Espacios vectoriales y Aplicaciones lineales

Pedro Montero

Departamento de Matemática, Universidad Técnica Federico Santa María

Estas notas son material complementario al curso de Álgebra Lineal para estudiantes de segundo año en Matemáticas, y tienen como objetivo recordar y recopilar resultados vistos durante primer año que serán usados directamente en el curso.

Valparaíso, Marzo 2020

Índice

1. Grupos, anillos y cuerpos

1.1. Definiciones generales

Definición 1.1 (grupo)**.** Un **grupo** es un conjunto no-vacío G dotado de una ley de composición interna

$$\begin{aligned} G \times G &\longrightarrow G \\ (g_1, g_2) &\longmapsto g_1 g_2 \end{aligned}$$

que satisface las siguientes condiciones:

1. **asociatividad:** para todos $g_1, g_2, g_3 \in G$ tenemos que
$$(g_1 g_2) g_3 = g_1 (g_2 g_3);$$
2. **elemento neutro:** existe un elemento $e \in G$ (necesariamente único) tal que para todo $g \in G$ tenemos que
$$ge = eg = g;$$
3. **inverso:** para todo $g \in G$ existe un elemento $g^{-1} \in G$ (necesariamente único) tal que
$$gg^{-1} = g^{-1}g = e.$$

Observación 1.2. *Un conjunto no vacío S dotado de una ley de composición interna asociativa y tal que existe un elemento neutro (es decir, que verifica las condiciones (1) y (2)) es llamado un* **semi-grupo**.

Decimos que el grupo G es **abeliano** (o **conmutativo**) si para todos $g_1, g_2 \in G$ tenemos que $g_1 g_2 = g_2 g_1$. En cuyo caso, la ley de composición interna es generalmente escrita de forma aditiva $g_1 + g_2$, el elemento neutro es denotado 0, y el inverso de g es llamado el elemento **opuesto**, el cual es denotado $-g$.

Definición 1.3 (anillo y cuerpo)**.** Sea $(A, +, \cdot)$ un conjunto no-vacío con dos leyes de composición interna. Se dice que A es un **anillo** si:

1. $(A, +)$ es un grupo abeliano.
2. $(A, \cdot)$ es un semi-grupo.
3. Para todos $a, b, c \in A$ se tiene que $a(b + c) = ab + ac$ y $(b + c)a = ba + ca$.

Además, se dice que A es un **anillo abeliano** si $ab = ba$ para todos $a, b \in A$. Finalmente, diremos que un anillo abeliano k es un **cuerpo** si $k \neq \{0\}$ y si $(k \setminus \{0\}, \cdot)$ es un grupo.

Ejemplo 1.4.

1. Los enteros con la suma $(\mathbb{Z}, +)$ forman un grupo abeliano.
2. Si k es un cuerpo (como $\mathbb{Q}, \mathbb{R}$ o $\mathbb{C}$), $(k, +)$ y $(k \setminus \{0\}, \cdot)$ son grupos abelianos. Más generalmente, para un anillo A tenemos el grupo abeliano $(A, +)$ y el grupo multiplicativo $(A^\times, \cdot)$ de **unidades** de A (los elementos de A que son inversibles respecto a la multiplicación). En particular, si k es un cuerpo entonces $k^\times = k \setminus \{0\}$.

1.2. El cuerpo $\mathbb{F}_p$

El objetivo de esta sección es recordar la construcción del cuerpo finito $\mathbb{F}_p$.

Definición 1.5 (relación de equivalencia)**.** Sea A un conjunto y sea $\mathscr{R}$ una relación en A (es decir, un subconjunto $\mathscr{R} \subseteq A \times A$). Si para todo $(a, b) \in A \times A$ tal que $(a, b) \in \mathscr{R}$ escribimos $a \sim b$, entonces decimos que $\mathscr{R}$ es una **relación de equivalencia** si es:

1. **reflexiva:** $a \sim a$ para todo $a \in A$,
2. **simétrica:** $a \sim b$ si y sólo si $b \sim a$ para todos $a, b \in A$,
3. **transitiva:** si $a \sim b$ y $b \sim c$ entonces $a \sim c$, para todos $a, b, c \in A$.

Definición 1.6 (clase de equivalencia)**.** Sea $\mathscr{R}$ una relación de equivalencia en A. Para todo $a \in A$ diremos que el conjunto

$$[a]_{\mathscr{R}} = \{b \in A \mid a \sim b\} = \{b \in A \mid b \sim a\}$$

es la **clase de equivalencia** de $a \in A$ respecto a $\mathscr{R}$, el cual es también denotado $a \mod \mathscr{R}$. En caso que la relación $\mathscr{R}$ sea clara en el contexto, escribiremos simplemente $[a]$ o bien $\overline{a}$.

Definición 1.7 (cociente)**.** Sea $\mathscr{R}$ una relación de equivalencia en A. El conjunto cuyos elementos son todas las clases de equivalencia es llamado **conjunto cociente** de A por $\mathscr{R}$, y será denotado $A/\mathscr{R}$ (o simplemente $A/\sim$ si la relación $\mathscr{R}$ es clara en el contexto). Explícitamente,

$$A/\mathscr{R} = \{[a]_{\mathscr{R}},\ a \in A\}.$$

Las relaciones de equivalencia satisfacen las siguientes propiedades.

Proposición 1.8. *Sea $\mathscr{R}$ una relación de equivalencia en A. Entonces:*

1. *Para todo $a \in A, a \in [a]_{\mathscr{R}}$. En particular, $[a]_{\mathscr{R}} \neq \emptyset$.*
2. *Si $b \in [a]_{\mathscr{R}}$ entonces $a \in [b]_{\mathscr{R}}$. Además, en este caso tenemos que $[a]_{\mathscr{R}} = [b]_{\mathscr{R}}$.*
3. *Para todos $a, b \in A$ ya sea $[a]_{\mathscr{R}} = [b]_{\mathscr{R}}$ o bien $[a]_{\mathscr{R}} \cap [b]_{\mathscr{R}} = \emptyset$. En particular, A es la unión disjunta de las clases $[a]_{\mathscr{R}}$.*

Demostración. Ejercicio al lector. □

Uno de los principales ejemplos de relación de equivalencia es la **congruencia módulo** $n \in \mathbb{N}^{\geq 1}$.

Ejemplo 1.9. Sea $n \in \mathbb{N}^{\geq 1}$. Consideremos la relacion en $\mathbb{Z}$ dada por

$$\begin{aligned} a \sim b &\Leftrightarrow n \text{ divide } a-b \\ &\Leftrightarrow \exists k \in \mathbb{Z} \text{ tal que } a-b = nk. \end{aligned}$$

No es difícil verificar que la relación anterior es en efecto una relación de equivalencia. Utilizaremos la siguiente notación en lo que sigue:

- Si $a \sim b$, escribimos $a \equiv b \pmod{n}$ y diremos que "a es congruente con b módulo n".
- La clase de equivalencia de a módulo n está dada por

$$[a]_n = \{b \in \mathbb{Z} \mid a \equiv b \pmod{n}\}.$$

- El conjunto de clases de equivalencia, "$\mathbb{Z}$ módulo $n\mathbb{Z}$", es denotado por

$$\mathbb{Z}/n\mathbb{Z} = \{[a]_n,\ a \in \mathbb{Z}\}.$$

Por ejemplo, si $n = 2$ entonces observamos que

$$[0]_2 = \{a \in \mathbb{Z} \mid a \equiv 0 \pmod{2}\} = \{a \in \mathbb{Z} \mid \exists k \in \mathbb{Z},\ a = 2k\}$$

es el conjunto de enteros pares. De manera similar, $[1]_2$ es el conjunto de enteros impares. Luego, el cociente $\mathbb{Z}/2\mathbb{Z} = \{[0]_2, [1]_2\}$ es un conjunto con 2 elementos.

Recuerdo (division euclideana): Sean $a, b \in \mathbb{Z}$ con $b \neq 0$. Entonces existen únicos enteros $q, r \in \mathbb{Z}$ tales que $a = bq + r$ y $0 \leq r < |b|$.

Ejercicio 1.10.

a) Utilizando la división euclideana en $\mathbb{Z}$, demostrar que

$$\mathbb{Z}/n\mathbb{Z} = \{[0]_n, [1]_n, \ldots, [n-1]_n\}.$$

b) Demostrar que para todo $a, b \in \mathbb{Z}$ se tiene que $[a+b]_n$ y $[ab]_n$ dependen sólamente de $[a]_n$ y $[b]_n$, es decir, si $a \equiv a' \pmod{n}$ y $b \equiv b' \pmod{n}$ entonces $[a+b]_n = [a'+b']_n$ y $[ab]_n = [a'b']_n$. En particular, la suma $[a]_n + [b]_n := [a+b]_n$ y el producto $[a]_n \cdot [b]_n := [ab]_n$ de clases de equivalencias están bien definidos.

Lema 1.11 (Bézout)**.** *Sean $a, b \in \mathbb{Z}$ no nulos. Entonces existen $x, y \in \mathbb{Z}$ tales que $ax + by = \mathrm{mcd}(a, b)$.*

Demostración. Sea $S = \{ax + by \mid x, y \in \mathbb{Z}\}$ y sea d el menor elemento en $S \cap \mathbb{N}^{\geq 1}$. Observemos que d divide a a. En efecto, la división euclidiana permite escribir:

$$a = qd + r, \quad \text{con } q, r \in \mathbb{Z}, 0 \leq r < d$$

Dado que $a, d \in S$, tenemos que $r = a - qd \in S$. Por otro lado, como d es el mínimo de $S \cap \mathbb{N}^{\geq 1}$ tenemos necesariamente que $r = 0$ (de lo contrario, se obtiene una contradicción con la minimalidad de d). Por lo tanto, $d \mid a$ ("d divide a a"). Análogamente, $d \mid b$.
Finalmente, si d' un divisor en común de a y de b entonces d' divide a todos los elementos de S. En particular, $d' \mid d$ y luego $d' \leq d$. Se concluye de esta manera que $d = \mathrm{mcd}(a, b) = ax_0 + by_0$ para ciertos $x_0, y_0 \in \mathbb{Z}$. □

Corolario 1.12. *Sea p un número primo. Entonces para todo $a \in \mathbb{Z}$ tal que $p \nmid a$, existe $b \in \mathbb{Z}$ tal que $p \nmid b$ tal que $ab \equiv 1 \pmod{p}$.*

Demostración. Dado que p no divide a $a \in \mathbb{Z}$ se tiene que $\mathrm{mcd}(a, p) = 1$, pues p es primo. Entonces, el lema de Bézout implica que existen $x, y \in \mathbb{Z}$ tales que $ax + py = 1$. Equivalentemente,

$$ax - 1 = p(-y)$$

Si definimos $b = x$, entonces $ab \equiv 1 \pmod{p}$. □

Una consecuencia del corolario anterior es que para todo número primo p, el conjunto

$$\mathbb{Z}/p\mathbb{Z} = \{[0]_p, [1]_p, ..., [p-1]_p\}$$

es un **cuerpo** (ver Definición 1.3). En efecto, para todo $[a]_p \neq [0]_p$ existe $[a]_p^{-1}$ tal que $[a]_p \cdot [a]_p^{-1} = [1]_p$.

Definición 1.13. Sea p un número primo. Denotaremos por $\mathbb{F}_p$ al **cuerpo de p elementos** $(\mathbb{Z}/p\mathbb{Z}, +, \cdot)$.

Ejercicio 1.14. Calcular las tablas de suma y multiplicación en $\mathbb{F}_3$.

Ejercicio 1.15. Sea p un número primo.

a) Sea $k \in \{1, \ldots, p-1\}$. Probar que p divide a $\binom{p}{k} = \frac{p!}{k!(p-k)!}$.

b) Probar que para todos $x, y \in \mathbb{F}_p$ se tiene $(x+y)^p = x^p + y^p$.

2. Espacios vectoriales: definición y ejemplos

Antes de dar la defínción formal de espacio vectorial, veamos tres ejemplos importantes.

Ejemplo 2.1.

1. Un ejemplo de un espacio vectorial sobre $\mathbb{R}$ es el espacio de dimensión 3

$$\mathbb{R}^3 = \{(x, y, z) \mid x, y, z \in \mathbb{R}\}.$$

   En este caso, el espacio vectorial nos es presentado como un conjunto de n-tuplas (aquí $n = 3$) de «coordenadas».

2. Sean $a, b \in \mathbb{R}$. El conjunto $\mathscr{S}(a, b)$ de sucesiones $(u_n)_{n\in\mathbb{N}}$ de números reales verificando la relación de recurrencia lineal

$$u_{n+2} = au_{n+1} + bu_n \qquad (\star)$$

   es un $\mathbb{R}$-espacio vectorial. Dicho espacio es de dimensión 2, pues toda sucesión verificando $(\star)$ está determinada por sus términos iniciales u_0 y u_1, que pueden ser escogidos arbitrariamente.

3. Sean $a, b, t_0 \in \mathbb{R}$. El conjunto $\mathscr{S}(a, b)$ de funciones $f : \mathbb{R} \to \mathbb{R}$ de clase $\mathscr{C}^\infty$ verificando la ecuación diferencial ordinaria lineal

$$f''(t) = af'(t) + bf(t) \qquad (E)$$

   para todo $t \in \mathbb{R}$, es un $\mathbb{R}$-espacio vectorial. Dicho espacio es de dimensión 2, pues toda solución de (E) está determinada por las «condiciones iniciales» $f(t_0)$ y $f'(t_0)$, que pueden ser escogidas arbitrariamente.

En los últimos dos ejemplos, la elección de «coordenadas» para el espacio $\mathscr{S}(a, b)$ no es para nada evidente. Una de las gracias del álgebra lineal es que ella permite describir simplemente todos los elementos de $\mathscr{S}(a, b)$, una vez que hemos elegido una base apropiada de dicho espacio.

Importante: La noción de espacio vectorial está definida para todo cuerpo k (tales como $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ o $\mathbb{F}_p$).

Definición 2.2 (espacio vectorial)**.** Sea k un cuerpo. Un **k-espacio vectorial** V es un grupo abeliano $(V, +)$ dotado de una operación «multiplicación por escalares» $(\lambda, v) \mapsto \lambda \cdot v$ de k sobre V verificando las dos condiciones siguientes:

(i) $1 \cdot v = v$ y $\lambda \cdot (\lambda' \cdot v) = (\lambda\lambda') \cdot v$.

(ii) $(\lambda + \lambda') \cdot v = \lambda \cdot v + \lambda' \cdot v$ y $\lambda \cdot (v + v') = \lambda \cdot v + \lambda \cdot v'$.

Podemos memorizar la condición (i) diciendo que «1 actúa como la identidad» y que la «multiplicación por escalares es asociativa», y la condición (ii) diciendo que la «multiplicación por escalares es compatible con la suma».

Observación 2.3 (vector nulo)**.** *Dado que $(V,+)$ es un grupo abeliano, V posee un elementro neutro, que denotaremos provisoriamente 0_V. Por ejemplo, si $V=\mathbb{R}^3$ entonces $0_V=(0,0,0)$.*

Sea 0 el elemento cero del cuerpo k. Entonces la condición (ii) implica que para todo $v\in V$ se tiene

$$0\cdot v=(0+0)\cdot v=0\cdot v+0\cdot v,$$

de donde se deduce que $0\cdot v=0_V$. En consecuencia, el vector nulo 0_V será denotado simplemente (y por abuso de notación) como 0. Así, denotamos por $\{0\}$ al espacio vectorial nulo. Por ejemplo, en $\mathbb{R}^2$ el espacio de soluciones del sistema

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

es el espacio vectorial nulo $\{0\}=\{(0,0)\}$.

Terminología: Sea k un cuerpo. Si V es un k-espacio vectorial, entonces decimos que k es el *cuerpo de escalares* de V, que los elementos de k son *escalares*, y que los elementos de V son *vectores*.

Ejemplo 2.4. Sea k un cuerpo.

1. Para todo $n\in\mathbb{N}^{\geq 1}$, el conjunto
$$k^n:=\{(x_1,\ldots,x_n)\mid x_i\in k\}$$
es un k-espacio vectorial.
2. El conjunto $M_{m\times n}(k)$ de matrices con m filas y n columnas con coeficientes en k es un k-espacio vectorial. Si $m=n$, lo denotaremos simplemente $M_n(k)$.
3. El anillo de polinomios en una variable $k[X]$ es un k-espacio vectorial.

Definición 2.5 (sub-espacio)**.** Sea V un k-espacio vectorial. Un **sub-espacio vectorial** W de V es un sub-conjunto de V que es un sub-grupo[1] y que es estable por la multiplicación por escalares.

En otras palabras, $W\subseteq V$ es un sub-espacio vectorial si $W\neq\emptyset$ y si para todos $w_1,w_2\in W$ y todo $\lambda\in k$, se tiene que $\lambda\cdot w_1+w_2\in W$.

Ejemplo 2.6.

1. El «plano horizontal»
$$\Pi=\{(x,y,0)\mid x,y\in\mathbb{R}\}$$
dado por la ecuación $z=0$, es un sub-espacio vectorial de $\mathbb{R}^3$.
2. El conjunto de matrices $(a_{ij})_{1\leq i,j\leq 3}\in M_3(\mathbb{R})$ que son triangulares superiores, es decir, tales que $a_{21}=a_{31}=a_{32}=0$, es un sub-espacio vectorial de $M_3(\mathbb{R})$.
3. El conjunto $k_d[X]$ de polinomios en X de grado $\leq d$ es un sub-espacio vectorial de $k[X]$.

Definición 2.7 (aplicación lineal)**.** Sea k un cuerpo y sean V y W dos k-espacios vectoriales. Decimos que una función $\varphi:V\to W$ es una **aplicación lineal**[2] si preserva la suma y la multiplicación por escalares, es decir, si para todos $v_1,v_2\in V$ y $\lambda\in k$ se tiene que

$$\varphi(v_1+v_2)=\varphi(v_1)+\varphi(v_2),\qquad \varphi(\lambda\cdot v_1)=\lambda\cdot\varphi(v_1).$$

Observación 2.8. *Cabe notar que podemos reagrupar estas dos condiciones en una sola:*

$$\varphi(\lambda\cdot v_1+v_2)=\lambda\cdot\varphi(v_1)+\varphi(v_2),$$

que a su vez es equivalente a la condición siguiente:

$$\varphi(\lambda_1\cdot v_1+\lambda_2\cdot v_2)=\lambda_1\cdot\varphi(v_1)+\lambda_2\cdot\varphi(v_2)$$

para todos $v_1,v_2\in V$ y $\lambda_1,\lambda_2\in k$.

[1] Explícitamente, se verifica que (1) $0\in W$, (2) si $w_1,w_2\in W$ entonces $w_1+w_2\in W$, y (3) si $w\in W$ entonces $-w\in W$.
[2] También se usan comunmente los términos **transformación lineal** y **(homo)morfismo** entre espacios vectoriales.

Definición 2.9 (endomorfismo)**.** Sea k un cuerpo y V un k-espacio vectorial. Una aplicación lineal $\varphi : V \to V$ de V en sí mismo es llamada un **endomorfismo**.

Ejemplo 2.10.

1. Sea $V = k[X]$. La aplicación $d : V \to V$ que asocia a todo polinomio $P(X) = a_n X^n + \ldots + a_1 X + a_0$ su derivada $P'(X) = n a_n X^{n-1} + \ldots + a_1$ es una aplicación lineal (endomorfismo).

2. Sea $V = \mathscr{C}([0,1], \mathbb{R})$ el $\mathbb{R}$-espacio vectorial de funciones continuas $f : [0,1] \to \mathbb{R}$. Entonces la aplicación
$$V \to \mathbb{R}, \qquad f \mapsto \int_0^1 f(x)\, \mathrm{d}x$$
es lineal, pero la aplicacíón $f \mapsto \int_0^1 f(x)^2\, \mathrm{d}x$ no lo es.

Definición 2.11 (isomorfismo)**.** Sean V y W dos k-espacios vectoriales, y sea $\varphi : V \to W$ una aplicación lineal. Decimos que φ es un **isomorfismo** (de espacios vectoriales) si φ es una función biyectiva y si su función inversa $\varphi^{-1} : W \to V$ es lineal.

Observación 2.12. *Es importante notar que la segunda condición es automaticamente verificada. En efecto, si escribimos $\psi = \varphi^{-1} : W \to V$ y consideramos $w_1, w_2 \in W$ y $\lambda \in k$, entonces el hecho que φ es biyectiva implica que existen $v_1, v_2 \in V$ únicos tales que $\varphi(v_1) = w_1$ y $\varphi(v_2) = w_2$. Más aún, dado que φ es lineal tenemos que*
$$\varphi(\lambda \cdot v_1 + v_2) = \lambda \cdot \varphi(v_1) + \varphi(v_2) = \lambda \cdot w_1 + w_2.$$
Aplicando ψ a esta última igualdad obtenemos
$$\psi(\lambda \cdot w_1 + w_2) = \lambda \cdot v_1 + v_2 = \lambda \cdot \psi(w_1) + \psi(w_2).$$
En conclusión, toda aplicación lineal biyectiva es un isomorfismo.

Notación: Sean V y W dos k-espacios vectoriales, y sea $\varphi : V \to W$ una aplicación lineal. Si φ es

1. *inyectiva*, escribimos $\varphi : V \hookrightarrow W$.
2. *sobreyectiva*, escribimos $\varphi : V \twoheadrightarrow W$.
3. *biyectiva*, escribimos $\varphi : V \xrightarrow{\sim} W$.

Decimos que V y W son **isomorfos** si existe un isomorfismo $\varphi : V \xrightarrow{\sim} W$ entre ellos. En este caso escribimos $V \cong W$.

Ejemplo 2.13. El conjunto $k^{\mathbb{N}}$ de todas las sucesiones $(u_n)_{n\in\mathbb{N}}$ de elementos $u_n \in k$ dotado de la suma y multiplicación por escalares «término a término», es decir,
$$(u_n)_{n\in\mathbb{N}} + (v_n)_{n\in\mathbb{N}} := (u_n + v_n)_{n\in\mathbb{N}}, \qquad \lambda \cdot (u_n)_{n\in\mathbb{N}} := (\lambda u_n)_{n\in\mathbb{N}},$$
es un k-espacio vectorial. Sean $a, b \in k$ y sea $\mathscr{S}(a,b)$ el sub-conjunto de $k^{\mathbb{N}}$ formado por las sucesiones $(u_n)_{n\in\mathbb{N}}$ verificando la relación de recurrencia lineal
$$u_{n+2} = a u_{n+1} + b u_n. \tag{$\star$}$$
Entonces $\mathscr{S}(a,b)$ es un sub-espacio vectorial de $k^{\mathbb{N}}$. Más aún, la aplicación $\varphi : \mathscr{S}(a,b) \to k^2$ que a toda sucesión $(u_n)_{n\in\mathbb{N}}$ le asocia el par (u_0, u_1) es lineal (ejercicio); ella es sobreyectiva (pues podemos escoger arbitrariamente u_0 y u_1), e inyectiva (pues los u_n están determinados a partir de u_0 y u_1 gracias a la fórmula ($\star$)). Luego, $\varphi : \mathscr{S}(a,b) \xrightarrow{\sim} k^2$ es un isomorfismo de espacios vectoriales.

Ejercicio 2.14. Sea k un cuerpo y sea $k_d[X]$ el k-espacio vectorial de polinomios con coeficientes en k en la variable X de grado $\leq d$. Demostrar que $k_d[X] \cong k^{d+1}$.

3. Familias generadoras, libres, bases y dimensión

Definición 3.1 (espacio generado)**.** Sea V un k-espacio vectorial. Sea S un sub-conjunto (finito o infinito) no-vacío de V. Denotamos[3] por $\mathrm{Vect}_k(S)$ al conjunto de todas las *combinaciones lineales* finitas de elementos de S, es decir, al conjunto con elementos de la forma

$$\sigma = \lambda_1 v_1 + \ldots \lambda_r v_r, \quad \text{donde } r \in \mathbb{N}^{\geq 1},\ v_i \in S,\ \lambda_i \in k.$$

Dicho conjunto es un sub-espacio vectorial de V (ejercicio). Más aún, si W es un sub-espacio vectorial de V conteniendo a S, entonces W contiene toda combinación lineal σ, es decir, contiene $\mathrm{Vect}_k(S)$. Así, $\mathrm{Vect}_k(S)$ *es el sub-espacio vectorial más pequeño de* V *conteniendo* S, y lo llamamos el **sub-espacio vectorial generado por** S.

Notación: Si $S = \{v_1, \ldots, v_n\}$ es un conjunto finito de vectores, escribimos simplemente

$$\mathrm{Vect}_k(S) = \mathrm{Vect}_k(v_1, \ldots, v_n) = \{\lambda_1 v_1 + \ldots + \lambda_n v_n \mid \lambda_1, \ldots, \lambda_n \in k\}.$$

Definición 3.2 (familia generadora)**.** Sea V un k-espacio vectorial. Sea S un sub-conjunto (finito o infinito) no-vacío de V. Decimos que S es un **conjunto de generadores** (o bien, una **familia generadora**) de V si $\mathrm{Vect}_k(S) = V$, es decir, si todo elemento de V se escribe como una combinación lineal de elementos de S. Más aún, diremos que V es **finitamente generado** si posee un conjunto de generadores finitos, es decir, si $V = \mathrm{Vect}_k(v_1, \ldots, v_n)$ para ciertos $v_1, \ldots, v_n \in V$.

Una consecuencia inmediata de la definición anterior es que *toda familia conteniendo una familia generadora es generadora.*

Ejemplo 3.3. Sea k un cuerpo.

1. Para todo $n \in \mathbb{N}^{\geq 1}$, el espacio k^n está generado por los vectores
$$e_1 = (1, 0, \ldots, 0),\ e_2 = (0, 1, 0, \ldots, 0), \ldots,\ e_n = (0, \ldots, 0, 1).$$
2. Los monomios X^n, donde $n \in \mathbb{N}$, generan el espacio vectorial $k[X]$. En efecto, todo polinomio $P \in k[X]$ se escribe como una combinación lineal finita: $P = a_0 + a_1 X + \ldots + a_d X^d$.

Definición 3.4 (dependencia lineal)**.** Sea V un k-espacio vectorial. Sea S un sub-conjunto (finito o infinito) no-vacío de V. Decimos que los elementos de S son **linealmente independientes** (o bien, que S es una **familia libre**) si no existen relaciones lineales no triviales entre los elementos de S, es decir, si la condición siguiente es verificada:

> Para todo $r \in \mathbb{N}^{\geq 1}$, todo conjunto de vectores $v_1, \ldots, v_r \in S$ distintos, y todo conjunto de escalares $\lambda_1, \ldots, \lambda_r \in k$ tenemos que: si se verifica la relación $\lambda_1 v_1 + \ldots + \lambda_r v_r = 0$, entonces $\lambda_1 = \ldots = \lambda_r = 0$.

En caso contrario, diremos que los elementos de S son **linealmente dependiente**. En otras palabras, si existe una relación lineal no trivial entre los elementos de S, es decir, si existe un entero $r \in \mathbb{N}^{\geq 1}$, vectores $v_1, \ldots, v_r \in S$ distintos, y escalares $\lambda_1, \ldots, \lambda_r \in k$ *no todos* iguales a 0, tales que $\lambda_1 v_1 + \ldots + \lambda_r v_r = 0$.

Caso particular: La definición anterior se simplifica si el conjunto $S = \{v_1, \ldots, v_n\}$ es *finito*. En tal caso, los vectores $v_1, \ldots, v_n$ son linealmente independientes si:

> Para todo conjunto de escalares $\lambda_1, \ldots, \lambda_n \in k$ tenemos que: si se verifica la relación lineal $\lambda_1 v_1 + \ldots + \lambda_n v_n = 0$, entonces $\lambda_1 = \ldots = \lambda_n = 0$.

En caso contrario, tenemos que los vectores $v_1, \ldots, v_n$ son linealmente dependientes si existen escalares $\lambda_1, \ldots, \lambda_n \in k$ no todos nulos, tales que $\lambda_1 v_1 + \ldots + \lambda_n v_n = 0$. En este último caso, si por ejemplo $\lambda_i \neq 0$, podemos expresar v_i en función de los otros v_j con $j \neq i$ mediante

$$v_i = -\sum_{j \neq i} \frac{\lambda_j}{\lambda_i} v_j.$$

[3] También se usa comunmente la notación inglesa $\mathrm{Span}(S)$.

Observación 3.5. *A partir de la definición anterior se tiene que:*

1. *Toda sub-familia de una familia libre es libre.*
2. *Todo conjunto conteniendo un sub-conjunto linealmente dependiente es linealmente dependiente.*

Ejemplo 3.6. Sea k un cuerpo.

1. En k^n, el conjunto $\{e_1, \ldots, e_n\}$ es linealmente independiente. En efecto, para todos $\lambda_1, \ldots, \lambda_n \in k$ tenemos que
$$\lambda_1 e_1 + \ldots + \lambda_n e_n = (\lambda_1, \ldots, \lambda_n),$$
por lo que si la suma a la izquierda es nula, entonces $\lambda_1 = \ldots = \lambda_n = 0$.
2. En $k[X]$, la familia de monomios $\{X^n\}_{n\in\mathbb{N}}$ es linealmente independiente. En efecto, sean $i_1 < \ldots < i_r$ en $\mathbb{N}$ y sea $\lambda_1, \ldots, \lambda_r \in k$ todos no nulos, entonces el polinomio
$$\lambda_1 X^{i_1} + \ldots + \lambda_r X^{i_r}$$
es no nulo. Esto demuestra que $\{X^n\}_{n\in\mathbb{N}}$ es una familia libre.
3. Si $k = \mathbb{Q}$ y $V = k^2$ entonces los vectores $v_1 = (1, 0)$ y $v_2 = (-1, 0)$ son linealmente dependientes. En efecto,
$$1 \cdot v_1 + 1 \cdot v_2 = 0,$$
pero $(\lambda_1, \lambda_2) = (1, 1) \neq (0, 0)$. Cabe notar que si $k = \mathbb{F}_2$ entonces $v_1 = v_2$.

El siguiente lema caracteriza los conjuntos linealmente independientes infinitos.

Lema 3.7. *Sea V un k-espacio vectorial y $S = \{v_i\}_{i\in I}$ una familia de vectores indexada por un conjunto de índices I infinito. Entonces S es libre si y sólo si la siguiente condición de unicidad es verificada:*

Si tenemos una igualdad
$$\sum_{j\in J} \lambda_j v_j = \sum_{p\in P} \mu_p v_p, \quad \lambda_j \in k,\ \mu_p \in k,$$
donde J, P son dos sub-conjuntos finitos de I, entonces los conjuntos $\{j \in J \mid \lambda_j \neq 0\}$ y $\{p \in P \mid \mu_p \neq 0\}$ son iguales y, denotando por L dicho conjunto, tenemos que $\lambda_\ell = \mu_\ell$ para todo $\ell \in L$.

Demostración. Supongamos que la condición de unicidad es verificada. Si tenemos una igualdad de la forma $\sum_{j\in J} \lambda_j v_j = 0$ (aquí, el término de la derecha corresponde a $P = \emptyset$: una suma indexada por $\emptyset$ vale 0), entonces $\{j \in J \mid \lambda_j \neq 0\}$ es vacío, es decir, todos los λ_j son nulos, de donde se deduce que S es una familia libre.
Recíprocamente, supongamos que S es una familia libre, y que tenemos una igualdad de la forma $\sum_{j\in J} \lambda_j v_j = \sum_{p\in P} \mu_p v_p$, como en la condición de unicidad. Entonces, tenemos que
$$0 = \sum_{j\in J\setminus P} \lambda_j v_j + \sum_{i\in J\cap P} (\lambda_i - \mu_i) v_i - \sum_{p\in P\setminus J} \mu_p v_p,$$
y como S es libre, esto implica que $\lambda_j = 0 = \mu_p$ para $j \in J \setminus P$ y $p \in P \setminus J$, y que $\lambda_i = \mu_i$ para todo $i \in J \cap P$, de donde se concluye que la condición de unicidad es verificada. □

Definición 3.8 (base)**.** Sea V un k-espacio vectorial y $\{v_i\}_{i\in I}$ una familia (finita o infinita) de vectores de V. Decimos que $\mathscr{B} = \{v_i\}_{i\in I}$ es una **base** de V si todo vector $v \in V$ *se escribe de manera única como combinación lineal de los v_i*, es decir, si:

1. $\mathscr{B}$ es una familia generadora, es decir, para todo $v \in V$, existe un sub-conjunto *finito* J de I (que depende de v) y escalares $\lambda_j \in k$, con $j \in J$, tales que $v = \sum_{j\in J} \lambda_j v_j$.

2. $\mathscr{B}$ verifica la condición de unicidad siguiente: si tenemos un segundo sub-conjunto finito P de I y escalares μ_p, con $p \in J$, tales que

$$v = \sum_{j \in J} \lambda_j v_j = \sum_{p \in P} \mu_p v_p,$$

entonces el sub-conjunto $L = \{j \in J \mid \lambda_j \neq 0\}$ coincide con $\{p \in P \mid \mu_p \neq 0\}$ y para todo $\ell \in L$ tenemos $\lambda_\ell = \mu_\ell$.

Gracias al lema anterior, esto quiere decir que $\mathscr{B}$ es una familia libre y generadora.

Caso particular: La definición anterior se simplifica si el conjunto $\mathscr{B} = \{v_1, \ldots, v_n\}$ es *finito*. En tal caso, los vectores $v_1, \ldots, v_n$ forman una base de V si para todo $v \in V$ existe una única n-tupla $(\lambda_1, \ldots, \lambda_n) \in k^n$ tal que $v = \lambda_1 v_1 + \ldots + \lambda_n v_n$.

Ejemplo 3.9. Sea k un cuerpo.

1. Si $V = k^n$, la familia $\{e_1, \ldots, e_n\}$ dada por

$$e_1 = (1, 0, \ldots, 0),\ e_2 = (0, 1, 0, \ldots, 0), \ldots,\ e_n = (0, \ldots, 0, 1),$$

es una base, llamada la **base canónica** de k^n.

2. Si $V = M_{m \times n}(k)$ es el k-espacio vectorial de matrices con m filas y n columnas con coeficientes en k, entonces denotamos por E_{ij} la **matriz elemental** cuyos coeficientes son todos nulos, salvo aquel de índice (i, j) (es decir, aquel situado sobre la línea i y la columna j), que vale 1. Entonces, toda matriz $A \in M_{m \times n}(k)$ *se escribe de manera única* como combinación lineal de las E_{ij}:

$$A = \sum_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} a_{ij} E_{ij},$$

donde a_{ij} es el coeficiente de índice (i, j) de A. Así, la familia $\{E_{ij}\}_{1 \leq i \leq m,\ 1 \leq j \leq n}$ es una base de $M_{m \times n}(k)$.

3. La familia $\{1, X, \ldots, X^d\}$ es una base del espacio vectorial $k_d[X]$ de polinomios de grado $\leq d$. En efecto, todo polinomio $P \in k_d[X]$ se escribe de manera única como

$$P = a_0 \cdot 1 + a_1 \cdot X + \ldots + a_d \cdot X^d,$$

donde $a_i \in k$. De manera similar, la familia de monomios $\{X_n\}_{n \in \mathbb{N}}$ forma una base de $k[X]$.

Recordemos que un k-espacio vectorial V es *finitamente generado* si posee un conjunto de generadores finito, es decir, si existen $v_1, \ldots, v_n \in V$ tales que $V = \mathrm{Vect}_k(v_1, \ldots, v_n)$. El siguiente resultado, cuya demostración será discutida en el Apéndice de este texto, es uno de los resultados fundamentales sobre espacios vectoriales finitamente generados.

Teorema 3.10. *Sea V un k-espacio vectorial finitamente generado. Entonces:*

1. *Existen bases de V, y todas tienen el mismo cardinal n; este entero se llama la* dimensión *de V sobre k y es denotada* $\dim_k(V)$ *o simplemente* $\dim(V)$.

2. *De toda familia generadora $\mathscr{F}$ podemos extraer una base, en particular $\mathscr{F}$ es de cardinal* $\geq \dim_k(V)$. *Más aún, si* $\mathrm{card}(\mathscr{F}) = \dim_k(V)$ *entonces $\mathscr{F}$ es una base de V.*

3. *Toda familia linealmente independiente es de cardinal* $\leq \dim_k(V)$. *Más aún, toda familia linealmente independiente de cardinal* $\dim_k(V)$ *es una base de V.*

4. Teorema de la base incompleta: *Toda familia linealmente independiente puede ser completada en una base de V.*

5. *Todo sub-espacio W de V es de dimensión finita* $\leq \dim_k(V)$. *Más aún, si* $\dim_k(W) = \dim_k(V)$ *entonces $W = V$. En otras palabras, todo sub-espacio vectorial distinto de V es de dimensión* $< \dim_k(V)$.

Terminología: En vista del teorema anterior, diremos a partir de ahora «k-espacio vectorial de dimensión finita», en lugar de «k-espacio vectorial finitamente generado». Si $n = \dim_k(V)$, diremos que V es de dimensión n.

Ejemplo 3.11. Sea k un cuerpo. Entonces

1. $\dim_k(k^n) = n$.
2. $\dim_k(M_{m\times n}(k)) = mn$.
3. $\dim_k(k_d[X]) = d + 1$.
4. $k[X]$ *no* es de dimensión finita.

Definición 3.12 (coordenadas)**.** Sea V un k-espacio vectorial de dimensión n y sea $\mathscr{B} = (v_1, \ldots, v_n)$ una base (ordenada) de V. Entonces todo vector $v \in V$ se escribe de manera única como

$$v = x_1v_1 + \ldots + x_nv_n;$$

decimos que $(x_1, \ldots, x_n)$ son las **coordenadas** de v respecto a la base $\mathscr{B} = (v_1, \ldots, v_n)$. Así, el darse una base $\mathscr{B}$ provee un isomorfismo de k-espacios vectoriales

$$\varphi_{\mathscr{B}} : k^n \xrightarrow{\sim} V,\ (x_1, \ldots, x_n) \mapsto x_1v_1 + \ldots + x_nv_n.$$

Importante: En la definición anterior, y en lo que sigue del texto, una base $\mathscr{B} = (v_1, \ldots, v_n)$ es una n-tupla *ordenada*. Por ejemplo, si $\mathscr{B} = (v_1, v_2)$ es una base de V, entonces $\mathscr{C} = (v_2, v_1)$ es una base de V distinta de $\mathscr{B}$: la imagen de $(1, 2) \in \mathbb{R}^2$ por $\varphi_{\mathscr{B}}$ es el vector $v_1 + 2v_2$, mientras que su imagen por $\varphi_{\mathscr{C}}$ es el vector $v_2 + 2v_1 \neq v_1 + 2v_2$. En caso que querramos pensar a $\mathscr{B}$ como un conjunto no-ordenado escribiremos $\mathscr{B} = \{v_1, \ldots, v_n\}$ en lugar de $\mathscr{B} = (v_1, \ldots, v_n)$.

Proposición 3.13. *Sea $f : V \to W$ una aplicación lineal entre k-espacios vectoriales.*

1. *Si f es inyectiva y si $\mathscr{F} = (v_1, \ldots, v_n)$ es una familia linealmente independiente de V, entonces $f(\mathscr{F})$ es linealmente independiente en W.*
2. *Si f es sobreyectiva y si $\mathscr{F} = (v_1, \ldots, v_n)$ es una familia generadora de V, entonces $f(\mathscr{F})$ genera W.*
3. *Si f es biyectiva y si $\mathscr{B} = (v_1, \ldots, v_n)$ es una base de V, entonces $f(\mathscr{B})$ es una base de W. En particular,* $\dim_k(V) = \dim_k(W) = n$.

Demostración. Para ver (1) supongamos que $f : V \hookrightarrow W$ es inyectiva y sea $\mathscr{F} = (v_1, \ldots, v_n)$ es una familia linealmente independiente de V. Supogamos que en W existe una relación lineal

$$\lambda_1 f(v_1) + \ldots + \lambda_n f(v_n) = 0,$$

donde $\lambda_i \in k$. Entonces, dado que f es lineal, tenemos que $0 = f(\lambda_1v_1 + \ldots + \lambda_nv_n)$. Como f es inyectiva, $0 = \lambda_1v_1 + \ldots + \lambda_nv_n$. Así, dado que $\mathscr{F}$ es una familia libre, tenemos que $\lambda_1 = \ldots = \lambda_n = 0$, de donde concluimos que $f(\mathscr{F})$ es libre.
Para probar (2) supongamos que $f : V \twoheadrightarrow W$ es sobreyectiva y sea $\mathscr{F} = (v_1, \ldots, v_n)$ es una familia generadora de V. Sea $w \in W$. Dado que f es sobreyectiva, existe $v \in V$ tal que $f(v) = w$. Como $\mathscr{F}$ genera V, existen $\lambda_1, \ldots, \lambda_n \in k$ tales que $v = \lambda_1v_1 + \ldots + \lambda_nv_n$, de donde se concluye que

$$w = f(v) = \lambda_1 f(v_1) + \ldots + \lambda_n f(v_n).$$

En otras palabras, $f(\mathscr{F})$ genera W.
Finalmente, para demostrar (3) supongamos que $f : V \xrightarrow{\sim} W$ es biyectiva y sea $\mathscr{B} = (v_1, \ldots, v_n)$ una base de V. Gracias a (1) y (2), $f(\mathscr{B})$ es una familia libre y generadora de W, luego una base de W. En particular, $\dim_k(W) = n = \dim_k(V)$. □

Corolario 3.14. *Sea k un cuerpo.*

1. *Todo k-espacio vectorial V de dimensión finita es isomorfo (de manera no canónica) a k^n, para un único n igual a $\dim_k(V)$.*

2. *Dos k-espacios vectoriales de dimensión finita V y W son isomorfos si y sólamente si tienen la misma dimensión.*

Demostración. Para probar (1) notamos que si V es de dimensión n entonces, mediante la elección de una base $\mathscr{B}$, es isomorfo a k^n via la aplicación $\varphi_{\mathscr{B}}$ introducida anteriormente[4]. Recíprocamente, si $V \cong k^m$, entonces la proposición anterior implica que $\dim_k(V) = \dim(k^m) = m$, de donde $m = n$. En particular, $k^m \ncong k^n$ no son isomorfos como k-espacios vectoriales si $m \neq n$.
Para probar (2) notamos que si $V \cong W$, entonces $\dim_k(V) = \dim_k(W)$, gracias a la proposición anterior. Recíprocamente, si $\dim_k(V) = \dim_k(W) = n$, entonces V y W son ambos isomorfos a k^n. □

Ejemplo 3.15.

1. La recta de ecuación $x_1 + x_2 = 0$ en $\mathbb{R}^2$ admite como base el vector $e_1 - e_2$, pero también podríamos haber escogido el vector $e_2 - e_1$.

2. El plano de ecuación $x_1 + x_2 + x_3 = 0$ en $\mathbb{R}^3$ admite como base $(e_1 - e_2, e_2 - e_3)$, pero también podríamos haber escogido $(e_1 - e_3, e_2 - e_3)$, o bien $(e_1 - e_2, e_1 + e_2 - 2e_3)$, etc.

Ejemplo 3.16. Retomemos el espacio $\mathscr{S}(a, b)$ de sucesiones $(u_n)_{n\in\mathbb{N}}$ de elementos de k verificando la relación de recurrencia lineal $u_{n+2} = au_{n+1} + bu_n$. Vimos anteriormente que $\mathscr{S}(a, b) \cong k^2$ via la aplicación lineal $(u_n)_{n\in\mathbb{N}} \mapsto (u_0, u_1)$, por lo tanto $\dim_k \mathscr{S}(a, b) = 2$. Supongamos que el polinomio $P = X^2 - aX - b$ tenga dos raíces distintas $\lambda \neq \mu$ en el cuerpo k. Consideremos los elementos $\mathbf{u} = (u_n)_{n\in\mathbb{N}}$ y $\mathbf{v} = (v_n)_{n\in\mathbb{N}}$ de $\mathscr{S}(a, b)$ definidos por

$$u_0 = v_0 = 1,\ u_1 = \lambda,\ v_1 = \mu.$$

Entonces la familia $\mathscr{B} = (\mathbf{u}, \mathbf{v})$ es linealmente independiente (puesto que si $s\mathbf{u} + t\mathbf{v} = 0$, obtenemos que $s + t = 0 = s\lambda + t\mu$, de donde $t = -s$ y $s(\lambda - \mu) = 0$, por lo que $s = 0$), por lo que es una base de $\mathscr{S}(a, b)$. En consecuencia, todo elemento $\mathbf{w} = (w_n)_{n\in\mathbb{N}}$ de $\mathscr{S}(a, b)$ se escribe de manera única como

$$\mathbf{w} = s\mathbf{u} + t\mathbf{v},$$

donde s, t están únicamente determinados por las condiciones $s + t = w_0$ y $s\lambda + t\mu = w_1$.

4. Núcleo, imagen y teorema del rango

Definición 4.1 (núcleo, imagen y rango)**.** Sea $f : V \to W$ una aplicación lineal entre k-espacios vectoriales. Definimos su **núcleo** (o **kernel**, en alemán) como

$$\ker(f) = \{v \in V \mid f(v) = 0\},$$

el cual es un sub-espacio vectorial de V. Notar que f es inyectiva si y sólamente si $\ker(f) = \{0\}$[5].
Por otro lado, definimos su **imagen** como

$$\mathrm{Im}(f) = f(V) = \{f(v) \mid v \in V\} \subseteq W,$$

el cual es un sub-espacio vectorial de W. Notar que f es sobreyectiva si y sólamente si $\mathrm{Im}(f) = W$.
Cuando $\mathrm{Im}(f)$ es de *dimensión finita* $r \in \mathbb{N}$ (esto ocurre, por ejemplo, si W o V son de dimensión finita), el entero $r = \dim_k \mathrm{Im}(f)$ es llamado el **rango** de f y es denotado $\mathrm{rg}(f)$ o bien $\mathrm{rango}(f)$.

Teorema 4.2 (Teorema del rango)**.** *Sea $f : V \to W$ una aplicación lineal entre k-espacios vectoriales. Supongamos que V es de dimensión finita. Entonces,*

$$\dim_k(V) = \dim_k \ker(f) + \mathrm{rg}(f).$$

En particular, f es sobreyectiva si y sólo si W es de dimensión $\dim_k(V) - \dim_k \ker(f)$.

[4] Este isomorfismo **no es canónico**, pues depende de la elección de una base.
[5] En efecto, tenemos que $f(v_1) = f(v_2)$ si y sólamente si $f(v_1 - v_2) = 0$.

Demostración. Sea $n = \dim_k(V) \in \mathbb{N}$. Dado que V es de dimensión finita, $\ker(f) \subseteq V$ es de dimensión finita $d \leq n$. Sea $\mathscr{K} = (e_1, \ldots, e_d)$ una base de $\ker(f)$. Completemos $\mathscr{K}$ en una base $\mathscr{B} = (e_1, \ldots, e_d, e_{d+1}, \ldots, e_n)$ de V. Entonces, $\mathrm{Im}(f) = f(V)$ es generada por $f(\mathscr{B})$ o, equivalentemente, por los vectores $f(e_{d+1}), \ldots, f(e_n)$ puesto que $f(e_i) = 0$ para $i \leq d$. Veamos que dichos son linealmente independientes: supongamos que existe una relación de dependencia lineal

$$0 = \lambda_1 f(e_{d+1}) + \ldots + \lambda_{n-d} f(e_n) = f(\lambda_1 e_{d+1} + \ldots + \lambda_{n-d} e_n)$$

entonces el vector $\lambda_1 e_{d+1} + \ldots + \lambda_{n-d} e_n$ pertenece a $\ker(f)$, y por ende es una combinación lineal de $e_1, \ldots, e_d$, de donde obtenemos

$$\lambda_1 e_{d+1} + \ldots + \lambda_{n-d} e_n - \mu_1 e_1 - \ldots - \mu_d e_d = 0.$$

Como $(e_1, \ldots, e_n)$ es una base de V, esto último implica que $\lambda_i = \mu_j = 0$ para todos i, j. Así, los vectores $f(e_{d+1}), \ldots, f(e_n)$ son linealmente independientes y luego forman una base de $f(V)$, de donde concluimos que $\dim_k f(V) = n - d$ o, equivalentemente, que $\mathrm{rg}(f) = \dim_k f(V) = \dim_k(V) - \dim_k \ker(f)$. □

Veamos a continuación una demostración alternativa del teorema del rango, que tiene como ventaja demostrar el hecho siguiente:

> si $(e_1, \ldots, e_d)$ es una base de $\ker(f)$, si $(w_1, \ldots, w_r)$ es una base de $\mathrm{Im}(f)$, y si $v_1, \ldots, v_r \in V$ satisfacen $f(v_i) = w_i$ para $i = 1, \ldots, r$, entonces $(e_1, \ldots, e_d, v_1, \ldots, v_r)$ es una base de V.

Demostración alternativa del Teorema del rango. Sea $n = \dim_k(V) \in \mathbb{N}$. Dado que V es de dimensión finita, $\ker(f) \subseteq V$ es de dimensión finita $d \leq n$. Sea $\mathscr{K} = (e_1, \ldots, e_d)$ una base de $\ker(f)$.
Como $\mathrm{Im}(f) = f(V)$ está generado por n vectores (las imagenes de una base de V), es de dimensión finita $r \leq n$. Sea $(w_1, \ldots, w_r)$ una base de $\mathrm{Im}(f)$ y para $i = 1, \ldots, r$ sea $v_i \in V$ un vector tal que $f(v_i) = w_i$. Entonces la familia

$$\mathscr{B} = (e_1, \ldots, e_d, v_1, \ldots, v_r)$$

es una base de V. En efecto, ella es generadora: sea $v \in V$ arbitrario, su imagen $f(v)$ se escribe como $f(v) = \lambda_1 w_1 + \ldots + \lambda_r w_r$, de donde $f(v - \lambda_1 w_1 - \ldots - \lambda_r w_r) = 0$. Luego, $v - \lambda_1 w_1 - \ldots - \lambda_r w_r \in \ker(f)$ y por ende podemos se escribe como $\mu_1 e_1 + \ldots + \mu_d e_d$, de donde obtenemos

$$v = \lambda_1 v_1 + \ldots + \lambda_r v_r + \mu_1 e_1 + \ldots + \mu_d e_d.$$

Esto demuestra que $\mathscr{B}$ es una familia generadora. Veamos que ella es linealmente independiente: si

$$\lambda_1 v_1 + \ldots + \lambda_r v_r + \mu_1 e_1 + \ldots + \mu_d e_d = 0$$

entonces $0 = f(0) = \lambda_1 w_1 + \ldots + \lambda_r w_r$, de donde concluímos que cada λ_i es nulo (pues $(w_1, \ldots, w_r)$ son linealmente independientes), y luego $0 = \mu_1 e_1 + \ldots + \mu_d e_d$, de donde concluímos que cada μ_i es nulo (pues $(e_1, \ldots, e_d)$ son linealmente independientes). De este modo, tenemos que $\mathscr{B}$ es linealmente independiente, y por ende una base de V. Finalmente, se tiene que $\dim_k(V) = d + r = \dim_k \ker(f) + \dim_k \mathrm{rg}(f)$. □

Proposición 4.3. *Sea $f : V \to W$ una aplicación lineal entre k-espacios vectoriales. Supongamos que $\dim_k(V) = \dim_k(W) = n$. Entonces las condiciones siguientes son equivalentes:*

1. *f es biyectiva.*
2. *f es inyectiva.*
3. *f es sobreyectiva.*

Demostración. Claramente (1) implica (2) y (3). Recíprocamente, si f es inyectiva, es decir, si $\ker(f) = \{0\}$ (resp. sobreyectiva, es decir, $\mathrm{Im}(f) = W$), entonces el teorema del rango implica que f es también sobreyectiva (resp. inyectiva), y luego biyectiva. □

5. Aplicaciones lineales y matrices

Definición 5.1 (espacio de aplicaciones lineales)**.** Sean V, W dos k-espacios vectoriales. Denotaremos por $\mathrm{Hom}_k(V,W)$ o bien $\mathscr{L}(V,W)$ al conjunto de las aplicaciones lineales de V en W. Si $\varphi, \psi \in \mathrm{Hom}_k(V,W)$ son dos aplicaciones lineales y si $\lambda \in k$, definimos las aplicaciones $\varphi + \psi$ y $\lambda \cdot \varphi$ de la manera siguiente: para todo $v \in V$,

$$(\varphi + \psi)(v) := \varphi(v) + \psi(v), \quad (\lambda \cdot \varphi)(v) := \lambda \cdot \varphi(v). \tag{$*$}$$

Ellas también son aplicaciones *lineales* $V \to W$. En efecto, si $v_1, v_2 \in V$ y $\mu \in k$, entonces

$$\begin{aligned}(\varphi+\psi)(\mu \cdot v_1 + v_2) &= \varphi(\mu \cdot v_1 + v_2) + \psi(\mu \cdot v_1 + v_2) && \text{(por definición)}\\ &= \mu \cdot \varphi(v_1) + \varphi(v_2) + \mu \cdot \psi(v_1) + \psi(v_2) && \text{(pues } \varphi, \psi \text{ lineales)}\\ &= \mu \cdot (\varphi+\psi)(v_1) + (\varphi+\psi)(v_2) && \text{(por definición)}\end{aligned}$$

y también

$$(\lambda \cdot \varphi)(\mu \cdot v_1 + v_2) = \lambda\varphi(\mu \cdot v_1 + v_2) = \lambda\mu \cdot \varphi(v_1) + \lambda \cdot \varphi(v_2) = \mu \cdot (\lambda \cdot \varphi)(v_1) + (\lambda \cdot \varphi)(v_2).$$

Así, $(*)$ dota al conjunto $\mathrm{Hom}_k(V,W) = \mathscr{L}(V,W)$ de una estructura de k-espacio vectorial. Decimos que es el **espacio de aplicaciones lineales** de V en W.

Supongamos que V es de *dimensión finita* n y sea $(e_1, \ldots, e_n)$ una base de V (por ejemplo, $V = k^n$ y $(e_1, \ldots, e_n)$ la base canónica). Sea $\varphi \in \mathrm{Hom}_k(V,W)$, y definamos $w_i := \varphi(e_i)$ para $i = 1, \ldots, n$. Entonces, todo vector $v \in V$ se escribe de manera única como $v = x_1e_1 + \ldots + x_ne_n$, de donde obtenemos

$$\varphi(v) = x_1w_1 + \ldots + x_nw_n \tag{$\star$}$$

y luego φ está determinada por una colección de n vectores $w_1, \ldots, w_n \in W$. Recíprocamente, por toda n-tupla $(w_1, \ldots, w_n) \in W^n$, la aplicación $\varphi : V \to W$ definida por la fórmula $(\star)$ es lineal. Así, hemos demostrado el siguiente resultado.

Proposición 5.2. *Si $(e_1, \ldots, e_n)$ es una base de V, darse una aplicación lineal $\varphi : V \to W$* «es la misma cosa» *que darse una n-tupla $(w_1, \ldots, w_n) \in W^n$.*

Supongamos además que W es de dimensión finita m y sea $(f_1, \ldots, f_m)$ una base de W. Entonces, cada $w_j = \varphi(e_j)$ se escribe de manera única como

$$w_j = a_{1j}f_1 + \ldots + a_{mj}f_m,$$

lo cual es representado por el **vector columna**:

$$w_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

y luego φ está determinada por la matriz siguiente:

$$\mathrm{Mat}_{(f_i),(e_j)}(\varphi) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

que expresa los vectores $\varphi(e_j)$ (las columnas) en función de $f_1, \ldots, f_m$.

Importante: Notar que la dimensión n del espacio de *partida* V es el número de *columnas*, y la dimensión m del espacio de *llegada* es el número de *filas*.

Recíprocamente, dada cualquier matriz A como más arriba, sus columnas definen de manera única n vectores $w_1, \ldots, w_n \in W$. Explícitamente,

$$w_j = a_{1j}f_1 + \ldots + a_{mj}f_j,$$

y dicha n-tupla $(w_1, \ldots, w_n) \in W^n$ define una aplicación lineal $\varphi : V \to W$ cuya matriz asociada es A. En conclusión, tenemos una *biyección*:

$$\mathrm{Hom}_k(V, W) \overset{\sim}{\leftrightarrow} M_{m\times n}(k).$$

Más aún, verificamos fácilmente (ejercicio) que si A (resp. B) es la matriz asociada a φ (resp. ψ), entonces $\lambda A + B$ es la matriz asociada a $\lambda\varphi + \psi$, para todo $\lambda \in k$. Así, la biyección anterior es en realidad un isomorfismo de espacios vectoriales: $\mathrm{Hom}_k(V, W) \cong M_{m\times n}(k)$.

Teorema 5.3 (aplicaciones lineales y matrices)**.** *Sea $\mathscr{B} = (e_1, \ldots, e_n)$ una base de V, y sea $\mathscr{C} = (f_1, \ldots, f_m)$ una base de W. Entonces*

1. *Una aplicación lineal $V \to W$* «es la misma cosa» *que una matriz de m filas y n columnas, es decir, la aplicación*
$$\mathrm{Hom}_k(V, W) \xrightarrow{\sim} M_{m\times n}(k), \quad \varphi \mapsto \mathrm{Mat}_{\mathscr{C},\mathscr{B}}(\varphi)$$
es un isomorfismo de espacios vectoriales.

2. *Esto último transforma la composición de aplicaciones lineales en el producto de matrices: si U es otro k-espacio vectorial de base $\mathscr{A} = (d_1, \ldots, d_p)$ y si $\psi \in \mathrm{Hom}_k(U, V)$, entonces*
$$\mathrm{Mat}_{\mathscr{C},\mathscr{A}}(\varphi \circ \psi) = \mathrm{Mat}_{\mathscr{C},\mathscr{B}}(\varphi) \cdot \mathrm{Mat}_{\mathscr{B},\mathscr{A}}(\psi).$$

Demostración. El punto (1) sigue de la discusión anterior, veamos (2) a continuación. Recordemos que si $A \in M_{m\times n}(k)$ y $B \in M_{n\times p}(k)$ están dadas por

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{pmatrix}$$

entonces la *matriz producto* $C = AB \in M_{m\times p}(k)$ está definida por

$$C = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1p} \\ c_{21} & c_{22} & \cdots & c_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mp} \end{pmatrix},$$

donde el coeficiente c_{ik} está dado por la fórmula

$$c_{ik} = a_{i1}b_{1k} + a_{i2}b_{2k} + \cdots + a_{in}b_{nk} = \sum_{j=1}^{n} a_{ij}b_{jk},$$

para $i = 1, \ldots, m$ y $k = 1, \ldots, p$. En nuestro contexto, sean

$$A = \mathrm{Mat}_{\mathscr{C},\mathscr{B}}(\varphi) = (a_{ij})_{\substack{1\le i\le m \\ 1\le j\le n}} \qquad B = \mathrm{Mat}_{\mathscr{B},\mathscr{A}}(\psi) = (b_{jk})_{\substack{1\le j\le n \\ 1\le k\le p}}$$

las matrices asociadas a φ y ψ. Entonces, para todo $k = 1, \ldots, p$ tenemos que

$$(\varphi \circ \psi)(d_k) = \varphi\left(\sum_{j=1}^{n} b_{jk}e_j\right) = \sum_{j=1}^{n} b_{jk}\varphi(e_j) = \sum_{j=1}^{n}\sum_{i=1}^{m} b_{jk}a_{ij}f_i = \sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{ij}b_{jk}\right) f_i.$$

Así, el coeficiente de índice (i, k) de $M = \mathrm{Mat}_{\mathscr{C},\mathscr{A}}(\varphi \circ \psi)$ es $\sum_{j=1}^{n} a_{ij}b_{jk}$, y luego $M = AB$. □

Observación 5.4. *Sea $A = (a_{ij}) \in M_{m\times n}(k)$ y sean $(e_1, \ldots, e_n)$ y $(f_1, \ldots, f_m)$ las bases canónicas de k^n y k^m, respectivamente. Entonces gracias al isomorfismo precedente, A corresponde a la aplicación lineal $u : k^n \to k^m$ tal que, para todo $j = 1, \ldots, n$,*

$$u(e_j) = \sum_{i=1}^{m} a_{ij}f_i.$$

En lo que sigue, identificaremos cada vez que sea útil a la matriz A con la aplicación lineal $u : k^n \to k^m$ así definida. En otras palabras, la i-ésima columna de A es la imagen del i-ésimo vector de la base canónica de k^n.

Corolario 5.5. *Sean $A \in M_{m\times n}(k)$ y $B \in M_{n\times p}(k)$, y sean $u : k^n \to k^m$ y $v : k^p \to k^n$ las aplicaciones lineales asociadas. Entonces AB es la matriz asociada a $(u \circ v) : k^p \to k^m$.*

Importante: $A \in M_{m\times n}(k)$ y $B \in M_{n\times p}(k)$. Si denotamos por $B_1, \ldots, B_p \in k^n$ las columnas de la matriz B, entonces las columnas de AB son los vectores $AB_1, \ldots, AB_p$. En efecto, si $(e_1, \ldots, e_p)$ es la base canónica de k^p, entonces la matriz B corresponde a la aplicación lineal que envía cada e_j en el vector $Be_j = B_j \in k^n$, y AB corresponde a la aplicación lineal que envía cada e_j en el vector $A(Be_j) = AB_j \in k^m$.

¡Atención! Sólo podemos efectuar el producto AB de dos matrices $A \in M_{m\times n}(k)$ y $B \in M_{\ell\times p}(k)$ cuando $n = \ell$, es decir, cuando el número de columnas de A es igual al número de filas de B.

Caso particular (endomorfismos): Dada la importancia del Teorema anterior, repetiremos y reformularemos dicho resultado en el caso particular en que *el espacio de partida es el mismo que el espacio de llegada*, es decir, en el caso en que consideramos *endomorfismos* de un espacio V de dimensión finita n, o matrices *cuadradas* de tamaño n.

Teorema 5.6. *El k-espacio vectorial $M_n(k)$ de matrices cuadradas de tamaño n es un anillo*[6]*. Más aún, $M_n(k)$ es una k-álgebra*[7]*. Del mismo modo, si V es un k-espacio vectorial de dimensión n, el espacio de endomorfismos $\mathrm{End}_k(V)$ es una k-álgebra (donde el producto está dado por la composición de endomorfismos). Adicionalmente, si escogemos una base $\mathscr{B} = (e_1, \ldots, e_n)$ de V, la aplicación*

$$\mathrm{End}_k(V) \to M_n(k), \quad u \mapsto \mathrm{Mat}_{\mathscr{B}}(u)$$

es un isomorfismo de anillos y de k-espacios vectoriales, es decir, un isomorfismo de k-álgebras.

Definición 5.7 (núcleo, imagen y rango de una matriz)**.** Sea $A \in M_{m\times n}(k)$. Definimos su núcleo $\ker(A)$, su imagen $\mathrm{Im}(A)$ y su rango $\mathrm{rg}(A)$ como el núcleo, imagen y rango de la aplicación lineal $u : k^n \to k^m$ asociada.

Observación 5.8. *Sea $A \in M_{m\times n}(k)$ y sea $u : k^n \to k^m$ la aplicación lineal asociada. Entonces $\mathrm{rg}(A) = \mathrm{rg}(u) \leq n$ (gracias al teorema del rango), y $\mathrm{rg}(A) = \mathrm{rg}(u) \leq m$ (pues $\mathrm{Im}(u)$ es un sub-espacio de k^m). Luego,*

$$\mathrm{rg}(A) = \mathrm{rg}(u) \leq \mathrm{mín}(m, n).$$

Alternativamente, la imagen de u es el sub-espacio de k^m generado por los vectores columna *$C_1, \ldots, C_n$ de A, luego por definición $\mathrm{rg}(A)$ es el número máximo de* columnas *de A linealmente independientes, y luego $\mathrm{rg}(A) \leq \mathrm{mín}(m, n)$. Veremos más adelante que $\mathrm{rg}(A)$ es también el número máximo de* filas *linealmente independientes.*

Definición 5.9 (matriz transpuesta)**.** Sea $A \in M_{m\times n}(k)$ dada por

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix},$$

definimos la **matriz transpuesta** de A como la matriz ${}^{\mathrm{t}}A \in M_{n\times m}(k)$ definida por

$${}^{\mathrm{t}}A = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix},$$

[6]El anillo $M_n(k)$ es **no conmutativo** si $n \geq 2$ pues $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, pero $\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$.

[7]Una **k-álgebra** es un k-espacio vectorial A que ademas es un anillo (no necesariamente conmutativo), verificando la condición de compatibilidad siguiente: para todos $a, b \in A$ y todo $\lambda \in k$, se tiene que $(\lambda \cdot a) \cdot b = \lambda \cdot (ab) = a(\lambda \cdot b)$.

es decir, la j-ésima columna de A corresponde a la j-ésima fila de ${}^{\mathrm{t}}A$. Equivalentemente, $({}^{\mathrm{t}}A)_{ij} = a_{ji}$. En particular, ${}^{\mathrm{t}}({}^{\mathrm{t}}A) = A$.

Notación: La notación A^{t} es comunmente utilizada también. Sin embargo, veremos más adelante que $({}^{\mathrm{t}}A)^{-1} = {}^{\mathrm{t}}(A^{-1})$, por lo que la notación ${}^{\mathrm{t}}A^{-1}$ está bien definida, y es más cómoda que $(A^{\mathrm{t}})^{-1}$.

Proposición 5.10. *La aplicación*

$$M_{m\times n}(k) \to M_{n\times m}(k), \quad A \mapsto {}^{\mathrm{t}}A$$

es lineal. Más aún, si $B \in M_{n\times p}(k)$, entonces $AB \in M_{m\times p}(k)$ y tenemos en $M_{p\times m}(k)$ la siguiente identidad

$${}^{\mathrm{t}}(AB) = {}^{\mathrm{t}}B\,{}^{\mathrm{t}}A.$$

Demostración. Sean $A = (a_{ij})$ y $A' = (a'_{ij})$ en $M_{m\times n}(k)$, y sea $\lambda \in k$. Entonces, $\lambda A + A'$ está dada por $(\lambda a_{ij} + a'_{ij})$ en $M_{m\times n}(k)$, y su transpuesta está dada por la matriz C tal que, para todo (i,j),

$$C_{ji} = (\lambda A + A')_{ij} = \lambda a_{ij} + a'_{ij} = \lambda({}^{\mathrm{t}}A)_{ji} + ({}^{\mathrm{t}}A')_{ji},$$

de donde se concluye la linealidad. Finalmente, si $B = (b_{jk}) \in M_{n\times p}(k)$, entonces para todo (i,k) se tiene que

$$({}^{\mathrm{t}}(AB))_{ki} = (AB)_{ik} = \sum_{\ell=1}^{n} a_{i\ell}b_{\ell k} = \sum_{\ell=1}^{n} ({}^{\mathrm{t}}B)_{k\ell}\cdot({}^{\mathrm{t}}A)_{\ell i} = ({}^{\mathrm{t}}B\,{}^{\mathrm{t}}A)_{ki},$$

lo que prueba que ${}^{\mathrm{t}}(AB) = {}^{\mathrm{t}}B\,{}^{\mathrm{t}}A$. □

6. Cambios de base

Definición 6.1 (automorfismos y matrices invertibles)**.** Sea V un k-espacio vectorial. Decimos que un endomorfismo $f : V \to V$ es un **automorfismo** si posee una inversa en $\mathrm{End}_k(V)$, es decir, si existe un endomorfismo $f^{-1} : V \to V$ tal que $f \circ f^{-1} = f^{-1} \circ f = \mathrm{id}_V$. Esto equivale a decir que $f \in \mathrm{End}_k(V)$ es una aplicación biyectiva, puesto que ya hemos observado que en tal caso la inversa f^{-1} es necesariamente lineal.

De manera similar, si $A \in M_n(k)$ es una matriz cuadrada de tamaño n, decimos que A es **invertible** si existe $B \in M_n(k)$ tal que $AB = BA = \mathrm{I}_n$, donde I_n denota la matriz identidad de tamaño n. En este caso B es denotada A^{-1}, la matriz inversa de A.

Notación: Sea V un k-espacio vectorial. Denotamos por

$$\mathrm{GL}(V) = \{f : V \to V \text{ automorfismo}\}$$

al conjunto de todos los automorfismos de V. Dicho conjunto es un *grupo* respecto a la composición de endomorfismos. En efecto, la composición de endomorfismos es asociativa, la aplicación identidad id_V es el elemento neutro, y si f, g son automorfismos entonces $f \circ g$ lo es también, y su inversa está dada por $g^{-1} \circ f^{-1}$ (puesto que $f \circ g \circ g^{-1} \circ f^{-1} = \mathrm{id}_V = g^{-1} \circ f^{-1} \circ f \circ g$).[8]

El grupo $\mathrm{GL}(V)$ es llamado el **grupo general lineal** del espacio vectorial V. De manera completamente análoga, definimos

$$\mathrm{GL}_n(k) = \{A \in M_n(k) \text{ invertible}\},$$

el grupo general lineal de matrices invertibles de tamaño n. Dado que la correspondencia biyectiva $\mathrm{End}_k(k^n) \to M_n(k)$ transforma la composición de endomorfismos en el producto de matrices, tenemos que una matriz A es invertible si y sólo si el endomorfismo $u : k^n \to k^n$ asociado es biyectivo, en cuyo caso A^{-1} es la matriz de u^{-1}. En general, la elección de una base $\mathscr{B} = (e_1, \ldots, e_n)$ de V induce un isomorfismo de grupos $\mathrm{GL}(V) \cong \mathrm{GL}_n(k)$.

Proposición 6.2. *Sea k un cuerpo.*

[8] Notar la inversión en el orden de los factores: $(f \circ g)^{-1}$) es igual a $g^{-1} \circ f$, mientras que $(g \circ f)^{-1}$ es igual a $f^{-1} \circ g^{-1}$.

(i) Sea V un k-espacio vectorial de dimensión finita *y sean $u, v \in \mathrm{End}_k(V)$ tales que $u \circ v = \mathrm{id}_V$. Entonces $u, v \in \mathrm{GL}(V)$ y $u = v^{-1}$.*

(i') Sean $A, B \in M_n(k)$ tales que $AB = \mathrm{I}_n$. Entonces tenemos que $BA = \mathrm{I}_n$ y luego $A, B \in \mathrm{GL}_n(k)$ y son inversas una de la otra.

(ii) Si A es invertible entonces ${}^{\mathrm{t}}A$ es invertible, y tenemos $({}^{\mathrm{t}}A)^{-1} = {}^{\mathrm{t}}(A^{-1})$.

Demostración. Para ver (i) notar que para todo vector $x \in V$ se tiene que

$$x = u(v(x)),$$

por lo que u es sobreyectiva, y v es inyectiva. Luego, tanto u como v son biyectivas. Luego, multiplicando la igualdad $u \circ v = \mathrm{id}_V$ a la izquierda por u^{-1} obtenemos $v = u^{-1}$.

Para probar (i') denotemos por u (resp. v) al endomorfismo $k^n \to k^n$ asociado a A (resp. B). Como $AB = \mathrm{I}_n$ equivale a $u \circ v = \mathrm{id}_{k^n}$ tenemos, gracias a (i), que u y v son biyectivas e inversas una de la otra. Así, esto último también es cierto para A y B, de donde concluimos que $B = A^{-1}$ y luego $BA = \mathrm{I}_n$.

Finalmente, supongamos que A es invertible y sea $B \in M_n(k)$ tal que $AB = BA = \mathrm{I}_n$. Tomando la transpuesta de dichas matrices, y utilizando el hecho que ${}^{\mathrm{t}}\mathrm{I}_n = \mathrm{I}_n$, deducimos que

$${}^{\mathrm{t}}B\,{}^{\mathrm{t}}A = {}^{\mathrm{t}}(AB) = \mathrm{I}_n = {}^{\mathrm{t}}(BA) = {}^{\mathrm{t}}A\,{}^{\mathrm{t}}B.$$

Esto último demuestra que ${}^{\mathrm{t}}A$ es invertible, y que su inversa está dada por ${}^{\mathrm{t}}B = {}^{\mathrm{t}}(A^{-1})$. □

¡Atención! Sea $V = \mathbb{R}[X]$ y consideremos el operador $I : \mathbb{R}[X] \to \mathbb{R}[X]$ de *integración* que envía cada monomio X^n en $X^{n+1}/(n+1)$, y el operador de *derivación* $D : \mathbb{R}[X] \to \mathbb{R}[X]$ que envía cada polinomio P sobre su derivada P'. Entonces, $D \circ I = \mathrm{id}_V$, por lo que D es sobreyectivo y I es inyectivo. Sin embargo, D *no es* inyectivo puesto que $D(1) = 0$, y I *no es* sobreyectivo puesto que su imagen está formada por polinomios con término constante nulo. En otras palabras, si V es un k-espacio vectorial de dimensión *infinita* y $u, v \in \mathrm{End}_k(V)$ verifican $u \circ v = \mathrm{id}_V$, entonces u y v no son necesariamente biyectivos.

Lema 6.3. *Sea V un k-espacio vectorial de dimensión finita n, $\mathscr{B} = (v_1, \ldots, v_n)$ una base de V, y $f : V \to V$ un endomorfismo. Si denotamos $w_i = f(v_i)$ y si $(w_1, \ldots, w_n)$ es una base de V, entonces f es biyectivo. Más aún, su inversa $g : V \to V$ es el endomorfismo de V definido por $g(w_i) = v_i$ para $i = 1, \ldots, n$.*

Demostración. Supogamos que $\mathscr{C} = (w_1, \ldots, w_n)$ es una base de V. Entonces f es sobreyectivo, puesto que para todo $w \in V$ existen $\lambda_1, \ldots, \lambda_n \in k$ tales que

$$w = \lambda_1 w_1 + \ldots + \lambda_n w_n = f(\lambda_1 v_1 + \ldots + \lambda_n v_n).$$

Así, f es biyectivo dado que V es de dimensión finita. Finalmente, sea $g : V \to V$ el endomorfismo de V definido por $g(w_i) = v_i$ para $i = 1, \ldots, n$. Entonces, por un lado tenemos que $(g \circ f)(v_i) = v_i$ para todo i, de donde se tiene $g \circ f = \mathrm{id}_V$, y por otro lado tenemos que $(f \circ g)(w_i) = w_i$ para todo i, de donde se concluye que $f \circ g = \mathrm{id}_V$. □

Definición 6.4 (matriz de cambio de base)**.** Sea V un k-espacio vectorial de dimensión n, y sean $\mathscr{B} = (e_1, \ldots, e_n)$ y $\mathscr{B}' = (v_1, \ldots, v_n)$ dos bases de V. Sea P la matriz $n \times n$ expresando la base $\mathscr{B}'$ en términos de la base $\mathscr{B}$, es decir, cada v_j se escribe de manera única como

$$v_j = p_{1j}e_1 + \ldots + p_{nj}e_n$$

y consideramos la matriz

$$P = \begin{pmatrix} p_{11} & \cdots & p_{1j} & \cdots & p_{1n} \\ \vdots & & \vdots & & \vdots \\ p_{i1} & \cdots & p_{ij} & \cdots & p_{in} \\ \vdots & & \vdots & & \vdots \\ p_{n1} & \cdots & p_{nj} & \cdots & p_{nn} \end{pmatrix}$$

donde las columnas son los vectores $v_1, \ldots, v_n$ expresados en la base $(e_1, \ldots, e_n)$. Entonces P se llama la **matriz de cambio de base** de la base $\mathscr{B}$ a la base $\mathscr{B}'$, y se denota $\mathrm{Mat}_{\mathscr{B}}(\mathscr{B}')$. La matriz P es invertible, y su inversa $P^{-1} = \mathrm{Mat}_{\mathscr{B}'}(\mathscr{B})$ está dada por la matriz que expresa $\mathscr{B} = (e_1, \ldots, e_n)$ en la base $\mathscr{B}' = (v_1, \ldots, v_n)$.

Ejercicio 6.5. Verificar que P puede ser vista como la matriz asociada a la aplicación identidad id_V, expresada en las base $\mathscr{B}' = (v_1, \ldots, v_n)$ en la partida y $\mathscr{B} = (e_1, \ldots, e_n)$ en la llegada. En otras palabras, $P = \mathrm{Mat}_{\mathscr{B},\mathscr{B}'}(\mathrm{id}_V)$. Del mismo modo, verificar que $P^{-1} = \mathrm{Mat}_{\mathscr{B}',\mathscr{B}}(\mathrm{id}_V)$.

Utilizando la notación precedente, todo vector $v \in V$ se escribe de manera única como

$$v = x_1 e_1 + \ldots + x_n e_n \qquad \text{y} \qquad v = x'_1 v_1 + \ldots + x'_n v_n,$$

y las x_i (resp. x'_i) se llaman las **coordenadas** de v respecto a la base $\mathscr{B}$ (resp. $\mathscr{B}'$). Así, relativamente a la base $\mathscr{B}$ (resp. $\mathscr{B}'$), podemos representar $v \in V$ por el vector columna

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \qquad \text{resp.} \qquad X' = \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$$

Proposición 6.6 (cambio de coordenadas)**.** *La fórmla de cambio de coordenadas, para el cambio de base $\mathscr{B} \to \mathscr{B}'$ dado por la matriz de cambio de base P, está dada por*

$$X = PX'.$$

Es decir, esta fórmula expresa las antiguas *coordenadas X en función de las* nuevas *coordenadas X'.*

Demostración. En efecto, escribiendo $v_j = \sum_{i=1}^n p_{ij} e_i$ y observando que

$$v = \sum_{j=1}^n x'_j v_j = \sum_{j=1}^n \sum_{i=1}^n x'_j p_{ij} e_i = \sum_{i=1}^n \left(\sum_{j=1}^n p_{ij} x'_j \right) e_i,$$

obtenemos al comparar con $v = \sum_{i=1}^n x_i e_i$ que necesariamente $x_i = \sum_{j=1}^n p_{ij} x'_j$ para $i = 1, \ldots, n$. Así, se concluye que $X = PX'$. □

Teorema 6.7 (cambio de base para aplicaciones lineales)**.** *Sea $A = \mathrm{Mat}_{\mathscr{C},\mathscr{B}}(u) \in M_{m\times n}(k)$ la matriz asociada a una aplicación lineal $u : V \to W$, respecto a las bases $\mathscr{B} = (e_1, \ldots, e_n)$ de V y $\mathscr{C} = (f_1, \ldots, f_m)$ de W. Sea $\mathscr{B}' = (v_1, \ldots, v_n)$ (resp. $\mathscr{C}' = (w_1, \ldots, w_n)$) una segunda base de V (resp. de W), y sea $P \in M_n(k)$ (resp. $Q \in M_m(k)$) la matriz de cambio de base correspondiente. Entonces, la matriz de u en las bases $\mathscr{B}'$ y $\mathscr{C}'$ está dada por*

$$\mathrm{Mat}_{\mathscr{C}',\mathscr{B}'}(u) = Q^{-1}AP.$$

Demostración. Gracias a la correspondencia entre aplicaciones lineales y matrices,

$$\mathrm{Mat}_{\mathscr{C},\mathscr{B}'}(u) = \mathrm{Mat}_{\mathscr{C},\mathscr{B}'}(u \circ \mathrm{id}_V) = \mathrm{Mat}_{\mathscr{C},\mathscr{B}}(u) \cdot \mathrm{Mat}_{\mathscr{B},\mathscr{B}'}(\mathrm{id}_V),$$

de donde obtenemos que la matriz de u en las bases $\mathscr{B}'$ y $\mathscr{C}$ es AP. De manera similar, si consideramos la segunda base $\mathscr{C}' = (w_1, \ldots, w_n)$ de W, con matriz de cambio de base Q de $\mathscr{C}$ a $\mathscr{C}'$, entonces $Q = \mathrm{Mat}_{\mathscr{C},\mathscr{C}'}(\mathrm{id}_W)$ y $Q^{-1} = \mathrm{Mat}_{\mathscr{C}',\mathscr{C}}(\mathrm{id}_W)$. Luego, obtenemos de manera análoga

$$\mathrm{Mat}_{\mathscr{C}',\mathscr{B}'}(u) = \mathrm{Mat}_{\mathscr{C}',\mathscr{C}}(\mathrm{id}_W) \cdot \mathrm{Mat}_{\mathscr{C},\mathscr{B}'}(u),$$

de donde se concluye que $\mathrm{Mat}_{\mathscr{C}',\mathscr{B}'}(u) = Q^{-1}AP$. □

Este último resultado puede ser fácilmente visualizado gracias al siguiente **diagrama conmutativo**:

$$\begin{array}{ccc} (V,\mathscr{B}) & \xrightarrow{A=\mathrm{Mat}_{\mathscr{C},\mathscr{B}}(u)} & (W,\mathscr{C}) \\ {\scriptstyle \mathrm{Mat}_{\mathscr{B}}(\mathscr{B}')=P}\uparrow & & \downarrow{\scriptstyle Q^{-1}=\mathrm{Mat}_{\mathscr{C}'}(\mathscr{C})} \\ (V,\mathscr{B}') & \xrightarrow[Q^{-1}AP=\mathrm{Mat}_{\mathscr{C}',\mathscr{B}'}(u)]{} & (W,\mathscr{C}') \end{array}$$

Observación 6.8. *Sea $u : V \to W$ una aplicación lineal. El teorema anterior es compatible con la fórmula de cambio de coordenadas: si denotamos por X (resp. X') las coordenadas de un vector $v \in V$ respecto a la base $\mathscr{B}$ (resp. a la base $\mathscr{B}'$), y por Y (resp. Y') las coordenadas del vector $u(v) \in W$ respecto a la base $\mathscr{C}$ (resp. $\mathscr{C}'$), entonces tenemos que*

$$Y = AX, \qquad X = PX', \qquad Y = QY',$$

de donde se obtiene $Y' = Q^{-1}Y = Q^{-1}AX = Q^{-1}APX'$.

Cabe destacar que incluso si nos interesamos en una matriz $A \in M_{m\times n}(k)$, es frecuentemente útil de considerar a A como una aplicación lineal $u : k^n \to k^m$ (definida por $u(e_j) = a_{1j}f_1 + \ldots + a_{mj}f_m$, donde $(e_1, \ldots, e_n)$, resp. $(f_1, \ldots, f_m)$, es la base canónica de k^n, resp. k^m). Por ejemplo, el teorema anterior tiene el corolario siguiente caracterizando el rango de una matriz.

Corolario 6.9. *Sea $A \in M_{m\times n}(k)$ y sea $r = \mathrm{rg}(A)$. Entonces,*

1. *Existen matrices invertible $P \in \mathrm{GL}_n(k)$ y $Q \in \mathrm{GL}_m(k)$ tales que*

$$Q^{-1}AP = \begin{pmatrix} \mathrm{I}_r & \mathbf{0}_{r,n-r} \\ \mathbf{0}_{m-r,r} & \mathbf{0}_{m-r,n-r} \end{pmatrix},$$

   donde I_r es la matriz identidad de tamaño r y donde $\mathbf{0}_{p,q}$ denota la matriz nula de p filas y q columnas.

2. *Recíprocamente, si existen matrices invertibles $P \in \mathrm{GL}_n(k)$ y $Q \in \mathrm{GL}_m(k)$ y un entero $s \in \mathbb{N}$ tales que*

$$Q^{-1}AP = \begin{pmatrix} \mathrm{I}_s & \mathbf{0}_{s,n-s} \\ \mathbf{0}_{m-s,s} & \mathbf{0}_{m-s,n-s} \end{pmatrix},$$

   entonces $s = \mathrm{rg}(A)$.

Demostración. Sean $(e_1, \ldots, e_n)$ y $(f_1, \ldots, f_m)$ las bases canónicas de k^n y k^m, y sea $u : k^n \to k^m$ la aplicación lineal asociada a A. Por definición, $r = \mathrm{rg}(A)$ es la dimensión de $\mathrm{Im}(u)$. Sea entonces $(w_1, \ldots, w_r)$ una base de $\mathrm{Im}(u)$, la cual puede ser completada en una base $\mathscr{C} = (w_1, \ldots, w_m)$ de k^m. Denotemos por $Q \in \mathrm{GL}_m(k)$ la matriz de cambio de base de $(f_1, \ldots, f_m)$ a $\mathscr{C}$.
Sean $v_1, \ldots, v_r$ vectores de k^n tales que $u(v_j) = w_j$ para todo $j = 1, \ldots, r$, y sea $(\varepsilon_1, \ldots, \varepsilon_d)$ una base de $\ker(u)$. Gracias a la Demostración alternativa del Teorema del rango, tenemos que $\mathscr{B} = (v_1, \ldots, v_r, \varepsilon_1, \ldots, \varepsilon_d)$ es una base de k^n. Luego, la matriz de u en las bases $\mathscr{B}$ y $\mathscr{C}$ está dada por

$$\mathrm{Mat}_{\mathscr{C},\mathscr{B}}(u) = \begin{pmatrix} \mathrm{I}_r & \mathbf{0}_{r,n-r} \\ \mathbf{0}_{m-r,r} & \mathbf{0}_{m-r,n-r} \end{pmatrix}.$$

Por otra parte, si P denota la matriz de cambio de base de $(e_1, \ldots, e_n)$ a $\mathscr{B}$, entonces

$$\mathrm{Mat}_{\mathscr{C},\mathscr{B}}(u) = Q^{-1} \cdot \mathrm{Mat}_{(f_i),(e_j)}(u) \cdot P = Q^{-1}AP,$$

de donde se deduce (1).
Recíprocamente, si suponemos que existen matrices invertibles $P \in \mathrm{GL}_n(k)$ y $Q \in \mathrm{GL}_m(k)$ y un entero $s \in \mathbb{N}$ tales que

$$Q^{-1}AP = \begin{pmatrix} \mathrm{I}_s & \mathbf{0}_{s,n-s} \\ \mathbf{0}_{m-s,s} & \mathbf{0}_{m-s,n-s} \end{pmatrix},$$

entonces existen bases $(v_1, \ldots, v_n)$ de k^n y $(w_1, \ldots, w_m)$ de k^m tales que $u(v_i) = w_i$ para todo $i = 1, \ldots, s$ y $u(v_j) = 0$ para todo $j = s+1, \ldots, n$. Así, $\mathrm{Im}(u) = \mathrm{Vect}_k(w_1, \ldots, w_s)$ es de dimensión s, de donde obtenemos $s = \mathrm{rg}(A)$. □

Como consecuencia de la caracterización anterior del rango de una matriz obtenemos la proposición siguiente.

Proposición 6.10. *Sea $A \in M_{m\times n}(k)$. Entonces,*

$$\mathrm{rg}(A) = \mathrm{rg}({}^{\mathrm{t}}A).$$

En particular, el rango de A es el número máximo de filas *de A que son linealmente independientes.*

Demostración. Gracias al corolario anterior, existen $P \in \mathrm{GL}_n(k)$ y $Q \in \mathrm{GL}_m(k)$ tales que

$$Q^{-1}AP = \begin{pmatrix} \mathrm{I}_r & \mathbf{0}_{r,n-r} \\ \mathbf{0}_{m-r,r} & \mathbf{0}_{m-r,n-r} \end{pmatrix},$$

donde $r = \mathrm{rg}(A)$. Entonces,

$${}^{\mathrm{t}}P\,{}^{\mathrm{t}}A\,{}^{\mathrm{t}}Q^{-1} = \begin{pmatrix} \mathrm{I}_r & \mathbf{0}_{r,m-r} \\ \mathbf{0}_{n-r,r} & \mathbf{0}_{n-r,m-r} \end{pmatrix}.$$

Luego, dado que ${}^{\mathrm{t}}P \in \mathrm{GL}_n(k)$ y ${}^{\mathrm{t}}Q^{-1} \in \mathrm{GL}_m(k)$, el corolario anterior implica que $r = \mathrm{rg}({}^{\mathrm{t}}A)$. □

La discusión anterior motiva la definición siguiente.

Definición 6.11 (matrices equivalentes)**.** Sean $A, B \in M_{m\times n}(k)$. Decimos que A y B son **equivalentes** si existen matrices invertibles $P \in \mathrm{GL}_n(k)$ y $Q \in \mathrm{GL}_m(k)$ tales que $Q^{-1}AP = B$. En otras palabras, dos matrices A y B son equivalentes si y sólo si tienen el mismo rango.

Ejercicio 6.12. Sean $m, n \in \mathbb{N}^{\geq 1}$ y sea k un cuerpo. Demostrar que la relación en $M_{m\times n}(k)$ dada por

$$A \sim B \Leftrightarrow \exists P \in \mathrm{GL}_n(k), Q \in \mathrm{GL}_m(k) \text{ tal que } Q^{-1}AP = B$$

es una relación de equivalencia. Describir el conjunto cociente y determinar su cardinal.

Caso particular (endomorfismos): El teorema de cambio de base de aplicaciones lineales trata el caso general de una aplicación lineal $u : V \to W$, donde V y W son *a priori* distintos. En tal caso, si autorizamos cambios de base arbitrarios tanto en V como en W, hemos observado que el único invariante de u es su rango, que es un entero $r \in \{0, \ldots, \mathrm{mín}(\dim_k(V), \dim_k(W))\}$.

Sin embargo, si $V = W$ y si nos interesamos a la *naturaleza geométrica* de un endomorfismo $u : V \to V$ (es decir, cuando queremos comparar $u(x)$ y x, donde x varía en V), entonces para poder hacer la comparación nos gustaría expresar tanto x como $u(x)$ en *la misma base*. Por esta razón, en el caso donde $V = W$, escribimos la matriz de u en la *misma base* $\mathscr{B}$ de V tanto en la partida como en la llegada.

Por ejemplo, si $V = W = k$ es un k-espacio vectorial de dimensión 1, los automorfismos de k como k-espacio vectorial son las **homotecias**[9] $h_\lambda : k \to k,\ x \mapsto \lambda x$ con $\lambda \neq 0$. Si consideramos $\mathscr{B} = \{1\}$ como la base de partida y $\mathscr{B}' = \{\lambda\}$ como la base de llegada, entonces la matriz de h_λ es $(1) \in \mathrm{GL}_1(k) \cong k^\times$, es decir, hemos «perdido» el factor λ de la homotecia (que describe su geometría). Sin embargo, si matenemos la misma base $\mathscr{B}$ a la partida y a la llegada, entonces la matriz de h_λ es $(\lambda) \in \mathrm{GL}_1(k) \cong k^\times$.

El teorema de cambio de base de aplicaciones lineales se reduce al siguiente resultado.

Teorema 6.13 (cambio de base para endomorfismos)**.** *Sea A la matriz de un endomorfismos $u : V \to V$ respecto a una base $\mathscr{B}$ de V. Si $\mathscr{B}'$ es una segunda base de V, y si P es la matriz de cambio de base de $\mathscr{B}$ a $\mathscr{B}'$, entonces la matriz de u en la base $\mathscr{B}'$ está dada por*

$$\mathrm{Mat}_{\mathscr{B}'}(u) = P^{-1}AP.$$

El resultado anterior motiva la siguiente definición.

Definición 6.14 (matrices semejantes)**.** Sean $A, B \in M_n(k)$ matrices *cuadradas* de tamaño n. Decimos que A y B son **semejantes** si existe una matriz invertible $P \in \mathrm{GL}_n(k)$ tal que $P^{-1}AP = B$. En este caso, diremos que A y B están en la misma **clase de semejanza**.

Importante: Se puede probar (cf. equivalencia de matrices) que la relación en $M_n(k)$ dada por

$$A \sim B \Leftrightarrow \exists P \in \mathrm{GL}_n(k) \text{ tal que } P^{-1}AP = B$$

es una relación de equivalencia. Así, las clases de semejanza corresponden a las clases de equivalencia respecto a esta relación. Cabe destacar que:

[9] En general, si V es un k-espacio vectorial, una **homotecia** es un endomorfismo de la forma $h_\lambda : V \to V,\ v \mapsto \lambda v$ donde $\lambda \in k$ es llamado el **factor** de la homotecia.

1. Dos matrices A y B son semejantes si y sólamente si ellas representan, en bases diferentes, el mismo endomorfismo $u : k^n \to k^n$ de k^n.

2. Si $A, B \in M_n(k)$ son semejantes, entonces ellas son equivalentes. Sin embargo, el recíproco está lejos de ser verdad, tal como se observa ya en el caso de $n = 1$ que acabamos de discutir. De hecho, las clases de semejanza forman una partición de $M_n(k)$ *mucho más fina* que aquella dada por el rango, tal como discutiremos más adelante en el curso.

7. Operaciones elementales sobre filas y columnas

7.1. Operaciones sobre columnas

El objetivo de esta sección es recordar los métodos algorítmicos para calcular el rango de una aplicación lineal $u : k^n \to k^m$.

Concretamente, sea $A \in M_{m\times n}(k)$ y denotemos por $(e_1, \ldots, e_n)$ la base canónica de k^n. Procederemos a calcular una base de $\mathrm{Im}(A)$ y de $\ker(A)$ de la manera siguiente.

Observemos que intercambiar las columnas C_i y C_j de A es equivalente a permutar, antes de aplicar A, los vectores e_i y e_j de la base canónica de k^n, lo que equivale a su vez a multiplicar A *por la derecha* por la **matriz de permutación** $P(i, j)$ dado por el automorfismo de k^n que intercambia e_i y e_j, y que deja fijo cada e_ℓ para $\ell \neq i, j$. Por ejemplo, para $n = 4$ y $(i, j) = (1, 4)$ se tiene que

$$P(1,4) = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

Del mismo modo, multiplicar una columna C_j por un escalar $\lambda_j \neq 0$ es equivalente a reemplazar el vector e_j por el vector $\lambda_j e_j$, es decir, a multiplicar A *por la derecha* por la matriz diagonal donde todos los términos diagonales valen 1, excepto por el j-ésimo que vale λ_j.

Por otro lado, para todo $\lambda \in k$ e $i \neq j$, adicionar λC_i a C_j equivale a reemplazar, antes de aplicar A, el vector e_j por el vector $e_j + \lambda e_i$, lo que equivale a su vez a multiplicar A *por la derecha* por la matriz $B_{ij}(\lambda)$ asociada al automorfismo de k^n que deja fijo e_ℓ para $\ell \neq j$ y que envía e_j en $e_j + \lambda e_i$. En otras palabras,

$$B_{ij}(\lambda) = \mathrm{I}_n + \lambda E_{ij}.$$

Definición 7.1 (operaciones elementales sobre columnas)**.** Llamaremos **operaciones elementales sobre las columnas** a las operaciones precedentes: intercambio de columnas, multiplicación de una columna por un escalar no-nulo, o adicionar λC_j a C_i con $j \neq i$. Gracias a la discusión anterior, observamos que efectuar estas operaciones sobre las columnas equivale a aplicar *automorfismos* sobre el espacio de *partida* k^n, por lo que esto no cambia la imagen de A ni su rango.

Luego, podemos calcular $\mathrm{Im}(A)$ de la manera siguiente:

(1°) Sea i_1 el índice de la primera *fila* no-nula. Entonces, permutando columnas, podemos suponer que $a_{i_1,1} \neq 0$ y luego, multiplicando por $a_{i_1,1}^{-1}$ nos reducimos al caso $a_{i_1,1} = 1$.

(2°) A continuación, sustrayendo $a_{i_1,j}C_1$ de C_j nos reducimos al caso $a_{i_1,j} = 0$ para $j \geq 2$.

(3°) Sea i_2 el índice más pequeño tal que existe $j \geq 2$ con $a_{i_2,j} \neq 0$. Procediendo como en los pasos anteriores, nos reducimos al caso $a_{i_2,2} = 1$ y $a_{i_2,j} = 0$ para $j \geq 3$.

(4°) Sea i_3 el índice más pequeño tal que existe $j \geq 3$ con $a_{i_3,j} \neq 0$. Repitiendo el proceso anterior, obtenemos una matriz de la forma siguiente:

$$A' = \begin{pmatrix} 0 & 0 & 0 & 0 & \cdots & 0 & \cdots & 0 \\ 1 & 0 & 0 & 0 & \cdots & 0 & \cdots & 0 \\ a_{2,1} & 0 & 0 & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & 0 & 0 & 0 & \cdots & 0 & \cdots & 0 \\ a_{i_2,1} & 1 & 0 & 0 & \cdots & 0 & \cdots & 0 \\ * & * & 0 & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & * & 0 & 0 & \cdots & 0 & \cdots & 0 \\ a_{i_3,1} & a_{i_3,2} & 1 & 0 & \cdots & 0 & \cdots & 0 \\ * & * & * & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ a_{i_r,1} & a_{i_r,2} & a_{i_r,3} & \cdots & 1 & 0 & \cdots & 0 \\ * & * & * & \cdots & * & 0 & \cdots & 0 \\ * & * & * & \cdots & * & 0 & \cdots & 0 \end{pmatrix}$$

es decir, las columnas de índice $> r$ son nulas y, para $j = 1, \ldots, r$, las columnas C'_j satisfacen

$$C'_j = e_{i_j} + \sum_{\ell > i_j} a_{\ell,j} e_\ell,$$

con $i_1 < i_2 < \cdots < i_r$. Así, los vectores $C'_1, \ldots, C'_r \in k^m$ son linealmente independientes, y luego forman una base de $\mathrm{Im}(A)$. En particular, $r = \mathrm{rg}(A)$.

El proceso anterior se llama **reducción de columnas**, la matriz A' obtenida de esta forma verifica

$$A' = AP$$

donde P es una matriz invertible, correspondiendo a las operaciones elementares sobre las columnas que hemos efectuado. Hemos visto que $\mathrm{Im}(A') = \mathrm{Im}(A)$; por otro lado, el núcleo de A' es el sub-espacio V de k^n generado por los vectores $e_{r+1}, \ldots, e_n$ de la base canónica. Luego, la igualdad $A' = AP$ implica que

$$\ker(A) = P(V) = \mathrm{Vect}_k(P_{r+1}, \ldots, P_n),$$

donde $P_{r+1}, \ldots, P_n$ denotan las últimas $(n - r)$ columnas de P. En efecto, si $v \in V$ entonces $0 = A'v = APv$ por lo que $Pv \in \ker(A)$. Recíprocamente, como P es invertible, tenemos que $A = A'P^{-1}$ por lo que si $x \in \ker(A)$ entonces $0 = Ax = A'P^{-1}x$, de donde se obtiene que $P^{-1}x \in \ker(A') = V$ y luego $x \in P(V)$.

La igualdad $\ker(A) = \mathrm{Vect}_k(P_{r+1}, \ldots, P_n)$ permite determinar explícitamente una base de $\ker(A)$ de la manera siguiente. Las operaciones sobre las columnas que hemos hecho corresponden a multiplicar A a la derecha por ciertas matrices invertibles de un tipo particular, y P es el producto de dichas matrices. Luego, la misma serie de operaciones sobre las columnas de la matriz identidad I_n nos permite encontrar la matriz P.

En la práctica: escribimos A sobre la matriz identidad I_n, donde n es el número de columnas de A (es decir, la dimensión del espacio de partida) y, a cada etapa, efectuamos las mismas operaciones elementales sobre las columnas de las dos matrices. Obtenemos así al final del proceso tanto la matriz $A' = AP$ como la matriz $P = \mathrm{I}_n P$.[10]

Consideremos el siguiente ejemplo:

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 6 & 6 \\ 3 & 8 & 15 & 16 \end{pmatrix}.$$

Escribamos I_4 bajo la matriz A y, denotando por C_1, C_2, C_3 y C_4 las columnas de A, realicemos las siguientes operaciones sobre las columnas:

[10] Más aún, si $m = n$, es decir, cuando consideramos matrices cuadradas, podemos utilizar este proceso para determinar si A es invertible y calcular su inversa. Veremos esto más adelante.

$$\left(\begin{array}{cccc} 1 & 2 & 3 & 4 \\ 1 & 3 & 6 & 6 \\ 3 & 8 & 15 & 16 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right) \xrightarrow[C_4 \mapsto C_4 - 4C_1]{\substack{C_2 \mapsto C_2 - 2C_1 \\ C_3 \mapsto C_3 - 3C_1}} \left(\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 1 & 1 & 3 & 2 \\ 3 & 2 & 6 & 4 \\ \hline 1 & -2 & -3 & -4 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right) \xrightarrow[C_4 \mapsto C_4 - 2C_2]{C_3 \mapsto C_3 - 3C_2} \left(\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 3 & 2 & 0 & 0 \\ \hline 1 & -2 & 3 & 0 \\ 0 & 1 & -3 & -2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right)$$

Luego, $\mathrm{Im}(A)$ y $\ker(A)$ poseen las siguientes bases:

$$\mathrm{Im}(A) = \mathrm{Vect}_k \left\langle \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} \right\rangle \quad \text{y} \quad \ker(A) = \mathrm{Vect}_k \left\langle \begin{pmatrix} 3 \\ -3 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -2 \\ 0 \\ 1 \end{pmatrix} \right\rangle.$$

En resumen, la «reducción de columnas» de una matriz $A \in M_{m\times n}(k)$ nos otorga *bases* de $\mathrm{Im}(A)$ y $\ker(A)$. Sin embargo, dado un vector $Y = (y_1, \ldots, y_m) \in k^m$ (resp. $X = (x_1, \ldots, x_n) \in k^n$), no es necesariamente claro si $Y \in \mathrm{Im}(A)$ (resp. si $X \in \ker(A)$). Por lo anterior, es a veces más cómodo tener *ecuaciones* que definan $\mathrm{Im}(A)$ y $\ker(A)$. Veremos en seguida que dichas ecuaciones se obtienen gracias al proceso de *reducción de filas* de A.

Importante: En el ejemplo anterior, el primer *pivote* estaba en posición (1,1), y luego el segundo en posición (2,2). Evidentemente, este no es siempre el caso; en teoría, uno siempre se puede reducirse a dicho caso intercambiando columnas, pero con un poco de práctica no es necesario hacer cambios de columnas: basta reducirse al caso de una matriz A' donde las columnas sean escalonadas *módulo permutación de columnas*. En tal caso, las columnas no-nulas de A' forman una base de $\mathrm{Im}(A)$ y, en la matriz inferior (obtenida a partir de I_n), las columnas bajo las columnas nulas de A' forman una base de $\ker(A)$.

Por ejemplo, el cálculo siguiente

$$\left(\begin{array}{cccc} 3 & 2 & 4 & 1 \\ 6 & 3 & 6 & 1 \\ 15 & 8 & 16 & 3 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right) \xrightarrow[C_1 \mapsto C_1 - 3C_4]{\substack{C_3 \mapsto C_3 - 4C_4 \\ C_2 \mapsto C_2 - 2C_4}} \left(\begin{array}{cccc} 0 & 0 & 0 & 1 \\ 3 & 1 & 2 & 1 \\ 6 & 2 & 4 & 3 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -3 & -2 & -4 & 1 \end{array}\right) \xrightarrow[C_3 \mapsto C_3 - 2C_2]{C_1 \mapsto C_1 - 3C_2} \left(\begin{array}{cccc} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 2 & 0 & 3 \\ \hline 1 & 0 & 0 & 0 \\ -3 & 1 & -2 & 0 \\ 0 & 0 & 1 & 0 \\ 3 & -2 & 0 & 1 \end{array}\right)$$

nos permite deducir que $\mathrm{Im}(A)$ y $\ker(A)$ están dados por

$$\mathrm{Im}(A) = \mathrm{Vect}_k \left\langle \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix} \right\rangle \quad \text{y} \quad \ker(A) = \mathrm{Vect}_k \left\langle \begin{pmatrix} 1 \\ -3 \\ 0 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 \\ -2 \\ 1 \\ 0 \end{pmatrix} \right\rangle.$$

7.2. Operaciones sobre filas

Sea $A \in M_{m\times n}(k)$. Podemos calcular el rango de A y, más específicamente, *ecuaciones* de $\ker(A)$ y de $\mathrm{Im}(A)$ realizando operaciones sobre las *filas* de A.

Observemos que multiplicar una fila F_i de A por un escalar $\lambda_i \neq 0$ equivale a multiplicar *por la izquierda* por la matriz diagonal invertible cuyos términos diagonales valen 1, excepto por el i-ésimo que vale λ_i.

Del mismo modo, para todo $\lambda \in k$ e $i \neq j$, adicionar λF_i a F_j equivale a multiplicar A *por la izquierda* por la matriz invertible

$$B_{ji}(\lambda) = \mathrm{I}_m + \lambda E_{ji}$$

(cuya inversa es $B_{ji}(-\lambda)$).

Finalmente, intercambiar las filas F_i y F_j de A equivale a multiplicar *por la izquierda* por la matriz asociada al automorfismo de k^m que intercambia los vectores f_i y f_j de la base canónica $(f_1, \ldots, f_m)$, y que deja fijo cada f_ℓ para $\ell \neq i, j$ (esta matriz es igual a su inversa).

Definición 7.2 (operaciones elementales sobre filas)**.** Llamareoms **operaciones elementales sobre las filas** a las operaciones precedentes: intercambio de filas, multiplicación de una fila por un escalar no-nulo, o adicionar λF_i a F_j con $j \neq i$. Gracias a la discusión anterior, observamos que efectuar estas operaciones sobre las filas equivale a aplicar *automorfismos* sobre el espacio de *llegada* k^m, por lo que esto no cambia el núcleo de A ni su rango.

Luego, podemos calcular ecuaciones de $\ker(A)$ e $\mathrm{Im}(A)$ de la manera siguiente:

(1°) Sea j_1 el índice de la primera *columna* no-nula. Entonces, permutando filas, podemos suponer que $a_{1,j_1} \neq 0$ y luego, multiplicando la primera fila por a_{1,j_1}^{-1} nos reducimos al caso $a_{1,j_1} = 1$.

(2°) A continuación, sustrayendo $a_{i,j_1}F_1$ de F_i para todo $i \geq 2$, nos reducimos al caso $a_{i,j_1} = 0$ para $i \geq 2$.

(3°) Sea j_2 el índice más pequeño tal que existe $i \geq 2$ con $a_{i,j_2} \neq 0$. Procediendo como en los pasos anteriores, nos reducimos al caso $a_{2,j_2} = 1$ y $a_{i,j_2} = 0$ para $i \geq 3$.
(4°) Sea j_3 el índice más pequeño tal que existe $i \geq 3$ con $a_{i,j_3} \neq 0$. Repitiendo el proceso anterior, obtenemos una matriz de la forma siguiente:

$$A'' = \begin{pmatrix} 0 & 1 & * & a_{1,j_2} & * & a_{1,j_3} & \cdots & a_{1,j_r} & * & * \\ 0 & 0 & 0 & 1 & * & a_{2,j_3} & \cdots & a_{2,j_r} & * & * \\ 0 & 0 & 0 & 0 & 0 & 1 & * & a_{3,j_r} & * & * \\ \vdots & \vdots & \vdots & \vdots & \vdots & 0 & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \end{pmatrix}$$

Tenemos que $A'' = QA$, donde $Q \in \mathrm{GL}_m(k)$ es invertible, por lo que $\ker(A) = \ker(A'')$. En particular, $\mathrm{rg}(A) = \mathrm{rg}(A'') = r$.

Más aún, A'' provee directamente *ecuaciones* de $\ker(A'')$. En efecto, considerando el sistema lineal

$$A'' \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 0$$

observamos que podemos escoger arbitrariamente x_i para $i \notin \{j_1, \ldots, j_r\}$, y que cada x_{j_ℓ} se expresa en función de los x_i con $i > j_\ell$.

Por otro lado, $\mathrm{Im}(A'')$ es el sub-espacio W de k^m generado por los vectores $f_1, \ldots, f_m$, de donde deducimos que $\mathrm{Im}(A) = Q^{-1}(W)$. Sin embargo, no es necesario calcular la inversa de la matriz Q para tener *ecuaciones* de $\mathrm{Im}(A)$. En efecto, si escribimos la matriz A de lado de un vector columna arbitrario

$$Y = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} \in k^m$$

y aplicamos tanto a A como a Y las mismas operaciones sobre las filas, obtenemos al final

$$(A'' = QA \mid Y'' = QY),$$

de donde se deduce que $Y \in \mathrm{Im}(A)$ si y sólo si $Y'' \in \mathrm{Im}(A'')$ y, dado que $\mathrm{Im}(A'') = \mathrm{Vect}_k(f_1, \ldots, f_r)$, tenemos que $Y'' \in \mathrm{Im}(A'')$ si y sólamente si las últimas $(n-r)$ coordenadas $y''_{r+1}, \ldots, y''_n$ de Y'' son nulas. En conclusión, $\mathrm{Im}(A)$ *está determinada por las ecuaciones* $y''_{r+1} = 0, \ldots, y''_n = 0$.

Consideremos el siguiente ejemplo:

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 6 & 6 \\ 3 & 8 & 15 & 16 \end{pmatrix}.$$

Apliquemos la **reducción de filas** a la matriz A y a un vector columna arbitrario Y:

$$\left(\begin{array}{cccc|c} 1 & 2 & 3 & 4 & y_1 \\ 1 & 3 & 6 & 6 & y_2 \\ 3 & 8 & 15 & 16 & y_3 \end{array}\right).$$

Primero, las operaciones $F_2 \mapsto F_2 - F_1$ y $F_3 \mapsto F_3 - 3F_1$ nos permiten obtener

$$\left(\begin{array}{cccc|c} 1 & 2 & 3 & 4 & y_1 \\ 0 & 1 & 3 & 2 & y_2 - y_1 \\ 0 & 2 & 6 & 4 & y_3 - 3y_1 \end{array}\right),$$

y luego la operación $F_3 \mapsto F_3 - 2F_2$ nos da

$$\left(\begin{array}{cccc|c} 1 & 2 & 3 & 4 & y_1 \\ 0 & 1 & 3 & 2 & y_2 - y_1 \\ 0 & 0 & 0 & 0 & y_3 - 3y_1 - 2(y_2 - y_1) = y_3 - y_1 - 2y_2 \end{array}\right).$$

Así, las ecuaciones de $\ker(A)$ están dadas por

$$\left\{\begin{array}{r} x_1 + 2x_2 + 3x_3 + 4x_4 = 0 \\ x_2 + 3x_3 + 2x_4 = 0 \end{array}\right.$$

Dichas ecuaciones expresan tanto x_2 como x_1 en función de x_3 y x_4, que a su vez podemos escorger arbitrariamente. Eligiendo $x_3 = 1$ y $x_4 = 0$ (resp. $x_3 = 0$ y $x_4 = 1$), obtenemos los vectores

$$u = \begin{pmatrix} 3 \\ -3 \\ 1 \\ 0 \end{pmatrix} \quad \text{resp.} \quad v = \begin{pmatrix} 0 \\ -2 \\ 0 \\ 1 \end{pmatrix}$$

que forman por lo tanto una base de $\ker(A)$.
Por un lado, como Q es invertible, las soluciones de la ecuación $AX = Y$ son las mismas que las soluciones de la ecuación $QAX = QY$, es decir, $A''X = Y''$. Por otro lado, el sistema

$$\left\{\begin{array}{rcl} x_1 + 2x_2 + 3x_3 + 4x_4 & = & y_1 \\ x_2 + 3x_3 + 2x_4 & = & y_2 - y_1 \\ 0 & = & y_3 - y_1 - 2y_2 \end{array}\right.$$

posee soluciones si y sólamente si

$$y_3 = y_1 + 2y_2. \qquad (\star)$$

Obtenemos así que $(\star)$ es una ecuación de $\mathrm{Im}(A)$. Por ejemplo, escogiendo $y_1 = 1$ y $y_2 = 0$ (resp. $y_1 = 0$ y $y_2 = 1$), obtenemos que los vectores

$$\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad \text{resp.} \quad \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$$

forman una base de $\mathrm{Im}(A)$.

Importante: En el ejemplo anterior, el primer *pivote* estaba en posición (1,1), y luego el segundo en posición (2,2). Evidentemente, este no es siempre el caso; en teoría, uno siempre puede reducirse a dicho caso intercambiando filas, pero con un poco de práctica no es necesario hacer cambios de filas: basta reducirse al caso de una matriz A'' donde las filas sean escalonadas *módulo permutación de filas.* En tal caso, las filas no-nulas de A'' dan ecuaciones de $\ker(A)$, y las filas de lado de las filas nulas de A'' dan ecuaciones de $\mathrm{Im}(A)$.

Por ejemplo, el cálculo siguiente

$$\left(\begin{array}{cccc|c} 3 & 15 & 8 & 16 & y_1 \\ 1 & 6 & 3 & 6 & y_2 \\ 1 & 3 & 2 & 4 & y_3 \end{array}\right) \xrightarrow[F_2 \mapsto F_2 - F_3]{F_1 \mapsto F_1 - 3F_3} \left(\begin{array}{cccc|c} 0 & 6 & 2 & 4 & y_1 - 3y_3 \\ 0 & 3 & 1 & 2 & y_2 - y_3 \\ 1 & 3 & 2 & 4 & y_3 \end{array}\right) \xrightarrow{F_1 \mapsto F_1 - 2F_2} \left(\begin{array}{cccc|c} 0 & 0 & 0 & 0 & y_1 - 2y_2 - y_3 \\ 0 & 3 & 1 & 2 & y_2 - y_3 \\ 1 & 3 & 2 & 4 & y_3 \end{array}\right)$$

nos permite deducir las siguientes ecuaciones para $\ker(A)$

$$\begin{cases} 3x_2 + x_3 + 2x_4 = 0 \\ x_1 + 3x_2 + 2x_3 + 4x_4 = 0 \end{cases}$$

y la ecuación $y_1 - 2y_2 - y_3 = 0$ para $\mathrm{Im}(A)$.

Relación con los sistemas lineales: La reducción de filas es equivalente a la teoría de sistemas lineales estudiada durante el primer año. En efecto, si denotamos por X al vector columna

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in k^n,$$

asociamos a la matriz $A \in M_{m\times n}(k)$ el sistema lineal $AX = 0$, es decir, las m ecuaciones

$$a_{i1}x_1 + \ldots + a_{in}x_n = 0 \qquad (F_i)$$

dada por las filas $F_1, \ldots, F_m$ de A. Sea s el rango de dicho sistema, es decir, el número máximo de filas de A que son linealmente independientes (es decir, $s = \mathrm{rg}({}^{\mathrm{t}}A)$). El método general de reducción de filas muestra que, haciendo operaciones elementales sobre las filas de A, este sistema posee las mismas soluciones que el sistema escalonado $A''X = 0$, donde A'' es la matriz con lineas escalonadas que obtuvimos anteriormente. Si $j_1 < \ldots < j_s$ denotan las columnas donde se encuentran los pivotes sobre las filas $1 \ldots, s$ de A, podemos escorger arbitrariamente x_i para $i \notin \{j_1, \ldots, j_s\}$, y cada x_{j_ℓ} se expresa en función de los x_i para $i > j_\ell$. Luego, el espacio de soluciones de este sistema, que no es nada más que $\ker(A)$, es de dimensión igual a $n - s$. Por otra parte, como $\dim_k \ker(A) = n - \mathrm{rg}(A)$ gracias al teorema del rango, obtenemos el corolario siguiente (que ya demostramos anteriormente usando otros métodos):

Corolario 7.3. *Para toda matriz $A \in M_{m\times n}(k)$ se tiene* $\mathrm{rg}({}^{\mathrm{t}}A) = \mathrm{rg}(A)$.

7.3. Cálculo de la inversa de una matriz cuadrada

Sea $A \in M_n(k)$ una matriz cuadrada. Podemos utilizar el algoritmo de reducción de columnas para determinar si A es invertible y, en caso de serlo, calcular su inversa.

Denotemos por $(e_1, \ldots, e_n)$ la base canónica de k^n. Si la primera fila de A es nula, entonces A no es invertible, puesto que en tal caso la imagen de A estaría contenida en el sub-espacio generado por $e_2, \ldots, e_n$. Podemos por ende suponer que la primera línea de A es no-nula. Así, la primera etapa del algorítmo de reducción de columnas nos otorga una matriz $A_1 = AP_1$ cuya primera línea es

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

Sea $t \in \{2, \ldots, n\}$ y supongamos que luego de $t-1$ etapas obtenemos una matriz

$$A_{t-1} = AP_1 \cdots P_{t-1}$$

donde las primeras $t-1$ filas son de la forma

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

Si la t-ésima fila tiene todos sus coeficientes de índice $\geq t$ nulos, entonces $\mathrm{Im}(A)$ está contenido en el sub-espacio generado por las columnas $A_1, \ldots, A_{t-1}$ y por los vectores $e_{t+1}, \ldots, e_n$ de la base canónica, por lo que A no es invertible.

Obtenemos así las alternativas siguientes: o bien durante el proceso de reducción de columnas obtenemos una matriz de la forma

$$A_t = \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & \cdots & 0 \\ * & 1 & 0 & 0 & \cdots & 0 \\ * & * & 1 & 0 & \cdots & 0 \\ \hline * & * & * & 0 & \cdots & 0 \\ * & * & * & * & \cdots & * \\ * & * & * & * & \cdots & * \end{array}\right)$$

en cuyo caso A no es invertible (gracias a la discusión anterior), o bien obtenemos una matriz triangular inferior con 1 sobre la diagonal:

$$A_n = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ * & 1 & 0 & \cdots & 0 \\ * & * & 1 & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & 0 \\ a_{n1} & a_{n2} & \cdots & a_{n,n-1} & 1 \end{pmatrix}$$

por lo que sustrayendo $a_{nj}C_n$ a la j-ésima columna, para $j = 1, \ldots, n-1$, obtenemos una matriz

$$A_{n+1} = A_n P_{n+1} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ * & 1 & 0 & \cdots & 0 \\ * & * & 1 & \ddots & 0 \\ a_{n-1,1} & a_{n-1,2} & \cdots & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

por lo que sustrayendo a continuación $a_{n-1,j}C_{n-1}$ a la j-ésima columnas, para $j = 1, \ldots, n-2$, obtenemos una matriz

$$A_{n+2} = A_n P_{n+1} P_{n+2} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ * & 1 & 0 & \cdots & 0 \\ * & * & 1 & \ddots & 0 \\ 0 & \cdots & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

etc. Continuando de esta forma, llegamos a una matriz $A' = AP$ que es igual a la matriz identidad I_n. Luego, $P = A^{-1}$. Más aún, tal como observamos anteriormente, la matriz P se obtiene escribiendo al comienzo del proceso la matriz identidad I_n bajo la matriz A, y efectuando en cada etapa *las mismas* operaciones elementales sobre las columnas de ambas matrices. Así, al final del proceso obtenemos arriba a la matriz $A' = AP = \mathrm{I}_n$, y abajo la matriz $\mathrm{I}_n P = P = A^{-1}$.

Sea $k = \mathbb{Q}$ y consideremos el ejemplo siguiente:

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

Luego,

$$\left(\begin{array}{ccc} 2 & -1 & 1 \\ 1 & 0 & 3 \\ 1 & 2 & 0 \\ \hline 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right) \xrightarrow[C_3 \mapsto C_3 - \frac{1}{2}C_1]{\substack{C_1 \mapsto \frac{1}{2}C_1 \\ C_2 \mapsto C_2 + \frac{1}{2}C_1}} \left(\begin{array}{ccc} 1 & 0 & 0 \\ 1/2 & 1/2 & 5/2 \\ 1/2 & 5/2 & -1/2 \\ \hline 1/2 & 1/2 & -1/2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right) \xrightarrow[C_3 \mapsto C_3 - 5C_2]{C_2 \mapsto 2C_2} \left(\begin{array}{ccc} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 1/2 & 5 & -13 \\ \hline 1/2 & 1 & -3 \\ 0 & 2 & -5 \\ 0 & 0 & 1 \end{array}\right)$$

$$\xrightarrow[C_3 \mapsto -\frac{1}{13}C_3]{\substack{C_1 \mapsto C_1 + \frac{1}{26}C_3 \\ C_2 \mapsto C_2 + \frac{5}{13}C_3}} \left(\begin{array}{ccc} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 0 & 0 & 1 \\ \hline 5/13 & -2/13 & 3/13 \\ -5/26 & 1/13 & 5/13 \\ 1/26 & 5/13 & -1/13 \end{array}\right) \xrightarrow{C_1 \mapsto C_1 - \frac{1}{2}C_2} \left(\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ \hline 6/13 & -2/13 & 3/13 \\ -3/13 & 1/13 & 5/13 \\ -2/13 & 5/13 & -1/13 \end{array}\right)$$

de donde obtenemos que A es invertible y que

$$A^{-1} = \begin{pmatrix} 6/13 & -2/13 & 3/13 \\ -3/13 & 1/13 & 5/13 \\ -2/13 & 5/13 & -1/13 \end{pmatrix}.$$

Notemos que, en lugar de la serie de operaciones elementales sobre las columnas que utilizamos anteriormente, podríamos elegir cualquier serie de operaciones sobre las columnas de tal suerte que se obtengan los cálculos más simples posibles. Así, en el ejemplo anterior, es más conveniente realizar las operaciones siguientes:

$$\left(\begin{array}{ccc} 2 & -1 & 1 \\ 1 & 0 & 3 \\ 1 & 2 & 0 \\ \hline 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right) \xrightarrow[C_3 \mapsto C_3 + C_2]{\substack{C_1 \mapsto -C_2 \\ C_2 \mapsto C_1 + 2C_2}} \left(\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 3 \\ -2 & 5 & 2 \\ \hline 0 & 1 & 0 \\ -1 & 2 & 1 \\ 0 & 0 & 1 \end{array}\right) \xrightarrow{C_3 \mapsto C_3 - 3C_2} \left(\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 5 & -13 \\ \hline 0 & 1 & -3 \\ -1 & 2 & -5 \\ 0 & 0 & 1 \end{array}\right)$$

$$\xrightarrow[C_3 \mapsto -\frac{1}{13} C_3]{\substack{C_1 \mapsto C_1 - \frac{2}{13} C_3 \\ C_2 \mapsto C_2 + \frac{5}{13} C_3}} \left(\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ \hline 6/13 & -2/13 & 3/13 \\ -3/13 & 1/13 & 5/13 \\ -2/13 & 5/13 & -1/13 \end{array}\right).$$

Importante: En general, podemos utilizar también el proceso de reducción de filas, tal como se ilustra a continuación.

Sea $A \in M_n(k)$. Si la primera columnas de A es nula, entonces A no es invertible. Luego, podemos suponer que la primera columna de A es no-nula. Así, la primera etapa del algorítmo de reducción de filas nos otorga una matriz $A_1 = QA$ cuya primera columna es

$$\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Sea $t \in \{2, \ldots, n\}$ y supongamos que luego de $t-1$ etapas obtenemos una matriz

$$A_{t-1} = Q_{t-1} \cdots Q_1 A$$

donde las primeras $t-1$ columnas son de la forma

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

Si la t-ésima columna tiene todos sus coeficientes de índice $\geq t$ nulos, entonces $\mathrm{Im}(A)$ está contenido en el sub-espacio generado por los vectores $e_1, \ldots, e_{t-1}$ de la base canónica y por las columnas $A_{t+1}, \ldots, A_n$, por lo que A no es invertible.

Obtenemos así las alternativas siguientes: o bien durante el proceso de reducción de filas obtenemos una matriz de la forma

$$A_t = \left(\begin{array}{ccc|ccc} 1 & * & * & * & \cdots & * \\ 0 & 1 & * & * & \cdots & * \\ 0 & 0 & 1 & * & \cdots & * \\ \hline 0 & 0 & 0 & 0 & * & * \\ \vdots & \vdots & \vdots & 0 & \vdots & * \\ 0 & 0 & 0 & 0 & * & * \end{array}\right)$$

en cuyo caso A no es invertible (gracias a la discusión anterior), o bien obtenemos una matriz triangular

superior con 1 sobre la diagonal:

$$A_n = \begin{pmatrix} 1 & * & * & \cdots & a_{1n} \\ 0 & 1 & * & \cdots & a_{2n} \\ 0 & 0 & 1 & * & * \\ \vdots & \ddots & \ddots & \ddots & a_{n-1,n} \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

por lo que sustrayendo $a_{in}F_n$ a la i-ésima fila, para $i = 1, \ldots, n-1$, obtenemos una matriz

$$A_{n+1} = Q_{n+1}A_n = \begin{pmatrix} 1 & * & * & a_{1,n-1} & 0 \\ 0 & 1 & * & a_{2,n-1} & 0 \\ 0 & 0 & 1 & \vdots & 0 \\ 0 & \ddots & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

por lo que sustrayendo $a_{i,n-1}F_{n-1}$ a la i-ésima fila, para $i = 1, \ldots, n-2$, obteneoms una matriz

$$A_{n+2} = Q_{n+2}Q_{n+1}A_n = \begin{pmatrix} 1 & * & * & 0 & 0 \\ 0 & 1 & * & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & \ddots & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

etc. Continuando de esta forma, llegamos a una matriz $A'' = QA$ que es igual a la matriz identidad I_n. Luego, $Q = A^{-1}$. Más aún, tal como observamos anteriormente, la matriz Q se obtiene escribiendo al comienzo del proceso la matriz identidad I_n al lado de la matriz A, y efectuando en cada etapa *las mismas* operaciones elementales sobre las filas de ambas matrices. Así, al final del proceso obtenemos a la izquierda la matriz $A'' = QA = \mathrm{I}_n$, y a la derecha la matriz $Q\mathrm{I}_n = Q = A^{-1}$.

Sea $k = \mathbb{Q}$ y consideremos el ejemplo precedente:

$$\left(\begin{array}{ccc|ccc} 2 & -1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 3 & 0 & 1 & 0 \\ 1 & 2 & 0 & 0 & 0 & 1 \end{array}\right) \xrightarrow[F_3 \mapsto F_3 - \frac{1}{2}F_1]{\substack{F_1 \mapsto \frac{1}{2}F_1 \\ F_2 \mapsto F_2 - \frac{1}{2}F_1}} \left(\begin{array}{ccc|ccc} 1 & -1/2 & 1/2 & 1/2 & 0 & 0 \\ 0 & 1/2 & 5/2 & -1/2 & 1 & 0 \\ 0 & 5/2 & -1/2 & -1/2 & 0 & 1 \end{array}\right)$$

$$\xrightarrow[F_3 \mapsto F_3 - 5F_2]{F_2 \mapsto 2F_2} \left(\begin{array}{ccc|ccc} 1 & -1/2 & 1/2 & 1/2 & 0 & 0 \\ 0 & 1 & 5 & -1 & 2 & 0 \\ 0 & 0 & -13 & 2 & -5 & 1 \end{array}\right)$$

$$\xrightarrow[F_3 \mapsto -\frac{1}{13}F_3]{\substack{F_1 \mapsto F_1 + \frac{1}{26}F_3 \\ F_2 \mapsto F_2 + \frac{5}{13}F_3}} \left(\begin{array}{ccc|ccc} 1 & -1/2 & 0 & 15/26 & -5/26 & 1/26 \\ 0 & 1 & 0 & -3/13 & 1/13 & 5/13 \\ 0 & 0 & 1 & -2/13 & 5/13 & -1/13 \end{array}\right)$$

$$\xrightarrow{F_1 \mapsto F_1 + \frac{1}{2}F_2} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 6/13 & -2/13 & 3/13 \\ 0 & 1 & 0 & -3/13 & 1/13 & 5/13 \\ 0 & 0 & 1 & -2/13 & 5/13 & -1/13 \end{array}\right),$$

de donde se concluye que A es invertible y que

$$A^{-1} = \begin{pmatrix} 6/13 & -2/13 & 3/13 \\ -3/13 & 1/13 & 5/13 \\ -2/13 & 5/13 & -1/13 \end{pmatrix}.$$

Notemos nuevamente que, en lugar de la serie de operaciones elementales sobre las filas que utilizamos anteriormente, podríamos elegir cualquier serie de operaciones sobre las filas de tal suerte que se obtengan los cálculos más simples posibles. Así, en el ejemplo anterior, es más conveniente realizar las operaciones

siguientes:

$$\left(\begin{array}{ccc|ccc} 2 & -1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 3 & 0 & 1 & 0 \\ 1 & 2 & 0 & 0 & 0 & 1 \end{array}\right) \xrightarrow[F_3 \mapsto F_3 - F_2]{\substack{F_1 \mapsto F_2 \\ F_2 \mapsto F_1 - 2F_2}} \left(\begin{array}{ccc|ccc} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & -1 & 5 & 1 & -2 & 0 \\ 0 & 2 & -3 & 0 & -1 & 1 \end{array}\right)$$

$$\xrightarrow[F_3 \mapsto F_3 + 2F_2]{F_2 \mapsto -F_2} \left(\begin{array}{ccc|ccc} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 5 & -1 & 2 & 0 \\ 0 & 0 & -13 & 2 & -5 & 1 \end{array}\right)$$

$$\xrightarrow[F_3 \mapsto -\frac{1}{13} F_3]{\substack{F_1 \mapsto F_1 + \frac{3}{13} F_3 \\ F_2 \mapsto F_2 + \frac{5}{13} F_3}} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 6/13 & -2/13 & 3/13 \\ 0 & 1 & 0 & -3/13 & 1/13 & 5/13 \\ 0 & 0 & 1 & -2/13 & 5/13 & -1/13 \end{array}\right).$$

¡Atención! En este algoritmo para calcular A^{-1} hay que escoger si se hacen operaciones sobre las columnas *o bien* sobre las filas, pero *no hay que mezclarlas* (!).
En efecto, si hicieramos a la vez operaciones sobre las filas y columnas de A, y si hicieramos las mismas operaciones sobre la matriz identidad I_n, llegaríamos al final del proceso a un par de matrices

$$(QAP = \mathrm{I}_n \mid Q\mathrm{I}_n P = QP),$$

y por ende la primera igualdad nos daría $A = Q^{-1}P^{-1}$ de donde obtenemos $A^{-1} = PQ$. Sin embargo, en el lado derecho habríamos calculado QP en lugar de PQ.

Ejercicio 7.4. Sea k un cuerpo. Supongamos que una matriz invertible $A \in \mathrm{GL}_n(k)$ conmuta con todas las matrices invertibles, es decir, supongamos que

$$AB = BA$$

para toda $B \in \mathrm{GL}_n(k)$. Probar que $A = \lambda \mathrm{I}_n$ es una homotecia, para cierto $\lambda \neq 0$.

8. Apéndice: existencia de bases (de Hamel)

El presente apéndice tiene dos objetivos: el primero es demostrar el teorema fundamental sobre espacios vectoriales finitamente generados (ver Teorema 3.10), y el segundo es probar el hecho que *todo* espacio vectorial (no necesariamente finitamente generado) admite una base. Si bien que este último hecho es una generalización del primero, por motivos sicológicos he decidido separar la discusión en dos partes relativamente independientes.

8.1. Teorema fundamental de espacios vectoriales finitamente generados

Referimos al lector a la seccion §3 para recordar los conceptos de familias generadoras y familias linealmente independientes, así como la caracterización de conjuntos linealmente independientes infinitos dada por el Lema 3.7.

Recordemos que un k-espacio vectorial V es *finitamente generado* si posee un conjunto de generadores finito, es decir, si existen $v_1, \ldots, v_n \in V$ tales que $V = \mathrm{Vect}_k(v_1, \ldots, v_n)$. El objetivo de esta sección es probar el siguiente resultado, anunciado anteriormente en el Teorema 3.10.

Teorema 8.1. *Sea V un k-espacio vectorial finitamente generado. Entonces:*

1. *Existen bases de V, y todas tienen el mismo cardinal n; este entero se llama la* dimensión *de V sobre k y es denotada* $\dim_k(V)$ *o simplemente* $\dim(V)$.
2. *De toda familia generadora $\mathscr{F}$ podemos extraer una base, en particular $\mathscr{F}$ es de cardinal* $\geq \dim_k(V)$. *Más aún, si* $\mathrm{card}(\mathscr{F}) = \dim_k(V)$ *entonces $\mathscr{F}$ es una base de V.*
3. *Toda familia linealmente independiente es de cardinal* $\leq \dim_k(V)$. *Más aún, toda familia linealmente independiente de cardinal* $\dim_k(V)$ *es una base de V.*

4. Teorema de la base incompleta: *Toda familia linealmente independiente puede ser completada en una base de V.*

5. *Todo sub-espacio W de V es de dimensión finita $\leq \dim_k(V)$. Más aún, si $\dim_k(W) = \dim_k(V)$ entonces $W = V$. En otras palabras, todo sub-espacio vectorial distinto de V es de dimensión $< \dim_k(V)$.*

Comencemos por introducir la siguiente terminología.

Definición 8.2. Sea V un k-espacio vectorial, y sea $\mathscr{F} = (v_i)_{i\in I}$ una familia de vectores de V.

1. Decimos que $\mathscr{F}$ es una **familia generadora minimal** si es una familia generadora y si «*no podemos hacerla más pequeña*», es decir, si para todo sub-conjunto $J \subsetneq I$, la familia $(v_i)_{i\in J}$ *no es* generadora.

2. Decimos que $\mathscr{F}$ es una **familia linealmente independiente maximal** si es una familia linealmente independiente y si «*no podemos hacerla más grande*», es decir, si para todo $v \in V \setminus \mathscr{F}$, la familia $\mathscr{F} \cup \{v\}$ es linealmente dependiente.

Proposición 8.3. *Sea V un k-espacio vectorial y sea $\mathscr{F} = (v_i)_{i\in I}$ una familia no-vacía de vectores de V. Las condiciones siguientes son equivalentes:*

1. *$\mathscr{F}$ es una base de V.*

2. *$\mathscr{F}$ es una familia generadora minimal.*

3. *$\mathscr{F}$ es una familia linealmente independiente maximal.*

Demostración. Supongamos que $(v_i)_{i\in I}$ es una base de V, es decir, que es una familia generadora y linealmente independiente. Entonces, para todo $i \in I$, la familia $(v_j)_{j\in I\setminus\{i\}}$ *no es* generadora: en caso contrario, v_i se escribiría como combinación lineal de los v_j, donde $j \neq i$, de donde obtendríamos una relación lineal no-trivial $v_i - (\lambda_1 v_{j_1} + \ldots + \lambda_r v_{j_r}) = 0$. Esto prueba que $(v_i)_{i\in I}$ es una familia generadora minimal.
Más aún, todo $v \in V$ se escribe como una combinación lineal (de un número finito) de v_i, por lo que si $v \notin \mathscr{F}$, la familia estrictamente mayor $\mathscr{F} \cup \{v\}$ es linealmente dependiente. Esto prueba que $(v_i)_{i\in I}$ es una familia linealmente independiente maximal. En otras palabras, tenemos que (1) implica (2) y (3).
Veamos que (2) implica (1). Para esto, supongamos que $(v_i)_{i\in I}$ es una familia generadora minimal y probemos que ella es linealmente independiente. En caso contrario, tendríamos una relación lineal no-trivial

$$\lambda_1 v_{i_1} + \ldots + \lambda_r v_{i_r} = 0$$

donde $r \geq 1$ y donde, sin pérdida de generalidad, los $\lambda_j \neq 0$. Así, $v_{i_r} = -\lambda_r^{-1} \sum_{j\neq i} \lambda_j v_{i_j}$ y luego la familia $(v_i)_{i\in I\setminus\{i_r\}}$ ya sería generadora, contradiciendo la hipótesis de minimalidad. Concluímos que (2) implica (1).
Finalmente, veamos que (3) implica (1). Para esto supongamos que $(v_i)_{i\in I}$ es una familia linealmente minimal y probemos que ella es generadora. Sea $v \in V \setminus \mathscr{F}$, entonces la familia $\mathscr{F} \cup \{v\}$ es linealmente dependiente, por lo que existe una relación lineal

$$\mu v + \lambda_1 v_{i_1} + \ldots + \lambda_r v_{i_r} = 0$$

que es *no-trivial* (es decir, μ y los λ_j *no son todos nulos*). Notar que es imposible que $\mu = 0$, pues en tal caso tendríamos que $\lambda_1 = \ldots = \lambda_r = 0$, pues los v_i son linealmente independientes. Así, $\mu \neq 0$ y por ende $v = -\mu^{-1}(\lambda_1 v_{i_1} + \ldots + \lambda_r v_{i_r})$. Luego, $\mathscr{F}$ es una familia generadora, de donde concluímos que (3) implica (1). □

Corolario 8.4. *Todo k-espacio vectorial finitamente generado posee una base*[11].

[11] Aquí, y en todo el texto, adoptamos la convención que el conjunto vacío $\emptyset$ es una base del espacio nulo $\{0\}$. En particular, $\mathrm{Vect}_k(\emptyset) = \{0\}$.

Demostración. Sea V un k-espacio vectorial finitamente generado. Por definición, V es generado por una familia *finita* $\{v_1, \ldots, v_r\}$. Esta familia contiene una sub-familia generadora $\mathscr{B}$ de cardinalidad minimal, la cual es por ende una familia generadora minimal. Gracias a la proposición anterior, dicha familia $\mathscr{B}$ es una base de V. □

Para probar el resultado principal de esta sección necesitaremos el lema siguiente.

Lema 8.5. *Sea V un k-espacio vectorial y sean $m > n$ dos enteros positivos. Sea $v_1, \ldots, v_n \in V$ un conjunto de vectores, y sean $u_1, \ldots, u_m$ vectores de V que son combinaciones lineales de los $v_1, \ldots, v_n$. Entonces, los vectores $u_1, \ldots, u_m$ son linealmente dependientes.*

Demostración. Procedemos por inducción en n. El resultado es claramente cierto para $n = 1$, por lo que supondremos $n \geq 2$ y que le resultado es cierto para $n - 1$. Escribamos

$$\begin{cases} u_1 = a_{11}v_1 + \ldots + a_{1n}v_n \\ u_2 = a_{21}v_1 + \ldots + a_{2n}v_n \\ \vdots \qquad\qquad \vdots \\ u_m = a_{m1}v_1 + \ldots + a_{mn}v_n. \end{cases}$$

Si todos los u_i son combinación lineal de los $n-1$ vectores $v_2, \ldots, v_n$, entonces los u_i son linealmente independientes, gracias a la hipótesis de inducción. En caso contrario, módulo reordenar los índices si fuese necesario, podemos suponer que $a_{11} \neq 0$. Luego, para $i = 2, \ldots, m$, los $m-1$ vectores

$$u_i' = u_i - a_{i1}a_{11}^{-1}u_1$$

son combinación lineal de los $n-1$ vectores $v_2, \ldots, v_n$ y por ende son linealmente independientes, gracias a la hipótesis de inducción. Así, existen escalares $\lambda_2, \ldots, \lambda_m$ no todos nulos, tales que

$$0 = \lambda_2 u_2' + \ldots + \lambda_m u_m' = \lambda_2 u_2 + \ldots + \lambda_m u_m - \left(\sum_{i=2}^{m} \lambda_i a_{i1} a_{11}^{-1} \right) u_1,$$

de donde concluímos que los vectores $u_1, \ldots, u_m$ son linealmente dependientes. □

Observación 8.6. *El lema precedente tiene las siguientes importantes consecuencias: Sea V un k-espacio vectorial generado por un número finito de elementos $x_1, \ldots, x_N$. Gracias al corolario anterior, V posee una base $\mathscr{B} = (v_1, \ldots, v_n)$ formada de $n \leq N$ elementos. Dado que $\mathscr{B}$ es una familia linealmente independiente, el lema precedente implica que:*

(a) *Toda familia generadora de V posee al menos n elementos.*

(b) *Toda familia generadora de V constituída por exactamente n elementos es* minimal*, y luego es una base de V.*

De mismo modo, dado que $\mathscr{B}$ es una familia generadora de V, el lema precedente implica que:

(a') *Toda familia linealmente independiente de V posee a lo más n elementos.*

(b') *Toda familia linealmente independiente de V constituída por exactamente n elementos es* maximal*, y luego es una base de V.*

En conclusión, toda base de V, siendo a la vez una familia generadora y linealmente independiente, está constituída exactamente por n elementos.

Veamos a continuación la prueba del teorema fundamental de espacios vectoriales finitamente generados.

Demostración del Teorema 8.1. Ya hemos visto los puntos (1), (2) y (3). El Teorema de la base incompleta (4) se deduce del hecho que toda familia linealmente independiente puede completarse en una familia linealmente independiente maximal, es decir, en una base de V.
Finalmente, para probar (5) consideremos un sub-espacio W de V. Gracias a (3), toda familia linealmente independiente de elementos de W es de cardinal $\leq n = \dim_k(V)$, por lo que W posee una familia

linealmente independiente *maximal* $\mathscr{C}$, de cardinal $m \leq n$. Luego, $\mathscr{C}$ es una base de W, pues $\mathscr{C}$ es una familia linealmente independiente maximal, de donde concluimos que W es finitamente generado, y de dimensión $m \leq n$. Si además se tiene $m = n$ entonces, gracias a (3), $\mathscr{C}$ es una base de V (y luego genera V), de donde se concluye que $W = V$. □

8.2. Lema de Zorn y existencia de bases de Hamel

Hemos visto que todo espacio vectorial finitamente generado posee una base. Pero, ¿qué sucede con los espacios vectoriales que no son finitamente generados (tales como el espacio $\mathscr{C}([0,1],\mathbb{R})$ de funciones continuas de $[0,1]$ con valores reales)?, ¿posee un tal espacio vectorial una base?.

Veamos qué ocurre en algunos ejemplos sencillos.

Ejemplo 8.7.

1. Consideremos el espacio $\mathbb{R}[X]$ de polinomios en la variable X con coeficientes reales. Tal como observamos anteriormente, dicho $\mathbb{R}$-espacio vectorial *no* está generado por ningún conjunto finito. Por otra parte, todo polinomio puede ser escrito como una única una combinación lineal *finita* de los monomios $(X^n)_{n\in\mathbb{N}}$. En otras palabras, $(X^n)_{n\in\mathbb{N}}$ es una base de $\mathbb{R}[X]$.

2. Sea $\mathbb{R}^\infty$ (o también $\mathbb{R}^\mathbb{N}$) el espacio vectorial de todas las sucesiones infinitas $(a_0, a_1, a_2, \ldots)$ de números reales, con la suma y multiplicación por escalares componente a componente. ¿Qué podría ser una base de $\mathbb{R}^\infty$?

   Una familia natural a considerar es $\{e_0, e_1, e_2, \ldots\}$ generalizando la base canónica de $\mathbb{R}^n$. En otras palabras, e_i es la sucesión compuesta sólo de ceros, excepto en la posición i-ésima donde aparece un 1. Dicha familia es linealmente independiente, sin embargo no es generadora: el espacio generado por ellos es por definición (!) el conjunto de sucesiones tales que sólo un número *finito* de términos no son cero. Así, el vector $v = (1, 1, 1, \ldots)$ no puede ser escrito como una combinación lineal finita de los e_i.

   Dado que el vector v no puede ser escrito como combinación lineal de los e_i, la familia formada por los e_i y por v es linealmente inpedendiente. Lamentablemente, dicho conjunto tampoco es una base: el vector $(1, 2, 3, 4, \ldots)$ no es una combinación lineal de los e_i y v. Podríamos intentar repetir este proceso agregando cada vez un vector nuevo y esperando que eventualmente termine para así obtener una base de $\mathbb{R}^\infty$. Sin embargo, es posible probar que ningún conjunto *numerable* de vectores de $\mathbb{R}^\infty$ puede ser una base de $\mathbb{R}^\infty$, por lo que incluso de $\mathbb{R}^\infty$ posee una base, no seremos capaces de construirla simplemente agregando elementos uno por uno a la familia canónica $(e_i)_{i\in\mathbb{N}}$.

3. El mismo argumento anterior, implica que si consideramos V como el sub-espacio de $\mathbb{R}^\infty$ dado por las sucesiones *convergentes* de números reales, entonces $(e_0, e_1, e_2, \ldots)$ es una familia linealmente independiente que *no es* generadora. En efecto, la sucesión constante $(1, 1, 1, \ldots) \in V$ es convergente.

4. Sea $\ell^2(\mathbb{R})$ el sub-espacio de $\mathbb{R}^\infty$ formado por todas las sucesiones *cuadrado sumables*, es decir, aquellos vectores $(a_0, a_1, a_2, \ldots)$ tales que la serie $\sum_{i=0}^\infty a_i^2$ converge. Este espacio incluye a los elementos e_i de la familia canónica, pero no está generado por ellos. En efecto, el vector $(1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \ldots) \in \ell^2(\mathbb{R})$ (la suma anterior vale 4/3), pero no es una combinación lineal finita de los e_i. ¿Posee $\ell^2(\mathbb{R})$ una base?

5. El conjunto $\mathbb{R}$ de los números reales puede ser considerado como un espacio vectorial sobre el cuerpo $\mathbb{Q}$ de los números racionales. ¿Cuál podría ser una base?

   El posible probar que los elementos $\sqrt{2}, \sqrt{3}, \sqrt{5}, \sqrt{6}, \ldots$ son linealmente independientes, pero ellos no generan a $\mathbb{R}$ puesto que también se necesitan elementos tales como $\pi, \pi^2, \pi^3, \ldots$, que también forman un conjunto linealmente independiente. En efecto, dado que $\mathbb{Q}$ es un conjunto numerable, uno puede probar que cualquier sub-espacio de $\mathbb{R}$ generado por un sub-conjunto numerable de $\mathbb{R}$ debe ser también numerable. Dado que $\mathbb{R}$ *no es* numerable, ningún conjunto numerable puede ser una base de $\mathbb{R}$ sobre $\mathbb{Q}$. Esto significa que una base de $\mathbb{R}$ sobre $\mathbb{Q}$, si llegase a existir, va a ser difícil de describir.

Los ejemplos anteriores nos señalan que incluso aunque *pudiesemos* demostrar que todo espacio vectorial admite una base, esto no quiere decir que dicha base sea fácil de encontrar o de describir. El principal resultado de esta sección es el siguiente.

Teorema 8.8 (Bleicher (1964), Halpern (1966))**.** *Todo k-espacio vectorial admite una base.*

Para probar este importante resultado es necesario utilizar el *Lema de Zorn*[12], que recordamos a continuación.

Recuerdo (Lema de Zorn): Sea P un conjunto y sea $\mathscr{R}$ una relación en P. Si para todo $(a, b) \in P \times P$ tal que $(a, b) \in \mathscr{R}$ escribimos $a \preccurlyeq b$, entonces decimos que $\mathscr{R}$ es un **orden parcial** si es:

1. **reflexiva:** $a \preccurlyeq a$ para todo $a \in P$,
2. **anti-simétrica:** si $a \preccurlyeq b$ y $b \preccurlyeq a$ entonces $a = b$, para todos $a, b \in P$,
3. **transitiva:** si $a \preccurlyeq b$ y $b \preccurlyeq c$ entonces $a \preccurlyeq c$, para todos $a, b, c \in P$.

El par $(P, \preccurlyeq)$ es llamado un **conjunto parcialmente ordenado**.

Sea $(P, \preccurlyeq)$ un conjunto parcialmente ordenado y sea $T \subseteq A$ un sub-conjunto. Decimos que T es **totalmente ordenado** o una **cadena** si para todos $a, b \in T$ se tiene que $a \preccurlyeq b$ o bien $b \preccurlyeq a$.

Sea $T \subseteq P$ una cadena y sea $s \in P$. Decimos que s es una **cota superior** de T si $t \preccurlyeq s$ para todo $t \in T$. Un conjunto parcialmente ordenado no vacío tal que toda cadena posee una cota superior es llamado un **conjunto inductivo**.

Sea $m \in P$. Decimos que m es un **elemento maximal** si para todo $a \in P$ se tiene que $m \preccurlyeq a$ implica que $m = a$.

Finalmente, el **lema de Zorn** afirma que *todo conjunto inductivo posee al menos un elemento maximal.*

Demostración del Teorema 8.8. Sea V un k-espacio vectorial. Sea P el conjunto de *todas* las familias $\mathscr{F}$ linealmente independientes de V, parcialmente ordenado por la inclusión de conjuntos. Dicho conjunto es no-vacío pues cualquier vector $v \neq 0$ de V define una familia linealmente independiente $\{v\}$. Sea

$$\mathscr{F}_1 \subseteq \mathscr{F}_2 \subseteq \mathscr{F}_3 \subseteq \cdots$$

una cadena en P. El conjunto $\mathscr{S} = \bigcup_{i \geq 1} \mathscr{F}_i$ pertenece a P, puesto que la unión *creciente* de familias linealmente independientes es linealmente independiente, y es claramente una cota superior para la cadena anterior. Así, el Lema de Zorn implica que existe una familia linealmente independiente maximal. Luego, la Proposición 8.3 implica que $\mathscr{M}$ es una base de V. □

Importante: El mayor inconveniente en la demostración del teorema anterior es que, a menos que la dimensión de V sea finita (o al menos numerable), no da ningún método para calcular concretamente alguna base. Así, a pesar de que el Lema de Zorn nos asegura la existencia de bases, en la práctica, este hecho puede ser un poco inútil si no tenemos un procedimiento para hallar una.

Por esta razón, las bases no son comúnmente usadas en espacios vectoriales de dimensión infinita y es necesario encontrar otros reemplazos. Debido a lo anterior, una base de un espacio vectorial de dimensión finita es llamada una **base de Hamel** para poder distinguirla de otras nociones más útiles que usualmente (¡y confusamente!) son llamadas bases.

Por ejemplo, en *Análisis de Fourier* se construyen ciertas «bases» ortonormales que permiten expresar vectores como combinaciones lineales *infinitas*, que son llamadas *series de Fourier*. En la práctica, a pesar de que un espacio vectorial de dimensión infinita posea una base ortonormal numerable (como por ejemplo $\mathscr{C}([0, 1], \mathbb{R})$), puede ocurrir perfectamente que ninguna base de Hamel sea numerable. Otra importante diferencia entre dichas bases ortonormales y las bases de Hamel es que no es cierto que *toda* combinación lineal infinita de elementos de la base ortonormal pertenezca a nuestro espacio vectorial (esto ocurre en $\mathscr{C}([0, 1], \mathbb{R})$, por ejemplo).

[12] Es un hecho conocido, que el Lema de Zorn es equivalente al Axioma de elección. En 1984, Andreas Blass publica el artículo *Existence of bases implies the axiom of choice*, donde prueba que la existencia de bases de todo espacio vectorial es equivalente al Axioma de elección.

Índice alfabético