda I)$ de $\mathbb{C}^n$ y, para cada j, notemos

$$u_j = \mathcal{R}(w_j), \quad v_j = \mathcal{I}(w_j)$$

entonces

(1) $\{u_1, v_1, \dots, u_p, v_p\}$ es una $\mathbb{R}$-base del $\mathbb{R}$-subespacio $\operatorname{Ker} q_\lambda(A)$ de $\mathbb{R}^n$.
(2) La restricción de L_A a $\operatorname{Ker} q_\lambda(A)$ define un endomorfismo de $\operatorname{Ker} q_\lambda(A)$ que en la base $\{u_1, v_1, \dots, u_p, v_p\}$ tiene por matriz

$$\begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix} \oplus \cdots \oplus \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}$$

Demostración. Sean $W = \operatorname{Ker}(A^{\mathbb{C}} - \lambda I)$ y $V = \operatorname{Ker} q_\lambda(A)$. Por la proposición anterior basta comprobar que $W \cap \overline{W} = 0$ y $V = \mathbb{R}^n \cap (W \oplus \overline{W})$.

En efecto, se tiene $\overline{W} = \operatorname{Ker}(A^{\mathbb{C}} - \overline{\lambda} I)$, y de $q_\lambda(t) = (t - \lambda)(t - \overline{\lambda})$, se deduce

$$\operatorname{Ker} q_\lambda(A^{\mathbb{C}}) = \operatorname{Ker}(A^{\mathbb{C}} - \lambda I) \oplus \operatorname{Ker}(A^{\mathbb{C}} - \overline{\lambda} I).$$

Por otra parte, tenemos la ecuación

$$\operatorname{Ker} q_\lambda(A) = \operatorname{Ker} q_\lambda(A^{\mathbb{C}}) \cap \mathbb{R}^n.$$

Finalmente

$$A \cdot u_j = \frac{1}{2}(A(w_j) + A(\overline{w_j})) = \frac{1}{2}(\lambda w_j + \overline{\lambda}\overline{w_j})) = \mathcal{R}(\lambda w_j) = \mathcal{R}(\lambda) u_j - \mathcal{I}(\lambda) v_j,$$
$$A \cdot v_j = \frac{1}{2}(A(w_j) - A(\overline{w_j})) = \frac{1}{2}(\lambda w_j - \overline{\lambda}\overline{w_j})) = \mathcal{I}(\lambda w_j) = \mathcal{I}(\lambda) u_j + \mathcal{R}(\lambda) v_j.$$

$\square$

El resultado anterior se puede extender a los endomorfismos de un $\mathbb{R}$-espacio vectorial E. Para ello definimos el espacio vectorial complejo asociado a E por $E^{\mathbb{C}} = E \times E$ junto con la multiplicación por números complejos

$$\mathbb{C} \times E^{\mathbb{C}} \longrightarrow E^{\mathbb{C}},$$

definida por

$$(a + bi) \cdot (u, v) = (au - bv, av + bu).$$

Se definen la parte real, la parte imaginaria y la conjugación, respectivamente, por

$$\mathcal{R}(u, v) = (u, 0), \quad \mathcal{I}(u, v) = (v, 0), \quad c(u, v) = (u, -v).$$

Identificamos E con el $\mathbb{R}$-subespacio de $E^{\mathbb{C}}$ formado por los vectores de la forma $(u, 0)$, con $u \in E$.

Si $f : E \longrightarrow E$ es un $\mathbb{R}$-endomorfismo de E, se define $f^{\mathbb{C}} : E^{\mathbb{C}} \longrightarrow E^{\mathbb{C}}$ por $f^{\mathbb{C}}(u, v) = (f(u), f(v))$. La aplicación $f^{\mathbb{C}}$ es $\mathbb{C}$-lineal y extiende a f.

Teorema 3.7.6. Sea f un endomorfismo de un espacio vectorial real E de dimensión finita. Para cada valor propio no real $\lambda = \alpha + i\beta$ de $f^{\mathbb{C}}$, sea $q_\lambda(t) = t^2 - 2\alpha + |\lambda|^2$.

Sea $w_1, \ldots, w_p$ una $\mathbb{C}$-base del $\mathbb{C}$-subespacio $\operatorname{Ker}(f^{\mathbb{C}} - \lambda I)$ de $E^{\mathbb{C}}$ y, para cada j, notemos

$$u_j = \mathcal{R}(w_j), \quad v_j = \mathcal{I}(w_j)$$

entonces $\{u_1, v_1, \ldots, u_p, v_p\}$ es una $\mathbb{R}$-base del $\mathbb{R}$-subespacio $\operatorname{Ker} q_\lambda(f)$ de E.

La restricción de f a $\operatorname{Ker} q_\lambda(f)$ define un endomorfismo de $\operatorname{Ker} q_\lambda(f)$ que en la base $\{u_1, v_1, \ldots, u_p, v_p\}$ tiene por matriz

$$\begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix} \oplus \cdots \oplus \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}$$

Corolario 3.7.7. Sea f un endomorfismo de un espacio vectorial real E tal que $f^{\mathbb{C}}$ es diagonalizable. Sean $\lambda_1, \ldots, \lambda_r$ los valores propios reales o complejos de f, contados con sus multiplicidades algebraicas. Entonces existe una $\mathbb{R}$-base de E tal que la matriz de f en dicha base es de la forma

$$D_{\lambda_1} \oplus \cdots \oplus D_{\lambda_r}.$$

Ejemplo 3.7.8. Sea E el espacio vectorial $\mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$. El espacio complejo asociado $E^{\mathbb{C}}$ se identifica con el espacio de funciones $\mathcal{C}^\infty(\mathbb{R}, \mathbb{C})$, donde las operaciones de la estructura compleja se hacen sobre los valores de la función. El operador de derivación $D : E \longrightarrow E$ se extiende a un operador complejo

$$E^{\mathbb{C}} \longrightarrow E^{\mathbb{C}}, \quad (u, v) \mapsto (Du, Dv),$$

que denotaremos igualmente por D.

Para cada $\lambda \in \mathbb{C}$, se define la función $e^{\lambda x} \in E^{\mathbb{C}}$ del modo siguiente. Si $\lambda = \alpha + i\beta$, entonces

$$e^{\lambda x} := (e^{\alpha x} \cos \beta x, e^{\alpha x} \sin \beta x).$$

Entonces $D(e^{\lambda x}) = \lambda \cdot e^{\lambda x}$, y $\operatorname{Ker}(D - \lambda \mathbf{1}_{E^{\mathbb{C}}})$ es un subespacio vectorial complejo de $E^{\mathbb{C}}$ de dimensión 1, generado por la función $e^{\lambda x}$.

Ejemplo 3.7.9. Sean $\omega \in \mathbb{R}$ no nulo y F el $\mathbb{R}$-subespacio vectorial de $E = \mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$ definido por

$$F = \{f \in E;\ D^2 f + \omega^2 f = 0\ \} = \operatorname{Ker}(D^2 + \omega^2 \mathbf{1}_E).$$

Puesto que $t^2 + \omega^2 = (t - i\omega)(t + i\omega)$ y $\operatorname{Ker}(D - i\omega \mathbf{1}_{E^{\mathbb{C}}})$ está generado por $e^{i\omega x} = (\cos \omega x, \sin \omega x)$, aplicando el teorema anterior deducimos que $\{\cos \omega x, \sin \omega x\}$ es una base de F.

Más generalmente, si las raíces de $t^2 + at + b$ son $\alpha \pm \beta i$, con $\beta \neq 0$, entonces las soluciones reales de la ecuación

$$f'' + af' + bf = 0,$$

son de la forma

$$c_1 e^{\alpha x} \cos(\beta x) + c_2 e^{\alpha x} \sin(\beta x).$$

Ejemplo 3.7.10. Para una sucesión recurrente de la forma

$$s_{n+2} + a s_{n+1} + b s_n = 0$$

en que las raíces de $t^2 + at + b$ sean $\alpha \pm \beta i$, con $\beta \neq 0$, las soluciones son de la forma

$$s = c_1 \rho^n \cos(n\theta) + c_2 \rho^n \sin(n\theta),$$

donde $\alpha + \beta i = \rho e^{i\theta}$ es la forma polar de $\alpha + \beta i$.