Pedro Montero

con la colaboración de A. Carrasco, S. Fuentes & E. Treumún

ÁLGEBRA LINEAL

Pedro MONTERO,
con la colaboración de A. CARRASCO*, S.* FUENTES *& E.* TREUMÚN

ÁLGEBRA LINEAL

Pedro Montero

con la colaboración de A. Carrasco, S. Fuentes & E. Treumún

CONTENIDOS

AGRADECIMIENTOS

El presente texto tiene como principal objetivo complementar el curso de *Álgebra Lineal* en la Universidad Técnica Federico Santa María.

Quisiera expresar mi gratitud a Alonso Carrasco, Sebastián Fuentes y Ernesto Treumún por su ayuda en traspasar pacientemente a LaTeX parte de mis apuntes escritos a mano, así como por sus comentarios y observaciones a versiones previas de este texto.

Agradezco de antemano por cualquier tipo de comentario, sugerencia o corrección, para lo cual pueden comunicarse directamente conmigo al correo electrónico `pedro.montero@usm.cl`.

Pedro Montero
Departamento de Matemática
Universidad Técnica Federico Santa María
Valparaíso, Diciembre 2023

CAPÍTULO 1

ESPACIOS VECTORIALES Y APLICACIONES LINEALES

1.1. Definiciones generales

***Definición 1.1.1* (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 1.1.2. — 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 1.1.3* (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 1.1.4. —

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 $k^\times = k \setminus \{0\}$.

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

***Definición 1.1.5* (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 1.1.6* (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 1.1.7* (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 1.1.8. — *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 1.1.9. — 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 1.1.10. —

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 1.1.11 **(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 $x_0, y_0 \in \mathbb{Z}$. $\square$

Corolario 1.1.12. — *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 $ab \equiv 1 \pmod{p}$. $\square$

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 1.1.3). En efecto, para todo $[a]_p \neq [0]_p$ existe $[a]_p^{-1}$ tal que $[a]_p \cdot [a]_p^{-1} = [1]_p$.

Definición 1.1.13. — Sea p un número primo. Denotaremos por $\mathbb{F}_p$ al **cuerpo de p elementos** $(\mathbb{Z}/p\mathbb{Z}, +, \cdot)$.

Ejercicio 1.1.14. — Calcular las tablas de suma y multiplicación en $\mathbb{F}_3$.

Ejercicio 1.1.15. — 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 $(x+y)^p = x^p + y^p$.

1.2. Espacios vectoriales: definición y ejemplos

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

Ejemplo 1.2.1. —

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

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

   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

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

   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 1.2.2* (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 1.2.3* (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 1.2.4. — 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 1.2.5* (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 1.2.6.* —**

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 1.2.7* (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 1.2.8.* —** 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 $\lambda_1, \lambda_2 \in k$.

***Definición 1.2.9* (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**.

[(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.

Ejemplo 1.2.10. —

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 1.2.11 **(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 1.2.12. — 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 1.2.13. — 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

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

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 1.2.14. — 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 $k_d[X] \cong k^{d+1}$.

1.3. Familias generadoras, libres, bases y dimensión

***Definición 1.3.1* (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 1.3.2* (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**

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

si posee un conjunto de generadores finitos, es decir, si $V = \mathrm{Vect}_k(v_1, \dots, v_n)$ para ciertos $v_1, \dots, v_n \in V$.

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

Ejemplo 1.3.3. — Sea k un cuerpo.

1. Para todo $n \in \mathbb{N}^{\geq 1}$, el espacio k^n está generado por los vectores
$$e_1 = (1, 0, \dots, 0),\ e_2 = (0, 1, 0, \dots, 0), \dots,\ e_n = (0, \dots, 0, 1).$$
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: $P = a_0 + a_1X + \dots + a_dX^d$.

***Definición 1.3.4* (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 $v_1, \dots, v_r \in S$ distintos, y todo conjunto de escalares $\lambda_1, \dots, \lambda_r \in k$ tenemos que: si se verifica la relación $\lambda_1 v_1 + \dots + \lambda_r v_r = 0$, entonces $\lambda_1 = \dots = \lambda_r = 0$.

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 $v_1, \dots, v_r \in S$ distintos, y escalares $\lambda_1, \dots, \lambda_r \in k$ *no todos* iguales a 0, tales que $\lambda_1 v_1 + \dots + \lambda_r v_r = 0$.

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

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

En caso contrario, tenemos que los vectores $v_1, \dots, v_n$ son linealmente dependientes si existen escalares $\lambda_1, \dots, \lambda_n \in k$ no todos nulos, tales que $\lambda_1 v_1 + \dots + \lambda_n v_n = 0$. 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

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

Observación 1.3.5. — 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 1.3.6. — 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 1.3.7. — *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 1.3.8* (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 $\mathscr{B} = \{v_1, \dots, v_n\}$ es *finito*. En tal caso, los vectores $v_1, \dots, v_n$ forman una base de V si para todo $v \in V$ existe una única n-tupla $(\lambda_1, \dots, \lambda_n) \in k^n$ tal que $v = \lambda_1 v_1 + \dots + \lambda_n v_n$.

Ejemplo 1.3.9. — Sea k un cuerpo.

1. Si $V = k^n$, la familia $\{e_1, \dots, e_n\}$ dada por
$$e_1 = (1, 0, \dots, 0),\ e_2 = (0, 1, 0, \dots, 0), \dots,\ e_n = (0, \dots, 0, 1),$$

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 1.3.10. — *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 1.3.11. — 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 1.3.12* (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 1.3.13. — *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 $\mathscr{F} = (v_1, \dots, v_n)$ es una familia linealmente independiente de V. Supogamos que en W existe una relación lineal

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

donde $\lambda_i \in k$. Entonces, dado que f es lineal, tenemos que $0 = f(\lambda_1 v_1 + \dots + \lambda_n v_n)$. Como f es inyectiva, $0 = \lambda_1 v_1 + \dots + \lambda_n v_n$. Así, dado que $\mathscr{F}$ es una familia libre, tenemos que $\lambda_1 = \dots = \lambda_n = 0$, de donde concluimos que $f(\mathscr{F})$ es libre.

Para probar (2) supongamos que $f : V \twoheadrightarrow W$ es sobreyectiva y sea $\mathscr{F} = (v_1, \dots, v_n)$ 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 $\lambda_1, \dots, \lambda_n \in k$ tales que $v = \lambda_1 v_1 + \dots + \lambda_n v_n$, de donde se concluye que

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

En otras palabras, $f(\mathscr{F})$ genera W.

Finalmente, para demostrar (3) supongamos que $f : V \xrightarrow{\sim} W$ es biyectiva y sea $\mathscr{B} = (v_1, \dots, v_n)$ 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 1.3.14. — *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 1.3.15. —

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 $e_2 - e_1$.

[(4)] Este isomorfismo **no es canónico**, pues depende de la elección de una base.

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 1.3.16. — 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 $s\lambda + t\mu = w_1$.

1.4. Núcleo, imagen y teorema del rango

***Definición 1.4.1* (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

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

el cual es un sub-espacio vectorial de W. Notar que f es sobreyectiva si y sólamente si $\operatorname{Im}(f) = W$.

Cuando $\operatorname{Im}(f)$ es de *dimensión finita* $r \in \mathbb{N}$ (esto ocurre, por ejemplo, si W o V son de dimensión finita), el entero $r = \dim_k \operatorname{Im}(f)$ es llamado el **rango** de f y es denotado $\operatorname{rg}(f)$ o bien $\operatorname{rango}(f)$.

(5) En efecto, tenemos que $f(v_1) = f(v_2)$ si y sólamente si $f(v_1 - v_2) = 0$.

***Teorema 1.4.2* (Teorema del rango).** — *Sea $f : V \to W$ una aplicación lineal entre k-espacios vectoriales. Supongamos que V es de dimensión finita. Entonces,*

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

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

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, $\operatorname{Im}(f) = f(V)$ 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 $\operatorname{rg}(f) = \dim_k f(V) = \dim_k(V) - \dim_k \ker(f)$. □

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 $\operatorname{Im}(f)$, 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 $\operatorname{Im}(f) = f(V)$ 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 $\operatorname{Im}(f)$ 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 $\dim_k(V) = d + r = \dim_k \ker(f) + \dim_k \operatorname{rg}(f)$. □

Proposición 1.4.3. — *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, $\operatorname{Im}(f) = W$), entonces el teorema del rango implica que f es también sobreyectiva (resp. inyectiva), y luego biyectiva. □

1.5. Aplicaciones lineales y matrices

Definición 1.5.1 **(espacio de aplicaciones lineales)**

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

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

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

$$(\star) \qquad \varphi(v) = x_1w_1+\ldots+x_nw_n$$

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 1.5.2. — *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 1.5.3 **(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 1.5.4. — 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 1.5.5. — *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 $B_1, \dots, B_p \in k^n$ las columnas de la matriz B, entonces las columnas de AB son los vectores $AB_1, \dots, AB_p$. En efecto, si $(e_1, \dots, e_p)$ 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 1.5.6. — *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 $\mathscr{B} = (e_1, \dots, e_n)$ de V, la aplicación*

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

---

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

***Definición 1.5.7* (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 1.5.8. — 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,

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

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

***Definición 1.5.9* (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

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

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 1.5.10. — *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 ${}^{\mathrm{t}}(AB) = {}^{\mathrm{t}}B\,{}^{\mathrm{t}}A$. □

1.6. Cambios de base

Definición 1.6.1 **(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 1.6.2. — *Sea k un cuerpo.*

(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

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

[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 $f^{-1} \circ g^{-1}$.

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

¡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 1.6.3. — *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 $f \circ g = \mathrm{id}_V$. $\square$

***Definición 1.6.4* (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 $v_1, \dots, v_n$ expresados en la base $(e_1, \dots, e_n)$. 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 $\mathscr{B} = (e_1, \dots, e_n)$ en la base $\mathscr{B}' = (v_1, \dots, v_n)$.

Ejercicio 1.6.5. — Verificar que P puede ser vista como la matriz asociada a la aplicación identidad id_V, expresada en las base $\mathscr{B}' = (v_1, \dots, v_n)$ en la partida y $\mathscr{B} = (e_1, \dots, e_n)$ 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 1.6.6 **(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 1.6.7* (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**:

$$\begin{tikzcd}
(V,\mathscr{B}) \arrow[r, "A=\mathrm{Mat}_{\mathscr{C},\mathscr{B}}(u)"] & (W,\mathscr{C}) \arrow[d, "Q^{-1}=\mathrm{Mat}_{\mathscr{C}'}(\mathscr{C})"] \\
(V,\mathscr{B}') \arrow[u, "\mathrm{Mat}_{\mathscr{B}}(\mathscr{B}')=P"] \arrow[r, "Q^{-1}AP=\mathrm{Mat}_{\mathscr{C}',\mathscr{B}'}(u)"'] & (W,\mathscr{C}')
\end{tikzcd}$$

Observación 1.6.8. — 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 1.6.9. — *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 1.6.10. — *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 1.6.11* (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 1.6.12. — 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 $r \in \{0, \ldots, \min(\dim_k(V), \dim_k(W))\}$.

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 1.6.13* (cambio de base para endomorfismos)**
Sea A la matriz de un endomorfismos $u : V \to V$ respecto a una base $\mathscr{B}$

(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.

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 1.6.14* (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:

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.

1.7. Operaciones elementales sobre filas y columnas

1.7.1. 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 $(e_1, \dots, e_n)$ 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 1.7.1 **(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:

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

[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.

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 $Y = (y_1, \dots, y_m) \in k^m$ (resp. $X = (x_1, \dots, x_n) \in k^n$), 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 $\operatorname{Im}(A)$ y $\ker(A)$ están dados por

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

1.7.2. 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 $\operatorname{Im}(A)$ 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 1.7.2 **(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 $\operatorname{Im}(A)$ 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

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

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

$$\begin{cases} x_1 + 2x_2 + 3x_3 + 4x_4 &= y_1 \\ x_2 + 3x_3 + 2x_4 &= y_2 - y_1 \\ 0 &= y_3 - y_1 - 2y_2 \end{cases}$$

posee soluciones si y sólamente si

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

Obtenemos así que $(\star)$ es una ecuación de $\operatorname{Im}(A)$. 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 $\operatorname{Im}(A)$.

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 $\operatorname{Im}(A)$.

Por ejemplo, el cálculo siguiente

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

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

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

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

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

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

1.7.3. 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 $j = 1, \dots, n-1$, 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 $j = 1, \dots, n-2$, 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 $i = 1, \ldots, n-2$, obtenemos 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

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

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 1.7.4. — 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 $\lambda \neq 0$.

CAPÍTULO 2

DETERMINANTES

En el capítulo anterior, se estudió la relación entre aplicaciones lineales y matrices, concluyendo que toda aplicación lineal tiene siempre una matriz asociada. Sin embargo, dicha matriz está sujeta a la elección de bases para los espacios vectoriales en los cuales está definida dicha aplicación. Es por esto que durante este capítulo el objetivo será encontrar un invariante para las aplicaciones lineales, es decir, un objeto matemático intrínsecamente asociado (no dependiente de la elección de bases) a cada aplicación.

2.1. Permutaciones

Definición 2.1.1. — Sea $n \in \mathbb{N}^{\geq 1}$, y denotemos $I_n = \{1, \ldots, n\}$. Una **permutación** de I_n es una biyección $\sigma : I_n \to I_n$. Usualmente las permutaciones se denotan por

$$\sigma = \begin{pmatrix} 1 & 2 & \ldots & n \\ \sigma(1) & \sigma(2) & \ldots & \sigma(n) \end{pmatrix} \qquad \sigma = (\sigma(1), \ldots, \sigma(n))$$

Dado un $n \in \mathbb{N}^{\geq 1}$, el conjunto de permutaciones de I_n se conoce como el **grupo simétrico** y se denota por

$$S_n = \{\sigma : I_n \to I_n \mid \sigma \text{ es biyectiva}\}$$

Su nombre se debe a que efectivamente S_n es un grupo bajo la composición, ya que la identidad en I_n es claramente biyectiva, la composición de biyecciones es también biyectiva, y su inversa claramente también lo es.

Proposición 2.1.2. — *Sea S_n el grupo simétrico de orden n. Entonces $|S_n| = n!$.*

Demostración. — Probaremos esto por inducción. Si $n = 1$, entonces $S_1 = \{(1)\}$, y se tiene la propiedad. A continuación, suponer que $|S_n| = n!$ para cierto n y consideremos el conjunto S_{n+1}. Sea $k \in I_{n+1}$ y $A_k = \{\sigma = (a_1, \dots, a_{n+1}) \in S_{n+1} \mid a_k = n+1\}$. Ahora, dado que se tiene una biyección $A_k \overset{\sim}{\longleftrightarrow} S_n$ dada explícitamente por

$$\begin{aligned} \sigma &\mapsto (a_1, \dots, a_{k-1}, a_{k+1}, \dots, a_{n+1}) \in S_n \\ (b_1, \dots, b_{k-1}, n+1, b_k, \dots, b_n) &\hookleftarrow (b_1, \dots, b_n) \in S_n \end{aligned}$$

se tiene que $|A_k| = |S_n| = n!$, y entonces

$$|S_{n+1}| = \sum_{k=1}^{n+1} |A_k| = \sum_{k=1}^{n+1} n! = (n+1)!$$

□

***Definición 2.1.3* (Transposición).** — Sea $\tau \in S_n$. Se dice que τ es una **transposición** si únicamente cambia dos elementos de I_n, es decir, existen $i, j \in I_n$ tal que $\tau(i) = j, \tau(j) = i$, y $\tau(k) = k \quad \forall k \neq i, j$.
La transposición que cambia los elementos i, j se denota como $\tau = (i, j)$.
Además, se dice que τ invierte los elementos i, j, si $i > j$ y $\tau(i) < \tau(j)$.

Ejemplo 2.1.4. —

1. La permutación $id = (1, 2, \dots, n) \in S_n$ no posee inversiones.
2. La tranposición $(1, 2) \in S_n$ tiene una única inversión.
3. Las permutaciones $(2, 3, 1), (3, 1, 2) \in S_3$ tiene 2 inversiones.

Definición 2.1.5. — Sea $\sigma \in S_n$. Se define como signatura de σ al número $\epsilon(\sigma) := (-1)^m$ donde m es el número de inversiones de σ. Se dice que σ es **par** si $\epsilon(\sigma) = 1$. En caso contrario se dice que σ es **impar**.

Ejemplo 2.1.6. —

1. La permutación identidad es par, ya que $\epsilon(id) = (-1)^0 = 1$.
2. Si $\sigma \in S_n$ es la transposición (i, j), entonces $\epsilon((i, j)) = (-1)^{2|i-j|-1} = -1$, por lo tanto es impar.
3. En S_3 hay 3 permutaciones pares y 3 impares.

Lema 2.1.7. — *Sea $\sigma \in S_n$, entonces* $\epsilon(\sigma) = \displaystyle\prod_{1 \le i < j \le n} \frac{\sigma(j) - \sigma(i)}{j - i}$

Demostración. — Notar que $j - i > 0$ siempre que $j > i$, y luego el producto $\prod_{i<j} \frac{\sigma(j)-\sigma(i)}{j-i}$ y $\epsilon(\sigma)$ poseen el mismo signo pues $\sigma(j) - \sigma(i) < 0$ si y solo si j, i se encuentran invertidos. Por otra parte

$$\left(\prod_{1\leq i<j\leq n} \frac{\sigma(j)-\sigma(i)}{j-i}\right)^2 = \prod_{1\leq i<j\leq n} \left(\frac{\sigma(j)-\sigma(i)}{j-i}\right)^2 = \prod_{i\neq j} \frac{\sigma(j)-\sigma(i)}{j-i} \overset{(*)}{=} 1$$

$(*)$ σ es una biyección.

Luego el producto del enunciado es exactamente $\epsilon(\sigma)$. □

Proposición 2.1.8. — *La signatura* $\epsilon : S_n \to \{\pm 1\}$ *satisface que* $\epsilon(\sigma\tau) = \epsilon(\sigma)\epsilon(\tau)$ *para todos* $\sigma, \tau \in S_n$.

Demostración. — Para demostrar esta propiedad basta notar que

$$\begin{aligned}
\epsilon(\sigma\tau) &= \prod_{1\leq i<j\leq n} \frac{\sigma(\tau(j))-\sigma(\tau(i))}{j-i} \\
&= \prod_{1\leq i<j\leq n} \frac{\sigma(\tau(j))-\sigma(\tau(i))}{\tau(j)-\tau(i)} \cdot \prod_{1\leq i<j\leq n} \frac{\tau(j)-\tau(i)}{j-i} \\
&= \epsilon(\sigma)\epsilon(\tau)
\end{aligned}$$

□

La proposición anterior motiva el hecho de descomponer permutaciones en otras más simples. De hecho, toda permutación se puede descomponer, de forma no única, como producto de transposiciones. Más aún, toda permutación par (impar) se descompone en un número par (respectivamente impar) de transposiciones.

Ejemplo 2.1.9. — Para descomponer una permutación, la idea es ir cambiando parejas de elementos hasta fijarlos en su posición respectiva. Por ejemplo, consideremos la permutación $(2, 4, 1, 5, 3) \in S_5$, la cual se puede descomponer como:

$$\begin{aligned}
\sigma = (2, 4, 1, 5, 3) &\Rightarrow (1, 2)\sigma = (1, 4, 2, 5, 3) \Rightarrow (2, 4)(1, 2)\sigma = (1, 2, 4, 5, 3) \\
&\Rightarrow (3, 4)(2, 4)(1, 2)\sigma = (4, 5) \\
&\Rightarrow \sigma = (1, 2)^{-1}(2, 4)^{-1}(3, 4)^{-1}(4, 5) \\
&\Rightarrow \sigma = (1, 2)(2, 4)(3, 4)(4, 5)
\end{aligned}$$

Notar además que si denotamos $\sigma = \tau_1 \cdots \tau_n \Rightarrow \epsilon(\sigma) = \epsilon(\tau_1) \cdots \epsilon(\tau_n) = (-1)^n$ y entonces σ y n poseen la misma paridad.

2.2. Formas multilineales alternadas

***Definición 2.2.1* (característica).** — Sea k un cuerpo. Se define la característica de k, denotada por $\mathrm{car}(k)$ como el menor $n \in \mathbb{N}^{\geq 2}$ en k tal que $n \cdot 1 = 1 + \ldots + 1 = 0$. Si no existe dicho n, entonces car(k)=0.

Ejemplo 2.2.2. — Los cuerpos de uso frecuente, vale decir $\mathbb{Q}, \mathbb{R}, \mathbb{C}$, son de característica 0, sin embargo, $\mathrm{car}(\mathbb{F}_p) = p$

En este curso, los cuerpos en los cuales se trabajará serán de $\mathrm{car}(k) \neq 2$.

Definición 2.2.3. — Sea V un k-espacio vectorial. Una **forma m-multilineal** es una función de la forma

$$F : V^m = \underbrace{V \times \cdots \times V}_{\text{n veces}} \to k \qquad (x_1, \ldots, x_m) \mapsto F(x_1, \ldots, x_m)$$

la cual es lineal en cada variable. Además, se dice que F es **alternada** cuando $F(x_1, \ldots, x_m) = 0$ si $x_i = x_j$ para algún par (i, j).

Ejemplo 2.2.4. — Sea $F : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ una forma **bilineal**, y (e_1, e_2) la base canónica de $\mathbb{R}^2$. Entonces:

$$F((x_{11}, x_{21}), (x_{12}, x_{22})) = F(x_{11}e_1 + x_{21}e_2, x_{12}e_1 + x_{22}e_2) =$$
$$x_{11}F(e_1, x_{12}e_1 + x_{22}e_2) + x_{21}F(e_2, x_{12}e_1 + x_{22}e_2) =$$
$$x_{11}(x_{12}F(e_1, e_1) + x_{22}F(e_1, e_2)) + x_{21}(x_{12}F(e_2, e_1) + x_{22}F(e_2, e_2))$$

Ahora, si F es alternada y $(x_{11}, x_{21}) = (x_{12}, x_{22})$ se obtiene que

$$F((x_{11}, x_{21}), (x_{11}, x_{21})) = 0 \Rightarrow F(e_1, e_1) = F(e_2, e_2) = 0 \quad \wedge \quad F(e_1, e_2) = -F(e_2, e_1)$$

De esta forma

$$F((x_{11}, x_{21}), (x_{12}, x_{22})) = F(e_1, e_2)(x_{11}x_{22} - x_{12}x_{21})$$

Proposición 2.2.5. — *Sea V k-espacio vectorial y $F : V^m \to k$ multilineal. Entonces F es alternada si y solo si:*

$$F(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_m) = -F(x_1, \ldots, x_j, \ldots, x_i, \ldots, x_m)$$

para todos $x_1, \ldots, x_m \in V$.

Demostración. — ($\Leftarrow$) Si $x_i = x_j = x$ entonces por hipótesis

$$F(x_1, \ldots, x, \ldots, x, \ldots, x_m) = -F(x_1, \ldots, x, \ldots, x, \ldots, x_m)$$

y considerando que $\mathrm{car}(k) \neq 2$ entonces

$$F(x_1, \ldots, x, \ldots, x, \ldots, x_m) = 0$$

y F es alternada.
($\Rightarrow$) Suponer que F es alternada. Luego

$$F(x_1, \ldots, x_i + x_j, \ldots, x_i + x_j, \ldots, x_m) = 0$$

y en consecuencia

$$F(x_1, \ldots, x_i, \ldots, x_i, \ldots, x_m) + F(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_m) +$$
$$+ F(x_1, \ldots, x_j, \ldots, x_i, \ldots, x_m) + F(x_1, \ldots, x_j, \ldots, x_j, \ldots, x_m) = 0$$

de lo cual se obtiene que $F(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_m) = -F(x_1, \ldots, x_j, \ldots, x_i, \ldots, x_m)$.
$\square$

Una forma de enunciar el resultado anterior, en términos del lenguaje de permutaciones introducido en la sección anterior, es que F es alternada $\Longleftrightarrow$ para toda transposición $\tau \in S_m$ se tiene que

$$F(x_{\tau(1)}, \ldots, x_{\tau(m)}) = -F(x_1, \ldots, x_m)$$

Proposición 2.2.6. — *Sea V un k-espacio vectorial y F multilineal alternada. Entonces:*

1. *Para toda $\sigma \in S_m$ se verifica $F(x_{\sigma(1)}, \ldots, x_{\sigma(m)}) = \epsilon(\sigma) F(x_1, \ldots, x_m)$*
2. *$F(x_1, \ldots, x_m)$ permanece constante si se añade a una variable una combinación lineal de las otras.*
3. *$F(x_1, \ldots, x_m) = 0$ si los x_i son linealmente dependientes. En particular, si $m > \dim_k(V)$ entonces $F = 0$.*

Demostración. — Para probar 1., basta descomponer σ en un producto de transposiciones, es decir, $\sigma = \tau_1 \cdots \tau_k$. Dado que $(-1)^k = \epsilon(\sigma)$ se obtiene 1. por inducción en k.
Para 2., basta observar que

$$F\left(x_1, \ldots, x_i + \sum_{j \neq i} \lambda_j x_j, x_{i+1}, \ldots, x_n)\right) =$$
$$= F(x_1, \ldots, x_m) + \sum_{j \neq i} \lambda_j F(x_1, \ldots, x_j, \ldots, x_j, \ldots, x_m) = F(x_1, \ldots, x_m)$$

Para la propiedad 3., notar que si los x_i son linealmente independientes, alguno de ellos es combinación lineal de las otras variables, es decir, existe x_i tal que

$$x_i = \sum_{j \neq i} \lambda_j x_j$$

y por la propiedad 2., es posible adicionar $-x_i$, obteniendo

$$
\begin{aligned}
F(x_1, \dots, x_i, \dots, x_m) &= F\left((x_1, \dots, x_i - \sum_{j \neq i} \lambda_j x_j, \dots, x_m)\right) \\
&= F(x_1, \dots, 0, \dots, x_m) = 0
\end{aligned}
$$

□

2.3. Determinantes

Durante esta sección el objetivo será definir el invariante mencionado al comienzo del capítulo, el cual es conocido como determinante. Para ello, el primer objetivo es expandir la forma multilineal alternada $F : V^n \to k$ como en el ejemplo 2.2.4.

Comenzamos considerando una base $\mathscr{B} = (e_1, \dots, e_n)$ de V. Entonces, escribiendo los x_i en términos de $\mathscr{B}$, por la mutilinealidad se tiene que:

$$
F(x_1, \dots, x_n) = F\left(\sum_{i_1=1}^{n} x_{i_1 1} e_{i_1}, \sum_{i_2=1}^{n} x_{i_2 2} e_{i_2}, \dots, \sum_{i_n=1}^{n} x_{i_n n} e_{i_n}\right) = \; = \sum_{i_1=1}^{n} \cdots \sum_{i_n=1}^{n} x_{i_1 1} x_{i_2 2} \cdots x_{i_n n} F(e_{i_1}, \dots
$$

Como F es alternada, los términos donde existen índices iguales se anulan. De esta forma, los únicos términos importantes corresponden a los que poseen índices distintos, o dicho de otra forma, aquellos donde la aplicación $k \mapsto i_k$ sea biyectiva, es decir, una permutación. Entonces, por la proposición 2.2.6, se tiene que

$$
\begin{aligned}
F(x_1, \dots, x_n) &= \sum_{\sigma \in S_n} x_{\sigma(1)1} \cdots x_{\sigma(n)n} F(e_{\sigma(1)}, \dots, e_{\sigma(n)}) \\
&= F(e_1, \dots, e_n) \sum_{\sigma \in S_n} \epsilon(\sigma) x_{\sigma(1)1} \cdots x_{\sigma(n)n}
\end{aligned}
$$

Definición 2.3.1. — Sea V un k-espacio vectorial de $\dim_k(V) = n$ y $\mathscr{B} = (e_1, \dots, e_n)$ base de V. Se define el **determinante de los vectores** $x_1, \dots, x_n \in V$ **en la base** $\mathscr{B}$ como

$$
\det_{\mathscr{B}}(x_1, \dots, x_n) = \sum_{\sigma \in S_n} \epsilon(\sigma) x_{\sigma(1)1} \cdots x_{\sigma(n)n}
$$

donde $x_j = \sum_{i=1}^{n} x_{ij} e_i$ para todo $j \in I_n$.

Observación 2.3.2. — Notar que por la definición anterior, para toda forma n-multilineal alternada F, se cumple que

$$F(x_1, \ldots, x_n) = F(e_1, \ldots, e_n) \det_{\mathscr{B}}(x_1, \ldots, x_n)$$

De esta forma, toda forma n-multilineal alternada (con $n = \dim_k(V)$) es proporcional a la función $\det_{\mathscr{B}} : V^n \to k$.

Ejemplo 2.3.3. — Si $\mathscr{B} = ((1,0),(0,1))$ es la base canónica de k^2, entonces

$$\det_{\mathscr{B}}((x_{11}, x_{21}), (x_{12}, x_{22})) = x_{11}x_{22} - x_{21}x_{12}$$

resultado que ya había sido encontrado en el Ejemplo 2.2.4

Proposición 2.3.4. — *Sea $\mathscr{B} = (e_1, \ldots, e_n)$ una base de V k-espacio vectorial de dimensión $\dim_k(V) = n$. La aplicación $\det_{\mathscr{B}} : V^n \to k$ es una forma n-multilineal alternada, y además $\det_{\mathscr{B}}(e_1, \ldots, e_n) = 1$.*

Demostración. — En primer lugar, es posible observar que la cantidad $\epsilon(\sigma)x_{\sigma(1)1} \cdots x_{\sigma(n)n}$ es multilineal para toda $\sigma \in S_n$. Así, la cantidad $\det_{\mathscr{B}}(x_1, \ldots, x_n)$ resulta ser también multilineal.
Veamos que es también alternada. Sea $\tau = (i,j)$ transposición. Luego se observa que

$$\det_{\mathscr{B}}(x_1, \ldots, x_j, \ldots, x_i, \ldots, x_n) = \sum_{\sigma \in S_n} \epsilon(\sigma) x_{\sigma(1)1} \cdots x_{\sigma(i)j} \cdots x_{\sigma(j)i} \cdots x_{\sigma(n)n} =$$

$$= \sum_{\sigma \in S_n} \epsilon(\sigma) x_{\sigma\circ\tau(1)1} \cdots x_{\sigma\circ\tau(j)i} \cdots x_{\sigma\circ\tau(i)j} \cdots x_{\sigma\circ\tau(n)n} =$$

$$= \sum_{\sigma \in S_n} -\epsilon(\sigma\circ\tau) x_{\sigma\circ\tau(1)1} \cdots x_{\sigma\circ\tau(j)i} \cdots x_{\sigma\circ\tau(i)j} \cdots x_{\sigma\circ\tau(n)n} = -\det_{\mathscr{B}}(x_1, \ldots, x_n)$$

y entonces $\det_{\mathscr{B}} : V^n \to k$ alternada. Además, tomando $(x_1, \ldots, x_n) = (e_1, \ldots, e_n)$, la cantidad $x_{\sigma(1)1} \cdots x_{\sigma(n)n} \neq 0 \iff \sigma(i) = i$ para todo $i \in I_n$, o sea, $\sigma = id$. Por lo anterior, $\det_{\mathscr{B}}(e_1, \ldots, e_n) = 1$. □

Teorema 2.3.5. — *Sea V un k-espacio vectorial de $\dim_k(V) = n$ y $\mathscr{B} = (e_1, \ldots, e_n)$ base de V. Entonces:*

1. *Si $\mathscr{B}'$ es base de V entonces*

$$\det_{\mathscr{B}'}(x_1, \ldots, x_n) = \det_{\mathscr{B}'}(e_1, \ldots, e_n) \det_{\mathscr{B}}(x_1, \ldots, x_n)$$

2. $\det_{\mathscr{B}}(x_1, \ldots, x_n) \neq 0 \iff (x_1, \ldots, x_n)$ *es base de V.*

Demostración. — Notar que por la discusión previa toda forma multilineal alternada es proporcional a a su determinante, es decir,

$$F(x_1, \dots, x_n) = F(e_1, \dots, e_n) \det_{\mathscr{B}}(x_1, \dots, x_n)$$

Entonces, en particular tomando $F = \det_{\mathscr{B}'}$ se demuestra la primera proposición.
Por otra parte, si $\mathscr{B}' = (x_1, \dots, x_n)$ es base de V, por la proposición anterior y la proposición 2.3.4 se tiene que

$$\det_{\mathscr{B}'}(e_1, \dots, e_n) \det_{\mathscr{B}}(x_1, \dots, x_n) = \det_{\mathscr{B}'}(x_1, \dots, x_n) = 1 \Rightarrow \det_{\mathscr{B}}(x_1, \dots, x_n) \neq 0$$

Además, dado que $\det_{\mathscr{B}}$ es multilineal alternado, si $\det_{\mathscr{B}}(x_1, \dots, x_n) \neq 0$, por la Proposición 2.2.6(más específicamente, el recíproco de la propiedad 3.) los x_i son linealmente independientes. Luego, $\mathscr{B}$ es una base de V. □

En dimensiones 2 y 3, el determinante tiene una interpretación geométrica muy particular. Si $\mathscr{B}$ es la base canónica de $\mathbb{R}^n$, entonces en particular, para $n = 2$, el determinante de dos vectores (en valor absoluto) corresponde al área encerrada por el paralelogramo generado por dichos vectores. Asimismo, para $n = 3$ el determinante corresponde al volumen del paralelepípedo generado por dichos vectores. De forma más general, en dimensión arbitraria el determinante de una familia de vectores puede ser interpretado como el "volumen" del paralelepípedo n-dimensional generado. Así, cuando dicha familia es linealmente dependiente, $\det_{\mathscr{B}}(x_1, \dots, x_n) = 0$ significa que no posee volumen (el paralelepípedo es plano).

2.3.1. Determinante de un endomorfismo. —

***Definición 2.3.6* (determinante de un endomorfismo)**
Sea V k-espacio vectorial tal que $\dim_k(V) = n$, $\mathscr{B} = (e_1, \dots, e_n)$ base de V y $u \in \mathrm{End}_k(V)$. Se define el determinante de u como

$$\det(u) = \det_{\mathscr{B}}(u(e_1), \dots, u(e_n))$$

En primer lugar, en la definición anterior se debe advertir que $\det(u)$ en primera instancia puede no tener sentido, ya que perfectamente puede depender de la elección de la base. Sin embargo, no es así, hecho que se prueba a continuación.

Teorema 2.3.7. — *Sea* $u \in \mathrm{End}_k(V)$. *Entonces* $\det(u)$ *está bien definido. Más aún, para todos* $x_1, \ldots, x_n \in V$ *y toda base* $\mathscr{B}$ *de* V *se verifica que*

$$\det_{\mathscr{B}}(u(x_1), \ldots, u(x_n)) = \det(u) \det_{\mathscr{B}}(x_1, \ldots, x_n)$$

Demostración. — Debido a que u es lineal y $\det_{\mathscr{B}}$ es multilineal alternada, se tiene que la composición definida por

$$(x_1, \ldots, x_n) \mapsto \det_{\mathscr{B}}(u(x_1), \ldots, u(x_n))$$

es también multilineal alternada. Por ello, dicha composición es proporcional a su determinante, o mejor dicho, se tiene que

$$\det_{\mathscr{B}}(u(x_1), \ldots, u(x_n)) = \det_{\mathscr{B}}(u(e_1), \ldots, u(e_n)) \det_{\mathscr{B}}(x_1, \ldots, x_n)$$

Ahora, solamente resta ver que dicha cantidad es independiente de la base $\mathscr{B}$ escogida. Sea entonces $\mathscr{B}' = (e'_1, \ldots, e'_n)$ otra base de V. Luego gracias al Teorema 2.3.5.1 y la Proposición 2.3.4 se tiene que

$$\det_{\mathscr{B}'}(e'_1, \ldots, e'_n) = \det_{\mathscr{B}'}(e_1, \ldots, e_n) \det_{\mathscr{B}}(e'_1, \ldots, e'_n) = 1$$

Entonces, haciendo uso de las igualdades anteriores, y el hecho de que $\det_{\mathscr{B}'}$ es alternada, se obtiene que

$$\begin{aligned}
\det_{\mathscr{B}'}(u(e'_1), \ldots, u(e'_n)) &= \det_{\mathscr{B}'}(e_1, \ldots, e_n) \det_{\mathscr{B}}(u(e'_1), \ldots, u(e'_n)) \\
&= \det_{\mathscr{B}'}(e_1, \ldots, e_n) \det_{\mathscr{B}}(u(e_1), \ldots, u(e_n)) \det_{\mathscr{B}}(e'_1, \ldots, e'_n) \\
&= \det_{\mathscr{B}}(u(e_1), \ldots, u(e_n))
\end{aligned}$$

Por lo anterior se puede concluir que $\det(u)$ es independiente de la base escogida. Así, el determinante de un endomorfismo está bien definido. □

Ejemplo 2.3.8. — Un endomorfismo de la forma $h_\lambda : V \to V$ tal que $x \mapsto \lambda x$ se conoce como **homotecia**, y λ se conoce como **factor de homotecia**. Sea $\mathscr{B} = (e_1, \ldots, e_n)$ una base de V, y calculemos su determinante:

$$\det(h_\lambda) = \det_{\mathscr{B}}(h_\lambda(e_1), \ldots, h_\lambda(e_n)) = \det_{\mathscr{B}}(_1, \ldots, \lambda e_n) = \lambda^n \det_{\mathscr{B}}(e_1, \ldots, e_n) = \lambda^n$$

En palabras sencillas, si $0 < \lambda < 1$, el efecto de la homotecia es "comprimir" el espacio V en todas direcciones en el factor λ, mientras que si $\lambda > 1$ el espacio se "expande" en todas direcciones. Por lo discutido con anterioridad, el determinante puede ser interpretado como el volumen del paralelepípedo generado por los vectores de la base $\mathscr{B}$. Como la homotecia "estira" dichos vectores en un factor λ, tiene sentido decir que dicho volumen será de λ^n veces el volumen original.

Teorema 2.3.9. — *Sean* $u, v \in \mathrm{End}_k(V)$. *Entonces se tiene que*

1. $\det(u \circ v) = \det(u)\det(v) = \det(v \circ u)$
2. $u \in \mathrm{GL}(V) \iff \det(u) \neq 0$. *Además,* $\det(u^{-1}) = 1/\det(u)$

Demostración. — Para probar la primera propiedad basta notar que dada $\mathscr{B} = (e_1, \ldots, e_n)$ base de V, entonces

$$\det(u \circ v) = \det_{\mathscr{B}}(u(v(e_1)), \ldots, u(v(e_n))) = \det(u)\det_{\mathscr{B}}(v(e_1), \ldots, v(e_n)) = \det(u)\det(v)$$

Notar además, que el hecho de que u sea invertible, es equivalente a que $u(\mathscr{B})$ sea una base de V. Luego por la Proposición 2.3.5, y por la definición de determinante se tiene que

$$\det(u) = \det_{\mathscr{B}}(u(e_1), \ldots, u(e_n)) \neq 0 \iff u(\mathscr{B}) \text{ es base de V}$$

Por último, es claro que, gracias a la primera proposición

$$u \circ u^{-1} = id_V \Rightarrow \det(u)\det(u^{-1}) = \det(id_V) = 1$$

concluyendo así la demostración. □

Es importante notar que la función definida por $\det : \mathrm{End}_k(V) \to k$ no es lineal. Esto debido a que en general $\det(u + v) \neq \det(u) + \det(v)$, y además $\det(\lambda u) = \lambda^n \det(u)$. Por esto, es fundamental hacer el contraste entre el determinante de un conjunto de vectores (el cual es multilineal alternado) y el determinante de un endomorfismo.

Ejercicio 2.3.10. — Sea $m \in \mathbb{N}^{\geq 1}$. Se dice que el endomorfismo $u : V \to V$ tiene orden m si $u^m = \underbrace{u \circ \ldots \circ u}_{\text{m veces}} = \mathrm{id}_V$. Probar que si $u \in \mathrm{End}_k(V)$ es de orden m entonces $\det(u)$ es una raíz $m-$ésima de 1 y por lo tanto $u \in \mathrm{GL}(V)$.

2.3.2. Determinante de una matriz. — A continuación, es de nuestro interés, tal como se ha venido haciendo con los conceptos anteriores, extender la noción de determinante a matrices.

Definición 2.3.11 **(determinante de una matriz)**

Sea $A \in M_n(k)$ una matriz cuadrada. Se define entonces el determinante de la matriz A, denotado por $\det(A)$ como

$$\det(A) := \sum_{\sigma \in S_n} \epsilon(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n}$$

La definición anterior tiene una interpretación clave. El determinante de una matriz corresponde al determinante de sus columnas (vistas como un conjunto de vectores), por lo que la definición resulta ser la misma que la estudiada en la sección anterior. También es posible relacionar esto con el determinante de un endomorfismo notando que el determinante de A es el determinante de cualquier endomorfismo u tal que $\mathrm{Mat}_{\mathscr{B}}(u) = A$ para cierta base $\mathscr{B}$ de V.

Proposición 2.3.12. — *Sea* $A \in M_n(k)$. *Entonces* $\det(A) = \det({}^{\mathrm{t}}A)$.

Demostración. — Denotemos $A = (a_{ij})$ y ${}^{\mathrm{t}}A = (b_{ij})$. En primer lugar, es importante notar que para todo $\sigma \in S_n$ se cumple que $\epsilon(\sigma) = \epsilon(\sigma^{-1})$ (la demostración de este hecho queda propuesta como ejercicio), y por ende se cumple la igualdad

$$\epsilon(\sigma)b_{\sigma(1)1}\cdots b_{\sigma(n)n} = \epsilon(\sigma)a_{1\sigma(1)}\cdots a_{n\sigma(n)} = \epsilon(\sigma^{-1})a_{\sigma^{-1}(1)1}\cdots a_{\sigma^{-1}(n)n}$$

Esto último se puede hacer ya que todo $\sigma \in S_n$ es biyección. Entonces

$$\begin{aligned} \det({}^{t}A) &= \sum_{\sigma \in S_n} \epsilon(\sigma^{-1})a_{\sigma^{-1}(1)1}\cdots a_{\sigma^{-1}(n)n} = \\ &= \sum_{\tau^{-1} \in S_n} \epsilon(\tau)a_{\tau(1)1}\cdots a_{\tau(n)n} = \\ &\overset{(*)}{=} \sum_{\tau \in S_n} \epsilon(\tau)a_{\tau(1)1}\cdots a_{\tau(n)n} = \det(A) \end{aligned}$$

(*) Este paso es posible ya que S_n es un grupo. □

Teorema 2.3.13. — *Sea* $A \in M_n(k)$. *Entonces se verifica que:*

1. $\det(A)$ *depende linealmente de sus filas y columnas.*
2. $\det(A) = 0$ *si sus filas o columnas son linealmente dependientes.*
3. $\det(A)$ *cambia de signo al intercambiar filas o columnas.*

Demostración. — Sea $u : k^n \to k^n$ un endomorfismo asociado a A (vale decir, $Mat_{\mathscr{B}}(u) = A$). Entonces $\det(A) = \det(u)$, y las proposiciones de este teorema se obtienen directamente de las demostradas anteriormente para el determinante de un endomorfismo. Con respecto a las filas, es suficiente hacer uso de la proposición 2.3.12 y aplicar estos resultados a ${}^{\mathrm{t}}A$. □

Es importante notar que las propiedades enunciadas en el teorema anterior nos indican cómo afectan las operaciones fila y operaciones columnas, vistas en el capítulo 1, en el valor del determinante de una matriz.

Teorema 2.3.14. — *Sean* $A, B \in M_n(k)$. *Entonces se cumple que*

1. $\det(AB) = \det(A)\det(B)$.
2. $A \in GL_n(k) \iff \det(A) \neq 0$. *Más aún,* $\det(A^{-1}) = 1/\det(A)$.

Demostración. — Sean $u, v : k^n \to k^n$ los endomorfismos asociados a las matrices A, B. Luego $\det(A) = \det(u), \det(B) = \det(v)$, y entonces es suficiente aplicar el Teorema 2.3.9 para observar que

$$\det(AB) = \det(u \circ v) = \det(u)\det(v) = \det(A)\det(B)$$

El segundo punto viene directamente del hecho que $u \in \mathrm{GL}(V) \iff \det(u) \neq 0$ y $\det(A^{-1}) = \det(u^{-1})$. □

Recuerdo 2.3.15. — Durante el Capítulo 2 se definió el concepto de matriz semejante. Si $A, B \in M_n(k)$ son semejantes entonces $\exists P \in GL_n(k)$ tal que $B = P^{-1}AP$. Esto significa claramente, que dos matrices semejantes representan el mismo endomorfismo en bases diferentes del espacio en cuestión. Dicho de otra forma, la semejanza de matrices implica la existencia de una matriz P de cambio de base entre las matrices A y B.

Corolario 2.3.16. — *Sean $A, B \in M_n(k)$ matrices semejantes. Entonces* $\det(A) = \det(B)$.

Demostración. — Sean $\mathscr{B} = (e_1, \ldots, e_n), \mathscr{B}' = (e'_1, \ldots, e'_n)$ bases del espacio k^n tales que $A = \mathrm{Mat}_{\mathscr{B}}(u)$ y $B = \mathrm{Mat}_{\mathscr{B}'}(u)$ para cierto endomorfismo $u : k^n \to k^n$. Entonces, dado que el determinante de un endomorfismo es independiente de la base, se tiene que

$$\det(u) = \det_{\mathscr{B}}(u(e_1), \ldots, u(e_n)) = \det_{\mathscr{B}'}(u(e'_1), \ldots, u(e'_n))$$

De forma alternativa, se puede notar que si $B = P^{-1}AP$, entonces

$$\det(B) = \det(P^{-1}AP) = \det(P^{-1})\det(A)\det(P) = \frac{1}{\det(P)}\det(A)\det(P) = \det(A)$$

□

Así como en la sección anterior se hizo la acotación acerca de la no-linealidad del determinante de un endomorfismo, es también importante recalcar que para el caso del determinante de una matriz, ocurre lo mismo. Esto es claro pues $\det : M_n(k) \to k$ verifica que $\det(A + B) \neq \det(A) + \det(B)$ y $\det(\lambda A) = \lambda^n \det(A)\ \forall \lambda \in k$.

Ejercicio 2.3.17. — Una matriz A se dice **antisimétrica** si ${}^{t}A = -A$. Sea $A \in M_n(k)$ antisimétrica. Probar que A no es invertible si n es impar.

Ejercicio 2.3.18. — Sea V un espacio vectorial real de dimensión $\dim_{\mathbb{R}}(V) = n$. Un edndomorfismo $J : V \to V$ es una **estructura compleja** si $J^2 = -id_V$. Demostrar que si V tiene una estructura compleja entonces n es par y además $J \in GL(V)$.

2.4. Cálculo de determinantes

Para finalizar este capítulo, se presentarán algunas propiedades sobre los determinantes, que resultan muy útiles para el cálculo de estos, y que resultarán útiles también en las demostraciones posteriores.

Recuerdo 2.4.1. — Para comenzar, se recordarán algunos tipos especiales de matrices, los cuales serán importantes de aquí en adelante. Sea $A \in M_n(k)$. Entonces A se dice

1. **triangular superior** con términos diagonales $\lambda_1, \ldots, \lambda_n$ si es de la forma:
$$A = \begin{pmatrix} \lambda_1 & * & * & * \\ 0 & \lambda_2 & * & * \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & 0 & \lambda_n \end{pmatrix}$$
2. **triangular inferior** con términos diagonales $\lambda_1, \ldots, \lambda_n$ si corresponde a una matriz de la forma:
$$A = \begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ * & \lambda_2 & 0 & 0 \\ * & * & \ddots & \vdots \\ * & * & * & \lambda_n \end{pmatrix}$$
3. **triangular superior(inferior) por bloques** si tiene la forma:
$$A = \begin{pmatrix} A_1 & * & * & * \\ 0 & A_2 & * & * \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & 0 & A_m \end{pmatrix} \quad B = \begin{pmatrix} A_1 & 0 & 0 & 0 \\ * & A_2 & 0 & 0 \\ * & * & \ddots & \vdots \\ * & * & * & A_m \end{pmatrix}$$
donde las matrices $A_i \in M_{n_i}(k)$ y se cumple que $n_1 + \ldots + n_m = n$. A dichas matrices se les llamarán los bloques de A.

Proposición 2.4.2. — *Sea* $A = \begin{pmatrix} A' & B \\ 0 & A'' \end{pmatrix} \in M_n(k)$ *matriz triangular superior por bloques, con* $A' \in M_p(k), A'' \in M_q(k)$. *Entonces,*
$$\det(A) = \det(A')\det(A'')$$

Demostración. — Denotemos $A = (a_{ij})$. Entonces su determinante viene dado por

$$\det(A) = \sum_{\sigma \in S_n} \epsilon(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n}$$

Notar entonces que los términos no-nulos corresponde a aquellos en donde $\sigma(1), \ldots, \sigma(p) \in \{1, \ldots, p\}$ (ie, que pertenezcan a la matriz A'), y entonces necesariamente, $\sigma(p+1), \ldots, \sigma(n) \in$
$\{p+1, \ldots, n\}$. De forma más precisa, los términos no-nulos son aquellos relacionados con $\sigma \in S_n$ para los cuales existen $\sigma' \in S_p, \sigma'' \in S_q$ tales que $\sigma(j) = \sigma'(j)$ si $1 \leq j \leq p$ y $\sigma(p+k) = p + \sigma''(k)$ si $1 \leq k \leq q$. Si denotamos $\sigma = (\sigma', \sigma'')$ (haciendo un abuso de notación), utilizando lo anterior en la definición del determinante se obtiene:

$$\det(A) = \sum_{\substack{\sigma \in S_n \\ \sigma = (\sigma', \sigma'')}} \epsilon(\sigma) a'_{\sigma'(1)1} \cdots a'_{\sigma'(p)p} a''_{\sigma''(1)1} \ldots a''_{\sigma''(q)q}$$

Además, es posible notar, por la construcción de σ y las propiedades de la signatura, que $\epsilon(\sigma) = \epsilon(\sigma')\epsilon(\sigma'')$. Finalmente, se obtiene

$$\det(A) = \left(\sum_{\sigma' \in S_p} \epsilon(\sigma') a'_{\sigma'(1)1} \cdots a'_{\sigma'(p)p} \right) \left(\sum_{\sigma'' \in S_q} \epsilon(\sigma'') a''_{\sigma''(1)1} \ldots a''_{\sigma''(q)q} \right) = \det(A') \det(A'')$$

□

Corolario 2.4.3. — *Sea $A \in M_n(k)$ triangular superior (o inferior) por bloques dados por $A_i \in M_{n_i}(k)$ tales que $n_1 + \ldots + n_m = n$. Entonces,*

$$\det(A) = \det(A_1) \cdots \det(A_m)$$

En particular, si $n_1 = \ldots = n_m = 1$ (ie, A es triangular superior (respectivamente inferior) con términos diagonales $\lambda_1, \ldots, \lambda_n$) entonces $\det(A) = \lambda_1 \cdots \lambda_n$

Demostración. — Se definen las matrices $A' = A_1$ y

$$A'' = \left(\begin{array}{c:c:c} A_2 & * & * \\ \hdashline 0 & \ddots & * \\ \hdashline 0 & 0 & A_m \end{array} \right)$$

Entonces la matriz A adopta la forma

$$A = \begin{pmatrix} A' & B \\ 0 & A'' \end{pmatrix}$$

y por la Proposición 2.4.2, $A = \det(A')\det(A'')$. Luego, utilizando inducción sobre el número de bloques de la matriz, se obtiene que

$$\det(A) = \det(A_1)\cdots\det(A_m)$$

En el caso particular en que A es triangular superior, podemos considerar sus términos diagonales como bloques de tamaño 1×1, y se tiene que $\det(A) = \lambda_1\cdots\lambda_n$.
Finalmente, si A es una matriz triangular inferior por bloques, entonces ${}^{\mathrm{t}}A$ es una matriz triangular superior por bloques, cuyos bloques quedan traspuestos y entonces por la proposición 2.3.12

$$\det(A) = \det({}^{\mathrm{t}}A) = \det({}^{\mathrm{t}}A_1)\cdots\det({}^{\mathrm{t}}A_m) = \det(A_1)\cdots\det(A_m)$$

□

☡**¡Importante!** Se debe advertir que dada una matriz

$$A = \begin{pmatrix} A_1 & B \\ C & A_2 \end{pmatrix} \in M_{2n}(k)$$

con $A_1, A_2, B, C \in M_n(k)$ no es cierto que $\det(A) = \det(A_1)\det(A_2) - \det(B)\det(C)$.

Ejemplo 2.4.4. — Considere el siguiente cálculo de un determinante, utilizando la proposición anterior:

$$\det(A) = \begin{vmatrix} 11 & 0 & 12 & 0 \\ 3 & 1 & 4 & 2 \\ 9 & 0 & 10 & 0 \\ 7 & 5 & 8 & 6 \end{vmatrix} \overset{C_2 \mapsto C_3}{=} - \begin{vmatrix} 11 & 12 & 0 & 0 \\ 3 & 4 & 1 & 2 \\ 9 & 10 & 0 & 0 \\ 7 & 8 & 5 & 6 \end{vmatrix} \overset{F_2 \mapsto F_3}{=} \begin{vmatrix} 11 & 12 & 0 & 0 \\ 9 & 10 & 0 & 0 \\ 3 & 4 & 1 & 2 \\ 7 & 8 & 5 & 6 \end{vmatrix}$$

$$= \begin{vmatrix} 11 & 12 \\ 9 & 10 \end{vmatrix} \cdot \begin{vmatrix} 1 & 2 \\ 5 & 6 \end{vmatrix} = (110-108)(6-10) = -8$$

Definición 2.4.5. — Sea $A = (a_{ij}) \in M_n(k)$ y denotemos por $A_{ij} \in M_{n-1}(k)$ a la matriz cuadrada obtenida de eliminar la i-ésima fila y j-ésima columna de A. Se define entonces

$$C_{ij} = (-1)^{i+j}\det(A_{ij})$$

cantidad a la cual nombraremos **cofactor** de índice (i,j) de la matriz A.

Teorema 2.4.6. — *Sea $A \in M_n(k)$. Entonces $\forall i, j \in \{1,\dots,n\}$ se verifica que:*

1. $\det(A) = a_{1j}C_{1j} + \ldots + a_{nj}C_{nj}$ *(desarrollo de la j-ésima columna)*
2. $\det(A) = a_{i1}C_{i1} + \ldots + a_{in}C_{in}$ *(desarrollo de la i-ésima fila)*

Demostración. — Para esta demostración basta probar únicamente la primera propiedad, ya que la segunda se deriva directamente de esta por el hecho de que $\det(A) = \det({}^{t}A)$. Denotemos por B_{ij} a la matriz obtenida de reemplazar las entradas de la j-esima columna de la matriz A por 0, a excepción de a_{ij}, es decir,

$$B_{ij} = \begin{pmatrix} a_{11} & \cdots & 0 & \cdots & a_{n1} \\ \vdots & & \vdots & & \vdots \\ \vdots & & a_{ij} & & \vdots \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & 0 & \cdots & a_{nn} \end{pmatrix}$$

Si $\mathscr{B} = (e_1, \ldots, e_n)$ es la base canónica de k^n y $x_1, \ldots, x_n \in k^n$ son las columnas de A, entonces $x_j = a_{1j}e_1 + \ldots + a_{nj}e_n$ y se obtiene que

$$\begin{aligned} \det(A) = \det_{\mathscr{B}}(x_1, \ldots, x_n) &= \det(x_1, \ldots, x_{j-1}, a_{1j}e_1 + \ldots + a_{nj}e_n, x_{j+1}, x_n) \\ &\overset{(*)}{=} \det(B_{1j}) + \ldots + \det(B_{nj}) \end{aligned}$$

$(*)$ Debido a la multilinealidad del determinante.
Ahora, el objetivo es calcular los determinantes B_{ij}. Para ello, notar que si cambiamos la j-ésima columna a la primera, sin alterar el orden relativo de las otras columnas, se realizan $j-1$ intercambios de columna, y el determinante se multiplica por un factor de $(-1)^{j-1}$. Asimismo, si se realiza la misma operación con las filas, el determinante se multiplica por un factor de $(-1)^{i-1}$. Por endex

$$\det(B_{ij}) = (-1)^{i+j} \left| \begin{array}{c|ccc} a_{ij} & a_{i1} & \cdots & a_{in} \\ \hline 0 & & & \\ \vdots & & A_{ij} & \\ 0 & & & \end{array} \right| = a_{ij}(-1)^{i+j}\det(A_{ij}) = a_{ij}C_{ij}$$

Reemplazando esta igualdad en la expresión obtenida anteriormente, se obtiene lo pedido. □

El teorema anterior se conoce como desarrollo por cofactores del determinante, y otorga gran libertad en el cálculo de estos, ya que el teorema anterior considera el desarrollo a través de cualquier fila o columna. Notar entonces que la existencia de entradas nulas en una matriz simplifica en gran medida el cálculo de determinantes mediante este método.

Definición 2.4.7. — Sea $A \in M_n(k)$. Se define la **comatriz** de A por $\mathrm{com}(A) := (C_{ij}) \in M_n(k)$, ie, la matriz cuyas entradas son sus cofactores respectivos.

Proposición 2.4.8. — *Sea* $A \in \mathrm{GL}_n(k)$. *Entonces* $A^{-1} = \dfrac{1}{\det(A)}\,{}^{\mathrm{t}}\mathrm{com}(A)$.

Demostración. — Notar que la entrada (i,j) del producto $A\,{}^{\mathrm{t}}\mathrm{com}(A)$ corresponde a $\sum_{k=1}^{n} a_{ik}C_{jk}$, Entonces, si $i = j$, por el Teorema 2.4.6

$$\sum_{k=1}^{n} a_{ik}C_{ik} = a_{i1}C_{i1} + \ldots + a_{in}C_{in} = \det(A)$$

Si $i \neq j$, entonces consideremos la matriz M_{ij} obtenida de copiar la i-ésima fila en la j-ésima fila. Luego, desarrollando la j-ésima fila de dicha matriz, se obtiene:

$$\det(M_{ij}) = m_{i1}C_{j1} + \ldots + m_{in}C_{jn} = (A\,{}^{\mathrm{t}}\mathrm{com}(A))_{ij} = 0$$

Notar que lo anterior es igual a 0 ya que $\det : V^n \to k$ es multilineal <u>alternado</u>, y M_{ij} por definición posee dos filas idénticas. De todo esto se obtiene la identidad

$$A\,{}^{\mathrm{t}}\mathrm{com}(A) = \det(A)I_n$$

y si $A \in \mathrm{GL}_n(k)$, entonces

$$A^{-1} = \frac{1}{\det(A)}\,{}^{\mathrm{t}}\mathrm{com}(A)$$

$\square$

CAPÍTULO 3

REDUCCIÓN DE ENDOMORFISMOS

3.1. Valores y vectores propios

La motivación de esta sección es, dado un endomorfismo $u : V \to V$, encontrar una base $\mathscr{B}$ "lo más sencilla posible", e.g., una base la cual nos permita calcular potencias fáciles de u, hacer cálculos "sencillos" de determinantes, entre otros. En particular, las **matrices diagonales** cumplen con esta propiedad. No obstante, no siempre se puede encontrar una base que cumpla con estas características. Con motivación en base a la idea anterior, se introduce el siguiente lenguaje.

Definición 3.1.1 **(valores y vectores propios).** — Sea $u : V \to V$ un endomorfismo. Si existe $\lambda \in k$ y existe $v \neq 0 \in V$ tal que $u(v) = \lambda v$, diremos que λ es un **valor propio** de u (o eigenvalor) y que v es un **vector propio** de u asociado al escalar λ.

Observación 3.1.2. — Todo múltiplo no nulo de un vector propio es también un vector propio asociado al mismo escalar, ie, los vectores propios **no** son únicos.

Ejemplo 3.1.3. — 1. Para $V = \mathbb{R}^2$ y $r_\pi : V \to V$ rotación en π, $A_\pi = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$, ie, $r_\pi(x, y) = (-x, -y)$. En particular, $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ y $e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ son vectores propios asociados al valor propio $\lambda = -1$.

2. Para $V = \mathbb{R}^2$ y $S_L : V \to V$, $(x, y) \mapsto (-x, y)$ simetría respecto a la recta $L = \mathrm{Vect}_{\mathbb{R}}(e_2)$. En este caso, e_1 es un vector propio asociado a $\lambda_1 = -1$ y e_2 asociado a $\lambda_2 = 1$.
3. Para $V = \mathscr{C}^\infty(\mathbb{R}, \mathbb{R}) = \{f : \mathbb{R} \to \mathbb{R} \mid \text{clase } \mathscr{C}^\infty\}$ y $u : V \to V$ tal que $f \mapsto f'$, la función $f(t) = e^{\lambda t}$ es un vector propio asociado al valor propio λ, pues, $f' = \lambda f$.
4. $V = \mathscr{C}^\infty(\mathbb{R}, \mathbb{R})$ y $u : V \to V$, $f \mapsto f + f''$. Dados $a, b \in \mathbb{R}$ (no ambos nulos), la función $f_{a,b}(t) = a\cos(t) + b\sin(t)$ es un valor propio asociado a $\lambda = 0$, ie, $f_{a,b} \in \ker(u)$. En cursos de cálculo en varias variables se verá que $\ker(u) = \mathrm{Vect}_{\mathbb{R}}\{\cos(t), \sin(t)\} \cong \mathbb{R}^2$.

Terminología Sea $u : V \to V$ un endomorfismo. El conjunto (eventualmente vacío) de valores propios de u es llamado el **espectro** de u y es denotado como $\sigma(u) \subseteq k$, ie:

$$\sigma(u) := \{\lambda \in k \mid \exists v \neq 0, u(v) = \lambda v\}$$

Ejercicio 3.1.4. — Sea $u \in \mathrm{End}_k(V)$ y v un vector propio de u asociado al valor propio $\lambda \in k$. Pruebe que para todo $Q \in k[x]$ se tiene $Q(u)(v) = Q(\lambda)v$.

***Definición 3.1.5* (subespacio propio).** — Sea $u : V \to V$ un endomorfismo y $\lambda \in k$ un valor propio de u. Se define el **subespacio propio** asociado a λ como:

$$V_\lambda := \ker(\lambda id_V - u) \subseteq V.$$

Observación 3.1.6. — Note que

$$u(v) = \lambda v = \lambda id_V(v) \iff (\lambda id_V - u)(v) = 0 \iff v \in V_\lambda$$

En particular, por definición de valor propio se tendrá $V_\lambda \neq \{0\}$.

¡Importante! En todo lo que sigue, supondremos que V es un e.v. de dimensión finita. La teoría espectral de dimensión infinita se estudiará en el curso de Análisis Funcional.

Proposición 3.1.7. — *Sea $u : V \to V$ un endomorfismo y $\lambda \in k$. Entonces, λ es un valor propio de u si y solo si* $\det(\lambda id_V - u) = 0$.

Demostración. — ($\Rightarrow$) Si $\lambda \in \sigma(u)$ entonces existe $v \neq 0 \in V_\lambda$, es decir, $\lambda id_V(v) - u(v) = 0$, luego el $\ker(\lambda id_v - u)$ es no trivial y por tanto $\det(\lambda id_V - u) = 0$.
($\Leftarrow$) Puesto que V es de dimensión finita, $\lambda id_V - u$ es automorfismo $\iff$ es inyectivo. Como $\det(\lambda id_V - u) = 0$, entonces no es inyectivo, es decir, existe $v \neq 0$ tal que $\lambda id_V(v) - u(v) = 0$, ie, $u(v) = \lambda v$. $\square$

***Definición 3.1.8* (suma directa).** — Si $V_1, ..., V_p \subseteq V$ son sub-espacios vectoriales, diremos que V es **suma directa** de $V_1, ..., V_p$ si todo elemento $v \in V$ se puede escribir de manera única como $v_1 + ... + v_p$ con $v_k \in V_k$. En tal caso, escribiremos $V = \bigoplus_{k=1}^{p} V_k$. En otra palabras, la aplicación $g : V_1 \times .... \times V_p \to V$, $(v_1, ..., v_p) \mapsto v_1 + ... + v_p$ es una biyección.
Por otro lado, si la aplicación g es inyectiva, diremos que $V_1, ..., V_p$ **están en suma directa**.

Proposición 3.1.9. — *$V = V_1 \oplus V_2$ si y sólo si se cumplen las siguientes condiciones simultáneamente:*

a) $V = \mathrm{Vect}_k(V_1, V_2)$
b) $V_1 \cap V_2 = \{0\}$

Demostración. — ($\Rightarrow$) Por definición, si $V = V_1 \oplus V_2$ entonces la aplicación lineal $g : V_1 \oplus V_2 \to V, (v_1, v_2) \mapsto v_1 + v_2$ es biyectiva. Puesto que g es sobreyectiva, se tiene que $V = \mathrm{Vect}_k(V_1, V_2)$. Por otro lado, si $V_1 \cap V_2 \neq \{0\}$, entonces existiría al menos un elemento $p_1 \in V_1 \cap V_2$, luego, $g(0, p_1) = g(p_1, 0) = p_1$ lo cual contradice la inyectividad de g.
($\Leftarrow$) Si $V = \mathrm{Vect}_k(V_1, V_2)$ entonces la aplicación lineal $g : V_1 \times V_2, (v_1, v_2) \mapsto v_1 + v_2$ es sobreyectiva. Además, dado $(v_1, v_2) \in V_1 \times V_2$ se tiene que

$$g(v_1, v_2) = v_1 + v_2 = 0 \iff v_1 = -v_2$$

lo cual ocurre sólo si $v_1 = v_2 = 0$, pues, por hipótesis, los subespacios sólo tienen intersección nula, lo cual nos dice que g es inyectiva. Luego g es biyectiva y por tanto $V = V_1 \oplus V_2$. $\square$

Ejercicio 3.1.10. — Para el mismo caso anterior, definimos $V_1 + V_2 := \mathrm{Im}(g)$. Pruebe que $V_1 + V_2 = \mathrm{Vect}_k(V_1, V_2)$ y que $\ker(g) = V_1 \cap V_2$. Deducir (eg, usando el teorema del rango) que

$$\dim_k(V_1 + V_2) = \dim_k(V_1) + \dim_k(V_2) - \dim_k(V_1 \cap V_2).$$

***Definición 3.1.11* (suplementario).** — Sean V_1, V_2 sub-e.v. de V. Diremos que V_1, V_2 son **suplementarios** si $V = V_1 \oplus V_2$.

Lema 3.1.12. — *Todo sub-e.v. de V admite un suplementario.*

Demostración. — Supongamos que $\dim_k(V) = n$ y sea $W \subseteq V$ sub-ev. Si $W = \{0\}$, entonces su suplementario será V. En otro caso, si $\dim_k(W) = p$, tomando una base $(e_1, ..., e_p)$ de W, la podemos completar en una base $(e_1, ..., e_p, e_{p+1}, ..., e_n)$ de V. Definimos $W' := \mathrm{Vect}_k(e_{p+1}, ..., e_n)$. Así, $V = \mathrm{Vect}_k(W, W')$ y $W \cap W' = \{0\}$. $\square$

Ejemplo 3.1.13. — Sea $f : V \to k$ una forma lineal en V. Supongamos que $f \neq 0$ y definamos $H := \ker(f)$. Como $f \neq 0$, existe $v \in V$ tal que $f(v) \neq 0$. En particular, si $f(v) = \lambda$, entonces $f(\frac{v}{\lambda}) = 1$, luego definimos $w := \frac{v}{\lambda}$. Luego, para todo $u \in V$, se tendrá

$$f(u - f(u)w) = f(u) - f(u)\underbrace{f(w)}_{=1} = 0 \Rightarrow u - f(u)w \in H.$$

Note que $u = h + f(u)w$, con $h := u - f(u)w \in H$, es decir, $V = \mathrm{Vect}_k(H, w)$. Por otro lado, si $\lambda w \in H$, entonces $0 = f(\lambda w) = \lambda f(w) = \lambda$, ie, $H \cap \mathrm{Vect}_k(w) = \{0\}$. Concluimos entonces que $V = H \oplus \mathrm{Vect}_k(w)$. Diremos (más adelante) que H es un **hiperplano** en V, pues es un sub-ev de codimensión 1, o en otras palabras, su suplementario es de dimensión 1.

Proposición 3.1.14. — *Sean $V_1, ..., V_p \subseteq V$ tal que $V = V_1 \oplus ... \oplus V_p$. Entonces:*

1. $\dim_k(V) = \dim_k(V_1) + ... + \dim_k(V_p)$.
2. *Si $\mathscr{B}_j$ es una base de V_j, con $j \in \{1, ..., p\}$, entonces $(\mathscr{B}_1, ..., \mathscr{B}_p)$ es una base de V.*

Demostración. — (1) Dado que $g : V_1 \times ... \times V_p \to V$ es un isomorfismo, tenemos

$$\dim_k(V) = \dim_k(V_1 \times ... \times V_p) = \dim_k(V_1) + ... + \dim_k(V_p).$$

(2) $\mathscr{B}$ es generadora, y por (1) tiene el mismo cardinal, $\dim_k(V)$, de una base de V, luego $\mathscr{B}$ es base de V. □

Proposición 3.1.15. — *Sean $V_1, ..., V_p \subseteq V$ sub e.v. tal que $V = V_1 \oplus ... \oplus V_p$ y sea $\mathscr{B}_j$ base de V_j para cada $j \in \{1, ..., p\}$. Sea $u : V \to V$ endomorfismo tal que $u(V_j) \subseteq V_j$ $\forall j \in \{1, ..., p\}$. Probar que*

$$A = Mat_{\mathscr{B}}(u) = \left(\begin{array}{c|c|c} A_1 & 0 & 0 \\ \hline 0 & \ddots & 0 \\ \hline 0 & 0 & A_p \end{array}\right)$$

es una matriz diagonal por bloques, donde $\mathscr{B} = (\mathscr{B}_1, ..., \mathscr{B}_p)$ y donde $A_j \in M_{n_j}(k)$ con $n_j = dim_k(V_j)$

Demostración. — Dado que $u(V_j) \subseteq V_j$, entonces, la imagen bajo u de cada vector de la base, sigue perteneciendo a V_j. Supongamos $\mathscr{B}_j = \{b_1, ..., b_n\}$ es de dimensión n. Notamos que la imagen de cierto vector b_i se puede escribir como combinación lineal de los vectores de $\mathscr{B}_j$, esto es, el resto de coordenadas de los demás $\mathscr{B}_k \subseteq \mathscr{B}$ son 0, por lo cual su bloque asociado sería

$A_i = (u(b_1), \cdots, u(b_n))$. Si hacemos esto con cada una de las bases $\mathscr{B}_j$ se concluye. $\square$

Proposición 3.1.16. — *Los sub-e.v. de un endomorfismo asociado a valores propios distintos están en suma directa.*

Demostración. — Por inducción en el número p de valores propios del endomorfismo. Si $p = 1$ no hay nada que probar. Sean $\lambda_1, ..., \lambda_p$ valores propios de $u : V \to V$ endomorfismo, tales que $\lambda_i \neq \lambda_j$ para $i \neq j$. Queremos probar que la aplicación $g : V_1 \times ... \times V_p \to V$, $(v_1, ..., v_p) \mapsto v_1 + ... + v_p$ es inyectiva, ie,

$$v_1 + \ldots + v_p = 0 \Longrightarrow v_1 = ... = v_p = 0$$

Si aplicamos u a la ecuación, obtendremos $0 = \lambda_1 v_1 + ... + \lambda_p v_p$ y restando $0 = \lambda_1(v_1 + ... + v_p)$ nos quedará

$$0 = (\lambda_2 - \lambda_1)v_2 + ... + (\lambda_p - \lambda_1)v_p$$

Por hipótesis inductiva, $V_{\lambda_2}, ..., V_{\lambda_p}$ están en suma directa, luego $(\lambda_j - \lambda_1)v_j = 0$, para $j = 2, ..., n$. Como $\lambda_1 \neq \lambda_j$, entonces $v_2 = ... = v_p = 0$. Sigue que $v_1 = 0$. $\square$

Observación 3.1.17. — Sean $\lambda_1, ..., \lambda_p$ los valores propios (distintos) de $u : V \to V$. Entonces $V_{\lambda_1} + ... + V_{\lambda_p} := \mathrm{Vect}_k(V_{\lambda_1}, ..., V_{\lambda_p}) = \oplus_{k=1}^{p} V_{\lambda_k}$, pero no necesariamente $V = \oplus_{k=1}^{p} V_{\lambda_k}$. Es decir, $V_{\lambda_1}, ..., V_{\lambda_p}$ **están** en suma directa, pero no necesariamente **son** suma directa.

Corolario 3.1.18. — *El número p de valores propios de $u : V \to V$ es $\leq \dim_k(V)$.*

Demostración. — Sean $\lambda_1, ..., \lambda_p$ los valores propios de u. Luego, $\oplus_{k=1}^{p} V_{\lambda_k} \subseteq V$, es decir

$$\dim_k(V_{\lambda_1}) + ... + \dim_k(V_{\lambda_p}) = \dim_k(\oplus_{k=1}^{p} V_{\lambda_k}) \leq \dim_k(V).$$

Puesto que $\dim_k(V_{\lambda_j}) \geq 1$ (pues existe $v \neq 0$ tal que $v \in V_{\lambda_j}$), se tiene $p \leq \dim_k(V)$. $\square$

3.2. Endomorfismos diagonalizables

Definición 3.2.1 **(endomorfismo diagonalizable)**

Un endomorfismo $u : V \to V$ se dice **diagonalizable** si existe una base de V formada por vectores propios de u.

Observación 3.2.2. — Si $\mathscr{B} = (v_1, ..., v_n)$ es una base de V formada por vectores propios y $\lambda_1, ..., \lambda_n$ son los valores propios asociados (no necesariamente distintos), entonces

$$\mathrm{Mat}_{\mathscr{B}}(u) = \left(\begin{array}{c:c:c} \lambda_1 & 0 & 0 \\ \hdashline 0 & \ddots & 0 \\ \hdashline 0 & 0 & \lambda_n \end{array}\right).$$

Recíprocamente, si existe una base $\mathscr{B}$ tal que $\mathrm{Mat}_{\mathscr{B}}(u)$ es diagonal, entonces necesariamente los vectores de $\mathscr{B}$ son vectores propios (¿Por qué?).

Proposición 3.2.3. — *Sea $u : V \to V$ un endomorfismo y sean $\lambda_1, ..., \lambda_p$ valores propios distintos entre sí. Son equivalentes:*

1. *u es diagonalizable.*
2. $V = V_{\lambda_1} \oplus .... \oplus V_{\lambda_p}$.
3. $\dim_k(V_{\lambda_1}) + ... + \dim_k(V_{\lambda_p}) \geq \dim_k(V)$.

Demostración. — (3) $\Rightarrow$ (2) : En la Proposición 3.1.16 se probó que $g : V_{\lambda_1} \times \ldots \times V_{\lambda_p} \to V$ es inyectiva. El hecho de que $\sum_j \dim_k(V_{\lambda_j}) \geq \dim_k(V)$ implica que g es también sobreyectiva de lo que se concluye (2).
(2) $\Rightarrow$ (1) : Si $\mathscr{B}_j$ es una base de V_{λ_j}, entonces $\mathscr{B} = \{\mathscr{B}_1, ..., \mathscr{B}_p\}$ es base de V formada por vectores propios.
(1) $\Rightarrow$ (3) : Sea $\mathscr{B}$ la base de vectores propios de u y sea n_j el número de vectores de $\mathscr{B}$ asociados a λ_j. Luego $\dim_k(V_{\lambda_j}) \geq n_j$ y por tanto $\dim_k(V_{\lambda_1}) + ... + \dim_k(V_{\lambda_p}) \geq \dim_k(V)$. □

Corolario 3.2.4. — *Supogamos que* $\dim_k(V) = n$ *y sea* $u : V \to V$ *endomorfismo. Si u posee n valores propios distintos, entonces es diagonalizable* [(1)].

Demostración. — En efecto, sean $\lambda_1, ..., \lambda_n$ los valores propios distintos de u, luego $\dim_k(V_{\lambda_j}) \geq 1$ y por tanto $\dim_k(V_{\lambda_1}) + ... + \dim_k(V_{\lambda_n}) \geq \dim_k(V)$. □

Ejercicio 3.2.5. — Sea $u : V \to V$ un endomorfismo diagonalizable. Pruebe que $\lambda \in k$ es el único valor propio de u si y solo si $u = \lambda id_V$.

[(1)] El recíproco no es cierto, si considera la aplicación $u = \lambda id_V$, se tendrá $V_\lambda = V$, pero su único valor propio es λ.

3.3. Polinomio característico

3.3.1. Polinomio característico de una matriz cuadrada. — Si $A = (a_{ij}) \in \mathscr{M}_n(k)$ es una matriz de orden n, entonces $u := u_A : k^n \to k^n$, $e_j \mapsto u(e_j) = \sum_{i=1}^{n} a_{ij}e_i$ es el endomorfismo de k^n asociado a A, con $(e_1, ..., e_n)$ la base canónica de k^n. Esta identificación, como ya se ha hecho anteriormente, nos servirá para extrapolar los conceptos vistos previamente al contexto de matrices, y a través de ellas se construirá un nuevo objeto con el cual se facilitará en gran medida el cálculo de valores propios.

Definición 3.3.1. — Sea $A \in \mathscr{M}_n(k)$. Diremos que $\lambda \in k$ es un **valor propio** de A si lo es para el endomorfismo $u : k^n \to k^n$ asociado a A. En otras palabras, λ es un valor propio de A si y solamente si existe $v \in k^n$ tal que $Av = \lambda v$. Además, diremos que v es un vector propio de A asociado al valor propio λ.

Ejemplo 3.3.2. — Para

$$A = \begin{pmatrix} 1 & 1 \\ 2 & -4 \end{pmatrix} \quad \text{y} \quad v = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

se tiene

$$Av = \begin{pmatrix} 1 & 1 \\ 2 & -4 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 \\ 6 \end{pmatrix} = 3 \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

. Luego v es un vector propio de A asociado al valor propio $\lambda = 3$.

Definición 3.3.3. — Decimos que $A \in \mathscr{M}_n(k)$ es diagonalizable si $u_A : k^n \to k^n$ lo es, es decir, si existe $P \in \mathrm{GL}_n(k)$ tal que $P^{-1}AP$ es una matriz diagonal.

Observación 3.3.4. — En general, todas las propiedades discutidas en la sección anterior se extrapolan para matrices, como por ejemplo, que $\lambda \in k$ es un valor propio de A si y solo si $\det(\lambda I_n - A) = 0$.

Ejemplo 3.3.5. — Sea

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

y $\lambda \in \mathbb{R}$. Luego,

$$\det(\lambda I_n - A) = \begin{vmatrix} \lambda & -1 \\ -1 & \lambda \end{vmatrix} = \lambda^2 - 1 = (\lambda - 1)(\lambda + 1)$$

y así λ es valor propio si y solamente si $(\lambda - 1)(\lambda + 1) = 0$, ie, $\lambda = \pm 1$. Se sigue que A es diagonalizable sobre $\mathbb{R}$, pues posee dos valores propios distintos. Alternativamente, notamos que $u_A(x, y) = (y, x)$, pues

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} y \\ x \end{pmatrix}$$

Geométricamente, A (o siendo estrictos, u_A) es la reflexión respecto a la recta $y = x$ en $\mathbb{R}^2$. Notamos que (geométricamente), $v_1 = (1, 1)$ y $v_2 = (1, -1)$ son vectores propios de A asociados a los valores propios $\lambda_1 = 1$ y $\lambda_2 = -1$ respectivamente. En particular, $\mathscr{B} = (v_1, v_2)$ es una base de $\mathbb{R}^2$, pues $\det(v_1, v_2) \neq 0$.

En el ejemplo anterior encontramos una forma muy útil de encontrar los valores propios de una matriz $A \in \mathscr{M}_n(k)$, donde debíamos resolver una ecuación polinomial. Esto se puede formalizar de la siguiente manera:
A partir del anillo de polinomios $k[X]$ se puede construir un cuerpo de fracciones racionales (véase la construcción de $\mathbb{Q}$ a partir de $\mathbb{Z}$) cuyos elementos son de la forma $\frac{P(X)}{Q(X)}$ con $P, Q \in k[X]$ y donde se define la igualdad

$$\frac{P(X)}{Q(X)} = \frac{R(X)}{S(X)} \iff PS = RQ \in k[X]$$

Dicho cuerpo es llamado **cuerpo de funciones racionales en la variable** X y se denota como $k(X)$. Todo lo que hemos discutido hasta ahora es válido para cualquier cuerpo k, en particular lo es para el cuerpo $K = k(X)$. Note que $k \subseteq k(X)$ es un sub-cuerpo.

Lema 3.3.6. — *Sea $P \in k[X]$ dado por*

$$P(X) = (X - \lambda_1) \cdots (X - \lambda_n) = X^n + a_{n-1}X^{n-1} + ... + a_2 x^2 + a_1 X + a_0$$

Entonces:

$$
\begin{aligned}
& a_{n-1} = -\lambda_1 - \lambda_2 - .... - \lambda_n \\
& a_{n-2} = \lambda_1\lambda_2 + \lambda_1\lambda_3 + ... + \lambda_2\lambda_3 + ... + \lambda_{n-1}\lambda_n = \sum_{1\leq i<j\leq n} \lambda_i\lambda_j \\
& \vdots \\
& a_1 = (-1)^{n-1}(\lambda_2\lambda_3\cdots\lambda_n + \lambda_1\lambda_3\lambda_4\cdots\lambda_n + ... + \lambda_1\lambda_2\cdots\lambda_{n-1}) \\
& \quad = (-1)^{n-1}\sum_{i=1}^{n}\prod_{j\neq i}\lambda_j \\
& a_0 = (-1)^n\lambda_1\lambda_2\cdots\lambda_n.
\end{aligned}
$$

Demostración. — Compare los coeficientes. □

***Definición 3.3.7* (Polinomio característico).** — Sea $A \in \mathscr{M}_n(k)$. Al polinomio $\det(XI_n - A)$ lo llamaremos el **polinomio característico** de A y lo denotaremos por P_A.

Teorema 3.3.8. — *Sea $A \in \mathscr{M}_n(k)$. El polinomio característico P_A es un polinomio unitario (ie, de coeficiente principal 1) y de grado n. Más aún, $P_A(X) = X^n - tr(A)X^{n-1} + ... + (-1)^n \det(A)$, donde $tr(A) = a_{11} + ... + a_{nn}$ es la* **traza** *de A, ie, la suma de sus términos diagonales.*

Demostración. — Escribimos $XI_n - A = b_{ij} \in \mathscr{M}_n(k(X))$, i.e.,

$$
b_{ij} = \begin{cases} -a_{ij} & si \quad i \neq j \\ X - a_{ii} & si \quad i = j \end{cases}
$$

Luego, $P_A(X) = \sum_{\sigma\in S_n} \epsilon(\sigma)b_{1\sigma(1)}\cdots b_{n\sigma(n)}$. Puesto que b_{ij} es un polinomio, $P_A(X)$ lo es. Más aún, el término maximal en X es aquel que corresponde al producto de la diagonal de $XI_n - A$, es decir, para $\sigma = id$, es decir, al término

$$
\epsilon(\sigma)b_{11}\cdots b_{nn} = (X - a_{11})\cdots(X - a_{nn}) = X^n + \ldots
$$

donde se observa que P_A es unitario de grado n. Por otro lado, el coeficiente asociado al monomio X^{n-1} se obtiene al multiplicar $n-1$ términos diagonales, es decir, para los términos en la sumatoria de la forma:

$$
\epsilon(\sigma)b_{11}\cdots b_{i-1,i-1}b_{i\sigma(i)}b_{i+1,i+1}\cdots b_{nn}.
$$

Puesto que $\sigma(j) = j$ para $j \neq i$, y que $\sigma \in S_n$ es biyección, necesariamente $\sigma(i) = i$, ie, $\sigma = id$, se sigue entonces que

$$\begin{aligned}
&\epsilon(\sigma) b_{11} \cdots b_{i-1,i-1} b_{i\sigma(i)} b_{i+1,i+1} \cdots b_{nn} = (X - a_{11}) \cdots (X - a_{nn}) \\
\underset{\text{Lema 3.3.6}}{=} \quad & X^n - \underbrace{(a_{11} + a_{22} + ... + a_{nn})}_{tr(A)} X^{n-1} + ... + (-1)^n a_{11} \cdots a_{nn}.
\end{aligned}$$

Finalmente, si escribimos $P_A(X) = X^n + a_{n-1}X^{n-1} + ... + a_1 X + a_0$, notamos que $a_0 = P_A(0) = \det(0Id_n - A) = (-1)^n \det(A)$. $\square$

Ejemplo 3.3.9. —

1. Sea
$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \in \mathscr{M}_n(k)$$
reflexión respecto a la recta $y = x \in k^2$. Luego,
$$P_A(X) = \det(XI_n - A) = \begin{vmatrix} X & -1 \\ -1 & X \end{vmatrix} = X^2 - 1 = (X-1)(X+1)$$
tiene raices $\lambda_1 = 1$ y $\lambda_2 = -1$ sobre todo cuerpo k. Si $\mathrm{car}(k) = 2$, entonces $\lambda_1 = \lambda_2$, pero si $\mathrm{car}(k) \neq 2$, entonces $\lambda_1 \neq \lambda_2$ y A es diagonalizable sobre k.
2. Sea $r_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ rotación en un ángulo $\theta \in \mathbb{R}$, cuya matriz asociada a la base canónica es
$$A_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \in \mathscr{M}_2(\mathbb{R})$$
Se sigue que
$$P_{A_\theta} = \begin{vmatrix} X - \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & X - \cos(\theta) \end{vmatrix} = (X^2 - \cos(\theta))^2 + \sin^2(\theta).$$
Entonces, $\triangle = b^2 - 4ac = 4\cos^2(\theta) - 4 = -4\sin^2(\theta) \leq 0$, donde observamos que P_{A_θ} posee raices reales si y solamente si $\sin(\theta) = 0$, ie $\theta \equiv 0 \pmod{\pi}$. No obstante, si pensamos $A_\theta = \mathscr{M}_2(\mathbb{C})$ (ie, con coeficientes en $\mathbb{C}$), observamos que las raices de P_{A_θ} estarán dadas por
$$\frac{-b \pm \sqrt{\triangle}}{2a} = \frac{2\cos(\theta) \pm 2i\sin(\theta)}{2} = e^{\pm i\theta}$$

ie, $\lambda_1 = e^{i\theta}$ y $\lambda_2 = e^{-i\theta}$ son los valores propios. Observe que dos vectores propios de A_θ son $v_1 = \begin{pmatrix} 1 \\ i \end{pmatrix}$ y $v_2 = \begin{pmatrix} 1 \\ -i \end{pmatrix}$ (ejercicio), y luego A_θ diagonalizable sobre $\mathbb{C}$.

3. Si $A = (a_{ij}) \in \mathscr{M}_n(k)$ es una matriz triangular (superior o inferior), entonces $XI_n - A$ también lo es, luego $P_A(X) = (X - a_{11}) \cdots (X - a_{nn})$ y por tanto sus valores propios son sus términos diagonales. En particular, si todos los términos de la diagonal son distintos, entonces A es diagonalizable.

Recordando resultados vistos en un primer curso de álgebra (cf MAT021 o bien MAT060), enunciaremos el siguiente teorema, el cual es demostrado en el apéndice de este apunte:

Teorema fundamental del Álgebra(Argand, 1806): Todo polinomio $P \in \mathbb{C}[X]$ con coeficientes complejos de grado ≥ 1 posee al menos una raíz en $\mathbb{C}$.

Corolario 3.3.10. — *Toda matriz compleja $A \in \mathscr{M}_n(\mathbb{C})$ posee al menos un valor propio en $\mathbb{C}$.*

Demostración. — $P_A(X) = \det(XI_n - A) \in \mathbb{C}[X]$ es de grado ≥ 1, luego posee una raiz, ie, existe $\lambda \in \mathbb{C}$ tal que $P_A(\lambda) = 0$. Se sigue que λ es valor propio de A. $\square$

Observación 3.3.11. — Si k es un cuerpo arbitrario, el Teorema de Steinitz (1910), afirma que existe un cuerpo $\overline{k}$ tal que $k \subseteq \overline{k}$ y tal que todo $P \in \overline{k}[X]$ no constante posee una raíz en $\overline{k}$. A $\overline{k}$ se le conoce como la **clausura algebraica** de k.

3.3.2. Polinomio característico de un endomorfismo. — Sea $u : V \to V$ un endomorfismo y sea $\mathscr{B}$ una base de V. Si denotamos $A = \mathrm{Mat}_{\mathscr{B}}(u) \in \mathscr{M}_n(k)$, entonces $\det(\lambda id_V - u) = \det(\lambda I_n - A)$ para todo $\lambda \in k$. En particular, los valores propios de u son los valores propios de A respecto a cualquier base.
Más aún, sabemos que A, B son matrices semejantes si y solo si representan un mismo endomorfismo (ie, existe $u : V \to V$ y bases $\mathscr{B}$, $\mathscr{C}$ tales que $A = \mathrm{Mat}_{\mathscr{B}}(u)$ y $B = \mathrm{Mat}_{\mathscr{C}}(u)$), es decir, A y B tienen los mismos valores propios. El resultado siguiente generalizará lo recién mencionado:

Teorema 3.3.12. — *Matrices semejantes tienen el mismo polinomio característico.*

Demostración. — Basta probar que si $A, B \in \mathscr{M}_n(k)$ son semejantes, entonces lo son $XI_n - A$ y $XI_n - B$ (¿Por qué?). En efecto, si A y B son semejantes, entonces existe $P \in \mathrm{GL}_n(k)$ tal que $B = P^{-1}AP$, se sigue que

$$P^{-1}(XI_n - A)P = P^{-1}XI_nP - P^{-1}AP = XP^{-1}I_nP - P^{-1}AP = XI_n - B.$$

□

Corolario 3.3.13. — *Sea $u : V \to V$ un endomorfismo, $\mathscr{B}$ una base de V y $A = (a_{ij}) = \mathrm{Mat}_{\mathscr{B}}(u)$. Entonces, el polinomio característico de u, dado por $P_u(X) = X^n - \mathrm{tr}(u)X^{n-1} + ... + (-1)^n \det(U)$ no depende de la base $\mathscr{B}$, es decir, está bien definido. En particular, $\mathrm{tr}(u)$ no depende de la base.*

Demostración. — Es directo. □

Ejercicio 3.3.14. — Sean $A, B \in \mathscr{M}_n(k)$. Pruebe que $\mathrm{tr}(AB) = \mathrm{tr}(BA)$.

3.4. División de polinomios

Los polinomios probarán ser de gran utilidad en las discusiones que están por venir dentro del capítulo. Por tal razón, en esta sección se recordarán algunas de sus propiedades y se demostrarán lemas y resultados de importancia que aparecerán nuevamente en las demostraciones de algunos de los teoremas que más adelante estudiaremos.

Recuerdo 3.4.1. — Sean $A, B \in k[X]$ polinomios. Decimos que B **divide** a A (o que A es **múltiplo** de B) si existe $Q \in k[X]$ tal que $A = QB$.

***Teorema 3.4.2* (División euclideana).** — *Sean $A, B \in k[X]$. Si $B \neq 0$, entonces existen polinomios Q (cociente) y R (resto) **únicos** tales que*

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

Demostración. — Existencia: Usaremos inducción en $n := gr(A)$. Si $A = 0$ basta tomar $Q = R = 0$. Si $n < gr(B)$ consideramos $Q = 0$ y $R = A$. En caso contrario, definimos $m := gr(B) \leq n$ y escribimos $A = a_nX^n + A_1(X)$ y $B = b_mX^m + B_1(X)$ con $a_n \neq 0$, $b_m \neq 0$ y $gr(A_1) < n$, $gr(B_1) < m$. Notemos que el polinomio $A - \frac{a_n}{b_m}X^{n-m}B = A_1 - \frac{a_n}{b_m}X^{n-m}B_1$ es de grado menor a n.

Luego, utilizando la hipótesis de indución, podemos encontrar dos polinomios Q y R tales que

$$A - (\frac{a_n}{b_m})X^{n-m}B = QB + R \qquad \text{con } gr(R) < gr(Q)$$

y escribir $A = (\frac{a_n}{b_m}X^{n-m} + Q)B + R$, que es lo que se quería demostrar.

Unicidad: Si $A = BQ + R = BQ' + R'$, tenemos que $R - R' = B(Q' - Q)$ y por tanto B divide a $R - R'$, pero como $gr(R - R') < gr(B)$ necesariamente $R - R' = 0$. Finalmente, $B \neq 0 \Rightarrow Q' - Q = 0$ y por ende $R = R'$ y $Q = Q'$. □

Ejemplo 3.4.3. — La división euclideana de $A(X) = X^3 + X + 1$ por $B(X) = X + 1$:

$$\begin{array}{l} X^3 + X + 1 \ : \ (X+1) \ = \ X^2 - X + 2 \\ \underline{-(X^3 + X^2)} \\ -X^2 + X + 1 \\ \underline{-(-X^2 - X)} \\ 2X + 1 \\ \underline{-(2X + 2)} \\ \qquad -1 \end{array}$$

$\Rightarrow A(X) = (X^2 - X + 2)B(X) + (-1)$

Corolario 3.4.4. — *Sea $P \in k[X]$. Entonces, $a \in k$ es una raíz de P (i.e., $P(a) = 0$) si y solo si P es divisible por $(X - a)$.*

Demostración. — La división euclideana de P por $(X-a)$ nos da $P = (X-a)Q+R$, donde $gr(R) < 1$, ie, $R \in k$. Al evaluar en a obtenemos $P(a) = 0 + R = R$, de donde se concluye el resultado. □

Recuerdo 3.4.5. — La **multiplicidad de la raíz** a de un polinomio P es el mayor $m \in \mathbb{N}$ tal que $(X - a)^m$ divide a P. Si $m = 1$ (resp. 2, resp 3, etc...) decimos que a es una **raíz simple** (resp. doble, resp. triple, etc...).

Ejemplo 3.4.6. — Sea $P \in \mathbb{C}[X]$ dado por $P(X) = X^n - 1$ para $n \in \mathbb{N}^{\geq 1}$. Entonces

$$P(X) = (X-1)(X-\omega)(X-\omega^2)\dots(X-\omega^{n-1}) \qquad \text{con } \omega = e^{\frac{2\pi i}{n}} \quad \text{(raíces simples)}$$

Ejercicio 3.4.7. — Sea $P \in \mathbb{C}[X]$ y $a \in \mathbb{C}$. Probar que a es una raíz **múltiple** de P (i.e., de multiplicidad mayor a 1) si y solo si $P(a) = P'(a) = 0$ (derivada).

A continuación, se introduce la definición de **ideal**, concepto clave en teoría de anillos, para el caso del anillo de polinomios $k[X]$, el cual resultará de gran utilidad en resultados próximos.

Definición 3.4.8. — Sea $I \subseteq k[X]$. Decimos que I es un **ideal** de $k[X]$ si:

1. I es un sub-e.v. de $k[X]$.
2. para todo $P \in I$ y para todo $Q \in k[X]$ se tiene que $PQ \in I$.

Ejemplo 3.4.9. — Sean $P_1, \ldots, P_r \in k[X]$ y sea $I = \{P_1Q_1 + \cdots + P_rQ_r : \ Q_i \in k[X]\}$. El conjunto I es el ideal más pequeño que contiene a $P_1, \ldots, P_r$ y es llamado el **ideal generado** por los polinomios $P_1, \ldots, P_r$.

Definición 3.4.10. — Decimos que un ideal I es un **ideal principal** si existe $P \in k[X]$ tal que $I = \langle P \rangle$ i.e., si I está formado por todos los múltiplos de P.

Teorema 3.4.11. — *Todo ideal I de $k[X]$ es principal. Más aún, si $I \neq \{0\}$ es un ideal, entonces existe un único polinomio unitario $P \in I$ tal que $I = \langle P \rangle$.*

Demostración. — $I = \{0\}$ es generado por el polinomio nulo. Supongamos que $I \neq \{0\}$. En este caso, el conjunto

$$\{gr(Q), \ Q \in I - \{0\}\} \subseteq \mathbb{N}$$

es un sub-conjunto no vacío de $\mathbb{N}$ y, como tal, posee un elemento minimal d. Sea $P \in I - \{0\}$ con $gr(P) = d$. Podemos asumir que P es unitario (si no, basta multiplicarlo por el inverso de su coeficiente principal) Vamos que $I = \langle P \rangle$. Sea $F \in I$ arbitrario. La división euclideana de F por P nos da $F = PQ + R$, con $gr(R) < gr(P) = d$. Como I es un ideal, $R \in I$ ya que es la resta de dos elementos en I. Además, la minimalidad de d nos permite deducir que $R = 0$. Así, $F = PQ \in \langle P \rangle$ y por lo tanto $I = \langle P \rangle$.
Ahora veamos que P es único. Si $I = \langle P \rangle = \langle P' \rangle$ con P' unitario, entonces existen $Q, Q' \in k[X]$ tales que $P' = PQ$ (pues $I = \langle P \rangle$) y $P = P'Q'$ (pues $I = \langle P' \rangle$). Pero de esto se sigue que $gr(Q) = gr(Q') = 0$, i.e., $Q, Q' \in k$. Como P y P' son unitarios, la igualdad $P' = PQ$ implica que $Q = 1$ y por lo tanto que $P = P'$. □

Definición 3.4.12. — Sean $A, B \in k[X]$ polinomios no nulos. Decimos que $D \in k[X]$ es un **máximo común divisor** (mcd) de A y B si D divide a A y a B y si todo divisor común de A y de B divide a D. Diremos que A y B son **primos entre sí** si $1 \in k$ es un mcd de A y B.

Observación importante 3.4.13. — El mcd **no** es único: si D es un mcd de A y B, entonces λD también lo es para todo escalar $\lambda \neq 0$. Por maximalidad, todos los mcd (si existen) se obtienen de este modo y, en particular, existe un único mcd de A y B que es **unitario**. De existir, denotamos a este polinomio como $\mathrm{mcd}(A, B)$.

Proposición 3.4.14. — *Sean $A, B \in k[X] \setminus \{0\}$ no nulos. Entonces, existe $D \in k[X]$ mcd de A y B. Más aún, existen $U, V \in k[X]$ tales que D se escribe como $D = AU + BV$. En particular, existe un único* $\mathrm{mcd}(A,B) \in k[X]$.

Demostración. — El conjunto $I = \langle A, B\rangle = \{AP + BQ : \ P, Q \in k[X]\}$ es un ideal distinto de $\{0\}$. Por lo tanto, existe un único polinomio unitario D tal que $I = \langle D\rangle$. En particular, como $D \in I = \langle A, B\rangle$, podemos escribir $D = AU + BV$ para algunos $U, V \in k[X]$, y asimismo, como $A, B \in I = \langle D\rangle$ son ambos múltiplos de D. Escribiendo a D de esta forma podemos ver que cualquier polinomio que divida tanto a A como a B también divide a D, de donde D es un mcd de A y B. $\square$

Teorema 3.4.15 **(Teorema de Bézout).** — *Sean $A, B \in k[X]$ no nulos. Entonces, A y B son primos entre sí si y solo si $\exists U, V \in k[X]$ tales que $AU + BV = 1$.*

Demostración. — A y B son primos si y solo si $\mathrm{mcd}(A, B) = 1$. Gracias a la proposición anterior, podemos asegurar la existencia de $U, V \in k[X]$ tales que $AU + BV = 1$.
Recíprocamente, si $AU + BV = 1$, entonces todo mcd de A y B divide a 1 y es por lo tanto constante. $\square$

Teorema 3.4.16 **(Lema de Gauss).** — *Sean $A, B, C \in k[X]$ polinomios no nulos. Entonces:*

1. $\mathrm{mcd}(A, B) = 1$ *y A divide a $BC \Rightarrow A$ divide a C.*
2. $\mathrm{mcd}(A, B) = 1$ *y* $\mathrm{mcd}(A, C) = 1 \Rightarrow \mathrm{mcd}(A, BC) = 1$.

Demostración. — Para (1), notamos que como $\mathrm{mcd}(A, B) = 1$, por el Teorema de Bézout $\exists U, V \in k[X]$ tales que $AU + BV = 1$. Como A divide a BC, existe $P \in k[X]$ tal que $BC = AP$. Así:

$$C = (AU + BV)C = ACU + BCV = ACU + APV = A\underbrace{(CU + PV)}_{\in k[X]}$$

de donde A divide a C.
(2) Sea $D = \mathrm{mcd}(A, BC)$. Notar que $\mathrm{mcd}(B, D) = 1$, pues todo divisor

común de D y B divide A y B, y estos últimos son primos entre sí. Como $\mathrm{mcd}(B, D) = 1$ y D divide a BC por (1), D divide a C. Luego, D divide A y D divide C, de donde $D = 1$. □

Corolario 3.4.17. — *Sean $P, P_1, \ldots, P_m \in k[X]$ polinomios no nulos. Entonces:*

1. *Si* $\mathrm{mcd}(P_i, P_j) = 1$ *para todo $i \neq j$ y si P_i divide a P para todo i, entonces el producto $P_1 \cdots P_m$ divide a P.*
2. *Si* $\mathrm{mcd}(P, P_i) = 1$ *para todo i, entonces* $\mathrm{mcd}(P, P_1 \cdots P_m) = 1$.

Demostración. — Las demostraciones de ambos puntos son solamente extender el lema de Gauss usando inducción sobre m. □

El lema de Gauss extendido nos permite discutir la factorización de polinomios. Ya vimos que para $P \in k[X]$ que $a \in k$ sea una raíz de P equivale a que $(X-a)$ divida a P. Ahora, si a *no* es una raíz de P entonces $\mathrm{mcd}(P, X-a) = 1$ pues el grado de $(X - a)$ es 1 y, usando el corolario anterior (2), llegamos a que P y $(X - a)^n$ son primos entre sí para todo $n \in \mathbb{N}^{\geq 1}$. Por otro lado, si $\lambda_1, \ldots, \lambda_p \in k$ son raíces distintas de P con multiplicidades $m_1, \ldots, m_p$ respectivamente, entonces haciendo uso del corolario anterior (1) podemos concluir que P es divisible por el producto

$$\prod_{j=1}^{p}(X - \lambda_j)^{m_j} = (X - \lambda_1)^{m_1}(X - \lambda_2)^{m_2} \cdots (X - \lambda_j)^{m_p}$$

Definición 3.4.18. — Sea $P \in k[X]$ un polinomio no nulo. Decimos que P **escinde sobre** k si se factoriza (sobre k) como un producto de polinomios de grado 1:

$$P(X) = a\prod_{j=1}^{p}(X - \lambda_j)^{m_j} = a(X - \lambda_1)^{m_1} \ldots (X - \lambda_j)^{m_p}$$

donde $a \in k$ y $\lambda_1, \ldots, \lambda_p \in k$. En otras palabras, P escinde sobre k si y solo si la suma $S(P) = m_1 + \ldots m_p$ de las multiplicidades de sus raíces en k es igual al grado de P.

La palabra escindir significa cortar, dividir, separar y es traducción de la palabra en inglés *split*. Más adelante trabajaremos extensamente con polinomios que escinden. Por eso, introducimos aquí el siguiente lema, que nos da un criterio para averiguar si un polinomio escinde.

***Lema 3.4.19* (Lema útil). —** *Sean $P, Q \in k[X]$ no nulos. Si P escinde sobre k y Q divide a P, entonces Q escinde sobre k.*

Demostración. — Escribiendo $P = AQ$, vemos que $\mathrm{gr}(P) = S(P) = S(A)+S(Q)$. Como $\mathrm{gr}(P) = \mathrm{gr}(A) + \mathrm{gr}(Q)$ También, siempre se tiene que $S(A) \leq \mathrm{gr}(A)$ y $S(Q) \leq \mathrm{gr}(Q)$. Sin embargo, $P = AQ \Longrightarrow \mathrm{gr}(P) = S(P) = S(A) + S(Q)$, de donde se deduce que necesariamente $S(Q) = \mathrm{gr}(Q)$ y que, por lo tanto, Q escinde sobre k. □

Este nuevo término nos permite refrasear el Teorema Fundamental del Álgebra de la siguiente manera:

Teorema Fundamental del Álgebra: Todo polinomio no nulo con coeficientes complejos $P \in \mathbb{C}[X]$ escinde sobre $\mathbb{C}$.

Esto **no** es cierto para otros cuerpos. Por ejemplo:

1. $P \in \mathbb{R}[X]$ dado por $P = X^3 - 1 = (X-1)(X^2+X+1)$ no escinde sobre $\mathbb{R}$ ya que $X^2 + X + 1$ no tiene raíces *reales.*
2. Pequeño teorema de Fermat (1636): Sea $a \in \mathbb{N}$ y sea p un número primo. Entonces
$$a^p \equiv a \pmod p$$
Como consecuencia, el polinomio $P \in \mathbb{F}_p[X]$ dado por $P(X) = X^p - X + 1$ no posee raíces en $\mathbb{F}_p$ ya que $P(a) = 1$ para todo $a \in \mathbb{F}_p$. Luego, P no escinde sobre $\mathbb{F}_p$.

3.5. Trigonalización de endomorfismos y matrices

En secciones anteriores estudiamos el caso de endomorfismos que son representables mediante una matriz diagonal (endomorfismos diagonalizables). Las matrices diagonales son muy cómodas de trabajar, pero no todos los endomorfismos son diagonalizables. Por eso, en esta sección y en las que siguen estudiaremos diferentes representaciones matriciales que tienen los endomorfismos. Estableceremos cuáles son estas representaciones, qué endomorfismos las poseen y cómo encontrar las bases adecuadas para utilizarlas.

***Definición 3.5.1.* —** Sea $u : V \to V$ un endomorfismo donde $\dim_k(V) = n$. Decimos que u es **trigonalizable** si existe una base $\mathscr{B}$ de V en la cual la matriz $A = \mathrm{Mat}_{\mathscr{B}}(u) \in M_n(k)$ es triangular superior. Diremos que una matriz $A \in M_n(k)$ es trigonalizable si la aplicación lineal asociada $u_A : k^n \to k^n$, $x \mapsto Ax$ lo es, i.e., si existe $P \in \mathrm{GL}_n(k)$ tal que $P^{-1}AP$ es triangular superior.

Ejercicio 3.5.2. — Sea $u : V \to V$ un endomorfismo y $\mathscr{B} = (e_1, \dots, e_n)$ una base de V. Probar que $\mathrm{Mat}_{\mathscr{B}}(u)$ es triangular superior $\iff$ $\mathrm{Mat}_{\mathscr{C}}(u)$ es triangular inferior, donde $\mathscr{C} = (e_n, e_{n-1}, \dots, e_1)$.

El ejercicio anterior significa que podemos reemplazar sin ningún problema el término "superior" por "inferior" en la definición anterior.

Muchos cálculos resultan más sencillos en matrices triangulares superiores, como se observa en el siguiente ejemplo.

Ejemplo 3.5.3. — Sea u un endomorfismo trigonalizable y sea $\mathscr{B}$ una base de V tal que la matriz $A = \mathrm{Mat}_{\mathscr{B}}(u)$ es triangular superior, i.e.,

$$A = \begin{pmatrix} a_{11} & * & * & * & * \\ & a_{22} & * & * & * \\ & & \ddots & * & * \\ & 0 & & \ddots & * \\ & & & & a_{nn} \end{pmatrix} \tag{1}$$

Luego $P_A(X) = \det(XI_n - A) = (X - a_{11})(X - a_{22}) \dots (X - a_{nn})$ y los valores propios de A (o, equivalentemente, de u) son los términos en la diagonal. En particular, el polinomio característico de u **escinde** sobre k.

El ejemplo anterior motiva el teorema presentado a continuación, el cual caracteriza los endomorfismos trigonalizables.

Teorema 3.5.4. — *Sea* $\dim_k(V) = n$ *y sea* $u : V \to V$ *un endomorfismo. Entonces,* u *es trigonalizable si y solo si su polinomio característico* $P_u \in k[X]$ *escinde sobre* k.

Demostración. — El ejemplo anterior demuestra una de las implicancias. Para probar la otra (escinde $\Rightarrow$ trigonalizable) procederemos por inducción en la dimensión de V. Si $n = 1$, toda matriz es triangular superior, lo que prueba el caso base.
Ahora consideremos $n \geqslant 2$. La estrategia será construir un subespacio de V de dimensión $n - 1$ estable bajo u en el cual aplicar la hipótesis de inducción. Como el polinomio cracterístico de u escinde, podemos escribir

$$P_u(X) = \prod_{j=1}^{n} (X - \lambda_j) \qquad \lambda_1, \dots \lambda_n \in k$$

donde $\lambda_1, \dots \lambda_n$ son valores propios de u (no necesariamente distintos). Como $\lambda_n \in k$ es un valor propio, el endomorfismo $u - \lambda_n id_V$ no es inyectivo y, al ser V de dimensión finita, tampoco es sobreyectivo. Esto implica que el espacio vectorial $W := \mathrm{Im}(u - \lambda_n id_V) \subseteq V$ es tal que $r := \dim_k(W) < n$. Si el espacio propio de λ_n tuviese dimensión 1 podríamos aplicar la hipótesis de inducción directamente en W porque este tendría dimensión $n-1$ y su estabilidad bajo $u - \lambda_n id_V$ permite demostrar fácilmente su estabilidad bajo u. Este no siempre es el caso y en general los siguientes pasos son necesarios:
Sea $(e'_1, \dots, e'_r)$ una base de W que completamos en una base $\mathscr{B}' = (e'_1, \dots, e'_n)$ de V. Definamos $U := \mathrm{Vect}_k(e'_1, \dots, e'_{n-1})$. Notar que $W \subseteq U$ (y que $\dim_k(W) = n-1 \Rightarrow U = W$). Veamos que U es estable. Para todo $j \in \{1, \dots, n-1\}$ tenemos que

$$u(e'_j) = u(e'_j) - \lambda_n e'_j + \lambda_n e'_j = \underbrace{(u - \lambda_n id_V)(e'_j)}_{\in W \subseteq U} + \underbrace{\lambda_n e'_j}_{\in U}$$

Luego, $u(e'_j) \in U$ al ser la suma de elementos en U y se concluye la estabilidad de este conjunto bajo u. Adicionalmente, notemos que podemos escribir $u(e'_n) = (u - \lambda_n id_V)(e'_n) + \lambda_n e'_n$ por lo que la última columna (la que coresponde al vector $u(e'_n)$) de la matriz $\mathrm{Mat}_{\mathscr{B}'}(u)$ termina en λ_n. Esto nos permite escribir

$$\mathrm{Mat}_{\mathscr{B}'}(u) = \left(\begin{array}{ccc:c} & & & * \\ & \mathrm{A} & & \vdots \\ & & & * \\ \hdashline 0 & \cdots & 0 & \lambda_n \end{array}\right)$$

Donde $A \in \mathscr{M}_{n-1}(k)$ es una matriz cuadrada de tamaño $n-1$. El polinomio característico de u se puede escribir como $P_u(X) = P_A(X)(X - \lambda_n)$. Como P_u escinde, P_A también escinde sobre k y así, por hipótesis de inducción, $u|_U : U \to U$ es trigonalizable, i.e., existe una base $(e_1, \dots, e_{n-1})$ en donde $A' = \mathrm{Mat}_{(e_1, \dots, e_{n-1})}(u|_U)$ es triangular superior. Finalmente, $\mathrm{Mat}_{\mathscr{B}}(u)$ es triangular superior en la base $\mathscr{B} = (e_1, \dots, e_{n-1}, e'_n)$, lo que demuestra el teorema. □

Corolario 3.5.5. — *Sean $A \in \mathscr{M}_n(\mathbb{C})$ y $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ los valores propios de A (no necesariamente distintos). Entonces,*

$$\det(A) = \prod_{j=1}^{n} \lambda_j \quad y \quad \mathrm{tr}(A) = \sum_{j=1}^{n} \lambda_j$$

Demostración. — Como el polinomio $P_A(X) = (X-\lambda_1)\cdots(X-\lambda_n)$ escinde sobre $\mathbb{C}$ (Teo. Fund. del Álgebra) A es trigonalizable, i.e., existe $P \in \mathrm{GL}_n(\mathbb{C})$ tal que $B := P^{-1}AP$ es triangular superior. Vimos anteriormente que $P_B = P_A$ (3.3.12) y entonces

$$(2) \qquad B = \begin{pmatrix} \lambda_1 & * & * & * & * \\ & \lambda_2 & * & * & * \\ & & \ddots & * & * \\ & 0 & & \ddots & * \\ & & & & \lambda_n \end{pmatrix} \quad \Longrightarrow \quad \mathrm{tr}(A) = \mathrm{tr}(B) = \sum_{j=1}^{n} \lambda_j$$

y también

$$\det(A) = \det(B) = \prod_{j=1}^{n} \lambda_j$$

□

Ejercicio 3.5.6. — Sea $A \in \mathrm{GL}_n(k)$ una matriz triagular superior invertible con términos diagonales $\lambda_1, \ldots \lambda_n \in k$. Probar que A^{-1} es triangular superior y determinar sus términos diagonales. Deducir que si $u : V \to V$ es un automorfismo de V tal que $P_u(X) = \prod_{j=1}^{n}(X - \lambda_j)$ escinde sobre k, entonces

$$\det(u^{-1}) = \prod_{j=1}^{n} \frac{1}{\lambda_j} \quad \text{y} \quad \mathrm{tr}(u^{-1}) = \sum_{j=1}^{n} \frac{1}{\lambda_j}.$$

3.6. Polinomio minimal y teorema de Cayley-Hamilton

Recuerdo 3.6.1. — Dados $u \in \mathrm{End}_k(V)$ y $P \in k[X]$ polinomio de la forma $P(X) = a_0 + a_1X + \cdots + a_dX^d$ podemos definir $P(u) \in \mathrm{End}_k(V)$ como $P(u) = a_0 id_V + a_1u + \cdots + a_du^d$. Más aún, si V es de dimensión finita sabemos que para todo $u \in \mathrm{End}_k(V)$ existe $P \in k[X] - \{0\}$ tal que $P(u) = 0$.

La siguiente proposición introduce el concepto de **polinomio minimal**, el cual cobrará gran importancia de aquí en adelante.

Proposición 3.6.2. — *Sea $u : V \to V$ un endomorfismo. Entonces, existe un único polinomio unitario $m_u \in k[X]$ tal que para todo polinomio $P \in k[X]$ se cumple que:*

$$P(u) = 0 \Longleftrightarrow m_u \text{ divide a } P.$$

Diremos que m_u es el ***polinomio minimal de u****.*

Demostración. — Consideremos al conjunto $I = \{P \in k[X] : P(u) = 0\}$. Este conjunto es un **ideal** pues satisface:

1. Si $\lambda \in k$ y $P, Q \in I$, $\lambda P+Q \in I$. Esto pues $(\lambda P+Q)(u) = \lambda P(u)+Q(u) = 0$. Luego, I es un sub-ev. de $k[X]$.
2. Si $P \in I$ y $Q \in k[X]$, $PQ \in I$. Esto pues $(PQ)(u) = P(u)Q(u) = 0 \cdot Q(u) = 0$

Más aún, $I \neq \{0\}$ pues sabemos que siempre existe algún polinomio que anula a u y, como vimos en la sección de polinomios, esto es suficiente para concluir que existe un único $m_u \in k[X]$ unitario tal que $I = \langle m_u \rangle$, lo que demuestra la proposición. $\square$

Como ya es costumbre, definimos el polinomio minimal de $A \in \mathscr{M}_n(k)$ como el polinomio mimimal $m_A \in k[X]$ del endomorfismo $u_A : k^n \to k^n$ definido por $x \mapsto Ax$.

Ejercicio Importante 3.6.3. — Sea $u : V \to V$ un endomorfismo. Probar que $\mathrm{gr}(m_u) = 1$, i.e., $m_u = X - \lambda$ para cierto $\lambda \in k$, equivale a que $u = \lambda id_V$, esto es, u es una homotecia. En particular, $m_u(X) = X \iff u = 0$.

Ejemplo 3.6.4. — La matriz

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

verifica $A^2 = 0$. Luego, m_A divide a $P(X) = X^2$. Dado que $A \neq 0, m_A \neq X$ y entonces $m_A = X^2$.

Los siguientes teoremas relacionan al polinomio minimal con el polinomio característico. Ninguno de estos dos objetos depende de elecciones de base, por lo que serán de utilidad para saber cuándo un endomorfismo admite ciertas representaciones matriciales. Ya vimos un ejemplo de esto último con la trigonalización y pronto demostraremos más teoremas de ese espíritu.

Teorema 3.6.5. — *Sea $u : V \to V$ un endomorfismo. Entonces, todo valor propio de u es raíz del polinomio minimal $m_u \in k[X]$.*

Demostración. — Sea λ un valor propio de u y sea $v \neq 0$ un vector propio asociado. La división euclideana de m_u por $(X - \lambda)$ nos da

$$m_u(X) = Q(X)(X - \lambda) + c \qquad \text{donde } c \in k$$

Luego,

$$0 = m_u(u) = Q(u)(u - \lambda id_V) + c \cdot id_v \qquad \text{en } \mathrm{End}_k(V)$$

Aplicando este endomorfismo a v:

$$0 = Q(u)(\underbrace{(u - \lambda id_V)(v)}_{=0}) + c\,\underbrace{id_V(v)}_{=v} = \underbrace{Q(u)(0)}_{=0} + cv = cv$$

Como $v \neq 0$, $cv = 0 \Longrightarrow c = 0$. Así, $m_u(X) = Q(X)(X - \lambda)$ y λ es una raíz de m_u □

El recíproco de este teorema también es verdadero, lo cual implica que m_u divide a P_u. Sin embargo, es usual enunciar este resultado como un teorema propio.

***Teorema 3.6.6* (Teorema de Cayley-Hamilton)**

Sea $u : V \to V$ un endomorfismo y $P_u \in k[X]$ su polinomio característico. Entonces $P_u(u) = 0$ o, equivalentemente, el polinomio minimal m_u divide a P_u.

Demostración. — Sea $\mathscr{B}$ una base de V y sea

$$A = \mathrm{Mat}_{\mathscr{B}}(u) = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \in M_n(k).$$

Usaremos que $P_A(A) = 0 \Leftrightarrow P_u(u) = 0$. Consideremos:

$$B := XI_n - A \in \mathscr{M}_n(k(X)), \text{i.e., } B = (b_{ij}) = \begin{cases} X - a_{ij} & \text{si } i = j \\ -a_{ij} & \text{si } i \neq j \end{cases}$$

Y sea $C := {}^t\mathrm{com}(B)$ su comatriz transpuesta. Dado que cada cofactor de B es un polinomio de grado $\leq n - 1$ (al ser los determinantes de matrices de tamaño $n - 1$) podemos escribir:

$$C = \sum_{j=0}^{n-1} X^j C_j = C_0 + XC_1 + X^2 C_2 + \cdots + X^{n-1} C_{n-1}$$

Donde cada $C_j \in \mathscr{M}_n(k)$ es una matriz. Por otro lado, sabemos que

$$B \cdot \underbrace{{}^t\mathrm{com}(B)}_{=C} = \underbrace{\det(B)}_{=P_A(X)} I_n$$

Luego, si $P_A(X) = X^n + a_{n-1}X^{n-1} + \cdots + a_1 X^1 + a_0$ entonces:

$$\begin{aligned} BC &= (XI_n - A)(C_0 + XC_1 + X^2C_2 + \cdots + X^{n-1}C_{n-1}) \\ &= X^n C_{n-1} + X^{n-1}(C_{n-2} - AC_{n-1}) + \cdots + X(C_0 - AC_1) - AC_0 \\ &= P_A(X) I_n = X^n I_n + X^{n-1} a_{n-1} I_n + \cdots + X a_1 I_n + a_0 I_n \end{aligned}$$

Y al igualar coeficientes:
$a_0 I_n = -AC_0$; $a_1 I_n = (C_0 - AC_1)$; ... ; $a_{n-1} I_n = (C_{n-2} - AC_{n-1})$; $a_n I_n = I_n = C_{n-1}$.
Por lo tanto, al evaluar A obtenemos:

$$P_A(A) = \sum_{j=0}^{n} A^j (a_j I_n) = -AC_0 + A(C_0 - AC_1) + A^2(C_1 - AC_2) + \dots$$
$$\dots + A^{n-1}(C_{n-2} - AC_{n-1}) + A^n C_{n-1}$$

Finalmente, resta observar que lo anterior es una suma telescópica cuyo valor resulta ser cero. □

Observación 3.6.7. — *Como m_u divide a P_u para todo endomorfismo u, tenemos que P_u escinde sobre $k \implies m_u$ escinde sobre k.*

Ejemplo 3.6.8. —

1. Sea
$$A = \begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix} \in M_2(\mathbb{R})$$
Entonces
$$P_A(X) = \begin{vmatrix} X & -1 \\ -2 & X-3 \end{vmatrix} = X^2 - 3X - 2$$
El teorema de Cayley-Hamilton implica que $A^2 - 3A - 2I_2 = 0$. En efecto,
$$A^2 = \begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 6 & 11 \end{pmatrix}$$
$$\Rightarrow A^2 - 3A - 2I_2 = \begin{pmatrix} 2 & 3 \\ 6 & 11 \end{pmatrix} - 3\begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix} - 2\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$
Un uso típico de Cayley-Hamilton es expresar de forma sencilla las potencias de una matriz. En este caso particular se verifica $A^2 - 3A - 2I_2 = 0 \Rightarrow A^2 = 3A + 2I_2$, y multiplicando por A se obtiene
$$A^3 = 3A^2 + 2A = 3(3A + 2I_2) + 2A = 11A + 6I_2, \text{ etc.}$$
Notar que
$$\det(A) = \begin{vmatrix} 0 & 1 \\ 2 & 3 \end{vmatrix} = -2 \neq 0$$

i.e., A es invertible. Al multiplicar por A^{-1} se obtiene

$$A = 3I_2 + 2A^{-1} \Rightarrow A^{-1} = \frac{1}{2}(A - 3I_2) = \begin{pmatrix} -\frac{3}{2} & \frac{1}{2} \\ 1 & 0 \end{pmatrix}$$

Como A no es una homotecia pues $\mathrm{gr}(m_A) \geq 2 \Longrightarrow m_A(X) = P_A(X)$ gracias al teorema de Cayley-hamilton.

2. Supongamos que $\dim_k(V) = n$. Entonces, para todo endomorfismo $u : V \to V$ se tiene $\mathrm{gr}(P_u) = n$. Por Cayley-Hamilton, m_u divide a $P_u \Longrightarrow gr(m_u) \leq n$. En particular, si $u : V \to V$ tiene n valores propios distintos entonces, como todo valor propio es raíz de m_u, necesariamente $gr(m_u) \geq n$ y por lo tanto $m_u = P_u$.
3. Sean
$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \in \mathscr{M}_3(\mathbb{R})$$
y
$$B := A^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \in \mathscr{M}_3(\mathbb{R})$$
Tenemos $P_A(X) = P_B(X) = X^3$. Como A y B no son homotecias, $\mathrm{gr}(m_A), \mathrm{gr}(m_B) \geq 2$. Como $A^3 = 0$ y $A^2 = B \neq 0$, $m_A(X) = P_A(X) = X^3$. Además $B^2 = A^4 = 0$ y así
$$m_B(X) = X^2 \neq P_B(X)$$

3.7. Reducción de endomorfismos y descomposición de Dunford

Comenzaremos por probar el siguiente resultado útil:

***Teorema 3.7.1* (Lema de los kernel).** — *Sea $u : V \to V$ un endomorfismo y sea $P \in k[X]$ un polinomio. Entonces:*

1. *El sub-e.v.* $\ker(P(u))$ *es **estable bajo u**, ie,*
$$u(\ker(P(u)) \subseteq \ker(P(u))$$
2. *Si $P = AB$ con A y B primos entre sí (ie,* $\mathrm{mcd}(A, B) = 1$*) entonces:*
$$\ker(P(u)) = \ker(A(u)) \oplus \ker(B(u)).$$

Demostración. — Para probar (1), notemos que u y $P(u)$ conmutan, i.e., $P(u) \circ u = u \circ P(u)$. Así, si $v \in \ker(P(u))$ entonces

$$P(u)(v) = 0 \Longrightarrow P(u)(u(v)) = u(\underbrace{P(u)(v)}_{=0}) = u(0) = 0 \Longrightarrow u(v) \in \ker(P(u))$$

Para (2) usaremos el teorema de Bézout y el hecho de que polinomios evaluados en u conmutan. Como A y B son primos entre sí, existen $U, V \in k[X]$ tales que $AU + BV = 1$. En términos de endomorfismos, esto es $A(u) \circ U(u) + B(u) \circ V(u) = id_V$. Evaluemos $v \in \ker(A(u)) \cap \ker(B(u))$:

$$id_V(v) = v = (A(u) \circ U(u))(v) + (B(u) \circ V(u))(v) = U(u)(\underbrace{A(u)(v)}_{=0}) + V(u)(\underbrace{B(u)(v)}_{=0}) = 0$$

De donde $\ker(A(u)) \cap \ker(B(u)) = \{0\}$. Al ser esto así, $\ker(A(u))$ y $\ker(B(u))$ están en suma directa. Falta ver que generan a $\ker(P(u))$.
Partamos viendo qué pasa si $v \in \ker(A(u))$:

$$P(u)(v) = (A(u) \circ B(u))(v) = B(u)(\underbrace{A(u)(v)}_{=0}) = 0 \Longrightarrow v \in \ker(P(u))$$

Luego, $\ker(A(u)) \subseteq \ker(P(u))$y, por razones análogas, $\ker(B(u)) \subseteq \ker(P(u))$.
Por lo tanto $\ker(A(u)) \oplus \ker(B(u)) \subseteq \ker(P(u))$.
Ahora sea $v \in \ker(P(u))$. Si escribimos

$$v = id_V(v) = (A(u) \circ U(u))(v) + (B(u) \circ V(u))(v)$$

bastaría probar que $(A(u) \circ U(u))(v) \in \ker(B(u))$ y que $(B(u) \circ V(u))(v) \in \ker(A(u))$ para concluir que $\ker(P(u)) = \ker(A(u)) \oplus \ker(B(u))$. Para esto, notar que:

$$\underbrace{B(u) \circ (A(u)}_{=P(u)} \circ U(u))(v) = (P(u) \circ U(u))(v) = U(u)(\underbrace{P(u)(v)}_{=0}) = 0.$$

Y por lo tanto $(A(u) \circ U(u))(v) \in \ker(B(u))$.
De manera similar, comprobamos que $A(u) \circ (B(u) \circ V(u))(v) = 0$ y concluimos el resultado. □

El lema de los kernel (2) se generaliza a factorizaciones de P que involucren más de dos polinomios todos primos entre sí. Esto se logra usando inducción y se presenta en el siguiente corolario.

Corolario 3.7.2. — *Sea $u : V \to V$ un endomorfismo y sean $P_1, \ldots, P_m \in k[X]$ polinomios tales que mcd(P_i, P_j) $= 1$ si $i \neq j$. Si denotamos $P = P_1 \cdots P_m$, entonces:*

$$\ker(P(u)) = \ker(P_1(u)) \oplus \cdots \oplus \ker(P_m(u))$$

Demostración. — Inducción sobre m. Ya vimos el caso $m = 2$. Si $P_1, \ldots, P_m \in k[X]$ son dos a dos primos entre sí (ie, $\mathrm{mcd}(P_i, P_j) = 1$ si $i \neq j$), entonces el (corolario del) lema de Gauss nos dice que P_1 y $P_2 \cdots P_m$ son primos entre sí. Luego, podemos aplicar el lema de los kernel y obtener

$$\ker(P(u)) = \ker(P_1(u)) \oplus \ker(P_2 \cdots P_m)(u))$$

Por hipótesis de inducción, tenemos que

$$\ker(P(u)) = \ker(P_1(u)) \oplus \cdots \oplus \ker(P_m(u))$$

□

Ejercicio 3.7.3. — Construir $A \in M_2(\mathbb{R})$ tal que $m_A(X) = P_A(X) = X^2+1$.

Este teorema permite encontrar más fácilmente subespacios estables bajo u que se hallen en suma directa. Sabemos que estos son de particular utilidad porque dan paso a representaciones matriciales de u que son **diagonales por bloques**, que son útiles para hacer cálculos pues podemos trabajar cada bloque *por separado* al realizar operaciones como multiplicar, invertir o calcular determinantes.

El lema de los kernel es mejor empleado cuando trabajamos con polinomios $P \in k[X]$ que anulan a u,puesto que en dichos casos el kernel de $P(u)$ es *todo* el espacio V. Ejemplos de estos polinomios ya nos son conocidos: tenemos por ejemplo al polinomio característico de u, $P_u(X)$ y también al polinomio minimal $m_u(X)$.

Recuerdo 3.7.4. —

1. Las raíces del polinomio minimal m_u del endomorfismo $u : V \to V$ son los valores propios $\lambda_1, \ldots, \lambda_p$ (distintos) de u. Luego, *si suponemos* que m_u escinde sobre k, entonces:

$$m_u(X) = \prod_{j=1}^{p} (X - \lambda_j)^{m_j}$$

donde $m_j \geq 1$ para todo $j \in \{1, \ldots, p\}$.

2. Si $\lambda \in k$ es un valor propio de $u : V \to V$, el sub-e.v.

$$V_\lambda = \ker(u - \lambda id_V) \subseteq V$$

es llamado el **(sub)espacio propio** asociado a λ.

Definición 3.7.5. — Sea $u : V \to V$ tal que $P_u \in k[X]$ escinde sobre k, ie,

$$P_u(X) = \prod_{j=1}^{p} (X - \lambda_j)^{n_j}$$

1. El entero $mult_{alg}(\lambda_j) := n_j \geq 1$ es la **multiplicidad algebraica** de λ_j.
2. El entero $mult_{geom}(\lambda_j) := \dim_k(V_{\lambda_j})$ es la **multiplicidad geométrica** de λ_j.

Notar que el teorema de Cayley-Hamilton implica que $m_j \leq n_j = mult_{alg}(\lambda_j)$, donde m_j es la multiplicidad de λ_j como raíz del polinomio minimal.

Teorema 3.7.6. — *Sea $u : V \to V$ un endomorfismo en V con $\dim_k(V) = n$ y sean $\lambda_1, \dots, \lambda_p \in k$ sus valores propios distintos. Supongamos que P_u escinde sobre k. Entonces:*

1. $mult_{geom}(\lambda_i) \leq mult_{alg}(\lambda_i)$ *para todo* $i \in \{1, \dots p\}$.
2. *u es **diagonalizable** si y solo si $mult_{alg}(\lambda_i) = mult_{geom}(\lambda_i)$ para todo $i \in \{1, \dots, p\}$.*

Demostración. — Para (1), amenicemos la notación escribiendo

$$P_u(X) = \prod_{j=1}^{p} (X - \lambda_j)^{n_j} \quad \text{y} \quad d_j = mult_{geom}(\lambda_j) := \dim_k(V_{\lambda_j}) \text{ para todo } i \in \{1, \dots, p\}.$$

Queremos probar que $d_j \leq n_j$. Sea $\mathscr{C}_j$ una base de V_{λ_j}. Dado que los V_{λ_j} están en suma directa, tenemos que $\mathscr{C} = (\mathscr{C}_1, \dots, \mathscr{C}_p)$ es una familia l.i. y por ende se puede completar en una base $\mathscr{B}$ de V. Luego,

$$A := \mathrm{Mat}_{\mathscr{B}}(u) = \left(\begin{array}{c:c:c:c} \lambda_1 I_{d_1} & 0 & 0 & * \\ \hdashline 0 & \ddots & 0 & * \\ \hdashline 0 & 0 & \lambda_p I_{d_p} & * \\ \hdashline 0 & 0 & 0 & B \end{array}\right)$$

donde $B \in \mathscr{M}_l(k)$ con $l = n - (d_1 + \cdots + d_p)$. En particular, A es triangular superior por bloques. Por lo tanto:

$$P_u(X) = \det(XI_n - A) = P_B(X) \prod_{j=1}^{p} (X - \lambda_j)^{d_j}$$

$\Longrightarrow d_j \leq n_j$ para todo $j \in \{1, \dots, p\}$.
Para probar (2), notemos que u diagonalizable si y solo si se cumple que

$$l = 0 \iff P_B(X) = 1 \iff d_j = n_j \quad \forall j \in \{1, \dots, p\}$$

□

Ejemplo 3.7.7. — Si un endomorfismo $u : V \to V$ con $\dim_k(V) = n$, tiene exactamente n valores propios $\lambda_1, \ldots, \lambda_n$ distintos, entonces

$$P_u(X) = m_u(X) = \prod_{j=1}^{n}(X - \lambda_j)$$

$\implies \mathrm{mult}_{\mathrm{alg}}(\lambda_j) = 1 \quad \forall j \in \{1, \ldots, n\}$. Como además

$$0 < mult_{geom}(\lambda_j) \leq \mathrm{mult}_{\mathrm{alg}}(\lambda_j) = 1$$

tenemos que $\mathrm{mult}_{\mathrm{geom}}(\lambda_j) = \mathrm{mult}_{\mathrm{alg}}(\lambda_j) = 1 \forall j \in \{1, \ldots, n\}$ y el teorema anterior nos dice correctamente que u es diagonalizable.
Podemos llegar a este mismo resultado de otra forma considerando el polimonio minimal de u:

$$m_u(X) = \prod_{j=1}^{n}(X - \lambda_j)^{m_j} \quad \text{con cada } m_j = 1$$

y utilizar el lema de los kernel para concluir que $\ker(m_u(u)) = \ker(0) = V$ se descompone como

$$V = \ker(u - \lambda_1 id_V) \oplus \cdots \oplus \ker(u - \lambda_n id_V) = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_n}$$

i.e., que u es diagonalizable. La gracia de este enfoque es que podemos usar la misma estrategia con cualquier endomorfismo cuyo polinomio minimal escinda. La única complicación es: ¿qué pasa si algún m_j es mayor que 1?

Definición 3.7.8. — Sea $u : V \to V$ un endomorfismo y sean $\lambda_1, \ldots, \lambda_p \in k$ sus valores propios distintos. Si suponemos que $m_u \in k[X]$ escinde sobre k escribimos

$$m_u(X) = \prod_{j=1}^{p}(X - \lambda_j)^{m_j} \qquad \text{con } m_j \geq 1$$

y definimos el sub-e.v.

$$V_{(\lambda_j)} := \ker[(u - \lambda_j id_V)^{m_j}] \subseteq V$$

como el (sub-)**espacio característico** o el **espacio propio generalizado** asociado al valor propio λ_j.

Observación 3.7.9. —

1. *El lema de los kernel implica que* $u(V_{(\lambda)}) \subseteq V_{(\lambda)}$ *,es decir, que* $V_{(\lambda)}$ *es estable bajo* u. Para observar ello basta notar que si $P_j := (X - \lambda_j)^{m_j}$ entonces $\ker(P_j(u)) = V_{(\lambda_j)}$
2. $V_\lambda \subseteq V_{(\lambda)}$ *para todo valor propio* λ.

3. *Un vector no nulo $v \in V \setminus \{0\}$ pertenece a $V_{(\lambda)}$ si y solo si existe $k \in \{1, \dots, m_\lambda\}$ tal que $(u - \lambda id_V)^k(v) = 0$, donde m_λ es la multiplicidad de λ* como raíz del polinomio minimal. En este caso, diremos que *v es un **vector propio generalizado*** asociado al valor propio $\lambda \in k$.

Ejemplo 3.7.10. — Sea

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \in \mathscr{M}_3(\mathbb{R})$$

traingular superior. Por lo tanto $P_A(X) = (X-2)^2(X-3)$ y del teorema de Cayley-Hamilton se desprende que $m_A(X) = (X-2)^m(X-3)$ donde $m = 1$ o $m = 2$. Si $m = 1$ entonces se tendría que $(A - 2I_3)(A - 3I_3) = 0$, pero con un cálculo rápido se puede verificar que dicha identidad no se cumple. Así $m = 2$. De forma sencilla se puede obtener que los espacios propios de la matriz corresponden a $V_2 = \ker(A - 2I_3) = \langle e_1 \rangle$ y $V_3 = \ker(A - 3I_3) = \langle e_3 \rangle$. Sin embargo, es claro que $V_2 \oplus V_3$ no genera $\mathbb{R}^3$ y en consecuencia A no es diagonalizable. Por otro lado, se pueden obtener los espacios propios generalizados los cuales son $V_{(3)} = V_3$ (3 es raíz simple) y $V_{(2)} = \ker[(A - 2I_3)^2] = \langle e_1, e_2 \rangle$ y en este caso sí se tiene $\mathbb{R}^3 = V_{(2)} \oplus V_{(3)}$.

Teorema 3.7.11. — *Sea V un k-e.v. de dimensión n, sea $u : V \to V$ un endomorfismo y sean $\lambda_1, \dots, \lambda_p$ sus valores propios distintos. Si el polinomio característico $P_u \in k[X]$ escinde sobre k, i.e.,*

$$P_u(X) = \prod_{j=1}^{p} (X - \lambda_j)^{n_j} \qquad \textit{con } n_j \geq 1$$

(y en particular $m_u(X) = \prod_{j=1}^{p} (X - \lambda_j)^{m_j}$ también escinde), entonces:

1. *V es suma directa de los espacios característicos de u, ie,*

$$V = \bigoplus_{j=1}^{p} V_{(\lambda_j)} \qquad \textit{en donde } \; V_{(\lambda_j)} = \ker[(u - \lambda_j id_V)^{m_j}]$$

2. $\dim_k(V_{(\lambda_j)}) = mult_{alg}(\lambda_j) = n_j \;\; \forall j \in \{1, \dots, p\}$. *En particular, dado que $V_{\lambda_j} \subseteq V_{(\lambda_j)}$, se tiene que $mult_{geom}(\lambda_j) = \dim_k(V_{\lambda_j}) \leq \dim_k(V_{(\lambda_j)}) = mult_{alg}(\lambda_j)$.*

Demostración. — (1) es consecuencia directa del (corolario del) lema de los kernel:

Sea $P_j(X) := (X - \lambda_j)^{m_j}$ para $j \in \{1, \dots, p\}$. Como $\mathrm{mcd}(P_i, P_j) = 1$ si $i \neq j$ y $m_u = P_1 \cdots P_p$, podemos concluir, gracias a que $m_u(u) = 0$, que

$$\ker(0) = V = \underbrace{\ker(P_1(u))}_{V_{(\lambda_1)}} \oplus \cdots \oplus \underbrace{\ker(P_p(u))}_{V_{(\lambda_p)}}$$

(2) Como los espacios característicos son estables bajo u y dado que $V = V_{(\lambda_1)} \oplus \cdots \oplus V_{(\lambda_p)}$, podemos escoger una base $\mathscr{B}_j$ de cada $V_{(\lambda_j)}$ y formar una base $\mathscr{B} = (\mathscr{B}_1, \dots, \mathscr{B}_p)$ de V tal que $\mathrm{Mat}_{\mathscr{B}}(u)$ es una matriz diagonal por bloques:

$$A = \left(\begin{array}{c:c:c} A_1 & 0 & 0 \\ \hdashline 0 & \ddots & 0 \\ \hdashline 0 & 0 & A_p \end{array}\right)$$

donde $A_j \in \mathscr{M}_{s_j}(k)$, con $s_j = \dim_k(V_{(\lambda_j)})$. Queremos probar que $s_j = n_j$. Como

$$P_u(X) = P_A(X) = \det(XI_n - A) = P_{A_1}(X) \cdots P_{A_p}(X)$$

y $P_u(X)$ escinde por hipótesis, vemos que cada P_{A_j} escinde sobre k. Ahora, para cualquier $j \in \{1, \dots, p\}$, consideremos al endomorfismo $(\lambda_j id_V - u)|_{V_{(\lambda_j)}} : V_{(\lambda_j)} \to V_{(\lambda_j)}$, que es la restricción de $(\lambda_j id_V - u)$ al sub-espacio invariante bajo u (de ahí que sea en efecto un endomorfismo) $V_{(\lambda_j)}$. Matricialmente y en la base $\mathscr{B}$, a esta aplicación lineal le corresponde la matriz $(\lambda_j I_{s_j} - A_j)$. Recordemos que por definición $V_{(\lambda_j)} = \ker[(\lambda_j id_V - u)^{m_j}]$. Luego, $[(\lambda_j id_V - u)|_{V_{(\lambda_j)}}]^{m_j} = 0$, de donde

$$(\lambda_j I_{s_j} - A_j)^{m_j} = 0 \implies \det(\lambda_j I_{s_j} - A_j) = 0 \implies \lambda_j \text{ es valor propio de } A_j.$$

Más aún, este es el *único* valor propio de A_j. De tener otro, λ_i digamos, cualquier vector propio no nulo asociado estaría a la vez en $V_{(\lambda_j)}$ (que es el dominio de A_j) y en $V_{(\lambda_i)}$ (pues sería también un vector propio de A), lo cual es imposible pues estos dos espacios, al estar en suma directa, tienen como intersección al conjunto $\{0\}$.

Como P_{A_j} escinde sobre k y $\mathrm{gr}(P_{A_j}) = s_j$, concluimos que $P_{A_j}(X) = (X - \lambda_j)^{s_j}$. Así,

$$P_u(X) = P_{A_1}(X) \cdots P_{A_p}(X) = \prod_{j=1}^{p} (X - \lambda_j)^{s_j}$$

y por lo tanto $s_j = n_j$ para todo $j \in \{1, \dots, p\}$. $\square$

Observación 3.7.12. — *La demostación anterior también demuestra que toda matriz cuyo polinomio carcterístico escinda sobre k es semejante a una matriz diagonal por bloques*

$$\left(\begin{array}{c|c|c} A_1 & 0 & 0 \\ \hline 0 & \ddots & 0 \\ \hline 0 & 0 & A_p \end{array}\right)$$

donde cada bloque A_j posee solo un valor propio λ_j y verifica $(A_j - \lambda_j I_{n_j})^{m_j} = 0$.

Los resultados anteriores exploran la estructura de endomorfismos con vectores propios l.i. asociados a un mismo valor propio. Para esto relacionamos los espacios propios de un endomorfismo con sus polinomios característico y minimal (si escinden) por medio de las multiplicidades geométricas y algebraicas, que nos permiten "medir" el grado en que un endomorfismo falla para ser diagonalizable (si coinciden no falla, si no coinciden, falla). El siguiente teorema cierra el tema al darnos condiciones necesarias y suficientes para la diagonalización.

Teorema 3.7.13. — *Sea $u : V \to V$ un endomorfismo. Son equivalentes:*

1. *u es diagonalizable.*
2. *El polinomio característico de u, P_u, escinde sobre k y $mult_{geom}(\lambda) = mult_{alg}(\lambda)$ para todo valor propio λ.*
3. *$V_\lambda = V_{(\lambda)}$ para todo valor propio λ de u.*
4. *El polinomio minimal de u, m_u, escinde sobre k y solo tiene raíces simples.*

Demostración. — Primero, observemos que la implicancia (2) $\Longrightarrow$ (1) ya fue probada. Además, podemos probar rápidamente que (2) $\Longleftrightarrow$ (3) pues por definición $mult_{geom}(\lambda) = \dim_k(V_\lambda)$ y $mult_{alg}(\lambda) = \dim_k(V_{(\lambda)})$ y por construcción $V_\lambda \subseteq V_{(\lambda)}$.
Supongamos (1). Sea $\mathscr{B}$ una base de V en la que $A = \mathrm{Mat}_{\mathscr{B}}(u)$ es diagonal y en donde los valores propios repetidos están agrupados:

$$A = \left(\begin{array}{c|c|c} \lambda_1 I_{n_1} & 0 & 0 \\ \hline 0 & \ddots & 0 \\ \hline 0 & 0 & \lambda_p I_{n_p} \end{array}\right)$$

Como $P_u(X) = \prod_{j=1}^{p}(X-\lambda_j)^{n_j}$ y $n_j = \dim_k(V_{\lambda_j})$, obtenemos (2). Suponiendo también (1), si consideramos el polinomio $Q(X) = \prod_{j=1}^{p}(X-\lambda_j)$. se tiene que $Q(A) = 0$: cada factor anula uno de los bloques de A y al multiplicar todos los $(A-\lambda_j I_n)$ bloque a bloque terminamos con la matriz nula. Ahora, $Q(A) = 0 \implies m_u$ divide a Q, pero como las raíces de $m_u(X)$ son exactamente $\lambda_1, \ldots, \lambda_p$, necesariamente $m_u = Q$ y concluimos (1) $\implies$ (4). Finalmente, si asumimos (4), i.e., si $m_u(X) = \prod_{j=1}^{p}(X-\lambda_j)$, entonces se tiene la igualdad $V_{(\lambda_j)} = \ker(u-\lambda_j id_V) = V_{\lambda_j},\ j \in \{1,\ldots,p\}$, obteniendo (3) y por lo anterior u es diagonalizable. $\square$

Terminamos esta sección con una consecuencia importante de los resultados previos: la descomposición de Dunford. Este resultado nos asegura que siempre podemos expresar un endomorfismo u como la suma de un endomorfismo diagonalizable y otro "nilpotente", los que además conmutan entre ellos. Partiremos por introducir el concepto de endomorfismo nilpotente y luego demostraremos resultados sobre endomorfismos que conmutan:

Definición 3.7.14. — Sea $u : V \to V$ un endomorfismo. Diremos que u es **nilpotente** si existe $r \in \mathbb{N}^{\geq 1}$ tal que $u^r = 0$. En tal caso, el menor r con esta propiedad será llamado el **índice de nilpotencia** de u.

Ejemplo 3.7.15. —

$$B = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \in \mathscr{M}_3(\mathbb{R}) \Rightarrow B^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \Rightarrow B^3 = 0$$

y así B es nilpotente con un índice de nilpotencia igual a 3.

Lema 3.7.16. — *Sea $u : V \to V$ un endomorfismo donde $\dim_k(V) = n$. Entonces:*

1. *u es nilpotente si y solo si $m_u(X) = X^r$ para algún $r \in \mathbb{N}^{\geq 1}$. Más aún, r es su índice de nilpotencia y satisface $r \leq n$.*
2. *El único valor propio de un endomorfismo nilpotente es el 0.*
3. *Si u es nilpotente y diagonalizable, entonces $u = 0$.*

Demostración. — Para (1) veamos que

$$u^r = 0 \implies m_u(X) \text{ divide a } X^r \implies m_u(X) = X^m \text{ con } m \leq r$$

Si r es el índice de nilpotencia de u, $m_u(u) = u^m = 0 \implies m = r$ por minimalidad de r. Recíprocamente, $m_u(X) = X^r \implies u^r = 0$ y luego u es nilpotente. Dado que m_u divide a P_u, $r = \mathrm{gr}(m_u) \leq \mathrm{gr}(P_u) = n$.
De lo anterior deducimos (2) fácilmente pues los valores propios de u son las raíces de $m_u(X) = X^r$.
Finalmente, sabemos que un endomorfismo es diagonalizable si y solo si m_u escinde sobre k y tiene solo raíces simples. Bajo estas condiciones, $m_u(X) = X$ (pues $r = 1$) y $m_u(u) = u = 0$. □

Definición 3.7.17. — Sean $u : V \to V$ y $v : V \to V$ endomorfismos. Decimos que u y v **conmutan** si $u \circ v = v \circ u$.

Observación 3.7.18. — Si u y v conmutan, el teorema del binomio de Newton implica que para todo $n \in \mathbb{N}^{\geq 1}$

$$(u+v)^n = \sum_{k=0}^{n} \binom{n}{k} u^{n-k} \circ v^k$$

Esto es **falso** si u y v no conmutan.

Proposición 3.7.19. — *Sean $u : V \to V$ y $v : V \to V$ endomorfismos que conmutan. Entonces:*

1. *Si $\lambda \in k$ un valor propio de u, entonces V_λ y $V_{(\lambda)}$ son estables bajo v, i.e., $v(V_\lambda) \subseteq V_\lambda$ y $v(V_{(\lambda)}) \subseteq V_{(\lambda)}$.*
2. *Si u y v son diagonalizables, existe una base $\mathscr{B}$ de V formada por vectores que son vectores propios de u y de v simultáneamente. En particular, $u+v$ y $u \circ v$ son diagonalizables.*
3. *Si u y v son nilpotentes, entonces $u + v$ y $u \circ v$ también son nilpotentes.*

Demostración. — (1) Sea $x \in V_\lambda$. Entonces

$$u(v(x)) = v(u(x)) = v(\lambda x) = \lambda v(x) \implies v(x) \in V_\lambda$$

Del mismo modo, v y $P(u) = (u - \lambda id_V)^{m_\lambda}$ conmutan, por lo que

$$x \in V_{(\lambda)} \iff x \in \ker(P(u)) \implies P(u)(v(x)) = v(\underbrace{P(u)(x)}_{=0}) = 0 \implies v(x) \in V_{(\lambda)}$$

(2) Sean $\lambda_1, \ldots, \lambda_p \in k$ los valores propios distintos de u y escribamos $V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_p}$. Fijemos $j \in \{1 \ldots, p\}$ y denotemos $\lambda = \lambda_j$. Esto simplifica el problema a demostrar que V_λ admite una base $\mathscr{B}_\lambda$ formada por vectores propios de v. Para ello, sean $\mu_1, \ldots, \mu_r \in k$ los valores propios distintos de v y realizemos la demostración por inducción en r (i.e., en el numero de valores propios de v). Si $r = 1$, v es una homotecia (pues es diagonalizable) y el resultado

se sigue de que todo vector en V es vector propio de v. Para $r \geq 2$, definamos al subespacio estable $W := V_{\mu_1} \oplus \cdots \oplus V_{\mu_{r-1}}$, esto es, la suma directa de todos los espacios propios de v exceptuando a V_{μ_r}. Recordemos que v también es diagonalizable y por lo tanto todo $x \in V$ (y en particular todo $x \in V_\lambda$) se escribe de manera única como $x = x_1 + \cdots + x_r$, $x_j \in V_{\mu_j}$. Ahora, consideremos al endomorfismo $(v - \mu_r Id_V) : V_\lambda \to V_\lambda$. Notar que (1) nos asegura que esta aplicación lineal es en efecto un endomorfismo, además de que es diagonalizable y por lo tanto podemos escribir $V = \operatorname{Im}(u - \mu_r Id_V) \oplus \ker(u - \mu_r id_V)$. Ahora, escribiendo $x \in V_\lambda$ como $x = x_1 + \cdots + x_r$, $x_j \in V_{\mu_j}$ tenemos que:

$$(v-\mu_r id_V)(x) = \sum_{j=1}^{r}(v-\mu_r id_V)(x_j) = \sum_{j=1}^{r}(v(x_j)-\mu_r x_j) = (\underbrace{\mu_1 - \mu_r}_{\neq 0})x_1+\cdots+(\underbrace{\mu_{r-1} - \mu_r}_{\neq 0})x_{r-1}$$

pues el término $j = r$ se anula. De esto se desprenden dos cosas importantes. La primera es que $\ker(u - \mu_r Id_V) = V_{\mu_r} \cap V_\lambda$ y que una base $(z_1, \ldots, z_k)$ de $\ker(u - \mu_r Id_V)$ contiene solo vectores propios v asociados a μ_r. La segunda es que $(v - \mu_r Id_V)(x) \in W$ y por lo tanto $\operatorname{Im}(v - \mu_r Id_V) \subseteq W$. Notemos ahora que, sobre el subespacio $\operatorname{Im}(v - \mu_r Id_V)$, u conmuta con $v|_W$, que tiene $r - 1$ valores propios pues W no contiene vectores propios asociados a μ_r. Luego, podemos aplicar la hipótesis de inducción y construir una base $\mathscr{B}'_\lambda = (y_1, \ldots, y_s)$ de $\operatorname{Im}(v - \mu_r Id_V)$ formada con vectores propios de v asociados a alguno de los valores propios $\mu_1, \ldots, \mu_{r-1}$. Juntando esta base con la de $\ker(u - \mu_r Id_V)$ obtenemos una base de V del tipo

$$\mathscr{B}_\lambda = (y_1, \ldots, y_s, w_1, \ldots, w_k)$$

formada por vectores propios de v

Haciendo esto para cada λ_j podemos formar una base $\mathscr{B} = (\mathscr{B}_{\lambda_1}, \ldots, \mathscr{B}_{\lambda_p})$ de V formada de vectores propios de u y de v simultáneamente. En esta base, las matrices de u y de v son diagonales. En particular, $u+v$ y $u \circ v = v \circ u$ son también matrices diagonales, por lo que los endomorfismos que representan (i.e., $u + v$ y $u \circ v$) son diagonalizables.

(3) Supongamos que $u^r = 0$ y que $v^s = 0$ para ciertos $r, s \in \mathbb{N}^{\geq 1}$. Como u y v conmutan, $(u \circ v)^n = 0 = u^n \circ v^n$ si $n = \max(r, s)$. Por otro lado,

$$(u + v)^n = \sum_{k=0}^{n} \binom{n}{k} u^k \circ v^{n-k}$$

para todo $n \in \mathbb{N}^{\geq 1}$. Luego, $u^k = 0$ (resp. $v^{n-k} = 0$) si $k \geq r$ (resp. $n - k \geq s$). Así, si tomamos $n = r + s + 1$ en cada sumando tenemos una potencia de u o de v que es igual a 0, de donde $(u + v)^n = 0$ y $u + v$ es nilpotente. □

Ejercicio 3.7.20. —

1. Encuentre dos endomorfismos (o matrices) A y B tales que ambos son diagonalizables, pero su suma $A + B$ y su composición $A \circ B$ no lo son.
2. Haga lo mismo, pero esta vez con endomorfismos nilpotentes.

Teorema 3.7.21 **(Descomposición de Dunford)**

Sea $u : V \to V$ un endomorfismo tal que su polinomio característico P_u escinde sobre k. Entonces, u se descompone de manera única *como:*

$$u = u_s + u_n$$

donde u_s es diagonalizable (o "semi-simple") y u_n es nilpotente. Más aún, u_s y u_n conmutan.

Demostración. — **Existencia:** Vimos que todo endomorfismo $u : V \to V$ cuyo polinomio característico escinde sobre k admite una base $\mathscr{B}$ de V tal que $A = \mathrm{Mat}_{\mathscr{B}}(u)$ es diagonal por bloques:

$$A = \left(\begin{array}{c|c|c} A_1 & 0 & 0 \\ \hline 0 & \ddots & 0 \\ \hline 0 & 0 & A_p \end{array}\right)$$

donde cada bloque A_j posee un único valor propio $\lambda_j \in k$. Sabemos además que se verifica que $(A_j - \lambda_j I_{n_j})^{m_j} = 0$ para $m_j \in \mathbb{N}^{\geq 1}$ (n_j y m_j representan lo usual).

Sea $u_s : V \to V$ el endomorfismo tal que $S = \mathrm{Mat}_{\mathscr{B}}(u_s)$ es la matriz diagonal

$$S = \left(\begin{array}{c|c|c} \lambda_1 I_{n_1} & 0 & 0 \\ \hline 0 & \ddots & 0 \\ \hline 0 & 0 & \lambda_p I_{n_p} \end{array}\right)$$

por lo que u_s es claramente diagonalizable. Sea $u_n = u - u_s$, cuya matriz $N = \mathrm{Mat}_{\mathscr{B}}(u_n)$ vendría dada por

$$N = \left(\begin{array}{c|c|c} A_1 - \lambda_1 I_{n_1} & 0 & 0 \\ \hline 0 & \ddots & 0 \\ \hline 0 & 0 & A_p - \lambda_p I_{n_p} \end{array}\right)$$

Dado que $(A_j - \lambda_j I_{n_j})^{m_j} = 0$, se tiene que $u_n^r = 0$ para $r \geq \max(m_1, \dots, m_p)$ y así u_n es nilpotente.
Probemos ahora que conmutan. Como $u_s|_{V_{(\lambda_j)}} = \lambda_j id_{V_{(\lambda_j)}}$ se tiene que si $v \in V_{(\lambda_j)} \Rightarrow u(v) \in V_{(\lambda_j)}$. Luego, para $v \in V_{(\lambda_j)}$

$$u(u_s(v)) = u(\lambda_j v) = \lambda_j u(v) \quad \text{y que} \quad u_s(\underbrace{u(v)}_{\in V_{(\lambda_j)}}) = \lambda_j u(v)$$

es decir, u y u_s conmutan si tomamos vectores en $V_{(\lambda_j)}$. Dado que $V = \bigoplus_{j=1}^p V_{(\lambda_j)}$, tenemos que u y u_s conmutan en todo V. Así,

$$u_n u_s = (u - u_s)u_s = uu_s - u_s^2 = u_s u - u_s^2 = u_s(u - u_s) = u_s u_n$$

Unicidad: Sea $u = u'_s + u'_n$ otra descomposición con las mismas propiedades. Como u'_s y u'_n conmutan, $u'_s u = uu'_s$ y $u'_n u = uu'_n$. En particular, por las propiedades de endomorfismos que conmutan, tenemos que $u'_s(V_{(\lambda_j)}) \subseteq V_{(\lambda_j)}$ y $u'_n(V_{(\lambda_j)}) \subseteq V_{(\lambda_j)}$, i.e., el espacio característico $V_{(\lambda_j)}$ de u es estable bajo u'_s y bajo u'_n. Dado que $u_s|_{V_{(\lambda_j)}} = \lambda_j id_{V_{(\lambda_j)}}$ es una homotecia (y por lo tanto conmuta con cualquier endomorfismo) y que $V = \bigoplus_{j=1}^p V_{(\lambda_j)}$, necesariamente $u_s u'_s = u'_s u_s$ y $u_s u'_n. = u'_s u_s$. A partir de esto podemos concluir que u'_s y u'_n conmutan con u_n. Finalmente, como u_s y u'_s (resp. u_n y u'_n) conmutan y son diagonalizables (resp. nilpotentes), tenemos que $u_s - u'_s$ es diagonalizable y que $u'_n - u_n$ es nilpotente. Pero también tenemos que $u = u_s + u_n = u'_s + u'_n \Rightarrow u_s - u'_s = u'_n - u_n$. es un endomorfismo diagonalizable y nilpotente a la vez, y en consecuencia debe ser nulo. Así, $u_s = u'_s$ y $u_n = u'_n$. □

3.8. Cálculos de descomposición de Dunford

Caso particular importante: $k = \mathbb{R}$
Sea $A \in \mathscr{M}_n(\mathbb{R})$ y $P_A \in \mathbb{R}[X]$ su polinomio característico. Las raíces **complejas** de P_A pueden separarse en raíces reales (valores propios reales) y en pares de raíces complejas no-reales conjugadas: $\mu_1, \overline{\mu_1}, \dots, \mu_s, \overline{\mu_s}$. La multiplicidad de μ_j y $\overline{\mu_j}$ es la misma y los espacios propios V_{μ_j} y $V_{\overline{\mu_j}}$, así como los espacios carcaterísticos $V_{(\mu_j)}$ y $V_{(\overline{\mu_j})}$, son **conjugados**, ie, $\overline{V_{\mu_j}} = V_{\overline{\mu_j}}$ y $\overline{V_{(\mu_j)}} = V_{(\overline{\mu_j})}$
Esto permite reducir cálculos.
Consideremos los polinomios:

1. <u>Minimal real:</u> $m_A \in \mathbb{R}[X]$ definido como el polinomio **real** unitario de grado más pequeño tal que $m_A(A) = 0$ en $\mathscr{M}_n(\mathbb{R})$.

2. minimal complejo: $M_A \in \mathbb{C}[X]$ definido como el polinomio **complejo** unitario de grado más pequeño tal que $M_A(A) = 0$ en $\mathscr{M}_n(\mathbb{C})$.

Proposición 3.8.1. — *$m_A = M_A \in \mathbb{R}[X]$ (y entonces no necesariamente escinde sobre $\mathbb{R}$).*

Demostración. — Sabemos que M_A divide a m_A. Sea $M_A(X) = \sum_{j=0}^{m} a_j X^j,\ a_j \in \mathbb{C}$. Dado que $M_A(A) = 0$ debe cumplir que $Re(M_A)(A) = 0$, i.e.,

$$Re(M_A)(A) = \sum_{j=0}^{m} Re(a_j) A^j = 0$$

de donde m_A divide a $Re(M_A)$ y luego:

$$m = \mathrm{gr}(M_A) \underbrace{\leq}_{M_A \text{ divide a } m_A} \mathrm{gr}(m_a) \underbrace{\leq}_{m_A \text{ divide a } Re(M_A)} \mathrm{gr}(Re(M_A)) = m$$

Se concluye que $M_A = m_A$. □

Ejercicio 3.8.2. — Sea $A \in \mathscr{M}_n(\mathbb{R})$. Suponer que existe $P \in \mathrm{GL}_n(\mathbb{C})$ tal que $P^{-1}AP$ es diagonal con coeficientes **reales**. Probar que existe $Q \in \mathrm{GL}_n(\mathbb{R})$ tal que $Q^{-1}AQ$ es diagonal.

Ejemplo 3.8.3. — Sea $A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix} \in \mathscr{M}_4(\mathbb{C})$ (matriz de permutación).

$$\Rightarrow P_A(X) = \det \begin{pmatrix} X & -1 & 0 & 0 \\ 0 & X & -1 & 0 \\ 0 & 0 & X & -1 \\ -1 & 0 & 0 & X \end{pmatrix} = X \cdot \det \begin{pmatrix} X & -1 & 0 \\ 0 & X & -1 \\ 0 & 0 & X \end{pmatrix} - (-1) \det \begin{pmatrix} -1 & 0 & 0 \\ X & -1 & 0 \\ 0 & X & -1 \end{pmatrix} = X^4 - 1.$$

Luego los valores propios de A son $\lambda_1 = 1,\ \lambda_2 = -1\ \lambda_3 = i.\ \lambda_4 = -i$. Dado que los 4 valores priopios de A son distintos, A es diagonalizable sobre $\mathbb{C}$ y $P_A(X) = m_A(X)$.

Para calcular los espacios propios:

$\lambda_1 = 1$:

$$A - \lambda_1 I_4 = \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 1 & 0 & 0 & -1 \end{pmatrix} \longrightarrow \text{para encontrar una base de } V_1 = \ker(A - I_4)$$

usamos . operaciones columna:

$$\left(\begin{array}{cccc} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 1 & 0 & 0 & -1 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right) \xmapsto{C_2 \mapsto C_2 + C_1} \left(\begin{array}{cccc} -1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 1 & 1 & 0 & -1 \\ \hline 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right) \xmapsto{C_3 \mapsto C_3 + C_2}$$

$$\left(\begin{array}{cccc} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 1 & 1 & 1 & -1 \\ \hline 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right) \xmapsto{C_4 \mapsto C_4 + C_3} \left(\begin{array}{cccc} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 1 & 1 & 1 & 0 \\ \hline 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{array}\right) \Rightarrow V_1 = \mathrm{Vect}_{\mathbb{C}} \left\langle \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \right\rangle$$

Similarmente,

$$V_{-1} = \mathrm{Vect}_{\mathbb{C}} \left\langle \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix} \right\rangle, \; V_i = \mathrm{Vect}_{\mathbb{C}} \left\langle \begin{pmatrix} 1 \\ i \\ -1 \\ i \end{pmatrix} \right\rangle, \; V_{-i} = \mathrm{Vect}_{\mathbb{C}} \left\langle \begin{pmatrix} 1 \\ -i \\ -1 \\ i \end{pmatrix} \right\rangle$$

Importante: Gracias a la discusión anterior, basta calcular V_i para obtener $\overline{V_i} = V_{-i}$, pues son conjugados y $A \in \mathscr{M}_4(\mathbb{R})$ tiene coeficientes **reales**. Finalmente, dado que A es diagonalizable, su descomposición de Dunford es

$A = S + N$ donde $S = A$ y $N = 0$.

Ejemplo 3.8.4. — Sea $A = \begin{pmatrix} 2 & -1 & 2 \\ 10 & -5 & 7 \\ 4 & -2 & 2 \end{pmatrix} \in \mathscr{M}_3(\mathbb{R})$.

$$\Rightarrow P_A(X) = \det \begin{pmatrix} X-2 & 1 & -2 \\ -10 & X+5 & -7 \\ -4 & 2 & X-2 \end{pmatrix} \underset{C_1 \mapsto C_1+2C_2}{\overset{\uparrow}{=}} \det \begin{pmatrix} X & 1 & -2 \\ 2X & X+5 & -7 \\ 0 & 2 & X-2 \end{pmatrix}$$

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

$\Rightarrow P_A(X) = X^2(X+1)$. Luego, los valores propios de A son $\lambda_1 = -1$ (raíz simple) y $\lambda_2 = 0$ (raíz doble): $mult_{alg}(\lambda_1) = 1$, $mult_{geom}(\lambda_2) = 2$. Dado que $mult_{geom}(\lambda) \leq mult_{alg}(\lambda)$ tenemos que $\dim_{\mathbb{R}}(V_{-1}) = 1$ y que $\dim_{\mathbb{R}}(V_0) = 1$ ó 2. Ahora buscaremos bases para los espacios propios.
$\underline{\lambda_1 = -1}$: Ya sabemos que su dimensión es 1, por lo que basta encontrar solo un vector $v_1 \in \mathbb{R}^3$ solución del sistema $(A + I_3)(v_1) = 0$. Basta resolver:

$$\begin{aligned} 3x - y + 2z &= 0 \\ 10x - 4y + 7z &= 0 \\ 4x - 2y + 3z &= 0 \end{aligned}$$

Tomando $x = 1$ y, realizando operaciones fila convenientemente, llegamos a los valores $y = -1$ y $z = -2 \implies v_1 = \begin{pmatrix} 1 \\ -1 \\ -2 \end{pmatrix}$ genera a V_{-1}.

$\underline{\lambda_2 = 0}$: $V_0 = \ker(A)$. Haciendo operaciones columna:

$$\left(\begin{array}{ccc} 2 & -1 & 2 \\ 10 & -5 & 7 \\ 4 & -2 & 2 \\ \hline 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right) \xmapsto[C_2 \leftrightarrow C_1]{C_2 \mapsto -C_2} \left(\begin{array}{ccc} 1 & 2 & 2 \\ 5 & 10 & 7 \\ 2 & 4 & 2 \\ \hline 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{array}\right) \xmapsto[C_3 \mapsto C_3 - 2C_1]{C_3 \mapsto C_3 - 2C_1} \left(\begin{array}{ccc} 1 & 0 & 0 \\ 5 & 0 & 3 \\ 2 & 0 & -2 \\ \hline 0 & 1 & 0 \\ -1 & 2 & 2 \\ 0 & 0 & 1 \end{array}\right) \Rightarrow \ker(A) = \mathrm{Vect}_{\mathbb{R}}(v_2)$$

Con $v_2 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} \in \mathbb{R}^3$. En particular, $mult_{geom}(0) < mult_{alg}(0)$ y A **no** es diagonalizable. Más aún, comoen este caso su polonomio minimal no escinde en raíces simples deducimos $m_A(X) = X^2(X+1)$.

Veamos el espacio característico $V_{(0)} = \ker(A^2)$. Calculamos $A^2 = \begin{pmatrix} 2 & -1 & 1 \\ -2 & 1 & -1 \\ -4 & 2 & -2 \end{pmatrix}$.

Ahora:

$$\left(\begin{array}{ccc} 2 & -1 & 1 \\ -2 & 1 & -1 \\ -4 & 2 & -2 \\ \hline 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right) \xmapsto[c_3 \mapsto c_3 + c_2]{c_1 \mapsto c_1 + 2c_2} \left(\begin{array}{ccc} 0 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 0 \\ \hline 1 & 0 & 0 \\ 2 & 1 & 1 \\ 0 & 0 & 1 \end{array}\right) \Rightarrow \ker(A^2) = \mathrm{Vect}_{\mathbb{R}}(v_2, v_3)$$

Con $v_2 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$ un vector propio y $v_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$ un vector propio generalizado.

Sea $\mathscr{B} = (v_1, v_2, v_3)$ y $P = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 2 & 1 \\ -2 & 0 & 1 \end{pmatrix}$. Vimos que $P^{-1}AP = \left(\begin{array}{c|c} A_1 & 0 \\ \hline 0 & A_2 \end{array}\right)$.

Donde $A_1 \in M_1(\mathbb{R})$ posee solo a $\lambda_1 = -1$ como valor propio y $A_2 \in M_2(\mathbb{R})$

posee solo a $\lambda_2 = 0$ como valor propio.
Notar que $Av_1 = (-1)v_1 \;\Rightarrow\; A_1 = (-1) \in M_1(\mathbb{R})$. También:

$$Av_2 = 0 \text{ y } Av_3 = \begin{pmatrix} 2 & -1 & 2 \\ 10 & -5 & 7 \\ 4 & -2 & 2 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} = v_2 \implies A_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

Así, $P^{-1}AP = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$. Más aún, al invertir: $P^{-1} = \begin{pmatrix} 2 & -1 & 1 \\ -1 & 1 & -1 \\ 4 & -2 & 3 \end{pmatrix}$

Finalmente, para encontrar la descomposición de Dunford $A = S + N$ procedemos como en la demostración del Teorema de la descomposición de Dunford:

A la matriz por bloques $A' := \left(\begin{array}{c|c} A_1 & 0 \\ \hline 0 & A_2 \end{array}\right) = \left(\begin{array}{c|cc} -1 & 0 & 0 \\ \hline 0 & 0 & 1 \\ 0 & 0 & 0 \end{array}\right)$ la descomponemos en

$$A' = \underbrace{\left(\begin{array}{c|c} \lambda_1 I_1 & 0 \\ \hline 0 & \lambda_2 I_2 \end{array}\right)}_{S'} + \underbrace{\left(\begin{array}{c|c} A_1 - \lambda_1 I_1 & 0 \\ \hline 0 & A_2 - \lambda_2 I_2 \end{array}\right)}_{N'} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

Al revertir el cambio de base:

$$S = PS'P^{-1} = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 2 & 1 \\ -2 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 2 & -1 & 1 \\ -1 & 1 & -1 \\ 4 & -2 & 3 \end{pmatrix} = \begin{pmatrix} -2 & 1 & -1 \\ 2 & -1 & 1 \\ 4 & -2 & 2 \end{pmatrix}$$

$$\text{y } N = PN'P^{-1} = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 2 & 1 \\ -2 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 2 & -1 & 1 \\ -1 & 1 & -1 \\ 4 & -2 & 3 \end{pmatrix} = \begin{pmatrix} 4 & -2 & 3 \\ 8 & -4 & 6 \\ 0 & 0 & 0 \end{pmatrix}$$

Ejercicio 3.8.5. — Verificar que $A = S + N$, que S es diagonalizable, que N es nilpotente y que S y N conmutan.

3.9. La forma canónica de Jordan

Incluso sobre $\mathbb{C}$, no toda matriz es diagonalizable. Sin embargo, vimos en la sección anterior que al descomponer el k-espacio vectorial V como suma directa de sus espacios característicos $V_{(\lambda)}$ entonces toda matriz $M_n(k)$ cuyo polinomio característico escinda sobre k es semejante a una matriz diagonal por bloques

$$A = \left(\begin{array}{c|c|c} A_1 & 0 & 0 \\ \hline 0 & \ddots & 0 \\ \hline 0 & 0 & A_p \end{array}\right)$$

Donde cada bloque A_i posee un único valor propio λ_i y $A_i - \lambda_i I_{n_j}$ es nilpotente.
Una consecuencia de esto es la descomposición de Dunford.
En esta sección, veremos un resultado imortante debido a Camille Jordan (1838-1922) que describe de manera más precisa a los bloques A_i.

Definición 3.9.1. — Sea $m \in \mathbb{N}^{\geq 1}$ y $\lambda \in k$. Un *bloque de Jordan* de tamaño m asociado a λ es una matriz en $M_m(k)$ dada por

$$J_m(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ 0 & 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & \lambda \end{pmatrix}$$

Un caso particular importante es el *bloque de Jordan nilpotente* de tamaño m que denotaremos por $J_m = J_m(0)$.

Observación 3.9.2. — El índice de nilpotencia de una matriz J_m es m. Además, $J_m + \lambda I_m$ es la descomposición de Dunford de $J_m(\lambda)$.

Definición 3.9.3. — Sea $n \in \mathbb{N}^{\geq 1}$. Una *partición* de n es una sucesión finita no decreciente de enteros positivos $\underline{n} = (n_1, \ldots, n_s)$ tales que

$$n = n_1 + \cdots + n_s$$

Definición 3.9.4. — Sea $\lambda \in k$, $m \in \mathbb{N}^{\geq 1}$ y $\underline{m} = (m_1, \ldots, m_s)$ un partición de m. La *matriz de Jordan* de tamaño m asociada a λ y a $\underline{m}$ es la matriz diagonal por bloques dada por

$$J_{\underline{m}}(\lambda) = \left(\begin{array}{c|c|c} J_{m_1}(\lambda) & 0 & 0 \\ \hline 0 & \ddots & 0 \\ \hline 0 & 0 & J_{m_s}(\lambda) \end{array}\right)$$

Teorema 3.9.5. — *Sea $u : V \to V$ un endomorfismo nilpotente y sea $v \in V$ tal que $u^r(v) = 0$, pero $u^{r-1}(v) \neq 0$. Entonces los vectores de la sucesión*

$$v,\ u(v),\ u^2(v),\ \ldots,\ u^{r-1}(v)$$

son todos linealmente independientes.

Demostración. — Sean $\lambda_1, \ldots, \lambda_r$ escalares tales que

$$\lambda_1 v + \cdots + \lambda_r u^{r-1}(v) = 0$$

Aplicando u $r-1$ veces obtenemos la igualdad $\lambda_1 u^{r-1}()v = 0$, de donde concluimos que $\lambda_1 = 0$ (ya que $v \neq 0$). Usando esto, la suma se reduce a

$$\lambda_2 u(v) + \cdots + \lambda_r u^{r-1}(v) = 0.$$

Similarmente, al aplicar u $r-2$ veces llegamos a que $\lambda_2 = 0$. Iterando este proceso concluimos que todos los λ son iguales a cero y por lo tanto estos vectores son l.i. □

Corolario 3.9.6. — *Si $\dim_k(V) = n$ y el índice de nilpotencia de un endomorfismo $u : V \to V$ es n, entonces existe una base de V en la cual la matriz de u es un bloque de Jordan nilpotente de tamaño n.*

Demostración. — Basta tomar un vector $v \in V$ tal que $u^{n-1}(v)$ sea distinto de cero y considerar la base $\mathscr{B} = (u^{n-1}(v), u^{n-2}(v), \ldots, u(v), v)$ □

En el caso de un endomorfismo nilpotente con un índice de nilpotencia $k < n$, también es posible obtener "cadenas" de la forma

$$u^r(v), u^{r-1}(v), \ldots, u(v), v \tag{1}$$

Compuestas de vectores no nulos l.i. donde r es a lo más $k-1$ (ya que $u^k(v) = 0$ siempre).
Inspirándonos en el corolario anterior y haciendo la importante observación de que los sub-espacios generados por los vectores en estas cadenas son invariantes bajo u, vamos a buscar la manera de encontrar representaciones matriciales para u compuestas por bloques de Jordan nilpotentes a lo largo de la diagonal. La idea es formar cadenas como en (1) y luego armar una base de V con los vectores en ellas, ordenados igual que en el corolario. Una de las complicaciones que se presentan es la de evitar que estas cadenas terminen en un vector v tal que $v \in \mathrm{Im}(u)$. De ser este el caso, podemos hallar $w \in V$ tal que $v = u(w)$ y la cadena (1) de v estaría contenida en la cadena más larga

$$u^{r+1}(w), u^r(w), \ldots, u(w), w$$

Una manera no obvia de solventar esta cuestion es utilizando un resultado, que juega un rol importante en la demostración del teorema del rango y que ahora recordaremos como un lema.

Lema 3.9.7. — *Si* $(u(v_1), \dots, u(v_p))$ *es una base de la imagen del endomorfismo* $u : V \to V$ *y* $(k_1, \dots, k_q)$ *es una base de su kernel, entonces* $(v_1, \dots, v_p, k_1, \dots, k_q)$ *es una base de* V*.*

Teorema 3.9.8. — *Sea* V *un* k*-ev de dimensión* n *y sea* $u : V \to V$ *un endomorfismo nilpotente de índice* $k \leq n$*. Entonces existe una base* $\mathscr{B}$ *de* V *tal que* $\mathrm{Mat}_{\mathscr{B}}(u) = J_{\underline{n}} = J_{\underline{n}}(0)$ *para una única partición* $\underline{n}$ *de* n *completamente determinada por* u*.*

Demostración. — Por simplicidad notacional, definamos $u_i := u|_{\mathrm{Im}(u^i)}$. Consideremos la restricción de u al sub-espacio invariante $\mathrm{Im}(u^{k-2})$, es decir, al endomorfismo nilpotente de grado dos $u_{k-2} : \mathrm{Im}(u^{k-2}) \to \mathrm{Im}(u^{k-2})$.
Sea

$$(u^{k-1}(a_1), \dots, u^{k-1}(a_p)) \tag{1}$$

una base de $\mathrm{Im}(u^{k-1}) = \mathrm{Im}(u_{k-2})$. Observemos que cada uno de estos vectores pertenece al conjunto $\ker(u_{k-2})$ debido a que $u^k(v) = 0$ para todo v en V y a la inclusión $\mathrm{Im}(u^i) \subseteq \mathrm{Im}(u^j)$ si $i > j$. Al ser esto así, podemos completar (1) en un base de $\ker(u_{k-2})$ y, gracias al lema, utilizar algún conjunto de sus preimágenes para hallar una base de $\mathrm{Im}(u^{k-2})$. Elijamos como preimágenes a los vectores $u^{k-2}(a_1), \dots, u^{k-2}(a_p)$ y ordenemos así la base resultante:

$$(u^{k-1}(a_1), u^{k-2}(a_1), \dots, u^{k-1}(a_p), u^{k-2}(a_p), w_1, \dots, w_q) \tag{2}$$

Donde cada w_i pertenece a $\ker(u_{k-2})$. Explícitamente, esto significa que $w_i = u^{k-2}(b_i)$ para algún $b_i \in V$ (recordar que el dominio de u_{k-2} es $\mathrm{Im}(u^{k-2})$) y además que $u(w_i) = u^{k-1}(b_i) = 0$.

La intención es que los vectores $a_1, \dots, a_p$ formen las cadenas de largo $k-1$ (el máximo posible) y los vectores $b_1, \dots, b_q$ las de largo $k-2$. No es posible incluir la cadena de algún b_i en una más grande pues $b_i \notin \mathrm{Im}(u)$. Si lo estuviese, existiría $c \in V$ tal que $u(c) = b_i$. Luego, $u^{k-2}(b_i) = u^{k-1}(c) = w_i$, de donde $w_i \in \mathrm{Im}(u^{k-1})$, lo cual es imposible porque usamos a w_i para completar una base de $\mathrm{Im}(u^{k-1})$ en una base $\ker(u_{k-2})$.

El siguiente paso es considerar la restricción de u al subespacio invariante $\mathrm{Im}(u^{k-3})$, es decir, a $u_{k-3} : \mathrm{Im}\, u^{k-3} \to \mathrm{Im}\, u^{k-3}$. Como base de

$\mathrm{Im}(u^{k-2}) = \mathrm{Im}(u_{k-3})$ usaremos a (2), pero reescribiendo los w_i:

(3) $\quad (u^{k-1}(a_1), u^{k-2}(a_1), \dots, u^{k-1}(a_p), u^{k-2}(a_p), u^{k-2}(b_1), \dots, u^{k-2}(b_q))$

En este caso no todos estos los vectores se anula si aplicamos u otra vez. Tan solo lo hace uno por cadena: el de exponente más alto. Podemos completar el conjunto de tales vectores en una base de $\ker(u_{k-3})$ y unirla a un conjunto de preimágenes de (3) para obtener una base de $\mathrm{Im}(u^{k-3})$ (Lema). Para recuperar las cadenas, elegimos $u^{j-1}(v)$ como preimágen $u^j(v)$ para todos los vectores. Obtenemos una base con la siguiente estructura:

(4) $\quad (u^{k-1}(a_1), u^{k-2}(a_1), u^{k-3}(a_1), \dots, u^{k-2}(b_1), u^{k-3}(b_1), \dots, f_1, \dots, f_r))$

Donde cada f_i pertenece a $\ker(u_{k-3})$. Reescribimos $f_i = u^{k-3}(c_i)$, donde $u^{k-2}(c_i) = 0$ y $c_i \notin \mathrm{Im}(u)$ por las mismas razones que utilizamos para argumentar que $b_i \notin \mathrm{Im}(u)$ en el paso anterior. Teniendo ya las cabezas de las cadenas de largo k-3 (los $f_i = u^{k-3}(c_i)$), repetimos los mismos pasos hasta hallar una base de $\mathrm{Im}(u)$ con esta misma estructura. En ese punto solo hay que repetir el proceso una vez más para obtener una base de V de la forma:

$$\mathscr{B} = (u^{k-1}(a_1), \dots, u(a_1), a_1, \dots, u^{k-2}(b_1), \dots, u(b_1), b_1, \dots\dots, u(d_1), d_1, \dots, e_1, \dots, e_s)$$

Donde los e_i pertenecen a $\ker(u)$.
Como los sub-espacios generados por cada cadena son invariantes bajo u, corresponden, cada una, a un bloque nilpotente de Jordan sobre la diagonal de $\mathrm{Mat}_{\mathscr{B}}(u)$ (ver corolario). Así, $\mathrm{Mat}_{\mathscr{B}}(u) = J_{\underline{n}}$ donde la partición $\underline{n}$ viene determinada por la cantidad de cadenas de un cierto largo que existen dentro de $\mathscr{B}$. Como cada cadena empieza con un vector que se anula si aplicamos u una vez más, el entero $\dim(\ker(u))$ nos da el número total de cadenas. Similarmente, el entero $\dim(\ker(u^2))$ nos dice el número de vectores que se anulan al aplicar u una o dos veces más. Por lo tanto, $\dim(\ker(u^2)) - \dim(\ker(u))$ es la cantidad vectores en $\mathscr{B}$ que se anulan al aplicar u exactamente 2 veces o, equivalentemente, la cantidad de cadenas de tamaño mayor o igual a 2 que existen dentro de $\mathscr{B}$. En general, la cantidad de cadenas de tamaño mayor o igual a r viene dada por el entero $\dim(\ker(u^r)) - \dim(\ker(u^{r-1}))$. Como estos valores dependen exclusivamente de u, la partición $\underline{n}$ viene completamente determinada por u también y es única al seguir la convención de que sus términos son no decrecientes. $\square$

***Teorema 3.9.9* (Jordan, 1870).** — *Sea V un ke.v. con $\dim_k(V) = n$ y sea $u : V \to V$ un endomorfismo tal que su polinomio característico escinde sobre* *k, i.e,*

$$P_u(X) = \prod_{j=1}^{p} (X - \lambda_j)^{n_j}$$

Con $n_j \geq 1$ y $\lambda_i \neq \lambda_j$ si $i \neq j$. Entonces, existe una base de $\mathscr{B}$ de V en la cual la matriz $\mathrm{Mat}_{\mathscr{B}}(u)$ es de la forma

$$J = \left(\begin{array}{c|c|c} J_{\underline{n_1}}(\lambda_1) & 0 & 0 \\ \hline 0 & \ddots & 0 \\ \hline 0 & 0 & J_{\underline{n_p}}(\lambda_p) \end{array}\right)$$

Donde $J_{\underline{n_j}}(\lambda) \in M_{n_j}(k)$ es una matriz de Jordan de tamaño n_j asociada a λ_j y a la partición $\underline{n_j}$, que es única exceptuando permutaciones de los bloques.

Demostración. — Para cada $i \in \{1, \dots, p\}$ denotemos $u_j := u|_{(V_{\lambda_j})} : V_{\lambda_j} \to V_{\lambda_j}$. Ya que por definición $V_{\lambda_j} = \ker((u - \lambda_j Id_v)^{m_j})$, tenemos que $u_j - \lambda_j Id_{V_{\lambda_j}}$ es un endomorfismo nilpotente de grado m_j. Ahora, el teorema anterior nos asegura la existencia de una base $\mathscr{B}_j$ de V_{λ_j} tal que $\mathrm{Mat}_{\mathscr{B}_j}(u_j - \lambda_j Id_{V_{\lambda_j}}) = J_{\underline{n_j}}$ para una única partición $\underline{n_j}$ de $n_j = mult_{alg}(\lambda_j)$.
Finalmente,

$$\mathrm{Mat}_{\mathscr{B}_j}(u_j) = J_{\underline{n_j}} + \lambda_j I_{n_j} = J_{\underline{n_j}}(\lambda_j)$$

Dado que $V = \bigoplus_{j=1}^{p} V_{\lambda_j}$, si consideramos la base $\mathscr{B} = (\mathscr{B}_1, \dots, \mathscr{B}_p)$ obtenemos para $\mathrm{Mat}_{\mathscr{B}}(u)$ una matriz diagonal por bloques de Jordan.
Veamos la unicidad módulo permutación de bloques: supongamos que existe una base $\mathscr{B}'$ de V tal que

$$J = \left(\begin{array}{c|c|c} J_{\underline{n'_1}}(\mu_1) & 0 & 0 \\ \hline 0 & \ddots & 0 \\ \hline 0 & 0 & J_{\underline{n'_s}}(\mu_s) \end{array}\right)$$

Calculando su polinomio carcaterístico obtenemos $P_u(X) = \prod_{j=1}^{s}(X - mu_j)^{n'_j}$. Como el polinomio carcaterístico es invariante ante cambios de base, debe existir $\sigma \in S_p$ tal que $\mu_j = \lambda_{\sigma(j)}$ y $n'_j = n_{\sigma(j)}$ para todo $j \in \{1, \dots, p\}$ (recordar que S_p es el grupo de permutaciones para un conjunto de p elementos). Resta probar que las particiones $\underline{n'_j}$ y $\underline{n_{\sigma(j)}}$ son iguales. Podemos reordenar los μ_j y suponer que $\sigma = Id$ en S_p para simplificar.
Sea $\mathscr{B}'_j$ el conjunto de vectores de $\mathscr{B}'$ que corresponden al bloque de Jordan $J_{\underline{n'_j}}(\lambda_j)$. Sea $W_j = \mathrm{Vect}_k(\mathscr{B}'_j)$. Se verifica que W_j es invariante bajo u y

que $\dim_k(W_j) = n_j$ pues $n_j = n'_j$. Además, si $\mathrm{Mat}_{\mathscr{B}'_j}(u|_{W_j}) = J_{\underline{n'_j}}(\lambda_j)$ por construcción.
Como $J_{\underline{n'_j}}(\lambda_j)$ es triangular superior con ceros en la diagonal y de tamaño $n_j \times n_j$, tenemos que

$$J_{\underline{n'_j}}^{n_j} = (J_{\underline{n'_j}}(\lambda_j) - \lambda_j I_{n_j})^{n_j} = 0 \implies (u|_{W-J} - \lambda Id_{w_j})^{n_j} = 0$$

De donde $W_j \subseteq V_{(\lambda_j)}$. Como tienen la misma dimensión (n_j), concluimos que en realidad $W_j = V_{(\lambda_j)}$.
Luego, $\mathscr{B}'$ y $\mathscr{B}$ son bases del mismo sub-espacio y verifican

$$\mathrm{Mat}_{\mathscr{B}_j}(u_j - \lambda_j Id_{V_{(\lambda_j)}}) = J_{\underline{n_j}} \text{ y } \mathrm{Mat}_{\mathscr{B}'_j}(u_j - \lambda_j Id_{V_{(\lambda_j)}}) = J_{\underline{n'_j}}$$

La unicidad de la partición en el caso nilpotente implica que $\underline{n_j} = \underline{n'_j}$. $\square$

Observación 3.9.10. — *En términos matriciales*: Sea $A \in M_n(k)$ tal que $P_A(X) \in k[X]$ escinde sobre k. Entonces, existe $P \in \mathrm{GL}_n(k)$ tal que

$$P^{-1}AP = \left(\begin{array}{c|c|c} J_{\underline{n_1}}(\lambda_1) & 0 & 0 \\ \hline 0 & \ddots & 0 \\ \hline 0 & 0 & J_{\underline{n_p}}(\lambda_p) \end{array}\right) =: J \qquad \text{“\textit{Forma canónica de Jordan}”}$$

Más aún, J es única módulo permutación de los bloques de Jordan $J_{\underline{n_j}}(\lambda_j)$.

Del teorema anterior y de la unicidad en el caso nilpotente obtenemos:

Corolario 3.9.11. — *Sean $A, B \in M_n(k)$ tales que P_Ay $P_B \in k[X]$ escinden sobre k. Son equivalentes:*

1. *A y B son semejantes (i.e., existe $P \in \mathrm{GL}_n(k)$ tal que $P^{-1}AP = B$).*
2. *A y B tienen el mmismo polinomio característico $P_A(X) = P_B(X) = \prod_{j=1}^{p}(X-\lambda_j)^{n_j}$ (con $\lambda_i \neq \lambda_j$ si $i \neq j$) y $\dim_k \ker[(A - \lambda I_n)^m] = \dim_k \ker[(B - \lambda I_n)^m]$ para todo valor propio λ y $m \in \mathbb{N}^{\geq 1}$.*
3. *A y B tienen la misma forma canónica de Jordan (módulo permutación).*

Ejercicio 3.9.12. — Sean $\lambda_1, \ldots, \lambda_p \in \mathbb{C}$ dos a dos distintos. Sea $P(X) = \prod_{j=1}^{p}(X - \lambda_j)^{n_j}$ con $n_j \in \mathbb{N}^{\geq 1}$ y $n_1 + \cdots + n_p = n$. El corolario anterior implica que el número de clases de semejanza de matrices $A \in M_n(\mathbb{C})$ con $P_A = P$ es igual a $p(n_1) \cdots p(n_p)$, donde $p(n_j)$ es el número de particiones de n_j.

1. Sea $P(X) = X^7 - 3X^6 + 3X^5 - X^4$. ¿Cuál es el número de clases de semejanza de matrices en $M_7(\mathbb{C})$ con polinomio característico P?

2. Sea $P(X) = \prod_{j=1}^{p}(X-\lambda_j)^{n_j}$ y $C(A) = \{A \in M_n(\mathbb{C}) : P_A(X) = P(X)\}$. ¿Bajo qué condición sobre P el conjunto $C(A)$ está formado por una sola clase de semejanza?

<u>**Importante**</u> : Presentamos ahora una manera algorítmica de calcular la forma canónoca de Jordan:
Sea $A \in M_n(k)$ tal que $P_A(X) = \prod_{j=1}^{p}(X-\lambda_j)^{n_j}$ escinde sobre k. Para cada λ_j:

1. Calculamos $d_m := \dim_k \ker[(A-\lambda_j I_n)^m]$ para $m = 1, 2, \ldots$ hasta un entero r_j tal que $d_{r_j} = n_j = \dim_k V_{(\lambda_j)}$.
   Observaciones: Si conocemos $m_A = \prod_{j=1}^{p}(X-\lambda_j)^{m_j}$ entonces $r_j = m_j$. Además, el teorema del rango establece que $d_m = n - \mathrm{rg}[(A-\lambda_j I_n)^m]$, por lo que el resultado de hacer esto es una secuencia $0 := d_0 < d_1 < \cdots < d_{r_j} = n_j$.
2. La secuencia $d_0, \ldots, d_{r_j}$ determina cuántos bloques de Jordan $J_m(\lambda_j)$ hay en la matriz de Jordan $J_{\underline{n_j}}(\lambda_j)$: $d_1 = d_1 - d_0 = \dim_k \ker[(A-\lambda_j I_n)]$ es el número de bloques de Jordan de tamaño mayor o igual a 1 que hay dentro de $J_{\underline{n_j}}(\lambda_j)$ (**observación**: si $d_1 = 1$, entonces $J_{\underline{n_j}}(\lambda_j) = J_{n_j}(\lambda_j)$ pues consiste de un único bloque). $d_2 - d_1$ es el número de bloques de Jordan de tamaño mayor o igual a 2 que hay dentro de $J_{\underline{n_j}}(\lambda_j)$. $d_3 - d_2$ es el número de bloques de Jordan de tamaño mayor o igual a 3 que hay dentro de $J_{\underline{n_j}}(\lambda_j)$, etc...

Como conocemos la estructura de los bloques de Jordan, esta información es suficiente para obtener la matriz de Jordan $J_{\underline{n_j}}(\lambda_j) \in M_{n_j}(k)$ asociada a λ_j.
Finalmente,

$$J = \left(\begin{array}{c|c|c} J_{\underline{n_1}}(\lambda_1) & 0 & 0 \\ \hline 0 & \ddots & 0 \\ \hline 0 & 0 & J_{\underline{n_p}}(\lambda_p) \end{array}\right) \in M_n(k).$$

Veremos en ejemplos cómo obtener $P \in \mathrm{GL}_n(k)$ a partir de A y de J.

Ejemplo 3.9.13. — Sea $A = \begin{pmatrix} 0 & 1 & -1 & 0 \\ -3 & -1 & 3 & 1 \\ -2 & 1 & 1 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix} \in M_4(\mathbb{R})$. Calculemos $P_A(X)$:

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

$$= (X-2)^2\begin{vmatrix} X & -1 & 0 \\ 3 & X+1 & 1 \\ 2 & -1 & 1 \end{vmatrix} = (X-2)^2\begin{vmatrix} X & -1 & 0 \\ 1 & X+2 & 0 \\ 2 & -1 & 1 \end{vmatrix} = (X-2)^2\begin{vmatrix} X & -1 \\ 1 & X+2 \end{vmatrix} = (X-2)^2(X+1)^2$$

Sea $\lambda_1 = -1$ y $\lambda_2 = 2$ y sea $n_1 = mult_{alg}(-1) = 2$ y $n_2 = mult_{alg}(2) = 2$.

$\underline{\lambda_1 = -1}$: Sea $B = A - \lambda_1 I_4 = A + I_4 = \begin{pmatrix} 1 & 1 & -1 & 0 \\ -3 & 0 & 3 & 1 \\ -2 & 1 & 2 & 1 \\ 0 & 0 & 0 & 3 \end{pmatrix}$. Nos interesan $\ker(B)$ y $\ker(B^2)$.

Para calcular una base de $\ker(B) = V_{-1}$ resolvemos el sistema $B\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$.

Obtenemos $\ker(B) = \mathrm{Vect}_{\mathbb{R}}(v_1)$ con $v_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$ (en particular $mult_{geom}(-1) = 1$ y A no es diagonalizable).

Como $\dim_{\mathbb{R}} \ker(B) = 1$, $J_{\underline{n_1}}(-1)$ posee 1 bloque de Jordan $J_{\underline{n_1}}(-1) = J_2(-1) = \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix}$.

Más aún, $J_{\underline{n_1}}(-1)$ nos dice cómo escoger una "buena" base de $V_{(-1)}$: $\mathscr{B}_1 = (v_1, v_2)$ con $Av_1 = (-1)v_1$, i.e., $(A+I_4)v_1 = 0$ y $Av_2 = v_1 + (-1)v_2 \iff (A+I_4)v_2 = v_1$.

Esto es, tomamos $v_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$ y v_2 tal que $Bv_2 = v_1$. Por ejemplo, resolvemos el sistema y escogemos $v_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$.

(Otra forma: calcular B^2 y una base de $\ker(B^2)$ dada por $\mathscr{B}_1 = (v_1, v_2)$ con $\ker(B) = \mathrm{Vect}_{\mathbb{R}}(v_1)$ y $\ker(B^2) = \ker(B) \oplus \mathrm{Vect}_{\mathbb{R}}(v_2)$. <u>Ejercicio</u>)

$\underline{\lambda_2 = 2}$: Sea $C = A - \lambda_2 I_4 = A - 2I_4 = \begin{pmatrix} -2 & 1 & -1 & 0 \\ -3 & -3 & 3 & 1 \\ -2 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$. Nos interesan $\ker(C)$ y $\ker(C^2)$.

Verificamos mediante operaciones columna que $\ker(C) = \mathrm{Vect}_{\mathbb{R}}(v_3)$ con $v_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}$. Dado que $\dim_{\mathbb{R}} \ker(C) = 1$, $J_{\underline{n_2}}(2)$ posee 1 bloque de Jordan $J_{\underline{n_2}}(2) = J_2(2) = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$ y luego al calcular una solución de $(A - 2I_4)v_4 = v_3 \iff Cv_4 = v_3$ como por ejemplo $v_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$ obtenemos una base $\mathscr{B}_2 = (v_3, v_4)$ de $V_{(2)}$.

Sea $\mathscr{B} = (v_1, v_2, v_3, v_4)$ y $P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$.

$$PAP^{-1} = J = \left(\begin{array}{cc|cc} -1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ \hline 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{array}\right) \qquad \text{(forma canónica de Jordan de A)}$$

Ejercicio importante: Sea $u : V \to V$ un endomorfismo tal que P_u escinde

sobre k y sea $u = u_s + u_n$ su descomposición de Dunford. Sea $\mathscr{B}$ una base de V tal que

$$J = \mathrm{Mat}_{\mathscr{B}}(u) = \left(\begin{array}{c|c|c} J_{\underline{n_1}(\lambda_1)} & 0 & 0 \\ \hline 0 & \ddots & 0 \\ \hline 0 & 0 & J_{\underline{n_p}(\lambda_p)} \end{array}\right)$$

la forma canónica de Jordan de u. Probar que $S = \mathrm{Mat}_{\mathscr{B}}(u_s)$ y $N = \mathrm{Mat}_{\mathscr{B}}(u_n)$ están dadas por:

$$S = \left(\begin{array}{c|c|c} \lambda_1 I_{n_1} & 0 & 0 \\ \hline 0 & \ddots & 0 \\ \hline 0 & 0 & \lambda_p I_{n_p} \end{array}\right) \qquad \text{y} \qquad N = \left(\begin{array}{c|c|c} J_{\underline{n_1}} & 0 & 0 \\ \hline 0 & \ddots & 0 \\ \hline 0 & 0 & J_{\underline{n_p}} \end{array}\right)$$

CAPÍTULO 4

EXPONENCIAL DE MATRICES Y SUS APLICACIONES

4.1. Espacios vectoriales normados

Recordemos que una sucesión $(x_n)_{n\in\mathbb{N}}$ de naturales converge a L si para cada $\epsilon > 0$ existe un $N \in \mathbb{N}$ tal que para $n \geq N$ se tendrá $|x_m - L| < \epsilon$. Esto último nos dice que existe un instante de tiempo tal que la sucesión está tan cerca del límite L como queramos.
La idea de esta sección es extender esta idea a sucesiónes $(v_n)_{n\in\mathbb{N}}$ donde $v_m \in V$ es un vector, y para ello necesitamos una noción de "distancia".
A modo de ejemplo, podemos considerar $V = \mathbb{R}^2$: Si $v_1 = (x_1, y_1)$ y $v_2 = (x_2, y_2)$ definimos la distancia <u>euclidiana</u> entre v_1 y v_2 por $\|v_1 - v_2\|_2$, donde

$$\|(x, y\|)_2 := \sqrt{x^2 + y^2} \qquad \forall (x, y) \in \mathbb{R}^2.$$

Es natural decir entonces que una sucesión de vectores $(v_n)_{n\in\mathbb{N}}$ en $\mathbb{R}^2$ converge a L si para cada $\epsilon > 0$ existe $N \in \mathbb{N}$ tal que si $n \geq N$, entonces $\|v_n - L\|_2 < \epsilon$.
Por otro lado, si escribimos las coordenadas $v_n = (x_n, y_n) \in \mathbb{R}^2$ y $L = (a, b) \in \mathbb{R}^2$, es natural también decir que $(v_n)_{n\in\mathbb{N}}$ converge a L si $(x_m)_{n\in\mathbb{N}}$ converge a a e $(y_m)_{n\in\mathbb{N}}$ converge a b. No es difícil notar que esto es equivalente a decir que para cada $\epsilon > 0$ existe $N \in \mathbb{N}$ tal que si $n \geq N$, entonces $\|v_n - L\|_\infty < \epsilon$, donde

$$\|(x, y)\|_\infty = \max\{|x|, |y|\} \qquad \forall (x, y) \in \mathbb{R}^2.$$

¿Cómo comparar ambas nociones de convergencia?

Ejercicio 4.1.1. — Probar que para todo $(x, y) \in \mathbb{R}^2$ se tiene

$$\frac{1}{\sqrt{2}}\|(x, y)\|_2 \leq \|(x, y)\|_\infty \leq \|(x, y)\|_2.$$

Deducir, por lo tanto, que las nociones de convergencia anteriores coinciden, es decir, si $(v_n)_{n\in\mathbb{N}}$ es una sucesión de vectores en $\mathbb{R}^2$ y $\ell \in \mathbb{R}^2$, entonces

$$v_n \xrightarrow{\|\cdot\|_2} \ell \iff v_n \xrightarrow{\|\cdot\|_\infty} \ell.$$

Conclusión: Necesitamos una noción de "distancia" y, aunque no haya una única manera de medir distancias, la noción de convergencia es la misma si podemos compararlas como en el ejercicio anterior.
En esta sección consideraremos $k = \mathbb{R}$ o $\mathbb{C}$, y $|\cdot|$ denota al valor absoluto o módulo, ie, $|x| = \max(x, -x)$ si $k = \mathbb{R}$ y $|x| = \sqrt{z\overline{z}} = \sqrt{a^2 + b^2}$ si $x = a+ib \in \mathbb{C}$ y $k = \mathbb{C}$.

Definición 4.1.2. — Sea V un k-espacio vectorial. Una norma de V es una función $\|\cdot\| : V \times V \to \mathbb{R}^{\geq 0}$ tal que para todos $x, y \in V$ y $\lambda \in \mathbb{R}$ se cumplen las siguientes proposicones:
1. $\|x\| = 0 \iff x = 0$.
2. $\|\lambda x\| = |\lambda|\|x\|$.
3. $\|x + y\| \leq \|x\| + \|y\|$ ("desigualdad triangular").

En tal caso, decimos que $(V, \|\cdot\|)$ (o que V) es un **espacio vectorial normado**
.

Ejemplo: Sea $V = \mathbb{R}^n$ o $\mathbb{C}^n$, definimos las normas:
1. $\|(x_1, ..., x_n)\|_1 = |x_1| + ... + |x_n|$.
2. $\|(x_1, ..., x_n\|_2 = \sqrt{|x_1|^2 + ... + |x_m|^2}$.
3. $\|(x_1, ..., x_m)\|_\infty = \max\{|x_1| + ... + |x_n|\}$.

Las propiedades 1,2,3 no son difíciles de probar para $\|\cdot\|_1$ y $\|\cdot\|_\infty$, pero la desigualdad triangular para la norma $\|\cdot\|_2$ es más sutil y la admitiremos de momento (aunque más adelante veremos que es una consecuencia de la "desigualdad de Cauchy Schwarz").

Ejercicio 4.1.3. — Probar que

$$\|x\|_\infty \leq \|x\|_2 \leq \|x\|_1 \leq n\|x\|_\infty,$$

para todo $x \in V = \mathbb{R}^n$ o $\mathbb{C}^n$.

Definición 4.1.4. — Sea V un $k-$espacio vectorial y sean $\|\cdot\|_a, \|\cdot\|_b$ dos normas sobre V. Decimos que las normas anteriores son **equivalentes** si existen $c_1, c_2 \in \mathbb{R}^{>0}$ tal que

$$c_1\|x\|_a \leq \|x\|_b \leq c_2\|x\|_a, \qquad \forall x \in V.$$

Hecho (Sin prueba): Si la dimensión de nuestro espacio es de k-dimensión finita, es decir, si $\dim_k(V) < +\infty$, ¡todas las normas son equivalentes! (Análisis I).
En todo lo que sigue, supondremos que V es un espacio vectorial de dimensión finita. En particular, la siguiente definición no depende de la norma escogida.

Definición 4.1.5. — Sea V un k-espacio vectorial y sea $(v_m)_{m\in\mathbb{N}}$ una sucesión de elementos de V. Diremos que $(v_m)_{m\in\mathbb{N}}$ converge a $\ell \in V$ si: Para todo $\epsilon > 0$ existe un $N \in \mathbb{N}$ tal que para cualquier $n \geq N$ se tiene que $\|v_n - \ell\| < \epsilon$.

Observación importante: Si $\mathscr{B}$ es una base de V, todo $v \in V$ se escribe de manera única como $v = x_1v_1 + ... + x_nv_n$ donde $(x_1, ..., x_n) \in k^n$ son las coordenadas de v respecto a la base $B = (v_1, ..., v_n)$. En otras palabras, la aplicación $\varphi_{\mathscr{B}} : k^n \to V$, $(x_1, ..., x_n) \mapsto x_1v_1 + ... + x_nv_n$ es un isomorfismo. Luego, toda norma $\|\cdot\|$ en V define una norma en k^n de la siguiente forma: Si $x \in k^n$, se puede definir $\|x\|_{k^n} := \|\varphi_{\mathscr{B}}(x)\|_{\mathscr{B}}$, donde $\|\cdot\|_{\mathscr{B}}$ es la norma que posee V. El proceso anterior es recíproco.
En particular, si escogemos la norma $\|\cdot\|_\infty$ en k^n tenemos que una sucesión $((x_{m,1}, ..., x_{m,n}))_{m\in\mathbb{N}}$ de elementos en k^n convergea a $\ell = (\ell_1, ..., \ell_n) \in k^n$ si y solo si la sucesión converge en cada componentes, ie, $(x_{m,k})_{m\in\mathbb{N}} \to \ell_k$ (ejercicio).
¡Importante!: La observación anterior implica que la mayoría de los resultados sobre sucesiones de números reales o complejos son también válidos para sucesiones vectoriales. Por ejemplo, una sucesión convergente es acotada, el límite de la suma es la suma de los límites, aunque esto no funciona con el producto, ya que aún no hemos definido un "producto" en espacios vectoriales de dimensión mayor.

Por otro lado, es importante notar que la definición de convergencia tiene el inconveniente de requerir conocer a priori el valor del límite $\ell \in V$. Es por ello que es útil conocer criterios de convergencia que puedan verificarse solamente analizando la sucesión $(v_m)_{m\in\mathbb{N}}$.

***Definición 4.1.6* (Sucesión de Cauchy).** — Sea $(V, \|\cdot\|)$ un espacio vectorial normado. Decimos que una sucesión $(v_m)_{m\in\mathbb{N}}$ es una sucesión de **Cauchy** si para cada $\epsilon > 0$ existe un $N \in \mathbb{N}$ tal que para cualquier $p, q \geq N$ se tiene que $\|v_p - v_q\| < \epsilon$.

Ejemplo: Si $(v_m)_{m\in\mathbb{N}}$ es una sucesión convergente, entonces es de Cauchy. En efecto, si $\ell \in V$ es el límite de la sucesión anterior, dado $\epsilon > 0$, por la convergencia sabemos que para cada $\epsilon/2 > 0$ existirá un $N \in \mathbb{N}$ tal que para cualquier $m \geq N$ se tiene que $\|v_m - \ell\| < \epsilon/2$. Por lo tanto, si $p, q \geq N$, de lo anterior tendremos que

$$\|v_p - v_q\| = \|(v_p - \ell) + (\ell - v_q)\| \underbrace{\leq}_{\text{Desigualdad triangular}} \|v_p - \ell\| + \|\ell - v_q\| \leq \epsilon/2 + \epsilon/2 = \epsilon,$$

verificando así la desigualdad deseada.
Conclusión: Toda sucesión convergente es de Cauchy. En general, el recíproco no es cierto, aunque existe un resultado muy importante en el cual sí se cumple el reciproco (Análisis I):
Hecho (Sin prueba): Si $\dim_k(V)$ es **finita**, toda sucesión de Cauchy es convergente.

Caso particular importante: Sea $(v_m)_{m\in\mathbb{N}}$ una sucesión en V. Definimos la sucesión de sumas parciales $(S_m)_{m\in\mathbb{N}}$ mediante:

$$S_m = \sum_{k=0}^{m} v_k = v_0 + v_1 + ... + v_m \in V.$$

Decimos que la serie vectorial $\sum_{k=0}^{\infty} v_k$ es convergente si $(S_m)_{m\in\mathbb{N}}$ converge. Gracias al criterio de Cauchy, la serie $\sum_{k=0}^{\infty}$ converge si y solo si para cada $\epsilon > 0$ existe $N \in \mathbb{N}$ tal que para $p, q \geq N$ se tiene

$$\|S_p - S_q\| = \|v_{q_1} + v_{q+2} + ... + v_p\| < \epsilon.$$

***Definición 4.1.7* (convergencia absoluta de series)**
Sea V un k-espacio vectorial (con $k = \mathbb{R}$ o $\mathbb{C}$). Decimos que una serie vectorial $\sum_{k=0}^{\infty} v_k$ es **absolutamente convergente** si la serie **real** $\sum_{k=0}^{\infty}\|v_k\|$ converge.

Proposición 4.1.8. — *Toda serie absolutamente convergente es convergente.*

Demostración. — Sea $\sum_{k=0}^{\infty} v_k$ una serie absolutamente convergente y sea $\epsilon > 0$. Usaremos el criterio de Cauchy: Dado que $\sum_{k=0}^{\infty}\|v_k\|$ es convergente (hipótesis), esta es de Cauchy, por lo que existe $n \in \mathbb{N}$ tal que para $p > q \geq N$ se tiene que $\sum_{k=q+1}^{p}\|v_k\| < \epsilon$. Aplicando la desigualdad triangular podemos obtener que

$$\left\| \sum_{k=q+1}^{p} v_k \right\| \leq \sum_{k=q+1}^{n} \|v_k\| < \epsilon.$$

Por tanto, la serie $(\sum_{k=0}^{m})_{m\in\mathbb{N}}$ es de Cauchy, y luego es convergente. $\square$

Caso particular importante: Si $V = \mathscr{M}_n(k)$ ($k = \mathbb{R}$ o $\mathbb{C}$), podemos considerar normas sobre $\mathscr{M}_n(k) \cong k^{n^2}$, sucesiones de matrices, series de matrices $\sum_{k=0}^{\infty} A_k$, etc.

Ejercicio 4.1.9. — Sea $(A_m)_{m\in\mathbb{N}}$ la sucesión de matrices en $\mathscr{M}_2(k)$ dada por

$$A_m = \begin{pmatrix} m\sin(1/m) & \frac{m}{m^2+1} \\ \frac{2m+3}{m} & 4 \end{pmatrix}.$$

Calcular (en caso de existir) $\lim\limits_{m\to\infty} A_m$.

¡Importante!: Notar que el producto en $\mathscr{M}_n(k)$ **no** es conmutitativo, y por lo tanto vale la pena redemostrar propiedades del límite del producto en este contexto.

Lema 4.1.10. — *Para todas $A, B \in \mathscr{M}_n(k)$ se tiene que $\|AB\|_1 \leq \|A\|_1\|B\|_1$.*

Demostración. — Si $A = (a_{ij})$ y $B = (b_{ij})$, entonces

$$\|AB\|_1 = \sum_{i,j}\left|\sum_k a_{ik}b_{kj}\right| \underbrace{\leq}_{\text{des. triang.}} \sum_{i,j,k}|a_{ik}||b_{kj}| \leq \left(\sum_{i,k}|a_{i,k}|\right)\left(\sum_{l,j}|b_{l,j}|\right) = \|A\|_1\|B\|_1.$$

$\square$

Observación 4.1.11. — Dado que $\mathscr{M}_n(k)$ es de dimensión finita, todas la normas son equivalentes, y por tanto podemos probar resultados de convergencia utilizando (sin pérdida de generalidad) la norma $\|\cdot\|_1$, la cual es útil por el lema anterior.

Lema 4.1.12. — *Sea $(V, \|\cdot\|)$ un espacio vectorial normado. Entonces, si $(x_n)_{n\in\mathbb{N}}$ es una sucesión convergente, esta es acotada, es decir, existe un real M tal que $\|x_n\| < M$ para cada natural n.*

Demostración. — Sea $(x_n)_{n\in\mathbb{N}}$ una sucesión convergente a x. Tomando $\epsilon = 1$, sabemos que existe un $N \in \mathbb{N}$ tal que para cualquier $m \geq N$ se tiene que $\|x - x_m\| < \epsilon = 1$. Esto nos dice que si $m \geq N$, entonces

$$\|x_m\| = \|(x_m - x) + x\| \leq \|x_m - x\| + \|x\| < 1 + \|x\|.$$

Finalmente, tomando $M = \max\{\|x_1\|, ..., \|x_N\|, 1 + \|x\|\}$ tendremos que si $n = \{1, ..., N\}$ entonces $\|x_n\| \leq M$ (pues M es mayor al máximo entre los $x_1, ..., x_N$), y si $n > N$, entonces $\|x_n\| \leq M$ ya que $M \geq 1 + \|x\|$ y este último término es una cota superior para los naturales mayores a N. $\square$

Proposición 4.1.13. — *Sean $(A_m)_{m\in\mathbb{N}}$ y $(B_m)_{m\in\mathbb{N}}$ dos sucesiones de matrices convergentes a A, B respectivamente. Entonces, la sucesión $(A_mB_m)_{m\in\mathbb{N}}$ converge a la matriz AB.*

Demostración. — Sea $\epsilon > 0$. Por el lema 4.1.12 las sucesiones $(A_m)_{m\in\mathbb{N}}$ y $(B_m)_{m\in\mathbb{N}}$ son acotadas, y luego existe un real M que acota simultáneamente a $(A_m)_m$ y $(B_m)_m$. En particular, dado que $\|A_m\|_1 \leq M$, tomando $m \to \infty$ se obtiene que $\|A\|_1 \leq M$. Así, por la convergencia de $(A_m)_m$ hacia A y $(B_m)_m$ hacia B, existe un natural N tal que si $m \geq N$, entonces

$$\|A_m - A\|_1 < \frac{\epsilon}{2M} \qquad y \qquad \|B_m - B\|_1 < \frac{\epsilon}{2M}$$

(esto tomando $\epsilon' := \frac{\epsilon}{2M}$ en la definición de convergencia). Luego, observe que si $m \geq N$, entonces

$$\begin{aligned}\|A_mB_m - AB\|_1 = \|(A_mB_m - AB_m) + (AB_m - AB)\|_1 &\underbrace{\leq}_{\text{des. triang}} \|(A_m - A)B_m\|_1 + \|A(B_m - B)\|_1 \\ &\underbrace{\leq}_{\text{lema}} \|A_m - A\|_1\|B_m\|_1 + \|A\|_1\|B_m - B\|_1 \\ &\leq \frac{\epsilon}{2M}\|B_m\|_1 + \frac{\epsilon}{2M}\|A\|_1 \\ &\leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.\end{aligned}$$

□

Ejercicio 4.1.14. — Sea $(A_m)_{m\in\mathbb{N}}$ una sucesión de matrices tales que $A_m \in \mathrm{GL}_n(k)$ para todo $m \in \mathbb{N}$. Supongamos que $(A_m)_{m\in\mathbb{N}}$ converge a $A \in \mathscr{M}_n(k)$ y que $(A_m^{-1})_{m\in\mathbb{N}}$ converge a $B \in \mathscr{M}_n(k)$. Probar que $A \in \mathrm{GL}_n(k)$ y que $A^{-1} = B$.

El producto de series convergentes de matrices es más delicado:

Proposición 4.1.15. — *Sean $\sum A_k$ y $\sum B_k$ series de matrices absolutamente convergentes con límites respectivos A y B. La series $\sum C_k$ de término general*

$$C_k := A_0B_k + A_1B_{k-1} + ... + A_kB_0$$

es absolutamente convergentes con límite AB.

Demostración. — Notar que para $m \in \mathbb{N}$ se tiene:

$$\sum_{k=0}^{m} \|C_k\| \leq \sum_{\substack{i,j\geq 0\\ i+j\geq m}} \|A_iB_j\|_1 \leq \sum_{\substack{i,j\geq 0\\ i+j\geq m}} \|A_i\|_i\|B_j\|_1 \leq \left(\sum_{k=0}^{m}\|A_i\|_1\right)\left(\sum_{k=0}^{m}\|B_j\|_1\right)$$
$$\leq \left(\sum_{k=0}^{\infty}\|A_i\|_1\right)\left(\sum_{k=0}^{\infty}\|B_j\|_1\right) < +\infty.$$

Es decir, las sumas parciales de la serie real $\sum_{k=0}^{m}\|C_k\|_1$ de términos positivos son acotados, y por tanto $\sum_{k=0}^{\infty}\|C_k\|$ es convergente, y luego $\sum_{k=0}^{\infty} C_k$ es absolutamente convergente. Finalmente, observemos que

$$\left\|\sum_{k=0}^{2m} C_k - \left(\sum_{i=0}^{m} A_i\right)\left(\sum_{j=0}^{m} B_j\right)\right\|_1 = \left\|\sum_{\substack{i>m \text{ ó } j>m\\ i+j\geq m}} A_iB_j\right\|_1 \leq \sum_{\substack{i>m \text{ ó } j>m\\ i+j\geq m}} \|A_i\|_1\|B_j\|_1$$
$$\leq \left(\sum_{i=0}^{2m}\|A_i\|_1\right)\left(\sum_{j=0}^{2m}\|B_j\|_1\right) - \left(\sum_{i=0}^{m}\|A_i\|_1\right)\left(\sum_{j=0}^{m}\|B_j\|_1\right).$$

Como el último término tiende a 0 cuando $m \to \infty$, y como el límite de la sucesión
$\left(\left(\sum_{i=0}^{m} A_i\right)\left(\sum_{j=0}^{m} B_j\right)\right)_{m\in\mathbb{N}}$ es AB, deduzimos que $\sum_{k=0}^{\infty} C_k = AB$.

$\square$

4.2. Exponencial de matrices

En cálculo en una y varias variables la exponencial juega un rol fundamental, sobre todo gracias a las propiedades de ser invariante bajo la derivada, es decir, que $\frac{d}{dt}e^{kt} = ke^{kt}$, o bien, que $\frac{d}{dt}Ce^{t} = Ce^{t}$. Es por ello que se busca generalizar la exponencial a otro tipo de espacios vectoriales, para así intentar obtener propiedades similares a las que ocurren en $\mathbb{R}$ o $\mathbb{C}$. No obstante, ¿Qué significiaría efectuar la operación e^B con B una matriz? A “priori” esto no tiene mucho sentido, pero haciendo uso de la serie de Taylor de la exponencial (recordando que $e^x = \sum_{n=0}^{\infty}\frac{x^n}{n!}$) se puede generalizar esta noción, para así rescatar propiedades interesantes que veremos a lo largo de la sección.
En lo que sigue, consideraremos k como el cuerpo de los números reales o complejos.

***Definición 4.2.1* (exponencial de matrices).** — Sea $A \in \mathscr{M}_n(k)$. Definimos la exponencial de la matriz A mediante la serie

$$\exp(A) = e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!} \in \mathscr{M}_n(k).$$

Observe que la exponencial de una matriz está bien definida, pues para todo $k \in \mathbb{N}$ se tiene que

$$\left\| \frac{A^k}{k!} \right\|_1 \leq \frac{\|A\|_1^k}{k!} \implies \sum_{k=0}^{\infty} \left\| \frac{A^k}{k!} \right\|_1 \leq \sum_{k=0}^{\infty} \frac{\|A\|_1^k}{k!} = e^{\|A\|_1}$$

y luego la serie es absolutamente convergente.

Ejemplo 4.2.2. —

1. Si $A = \left(\begin{array}{c:c:c} A_1 & 0 & 0 \\ \hdashline 0 & \ddots & 0 \\ \hdashline 0 & 0 & A_p \end{array}\right)$ es una matriz diagonal por bloques, entonces

$$A^k = \left(\begin{array}{c:c:c} A_1^k & 0 & 0 \\ \hdashline 0 & \ddots & 0 \\ \hdashline 0 & 0 & A_p^k \end{array}\right) \quad \forall k \in \mathbb{N} \implies e^A = \left(\begin{array}{c:c:c} e^{A_1} & 0 & 0 \\ \hdashline 0 & \ddots & 0 \\ \hdashline 0 & 0 & e^{A_p} \end{array}\right).$$

   En particular, si $A = \begin{pmatrix} \lambda_1 & & \text{\Large 0} \\ & \ddots & \\ \text{\Large 0} & & \lambda_n \end{pmatrix}$ es una matriz diagonal, entonces $e^A = \left(\begin{array}{c:c:c} e^{\lambda_1} & 0 & 0 \\ \hdashline 0 & \ddots & 0 \\ \hdashline 0 & 0 & e^{\lambda_n} \end{array}\right)$.

2. Si $A \in \mathscr{M}_n(k)$ es nilpotente de índice de nilpotencia $p \in \mathbb{N}^{\geq 1}$, ie, $A^p = 0$ pero $A^{p-1} \neq 0$, entonces $e^A = \sum_{k=0}^{p-1} \frac{A^k}{k!} = I_n + A + \frac{1}{2}A^2 + \dots + \frac{1}{(p-1)!}A^{p-1}$.

   Por ejemplo, si $A = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ cumple con $A^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ y $A^3 = 0$, por lo que $e^A = I_3 + A + \frac{1}{2}A^2 = \begin{pmatrix} 1 & 1 & 3/2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$. Si $x, y \in \mathbb{R}$ o $\mathbb{C}$, entonces $e^{x+y} = e^x e^y$. Esto último no es cierto en general para matrices: Si $e^{A+B} = e^A e^B$, entonces e^A y e^B deberían conmutar (Pues la suma $A + B = B + A$ es conmutativa). Sin embargo, considerando $\exp \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ y $\exp \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ no conmutan.

Proposición 4.2.3. — *Sean $A, B \in \mathscr{M}_n(k)$ dos matrices que conmutan, ie, $AB = BA$. Entonces,*

$$e^{A+B} = e^A e^B = e^B e^A$$

Demostración. — Dado que A y B conmutan, es válido el Teorela del Binomio de Newton, con lo que tenemos que

$$(A + B)^k = \sum_{i=0}^{k} \binom{k}{i} A^{k-i} B^i = \sum_{i=0}^{k} \frac{k!}{(k-i)!i!} A^{k-i} B^i.$$

Luego,

$$e^{A+B} = \sum_{k=0}^{\infty} \frac{(A + B)^k}{k!} = \sum_{k=0}^{\infty} \underbrace{\left(\sum_{i=0}^{k} \frac{A^{k-i}}{(k-j)!} \frac{B^i}{i!} \right)}_{:=C_k} = \sum_{k=0}^{\infty} C_k.$$

Visto (sección anterior): $\sum_{k=0}^{\infty} C_k$ converge al producto $\left(\sum_{i=0}^{\infty} \frac{A^i}{i!}\right) \left(\sum_{j=0}^{\infty} \frac{B^j}{j!}\right)$.

□

Corolario 4.2.4. — *Sea $A \in \mathscr{M}_n(k)$. Entonces $e^A \in \mathrm{GL}_n(k)$ y $(e^A)^{-1} = e^{-A}$.*

Demostración. — Basta aplicar la proposición a A y $B := -A$ (que conmutan), y notar que

$$I_n = e^{0_n} = e^{A+(-A)} = e^A + e^{-A}.$$

□

Proposición 4.2.5. — *Sean $A, B \in \mathscr{M}_n(k)$ dos matrices semejantes, ie, existe $P \in \mathrm{GL}_n(k)$ tal que $B = P^{-1}AP$. Entonces, e^A y e^B son semejantes. Más precisamente, $e^B = P^{-1}e^A P$.*

Demostración. — Mediante inducción no es difícil probar que $(P^{-1}AP)^k = P^{-1}A^kP$. Por lo tanto, $\sum_{k=0}^{m} \frac{(P^{-1}AP)^k}{k!} = P^{-1}\left(\sum_{k=0}^{m} \frac{A^k}{k!}\right)P$ y tomando $m \to \infty$ se tendrá que $e^B = P^{-1}AP$. □

Observación 4.2.6. — La proposición anterior implica que para todo endomorfismo $u : V \to V$ de un k-e.v. de dimensión finita ($k = \mathbb{R}$ o $\mathbb{C}$) podemos definir $\exp(u) = e^u \in \mathrm{End}_k(V)$, el cual es el endomorfismo asociado a $\exp(\mathrm{Mat}_{\mathscr{B}}(u))$ donde $\mathscr{B}$ es cualquier base de V.

Ejercicio 4.2.7. — Sea $A \in \mathscr{M}_n(k)$.

1. Probar que $\exp({}^tA) = {}^t\exp(A)$.
2. Probar que $\det(e^A) = e^{\mathrm{tr}(A)}$

Ejemplo 4.2.8. — Sea $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, entonces $P_A(x) = X^2-1 = (X-1)(X+1)$. Luego, A es diagonalizable ccon valores propios $\lambda_1 = -1$ y $\lambda_2 = 1$, y con vectores propios $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ y $v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ respectivamente. Además, $P = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$ y $P^{-1} = \begin{pmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \end{pmatrix}$ y luego $A = P\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}P^{-1}$. Esto implica, por la discusión anterior, que

$$e^A = P\begin{pmatrix} e^{-1} & 0 \\ 0 & e \end{pmatrix}P^{-1} = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} e^{-1} & 0 \\ 0 & e \end{pmatrix}\begin{pmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} e+e^{-1} & e-e^{-1} \\ e-e^{-1} & e+e^{-1} \end{pmatrix}$$

¡Atención! — Incluso si $A \in \mathscr{M}_n(\mathbb{C})$ no es diagonalizable, su polinomio característico escinde sobre $\mathbb{C}$, y por tanto admite una única descomposición de Dunford $A = S + N$ con S diagonalizable, N nilpotente y tal que $SN = NS$ (¡conmutan!). Luego, $e^A = e^{S+N} = e^S e^N$ donde e^S se calcula como en el ejemplo anterior y donde $e^N = I_n + N + \frac{1}{2}N^2 + ... + \frac{1}{(n-1)!}N^{n-1}$.

Caso particular: Si A posee solamente un valor propio λ, entonces su descomposición de Dunford es $A = \underbrace{\lambda I_n}_{S} + \underbrace{(A-\lambda I_n)}_{N}$, de donde obtemos que

$$e^A = e^{\lambda I_n}e^{A-\lambda I_n} = e^{\lambda}\left(I_n + (A-\lambda I_n) + \frac{(A-\lambda I_n)^2}{2} + ... + \frac{(A-\lambda I_n)^{n-1}}{(n-1)!}\right)$$

Caso general: A es semejante a una matriz diagonal por bloques, donde cada bloque posee solo un valor propio. Basta aplicar el procedimiento anterior en cada bloque.

Ejemplo 4.2.9. — Sea $A = \begin{pmatrix} -4 & 1 & 1 \\ 1 & -1 & -2 \\ -2 & 1 & -1 \end{pmatrix} \in \mathscr{M}_n(\mathbb{R})$. Notar que

$$\begin{aligned}
P_A(X) = \begin{vmatrix} X+4 & -1 & -1 \\ -1 & X+1 & 2 \\ 2 & -1 & X+1 \end{vmatrix} &\underset{c_3 \mapsto c_3 - c_2}{\overset{c_1 \mapsto c_1 + 2c_2}{=}} \begin{vmatrix} X+2 & -1 & 0 \\ 2X+1 & X+1 & -X+1 \\ 0 & -1 & X+2 \end{vmatrix} \\
&\underset{F_2 \mapsto F_2 + F_3}{\overset{F_1 \mapsto F_1 - F_3}{=}} \begin{vmatrix} X+2 & 0 & -(X+2) \\ 2X+1 & X & 3 \\ 0 & -1 & X+2 \end{vmatrix} \\
&= (X+2) \begin{vmatrix} 1 & 0 & -1 \\ 2X+1 & X & 3 \\ 0 & -1 & X+2 \end{vmatrix} \\
&= (X+2) \underbrace{[1(X^2+2X+3) + (-1)(-2X-1)]}_{X^2+4X+4} \\
&= (X+2)^3.
\end{aligned}$$

Se sigue que $\lambda = -2$ es el único valor propio de A. Por tanto,

$$e^A = e^{-2I_3} e^{A+2I_3} = e^{-2} \left(I_3 + (A+2I_3) + \frac{1}{2}(A+2I_3)^2 \right).$$

Calculamos $A + 2I_3 = \begin{pmatrix} -2 & 1 & 1 \\ 1 & 1 & -2 \\ -2 & 1 & 1 \end{pmatrix}$ y $(A+2I_3)^2 = \begin{pmatrix} 3 & 0 & -3 \\ 3 & 0 & -3 \\ 3 & 0 & -3 \end{pmatrix}$, de donde obtenemos $e^A = e^{-2} \begin{pmatrix} 1/2 & 1 & -1/2 \\ 5/2 & 2 & -7/2 \\ -1/2 & 1 & 1/2 \end{pmatrix}$.

Ejercicio 4.2.10. — Sea $A = \begin{pmatrix} 2 & -1 & 2 \\ 10 & -5 & 7 \\ 4 & -2 & 2 \end{pmatrix} \in \mathscr{M}_3(\mathbb{R})$. Calcular e^A (c.f. sección 3.8).

***Lema 4.2.11* (útil).** — *Sea* $J_m = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & 1 \\ 0 & \cdots & \cdots & 0 & 0 \end{pmatrix} \in \mathscr{M}_n(k)$ *bloque de Jordan nilpotente de tamaño* $m \times m$ *y sea* $t \in k$. *Entonces:*

$$\exp(tJ_m) = \begin{pmatrix} 1 & t & t^2/2 & \cdots & \frac{t^{m-1}}{(m-1)!} \\ 0 & 1 & t & \cdots & \frac{t^{m-2}}{(m-2)!} \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 1 & t \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix}.$$

Demostración. — Sea $\mathscr{B} = (e_1, ..., e_m)$ la base canónica de k^m. Entonces $Je_1 = 0$ y $JE_i = e_{i-1}$ para $i \in \{2, ..., m\}$. Se sigue que para cada $k \in \{1, ..., m-1\}$ tenemos que

$$J_m^k e_i = \begin{cases} 0 & si \quad i = 1, ..., k \\ e_{i-k} & si \quad i = k+1, ..., m \end{cases}$$

En otras palabras, $J_m^2 = \begin{pmatrix} 0 & 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \cdots & \cdots & 0 & 0 \\ 0 & \cdots & \cdots & 0 & 0 \end{pmatrix}$, ..., $J_m^{m-1} = \begin{pmatrix} 0 & \cdots & 0 & 1 \\ & \ddots & & 0 \\ \mathbf{0} & & & 0 \end{pmatrix}$, $J_m^m = 0$.

Esto nos dice que

$$\exp(tJ_m) = I_m + tJ_m + \frac{1}{2}(tJ_m)^2 + ... + \frac{1}{(m-1)!}(tJ_m)^{m-1}$$

$$= \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix} + \begin{pmatrix} 0 & t & & \\ & \ddots & \ddots & \\ & & \ddots & t \\ & & & 0 \end{pmatrix} + ... + \begin{pmatrix} 0 & \cdots & 0 & \frac{t^{m-1}}{(m-1)!} \\ & \ddots & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 \end{pmatrix}.$$

□

Proposición 4.2.12. — *Sea $m \in \mathbb{N}^{\geq 1}$ y $\lambda, t \in k$. Si consideramos el bloque de Jordan de tamaño m asociado a λ, ie,* $J_m(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & 1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{pmatrix}$.

Entonces:

$$\exp(tJ_m(\lambda)) = e^{\lambda t}\left(I_m + tJ_m + \frac{t^2}{2}J_m^2 + ... + \frac{t^{m-1}}{(m-1)!}J_m^{m-1}\right) = e^{\lambda t}\begin{pmatrix} 1 & t & t^2/2 & \cdots & \frac{t^{m-1}}{(m-1)!} \\ 0 & 1 & t & \cdots & \frac{t^{m-2}}{(m-2)!} \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 1 & t \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix}.$$

En particular, para toda matriz de Jordan $J = \left(\begin{array}{c:c:c} J_{\underline{n_1}}(\lambda_1) & 0 & 0 \\ \hdashline 0 & \ddots & 0 \\ \hdashline 0 & 0 & J_{\underline{n_p}(\lambda_p)} \end{array}\right)$ *se tiene que*

$$\exp(tJ) = \left(\begin{array}{c:c:c} e^{\lambda_1 t}\exp(J_{\underline{n_1}}) & 0 & 0 \\ \hdashline 0 & \ddots & 0 \\ \hdashline 0 & 0 & e^{\lambda_p t}\exp(J_{\underline{n_p}}) \end{array}\right)$$

Demostración. — La descomposición de Dunford de $J_n(\lambda)$ es $J_m(\lambda) = \lambda I_m + J_m$. Luego $\exp(tJ_m(\lambda)) = \exp(\lambda t I_m)\exp(tJ_m) = e^{\lambda t}\exp(tJ_m)$, por lo que la primera parte se obtiene gracias al lema útil. La última parte se obtiene al considerar cada bloque de Jordan por separado y ocupar lo anterior en cada uno de ellos. □

Terminaremos esta sección estudiando la continuidad y diferenciabilidad de la exponencial, lo cual será útil en aplicaciones de ella.

Recuerdo: Sea $\Omega \subseteq \mathbb{R}$ un subconjunto no vacío y $f : \Omega \to \mathbb{R}$, $t \mapsto f(t)$ una función real. Decimos que f es **continua** en $t_0 \in \Omega$ si "Para cada $\epsilon > 0$ existe un $\delta > 0$ tal que si $t \in \Omega$ y $|t - t_0| < \delta$, entonces $|f(t) - f(t_0)| < \epsilon$".
Si consideramos $\Omega \subseteq \mathbb{R}$ y $f : \Omega \to V$ una función con valores en el espacio vectorial normado $(V, \|\cdot\|)$, entonces la definición anterior se puede adaptar fácilmente de la siguiente manera:

***Definición 4.2.13* (función continua).** — Sea $\Omega \subseteq \mathbb{R}$ un sobconjunto no vacío y $(V, \|\cdot\|)$ un espacio vectorial normado. Diremos que $f : \Omega \to V$ es una función continua en $t_0 \subseteq \Omega$ si para cada $\epsilon > 0$ existe $\delta > 0$ tal que si $t \in \Omega$ y $|t - t_0| < \delta$, entonces $\|f(t) - f(t_0)\| < \epsilon$.
Diremos que f es continua si lo es para cada $t_0 \in V$.

Ejercicio 4.2.14. — Sabemos que si $\dim_k(V)$ es de dimensión finita entonces todas las normas son equivalentes. Pruebe, en tal caso, que la definición de continuidad no depende de la norma escogida en V.

Ejercicio 4.2.15. — Sea $V = \mathbb{R}^n$ ó $\mathbb{C}^n$. Pruebe que $f : \Omega \to V$, $t \mapsto f(t) = (f_1(t), f_2(t), ..., f_n(t))$ es continua si y solamente si cada una de las funciones componenetes f_i lo es.
Sugerencia: Utilice la norma $\|\cdot\|_\infty$.

Observación 4.2.16. — 1. Sea $V = \mathbb{R}^3$. Entonces la función $f(t) = (t^2, t^3, \sin(t))$ es continua en todo $\mathbb{R}$.

2. Sea $V = \mathscr{M}_2(\mathbb{R}) \cong \mathbb{R}^4$. Entonces $f(t) = \begin{pmatrix} t & t^2+1 \\ \sqrt{4} & e^t \end{pmatrix}$ es continua en $\Omega = [0, +\infty)$.

Observación 4.2.17. — La noción de límite es completamente análoga: Sea $\Omega \subseteq \mathbb{R}$ no vacío y $f : \Omega \to V$ y $t_0 \in \Omega$. Decimos que $\lim_{t \to t_0} f(t) = \ell$, para cierto $\ell \in V$ si: "Para todo $\epsilon > 0$ existe $\delta > 0$ tal que si $t \in \Omega$ y $|t - t_0| < \delta$, entonces $\|f(t) - \ell\| < \epsilon$".
Al igual que en el ejercicio 4.2.15, si $V = \mathbb{R}^n$ ó $\mathbb{C}^n$ y $\ell = (\ell_1, ..., \ell_n)$, al escoger la norma $\|\cdot\|_\infty$ se puede probar que si $f : \Omega \to V, t \mapsto f(t) = (f_1(t), ..., f(t))$, entonces $\lim_{t \to t_0} f(t) = \ell$ si y solamente si $\lim_{t \to t_0} f_i(t) = \ell_i$.
Este último hecho permite probar directamente que el límite de la suma es la suma de los límites.

Proposición 4.2.18. — *Sea $\Omega \subseteq \mathbb{R}$, y sean $F : \Omega \to \mathscr{M}_{m\times n}(\mathbb{C})$ y $G : \Omega \to \mathscr{M}_{n\times p}(\mathbb{C})$ funciones con valores matriciales tales que para $t_0 \in \Omega$ los límites $\lim\limits_{t\to t_0} F(t)$ y $\lim\limits_{t\to t_0} G(t)$ existen. Entonces la función $FG : \Omega \to \mathscr{M}_{m\times p}(\mathbb{C})$, $t \mapsto F(t)G(t)$ tiene un límite en $t_0 \in \Omega$ y además $\lim\limits_{t\to t_0} (F(t)G(t)) = \left(\lim\limits_{t\to t_0} F(t)\right)\left(\lim\limits_{t\to t_0} G(t)\right)$.*

Demostración. — La demostración es análoga a la hecha en la proposición 4.1.13. $\square$

Definición 4.2.19 **(diferenciabilidad de funciones)**
Sea $\Omega \subseteq \mathbb{R}$ no vacío y $(V, \|\cdot\|)$ un espacio vectorial normado. Decimos que f es diferenciable (o derivable) en $t_0 \in \Omega$ si $f : \Omega \to V$ es tal que el límite

$$\lim_{t\to t_0} \frac{f(t) - f(t_0)}{t - t_0}$$

existe. En tal caso, el valor del límite anterior lo denotamos como $f'(t_0) \in V$: la derivada de f en t_0.
Si f es diferenciable en cada punto, diremos simplemente que f es diferenciable.

Si $V = \mathbb{R}^n$ ó $\mathbb{C}^n$ y $f : \Omega \to V$, $t \mapsto f(t) = (f_1(t), ..., f_n(t))$, entonces la elección de la norma $\|\cdot\|_\infty$ permite demostrar que f es diferenciable en $t_0 \in \Omega$ si y solamente si las funciones $f_1, ..., f_n$ son diferenciables en t_0. En tal caso, $f'(t) = (f_1'(t), ..., f_n'(t))$.

Ejemplo 4.2.20. —

1. Sea $V = \mathbb{R}^3$ y $f(t) = (t^2, t^3, \sin(t))$, entonces f es derivable en todo $\mathbb{R}$ y $f'(t) = (2t, 3t^2, \cos(t))$.
2. Sea $V = \mathscr{M}_2(\mathbb{R}) \cong \mathbb{R}^4$ y $f(t) = \begin{pmatrix} t & t^2+1 \\ \sqrt{t} & e^t \end{pmatrix}$. Entonces f es derivable en $(0, +\infty)$ y $f'(t) = \begin{pmatrix} 1 & 2t \\ 1/(2\sqrt{t}) & e^t \end{pmatrix}$.

Observación 4.2.21. — Gracias a las propiedades de límites, se puede verificar directamente que la derivada de una suma es la suma de las derivadas.

Proposición 4.2.22. — *Sea $\Omega \subseteq \mathbb{R}$ no vacío y sean $F : \Omega \to \mathscr{M}_{m\times n}(\mathbb{C})$ y $G : \Omega \to \mathscr{M}_{n\times p}$ funciones con valores matriciales, definidas en un intervalo abierto en torno a $t_0 \in \Omega$ y derivable en t_0. Entonces FG es derivable en $t_0 \in \Omega$ y*

$$(FG)'(t_0) = F'(t_0)G(t_0) + F(t_0)G'(t_0).$$

Demostración. — Sea $\delta(t) = \frac{F(t)G(t)-F(t_0)G(t_0)}{t-t_0}$. Basta calcular $\lim_{t\to t_0} \delta(t)$, y lo haremos de la siguiente forma: Escribimos

$$\begin{aligned}\delta(t) &= \frac{F(t)G(t) - F(t_0)G(t) + F(t_0)G(t) - F(t_0)G(t_0)}{t-t_0} \\ &= \left(\frac{F(t)-F(t_0)}{t-t_0}\right) G(t) + F(t_0)\left(\frac{G(t)-G(t_0)}{t-t_0}\right),\end{aligned}$$

y luego, gracias al producto de límites de matrices obtenemos que $\lim\limits_{t\to t_0} \delta(t) = F'(t_0)G(t_0)+F(t_0)G'(t_0)$. □

☡**¡Muy importante!** Dado que el producto de matrices no es conmutativo, hay que prestar atención al orden de las matrices en la fórmula anterior.

Para estudiar la función exponencial haremos uso de la siguiente desigualdad:

Lema 4.2.23. — *Sea $A \in \mathscr{M}_n(k)$, entonces $\|e^A - I_n - A\| \leq \|A\|_1^2 e^{\|A\|_1}$.*

Demostración. — Tenemos que $e^A - I_n - A = \sum_{k=2}^{\infty} \frac{A^k}{k!}$. Considere el siguiente cálculo:

$$\|e^A - I_n - A\| \leq \sum_{k=2}^{\infty} \frac{\|A^k\|_1}{k!} = \sum_{k=0}^{\infty} \frac{\|A^{k+2}\|_1}{(k+2)!} \leq \|A\|_1^2 \sum_{k=0}^{\infty} \frac{\|A\|_1^k}{(k+2)!} \leq \|A\|_1^2 \sum_{k=0}^{\infty} \frac{\|A\|_1^k}{k!}$$
$$= \|A\|_1^2 e^{\|A\|_1}.$$

□

Teorema 4.2.24. — *Sea $A \in \mathscr{M}_n(k)$. La función $f : \mathbb{R} \to \mathscr{M}_n(k)$, $t \mapsto f(t) = e^{tA}$ es derivable en todo $\mathbb{R}$, y además $f'(t_0) = Ae^{t_0A} = e^{t_0A}A$. En particular, $f(t) = e^{tA}$ es de clase $\mathscr{C}^\infty$ y su derivada m-ésima es $f^{(m)}(t_0) = A^m e^{t_0A} = e^{t_0A}A$.*

Demostración. — Queremos probar que $\Delta(t) := \frac{e^{tA} - e^{t_0A}}{t - t_0} - Ae^{t_0A} = \frac{e^{tA} - e^{t_0A} - (t-t_0)Ae^{t_0A}}{t - t_0}$ tiende a 0 cuando $t \to t_0$. Dado que t_0A y $(t - t_0)A$ conmutan, tenemos que $e^{tA} = e^{(t-t_0)A + t_0A} = e^{(t-t_0)A}e^{t_0A}$. Sea $h := t - t_0$, entonces,

$$e^{tA} - e^{t_0A} - (t-t_0)Ae^{t_0A} = e^{(t-t_0)A}e^{t_0A} - e^{t_0A} - (t-t_0)Ae^{t_0A} = (e^{hA} - I_n - hA)e^{t_0A}.$$

El lema anterior aplicado a hA implica que

$$\|e^{tA} - e^{t_0A} - (t-t_0)Ae^{t_0A}\|_1 \leq \|e^{hA} - I_n - hA\|_1 \|e^{t_0A}\|_1 \leq h^2 \|A\|_1^2 e^{|h|\|A\|_1} \|e^{t_0A}\|_1.$$

Esto implica que $\|\Delta(t)\|_1 \leq h\|A\|_1^2 e^{|h|\|A\|_1} \|e^{t_0A}\|_1 \to 0$ cuando $h \to 0$, ie, $f'(t_0) = Ae^{t_0A}$. Por inducción obtenemos que $f^{(m)}(t_0) = A^m e^{t_0A}$ para todo $m \in \mathbb{N}$. Finalmente, notamos que

$$A^m \left(\sum_{k=0}^{N} \frac{(tA)^k}{k!} \right) = \left(\sum_{k=0}^{N} \frac{(tA)^k}{k!} \right) A^m \xrightarrow{N\to\infty} A^m e^{tA} = e^{tA} A^m.$$

□

4.3. Sistemas y EDOs lineales con coeficientes constantes

Las ecuaciones diferenciales son una de las ramas de la matemática más amplias y con grandes aplicaciones físicas. Algunos ejemplos de ellas son la ecuación del péndulo: $\ddot{\theta} + \frac{g}{\ell}\sin(\theta) = 0$, la ecuación de descarga de un condensador: $\dot{Q} - \frac{1}{RC}Q = 0$, la ecuación de caída libre: $\ddot{y} = g$, etc.
Es por ello que siempre se han buscado métodos para resolver ecuaciones diferenciales y, aunque esto no es siempre posible (de hecho, existen pocos casos en los que sí lo es), existen casos particulares muy importantes.
Por ejemplo, si $a \in \mathbb{C}$, resolver la EDO $x' = ax$ significa encontrar una función diferenciable $f : \mathbb{R} \to \mathbb{C}$ tal que para cada $t \in \mathbb{R}$ se cumpla que $f'(t) = af(t)$.

En este caso, si $f(t) = Ce^{at}$ con $C \in \mathbb{C}$, entonces $f'(t) = a(Ce^{at}) = af(t)$. Por otro lado, si $g(t)$ es una solución de $x' = ax$, definimos la función $F(t) = g(t)e^{-at}$, luego, $F'(t) = g'(t)e^{-at} - ag(t)e^{-at} = (g'(t) - ag(t))e^{-at} = 0$ para cada $t \in \mathbb{R}$. Esto dice que $F(t) = C$ con $C \in \mathbb{C}$ constante, y luego $g(t) = Ce^{at}$. El objetivo de esta sección es generalizar el cálculo anterior para dimensiones más altas.

***Definición 4.3.1* (sistema diferencial lineal).** — Un **sistema diferencial lineal** (homogéneo) con coeficientes constantes es un sistema de la forma

$$(S)\left\{\begin{array}{lcl} x_1'(t) & = & a_{11}x_1(t) + a_{12}x_2(t) + \ldots + a_{1n}x_n(t) \\ x_2'(t) & = & a_{21}x_1(t) + a_{22}x_2(t) + \ldots + a_{2n}x_n(t) \\ \quad\vdots & & \\ x_n'(t) & = & a_{n1}(t) + a_{n2}x_2(t) + \ldots + a_{nn}x_n(t) \end{array}\right.$$

donde los $a_{ij} \in \mathbb{C}$ son constantes y las funciones derivables $x_i : \mathbb{R} \to \mathbb{C}$ son las incógnitas.

El sistema (S) es equivalente a $X'(t) = AX(t)$, donde $A = (a_{ij}) \in \mathscr{M}_n(\mathbb{C})$ y $X(t) = \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix}$.

Teorema 4.3.2. — *Las soluciones de (S) son las funciones de la forma $X(t) = e^{tA}X_0$, donde $X_0 \in \mathbb{C}^n$. El conjunto de soluciones es un $\mathbb{C}$-e.v. de dimensión n.*

Demostración. — Veamos que toda función de la forma $X(t) = e^{tA}X_0$ es solución de (S): La fórmula para la derivada de un producto de matrices implica que

$$X'(t) = Ae^{tA}X_0 + e^{tA} \cdot 0 = Ae^{tA}X_0 = AX(t),$$

concluyendo lo deseado. Por otro lado, si $X(t)$ es una solución de (S), definimos la función $F(t) := e^{-tA}X(t)$. Luego,

$$\begin{aligned} F'(t) &= -Ae^{-tA}X(t) + e^{-tA}X'(t) \\ &= -e^{-tA}AX(t) + e^{-tA}X'(t) \\ &= e^{-tA}(AX(t) - X'(t)) \\ &= 0. \end{aligned}$$

Se sigue que $F(t) = C \in \mathbb{C}^n$. En particular, $C = F(0) = X(0)$, por lo que $e^{-tA}X(t) = X(0)$. Puesto que $e^{-tA} \in \mathrm{GL}_n(\mathbb{C})$ y $(e^{-tA})^{-1} = e^{tA}$, obtenemos que $X(t) = e^{tA}X(0)$.
Finalmente, sea $V := \{X : \mathbb{R} \to \mathbb{C}^n \text{ derivable tq } X' = AX\}$ conjunto de soluciones de (S). Entonces, la aplicación $\varphi : \mathbb{C}^n \to V$ dada por $\varphi(X_0) = e^{tA}X_0$ es lineal es inyectiva, con inversa dada por $X \mapsto X(0) \in \mathbb{C}^n$. Luego $V \cong \mathbb{C}^n$. $\square$

Corolario 4.3.3. — *Sean $X_0 \in \mathbb{C}^n$ y $t_0 \in \mathbb{R}$. La función $X(t) = e^{(t-t_0)A}X_0$ es la* **única** *solución de (S) que verifica $X(t_0) = X_0$.*

Demostración. — La función $X(t) = e^{tA}(e^{-t_0A}X_0)$ es solución de (S) gracias al teorema anterior, y ella verifica que $X(t_0) = e^{0_n}X_0 = I_nX_0 = X_0$.
Por otro lado, si $Y(t)$ es una solución de (S) tal que $Y(t_0) = X_0$, entonces la función $Y(t + t_0)$ es solución de (S) y vale X_0 en $t = 0$. La demostración anterioir implica que $Y(t + t_0) = e^{tA}X_0$, es decir, $Y(t) = e^{(t-t_0)A}X_0$. $\square$

Nos gustaría tener una descripción más explícita de las soluciones de (S). Veamos primero el caso más simple: **A es diagonalizable.**

Teorema 4.3.4. — *Supongamos que existe una base $(v_1, ..., v_n) \in \mathbb{C}^n$ formada por vectores propios de A asociados a valores propios $\lambda_1, ..., \lambda_n$ (no necesariamente distintos), ie, A es diagonalizable. Entonces las soluciones de (S) son de la forma*

$$X(t) = c_1e^{\lambda_1 t} + c_2e^{\lambda_2 t} + ... + c_ne^{\lambda_n t}v_n,$$

donde $c_1, ..., c_n \in \mathbb{C}^n$ son constantes.

Demostración. — Toda "condición inicial" $X_0 \in \mathbb{C}^n$ se escribe como $X_0 = \sum_{j=1}^n c_jv_j$. Luego, si $P \in \mathrm{GL}_n(\mathbb{C})$ es la matriz de cambio de base de la base canónica a la base $(v_1, ..., v_n)$, entonces $X_0 = P \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$. Por otro lado, $P^{-1}AP = D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$ matriz diagonal. Se sigue que

$$e^{tA}X_0 = Pe^{tD}P^{-1}X_0 = Pe^{tD}\begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = P\begin{pmatrix} c_1e^{\lambda_1 t} \\ \vdots \\ c_ne^{\lambda_n t} \end{pmatrix} = \sum_{j=1}^{n} c_je^{\lambda_j t}v_j.$$

(otra forma: En la base $(v_1,...,v_n)$ el sistema (S) se reduce a $y_j' = \lambda_j y_j$; $j = 1,..,n$). $\square$

Ejemplo: Consideremos el sistema

$$\begin{cases} x_1' &= 2x_2 - 2x_3 \\ x_2' &= -2x_1 + x_3 \\ x_3' &= 2x_1 - x_2 \end{cases} \iff \begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 2 & -2 \\ -2 & 0 & 1 \\ 2 & -1 & 0 \end{pmatrix}}_{A}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

Ejercicio $P_A(X) = X^3 + 9X = X(X+3i)(X-3i)$ y los vectores propios $v_1 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$, $v_2 = \begin{pmatrix} 4 \\ -1+3i \\ -1-3i \end{pmatrix}$ y $v_3 = \overline{v_2} = \begin{pmatrix} 4 \\ -1-3i \\ -1+3i \end{pmatrix}$ están asociados a $\lambda_1 = 0$, $\lambda_2 = 3i$ y $\lambda_3 = \overline{\lambda_2} = -3i$. Luego, las soluciones complejas del sistema son

$$\begin{cases} x_1(t) &= c_1 + 4c_2e^{3it} + 4c_3e^{-3it} \\ x_2(t) &= 2c_1 + (-1+3i)c_2e^{3it} + (-1-3i)c_3e^{-3it} \\ x_3(t) &= 2c_1 + (-1-3i)c_2e^{3it} + (-1+3i)e^{-3it} \end{cases}$$

donde $c_1, c_2, c_3 \in \mathbb{C}$, ie, una base del espacio de soluciones está dada por

$$X_1(t) = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \quad X_2(t) = \begin{pmatrix} 4 \\ -1+3i \\ -1-3i \end{pmatrix}e^{3it}, \quad X_3(t) = \overline{X_2(t)} = \begin{pmatrix} 4 \\ -1-3i \\ -1+3i \end{pmatrix}e^{-3it}.$$

❦ Para encontrar soluciones **reales** del sistema, notamos que las soluciones $X_1(t), (X_2(t) + X_3(t))/2 = \Re(X_2)(t)$ y $(X_2(t) - X_3(t))/2i = \Im(X_2)(t)$

son reales y linealmente independientes. Luego, toda solución real es combinación lineal con coeficientes reales de

$$X_1(t) = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \quad \Re(X_2)(t) = \begin{pmatrix} 4 \\ -1 \\ -1 \end{pmatrix} \cos(3t), \quad \Im(X_2)(t) = \begin{pmatrix} 4 \\ 3 \\ -3 \end{pmatrix} \sin(3t).$$

Veamos ahora el caso general. Como siempre, si λ_j es un valor propio de A entonces n_j es su multiplicidad como raíz del polinomio característico P_A y m_j su multiplicidad como raíz del polinomio minimal m_A, es decir,

$$m_A(X) = \prod_{j=1}^{p} (X - \lambda_j)^{m_j} \qquad y \qquad P_A(X) = \prod_{j=1}^{p} (X - \lambda_j)^{n_j}.$$

Más áun, $0 < m_j \leq n_j$ (Cayley-Hamilton).

Teorema 4.3.5. — *Sean $\lambda_1, ..., \lambda_p$ los valores propios (distintos) de A y sean $m_1, ..., m_p$ sus multiplicidades como raices del polinomio minimal m_A. Entonces, las soluciones de (S) son de la forma*

$$X(t) = e^{\lambda_1 t} P_1(t) + ... + e^{\lambda_p t} P_p(t),$$

donde $P_j(t)$ es una función (vectorial) polinomial de grado $< m_j$ con valores en el espacio característico $V_{(\lambda_j)}$ asociado al valor propio λ_j.

Demostración. — Recordemos que $V_{(\lambda_j)} = \ker[(A - \lambda_j I_n)^{m_j}]$ y que $\mathbb{C}^n = V_{(\lambda_1)} \oplus ... \oplus V_{(\lambda_p)}$. Por otra parte, toda solución $X(t)$ de (S) es de la forma $X(t) = e^{tA} X_0$ para cierto $X_0 \in \mathbb{C}^n$. Escribamos $X_0 = v_1 + ... + v_p$ para únicos vectores $v_j \in V_{(\lambda_j)}$. Notar que $e^{tA} v_j = e^{\lambda_j t} e^{A - \lambda_j I_n)t} v_j$ y luego, dado que $(A - \lambda_j I_n)^{m_j}$ anula a todo vector $v_j \in V_{(\lambda_j)}$, deducimos que:

$$e^{tA} v_j = e^{\lambda_j t} \sum_{k=0}^{m_j - 1} \frac{((A - \lambda_j I_n)t)^k}{k!} v_j,$$

por lo que $X(t) = \sum_{j=1}^{p} e^{\lambda_j t} P_j(t)$, donde $P_j(t) = \displaystyle\sum_{k=0}^{m_j - 1} \frac{t^k}{k!} (A - \lambda_j I_n)^k v_j$. □

Observación 4.3.6. — A es diagonalizable ssi $m_j = 1$ para todo j. En part, $V_{(\lambda_j)} = V_{\lambda_j}$ y $P_j(t)$ es un polinomio constante con valores en V_{λ_j} (cf. 4.3.4).

Ejemplo: Consideremos el sistema

$$\left\{\begin{array}{rcl} x_1' &=& 2x_1 - x_2 + 2x_3 \\ x_2' &=& 10x_1 - 5x_2 + 7x_3 \\ x_3' &=& 4x_1 - 2x_2 + 2x_3 \end{array}\right. \iff \begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix} = \underbrace{\begin{pmatrix} 2 & -1 & 2 \\ 10 & -5 & 7 \\ 4 & -2 & 2 \end{pmatrix}}_{=A} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

Vimos en 3.8 que $P_A(X) = X^2(X+1)$ y luego $\lambda_1 = -1$ y $\lambda_2 = 0$ son sus valores propios. Además, $V_{\lambda_1} = V_{(\lambda_1)} = \ker(A+I_3) = \mathrm{Vect}_{\mathbb{R}}(v_1)$, con $v_1 = (1, -1, -2)^t$ y $V_{\lambda_2} = \ker(A) = \mathrm{Vect}_{\mathbb{R}}(v_2)$ y $V_{(\lambda_2)} = \ker(A^2) = \mathrm{Vect}_{\mathbb{R}}(v_1, v_2)$, con $v_2 = (1, 2, 0)^t$ y $v_3 = (0, 1, 1)^t$.

Método 1: Gracias a la demostración del lema anterior, tenemos que una base para el espacio de soluciones del sistema está dado por:

$$X_1(t) = \underbrace{e^{\lambda_1 t} v_1}_{=e^{tA}v_1} = \begin{pmatrix} 1 \\ -1 \\ -2 \end{pmatrix} e^{-t}, \qquad X_2(t) = \underbrace{e^{\lambda_2 t} v_2}_{=e^{tA}v_2} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} \qquad y \qquad X_3(t) = e^{tA} v_3 = \underbrace{e^{\lambda_2 t}}_{=1}(v_3 + t \underbrace{Av_3}_{v_2})$$

$$= \begin{pmatrix} t \\ 2t+1 \\ 1 \end{pmatrix}$$

Luego, la solución del sistema es

$$\left\{\begin{array}{rcl} x_1(t) &=& c_1 e^{-t} + c_2 + c_3 t \\ x_2(t) &=& -c_1 e^{-t} + 2c_2 + c_3(2t+1) \\ x_3(t) &=& -2c_1 e^{-t} + c_3 \end{array}\right.$$

Método 2: Calcular explícitamente $X(t) = e^{tA} X_0$: Dado que $\mathscr{B} = (v_1, v_2, v_3)$ es una base, escribimos $X_0 = c_1 v_1 + c_2 v_2 + c_3 v_3$, por lo que $X_0 = P \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$, con

$$P = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 2 & 1 \\ -2 & 0 & 1 \end{pmatrix}. \text{ Más aún, } P^{-1}AP = \left(\begin{array}{c:cc} -1 & 0 & 0 \\ \hdashline 0 & 0 & 1 \\ 0 & 0 & 0 \end{array}\right) = J \text{ forma canónica}$$

de Jordan. Se sigue que

$$e^{tA} = Pe^{tJ}P^{-1}, \text{ donde } e^{tJ} = \left(\begin{array}{c|cc} e^{(-1)t} & & \\ \hline & e^{0t} & e^{0t}t \\ & 0 & e^{0t} \end{array}\right) = \left(\begin{array}{c|cc} e^{-t} & & \\ \hline & 1 & t \\ & 0 & 1 \end{array}\right) \text{ cf 4.2.11}$$

Luego, $e^{tA}X_0 = Pe^{tJ}P^{-1}X_0$, $Pe^{tJ}\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = P\begin{pmatrix} c_1e^{-t} \\ c_2 + c_3t \\ c_3 \end{pmatrix}$ y la solución del sistema estará dada por

$$\begin{cases} x_1(t) &= c_1e^{-t} + c_2 + c_3t \\ x_2(t) &= -c_1e^{-t} + 2c_2 + c_3(2t+1) \\ x_3(t) &= -2c_1e^{-t} + c_3 \end{cases}$$

Ejemplo: Consideremos el sistema

$$\begin{cases} x_1' &= -4x_1 + x_2 + x_3 \\ x_2' &= x_1 - x_2 - 2x_3 \\ x_3' &= -2x_1 + x_2 - x_3 \end{cases} \iff \begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix} = \underbrace{\begin{pmatrix} -4 & 1 & 1 \\ 1 & -1 & -2 \\ -2 & 1 & -1 \end{pmatrix}}_{=A} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

Vimos en 4.2 que $P_A(X) = (X+2)^3$ y luego $\lambda = -2$ es el único valor propio de A. Las soluciones del sistema están dadas por $X(t) = e^{tA}X_0$. Dado que $(A + 2I_3)^3 = 0$ (Cayley-Hamilton), tenemos que:

$$e^{tA} \underset{\substack{\uparrow \\ \text{Dunford}}}{=} e^{-2I_3t}e^{(A+2I_3)t} = e^{-2t}(I_3 + (A+2I_3)t + \frac{1}{2}t^2(A+2I_3)^2).$$

Dado que $A+2_3=\begin{pmatrix}-2 & 1 & 1\\ 1 & 1 & -2\\ -2 & 1 & 1\end{pmatrix}$ y $(A+2I_3)^2=\begin{pmatrix}3 & 0 & -3\\ 3 & 0 & -3\\ 3 & 0 & -3\end{pmatrix}$ obtenemos que la solución del sistema está dada por

$$\begin{cases} x_1(t) &= & e^{-2t}\left[\left((1-2t+\frac{3t^2}{2}\right)c_1+tc_2+\left(t-\frac{3t^2}{2}\right)c_3\right] \\ x_2(t) &= & e^{-2t}\left[\left(t+\frac{3t^2}{2}\right)c_1+(1+t)c_2-\left(2t+\frac{3t^2}{2}\right)c_3\right] \\ x_3(t) &= & e^{-2t}\left[\left(-2t+\frac{3t^2}{2}\right)c_1+tc_2+\left(1+t-\frac{3t^2}{2}\right)c_3\right] \end{cases}$$

Consejos prácticos: Para resolver el sistema diferencial $X'(t)=AX(t)$:

1. En primer lugar, calculamos los valores y vectores propios de A. Si los vectores propios forman una base (ie, A es diagonalizable) entonces el teorema 4.3.4 da la solución.
2. Si A **no** es diagonalizable pero sólo posee un valor propio λ, entonces
$$e^{tA} \underset{\substack{\uparrow \\ \text{Dunford}}}{=} e^{\lambda t}e^{t(A-\lambda I_n)} = e^{\lambda t}\sum_{k=0}^{n-1}(A-\lambda I_n)^k\frac{t^k}{k!}$$
3. Si A **no** es diagonalizable y posee varios valores propios $\lambda_1,\ldots,\lambda_p$ con multiplicidades algebraicas $n_1,\ldots,n_p$ entonces:
   a) Podemos buscar para cada λ_j un polinomio de grado $< n_j$, $P_j(t)$, tal que $e^{\lambda_j t}P_j(t)$ sea solución del sistema (obtendremos al reemplazar un sistema lineal en los coeficientes de P_j).

   Observación 4.3.7. — De hecho, el teorema 4.3.5 nos dice que basta considerar P_j de grado $< m_j$, la multiplicidad de λ_j en el polinomio minimal m_A.

   b) Podemos buscar una base $(v_j^1, v_j^2, \ldots)$ de $V_{(\lambda_j)} = \ker\left[(A-\lambda_j I_n)^{m_j}\right]$ de tal suerte que $v_j^\ell \in \ker\left[(A-\lambda_j I_n)^\ell\right]$ y luego calcular $e^{tA}v_j^\ell$ como en (2).

   Observación 4.3.8. — Un caso particular típico es $n_j = 2$, en cuyo caso hay que encontrar v_j vector propio y v_j' vector propio generalizado en $\ker\left[(A-\lambda_j I_n)^2\right]$ tal que v_j y v_j' sean l.i. En este caso obtenemos soluciones
$$X_j(t)=e^{\lambda_j t}v_j \quad y \quad X_j'(t)=e^{\lambda_j t}[v_j'+t(A-\lambda_j I_n)v_j'].$$

c) Podemos calcular la forma canónica de Jordan de $A = PJP^{-1}$ y obtener e^{tJ} usando lo visto anteriormente.

Observación 4.3.9. — Vale la pena señalar que **no** es necesario calcular la inversa P^{-1} para encontrar la solución general (c.f. 4.3).

4. Si la matriz $A \in \mathscr{M}_n(\mathbb{R})$ tiene coeficientes **reales**, sus valores propios pueden agruparse en valores propios reales $\lambda_1, ..., \lambda_p \in \mathbb{R}$ y en pares conjugados de valores propios complejos $\mu_1, \underline{\mu_1}, ..., \mu_q, \underline{\mu_q} \in \mathbb{C}$. Es importante (para ahorrar cálculos) recordar que espacios propios/característicos de valores propios conjugados son también conjugados. Luego, las soluciones del sistema serán del tipo $X_1(t), ..., X_p(t), Y_1(t), \overline{Y_1(t)}, ..., Y_q(t), \overline{Y_q(t)}$, con $X_j(t)$ real. Así, una base de soluciones **reales** es:
$$X_1(t), ..., X_p(t), (\Re Y_1)(t), (\Im Y_1)(t), ..., (\Re Y_q)(t), (\Im Y_q)(t).$$

Ejemplo 4.3.10. — Resolver el sistema diferencial
$$\begin{cases} x' &= \quad x + y \\ y' &= -4x - 3y \end{cases}$$

Definición 4.3.11. — Una EDO homogénea lineal de orden n con coeficientes constantes es
$$(E) \quad x^{(n)}(t) + a_{n-1}x^{(n-1)}(t) + ... + a_1 x'(t) + a_0 x(t) = 0,$$
donde los $a_j \in \mathbb{C}$ son constantes y $x : \mathbb{R} \to \mathbb{C}$ es una función n-veces derivable.

Si definimos $x_0 := x$, $x_1 = x'$, $x_2 = x''$,...,$x_{n-1} = x^{(n-1)}$ entonces $\underbrace{(E) \iff X'(t) = AX(t)}_{}$ donde

(¡¡Podemos aplicar todos los resultados anteriores!!)

$$X(t) = \begin{pmatrix} x_0(t) \\ \vdots \\ x_{n-1}(t) \end{pmatrix} \quad y \quad A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & & & & 1 \\ -a_0 & -a_1 & \cdots & & -a_{n-1} \end{pmatrix}.$$

Luego, el conjunto de soluciones de (E) es un $\mathbb{C}$-e.v. de dimensión n. Más aún, vimos en 3.8 que $P_A(X) = m_A(X) = X^n + a_{n-1}X^{n-1} + ... + a_1 X + a_0$, por lo que P_A es el "polinomio característico de (E)" y "$P_A(\lambda) = 0$" es la ecuación característica de (E).

Teorema 4.3.12. — *Sean $\lambda_1, ..., \lambda_p$ las raíces (distintas) de la ecuación característica y sean $n_1, ..., n_p$ sus multiplicidades. Entonces, las soluciones de (E) son* **todas** *las funciones de la forma*

$$X(t) = Q_1(t)e^{\lambda_1 t} + ... + Q_p(t)e^{\lambda_p t},$$

donde $Q_j \in \mathbb{C}[t]$ es un polinomio de grado $< n_j$.

Demostración. — Toda solución de $X' = AX$ es de la forma

$$X(t) = e^{\lambda_1 t}P_1(t) + ... + e^{\lambda_p t}P_p(t)$$

donde $P_j(t)$ es un polinomio de grado $< m_j = n_j$ (pues $m_A = P_A$). Al considerar la primera componente del vector $X(t)$, obtenemos que toda solución de (E) es de la forma $x(t) = e^{\lambda_1 t}P_{11}(t) + e^{\lambda_2 t}P_{21}(t) + ... + e^{\lambda_p t}P_{p1}(t)$, donde $P_{j1}(t)$ es un polinomio de grado $< n_j$. En otras palabras, el espacio de soluciones está contenido en el espacio vectorial generado por los $n = n_1 + ... + n_p$ funciones $t \mapsto t^k e^{\lambda_j t}$, para $0 \leq k < n_j$ y $j = \{1, ..., p\}$. Como el espacio de soluciones es de dimensión n, ambos espacios son iguales. $\square$

Ejemplo:

1. $x''' + 3x'' + 3x' + x = 0 \implies P(\lambda) = (\lambda + 1)^3$ y luego las soluciones son todas las funciones $x(t) = (at^2 + bt + c)e^{-t}$, con $a, b, c \in \mathbb{C}$.

2. $x^{(4)} + 2x'' + x = 0 \implies P(\lambda) = (\lambda^2 + 1)^2$ y luego las soluciones son todas las funciones $x(t) = (at + b)e^{it} + (ct + d)e^{-it}$ con $a, b, c, d \in \mathbb{C}$. En particular, las soluciones **reales** son $x(t) = (at + b)\cos(t) + (ct + d)\sin(t)$ con $a, b, c, d \in \mathbb{R}$.

Ejercicio: Encontrar todas las soluciones de la EDO $x^{(n)} = x$.

Hint: Tratar por separado el caso n par y el caso n impar.

CAPÍTULO 5

DUALIDAD Y FORMAS BILINEALES

5.1. Espacio vectorial cociente

Uno de los objetos matemáticos más importantes de este curso, el cual se ha venido estudiando desde el comienzo, es el de aplicación lineal. Esto debido a que corresponden a los **morfismos** de k-espacios vectoriales, es decir, a las funciones entre espacios de este tipo, que preservan su estructura propia-. Entonces, dada una aplicación lineal $f : V \to W$ vimos que $\ker(f) \subseteq V$ es un sub-espacio vectorial. Luego surge la interrogante, ¿será cierto que para todo sub-espacio vectorial existe una aplicación lineal cuyo núcleo sea dicho sub-espacio?
Sea entonces $f : V \to W$ lineal y denotemos $K = \ker(f) \subseteq V$. Considere $w \in \Im(f) \subseteq W$. Entonces por definición $\exists v \in V$ tal que $f(v) = w$. Si $v' \in V$ es otro vector tal que $f(v') = w$, por linealidad se tiene que $f(v) = f(v') \iff v - v' \in K$. Esto motiva definir la siguiente relación de equivalencia:

$$v \sim v' \iff (v - v') \in K$$

Así, se tienen las siguientes clases de equivalencia:

$$\begin{aligned}[v] &= \{v' \in V \mid v \sim v'\} = \{v' \in V \mid v' - v \in K\} =: v \pmod K \\ &= \{v + x \mid x \in K\} =: v + K\end{aligned}$$

Notar que las notaciones introducidas son bastante intuitivas. Primero, la notación de módulo viene a indicar que $v \in K \iff [v] = [0]$, y por otro lado, la notación $v + K$ viene del hecho que si $v' \in [v] \Rightarrow \exists x \in K$ tal que $v' = v + x$ (de alguna forma estamos sumando a cada vector todos los elementos de K).
En conclusión, cada $w \in \Im(f) \subseteq W$ se identifica con una única clase de equivalencia $[v]$ donde $f(v) = w$.

La idea es extender esta construcción a cualquier sub-espacio $U \subseteq V$. Para ello, comenzaremos por la construcción en grupos abelianos (ver 1.1.1), para luego extenderla. Sea entonces A grupo abeliano y $B \subseteq A$ sub-grupo. Se define en A la relación de equivalencia $\sim_B$ como sigue: $x \sim_B y \iff x - y \in B$. Entonces $\forall a \in A$ se tienen las clases

$$[a] = \{b \in A \mid b - a \in B\} = \{a + x\}_{x \in B} =: a \pmod{B} =: a + B$$

Entonces, denotamos por A/B al conjunto cociente por la relación de equivalencia descrita anteriormente. Más aún, es posible definir una función, llamada la **proyección canónica** como la función

$$\pi : A \to A/B \qquad a \mapsto [a] = a + B$$

Proposición 5.1.1. — *Sea A un grupo abeliano y $B \subseteq A$ sub-grupo. Entonces el conjunto cociente A/B admite una única estructura de grupo tal que la proyección canónica $\pi : A \to A/B$ es un* **morfismo de grupos***, es decir,*

$$\pi(a + b) = \pi(a) + \pi(b) \quad \forall a, b \in A$$

Demostración. — Debido a la condición que se desea obtener sobre π, buscamos que

$$[a] + [b] = (a + B) + (b + B) = \pi(a) + \pi(b) = \pi(a + b) = (a + b) + B = [a + b]$$

por lo que nos interesa definir $[a]+[b] := [a+b]$. Pero, esto no es necesariamente cierto (no tiene por qué ser cierto para todo $a' \in [a]$). Así, se debe probar que $[a] + [b] := [a + b]$ no depende de los elementos escogidos. En efecto, sean $a', b' \in A$ tal que $a' \sim_B a, b' \sim_B b$. Luego existen $x, y \in B$ tales que $a' = a + x, b' = b + y$. Debido a que $(A, +)$ es grupo abeliano, entonces:

$$a' + b' = (a + x) + (b + y) = (a + b) + \underbrace{(x + y)}_{\in B} \Rightarrow (a + b) \sim_B (a' + b')$$

De esta forma, $[a + b] =: [a] + [b] = [a'] + [b'] := [a' + b']$ está bien definida. También es posible notar que la suma en A/B hereda la asociatividad y conmutatividad directamente de la suma en A, y posee un neutro dado por $[0] = 0 + B = B$. Además, el opuesto aditivo corresponde a $[-a]$ (ya que $[-a] + [a] = [0]$). Todo lo anterior, dota al conjunto A/B de una estructura de grupo abeliano. □

Observación 5.1.2. — Dado que $\forall a \in A$ se tiene que $a \in [a]$, claramente se tiene que la proyección canónica $\pi : A \twoheadrightarrow A/B$ es siempre sobreyectiva. Sin embargo, esta función no siempre (de hecho, rara vez) es inyectiva, por lo que se invita al lector a reflexionar sobre esto.

Ejercicio 5.1.3. — Determinar cuando la proyección canónica $\pi : A \twoheadrightarrow A/B$ es inyectiva.

Teorema 5.1.4. — *Sea V un k-espacio vectorial y $U \subseteq V$ subespacio. Entonces:*

1. *Es posible dotar al grupo abeliano V/U de una única estructura de k-espacio vectorial tal que la proyección $\pi : V \to V/U$ sea lineal, ie, $\forall \lambda \in k$ y $v \in V$ se tiene que $\pi(\lambda v) = \lambda\pi(v)$. El espacio V/U se dirá* **espacio vectorial cociente** *de V por U*
2. *Si V es de dimensión finita, entonces* $\dim_k(V/U) = \dim_k(V) - \dim_k(U)$
3. *Si V es de dimensión finita y $W \subseteq V$ es el suplementario de U(ie, $V = U \oplus W$) entonces $\pi|_W : W \xrightarrow{\sim} V/U$ es un isomorfismo. En particular, si $\mathscr{C} = (w_1, \ldots, w_d)$ es una base de W, entonces $\pi(\mathscr{C})$ es una base de V/U.*

Demostración. — Gracias a la Proposición 5.1.1, ya sabemos que V/U es un grupo abeliano (pues $\pi(v + v') = \pi(v) + \pi(v')$. Entonces, basta verificar que $\forall\lambda \in k, v \in V$ la expresión

$$\lambda\pi(v) = \lambda(v + U) = \lambda v + U := \pi(\lambda v)$$

está bien definida: sea $v' \in V$ tal que $v \sim_U v'$ (ie, $v - v' \in U$), y $x \in U$ tal que $v = v' + x \Rightarrow \lambda v' = \lambda v' + \lambda x$. Como U es subespacio $\lambda x \in U$ y se tiene que $\lambda v \sim_U \lambda v'$. Entonces la expresión $\lambda[v] := [\lambda v] = [\lambda v'] =: \lambda[v']$ está bien definida.

Claramente, la proyección canónica $\pi : V \twoheadrightarrow V/U$ es sobreyectiva, y por otro lado se tiene que

$$\ker(\pi) = \pi^{-1}([0]) = \pi^{-1}(0 + U) = \{v \in V \mid v - 0 \in U\} = U$$

Entonces, suponiendo que V es de dimensión finita, se puede aplicar el teorema del rango para obtener

$$\dim_k(V) = \dim_k \ker(\pi) + \mathrm{rg}(\pi) = \dim_k(U) + \dim_k(V/U)$$
$$\Rightarrow \dim_k(V/U) = \dim_k(V) - \dim_k(U)$$

Sea W el subespacio suplementario de U y denotemos por $\pi_W = \pi|_W : W \to V/U$ la restricción de π a W. Notar entonces que $\ker(\pi_W) = \ker(\pi) \cap W = U \cap W = \{0\}$ (dado que $V = U \oplus W$), y entonces π_W resulta ser inyectiva. Nuevamente, debido a que $V = U \oplus W$ para todo $v \in V$ existem únicos $u \in U, w \in W$ tales que

$$v = u + w \Rightarrow \pi(v) = \underbrace{\pi(u)}_{=0} + \pi(w) = \pi(w) = \pi_W(w)$$

y así π_W es sobreyectiva, y entonces π_W es un isomorfismo. Por lo anterior, claramente la imagen de una base de W resultará ser base de V/U. $\square$

Ejercicio 5.1.5. — Sea $V = \mathbb{R}^2$ y $U = \mathrm{Vect}_{\mathbb{R}}(e_1)$ (ie, el eje x). Describir los elementos del cociente V/U.

Con respecto al teorema anterior, es importante notar que el suplementario $W \subseteq V$ no es único, por lo que el isomorfismo $W \cong V/U$ descrito resulta ser "no canónico".
El teorema siguiente, conocido como la **propiedad universal del cociente**, será una tónica recurrente en el estudio del álgebra, ya que muchos espacios que se construirán basarán gran parte de su importancia en alguna propiedad universal.

***Teorema 5.1.6* (propiedad universal).** — *Sea V k-espacio vectorial, $U \subseteq V$ y $f : V \to W$ una aplicación lineal tal que $U \subseteq \ker(f)$. Entonces $\exists! \hat{f} : V/U \to W$ tal que $f = \hat{f} \circ \pi$, donde $\pi : V \to V/U$ es la proyección canonica. Es decir, el siguiente diagrama*

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[d, "\pi"'] & W \\ V/U \arrow[ur, "\exists! \hat{f}"'] & \end{tikzcd}$$

es conmutativo.

Demostración. — Notar en primer lugar que si existiera la aplicación $\hat{f}$, entonces ésta cumpliría que $f(v) = \hat{f}(\pi(v)) = \hat{f}([v])$. Entonces, únicamente es necesario verificar que lo anterior está bien definido. Sea entonces $v, v' \in V$ tales que $\pi(v) = \pi(v')$ (ie, $v - v' \in U$), y sea además $x \in U$ tal que $v' = v + x \Rightarrow f(v') = f(v) + f(x) = f(v)$, dado que $x \in U \subseteq \ker(f)$. Así, $\hat{f} : V/U \to W$ tal que $[v] \mapsto f(v)$ está bien definida. Más aún, para todos $v, v' \in V, \lambda \in k$ se cumple que

$$\hat{f}(\lambda[v] + [v']) = \hat{f}([\lambda v + v']) = f(\lambda v + v') = \lambda f(v) + f(v') = \lambda \hat{f}([v]) + \hat{f}([v'])$$

y entonces se observa que $\hat{f}$ corresponde a una aplicación lineal. $\square$

De la propiedad universal se deduce un teorema conocido como **Teorema del Isomorfismo** (más específicamente el Primer Teorema del Isomorfismo) de Emmy Noether, el cual se demuestra a continuación.

Teorema 5.1.7 **(Noether (1927)).** — *Sean V, W k-espacios vectoriales y $f : V \to W$ una aplicación lineal. Entonces la aplicación f induce un isomorfismo:*

$$\hat{f} : V/\ker(f) \xrightarrow{\sim} \mathrm{Im}(f)$$

Demostración. — Es posible reemplazar la aplicación $f : V \to W$ por $f : V \to \mathrm{Im}(f)$, la cual resulta ser trivialmente sobreyectiva. Luego, se puede aplicar la propiedad universal del cociente sobre $U = \ker(f)$, lo cual implica la existencia de una única $\hat{f} : V/\ker(f) \to \mathrm{Im}(f)$ lineal tal que $\hat{f}([v]) = f(v)$ para todo $v \in V$. Claramente, como f es sobreyectiva $\Rightarrow$ $\hat{f}$ sobreyectiva[1]. Veamos que $\hat{f}$ es inyectiva. Sea entonces $x \in \ker(\hat{f})$ y sea $v \in V$ tal que $\pi(v) = [v] = x \Rightarrow 0 = \hat{f}([v]) = f(v) \Rightarrow v \in \ker(f)$, y así, $x = [0]$ (es 0 en el cociente), y por ende $\hat{f}$ es un isomorfismo. □

A continuación, se establecerá una importante relación existente entre los espacios cocientes, y la matrices triangulares por bloques.

Proposición 5.1.8. — *Sea V k-espacio vectorial tal que $\dim_k(V) = n$, $u : V \to V$ endomorfismo y $U \subseteq V$ subespacio estable por u ($u(U) \subseteq U$). Si denotamos por $u_U : U \to U$ la restricción $u|_U$ y $\pi : V \to V/U$ la proyección canónica, entonces se verifica que:*

1. *u induce un endomorfismo $\hat{u} : V/U \to V/U$ tal que $\hat{u}(\pi(v)) = \pi(u(v)) \quad \forall v \in V$.*
2. *Sea $\mathscr{B} = (e_1, \ldots, e_n)$ base de V tal que $\mathscr{C} = (e_1, \ldots, e_r)$ es base de U. Entonces la matriz*
$$\mathrm{Mat}_{\mathscr{B}}(u) = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix}$$
donde $A = \mathrm{Mat}_{\mathscr{C}}(u_U), B \in M_{r,n-r}(k)$ y $D \in M_{n-r}(k)$ corresponde a la matriz de $\hat{u}$ respecto a la base $\mathscr{D} = (\pi(e_{r+1}), \ldots, \pi(e_n))$.
3. $P_u(X) = P_{u_U}(X) P_{\hat{u}}(X)$

Demostración. — Dado que $u(U) \subseteq U$, se tiene que $(\pi \circ u)(U) = \{0\}$, y entonces por la propiedad universal del cociente se tiene la existencia de $\hat{u} : V/U \to V/U$ tal que $\pi \circ u = \hat{u} \circ \pi$ (ie, se considera $f = \pi \circ u$). Esto se

[1] Si $(g \circ f)$ es sobreyectiva entonces g es sobreyectiva también (f no nocesariamente lo es, en este caso si, ya que π es sobreyectiva).

puede observar en el siguiente diagrama:

$$\begin{tikzcd} V \arrow[r, "u"] \arrow[d, "\pi"'] \arrow[dr, "f"] & V \arrow[d, "\pi"'] \\ V/U \arrow[r, "\exists ! \hat{u}"'] & V/U \end{tikzcd}$$

Dado que $U = \mathrm{Vect}_k(e_1, \ldots, e_r)$ es estable por u, la matriz $M = \mathrm{Mat}_{\mathscr{B}}(u)$ tiene la forma indicada (es triangular superior). Entonces, si escribimos $B = (b_{ij}), D = (d_{lj})$ con $i = 1, \ldots, r$ y $j, l = 1, \ldots, n-r$, se tiene que

$$u(e_{r+j}) = \sum_{i=1}^{r} b_{ij}e_i + \sum_{l=1}^{r} d_{lj}e_{r+l} \Rightarrow \hat{u}(\pi(e_{r+j})) = \pi(u(e_{r+j})) = \sum_{l=1}^{r} d_{lj}\pi(e_{r+j})$$

por lo que D resulta ser la matriz de $\hat{u}$ respecto a la base $\mathscr{D} = (\pi(e_{r+1}), \ldots, \pi(e_n))$.
De lo anterior, se tiene que fácilmente que $P_u(X) = P_M(X) = P_A(X)P_D(X) = P_{u_U}(X)P_{\hat{u}}(X)$.

$\square$

Ejercicio 5.1.9. — Sea $F : V \to W$ lineal y $U \subseteq V$ subespacio. Sea además $T := u(U)$ subespacio de W. Probar que f induce una aplicación lineal

$$\hat{g} : V/U \to W/T$$

y describirla en términos de $f : V \to W, \pi_U : V \to V/U, \pi_T : W \to W/T$ (más aún, basta que $u(U) \subseteq T$ para obtener este resultado).

5.2. Espacio dual

Dado V un k.e-v, se dice que una forma lineal es una aplicación lineal de V a k.

Definición 5.2.1 **(Espacio dual).** — Sea V un $k-$espacio vectorial. Se define el **espacio dual** (o simplemente **dual**) al conjunto

$$V^* := \mathrm{Hom}_k(V, k) = \{f : V \to k \mid f \text{ es lineal } \}$$

Por otro lado, si $\dim_k(V) = n$, dada una base $\mathscr{B} = \{e_1, ..., e_n\}$ de V, se define su **base dual** de la base $\mathscr{B}$ donde, para cada e_j se le asocia la forma $e_j^* = e^j \in V^*$ tal que

$$e_j^*(e_k) = \begin{cases} 0 & si \quad k \neq j \\ 1 & si \quad k = j. \end{cases}$$

Lema La familia $\mathscr{B}^* = \{e_1^*, ..., e_n^*\}$ es efectivamente una base. Además $\dim_k(V) = \dim_k(V^*)$.

Demostración. — Note que la familia, efectivamente, es l.i. pues, si existen $\lambda_1, ..., \lambda_n$ tales que

$$\sum_{k=1}^{n} \lambda_k e_k^* = 0$$

entonces, para cada $j \in \{1, ..., n\}$ se tendrá

$$0 = \sum_{k=1}^{n} \lambda_k e_k^*(e_j) = \lambda_j$$

luego cada λ_k es nulo. Por otro lado, considere $f \in V^*$, luego, note que

$$f = \sum_{j=1}^{n} f(e_j) e_j^*$$

pues ambas expresiones son iguales para cada e_k. Luego f se puede escribir como combinación lineal de los elementos de $\mathscr{B}^*$, y por tanto$\mathscr{B}^*$ es base. En particular $\dim_k(V) = \dim_k(V^*)$. $\square$

***Definición 5.2.2* (Bidual).** — Se define el **bidual** de V como $V^{**} = (V^*)^* = \mathrm{Hom}_k(V^*, k)$ y se define la **evaluación canónica** $\psi : V \to V^{**}$ donde $x \mapsto \psi(x)$ que cumple con, para cada $f \in V^*$, $(\psi(x))(f) = \psi(x)(f) = f(x)$.

***Teorema 5.2.3* (Reflexividad).** — *La evaluación canónica $\psi : V \to V^{**}$ es lineal y biyectiva.*

Demostración. — Veamos que ψ es lineal. Sean $v, w \in V$, $\lambda \in k$ y $f \in V^*$, considere el siguiente cálculo:

$$\psi(v + \lambda w)(f) = f(v + \lambda w) = f(v) + \lambda f(w) = \psi(v)(f) + \lambda\psi(w)(f) = (\psi(v) + \lambda\psi(w))(f).$$

sigue que ψ es lineal. Por otro lado, como $\dim_k(V) = \dim_k(V^*) = \dim_k(V^{**})$, basta probar que ψ es inyectiva. Así, para $x \in V \setminus \{0\}$ existen $e_2, ..., e_n \in V$ tales que $\mathscr{B} := \{x, e_2, ..., e_n\}$ es base de V (teorema de la base incompleta). En particular, para $x^* \in V^*$ se tendrá $\psi(x)(x*) = x^*(x) = 1$, luego $\psi(x) \neq 0$. $\square$

***Definición 5.2.4* (Anulador).** — Sea V un k-e.v. y $U \subseteq V$ un conjunto no vacío. Se define el **anulador** u **ortogonal** de U como

$$U^\circ := \{f \in V^* \mid f/U = 0\} = \{f \in V* \mid f(u) = 0, \forall u \in U\}.$$

Observación 5.2.5. — 1. Notar que $U^\circ \subseteq V^*$ es un sub espacio vectorial, pues, para $f, g \in U\circ$ y $\lambda \in k$ se tendrá, para todo $x \in U$, que

$$(f + \lambda g)(x) = f(x) + \lambda g(x) = 0$$

es decir, $(f + \lambda g) \in U^\circ$.

2. **Importante:** Si $W \subseteq V^*$, es no vacío, entonces $W^\circ \subseteq V^{**}$. En particular, si $x \in V$ entonces $\psi(x) \in V^{**}$, donde, $\psi(x) \in W^\circ$ si y solo si para todo $f \in W$ se tiene $0 = \psi(x)(f) = f(x)$. En particular, si **identificamos** V y V^{**} mediante la evaluación ψ se tendrá

$$W^\circ \text{``}=\text{''} \{x \in V \mid \forall f \in W, f(x) = 0\} \text{``}\subseteq\text{''} V.$$

Teorema 5.2.6. — *Sea V un k-e.v. de $\dim_k(V) = n$ y $U \subseteq V$ un sub espacio vectorial. Entonces:*

$$\dim_k(U) + \dim_k(U^\circ) = \dim_k(V).$$

Demostración. — Si $U = \{0\}$, entonces $U^\circ = V^*$ teniéndose la igualdad. Si $U \neq \{0\}$, considere $\{e_1, ..., e_r\}$ una base de U (con $r \geq 1$), completada en una base $\mathscr{B} = (e_1, ..., e_r, e_{r+1}, .., e_n)$ de V. Sea $\mathscr{B}^* = (e_1^*, ..., e_n^*)$ la base dual de V^* relativa a $\mathscr{B}$. Note que si $f \in V^*$ pertenece a U° si y solo si $f(e_1) = ... = f(e_r) = 0$.
Luego, si escribimos $f = \sum_{j=1}^n \lambda_j e_j^*$ con $\lambda_j = f(e_j)$, lo mencionado anteriormente nos dice que $\lambda_1, ..., \lambda_r = 0$, es decir, $U^\circ = \mathrm{Vect}_k(e_{r+1}^*, ..., e_n^*\}$, es decir, $\dim_k(U^\circ) = n - r$. $\square$

***Definición 5.2.7* (aplicación traspuesta).** — Sea $u : V \to V$ lineal. Se define la **aplicación traspuesta** de u como la aplicación ${}^t u : W^* \to V^*$ dada por

$${}^t u(g) = g \circ u.$$

Observación 5.2.8. —

1. Notar que ${}^t u : W^* \to V^*$ es lineal (ejercicio).
2. Si se considera ${}^t({}^t u) : V^{**} \to W^{**}$ y si identificamos $V \cong V^{**}$, $W \cong W^{**}$ mediante la evaluación canónica, entonces ${}^t({}^t u)\text{``}=\text{''}u$, pues:

$${}^t({}^t u)(\psi_V(x))(g) \underbrace{=}_{\substack{\text{def de}\\ {}^t({}^t u)}} \psi_V(x)({}^t u(g)) \underbrace{=}_{\substack{\text{def de}\\ {}^t u}} \psi_V(x)(g \circ u) \underbrace{=}_{\substack{\text{def de}\\ \psi_V(x)}} (g \circ u)(x) = g(u(x)) = \psi_W(u(x))(g)$$

es decir, ${}^t({}^t u)(\psi_V(x)) = \psi_W(u(x))$.

Teorema 5.2.9. — *Sea $u : V \to W$ lineal entre k-e.v. de dimensión finita. Entonces:*

$$\ker({}^t u) = (\mathrm{Im}(u))^\circ.$$

en particular, u y ${}^t u$ tienen el mismo rango.

Demostración. — Dada $g \in W^*$, esta pertenece a $\ker({}^t u)$ ssi ${}^t u(g) = 0$, es decir, para todo $x \in V$, $g(u(x)) = 0$, lo que es equivalente a que $g/\,\mathrm{Im}(u) = 0$, ie, $g \in (\mathrm{Im}(u))^\circ$. Sigue que $\ker({}^t u) = (\mathrm{Im}(u))^\circ$.
Por otro lado, por teorema del rango se tiene $\dim_k(W^*) = \dim_k(\ker({}^t u)) + \dim_k(\mathrm{Im}({}^t u))$ y por teorema anterior se tiene $\dim_k(W) = \dim_k(\mathrm{Im}(u)) + \dim_k(\mathrm{Im}(u)^\circ)$. Luego, como W, V son de dimensión finita, se tiene $\dim_k(W) = \dim_k(W^*)$, además, por lo visto anteriormente también tenemos $\dim_k(\ker({}^t u)) = \dim_k(\mathrm{Im}(u)^\circ)$, obteniéndose la igualdad $\mathrm{rg}(u) = \mathrm{rg}({}^t u)$. □

Corolario 5.2.10. — *Sea $u : V \to W$ lineal entre $k-$e.v. de dimensión finita. Entonces:*

1. *${}^t u$ es inyectiva ssi u es sobreyectiva.*
2. *${}^t u$ es sobreyectiva ssi u es inyectiva.*

Demostración. — (a) Es consecuencia del teorema anterior y (b) es consecuencia de (a) aplicado en ${}^t u$. □

Proposición 5.2.11. — *Sean V, W k- espacios vectoriales, $\mathscr{B} = (e_1, ..., e_n)$ base de V, $\mathscr{C} = \{f_1, ..., f_n\}$ base de W. Sea, además, $u : V \to W$ una aplicación lineal. Entonces, la matriz de ${}^t u : W^* \to V^*$ respecto a las bases duales $\mathscr{C}^*$ y $\mathscr{B}^*$ es:*

$$\mathrm{Mat}_{\mathscr{B}^*,\mathscr{C}^*}({}^t u) = {}^t\,\mathrm{Mat}_{\mathscr{C},\mathscr{B}}(u) \qquad \textit{(matriz traspuesta)}$$

Demostración. — Sea $A = \mathrm{Mat}_{\mathscr{C},\mathscr{B}}(u) \in \mathscr{M}_{m\times n}(k)$ dada por $u(e_j) = \displaystyle\sum_{j=1}^{m} a_{ij} f_i$, donde induce a $A = (a_{ij})$. Sigue que

$${}^t u(f_i^*)(e_j) = f_i^*(u(e_j)) = \sum_{k=1}^{m} a_{kj} f_i^*(f_k) = a_{ij}$$

así, para cada $j = 1, ..., m$ se tendrá

$${}^t u(f_i^*) = \sum_{j=1}^{n} {}^t u(f_i^*)(e_j) e_j^* = \sum_{j=1}^{n} a_{ij} e_j^* \quad \in V^*$$

Finalmente, la matriz $B = \mathrm{Mat}_{B^*,C^*}({}^tu) \in \mathscr{M}_{n\times m}(k)$ está dada por $B = (b_{ij})$ tal que

$$ {}^tu(f_i^*) = \sum_{j=1}^{n} b_{ji}e_j^* \Longrightarrow b_{ji} = a_{ij} \Longrightarrow B = {}^tA $$

□

Corolario 5.2.12. — *Sea* $A \in \mathscr{M}_{m\times n}(k)$*, entonces* $\mathrm{rg}(A) = \mathrm{rg}({}^tA)$.

Demostración. — Considerar $u = u_A : k^n \to k^m$, $x \mapsto Ax$. Si $\mathscr{B}$ (resp. $\mathscr{C}$) es la base canónica de k^n (resp. k^m), entonces $\mathrm{Mat}_{\mathscr{C},\mathscr{B}}(u_A) = A$ y $\mathrm{Mat}_{\mathscr{B}^*,\mathscr{C}^*}({}^tu_A) = {}^tA$. Más aún, sabemos que

$$ rg(u_A) = rg({}^tu_A) \iff rg(A) = rg({}^tA). $$

□

5.3. Formas bilineales y formas cuadráticas

En esta sección introduciremos el concepto de forma bilineal (el cual no es más que un caso particular de las formas multilineales estudiadas en el capítulo 2), y su forma cuadrática asociada los cuales resultarán de vital importancia en todos los tópicos que restan. Supondremos de ahora en adelante que $\mathrm{car}(k) \neq 2$.

Definición 5.3.1. — Sea V un k-espacio vectorial. Una **forma bilineal** sobre V es una forma 2-multilineal, ie, una aplicación $B : V \times V \to k$ que verifica

1. La aplicación $B^y : V \to k, \quad x \mapsto B(x, y)$ es una forma lineal $\forall y \in V$.
2. La aplicación ${}^xB : V \to k, \quad y \mapsto B(x, y)$ es una forma lineal $\forall x \in V$.

Observación 5.3.2. — De forma más explícita, la aplicación $B : V \times V \to k$ es bilineal si $\forall x, y, z \in V, \lambda \in k$ se tiene que

$$ B(\lambda x, y) = B(x, \lambda y) = \lambda B(x, y), \quad B(x+y, z) = B(x, z)+B(y, z), \quad B(x, y+z) = B(x, y)+B(x, z) $$

Ejemplo 5.3.3. — La aplicación

$$ B((x_1, \ldots, x_n), (y_1, \ldots, y_n)) = x_1y_1 + \ldots + x_ny_n $$

es una forma bilineal sobre k^n.

Veremos ahora que, así como ocurría con las aplicaciones lineales, es posible asociar una representaión matricial a las formas bilineales, lo cual nos permitirá estudiarlas desde una nueva perspectiva.

Consideremos $B : V \times V \to k$ bilineal. Sea $\mathscr{B} = (e_1, \ldots, e_n)$ base de V, y $x, y \in V$ vectores. Entonces podemos escribir $x = \sum_{i=1}^n x_i e_i, y = \sum_{j=1}^n y_j e_j$ y luego

$$B(x,y) = B\left(\sum_{i=1}^n x_i e_i, \sum_{j=1}^n y_j e_j\right) = \sum_{1\le i,j\le n} B(e_i,e_j)x_i y_j$$

y de esta forma B estará completamente determinada por la matriz $(B(e_i, e_j))_{ij}$

Definición 5.3.4. — Sea $\mathscr{B} = (e_1, \ldots, e_n)$ base de V y $B : V \times V \to k$ una forma bilineal. Se define la **matriz de la forma bilineal** B **respecto a la base** $\mathscr{B}$ como

$$\mathrm{Mat}_{\mathscr{B}}(B) := (B(e_i, e_j))_{ij} \in M_n(k)$$

Si es claro de qué forma bilineal estamos hablando, escribiremos simplemente $A_{\mathscr{B}} := \mathrm{Mat}_{\mathscr{B}}(B)$.

Denotemos por

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in k^n$$

Luego

$$B(x,y) = \sum_{1\le i,j\le n} B(e_1,e_j)x_i y_j = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} A_{\mathscr{B}} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = {}^{\mathrm{t}}X A_{\mathscr{B}} Y$$

Ahora, si $\mathscr{B}'$ es otra base de V y $P = \mathrm{Mat}_{\mathscr{B}}(\mathscr{B}')$ es la matriz de cambio de base y $X', Y' \in k^n$ son los vectores de coordenadas de x e y respectivamente, respecto a la base $\mathscr{B}'$ entonces $X = PX', Y = PY'$ y luego

$$B(x,y) = {}^{\mathrm{t}}X A_{\mathscr{B}} Y = {}^{\mathrm{t}}PX' A_{\mathscr{B}}(PY') = {}^{\mathrm{t}}X' \, {}^{\mathrm{t}}P A_{\mathscr{B}} P Y' = {}^{\mathrm{t}}X' A_{\mathscr{B}'} Y'$$

y entonces hemos obtenido una fórmula para el cambio de base de formas bilineales, la cual resulta ser

$$A_{\mathscr{B}'} = {}^{\mathrm{t}}PA_{\mathscr{B}}P$$

Notar además que

$$\det(A'_{\mathscr{B}}) = \det(P)^2 \det(A_{\mathscr{B}})$$

de lo que se deduce que $\det(A_{\mathscr{B}})$ **sí** depende de la base escogida.

Definición 5.3.5. — Sea V un k-espacio vectorial de dimensión finita y $B : V \times V \to k$ una forma bilineal. Definimos las siguientes aplicaciones lineales

$$\begin{aligned} \widehat{B} :&V \to V^* \\ y &\mapsto B^y = B(\cdot, y) \end{aligned} \qquad \begin{aligned} \breve{B} :&V \to V^* \\ x &\mapsto {}^xB = B(x, \cdot) \end{aligned}$$

Lema 5.3.6. — *Sea $B : V \times V \to k$ una forma bilineal. Si identificamos V y V^{**} mediante la evaluación canónica $\psi : V \to V^{**}$ entonces la aplicación traspuesta ${}^{\mathrm{t}}\breve{B} : V^{**} \to V$ verifica que ${}^{\mathrm{t}}\breve{B} = \widehat{B}$. Además* $\mathrm{rg}(\widehat{B}) = \mathrm{rg}(\breve{B})$

Demostración. — Sean $x, y \in V$. Luego basta notar que

$${}^{\mathrm{t}}\breve{B}(\psi(y))(x) = \psi(y)(\breve{B}(x)) = \breve{B}(x)(y) = B(x, y) = \widehat{B}(y)$$

Además, dado que ψ es un isomorfismo, y $\mathrm{rg}(u) = \mathrm{rg}({}^{\mathrm{t}}u)$ (ver corolario 5.2.12) se tiene que $\mathrm{rg}(\breve{B}) = \mathrm{rg}(\widehat{B})$. $\square$

Definición 5.3.7. — Sea $B : V \times V \to k$ forma bilineal. Se define el **rango** de B, denotado $\mathrm{rg}(B)$, como $\mathrm{rg}(B) = \mathrm{rg}(\widehat{B}) = \mathrm{rg}(\breve{B})$. Diremos que B es **no degenerada** si $\mathrm{rg}(B) = \dim_k(V)$, ie, $\widehat{B}, \breve{B}$ son biyectivas.

Si B es una forma bilineal no degenerada, también se dice que $B : V \times V \to k$ es un **emparejamiento perfecto**, pues dicha forma bilineal induce un isomorfismo $V \cong V^*$ al fijar un vector $x \in V$. Sobre esto, podemos enunciar el siguiente resultado.

Proposición 5.3.8. — *Sea $B : V \times V \to k$ forma bilineal, $\mathscr{B}$ base de V y $\mathscr{B}^*$ su base dual asociada. Sea $A_{\mathscr{B}} \in M_n(k)$ la matriz de B en la base $\mathscr{B}$ y $\widehat{B} : V \to V^*$ su aplicación lineal asociada. Entonces*

$$\mathrm{Mat}_{\mathscr{B}^*,\mathscr{B}}(\widehat{B}) = A_{\mathscr{B}}$$

Demostración. — Escribamos $\mathscr{B} = (e_1, \ldots, e_n)$, $\mathscr{B}^* = (e_1^*, \ldots, e_n^*)$ y $M = \mathrm{Mat}_{\mathscr{B}^*, \mathscr{B}}(\widehat{B}) = (b_{ij}) \in M_n(k)$. Luego basta observar que

$$B(e_i, e_j) = \widehat{B}(e_j)(e_i) = \left(\sum_{l=1}^{n} b_{lj} e_l^* \right)(e_i) = \sum_{l=1}^{n} b_{lj} \underbrace{e_l^*(e_i)}_{\delta_{li}} = b_{ij}$$

□

☡ Notar que $\mathrm{rg}(B) = \mathrm{rg}(A_{\mathscr{B}})$ para toda base $\mathscr{B}$ de V. Entonces

$$\begin{aligned} B \text{ es no degenerada } &\iff A_{\mathscr{B}} \in GL_n(k) \text{ para toda base } \mathscr{B} \\ &\iff \det(A_{\mathscr{B}}) \neq 0 \text{ para toda base } \mathscr{B} \\ &\iff \forall y \in V \setminus \{0\}, \exists x \in V \text{ tal que } B(x, y) \neq 0 \end{aligned}$$

Un tipo muy particular (y especialmente importante) de formas bilineales son las llamadas "simétricas", las cuales se definen a continuación, y nos acompañarán durante un buen tiempo a lo largo de nuestro estudio.

***Definición 5.3.9* (forma bilineal simétrica).** — Sea $B : V \times V \to k$ una forma bilineal. Entonces B se dice **simétrica** si

$$B(x, y) = B(y, x) \qquad \forall x, y \in V$$

Además, una matriz $A \in M_n(k)$ se dice **simétrica** si $A = {}^{\mathrm{t}}A$.

Ejemplo 5.3.10. — La forma bilineal $B : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ definida por

$$((x_1, \ldots, x_n), (y_1, \ldots, y_n)) \mapsto x_1 y_1 + \ldots + x_n y_n$$

es simétrica.

Observación 5.3.11. — Si $B : V \times V \to k$ es simétrica entonces $\widehat{B} = \check{B} : V \to V^*$. En dicho caso definimos el **kernel** de $\mathscr{B}$ como $\ker(B) := \ker(\widehat{B}) = \ker(\check{B}) \subseteq V$, ie,

$$\ker(B) = \{x \in V : B(x, y) = 0 \quad \forall y \in V\}$$

Proposición 5.3.12. — *Sea $B : V \times V \to k$ una forma bilineal. Entonces*

$$B \text{ es simétrica } \iff A_{\mathscr{B}} = \mathrm{Mat}_{\mathscr{B}}(B) \text{ es simétrica para toda base } \mathscr{B} \text{ de } V$$

Demostración. — ($\Longrightarrow$) Suponer que B es simétrica. Luego $A_{\mathscr{B}} = (B(e_i, e_j))_{ij}$ verifica que ${}^{\mathrm{t}}A_{\mathscr{B}} = (B(e_j, e_i)) = (B(e_i, e_j)) = A_{\mathscr{B}}$ y entonces $A_{\mathscr{B}}$ es simétrica. ($\Leftarrow$) Sea $\mathscr{B} = (e_1, \ldots, e_n)$ base de V de tal forma que $A_{\mathscr{B}} = (B(e_i, e_j))$ es una matriz simétrica. Entonces $B(e_i, e_j) = B(e_j, e_i)$ para todos $i, j \in \{1, \ldots, n\}$. Luego para todo $x, y \in V$

$$B(x, y) = \sum_{1 \le i,j \le n} B(e_i, e_j) x_i y_j = \sum_{1 \le i,j \le n} B(e_j, e_i) y_j x_i = B(y, x)$$

y B es simétrica. □

***Definición 5.3.13* (forma cuadrática).** — Sea V un k-espacio vectorial. Una aplicación $Q : V \to k$ se dice **forma cuadrática** si existe $B : V \times V \to k$ bilineal de tal forma que $Q(x) = B(x, x)$ para todo $x \in V$.

Observación 5.3.14. — Considerar $\mathscr{B} = (e_1, \ldots, e_n)$ base de V y $x = \sum_{j=1}^{n} x_j e_j \in V$. Entonces

$$Q(x) = \sum_{1 \le i,j \le n} B(e_i, e_j) x_i x_j$$

y podemos notar que las formas cuadráticas se comportan de la misma forma que las funciones polinomiales homogéneas de grado 2, puesto que verifican que $Q(\lambda x) = \lambda^2 Q(x)$ para todo $x \in V, \lambda \in k$. Por tanto vemos claramente que las formas cuadráticas no son lineales.

***Teorema 5.3.15* (fórmula de polarización).** — *Sea $Q : V \to k$ una forma cuadrática. Entonces existe una única forma bilineal simétrica $B : V \times V \to k$ de tal forma que $Q(x) = B(x, x)$ para todo $x \in V$. Aún más, B está dada explícitamente por la* **fórmula de polarización**

$$B(x, y) = \frac{1}{2}(Q(x + y) - Q(x) - Q(y))$$

Decimos que B es la forma bilineal simétrica asociada a Q.

Demostración. — Sea $B : V \times V \to k$ bilineal tal que $Q(x) = B(x, x)$. Entonces

$$\begin{aligned} Q(x + y) = B(x + y, x + y) &= B(x, x) + B(x, y) + B(y, x) + B(y, y) \\ &= Q(x) + B(x, y) + B(y, x) + Q(y) \end{aligned}$$

y luego si B es simétrica $B(x, y) = B(y, x)$ y se obtiene directamente el resultado. □

Ejemplo 5.3.16. — A continuación veremos algunos ejemplos de formas cuadráticas:

1. En general, si $Q : V \to k$ es una forma cuadrática existen infinitas formas bilineales B tal que $Q(x) = B(x, x)$. Como ejemplo, sea $V = \mathbb{R}^2$ y $Q(x_1, x_2) = 2x_1x_2$. Su forma bilineal simétrica asociada es
$$B((x_1, x_2), (y_1, y_2)) = \frac{1}{2}(2(x_1 + y_1)(x_2 + y_2) - 2x_1x_2 - 2y_1y_2) = x_1y_2 + x_2y_1$$
No obstante, para todo $\lambda \in \mathbb{R}$ la forma bilineal $B_\lambda((x_{1,2}), (y_1, y_2)) = \lambda x_1y_2+(2-\lambda)x_2y_1$ verifica $B_\lambda(x, x) = Q(x)$. Matricialmente, si $\mathscr{B} = (e_1, e_2)$ es la base canónica de $\mathbb{R}^2$ entonces
$$\mathrm{Mat}_{\mathscr{B}}(B_\lambda) = \begin{pmatrix} B_\lambda(e_1, e_1) & B_\lambda(e_1, e_2) \\ B_\lambda(e_2, e_1) & B_\lambda(e_2, e_2) \end{pmatrix} = \begin{pmatrix} 0 & \lambda \\ 2-\lambda & 0 \end{pmatrix}$$
que resulta ser simétrica si y sólo si $\lambda = 1$.
2. Consideremos $Q : k^n \to k$ forma cuadrática definida por
$$Q(x_1, \dots, x_n) = \sum_{1 \le i,j \le n} a_{ij}x_ix_j$$
Entonces la matriz $(b_{ij}$ de la forma bilineal <u>simétrica</u> asociada a Q en la base canónica viene dada por
$$b_{ii} = a_{ii}, \quad b_{ij} = \frac{1}{2}a_{ij} \text{ si } i < j, \quad bij = \frac{1}{2}a_{ji} \text{ si } i > j$$
3. Sea $f \in V^* \setminus \{0\}$ una forma lineal no nula. Luego $Q(x) := f(x)^2$ es una forma cuadrática sobre V cuya forma bilineal asociada corresponde a $B(x, y) = f(x)f(y)$. Además, $\widehat{B}(y) = f(y)f \in V^*$ es una forma lineal y entonces $\Im(\widehat{B}) = \mathrm{Vect}_k(f) \subseteq V^*$ corresponde a la “recta” generada por f, y por tanto $\mathrm{rg}(Q) := \mathrm{rg}(B) = 1$.

***Definición 5.3.17* (rango de una forma cuadrática)**

Sea $Q : V \to k$ forma cuadrática y $B : V \times V \to k$ su forma bilineal simétrica asociada. Se define el **rango de la forma cuadrática** Q como $\mathrm{rg}(Q) := \mathrm{rg}(B)$.

Ejercicio 5.3.18. — Demostrar que toda forma cuadrática de rango 1 es proporcional a una forma lineal $f \in V^*$ (similar a lo visto en (3) del ejemplo 5.3.16).

5.4. Ortogonalidad respecto a una forma cuadrática

***Definición 5.4.1* (vectores ortogonales).** — Sea $B : V \times V \to k$ una forma bilineal simétrica. Decimos que:

1. Los vectores $x, y \in V$ son **ortogonales** si $B(x, y) = 0$.
2. Si $U \subseteq V$ es un subconjunto no vacío, se define el conjunto **ortogonal** de U al conjunto
$$U^{\perp} := \{y \in U \mid \forall x \in U, B(x, y) = 0\}.$$
3. Un vector $x \in V$ se dice **isótropo** si $Q(x) = B(x, x) = 0$ (ie, es ortogonal a sí mismo).

Ejemplo 5.4.2. —

1. Sea $V = \mathbb{R}^2$ y $Q(x_1, x_2) = x_1^2 - x_2^2$ forma cuadrática. Su **única** forma bilineal simétrica asociada es $B((x_1, x_2)(y_1, y_2)) = x_1y_1 - x_2y_2$, entonces:
   a) Los vectores $e_1 = (1, 0)$ y $e_2 = (0, 1)$ son ortogonales.
   b) Si $L = \mathrm{Vect}_{\mathbb{R}}(e_1)$, entonces $\mathscr{L}^{\perp}\{(y_1, y_2) \in \mathbb{R}^2 \mid y_1 = 0\} = \mathrm{Vect}_{\mathbb{R}}(e_2)$.
   c) El vector $(1, 1)$ es isótropo.
2. Sea $B : V \times V \to k$ bilineal simétrica. Recuerde que
$$\ker(B) = \{x \in V \mid B(x, y) = 0 \quad \forall y \in V\}$$
en particular, $\ker(B) = V^{\perp}$.

Ejercicio 5.4.3. — Dado un $U \subseteq V$ no vacío, pruebe qu $U^{\perp}$ es un sub espacio vectorial de V.

Lema 5.4.4. — *Sea $B : V \times V \to k$ forma bilineal simétrica, $\widehat{B} : V \times V \to k$, $y \mapsto B^y = B(\cdot, y)$ y sea $U \subseteq V$ no vacío. Entonces:*
$$U^{\perp} = \widehat{B}^{-1}(U^{\circ}).^{(2)}$$

Demostración. — Considere el siguiente cálculo:
$$\begin{aligned} U^{\perp} = \{y \in V \mid \forall x \in U, B(x, y) = 0\} &= \{y \in V \mid \forall x \in U, \widehat{B}(y)(x) = 0\} \\ &= \{y \in V \mid \widehat{B}(y) = 0\} \\ &= \widehat{B}^{-1}(U^{\circ}). \end{aligned}$$
□

Teorema 5.4.5. — *Sea $B : V \times V \to k$ una forma bilineal* **simétrica** *y $U \subseteq V$ un sub espacio vectoria. Entonces*
$$\dim_k(U) + \dim_k(U^{\perp}) \geq \dim_k(V) \qquad y\ U \subseteq (U^{\perp})^{\perp}.$$

(2) Importante, note que, para un conjunto $A \subseteq V^*$, $\widehat{B}^{-1}(A)$ es el conjunto de puntos x tales que $B(\cdot, x) \in A$, en este caso, si nos apegamos a la definición, el conjunto $\widehat{B}^{-1}(U^{\circ})$ son el conjunto de puntos x tales que $B(\cdot, x)$ se anulan en U, es decir, que están en el ortogonal.

Además, si B es **no degenerada**, *se tiene la igualdad en ambos casos.*

Demostración. — Sea $(e_1, ..., e_r)$ una base de U. Note que el espacio vectorial $U^\perp$ está definido por las ecuaciones lineales $B(e_1, y) = B(e_2, y) = ... = B(e_r, y) = 0$, ie. $\dim_k(U^\perp) \geq \dim_k(V) - r$ (pues, pueden haber ecuaciones l.d.). En particular, si B es no degenerada, entonces $\dim_k(U^\perp) = \dim_k(U^\circ) = \dim_k(V) - \dim_k(U)$. Por otro lado, si $x \in U$, entonces, para cualquier $y \in U^\perp$, se tendrá $0 = B(x, y) = B(y, x)$, i.e. $x \in (U^\perp)^\perp$ y por tanto $U \subseteq (U^\perp)^\perp$. Finalmente, si B es no degenerada, el cálculo de las dimensiones prueba la igualdad entre estos dos últimos conjuntos. □

Observación 5.4.6. — Incluso si B es no degenerada, no necesariamente $U \oplus U^\perp$, pues, si un vector v es isótropo, este pertence a ambos conjuntos y, por tanto, $U \cap U^\perp \neq \{0\}$. Por ejemplo, para $V = \mathbb{R}^2$ y $B((x_1, x_2), (y_1, y_2)) = x_1 y_1 - x_2 y_2$, se tendrá que la recta $U = \{(x_1, y_2) \in \mathbb{R}^2 \mid x_1 = x_2\}$ cumple con $U = U^\perp$ (son isótropos). Luego $\dim_\mathbb{R}(U) + \dim_\mathbb{R}(U^\perp) = 2$, pero ambos conjuntos son iguales.

***Definición 5.4.7* (base ortogonal).** — Sea $B : V \times V \to k$ una forma bilineal simétrica. Dada una base $\mathscr{B} = \{e_1, .., e_n\}$ de V, diremos que es **ortogonal** si $B(e_j, e_k) = 0$ cuando $j \neq k$.

Observación 5.4.8. — Si identificaramos a $A_\mathscr{B} = \mathrm{Mat}_\mathscr{B}(B)$, entonces la base $\mathscr{B}$ es ortogonal si la matriz $A_\mathscr{B}$ es diagonal, con

$$A_\mathscr{B} = \left(\begin{array}{c:c:c} B(e_1, e_1) & 0 & 0 \\ \hdashline 0 & \ddots & 0 \\ \hdashline 0 & 0 & B(e_n, e_n) \end{array}\right)$$

Teorema 5.4.9. — *Toda forma bilineal simétrica posee una base ortogonal.*

Demostración. — Por inducción en $n = \dim_k(V)$. Si $n = 1$ entonces, toda base es ortogonal. Supondremos $n \geq 2$:
Si $B = 0$ entonces toda base es ortogonal. Supondremos $B \neq 0$. Por la fórmula de polarización

$$B(x, y) = \frac{1}{2}\left(Q(x + y) - Q(x) - Q(y)\right)$$

necesariamente existe e_1 tal que $Q(e_1) \neq 0$. Consideremos $L = \mathrm{Vect}_k(e_1)$. Note que $L \cap L^\perp = \{0\}$ (hipótesis), luego $V = L \oplus L^\perp$, en particular, $\dim_k(L^\perp) = n - 1$, luego, por hipótesis inductiva, existe una base $(e_2, ..., e_n)$ ortogonal de $L^\perp$ para la restrición $B_{L^\perp} : L^\perp \times L^\perp \to k$. Sigue que

$\mathscr{B} = (e_1, e_2, ..., e_n)$ es una base ortogonal de V, pues $B(e_1, e_k) = 0$ para todo $k \in \{2, ...n\}$ y los otros casos se tienen por definición. $\square$

Corolario 5.4.10. — *Toda forma cuadrática en un $k-$ e.v. de dimensión n se puede escribir como combinación lineal de n cuadrados de formas lineales. Es decir, dada Q forma cuadrática, entonces existen $f_1, ...f_n$ formas lineales y $\lambda_1, ..., \lambda_n$ tales que*

$$Q = \lambda_1(f_1)^2 + ... + \lambda_n(f_n)^2$$

Demostración. — Sea $Q : V \to k$ una forma cuadrática, $B : V \times V \to k$ su forma bilineal simétrica sociada y sea $\mathscr{B} = (e_1, ..., e_n)$ una base de V. Para $x \in V$, note que x se puede escribir como $x = \sum_{k=1}^{n} x_j e_j$ para ciertos escalares $x_1, ..., x_n$. Así, observe que

$$Q(x) = B(x, x) = \sum_{j=1}^{n}\sum_{i=1}^{n} B(e_i, e_j)x_i x_j \underbrace{=}_{\substack{\text{la base es}\\ \text{ortogonal}}} \sum_{j=1}^{n} B(e_j, e_j)(x_j)^2$$

Por otro lado, cada cada j se tiene $x_j = e_j^*(x)$ con e_j^* elemento de $\mathscr{B}^* = (e_1^*, ..., e_n^*)$ la base dual asociada a $\mathscr{B}$. Se desprenden entonces:

$$Q = \sum_{j=1}^{n} B(e_j, e_j)(e_j^*)^2.$$

$\square$

5.5. Clasificación de formas cuadráticas reales y complejas

En esta sección se clasificarán de manera general, las formas cuadrátricas, en el caso en que $k = \mathbb{R}, \mathbb{C}$. Para ello se comenzará definiendo algunos conceptos, como el de endomorfismo ortogonal.

Definición 5.5.1. — Sean $Q : V \to k, Q' : V \to k$ formas cuadráticas. Q y Q' se dicen equivalentes si existe $u \in \mathrm{GL}(V)$ tal que $Q(x) = Q'(u(x))$ para todo $x \in V$

La definición anterior también se puede formular en términos de otros objetos matemáticos. Las definiciones siguientes entonces, son heredadas de la anterior:

- Dos formas bilineales B, B' se dicen equivalentes si existe $u \in \mathrm{GL}(V)$ tal que $B(x, y) = B'(u(x), u(y))$ para todos $x, y \in V$.

- Dos matrices asociadas a B', B formas bilineales se dicen equivalentes si existen bases $\mathscr{B}, \mathscr{B}'$ de V tales que $\mathrm{Mat}_{\mathscr{B}}(B) = \mathrm{Mat}_{\mathscr{B}'}(B')$. Por la fórmula de cambio de base para formas bilineales, esto significa que existe $P \in GL_n(k)$ tal que $\mathrm{Mat}_{\mathscr{B}}(B') = {}^tP\,\mathrm{Mat}_{\mathscr{B}}(B)P$ (notar que en este caso $P = \mathrm{Mat}_{\mathscr{B}'}(\mathscr{B})$, y se están cambiando las matrices de la igualdad).

Observación 5.5.2. — Es importante notar que dos formas cuadrátricas equivalentes tienen el mismo rango. Esto es debido a que si Q es equivalente con Q' existe $u \in \mathrm{GL}(V)$ tal que $Q = Q \circ u$, y como u es invertible, entonces preserva el rango.

Teorema 5.5.3. — *Toda forma cuadrática de rango r en un $\mathbb{C}$ espacio vectorial de dimensión finita se puede escribir como*

$$Q = f_1^2 + \ldots + f_r^2$$

con $f_1, \ldots, f_r \in V^*$ linealmente independientes.

Demostración. — Sea $Q : V \to \mathbb{C}$ forma cuadrática. y sea $\mathscr{B} = (e_1, \ldots, e_n)$ base ortogonal (respecto a Q) de V. Entonces dado un $x \in V$ arbitrario, entonces

$$Q(x) = Q\left(\sum_{j=1}^{n} x_j e_j\right) = \sum_{j=1}^{n} \lambda_j x_j^2$$

donde $\lambda = B(e_j, e_j) \in \mathbb{C}$. Por lo anterior, la matriz de Q viene dada por

$$\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

Notar a continuación que el rango de Q viene determinado por el número de $\lambda_j \neq 0$ (cantidad de columnas linealmente independientes de la matriz asociada).

Es posible reordenar la base $\mathscr{B}$ de tal forma que $\lambda_j \neq 0$ para todo $j \leq r$ y λ_j para todo $j > r$. Dado que V es un $\mathbb{C}$ espacio vectorial, es posible definir $\mu_j^2 = \lambda_j$ para todo $j \leq r$.

Por la demostración del corolario 5.4.10, se tiene que Q se puede escribir de la siguiente forma

$$Q = \sum_{j=1}^{n} Q(e_j)(e_j^*)^2$$

Entonces, definiendo $(f_1, \ldots, f_n)$ como la base dual de la base $(e_1/\mu_1, \ldots, e_r/\mu_r, e_{r+1}, \ldots, e_n)$ se obtiene que $Q = f_1^2 + \ldots + f_r^2$. □

Observación 5.5.4. — La descomposición en formas lineales dada por el teorema anterior no es única. Por ejemplo, la forma $Q : \mathbb{C}^2 \to \mathbb{C}$ dada por $Q(x, y) = 2x^2 + 2y^2$, se puede escribir de las siguientes maneras

$$Q(x, y) = (x\sqrt{2})^2 + (y\sqrt{2})^2 = (x + y)^2 + (x - y)^2$$

Además, la conclusión del teorema anterior no es cierta en $\mathbb{R}$, ya que de lo contrarior $Q(x) \geq 0$ para todo $x \in V \equiv \mathbb{R}^n$.

Corolario 5.5.5. — *Sea $Q : V \to \mathbb{C}$ y $Q' : V \to \mathbb{C}$ formas cuadráticas en un $\mathbb{C}$ espacio vectorial. Entonces $Q \sim Q' \iff \mathrm{rg}(Q) = \mathrm{rg}(Q')$.*

Demostración. — Anteriormente se observó que $Q \sim Q' \Rightarrow \mathrm{rg}(Q) = \mathrm{rg}(Q')$. Además, por la demostración del teorema anterior existen bases de V tales que Q, Q' tienen como matriz asociada

$$\left(\begin{array}{c:c} I_r & 0 \\ \hdashline 0 & 0_{n-r} \end{array}\right)$$

, de lo cual se concluye que $Q \sim Q'$. □

Definición 5.5.6. — Sea $Q : V \to \mathbb{R}$ forma cuadrática donde V es un $\mathbb{R}$ espacio vectorial de dimensión finita. La forma cuadrátrica Q se dice:

- **Definida positiva** si $Q(x) > 0$ para todo $x \in V \setminus \{0\}$
- **Semi-definida positiva** si $Q(x) \geq 0$ para todo $x \in V$.
- **Definida negativa** si $Q(x) < 0$ para todo $x \in V \setminus \{0\}$
- **Semi-definida negativa** si $Q(x) \leq 0$ para todo $x \in V$.

Como es costumbre, las definiciones anteriores se extienden a la forma bilineal asociada a Q.
El siguiente teorema permite descomponer las formas cuadráticas reales en formas positivas y negativas.

***Teorema 5.5.7* (Sylvester,1852).** — *Toda forma cuadrática $Q : V \to \mathbb{C}$ de rango r se puede escribir como:*

$$Q = f_1^2 + \ldots, f_p^2 - f_{p+1}^2 - \ldots - f_r^2$$

donde $f_1, \ldots, f_r \in V^*$ son linealmente independientes. Además, $p \in \mathbb{N}$ depende únicamente de Q, y se define el par (p, q) con $q := r - p$ como la signatura de Q.

Demostración. — La primera parte de esta demostración es muy similar a la del teorema 5.5.3. Se define entonces $\mathscr{B}$ base ortogonal (respecto a Q) de V. Entonces para todo $x \in V$ se tiene la igualdad:

$$Q(x) = \sum_{j=1}^{n} \lambda_j x_j^2$$

con $\lambda_j = Q(e_j) \in \mathbb{R}$. Se reordena entonces la base $\mathscr{B}$ de tal forma que $\lambda_1, \ldots, \lambda_r \neq 0$ y $\lambda_j = 0$ para $j > r$. Además, la reordenación se hace de tal forma que $\lambda_1, \lambda_p > 0$ y $\lambda_{p+1}, \ldots, \lambda_r < 0$. Luego se define $\mu_j := \sqrt{|\lambda_j|}$ para todo $j \in \{1, \ldots, r\}$. Entonces tomando $(f_1, \ldots, f_n)$ la base dual de $(e_1/\mu_1, \ldots, e_r/\mu_r, e_{r+1}, e_n)$ se obtiene que $Q = f_1^2 + \ldots + f_p^2 - f_{p+1}^2 - \ldots - f_r^2$.

A continuación, se demostrará la unicidad de la signatura de Q. Considere los subespacios $\mathscr{U} = \mathrm{Vect}_{\mathbb{R}}(e_1, \ldots, e_p)$ y $W = \mathrm{Vect}_{\mathbb{R}}(e_{p+1}, \ldots, e_n)$. Luego, las restricciones $Q|_{\mathscr{U}} : \mathscr{U} \to \mathbb{R}, Q|_W : W \to \mathbb{R}$ son definida positiva y semi-definida negativa respectivamente. Si $P \subseteq V$ es subespacio tal que $Q|_P$ es definida positiva entonces $P \cap W = \{0\}$ y entonces $\dim_{\mathbb{R}}(P) + dim_{\mathbb{R}}(W) \leq n \Rightarrow \dim_{\mathbb{R}}(P) \leq p$. Entonces $p = \max\{dim_{\mathbb{R}}(P) \mid \quad Q|_P \text{ es definida positiva}\}$ está completamente definido por Q. □

Corolario 5.5.8. — *Sean $Q : V \to \mathbb{R}, Q' : V \to \mathbb{R}$ formas cuadráticas con V $\mathbb{R}$ espacio vectorial, y sean $(p, q), (p', q')$ sus signaturas. Entonces $Q \sim Q' \iff (p, q) = (p', q')$.*

Demostración. — Como $Q \sim Q'$ entonces $\mathrm{rg}(Q) = \mathrm{rg}(Q')$. □

Observación 5.5.9. — Sea $V \equiv \mathbb{R}^n$ y $Q : V \to \mathbb{R}$ forma cuadrática de signatura (p, q), entonces:

- Q es semi-definida positiva(respectivamente negativa) $\iff$ $q = 0$(repectivamente $p = 0$).
- Q es definida positiva(respectivamente negativa) $\iff$ $p = n$(respectivamente $q = n$).

Ejemplo 5.5.10. — Sea $c \in \mathbb{R}^{>0}$. La forma cuadrática de **Lorentz-Minkowski** $Q : \mathbb{R}^4 \to \mathbb{R}$ dada por $Q(x, y, z, t) = x^2 + y^2 + z^2 - c^2t^2$ tiene signatura $(3, 1)$.

Si $Q : V \to \mathbb{R}$ tiene signatura (p, q) entonces la forma $-Q : V \to \mathbb{R}$ tal que $x \mapsto -Q(x)$ tiene signatura (q, p).

5.6. Método de reducción de Gauss

El método de reducción de Gauss de formas cuadráticas permite descomponer de manera efectiva una forma cuadrática real o compleja como combinación lineal de cuadrados de formas lineales independientes, permitiéndonos calcular su rango (y su signatura en el caso real): Hay dos casos a considerar:

Caso 1: Si Q "contiene un cuadrado', ie, un término de la forma x_j^2:
Reordenando, en caso de ser necesario, podemos suponer que Q se escribe como

$$Q(x_1, ..., x_n) = \lambda_1 x_1^2 + x_1 f(x_2, ..., x_n) + R(x_2, ..., x_n)$$

donde $\lambda_1 \neq 0$, f forma lineal y R forma cuadrática. Formamos un cuadrado de binomio al escribir

$$Q(x_1, ..., x_n) = \lambda_1 \left(x_1 + \frac{f}{2\lambda_1} \right)^2 - \frac{f^2}{4\lambda_1^2} + R; \text{ y definimos } f_1 := x_1 + \frac{f}{2\lambda_1}.$$

Por inducción en el número de variables, podemos escribir la forma cuadrática $\frac{-f}{4\lambda_1} + R$ como $\lambda_2 f_2^2 + ... + \lambda_r f_r^2$ son formas lineales en las variables $(x_2, ..., x_m)$, las cuales son l.i.. De esta forma, dado que f_1 contiene un término no nulo en la variable x_1, f_1 no es combinación lineal de $f_2, ..., f_r$. Luego $Q = \lambda_1 f_1^2 + ... + \lambda_r f_r^2$ con $f_1, ..., f_r$ formas linealmente independientes.

Caso 2: Si Q "no contiene cuadrados":
Reordenando si fuera necesario, podemos suponer que Q se escribe como

$$Q(x_1, ..., x_n) = a x_1 x_2 + x_1 g_1(x_3, ..., x_n) + x_2 g_2(x_3, ..., x_n) + R(x_3, ..., x_n),$$

donde $a \neq 0$, g_1 y g_2 formas lineales y R forma cuadrática. Utilizamos la identidad $uv = \frac{1}{4}(u+v)^2 - \frac{1}{4}(u-v)^2$ para escribir:

$$\begin{aligned} Q(x_1, ..., x_n) &= a \left(x_1 + \frac{g_2}{a} \right) \left(x_2 + \frac{g_1}{a} \right) - \frac{g_1 g_2}{a} + R \\ &= \frac{a}{4} \left(x_1 + x_2 + \frac{g_2 + g_1}{a} \right)^2 - \frac{a}{4} \left(x_1 - x_2 + \frac{g_2 - g_1}{a} \right)^2 - \frac{g_1 g_2}{a} + R \end{aligned}$$

y definimos $f_1 := x_1 + x_2 + \frac{g_2+g_1}{a}$, $f_2 := x_1 - x_2 + \frac{g_2-g_1}{a}$. Al igual que antes, aplicamos la hipótesis de inducción a la forma cuadrática dada por $\frac{-1}{a} g_1 g_2 + R$ y la escribimos como $\lambda_3 f_3^2 + ... + \lambda_r f_r^2$, donde $f_3, ..., f_r$ son formas lineales en las variables $(x_3, ..., x_n)$ que son linealmente independientes. Al igual que antes:

$$Q = \lambda_1 f_1^2 + ... + \lambda_r f_r^2; \quad \text{con } f_1, ..., f_r \text{ formas l.i.}$$

Ejemplo 5.6.1. —

1. Sea $Q:\mathbb{R}^3\to\mathbb{R}$ la forma cuadrática dada por $Q(x,y,z)=xy+yz+zx$ ("sin cuadrados"). Luego:

$$Q(x,y,z)=xy+x\underbrace{z}_{g_1}+y\underbrace{z}_{g_2}=(x+z)(y+z)-z^2=\frac{1}{4}(x+y+2z)^2-\frac{1}{4}(x-y)^2-z^2.$$

   Se sigue que la signatura de Q es $(1,2)$ y rango $\operatorname{rg}(Q)=3$. En particular es no degenerada.

2. Sea $Q:\mathbb{R}^4\to\mathbb{R}$ dada por $Q(x,y,z,t)=xy+yz+zt+tx$ ("sin cuadrados"). Luego:

$$\begin{aligned}Q(x,y,z,t)=xy+x\underbrace{t}_{g_1}+y\underbrace{z}_{g_2}+\underbrace{zt}_{R}&=(x+z)(y+t)-\cancel{tz}+\cancel{zt}\\&=\frac{1}{4}(x+y+z+t)^2-\frac{1}{4}(x-y+z-t)^2.\end{aligned}$$

   Se sigue que la signatura de Q es $(1,1)$ y su rango es $\operatorname{rg}(Q)=2$. En particular es degenerada.

3. Sea $Q:\mathbb{R}^4\to\mathbb{R}$ dada por $Q(x,y,z,t)=x^2+2y^2+3z^2-2zt+tx+3xy-yt$ ("con cuadrados"). Es conveniente comenzar con el término cuadrado de la variable que aparece menos veces: z en este caso. Luego:

$$Q(x,y,z,t)=3z^2-2zt+\underbrace{(x^2+2y^2+tx+3xy-yt)}_{R_1(x,y,t)}=3\left(z-\frac{t}{3}\right)^2-\frac{t^2}{3}+R_1(x,y,r).$$

   En la expresión $\frac{-t^2}{3}+R_1=x^2+2y^2+tx+3xy-yt-\frac{t^2}{3}$ escogemos la variable x:

$$x^2+x(t+3y)+(2y^2-yt-\frac{t^2}{3})=\left(x+\frac{1}{2}(t+3y)\right)^2-\frac{1}{4}(t+3y)^2+2y^2-yt-\frac{t^2}{3}.$$

   En la expresión $\frac{1}{4}(t+3y)^2+2y^2-yt-\frac{t^2}{3}=\frac{-1}{4}y^2-\frac{7}{12}t^2-\frac{5}{2}yt$ escogemos la variable y:

$$-\frac{1}{4}y^2-\frac{7}{12}t^2-\frac{5}{2}yt=-\frac{1}{4}(y^2+10yt)-\frac{7}{12}t^2=-\frac{1}{4}(y+5t)^2+\frac{17}{3}t^2.$$

   Finalmente,

$$Q(x,y,z,t)=3\left(z-\frac{t}{3}\right)^2+\left(x+\frac{t}{2}+\frac{3}{2}y\right)^2-\frac{1}{4}(y+5t)^2+\frac{17}{3}t^2.$$

   Observamos que la signatura de Q es $(3,1)$ y $\operatorname{rg}(Q)=4$ (no-degenerada).

4. Sea $Q : \mathbb{R}^3 \to \mathbb{R}$ dada por $Q(x, y, z) = x^2 + y^2 + 3z^2 + 4xy + 2xz + 2yz$.
   Escogemos x:

$$\begin{aligned} Q(x, y, z) = x^2 + x(4y + 3z) + (y^2 + 3z^2 + 2yz) &= (x + 2y + z)^2 - (2y + z)^2 + y^2 + 3z^2 + 2yz \\ &= (x + 2y + z)^2 - 3y^2 + 2z^2 - 2yz. \end{aligned}$$

   En la expresión $-3y^2 + 2z^2 - 2yz$ escogemos y:

$$-3y^2 + 2z^2 - 2yz = -3\left(y^2 + \frac{2}{3}yz\right) + 2z^2 = -3\left(y + \frac{z}{3}\right)^2 + \frac{7}{3}z^2.$$

   Finalmente,

$$Q(x, y, z) = (x + 2y + z)^2 - 3\left(y + \frac{z}{3}\right)^2 + \frac{7}{3}z^2.$$

   Luego, la signatura de Q es $(2, 1)$ y $\mathrm{rg}(Q) = 3$ (no-degenerada).

Ejercicio 5.6.2. — Sea $Q : \mathbb{R}^3 \to \mathbb{R}$ dada por $Q(x, y, z) = x^2 + y^2 + z^2 - 4(xy + yz + zx)$.

1. Descomponer Q como combinación lineal de cuadrados de formas lineales independientes. Determinar la signatura de Q y $\mathrm{rg}(Q)$.
2. Determinar una base de $\mathbb{R}^3$ ortogonal respecto a Q.

5.7. Grupo ortogonal

En esta sección se discutirá acerca de la ortogonalidad de endomorfismos respecto a una forma cuadrática, y sobre algunos importantes conjuntos que esta idea de ortogonalidad define, como lo es el caso del grupo ortogonal. En adelante $B : V \times V \to k$ denotará la forma bilineal simétrica asociada a una forma cuadrática Q.

Lema 5.7.1. — *Sea $u \in \mathrm{GL}(V)$. Entonces las siguientes afirmaciones son equivalentes:*

1. *u preserva la forma cuadrática, es decir, $Q(u(x)) = Q(x) \quad \forall x \in V$.*
2. *u preserva la forma bilineal simétrica asociada a Q, es decir, $B(u(x), u(y)) = B(x, y)$ $\forall x, y \in V$.*

Demostración. — Notar que si u preserva B, claramente se tiene que u preserva Q, por lo que dicha implicación es trivial. Ahora, si u preserva Q, por la fórmula de polarización se tiene que

$$B(u(x), u(y)) = \frac{1}{2}(Q(u(x) + u(y)) - Q(x) - Q(y)) =$$

$$= \frac{1}{2}(Q(u(x+y)) - Q(u(x)) - Q(u(y))) = \frac{1}{2}(Q(x+y) - Q(x) - Q(y)) = B(x,y)$$

□

Definición 5.7.2. — Sea $Q : V \to k$ una forma cuadrática. Un automorfismo $u \in \mathrm{GL}(V)$ se dice ortogonal respecto a Q si preserva Q, es decir, $Q(u(x)) = Q(x) \quad \forall x \in V$.

De la definición anterior, se puede notar que un endomorfismo se dice ortogonal respecto de Q si preserva Q, y además es invertible. Así, la noción de ortogonalidad para endomorfismos está restringida únicamente al mundo de los automorfismos. Se define además el conjunto

$$O(Q) := \{u \in \mathrm{GL}(V) \mid Q(u(x)) = Q(x) \quad \forall x \in V\} \subseteq \mathrm{GL}(V)$$

Dicho conjunto resultará tener estructura de grupo, y se le conocerá como el **grupo ortogonal de Q**.

Proposición 5.7.3. — *Sea $Q : V \to k$ forma cuadrática. Entonces el conjunto $O(Q)$ es un subgrupo de* $\mathrm{GL}(V)$ *(bajo la composición de automorfismos).*

Demostración. — En primer lugar $Id_V \in O(Q)$, ya que trivialmente preserva Q. Sean $u, v \in O(Q)$. Entonces claramente $u \circ v$ está en $O(Q)$ ya que

$$Q((u \circ v)(x)) = Q(u(v(x))) = Q(v(x)) = Q(x)$$

Además, dado $u \in O(Q)$, claramente u^{-1} está en $O(Q)$ pues denotando $x' = u^{-1}(x)$ para un $x \in V$ arbitrario se obtiene

$$Q(u^{-1}(x)) = Q(x') = Q(u(x)) = Q(x)$$

Así, el conjunto $O(Q)$ posee estructura de grupo. □

Observación 5.7.4. — Si $Q : V \to k$ es una forma cuadrática no-degenerada, entonces todo endomorfismo que preserva Q es biyectivo. Esto es así, ya que si $u : V \to V$ preserva Q y $u(x) = 0$, por el lema 5.7.1 u preserva B y entonces:

$$B(x,y) = B(u(x), u(y)) = 0 \quad \forall y \in V \Rightarrow x \in \ker(B) \Rightarrow x = 0$$

De esta forma, se tiene que u es inyectivo y, por ende también biyectivo.

Proposición 5.7.5. — *Sea $Q : V \to k$ forma cuadrática, $\mathscr{B}$ base de V y $A_{\mathscr{B}} = \mathrm{Mat}_{\mathscr{B}}(Q)$. Sea además $u \in \mathrm{GL}(V)$ y $M = \mathrm{Mat}_{\mathscr{B}}(u)$. Entonces se tiene que*

1. $u \in O(Q) \iff {}^tMA_{\mathscr{B}}M = A_{\mathscr{B}}$.
2. *Si Q es no-degenerada y $u \in O(Q) \Rightarrow \det(u) = \pm 1$.*

Demostración. — Sea $x, y \in V$, y $X, Y \in k^n$ sus vectores de coordenadas en la base $\mathscr{B}$. Entonces

$$B(u(x), u(y)) = {}^t(MX)A_{\mathscr{B}}(MY) = {}^tX^tMA_{\mathscr{B}}MY$$

De esta forma, se concluye que $B(u(x), u(y)) = B(x, y)$ si y solamente si ${}^tMA_{\mathscr{B}}M = A_{\mathscr{B}}$.
Supongamos ahora que Q es no-degenerada y $u \in O(Q)$. Entonces $\det(A_{\mathscr{B}}) \neq 0$. Aplicando la propiedad anterior

$$\det({}^{\mathrm{t}}M)\det(A_{\mathscr{B}})\det(M) = \det(M)^2\det(A_{\mathscr{B}}) = \det(A_{\mathscr{B}}) \Rightarrow \det(M)^2 = 1$$

Como $car(k) \neq 2 \Rightarrow \det(M) = \pm 1$. □

Debido a la proposición anterior, todo endomorfismo ortogonal a una forma cuadrática cumple que $\det(u) = \pm 1$. Esto motiva definir el conjunto

$$SO(Q) = \{u \in O(Q) \mid \det(u) = 1\} \subseteq O(Q)$$

el cual goza de propiedades importantes, partiendo por el hecho de que $SO(Q)$ es subgrupo de $O(Q)$. Este grupo se conoce como el **grupo especial ortogonal de Q**

Proposición 5.7.6. — *Si Q es no-degenerada, entonces el conjunto $SO(Q)$ es un subgrupo de $O(Q)$.*

Demostración. — En primer lugar, $Id_V \in SO(Q)$, ya que $\det(Id_V) = 1$.
SI $u, v \in SO(Q)$, claramente $\det(u \circ v) = \det(u)\det(v) = 1$ y $u \circ v \in SO(Q)$.
Además, si $u \in SO(Q)$ entonces $\det(u^{-1}) = 1/\det(u) = 1$ y $u^{-1} \in SO(Q)$. De esta forma, $SO(Q)$ es un subgrupo de $O(Q)$. □

Proposición 5.7.7. — *Sean $Q : V \to k, Q' : V \to k$ formas cuadráticas. Entonces si $Q \sim Q' \Rightarrow O(Q) \cong O'(Q)$.*

Demostración. — Supongamos que $Q \sim Q'$. Luego existe $u \in \mathrm{GL}(V)$ tal que $Q(x) = Q'(u(x))$ para todo $x \in V$. Sea entonces $v \in O(Q), x \in V$. Entonces se observa que

$$Q'(x) = Q'((u \circ u^{-1})(x)) = Q(u^{-1}(x)) = Q((v \circ u^{-1})(x)) = Q'((u \circ v \circ u^{-1})(x))$$

Por lo anterior, el automorfismo $u \circ v \circ u^{-1} \in O(Q')$. Entonces el endomorfismo $\varphi : O(Q) \to O(Q')$ dado por $v \mapsto u \circ v \circ u^{-1}$ es un morfismo de grupos, el cual resulta ser biyectivo. De esta forma, $O(Q) \cong O(Q')$. □

Importante: Gracias al Teorema de Sylvester, el estudio del grupo $O(Q)$ para formas cuadráticas $Q : V \to k$ no degeneradas se reduce a los siguientes casos:

1. Sobre $k = \mathbb{C}$: Toda forma no degenerada $Q : V \to \mathbb{C}$ puede ser reducida, en una base conveniente, a
$$Q(x) = x_1^2 + \ldots + x_n^2$$
Al grupo de matrices que preserva dicha forma cuadrática, llamado **grupo ortogonal complejo**, se le denotará por $O(n, \mathbb{C}) = O_n(\mathbb{C})$.
2. Sobre $k = \mathbb{R}$: Toda forma no degenerada $Q : V \to \mathbb{R}$ puede ser escrita como
$$Q(x) = x_1^2 + \ldots + x_p^2 - x_{p+1}^2 - \ldots - x_{p+q}^2$$
donde $n = p + q$, y a su grupo de automorfismos ortogonales, el **grupo ortogonal real**, se le denotará como $O(p, q, \mathbb{R}) = O(p, q)$. En particular, si $(p, q) = (n, 0)$ (Q es definida positiva) entonces escribimos $O(n, 0) = O(n, \mathbb{R}) = O_n(\mathbb{R})$.

Ejercicio 5.7.8. — Demuestre que $O(p, q) \cong O(q, p)$. Notar que en particular, no se tiene que $O(Q) \cong O(Q') \Rightarrow Q \sim Q'$. Probar además que $O_n(k) = \{A \in M_n(k) : {}^{\mathrm{t}}AA = I_n\}$ con $k = \mathbb{R}, \mathbb{C}$.

5.8. Espacios euclideanos

Para efectos de esta sección, consideraremos $V \cong \mathbb{R}^n$ un $\mathbb{R}$-e.v. de dimensión finita.

***Definición 5.8.1* (producto escalar).** — Sea V un $\mathbb{R} - e.v.$ de dimensión finita. Diremos que una forma bilineal simétrica es un **producto escalar** (también llamado "producto interno" o "producto punto") en V si es definida positiva. Un **espacio euclideano** es un espacio vectorial real dotado de un producto escalar.

Notación: El producto escalar entre dos vectores $x, y \in V$ se usualmente como $\langle x, y \rangle$. Así, la forma bilineal $\langle \cdot, \cdot \rangle : V \times V \to k$ cumple con ser:
1. **Simétrica**, ie, $\langle x, y \rangle = \langle y, x \rangle$.
2. **Definida positiva**, ie, $\langle x, x \rangle \geq 0$ y es solo 0 cuando $x = 0$ (es decir, el único vector isótropo es el 0).

Observación 5.8.2. — En particular, $\langle \cdot, \cdot \rangle$ es siempre **no degenerado**. Además, dado que su signatura es $(n, 0)$ con $n = \dim_{\mathbb{R}}(V)$, existe una base ortogonal $\mathscr{B} = (e_1, ..., e_n)$ tal que, para todo $x = \sum_{k=1}^n x_j e_j \in V$, se tenga
$$\langle x, x \rangle = (x_1)^2 + ... + (x_n)^2.$$

Ejemplo 5.8.3. — Para $V = \mathbb{R}^n$, la forma

$$\langle (x_1, ..., x_n), (y_1, ..., y_n) \rangle := x_1 y_1 + ... + x_n y_n$$

define un producto escalar **canónico** de $\mathbb{R}^n$, dotándolo así de una estructura de espacio euclidiano.

Ejercicio 5.8.4. — Sea $V = \mathbb{R}_d[X]$. Pruebe que la relación

$$\langle P, Q \rangle = \int_{-1}^{1} P(t)Q(t)dt$$

es un producto interno en V.

***Lema 5.8.5* (Desigualdad de Cauchy-Schwarz)**

Sea V un espacio euclideano. Para $x, y \in V$, siempre se cumplirá

$$|\langle x, y \rangle| \leq |\langle x, x \rangle|^{\frac{1}{2}} |\langle y, y \rangle|^{\frac{1}{2}}$$

Teniéndose la igualdad solamente cuando x e y son colineales.

Demostración. — Si x o y son iguales a 0 no hay nada que probar. Supondremos que ambos son no nulos y además cumple con $\alpha x \neq \lambda y$ para todos $\alpha, \lambda \in \mathbb{R}$. En tal caso, tendremos entonces

$$\langle \alpha x - \lambda y, \alpha x - \lambda y \rangle > 0 \iff \alpha^2 \langle x, x \rangle + \lambda^2 \langle y, y \rangle - 2\alpha\lambda \langle x, y \rangle > 0$$

Tomando el caso particular $\alpha = (\langle y, y \rangle)^{\frac{1}{2}}$ y $\lambda = (\langle x, x \rangle)^{\frac{1}{2}}$ tendremos

$$\langle y, y \rangle \langle x, x \rangle - 2\langle x, x \rangle^{\frac{1}{2}} \langle y, y \rangle)^{\frac{1}{2}} \langle x, y \rangle + \langle x, x \rangle \langle y, y \rangle > 0 \iff 2\langle x, x \rangle \langle y, y \rangle > 2\langle x, x \rangle^{\frac{1}{2}} \langle y, y \rangle^{\frac{1}{2}} \langle x, y \rangle$$

Dividiendo por $2\langle x, x \rangle^{\frac{1}{2}} \langle y, y \rangle^{\frac{1}{2}}$ obtedremos

$$\langle x, y \rangle < \langle x, x \rangle^{\frac{1}{2}} \langle y, y \rangle^{\frac{1}{2}}$$

Reemplazando x por $-x$ en la anterior desigualdad obtendremos

$$-\langle x, x \rangle^{\frac{1}{2}} \langle y, y \rangle^{\frac{1}{2}} < \langle x, y \rangle$$

juntando ambos resultados obtendremos la desigualdad pedida. Si $x = \lambda y$ entonces

$$|\langle x, y \rangle| = |\langle \lambda y, y \rangle| = |\lambda| \langle y, y \rangle = \langle \lambda y, \lambda y \rangle^{\frac{1}{2}} \langle y, y \rangle^{\frac{1}{2}} = \langle x, x \rangle^{\frac{1}{2}} \langle y, y \rangle^{\frac{1}{2}}.$$

□

Teorema 5.8.6. — *Sea V un espacio euclideano. Entonces, la aplicación $\|\cdot\| : V \to k$ definida por*

$$\|x\| = \langle x, x \rangle^{\frac{1}{2}}$$

es una norma [3]*, es decir, verifica:*

1. $\|x\| \geq 0$ *y* $\|x\| = 0$ *ssi* $x = 0$.
2. $\|\lambda x\| = |\lambda|\|x\|$.
3. $\|x + y\| \leq \|x\| + \|y\|$ *(desigualdad triangular).*

Demostración. — La propiedades (1) y (2) se obtienen por la definición de producto escalar. Para (3) considere el siguiente cálculo:

$$\begin{aligned}\|x+y\|^2 = \langle x+y, x+y\rangle^{\frac{1}{2}} &= \|x\|^2 + 2\langle x, y\rangle + \|y\|^2 \\ &\underbrace{\leq}_{\text{C-S}} \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 \\ &= (\|x\| + \|y\|)^2\end{aligned}$$

Como ambos términos de la desigualdad son positivos, se obtiene $\|x+y\| \leq \|x\|+\|y\|$.

□

Observación 5.8.7. —

1. De la demostración anterior, podemos ver que, para la norma inducida por el producto escalar, si x, y son ortogonales, entonces podemos ver una generalización del Teorema de Pitágoras, pues:
$$\|x+y\|^2 = \|x\|^2 + 2\langle x, y\rangle + \|y\|^2 = \|x\|^2 + \|y\|^2$$
2. De la demostración de C-S, si $x, y \in V \setminus \{0\}$ se puede observar que:
$$-1 \leq \frac{\langle x, y\rangle}{\|x\|\|y\|} \leq 1$$
lo que motiva la siguiente definición.

***Definición 5.8.8* (ángulo entre vectores).** — Sea V un espacio euclidiano. Sean $x, y \in V \setminus \{0\}$, definimos el **ángulo no orientado** entre x e y como el único real $\theta = \theta(x, y) \in [0, \pi]$ tal que

$$\cos(\theta) = \frac{\langle x, y\rangle}{\|x\|\|y\|}.$$

Decimos que $\theta(x, y)$ es **no orientado**, pues $\theta(x, y) = \theta(y, x)$. Además, gracias a la desigualdad de Cauchy-Schwarz, tendremos las siguientes propiedades:

1. $\theta(x, y) = 0 \iff x = \lambda y$ con $\lambda > 0$.
2. $\theta(x, y) = \pi \iff x = \lambda y$ con $\lambda < 0$.

---

[3] En general, se dice que V es un Pre-Hilbert si V es un $\mathbb{R}$ espacio vectorial (no necesariamente de dimensión finita) con un producto interno. Si V es completo bajo la norma inducida por su producto interno, diremos que V es un **Espacio de Hilbert**.

3. $\theta(x,y) = \frac{\pi}{2} \iff x$ e y son ortogonales, ie, $\langle x, y\rangle = 0$.

5.9. Bases ortonormales y proceso de Gram-Schmidt

Recordemos que si B es una forma bilineal simétrica definida positiva, por definición, el único vector ortogonal a sí mismo es el vector nulo. Esta afirmación, nos motiva a enunciar la siguiente propiedad (muy importante por lo demás):

Proposición 5.9.1. — *Sea V un espacio euclidiano. Entonces, para todo $U \subseteq V$ sub-ev, se tiene que $V = U \oplus U^{\perp}$.*

Demostración. — Por la observación anterior, sabemos que $U \cap U^{\perp} = \{0\}$, pues, el único vector ortogonal a sí mismo es el 0. Por otro lado, se tiene que $\dim_{\mathbb{R}}(U) + \dim_{\mathbb{R}}(U^{\perp}) = \dim_{\mathbb{R}}(V)$, luego, necesariamente $V = U \oplus U^{\perp}$. □

Observación 5.9.2. — Sabemos que V admite una base ortogonal $\mathscr{B} = (e_1, ..., e_n)$ respecto a la forma bilineal $\langle\cdot,\cdot\rangle$. En particular, como el producto interno es definido positivo, ie, de signatura $(n, 0)$, entonces $\langle e_j, e_j\rangle > 0$En particular, si definimos $e_i :=$, se tendrá $\|e_i\| = 1$. Esta pequeña observación nos motiva a hacer la siguiente definición.

***Definición 5.9.3* (base ortonormal).** — Sea V un espacio euclideano y $\mathscr{B} = (e_1, ..., e_n)$ una base de V. Diremos que $\mathscr{B}$ es una **base ortonormal** (o bien, ortonormada), si $\langle e_i, e_j\rangle = \delta_{ij}$ para todos $i, j \in \{1, ..., n\}$. De manera más general, diremos que una familia de vectores $x_1, ..., x_m$ es ortonormal, si $\langle x_j, x_k\rangle = \delta_{jk}$ para $j, k \in \{1, ..., m\}$.

Observación 5.9.4. — 1. De la observación anterior, vimos que todo espacio euclidiano posee uan base ortonormal.
2. La matris $A_{\mathscr{B}}$ de la forma bilineal $\langle\cdot,\cdot\rangle$ respecto a una base $\mathscr{B}$ ortonormal es la identidad.
3. Toda familia ortonormal $x_1, ..., x_m$ es linealmente independiente, con $m \leq n$. Pues si $\lambda_1 x_1 + ... + \lambda_m x_m = 0$, entonces $0 = \langle \lambda_1 x_1 + ... + \lambda_n x_n, x_j\rangle = \lambda\langle x_j, x_j\rangle = \lambda$ para cualquier j.

Proposición 5.9.5. — *Sea V un espacio euclideano y sea $\mathscr{B} = (e_1, ..., e_n)$ una base ortonormal de V. Entonces, para cada $x \in V$ se tendrá*

$$x = \sum_{j=1}^{n} \langle x, e_j\rangle e_j \qquad y \qquad \|x\|^2 = \sum_{j=1}^{n} \langle x, e_j\rangle^2.$$

Demostración. — Si escribimos $x = \sum_{j=1}^{n} x_j e_j$, tendremos

$$\langle x, e_j\rangle = \langle x_1 e_1 + ... + x_n e_n, e_j\rangle = x_j \langle e_j, e_j\rangle = x_j.$$

Luego, $x = \sum_{j=1}^{n} \langle x, e_j\rangle e_j$.
Por otro lado, note que

$$\begin{aligned}\|x\|^2 = \langle \sum_{j=1}^{n} \langle x, e_j\rangle e_j, \sum_{j=1}^{n} \langle x, e_j\rangle e_j\rangle &= \langle x, e_1\rangle^2 \|e_1\|^2 + ... + \langle x, e_n\rangle^2 \|e_n\|^2 \\ &= \langle x, e_1\rangle^2 + ... + \langle x, e_n\rangle^2.\end{aligned}$$

□

Recuerdo: En la sección de grupo ortogonal se vio que, dada $Q : V \to k$ una forma cuadrática y $B : V \times V \to k$ es su forma bilineal simétrica asociada, un endomorfismo u es ortogonal respecto a Q si y solo si es ortogonal respecto a B. Además, si B es no degenerada, entonces u necesariamente pertenece a $\mathrm{GL}(V)$.

Proposición 5.9.6. — *Sea V un espacio euclideano y $u : V \to V$ un endomorfismo. Las siguientes propiedades son equivalentes:*

1. $\langle u(x), u(y)\rangle = \langle x, y\rangle$ *para todos* $x, y \in V$.
2. $\|u(x)\| = \|x\|$ *para todo* $x \in V$.
3. *Existe una base $\mathscr{B}$ ortonormal de V tal que $u(\mathscr{B})$ es una base ortonormal.*
4. *Para toda base ortonormal $\mathscr{B}$ de V, $u(\mathscr{B})$ es ortonormal.*
5. *La matriz $\mathrm{Mat}_{\mathscr{B}}(u)$ respecto a una base $\mathscr{B}$ ortonormal, es una* **matriz ortogonal**, *ie,* ${}^tAA = I_n$.

Demostración. — Hemos discutido que (1) $\iff$ (2). Por otro lado, (1) $\to$ (4) $\to$ (4). Veamos que (3) $\to$ (1): Sea $\mathscr{B} = (e_1, ..., e_n)$ una base ortonormal de V tal que $u(\mathscr{B}) = (u(e_1), ..., u(e_n))$ es una base ortonormal. Sea $x = x_1 e_1 + ... + x_n e_n \in V$, luego

$$u(x) = \sum_{j=1}^{n} x_j u(e_j) \tag{3}$$

Por otro lado, vimos que

$$u(x) = \sum_{j=1}^{n} \langle u(x), u(e_j)\rangle u(e_j) \tag{4}$$

juntando 3 y 4 tendremos $\langle u(x_j), u(e_j)\rangle = x_j = \langle x, e_j\rangle$. Luego

$$\|x\|^2 = \sum_{j=1}^{n} \langle x, e_j\rangle^2 = \sum_{j=1}^{n} \langle u(x), u(e_j)\rangle^2 = \|u(x)\|^2.$$

Finalmente, veamos que (1) $\iff$ (5). En efecto, vimos que un endomorfismo u preserva una forma bilineal simétrica B si y solo si para toda base $\mathscr{B}$ de V se tiene que ${}^tMA_{\mathscr{B}}M = A_{\mathscr{B}}$, para $M = \mathrm{Mat}_{\mathscr{B}}(u)$ y $A_{\mathscr{B}} = \mathrm{Mat}_{\mathscr{B}}(u)$. En particular, si $\mathscr{B}$ es una base ortonormal, entonces, para $B = \langle\cdot,\cdot\rangle$, tendremos $A_{\mathscr{B}} = I_n$. Se sigue inmediatamente que $I_n = {}^tMI_nM = {}^tMM$. $\square$

***Definición 5.9.7* (isometría).** — Dado un espacio euclideano V, diremos que u es una isometría si preserva $\langle\cdot,\cdot\rangle$, en particular, cumple con las 5 propiedades de la proposición anterior.

Definición 5.9.8. — Si V es un espacio euclidiano, se puede definir la **distancia** entre x e y como

$$d(x, y) = \|x - y\|.$$

Si $u : V \to V$ es un endomorfismo que preserva $\langle\cdot,\cdot\rangle$, entonces

$$d(u(x), u(y)) = \|u(x) - u(y)\| = \|u(x - y)\| = \|x - y\| = d(x, y).$$

Es decir, u preserva la distancia. Es por ello, que se prefiere llamar a u una **isometría** (en lugar de endomorfismo ortogonal) en un contexto de espacio euclidiano. Además, sabemos que si u preserva $\langle\cdot,\cdot\rangle$, necesariamente $\det(u) = \pm$. En particular, diremos que u es una:

1. **Isometría directa** si $\det(u) = 1$ (ie, $u \in SO(Q)$, con $Q = \|\cdot\|$).
2. **Isometría indirecta** si $\det(u) = -1$.

Observación 5.9.9. — Más generalmente, dado un conjunto X no vacío, y $x, y \in X$, se dice que una función $d : X \times X \to \mathbb{R}$ es una **métrica** (es decir, mide la distancia entre x e y) si cumple con:

1. $d(x, y) \geq 0$ y $d(x, y) = 0$ si y solo si $x = y$.
2. $d(x, y) = d(y, x)$, es decir, d es **simétrica**.
3. Para $z \in X$, $d(x, y) \leq d(x, z) + d(z, y)$ (desigualdad triangular).

Si d es una métrica, diremos que (X, d) es un espacio métrico. Por otro lado, se puede probar que si $(X, \|\cdot\|_X)$ es un espacio vectorial normado, entonces $d : V \times V \to \mathbb{R}$, $(x, y) \mapsto d(x, y) = \|x - y\|$ es efectivamente una métrica (ejercicio).

Corolario 5.9.10. — *Sea V un espacio euclidiano y $\mathscr{B}$ una base ortonormal. Entonces, una base $\mathscr{B}'$ de V es ortonormal si y solo si la matriz $P = \mathrm{Mat}_{\mathscr{B}}(\mathscr{B}')$ es ortogonal.*

Demostración. — Sea $\mathscr{B} = (e_1, ..., e_n)$ y $\mathscr{B}' = (e_1', ..., e_n')$. Si consideramos el endomorfismo $u : V \to V$ dado por $u(e_j) = e_j'$ para todo $j = \{1, ..., n\}$. Entonces, $\mathrm{Mat}_{\mathscr{B}}(u) = P$ (en efecto, pues, en las columnas de $\mathrm{Mat}_{\mathscr{B}}(u)$, están representados los vectores de la base $\mathscr{B}'$ en función de la base $\mathscr{B}$.), y luego $\mathscr{B}'$ es ortonormal si y solo si ${}^tPP = I_n$ (pues (4) $\iff$ (5)). □

Proposición 5.9.11. — *Sea V un espacio euclideano y sea $u : V \to V$ una isometría. Entonces, para todo $U \subseteq V$, sub espacio vectorial se tiene que:*

1. *$u(U^\perp) = u(U)^\perp$.*
2. *Si $u(U) \subseteq U$, entonces $u(U^\perp) \subseteq U^\perp$.*

Demostración. — (1.)Sea $y \in U^\perp$, entonces para cada $x \in U$ se tiene $\langle x, y\rangle = \langle u(x), u(y)\rangle = 0$, ie, $u(y)$ es ortogonal a todo elemento de $u(U)$, esto es, $u(y) \in u(U)^\perp$. Se sigue que $u(U^\perp) \subseteq u(U)^\perp$. Por otro lado, dado que $u \in \mathrm{GL}(V)$ es un automorfismo:

$$\dim_{\mathbb{R}}(u(U^\perp)) = \dim_{\mathbb{R}}(U^\perp) = \dim_{\mathbb{R}}(V) - \dim_{\mathbb{R}}(U) = \dim_{\mathbb{R}}(V) - \dim_{\mathbb{R}}(u(U)) = \dim_{\mathbb{R}}(u(U)^\perp),$$

por lo que $u(U^\perp) = u(U)^\perp$.
(2.) Dado que $u \in \mathrm{GL}(V)$, $\dim_{\mathbb{R}}(U) = \dim_{\mathbb{R}}(u(U))$. Luego, si $u(U) \subseteq U$, entonces $u(U) = U$, y por (1.) tenemos $u(U)^\perp = u(U^\perp) = U^\perp$. □

A continuación describiremos un algoritmo, llamado usualmente "**proceso de Gram - Schmidt**" que permite construir una base ortonormal a partir de cualquier base. Este método fue descubierto por Laplace en 1816, pero usualmente se le atribuye (erróneamente) a Gram (1883) y a Schmidt (1907).

Teorema 5.9.12. — *Sea V un espacio euclideano. Dado $(e_1, ..., e_n)$ base de V, existe una* **única** *base ortonormal $(e_1', ..., e_n')$ de V tal que para todo $j \in \{1, ..., n\}$ se tiene que*

1. $\mathrm{Vect}_{\mathbb{R}}(e_1, ..., e_j) = \mathrm{Vect}_{\mathbb{R}}(e_1', ..., e_j')$.
2. $\langle e_j, e_j'\rangle = 0$.

Demostración. — En la práctica, primero construiremos una base $(e_1'', ..., e_n'')$ que es sólo ortogonal y cumple (1) y (2), para luego "normalizarla":

1. Definimos $e_1'' := e_1$.

2. Suponemos (inductivamente) que $e''_1, ..., e''_{j-1}$ están ya construidos y buscamos e''_j de la forma
$$e''_j = e_j + \lambda_1 e''_1 + ... + \lambda_{j-1} e''_{j-1}.$$
Para $k \in \{1, ..., j-1\}$ se tiene que $\langle e''_j, e''_k \rangle = 0 \iff \langle e_j, e''_k \rangle + \lambda_k \|e''_k\|^2 = 0$, lo cual determina $\lambda_1, ..., \lambda_{j-1}$ y luego determina a e''_j.
3. Por inducción construimos $(e''_1, ..., e''_n)$ base ortonormal, y notamos que verifica (1) por construcción, y $\langle e_j, e''_j \rangle = \langle e''_j, e''_j \rangle = \|e''_j\|^2 > 0$, por lo que también verifica (2).
4. Definimos $e'_j := \frac{e''_j}{\|e_j\|''}$ para $j \in \{1, ..., n\}$ y luego $(e'_1, ..., e'_n)$ es una base **ortonormal** que verifica (1) y (2), cuya unicidad se obtiene por construcción.

$\square$

***Corolario 5.9.13* (Factorización QR).** — *Sea $A \in \mathrm{GL}_n(\mathbb{R})$ matriz invertible. Entonces, un único par (Q, R) de matrices, donde:*
1. *Q es ortogonal (ie, ${}^tQQ = I_n$).*
2. *R es triangular superior con coeficientes diagonales positivos.*
3. *$A = QR$.*

Demostración. — Los vectores columna $e_1, ..., e_n$ de A forman una base $\mathscr{B}$ de $\mathbb{R}^n$, a la cual le aplicamos Gram-Schmidt para obtener una base ortonormal $\mathscr{B}'$ de $\mathbb{R}^n$. Sea $\mathscr{C}$ la base canónica de $\mathbb{R}^n$, entonces (por definición): $A = \mathrm{Mat}_{\mathscr{C}}(\mathscr{B})$. Además, la demostración del proceso de Gram-Schmidt nos dice precisamente que $\mathrm{Mat}_{\mathscr{B}}(\mathscr{B}')$ es triangular superior con coeficientes diagonales estrictamente positivos. Por otro lado, sabemos que $\mathscr{B}'$ es ortonormal si y solo si $Q := \mathrm{Mat}_{\mathscr{C}}(\mathscr{B}')$ es ortogonal. Finalmente, $A = \mathrm{Mat}_{\mathscr{C}}(\mathscr{B}) = \mathrm{Mat}_{\mathscr{C}}(\mathscr{B}')\,\mathrm{Mat}_{\mathscr{B}'}(\mathscr{B}) = Q\,\mathrm{Mat}_{\mathscr{B}'}(\mathscr{B}) = Q\,\mathrm{Mat}_{\mathscr{B}}(\mathscr{B}')^{-1}$, donde $R := \mathrm{Mat}_{\mathscr{B}}(\mathscr{B}')^{-1}$ es triangular superior con coeficientes diagonales > 0. $\square$

Ejemplo: Sea $(e_1, e_2, e_3) = ((1,1,1), (1,1,0), (1,0,0))$ base de $\mathbb{R}^3$. Gram-Schmidt:
$$e''_1 := e_1 = (1,1,1), \qquad e''_2 := e_2 - \frac{\langle e_2, e''_1 \rangle}{\|e''_1\|^2} e''_1 = (1,1,0) - \frac{2}{3}(1,1,1) = \left(\frac{1}{3}, \frac{1}{3}, -\frac{2}{3}\right)$$
$$e''_3 := e_3 - \frac{\langle e_3, e''_1 \rangle}{\|e''_1\|^2} e''_1 - \frac{\langle e_3, e''_2 \rangle}{\|e''_2\|^2} e''_2 = (1,0,0) - \frac{1}{3}(1,1,1) - \frac{1}{2}\left(\frac{1}{3}, \frac{1}{3}, -\frac{2}{3}\right) = \left(\frac{1}{2}, -\frac{1}{2}, 0\right)$$
$$\implies e'_1 = \frac{e''_1}{\|e''_1\|} = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right), \quad e'_2 = \frac{e''_2}{\|e''_2\|} = \frac{1}{\sqrt{6}}(1,1,-2), \quad e'_3 = \frac{e''_3}{\|e''_3\|} = \frac{1}{\sqrt{2}}(1,-1,0).$$

La base obtenida **depende** del orden de los vectores de la base original. Por ejemplo, la base (e_3, e_2, e_1) nos da $e'_1 = (1, 0, 0)$, $e'_2 = (0, 1, 0)$ y $e'_3 = (0, 0, 1)$.

En términos de factorización QR: Si $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$ entonces la matriz

$Q = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} & 0 \end{pmatrix}$ es ortogonal y $R = Q^{-1}A = {}^tQA = \begin{pmatrix} \sqrt{3} & \frac{2}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{pmatrix}$.

Luego,

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} = QR = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} & 0 \end{pmatrix} \begin{pmatrix} \sqrt{3} & \frac{2}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{pmatrix}.$$

Observación 5.9.14. — Notar que el proceso de ortogonalización permite asociar a una familia linealmente independiente $(e_1, ..., e_j)$ (no necesariamente una base), una familia ortonormal.

5.10. Endomorfismos adjuntos y simétricos

Proposición 5.10.1. — *Sea V un espacio euclideano, entonces, para todo endomorfismo $u : V \to V$ existe un único endomorfismo $u^* : V \to V$ que cumple con*

$$\langle u(x), y \rangle = \langle x, u^*(y) \rangle \qquad \forall x, y \in V$$

Además, $\mathrm{rg}(u) = \mathrm{rg}(u^*)$.

Demostración. — Si escribimos $B = \langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ producto escalar, denotamos por $\widehat{B} : V \to V^*$ tal que $y \mapsto \langle y, \cdot \rangle = \langle \cdot, y \rangle$. Nuestra identidad buscada se puede replantear como

$$\widehat{B}(y)(u(x)) = \widehat{B}(u^*(y))(x) \text{ o bien } {}^tu \circ \widehat{B} = \widehat{B} \circ u^*.$$

Como B es no degenerada (definición), $\widehat{B}$ es un isomorfismo, luego $u^* = \widehat{B}^{-1} \circ {}^tu \circ \widehat{B}$, existe y es único. En particular, $\mathrm{rg}(u^*) = \mathrm{rg}({}^tu) = \mathrm{rg}(u)$. $\square$

La proposición anterior nos motiva a hacer la siguiente definición:

***Definición 5.10.2* (endomorfismo adjunto). —** Dado $u : V \to V$, definimos al **adjunto** de u como el único endomorfismo $u^* : V \to V$ tal que, para cualquier $x, y \in V$, se tiene:

$$\langle u(x), y\rangle = \langle x, u^*(y)\rangle.$$

***Ejemplo 5.10.3.* —** Si $u : V \xrightarrow{\sim} V$ es un automorfismo, ie $u \in GL(V)$. Entonces

$$\begin{aligned} u \in GL(n) &\iff \langle u(x), u(y)\rangle = \langle x, y\rangle \qquad \forall x, y \in V \\ &\iff \langle x, (u^* \circ u)(y)\rangle = \langle x, y\rangle \qquad \forall x, y \in V \\ &\iff u^* \circ u = u \circ u^* = Id_V \\ &\iff u^* = u^{-1}. \end{aligned}$$

***Proposición 5.10.4.* —** *Sea V un espacio euclideano y sea $\mathscr{B}$ una base ortonormal de V. Entonces, para todo endomorfismo $u : V \to V$ se tiene:*

$$\mathrm{Mat}_{\mathscr{B}}(u^*) = {}^t\,\mathrm{Mat}_{\mathscr{B}}(u).$$

Demostración. — Sea $\mathscr{B} = (e_1, ..., e_n)$ base ortonormal y $(a_{ij}) = \mathrm{Mat}_{\mathscr{B}}(u)$. Escribimos $u(e_j) = \sum_{i=1}^n a_{ij}e_i$. Como $\mathscr{B}$ es ortonormal, se tendrá, aplicando la propiedad de la sección anterior en $u(e_j)$, para a_{ij} (la iésima coordenada del vector $u(e_j)$) que $a_{ij} = \langle u(e_j), e_i\rangle = \langle e_j, u^*(e_i)\rangle$. De la misma manera, tendremos

$$u^*(e_j) = \sum_{i=1}^n \langle u^*(e_j), e_i\rangle e_i = \sum_{i=1}^n \langle e_j, u(e_i)\rangle e_i = \sum_{i=1}^n \langle u(e_i), e_j\rangle e_i = \sum_{i=1}^n a_{ji}e_i.$$

□

***Ejercicio 5.10.5.* —** Sean u, v endomorfismos de un espacio euclideano V. Pruebe que $(u \circ v)^* = v^* \circ u^*$ y que $(u^*)^* = u$.

***Definición 5.10.6* (auto-adjunto). —** Sea V un espacio euclideano y sea $u : V \to V$ un endomorfismo. Diremos que u es **simétrico** (o bien, auto-adjunto) si

$$\langle u(x), y\rangle = \langle x, u(y)\rangle \qquad \forall x, y \in V$$

***Observación 5.10.7.* —** Sea $\mathscr{B}$ una base ortonormal de V. Entonces, un endomorfismo $u : V \to V$ es simétrico si y solo si $A := \mathrm{Mat}_{\mathscr{B}}(u)$ es simétrica, ie, $A = {}^tA$.

Ejercicio: Sea $V \cong \mathbb{R}^n$ espacio euclideano. Pruebe que el conjunto

$S \subseteq \mathrm{End}(V)$ dado por $S := \{u : V \to V \mid u$ es endomorfismo simétrico $\}$ es un sub-ev y que $\dim_{\mathbb{R}}(S) = \frac{n(n+1)}{2}$.

Recuerdo 5.10.8. — Dado V un espacio euclideano, sabemos que para todo $U \subseteq V$ sub-e.v., se cumple que $V = U \oplus U^{\perp}$, en particular, todo vector $v \in V$ se escribe de manera única como $v = a + b$ con $a \in U$ y $b \in U^{\perp}$. Este pequeño recuerdo, nos motiva a hacer la siguiente definición.

***Definición 5.10.9* (proyección ortogonal).** — Sea $U \subseteq V$ un sub-e.v. de un espacio euclideano. Se define la **proyección ortogonal** sobre U como:

$$P_U : V \to V, v \mapsto a.$$

donde $v = a + b$ con $a \in U$ y $b \in U^{\perp}$. En particular, se cumple que $P_U + P_{U^{\perp}} = Id_V$.

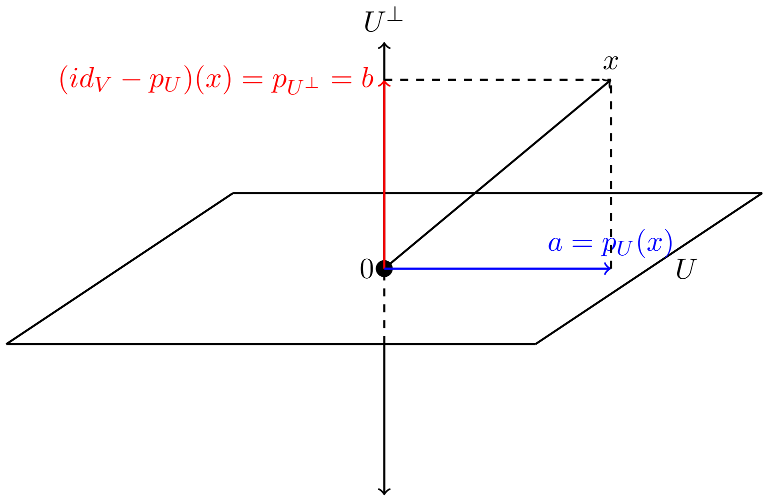

Imagen 1. En esta imagen se puede apreciar la proyección ortogonal del vector x sobre el subespacio U y sobre su ortogonal.

***Definición 5.10.10* (distancia a un conjunto).** — Sea $U \subseteq V$ un sub ev de un espacio vectorial normado[(4)] (en particular, V puede ser un espacio euclideano). Dado $x \in V$, se define la distancia entre x y U como

$$d(x,U) = \inf_{y\in U} d(x,y) = \inf_{y\in U} \|x-y\|.$$

Teorema 5.10.11. — *Sea V un espacio euclideano y $U \subseteq V$ un sub ev. Entonces, la proyección ortogonal sobre U es lineal y cumple con $P_U^2 = P_U$ y con que ss su kernel es* $\ker(P_U) = U^\perp$ *y su imagen es* $\operatorname{Im}(P_U) = U$. *Más aún, P_U es un endomorfismo auto-adjunto (ie, simétrico: $P_U^* = P_U$). Además, para todo $x \in V$ y todo $y \in U$ se tiene:*

$$d(x,U) = \|x - P_u(x)\| \leq \|x-y\|.$$

Demostración. — Sean $v, w \in V$ y $\lambda \in \mathbb{R}$. Escribimos $v = a + b$ y $w = a' + b'$ con $a, a' \in U$ y $b, b' \in U^\perp$, luego, $v + \lambda w = \underbrace{(a+\lambda a')}_{\in U} + \underbrace{(b+\lambda b')}_{\in U^\perp}$, por lo que $P_U(v+\lambda w) = a + \lambda a' = P_U(v) + \lambda P_U(w)$. Por otro lado, note que $V = U \oplus U^\perp$, se sigue entonces que, dado $b \in U^\perp$, tendremos $b = \underbrace{b}_{\in U^\perp} + \underbrace{0}_{\in U}$, ie, $P_U(b) = 0$, y dado $a \in U$, se tiene $P_U(a) = a$, donde observamos que $\ker(P_U) = U^\perp$ e $\operatorname{Im}(P_U) = U$, en particular, $P_U(P_U(v)) = P_U(a) = a$, ie, $P_U^2 = P_U$. Por otro lado, observe que

$$\langle P_U(v), w\rangle = \langle a, w\rangle = \langle a, a'+b'\rangle = \langle a, a'\rangle$$

y que

$$\langle v, P_U(w)\rangle = \langle v, a'\rangle = \langle a+b, a'\rangle = \langle a, a'\rangle$$

es decir, $P_U^* = P_U$ (P_U es auto-adjunto). Finalmente, si $y \in U$ entonces

$$\|v-y\|^2 = \|(a-y)+b\|^2 \underbrace{=}_{\text{Pitágoras}} \|a-y\|^2 + \|b\|^2 \geq \|b\|^2 = \|a+b-a\|^2 = \|v-a\|^2 = \|v-P_U(v)\|^2.$$

□

¡Importante! Dado V un espacio euclideano y $U \subseteq V$ sub e.v. tal que $\dim_{\mathbb{R}}(U) = r$ y $(e_1, ..., e_r)$ es una base ortogonal de U, entonces para todo vector $x \in V$ se tiene

$$P_U(x) = \langle x, e_1\rangle \frac{e_1}{\|e_1\|^2} + ... + \langle x, e_r\rangle \frac{e_r}{\|e_r\|^2}.$$

---

(4) Esta definición se puede extender a U como un conjunto cualquiera, y V un espacio métrico.

En efecto, sea $a := \langle x, e_1\rangle \frac{e_1}{\|e_1\|^2} + ... + \langle x, e_r\rangle \frac{e_r}{\|e_r\|^2} \in U$ y sea $b := x - a$. Luego, para todo $j \in \{1, ..., r\}$ se tiene que:

$$\langle b, e_j\rangle = \langle x - a, e_j\rangle = \langle x, e_j\rangle - \langle a, e_j\rangle = \langle x, e_j\rangle - \frac{\langle x, e_j\rangle}{\|e_j\|^2}\langle e_j, e_j\rangle = 0.$$

Así, $b \in U^\perp$ y luego necesariamente $P_U(x) = a$. De hecho, si $(e_1, ..., e_r)$ es una base **ortonormal** de U, entonces

$$P_U(x) = \langle x, e_1\rangle e_1 + ... + \langle x, e_r\rangle e_r.$$

Note que esta fórmula ya la mencionamos anteriormente en la fórmula de ortogonalización de Gram-Schmidt. En efecto, usando la misma notación que en aquella sección, tenemos que

$$e''_j = e_j - \langle e_j, e''_1\rangle \frac{e''_1}{\|e''_1\|^2} - .... - \langle e_j\rangle = e_j - P_{U_{j-1}},$$

donde $U_{j-1} = \mathrm{Vect}_{\mathbb{R}}(e_1, ..., e_{j-1}) = \mathrm{Vect}_{\mathbb{R}}(e''_1, ..., e''_{j-1})$. Dado que $P_U + P_{U^\perp} = Id_V$, tenemos que $e''_j = e_j - P_{U_{j-1}}(e_j) = (Id_V - P_{U_{j-1}})(e_j) = P_{U^\perp_{j-1}}(e_j)$, por lo que e''_j es la proyección ortogonal de e_j a $U^\perp_{j-1}$, y en particular, e''_j es ortogonal a U_{j-1} (tal como discutimos en la sección anterior).

***Definición 5.10.12* (simetría ortogonal).** — Sea V un espacio euclideano y $U \subseteq V$ un sub ev. Definimos el endomorfismo $S_U : V \to V$ dado por:

$$x \mapsto S_U(x) = 2P_U(x) - x \iff S_U = 2P_U - Id_V \iff P_U = \frac{S_U + Id_V}{2},$$

que llamaremos **siimetría ortogonal** respecto a U. Si U es un hiperplano , ie $\dim_{\mathbb{R}}(U) = n - 1$, diremos que S_U es una **reflexión** respecto a U.

La terminología usada anteriormente está justificada por las siguientes propiedades:

Proposición 5.10.13. — *Sea V un espacio euclideano y $U \subseteq V$ un sub-ev. Entonces, la simetría ortogonal $S_U : V \to V$ es un* **automorfismo** *(ie, $S_U \in \mathrm{GL}(v)$) que es* **ortogonal** *(ie, $S_U \in O(n)$) y simétrico (ie, $S_U^* = S_U$), que además verifica $S_U^2 = Id_V$.*

Demostración. — La simetría $S_U := 2P_U - Id_V$ es combinación lineal de endomorfismos simétricos y luego es simétrico, pues, $S_U^* = 2P_U^* - Id_V^* = 2P_U - Id_V = S_U$. Más aún,

$$S_U^2 = S_U \circ S_U = (2P_U - Id_V) \circ (2P_U - Id_V) = 4\underbrace{P_U^2}_{=P_U} - 4P_U + Id_V = Id_V$$

En particular, $\det(S_U)^2 = 1 \neq 0$ y luego $S_U \in \mathrm{GL}(V)$. Finalmente, dado que S_U es simétrico, $S_U \circ S_U^* = S_U \circ S_U = Id_V$ y luego $S_U^* = S_U^{-1}$ y por tanto S_U es una isometría (ie, $S_U \in O(n)$). $\square$

Observación 5.10.14. — 1. Notar que $P_U + P_{U^\perp} = Id_V$ implica que:

$$S_{U^\perp} = 2P_{U^\perp} - Id_V = 2(Id_V - P_U) - Id_V = Id_V - 2P_U = -S_U$$

es decir, $S_{U^\perp} = -S_U$.

2. Geométricamente: Dado $x \in V$, el vector $S_U(x)$ está caracterizado por:
   - $x - S_U(x)$ es **ortogonal** a U.
   - El "punto medio" $\frac{x+S_U(x)}{2} = P_U(x)$ pertenece a U.

$$\begin{aligned} S_U : V &\to V \\ x = a+b &\mapsto S_U(x) = a-b \end{aligned}$$

(c.f. conjugación $z \mapsto \overline{z}$ en $\mathbb{C}$)

Ejemplo 5.10.15. — 1. Sea $L := \mathrm{Vect}_{\mathbb{R}}\langle(1,2)\rangle$ recta en $\mathbb{R}^2$ y $e := (1,2)$. Gracias a la fórmula de proyección, tenemos que

$$P_L((1,0)) = \langle(1,0),e\rangle\frac{e}{\|e\|^2} = \frac{1}{5}e \qquad \text{y} \qquad P_L((0,1)) = \langle(0,1),e\rangle\frac{e}{\|e\|^2} = \frac{2}{5}e$$

Luego, la matriz A de P_L asociada a la base canónica será

$$A = \begin{pmatrix} \frac{1}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{4}{5} \end{pmatrix} \text{ y la matriz de } S_L \text{ es } 2A - I_2 = \begin{pmatrix} \frac{-3}{5} & \frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{pmatrix}$$

donde notamos que S_L es simétrica y además es ortogonal (tal y como predice la teoría). Además, $\det(S_L) = -1$ y luego $S_L \in O(n)$ pero $S_L \notin SO(n)$.

2. Sea $\Pi \subseteq \mathbb{R}^3$ el plano de ecuación $3x-y+2z = 0$. Para encontrar la matriz de P_Π respecto a la base canónica, es más sencillo deducirla a partir $P_{\Pi^\perp}$. Note que la recta $L := \Pi^\perp = \mathrm{Vect}_{\mathbb{R}}((3,-1,2))$ y luego

$$P_L((1,0,0)) = \frac{3}{14}e \qquad P_L((0,1,0)) = \frac{-1}{14}e \qquad P_L((0,0,1)) = \frac{2}{14}e \Longrightarrow A = \frac{1}{14}\begin{pmatrix} 9 & -3 & 6 \\ -3 & 1 & -2 \\ 6 & -2 & 4 \end{pmatrix}$$

donde, la matriz B de P_Π será

$$B = I_3 - A = \begin{pmatrix} \frac{5}{14} & \frac{3}{14} & \frac{-3}{14} \\ \frac{3}{14} & \frac{7}{14} & \frac{1}{7} \\ \frac{-3}{7} & \frac{1}{7} & \frac{5}{7} \end{pmatrix}$$

y la matriz S_Π es

$$S_\Pi = 2B - I_3 = \frac{1}{7}\begin{pmatrix} -2 & 3 & -6 \\ 3 & 6 & 2 \\ -6 & 2 & 7 \end{pmatrix}.$$

Observación 5.10.16. — Sea V un espacio euclideano y $U \subseteq V$ un sub-ev. Entonces:

1. P_U es **diagonalizable**. En efecto, $P_U^2 - P_U = 0$ y luego el polinomio minimal m_{P_U} divide a $P(X) = X^2 - X = X(X-1)$ escinde sobre $\mathbb{R}$ y tiene raices simples. Más aún, $V = V_0 \oplus V_1$ donde $V_0 = \ker(P_U) = U^\perp$ y $V_1 = \ker(Id_V - P_U) = \ker(P_{U^\perp}) = U$.
2. S_U es **diagonalizable**. Similar, dado que $S_U^2 = Id_V$ implica que m_{S_U} escinde sobre $\mathbb{R}$ con raices simples. Más aún, $V = V_{-1} \oplus V_1$ donde $V_{-1} = \ker(Id_V + S_U) = \ker(P_U) = U^\perp$ y $V_1 = \ker(S_U - Id_V) = \ker(P_{U^\perp}) = U$.

Teorema 5.10.17. — *Sea V un espacio euclideano. Entonces, todo endomorfismo $u : V \to V$ simétrico (ie, auto-adjunto: $u^* = u$) es diagonalizable en una base ortonormal.*

Demostración. — Apliquemos inducción en $n = \dim_k(V)$. Sea u un endomorfismo simétrico y $\mathscr{B}$ una base ortonormal de V. Entonces, $A = \mathrm{Mat}_{\mathscr{B}}(u)$ es una matriz simétrica. Veamos que posee solo valores propios reales: Sea $\lambda \in \mathbb{C}$ valor propio de A y sea $X \in \mathbb{C}^n \setminus \{0\}$ vector columna tal que $AX = \lambda X$. Puesto que $A \in \mathscr{M}_n(\mathbb{R})$ es real, tenemos que $\overline{AX} = A\overline{X} = \overline{\lambda X}$ y luego, si escribimos $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ entonces

$$
\begin{aligned}
{}^tX\overline{X} = {}^t\overline{X}X = (x_1 \cdots x_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} &= x_1\overline{x_1} + ... + x_n\overline{x_n} \\
&= |x_1|^2 + ... + |x_n|^2 = \underbrace{\sum_{i=1}^{n} |x_i|^2}_{\in\mathbb{R}!!} > 0
\end{aligned}
$$

Se sigue que

$$\overline{\lambda}\,{}^tX\overline{X} = {}^tX(\overline{\lambda X}) = {}^tXA\overline{X} \underbrace{=}_{\text{Matriz } 1\times 1} {}^t({}^tXA\overline{X}) = {}^t\overline{X}\,{}^tAX = {}^t\overline{X}AX = \lambda\,{}^t\overline{X}X$$

y dado que $\overline{X}X$ es un real no nulo, concluimos que $\lambda = \overline{\lambda}$, y luego $\lambda \in \mathbb{R}$. Finalmente, si $\lambda \in \mathbb{R}$ es un valor propio de u y si $x \in V \setminus \{0\}$ es un vector propio asociado a λ, entonces consideramos la recta $L = \mathrm{Vect}_{\mathbb{R}}(x)$ y $H := L^{\perp}$ el hiperplano ortogonal a L. Observe que para cada $y \in H$ se tiene que

$$\langle u(y), x\rangle = \langle y, u^*(x)\rangle = \langle y, u(x)\rangle = \lambda = 0$$

por lo que $u(y) \in L^{\perp} = H$, ie, $u(H) \subseteq H$ (H es estable bajo u). Notamos que $H \cong \mathbb{R}^{n-1}$ es un espacio euclideano y que la restricción $u|_H : H \to H$ es un endomorfismo simétrico. Luego, la hipótesis inductiva implica que $u|_H$ es diagonalizable en cierta base ortonormal $(e_2, ..., e_n)$ de $H = L^{\perp}$. Finalmente, u es diagonalizable ortonormalmente en la base $(\frac{x}{\|x\|}, e_2, ..., e_n)$ de V. $\square$

¡Importante! Sea V un espacio euclideano y $u : V \to V$ un endomorfismo simétrico. La demostración anterior dice, en particular, que "vectores propios asociados a valores propios **distintos** son ortogonales". La afirmación anterior también se puede hacer "a mano" de la siguiente manera: Si $\lambda \neq \mu$ y si $x, y \in V \setminus \{0\}$ cumplen $u(x) = \lambda x$ y $u(y) = \mu y$. Luego $\lambda\langle x, y\rangle = \langle u(x), y\rangle = \langle x, u(y)\rangle = \mu\langle x, y\rangle$, y por tanto $\langle x, y\rangle = 0$ (ya que $\lambda \neq \mu$).

Comunmente se dice que V es la **suma directa ortogonal** de los espacios propios $V_{\lambda_1}, ..., V_{\lambda_p}$ de u, y escribimos

$$V = V_{\lambda_1} \oplus^{\perp} V_{\lambda_2} \oplus^{\perp} ... \oplus^{\perp} V_{\lambda_p} = \bigoplus_{j=1}^{p}{}^{\perp} V_{\lambda_j}.$$

Corolario 5.10.18. — *Sea $A \in \mathscr{M}_n(\mathbb{R})$ matriz simétrica real. Entonces, existe una matriz $P \in \mathrm{GL}_n(\mathbb{R})$* **ortogonal** *(ie, ${}^tPP = I_n$) y una matriz diagonal $D \in \mathscr{M}_n(\mathbb{R})$* **diagonal real** *tal que $A = PDP^{-1} = PDP^t$.*

Demostración. — Sea $u = u_A : \mathbb{R}^n \to \mathbb{R}^n$, $x \mapsto Ax$ el endomorfismo asociado a A. Si $\mathscr{C}$ es la base canónica de $\mathbb{R}^n$, entonces $\mathrm{Mat}_{\mathscr{C}}(u) = A$ matriz simétrica, por lo que $u^* = u$. Gracias al teorema anterior, existe $\mathscr{B}$ base ortonormal de $\mathbb{R}^n$ tal que $D = \mathrm{Mat}_{\mathscr{B}}(u)$ es diagonal real. Por otra parte, la matriz cambio de base $P = \mathrm{Mat}_{\mathscr{C}}(\mathscr{B})$ es ortogonal y $\mathrm{Mat}_{\mathscr{B}}(u) = \mathrm{Mat}_{\mathscr{B}}(\mathscr{C})\,\mathrm{Mat}_{\mathscr{C}}(u)\,\mathrm{Mat}_{\mathscr{C}}(B)$, es decir, $D = P^{-1}AP$. □

Ejemplo 5.10.19. — Sea $A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 3 & 0 \\ 2 & 0 & 1 \end{pmatrix} \in \mathscr{M}_3(\mathbb{R})$ matriz simétrica real.

Su polinomio característico es

$$\begin{aligned} P_A(x) = \det(XI_3 - A) = \begin{vmatrix} X-1 & 0 & -2 \\ 0 & X-3 & 0 \\ -2 & 0 & X-1 \end{vmatrix} &= (X-3)\begin{vmatrix} X-1 & -2 \\ -2 & X-1 \end{vmatrix} \\ &= (X-3)(X^2 - 2X + 1 - 4) \\ &= (X-3)^2(X+1). \end{aligned}$$

Se sigue que $\lambda = -1$ y $\mu = 3$ son sus valores propios. Encontremos los vectores propios:

$\boxed{\lambda = -1}$:

$$A + I_3 = \begin{pmatrix} 2 & 0 & 2 \\ 0 & 4 & 0 \\ 2 & 0 & 2 \end{pmatrix} \implies \begin{pmatrix} 2 & 0 & 2 \\ 0 & 4 & 0 \\ 2 & 0 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x+2z \\ 4y \\ 2x+2z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Así, obtenemos que $y = 0$ y $x = -z$, por lo que $v = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$ es un vector propios. **Normalizamos:** $v_1 = \begin{pmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{pmatrix}$ genera V_λ y $\|v_1\| = 1$.

$\boxed{\mu = 3}$:

$$A - 3I_3 = \begin{pmatrix} -2 & 0 & 2 \\ 0 & 0 & 0 \\ 2 & 0 & -2 \end{pmatrix} \Longrightarrow \begin{pmatrix} -2 & 0 & 2 \\ 0 & 0 & 0 \\ 2 & 0 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2z - 2x \\ 0 \\ 2x - 2z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Longrightarrow x - z = 0.$$

❦ Notar que el plano $x - z = 0$ es precisamente ortogonal a V_λ, ie $V_\mu = V_\lambda^\perp$ (de hecho, se podría haber calculado de esta forma). Luego V_μ está generado por $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ y $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$. **Normalizamos:** $v_2 = \begin{pmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{pmatrix}$ y $v_3 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$.

¡Atención! — En general, se busca una base cualquiera del espacio propio, y se le aplica el proceso de Gram-Schmidt para ortornormalizarla. Luego, obtenemos que

$$\underbrace{\begin{pmatrix} 1 & 0 & 2 \\ 0 & 3 & 0 \\ 2 & 0 & 1 \end{pmatrix}}_{A} = \underbrace{\begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \\ -1/\sqrt{2} & 1/\sqrt{2} & 0 \end{pmatrix}}_{P} \underbrace{\begin{pmatrix} -1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}}_{D} \underbrace{\begin{pmatrix} 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 1 & 0 \end{pmatrix}}_{{}^tP = P^{-1}}.$$

CAPÍTULO 6

ISOMETRÍAS

En esta sección estudiaremos en detalle las funciones ya conocidas que anteriormente nombramos como isometrías. Estas funciones son especiales ya que como se observó anteriormente, preservan la norma del espacio euclideano, o geométricamente, preservan las distancias entre todos los elementos del espacio en cuestión. Para estudiar las isometrías de una manera completa y adecuada, comenzaremos analizando las isometrías del espacio euclídeo de dimensión 2, para luego proseguir con espacios de dimensión 3 y finalmente, extender estas ideas a dimensión finita arbitraria. La noción de orientación del espacio será clave en este proceso. Se recomienda, para entender mentalmente los conceptos siguientes, fijar la atención en el espacio $\mathbb{R}^n$.

6.1. Isometrías del plano y orientación

Sea $V \cong \mathbb{R}^2$ un espacio euclideano (ie, $\dim_{\mathbb{R}}(V) = 2$), al cual nos referiremos en esta sección como plano euclideano, y $\mathscr{B}$ una base ortonormal de V. Notar que las isometrías de V corresponden a los elementos del grupo $\mathrm{O}(2)$ (ver Sección 5.7). Sea entonces $u : V \cong V$ con $u \in \mathrm{O}(2)$. Claramente $A = \mathrm{Mat}_{\mathscr{B}}(u)$ es ortogonal, por lo que si denotamos:

$$A = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \in \mathrm{GL}_2(\mathbb{R})$$

y $\delta := \det(A) = \pm 1$ entonces se tiene que:

$$A = {}^{\mathrm{t}}(A)^{-1} = \frac{1}{\delta}\, {}^{\mathrm{t}}\!\begin{pmatrix} d & -c \\ -b & a \end{pmatrix} = \frac{1}{\delta}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

lo cual significa que $d = \delta a, c = -\delta b$ y $\delta = ad - bc = \delta(a^2 + b^2)$.

Proposición 6.1.1. — *Sea $A \in \mathrm{GL}_2(\mathbb{R})$ la matriz de una isometría indirecta (ie, $\delta = -1$) $u : V \to V$ respecto a una base ortonormal. Entonces la matriz A tiene la forma*

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

con $a^2 + b^2 = 1$. Además, existe un subespacio $L \subseteq V$ con $\dim_{\mathbb{R}}(L) = 1$ tal que $u = S_L$ es la reflexión del espacio V respecto a L.

Demostración. — Por la discusión anterior, basta notar que si $\delta = -1$ entonces $d = -a, c = b$ y $a^2 + b^2 = 1$. Debido a las igualdades entre coeficientes, es sencillo notar que $A = {}^{\mathrm{t}}A$ y entonces $A\,{}^{\mathrm{t}}A = A^2 = I_2$. De esto se deduce que los valores propios de A son 1 y -1, por lo que tomando $L = V_1$, u resulta ser la simetría ortogonal respecto al subespacio L. □

Dado que es posible imaginar el espacio V como un plano, y L como una recta en dicho espacio, la idea clave de la proposición anterior es que una isometría indirecta en dimensión 2 no es más que una reflexión con respecto a alguna recta.

Corolario 6.1.2. — *Toda isometría de un plano euclideano es producto de a lo más dos reflexiones.*

Demostración. — Sea $V \cong \mathbb{R}^2$ espacio euclideano y $u \in \mathrm{O}(2)$. Si $\det(u) = -1$ entonces por la proposición anterior u es una reflexión. Si $\det(u) = 1$ y $s \in O(2)$ es una reflexión, entonces $\det(s \circ u) = \det(s)\det(u) = -1$ y $s \circ u$ resulta ser una reflexión. Dado que $s^2 = id_V$ se tiene que $u = s \circ (s \circ u)$ se escribe como producto de dos reflexiones. □

Sea $U \subseteq \mathbb{C}$ el círculo unitario complejo, es decir,

$$U = \{z \in \mathbb{C} \mid |z| = 1\} = \{a + ib \in \mathbb{C} \mid a, b \in \mathbb{R} \wedge a^2 + b^2 = 1\}$$

Sin mayor complicación se puede mostrar que U es un grupo abeliano respecto a la multiplicación. Esto motiva la siguiente proposición.

Proposición 6.1.3. — *Sea $A \in \mathrm{GL}_2(\mathbb{R})$ la matriz de una isometría directa $u : V \to V$ respecto a una base ortonormal. Entonces la matriz A tiene la* *forma*

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

con $a^2+b^2=1$. *Más aún,* $\varphi : U \xrightarrow{\sim} SO(2)$, $a+ib \mapsto \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ *es un morfismo de grupos. En particular,* $SO(2)$ *es un grupo abeliano.*

Demostración. — Si $\delta = 1$ entonces $d = a, c = -b$ y $a^2 + b^2 = 1$, de lo cual se deduce la forma de la matriz. Por definición, claramente la aplicación φ definida en el enunciado es biyectiva. Además, si $z, z' \in U$ con un sencillo cálculo se puede verificar que $\varphi(zz') = \varphi(z)\varphi(z')$, por lo que φ es un morfismo de grupos. □

Definición 6.1.4. — Sea V un plano euclideano. Se designan los elementos del grupo SO(2) como **rotaciones**. Como SO(2) es un grupo abeliano, si $r_1, r_2 \in \mathrm{SO}(2)$ son rotaciones se verifica que $r_1 \circ r_2 = r_2 \circ r_1$.

Observación 6.1.5. — Sea $r : V \to V$ una rotación del plano euclideano y $\mathscr{B}$ base ortonormal de V. Entonces por la proposición anterior la matriz $R = \mathrm{Mat}_{\mathscr{B}}(r)$ tiene la forma

$$R = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

con $a^2 + b^2 = 1$. Es inevitable, debido a la forma de la matriz R, no asociarla con las matrices de rotación de la forma

$$R_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

que de seguro el lector ya ha estudiado o ha visto previamente. Es imposible entonces (más aún por el nombre de rotaciones introducido) resistirse a la idea de asociar un "ángulo" $\theta \in \mathbb{R}$ de tal forma que $R = R_\theta$. Esto significa que $a = cos(\theta)$ y $b = sin(\theta)$, y lo más importante, es que $2a = \mathrm{tr}(r) \iff a = \frac{1}{2}\mathrm{tr}(r)$, y dado que la traza es una propiedad intrínseca de cada matriz, el valor de a está bien definido para todo $r \in \mathrm{SO}(2)$. A pesar de este increíble hecho, es crucial notar que el coeficiente b viene dado por la expresión $a^2 + b^2 = 1$ y entonces dicho coefieciente no está definido de manera única (del cálculo en una variable es conocido que $sin(\theta)$ es una función impar, por lo que se requiere escoger un "signo"). Analizemos en profundidad este hecho. Sea $\mathscr{B}'$ otra base ortonormal de V y $P = \mathrm{Mat}_{\mathscr{B}}(\mathscr{B}')$. Recordar que dicha matriz resulta ser ortogonal. Debemos tener en cuenta los siguientes casos:

1. Si $\det(P) = 1$ entonces $P \in SO(2)$, y por la conmutatividad de dicho grupo $PR = RP$ y se deduce que $R' = \mathrm{Mat}_{\mathscr{B}'}(r) = P^{-1}RP = P^{-1}PR = R$.
2. Si $\det(P) = -1$ por la Proposición 6.1.1 P es una reflexión y $\det(PR) = -1$, siendo PR también una reflexión y entonces $P^2 = (PR)^2 = I_2$. Entonces

$$R' = P^{-1}RP = PRPRR^{-1} = (PR)^2R^{-1} = R^{-1} = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$

De la discusión anterior, se puede concluir que al cambiar la base del plano, existen únicamente dos opciones que pueden ocurrir con las rotaciones. En primer lugar, el cambio de base puede preservar el sentido de la rotación, o en última instancia, este puede invertir el sentido de rotación. No obstante, se puede apreciar que la magnitud del ángulo de rotación permanece inalterado.

Proposición 6.1.6. — *Sea $V \cong \mathbb{R}^n$. Se define la relación de equivalencia en el conjunto de bases del espacio V:*

$$\mathscr{B} \sim \mathscr{B}' \iff \det_{\mathscr{B}}(\mathscr{B}') > 0$$

Esta relación define únicamente dos clases de equivalencia las cuales se denominarán **orientaciones** *de V.*

Demostración. — Dado que para toda base $\mathscr{B}$ se tiene que $\det_{\mathscr{B}}(\mathscr{B}) = 1 \Rightarrow \mathscr{B} \sim \mathscr{B}$. Si $\mathscr{B}, \mathscr{B}', \mathscr{B}''$ son bases de V, entonces $\mathrm{Mat}_{\mathscr{B}}(\mathscr{B}'') = \mathrm{Mat}_{\mathscr{B}}(\mathscr{B}')\,\mathrm{Mat}_{\mathscr{B}'}(\mathscr{B}'') \Rightarrow \det_{\mathscr{B}}(\mathscr{B}'') = \det_{\mathscr{B}}(\mathscr{B}')$ d y se verifica que $\mathscr{B} \sim \mathscr{B}' \wedge \mathscr{B}' \sim \mathscr{B}'' \Rightarrow \mathscr{B} \sim \mathscr{B}''$. Si se toma $\mathscr{B}'' = \mathscr{B}$ en la relación anterior se obtiene $\det_{\mathscr{B}}(\mathscr{B}')\det_{\mathscr{B}'}(\mathscr{B}) = 1$, lo cual significa que $\mathscr{B} \sim \mathscr{B}' \iff \mathscr{B}' \sim \mathscr{B}$. Finalmente, $\sim$ es una relación de equivalencia.
Sean ahora $\mathscr{B}_0 = (e_1, \ldots, e_n)$ y $\mathscr{C}_0 := (-e_1, e_2, \ldots, e_n)$ bases de V. Notar que $\det_{\mathscr{B}_0}(\mathscr{C}_0) = -1$, por lo que $\mathscr{B}_0 \nsim \mathscr{C}_0$. Ahora veremos que $[\mathscr{B}_0]$ y $[\mathscr{C}_0]$ son las únicas clases de equivalencia. Sea $\mathscr{B}$ una base de V. Si $\det_{\mathscr{B}_0}(\mathscr{B}) > 0$, por definición $\mathscr{B} \in [\mathscr{B}_0]$. Supongamos entonces que $\det_{\mathscr{B}_0}(\mathscr{B}) < 0$. Entonces:

$$\det_{\mathscr{C}_0}(\mathscr{B}) = \underbrace{\det_{\mathscr{C}_0}(\mathscr{B}_0)}_{=-1}\underbrace{\det_{\mathscr{B}_0}(\mathscr{B})}_{<0} > 0 \Rightarrow \mathscr{B} \in [\mathscr{C}_0]$$

$\square$

Observación 6.1.7. — Sea $V \cong \mathbb{R}^n$ un $\mathbb{R}$-espacio vectorial. De la proposición anterior notamos que orientar el espacio V es un proceso siempre dependiente de la elección de una base. Dada una base $\mathscr{B}_0$ de V, elegir una orientación entonces, corresponde a seleccionar la clase de equivalencia de bases de V tales que $\det_{\mathscr{B}_0}(\mathscr{B}) > 0$, las cuales denominaremos **bases directas**, o por el contrario, la otra clase, cuyos elementos se conocerán como **bases indirectas**. Si en V se ha seleccionado una orientación, nos referiremos a V como un **espacio orientado**.
En general, se conocerá como **orientación canónica** al conjunto de bases tal que la base canónica es directa.

Definición 6.1.8. — Sea $V \cong \mathbb{R}^n$ un plano euclideano orientado y $r : V \xrightarrow{\sim} V$ una rotación de V. Entonces para toda base ortonormal directa $\mathscr{B}$ de V existe $\theta \in \mathbb{R}$ (el cual es único módulo $2\pi\mathbb{Z}$) tal que

$$R = \mathrm{Mat}_{\mathscr{B}}(r) = R_\theta = \begin{pmatrix} cos(\theta) & -sin(\theta) \\ sin(\theta) & cos(\theta) \end{pmatrix}$$

Nombraremos a θ como el **ángulo de rotación** y utilizaremos la notación $r = r_\theta$ para hacer referencia a éste.

Ejercicio 6.1.9. — Sea V un plano euclideano orientado y $\theta, \theta' \in \mathbb{R}$. Probar que las rotaciones $r_\theta, r_{\theta'} \in \mathrm{SO}(2)$ verifican las relaciones $r_\theta \circ r_{\theta'} = r_{\theta+\theta'}, r_0 = id_V, r_\pi = -id_V$.

Lema 6.1.10. — *Sea V un plano euclideano, y $x, y \in V \setminus \{0\}$ tales que $\|x\| = \|y\|$. Entonces existe una única rotación $r : V \to V$ tal que $r(x) = y$.*

Demostración. — Definamos el vector unitario $e_1 = \frac{x}{\|x\|}$, y sea $e_2 \in V$ tal que el conjunto $\mathscr{B} = (e_1, e_2)$ es una base ortonormal de V. Luego existen $a, b \in \mathbb{R}$ tales que se tiene la igualdad $\frac{y}{\|y\|} = ae_1 + be_2$. Dado que dicho vector es unitario claramente $a^2 + b^2 = 1$. Entonces la siguiente matriz corresponde a una rotación:

$$R = \mathrm{Mat}_{\mathscr{B}}(r) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

de la cual es posible notar que $r(e_1) = ae_1 + be_2 \iff r(\frac{x}{\|x\|}) = \frac{y}{\|y\|}$, y dado que $\|x\| = \|y\|$ se tiene $r(x) = y$.
Para la unicidad, suponer que existe $r' \in \mathrm{SO}(2)$ tal que $r'(x) = y$. Como r' es un isomorfismo $((r')^{-1} \circ r)(x) = x$, y por ende $\lambda_1 = 1$ es valor propio de

$(r')^{-1} \circ r$. En consecuencia,

$$\det((r')^{-1} \circ r) = 1 = \lambda_1 \lambda_2 \Rightarrow \lambda_2 = 1$$

Como los dos valores propios de $(r')^{-1} \circ r$ son 1, se concluye que $(r')^{-1} \circ r = id_V$, y por ende $r' = r$. □

Definición 6.1.11. — Sea $V \cong \mathbb{R}^2$ un plano euclideano orientado y $x, y \in V \setminus \{0\}$. Si r_θ es la única rotación tal que $r(\frac{x}{\|x\|}) = \frac{y}{\|y\|}$ (lo cual está bien definido gracias al lema anterior), se define el **ángulo orientado** entre x e y como:

$$\measuredangle(x, y) := \theta$$

donde $\theta \in \mathbb{R}$ es único módulo $2\pi\mathbb{Z}$. Si se considera $\theta \in [-\pi, \pi]$, entonces el ángulo resulta ser único.

Observación 6.1.12. — Como consecuencia de la demostración del lema 6.1.10, se observa que:

$$\frac{y}{\|y\|} = ae_1 + be_2 \Rightarrow a = cos\measuredangle(x, y) = \left\langle e_1, \frac{y}{\|y\|} \right\rangle = \frac{\langle x, y\rangle}{\|x\|\|y\|} \Rightarrow cos\measuredangle(x, y) = \frac{\langle x, y\rangle}{\|x\|\|y\|}$$

y además si $\mathscr{B}'$ es cualquier base ortonormal directa:

$$b = \begin{vmatrix} 1 & a \\ 0 & b \end{vmatrix} = \det_{\mathscr{B}}\left(e_1, \frac{y}{\|y\|}\right) = \frac{\det_{\mathscr{B}}(x, y)}{\|x\|\|y\|} = \underbrace{\det_{\mathscr{B}}(\mathscr{B}')}_{=1} \frac{\det_{\mathscr{B}'}}{\|x\|\|y\|} \Rightarrow sin(\measuredangle(x, y)) = \frac{\det_{\mathscr{B}}(x, y)}{\|x\|\|y\|}$$

Si consideramos $-\pi < \measuredangle(x, y) < \pi$, entonces el ángulo no orientado definido en 5.8.8 coincide con $|\measuredangle(x, y)|$.
Si consideramos una rotación $r \in \mathrm{SO}(2)$, entonces $(r \circ r_\theta)(x) = r(y)$, dado que $\mathrm{SO}(2)$ es conmutativo $r_\theta \circ r = r \circ r_\theta$ entonces $(r \circ r_\theta)(x) = r_\theta(r(x))$ para todo $x, y \in V$ tal que $r_\theta(x) = y$. De esto notamos que $\measuredangle(r(x), r(y)) = \measuredangle(x, y)$ (mod $2\pi\mathbb{Z}$), lo cual no significa nada más que las rotaciones preservan los ángulos entre vectores, en el sentido que el ángulo definido por cada rotación es único.

Ejercicio 6.1.13. — Sea $s : V \to V$ una reflexión. Demostrar que $\measuredangle(s(x), s(y)) = -\measuredangle(x, y) \pmod{2\pi\mathbb{Z}}$. Probar además que, dados $x, y, z \in V \setminus \{0\}$, entonces

$$\measuredangle(x, y) + \measuredangle(y, z) = \measuredangle(x, z), \quad \measuredangle(y, x) = \measuredangle(x, -y), \quad \measuredangle(x, -y) = \measuredangle(x, y) + \pi \pmod{2\pi\mathbb{Z}}$$

[**Indicación:** utilizar ejercicio 6.1.9].

***Definición 6.1.14* (ángulo entre rectas).** — Sean $L_1, L_2 \subseteq V$ rectas en un plano euclideano orientado V, y considere e_1, e_2 vectores directores unitarios respectivos. Se define el **ángulo entre rectas** como la cantidad $\measuredangle(L_1, L_2) := \measuredangle(e_1, e_2) \pmod{\pi\mathbb{Z}}$. Si consideramos $\measuredangle(L_1, L_2) \in [0, \pi[$, entonces dicho ángulo es único.

Proposición 6.1.15. — *Sea V un plano euclideano orientado y $L_1, L_2 \subseteq$ rectas. Entonces se cumple que:*

$$s_{L_2} \circ s_{L_1} = r_{2\measuredangle(L_1, L_2)}$$

Demostración. — Considere los vectores directores unitarios e_1, e_2 de las rectas L_1, L_2. Dado que estamos considerando reflexiones, $\det(s_{L_2} \circ s_{L_1}) = (-1)^2 = 1$ y se deduce que $r_\theta := s_{L_2} \circ s_{L_1} \in \mathrm{SO}(2)$ es una rotación. Luego es suficiente calcular el ángulo de rotación θ, que corresponde a:

$$\begin{aligned} \theta = \measuredangle(e_1, s_{L_2}(s_{L_1}(e_1))) &= \measuredangle(e_1, s_{L_2}(e_1)) = \\ \underbrace{=}_{\text{Ejercicio 6.1.13}} \measuredangle(e_1, e_2) + \measuredangle(e_2, s_{L_2}(e_1)) &= \measuredangle(e_1, e_2) - \measuredangle(e_2, e_1) = 2\measuredangle(e_1, e_2) \end{aligned}$$

□

6.2. Isometrías del espacio

En esta sección consideraremos un espacio euclideano $V \cong \mathbb{R}^3$, ie, de dimensión 3. El objetivo será entonces describir las isometrias de un espacio tridimensional, para lo cual emplearemos principalmente ideas y argumentos geométricos.

Teorema 6.2.1. — *Sea $V \cong \mathbb{R}^3$ espacio euclideano y sea $u : V \to V$ una isometría. Entonces existe una base ortonormal $\mathscr{B}$ de V de tal forma que $A = \mathrm{Mat}_{\mathscr{B}}(u)$ es de la forma*

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

donde se verifica que $a^2 + b^2 = 1$ y $\delta = \det(u) = \pm 1$.

Demostración. — En primer lugar, notar que el polinomio característico $P_u \in \mathbb{R}[X]$ es de grado 3, y entonces por continuidad es posible asegurar la existencia de una raíz real $\lambda \in \mathbb{R}$. Sea ahora $x \in V$ vector propio asociado a λ. Entonces dado que u es isometría

$$\|x\| = \|u(x)\| = \|\lambda x\| = |\lambda|\|x\| \underset{x\neq 0}{\Rightarrow} |\lambda| = 1$$

Por lo anterior podemos distinguir dos casos distintos:

1. Suponer $\lambda = 1$. Entonces existe $x \neq 0$ tal que $u(x) = x$ el cual podemos suponer unitario, es decir, $\|x\| = 1$. Considerar $L = \mathrm{Vect}_{\mathbb{R}}(x)$ la cual es estable por u ($u(L) = L$). Sea $\Pi = L^{\perp}$ su plano ortogonal el cual es también estable ya que $u(L^{\perp}) = u(L)^{\perp}$. Notar que $\Pi \cong \mathbb{R}^2$. Entonces consideramos $u_\Pi := u|_\Pi : \Pi \xrightarrow{\sim} \Pi$ isometría. Es claro que $\det(u) = \lambda \det(u_\Pi) = \det(u_\Pi)$. Entonces podemos distinguir dos casos:
   a) Si $\det(u_\Pi) = -1$ entonces $u_\Pi \in O(2)$ es una reflexión y existe una base ortonormal $\mathscr{B}' = (e_1, e_2)$ de vectores propios de Π de tal forma que
   $$B := \mathrm{Mat}_{\mathscr{B}'}(u_\Pi) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$
   Luego tomando la base $\mathscr{B} = (e_2, e_1, x)$ de V se tiene que
   $$A = \mathrm{Mat}_{\mathscr{B}}(u) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$
   b) Si $\det(u_\Pi) = 1$ entonces u_Π es una rotación. Luego existe una base ortonormal $\mathscr{B}' = (e_1, e_2)$ de Π tal que
   $$B = \mathrm{Mat}_{\mathscr{B}'}(u_\Pi) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$
   donde $a^2 + b^2 = 1$. Entonces si consideramos la base ortonormal $\mathscr{B} = (x, e_1, e_2)$ se verifica que
   $$A = \mathrm{Mat}_{\mathscr{B}}(u_\Pi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & a & -b \\ 0 & b & a \end{pmatrix}$$
2. Suponer ahora $\lambda = -1$. Entonces $-u \in O(3)$ verifica que $\det(-u) = (-1)^3 \det(u) = -\det(u)$ y basta aplicar el caso anterior a la isometría $-u$.

$\square$

Observación 6.2.2. — Del teorema anterior es posible notar que $tr(u) = 2a+\delta$ y entonces $a = (\mathrm{tr}(u) - \det(u))/2$ está definido completamente por u. No obstante, esto no es cierto para b, dado que b está determinado por $a^2+b^2 = 1$, y entonces se debe "escoger un signo" (de la misma forma que ocurría con las isometrías del plano!).

Proposición 6.2.3. — *Sea V espacio euclideano y $u \in SO(3)$ isometría directa. Entonces se verifica que $-1 \leq \mathrm{tr}(u) \leq 3$ y $\mathrm{tr}(u) = 3 \iff u = id_V$. Si* $\mathrm{tr}(3)$ *entonces existe una única recta invariante $L \subseteq V$ de tal forma que $u|_\Pi$ (donde $\Pi = L^\perp$) es una rotación. Diremos en este caso que u es una* **rotación de eje L.**

Demostración. — Por el teorema anterior existe $\mathscr{B}$ base ortonormal de V tal que

$$A = \mathrm{Mat}_{\mathscr{B}}(u) = \begin{pmatrix} \delta & a & -b \\ 0 & b & a \end{pmatrix}$$

donde $\delta = 1$ y $a^2 + b^2 = 1$. Entonces $-1 \leq a \leq 1$ y

$$\mathrm{tr}(u) = 1 + 2a \Rightarrow 1 - 2 = -1 \leq \mathrm{tr}(u) \leq 3 = 1 + 2$$

y si $\mathrm{tr}(u) = 3 \iff a = 1 \wedge b = 0 \iff u = id_V$. A continuación notamos del polinomio característico que

$$P_u(x) = (x - \delta)(x^2 - 2ax + 1) = (x - 1)(x^2 - (\mathrm{tr}(u) - 1)x + 1)$$

y que $\lambda = 1$ es una raíz del polinomio $x^2-(\mathrm{tr}(u)-1)x+1 \iff 1-\mathrm{tr}(u)+2 = 3-\mathrm{tr}(u) = 0 \iff \mathrm{tr}(u) =$ Ahora, si $\mathrm{tr}(u) < 3$ entonces $L := V_1 = \ker(u - id_v)$ es una recta invariante de u (pues $\lambda = 1$ es valor propio por lo anterior). Consideramos $\Pi := L^\perp$ su plano ortogonal y entonces $u(\Pi) = \Pi$ y $\det(u) = \det(u_\Pi) = 1$, de donde vemos que $u|_\Pi : \Pi \xrightarrow{\sim} \Pi$ es una rotación del plano Π. □

Es importante notar que si orientamos el plano Π entonces $u|_\Pi = r_\theta$ para un único $\theta \in \mathbb{R} \pmod{2\pi\mathbb{Z}}$, y que además verifica que $\mathrm{tr}(u) = 2\cos(\theta) + 1$.

Observación 6.2.4. — Es fundamental notar que no es suficiente con escoger una orientación para el espacio $V \cong \mathbb{R}^3$ para determinar el ángulo de rotación θ. Para ello es necesario también definir una orientación para $\Pi = L^\perp$.
Para solucionar dicho problema, utilizaremos el siguiente convenio. Sea $V \cong \mathbb{R}^3$ espacio euclideano orientado, esto es, se ha fijado una orientación con antelación, diremos que una base (e_1, e_2) de Π es directa si la base (e, e_1, e_2) de

V es directa, donde e es tal que $L = \mathrm{Vect}_k(e)$. De esta forma, orientar L[1] induce una orientación natural en el plano Π [2].

Ejercicio 6.2.5. — Demostrar que $SO(3)$ no es conmutativo. (**Indicación: Considere $V \cong \mathbb{R}^3$ y las rotaciones r_1, r_2 de los planos XY e YZ respectivamente en $\pi/2$ y observe que $r_1r_2 \neq r_2r_1$.**

Corolario 6.2.6. — *Sea $V \cong \mathbb{R}^3$ espacio euclideano y $u \in SO(3)$ isometría directa. Entonces u es el producto de dos reflexiones.*

Demostración. — Si $u = id_v$ entonces $id_V = s \circ s = s^2$ para toda reflexión $s_V \to V$. Suponer entonces que $u \neq id_v$. Entonces u es una rotación de eje L y su restricción al plano ortogonal $\Pi = L^{\perp}$ es una rotación del plano $\Pi \cong \mathbb{R}^2$. Entonces por el Corolario 6.1.2dicha restricción es producto de dos reflexiones. De forma explícita, si Π_1 es un plano que contiene a L, $S_{\Pi_1} \circ u$ es una reflexión respecto a un plano Π_2 conteniendo a L. Así

$$S_{\Pi_1} \circ u = S_{\Pi_2} \implies u = S_{\Pi_1}^{-1} \circ S_{\Pi_2} = S_{\Pi_1} \circ S_{\Pi_2}$$

□

Proposición 6.2.7. — *Sea $V \cong \mathbb{R}^3$ plano euclideano y $u \in O(3)$ una isometría indirecta ($\det(u) = \delta = -1$). Entonces existe una recta $L \subseteq V$ invariante de tal forma que $u = r_L \circ S_\Pi = S_\Pi \circ r_L$ es el producto conmutativo de una rotación $r_L \in SO(3)$ de eje L y una reflexión s_Π respecto al plano ortogonal $\Pi = L^{\perp}$. Además, u es el producto de tres reflexiones.*

Demostración. — Sea $\mathscr{B} = (e_1, e_2, e_3)$ una base ortonormal de V tal que

$$A = \mathrm{Mat}_{\mathscr{B}}(u) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & a & -b \\ 0 & b & a \end{pmatrix}$$

con $a^2 + b^2 = 1$. Notar que

(1) Geométricamente orientar la recta L es escoger un vector director de L que esté en dirección hacia 'arriba' o 'abajo'.

(2) Esto es lo que generalmente se conoce como la 'regla de la mano derecha', pues definir una orientación en el plano Π equivale a escoger como positivo el sentido 'antihorario' u 'horario'.

$$\underbrace{\begin{pmatrix} -1 & 0 & 0 \\ 0 & a & -b \\ 0 & b & a \end{pmatrix}}_{A} = \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & a & -b \\ 0 & b & a \end{pmatrix}}_{R} \underbrace{\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{S} = \underbrace{\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{S} \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & a & -b \\ 0 & b & a \end{pmatrix}}_{R}$$

Si $L := \mathrm{Vect}_{\mathbb{R}}(e_1)$ entonces claramente $\Pi = \mathrm{Vect}_{\mathbb{R}}(e_2, e_3)$ y $R = \mathrm{Mat}_{\mathscr{B}}(r_L)$ es una rotación de eje L, mientras que $S = \mathrm{Mat}_{\mathscr{B}}(S_\Pi)$ es la reflexión respecto al plano Π. Del corolario anterior sabemos que r_L es producto de dos reflexiones, por lo que A es producto de 3 reflexiones. $\square$

Para finalizar esta sección se dará una definición formal del **producto vectorial** o **producto cruz**, el cual será de utilidad para determinar los ángulos de rotación de los elementos de $\mathrm{SO}(3)$.
Sea $V \cong \mathbb{R}^3$ espacio euclideano orientado y $\mathscr{B}$ una base ortonormal de $\mathscr{B}$. Definimos la aplicación lineal

$$f : V \to \mathbb{R} \quad , \qquad z \mapsto f(z) = \det_{\mathscr{B}}(x, y, z)$$

en donde $x, y \in V$ son vectores fijos. Notar que dicha aplicación no depende de la base ortonormal directa escogida, pues si $\mathscr{B}'$ es ortonormal directa entonces $\det_{\mathscr{B}'}(\mathscr{B}) = 1$. Si denotamos por $B = \langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ al producto escalar entonces $\widehat{B} : V \xrightarrow{\sim} V^*$ tal que $\widehat{B}(y) = \langle \cdot, y \rangle$ es un isomorfismo. Por lo tanto, dados $x, y \in V$ existe un único vector en V asociado a $\widehat{B}^{-1}(f)$. Estas dos observaciones nos permiten definir el siguiente objeto.

Definición 6.2.8. — Sea $V \cong \mathbb{R}^3$ espacio euclideano orientado, y $x, y \in V$. Se define el **producto cruz** entre x e y como el único vector denotado por $x \times y \in V$ de tal forma que para toda $\mathscr{B}$ base ortonormal directa se verifica que

$$\det_{\mathscr{B}}(x, y, z) = \langle x \times y, z \rangle \quad \forall z \in V$$

Proposición 6.2.9. — *Sea $V \cong \mathbb{R}^3$ espacio euclideano orientado. Entonces*

1. *La aplicación*

$$V \times V \to V, \qquad (x, y) \mapsto x \times y$$

   es bilineal alternada.
2. *El producto $x \times y$ es ortogonal a x e y. Además $x \times y = 0 \iff x$ e y son linealmente dependientes. Si x e y no son colineales entonces $(x, y, x \times y)$ es una base directa de V.*

3. *Si* $\theta = \theta(x,y)$ *es el ángulo no orientado entre* x *e* y *entonces* $\|x \times y\| = \|x\|\|y\||\sin(\theta)|$

Demostración. — Para probar *1.* es suficiente notar que para todo $z \in V$

$$\langle y \times x, z\rangle = \det_{\mathscr{B}}(y,x,z) = -\det_{\mathscr{B}}(x,y,z) = -\langle x \times y, z\rangle \Rightarrow y \times x = -x \times y$$

de donde se deduce que la aplicación es alternada. La bilinealidad es directa de las propiedades del determinante.
Para la propiedad *2.* notar que $\langle x \times y, x\rangle = \det_{\mathscr{B}}(x,y,x) = 0$ (pues es $\det_{\mathscr{B}}$ es una forma alternada) y $\langle x \times y, y\rangle = \det_{\mathscr{B}}(x,y,y) = 0$, y así x e y son ortogonales a $x \times y$. Si x, y son linealmente independientes entonces es posible completar en una base (x,y,z) de V y entonces $\langle x \times y, z\rangle = \det_{\mathscr{B}}(x,y,z) \neq 0 \Rightarrow x \times y \neq 0$. Además

$$\|x \times y\|^2 = \langle x \times y, x \times y\rangle = \det_{\mathscr{B}}(x,y,x \times y)$$

y entonces si x, y son colineales entonces $\|x \times y\|^2 = 0 \Rightarrow x \times y = 0$. De lo anterior también se deduce que si x, y son linealmente independientes $\det_{\mathscr{B}}(x,y,x \times y) > 0$ entonces $(x,y,x \times y)$ es una base ortonormal directa.
A continuación probaremos *3.* Para comenzar supondremos que x e y son ortogonales y unitarios (esto es, $\|x\| = \|y\| = 1$ y $\langle x,y\rangle = 0$). Sea $e \in V$ tal que $\mathscr{B} = (x,y,e)$ es una base ortonormal directa. Entonces para todo $z \in V$ por la Proposición 5.9.5 se tiene que

$$z = \langle z,x\rangle x + \langle z,y\rangle y \Rightarrow \det_{\mathscr{B}}(x,y,z) = \langle z,e\rangle \det_{\mathscr{B}}(x,y,e) = \langle e,z\rangle \Rightarrow e = x \times y$$

y además se tiene que $\|e\| = \|x \times y\|$. Además como $\langle x,y\rangle = 0$ entonces $\theta(x,y) = \pi/2 \Rightarrow |\sin(\theta)| = 1$. Por lo tanto $\|x \times y\| = \|x\|\|y\||\sin(\theta)| = 1$. Ahora supongamos que x e y no son unitarios. Entonces $x/\|x\|$ y $y/\|y\|$ son unitarios y

$$\left\|\frac{x}{\|x\|} \times \frac{y}{\|y\|}\right\| = 1 \iff \|x \times y\| = \|x\|\|y\|$$

Ahora, si x e y son unitarios, es claro que x y $x\langle x,y\rangle - y$ son ortogonales pues

$$\langle x, x\langle x,y\rangle - y\rangle = \langle x,x\rangle\langle x,y\rangle - \langle x,y\rangle = 0$$

Entonces utilizando el caso anterior y el hecho que $x \times x = 0$ notamos que

$$\|x \times y\| = \|x \times (x\langle x,y\rangle - y)\| = \|x\|\|x\langle x,y\rangle - y\| = \|x\langle x,y\rangle - y\|$$

además de que

$$\|x\langle x,y\rangle - y\|^2 = \langle x,y\rangle^2 - 2\langle x,y\rangle^2 + \|y\| = 1 - \langle x,y\rangle^2 = 1 - \cos^2(\theta) = \sin^2(\theta)$$

y deducimos que $\|x \times y\| = |\sin(\theta)|$. Con todo lo anterior, para el caso general se tendrá que

$$\left\| \frac{x}{\|x\|} \times \frac{y}{\|y\|} \right\| = |\sin(\theta)| \Rightarrow \|x \times y\| = \|x\| \|y\| |\sin(\theta)|$$

□

Observación 6.2.10. — 1. Sea $V \cong \mathbb{R}^3$ espacio euclideano orientado y $\mathscr{B}$ base ortonormal directa. Si $(x_1, x_2, x_3), (y_1, y_2, y_3) \in \mathbb{R}^3$ son las coordenadas de x e y en la base $\mathscr{B}$ respectivamente, entonces las coordenadas de $x \times y \in V$ son

$$\left(\begin{vmatrix} x_2 & y_2 \\ x_3 & y_3 \end{vmatrix}, - \begin{vmatrix} x_1 & y_1 \\ x_3 & y_3 \end{vmatrix}, \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} \right)$$

dado que son los coeficientes del desarrollo de la última columna de la forma lineal

$$\langle x \times y, z \rangle = \det_{\mathscr{B}}(x, y, z) = \begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix}$$

2. Sea $u = r_L \in SO(3)$ una rotación de eje L. Podemos orientar la recta L al escoger un vector director $e \in L$. Entonces se induce una orientación natural en el plano $\Pi = L^\perp$ de tal forma que $u|_\Pi = r_\theta \in SO(2)$ posee un único ángulo de rotación $\theta \in]-\pi, \pi]$. Sea $y \in V \setminus \{0\}$ vector ortogonal a L y $e' := e/\|e\|$ $y' := y/\|y\|$. Entonces $\mathscr{B}' = (e', y', e' \times y')$ es una base ortonormal directa de V y

$$\mathrm{Mat}_{\mathscr{B}'}(u) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \langle u(y'), y' \rangle & \langle u(e' \times y'), y' \rangle \\ 0 & \langle u(y'), e' \times y' \rangle & \langle u(e' \times y'), e' \times y' \rangle \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{pmatrix}$$

$$\Rightarrow \sin(\theta) = \langle u(y'), e' \times y' \rangle = \langle e' \times y', u(y') \rangle = \det_{\mathscr{B}}(e', y', u(y')) = \frac{1}{\|y\|^2 \|e\|} \det_{\mathscr{B}}(e, y, u(y))$$

Notemos ahora que si $x \in V \setminus \{0\}$ es no colineal a e entonces $y := x - \langle x, e' \rangle e'$ es no nulo y ortogonal a e. Entonces $\sin(\theta)$ posee el mismo signo que

$$\begin{aligned}\det_{\mathscr{B}}(e,y,u(y)) &= \det_{\mathscr{B}}(e,x-\langle x,e'\rangle e',u(x-\langle x,e'\rangle e'))\\ &= \det_{\mathscr{B}}(e,x,u(x)-\langle x,e'\rangle \underbrace{u(e')}_{e'})\\ &= \det_{\mathscr{B}}(e,x,u(x))\end{aligned}$$

De todo lo anterior podemos concluir que si $u = r_{L,\theta} \in SO(3)$ es una rotación de eje L, orientado por un vector director $e \in V$, y ángulo $\theta \in]-\pi,\pi]$, entonces

1. $\mathrm{tr}(u) = 2\cos(\theta) + 1$
2. $\sin(\theta)$ posee el mismo signo que $\det_{\mathscr{B}}(e,x,u(x))$ donde $x \notin L$ y $\mathscr{B}$ es cualquier base directa.

Ejemplo 6.2.11. — Veamos a continuación dos ejemplos de lo anterior.

1. Sea
$$A = \begin{pmatrix} 0 & 0 & 1\\ 1 & 0 & 0\\ 0 & 1 & 0\end{pmatrix} \in M_3(\mathbb{R})$$
Es sencillo notar que A es ortogonal pues $A\,{}^{\mathrm{t}}A = I_3$ y $\det(A) = 1$, esto es, $A \in SO(3)$ y es la matriz de una rotación $r_{L,\theta}$ cuyo eje L está orientado por un vector propio asociado a $\lambda = 1$. Es directo calcular que $L = V_1 = \ker(A - I_3) = \mathrm{Vect}_k((1,1,1))$. Entonces es posible calcular θ notando que
   - $\mathrm{tr}(A) = 0 = 2\cos(\theta)+1 \Rightarrow \cos(\theta) = -1/2$ y $\theta = 2\pi/3$ o $\theta = -2\pi/3$.
   - El vector $x = (1,0,0)$ no se encuentra sobre L y $r_{L,\theta}(x) = (0,1,0)$. Luego
   $$\det_{\mathscr{B}}(e,x,r_{L,\theta}(x)) = \begin{vmatrix} 1 & 1 & 0\\ 1 & 0 & 1\\ 1 & 0 & 0\end{vmatrix} > 0$$
   y se deduce que $\sin(\theta) > 0 \Rightarrow \theta = 2\pi/3$.
   - Sea $V = \mathbb{R}^3$ y r la rotación de eje generado y orientado por el vector $e_1 = (36,-48,25)$ y ángulo $\theta = \pi/2$. Si $L = \mathrm{Vect}_{\mathbb{R}}(e_1)$ entonces el plano $\Pi = L^{\perp}$ viene dado por la ecuación $36x - 48y + 25z = 0$. En particular $e_2 = (4,3,0) \in \Pi$. Luego si $e_3 := e_1 \times e_2 = (-75,100,300)$

entonces (e_1, e_2, e_3) es una base directa. Normalizamos

$$e'_1 = \frac{e_1}{\|e_1\|\|} = \left(\frac{-36}{65}, \frac{-48}{65}, \frac{25}{65}\right) \quad e'_2 = \frac{e_2}{\|e_2\|\|} = \left(\frac{4}{5}, \frac{-3}{5}, 0\right) \quad e'_3 = \frac{e_3}{\|e_3\|\|} = \left(\frac{-3}{13}, \frac{4}{13}, \frac{12}{13}\right)$$

para obtener la base ortonormal directa $\mathscr{B} = (e'_1, e'_2, e'_3)$. Notamos a continuación que en dicha base

$$\mathrm{Mat}_{\mathscr{B}}(r) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$$

Además se tiene que la matriz cambio de base a la base canónica es

$$P = \frac{1}{65} \begin{pmatrix} 36 & 52 & -15 \\ -48 & 39 & 20 \\ 25 & 0 & 60 \end{pmatrix}$$

y entonces

$$PAP^{-1} = PA\,{}^{\mathrm{t}}P = \frac{1}{4225} \begin{pmatrix} 1296 & -3353 & -2220 \\ -103 & 2309 & -3540 \\ 4020 & 1140 & 625 \end{pmatrix}$$

es la matriz de r en la base canónica de $\mathbb{R}^3$. Además notar que la traza verifica que $\mathrm{tr}(r) = (1296 + 2304 + 625)/4225 = 1 = 1 + 2\cos(\pi/2)$.

2. Consideremos la rotación r de $V = \mathbb{R}^3$ de eje generado y **orientado** por el vector $e_1 = (36, -48, 25)$ y de ángulo $\theta = \pi/2$. Si $L = \mathrm{Vect}_{\mathbb{R}}(e_1)$, el plano $\Pi = L^{\perp}$ está dado por la ecuación $36x - 48y + 25z = 0$. En particular, el vector $e_2 = (4, 3, 0) \in \Pi$. Sea $e_3 = e_1 \times e_2 = (-75, 100, 300)$ producto cruz. Por construcción, (e_1, e_2, e_3) es una base directa.
Al normalizar, obtenemos:

$$e'_1 = \frac{e_1}{\|e_1\|} = \left(\frac{36}{65}, \frac{-48}{65}, \frac{25}{65}\right); \quad e'_2 = \frac{e_2}{\|e_2\|} = \left(\frac{4}{5}, \frac{3}{5}, 0\right); \quad e'_3 = \frac{e_3}{\|e_3\|} = \left(\frac{-3}{13}, \frac{4}{13}, \frac{12}{13}\right),$$

y luego $\mathscr{B} = (e_1, e_2, e_3)$ es una base ortonormal directa. En esta base se tiene que:

$$A = \mathrm{Mat}_{\mathscr{B}}(r) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}.$$

La matriz cambio de base a la base canónica es

$$P = \begin{pmatrix} \frac{36}{65} & \frac{4}{5} & \frac{-3}{13} \\ \frac{-48}{65} & \frac{3}{5} & \frac{4}{13} \\ \frac{25}{65} & 0 & \frac{12}{13} \end{pmatrix} = \frac{1}{65} \begin{pmatrix} 36 & 52 & -15 \\ -48 & 39 & 20 \\ 25 & 0 & 60 \end{pmatrix}$$

$$\Longrightarrow PAP^{-1} = PAP^t = \frac{1}{4225} \begin{pmatrix} 1296 & -3353 & -2220 \\ -103 & 2309 & -3540 \\ 4020 & 1140 & 625 \end{pmatrix}$$

matriz de r en la base canónica de $\mathbb{R}^3$. Notar que la traza $\operatorname{tr}(A) = \frac{1}{4225}(1296{+}2304{+}625) = 1 = 1+$

6.3. Isometrías en dimensión arbitraria y Teorema de Cartan-Dieudonné

En las secciones previas se demostró que toda isometría de $V \cong \mathbb{R}^2$ es el producto de a lo más 2 reflexiones, y de forma análoga que toda isometría de $V \cong \mathbb{R}^3$ es producto de a lo más 3 reflexiones. Como el lector podrá imaginarse, este resultado puede ser generalizado a espacios de dimensión finita arbitraria. Establecer este resultado, demostrado por Élie Cartan (1869-1951) y Jean Dieudonné (1906-1992), será nuestro objetivo durante las próximas páginas.

Lema 6.3.1. — *Sea $V \cong \mathbb{R}^n$ espacio euclideano y $x, y \in V \setminus \{0\}$ tales que $x \neq y$ y $\|x\| = \|y\|$. Sea $H := (x-y)^{\perp}$ el hiperplano ortogonal a la recta generada por el vector $x-y$. Entonces $s_H(x) = y$.*

Demostración. — Nos interesa encontrar vectores $a \in H$ y $b \in H^{\perp} = \operatorname{Vect}_{\mathbb{R}}(x-y)$ de tal forma que $x = a + b$. Para ello notar que

$$\langle x+y, x-y\rangle = \|x\|^2 - \|y\|^2 = 0 \Rightarrow a := \frac{x+y}{2} \in H \wedge b := x - a = \frac{x-y}{2}$$

Luego es suficiente observar que

$$s_H(x) = \left(\frac{x+y}{2}\right) - \left(\frac{x-y}{2}\right) = y$$

□

***Teorema 6.3.2* (Cartán-Dieudonné).** — *Toda isometría de un espacio euclideano V de $\dim_{\mathbb{R}}(V) = n \geq 2$ es producto de a lo más n reflexiones.*

Demostración. — Utilizaremos inducción sobre $\dim_{\mathbb{R}}(V) = n$. Si $n = 2$ entonces este caso ya ha sido probado. Sea $u \in O(n)$ isometría de $V \cong \mathbb{R}^n$. Si $u = Id_V$ entonces $u = s \circ s$ para toda reflexión $s \in O(n)$. Suponer entonces que $u \neq Id_V$, por lo que existirá $x \in V$ vector no nulo de tal forma que $u(x) \neq x$. Como $u \in O(n)$ se tendrá que $\|x\| = \|u(x)\|$ y por el lema 6.3.1 existe una reflexión $s \in O(n)$ tal que $s(x) = u(x) \Rightarrow s^2(x) = x = (s \circ u)(x)$. Por ende vemos que la recta $L := \mathrm{Vect}_{\mathbb{R}}(x)$ y su hiperplano ortogonal $H := L^{\perp}$ son subespacios invariantes de la isometría $s \circ u \in O(n)$.
Por hipótesis de inducción se tendrá que $(s \circ u)|_H$ posee una descomposición en a lo más $n - 1$ reflexiones, esto es,

$$(s \circ u)|_H = s_{H_1'} \circ \cdots s_{H_m'}, \qquad m \leq n - 1$$

donde $H_1', \ldots, H_m'$ son hiperplanos en H ($\dim_{\mathbb{R}} H_j' = n - 2$ $\forall 1 \leq j \leq m$). Para cada $j \in \{1, \ldots, m\}$ definimos $H_j := \mathrm{Vect}_{\mathbb{R}}(H_j', L)$ el hiperplano en V generado por H_j' y la recta L. Entonces es posible notar que $s_{H_1} \circ \cdots \circ s_{H_m}|_L = Id_L$ y que además dicha isometría conincide claramente con $s \circ u$ en H. Debido a que $V = H \circ L$ entonces

$$s \circ u = s_{H_1} \circ \cdots \circ s_{H_m}$$

y finalmente

$$u = s^{-1} \circ s_{H_1} \circ \cdots s_{H_m} = s \circ s_{H_1} \circ \cdots \circ s_{H_m}$$

□

***Definición 6.3.3* (función isométrica).** — Sea $V \cong \mathbb{R}^n$ espacio euclideano. Se dice que $\varphi : V \to V$ es una **función isométrica** si $\forall x, y \in V$

$$d(\varphi(x), \varphi(y)) = d(x, y) \overset{def}{\iff} \|\varphi(x) - \varphi(y)\| = \|x - y\|$$

Por definición, una función isométrica es aquella que preserva la norma inducida por el producto escalar en un espacio euclideano, lo cual intuitivamente significa que dicha función mantiene intactas las distancias entre los distintos elementos del espacio (de ahí que 'preserve la métrica').
Por lo estudiado hasta ahora, claramente toda isometría $u \in O(n)$ es a su vez una función isométrica. Sin embargo, no toda isometría es lineal. Esto es claro notando que si $a \in V$ es un vector fijo, la **traslación en a**, definida como

$$t_a : V \to V, \qquad x \mapsto t_a(x) = x + a$$

es claramente una función isométrica la cual no es lineal para $a \neq 0$.

Teorema 6.3.4. — *Sea $V \cong \mathbb{R}^n$ espacio euclideano y $\varphi : V \to V$ función isométrica. Entonces existe un único $a \in V$ y una única isometría $uinO(n)$ de tal forma que*

$$\varphi = t_a \circ u \iff \varphi(x) = u(x) + a \quad \forall x \in V$$

Demostración. — Considerar $a := \varphi(0) \in V$. Note que la función $\psi := t_{-a} \circ \varphi : V \to V$ verifica que $\psi(0) = 0$ y $d(\psi(x), psi(0)) = d(x, 0) \iff \|\varphi(x)\| = \|x\|$ para todo $x \in V$. Consideremos $(e_1, \ldots, e_n)$ base ortonormal de V. Entonces para $i \neq j$

$$\|\psi(e_i) - \psi(e_j)\| = d(\psi(e_i), \psi(e_j)) = d(e_i, e_j) = \|e_i - e_j\| = \sqrt{2}$$

Además

$$\underbrace{\|\psi(e_i) - \psi(e_j)\|^2}_{=2} = \underbrace{\|\psi(e_i)\|^2}_{=1} + \underbrace{\|\psi(e_j)\|^2}_{=1} - 2\langle\psi(e_i), \psi(e_j)\rangle \Rightarrow \langle\psi(e_i), \psi(e_j)\rangle = 0$$

y entonces $(\psi(e_1), \ldots \psi(e_n))$ es una base ortonormal de V. Sea $x \in V$ arbitrario. Luego podemos escribir x y $\psi(x)$ como

$$x = \sum_{j=1}^{n} x_j e_j, \quad \psi(x) = \sum_{j=1}^{n} y_j \psi(e_j)$$

Entonces

$$d(x, e_k)^2 = \|x - e_k\|^2 = \|x\|^2 - 2\langle x, e_k\rangle + \|e_k\|^2 = \|x\|^2 x_k + 1$$

y además

$$d(\psi(x), \psi(e_k))^2 = \|\psi(x) - \psi(e_k)\|^2 = \|\psi(x)\|^2 - 2y_k + 1 = \|x\|^2 - 2y_k + 1$$

Dado que $d(\psi(x), \psi(e_k)) = d(x, e_k)$ se obtiene que $x_k = y_k$ $\forall k$ y así

$$\psi\left(\sum_{j=1}^{n} x_j e_j\right) = \sum_{j=1} x_j \psi(e_j)$$

de donde deducimos que ψ es lineal. Además, vimos que ψ preserva la norma, lo que permite afirmar que $\psi \in O(n)$. De esta forma $\varphi = t_a \circ u$.
Para concluir veamos la unicidad. Sean $a, b \in V$ y $u, v \in O(n)$ tales que $\varphi = t_a \circ u = t_b \circ v$. Entonces

$$\varphi(0) = \underbrace{u(0)}_{=0} + a = \underbrace{v(0)}_{=0} + b \Rightarrow a = b \wedge u = t_{-a} \circ \varphi = v$$

$\square$

6.4. Reducción simultánea de formas cuadráticas e hipersuperficies cuádricas

Consideremos $V \cong \mathbb{R}^n$ espacio euclideano. Gracias al Teorema de Sylvester ya estudiado, es posible afirmar la existencia de una base ortogonal de tal manera que la forma cuadrática $\|x\|^2$ definida por la norma se reduce a la expresión $\|x\|^2 = x_1^2 + \cdots + x_n^2$. De forma general, toda forma cuadrática $Q : V \to \mathbb{R}$ posee una escritura de la forma

$$Q(x) = \lambda_1 x_1^2 + \cdots + \lambda_p x_p^2 - \lambda_{p+1} x_{p+1}^2 - \cdots - \lambda_{p+q} x_{p+q}^2$$

donde (p, q) es la **signatura** de Q y $(x_1, \ldots, x_n)$ son las coordenadas de x en una base ortogonal. El siguiente teorema nos permitirá hacer la reducción simultánea de las formas cuadráticas Q y $\|\cdot\|^2$.

Teorema 6.4.1. — *Sea $V \cong \mathbb{R}^n$ espacio euclideano y $Q : V \to \mathbb{R}$ forma cuadrática. Entonces existe una base ortonormal de V la cual es ortogonal respecto a Q.*

Demostración. — Por simplicidad denotaremos durante esta demostración $E := \langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ el producto escalar de V. y sea $B : V \times V \to \mathbb{R}$ la forma bilineal simétrica asociada a Q. EL objetivo será construir un endomorfismo $u : V \to V$ tal que $B(x, y) = \langle x, u(y) \rangle \ \forall x, y \in V$. Utilizando la notación $B(x, y) = \widehat{B}(y)(x)$ y $\langle x, u(y) \rangle = E(x, u(y)) = \widehat{E}(u(y))(x)$ entonces $\widehat{B} = \widehat{E} \circ u$. Como E es no degenerado $\widehat{E} : V \xrightarrow{\sim} V^*$ es un isomorfismo y por tanto es posible definir $u := \widehat{E}^{-1} \circ \widehat{B} : V \to V$. Luego

$$\langle x, u(y) \rangle = B(x, y) = B(y, x) = \langle y, u(x) \rangle$$

de donde vemos que $u = u^*$, es decir, u es simétrico. Por ende, existe una base ortonormal $\mathscr{B} = (e_1, \ldots, e_n)$ en la cual u es diagonalizable, esto es, $u(e_j) = \lambda_j e_j$ para ciertos $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$.

$$B(e_i, e_j) = \langle e_i, u(e_j) \rangle = \lambda_j \langle e_i, e_j \rangle = \langle \delta_{ij}$$

de donde se deduce que $\mathscr{B}$ es ortogonal respecto a Q. $\square$

A continuación veremos una de las aplicaciones del resultado anterior. En concreto, usaremos esto para estudiar las llamadas **hipersuperficies cuádricas** o simplemente **cuádricas**. Comenzamos con la siguiente definición.

Definición 6.4.2. — Sea $V \cong \mathbb{R}^n$ espacio euclideano, $Q : V \to \mathbb{R}$ forma cuadrática, $f : V \to \mathbb{R}$ forma lineal y $c \in \mathbb{R}$ una constante. Decimos que el conjunto definido por

$$X = \{x \in V : Q(x) + f(x) + c = 0\} \subseteq V$$

es una **cuádrica** en V. Además, diremos que la cuádrica $X \subseteq V$ es **hiper-superficie cuádrica!suave** si la forma cuadrática $\widetilde{Q} : V \times \mathbb{R} \to \mathbb{R}$ definida por

$$\widetilde{Q}(x,t) = Q(x) + f(x)t + ct^2$$

en $V \times \mathbb{R}^n \cong \mathbb{R}^{n+1}$ es no degenerada, de lo contrario diremos que $X \subseteq V$ es **singular**.

Observación 6.4.3. — Sea $p \in V$ y $t_{-p} : V \to V$, $x \mapsto x - p$ la traslación en $-p$. Entonces si $X \subseteq V$ es una cuádrica el conjunto $t_{-p}(X)$ corresponde al conjunto de $x \in V$ tales que

$$\begin{aligned} 0 =& Q(x+p) + f(x+p) + c = Q(x) + 2B(x,p) + f(x) + Q(p) + f(p) + c \\ =& Q(x) + (2\widehat{B}(p) + f)(x) + (Q(p) + f(p) + c) \end{aligned}$$

donde $B : V \times V \to \mathbb{R}$ es la forma bilineal simétrica asociada a Q. De esta forma $t_{-p}(X)$ es una cuádrica asociada a la forma cuadrática Q, a la forma lineal $(2\widehat{B}(p) + f) \in V^*$ y a la constante $c' := Q(p) + f(p) + c$.
Ahora, si Q es no degenerada entonces $\widehat{B} : V \xrightarrow{\sim} V^*$ es un isomorfismo y es posible escoger $p := -\widehat{B}^{-1}(f/2)$. Luego $t_{-p}(X)$ adquiere la forma $Q(x) + c' = 0$ y $t_{-p}(X)$ resulta ser simétrica respecto del origen (puesto que $x \in t_{-p}(X) \iff -x \in t_{-p}(X)$).

Definición 6.4.4. — Sea $V \cong \mathbb{R}^n$ espacio euclideano y $X = \{x \in V : Q(x)+f(x)+c = 0\} \subseteq V$ cuádrica. Entonces la cuádrica X se dice **centrada** si la forma cuadrática asociada $Q : V \to \mathbb{R}$ es no degenerada. Más aún, X se dice **centrada en el origen** si $f = 0$.

Gracias a la observación anterior, es claro que toda cuádrica centrada puede ser trasladada a una cuádrica centrada en el origen. El punto $p = -\widehat{B}^{-1}(f/2)$ se dirá el **centro** de la cuádrica centrada $X = \{x \in V : Q(x)+f(x)+c = 0\} \subseteq V$.

Proposición 6.4.5. — *Sea $V \cong \mathbb{R}^n$ espacio euclideano y $X = \{x \in V : Q(x)+c = 0\}$ una cuádrica centrada en el origen. Entonces existe una base $\mathscr{B} = (e_1, \ldots, e_n)$ de tal forma que*

$$x = x_1e_1 + \ldots + x_ne_n \in X \iff \lambda_1 x_1^2 + \ldots \lambda_n x_n^2 + c = 0$$

donde $\lambda_1, \ldots, \lambda_n \in \mathbb{R}\backslash\{0\}$ son no nulos. En particular X es suave $\iff c \neq 0$.

Demostración. — Del Teorema 6.4.1 existe $\mathscr{B} = (e_1, \dots, e_n)$ base ortonormal de V la cual es ortogonal respecto a Q. Entonces la matriz $A_{\mathscr{B}}$ de Q en dicha base es

$$A_{\mathscr{B}} = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}$$

donde $\det(A_{\mathscr{B}}) = \lambda_1 \cdots \lambda_n \neq 0$ pues Q es no degenerada. Entonces si $(x_1, \dots, x_n) \in \mathbb{R}^n$ son las coordenadas del vector $x \in V$ respecto a la base $\mathscr{B}$ entonces $Q(x) = \sum_{j=1}^n \lambda_j x_j^2$ obteniendo el resultado deseado. Además, si consideramos $\widetilde{Q}(x,t) = Q(x) + ct^2$ su matriz corresponde a

$$\begin{pmatrix} \lambda_1 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & \lambda_n & 0 \\ 0 & \cdots & 0 & c \end{pmatrix}$$

por lo tanto, X es suave $\iff c \neq 0$. $\square$

La ecuación de la proposición anterior se conoce como "ecuación reducida" de la cuádrica X. Utilizaremos la siguiente terminología:

1. Un eje de la cuádrica $X \subseteq V$ es cualquier recta $L \subseteq V$ generada por un vector propio asociado a uno de los valores propios $\lambda_1, \dots, \lambda_n \in \mathbb{R}$. Si $p \in X$ se encuentra sobre un eje, es decir, $p \in X \cap L$ para algún eje L de X, entonces p es un **vértice** de X.
2. Una cuádrica en el plano $V \cong \mathbb{R}^2$ es denominada **cónica**.
3. Una cuadríca en el espacio $V \cong \mathbb{R}^3$ es denominada **superficie cuádrica**

Ejemplo 6.4.6. — A continuación estudiaremos tres ejemplos en donde se aplicarán los conceptos y resultados previos.

1. La cónica en $V = \mathbb{R}^2$ dada por la ecuación $y = x^2 + 1$ no corresponde a una cónica centrada pues la forma cuadrática $Q(x,y) = x^2$ es degenerada. No obstante, la forma cuadrática $\widetilde{Q}(x,y,t) = x^2 - yt + t^2$ es no degenerada y luego dicha cónica es suave.
2. Considerar la cónica $C \subseteq \mathbb{R}^2$ definida por la ecuación $3x^2 + 2xy + 3y^2 + 4x - 4y = 0$. La matriz de $Q(x,y) = 3x^2 + 2xy + 3y^2$ corresponde a

$$A = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} \Rightarrow \det(A) = 8 \neq 0$$

siendo entonces Q no degenerada y C centrada. Considerar su centro $p = (a, b) \in \mathbb{R}^2$ y $f(x, y) = 4x - 4y$. Hallaremos el centro de la cónica imponiendo la anulación del término lineal

$$\begin{aligned}&3(x+a)^2 + 2(x+a)(y+b) + 3(y+b)^2 + 4(x+a) - 4(y+b) = 0\\ &\qquad\iff 6ax + 2ay + 2bx + 6by + 4x - 4y = 0 \quad \forall x, y \in \mathbb{R}\end{aligned}$$

y obtenemos un sistema lineal cuya solución corresponde a $p = (-1, 1)$. Analizemos ahora la cónica $C' = t_{-p}(C)$ trasladada en p. Entonces

$$\begin{aligned}C' =&\{(x,y) \in \mathbb{R}^2 : 3x^2 + 2xy + 3y^2 + \overbrace{Q(-1,1)}^{=4} + \overbrace{f(-1,1)}^{=-8} = 0\}\\ =&\{(x,y) \in \mathbb{R}^2 : 3x^2 + 2xy + 3y^2 = 4\}\end{aligned}$$

Calculamos el polinomio característico

$$P_A(x) = \begin{vmatrix} x-3 & -1 \\ -1 & x-3 \end{vmatrix} = (x-3)^2 - 1 = x^2 - 6x + 8 = (x-2)(x-4)$$

y notamos que A tiene valores propios $\lambda_1 = 2, \lambda_2 = 4$. Encontramos un vector propio asociado a λ_1

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x+y \\ x+y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow x = -y$$

y entonces

$$e_1 := \left(\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}\right)$$

es un vector propio unitario. Como $V_4 = V_2^{\perp}$ es claro que

$$e_2 := \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$$

es un vector propio unitario asociado a λ_2. Entonces, en la base ortonormal $\mathscr{B} = (e_1, e_2)$ obtenemos la ecuación reducida

$$2X^2 + 4Y^2 = 4$$

la cual corresponde a una elipse. Notar que los ejes de C' corresponden a las rectas generadas por e_1 y e_2. Además sus vértices son $\pm\sqrt{2}e_1$ y $\pm e_2$.

3. Estudiaremos ahora la superficie cuádrica $S \subseteq \mathbb{R}^3$ determinada por $x^2 + 3y^2 + z^2 + 4xz = 3$. La matriz de la forma cuadrática asociada

$Q(x, y, z) = x^2 + 3y^2 + z^2 + 4xz$ corresponde a

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

Dicha matriz posee valores propios $\lambda_1 = -1$ y $\lambda_2 = 3$ con multiplicidad 2. Luego los vectores

$$v_1 = \begin{pmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{pmatrix} \in V_{\lambda_1}, \quad v_2 = \begin{pmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \in V_{\lambda_2} \end{pmatrix}, \quad v_3 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \in V_{\lambda_2}$$

forman una base ortonormal $\mathscr{B}$ de $\mathbb{R}^3$ en la cual la ecuación reducida de S corresponde a

$$-X^2 + 3Y^2 + 3Z^2 = 3$$

y por lo tanto S es un "hiperboloide de una hoja". Además los ejes de S son las rectas generadas por v_1 y todas las rectas contenidas en $V_{\lambda_2} = V_{\lambda_1}^{\perp}$.

Definición 6.4.7. — Sea $V \cong \mathbb{R}^n$ un espacio vectorial real y $U \subseteq V$ subespacio. Sea $p \in V$. Definimos el **subespacio afín** de V **dirigido por** U **pasando poe** p como el conjunto

$$p + U = \{x \in V : x - p \in U\} = \{p + u : u \in U\}$$

Utilizando las ideas estudiadas sobre espacios vectoriales cocientes, es claro que el conjunto $p + U$ corresponde a la preimagen de $[p]$ en el espacio vectorial cociente V/U. De forma más explícita, si $\mathscr{B} = (e_1, \ldots, e_r)$ es una base de U entonces

$$p + U = \{p + \lambda_1 e_1 + \ldots + \lambda_r e_r : \lambda_1, \ldots, \lambda_r \in \mathbb{R}\}$$

Ejemplo 6.4.8. — A continuación se presentan dos ejemplos concretos de la definición anterior.

1. Sea $\mathscr{B} = (e_1, \ldots, e_n)$ base de $V \cong \mathbb{R}^n$ y considerar el hiperplano definido por

$$H = \left\{ x = \sum_{j=1}^{n} \in V : a_1 x_1 + \ldots + a_n x_n \right\}$$

Si $(p_1, \ldots, p_n)$ es el vector de coordendas de un punto $p \in V$, entonces el **hiperplano afín** dirigido por H pasando por p es el conjunto

$$\{x \in V : x - p \in H\} = \left\{ x = \sum_{j=1}^{n} x_j e_j \in V : a_1(x_1 - p_1) + \ldots + a_n(x_n - p_n) \right\}$$

2. Sea $V \cong \mathbb{R}^n$ espacio euclideano y $X = \{x \in V : Q(x) + f(x) + c = 0\}$ una cuádrica. Sea $L = p + \mathrm{Vect}_{\mathbb{R}}(e) = \{p + te : t \in \mathbb{R}\}$ la recta afín dirigida por e pasando por p. Entonces los puntos del conjunto $X \cap L$ verifican

$$0 = Q(p+te) + f(p+te) + c = Q(e)t^2 + (2B(e,p) + f(e))t + (Q(p) + f(p) + c)$$

Se obtiene entonces una ecuación de grado 2 en la variable t cuyo discriminante corresponde a

$$\Delta = (2B(e,p) + f(e))^2 - 4Q(e)(Q(p) + f(p) + c)$$

Por ende existen dos alternativas, $L \subseteq X$ o bien $X \cap L$ consiste en a lo más dos puntos. En caso que $Q(e) \neq 0$ pueden suceder
- $\Delta < 0$ en cuyo caso $X \cap L = \emptyset$
- $\Delta > 0$ y se tendrá que $X \cap L = \{p_1, p_2\}$
- $\Delta = 0$ y en tal caso $X \cap L = \{p\}$

Si ocurre que $\Delta = 0$ entonces se dirá que L es **tangente** a la cuádrica X, incluso si $Q(e) = 0$.

Definición 6.4.9 **(espacio tangente).** — Sea X cuádrica en el espacio euclideano $V \cong \mathbb{R}^n$ y $p \in X$. Definimos el **espacio tangente de X en p** por

$$T_pX := \{L \subseteq V : L \text{ es una recta afín pasando por } p \text{ tal que } L \text{ es tangente a } X\}$$

Podemos distinguir dos casos

1. $T_pX = V$ y en dicho caso X es una cuádrica singular.
2. T_pX es un hiperplano afín.

Más aún, se tendrá que X es suave $\iff$ T_pX es un hierplano afín $\forall p \in X$.

Demostración. — Dado que $p \in X$ la condición de tangencia $\Delta = 0$ se reduce a $2B(e,p) + f(e) = 0$, o equivalentemente $(2\widehat{B}(p) + f)(e) = 0$. Entonces la unión de las rectas tangentes a X en el punto p corresponde al subespacio afín dirigido por $H := \ker(2\widehat{B}(p) + f)$. El conjunto H corresponderá a un hiperplano excepto si $2\widehat{B}(p) + f = 0$, en cuyo caso $T_pX = V$. Consideremos la forma cuadrática $\widetilde{Q}(x,t) = Q(x) + f(x)t + ct^2$. Empleando la formula de polarización para encontrar su forma bilineal simétrica asociada

$$\widetilde{B}((x,t),(y,s)) = B(x,y) + \frac{1}{2}f(x)s + \frac{1}{2}f(y)t + cts$$

y además

$$\widetilde{B}((p,1),(y,0)) = B(p,y) + \frac{1}{2}f(y) = \frac{1}{2}(2\widehat{B}(p) + f)(y)$$

De lo anterior deducimos que la condición $2\widehat{B}(p)+f = 0$ significa que el vector $(p,1) \in V \times \mathbb{R}$ es ortogonal a $V \times \{0\}$ con respecto a la forma cuadrática $\widetilde{Q}$. Además es claro que

$$\widetilde{Q}(p,1) = Q(p) + f(p) + c = 0$$

pues $p \in X$. $\square$

Ejemplo 6.4.10. — Si x es una cuádrica centrada en el origen (esto es, Q es no degenerada y $f = 0$) entonces su espacio tangente T_pX está dirigido por $H = \ker(\widehat{B}(p)) = \{x \in V : B(x,p) = 0\} =: p^{\perp Q}$ ortogonal de p.

Ejemplo 6.4.11. — Sea $C \subseteq \mathbb{R}^2$ la elipse de ecuación $x^2/a^2 + y^2/b^2{=}1$. Entonces tenemos que

$$Q(x,y) = \frac{x^2}{a^2} + \frac{y^2}{b^2} \Rightarrow B((x_1,y_1),(x_2,y_2)) = \frac{x_1x_2}{a^2} + \frac{y_1y_2}{b^2}$$

Entonces si $p = (x_0,y_0) \in C$, como C está centrada en el origen el espacio tangente a C por el punto p está dirigido por la recta

$$L = p^{\perp Q} = \left\{(x,y) \in \mathbb{R}^2 : \frac{xx_0}{a^2} + \frac{yy_0}{b^2} = 0\right\}$$

y entonces

$$T_pC = \left\{(x,y) \in \mathbb{R}^2 : \frac{(x-x_0)y_0}{a^2} + \frac{(y-y_0)y_0}{b^2} = 0\right\} = \left\{(x,y) \in \mathbb{R}^2 : \frac{xx_0}{a^2} + \frac{yy_0}{b^2} = 1\right\}$$

Con las herramientas que hemos desarrollado, podemos hacer una clasificación bastante sencilla de superficies cuádricas en base a la signatura de la forma cuadrática asociada. Si $S \subseteq \mathbb{R}^3$ es una superficie cuádrica suave y centrada en el origen, su ecuación puede ser reducida, empleando una base conveniente, a uno de los siguientes casos

1. Si posee signatura (3,0) entonces S corresponde a un **elipsoide**, cuya ecuación es
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

2. Si su signatura es (2,1), S es un **hiperboloide de una hoja** de ecuación
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$$
3. Si tiene signatura $(1,2)$ entonces S es un **hiperboloide de dos hojas**
$$\frac{x^2}{a^2} - \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$$

Teorema 6.4.12. — *Sea $V \cong \mathbb{R}^n$ espacio euclideano y $S \subseteq V$ hiperboloide de una hoja. Entonces S es una* **superficie reglada***: para cada $p \in S$ existen dos rectas afines contenidas en S, definidas por la intersección de S y su plano tangente T_pS.*

Demostración. — Dado que S es suave por la Proposición 6.4.9 T_pS es un plano afín, que además está dirigido por $p^{\perp Q}$, pues S es centrada en el origen (ejemplo 6.4.10). Consideramos la forma cuadrática de su ecuación reducida
$$Q(x,y,z) = \frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2}$$
Considerar $p' \in p^{\perp Q}$ y $p + p' \in T_pS$. Luego $p + p' \in S$ si y sólo si
$$0 = Q(p+p') - 1 = \underbrace{Q(p)}_{=1} + 2\underbrace{B(p,p')}_{=0} + Q(p') - 1 = Q(p')$$
y entonces la intersección $T_pS \cap S$ corresponde a la traslación por el vector p del conjunto de vectores isótropos respecto a la restricción $Q' = Q|_{p^{\perp Q}}$. Debido a que $Q(p) = 1 > 0$ y Q tiene signatura $(2,1)$, entonces la forma cuadrática Q' posee signatura $(1,1)$. Finalmente, por el Teorema de Sylvester existe una base de $p^{\perp Q} \cong \mathbb{R}^2$ de tal forma que $Q'(u,v) = u^2 - v^2 = (u+v)(u-v)$. Por lo tanto, los vectores isótropos respecto a Q corresponden a la unión de las rectas $u = v$ y $u = -v$. □

Ejercicio 6.4.13. — Sea $V \cong \mathbb{R}^3$ un espacio euclideano y $S \subseteq V$ un elipsoide o un hiperboloide de dos hojas. Demostrar que para todo $p \in S$ $T_pS \cap S = \{p\}$. [Indicación: Analizar la signatura de Q'].

Ejercicio 6.4.14. — Considere la superficie cuádrica $S \subseteq \mathbb{R}^3$ definida por
$$xy + xz + yz + 2y + 1 = 0$$
Determinar una base ortonormal de $\mathbb{R}^3$ en la cual S adquiera una ecuación reducida. Determinar su ecuación y su naturaleza geométrica.

CAPÍTULO 7

FORMAS HERMITIANAS Y PRODUCTO ESCALAR COMPLEJO

En el capítulo 5, estudiamos en detalle los espacios euclideanos, los cuales no son más que un espacio vectorial real $V \cong \mathbb{R}^n$ junto a una forma bilineal simétrica no degenerada, que conocemos como producto escalar o producto interno. Durante este capítulo nos dedicaremos a extender esta noción de espacio euclideano a los espacios vectoriales ahora sobre el cuerpo $\mathbb{C}$ de los números complejos. En definitiva, trabajaremos sobre un espacio vectorial complejo $V \cong \mathbb{C}^n$ al cual dotaremos de un producto escalar.

7.1. Formas hermitianas

Para motivar nuestro estudio, tendremos como punto de partida el siguiente ejemplo, que resultará ser fundamental.

Ejemplo 7.1.1. — En $\mathbb{R}^2$, la norma euclideana $\|(x,y)\| = \sqrt{x^2+y^2}$ está definida a partir (de la raíz cuadrada) del producto escalar canónico $\langle (x_1,y_1),(x_2,y_2)\rangle = x_1x_2 + y_1y_2$. Si consideramos ahora los números complejos $z_1 = x_1 + iy_1$ y $z_2 = x_2 + iy_2$, nos gustaría construir una función $h : \mathbb{C} \times \mathbb{C} \to \mathbb{C}$ tal que para $z_1 = z_2$ obtengamos $h(z_1,z_1) = x_1^2 + y_1^2$.

Notar que si considerásemos simplemente el producto $z_1z_2 = (x_1+iy_1)(x_2+iy_2) = (x_1x_2-y_1y_2)+i(x_1y$ entonces para $z_1 = z_2$ obtendríamos $z_1^2 = (x_1^2 - y_1^2) + i2x_1y_1$, el cual **no** es necesariamente real y que no podemos usar para calcular el módulo de z_1.

Por otra parte, dado que $|z|^2 = z\overline{z}$, si consideramos cualquiera de las dos funciones

$$h_1(z_1,z_2) := z_1\overline{z_2} \quad \text{o} \quad h_2(z_1,z_2) := \overline{h_1(z_1,z_2)} = \overline{z_1}z_2,$$

entonces obtenemos para $z_1 = z_2$ el resultado esperado.

Es muy importante notar que h_1 y h_2 **no** son formas bilineales complejas, puesto que h_1 (resp. h_2) es $\mathbb{C}$–lineal en la primera variable (resp. segunda variable), pero **no** es lineal en la segunda (resp. en la primera). Por ejemplo, para todos $z_1, z_2, z_3 \in \mathbb{C}$ (pensados como vectores) y todo $\lambda \in \mathbb{C}$ (pensado como escalar) se tiene:

$$h_1(z_1+\lambda z_2, z_3) = h(z_1, z_3)+\lambda h(z_2, z_3) \text{ y } h_1(z_3, z_1+\lambda z_2) = h_1(z_3, z_1)+\overline{\lambda} h_1(z_3, z_2).$$

Luego, h_1 es *anti-lineal* (o *lineal conjugada*) respecto a la segunda variable: diremos que h_1 es **sesquilineal**.

En lo que sigue supondremos que los espacios vectoriales son de dimensión *finita*, a pesar de que la mayoría de las definiciones tienen sentido en dimensión arbitraria.

***Definición 7.1.2* (aplicación anti-lineal). —** Sean V, W dos $\mathbb{C}$-espacios vectoriales. Decimos que una función $\varphi : V \to W$ es una **aplicación anti-lineal** si para todos $v_1, v_2 \in V$ y todo $\lambda \in \mathbb{C}$ se tiene que

$$\varphi(v_1 + v_2) = \varphi(v_1) + \varphi(v_2), \quad \varphi(\lambda v_1) = \overline{\lambda}\varphi(v).$$

***Observación 7.1.3.* —** Es muy importante observar que si V es un espacio vectorial sobre $\mathbb{C}$, entonces podemos definir el **espacio vectorial conjugado** $\overline{V}$ de la manera siguiente: como <u>*grupo abeliano*</u> definimos $\overline{V} := V$ y, si $v \in V$ es un vector y $\lambda \in \mathbb{C}$ un escalar, entonces <u>*definimos*</u> la multiplicación por escalares en $\overline{V}$ mediante

$$\lambda \cdot v := \overline{\lambda} v,$$

donde el lado derecho lo definimos a partir de la estructura de espacio vectorial complejo de V. Luego, una función $\varphi : V \to W$ es anti-lineal si y sólo si la función $\varphi : V \to \overline{W}$ es una aplicación lineal (en el sentido usual).

***Definición 7.1.4* (forma sesquilineal). —** Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo. Decimos que una función $h : V \times V \to \mathbb{C},\ (x, y) \mapsto h(x, y)$ es una **forma sesquilineal** si:

1. h es $\mathbb{C}$-lineal en la *primera* variable, i.e., para todos $x, y, z \in V$ y $\lambda \in \mathbb{C}$ se tiene que
$$h(\lambda x, y) = \lambda h(x, y) \quad \text{y} \quad h(x + y, z) = h(x, z) + h(y, z).$$
2. h es anti-lineal en la *segunda* variable, i.e., para todos $x, y, z \in V$ y $\lambda \in \mathbb{C}$ se tiene que
$$h(x, \lambda y) = \overline{\lambda} h(x, y) \quad \text{y} \quad h(x, y + z) = h(x, y) + h(x, z).$$

Denotamos por $\mathrm{Sesq}(V)$ al espacio vectorial *complejo* de formas sesquilineales en V. Finalmente, dada una forma sesquilineal $h : V \times V \to \mathbb{C}$, definimos la **forma sesquilineal adjunta** $h^* : V \times V \to \mathbb{C}$ mediante la fórmula $h^*(x, y) := \overline{h(y, x)}$.

Ejercicio 7.1.5. — Probar que la forma sesquilineal adjunta $h^* : V \times V \to \mathbb{C}$ es una forma sesquilineal.

Observación 7.1.6. — En vista de la observación anterior, una forma sesquilineal $h : V \times V \to \mathbb{C}$ es lo mismo que una forma $\mathbb{C}$-bilineal $h : V \times \overline{V} \to \mathbb{C}$.

☡ Es importante señalar que, tal como lo vimos en el Ejemplo 7.1.1, podemos considerar también formas anti-lineales en la *primera* variable y $\mathbb{C}$-lineales en la *segunda* variable. Esta última convención es comúnmente usada en Física, puesto que es coherente con la notación *bra-ket* introducida por Dirac.

***Definición 7.1.7* (forma (anti-)hermitiana).** — Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo. Decimos que una forma sesquilineal $h : V \times V \to \mathbb{C}$ es:

1. una **forma hermitiana** si $h^* = h$, i.e., $h(x, y) = \overline{h(y, x)}$ para todos $x, y \in V$.
2. una **forma anti-hermitiana** si $h^* = -h$, i.e., $h(x, y) = -\overline{h(y, x)}$ para todos $x, y \in V$.

Denotaremos por $\mathrm{Herm}(V)$ (resp. $\mathrm{AntiHerm}(V)$) al espacio vectorial <u>real</u> de formas hermitianas (resp. anti-hermitianas) de V.

Proposición 7.1.8. — *Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo. Una forma sesquilineal h sobre V es hermitiana si y sólo si $h(x, x) \in \mathbb{R}$ es real para todo $x \in V$.*

Demostración. — Si h es hermitiana entonces $h(x, y) = \overline{h(y, x)}$ implica (tomando $x = y$) que $h(x, x) = \overline{h(x, x)}$, i.e., $h(x, x) \in \mathbb{R}$. Por otra parte, si la forma sesquilineal h verifica $h(x, x) \in \mathbb{R}$ para todo $x \in V$ entonces para todos $x, y \in V$ tenemos que

$$h(x, y) + h(y, x) = (h(x + y, x + y) - h(x - y, x - y))/2 \in \mathbb{R},$$

por lo que $h(x, y) - \overline{h(y, x)} = \underbrace{(h(x, y) + h(y, x))}_{\in \mathbb{R}} - \underbrace{(h(y, x) + \overline{h(y, x)})}_{= 2\Re(h(y,x)) \in \mathbb{R}} \in \mathbb{R}$. Si reemplazamos x por ix obtenemos

$$i(h(x, y) - \overline{h(y, x)}) = h(ix, y) + i\overline{h(y, x)} = h(ix, y) + \overline{\overline{(-i)}h(y, x)} = h(ix, y) - \overline{h(y, ix)} \in \mathbb{R}.$$

Por otra parte, si $a \in \mathbb{C}$ verifica $a \in \mathbb{R}$ y $ia \in \mathbb{R}$ entonces necesariamente $a = 0$, de donde concluimos que $h(x, y) - \overline{h(y, x)} = 0$ y luego h es hermitiana. $\square$

Observación 7.1.9. — Sea $h : V \times V \to \mathbb{C}$ forma sesquilineal en $V \cong \mathbb{C}^n$. Dado que $(ih)^* = -ih^*$, se tiene

$$h \text{ hermitiana} \Leftrightarrow ih \text{ anti-hermitiana}, \quad h \text{ anti-hermitiana} \Leftrightarrow ih \text{ hermitiana}.$$

En particular, h es anti-hermitiana si y sólo si $h(v, v) \in i\mathbb{R}$ (imaginario puro) para todo $v \in V$.

Notemos que toda forma sesquilineal h se descompone de manera única en una suma

$$h = \sigma + \alpha,$$

donde σ es hermitiana y α anti-hermitiana, donde

$$\sigma = \frac{1}{2}(h + h^*) \quad \text{y} \quad \alpha = \frac{1}{2}(h - h^*).$$

En particular, tenemos el siguiente resultado.

Proposición 7.1.10. — *Sea $V \cong \mathbb{C}^n$ espacio vectorial complejo. Entonces, tenemos una descomposición en suma directa de espacios vectoriales reales*

$$\mathrm{Sesq}(V) = \mathrm{Herm}(V) \oplus \mathrm{AntiHerm}(V),$$

donde

$$\dim_{\mathbb{R}} \mathrm{Herm}(V) = \dim_{\mathbb{R}} \mathrm{AntiHerm}(V) = \frac{1}{2} \dim_{\mathbb{R}} \mathrm{Sesq}(V) = \dim_{\mathbb{C}} \mathrm{Sesq}(V).$$

Demostración. — La discusión anterior implica directamente que $\mathrm{Sesq}(V) = \mathrm{Herm}(V) \oplus \mathrm{AntiHerm}(V)$ como $\mathbb{R}$-espacios vectoriales. Más aún, tenemos un isomorfismo $\mathbb{R}$-lineal

$$\mathrm{Herm}(V) \xrightarrow{\sim} \mathrm{AntiHerm}(V),\ h \mapsto ih,$$

de donde se deduce la relación entre las dimensiones. $\square$

Tal como en el caso de formas bilineales reales, podemos describir formas sesquilineales en términos matriciales. Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo y sea $\mathscr{B} = (e_1, \ldots, e_n)$ una base de V. Dados vectores

$$x = \sum_{j=1}^{n} x_j e_j \in V,\ y = \sum_{j=1}^{n} y_j e_j \in V,$$

la hipótesis de sesquilinealidad de $h : V \times V \to \mathbb{C}$ implica que

$$h(x, y) = h\left(\sum_{j=1}^{n} x_j e_j, y\right) = \sum_{j=1}^{n} x_j h(e_j, y) = \sum_{j=1}^{n} x_j h\left(e_j, \sum_{k=1}^{n} y_k e_k\right) = \sum_{j=1}^{n}\sum_{k=1}^{n} x_j \overline{y_k}\, h(e_j, e_k).$$

Luego, si definimos $h_{jk} := h(e_j, e_k) \in \mathbb{C}$ entonces

$$h(x,y) = \sum_{1 \leq j,k \leq n} h_{jk} x_j \overline{y_k}.$$

Recíprocamente, toda expresión de esta forma con coeficientes $h_{jk} \in \mathbb{C}$ arbitrarios define una forma sesquilineal $h : V \times V \to \mathbb{C}$. En particular, $\mathrm{Sesq}(V) \cong M_n(\mathbb{C})$ y luego $\dim_{\mathbb{C}} \mathrm{Sesq}(V) = n^2$.

Definición 7.1.11. — Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo y sea $\mathscr{B} = (e_1, \ldots, e_n)$ una base de V. Dada una forma sesquilineal $h : V \times V \to \mathbb{C}$, definimos la **matriz de h en la base $\mathscr{B}$** mediante

$$H = \mathrm{Mat}_{\mathscr{B}}(h) := (h_{jk})_{1 \leq j,k \leq n} = (h(e_j, e_k))_{1 \leq j,k \leq n}.$$

Más aún, si $A = (a_{jk}) \in M_{n \times p}(\mathbb{C})$ es una matriz compleja de n filas y p columnas, llamamos la **matriz adjunta** a la matriz $A^* \in M_{p \times n}(\mathbb{C})$ dada por la transpuesta conjugada

$$A^* = {}^{\mathrm{t}}\overline{A} = (\overline{a_{kj}}).$$

En particular, $\mathrm{Mat}_{\mathscr{B}}(h^*) = H^*$.

Ejercicio 7.1.12. — Sean $A \in M_{n \times p}(\mathbb{C})$ y $B \in M_{p \times r}(\mathbb{C})$ matrices complejas.

a) Probar que $(A^*)^* = A$ y que la aplicación $A \mapsto A^*$ es anti-lineal.
b) Probar que $(AB)^* = B^* A^*$.
c) Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo y sea $\mathscr{B} = (e_1, \ldots, e_n)$ una base de V. Si $h : V \times V \to \mathbb{C}$ es una forma sesquilineal de matriz $H = (h_{jk})$ respecto a $\mathscr{B}$, probar que para $x = \sum_{j=1}^n x_j e_j$ e $y = \sum_{k=1}^n y_k e_k$ se cumple que

$$h(x,y) = {}^{\mathrm{t}}XH\overline{Y},$$

donde

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{C}^n \quad \text{y} \quad Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{C}^n$$

son los vectores coordenadas de x e y respecto a $\mathscr{B}$. Deducir que si $\mathscr{B}'$ es otra base de V y $P = \mathrm{Mat}_{\mathscr{B}}(\mathscr{B}')$ es la matriz de cambio de base, entonces

$$H' = \mathrm{Mat}_{\mathscr{B}'}(h) = {}^{\mathrm{t}}PH\overline{P}.$$

❦ Usualmente en Física se utiliza la notación A^* para referirse a la matriz conjugada de A (que nosotros denotamos $\overline{A}$), y se usa $A^\dagger$ para referirse a la matriz adjunta de A (que nosotros denotamos A^*).

Una consecuencia inmediata de la definición de matriz adjunta es la siguiente caracterización matricial de las formas hermitianas y anti-hermitianas.

Proposición 7.1.13. — *Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo, y sea $h : V \times V \to \mathbb{C}$ una forma sesquilineal. Entonces, para toda base $\mathscr{B}$ de V se tiene que:*

1. *La forma h es* **hermitiana** *si y sólo $H^* = H$ (i.e., ${}^{\mathrm{t}}H = \overline{H}$).*
2. *La forma h es* **anti-hermitiana** *si y sólo si $H^* = -H$ (i.e., ${}^{\mathrm{t}}H = -\overline{H}$).*

donde $H = \mathrm{Mat}_{\mathscr{B}}(h)$ es la matriz de h en la base $\mathscr{B}$.

Demostración. — Sea $\mathscr{B} = (e_1, \dots, e_n)$ una base arbitraria de V. Gracias al Ejercicio 7.1.12 (c), si escribimos $x = \sum_{j=1}^n x_j e_j$ e $y = \sum_{k=1}^n y_k e_k$ entonces se cumple que

$$h(x, y) = {}^{\mathrm{t}}XH\overline{Y},$$

donde

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{C}^n \quad \text{y} \quad Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{C}^n$$

son los vectores coordenadas de x e y respecto a $\mathscr{B}$. En particular, $h^*(x, y) = {}^{\mathrm{t}}XH^*\overline{Y}$ y luego $h^* = h$ (resp. $h^* = -h$) si y sólo si $H^* = H$ (resp. $H^* = -H$). $\square$

***Definición 7.1.14* (matriz (anti-)hermitiana).** — Sea $A \in M_n(\mathbb{C})$ una matriz compleja. Decimos que A es una **matriz hermitiana** (resp. **matriz anti-hermitiana**) si $A^* = A$ (resp. $A^* = -A$). Denotaremos por $\mathrm{Herm}_n(\mathbb{C})$ (resp. $\mathrm{AntiHerm}_n(\mathbb{C})$) al espacio vectorial <u>real</u> de matrices hermitianas (resp. anti-hermitianas).

Observación 7.1.15. —

1. Si $A = (a_{jk}) \in M_n(\mathbb{C})$ es una matriz hermitiana, los coeficientes diagonales a_{jj} verifican $a_{jj} = \overline{a_{jj}}$, y luego $a_{jj} \in \mathbb{R}$. De manera análoga, si A es una matriz anti-hermitiana entonces sus coeficientes diagonales son imaginarios puros.
2. Notemos que $\dim_{\mathbb{C}} \mathrm{Sesq}(V) = n^2$ y la relación

$$\dim_{\mathbb{R}} \mathrm{Herm}(V) = \dim_{\mathbb{R}} \mathrm{AntiHerm}(V) = \frac{1}{2} \dim_{\mathbb{R}} \mathrm{Sesq}(V) = \dim_{\mathbb{C}} \mathrm{Sesq}(V)$$

implican que $\dim_{\mathbb{R}} \mathrm{Herm}_n(\mathbb{C}) = \dim_{\mathbb{R}} \mathrm{AntiHerm}_n(\mathbb{C}) = n^2$.

Definición 7.1.16 **(forma cuadrática hermitiana)**

Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo y sea $h : V \times V \to \mathbb{C}$ una forma hermitiana. La **forma cuadrática hermitiana** asociada a h es por definición la función $Q : V \to \mathbb{R}$ dada por

$$Q(x) = h(x,x) \quad \text{para todo } x \in V.$$

En particular, para todo $x \in V$ y todo $\lambda \in \mathbb{C}$ se cumple

$$Q(\lambda x) = h(\lambda x, \lambda x) = \lambda\overline{\lambda} h(x,x) = |\lambda|^2 Q(x).$$

Observación 7.1.17. — Dado que h es una forma hermitiana, tenemos que $Q(x) = h(x,x) \in \mathbb{R}$ para todo $x \in V$ (c.f. Proposición 7.1.8). Luego, Q está asociada también a la forma *bilineal real* $B : V \times V \to \mathbb{R}$ dada por $(x,y) \mapsto B(x,y) := \Re h(x,y)$, la cual es simétrica pues

$$B(y,x) = \Re h(y,x) = \Re\overline{h(x,y)} = \Re h(x,y) = B(x,y).$$

En particular, Q es una *forma cuadrática real* si pensamos $V \cong \mathbb{R}^{2n}$ como un espacio vectorial real. Más aún, la forma cuadrática real Q verifica la propiedad adicional

$$Q(\lambda x) = |\lambda|^2 Q(x) \quad \forall \lambda \in \mathbb{C},\ \forall x \in V.$$

Veremos que esta última propiedad permite distinguir a las formas cuadráticas hermitianas entre las formas cuadráticas reales (ver Proposición 7.1.19). Finalmente, es importante destacar que una forma cuadrática hermitiana **no es** una forma cuadrática compleja asociada a una forma bilineal compleja sobre $V \cong \mathbb{C}^n$, puesto que en el último caso deberíamos tener $Q(\lambda x) = \lambda^2 Q(x)$ para todo $\lambda \in \mathbb{C}$ (lo cual impide que Q tenga valores reales, a menos que $Q = 0$).

Tal como en el caso real, la **fórmula de polarización** nos permitirá deducir de manera única la forma hermitiana asociada a una forma cuadrática hermitiana.

Lema 7.1.18 **(polarización hermitiana).** — *Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo y sea $h : V \times V \to \mathbb{C}$ una forma hermitiana. Si denotamos por $Q : V \to \mathbb{R}$ la forma cuadrática hermitiana $Q(x) = h(x,x)$ asociada a h, entonces para todos $x, y \in V$ se tiene:*

1. $\Re h(x,y) = \frac{1}{2}(Q(x+y) - Q(x) - Q(y)) = \frac{1}{4}(Q(x+y) - Q(x-y))$.
2. $\Im h(x,y) = \frac{1}{2}(Q(x+iy) - Q(x) - Q(y)) = \frac{1}{4}(Q(x+iy) - Q(x-iy))$.

Luego,

$$h(x,y) = \frac{1}{4}(Q(x+y) - Q(x-y) + iQ(x+iy) - iQ(x-iy)).$$

Demostración. — Para todos $x, y \in V$ tenemos que

$$Q(x+y) = h(x+y, x+y) = Q(x) + Q(y) + h(x,y) + h(y,x)$$
$$Q(x-y) = h(x-y, x-y) = Q(x) + Q(y) - h(x,y) - h(y,x)$$

y, dado que $h(x,y)+h(y,x) = h(x,y) + \overline{h(x,y)} = 2\Re h(x,y)$, obtenemos[(1)] (1). Dado que para todo $z \in \mathbb{C}$ se tiene que $\Im(z) = \Re(-iz)$, obtenemos

$$\Im h(x,y) = \Re(-ih(x,y)) = \Re h(x, iy),$$

por lo que (2) se obtiene de (1) reemplazando y por iy, y usando que $Q(iy) = |i|^2 Q(y) = Q(y)$. Finalmente, dado que $h(x,y) = \Re h(x,y) + i\Im h(x,y)$, la fórmula de polarización hermitiana

$$h(x,y) = \frac{1}{4}(Q(x+y) - Q(x-y) + iQ(x+iy) - iQ(x-iy)).$$

se obtiene de (1) y (2). □

La siguiente proposición prueba que las formas cuadráticas reales sobre $V \cong \mathscr{C}^n \cong \mathbb{R}^{2n}$ que verifican la propiedad adicional $Q(\lambda x) = |\lambda|^2 Q(x)$ para todo $\lambda \in \mathbb{C}$ corresponden exactamente a formas cuadráticas hermitianas. Mejor aún, veremos que basta considerar $\lambda = i$.

Proposición 7.1.19. — *Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo. Si $Q : V \to \mathbb{R}$ es una forma cuadrática real en $V \cong \mathbb{R}^{2n}$ verificando adicionalmente que $Q(ix) = Q(x)$ para todo $x \in V$, entonces la fórmula de polarización hermitiana define una forma sesquilineal hermitiana h. En particular, Q es una forma cuadrática hermitiana.*

Demostración. — El hecho que $Q : V \to \mathbb{R}$ sea una forma cuadrática real implica que

$$B(x,y) = \frac{1}{4}(Q(x+y) - Q(x-y))$$

es una forma bilineal simétrica real. Deducimos que

$$h(x,y) := \frac{1}{4}(Q(x+y) - Q(x-y) + iQ(x+iy) - iQ(x-iy)) = B(x,y) - iB(x,iy)$$

es una forma $\mathbb{R}$-bilineal con valores complejos. Más aún, por una parte tenemos que

$$h(x,iy) = \frac{1}{4}(Q(x+iy) - Q(x-iy) + iQ(x-y) - iQ(x+y)) = -ih(x,y).$$

[(1)]Notemos que (1) es exactamente la fórmula de polarización *real* aplicada a la forma bilineal simétrica $(x,y) \mapsto \Re h(x,y)$.

Por otra parte, la hipótesis $Q(ix) = Q(x)$ implica que $Q(x) = Q(-x) = Q(-ix)$, de donde deducimos

$$\begin{aligned} h(ix, y) &= \frac{1}{4}(Q(ix + y) - Q(ix - y) + iQ(ix + iy) - iQ(ix - iy)) \\ &= \frac{1}{4}(Q(x - iy) - Q(x + iy) + iQ(x + y) - iQ(x - y)) = ih(x, y). \end{aligned}$$

De lo anterior y de la $\mathbb{R}$-bilinealidad de h se obtiene que $h : V \times V \to \mathbb{C}$ es una forma hermitiana. □

Terminemos esta sección definiendo el rango de una forma hermitiana o, equivalentemente, de una forma cuadrática hermitiana.

***Definición 7.1.20* (rango).** — Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo y sea $h : V \times V \to \mathbb{C}$ una forma hermitiana. Dada una base $\mathscr{B} = (e_1, \ldots, e_n)$ de V, definimos el **rango** de h (o de la forma cuadrática hermitiana $Q : V \to \mathbb{R}$ asociada a h) como $\mathrm{rg}(h) = \mathrm{rg}(Q) := \mathrm{rg}(H)$, donde $H = \mathrm{Mat}_{\mathscr{B}}(h)$ es la matriz de h respecto a la base[(2)] $\mathscr{B}$. Diremos que la forma hermitiana h (o que Q) es **no-degenerada** si $\mathrm{rg}(h) = \dim_{\mathbb{C}}(V)$, i.e., si la matriz H es invertible.

7.2. Ortogonalidad hermitiana y Teorema de Sylvester

En esta sección estudiaremos el análogo hermitiano de la noción de ortogonalidad respecto a una forma bilineal simétrica (o, equivalentemente, respecto a una forma cuadrática).

Definición 7.2.1. — Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo y sea $h : V \times V \to \mathbb{C}$ una forma hermitiana. Diremos que:

1. Dos vectores $x, y \in V$ son **ortogonales** respecto a h si $h(x, y) = 0$ (o equivalemente, si $h(y, x) = 0$).
2. Si $U \subseteq V$ es un sub-conjunto no-vacío, el conjunto[(3)]

$$U^{\perp} = \{y \in V \mid \forall x \in U,\ h(x, y) = 0\}$$

   es el **ortogonal de** U.

(2) Notamos que dicho rango **no depende** de la base escogida, pues si $\mathscr{B}'$ es otra base y $P = \mathrm{Mat}_{\mathscr{B}}(\mathscr{B}')$ es la matriz de cambio de base, entonces $H' = \mathrm{Mat}_{\mathscr{B}'}(h) = {}^{\mathrm{t}}PH\overline{P}$ y, dado que ${}^{\mathrm{t}}P, \overline{P} \in \mathrm{GL}_n(\mathscr{C})$, se tiene que $\mathrm{rg}(H') = \mathrm{rg}(H)$.

(3) En ocasiones, para insistir en el hecho que $U^{\perp}$ es el ortogonal respecto a h (o equivalentemente, respecto a la forma cuadrática hermitiana Q asociada a h), es que se denota también $U^{\perp_h}$ o $U^{\perp_Q}$.

3. El **kernel** de h es
$$\ker(h) := V^{\perp} = \{x \in V \mid \forall y \in V,\ h(x,y) = 0\}.$$
4. El conjunto de **vectores isótropos** de la forma cuadrática hermitiana $Q(x) = h(x,x)$ asociada es $\{x \in V \mid Q(x) = 0\}$.

Observación 7.2.2. — Cabe destacar que a pesar que $U \subseteq V$ pueda ser símplemente un sub-conjunto (y no necesariamente un *sub-espacio vectorial*) de V, el ortogonal $U^{\perp} \subseteq V$ es *siempre* un sub-espacio vectorial[(4)] de V. Por otro lado, el conjunto de vectores isótropos de una forma cuadrática hermitiana Q (asociada a la forma hermitiana h) es un **cono complejo**, i.e., si $x \in V$ es un vector isótropo entonces λx es un vector isótropo para todo $\lambda \in \mathbb{C}$. En general, se tiene la inclusión $\ker(h) \subseteq \{x \in V \mid Q(x) = 0\}$ pero dicha inclusión puede ser estricta, tal como lo muestra el siguiente ejemplo.

Ejemplo: Sea $V = \mathbb{C}^2$ y consideremos la forma cuadrática hermitiana
$$Q(z_1, z_2) = z_1\overline{z_1} - z_2\overline{z_2} = |z_1|^2 - |z_2|^2.$$
La fórmula de polarización hermitiana implica que $h((z_1,z_2),(w_1,w_2)) = z_1\overline{w_1} - z_2\overline{w_2}$ es la forma hermitiana asociada. La matríz H de la forma h respecto a la base canónica $\mathscr{B}$ de $\mathbb{C}^2$ está dada por
$$H = \mathrm{Mat}_{\mathscr{B}}(h) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$
la cual es una matriz invertible y luego h es no-degenerada. En particular,
$$\ker(h) = \{0\} \neq \{(z_1,z_2) \in \mathbb{C}^2 \mid Q(z_1,z_2) = 0\} = \{(z_1,z_2) \in \mathbb{C}^2 \mid |z_1| = |z_2|\}.$$

Ejercicio 7.2.3. — Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo y sea $h : V \times V \to \mathbb{C}$ una forma hermitiana. Sean $U, W \subseteq V$ sub-conjuntos no-vacíos de V.

a) Probar que si $U \subseteq V$ entonces $V^{\perp} \subseteq U^{\perp}$.
b) Probar que $U^{\perp} = \mathrm{Vect}_{\mathbb{C}}(U)^{\perp}$.
c) Supongamos que $U \subseteq V$ es un sub-espacio vectorial y que $(e_1, \ldots, e_r)$ es un conjunto generador de U (no necesariamente linealmente independiente), probar que
$$U^{\perp} = \{e_1, \ldots, e_r\}^{\perp} = \{x \in V \mid h(x, e_j) = 0\ \forall j = 1, \ldots, r\}.$$

[(4)] Si $y_1, y_2 \in U^{\perp}$ y $\lambda \in \mathbb{C}$ entonces para todo $x \in U$ se cumple $h(x, \lambda y_1 + y_2) = \overline{\lambda} h(x, y_1) + h(x, y_2) = 0$, i.e., $\lambda y_1 + y_2 \in U^{\perp}$.

Teorema 7.2.4. — *Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo y sea $h : V \times V \to \mathbb{C}$ una forma hermitiana. Entonces, para todo sub-espacio vectorial $U \subseteq V$ se tiene que:*

1. $U \subseteq (U^\perp)^\perp$ *y* $\dim_{\mathbb{C}}(U^\perp) \geq \dim_{\mathbb{C}}(V) - \dim_{\mathbb{C}}(U)$.
2. $\ker(h) = \{0\}$ *si y sólo si h es no-degenerada.*
3. *Si h es no-degenerada, entonces* $U = (U^\perp)^\perp$ *y* $\dim_{\mathbb{C}}(U^\perp) = \dim_{\mathbb{C}}(V) - \dim_{\mathbb{C}}(U)$.
4. *Si* $U \cap U^\perp = \{0\}$ *entonces* $V = U \oplus U^\perp$.

Demostración. — Sea $x \in U$, entonces para todo $y \in U^\perp$ se tiene $h(x,y) = \overline{h(y,x)} = 0$ por lo que $x \in (U^\perp)^\perp$, de donde se prueba la primera parte de (1). Para probar la segunda parte de (1) consideramos una base $(e_1, \dots, e_r)$ una base de U, donde $r = \dim_{\mathbb{C}}(U)$, la cual completamos en una base $\mathscr{B} = (e_1, \dots, e_n)$ de V. Sea $H = (h_{jk})$ la matriz de h respecto a la base $\mathscr{B}$ (i.e., $h_{jk} = h(e_j, e_k)$ para todo $j, k = 1, \dots, n$). Sea $x = x_1 e_1 + \ldots + x_n e_n \in V$, entonces $x \in U^\perp$ si y sólo si $h(x, e_j) = 0$ para $j = 1, \dots, r$. Dado que

$$h(x, e_j) = h(x_1 e_1 + \ldots + x_n e_n, e_j) = \sum_{k=1}^{n} x_k h(e_k, e_j) = \sum_{k=1}^{n} h_{kj} x_k,$$

la condición $x \in U^\perp$ equivale a decir que el vector de coordenadas $X = (x_1, \dots, x_n) \in \mathbb{C}^n$ es solución del sistema lineal

$$\begin{cases} a_{11}x_1 & + \cdots + & a_{n1}x_n = & 0 \\ \vdots & & \vdots & \\ a_{1r}x_1 & + \cdots + & a_{nr}x_n = & 0 \end{cases}$$

cuya matriz asociada $A \in M_{n \times r}(\mathbb{C})$ es la matriz obtenida al considerar las r primeras filas de ${}^{\mathrm{t}}H$. Dado que el espacio de soluciones del sistema es de dimensión $n - \mathrm{rg}(A)$, obtenemos

$$\dim_{\mathbb{C}}(U^\perp) = n - \mathrm{rg}(A) \geq n - r,$$

de donde obtenemos la segunda parte de (1). Más aún, en el caso particular donde $U = V$, tenemos que $A = {}^{\mathrm{t}}H$ y luego $\dim_{\mathbb{C}}(U^\perp) = n - \mathrm{rg}({}^{\mathrm{t}}H) = n - \mathrm{rg}(H)$. Luego, $\ker(h) = V^\perp$ es nulo si y sólo si $\mathrm{rg}(H) = n$, de donde obtenemos (2).

Para probar (3) supongamos que h es no-degenerada, i.e. las columnas de la matriz H son linealmente independientes. En particular, las primeras r columnas de H son linealmente independientes y luego $\mathrm{rg}(A) = r$, de donde deducimos que $\dim_{\mathbb{C}}(U^\perp) = n - r$. Reemplazando U por $U^\perp$ obtenemos la igualdad $\dim_{\mathbb{C}}(U^{\perp})^{\perp} = n - (n - r) = r$ y luego la inclusión $U \subseteq (U^{\perp})^{\perp}$ es una igualdad, de donde obtenemos (3).

Finalmente, si suponemos que $U \cap U^{\perp} = \{0\}$ (sin suponer necesariamente que h es no-degenerada), entonces U y $U^{\perp}$ están en suma directa y la dimensión del sub-espacio $U \oplus U^{\perp}$ de V es $d = r + \dim_{\mathbb{C}}(U^{\perp})$. Gracias a (1), tenemos que $d \geq n$ y luego $V = U \oplus U^{\perp}$ (en particular, $\dim_{\mathbb{C}}(U^{\perp}) = n - r$), lo que prueba (4). □

Un caso particular importante es cuando los vectores de una base son ortogonales respecto a una forma hermitiana.

***Definición 7.2.5* (base ortogonal).** — Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo, sea $h : V \times V \to \mathbb{C}$ una forma hermitiana y sea $Q : V \to \mathbb{R}$ la forma cuadrática hermitiana asociada. Decimos que una base $\mathscr{B} = (e_1, \ldots, e_n)$ de V es una **base ortogonal** respecto a h (o respecto a Q) si

$$h(e_j, e_k) = 0 \text{ para todos } j \neq k.$$

Equivalentemente, la matriz $H = \mathrm{Mat}_{\mathscr{B}}(h)$ es una matriz *diagonal*

$$H = \begin{pmatrix} \lambda_1 & 0 & \cdots & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 0 & \lambda_n \end{pmatrix}$$

donde los $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ son números *reales* (pues H es hermitiana), siendo esto último a su vez equivalente a que si $(z_1, \ldots, z_n)$ son coordenadas respecto a la base $\mathscr{B}$ entonces

$$Q(z_1, \ldots, z_n) = \lambda_1 |z_1|^2 + \ldots + \lambda_n |z_n|^2.$$

El siguiente resultado afirma que siempre podemos encontrar al menos una base ortogonal respecto a una forma hermitiana dada.

Lema 7.2.6. — *Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo y sea $h : V \times V \to \mathbb{C}$ una forma hermitiana. Entonces, existe una base $\mathscr{B}$ de V que es ortogonal respecto a h.*

Demostración. — Demostraremos la existencia de una base ortogonal por inducción en la dimensión $n = \dim_{\mathbb{C}}(V)$. Dado que el resultado es cierto si $n = 0$ o si $h = 0$, podemos suponer que el resultado es cierto para $n - 1$, y que además

$h \neq 0$. En particular, la forma cuadrática hermitiana $Q : V \to \mathbb{R}$ asociada a h es no-nula y luego existe un vector $e_1 \in V$ tal que $Q(e_1) \neq 0$.

Sea $U = \mathrm{Vect}_{\mathbb{C}}(e_1)$ la recta generada por e_1. Dado que $Q(e_1) = h(e_1, e_1) \neq 0$, tenemos que $U \cap U^{\perp} = \{0\}$ y luego $V = U \oplus U^{\perp}$. La hipótesis de inducción implica que existe una base $(e_2, \ldots, e_n)$ de $U^{\perp}$ tal que $h(e_j, e_k) = 0$ para todos $j \neq k$ en $\{2, \ldots, n\}$. Luego, $\mathscr{B} = (e_1, e_2, \ldots, e_n)$ es una base de V ortogonal respecto a h. $\square$

Concluímos esta sección con el análogo hermitiano del teorema de Sylvester de clasificación de formas cuadráticas reales.

***Teorema 7.2.7* (de Sylvester, caso hermitiano)**

Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo, sea $h : V \times V \to \mathbb{C}$ una forma hermitiana, y sea $Q : V \to \mathbb{R}$ la forma cuadrática hermitiana asociada. Sea $\mathscr{B} = (e_1, \ldots, e_n)$ una base de V ortogonal respecto a h y sea $H = \mathrm{Mat}_{\mathscr{B}}(h)$ la matriz diagonal real de coeficientes $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$. Entonces,

1. *Módulo permutación de los elementos de la base $\mathscr{B}$, podemos suponer que $\lambda_1, \ldots, \lambda_r \neq 0$ y que $\lambda_j = 0$ si $j > r$. En particular, si $(z_1, \ldots, z_n)$ son las coordenadas respecto a la base $\mathscr{B}$, entonces*
$$Q(z_1, \ldots, z_n) = \lambda_1 |z_1|^2 + \ldots + \lambda_r |z_r|^2.$$
2. *Sea $p \in \mathbb{N}$ (resp. $q \in \mathbb{N}$) el número de índices $j \in \{1, \ldots, n\}$ tales que $Q(e_j) > 0$ (resp. $Q(e_j) < 0$). Entonces, p y q* no dependen *de la base ortogonal escogida. El par (p, q) se llama la* **signatura**[5] *de la forma hermitiana h (o de Q), y se cumple que $p + q = r = \mathrm{rg}(h)$.*
3. *El* kernel *$\ker(h)$ es el sub-espacio $\mathrm{Vect}_{\mathbb{C}}(e_{r+1}, \ldots, e_n)$ dado por las ecuaciones $z_1 = \ldots = z_r = 0$.*

Más aún, podemos escoger $\mathscr{B}$ de tal suerte que
$$H = \left(\begin{array}{c:c:c} \mathrm{I}_p & 0 & 0 \\ \hdashline 0 & -\mathrm{I}_q & 0 \\ \hdashline 0 & 0 & \mathbf{0}_{n-r} \end{array}\right).$$

Demostración. — El hecho que $\mathscr{B}$ sea una base ortogonal respecto a h implica directamente (1), así como la igualdad $p + q = r = \mathrm{rg}(h)$ en (2).

[5] En ocasiones, se define también la signatura de h como el trío $(\sigma_+, \sigma_-, \sigma_0)$, donde $\sigma_+ = p$, $\sigma_- = q$ y $\sigma_0 = \dim_{\mathbb{C}} \ker(h)$.

Comencemos por probar (3). La fórmula de polarización hermitiana implica que, respecto a la base $\mathscr{B}$, la forma hermitiana h está dada por

$$h(x,y) = \lambda_1 x_1 \overline{y_1} + \ldots + \lambda_r x_r \overline{y_r},$$

donde $x = \sum_{j=1}^n x_j e_j$ e $y = \sum_{j=1}^n y_j e_j$. Supongamos que $x = \sum_{j=1}^n x_j e_j$ pertenece a $\ker(h)$ y consideremos $y = e_j$ para $j = 1, \ldots, r$ (i.e., $y_j = 1$ e $y_k = 0$ para $k \neq j$), de donde obtenemos que $x_j = 0$ y en particular $x \in \mathrm{Vect}_{\mathbb{C}}(e_{r+1}, \ldots, e_n)$. Recíprocamente, la fórmula de h en la base $\mathscr{B}$ implica que todo vector $x = \sum_{j=1}^n x_j e_j$ en $\mathrm{Vect}_{\mathbb{C}}(e_{r+1}, \ldots, e_n)$ (i.e., tal que $x_1 = \ldots = x_r = 0$) pertenece a $\ker(h)$, de donde se obtiene (3).

Veamos ahora (2), i.e., que el par (p,q) es independiente de la base ortogonal escogida. Recordemos que $r = \mathrm{rg}(h)$. Sean $\mathscr{B} = (e_1, \ldots, e_n)$ y $\mathscr{B}' = (e'_1, \ldots, e'_n)$ dos bases de V ortogonales respecto a h. Denotemos por p (resp. p') al número de índices $j \in \{1, \ldots, n\}$ tales que $Q(e_j) > 0$ (resp. $Q(e'_j) > 0$) y por q (resp. q') al número de índices $j \in \{1, \ldots, n\}$ tales que $Q(e_j) < 0$ (resp. $Q(e'_j) < 0$). Entonces,

$$p + q = r = p' + q'.$$

En particular, basta probar que $q = q'$ para obtener que $(p,q) = (p', q')$. Reordenando los elementos de $\mathscr{B}$ y $\mathscr{B}'$ si fuese necesario, podemos suponer que

$$(\star) \qquad \begin{cases} Q(e_j) > 0 & \text{si } j = 1, \ldots, p \\ Q(e_j) < 0 & \text{si } j = p+1, \ldots, p+q = r \\ Q(e_j) = 0 & \text{si } j > p+q = r \end{cases} \qquad \begin{cases} Q(e'_j) > 0 & \text{si } j = 1, \ldots, p' \\ Q(e'_j) < 0 & \text{si } j = p'+1, \ldots, p'+q' = r \\ Q(e'_j) = 0 & \text{si } j > p'+q' = r \end{cases}$$

Sea P_+ el sub-espacio vectorial de V generado por los $n - q$ vectores e_j tales que $Q(e_j) \geq 0$. En particular, $\dim_{\mathbb{C}} P_+ = n - q$. Sea $x \in P_+$ y escribamos $x = \sum_{j \in I} x_j e_j$ donde $J = \{1, \ldots, p\} \cup \{r+1, \ldots, n\}$. Luego, las relaciones $(\star)$ implican que

$$Q(x) = \sum_{j=1}^{p} |x_j|^2 Q(e_j) \geq 0.$$

Por otra parte, si P'_- es el sub-espacio vectorial de V generado por los q' vectores e'_j tales que $Q(e'_j) < 0$ entonces $\dim_{\mathscr{C}} P'_- = q'$. Sea $y \in P'_-$ un vector no-nulo y escribamos $y = \sum_{j=p'+1}^{r} y_j e'_j$, donde al menos uno de los $y_j \neq 0$ (pues $y \neq 0$). Entonces, las relaciones $(\star)$ implican que

$$Q(y) = \sum_{j=p'+1}^{r} |y_j|^2 Q(e'_j) < 0.$$

En consecuencia, tenemos que $P_+ \cap P'_- = \{0\}$ y por lo tanto

$$n = \dim_{\mathbb{C}}(V) \geq \dim_{\mathbb{C}}(P_+) + \dim_{\mathbb{C}}(P'_-) = n - q + q',$$

de donde deducimos que $q \geq q'$. Intercambiando los roles de $\mathscr{B}$ y $\mathscr{B}'$ en la discusión anterior, obtenemos del mismo modo que $q' \geq q$. Así, $q = q'$ y por ende $p = p'$, lo cual prueba (2).

Finalmente, si ordenamos los elementos de una base ortogonal $\mathscr{B} = (e_1, \ldots, e_n)$ como en $(\star)$ y consideramos el valor absoluto *real* $|Q(e_j)| > 0$ para $j = 1, \ldots, p+q = r$, entonces al reemplazar e_j por el vector $e_j/\sqrt{|Q(e_j)|}$ para cada $j = 1, \ldots, r$ obtenemos una base ortogonal $\mathscr{B}'$ tal que $\mathrm{Mat}_{\mathscr{B}'}(h)$ es de la forma

$$\left(\begin{array}{c:c:c} \mathrm{I}_p & 0 & 0 \\ \hdashline 0 & -\mathrm{I}_q & 0 \\ \hdashline 0 & 0 & \mathbf{0}_{n-r} \end{array}\right),$$

de donde se obtiene el resultado. $\square$

7.3. Grupo unitario $\mathbf{U}(p,q)$ y grupo especial unitario $\mathbf{SU}(p,q)$

Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo y sea $Q : V \to \mathbb{R}$ una forma cuadrática hermitiana. Denotamos por $h : V \times V \to \mathbb{C}$ la única forma hermitiana asociada a Q mediante la fórmula de polarización hermitiana.

Ejercicio 7.3.1. — Sea $u : V \to V$ un endomorfismo. Probar (usando la fórmula de polarización hermitiana) que las propiedades siguientes son equivalentes:

1. u preserva la forma hermitiana h, i.e., $h(u(x), u(y)) = h(x, y)$ para todos $x, y \in V$.
2. u preserva la forma cuadrática hermitiana Q, i.e., $Q(u(x)) = Q(x)$ para todo $x \in V$.

Definición 7.3.2 **(automorfismo unitario).** — Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo y sea $Q : V \to \mathbb{R}$ una forma cuadrática hermitiana. Decimos que un automorfismo $u : V \xrightarrow{\sim} V$ (i.e., $u \in \mathrm{GL}(V) \cong \mathrm{GL}_n(\mathbb{C})$) es **unitario** respecto a Q si u preserva Q (o equivalentemente, preserva la forma hermitiana asociada h). Denotamos por

$$\mathbf{U}(Q) = \{u \in \mathrm{GL}(V) \mid u \text{ unitario respecto a } Q\} \subseteq \mathrm{GL}(V)$$

al conjunto de automorfismos unitarios respecto a Q.

***Proposición 7.3.3* (grupo unitario).** — *Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo y sea $Q : V \to \mathbb{R}$ una forma cuadrática hermitiana. El conjunto* $\mathbf{U}(Q)$ *es un* sub-grupo *de* $\mathrm{GL}(V)$*, llamado el* **grupo unitario de** Q.

Demostración. — La identidad id_V preserva la forma cuadrática hermitiana Q y luego $\mathrm{id}_V \in \mathbf{U}(Q)$. Por otra parte, si $u, v \in \mathbf{U}(Q)$ preservan Q entonces para todo $x \in V$ se tiene

$$Q((u \circ v)(x)) = Q(u(v(x)) = Q(v(x)) = Q(x),$$

por lo que la composición $u \circ v \in \mathbf{U}(Q)$ preserva Q. Finalmente, si $u \in \mathbf{U}(Q)$ preserva Q entonces para todo $x \in V$ escribimos $y = u^{-1}(x)$ y calculamos

$$Q(u^{-1}(x)) = Q(y) = Q(u(y)) = Q(x),$$

por lo que $u^{-1} \in \mathbf{U}(Q)$ preserva Q. Así, concluímos que $\mathbf{U}(Q)$ es un sub-grupo de $\mathrm{GL}(V)$. □

Si la forma cuadrática hermitiana $Q : V \to \mathbb{R}$ es *no-degenerada*, entonces todo endomorfismo $u : V \to V$ que preserva Q es automáticamente biyectivo (i.e., $u \in \mathrm{GL}(V)$). En efecto, si denotamos por $h : V \times V \to \mathbb{C}$ la forma hermitiana asociada a Q y si $x \in V$ verifica $u(x) = 0$, entonces

$$h(x, y) = h(u(x), u(y)) = 0 \quad \text{para todo } y \in V.$$

En particular, $x \in V^\perp$ y por ende $x = 0$ (pues $V^\perp = \{0\}$ si Q es no-degenerada). Luego, u es inyectivo y por ende biyectivo (pues $V \cong \mathbb{C}^n$ es de dimensión finita).

Convención: Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo y $Q : V \to \mathbb{R}$ una forma cuadrática hermitiana. Si $\mathscr{B}$ es una base de V, diremos que la **matriz de** Q respecto a la base $\mathscr{B}$ es la matríz de la única forma hermitiana $h : V \times V \to \mathbb{C}$ asociada a Q mediante la fórmula de polarización hermitiana. En otras palabras,

$$\mathrm{Mat}_{\mathscr{B}}(Q) := \mathrm{Mat}_{\mathscr{B}}(h) = H \in M_n(\mathbb{C}).$$

***Proposición 7.3.4* (grupo especial unitario).** — *Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo y $Q : V \to \mathbb{R}$ una forma cuadrática hermitiana. Sea $\mathscr{B}$ una base de V y denotemos por H la matriz de Q respecto a la base $\mathscr{B}$. Dado un automorfismo $u : V \xrightarrow{\sim} V$ de V de matriz $A = \mathrm{Mat}_{\mathscr{B}}(u)$ respecto a $\mathscr{B}$, tenemos que:*

1. $u \in \mathbf{U}(Q)$ *(i.e., u es unitario) si y sólo si* ${}^{\mathrm{t}}AH\overline{A} = H$.

2. *Si Q es* no-degenerada *y $u \in \mathbf{U}(Q)$ entonces $|\det(u)| = 1$. Más aún, el conjunto*
$$\mathbf{SU}(Q) = \{u \in \mathbf{U}(Q) \mid \det(u) = 1\}$$
es un sub-grupo de $\mathbf{U}(Q)$, llamado el **grupo especial unitario de Q**.

Demostración. — Sean $x, y \in V$. Si representamos x e y por los vectores columna $X, Y \in \mathbb{C}^n$ dados por sus coordenadas respecto a la base $\mathscr{B}$, entonces sabemos que
$$h(x,y) = {}^{\mathrm{t}}XH\overline{Y},$$
por lo que $h(u(x),u(y)) = {}^{\mathrm{t}}(AX)H\overline{(AY)} = {}^{\mathrm{t}}X({}^{\mathrm{t}}AH\overline{A})\overline{Y}$. Luego, la matriz de la forma hermitiana $(x,y) \mapsto h(u(x),u(y))$ respecto a la base $\mathscr{B}$ es ${}^{\mathrm{t}}AH\overline{A}$. Así, u es unitario respecto a Q si y sólo si las formas hermitianas $h(\cdot,\cdot)$ y $h(u(\cdot),u(\cdot))$ coínciden, lo cual equivale a la igualdad $H = {}^{\mathrm{t}}AH\overline{A}$ en (1).

Si Q es no-degenerada entonces $H \in \mathrm{GL}_n(\mathbb{C})$ y luego $\det(H) \neq 0$. En particular, si $u \in \mathbf{U}(Q)$ entonces, gracias a (1), tenemos que $H = {}^{\mathrm{t}}AH\overline{A}$ y por ende $\det(H) = \det(A)\det(H)\det(\overline{A})$. Esto último implica que $\det(A)\det(\overline{A}) = \det(A)\overline{\det(A)} = |\det(A)|^2 = 1$, de donde deducimos que en este caso $|\det(u)| = 1$. Finalmente, dado que $\det(\mathrm{id}_V) = 1$ y que $\det(u \circ v) = \det(u)\det(v)$, concluímos directamente que $\mathrm{id}_V \in \mathbf{SO}(U)$, que $u, v \in \mathbf{SU}(Q)$ entonces $u \circ v \in \mathbf{SU}(Q)$, y que si $u \in \mathbf{SU}(Q)$ entonces $u^{-1} \in \mathbf{SU}(Q)$. En otras palabras, $\mathbf{SU}(Q)$ es un sub-grupo del grupo unitario $\mathbf{U}(Q)$, probando así (2). $\square$

Definición 7.3.5 **(equivalencia de formas cuadráticas hermitianas)** Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo, y sean $Q : V \to \mathbb{R}$ y $Q' : V \to \mathbb{R}$ dos formas cuadráticas hermitianas. Decimos que las formas Q y Q' son **equivalentes** si existe un automorfismo $v \in \mathrm{GL}(V)$ tal que $Q(x) = Q'(v(x))$ para todo $x \in V$, y escribiremos $Q \sim Q'$.

Proposición 7.3.6. — *Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo, y sean $Q : V \to \mathbb{R}$ y $Q' : V \to \mathbb{R}$ dos formas cuadráticas hermitianas. Si las formas $Q \sim Q'$ son equivalentes, entonces los grupos unitarios asociados $\mathbf{U}(Q) \cong \mathbf{U}(Q')$ son isomorfos.*

Demostración. — Sea $v \in \mathrm{GL}(V)$ automorfismo tal que $Q(x) = Q'(v(x))$ para todo $x \in V$. Dado $u : V \to V$ endomorfismo unitario respecto a Q (i.e., $u \in \mathbf{U}(Q)$) y dado $x \in V$, tenemos que:
$$Q'(x) = Q'((v \circ v^{-1})(x)) = Q(v^{-1}(x)) = Q((u \circ v^{-1}(x)) = Q'((v \circ u \circ v^{-1})(x)).$$

En particular, $v \circ u \circ v^{-1} \in \mathbf{U}(Q')$. Finalmente, la función

$$\varphi : \mathbf{U}(Q) \longrightarrow \mathbf{U}(Q'),\ u \longmapsto \varphi(u) = v \circ u \circ v^{-1}$$

es biyectiva (pues su inversa está dada explícitamente por $\varphi^{-1}(u') = v^{-1} \circ u' \circ v$ para todo $u' \in \mathbf{U}(Q')$) y preserva las estructuras de grupos, i.e., $\varphi(u_1 \circ u_2) = \varphi(u_1) \circ \varphi(u_2)$. Concluímos así que $\mathbf{U}(Q) \cong \mathbf{U}(Q')$. $\square$

Caso particular importante: Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo. Gracias al Teorema de Sylvester y a la proposición anterior, el estudio de los grupos unitarios de formas cuadráticas hermitianas **no-degeneradas** de signatura (p, q), con $p + q = n$ en este caso, se reduce a estudiar

$$Q(z_1, \ldots, z_n) = |z_1|^2 + \ldots + |z_p|^2 - |z_{p+1}|^2 - \ldots - |z_n|^2.$$

El grupo unitario correspondiente es denotado $\mathbf{U}(p, q)$ y el grupo especial unitario asociado es denotado $\mathbf{SU}(p, q)$. Explícitamente, si consideramos

$$H_{p,q} := \left(\begin{array}{c|c} \mathrm{I}_p & 0 \\ \hline 0 & -\mathrm{I}_q \end{array}\right) \in \mathrm{GL}_n(\mathbb{C}),$$

entonces $\mathbf{U}(p, q) \cong \{A \in \mathrm{GL}_n(\mathbb{C}) \mid {}^{\mathrm{t}}AH_{p,q}\overline{A} = H_{p,q}\}$ y $\mathbf{SU}(p, q) \cong \{A \in \mathbf{U}(p, q) \mid \det(A) = 1\}$.

Si $(p, q) = (n, 0)$ (i.e., Q es una forma cuadrática hermitiana **definida positiva**) escribimos $\mathbf{U}(n)$ y $\mathbf{SU}(n)$ en lugar de $\mathbf{U}(n, 0)$ y $\mathbf{SU}(n, 0)$. Explícitamente,

$$\mathbf{U}(n) \cong \{A \in \mathrm{GL}_n(\mathbb{C}) \mid {}^{\mathrm{t}}A\overline{A} = \mathrm{I}_n\} = \{A \in \mathrm{GL}_n(\mathbb{C}) \mid A^*A = \mathrm{I}_n\}.$$

7.4. Producto escalar complejo y espacios hermitianos

Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo. Recordemos que una forma hermitiana $Q : V \to \mathbb{R}$ es **definida positiva** si su signatura es $(n, 0)$. En otras palabras, existe una base $\mathscr{B}$ de V tal que, si $(z_1, \ldots, z_n)$ son las coordenadas respecto a dicha base, entonces se tiene

$$Q(z_1, \ldots, z_n) = |z_1|^2 + \ldots + |z_n|^2 = z_1\overline{z_1} + \ldots + z_n\overline{z_n}.$$

En particular, toda forma cuadrática hermitiana definida positiva es *no-degenerada*, i.e., si $x \in V$ entonces $Q(x) = 0$ implica que $x = 0$. Más aún, $Q(x) > 0$ para todo $x \in V \setminus \{0\}$ no-nulo.

***Definición 7.4.1* (producto escalar y espacio hermitiano)**
Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo. Un **producto escalar complejo** es una forma hermitiana $h : V \times V \to \mathbb{C}$ tal que la forma cuadrática

hermitiana asociada es *definida positiva.* Un **espacio hermitiano**[6] es un espacio vectorial complejo dotado de un producto escalar complejo.

Notación: El producto escalar entre dos vectores $x, y \in V \cong \mathbb{C}^n$ se denota usualmente como $\langle x, y\rangle$ en lugar de $h(x, y)$. Así, la forma hermitiana

$$\langle \cdot, \cdot\rangle : V \times V \longrightarrow \mathbb{C},\ (x, y) \longmapsto \langle x, y\rangle$$

es:

1. lineal en la *primera* variable, i.e., para todos $x, y, z \in V$ y $\lambda \in \mathscr{C}$ se tiene que
$$\langle \lambda x, y\rangle = \lambda\langle x, y\rangle \quad \text{y} \quad \langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle.$$
2. anti-lineal en la *segunda* variable, i.e., para todos $x, y, z \in V$ y $\lambda \in \mathscr{C}$ se tiene que
$$\langle x, \lambda y\rangle = \overline{\lambda}\langle x, y\rangle \quad \text{y} \quad \langle x, y + z\rangle = \langle x, y\rangle + \langle x, z\rangle.$$
3. hermitiana, i.e., para todos $x, y \in V$ se tiene que
$$\langle x, y\rangle = \overline{\langle y, x\rangle}.$$
4. definida positiva, i.e., $\langle x, x\rangle \in \mathbb{R}^{>0}$ para todo $x \in V \setminus \{0\}$ no-nulo. En otras palabras, el único vector isótropo es el vector nulo.

Ejemplo: Sea $V = \mathbb{C}^n$. Para todos $x = (x_1, \dots, x_n), y = (y_1, \dots, y_n) \in \mathbb{C}^n$, la relación

$$\langle x, y\rangle := x_1\overline{y_1} + x_2\overline{y_2} + \ldots + x_n\overline{y_n}$$

define un producto escalar complejo en $\mathscr{C}^n$, llamado el *producto escalar canónico.* Así, $\mathscr{C}^n$ es un espacio hermitiano.

Terminología: La generalización de los conceptos de espacio euclidiano y de espacio hermitiano a *dimensión infinita* es lo que conocemos como un **espacio de Hilbert**[7]. Así, un espacio euclidiano (resp. espacio hermitiano) no es nada más que un espacio de Hilbert real (resp. complejo) de dimensión finita.

***Definición 7.4.2* (base ortonormal).** — Sea $V \cong \mathbb{C}^n$ un espacio hermitiano. Diremos que

1. una familia de vectores $(e_j)_{j\in J}$, donde J es un conjunto de índices arbitrario, es **ortonormal** si $\langle e_j, e_k\rangle = \delta_{jk}$ para todos $j, k \in J$ (i.e., $\langle e_j, e_j\rangle = 1$ y $\langle e_j, e_k\rangle = 0$ si $j \neq k$).

[6] En honor al matemático francés Charles Hermite (1822–1901).
[7] En honor al matemático alemán David Hilbert (1862–1943).

2. una base $\mathscr{B} = (e_1, \ldots, e_n)$ es una **base ortonormal**[8] de V si $\mathscr{B}$ es una familia ortonormal.

Lema 7.4.3. — *Sea $V \cong \mathbb{C}^n$ un espacio hermitiano. Entonces, toda familia ortonormal de vectores es linealmente independiente. En particular, toda familia ortonormal $(e_1, \ldots, e_n)$ de cardinal $n = \dim_{\mathbb{C}}(V)$ es una base ortonormal de V.*

Demostración. — Sea $(e_j)_{j \in J}$ una familia ortonormal y supongamos que existe una relación lineal

$$\lambda_1 e_{j_1} + \ldots + \lambda_r e_{j_r} = 0$$

donde $j_1, \ldots, j_r \in J$ son índices distintos y $\lambda_1, \ldots, \lambda_r \in \mathbb{C}$. Al considerar el producto escalar de la igualdad anterior con el vector e_{j_k}, donde $k \in \{1, \ldots, r\}$, obtenemos

$$\langle e_{j_k}, \lambda_1 e_{j_1} + \ldots + \lambda_r e_{j_r} \rangle = \overline{\lambda_1} \langle e_{j_k}, e_{j_1} \rangle + \ldots + \overline{\lambda_r} \langle e_{j_k}, e_{j_r} \rangle.$$

Como $\langle e_{j_k}, e_{j_\ell} \rangle = 0$ si $k \neq \ell$, se tiene que $0 = \overline{\lambda_k} \langle e_{j_k}, e_{j_k} \rangle = \overline{\lambda_k}$, y luego $\lambda_k = 0$ para todo $k \in \{1, \ldots, r\}$. En otras palabras, $(e_j)_{j \in J}$ es una familia linealmente independiente. □

Teorema 7.4.4 **(existencia de bases ortonormales)**
Sea $V \cong \mathbb{C}^n$ un espacio hermitiano. Entonces V admite una base ortonormal.

Demostración. — Esto es consecuencia directa del Teorema de Sylvester en el caso hermitiano, y del hecho que la signatura de la forma cuadrática hermitiana asociada al producto escalar es de signatura $(n, 0)$. □

Observación 7.4.5 **(sub-espacios).** — Si $V \cong \mathbb{C}^n$ es un espacio hermitiano y $U \subseteq V$ es un sub-espacio vectorial. Entonces la restricción a U del producto escalar $\langle \cdot, \cdot \rangle|_U : U \times U \to \mathbb{C}$ es un producto escalar en U. Luego, todo sub-espacio de un espacio hermitiano es también un espacio hermitiano.

Teorema 7.4.6 **(proyección ortogonal).** — *Sea $V \cong \mathbb{C}^n$ un espacio hermitiano y sea $U \subseteq V$ un sub-espacio vectorial. Si denotamos por $U^\perp$ el ortogonal de U respecto a la forma hermitiana dada por el producto escalar de V, entonces:*

[8] Notar que la definición de ortogonalidad en el caso complejo es exactamente idéntica a la definición en el caso real

1. *Se tiene que* $V = U \oplus U^{\perp}$. *Más aún, la aplicación* $p_U : V \to V$ *definida por esta descomposición es lineal, donde* $\mathrm{Im}(p_U) = U$ *y* $\ker(p_U) = U^{\perp}$. *Decimos que* $p_U : V \to V$ *es la* **proyección ortogonal** *sobre* U.
2. *Sea* $(e_1, \ldots, e_r)$ *una base* ortonormal *de* U, *donde* $r = \dim_{\mathbb{C}}(U)$. *Entonces, para todo* $x \in V$ *se tiene que*
$$p_U(x) = \langle x, e_1 \rangle e_1 + \ldots + \langle x, e_r \rangle e_r,$$
donde los coeficientes $\langle x, e_j \rangle \in \mathbb{C}$ *son llamados los* **coeficientes de Fourier** *de* $x \in V$ *respecto a la base* $(e_1, \ldots, e_r)$.
3. *Se tiene que* $(U^{\perp})^{\perp}$. *En particular, la proyección ortonogonal* $p_{U^{\perp}} : V \to V$ *sobre* $U^{\perp}$ *coincide con* $\mathrm{id}_V - p_U$, *i.e.*, $\mathrm{id}_V = p_U + p_{U^{\perp}}$.

Demostración. — Gracias al Teorema 7.2.4, y dado que el producto escalar complejo es una forma hermitiana no-degenerada, tenemos que $U = (U^{\perp})^{\perp}$. Además, dado que no hay vectores isótropos no-nulos pues la forma cuadrática hermitiana asociada es definida positiva, tenemos que $V = U \oplus U^{\perp}$. Así, todo vector $x \in V$ se escribe de manera única como $x = a + b$, donde $a \in U$ y $b \in U^{\perp}$. Luego, si definimos $p_U(x) = a$ entonces $p_U : V \to V$ es lineal y verifica por construcción $\mathrm{Im}(p_U) = U$ y $\ker(p_U) = U^{\perp}$. Más aún, dado que $(U^{\perp})^{\perp} = U$, tenemos que $p_{U^{\perp}}$ está definido por $p_{U^{\perp}}(x) = b$ y luego $\mathrm{id}_V = p_U + p_{U^{\perp}}$. Con esto hemos probado (1) y (3).

Para probar (2) escribamos provisoriamente $a := \langle x, e_1 \rangle e_1 + \ldots + \langle x, e_r \rangle e_r \in U$. Entonces, para todo $j \in \{1, \ldots, r\}$ tenemos que
$$\langle x - a, e_j \rangle = \langle x, e_j \rangle - \left\langle \sum_{k=1}^{r} \langle x, e_k \rangle e_k, e_j \right\rangle = \langle x, e_j \rangle - \sum_{k=1}^{r} \langle x, e_k \rangle \underbrace{\langle e_k, e_j \rangle}_{\delta_{jk}} = \langle x, e_j \rangle - \langle x, e_j \rangle = 0,$$
por lo que $x = a + (x - a)$, donde $a \in U$ y $(x - a) \in U^{\perp}$. Dado que $V = U \oplus U^{\perp}$, la escritura anterior es única y luego $p_U(x) = a = \langle x, e_1 \rangle e_1 + \ldots + \langle x, e_r \rangle e_r$. □

Al igual que en el caso de espacios euclideanos, el producto escalar complejo nos permite definir una noción de *distancia* entre vectores en un espacio hermitiano. Mejor aún, nos permite definir una *norma*.

***Definición 7.4.7* (norma).** — Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo. Una **norma** $\| \cdot \|$ en V es una función $\| \cdot \| : V \to \mathbb{R}^{\geq 0}$, $x \mapsto \|x\|$ verificando las propiedades siguientes:

1. $\|x\| = 0$ si y sólo si $x = 0$.
2. $\|\lambda x\| = |\lambda| \|x\|$ para todos $x \in V$ y $\lambda \in \mathbb{C}$.

3. Para todos $x, y \in V$ se tiene que $\|x + y\| \leq \|x\| + \|y\|$ («desigualdad triangular»).

Un **espacio vectorial normado** es un espacio vectorial dotado de una norma.

Teorema 7.4.8 **(desigualdad de Cauchy-Schwarz y norma hermitiana)** *Sea $V \cong \mathbb{C}^n$ un espacio hermitiano, y sea $Q(x) = \langle x, x\rangle$ la forma cuadrática hermitiana asociada al producto escalar complejo. Entonces, para todos $x, y \in V$ se tiene la* desigualdad de Cauchy-Schwarz

$$|\langle x, y\rangle|^2 \leq Q(x)Q(y),$$

con igualdad si y sólo si x e y son colineales (i.e., son linealmente dependientes). En consecuencia, la aplicación $x \mapsto \|x\| := \sqrt{Q(x)} = \sqrt{\langle x, x\rangle}$ define una norma en V, llamada la **norma hermitiana** *asociada al producto escalar $\langle\cdot,\cdot\rangle$. Luego, la desigualdad de Cauchy-Schwarz se reescribe como*

$$|\langle x, y\rangle| \leq \|x\|\|y\|$$

para todos $x, y \in V$.

Demostración. — Ya observamos anteriormente que la forma cuadrática hermitiana $Q(x) = \langle x, x\rangle$ es también la forma cuadrática *real* asociada a la forma $\mathbb{R}$-bilineal simétrica $B(x, y) = \Re\langle x, y\rangle$ en $V \cong \mathbb{R}^{2n}$. Para $x, y \in V \cong \mathbb{R}^{2n}$ (visto como espacio vectorial real), la desigualdad de Cauchy-Schwarz real implica que

$$|\Re\langle x, y\rangle| \leq \sqrt{Q(x)}\sqrt{Q(y)}.$$

Consideremos la escritura del número complejo $z = \langle x, y\rangle \in \mathbb{C}$ en su forma polar:

$$\langle x, y\rangle = re^{i\theta},\ r = |\langle x, y\rangle|.$$

Entonces, dado que $e^{-i\theta} = \overline{e^{i\theta}}$ y $r \in \mathbb{R}$, se tiene que

$$r = e^{-i\theta}\langle x, y\rangle = \langle x, e^{i\theta}y\rangle = \Re\langle x, e^{i\theta}y\rangle,$$

por lo que

$$|\langle x, y\rangle| = |\Re\langle x, e^{i\theta}\rangle| \leq \sqrt{Q(x)}\sqrt{Q(e^{i\theta}y)} = \sqrt{Q(x)}\sqrt{|e^{i\theta}|^2Q(y)} = \sqrt{Q(x)}\sqrt{Q(y)}.$$

Finalmente, por el cálculo anterior y por el caso real de la desigualdad de Cauchy-Schwarz, tenemos la igualdad ocurre si y sólo si x e $e^{i\theta}y$ son $\mathbb{R}$-linealmente dependientes. Esto implíca en particular que x e y son $\mathbb{C}$-linealmente dependientes. Recíprocamente, si x e y son $\mathbb{C}$-linealmente dependientes y $x \neq 0$, entonces $y = \lambda x$ para cierto $\lambda \in \mathbb{C}$ y por ende

$$\langle x, y\rangle = \overline{\lambda}\langle x, x\rangle = \overline{\lambda}Q(x), \quad \sqrt{Q(x)}\sqrt{Q(y)} = |\lambda|\sqrt{Q(x)}\sqrt{Q(x)} = |\lambda|Q(x),$$

de donde $|\langle x, y\rangle = \sqrt{Q(x)}\sqrt{Q(y)}$ en este caso. El caso $x = 0$ es similar, pero más simple.

Finalmente, dado que para todo número complejo $z = a + ib$ se tiene $\Re(z) = a \leq \sqrt{a^2 + b^2} = |z|$, para todos $x, y \in V$ tenemos por la fórmula de polarización y por la desigualdad de Cauchy-Schwarz que:

$$\|x+y\|^2 = \|x\|^2+\|y\|^2+2\Re\langle x, y\rangle \leq \|x\|^2+\|y\|^2+2|\langle x, y\rangle| \leq \|x\|^2+\|y\|^2+2\|x\|\|y\| = (\|x\|^2+\|y\|^2),$$

de donde se deduce la desigualdad triangular. □

Ejemplo: Sea $V = \mathbb{C}^n$. La norma hermitiana asociada al producto escalar canónico de $\mathbb{C}^n$ está dada por

$$\|(z_1, \ldots, z_n)\| = \sqrt{|z_1|^2 + \ldots + |z_n|^2},$$

donde $(z_1, \ldots, z_n) \in \mathbb{C}^n$. En particular, si $(w_1, \ldots, w_n) \in \mathscr{C}^n$ es otro vector, entonces la desigualdad de Cauchy-Schwarz se escribe como

$$|z_1\overline{w_1} + \ldots + z_n\overline{w_n}| \leq \sqrt{\sum_{j=1}^{n} |z_j|^2} \cdot \sqrt{\sum_{j=1}^{n} |w_j|^2}.$$

Observación 7.4.9. — Vale la pena reescribir con la nueva notación las fórmulas de polarización discutidas anteriormente en Lema 7.1.18. Explícitamente, si $V \cong \mathbb{C}^n$ es un espacio hermitiano y $x, y \in V$, entonces:

1. $\Re\langle x, y\rangle = \frac{1}{2}(\|x + y\|^2 - \|x\|^2 - \|y\|^2) = \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2)$.
2. $\Im\langle x, y\rangle = \frac{1}{2}(\|x + iy\|^2 - \|x\|^2 - \|y\|^2) = \frac{1}{4}(\|x + iy\|^2 - \|x - iy\|^2)$.

Luego,

$$\langle x, y\rangle = \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2 + i\|x + iy\|^2 - i\|x - iy\|^2).$$

Ejercicio 7.4.10. — Sea $V \cong \mathbb{C}^n$ un espacio hermitiano. Probar que si $x_1, \ldots, x_m \in V$ son ortogonales, entonces se verifica el «Teorema de Pitágoras»:

$$\|x_1 + \ldots + x_m\|^2 = \|x_1\|^2 + \ldots + \|x_m\|^2.$$

Tal como en el caso euclideano, una clase importante de aplicaciones lineales son aquellas que preservan la norma hermitiana (o equivalentemente, el producto escalar complejo).

***Proposición 7.4.11* (isometrías).** — *Sean $V \cong \mathbb{C}^n$ un espacio hermitiano, y sea $u : V \to V$ un endomorfismo. Las siguientes propiedades son equivalentes:*

1. *u preserva la norma, i.e., para todo $x \in V$ se tiene que*
$$\|x\| = \|u(x)\|.$$
2. *u preserva el producto escalar, i.e., para todos $x, y \in V$ se tiene que*
$$\langle x, y\rangle = \langle u(x), u(y)\rangle.$$
3. Para toda *base ortonormal $\mathscr{B}$ de V, la imagen $u(\mathscr{B})$ es una base ortonormal de V.*
4. Existe *una base ortonormal $\mathscr{B}$ de V tal que la imagen $u(\mathscr{B})$ es una base ortonormal de V.*
5. *La matriz $A = \mathrm{Mat}_{\mathscr{B}}(u)$ respecto a una base* ortonormal *$\mathscr{B}$ de V es una* **matriz unitaria**, *i.e., $A^*A = \mathrm{I}_n$ (o equivalentemente, ${}^{\mathrm{t}}A\overline{A} = \mathrm{I}_n$).*

Un endomorfismo $u : V \to V$ que cumple estas propiedades es necesariamente biyectivo (i.e., $u \in \mathrm{GL}(V)$), y es llamado una **isometría** *de V. En particular, el conjunto de isometrías (lineales) de un espacio hermitiano $V \cong \mathbb{C}^n$ se identifica naturalmente al grupo unitario $\mathbf{U}(n)$.*

Demostración. — Vimos que (1) y (2) son equivalentes gracias a la fórmula de polarización hermitiana. Por otra parte, claramente (2) implica (3), y (3) implica (4). Veamos que (4) implica (1):

Sea $\mathscr{B} = (e_1, \ldots, e_n)$ una base ortonormal tal que $u(\mathscr{B}) = (u(e_1), \ldots, u(e_n))$ es una base ortonormal de V. Sea $x \in V$ y escribamos $x = \sum_{j=1}^n x_j e_j$. Entonces, $u(x) = \sum_{j=1}^n x_j u(e_j)$ y, dado que tanto $(e_1, \ldots, e_n)$ como $(u(e_1), \ldots, u(e_n))$ son bases ortonormales de V, tenemos que los coeficientes de Fourier respecto a dichas bases verifican
$$x_j = \langle x, e_j\rangle \quad \text{y} \quad x_j = \langle u(x), u(e_j)\rangle,$$
por lo que $\langle u(x), u(e_j)\rangle = \langle x, e_j\rangle$ para todo $j \in \{1, \ldots, n\}$. En particular, el Teorema de Pitágoras implica
$$\|x\|^2 = \sum_{j=1}\langle x, e_j\rangle^2 = \sum_{j=1}\langle u(x), u(e_j)\rangle^2 = \|u(x)\|^2,$$
de donde obtenemos $\|u(x)\| = \|x\|$ y por ende (1) se verifica.

Finalmente, sabemos por la Proposición 7.3.4 que un endomorfismo $u : V \to V$ preserva la forma hermitiana *no-degenerada* $h = \langle \cdot, \cdot\rangle$ (y luego, u es automáticamente biyectivo) si y sólo si ${}^{\mathrm{t}}AH\overline{A} = H$, donde $H = (h_{jk})$ con $h_{jk} = \langle e_j, e_k\rangle = \delta_{jk}$ si $\mathscr{B} = (e_1, \ldots, e_n)$ es una base ortonormal. En otras palabras, $H = \mathrm{I}_n$, y luego (1) es equivalente a (5). $\square$

Como consecuencia inmediata, obtenemos las siguientes caracterizaciones diferentes del grupo unitario.

Corolario 7.4.12. — *Sea $V \cong \mathbb{C}^n$ un espacio hermitiano y sea $\mathscr{B}$ una base ortonormal de V. Entonces, el grupo unitario de isometrías de V es isomorfo a*

$$\begin{aligned}\mathbf{U}(n) &= \{A \in M_n(\mathbb{C}) \mid A^*A = I_n\} = \{A \in M_n(\mathbb{C}) \mid {}^{\mathrm{t}}A\overline{A} = I_n\} = \{A \in \mathrm{GL}_n(\mathbb{C}) \mid A^{-1} = A^*\} \\ &= \{A \in M_n(\mathbb{C}) \mid \langle AX, AY\rangle = \langle X, Y\rangle \text{ para todos } X, Y \in \mathbb{C}^n\} \\ &= \{A \in M_n(\mathbb{C}) \mid \|AX\| = \|X\| \text{ para todo } X \in \mathbb{C}^n\} \\ &= \{A \in M_n(\mathbb{C}) \mid (Av_1, \dots, Av_n) \text{ es base ortonormal para toda base ortonormal } (v_1, \dots, v_n) \text{ de } \mathbb{C} \\ &= \{A \in M_n(\mathbb{C}) \mid (Ae_1, \dots, Ae_n) \text{ es base ortonormal, donde } (e_1, \dots, e_n) \text{ es la base canónica de } \mathbb{C} \\ &= \{A \in M_n(\mathbb{C}) \mid \text{las columnas de } A \text{ son ortonormales}\},\end{aligned}$$

donde $\langle \cdot, \cdot \rangle$ y $\|\cdot\|$ denotan el producto escalar y la norma hermitiana canónicos en $\mathscr{C}^n$. □

7.5. Endomorfismo adjunto y endomorfismos normales

Comencemos por introducir la noción de *adjunto* de un endomorfismo.

***Teorema 7.5.1* (adjunto).** — *Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo y sea $h : V \times V \to \mathbb{C}$ una forma hermitiana* no-degenerada. *Entonces, para todo endomorfismo $u : V \to V$ existe un único endomorfismo $u^* : V \to V$ que verifica*

$$h(u(x), y) = h(x, u^*(y)) \text{ para todos } x, y \in V,$$

llamado el **adjunto** *de u (respecto a h). Más aún, para toda base $\mathscr{B}$ de V se tiene que*

$$\mathrm{Mat}_{\mathscr{B}}(u^*) = \overline{H^{-1}}\,\mathrm{Mat}_{\mathscr{B}}(u)^*\overline{H},$$

donde $H = \mathrm{Mat}_{\mathscr{B}}(h)$ es la matriz de h respecto a $\mathscr{B}$. En particular, $(u^*)^* = u$.

Demostración. — Supongamos que existe un endomorfismo $u^* : V \to V$ verificando $h(u(x), y) = h(x, u^*(y))$ para todos $x, y \in V$. Sea $\mathscr{B}$ una base de V y denotemos $H = \mathrm{Mat}_{\mathscr{B}}(h)$, $A = \mathrm{Mat}_{\mathscr{B}}(u)$ y $B = \mathrm{Mat}_{\mathscr{B}}(u^*)$. Sean $x, y \in V$ arbitrarios y denotemos por $X, Y \in \mathbb{C}^n$ los vectores columnas asociadas (i.e., los vectores coordenadas respecto a la base $\mathscr{B}$). Entonces,

$$h(u(x), y) = {}^{\mathrm{t}}(AX)H\overline{Y} = {}^{\mathrm{t}}X\,{}^{\mathrm{t}}AH\overline{Y} = h(x, u^*(y)) = {}^{\mathrm{t}}XH\overline{BY},$$

por lo que ${}^{\mathrm{t}}AH = H\overline{B}$. Así, dado que H es invertible (puesto que la forma hermitiana h es no-degenerada), tenemos que $B = \overline{H^{-1}\,{}^{\mathrm{t}}AH} = \overline{H^{-1}}A^*\overline{H}$, donde por definición $A^* = {}^{\mathrm{t}}\overline{A}$ es la matriz adjunta de A. Esto último muestra que u^*, en caso de existir, verifica la relación

$$\mathrm{Mat}_{\mathscr{B}}(u^*) = \overline{H^{-1}}\,\mathrm{Mat}_{\mathscr{B}}(u)^*\overline{H}$$

y por ende está únicamente determinado por u y por h. Recíprocamente, y con la notación anterior, si denotamos por $u^* : V \to V$ el endomorfismo de V cuya matriz respecto a la base $\mathscr{B}$ está dada por $\overline{H^{-1}}A^*\overline{H}$ entonces para todos $x, y \in V$ se tiene que

$$h(x, u^*(y)) = {}^{\mathrm{t}}XH\overline{BY} = {}^{\mathrm{t}}XH(\overline{\overline{H^{-1}}A^*\overline{H}})\overline{Y} = {}^{\mathrm{t}}XHH^{-1}\overline{A^*}H\overline{Y} = {}^{\mathrm{t}}X\,{}^{\mathrm{t}}AH\overline{Y} = {}^{\mathrm{t}}(AX)H\overline{Y} = h(u(x), y),$$

por lo que $u^* : V \to V$ verifica la relación pedida. Esto demuestra la existencia y unicidad del endomorfismo adjunto. Finalmente, notamos que para todos $x, y \in V$ se tiene que

$$h(u^*(x), y) = \overline{h(y, u^*(x))} = \overline{h(u(y), x)} = h(x, u(y))$$

y por ende u es el adjunto de u^*, i.e., $(u^*)^* = u$. □

Ejercicio 7.5.2. — Sea $V \cong \mathbb{C}^n$ un espacio vectorial complejo. Probar que la aplicación

$$\mathrm{End}_{\mathbb{C}}(V) \longrightarrow \mathrm{End}_{\mathbb{C}}(V),\ u \longmapsto u^*$$

es anti-lineal.

En el caso de espacios hermitianos el teorema anterior se reescribe de la manera siguiente.

Teorema 7.5.3 **(adjunto en espacio hermitiano)**
Sea $V \cong \mathbb{C}^n$ un espacio hermitiano. Entonces, para todo endomorfismo $u : V \to V$ existe un único endomorfismo $u^* : V \to V$ que verifica

$$\langle u(x), y\rangle = \langle x, u^*(y)\rangle \text{ para todos } x, y \in V,$$

llamado el **adjunto** *de u. Más aún, para toda* base ortonormal *$\mathscr{B}$ de V se tiene que*

$$\mathrm{Mat}_{\mathscr{B}}(u^*) = \mathrm{Mat}_{\mathscr{B}}(u)^* = {}^{\mathrm{t}}\overline{\mathrm{Mat}_{\mathscr{B}}(u)}.$$

Demostración. — Si $\mathscr{B}$ es una base ortonormal de V, entonces la matriz H de la forma hermitiana $h = \langle\cdot,\cdot\rangle$ es la identidad I_n, de donde obtenemos el resultado. □

El siguiente lema sobre estabilidad de sub-espacios vectoriales será útil en lo que sigue. Recordemos que si $u : V \to V$ es un endomorfismo de un espacio vectorial V, entonces un sub-espacio vectorial $U \subseteq V$ es **estable** por u si $u(U) \subseteq U$, i.e., si para todo $x \in U$ se tiene que $u(x) \in U$.

***Lema 7.5.4* (estabilidad). —** *Sea $V \cong \mathbb{C}^n$ un espacio hermitiano y sea $u : V \to V$ un endomorfismo. Entonces, para todo sub-espacio vectorial $U \subseteq V$ se tiene que*

$$u(U) \subseteq U \Leftrightarrow u^*(U^\perp) \subseteq U^\perp.$$

En otras palabras, U es estable por u si y sólo si $U^\perp$ es estable por u^*.

Demostración. — Supongamos que $u(U) \subseteq U$ y sea $y \in U^\perp$. Entonces, para todo $x \in U$ se tiene

$$\langle x, u^*(y)\rangle = \langle u(x), y\rangle = 0,$$

donde la última igualdad se obtiene gracias a que $u(x) \in U$ e $y \in U^\perp$, por lo que $u^*(y) \in U^\perp$. Así, $u^*(U^\perp) \subseteq U^\perp$. Recíprocamente, supongamos que $u^*(U^\perp) \subseteq U^\perp$ y, por el mismo argumento anterior pero reemplazando u por u^* y U por $U^\perp$, tenemos en este caso que $(u^*)^*((U^\perp)^\perp) \subseteq (U^\perp)^\perp$. Dado que $(U^\perp)^\perp = U$ y $(u^*)^*$=u, la inclusión $(u^*)^*((U^\perp)^\perp) \subseteq (U^\perp)^\perp$ es equivalente a $u(U) \subseteq U$. □

Las siguientes clases de endomorfismos de un espacio hermitiano son especialmente importantes.

***Definición 7.5.5* (endomorfismo normal). —** Sea $V \cong \mathbb{C}^n$ un espacio hermitiano, y sea $u : V \to V$ un endomorfismo. Diremos que u es:

1. **auto-adjunto** (o **hermitiano**) si $u^* = u$.
2. **anti-hermitiano** si $u^* = -u$.
3. **unitario** si $u \circ u^* = \mathrm{id}_V$ (i.e., $u \in \mathrm{GL}(V)$ y $u^{-1} = u^*$).

De manera más general, diremos que u es un endomorfismo **normal** si conmuta con su adjunto u^*, i.e., $u \circ u^* = u^* \circ u$. En particular, esta noción engloba los tres casos precedentes.

En dimensión finita, el *espectro* de un endomorfismo es por definición el conjunto de sus valores propios. La *teoría espectral* es el estudio de valores propios y vectores propios de endomorfismos (eventualmente generalizando resultados de dimensión finita a dimensión infinita, para poder tomar en cuenta *operadores*, tales como el operador de Schrödinger en mecánica cuántica). En este texto introductorio, nos contentaremos con estudiar la teoría espectral de endomorfismos normales en dimensión finita (que cubre en particular el caso de endomorfismos simétricos y anti-simétricos reales, hermitianos y anti-hermitianos complejos, unitarios y ortogonales reales). El siguiente resultado es conocido como **teorema espectral**.

Teorema 7.5.6. — *Sea $V \cong \mathbb{C}^n$ un espacio hermitiano y sea $u : V \to V$ un* endomorfismo normal. *Entonces, u es diagonalizable y los espacios propios asociados a valores propios distintos son ortogonales. En particular, existe una base ortonormal $\mathscr{B}$ de V formada por vectores propios de u. Más aún:*

1. *Si u es* hermitiano*, entonces los valores propios de u son* reales *$\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ y luego*
$$\mathrm{Mat}_{\mathscr{B}}(u) = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix} \in M_n(\mathbb{R}).$$
2. *Si u es* anti-hermitiano*, entonces los valores propios de u son* imaginarios puros *$i\lambda_1, \ldots, i\lambda_n \in i\mathbb{R}$ y luego*
$$\mathrm{Mat}_{\mathscr{B}}(u) = \begin{pmatrix} i\lambda_1 & 0 & \ldots & 0 \\ 0 & i\lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & i\lambda_n \end{pmatrix} \in M_n(\mathbb{C}).$$
3. *Si u es* unitario*, entonces los valores propios de u son* de módulo 1*, i.e., $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ con $|\lambda_j| = 1$, y luego*
$$\mathrm{Mat}_{\mathscr{B}}(u) = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix} \in \mathrm{GL}_n(\mathbb{C}).$$

Demostración. — La demostración es por inducción en la dimensión $n = \dim_{\mathbb{C}}(V)$. Notamos que el resultado es cierto si $n = 1$, por lo que podemos suponer que $n \geq 2$ y que el resultado se verifica para $n-1$. Dado que $\mathbb{C}$ es algebraicamente cerrado, el polinomio característico $P_u(X)$ posee al menos una raíz λ en $\mathbb{C}$, y λ es un valor propio de u.

Sea V_λ el espacio propio asociado, el cual es *estable* por u^* (i.e., $u^*(V_\lambda) \subseteq V_\lambda$). En efecto, dado que u y u^* conmutan, tenemos que para todo $x \in V_\lambda$ se cumple que
$$u(u^*(x)) = u^*(u(x)) = u^*(\lambda x) = \lambda u^*(x),$$

por lo que $u^*(x) \in V_\lambda$. Luego, gracias al lema de estabilidad (ver Lema 7.5.4) tenemos que el ortogonal $V_\lambda^\perp$ es estable tanto por u como por u^*. Por otro lado, dado que $V = V_\lambda \oplus V_\lambda^\perp$ y $\dim_{\mathbb{C}}(V_\lambda) > 0$ (por definición de espacio propio), tenemos que $\dim_{\mathbb{C}}(V_\lambda^\perp) = \dim_{\mathbb{C}}(V) - \dim_{\mathbb{C}}(V_\lambda) < \dim_{\mathbb{C}}(V) = n$.

Denotemos por $v := u|_{V_\lambda^\perp} : V_\lambda^\perp \to V_\lambda^\perp$ la restricción de u a $V_\lambda^\perp$, y por v^* la restricción de u^* a $V_\lambda^\perp$. Entonces, para todos $x, y \in V_\lambda^\perp$ tenemos que $v(x) = u(x)$ y $v^*(y) = u^*(y)$, por lo que

$$\langle v(x), y\rangle = \langle u(x), y\rangle = \langle x, u^*(y)\rangle = \langle x, v^*(y)\rangle,$$

lo que muestra que el adjunto de v es v^*, vistos como endomorfismos de $V_\lambda^\perp$. Dado que u y u^* conmutan, las restricciones v y v^* también conmutan, i.e., v es un endomorfismo normal de $V_\lambda^\perp$. Luego, la hipótesis de inducción implica que v es diagonalizable y que espacios propios asociados a valores propios distintos son ortogonales. Dado que $V = V_\lambda \oplus V_\lambda^\perp$, lo anterior también se cumple para u. En particular, considerando bases ortonormales de cada espacio propio de u, obtenemos que existe una base ortonormal $\mathscr{B}$ de V formada por vectores propios de u.

Finalmente, si consideramos la matriz diagonal $D = \mathrm{Mat}_{\mathscr{B}}(u)$ con coeficientes diagonales $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$, entonces se tiene que si u es además

1. hermitiano, entonces $D^* = \overline{D} = D$, i.e., $\lambda_j \in \mathbb{R}$ para todo $j \in \{1, \ldots, n\}$.
2. anti-hermitiano, entonces $D^* = \overline{D} = -D$, i.e., $\lambda_j \in \mathbb{R}$ para todo $j \in \{1, \ldots, n\}$.
3. unitario, entonces $D^*D = \overline{D}D = \mathrm{I}_n$, i.e., $|\lambda_j| = 1$ para todo $j \in \{1, \ldots, n\}$.

de donde se prueba el resultado. □

Observación 7.5.7. — Tal como mencionamos anteriormente, el teorema espectral anterior cubre también el caso de matrices *reales* simétricas, anti-simétricas y ortogonales. En efecto, si $A \in M_n(\mathbb{R})$ es una matriz real y la pensamos como elemento de $M_n(\mathbb{C})$, entonces $A^* = {}^t\overline{A} = {}^tA$. Por lo que A es hermitiana (resp. anti-hermitiana, resp. unitaria) si y sólo si es simétrica (resp. anti-simétrica, resp. ortogonal).

CAPÍTULO 8

TENSORES

En este capítulo se dará una introducción al *álgebra multilineal* o, en términos modernos, al *álgebra tensorial*. Los tensores son muy útiles en matemática y en física, y fueron introducidos en 1846 por el matemático irlandés William Hamilton. La idea de los tensores es que, dado un cuerpo k, ellos permiten describir aplicaciones multilineales (i.e., lineales en cada variable) $T : V_1 \times \cdots \times V_p \to W$, donde $V_1, \dots, V_p, W$ son k-espacios vectoriales.

Durante todo este capítulo fijamos k un cuerpo arbitrario.

8.1. Producto tensorial de espacios vectoriales

Comencemos por el caso "2-multilineal", i.e., **bilineal**. La siguiente definición es una pequeña generalización de la noción de forma bilineal (ver Definición 5.3.1).

Definición 8.1.1. — Sean V, W y U tres k-espacios vectoriales. Una **aplicación bilineal** es una aplicación

$$B : V \times W \to U$$

que verifica

1. La aplicación $B^y : V \to U, \quad x \mapsto B(x, y)$ es una aplicación lineal $\forall y \in W$.
2. La aplicación ${}^xB : W \to U, \quad y \mapsto B(x, y)$ es una aplicación lineal $\forall x \in V$.

Observación 8.1.2. — En particular, una forma bilineal sobre un k-espacio vectorial V es una aplicación bilineal $B : V \times W \to U$ donde $W = V$ y $U = k$.

En términos prácticos, un *producto tensorial* entre dos k-espacios vectoriales V y W es un tercer espacio vectorial que cumple que toda aplicación <u>bilineal</u> cuyo dominio es $V \times W$ es equivalente a una aplicación <u>lineal</u> cuyo dominio es dicho producto tensorial entre V y W. Más formalmente, tenemos la siguiente definición.

***Definición 8.1.3* (producto tensorial). —** Sean V y W dos k-espacios vectoriales. Un **producto tensorial** de V y W es un par (T, t), donde T es un k-espacio vectorial y $t : V \times W \to T$ es una aplicación <u>bilineal</u> verificando la siguiente *propiedad universal*:

> Para todo k-espacio vectorial U y toda aplicación bilineal $B : V \times W \to U$, existe una *única* aplicación <u>lineal</u> $\widehat{B} : T \to U$ tal que $B = \widehat{B} \circ t$.

En otras palabras, el diagrama siguiente

$$\begin{tikzcd} V \times W \arrow[rr, "B"] \arrow[dr, "t"'] & & U \\ & T \arrow[ur, "\exists! \widehat{B}"'] & \end{tikzcd}$$

es conmutativo.

Una pregunta natural que surge de la discusión anterior es si los productos tensoriales existen y, en caso de existir, si estos son únicos. El siguiente resultado da una respuesta afirmativa a ambas preguntas. Como podrá apreciarse en la demostración, la construcción no es para nada explícita (y puede ser pasada por alto en una primera lectura sin mayor problema). Sin embargo, es importante desde un punto de vista filosófico pues nos dice por una parte que es más importante comprender la *propiedad universal* del producto tensorial que el espacio vectorial en si mismo, y por otra parte ejemplifica la importancia de los espacios vectoriales cocientes: estos nos permiten construir objetos imponiendo propiedades que queramos que se cumplan.

***Teorema 8.1.4.* —** *Sean V y W dos k-espacios vectoriales. Entonces, existe un producto tensorial (T, t) de V y W, el cuál es único módulo un único isomorfismo. Así, denotaremos $T =: V \otimes_k W$ (o símplemente*[1] *$V \otimes W$) y para $v \in V$ y $w \in W$ escribimos $t(v, w) =: v \otimes w$.*

[1] Siempre y cuando k sea claro en el contexto. Veremos más adelante que esto puede ser relevante en ocasiones.

Demostración. — Comencemos por probar la existencia. Para ello, consideremos el k-espacio vectorial $E := k^{(V\times W)}$ cuya base (no necesariamente finita) está dada por $\{e_{(v,w)}\}_{(v,w)\in V\times W}$, i.e., un elemento de E es de la forma

$$\sum_{\text{finita}} \lambda_{(v,w)} e_{(v,w)}, \text{ donde } \lambda_{(v,w)} \in k.$$

Es importante notar que la aplicación $V \times W \to E,\ (v,w) \mapsto e_{(v,w)}$ **no** es bilineal. Sin embargo, si $S \subseteq E$ es el k-subespacio vectorial generado por todos los vectores de la forma

$$e_{(v+v',w)} - e_{(v,w)} - e_{(v',w)}, \quad e_{(v,w+w')} - e_{(v,w)} - e_{(v,w')}, \quad e_{(\lambda v,w)} - \lambda e_{(v,w)}, \quad e_{(v,\lambda w)} - \lambda e_{(v,w)},$$

donde $v, v' \in V,\ w, w' \in W$ y $\lambda \in k$, entonces el k-espacio vectorial cociente $T := E/S$ está dotado de la aplicación *bilineal*

$$t : V \times W \longrightarrow T,\ (v,w) \mapsto [e_{(v,w)}]$$

que hace de (T, t) un producto tensorial de V y W. En efecto, si denotamos $T := V \otimes W$ y $t(v,w) = [e_{(v,w)}] := v \otimes w$, entonces para toda aplicación bilineal $B : V \times W \to U$ se tiene que la aplicación $B' : E \to U,\ e_{(v,w)} \mapsto B(v,w)$ es lineal y, por bilinealidad, se anula en S. Dado que $B'(S) = \{0_U\}$, la propiedad universal del cociente implica que existe una única aplicación lineal $\widehat{B} : E/S = V \otimes W \to U$ tal que $B = \widehat{B} \circ t$, lo cual equivale a decir que para todo $v \otimes w \in T$ se tiene que $\widehat{B}(v \otimes w) = B(v,w)$.

La unicidad (módulo único isomorfismo) es consecuencia de la propiedad universal del par (T,t). En efecto, si (T', t') es otro producto tensorial de V y W, entonces tendríamos dos diagramas conmutativos

$$\begin{tikzcd} V \times W \arrow[rr, "t'"] \arrow[dr, "t"'] & & T' \\ & T \arrow[ur, "\exists!\,\alpha"'] & \end{tikzcd} \qquad \begin{tikzcd} V \times W \arrow[rr, "t"] \arrow[dr, "t'"'] & & T \\ & T' \arrow[ur, "\exists!\,\beta"'] & \end{tikzcd}$$

donde $t' = \alpha \circ t$ y $t = \beta \circ t'$. En particular, $t' = \alpha \circ t = (\alpha \circ \beta) \circ t'$ y $t = \beta \circ t' = (\beta \circ \alpha) \circ t$. Dado que α y β son únicos, tenemos que necesariamente $\alpha \circ \beta = \mathrm{Id}_{T'}$ y $\beta \circ \alpha = \mathrm{Id}_T$, i.e., α y β son isomorfismos. $\square$

Recuerdo 8.1.5. — Sean U, V, W tres k-espacios vectoriales. Recordemos que

$$\mathrm{Hom}_k(V,W) = \{f : V \to W \text{ lineal}\},\ V^* = \mathrm{Hom}_k(V,k) \text{ (espacio dual), y}$$

$$\mathrm{Bil}_k(V \times W, U) = \{B : V \times W \to U \text{ aplicación } k\text{-bilineal}\}$$

son k-espacios vectoriales.

Con la notación anterior, tenemos la siguiente consecuencia inmediata del Teorema anterior.

Corolario 8.1.6. — *Sean U, V, W tres k-espacios vectoriales. Entonces, hay un isomorfismo*

$$\mathrm{Bil}_k(V \times W, U) \cong \mathrm{Hom}_k(V \otimes W, U).$$

En particular, $\mathrm{Bil}_k(V \times W, k) = \{B : V \times W \to k \text{ forma bilineal}\} \cong (V \otimes W)^*$.

Demostración. — La aplicación $B \longmapsto \widehat{B}$ es un isomorfismo. □

El siguiente ejemplo nos permite verificar "a mano" que una formal bilineal codifica la misma información que su forma lineal asociada.

Ejemplo 8.1.7. — Sea $V = \mathbb{R}^2$ y $\mathscr{B} = (e_1, e_2)$ su base canónica. Toda forma bilineal $B : V \times V \to \mathbb{R}$ es de la forma

$$B((x_1, y_1), (x_2, y_2)) = ax_1x_2 + bx_1y_2 + cx_2y_1 + dy_1y_2.$$

En particular, B está determinada por las constantes $a = B(e_1, e_1)$, $b = B(e_1, e_2)$, $c = B(e_2, e_1)$ y $d = B(e_2, e_2)$. Por otro lado, tenemos que

$$V \otimes V = \mathrm{Vect}_{\mathbb{R}}\langle e_1 \otimes e_1, e_1 \otimes e_2, e_2 \otimes e_1, e_2 \otimes e_2\rangle \cong \mathbb{R}^4,$$

donde por definición se cumple que $\lambda(e_i \otimes e_j) = (\lambda e_i) \otimes e_j = e_i \otimes (\lambda e_j)$ para todo $\lambda \in \mathbb{R}$. En particular, tenemos que $(x_ie_i \otimes y_je_j) = x_iy_j(e_i \otimes e_j)$ para todos $x_i, y_j \in \mathbb{R}$. Así, tenemos la igualdad

$$(x_1e_1+y_1e_2)\otimes(x_2e_1+y_2e_2) = x_1x_2(e_1\otimes e_1)+x_1y_2(e_1\otimes e_2)+x_2y_1(e_2\otimes e_1)+y_1y_2(e_2\otimes e_2).$$

Finalmente, observamos que $\widehat{B} : V \otimes V \to \mathbb{R}$ en $(V \otimes V)^*$ está determinada por las constantes $a' = \widehat{B}(e_1 \otimes e_1)$, $b' = \widehat{B}(e_1 \otimes e_2)$, $c' = \widehat{B}(e_2 \otimes e_1)$ y $d' = \widehat{B}(e_2 \otimes e_2)$. Dado que $\widehat{B}(e_i \otimes e_j) = B(e_i, e_j)$, tenemos que $\widehat{B} \in (V \otimes V)^*$ y $B \in \mathrm{Bil}_{\mathbb{R}}(V \times V, \mathbb{R})$ contienen la misma información.

Importante: En general, un vector en $V \otimes W$ **no** es de la forma $v \otimes w$ con $v \in V$ y $w \in W$, sino que es una *suma finita* de vectores de esta forma, i.e.,

$$\lambda_1(v_1 \otimes w_1) + \lambda_2(v_2 \otimes w_2) + \ldots + \lambda_r(v_r \otimes w_r)$$

para ciertos $v_i \in V$, $w_i \in W$ y $\lambda_i \in k$. Los elementos de $V \otimes W$ son llamados **tensores** . Los tensores de la forma $v \otimes w$ con $v \in V$ y $w \in W$ son llamados **tensores simples** o **tensores descomponibles**, estos tensores generan el k-espacio vectorial $V \otimes W$ por construcción.

Ejercicio 8.1.8. — Sea $V \cong \mathbb{R}^2$ de base (e_1, e_2) y $W \cong \mathbb{R}^2$ de base (f_1, f_2). Probar que el tensor

$$T = a(e_1 \otimes f_1) + b(e_1 \otimes f_2) + c(e_2 \otimes f_1) + d(e_2 \otimes f_2) \in V \otimes W$$

es un tensor simple sí y sólo si $ad - bc = 0$.

Lema 8.1.9. — *Sean V y W dos k-espacios vectoriales. Sea $\{v_i\}_{i \in I}$ una base de V y sea $\{w_j\}_{j \in J}$ una base de W. Entonces, $\{v_i \otimes w_j\}_{(i,j) \in I \times J}$ es una base de $V \otimes W$. En particular, si V y W son de dimensión* finita, *entonces*

$$\dim_k(V \otimes W) = \dim_k(V) \cdot \dim_k(W).$$

Demostración. — Por construcción de $V \otimes W$ se tiene que $\{v_i \otimes w_j\}_{(i,j) \in I \times J}$ es una familia generadora. Veamos que es linealmente independiente: Para $i \in I$ y $j \in J$ definimos las formas lineales $f_i \in V^*$ y $g_j \in W^*$ dadas por $f_i(v_k) := \delta_{ik}$ y $g_j(w_k) := \delta_{jk}$, donde

$$\delta_{ij} := \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

Consideremos la forma *bilineal* $\varphi_{ij} : V \times W \longrightarrow k$ dada por $\varphi_{ij}(v, w) = f_i(v) g_j(w)$. La propiedad universal del producto tensorial implica que existe una única aplicación *lineal* $\widehat{\varphi_{ij}} : V \otimes W \longrightarrow k$ tal que

$$\widehat{\varphi_{ij}}(v_k \otimes w_\ell) = \varphi_{ij}(v_k, w_\ell) = f_i(v_k) g_j(w_\ell) = \delta_{ik}\delta_{j\ell} = \begin{cases} 1 & \text{si } i = k \text{ y } j = \ell \\ 0 & \text{sino} \end{cases}$$

Aplicando las formas lineales $\widehat{\varphi_{ij}}$ a cualquier combinación lineal de elementos de $\{v_i \otimes w_j\}_{(i,j) \in I \times J}$, deducimos que dichos vectores son linealmente independientes. $\square$

El siguiente resultado muestra que podemos tensorizar aplicaciones lineales.

***Proposición 8.1.10* (functorialidad).** — *Sean $f : V \longrightarrow V'$ y $g : W \longrightarrow W'$ aplicaciones lineales entre k-espacios vectoriales. Entonces, existe una* única *aplicación lineal*

$$f \otimes g : V \otimes W \longrightarrow V' \otimes W'$$

tal que para todo tensor simple $v \otimes w \in V \otimes W$ se cumple que $(f \otimes g)(v \otimes w) = f(v) \otimes g(w)$. Más aún, el producto tensorial de aplicaciones lineales verifica que

$$(f' \otimes g') \circ (f \otimes g) = (f' \circ f) \otimes (g' \circ g).$$

Demostración. — Basta completar el diagrama conmutativo

$$\begin{tikzcd} V \times W \arrow[r, "{(f,g)}"] \arrow[d, "t"] \arrow[dr, "\text{bilineal}"] & V' \times W' \arrow[d, "t'"] \\ V \otimes W \arrow[r, dotted, "\exists!\, f\otimes g"'] & V' \otimes W' \end{tikzcd}$$

En otras palabras, $f \otimes g$ se obtiene al aplicar la propiedad universal del producto tensorial a la aplicación bilineal $t' \circ (f, g)$. La unicidad, así como la última parte, se deducen del mismo modo y se dejan como Ejercicio al lector. $\square$

A continuación, describimos las principales propiedades del producto tensorial.

Teorema 8.1.11. — *Sean U, V y W tres k-espacios vectoriales. Entonces, hay isomorfismos* canónicos*:*

1. $k \otimes V \xrightarrow{\sim} V,\ \lambda \otimes v \longmapsto \lambda v.$
2. $(U \oplus V) \otimes W \xrightarrow{\sim} (U \otimes W) \oplus (V \otimes W),\ (u + v) \otimes w \longmapsto u \otimes w + v \otimes w.$
3. $U \otimes V \xrightarrow{\sim} V \otimes U,\ u \otimes v \longmapsto v \otimes u.$
4. $U \otimes (V \otimes W) \xrightarrow{\sim} (U \otimes V) \otimes W,\ u \otimes (v \otimes w) \longmapsto (u \otimes v) \otimes w.$

Aquí, cada isomorfismo está descrito para tensores simples *y se extiende por linealidad para tensores arbitrarios (i.e., para sumas finitas de tensores simples).*

Demostración. — Para probar (1), notamos que la aplicación $k \times V \to V,\ (\lambda, v) \mapsto \lambda v$ es bilineal, y por ende existe una única aplicación lineal inducida $k \otimes V \to V,\ \lambda \otimes v \mapsto \lambda v$, cuya inversa está dada por $v \mapsto 1 \otimes v$.

Para probar (2), consideramos la aplicación bilineal

$$(U \oplus V) \times W \to (U \otimes W) \oplus (V \otimes W),\ (u + v, w) \mapsto u \otimes w + v \otimes w$$

que, tal como antes, induce una única aplicación lineal

$$\varphi : (U \oplus V) \otimes W \to (U \otimes W) \oplus (V \otimes W),\ (u + v) \otimes w \mapsto u \otimes w + v \otimes w.$$

Por otro lado, las inclusiones $\iota_1 : U \hookrightarrow U \oplus V,\ u \mapsto u + 0$ y $\iota_2 : V \hookrightarrow U \oplus V,\ v \mapsto 0 + v$ inducen, al tensorizar con $\mathrm{Id}_W : W \to W$, aplicaciones lineales

$$\iota_1 \otimes \mathrm{Id}_W : U \otimes W \to (U \oplus V) \otimes W \quad \text{y} \quad \iota_2 \otimes \mathrm{Id}_W : V \otimes W \to (U \oplus V) \otimes W.$$

Para concluir, basta notar que la aplicación lineal

$$\psi := (\iota_1 \otimes \mathrm{Id}_W, \iota_2 \otimes \mathrm{Id}_W) : (U \otimes W) \oplus (V \otimes W) \longrightarrow (U \oplus V) \otimes W$$

es la inversa de φ. En efecto, basta calcular para tensores simples que

$$(\psi \circ \varphi)((u+v) \otimes w) = \psi(u \otimes w + v \otimes w) = (\iota_1 \otimes \mathrm{Id}_W)(u \otimes w) + (\iota_2 \otimes \mathrm{Id}_W)(v \otimes w)$$
$$= (u+0) \otimes w + (0+v) \otimes w = (u+v) \otimes w,$$

y por ende $\psi \circ \varphi = \mathrm{Id}_{(U \oplus V) \otimes W}$. El mismo cálculo para $\varphi \circ \psi$ permite verificar el isomorfismo. Finalmente, los isomorfismos (3) y (4) se prueban de manera similar. $\square$

***Ejemplo 8.1.12* (Producto de Kronecker).** — Sean V_1, V_2, W_1 y W_2 espacios vectoriales de dimensión finita. Supongamos que $\{v_{1,j}\}_{j \in I_1}$ es una base de V_1, que $\{v_{2,i}\}_{i \in I_2}$ es una base de V_2, que $\{w_{1,l}\}_{l \in J_1}$ es una base de W_1, y que $\{w_{2,k}\}_{k \in J_2}$ es una base de W_2.

Si $f : V_1 \to V_2$ y $g : W_1 \to W_2$ son aplicaciones lineales dadas por matrices $A = (a_{ij})_{i \in I_2,\, j \in I_1}$ y $B = (b_{kl})_{k \in J_2,\, l \in J_1}$ respecto a las bases anteriores, entonces obtenemos por bilinealidad del producto tensorial que

$$(f \otimes g)(v_{1,j} \otimes w_{1,l}) = \sum_{\substack{i \in I_2 \\ k \in J_2}} a_{ij} b_{kl} v_{2,i} \otimes w_{2,k},$$

de tal suerte que la matriz de $f \otimes g$ en la base $\{v_{1,j} \otimes w_{1,l}\}_{(j,l) \in I_1 \times J_1}$ de $V_1 \otimes W_1$ y en la base $\{v_{2,i} \otimes w_{2,k}\}_{(i,k) \in I_2 \times J_2}$ de $V_2 \otimes W_2$ está dada por el **producto de Kronecker**

$$A \otimes B := (a_{ij} b_{kl})_{(i,k) \in I_2 \times J_2,\, (j,l) \in I_1 \times J_1}.$$

Por ejemplo, si todos los espacios vectoriales involucrados son de dimensión 2, entonces tenemos que $A \otimes B$ está dada por

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \otimes \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11}B & a_{12}B \\ a_{21}B & a_{22}B \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} & a_{11}b_{12} & a_{12}b_{11} & a_{12}b_{12} \\ a_{11}b_{21} & a_{11}b_{22} & a_{12}b_{21} & a_{12}b_{22} \\ a_{21}b_{11} & a_{21}b_{12} & a_{22}b_{11} & a_{22}b_{12} \\ a_{21}b_{21} & a_{21}b_{22} & a_{22}b_{21} & a_{22}b_{22} \end{pmatrix}.$$

En particular, notamos que en general $A \otimes B \neq B \otimes A$.

***Ejercicio 8.1.13.* —**

a) Probar que $\mathrm{tr}(A \otimes B) = \mathrm{tr}(A)\,\mathrm{tr}(B)$.
b) Probar que $\mathrm{rango}(A \otimes B) = \mathrm{rg}(A)\,\mathrm{rg}(B)$.
c) Si $A \in M_{a \times a}(k)$ y $B \in M_{b \times b}(k)$, probar que $\det(A \otimes B) = \det(A)^b \det(B)^a$.

El siguiente ejemplo permite identificar matrices con tensores.

Ejemplo 8.1.14. — Veamos que toda matriz puede ser pensada como un tensor. Más generalmente, consideremos la aplicación lineal

$$\begin{aligned} \varphi :\ V^* \otimes W &\longrightarrow \mathrm{Hom}_k(V, W) \\ f \otimes w &\longmapsto (v \mapsto f(v)w) \end{aligned}$$

y veamos que φ es inyectiva: si $\{w_j\}_{j\in J}$ es una base de W, entonces sabemos que todo elemento de $V^* \otimes W$ es una suma *finita* de tensores simples $\sum_{j\in J} f_j \otimes w_j$, donde cada $f_j \in V^*$ es una forma lineal en V. Así, si la imagen por φ de dicha suma es nula en $\mathrm{Hom}_k(V, W)$, entonces $\sum_{\text{finita}} f_j(v)w_j = 0$ para todo $v \in V$ y por ende $f_j(v) = 0$ para todo $v \in V$ y para todo j, dado que $\{w_j\}_{j\in J}$ es una base de W. Así, concluímos que en tal caso $f_j = 0$ para todo $j \in J$ y por ende φ es inyectivo.

Por otra parte, la aplicación φ no siempre es sobreyectiva: dado que tomamos combinaciones lineales *finitas*, su imagen en $\mathrm{Hom}_k(V, W)$ consiste en las aplicaciones lineales de rango finito. En particular, si V o W es de dimensión finita, entonces

$$\varphi : V^* \otimes W \xrightarrow{\sim} \mathrm{Hom}_k(V, W)$$

es un isomorfismo. Explícitamente, si $\dim_k(V) = n$ y $(v_1, \ldots, v_n)$ es una base de V, de base dual $(v_1^*, \ldots, v_n^*)$, entonces para todo $u \in \mathrm{Hom}_k(V, W)$ se tiene

$$\varphi^{-1}(u) = \sum_{j=1}^{n} v_j^* \otimes u(v_j).$$

Ejercicio 8.1.15. — Con la notación del Ejemplo anterior, determinar la imagen por φ del conjunto de tensores simples de $V^* \otimes W$.

Una aplicación típica de los tensores es utilizarlos para formalizar la idea de pasar de un k-espacio vectorial a un K-espacio vectorial, donde $k \subseteq K$ (eg. $\mathbb{R} \subseteq \mathbb{C}$).

***Ejemplo 8.1.16* (Extensión de escalares).** — Sean k y K cuerpos tales que $k \subseteq K$. Si V es un k-espacio vectorial entonces, dado que K también es un k-espacio vectorial, podemos considerar el producto tensorial

$$V_K := K \otimes_k V \cong V \otimes_k K.$$

Dicho espacio es un k-espacio vectorial, pero también puede ser visto como un K-espacio vectorial: si $\lambda \in K$, la multiplicación por λ dada por $m_\lambda : K \to K$, $x \mapsto \lambda x$ define un endomorfismo k-lineal de K. Luego, podemos definir la multiplicación por $\lambda \in K$ en V_K como el endomorfismo

$m_\lambda \otimes \mathrm{Id}_V$. Las propiedades de K-espacio vectorial son fáciles de verificar y se dejan como Ejercicio al lector.

En este caso, decimos que V_K se obtiene a partir de V por **extensión de escalares** de k a K. Más aún, $\dim_K(V_K) = \dim_k(V)$ puesto que si $\{v_i\}_{i\in I}$ es una base de V, entonces $\{1 \otimes v_i\}_{i\in I}$ es una base de V_K como K-espacio vectorial.

Por otra parte, un endomorfismo $u : V \to V$ en $\mathrm{End}_k(V)$ se extiende a $u_K : V_K \to V_K$ en $\mathrm{End}_K(V_K)$ al definir $u_K := \mathrm{Id}_K \otimes u$. Matricialmente, si la matriz de u es $A \in M_n(k)$ respecto a una base $\{v_i\}$ de V, entonces u_K posee *la misma* matriz $A \in M_n(K)$ respecto a la base $\{1 \otimes v_i\}$ de V_K.

Un caso particular importante es cuando $k = \mathbb{R}$ y $K = \mathbb{C}$. En tal caso, $V_{\mathbb{C}} = \mathbb{C} \otimes_{\mathbb{R}} V$ es la **complejificación** del $\mathbb{R}$-espacio vectorial V. En particular, $\mathbb{C} \otimes_{\mathbb{R}} \mathbb{R}^n \cong \mathbb{C}^n$.

Ejercicio 8.1.17. — Sean V y W dos espacios vectoriales complejos de dimensión finita ≥ 1. Probar que los espacios vectoriales *reales* $V \otimes_{\mathbb{C}} W$ y $V \otimes_{\mathbb{R}} W$ **no** son isomorfos.

Ejercicio 8.1.18. — Sean V y W dos k-espacios vectoriales de dimensión finita. Probar que

$$\psi : V^* \otimes W^* \xrightarrow{\sim} (V \otimes W)^*,\ f \otimes g \longmapsto (v \otimes w \mapsto f(v)g(w))$$

es un isomorfismo.

8.2. Álgebra tensorial

Comencemos por dar una generalización natural del concepto de m-forma multilineal.

Definición 8.2.1. — Sean $V_1, \ldots, V_d, U$ k-espacios vectoriales. Una **aplicación d-lineal** (o d-multilineal) es una función

$$f : V_1 \times \cdots \times V_d \longrightarrow U$$

que es lineal en cada variable. Denotaremos por

$$\mathrm{Mult}^d(V_1 \times \cdots \times V_d, U) := \{f : V_1 \times \cdots \times V_d \longrightarrow U \ d\text{-lineal}\}$$

al k-espacio vectorial correspondiente.

☙ Siguiendo exactamente el mismo procedimiento que en el caso bilineal, dados $V_1, \ldots, V_d$ k-espacios vectoriales construímos un k-espacio vectorial $V_1 \otimes \cdots \otimes V_d$ producto tensorial, junto con una aplicación d-lineal universal

$$t : V_1 \times \cdots \times V_d \longrightarrow V_1 \otimes \cdots \otimes V_d,$$

tal que $\mathrm{Mult}^d(V_1 \times \cdots \times V_d, U) \cong \mathrm{Hom}_k(V_1 \otimes \cdots \otimes V_d, U)$ para todo k-espacio vectorial U. Más aún, si tenemos d aplicaciones lineales $f_i : V_i \to W_i$ para $i \in \{1, \ldots, d\}$, la propiedad universad permite "tensorizarlas" y obtener una única aplicación lineal

$$f_1 \otimes \cdots \otimes f_d : V_1 \otimes \cdots \otimes V_d \longrightarrow W_1 \otimes \cdots \otimes W_d$$

tal que para todo tensor simple $v_1 \otimes \cdots \otimes v_d$ se verifica $(f_1 \otimes \cdots \otimes f_d)(v_1 \otimes \cdots \otimes v_d) = f_1(v_1) \otimes \cdots \otimes f_d(v_d)$.

Observación 8.2.2. — En la sección anterior vimos que $(V_1 \otimes V_2) \otimes V_3$ y $V_1 \otimes (V_2 \otimes V_3)$ son *canónicamente isomorfos* (i.e., existe un isomorfismo construído sin escoger bases). No es difícil verificar que ambos son canónicamente isomorfos a $V_1 \otimes V_2 \otimes V_3$.

Recuerdo 8.2.3. — Recordemos que una k-álgebra es un k-espacio vectorial A dotado de un producto

$$A \times A \longrightarrow A, \ (a, b) \longmapsto ab$$

que es una aplicación k-bilineal y que dota a A de estructura de *anillo*. Así, la multiplicación es asociativa pero **no** necesariamente conmutativa.

☙ Todas las álgebras que consideraremos poseen una *unidad*, i.e., existe un único elemento $1 \in A$ tal que $a \cdot 1 = 1 \cdot a = a$ para todo $a \in A$. Notar que $1 = 0$ si y sólo si $A = 0$.

Definición 8.2.4. — Sea A una k-álgebra. Decimos que A es una **álgebra graduada** si existe una descomposición (eventualmente infinita)

$$A = \bigoplus_{d \in \mathbb{N}} A_d = A_0 \oplus A_1 \oplus A_2 \oplus \cdots$$

en suma directa de sub-espacios vectoriales tales que para todo $d, e \in \mathbb{N}$ se tiene que $A_d \cdot A_e \subseteq A_{d+e}$. En particular, $1 \in A_0$ y luego $k := k \cdot 1 \subseteq A_0$. Más aún, un **morfismo de álgebras graduadas** es un morfismo de álgebras

$$f : A \longrightarrow B$$

(i.e., $f(a_1 a_2) = f(a_1) f(a_2)$ para todos $a_1, a_2 \in A$ y f es lineal) que además preserva la graduación, i.e., $f(A_d) \subseteq B_d$ para todo $d \in \mathbb{N}$.

Ejemplo 8.2.5. — El álgebra conmutativa $k[X]$ de polinomios en una variable está graduada (mediante el grado de cada polinomio):

$$k[X] = \bigoplus_{d\geq 0} \mathrm{Vect}_k\langle X^d\rangle.$$

Más generalmente, el álgebra $A = k[X_1,\ldots,X_r]$ de polinomios en r variables está graduada si definimos a A_d como el sub-espacio vectorial generado por los polinomios **homogéneos** de grado total d, i.e., generado por los monomios $X_1^{i_1}\cdots X_r^{i_r}$ con $i_1+\cdots+i_r = d$.

Observación 8.2.6. — Más generalmente, si $A = \bigoplus_{d\geq 0} A_d$ es una k-álgebra graduada. Un elemento no-nulo $x \in A\setminus\{0\}$ se dice **homogéneo** si existe $d \in \mathbb{N}$ (que es necesariamente único) tal que $x \in A_d$. Diremos que x **es de grado** d.

Definición 8.2.7. — Sea V un k-espacio vectorial. Definimos las **potencias tensoriales** de V como $T^0V := k$ y, para $d \geq 1$ como

$$T^dV := V^{\otimes d} := V \otimes \overset{d \text{ veces}}{\cdots} \otimes V.$$

En particular, por construcción, se tiene que $(T^dV)^* \cong \mathrm{Mult}^d(V^d, k)$ es el k-espacio vectorial de formas d-lineales en V. Más aún, definimos el **álgebra tensorial** de V por

$$TV := \bigoplus_{d\geq 0} T^dV = k \oplus V \oplus T^2V \oplus T^3V \oplus \cdots$$

donde el producto en TV está dado por

$$\begin{aligned} T^dV \times T^eV &\longrightarrow T^{d+e}V \\ (v_1\otimes\cdots\otimes v_d, w_1\otimes\cdots\otimes w_e) &\longmapsto v_1\otimes\cdots v_d\otimes w_1\otimes\cdots\otimes w_e, \end{aligned}$$

y cuya unidad está dada por $1 \in k = T^0V$. Así, TV es una k-álgebra graduada.

Observación 8.2.8. —

1. El álgebra tensorial TV **no** es conmutativa si $\dim_k(V) \geq 2$, pues $v_1\otimes v_2 \neq v_2\otimes v_1$ si v_1 y v_2 no son proporcionales.
2. Dado que $T^1V = V$, hay una inclusión canónica $\iota : V \hookrightarrow TV$.
3. Si $\{e_i\}_{i\in I}$ es una base de V, entonces TV tiene por base *todos* los tensores $e_{i_1}\otimes\cdots\otimes e_{i_d}$ para $d \in \mathbb{N}$ y $i_1,\ldots,i_d \in I$. En particular, incluso si V es de dimensión finita, el k-espacio vectorial TV es siempre de dimensión infinita cuando $V \neq 0$.

***Teorema 8.2.9* (propiedad universal).** — *El álgebra tensorial TV satisface la propiedad universal siguiente:*

> *Para toda aplicación lineal $f : V \to A$ hacia una k-álgebra con unidad A, existe un* único *morfismo de álgebras $\widehat{f} : TV \to A$ tal que $f = \widehat{f} \circ \iota$, donde $\iota : V \hookrightarrow TV$.*

En otras palabras, el diagrama

$$\begin{tikzcd} V \arrow[rr, "f"] \arrow[dr, hook, "\iota"'] & & A \\ & TV \arrow[ur, "\exists!\, \widehat{f}"'] & \end{tikzcd}$$

es conmutativo.

Demostración. — Como la aplicación $V^d \to A$, $(v_1, \ldots, v_d) \mapsto f(v_1) \cdots f(v_d)$ es d-lineal, la propiedad universal de $T^d V$ permite definir la aplicación lineal $\widehat{f}_d : T^d V \to A$ para todo $d \in \mathbb{N}$, y en particular definir

$$\widehat{f} := \bigoplus_{d \geq 0} \widehat{f}_d : TV \longrightarrow A.$$

Basta ver que $\widehat{f}$ es un morfismo de álgebras. Para esto último, es suficiente verificarlo para tensores simples (pues generan TV):

$$\begin{aligned} \widehat{f}((v_1 \otimes \cdots \otimes v_d) \otimes (w_1 \otimes \cdots \otimes w_e)) &= f(v_1) \cdots f(v_d) f(w_1) \cdots f(w_e) \\ &= \widehat{f}(v_1 \otimes \cdots \otimes v_d) \widehat{f}(w_1 \otimes \cdots \otimes w_e), \end{aligned}$$

de donde se concluye lo pedido. □

Observación 8.2.10. — Como en todos los casos anteriores, la propiedad universal implica que la construcción es "*functorial*", i.e., si $f : V \to W$ es una aplicación lineal entonces existe un *único* morfismo de álgebras $Tf : TV \to TW$ tal que el diagrama

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[d, hook, "\iota_V"] & W \arrow[d, hook, "\iota_W"] \\ TV \arrow[r, "Tf"'] & TW \end{tikzcd}$$

es conmutativo. En efecto, Tf no es nada más que $\bigoplus_{d \geq 0} T^d f$, donde $T^d f = f^{\otimes d} = f \otimes \overset{d \text{ veces}}{\cdots} \otimes f$. Más aún, se verifica la construcción es compatible con las composiciones[(2)] en el sentido que $T(f \circ g) = Tf \circ Tg$.

[(2)] En términos más sofisticados (usando el lenguaje de la *teoría de categorías*), T es un "*functor*" entre las *categorías* de k-espacios vectoriales y la de k-álgebras.

Es importante destacar que en Física (y en Geometría Diferencial) se consideran frecuentemente ciertos tensores particulares:

Sea V un k-espacio vectorial de $\dim_k(V) = n$, y sea $V^* = \mathrm{Hom}_k(V, k)$ su espacio dual. Durante este párrafo, y para simplificar la notación, si $\mathscr{B} = (e_1, \dots, e_n)$ es una base de V, entonces denotaremos por $\mathscr{B}^* = (e^1, \dots, e^n)$ la respectiva base dual de V^*. Así, por definición, tenemos que $e^i(e_j) = \delta_{ij}$ (que usualmente se reescribe como $e^i(e_j) = \delta^i_j$ en este contexto).

Definición 8.2.11. — Sea $V \cong k^n$ un k-espacio vectorial, y sea V^* su espacio dual. Un tensor **p-covariante y q-contravariante** es un elemento del producto tensorial

$$T^pV^* \otimes T^qV = V^* \otimes \overset{p \text{ veces}}{\cdots} \otimes V^* \otimes V \otimes \overset{q \text{ veces}}{\cdots} \otimes V.$$

Alternativamente, los vectores de $T^pV^* \otimes T^qV$ son llamados **tensores de tipo** (p, q). Explícitamente, si $(e_1, \dots, e_n)$ es una base de V y $(e^1, \dots, e^n)$ es la base dual de V^*, entonces todo tensor T de tipo (p, q) se escribe como

$$T = \sum_{\substack{i_1,\dots,i_q \\ j_1,\dots,j_p}} T^{i_1,\dots,i_q}_{j_1,\dots,j_p} e^{j_1} \otimes \cdots \otimes e^{j_p} \otimes e_{i_1} \otimes \cdots \otimes e_{i_q},$$

pues los vectores de la forma $e^{j_1} \otimes \cdots \otimes e^{j_p} \otimes e_{i_1} \otimes \cdots \otimes e_{i_q}$ forman una base de $T^pV^* \otimes T^qV$.

Observación 8.2.12 **(Notación de Einstein).** — La **notación de Einstein** (que en realidad es un caso particular del *cálculo de Ricci*, desarrollado 30 años antes de los trabajos de Einstein) consiste en omitir los símbolos de sumatorias. Así, un tensor de tipo (p, q) se escribe como

$$T = T^{i_1,\dots,i_q}_{j_1,\dots,j_p} e^{j_1} \otimes \cdots \otimes e^{j_p} \otimes e_{i_1} \otimes \cdots \otimes e_{i_q},$$

y decimos que los coeficientes $T^{i_1,\dots,i_q}_{j_1,\dots,j_p} \in k$ son las coordenadas del tensor T. Por ejemplo, la relación

$$y = \sum_{i=1}^{3} x^i e_i = x^1 e_1 + x^2 e_2 + x^3 e_3$$

se reduce a escribir $y = x^i e_i$.

Ejemplo 8.2.13. — Sea $v = v^i e_i$ un vector de V. Vimos en el Ejemplo 8.1.14 que todo endormofismo de V es un tensor de tipo $(1, 1)$, pues $\mathrm{End}_k(V) = \mathrm{Hom}_k(V, V) \cong V^* \otimes V$. Así, si $(e'_1, \dots, e'_n)$ es otra base de V

y $P = (P^i_j)$ es la matriz de cambio de base, entonces $e'_j = P^i_j e_i$. Luego, si escribimos $v = v'^j e'_j$ entonces

$$v = v'^j e'_j = v'^j P^i_j e_i,$$

y luego $v^i = P^i_j v'^j$, o bien, $v'^j = (P^{-1})^j_i v^i$. Así, observamos que las coordenadas de v se transforman *en sentido inverso* a los vectores de la base: de aquí viene el término *contravariante*.

Análogamente, para una forma lineal $f = f_j e^j = f'_i e'^i$ se tiene que

$$f'_i = f(e'_i) = f(P^j_i e_j) = P^j_i f_j,$$

i.e., las coordenadas de $f \in V^*$ se transforman *en el mismo sentido* que los vectores de la base: de aquí viene el término *covariante*.

Ejercicio 8.2.14. — Dado un tensor de tipo (p, q) de coordenadas $(T^{i_1,\dots,i_q}_{j_1,\dots,j_p})$, probar que las coordenadas en una nueva base $(e'_1, \dots, e'_n)$ están dadas por

$$(T')^{i'_1,\dots,i'_q}_{j'_1,\dots,j'_p} = (P^{-1})^{i'_1}_{i_1} \cdots (P^{-1})^{i'_q}_{i_q} P^{j_1}_{j'_1} \cdots P^{j_p}_{j'_p} T^{i_1,\dots,i_q}_{j_1,\dots,j_p}.$$

Observación 8.2.15. — Generalmente, los cálculos anteriores se hacen en presencia de una "*pseudo-métrica*", i.e., una forma bilineal $B : V \times V \to k$ no-degenerada (e.g. un producto escalar, o la forma de Lorentz), y decimos que B es un **tensor métrico**. En particular, dado que

$$\mathrm{Bil}(V \times V, k) \cong (V \otimes V)^* \cong V^* \otimes V^*,$$

podemos pensar a B como un tensor de tipo $(2, 0)$ (i.e., 2-covariante) y denotamos su matriz asociada respecto a la base $(e_1, \dots, e_n)$ de V mediante $g_{ij} := B(e_i, e_j)$.

Dado que B es una forma bilineal no-degenerada, tenemos que

$$\widehat{B} : V \xrightarrow{\sim} V^*,\ x \longmapsto B(x, \cdot\,)$$

es un isomorfismo que identifica vectores (i.e., tensores contravariantes) y formas lineales (i.e., tensores covariantes). Así, en coordenadas, tenemos que $\widehat{B}(e_j)(e_i) = B(e_j, e_i) = g_{ji}$ y por ende $\widehat{B}(v^j e_j) = v^j g_{ji} e^i$. En otras palabras, pasamos de coordenadas contravariantes (v^j) a coordenadas covariantes (v_i) mediante la fórmula $v_i = v^j g_{ji}$.

Finalmente, observamos que el isomorfismo $\widehat{B}$ permite transportar la métrica B a una métrica B^* en V^*. Explícitamente, la matriz $g^{ij} := B^*(e^i, e^j)$ de B^* respecto a la base dual $(e^1, \dots, e^n)$ de V^* es la inversa de la matriz (g_{ij}), i.e., $g_{ij} g^{jk} = \delta^k_i$. Este "*tensor métrico dual*" permite pasar de coordenadas covariantes a contravariantes mediante la fórmula $v^j = v_i g^{ij}$ y, en particular, se tiene que

$$B(u,v) = u^i v_i = u_i v^i = u_i v^i = g_{ij} u^i v^j = g^{ij} u_i v_j.$$

8.3. Álgebra exterior y formas multilineales alternadas

Por un lado, gracias a la construcción del álgebra tensorial $TV = \bigoplus_{d\geq 0} T^d V$, tenemos que los espacios vectoriales

$$(T^d V)^* \cong \mathrm{Mult}^d(V \times \overset{d \text{ veces}}{\cdots} \times V, k)$$

son canónicamente isomorfos. Por otro lado, recordemos que en el Capítulo 2 estudiamos formas multilineales alternadas: una forma d-lineal $\omega : V \times \overset{d \text{ veces}}{\cdots} \times V \to k$ es **alternada** si $\omega(v_1, \ldots, v_d) = 0$ cuando dos de los $v_1, \ldots, v_d$ son iguales. Más aún, si denotamos por

$$\mathrm{Alt}^d(V) = \{\omega : V \times \overset{d \text{ veces}}{\cdots} \times V \to k \text{ forma } d\text{-lineal alternada}\}$$

al k-espacio vectorial de formas d-lineales alternadas, sabemos que si $n = \dim_k(V)$ entonces *toda* forma n-lineal alternada es proporcional al determinante, i.e., $\mathrm{Alt}^n(V) = \mathrm{Vect}_k(\det)$.

El objetivo de esta sección es realizar una construcción análoga a $\mathrm{Alt}^d(V)$ en TV, y que será *dual* a $\mathrm{Alt}^d(V)$ en un sentido preciso. De manera más concreta, nos gustaría construir a partir de TV una nueva álgebra $\bigwedge V$ tal que a cada tensor $v_1 \otimes \cdots \otimes v_d$ en TV seamos capaces de asociale un "*símbolo*" $v_1 \wedge \cdots \wedge v_d$ (i.e., un vector en $\bigwedge V$) tal que $v_1 \wedge \cdots \wedge v_d = 0$ si dos de los $v_1, \ldots, v_d$ son iguales.

Como ya hemos visto anteriormente, los cocientes son útiles para construir objetos a partir de otros: basta cocientar por las "*relaciones*" que deseamos que se verifiquen.

Definición 8.3.1. — Sea V un k-espacio vectorial y TV su álgebra tensorial. Sea $I \subseteq TV$ el sub-espacio vectorial generado por todos los tensores de la forma

$$a \otimes v \otimes v \otimes b \text{ donde } v \in V \text{ y } a, b \in TV.$$

Definimos el **álgebra exterior** de V como $\bigwedge V := TV/I$, y denotamos por $v_1 \wedge \cdots \wedge v_d$ la imagen de $v_1 \otimes \cdots \otimes v_d$ en el cociente. Adicionalmente, la composición

$$V \hookrightarrow TV \twoheadrightarrow \bigwedge V$$

será denotada $\iota : V \longrightarrow \bigwedge V$.

Hecho (MAT214): Se puede probar que I es un *ideal* del álgebra TV. En términos prácticos, esto implica que el cociente $\bigwedge V = TV/I$ también es una k-álgebra, i.e., la multiplicación (no conmutativa) en TV define naturalmente una multiplicación en el álgebra exterior $\bigwedge V$, que llamamos **producto exterior**:

$$(v_1 \wedge \cdots \wedge v_d, w_1 \wedge \cdots \wedge w_e) \longmapsto v_1 \wedge \cdots \wedge v_d \wedge w_1 \wedge \cdots \wedge w_e.$$

***Teorema 8.3.2* (propiedad universal).** — *El álgebra exterior $\bigwedge V$ satisface la siguiente propiedad universal:*

> *Para toda aplicación lineal $f : V \to A$ a una $k-$álgebra con unidad A tal que $f(v)^2 = 0$ para todo $v \in V$ existe un único morfismo de álgebras $\widehat{f} : \bigwedge V \to A$ tal que $f = \widehat{f} \circ \iota$ con $\iota : V \to \bigwedge V$.*

En otras palabras, el diagrama

$$\begin{tikzcd}
V \arrow[rr, "f"] \arrow[dr, "\iota"'] & & A \\
& \bigwedge V \arrow[ur, "\exists ! \widehat{f}"'] &
\end{tikzcd}$$

Demostración. — Por la propiedad universal de TV, $\exists ! g : TV \to A$ inducido por f. Dado que

$$g(v \otimes v) = f(v) \otimes f(v) = f(v)^2 = 0$$

la aplicaicón lineal g se anula en el ideal I. Luego, la propiedad universal del cociente implica que $\exists ! \widehat{f} : \bigwedge V := TV/A \to A$ que además es un morfismo de álgebras en este caso (pues g lo es). En otras palabras, la demostración puede resumirse en el diagrama conmutativo:

$$\begin{tikzcd}
V \arrow[r, "f"] \arrow[d] \arrow[dd, bend right, "\iota"'] & A \\
TV \arrow[ur, dashed, "\exists ! g"'] \arrow[d] & \\
\bigwedge V \arrow[uur, dashed, bend right, "\exists ! \widehat{f}"'] &
\end{tikzcd}$$

□

Describamos ahora la “functorialidad” asociada a la propiedad universal anterior, cuya demostración se deja como ejercicio.

***Proposición 8.3.3* (functorialidad). —** *Sea $f : V \to W$ es una aplicación lineal entre $k-$ev. Entonces, existe un único morfismo de álgebras $\bigwedge f : \bigwedge V \to \bigwedge W$ tal que $\iota_W \circ f = \bigwedge f \circ \iota_V$, i.e., el diagrama*

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[d, "\iota_V"'] & W \arrow[d, "\iota_W"] \\ \bigwedge V \arrow[r, "\bigwedge f"'] & \bigwedge W \end{tikzcd}$$

Más aún, $\bigwedge(f \circ g) = \bigwedge f \circ \bigwedge g$.

Veamos ahora cómo describir el álgebra exterior $\bigwedge V$ concretamente. Recordemos que $I \subseteq TV$ es el sub$-$ev generado por tensores de la forma:

$$a \otimes v \otimes v \otimes b \quad \text{con} \quad v \in V \quad \text{y} \quad a, b \in TV$$

Para $d \in \mathbb{N}$, definimos $I_D := I \cap T^d V$ como el sub$-$ev de I generado por elementos homogéneos de grado d. Luego $I = \bigoplus_{d \geq 0} I_D$ y en particular podemos considerar para cada $d \in \mathbb{N}$ el cociente

$$\bigwedge^d V := T^d V / I_d$$

llamada la *$d-$ésima potencia exterior de V*. Así, el álgebra exterior también es uan $k-$álgebra graduada:

$$\bigwedge V = \bigoplus_{d \in \mathbb{N}} T^d V / I_d = \bigoplus_{d \in \mathbb{N}} \bigwedge^d V$$

Importante: Dado que los vectores no-nulos ed I contienen los elementos de grado 2 de la forma $v \otimes v$ tenemos que $T^0 V \cap I = \{0\}$ y $T^1 V \cap I = \{0\}$, asi que $\bigwedge^0 V \cong k$ y $\bigwedge^1 V \cong V$. En particular, $\iota : V \hookrightarrow \bigwedge V$ es *inyectivo*. Así,

$$\begin{array}{ccccccccccc} \bigwedge V & = & k & \oplus & V & \oplus & \bigwedge^2 V & \oplus & \bigwedge^3 V & \oplus & \cdots \\ & & \Cup & & \Cup & & \Cup & & \Cup & & \\ & & \lambda & & v & & v \wedge v & & v \wedge v \wedge v & & \end{array}$$

Luego, el $K-$ev $\bigwedge V$ está generado por los vectores de la forma $v_1 \wedge \cdots \wedge v_d$ con $d \in \mathbb{N}$ y donde $v_1 \wedge \cdots \wedge v_d = 0$ si dos de los $v_1, \ldots, v_d \in V$ son iguales. Además, el hecho que $\bigwedge V$ sea una álgebra graduada se escribe como

$$\left(\bigwedge^d V \right) \wedge \left(\bigwedge^e V \right) \subseteq \bigwedge^{d+e} V$$

Más aún. si $f : V \to W$ es una aplicación lineal, entonces $\bigwedge f = \bigwedge_{d\geq 0} \bigwedge^d f$, donde $\bigwedge^d f : \bigwedge^d V \to \bigwedge^d W$ está dada por

$$\left(\bigwedge^d f\right)(v_1 \wedge \cdots \wedge v_d) = f(v_1) \wedge \cdots \wedge f(v_d)$$

Proposición 8.3.4. — *Sea V un $k-$ev. Entonces, para todo $d \geq 2$ la aplicación $d-$lineal*

$$V^d = \underbrace{V \times \cdots \times V}_{a} \longrightarrow \bigwedge^d V$$

$$(v_1, \ldots, v_d) \longmapsto v_1 \wedge \cdots \wedge v_d$$

es alternada. En particular, para toda permutación $\sigma \in S_d$ y $v_1, \ldots, v_d \in V$ se tiene:

$$v_{\sigma(1)} \wedge \cdots \wedge v_{\sigma(d)} = \epsilon(\sigma) v_1 \wedge \cdots \wedge v_d$$

Demostración. — Por definición de I y de $\bigwedge V$, la expresión $v_1 \wedge \ldots \wedge v_d$ es nula si dos vectores consecutivos son iguales. El caso general se deduce al expresar una permutación arbitraria como producto de transposiciones. □

Corolario 8.3.5. — *El álgebra graduadad $\bigwedge V$ es* **anti-conmutativa***: si $\alpha \in \bigwedge^d V$ y $\beta \in \bigwedge^e V$ entonces se tiene que $\beta \wedge \alpha = (-1)^{de} \alpha \wedge \beta$.*

Demostración. — Basta probarlo en el caso que $\alpha = v_1 \wedge \cdots \wedge v_d$ u $\beta = w_1 \wedge \cdots \wedge w_e$ (el caso general es una combinación ñieal de elementos de esta forma). En tal caso, el resultado se obtiene usando repetidamente la identidad $w \wedge v = -v \wedge w$. □

El siguiente resultado resume las principales propiedades de las potencias exteriores $\bigwedge^d V$-

Teorema 8.3.6. — *Sea V un $k-$ev y sea $d \in \mathbb{N}$. Entonces*

1. *Para todo $d \geq 1$, la aplicación añternada $V^d \to \bigwedge^d V, (v_1, \ldots, v_d) \mapsto v_1 \wedge \cdots \wedge v_d$ satisface la propiedad universal siguiente:*

   Para todo $k-$ev U y toda aplicación $d-$lineal alternada $\omega : V^d \to U$, existe una **única** *aplicación lineal $\bigwedge \omega : \bigwedge^d V \to U$ tal que el diagrama*

$$\begin{tikzcd} V^d \arrow[r, "\omega"] \arrow[d] & U \\ \bigwedge^d V \arrow[ur, "\exists!\widehat{\omega}"'] & \end{tikzcd}$$

es conmutativo. En particular, $\mathrm{Alt}^d(V) \cong (\bigwedge^d V)^*$*,*

2. *Si* V *es de dimensión finita* n *y* $(e_1, \ldots, e_n)$ *es una base de* V*, entonces* $(e_{i_1} \wedge \cdots \wedge e_{i_d})_{1 \leq i_1 < \cdots < i_d \leq n}$ *es una base de* $\bigwedge^d V$ *para todo* $d \geq 1$*. En particular, tenemos* $\bigwedge^d V = 0$ *para* $d > n$ *y* $\dim_k(\bigwedge^d V) = \binom{n}{d}$.
3. *Si* V *es de dimensión finita* n *y* $0 \leq d \leq n$*, la forma bilineal*
$$\begin{aligned} \textstyle\bigwedge^d V \times \bigwedge^{n-d} V &\longrightarrow \textstyle\bigwedge^n V \cong k \\ (\alpha, \beta) &\longmapsto \alpha \wedge \beta \end{aligned}$$
es **no-degenerada**. *En particular, induce un isomorfismo canónico* $\bigwedge^n V \otimes \left(\bigwedge^d V\right)^* \cong \bigwedge^{n-d} V$ *y un isomorfismo* **no-***canónico (i.e., depende de la elección de un isomorfismo* $\bigwedge^n V \cong k$*)*
$$\left(\bigwedge^d V\right)^* \cong \bigwedge^{n-d} V$$
4. *La aplicación* $\bigwedge^d(V^*) \times \bigwedge^d V \to k, (f_1 \wedge \cdots \wedge f_d, v_1 \wedge \cdots \wedge v_d) \mapsto \det((f_i(v_j))_{1 \leq i,j \leq d})$ *es bilineal y no-degenerada. En particular, si* V *es de dimensión finita entonces hay un isomorfismo canónico* $\bigwedge^d V^* \cong \left(\bigwedge^d V\right)^*$.
5. *Si* V *es de dimensión finita* n *y* $u : V \to V$ *es un endomorfismo, entonces el endomorfismo* $\bigwedge^n u$ *de* $\bigwedge^n V \cong k$ *está dado por la multiplicación por* $\det(u)$*. Explícitamente,* $(\bigwedge^n u)(e_1 \wedge \cdots \wedge e_n) = \det(u) e_1 \wedge \cdots \wedge e_n$.

Demostración. —

1. La propiedad universal de TV implica que la aplicación d-lineal (alternada) $\omega : V^d \to U$ induce una única $g : T^d V \to U$ lineal tal que $\omega = g \circ \iota$ donde $\iota : V^d \to T^d V$.
   Pero, dado que ω es alternada, g se anula en $I_d = I \cap T^d V$ y luego la propiedad universal del cociente implica que $\exists! \widehat{\omega} : \bigwedge^d V = T^d V / I_d \to U$.
   En particular, si $U \cong k$ obtenemos
$$\mathrm{Alt}^d(V) \cong \left(\bigwedge^d V\right)^*$$

2 y 3. Dado que $v_{\sigma(1)} \wedge \cdots \wedge v_{\sigma(d)} = \epsilon(\sigma) v_1 \wedge \cdots \wedge v_d$, tenemos que los vectores $(e_{i_1} \wedge \cdots \wedge e_{i_d})_{1 \leq i_1 < \cdots < i_d \leq n}$ genera $\bigwedge^d V$ (podemos suponer los índices ordenados).

Veamos que son l.i.: si $d = n$ entonces sabemos que $\mathrm{Alt}^n(V) = \mathrm{Vect}_k(\det_{\mathscr{B}})$ es de dimensión 1, y luego $\bigwedge^n V \cong (\mathrm{Alt}^n(V))^* \cong$ es dedimensión 1, por lo que el vector no-nulo $e_1 \wedge e_2 \wedge \cdots \wedge e_n$ es una base de $\bigwedge^n V$. Supongamos que $d < n$ y fijemos índices $1 \leq j_1 < \cdots < j_{n-d} \leq n$, entonces el único caso en que

$$(e_{i_1} \wedge \cdots e_{i_d}) \wedge (e_{j_1} \wedge \cdots \wedge e_{j_{n-d}}) \neq 0 \quad \text{en} \quad \bigwedge^n V \cong k$$

es cuando $\{j_1, \ldots, j_{n-d}\}$ es el complemento del conjunto $\{i_1, \ldots, i_d\}$ en $\{1, \ldots, n\}$, pues no pueden repetirse índices. Así, concluimos que los $(e_{i_1} \wedge \cdots \wedge e_{i_d})_{1 \leq i_1 < \cdots < i_d \leq n}$ son l.i. (basta tomar producto exterior de cualquier combinación lineal con $e_{j_1} \wedge \cdots \wedge e_{j_{n-d}}$) y luego una base de $\bigwedge^d V$. Más aún, esto también prueba que $\bigwedge^d V \times \bigwedge^{n-d} V \to \bigwedge^n V, (\alpha, \beta) \mapsto \alpha \wedge \beta$ es bilineal no-degenerada.

En particular, "fijar la segunda variables" induce un isomorfismo canónico entre $\bigwedge^{n-d} V$ y

$$\mathrm{Hom}_k(\bigwedge^d V, \bigwedge^n V) \cong \left(\bigwedge^d V\right)^* \otimes \bigwedge^n V$$

2. Las propiedades del determinante vistas en el Capítulo 2 implican que la aplicación

$$\underbrace{V^* \times \cdots \times V^*}_{d \text{ veces}} \times \underbrace{V \times \cdots \times V}_{d \text{ veces}} \to k, \quad ((f_1, \ldots, f_d), (v_1, \ldots, v_d)) \mapsto \det((f_i(v_j))_{1 \leq i,j \leq d})$$

es alternada en los f_j y es alternada en los v_i, por lo que induce una aplicación bilineal

$$B : \bigwedge^d V^* \times \bigwedge^d V \to k$$

Si $(e_1, \ldots, e_n)$ es una base de V y $(e_1^*, \ldots, e_n^*)$ es ña base dual de V^*, y si $1 \leq i_1 < \cdots < i_d \leq n$ y $1 \leq j_1 < \cdots < j_d \leq n$ entonces (Ejercicio)

$$B(e_{i_1^* \wedge \cdots \wedge e_{i_d}^*, e_{j_1} \wedge \cdots \wedge e_{j_d})=}\begin{cases} 1 & \text{si} \quad i_k = j_k \quad \text{para todo} \quad k \\ 0 & \text{sino} \end{cases}$$

Así, la forma B es no-degenerada e induce una dualidad en la cual $(e_{i_1}^* \wedge \cdots \wedge e_{i_d}^*)_{1 \leq i_1 < \cdots < i_d \leq n}$ es la base dual de $(e_{i_1} \wedge \cdots e_{i_d})_{1 \leq i_1 < \cdots < i_d \leq n}$.

3. Sea $A = (a_{ij}) = \mathrm{Mat}_B(u)$ respecto a una base $\mathscr{B} = (e_1, \ldots, e_n)$ de V. Entonces

$$\left(\bigwedge^n u\right)(e_1 \wedge \cdots \wedge e_n) = u(e_1) \wedge \cdots \wedge u(e_n) = \left(\sum_i a_{i1}e_i\right) \wedge \cdots \wedge \left(\sum_i a_{in}e_i\right)$$
$$= \sum_{i_1,\ldots,i_n=1}^{n} a_{i_1 1} \cdots a_{i_n n} e_{i_1} \wedge \cdots \wedge e_{i_n}$$
$$= \sum_{\sigma \in S_n} a_{\sigma(1)1} \cdots a_{\sigma(n)n}(e_{\sigma(1)} \wedge \cdots \wedge e_{\sigma(n)})$$
$$= \det(u)e_1 \wedge \cdots \wedge e_n$$

□

El siguiente resultado es una aplicación, muy útil en la práctica, del Teorema anterior.

Corolario 8.3.7. — *Sea V un $k-$ev de dimensión finita. Entonces, los vectores $v_1, \ldots, v_d \in V$ son l.i. $\iff$ $v_1 \wedge \cdots \wedge v_d \neq 0$ en $\bigwedge^d V$. En particular, si los vectores $v_1 \wedge \cdots \wedge v_d \in V$ son l.i. y $v \in V$ es otro vector, entonces $v \in \mathrm{Vect}_k(v_1, \ldots, v_d) \iff v_1 \wedge \cdots \wedge v_d \wedge v = 0$ en $\bigwedge^{d+1} V$.*

Demostración. — Si los $v_1, \ldots, v_d$ son l.i., entonces se pueden completar en una base de V y luego $v_1 \wedge \cdots \wedge v_d$ es parte de una base de $\bigwedge^d V$, en particular es no-nulo. Recíprocamente, si $v_1, \ldots, v_d$ son l.d. y, por ejemplo, $v_d = \lambda v_1 + \ldots + \lambda_{d-1}v_{d-1}$ entonces

$$v_1 \wedge \cdots \wedge v_{d-1} \wedge v_d = v_1 \wedge \cdots \wedge v_{d-1} \wedge (\lambda_1 v_1 + \ldots + \lambda_{d-1}v_{d-1}) = 0$$

Finalmente, si los $v_1, \ldots, v_d$ son l.i. y $v \in V$ entonces $v \in \mathrm{Vect}_k(v_1, \ldots, v_d) \iff \{v_1, \ldots, v_d, v\}$ es l.d.. □

Ejemplo 8.3.8. — En $\mathbb{R}^4$ con base canónica e_1, e_2, e_3, e_4 los vectores $v_1 = (1, 0, 1, 0)$ y $v_2 = (0, 1, 0, 1)$ son l.i. pues

$$v_1 \wedge v_2 = (e_1 + e_3) \wedge (e_2 + e_4) = e_1 \wedge_2 + e_1 \wedge e_4 + e_3 \wedge e_2 + e_3 \wedge e_4$$
$$= e_1 \wedge e_2 + e_1 \wedge e_4 - e_2 \wedge e_3 + e_3 \wedge e_4 \neq 0 \quad \text{en} \quad \lambda^2\mathbb{R}^4 \wedge \mathbb{R}^6$$

Por otro lado, un vector $v \in \mathbb{R}^4$ pertenece al plano $\Pi = \mathrm{Vect}_{\mathbb{R}}(v_1, v_2) \cong \mathbb{R}^2$ si y solo si $v_1 \wedge v_2 \wedge v = 0$.

***Ejemplo 8.3.9* (producto cruz).** — Recordemos que si $V \cong \mathbb{R}^3$ es un espacio euclideano **orientado**, enonces el producto cruz $u \times v \in V$ de $u, v \in V$

está definido por la relación

$$\det_{\mathscr{B}}(u, v, w) = \langle u \times v, w \rangle \quad \forall w \in V$$

donde $\mathscr{B} = (e_1, e_2, e_3)$ es cualquier base ortonormal directa de V.
Por otro lado, el producto exterior $e_1 \wedge e_2 \wedge e_3 \in \bigwedge^3 V$ provee un generador canónico (gracias a la orientación) de $\bigwedge^3 V \cong \mathbb{R}$. Así, en este caso el isomorfismo $\bigwedge^{n-d} V \cong \left(\bigwedge^d V\right)^* \otimes \bigwedge^n V$ se reduce (donde $n = 3$ y $d = 1$) a $\bigwedge^2 V \cong V^*$. Más aún, si consideramo el isomorfismo $V \cong V^*$ dado por el producto escalar, obtenemos un isomorfismo $\bigwedge^2 V \cong V$ ly por ende $u \wedge v \in \bigwedge^2 V$ corresponde a un único vector en V. **Ejercicio:** Probar que dicho vector es $u \times v \in V$.

Caso particular importante: Supongamos que $\mathrm{car}(k) = 0$ (e.g. $k = \mathbb{Q}, \mathbb{R}$ o $\mathbb{C}$). En este caso, podemos pensar a $\bigwedge^d V$ como un **sub-espacio vectorial** de $T^d V$ de la manera siguiente: Para cada $\sigma \in S_d$ permutación, entonces definimos el endomorfismo $\widetilde{\sigma}$ de $T^d V$ mediante la fórmula

$$\widetilde{\sigma}(v_1 \otimes \cdots \otimes v_d) = v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(d)}$$

Definición 8.3.10. — Un tensor $T \in T^d V$ se dice **anti-simétrico** si $\widetilde{\sigma}(T) = \epsilon(\sigma) T$ para todo $\sigma \in S_d$. Denotamos por $A_d V \subseteq T^d V$ al sub-ev de tensores anti-simétricos. Más aún, definimos la **aplicación de anti-simetrización** mediante:

$$p : T^d V \to T^d V, \quad T \mapsto \frac{1}{d!} \sum_{\sigma \in S_d} \epsilon(\sigma) \widetilde{\sigma}(T)$$

Ejemplo 8.3.11. — Para $d = 2$ se tiene

$$p(v_1 \otimes v_2) = \frac{1}{2}(v_1 \otimes v_2 - v_2 \otimes v_1)$$

Proposición 8.3.12. — *La aplicación lineal p es un proyector (i.e., $p^2 = p$), de kernel $I_d = I \cap T^d V$ y de imagen $A^d V$. En particular, $A^d V \cong \bigwedge^d V$.*

Demostración. — Para todo $\tau \in S_d$ se tiene:

$$p(\widetilde{\tau}(T)) = \frac{1}{d!} \sum_{\sigma \in S_d} \epsilon(\sigma) \widetilde{\sigma} \widetilde{\tau}(T) \underset{\sigma' = \sigma\tau}{\underset{\uparrow}{=}} \frac{1}{d!} \sum_{\sigma' \in S_d} \epsilon(\sigma' \tau^{-1}) \widetilde{\sigma}'(T) = \epsilon(\tau^{-1}) p(T) = \epsilon(\tau) p(T)$$

Del mismo modo

$$\widetilde{\tau}(p(T)) = \frac{1}{d!} \sum_{\sigma ß S_d} \epsilon(\tau) \widetilde{\tau} \widetilde{\sigma}(T) = \frac{1}{d!} \sum_{\sigma' \in S_d} \epsilon(\tau^{-1} \sigma') \widetilde{\sigma}'(T) = \epsilon(T) p(T)$$

i.e., $p(T) \in A^d V$.
Además,

$$p^2(T) = p(p(T)) = \frac{1}{d!}\sum_{\tau \in S_d} \epsilon(\tau)\widetilde{\tau}(p(T)) = \frac{1}{d!}\sum_{\tau \in S_d} \epsilon(\tau)^2 p(T) = \frac{d!}{d!}p(T) = p(T)$$

de donde $p^2 = p$ y $\Im(p) \subseteq A^d V$. Veamos que $\Im(p) = A^d V$. Sea $T \in A^d V$. Entonces

$$p(T) = \frac{1}{d!}\sum_{\sigma \in S_d} \epsilon(\sigma) \underbrace{\widetilde{\sigma}}_{=\epsilon(\sigma)T} = \frac{1}{d!}\sum_{\tau \in S_d} \underbrace{\epsilon(\tau)^2}_{=1} T = \frac{d!}{d!}T$$

Así, $p^2 = p$ y $p|_{A^d V} = \mathrm{id}_{A^d V}$ (i.e., $\mathrm{Im}(p) = A^d V$).
Veamos que $I_d \subseteq \ker(p)$. Sabemos que I_d está generado por los tensores de la forma $T = v_1 \otimes \cdots \otimes v_d$ con $v_i = v_{i+1}$ para cierto $i \in \{1, \ldots, d-1\}$. Sea $\tau = (i, i+1) \in S_d$ transposición, entonces $\widetilde{\tau}(T) = T$ y luego $p(T) = p(\widetilde{\tau}(T)) = \epsilon(\tau)p(T) = -p(T)$ y como $\mathrm{car}(k) = 0$ tenemos que $p(T) = 0$.
Luego, la propiedad universal del cociente implica que $\exists ! \widehat{p} : \bigwedge^d V = T^d V / I_d \twoheadrightarrow A^d V$.
Si $\pi : T^d V \twoheadrightarrow \bigwedge^d V$ es la proyección al cociente entonces

$$\begin{aligned}
(\pi \circ \widehat{p})(v_1 \wedge \cdots \wedge v_d) &= (\pi \circ p)(v_1 \otimes \cdots \otimes v_d) \\
&= \pi\left(\frac{1}{d!}\sum_{\sigma \in S_d} \epsilon(\sigma) v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(d)}\right) \\
&= \frac{1}{d!}\sum_{\sigma \in S_d} \epsilon(\sigma) v_{\sigma(1)} \wedge \cdots \wedge v_{\sigma(d)} \\
&= \frac{1}{d!}\sum_{\sigma \in S_d} \epsilon(\sigma)^2 v_1 \wedge \cdots \wedge v_d \\
&= v_1 \wedge \cdots \wedge v_d
\end{aligned}$$

así que $\pi \circ \widehat{p} = \mathrm{id}_{\bigwedge^d V}$ y luego $\widehat{p}$ es inyectiva, i.e., $I_d = \ker(p)$.

□

8.4. Pffafiano y matrices antisimétricas

El objetivo de esta sección es aplicar el teorema espectral y las propiedades del álgebra exterior para estudiar matrices **anti-simétricas**, i.e., que cumplen ${}^t A = -A$.

¡Atención! — Durante toda esta sección asumiremos $\mathrm{car}(k) = 0$. en particular, $\mathbb{Q} \subseteq k$.

Motivación: Toda matriz anti-simétrica 2×2 se escribe como

$$A = \begin{pmatrix} 0 & \lambda \\ -\lambda & 0 \end{pmatrix} \qquad , \lambda \in k$$

En particular, $\det(A) = \lambda^2$ es el cuadrado de la función $P(A) := \lambda$. En esta sección veremos que toda matriz anti-simétrica real puede escribirse en términos de bloques 2×2 como el anterior, y que el determinante es el cuadrado de una función: el **pfaffiano**.

Observación 8.4.1. — Si $A \in M_m(k)$ anti-simétrica es invertible (i.e., $A \in \mathrm{GL}_m(k)$) entonces $m = 2n$ es par: ${}^tA = -A \Rightarrow \det(A) = \det({}^tA) = \det(-A) = (-1)^m \det(A) = 0$ si m es impar.

Teorema 8.4.2 **(Teorema espectral (caso antisimétrico))**
Sea $V \cong \mathbb{R}^b$ espacio euclideano y $u : V \to V$ endomorfismo anti-simétrico (i.e., $u^* = -u$. Entonces, existe una base ortonormal $\mathscr{B}$ de V y reales $\lambda_1, \dots, \lambda_s \in \mathbb{R}^{\geq 0}$ positivos tales que

$$\mathrm{Mat}_B(u) = \left(\begin{array}{cc|c|ccc|c} 0 & \lambda_1 & 0 & & \cdots & & 0 \\ -\lambda_1 & 0 & 0 & & \cdots & & 0 \\ \hline 0 & 0 & \ddots & 0 & & 0 & 0 \\ \hline \vdots & \vdots & 0 & 0 & & \lambda_s & \vdots \\ & & 0 & -\lambda_s & & 0 & \\ \hline 0 & 0 & 0 & 0 & & 0 & 0_{n-r} \end{array}\right)$$

y donde $\pm i\lambda_1, \dots, \pm i\lambda_s$ son los valores propios no-nulos de u.

Demostración. — Recordemos que si $U \subseteq V$ es un sub-ev, entonces $u(U) \subseteq U \iff u^*(U^\perp) \subseteq U^\perp$. En particular, $K = \ker(u)$ estable por u se verifica que $K^\perp$ es estable por $u^* = -u$, y luego $K^\perp$ es estable por u. Más aún, la restricción de u a $K^\perp$, $u|_{K^\perp}$, es anti-simétrica y de valores propios no-nulos, así que $\dim_R(K^\perp) = \mathrm{rg}(u) = r = 2s$ es par. Luego, el Teorema espectral implica en este caso que los valores propios de $u|_{K^\perp}$ son imaginarios puros no-nulos $\pm i\lambda_1, \dots, \pm i\lambda_s$ que vienen en pares conjugados pues u es un endomorfismo **real**.
Si $K = V$ entocnes $u = 0$ y el resulado es claro. Sino, consideramos el valor

propio $i\lambda_1$ y notamos que podemos asociarle un plano $\Pi_1 \cong \mathbb{R}^2$ en V que es estable por u, en efecto:
Para toda matriz real A de valor propio $i\lambda \in i\mathbb{R}$ no-nulo, consideramos $z \in \mathbb{C}^n \setminus \{0\}$ vector propio complejo y escribimos $Z = X + iY$ con $X, Y \in \mathbb{R}^n$ y luego

$$AZ = i\lambda Z \iff AX + iAY = -\lambda Y + i\lambda X \iff AX = -\lambda Y \quad \text{y} \quad AY = \lambda X$$

Así, $\Pi := \mathrm{Vect}_{\mathbb{R}}(X, Y) \cong \mathbb{R}^2$ es un plano invariante por A y $\langle X, Y\rangle = 0$.
El cálculo anterior implica que si $\mathscr{B}_1 = (a_1, b_1)$ es una base ortonormal de Π_1 tal que $u(a_1) = -\lambda_1 b_1$ y $u(b_1) = \lambda_1 a_1$, entonces

$$\mathrm{Mat}_{B_1}(u|_{\Pi_1}) = \begin{matrix} 0 & \lambda_1 \\ -\lambda_1 & 0 \end{matrix}$$

Más aún, intercambiando $i\lambda_1$ con $-i\lambda_1$ si fuera necesario, podemos suponer $\lambda_1 > 0$.
Finalmente, por inducción en $n = \dim_R(V)$ y considerando $V = \Pi_1 \oplus \Pi_1^\perp$, tenemos que necesariamente $K \subseteq \Pi_1^\perp$ y que (por hipótesis de inducci+on) $\Pi^\perp = \Pi_2 \oplus \cdots \oplus \Pi_s \oplus K$ con

$$\mathrm{Mat}_{\mathscr{B}_j}(u|_{\Pi_j}) = \begin{matrix} 0 & \lambda_j \\ -\lambda_j & 0 \end{matrix}$$

para $j = 2, \ldots, s$. $\square$

Corolario 8.4.3. — *Sea $M \in M_n(\mathbb{R})$ matriz anti-simétrica real tal que $r = \mathrm{rg}(A) = 2s$ y $\pm i\lambda_1, \ldots, \pm i\lambda_s \in i\mathbb{R}$ son los valores propios no-nulos de A. Entonces, existe $P \in O(n)$ matriz ortogonal tal que $M = PAP^{-1} = PA\,{}^{\mathrm{t}}P$, donde*

$$A = \left(\begin{array}{cc|c|ccc|c} 0 & \lambda_1 & 0 & & \cdots & & 0 \\ -\lambda_1 & 0 & 0 & & \cdots & & 0 \\ \hline 0 & 0 & \ddots & & 0 & 0 & 0 \\ \hline \vdots & \vdots & 0 & & 0 & \lambda_s & \vdots \\ & & 0 & & -\lambda_s & 0 & \\ \hline 0 & 0 & 0 & & 0 & 0 & 0_{n-r} \end{array}\right) \qquad \text{con} \quad \lambda_1, \ldots, \lambda_s > 0$$

El resto de la sección nos cocentraremos en matrices cuadradas anti-simétricas de tamaño par $2n \times 2n$.

Definición 8.4.4. — Sea $A = (a_{ij}) \in M_{2n}(k)$ matriz anti-simétrica y sea $(e_1, \dots, e_{2n})$ la base canónica de $V = k^{2n}$. Definimos entonces

$$\omega(A) := \sum_{1 \le i < j \le 2n} a_{ij} e_i \wedge e_j \qquad \text{en} \qquad \bigwedge^2 V$$

Si para todo $k \in \mathbb{N}^{\ge 1}$ denotamos

$$\omega(A)^k := \omega(A)^{\wedge k} = \underbrace{\omega(A) \wedge \cdots \wedge \omega(A)}_{k \text{ veces}} \in \bigwedge^{2k} V$$

Entonces, definimos el **pfaffiano de A** como el único escalar Pf(A) $\in k$ tal que:

$$\omega(A)^n = n! \, \mathrm{Pf}(A) e_1 \wedge \cdots e_{2n} \quad \iff \quad \frac{1}{n!}\omega(A)^n = \mathrm{Pf}(A) e_1 \wedge \cdots \wedge e_{2n} \in \bigwedge^{2n} V \cong k$$

Ejemplo 8.4.5. —

1. La matriz
$$A = \begin{matrix} 0 & a_{12} \\ -a_{12} & 0 \end{matrix}$$
cumple $\omega(A) = a_{12} e_1 \wedge e_2$. Aquí $2 = 2n \Rightarrow n = 2$ y luego $\mathrm{Pf}(A) = a_{12}/1! = a_{12}$
2. La matriz antisimétrica 4×4
$$A = \begin{matrix} 0 & a_{12} & a_{13} & a_{14} \\ -a_{12} & 0 & a_{23} & a24 \\ -a_{13} & -a_{23} & 0 & a_{34} \\ -a_{14} & a_{24} & -a_{34} & 0 \end{matrix}$$
cumple

$$\omega(A) \underset{\text{def}}{=} a_{12} e_1 \wedge e_2 + a_{13} e_1 \wedge e_3 + a_{14} e_1 \wedge e_4 + a_{23} e_2 \wedge e_3 + a_{24} e_2 \wedge e_4 + a_{34} e_3 \wedge e_4$$

Luego, el pfaffiano de A se calcula mediante $\omega(A) \wedge \omega(A) = 2! \, \mathrm{Pf}(A) e_1 \wedge e_2 \wedge e_3 \wedge e_4$, resultando

$$\begin{aligned} \omega(A) \wedge \omega(A) =& (a_{12} e_1 \wedge e_2) \wedge (a_{34} e_3 \wedge e_4) + (a_{13} e_1 \wedge e_3) \wedge (a_{24} e_2 \wedge e_4) + \\ & (a_{14} e_1 \wedge e_4) \wedge (a_{23} e_2 \wedge e_3) + (a_{23} e_2 \wedge e_3) \wedge (a_{14} e_1 \wedge e_4) + \\ & (a_{24} e_2 \wedge e_4) \wedge (a_{13} e_1 \wedge e_3) + (a_{34} e_3 \wedge e_4) \wedge (a_{12} e_1 \wedge e_2) \end{aligned}$$

puesto que $e_i \wedge e_j \wedge e_k \wedge e_l = 0$ si dos índices se repiten. Más aún, sabemos que $e_i \wedge e_j = -e_j \wedge e_i$, por lo que finalmente obtenemos

$$\omega(A) \wedge \omega(A) = (a_{12}a_{34} - a_{13}a_{24} + a_{14}a_{23}) e_1 \wedge e_2 \wedge e_3 \wedge e_4$$

y así

$$\mathrm{Pf}(A) = a_{12}a_{34} - a_{13}a_{24} + a_{14}a_{23}$$

es el pfaffiano de A en este caso.

Ejercicio 8.4.6. — Calcular $\mathrm{Pf}(A)$ para una matriz antisimétrica 6×6.

Ejemplo 8.4.7. — Consideremos la matriz antisimétrica

$$A = \left(\begin{array}{cc|c|cc|c} 0 & \lambda_1 & 0 & & \cdots & 0 \\ -\lambda_1 & 0 & 0 & & \cdots & 0 \\ \hline 0 & 0 & \ddots & 0 & 0 & 0 \\ \hline \vdots & \vdots & 0 & 0 & \lambda_s & \vdots \\ & & 0 & -\lambda_s & 0 & \\ \hline 0 & 0 & 0 & 0 & 0 & 0_{n-r} \end{array}\right) \in M_{2n}(k)$$

Entonces,

$$\omega(A) = \sum_{j=1}^{n} \lambda_j e_{2j-1} \wedge e_{2j} = \lambda_1 e_1 \wedge e_2 + \lambda_2 e_3 \wedge e_4 + \ldots + \lambda_n e_{2n-1} \wedge e_{2n}$$

Probemos por inducción en n que $\mathrm{Pf}(A) = \lambda_1 \cdots \lambda_n$, i.e., que

$$\omega(A)^n = n! \lambda_1 \cdots \lambda_n e_1 \wedge e_2 \cdots \wedge e_{2n-1} \wedge e_{2n}$$

Sea $\eta = \lambda_1 e_1 \wedge e_2 + \ldots \lambda_{n-1} e_{2n-3} \wedge e_{2n-2}$. Por hipótesis de inducción

$$\eta^{n-1} = (n-1)! \lambda_1 \cdots \lambda_{n-1} e_1 \wedge \cdots \wedge e_{2n-2}$$

y además $\eta^n = \eta^{n-1} \wedge \eta = 0$.

Recuerdo 8.4.8. — Si $\alpha \in \bigwedge^d V$ y $\beta\text{ß}\bigwedge^e V \Rightarrow \beta \wedge \alpha = (-1)^{de} \alpha \wedge \beta$. En particular, si $\alpha := \lambda_n e_{2n-1} \wedge e_{2n} \in \bigwedge^2 V$ entonces $\beta \wedge \alpha = \alpha \wedge \beta$ para todo β, así que la fórmula del binomio de Newton vale en este caso particular. Más aún, si $\alpha = \lambda_n e_{2n-1} \wedge e_{2n}$ entonces $\alpha^k = 0$ si $k \geq 2$.

Escribamos $\omega = \eta + \alpha$. Entonces

$$\omega^n = (\eta + \alpha)^n = \underbrace{\eta^n}_{=0} + n\eta^{n-1} \wedge \alpha + \underbrace{\binom{n}{2} \eta^{n-2} \wedge \alpha^2 + \ldots}_{=0}$$

obteniendo

$$\omega^n = n\eta^{n-1}\wedge\alpha = n\cdot(n-1)!\lambda_1\cdots\lambda_{n-1}e_1\wedge\cdots e_{2n-2}\wedge(\lambda_n e_{2n-1}\wedge e_{2n}) = n!\lambda_1\cdots\lambda_n e_1\wedge\cdots\wedge e_{2n}$$

y luego $\mathrm{Pf}(A) = \lambda_1\cdots\lambda_n$.

¡Atención! — Notar además que en el caso anterior $\det(A) = \lambda_1^2\cdots\lambda_n^2 = \mathrm{Pf}(A)^2$.

Ejercicio 8.4.9. — Sea $A = (a_{ij}) \in M_{2n}(k)$ matriz anti-simétrica. Probar (e.g. por inducción) la fórmula general

$$\mathrm{Pf}(A) = \frac{1}{2^n n!}\sum_{\sigma\in S_{2n}}\epsilon(\sigma)a_{\sigma(1),\sigma(2)}\cdots a_{\sigma(2n-1),\sigma(2n)} = \sum_{\sigma\in F_{2n}}\epsilon(\sigma)a_{\sigma(1),\sigma(2)}\cdots a_{\sigma(2n-1),\sigma(2n)}$$

donde

$$F_{2n} := \{\sigma\in S_{2n} : \sigma(1)<\sigma(3)<\cdots<\sigma(2n-1) \quad \text{y} \quad \sigma(1)<\sigma(2), \sigma(3)<\sigma(4),\ldots,\sigma(2n-1)<\sigma(2n)\}$$

Recuerdo 8.4.10. — Si $u : V \to V$ es un endomorfismo, entonces (por functorialidad) existe $\bigwedge^2 u : \bigwedge^2 V \to \bigwedge^2 V$ tal que $(\bigwedge^2 u)(e_i\wedge e_j) = u(e_i)\wedge u(e_j)$. Matricialmente, si P es una matriz entonces denotamos de igual modo

$$\left(\bigwedge^2 P\right)(e_i\wedge e_j) = P(e_i)\wedge P(e_j)$$

Lema 8.4.11. — *Sea $A \in M_{2n}(k)$ una matriz anti-simétrica y sea $P \in M_{2n}(k)$ una matriz. Entonces,*

$$\omega(PA\,{}^{\mathrm{t}}P) = \left(\bigwedge^2 P\right)(\omega(A)) \quad \text{en} \quad \bigwedge^2 k^{2n}$$

Demostración. — Si $P = (p_{ij}$ y $A = (a_{kl}$ entonces $PA\,{}^{\mathrm{t}}P = (c_{ij})$ con

$$c_{ij} = \sum_{k,l} p_{ik}a_{kl}p_{jl}$$

entonces

$$\begin{aligned}
\omega(PA\,{}^{\mathrm{t}}P = \sum_{i<j} c_{ij} e_i \wedge e_j) &= \sum_{i<j}\sum_{k,l} p_{ik}a_{kl}p_{jl}e_i \wedge e_j \\
&= \frac{1}{2}\sum_{i,j,k,l} p_{ik}a_{kl}p_{jl}e_i \wedge e_j \\
&= \frac{1}{2}\sum_{i,j,k,l} a_{kl}(p_{ik}e_i) \wedge (p_{jl}e_j) \\
&= \frac{1}{2}\sum_{k,l} a_{kl}P(e_k) \wedge P(e_l) \\
&= \sum_{k<l} a_{kl}P(e_k) \wedge P(e_l) \\
&= \left(\bigwedge^2 P\right)(\omega(A))
\end{aligned}$$

□

El Lema anterior permite probar la siguiente identidad útil.

Proposición 8.4.12. — *Sea $A \in M_{2n}(k)$ matriz anti-simétrica y sea $P \in M_{2n}(k)$ una matriz. Entonces,*

$$\mathrm{Pf}(PA\,{}^{\mathrm{t}}P) = \det(P)\,\mathrm{Pf}(A)$$

Demostración. — La identidad $\omega(PA\,{}^{\mathrm{t}}P) = (\bigwedge^2 P)(\omega(A))$ implica que

$$n!\,\mathrm{Pf}(PA\,{}^{\mathrm{t}}P)e_1 \wedge \cdots \wedge e_{2n} \underset{\text{def}}{=} \omega(PA\,{}^{\mathrm{t}}P)^n = \left(\bigwedge^2 P)(\omega(A)\right)^n$$

Por otro lado, si $V = k^{2n}$, entonces en $\bigwedge V = k \oplus V \oplus \bigwedge^2 V \oplus \cdots \bigwedge^{2n} V$ se tiene $\bigwedge P = 1 \oplus P \oplus \bigwedge^2 P \oplus \cdots \oplus \bigwedge^{2n} P$. Así, dado que $\omega(A) \in \bigwedge^2 V$ (homogéneo de grado 2) tenemos

$$(\bigwedge^2 P)(\omega(A)) = (\bigwedge P)(\omega(A))$$

Sin embargo, $\bigwedge P$ es un morfismo de $k-$álgebras (i.e., respeta el producto) y por lo tanto

$$\left(\bigwedge P\right)(\omega(A)^n) = \left(\bigwedge^{2n} P\right)(\omega(A)^n)$$

Finalmente, dado que $\bigwedge^{2n} P$ es la homotecia de factor $\det(P)$ en $\bigwedge^{2n} V \cong k$ obtenemos

$$\left(\bigwedge^2 P\right)(\omega(A))^n = \left(\bigwedge^{2n} P\right)(\omega(A)^n) = \det(P)\cdot\omega(A)^n = n!\det(P)\,\mathrm{Pf}(A)e_1\wedge\cdots\wedge e_{2n}$$

□

***Teorema 8.4.13* (Cayley, 1847). —** *Sea $A \in M_{2n}(k)$ matriz anti-simétrica. Entonces,*

$$\det(A) = \mathrm{Pf}(A)$$

Demostración. — Basta probarlo para $k = \mathbb{Q}$, pues en particular $k \subseteq \mathbb{R}$. El Teorema espectral implica que $A = P\widetilde{A}\,{}^{\mathrm{t}}P$ con $P \in O(n)$ y $\widetilde{A}$ de la forma dada en el Teorema 8.4.2. Finalmente,

$$\mathrm{Pf}(\widetilde{A})^2 = \det(\widetilde{A}) = \det(A) \quad \text{y} \quad \mathrm{Pf}(A) = \det(P)\,\mathrm{Pf}(\widetilde{A}) = \pm\,\mathrm{Pf}(\widetilde{A})$$

□

8.5. Álgebra simétrica y anillos de polinomios

Motivación: El álgebra exterior $\bigwedge V$ generaliza el concepto de formas multilineales alternadas. El objetivo de esta sección es estudiar el caso simétrico.

Recordemos que una forma bilineal $B : V \times V \to k$ es **simétrica** si $B(x,y) = B(y,x)$ para todos $x, y \in V$. Equivalentemente, si $\widehat{B} : V \otimes V \to k$ es la forma lineal asociada entonces

$$\widehat{B}(x\otimes y) = \widehat{B}(y\otimes x) \quad \Longleftrightarrow \quad \widehat{B}(x\otimes y - y\otimes x) = 0 \quad \forall x, y \in V$$

En particular, si denotamos por $J_2 \subseteq T^2V = V\otimes V$ al sub-ev generado por todos los tensores de la forma $x\otimes y - y\otimes x$ con $x, y \in V$, entonces $J_2 \subseteq \ker(\widehat{B})$.

A continuación construiremos a partir de TV una nueva álgebra SV tal que para cada tensor $v_1\otimes\cdots\otimes v_d$ en TV asociamos un "símbolo" $v_1v_2\cdots v_d$ (i.e., vector de SV) tal que $v_1\cdots v_d = v_{\sigma(1)}\cdots v_{\sigma(d)}$ para toda permutación $\sigma \in S_d$. En particular, $v_1v_2 = v_2v_1$.

***Definición 8.5.1.* —** Sea V un $k-$ev y V su álgebra tensorial. Sea $J \subseteq TV$ el sub-espacio generado por todos los tensores de la forma

$$a\otimes(v\otimes w - w\otimes v)\otimes b \quad \text{con} \quad v, w \in V \quad \text{y} \quad a, b \in TV$$

Definimos el **álgebra simétrica** de V como $SV := TV/J$ (o por $\mathrm{Sym}(V)$), y denotamos por $v_1 \cdots v_d$ la imagen de $v_1 \otimes v_d$ en el cociente. Más aún, la composición $V \hookrightarrow TV \twoheadrightarrow SV$ será denotada $j : V \to SV$.

Observación 8.5.2. — Tal como en el caso de $\bigwedge V$, el cociente $SV = TV/J$ es una $k-$álgebra. Más aún, la definición de J implica que la multiplicación

$$(v_1 \cdots v_d, w_1 \cdots w_e) \mapsto v_1 \cdots v_d w_1 \cdots w_e$$

es conmutativa.

***Teorema 8.5.3* (propiedad universal).** — *El álgebra simétrica SV satisface la siguiente propiedad universal:*

> *Para toda aplicación lineal $f : V \to A$ a una k-álgebra conmutativa con unidad A, existe un único morfismo de álgebras $\widehat{f} : SV \to A$ tal que el diagrama*
>
> $$\begin{tikzcd} V \arrow[rr, "f"] \arrow[dr, "j"'] & & W \\ & SV \arrow[ur, "\exists!\widehat{f}"'] & \end{tikzcd}$$
>
> *es conmutativo.*

Demostración. — Análoga a la propiedad universal de $\bigwedge V$ reemplazando $I \subseteq TV$ por $J \subseteq TV$. $\square$

***Corolario 8.5.4* (functorialidad).** — *Si $f : V \to W$ es una aplicación lineal entre $k-$ev. Entonces, existe un único morfismo de $k-$álgebras $Sf : SV \to SW$ tal que*

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[d, "j_V"'] & W \arrow[d, "j_W"] \\ SV \arrow[r, "Sf"] & SW \end{tikzcd}$$

Más aún, $S(f \circ g) = Sf \circ Sg$.

Veamos ahora cómo describir el álgebra simétrica SV concretamente.

Recordemos que $J \subseteq TV$ es el sub-ev generado por los tenores de la forma

$$a \otimes (v \otimes w - w \otimes v) \otimes b \quad \text{con} \quad v, w \in V \quad \text{y} \quad a, b \in TV$$

Para $d \in \mathbb{N}$, definimos $J_d := J \cap T^dV$ como el sub-ev de J generado por elementos homogéneos de grado d. Luego $J = \bigoplus_{d \geq 0} J_d$ y en particular

$$SV = \bigoplus_{d \in \mathbb{N}} S^dV \qquad \text{, donde} \quad S^dV := T^dV/J_d$$

es la **d-ésima potencia simétrica de** V. En particular, SV es una $k-$álgebra graduada.

¡Importante! — Dado que los vectores no-nulos de J contienen los elementos de grado 2 de la foma $v \otimes w - w \otimes v$, tenemos que $T^0V \cap J = \{0\}$ y $T^1V \cap J0\{0\}$, así que $S^0V \cong k$ y $S^1V \cong T^1V \cong V$. En particular, $j : V \hookrightarrow SV$ es inyectivo. Explícitamente,

$$\begin{array}{ccccccccccc} SV & = & k & \oplus & V & \oplus & S^2V & \oplus & S^3V & \oplus & \cdots \\ & & \Cup & & \Cup & & \Cup & & \Cup & & \\ & & \lambda & & v & & v_1v_2 & & v_1v_2v_3 & & \end{array}$$

Más aún, si $f : V \to W$ aplicación lineal, entonces $Sf = \bigoplus_{d \geq 0} S^df$ donde $S^df : S^dV \to S^dW$ está dada por

$$(S^df)(v_1 \cdots v_d) = f(v_1) \cdots f(v_d)$$

El siguiente resultado resumen las principales propiedades de las potencias simétricas S^dV.

Teorema 8.5.5. — *Sea V un $k-$ev y sea $d \in \mathbb{N}$. Entonces*

1. *La aplicación $d-$lineal $V^d \to S^dV$ dada por $(v_1, \ldots, v_d) \mapsto v_1 \cdots v_d$ es una aplicación $d-$lineal simétrica (i.e., $f(v_{\sigma(1)}, \ldots, v_{\sigma(d)}) = f(v_1, \ldots, v_d)$ para toda $\sigma \in S_d$. En particular, $(S^dV)^*$ es el $k-$ev de formas $d-$lineales simétricas en V.*
2. *Si V es de dimensión finita n, entonces $SV \cong k[X_1, \ldots, X_n]$ es isomorfa al álgebra de polinomios en n variables. Más aún, si $(e_1, \ldots, e_n)$ es una base de V, entonces una base de S^dV está dada por los $e_{i_1}^{k_1} \cdots e_{i_r}^{k_r}$ para toda $r-$tupla con $1 \leq i_1 < \ldots < i_r \leq n$ y enteros k_i tales que $k_1 + \ldots + k_r = d$ (c.f. polinomio homogéneos de grado d). Luego,*

$$\dim_k(S^dV) = \binom{n+d-1}{n-1} = \binom{n+d-1}{d}$$

Así, si $V \neq 0$, el álgebra simétrica SV es de dimensión infinita.

3. *Si* $\mathrm{car}(k) = 0$ *podemos pensar a* S^dV *como un sub-ev de* T^dV *generado por* **tensores simétricos** *(i.e.* $T \in T^dV$ *tal que* $\widetilde{\sigma}(T) = T$ *para toda* $\sigma \in S_d$. *En efecto, la* **aplicación de simetrización** *dada por*
$$\begin{aligned} q :T^dV &\longrightarrow T^dV \\ v_1 \otimes \cdots \otimes v_d &\longmapsto \frac{1}{d!} \sum_{\sigma \in S_d} v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(d)} \end{aligned}$$
tiene imagen $\mathscr{S}^d(V) \subseteq T^dV$ *el sub-ev de tensores simétricos y* $\ker(q) = J_d$. *En particular,* $\mathscr{S}^d(V) \cong S^d(V) = T^dV/J_d$

Demostración. — La demostración es análoga al caso de $\bigwedge^d V$ y queda como ejercicio. □

Observación 8.5.6. — Supongamos que $\mathrm{car}(k) = 0$. Entonces, toda matriz $M \in M_n(k)$ se escribe de manera única como suma de una matriz simétrica y anti-simétrica:
$$M = S + A \quad , \qquad S = \frac{M + {}^{\mathrm{t}}M}{2} y A = \frac{M - {}^{\mathrm{t}}M}{2}$$
Del mismo modo, todo tensor en $T^2V = V \otimes V$ puede escribirse como (suma de tensores de la forma):
$$v \otimes w = \frac{1}{2}(v \otimes w - w \otimes v) + \frac{1}{2}(v \otimes w + w \otimes v) = p(v \otimes w) + q(v \otimes w)$$
es decir, dichos tensores se descomponen en suma de un tensor simétrico y un tensor anti-simétrico, y obtenemos
$$T^2V = A^2V \oplus \mathscr{S}^2V$$
con $A^2V \cong \bigwedge^2 V$ y $\mathscr{S}^2V \cong S^2V$ (en este caso basta que $\mathrm{car}(k) \neq 2$).

En general, para $d \geq 3$ tenemos una inclusión $T^dV \supseteq A^dV \oplus \mathscr{S}^dV$ pero ella es estricta: si $d = 3$ calculamos
$$\dim_k(T^3V) = \dim_k(V \otimes V \otimes V) = n^3 > \dim_k(A^3V) + \dim_k(\mathscr{S}^3V) = \binom{n}{3} + \binom{n+2}{3}$$
La parte faltante es un sub-ev de T^3V de dimensión $2\frac{n(n^2-1)}{3}$. De hecho, es la suma directa de dos copias de cierto espacio vectorial denotado $S_{(2,1)(V)}$. En general, existe una descomposición canónica

$$T^dV = V^{\otimes d} = \bigoplus_{\substack{\lambda_1 \geq \cdots \geq \lambda_n \geq 0 \\ \lambda_1 + \ldots + \lambda_n = d}} S_{(\lambda_1,\ldots,\lambda_n)}(V)^{\oplus m_\lambda}$$

donde $m_\lambda \in \mathbb{N}^{\geq 1}$ y los $S_{(\lambda_1,\ldots,\lambda_n)}$ son los llamados **functores de Schur** (omitiendo los $\lambda_j = 0$), que verifican

1. $S_{(1,\ldots,1)}(V) \cong \bigwedge^d V$.
2. $S_{(d)}(V) \cong S^d V$.
3. $\dim_k S_{(\lambda_1,\ldots,\lambda_n)}(V) = \prod_{1 \leq i < j \leq n} \left(\frac{\lambda_i - \lambda_j + j - i}{j - i}\right)$

Aquí, la functorialidad significa que para todo $\lambda = (\lambda_1, \ldots, \lambda_n)$ y toda aplicación lineal $f : V \to W$ existe una única lineal inducida

$$S_\lambda(f) : S_\lambda(V) \to S_\lambda(W)$$

con $S_\lambda(\mathrm{id}) = \mathrm{id}$ y $S_\lambda(f \circ g) = S_\lambda(f) \circ S_\lambda(g)$.

Para los interesados, la construcción de los functores de Schur está relacionada a la “teoría de representaciones” y a objetos combinatorios llamados “tablas de Young”. Una introducción a la “teoría de representaciones de grupos finitos” puede ser hallada en el Capítulo 2 de este apunte.

ÍNDICE

BIBLIOGRAFIA