MAT210 : "Algebra lineal"

§0. Notación :

$\mathbb{N} = \{0, 1, 2, 3, \dots\}$ (naturales), $\mathbb{N}^{\geq 1} = \{1, 2, 3, \dots\}$

Abreviaciones (latín):

- cf : "comparar con"
- eg : "por ejemplo"
- ie : "es decir"

La notación $f : A \hookrightarrow B$ (resp. $f : A \twoheadrightarrow B$) indica que f es una función inyectiva (resp. sobreyectiva).

§1. Espacios vectoriales :

Sea k un cuerpo (eg. $k = \mathbb{Q}, \mathbb{R}, \mathbb{C}$ o $\mathbb{F}_p$)

Def: Un k-espacio vectorial (o k-e.v.) es un grupo abeliano $(V, +)$ dotado de una operación externa "multiplicación por escalares"

$$\begin{aligned} k \times V &\longrightarrow V \\ (\lambda, v) &\longmapsto \lambda v \end{aligned}$$

verificando

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

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

Diremos que los elementos $v \in V$ son vectores y los $\lambda \in k$ son escalares.

Ejemplos : son k-e.v. :

1) $k^n := \{(x_1, \dots, x_n) \text{ tq } x_i \in k\}$

2) $M_{m \times n}(k)$ matrices con m-filas y n columnas.

Notación : $M_n(k) := M_{n \times n}(k)$ matrices cuadradas $n \times n$.

3) $k[X]$ anillo de polinomios en una variable X.

Ejercicio Sea V un k-e.v. y sea X un conjunto no-vacío. Probar que

$$W = \{f : X \to V \text{ función}\}$$

es un k-e.v.

Def : Sea V un k-e.v. y $W \subseteq V$ subconjunto. Diremos que W es un sub-e.v. de V si :

1) $0 \in W$
2) $\forall\, w_1, w_2 \in W \Rightarrow w_1 + w_2 \in W$
3) $\forall\, w \in W \Rightarrow -w \in W$

} "$(W, +)$ sub-grupo de $(V, +)$"

4) $\forall\, w \in W$ y $\lambda \in k \Rightarrow \lambda w \in W$ } "estable por mult. por escalares".

Ejercicio $\forall\, W \subseteq V$ subconj no-vacío. Entonces W es sub-e.v. $\Leftrightarrow \forall w_1, w_2 \in W$ y $\forall \lambda \in k$ se tiene que $\lambda w_1 + w_2 \in W$.

Ejemplos: 1) Si $1 \leq m \leq n$, entonces $W = \{(x_1, \ldots, x_m, 0, \ldots, 0);\ x_1, \ldots, x_m \in k\}$ es un sub-e.v. de $V = k^n$.

2) $W = \left\{ \begin{pmatrix} * & * & * \\ 0 & * & * \\ 0 & 0 & * \end{pmatrix} \right\} \subseteq M_3(k)$ es un sub-e.v.

"matrices triangulares superiores"

3) Sea $d \in \mathbb{N}$, entonces $k_d[X] := \{P \in k[X] \text{ tq } gr(P) \leq d\}$ es un sub-e.v. de $k[X]$.

Def: Una función $\varphi: V \to W$ entre dos k-e.v. es una aplicación lineal (u homomorfismo) si preserva la suma y mult. por escalares, ie, si verifica:

i) $\varphi(v_1 + v_2) = \varphi(v_1) + \varphi(v_2) \quad \forall v_1, v_2 \in V \qquad (\Rightarrow \varphi(0_V) = 0_W)$

ii) $\varphi(\lambda v) = \lambda \varphi(v) \quad \forall v \in V, \forall \lambda \in k.$

Ejercicio: $\varphi: V \to W$ es lineal $\iff \varphi(\lambda v_1 + v_2) = \lambda \varphi(v_1) + \varphi(v_2) \quad \forall v_1, v_2 \in V, \forall \lambda \in k.$

Caso particular importante: Si $V = W$, una aplicación lineal $\varphi: V \to V$ es llamada un endomorfismo.

Ejemplos: 1) La derivada (formal) $d: k[X] \longrightarrow k[X]$ dada por

$P(X) = a_n X^n + \ldots + a_1 X + a_0 \longmapsto P'(X) := n a_n X^{n-1} + \ldots + 2a_2 X + a_1.$

es un endomorfismo.

2) $\varphi: k^2 \to k^2$ dada por $\varphi(x,y) = (-x, x+2y)$ es lineal, pero

$\psi: k^2 \to k^2$ dada por $\psi(x,y) = (-x, x+2y+1)$ no es lineal.

Def: Una aplicación lineal $\varphi: V \to W$ es un isomorfismo si φ es biyectiva y si $\varphi^{-1}: W \to V$ es lineal.

Lema: Sea $\varphi: V \to W$ lineal biyectiva, entonces $\varphi^{-1}: W \to V$ es lineal.

Dem: Sea $\psi := \varphi^{-1}$. Basta probar que $\psi(\lambda w_1 + w_2) = \lambda \psi(w_1) + \psi(w_2)$ para todos $w_1, w_2 \in W$ y $\lambda \in k$.

φ biyectiva $\Rightarrow \exists! v_1, v_2 \in V$ tq $\varphi(v_1) = w_1$ y $\varphi(v_2) = w_2$. $(\Rightarrow \psi(w_i) = v_i)$

φ lineal $\Rightarrow \varphi(\lambda v_1 + v_2) = \lambda \varphi(v_1) + \varphi(v_2) = \lambda w_1 + w_2 \qquad (*)$

Si aplicamos ψ a $(*)$ obtenemos: $(\psi \circ \varphi)(\lambda v_1 + v_2) = \psi(\lambda w_1 + w_2)$

$\Rightarrow \psi(\lambda w_1 + w_2) = \lambda v_1 + v_2 = \lambda \psi(w_1) + \psi(w_2)$ ∎

Def: Sean V y W dos k-e.v. Decimos que V y W son isomorfos si $\exists\, \varphi: V \xrightarrow{\sim} W$ isomorfismo. Escribiremos $V \cong W$ en este caso.

Ejercicio: 1) Probar que $W = \{(x_1, \ldots, x_m, 0, \ldots, 0) \in k^n \text{ tq } x_1, \ldots, x_m \in k\} \cong k^m$.

2) Probar que $k_d[X] \cong k^{d+1}$.

3) ¿Es $\varphi: k^2 \to k^2$, $(x,y) \mapsto (-x, x+2y)$ un isomorfismo?

§2. Familias generadoras, libres, bases y dimensión:

Sea V un k-e.v.

Def: Sea $S \subseteq V$ subconj no-vacío. Definimos el sub-e.v. generado por S como:

$$\mathrm{Vect}_k(S) := \{ v = \lambda_1 v_1 + \ldots + \lambda_r v_r \text{ donde } r \in \mathbb{N}^{\geq 1},\ v_i \in S \text{ y } \lambda_i \in k \}.$$

Es decir, $\mathrm{Vect}_k(S)$ es el conjunto de todas las combinaciones lineales (finitas) de elementos de S: es el sub-e.v. más pequeño de V que contiene S.

Convención: $\mathrm{Vect}_k(\emptyset) = \{0\}$.

Notación: Si $S = \{v_1, \ldots, v_n\}$ es finito, escribiremos

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

Def: Un sub-conj $S \subseteq V$ es una familia generadora si $\mathrm{Vect}_k(S) = V$, ie, si todo elemento de V es combinación lineal de elementos de S. Diremos que V es finitamente generado si $\exists S = \{v_1, \ldots, v_n\}$ conj finito tal que $V = \mathrm{Vect}_k(v_1, \ldots, v_n)$.

⚠ Toda familia conteniendo una familia generadora es generadora.

Ejemplo: 1) k^n está generado por los vectores

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

2) Los monomios $\{X^n\}_{n \in \mathbb{N}}$ generan $k[X]$, pues todo polinomio se escribe como combinación lineal finita $P = a_0 + a_1 X + \ldots + a_d X^d$.

Def: Sea $S \subseteq V$ sub-conj no vacío. Diremos que los elementos de S son linealmente independientes (o l.i.) y que S es una familia libre si:

$\forall r \in \mathbb{N}^{\geq 1}$ y $\forall v_1, \ldots, v_r \in S$ distintos y $\forall \lambda_1, \ldots, \lambda_r \in k$ se tiene que

"$\lambda_1 v_1 + \ldots + \lambda_r v_r = 0 \Rightarrow \lambda_1 = \ldots = \lambda_r = 0$"

En caso contrario, diremos que los elementos de S son linealmente dependientes (o l.d.). Esto último equivale a: $\exists r \in \mathbb{N}^{\geq 1}$ y $\exists v_1, \ldots, v_r \in S$ distintos y $\exists \lambda_1, \ldots, \lambda_r \in k$ escalares no todos nulos tq $\lambda_1 v_1 + \ldots + \lambda_r v_r = 0$. Por ejemplo, si $\lambda_i \neq 0$ entonces $v_i = -\sum_{j \neq i} \frac{\lambda_j}{\lambda_i} v_j$. es comb. lineal de los otros v_j con $j \neq i$.

Caso particular: Si $S = \{v_1, \ldots, v_n\}$ finito, entonces los vectores $v_1, \ldots, v_n$ son l.i. si:

$\forall \lambda_1, \ldots, \lambda_n \in k$ se tiene que "$\lambda_1 v_1 + \ldots + \lambda_n v_n = 0 \Rightarrow \lambda_1 = \ldots = \lambda_n = 0$".

⚠ Sup. que $S' \subseteq S \subseteq V$, entonces:

i) S libre $\Rightarrow S'$ libre

ii) S' l.d. $\Rightarrow S$ l.d.

Ejemplos: 1) Si $0 \in S \Rightarrow S$ es l.d.

2) $\{e_1, \dots, e_n\}$ en k^n es l.i., pues $\forall \lambda_1, \dots, \lambda_n \in k$ se tiene $\lambda_1 e_1 + \dots + \lambda_n e_n = (\lambda_1, \dots, \lambda_n) = 0 \Rightarrow \lambda_1 = \dots = \lambda_n = 0$.

3) $\{X^n\}_{n \in \mathbb{N}}$ en $k[X]$ es l.i. **Ejercicio**

Def: Una familia de vectores $\mathcal{B} = \{v_i\}_{i \in I} \subseteq V$ es una base de V si todo vector $v \in V$ se escribe de manera única como comb. lineal de los v_i, ie, si:

1) $\mathcal{B}$ es una familia generadora
2) $\mathcal{B}$ es libre.

En part, el conjunto finito de vectores $v_1, \dots, v_n$ forma una base de V si $\forall v \in V$, $\exists!\ (\lambda_1, \dots, \lambda_n) \in k^n$ tal que $v = \lambda_1 v_1 + \dots + \lambda_n v_n$.

Ejemplos: 1) $\{e_1, \dots, e_n\}$ en k^n es una base, llamada base canónica de k^n.

2) Si $V = M_{m \times n}(k)$, definimos la matriz elemental $E_{ij} \in M_{m \times n}(k)$ como la matriz cuyos coeficientes son todos 0, excepto por el coeficiente (i,j) en la línea i y columna j, que vale 1. Entonces, toda matriz $A \in M_{m \times n}(k)$ se escribe de manera única como

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

$\Rightarrow \{E_{ij}\}_{\substack{1 \le i \le m \\ 1 \le j \le n}}$ es una base de $M_{m \times n}(k)$.

3) $\{1, X, X^2, \dots, X^d\}$ es una base de $k_d[X]$

4) $\{X^n\}_{n \in \mathbb{N}}$ es una base de $k[X]$.

Recuerdo: V es finitamente generado $\Leftrightarrow \exists\, v_1, \dots, v_n \in V$ tq $V = \mathrm{Vect}_k(v_1, \dots, v_n)$.

Teorema: Sea V un k-e.v. finitamente generado.

1) Existen bases de V y todas tienen el mismo cardinal $\dim_k(V)$, llamado la dimensión de V sobre k.

2) De toda familia generadora $\mathcal{F}$ podemos extraer una base, en part. $\mathrm{card}(\mathcal{F}) \ge \dim_k(V)$. Más aún, si $\mathrm{card}(\mathcal{F}) = \dim_k(V) \Rightarrow \mathcal{F}$ es una base.

3) Si $\mathcal{F}$ es una familia l.i., entonces $\mathrm{card}(\mathcal{F}) \le \dim_k(V)$. Más aún, si $\mathrm{card}(\mathcal{F}) = \dim_k(V) \Rightarrow \mathcal{F}$ es una base.

4) (Teorema de la base incompleta): Toda familia l.i. puede completarse en una base.

5) Todo sub-e.v. $W \subseteq V$ es de dimensión finita $\le \dim_k(V)$. Más aún, si $\dim_k(W) = \dim_k(V)$ entonces $W = V$.

Terminología: En adelante, diremos que un k-e.v. es de "dimensión finita" en lugar de "finitamente generado".

Ejemplos: 1) $\dim_k(k^n) = n$

2) $\dim_k(M_{m\times m}(k)) = mn$

3) $\dim_k(k_d[X]) = d+1$

4) $k[X]$ no es de dimensión finita.

Def: Sea $\mathcal{B} = (v_1, \dots, v_n)$ una base ordenada de V. Entonces, todo $v \in V$ se escribe de forma única como

$$v = x_1 v_1 + \dots + x_n v_n .$$

Decimos que $(x_1, \dots, x_n) \in k^n$ son las coordenadas de v respecto a $\mathcal{B}$. Así, la elección de una base $\mathcal{B}$ provee un isomorfismo

$$\varphi_{\mathcal{B}} : k^n \xrightarrow{\sim} V$$

$$(x_1, \dots, x_n) \mapsto x_1 v_1 + \dots + x_n v_n .$$

⚠ En lo que sigue, todas las bases serán ordenadas.

Prop: Sea $f : V \to W$ lineal. Entonces:

1) Si f inyectiva y $\mathcal{F} = (v_1, \dots, v_m) \subseteq V$ es l.i. $\Rightarrow f(\mathcal{F}) \subseteq W$ es l.i.

2) Si f sobreyectiva y $\mathcal{F} = (v_1, \dots, v_m) \subseteq V$ generadora $\Rightarrow f(\mathcal{F})$ genera W.

3) Si f biyectiva y $\mathcal{B} = (v_1, \dots, v_m) \subseteq V$ base $\Rightarrow f(\mathcal{B})$ base de W.

En part, $\dim_k(V) = \dim_k(W) = n$.

Dem: 1) Sup. que existe una relación lineal $\lambda_1 f(v_1) + \dots + \lambda_m f(v_m) = 0$ en W.

f lineal $\Rightarrow f(\lambda_1 v_1 + \dots + \lambda_m v_m) = 0 \underset{\substack{f \text{ inyectiva} \\ \& f(0)=0}}{\Rightarrow} \lambda_1 v_1 + \dots + \lambda_m v_m = 0 \underset{\mathcal{F} \text{ libre}}{\Rightarrow} \lambda_1 = \dots = \lambda_m = 0$ ✓

2) Sea $w \in W$. f sobreyectiva $\Rightarrow \exists v \in V$ tq $f(v) = w$.

$\mathcal{F}$ genera $V \Rightarrow \exists \lambda_1, \dots, \lambda_m \in k$ tq $v = \lambda_1 v_1 + \dots + \lambda_m v_m$, de donde obtenemos

$w = f(v) \underset{\substack{\uparrow \\ f \text{ lineal}}}{=} \lambda_1 f(v_1) + \dots + \lambda_m f(v_m) \in \mathrm{Vect}_k(f(\mathcal{F}))$ ✓

3) Por ① y ②, $f(\mathcal{B})$ es libre y generadora, luego una base. En particular,

$$\dim_k(V) = \underbrace{\mathrm{card}(\mathcal{B})}_{=n} = \underbrace{\mathrm{card}(f(\mathcal{B}))}_{=n} = \dim_k(W) \quad \blacksquare$$

Corolario: 1) Si V es k-e.v. de dimensión finita $\Rightarrow V \cong k^n$, con $n = \dim_k(V)$.

2) Sean V, W dos k-e.v. de dimensión finita. Entonces: $V \cong W \iff \dim_k(V) = \dim_k(W)$.

Dem: 1) Si $\dim_k(V) = n$, la elección de una base $\mathcal{B}$ provee un isomorfismo $\varphi_{\mathcal{B}} : k^n \xrightarrow{\sim} V$. Recíprocamente, si $V \cong k^m$ para cierto $m \underset{\text{Prop ③}}{\Rightarrow} \dim_k(V) = \dim_k(k^m) = m$

$\Rightarrow m = n$ ✓

2) ($\Rightarrow$) Si $V \cong W$, entonces $\dim_k(V) = \dim_k(W)$ (Prop ③).

($\Leftarrow$) Si $\dim_k(V) = \dim_k(W) = n \underset{①}{\Rightarrow} V \cong k^n$ y $W \cong k^n \Rightarrow V \cong W$ $\blacksquare$

3. Núcleo, imagen y teorema del rango

Def: Sea $f: V \to W$ lineal.

1) El <u>núcleo</u> o <u>kernel</u> de f es el sub-e.v. dado por
$$\ker(f) := \{v \in V \text{ tq } f(v) = 0\} \subseteq V.$$
Notar que f inyectiva $\Leftrightarrow \ker(f) = \{0\}$.

2) La <u>imagen</u> de f es el sub-e.v. dado por
$$\mathrm{Im}(f) := f(V) = \{f(v), \text{ donde } v \in V\} \subseteq W.$$
Notar que f sobreyectiva $\Leftrightarrow \mathrm{Im}(f) = W$.

3) Si $\mathrm{Im}(f)$ es de dimensión <u>finita</u> $r \in \mathbb{N}$ (eg, si W o V son de dim. finita) el entero $r = \dim_k \mathrm{Im}(f)$ es llamado el <u>rango</u> de f, denotado $rg(f)$.

<u>Teorema del rango</u>: Sea $f: V \to W$ lineal. Si V es de dimensión <u>finita</u>, entonces
$$\dim_k(V) = \dim_k \ker(f) + rg(f).$$
En part, f sobreyectiva $\Leftrightarrow \dim_k(W) = \dim_k(V) - \dim_k \ker(f)$.

Dem: Sea $n := \dim_k(V)$. V dim. finita $\Rightarrow \ker(f) \subseteq V$ dim. finita.
Sea $d := \dim_k(\ker(f)) \leq n$, y sea $\mathcal{K} = (e_1, \dots, e_d)$ una base de $\ker(f)$.
Como $\mathrm{Im}(f) = f(V)$ está generada por n vectores (imagenes de una base de V)
$\Rightarrow r := \dim_k \mathrm{Im}(f) \leq n$.
Sea $(w_1, \dots, w_r)$ una base de $\mathrm{Im}(f)$, y para cada $i = 1, \dots, r$ sea $v_i \in V$ tq $f(v_i) = w_i$. Veamos que $\mathcal{B} := (e_1, \dots, e_d, v_1, \dots, v_r)$ es una <u>base</u> de V:

i) $\mathcal{B}$ genera: Sea $v \in V$ arbitrario. $\Rightarrow f(v) = \lambda_1 w_1 + \dots + \lambda_r w_r$

$\Rightarrow f(v - (\lambda_1 v_1 + \dots + \lambda_r v_r)) = 0 \Rightarrow v - \lambda_1 v_1 - \dots - \lambda_r v_r \in \ker(f) = \mathrm{Vect}_k(e_1, \dots, e_d)$ (lineal)

Luego, $v = \lambda_1 v_1 + \dots + \lambda_r v_r + \mu_1 e_1 + \dots + \mu_d e_d = \mathrm{Vect}_k(\mathcal{B})$ ✓

ii) $\mathcal{B}$ es libre: Si $\lambda_1 v_1 + \dots + \lambda_r v_r + \mu_1 e_1 + \dots + \mu_d e_d = 0$

$\Rightarrow 0 = f(0) = \lambda_1 w_1 + \dots + \lambda_r w_r \Rightarrow \lambda_1 = \dots = \lambda_r = 0$. ($(w_1, \dots, w_r)$ l.i.)

Luego, $0 = \mu_1 e_1 + \dots + \mu_d e_d \Rightarrow \mu_1 = \dots = \mu_d = 0$ ✓ ($(e_1, \dots, e_d)$ l.i.)

Finalmente: $\dim_k(V) = \mathrm{card}(\mathcal{B}) = d + r = \dim_k \ker(f) + rg(f)$ ∎

Prop: Sea $f: V \to W$ lineal. Si $\dim_k(V) = \dim_k(W) = n$, entonces son equivalentes:
1) f es biyectiva (isomorfismo).
2) f es inyectiva
3) f es sobreyectiva.

Dem: ① $\Rightarrow$ ② & ③ ✓ Si f es inyectiva (resp. sobreyectiva) entonces $\ker(f) = \{0\}$ (resp. $\mathrm{Im}(f) = W$), por lo que el teorema del rango implica que f es sobreyectiva (resp. inyectiva), ie, biyectiva. ∎

§4. Aplicaciones lineales y matrices

Sean V y W dos k-e.v.

Def: Definimos el k-e.v. de aplicaciones lineales de V en W mediante

$$\mathrm{Hom}_k(V,W) = \mathcal{L}(V,W) := \{ f : V \to W \text{ lineal} \}.$$

Ejercicio $\mathrm{Hom}_k(V,W)$ es un k-e.v.

Caso particular: Si $\dim_k(V) = n$ y $\mathcal{B} = (e_1, \dots, e_n)$ es una base de V, entonces dado $\varphi \in \mathrm{Hom}_k(V,W)$ definimos $w_i := \varphi(e_i) \in W$ para $i = 1, \dots, n$.

Todo vector $v \in V$ se escribe de manera única como $v = x_1 e_1 + \dots + x_n e_n$ y luego (por linealidad): $\varphi(v) = x_1 w_1 + \dots + x_n w_n$.

Conclusión: $\varphi : V \to W$ está determinada por los vectores $w_1, \dots, w_n \in W$.

Sub-caso particular: Si $\dim_k(W) = m$ y $\mathcal{C} = (f_1, \dots, f_m)$ es una base de W, entonces cada $w_j = \varphi(e_j)$ se escribe de manera única como

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

que representamos usando el vector columna $w_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}$.

Luego, φ está determinada por la matriz

$$\mathrm{Mat}_{\mathcal{C},\mathcal{B}}(\varphi) := \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \in M_{m \times n}(k).$$

Teorema: Sea $\mathcal{B} = (e_1, \dots, e_n)$ una base de V y sea $\mathcal{C} = (f_1, \dots, f_m)$ una base de W.

① Una aplicación lineal $\varphi : V \to W$ está determinada por una matriz de m filas y n columnas. Más precisamente, la aplicación

$$\begin{aligned} \mathrm{Hom}_k(V,W) &\xrightarrow{\sim} M_{m \times n}(k) \\ \varphi &\longmapsto \mathrm{Mat}_{\mathcal{C},\mathcal{B}}(\varphi) \end{aligned}$$

es un isomorfismo de k-e.v.

② Si $\mathcal{U}$ es otro k-e.v. y si $\mathcal{A} = (d_1, \dots, d_p)$ es una base de $\mathcal{U}$. Entonces para todos $\varphi \in \mathrm{Hom}_k(V,W)$ y $\psi \in \mathrm{Hom}_k(\mathcal{U},V)$ se tiene que

$$\mathrm{Mat}_{\mathcal{C},\mathcal{A}}(\varphi \circ \psi) = \mathrm{Mat}_{\mathcal{C},\mathcal{B}}(\varphi) \cdot \mathrm{Mat}_{\mathcal{B},\mathcal{A}}(\psi).$$

Dem: El punto ① sigue de la discusión anterior, mientras que el punto ② sigue precisamente de la definición de multiplicación de matrices ■

Obs: Sea $A = (a_{ij}) \in M_{m\times n}(k)$ y sean $(e_1,\dots,e_n)$ y $(f_1,\dots,f_m)$ las bases canónicas de k^n y k^m, resp. Entonces, A corresponde a la aplicación lineal $u: k^n \to k^m$ dada por

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

para todo $j = 1,\dots,n$. Es decir, la j-ésima columna de A es la imagen de e_j, el j-ésimo vector de la base canónica de k^n.

Corolario: 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 la aplicación $(u\circ v): k^p \to k^m$.

⚠ Si $A \in M_{m\times n}(k)$ y $B \in M_{l\times p}(k)$, entonces AB está definido si y sólo si $n = l$

Caso particular importante: $V = W$ (endomorfismos) y $\mathcal{B} = \mathcal{C}$.

Teorema: El k-e.v. $M_n(k)$ es una k-álgebra, es decir, es un anillo (no conmutativo si $n \geq 2$) compatible con la estructura de k-e.v. Del mismo modo si V es un k-e.v. entonces el espacio de endomorfismos $\mathrm{End}_k(V)$ es una k-álgebra. En particular, si $\dim_k(V) = n$ y si escogemos una base $\mathcal{B} = (e_1,\dots,e_n)$ de V, la aplicación

$$\begin{aligned} \mathrm{End}_k(V) &\longrightarrow M_n(k) \\ u &\longmapsto \mathrm{Mat}_{\mathcal{B}}(u) \end{aligned}$$

es un isomorfismo de anillos y de k-e.v., ie, un isomorfismo de k-álgebras.

La correspondencia entre aplicaciones lineales y matrices nos permite definir:

Def: 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 $u: k^n \to k^m$ asociada a A.

Observación: 1) $\mathrm{rg}(A) := \mathrm{rg}(u) \leq n$ (teorema del rango) y $\mathrm{rg}(u) \leq m$ (pues $\mathrm{Im}(u)$ es un sub-e.v. de k^m). Luego, $\mathrm{rg}(A) \leq \min(m,n)$.

2) $\mathrm{Im}(u)$ es el sub-e.v. de k^m generado por las columnas de A, luego $\mathrm{rg}(A)$ es el número máximo de columnas de A que son l.i.

3) Más adelante: $\mathrm{rg}(A)$ es el número máximo de filas de A que son l.i.

Para esto último, necesitamos definir:

Def: Sea $A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & & a_{mn} \end{pmatrix} \in M_{m\times n}(k)$, definimos su transpuesta como

$${}^tA := \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & & a_{mn} \end{pmatrix} \in M_{n\times m}(k)$$

, ie, $({}^tA)_{ij} = a_{ji}$. En part, ${}^t({}^tA) = A$.

Ejercicio importante Probar que la aplicación $M_{m\times n}(k) \to M_{n\times m}(k)$, $A \mapsto {}^tA$ es lineal, y que es un isomorfismo. Más aún, si $B \in M_{n\times p}(k)$ entonces ${}^t(AB) = {}^tB \cdot {}^tA$ en $M_{p\times m}(k)$.

§5. Cambios de base: Sea V un k-e.v.

Def: Un endomorfismo $f: V \to V$ es un <u>automorfismo</u> si es biyectivo o, equivalentemente, si $\exists f^{-1}: V \to V$ endomorfismo tq $f \circ f^{-1} = f^{-1} \circ f = id_V$.

De manera similar, una matriz cuadrada $A \in M_n(k)$ es <u>invertible</u> si existe $A^{-1} \in M_n(k)$ tq $AA^{-1} = A^{-1} \cdot A = I_n$; A^{-1} es la <u>inversa</u> de A.

Notación: $GL(V) := \{f: V \to V \text{ automorfismos}\}$ es el <u>grupo general lineal</u>.

Ejercicio Verificar que $GL(V)$ es un grupo (Ayuda: $(f\circ g)^{-1} = g^{-1} \circ f^{-1}$).

$GL_n(k) := \{A \in M_n(k) \text{ invertible}\}$.

Notar que la elección de una base $\mathcal{B}$ de V induce un isomorfismo de grupos $GL(V) \xrightarrow{\sim} GL_n(k)$, $f \mapsto Mat_{\mathcal{B}}(f)$, donde $n = \dim_k(V)$.

Prop: 1) Supg que V es de dimensión <u>finita</u>, y sean $u, v \in End_k(V)$ tales que $u \circ v = id_V \Rightarrow u, v \in GL(V)$ y $u = v^{-1}$.

2) Sean $A, B \in M_n(k)$ tq $AB = I_n \Rightarrow BA = I_n$, $A, B \in GL_n(k)$ y $B = A^{-1}$.

3) Si A invertible, entonces ${}^tA$ invertible. Más aún, $({}^tA)^{-1} = {}^t(A^{-1})$.

Dem: 1) $\forall x \in V$ se tiene $x = u(v(x)) \Rightarrow u$ sobreyectivo y v inyectivo.

$\Rightarrow u$ y v son isomorfismos (V dim finita). En part, $\underbrace{u^{-1} \circ u}_{id_V} \circ v = u^{-1} = v$ ✓

2) Sea u (resp. v) el endomorfismo $k^n \to k^n$ asociado a A (resp. B).

Luego, $AB = I_n \Rightarrow u \circ v = id_V \overset{①}{\Rightarrow} u, v$ isomorfismos y $u = v^{-1}$.

En part, $B = A^{-1}$ y $BA = I_n$.

3) Si $AB = BA = I_n$, entonces ${}^tB\,{}^tA = {}^tA\,{}^tB = {}^tI_n = I_n \Rightarrow {}^tA \in GL_n(k)$ y ${}^tB = {}^t(A^{-1})$ es la inversa de ${}^tA$. ■

Lema: Supg que $\dim_k(V) = n$ y sea $\mathcal{B} = (v_1, \dots, v_n)$ base de V. Sea $f: V \to V$ endomorfismo y denotemos $w_i := f(v_i)$. Si $(w_1, \dots, w_n)$ es una base de V, entonces f es un isomorfismo, cuya inversa $g: V \to V$ está dada por $g(w_i) = v_i$.

Dem: Supg. que $\mathcal{F} = (w_1, \dots, w_n)$ es una base de V.

$\Rightarrow \forall w \in V$, $\exists \lambda_1, \dots, \lambda_n \in k$ tq $w = \lambda_1 w_1 + \dots + \lambda_n w_n = \lambda_1 f(v_1) + \dots + \lambda_n f(v_n)$
$= f(\lambda_1 v_1 + \dots + \lambda_n v_n)$

es decir, f es sobreyectivo $\underset{V \text{ dim finita}}{\Rightarrow} f$ biyectivo ✓

Notar que $(g\circ f)(v_i) = v_i$ $\forall i$ y $(f \circ g)(w_i) = w_i$ $\forall i \Rightarrow g \circ f = f \circ g = id_V$ ■
($\mathcal{B}$ y $\mathcal{F}$ bases)

Def: Sup. que $\dim_k(V) = n$ y sean $\mathcal{B} = (e_1, \dots, e_n)$ y $\mathcal{B}' = (v_1, \dots, v_n)$ dos bases de V. Sea P la matriz (invertible) expresando $\mathcal{B}'$ en términos de $\mathcal{B}$, ie, si

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

entonces $P := \begin{pmatrix} p_{11} & \cdots & p_{1n} \\ \vdots & & \vdots \\ p_{n1} & \cdots & p_{nn} \end{pmatrix} \in GL_n(k)$; y es llamada la <u>matriz de cambio de base de $\mathcal{B}$ a $\mathcal{B}'$</u> y la denotaremos $Mat_{\mathcal{B}}(\mathcal{B}')$.

⚠ Por definición, $P^{-1} = Mat_{\mathcal{B}'}(\mathcal{B})$ es la matriz que expresa $\mathcal{B}$ en términos de $\mathcal{B}'$

Ejercicio importante Verificar que $P = Mat_{\mathcal{B}}(\mathcal{B}') = Mat_{\mathcal{B},\mathcal{B}'}(id_V)$ y que $P^{-1} = Mat_{\mathcal{B}'}(\mathcal{B}) = Mat_{\mathcal{B}',\mathcal{B}}(id_V)$. En otras palabras (¡y contra-intuitivamente!):

$$(V, \mathcal{B}') \xrightarrow[P]{id_V} (V, \mathcal{B}) \quad \text{y} \quad (V, \mathcal{B}) \xrightarrow[P^{-1}]{id_V} (V, \mathcal{B}').$$

Def: Todo vector $v \in V$ se escribe de manera única como

$$v = x_1 e_1 + \dots + x_n e_n \quad \text{y} \quad v = x_1' v_1 + \dots + x_n' v_n$$

y las x_i (resp. x_i') son las <u>coordenadas</u> de v respecto a $\mathcal{B}$ (resp. $\mathcal{B}'$). Luego, $v \in V$ se representa, respecto a la base $\mathcal{B}$ (resp. $\mathcal{B}'$), por el vector:

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \left(\text{resp. } X' = \begin{pmatrix} x_1' \\ \vdots \\ x_n' \end{pmatrix}\right).$$

Prop (cambio de coordenadas): $X = PX'$.

Dem: Si escribimos $v_j = \sum_{i=1}^{n} p_{ij} e_i$ y $v = \sum_{j=1}^{n} x_j' v_j = \sum_{j=1}^{n} \sum_{i=1}^{n} x_j' p_{ij} e_i = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} p_{ij} x_j' \right) e_i$ entonces al comparar con $v = \sum_{i=1}^{n} x_i e_i$, obtenemos que $x_i = \sum_{j=1}^{n} p_{ij} x_j'$, ie, $X = PX'$ ∎

Teorema (cambio de base): Sea $A \in Mat_{\mathcal{E},\mathcal{B}}(u) \in M_{m \times n}(k)$ la matriz de $u : V \to W$ respecto a la base $\mathcal{B} = (e_1, \dots, e_n)$ de V y $\mathcal{E} = (f_1, \dots, f_m)$ de W. Sea $\mathcal{B}'$ (resp. $\mathcal{E}'$) otra base de V (resp. W), y sea $P \in GL_n(k)$ (resp. $Q \in GL_m(k)$) la matriz de cambio de base. Entonces,

$$Mat_{\mathcal{E}',\mathcal{B}'}(u) = Q^{-1} A P$$

Dem: El diagrama

$$\begin{array}{ccc} (V, \mathcal{B}) & \xrightarrow[A = M_{\mathcal{E},\mathcal{B}}(u)]{u} & (W, \mathcal{E}) \\ {\scriptstyle P = Mat_{\mathcal{B}}(\mathcal{B}')} \uparrow {\scriptstyle id_V} & & {\scriptstyle id_W} \downarrow {\scriptstyle Q^{-1} = Mat_{\mathcal{E}'}(\mathcal{E})} \\ (V, \mathcal{B}') & \xrightarrow[Q^{-1}AP = Mat_{\mathcal{E}',\mathcal{B}'}(u)]{u} & (W, \mathcal{E}') \end{array}$$

es conmutativo ∎

Obs: En coordenadas, si $Y = AX$, $X = PX'$ y $Y = QY'$

$$\Rightarrow Y' = Q^{-1} Y = Q^{-1} A X = (Q^{-1} A P) X.$$

Notación: $O_{p,q} \in M_{p\times q}(k)$ denota la matriz nula de p filas y q columnas.

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

1) $\exists P \in Gl_n(k)$ y $Q \in Gl_m(k)$ tq $Q^{-1}AP = \begin{pmatrix} I_r & O_{r,n-r} \\ O_{m-r,r} & O_{m-r,n-r} \end{pmatrix}$.

2) Si $\exists P \in Gl_n(k)$ y $Q \in Gl_m(k)$ tq $Q^{-1}AP = \begin{pmatrix} I_s & O_{s,n-s} \\ O_{m-s,s} & O_{m-s,n-s} \end{pmatrix} \Rightarrow s = r = rg(A)$.

Dem: Sean $(e_1,\dots,e_n)$ y $(f_1,\dots,f_m)$ las bases canónicas de k^n y k^m, y sea $u: k^n \to k^m$ la aplicación lineal asociada a A.

1) Sea $r = rg(A) = \dim_k \mathrm{Im}(u)$ y sea $(w_1,\dots,w_r)$ base de $\mathrm{Im}(u)$, la cual completamos en una base $\mathcal{C} = (w_1,\dots,w_m)$ de k^m.
Consideramos $v_1,\dots,v_r \in k^n$ tq $u(v_j) = w_j$ $\forall j$ y sea $(\varepsilon_1,\dots,\varepsilon_d)$ una base de $\ker(u)$. Al probar el Teorema del rango, vimos que $\mathcal{B} := (v_1,\dots,v_r,\varepsilon_1,\dots,\varepsilon_d)$ es una base de k^n. Más aún, por definición, teníamos que:

$$\mathrm{Mat}_{\mathcal{C},\mathcal{B}}(u) = \begin{pmatrix} I_r & O_{r,n-r} \\ O_{m-r} & O_{m-r,n-r} \end{pmatrix}.$$

Así, el diagrama $\begin{array}{ccc} (k^n,\mathcal{B}) & \xrightarrow{\left(\begin{smallmatrix} I_r & 0 \\ 0 & 0 \end{smallmatrix}\right)} & (k^m,\mathcal{C}) \\ P^{-1}\uparrow\, id_{k^n} & & id_{k^m}\downarrow Q \\ (k^n,(e_j)) & \xrightarrow[u]{A} & (k^m,(f_i)) \end{array}$ implica que $\mathrm{Mat}_{\mathcal{C},\mathcal{B}}(u) = Q^{-1}AP$ ✓

2) La igualdad $Q^{-1}AP = \begin{pmatrix} I_s & O_{s,n-s} \\ O_{m-s,s} & O_{m-s,n-s} \end{pmatrix}$ implica que existen bases $(v_1,\dots,v_n)$ de k^n y $(w_1,\dots,w_m)$ de k^m tq $u(v_i) = w_i$ para $i = 1,\dots,s$ y $u(v_j) = 0$ para $j = s+1,\dots,n \Rightarrow \mathrm{Im}(u) = \mathrm{Vect}_k(w_1,\dots,w_s) \Rightarrow s = rg(A)$ ■

Prop: Sea $A \in M_{m\times n}(k)$, entonces $rg(A) = rg({}^tA)$. En part, $rg(A)$ es el número máximo de filas de A que son l.i.

Dem: Sea $r = rg(A) \underset{\text{corolario}}{\Rightarrow} \exists P \in Gl_n(k)$ y $Q \in Gl_m(k)$ tq $Q^{-1}AP = \begin{pmatrix} I_r & O_{r,n-r} \\ O_{m-r,r} & O_{m-r,n-r} \end{pmatrix}$

$\Rightarrow {}^tP\,{}^tA\,{}^tQ^{-1} = \begin{pmatrix} I_r & O_{r,m-r} \\ O_{n-r,r} & O_{n-r,m-r} \end{pmatrix} \underset{\text{corolario}}{\Rightarrow} r = rg({}^tA)$ ■

Def: Dos matrices $A, B \in M_{m\times n}(k)$ son equivalentes si $\exists P \in Gl_n(k)$ y $Q \in Gl_m(k)$ tq $Q^{-1}AP = B \underset{\text{corolario}}{\Leftrightarrow} rg(A) = rg(B)$.

Caso particular importante: $V = W$ (endomorfismos) y $\mathcal{C} = \mathcal{B}$.

Teorema: Sea $A = \mathrm{Mat}_{\mathcal{B}}(u) \in M_n(k)$ la matriz del endomorfismo $u: V \to V$ respecto a la base $\mathcal{B}$ de V. Si $\mathcal{B}'$ es otra base de V y si $P = \mathrm{Mat}_{\mathcal{B}}(\mathcal{B}')$ es la matriz de cambio de base de $\mathcal{B}$ a $\mathcal{B}'$, entonces

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

Def: Dos matrices $A, B \in M_n(k)$ son semejantes si $\exists P \in Gl_n(k)$ tq $P^{-1}AP = B$. Es decir, A y B representan (en bases diferentes) el mismo endomorfismo de k^n.

⚠ Veremos que la relación de semejanza es mucho más fina que aquella dada por el rango.

6. Operaciones elementales sobre filas y columnas:

Sea $A \in M_{m\times n}(k)$, y denotemos por $(e_1,\dots,e_n)$ y $(f_1,\dots,f_m)$ las bases canónicas de k^n y k^m, respectivamente.

Reducción de columnas: Consideremos las operaciones elementales dadas por

a) Intercambiar las columnas C_i y C_j de $A \Leftrightarrow$ Multiplicar a la derecha por la matriz de permutación $P(i,j)$ asociada al automorfismo de k^n dado por $\begin{cases} e_i \mapsto e_j \\ e_j \mapsto e_i \\ e_l \mapsto e_l \text{ si } l \neq i,j \end{cases}$. E.g. $(n=3)$: $P(2,3) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$

b) Multiplicar C_j por $\lambda_j \neq 0 \Leftrightarrow$ Mult. a la derecha por $\begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & \lambda_j & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix}$ (λ_j en la j-ésima columna)

c) Adicionar λC_i a $C_j \Leftrightarrow$ Mult. a la derecha por $B_{ij}(\lambda) := I_n + \lambda E_{ij}$, i.e., por $\begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & \lambda & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix}$ (λ en la (i,j)-ésima posición)

Ejercicio Las operaciones ⓐ, ⓑ y ⓒ son automorfismos de k^n. (e.v. de partida)

Recuerdo: Al realizar operaciones elementales sobre columnas obtenemos una matriz de la forma

$$A' = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 1 & \vdots & & \vdots & \vdots & & \vdots \\ * & 0 & & \vdots & \vdots & & \vdots \\ \vdots & 1 & & \vdots & \vdots & & \vdots \\ * & * & & * & 0 & & 0 \end{pmatrix}$$

donde las columnas C_j' son nulas si $j > r$, y donde $C_1', \dots, C_r' \in k^m$ son l.i. Es decir $\mathrm{Im}(A) = \mathrm{Im}(A') = \mathrm{Vect}_k\langle C_1', \dots, C_r' \rangle$ y $r = \mathrm{rg}(A)$.

La matriz A' obtenida por reducción de columnas verifica $A' = AP$, donde P es la matriz (invertible) que corresponde a las operaciones elementales efectuadas. Dado que $\ker(A') = \mathrm{Vect}_k\langle e_{r+1}, \dots, e_n \rangle$, la igualdad $A' = AP$ implica

$$\ker(A) = \mathrm{Vect}_k\langle P_{r+1}, \dots, P_n \rangle$$

es el espacio generado por las últimas $(n-r)$ columnas de P.

Ejemplo concreto: Sea $A = \begin{pmatrix} 0 & -1 & 4 \\ 2 & 2 & -3 \end{pmatrix} \in M_{2\times 3}(\mathbb{Q})$ asociada a la aplicación $u: \mathbb{Q}^3 \to \mathbb{Q}^2$, $(x,y,z) \mapsto (-y+4z, 2x+2y-3z)$. La reducción de columnas ("a la noté")

$$\left(\begin{array}{ccc} 0 & -1 & 4 \\ \hline 2 & 2 & -3 \\ \hline 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right) \xrightarrow{C_2 \mapsto \frac{1}{2}C_2} \left(\begin{array}{ccc} 0 & -1 & 4 \\ \hline 1 & 2 & -3 \\ \hline 1/2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right) \xrightarrow[C_3 \mapsto C_3+3C_1]{C_2 \mapsto C_2-2C_1} \left(\begin{array}{ccc} 0 & -1 & 4 \\ \hline 1 & 0 & 0 \\ \hline 1/2 & -1 & 3/2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right) \xrightarrow{C_2 \mapsto -C_2} \left(\begin{array}{ccc} 0 & 1 & 4 \\ \hline 1 & 0 & 0 \\ \hline 1/2 & 1 & 3/2 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{array}\right) \xrightarrow{C_3 \mapsto C_3-4C_2} \left(\begin{array}{ccc} 0 & 1 & 0 \\ \hline 1 & 0 & 0 \\ \hline 1/2 & 1 & -5/2 \\ 0 & -1 & 4 \\ 0 & 0 & 1 \end{array}\right) \begin{matrix} \} A' \\ \} P \end{matrix}$$

$\Rightarrow \mathrm{Im}(A) = \mathrm{Vect}_{\mathbb{Q}}\left\langle \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle = \mathbb{Q}^2$ (i.e., u es sobreyectiva).

Además, $\ker(A) = \mathrm{Vect}_{\mathbb{Q}}\left\langle \begin{pmatrix} -5/2 \\ 4 \\ 1 \end{pmatrix} \right\rangle = \mathrm{Vect}_{\mathbb{Q}}\left\langle \begin{pmatrix} -5 \\ 8 \\ 2 \end{pmatrix} \right\rangle$ (i.e., $\dim_{\mathbb{Q}} \ker(A) = 1$).

Reducción de filas: Consideremos las operaciones elementales dadas por:

a) Intercambiar las filas F_i y F_j de $A \Leftrightarrow$ Multiplicar a la izquierda por $P(i,j)$.

b) Multiplicar F_i por $\lambda_i \neq 0 \Leftrightarrow$ Mult. a la izquierda por $\begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & \lambda_i & \\ & & & \ddots \end{pmatrix}$ ← i-ésima fila

c) Adicionar λF_i a $F_j \Leftrightarrow$ Mult. a la izquierda por $B_{ji}(\lambda) = I_m + \lambda E_{ji}$.

De manera análoga que para la reducción de columnas, las operaciones (a), (b) y (c) son automorfismos de k^m (e.v. de llegada).

Recuerdo: Al realizar operaciones elementales sobre ~~columnas~~ filas obtenemos una matriz de la forma

$$A'' = \begin{pmatrix} 0 & 1 & * & * & \cdots & * \\ 0 & 0 & 0 & 1 & * & \cdots \\ \vdots & & & & & \\ 0 & \cdots & \cdots & \cdots & 1 & * \\ 0 & \cdots & \cdots & \cdots & 0 & \cdots \\ 0 & \cdots & \cdots & \cdots & 0 & \cdots \end{pmatrix}$$

donde las filas F_j'' de A'' son nulas si $j > r$ y las filas $F_1'', \ldots, F_r'' \in k^n$ son l.i.

La matriz A'' obtenida por reducción de filas verifica $A'' = QA$, donde $Q \in GL_m(k)$ corresp. a las operaciones fila ejecutadas. En part, $\ker(A) = \ker(A'')$ y $\operatorname{rg}(A) = \operatorname{rg}(A'') = r$.

Más aún, el sistema lineal $A''\begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} = 0$ provee ecuaciones para $\ker(A)$ y si denotamos $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} \in k^m$ y $Y'' = QY$, entonces $y_{r+1}'' = \ldots = y_m'' = 0$ provee ecuaciones para $\operatorname{Im}(A)$.

Ejemplo concreto: Sea $A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 6 & 6 \\ 3 & 8 & 15 & 16 \end{pmatrix}$ asociada a la aplicación

$u : \mathbb{R}^4 \to \mathbb{R}^3$, $(x,y,z,t) \mapsto (x+2y+3z+4t,\ x+3y+6z+6t,\ 3x+8y+15z+16t)$.

Reducción de filas:

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

$$\xrightarrow{F_3 \mapsto F_3 - 2F_2} \left(\begin{array}{cccc|c} 1 & 2 & 3 & 4 & y_1 \\ 0 & 1 & 3 & 2 & y_2 - y_1 \\ 0 & 0 & 0 & 0 & y_3 - y_1 - 2y_2 \end{array}\right).$$

A'' $\qquad$ Y''

Luego: $\ker(A) = \begin{cases} x_1 + 2x_2 + 3x_3 + 4x_4 = 0 \\ x_2 + 3x_3 + 2x_4 = 0 \end{cases}$. **Base**: x_1 y x_2 están determinadas por x_3 y x_4.

$\Rightarrow u = \begin{pmatrix} 3 \\ -3 \\ 1 \\ 0 \end{pmatrix}$ y $v = \begin{pmatrix} 0 \\ -2 \\ 0 \\ 1 \end{pmatrix}$ generan $\ker(A)$

$\uparrow$ $x_3 = 1, x_4 = 0$ $\qquad$ $\uparrow$ $x_3 = 0, x_4 = 1$

$y_3'' = 0 \Leftrightarrow y_3 = y_1 + 2y_2$ (Ecuación de $\operatorname{Im}(A) \subseteq \mathbb{R}^3$)

Base: y_3 está det. por y_1 & $y_2 \Rightarrow \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ y $\begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$ generan $\operatorname{Im}(A)$.

$\uparrow$ $y_1 = 1, y_2 = 0$ $\qquad$ $\uparrow$ $y_1 = 0, y_2 = 1$

⚠ Tal como se discutió en primer año, el proceso de reducción de columnas (resp. filas) permite calcular A^{-1} si $A \in Gl_n(k)$, puesto que (con la notación anterior) la reducción se reduce en este caso a la igualdad $A' = AP = I_n$ (resp. $A'' = QA = I_n$) de donde $P = A^{-1}$ (resp. $Q = A^{-1}$). Es muy importante escoger hacer operaciones sobre columnas o sobre filas, pero no mezclarlas !

Ejercicio Sea $A \in Gl_n(k)$ y supongamos que $AB = BA$ para toda $B \in Gl_n(k)$. Probar que $A = \lambda I_n$ para cierto $\lambda \neq 0$.

§7. Polinomios de endomorfismos :

Sea V un k-e.v. y sea $u : V \to V$ un endomorfismo.

Notación: $u^0 := id_V$, $u^1 := u$, $u^2 = u \circ u$, ..., $u^n := \underbrace{u \circ \dots \circ u}_{n \text{ veces}}$

De manera general, si $P \in k[X]$ es un polinomio dado por $P(X) = a_0 + a_1 X + a_2 X^2 + \dots + a_d X^d$, entonces definimos $P(u) \in End_k(V)$ mediante :

$$P(u) := a_0 id_V + a_1 u + a_2 u^2 + \dots + a_d u^d .$$

La definición anterior y la definición de composición de funciones implican :

Lema: La aplicación $k[X] \to End_k(V)$, $P \mapsto P(u)$ es un morfismo de k-álgebras, ie, es k-lineal y si $P, Q \in k[X]$ entonces $PQ \mapsto P(u)Q(u)$ ▩

Ejemplo : $P(X) = X^4 - 1$.

1) Si $k = \mathbb{R}$, $P(X) = (X-1)(X+1)(X^2+1)$, por lo que

$$u^4 - id_V = (u - id_V)(u + id_V)(u^2 + id_V) \quad \text{para todo } u \in End_{\mathbb{R}}(V).$$

2) Si $k = \mathbb{C}$, $P(X) = (X-1)(X+1)(X+i)(X-i)$, por lo que

$$u^4 - id_V = (u - id_V)(u + id_V)(u + i\, id_V)(u - i\, id_V) \quad \text{para todo } u \in End_{\mathbb{C}}(V).$$

Obs: La construcción anterior funciona del mismo modo si consideramos $A \in M_n(k)$ y definimos $P(A) := a_0 I_n + a_1 A + a_2 A^2 + \dots + a_d A^d$.

Prop: Para todo $u \in End_k(V)$ endomorfismo, donde $\dim_k(V) = n$ es finita, existe $P \in k[X]$ tq $P \neq 0$ y $P(u) = 0$.

Dem: Como $\dim_k End_k(V) = n^2$, los vectores $id_V, u, u^2, \dots, u^{n^2}$ de $End_k(V)$ son l.d. $\Rightarrow \exists\, a_0, a_1, \dots, a_{n^2} \in k$ no todos nulos tq $a_0 id_V + a_1 u + \dots + a_{n^2} u^{n^2} = 0$. Luego, basta escoger $P(X) := a_0 + a_1 X + \dots + a_{n^2} X^{n^2}$. ▩

Ejercicio Sea $A = \begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix} \in M_2(\mathbb{R})$. Hallar $P \in \mathbb{R}[X]$ no-nulo tq $P(A) = 0$.

§8. Permutaciones:

Def: Sea $n \in \mathbb{N}^{\geq 1}$. Una permutación del conjunto $\{1,2,\dots,n\}$ es una función biyectiva $\sigma : \{1,\dots,n\} \to \{1,\dots,n\}$, la cual denotaremos

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & & \sigma(n) \end{pmatrix} \quad \text{o bien} \quad \sigma = (\sigma(1), \sigma(2), \dots, \sigma(n)).$$

Ejemplo: La permutación $\sigma = (2,3,4,1)$ corresponde a $\sigma : \begin{cases} 1 \mapsto 2 \\ 2 \mapsto 3 \\ 3 \mapsto 4 \\ 4 \mapsto 1 \end{cases}$

Notación: El conjunto de todas las permutaciones de $\{1,\dots,n\}$

$$S_n = \{\sigma : \{1,\dots,n\} \to \{1,\dots,n\} \text{ biyección}\}$$

es llamado el **grupo simétrico**. Es conjunto es en efecto un grupo: si $\sigma, \tau \in S_n$ entonces $\sigma\tau := \sigma \circ \tau$ (composición) pertenece a S_n, la función identidad id pertenece a S_n, y si $\sigma \in S_n$ entonas su inversa $\sigma^{-1} \in S_n$.

Ejemplo: Si $\sigma = (2,3,4,1)$ y $\tau = (2,1,4,3)$, entonces $\tau\sigma = (1,4,3,2)$ y $\sigma\tau = (3,2,1,4)$. Además, podemos calcular σ^{-1} "gráficamente":

$$\sigma : \begin{cases} 1 \mapsto 2 \\ 2 \mapsto 3 \\ 3 \mapsto 4 \\ 4 \mapsto 1 \end{cases} \Rightarrow \sigma^{-1} : \begin{cases} 1 \mapsto 4 \\ 2 \mapsto 1 \\ 3 \mapsto 2 \\ 4 \mapsto 3 \end{cases}, \text{ ie, } \sigma^{-1} = (4,1,2,3).$$

Ejercicio Verifican que $S_3 = \{id = (1,2,3), (2,1,3), (3,2,1), (1,3,2), (2,3,1), (3,1,2)\}$

Prop: $\text{card}(S_n) = n!$

Dem: Inducción en n: $S_1 = \{(1)\}$ ✓. Sup. que $\text{card}(S_n) = n!$ para cierto $n \in \mathbb{N}^{\geq 1}$ y consideremos S_{n+1}:

Para $k \in \{1,\dots,n+1\}$, sea $A_k = \text{card}(\{\sigma = (a_1,\dots,a_{n+1}) \in S_{n+1} \text{ tq } a_k = n+1\})$.

La correspondencia biyectiva

$$\{\sigma = (a_1,\dots,a_{n+1}) \in S_{n+1} \text{ tq } a_k = n+1\} \xleftrightarrow{\sim} S_n$$

$$\sigma \longmapsto (a_1,\dots,a_{k-1},a_{k+1},\dots,a_{n+1}) \in S_n$$

$$(b_1,\dots,b_{k-1},n+1,b_k,\dots,b_n) \longleftarrow (b_1,\dots,b_n) \in S_n$$

y la hipótesis de inducción, implican que $A_k = |S_n| = n!$.

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

Def: Una permutación $\tau \in S_n$ que sólo cambia dos elementos de $\{1,\dots,n\}$ es llamada una **transposición**.

Notación: Para $n \geq 2$ y todos $i,j \in \{1,\dots,n\}$ con $i \neq j$, denotaremos por $\tau = (i,j)$ a la transposición $\tau : \begin{cases} i \mapsto j \\ j \mapsto i \\ k \mapsto k \quad \forall k \neq i,j \end{cases}$.

Con esta notación, tenemos que $(i,j) = (j,i) = (i,j)^{-1}$.

Ejemplo: $S_3 = \{id, (1,2), (2,3), (1,3), (2,3,1), (3,1,2)\}$.

Def: Sea $\sigma \in S_n$ y sean $i,j \in \mathbb{N}$ tq $1 \le i < j \le n$. Decimos que σ invierte i y j si $\sigma(i) > \sigma(j)$.

Ejemplo: 1) $id = (1,2,..,n) \in S_n$ tiene 0 inversiones.
2) La transposición $(1,2) \in S_n$ tiene 1 inversión.
3) $(2,3,1)$ y $(3,1,2)$ en S_3 tiene 2 inversiones.

Ejercicio Demostrar que la transposición $(i,j) \in S_n$ tiene $2|i-j|-1$ inversiones.

Def: Sea $\sigma \in S_n$. El número

$$\varepsilon(\sigma) := (-1)^{n^\circ \text{ de inversiones de } \sigma}$$

es llamado la signatura de σ. Decimos que σ es par (resp. impar) si $\varepsilon(\sigma) = 1$ (resp. $\varepsilon(\sigma) = -1$).

Ejemplo: 1) $\varepsilon(id) = (-1)^0 = 1$ (par)
2) Si $(i,j) \in S_n$ transposición, entonces $\varepsilon((i,j)) = (-1)^{2|i-j|-1} = -1$ (impar).
3) En S_3 hay 3 permutaciones pares y 3 impares.

Lema: Sea $\sigma \in S_n$, entonces $\varepsilon(\sigma) = \prod_{1 \le i < j \le n} \frac{\sigma(j) - \sigma(i)}{j - i}$.

Dem: La cantidad $\prod_{i<j} \frac{\sigma(j)-\sigma(i)}{j-i}$ tiene el mismo signo que $\varepsilon(\sigma)$. Además,

$$\left(\prod_{i<j} \frac{\sigma(j)-\sigma(i)}{j-i}\right)^2 = \prod_{i<j} \left(\frac{\sigma(j)-\sigma(i)}{j-i}\right)^2 = \frac{\prod_{i \ne j} (\sigma(j)-\sigma(i))^2}{\prod_{i \ne j} (j-i)^2} = 1$$

($\uparrow$ σ biyección) ■

Prop: La signatura $\varepsilon : S_n \to \{\pm 1\}$ verifica $\varepsilon(\sigma\tau) = \varepsilon(\sigma)\varepsilon(\tau)$ para todos $\sigma, \tau \in S_n$.

Dem: Sean $\sigma, \tau \in S_n$. Como $\frac{\sigma(j)-\sigma(i)}{j-i} = \frac{\sigma(i)-\sigma(j)}{i-j}$, esta cantidad no depende de i y j.

Luego, $\varepsilon(\sigma) = \prod_{i<j} \frac{\sigma(\tau(j)) - \sigma(\tau(i))}{\tau(j) - \tau(i)} \Rightarrow \varepsilon(\sigma\tau) = \prod_{i<j} \frac{\sigma(\tau(j)) - \sigma(\tau(i))}{j-i}$.

$$= \prod_{i<j} \frac{\sigma(\tau(j)) - \sigma(\tau(i))}{\tau(j) - \tau(i)} \cdot \prod_{i<j} \frac{\tau(j)-\tau(i)}{j-i}$$

$$= \varepsilon(\sigma)\,\varepsilon(\tau)$$ ■

Obs: Notar que los elementos $\{id, (1,2), (2,3), (1,3), (2,3,1), (3,1,2)\}$ de S_3 pueden ser escritos como productos de transposiciones: $id = (1,2)(1,2)$; $(2,3,1) = (1,3)(1,2)$ y $(3,1,2) = (1,2)(1,3) = (1,3)(2,3)$ (la escritura no es única). En general:

Prop: Toda $\sigma \in S_n$ se escribe (de forma no única) como producto de transposiciones. Toda permutación par (resp. impar) se descompone necesariamente en un número par (resp. impar) de transposiciones.

Idea: Dada $\sigma \in S_n$, multiplicar a la izquierda por transposiciones para ir obteniendo cada vez más elementos fijos. Por ejemplo:

$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 1 & 5 & 3 \end{pmatrix} \underset{\sigma(1)=2}{\Rightarrow} (1,2)\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 4 & 2 & 5 & 3 \end{pmatrix} \underset{\sigma(2)=4}{\Rightarrow} (2,4)(1,2)\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 4 & 5 & 3 \end{pmatrix}$

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

Finalmente, si $\sigma = \tau_1 \cdots \tau_r \Rightarrow \varepsilon(\sigma) = \varepsilon(\tau_1)\cdots\varepsilon(\tau_r) = (-1)^r \Rightarrow \sigma$ y r tienen misma paridad ■

§9. Formas multilineales alternadas

Vimos que a todo endomorfismo $u: V \to V$ de un k-e.v. de dimensión finita se le puede asociar una matriz en $M_n(k)$, pero ella depende de la elección de una base.

Objetivo: Asociar a u un invariante, ie, algo que sólo depende de u.

Obs: Para $n \in \mathbb{N}^{\geq 2}$ tenemos que $n \cdot 1 = \underbrace{1 + \ldots + 1}_{n \text{ veces}}$ en k.

Notar que $p \cdot 1 = 0$ en $\mathbb{F}_p$, pero $\nexists\, n \cdot 1 \neq 0$ en $\mathbb{Q}$ para todo $n \in \mathbb{N}^{\geq 2}$

Def: Sea k cuerpo. Si $\exists n \in \mathbb{N}^{\geq 2}$ tq $n \cdot 1 = 0$ en k, entonces la característica de k es el menor entero $\mathrm{car}(k) \in \mathbb{N}^{\geq 2}$ con esta propiedad. Por otro lado, si $n \cdot 1 \neq 0$ para todo $n \in \mathbb{N}^{\geq 2}$, decimos que $\mathrm{car}(k) = 0$.

Hecho (MAT214): Sea k un cuerpo. Entonces:
i) ya sea $\mathrm{car}(k) = 0$, y en tal caso $\mathbb{Q} \subseteq k$; o bien
ii) $\mathrm{car}(k) = p$ es un número primo, y en tal caso $\mathbb{F}_p \subseteq k$.

Ejemplo: $\mathrm{car}(\mathbb{F}_p) = p$ y $\mathrm{car}(\mathbb{Q}) = \mathrm{car}(\mathbb{R}) = \mathrm{car}(\mathbb{C}) = 0$.

⚠ En todo lo que sigue, supondremos siempre que $\mathrm{car}(k) \neq 2$, ie, $2 \neq 0 \iff 1 \neq -1$. (e.g. $k = \mathbb{Q}$).

Def: Sea V un k-e.v. Una forma m-multilineal sobre V es una función

$$F: V^m = \underbrace{V \times \ldots \times V}_{m \text{ veces}} \longrightarrow k$$
$$(x_1, \ldots, x_m) \longmapsto F(x_1, \ldots, x_m)$$

que es lineal en cada variable. Diremos que F es alternada si

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

cuando dos de los x_j son iguales.

Ejemplo: Sea $F: \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ forma bilineal (ie, $m=2$) sobre $\mathbb{R}^2$, y sea $e_1 = (1,0)$, $e_2 = (0,1)$ la base canónica. Luego:

$$F((x_{11}, x_{21}), (x_{12}, x_{22})) = F(x_{11}e_1 + x_{21}e_2,\ x_{12}e_1 + x_{22}e_2)$$
$$= x_{11} F(e_1, x_{12}e_1 + x_{22}e_2) + x_{21} F(e_2, x_{12}e_1 + x_{22}e_2) \quad [F \text{ lineal en } 1^a \text{ variable}]$$
$$= x_{11}\left(x_{12}\underbrace{F(e_1,e_1)}_{:=a} + x_{22}\underbrace{F(e_1,e_2)}_{:=b}\right) + x_{21}\left(x_{12}\underbrace{F(e_2,e_1)}_{:=c} + x_{22}\underbrace{F(e_2,e_2)}_{:=d}\right) \quad [F \text{ lineal } 2^a \text{ variable}]$$
$$= a x_{11}x_{12} + b x_{11}x_{22} + c x_{21}x_{12} + d x_{21}x_{22}$$

Recíprocamente, toda función de la forma anterior es bilineal.

Notar que F es alternada $\iff F((x_{11},x_{21}),(x_{11},x_{21})) = 0 \quad \forall (x_{11},x_{21}) \in \mathbb{R}^2$

$\iff a x_{11}^2 + (b+c) x_{11} x_{21} + d x_{21}^2 = 0 \ \forall \ (x_{11},x_{21}) \in \mathbb{R}^2$

$\iff a = b+c = d = 0$

Luego, F se escribe como $F((x_{11},x_{21}),(x_{12},x_{22})) = b(x_{11}x_{22} - x_{21}x_{12})$

o equivalentemente $F\left(\begin{pmatrix}x_{11}\\x_{21}\end{pmatrix},\begin{pmatrix}x_{12}\\x_{22}\end{pmatrix}\right) = b\,(x_{11}x_{22} - x_{21}x_{12})$ ¡"Determinante 2×2"!

Prop: Sea V un k-e.v. y $F: V^m \to k$ una forma multilineal. Entonces,

F alternada $\iff F(x_1,\dots,x_i,\dots,x_j,\dots,x_m) = -F(x_1,\dots,x_j,\dots,x_i,\dots,x_m)$

para todos $x_1,\dots,x_m \in V$ y $1 \le i < j \le m$.

Dem: ($\Leftarrow$) Si $x_i = x_j = x$, entonces $F(x_1,\dots,x,\dots,x,\dots,x_m) = -F(x_1,\dots,x,\dots,x,\dots,x_m)$

ie, $2F(x_1,\dots,x,\dots,x,\dots,x_m) \underset{\mathrm{car}(k)\neq 2}{\Rightarrow} F(x_1,\dots,x,\dots,x,\dots,x_m) = 0$, ie, F alternada.

($\Rightarrow$) $0 = F(x_1,\dots,x_i+x_j,\dots,x_i+x_j,\dots,x_m)$ [F alternada]

$= F(x_1,\dots,x_i,\dots,x_i,\dots,x_m) + F(x_1,\dots,x_i,\dots,x_j,\dots,x_m)$ [F bilineal]

$+ F(x_1,\dots,x_j,\dots,x_i,\dots,x_m) + F(x_1,\dots,x_j,\dots,x_j,\dots,x_m)$

$= F(x_1,\dots,x_i,\dots,x_j,\dots,x_m) + F(x_1,\dots,x_j,\dots,x_i,\dots,x_m)$ [F alternada]

$\Rightarrow F(x_1,\dots,x_i,\dots,x_j,\dots,x_m) = -F(x_1,\dots,x_j,\dots,x_i,\dots,x_m)$ ∎

Obs: La proposición anterior equivale a decir que F es alternada si y sólo si para toda <u>transposición</u> $\tau \in S_m$ se tiene

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

Prop: Sea V un k-e.v. y sea $F: V^m \to k$ una forma multilineal <u>alternada</u>. Entonces:

1) Para toda $\sigma \in S_m$, tenemos

$$F(x_{\sigma(1)},\dots,x_{\sigma(m)}) = \varepsilon(\sigma) F(x_1,\dots,x_m).$$

2) $F(x_1,\dots,x_m)$ <u>no</u> cambia si adicionamos a x_i una combinación lineal de las otras variables.

3) $F(x_1,\dots,x_m) = 0$ si los x_i son l.d. En particular, <u>toda</u> forma m-multilineal alternada es nula si $m > \dim_k(V)$.

Dem: Descomponemos $\sigma = \tau_1 \cdots \tau_k$ como producto de transposiciones, con $\varepsilon(\sigma) = (-1)^k$. El punto ① se obtiene por inducción en k. Para ②, notamos que

$$F(x_1,\dots,x_{i-1},x_i + \sum_{j\neq i}\lambda_j x_j, x_{i+1},\dots,x_m) \underset{F \text{ multilineal}}{=} F(x_1,\dots,x_m) + \sum_{j\neq i}\lambda_j F(x_1,\dots,x_j,\dots,x_j,\dots,x_m) \underset{F \text{ alternada}}{=} F(x_1,\dots,x_m)$$ ✓

Si los x_i son l.d., uno de ellos es combinación lineal de los otros. Por ②, podemos adicionar a dicho vector el opuesto aditivo de esta combinación lineal sin cambiar el valor de F, que es por lo tanto cero. Esto prueba ③ ∎

Sea V un k-e.v. de $\dim_k(V) = n$ y F una forma n-multilineal alternada.

Objetivo: Nos gustaría "expandir" $F(x_1, \dots, x_n)$ como en el ejemplo de $F: \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$.

Sea $\mathcal{B} = (e_1, \dots, e_n)$ una base de V y escribamos $x_j = \sum_{i=1}^{n} x_{ij} e_i$.

Por multilinealidad, tenemos que:

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

$$= \sum_{1 \le i_1, \dots, i_n \le n} x_{i_1 1} x_{i_2 2} \cdots x_{i_n n} F(e_{i_1}, e_{i_2}, \dots, e_{i_n})$$

Dado que F es alternada, los términos donde índices $i_1, \dots, i_n$ son los mismos se anulan. Luego, los únicos términos que quedan son aquellos en que todos los índices son distintos, i.e., para los cuales la aplicación $k \mapsto i_k$ es inyectiva, y por ende una permutación $\sigma \in S_n$. Así,

$$F(x_1, \dots, x_n) = \sum_{\sigma \in S_n} x_{\sigma(1)1} \cdots x_{\sigma(n)n} F(e_{\sigma(1)}, \dots, e_{\sigma(n)})$$

$$\overset{\text{Prop. anterior}}{=} \sum_{\sigma \in S_n} x_{\sigma(1)1} \cdots x_{\sigma(n)n} \varepsilon(\sigma) F(e_1, \dots, e_n)$$

$$= F(e_1, \dots, e_n) \sum_{\sigma \in S_n} \varepsilon(\sigma) x_{\sigma(1)1} \cdots x_{\sigma(n)n}.$$

Def: Sean V un k-e.v. de $\dim_k(V) = n$ y $\mathcal{B} = (e_1, \dots, e_n)$ una base de V. Definimos el <u>determinante de los vectores $x_1, \dots, x_n \in V$ en la base $\mathcal{B}$</u> como

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

donde $x_j = \sum_{i=1}^{n} x_{ij} e_i$ para $j = 1, \dots, n$.

Luego, la discusión anterior puede resumirse en:

Lema: Sea V un k-e.v. de $\dim_k(V) = n$ y sea $\mathcal{B} = (e_1, \dots, e_n)$ una base de V. Entonces para toda forma n-multilineal alternada F se tiene que

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

En otras palabras, <u>todas</u> las formas n-multilineales alternadas sobre V son proporcionales a $\det_{\mathcal{B}}: V^n \to k$. ■

Ejemplo: Sea $\mathcal{B} = ((1,0),(0,1))$ la base canónica de k^2, entonces:

$$\det_{\mathcal{B}}\big(\underbrace{(x_{11}, x_{21})}_{x_1 \in k^2}, \underbrace{(x_{12}, x_{22})}_{x_2 \in k^2}\big) = \sum_{\sigma \in S_2} \varepsilon(\sigma) x_{\sigma(1)1} x_{\sigma(2)2} = x_{11} x_{22} - x_{21} x_{12}.$$

Ejercicio Sea $\mathcal{B}$ la base canónica de k^3. Calcular $\det_{\mathcal{B}}(x_1, x_2, x_3)$.

Prop: La aplicación $\det_{\mathcal{B}} : V^n \to k$ es una forma n-multilineal alternada que cumple $\det_{\mathcal{B}}(e_1, \dots, e_n) = 1$.

Dem: Notar que cada una de las expresiones $\varepsilon(\sigma)\, x_{\sigma(1)1} \cdots x_{\sigma(n)n}$ que intervienen en la definición de $\det_{\mathcal{B}}(x_1, \dots, x_n)$ es n-multilineal en $(x_1, \dots, x_n)$ y luego su suma, $\det_{\mathcal{B}}$, lo es también. Veamos que es alternada:

Sea $\tau = (i, j)$ y consideremos $\det_{\mathcal{B}}(x_1, \dots, \underset{\substack{\uparrow \\ i\text{-ésima} \\ \text{posición}}}{x_j}, \dots, \underset{\substack{\uparrow \\ j\text{-ésima} \\ \text{posición}}}{x_i}, \dots, x_n)$ que vale:

$$\sum_{\sigma \in S_n} \varepsilon(\sigma)\, x_{\sigma(1)1} \cdots x_{\sigma(i)j} \cdots x_{\sigma(j)i} \cdots x_{\sigma(n)n}$$

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

$$= \sum_{\sigma \in S_n} -\varepsilon(\sigma\circ\tau)\, x_{\sigma\circ\tau(1)1} \cdots x_{\sigma\circ\tau(n)n} = -\det_{\mathcal{B}}(x_1, \dots, x_n)$$

y luego $\det_{\mathcal{B}} : V^n \to k$ es alternada. Finalmente, si $(x_1, \dots, x_n) = (e_1, \dots, e_n)$ entonces $x_{ij} = \delta_{ij} := \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \end{cases}$. Así, el término $x_{\sigma(1)1} \cdots x_{\sigma(n)n}$ es no-nulo sólo si $\sigma(i) = i \ \forall i$, o sea, $\sigma = \mathrm{id} \Rightarrow \det_{\mathcal{B}}(e_1, \dots, e_n) = 1$ ■

Teorema: Sea $\mathcal{B}$ una base de un k-e.v. V de dimensión n. Luego:

1. $\det_{\mathcal{B}}(x_1, \dots, x_n)$ es lineal en cada x_i.
2. $\det_{\mathcal{B}}(x_1, \dots, x_n) = 0$ si dos de los x_i son iguales.
3. Para toda $\sigma \in S_n$ se tiene $\det_{\mathcal{B}}(x_{\sigma(1)}, \dots, x_{\sigma(n)}) = \varepsilon(\sigma) \det_{\mathcal{B}}(x_1, \dots, x_n)$. En part, el determinante cambia de signo si intercambiamos dos de los x_i.
4. $\det_{\mathcal{B}}(x_1, \dots, x_n)$ no cambia cuando adicionamos a un x_i una combinación lineal de los otros.
5. Si $\mathcal{B}'$ es otra base de V, entonces
$$\det_{\mathcal{B}'}(x_1, \dots, x_n) = \det_{\mathcal{B}'}(e_1, \dots, e_n) \det_{\mathcal{B}}(x_1, \dots, x_n).$$
6. $\det_{\mathcal{B}}(x_1, \dots, x_n) \neq 0 \iff (x_1, \dots, x_n)$ es una base de V.

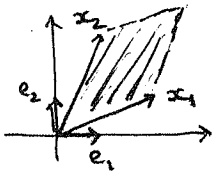
Dem: (1) y (2) son la definición de una aplicación multilineal alternada. (3) y (4) fueron probados en §9. El punto (5) sigue del Lema anterior aplicado a $F = \det_{\mathcal{B}'}$. Veamos (6): Si $(x_1, \dots, x_n)$ es una base $\mathcal{B}'$, entonces (5) da:

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

Recíprocamente, si los vectores $x_1, \dots, x_n$ son l.d., entonces $\det_{\mathcal{B}}(x_1, \dots, x_n) = 0$ pues $\det_{\mathcal{B}}$ es multilineal alternada ■

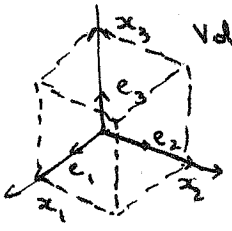
Interpretación geométrica: Sea $\mathcal{B}$ la base canónica de $\mathbb{R}^n$, y sean $x_1, \dots, x_n \in \mathbb{R}^n$ vectores. Entonces el valor absoluto del determinante, $|\det_{\mathcal{B}}(x_1, \dots, x_n)|$ es el **volumen** del paralelepípedo n-dimensional construido a partir de $x_1, \dots, x_n$.

$\boxed{n=2}$



$\text{Área} = |\det_{\mathcal{B}}(x_1, x_2)|$

$\boxed{n=3}$

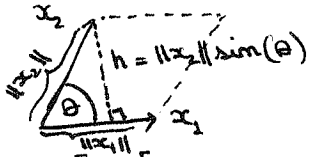


$\text{Volumen} = |\det_{\mathcal{B}}(x_1, x_2, x_3)|$

En particular, $\det_{\mathcal{B}}(x_1, \dots, x_n) = 0 \iff$ los $x_1, \dots, x_n$ son l.d. (el paralelepípedo es "plano").

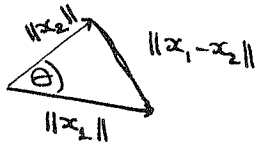
Idea $(n=2)$: Si $x_1 = (a,b)$ y $x_2 = (c,d)$ entonces:

$$\det_{\mathcal{B}}(x_1, x_2)^2 = (ad-bc)^2 = a^2d^2 - 2abcd + b^2c^2 = (a^2+b^2)(c^2+d^2) - (ac+bd)^2$$
$$= \|x_1\|^2 \|x_2\|^2 - (x_1 \cdot x_2)^2$$



$\theta \in [0, \pi[$

Por otra parte (teorema del coseno):



$$\|x_1 - x_2\|^2 = \|x_1\|^2 + \|x_2\|^2 - 2\|x_1\| \|x_2\| \cos(\theta)$$
$$\|x_1 - x_2\|^2 = (x_1 - x_2) \cdot (x_1 - x_2) = x_1 \cdot x_1 - 2(x_1 \cdot x_2) + x_2 \cdot x_2$$
$$= \|x_1\|^2 + \|x_2\|^2 - 2(x_1 \cdot x_2)$$

$$\Rightarrow (x_1 \cdot x_2) = \|x_1\| \|x_2\| \cos(\theta).$$

Luego, $|\det_{\mathcal{B}}(x_1, x_2)| = \sqrt{\|x_1\|^2 \|x_2\|^2 (1 - \cos^2(\theta))} = \|x_1\| \|x_2\| \sin(\theta) = 2 \cdot \left(\frac{1}{2} \|x_2\| h\right)$ ■

§11. Determinante de un endomorfismo

Sea V un $\mathbb{R}$-e.v. de $\dim_{\mathbb{R}}(V) = n$.

Teorema: Sea $u : V \to V$ un endomorfismo.

1. Sea $\mathcal{B} = (e_1, \dots, e_n)$ una base de V. Entonces, el escalar $\det_{\mathcal{B}}(u(e_1), \dots, u(e_n))$ no depende de la elección de $\mathcal{B}$; dicho escalar es denotado $\det(u)$ y es llamado el **determinante de u**.
2. Para todos $x_1, \dots, x_n \in V$ y toda base $\mathcal{B}$ de V, tenemos:
$$\det_{\mathcal{B}}(u(x_1), \dots, u(x_n)) = \det(u) \det_{\mathcal{B}}(x_1, \dots, x_n).$$

Dem: Para toda base $\mathcal{B} = (e_1, \dots, e_n)$ de V, escribamos $d(\mathcal{B}) := \det_{\mathcal{B}}(u(e_1), \dots, u(e_n))$.
Dado que u es lineal y $\det_{\mathcal{B}}$ multilineal alternada, la composición
$$(x_1, \dots, x_n) \mapsto \det_{\mathcal{B}}(u(x_1), \dots, u(x_n))$$
es n-multilineal alternada. Luego, el lema de §10 implica que
$$\det_{\mathcal{B}}(u(x_1), \dots, u(x_n)) = d(\mathcal{B}) \det_{\mathcal{B}}(x_1, \dots, x_n). \quad (*)$$
Por ende, basta probar que $d(\mathcal{B})$ es independiente de $\mathcal{B}$: Sea $\mathcal{B}'$ otra base y sea $\lambda := \det_{\mathcal{B}'}(e_1, \dots, e_n)$. El Teorema de §10 nos da, aplicado a $\mathcal{B}' = (e_1', \dots, e_n')$:
$$1 = \det_{\mathcal{B}'}(e_1', \dots, e_n') = \lambda \det_{\mathcal{B}}(e_1', \dots, e_n') \quad (**)$$
y
$$d(\mathcal{B}') \overset{\text{definición}}{=} \det_{\mathcal{B}'}(u(e_1'), \dots, u(e_n')) \overset{\text{Teo §10}}{=} \lambda \det_{\mathcal{B}}(u(e_1'), \dots, u(e_n')) \overset{(*)}{=} \lambda d(\mathcal{B}) \det_{\mathcal{B}}(e_1', \dots, e_n') \overset{(**)}{=} d(\mathcal{B})$$
■

Ejemplo: Una homotecia es un endomorfismo de la forma $h_\lambda : V \to V$, $x \mapsto \lambda x$, donde $\lambda \in k$ es llamado el factor de la homotecia.

Sea $\mathcal{B} = (e_1, \dots, e_n)$ una base de V, entonces:

$$\det(h_\lambda) = \det_{\mathcal{B}}(h_\lambda(e_1), \dots, h_\lambda(e_n)) = \det_{\mathcal{B}}(\lambda e_1, \lambda e_2, \dots, \lambda e_n) = \lambda^n \det_{\mathcal{B}}(e_1, \dots, e_n)$$

($\det_{\mathcal{B}}$ multilineal; $\det_{\mathcal{B}}(e_1, \dots, e_n) = 1$)

$$\Rightarrow \boxed{\det(h_\lambda) = \lambda^n}$$

Teorema: Sean $u, v \in \mathrm{End}_k(V)$ dos endomorfismos.

1. $\det(u \circ v) = \det(u)\det(v) = \det(v \circ u)$
2. u es un automorfismo ($u \in \mathrm{GL}(V)$) $\Leftrightarrow \det(u) \neq 0$. Más aún, $\det(u^{-1}) = \frac{1}{\det(u)}$.

Dem: Sea $\mathcal{B} = (e_1, \dots, e_n)$ una base de V. Entonces:

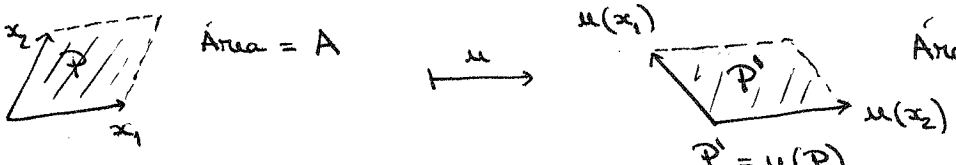
$$\det(u \circ v) = \det_{\mathcal{B}}(u(v(e_1)), \dots, u(v(e_n))) = \det(u) \det_{\mathcal{B}}(v(e_1), \dots, v(e_n)) = \det(u)\det(v)$$

de donde obtenemos (1).

Para que u sea biyectivo, es necesario y suficiente que $u(\mathcal{B}) = (u(e_1), \dots, u(e_n))$ sea una base de V. Por definición, $\det(u) = \det_{\mathcal{B}}(u(e_1), \dots, u(e_n))$. Por otro lado, el Teorema de §10 nos dice que $\det(u) \neq 0 \Leftrightarrow u(\mathcal{B})$ es una base. Finalmente, $u \circ u^{-1} = \mathrm{id}_V$ implica que $\det(u)\det(u^{-1}) = \det(\mathrm{id}_V) = 1$, gracias a (1). ■

Interpretación geométrica: Sea $\mathcal{B} = (e_1, \dots, e_n)$ la base canónica de $\mathbb{R}^n$ y sea $u : \mathbb{R}^n \to \mathbb{R}^n$ un endomorfismo. Entonces, $|\det(u)|$ es el volumen del paralelepípedo construido a partir de $u(e_1), \dots, u(e_n)$. El primer Teorema de esta sección dice que si consideramos la imagen de un paralelepípedo por u, su volumen se amplifica por $|\det(u)|$.

$n = 2$



⚠ Esto será muy útil para calcular integrales multiples

Importante: La función $\det : \mathrm{End}_k(V) \to k$ no es lineal:

En general, $\det(u+v) \neq \det(u) + \det(v)$. Además, $\det(\lambda u) = \lambda^n \det(u)$ para $\lambda \in k$ (pues $\lambda u = h_\lambda \circ u \Rightarrow \det(\lambda u) = \det(h_\lambda \circ u) = \underbrace{\det(h_\lambda)}_{\lambda^n}\det(u)$).

Ejercicio Sea $m \in \mathbb{N}^{\geq 1}$. Diremos que un endomorfismo $u : V \to V$ es de orden m si $u^m = \underbrace{u \circ \dots \circ u}_{m\text{-veces}}$ es igual a id_V. Probar que si $u \in \mathrm{End}_k(V)$ tiene orden m entonces $\det(u) \in k$ es una raíz m-ésima de 1, y deducir que $u \in \mathrm{GL}(V)$.

Def: Sea $A = (a_{ij})_{1 \le i,j \le n} \in M_n(k)$ matriz cuadrada. Definimos el **determinante de A**, denotado $\det(A)$ o $|A|$, por:

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

En otras palabras, $\det(A)$ es el determinante de los vectores columnas de A, ie, es el determinante de todo endomorfismo u t.q. $\mathrm{Mat}_{\mathcal{B}}(u) = A$ para cierta base $\mathcal{B}$.

Ejemplos:

1) $\det(I_n) = 1$ y $\det(\lambda I_n) = \lambda^n$ para $\lambda \in k$.

2) $\boxed{n=2}$ Si $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \Rightarrow \det(A) = a_{11}a_{22} - a_{12}a_{21}$

3) $\boxed{n=3}$ Si $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$ y $S_3 = \{ \underset{\varepsilon=1}{\mathrm{id}}, \underset{\varepsilon=-1}{(1,2)}, \underset{\varepsilon=-1}{(2,3)}, \underset{\varepsilon=-1}{(1,3)}, \underset{\varepsilon=1}{(2,3,1)}, \underset{\varepsilon=1}{(3,1,2)} \}$

$$\Rightarrow \det(A) = a_{11}a_{22}a_{33} - a_{21}a_{12}a_{33} - a_{11}a_{32}a_{23} - a_{31}a_{22}a_{13} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23}.$$

4) Como $\mathrm{card}(S_n) = n!$, si $A \in M_n(k)$ entonces $\det(A)$ tiene $n!$ términos !

Prop: Sea $A \in M_n(k)$. Entonces, $\det(A) = \det({}^tA)$.

Dem: Sea $A = (a_{ij})$ y ${}^tA = (b_{ij})$ con $b_{ij} = a_{ji}$. Luego, para $\sigma \in S_n$ tenemos

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

donde la última igualdad se deduce de $\varepsilon(\sigma) = \varepsilon(\sigma^{-1})$ $\boxed{\text{Ejercicio}}$ y de $a_{1\sigma(1)} \cdots a_{n\sigma(n)} = a_{\sigma^{-1}(1)1} \cdots a_{\sigma^{-1}(n)n}$ pues σ es una biyección. Luego:

$$\det({}^tA) = \sum_{\sigma \in S_m} \varepsilon(\sigma^{-1})\, a_{\sigma^{-1}(1)1} \cdots a_{\sigma^{-1}(n)n}$$

$$\overset{\text{cambio de variable } \tau = \sigma^{-1}}{=} \sum_{\tau^{-1} \in S_m} \varepsilon(\tau)\, a_{\tau(1)1} \cdots a_{\tau(n)n} \overset{\uparrow\ S_m \text{ es un grupo}}{=} \sum_{\tau \in S_m} \varepsilon(\tau)\, a_{\tau(1)1} \cdots a_{\tau(n)n}$$

$$= \det(A) \quad \blacksquare$$

Teorema (propiedades del determinante): Sea $A \in M_n(k)$. Entonces:

1. $\det(A)$ depende linealmente de cada fila y cada columna.
2. $\det(A) = 0$ si dos filas o si dos columnas son iguales.
3. $\det(A)$ no cambia si adicionamos a una fila, o a una columna, una combinación lineal de las otras.
4. $\det(A)$ cambia de signo si intercambiamos dos filas o dos columnas.

Dem: Si consideramos $u : k^n \to k^n$ la aplicación lineal asociada a A, entonces $\det(u) = \det(A)$ y todas las afirmaciones respecto a las columnas se obtienen directamente. Las afirmaciones respecto a las filas se obtienen a considerar ${}^tA$ y se deducen de la Proposición anterior. $\blacksquare$

Teorema: Sean $A, B \in M_n(k)$. Entonces:

① $\det(AB) = \det(A)\det(B)$

② A es invertible (ie, $A \in GL_n(k)$) $\iff \det(A) \neq 0$. Más aún, $\det(A^{-1}) = \frac{1}{\det(A)}$.

Dem: Sea $u: k^n \to k^n$ (resp. $v: k^n \to k^n$) el endomorfismo asociado a A (resp. B). Entonces, $\det(u) = \det(A)$ y $\det(v) = \det(B)$. Dado que AB es la matriz asociada a $u \circ v$, tenemos:

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

de donde se tiene ①. El punto ② se deduce de $\det(u) = \det(A)$ y del hecho que la matriz asociada a u^{-1} es A^{-1} ∎

Ejemplo: Sea $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(k)$ tq $\det(A) = ad - bc \neq 0$ (ie, $A \in GL_2(k)$). Un cálculo directo permite verificar que

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

Recuerdo: Dos matrices $A, B \in M_n(k)$ son semejantes si $\exists P \in GL_n(k)$ tal que $P^{-1}AP = B$, ie, A y B representan (en diferentes bases) el mismo endomorfismo $u: k^n \to k^n$.

Corolario: Si $A, B \in M_n(k)$ son semejantes, entonces $\det(A) = \det(B)$.

Dem: Sean $\mathcal{B}$ y $\mathcal{B}'$ bases de k^n tq $A = Mat_{\mathcal{B}}(u)$ y $B = Mat_{\mathcal{B}'}(u)$ para cierto endomorfismo $u: k^n \to k^n$. Entonces, si $\mathcal{B} = (e_1, \dots, e_n)$ y $\mathcal{B}' = (e_1', \dots, e_n')$:

$$\det(u) = \underbrace{\det_{\mathcal{B}}(u(e_1), \dots, u(e_n))}_{\det(A)} \overset{\text{Teorema de §11}}{=} \underbrace{\det_{\mathcal{B}'}(u(e_1'), \dots, u(e_n'))}_{\det(B)} \checkmark$$

Alternativamente, si $P^{-1}AP = B$ con $P \in GL_n(k)$, entonces

$$\det(B) = \det(P^{-1}AP) = \underbrace{\det(P^{-1})}_{1/\det(P)}\det(A)\det(P) = \det(A) \checkmark$$ ∎

Importante: La función $\det: M_n(k) \to k$ no es lineal: En general, $\det(A+B) \neq \det(A) + \det(B)$. Además, $\det(\lambda A) = \lambda^n \det(A)$ para todo $\lambda \in k$.

Ejercicio Sea $A \in M_n(k)$. Decimos que A es antisimétrica si ${}^tA = -A$. Sea $A \in M_n(k)$ matriz antisimétrica, probar que A no es invertible si n es impar.

Ejercicio Sea V un $\mathbb{R}$-e.v. de $\dim_{\mathbb{R}}(V) = n$. Una estructura compleja en V es un endomorfismo $J: V \to V$ tq $J^2 = -id_V$. Probar que si V posee una estructura compleja J entonces necesariamente n es par, y además $J \in GL(V)$.

§13. Cálculo de determinantes

Recuerdo: Una matriz $A \in M_n(k)$ es llamada

a) <u>triangular superior</u> con términos diagonales $\lambda_1, \dots, \lambda_n$ si es de la forma

$$A = \begin{pmatrix} \lambda_1 & * & * & * \\ & \lambda_2 & \ddots & \vdots \\ 0 & & \ddots & * \\ & & & \lambda_n \end{pmatrix}$$

b) <u>triangular inferior</u> con términos diagonales $\lambda_1, \dots, \lambda_n$ si es de la forma

$$A = \begin{pmatrix} \lambda_1 & & & 0 \\ * & \lambda_2 & & \\ * & \ddots & \ddots & \\ * & \cdots & * & \lambda_n \end{pmatrix}$$

a)' <u>triangular superior por bloques</u> si es de la forma

$$A = \left(\begin{array}{c|c|c|c} A_1 & * & * & * \\ \hline 0 & A_2 & * & * \\ \hline 0 & 0 & \ddots & * \\ \hline 0 & 0 & 0 & A_m \end{array}\right)$$

donde $A_i \in M_{n_i}(k)$ son matrices tq $n_1 + \dots + n_m = n$, llamadas los <u>bloques de A</u>.

b)' <u>triangular inferior por bloques</u> si es de la forma

$$A = \left(\begin{array}{c|c|c|c} A_1 & 0 & 0 & 0 \\ \hline * & A_2 & 0 & 0 \\ \hline * & * & \ddots & 0 \\ \hline * & * & * & A_m \end{array}\right)$$

donde $A_i \in M_{n_i}(k)$ son matrices tq $n_1 + \dots + n_m = n$, llamadas los <u>bloques de A</u>.

<u>Prop</u>: Sea $A = \begin{pmatrix} A' & B \\ 0 & A'' \end{pmatrix} \in M_n(k)$ una matriz triangular superior por bloques, donde $A' \in M_p(k)$, $A'' \in M_q(k)$ con $p+q = n$. Entonces,

$$\det(A) = \det(A')\det(A'')$$

<u>Dem</u>: Sea $A = (a_{ij})$, entonces $\det(A) = \sum_{\sigma \in S_n} \varepsilon(\sigma)\, a_{\sigma(1)1} \cdots a_{\sigma(n)n}$.

Dada la forma de A, para que el término $a_{\sigma(1)1} \cdots a_{\sigma(n)n}$ sea no-nulo, se requiere que $\sigma(1), \dots, \sigma(p)$ (n° de fila) pertenezca a $\{1, \dots, p\}$, y luego tenemos que necesariamente $\sigma(p+1), \dots, \sigma(n) \in \{p+1, \dots, n\}$.

En otras palabras, existen permutaciones $\sigma' \in S_p$ y $\sigma'' \in S_q$ tq $\sigma(j) = \sigma'(j)$ si $1 \le j \le p$ y $\sigma(p+k) = p + \sigma''(k)$ si $1 \le k \le q$. Por abuso de notación: $\sigma = (\sigma', \sigma'')$.

Luego, $\det(A) = \sum_{\substack{\sigma = (\sigma', \sigma'') \\ \sigma' \in S_p,\, \sigma'' \in S_q}} \varepsilon(\sigma)\, a'_{\sigma'(1)1} \cdots a'_{\sigma'(p)p}\, a''_{\sigma''(1)1} \cdots a''_{\sigma''(q)q}$

y por otro lado, tenemos que $\varepsilon(\sigma) = \varepsilon(\sigma')\,\varepsilon(\sigma'')$ **Ejercicio**

$$\Rightarrow \det(A) = \underbrace{\left(\sum_{\sigma' \in S_p} \varepsilon(\sigma')\, a'_{\sigma'(1)1} \cdots a'_{\sigma'(p)p}\right)}_{\det(A')} \underbrace{\left(\sum_{\sigma'' \in S_q} \varepsilon(\sigma'')\, a''_{\sigma''(1)1} \cdots a''_{\sigma''(q)q}\right)}_{\det(A'')}$$

$$= \det(A')\det(A'') \quad \blacksquare$$

Corolario: Sea $A \in M_n(k)$ una matriz triangular superior (resp. inferior) por bloques dados por $A_1, \dots, A_m$ con $A_i \in M_{n_i}(k)$ tq $n_1 + \dots + n_m = n$, entonces

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

En part, si $n_1 = \dots = n_m = 1$ y $A_i = (\lambda_i)$, ie, si A es triangular superior (resp. inferior) con términos diagonales $\lambda_1, \dots, \lambda_n$, entonces $\det(A) = \lambda_1 \cdots \lambda_n$.

Dem: Si A es triang. inferior por bloques, entonces ${}^tA$ es triang. superior por bloques y los bloques son los mismos. Así, dado que $\det(A) = \det({}^tA)$, podemos suponer que $A = \begin{pmatrix} A_1 & * & * \\ \hline 0 & \ddots & * \\ \hline 0 & 0 & A_m \end{pmatrix}$ es triang. superior por bloques.

Notar que si definimos $A' := A_1$ y $A'' = \begin{pmatrix} A_2 & * & * \\ \hline 0 & \ddots & * \\ \hline 0 & 0 & A_m \end{pmatrix}$, entonces $A = \begin{pmatrix} A' & B \\ 0 & A'' \end{pmatrix}$ y luego $\det(A) = \det(A')\det(A'')$. Usando inducción en el número de bloques m obtenemos que $\det(A) = \det(A_1) \cdots \det(A_m)$.

Finalmente, si $A = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$, ie, si todos los bloques son de tamaño 1×1, entonces $\det(A) = \lambda_1 \cdots \lambda_n$. ∎

⚠ Si $A = \begin{pmatrix} A_1 & B \\ C & A_2 \end{pmatrix} \in M_{2n}(k)$, con $A_1, A_2, B, C \in M_n(k)$, entonces no necesariamente $\det(A) = \det(A_1)\det(A_2) - \det(B)\det(C)$!

Por ejemplo: $A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \hline 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ cumple $\det(A_1) = \det(A_2) = \det(B) = \det(C) = 0$, pero $\det(A) = -1 \neq 0$.

Ejemplo: Sea $A = \begin{pmatrix} 11 & 0 & 12 & 0 \\ 3 & 1 & 4 & 2 \\ 9 & 0 & 10 & 0 \\ 7 & 5 & 8 & 6 \end{pmatrix}$. Entonces:

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

$$= \begin{vmatrix} 11 & 12 \\ 9 & 10 \end{vmatrix} \cdot \begin{vmatrix} 1 & 2 \\ 5 & 6 \end{vmatrix} = \underbrace{(110 - 108)}_{=2} \cdot \underbrace{(6 - 10)}_{=-4} = -8.$$

Def: Sea $A = (a_{ij}) \in M_n(k)$ y denotemos por $A_{ij} \in M_{n-1}(k)$ la matriz cuadrada obtenida al eliminar la i-ésima fila y la j-ésima columna de A. Diremos que

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

es el cofactor de índice (i,j) de A.

Ejemplo ($n = 2$)

Si $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ (con $+$ en a_{11}, $-$ en a_{12}, $-$ en a_{21}, $+$ en a_{22}) $\rightsquigarrow$

$C_{11} = +a_{22}$ $\quad C_{12} = -a_{21}$

$C_{21} = -a_{12}$ $\quad C_{22} = +a_{11}$

Teorema: Para todos $i,j \in \{1,\dots,n\}$ tenemos que:

① $\det(A) = a_{1j}C_{1j} + \dots + a_{nj}C_{nj}$ (desarrollo de la j-ésima columna).

② $\det(A) = a_{i1}C_{i1} + \dots + a_{in}C_{in}$ (desarrollo de la i-ésima fila).

Dem: Basta probar ①, pues $\det(A) = \det({}^tA)$. Fijemos $j \in \{1,\dots,n\}$ y veamos ①: Sea B_{ij} la matriz obtenida a partir de A reemplazando por 0 todos los elementos de la j-ésima columna, excepto el de la i-ésima fila:

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & & \vdots & & \\ & & a_{ij} & & \\ \vdots & & \vdots & & \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \\ \underbrace{}_{x_1} & & \underbrace{}_{x_j} & & \underbrace{}_{x_n} \end{pmatrix} \longmapsto B_{ij} := \begin{pmatrix} a_{11} & \cdots & 0 & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ & & 0 & & \\ & & a_{ij} & & \\ & & 0 & & \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & 0 & \cdots & a_{nn} \end{pmatrix}$$

Si $x_1,\dots,x_n \in k^n$ son las columnas de A y $x_j = a_{1j}e_1 + \dots + a_{nj}e_n$, donde $\mathcal{B} = (e_1,\dots,e_n)$ es la base canónica de k^n, entonces:

$$\det(A) = \det_{\mathcal{B}}(x_1,\dots,x_n) = \det(x_1,\dots,x_{j-1}, a_{1j}e_1 + \dots + a_{nj}e_n, x_{j+1},\dots,x_n)$$
$$\overset{\det \text{ es multilineal}}{=} \det(B_{1j}) + \det(B_{2j}) + \dots + \det(B_{nj}).$$

Si hacemos pasar la j-ésima columna de B_{ij} a la primera posición, ie, si hacemos $j-1$ transposiciones de columnas, el determinante se multiplica por $(-1)^{j-1}$. Si luego de esto hacemos pasar la i-ésima fila a la primera posición, el determinante se multiplica por $(-1)^{i-1}$. Luego:

$$\det(B_{ij}) = \underbrace{(-1)^{(i-1)+(j-1)}}_{(-1)^{i+j}} \begin{vmatrix} a_{ij} & a_{i1} & \cdots & a_{in} \\ 0 & & & \\ \vdots & & A_{ij} & \\ 0 & & & \end{vmatrix} = a_{ij}(-1)^{i+j}\det(A_{ij}) = a_{ij}C_{ij}. \quad \blacksquare$$

Ejemplo: Siempre conviene desarrollar una fila o columna con muchos ceros:

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 4 & 0 & 5 & 6 \\ 0 & 0 & 3 & 0 \\ 1 & 6 & 4 & 2 \end{vmatrix} = 3 \begin{vmatrix} 1 & 2 & 4 \\ 4 & 0 & 6 \\ 1 & 6 & 2 \end{vmatrix} = 3\left((-2)\underbrace{\begin{vmatrix} 4 & 6 \\ 1 & 2 \end{vmatrix}}_{=2} + (-6)\underbrace{\begin{vmatrix} 1 & 4 \\ 4 & 6 \end{vmatrix}}_{=-10}\right) = 3\cdot(-4+60) = 168.$$

Def: Sea $A \in M_n(k)$. Definimos la **comatriz** de A por $\operatorname{com}(A) := (C_{ij}) \in M_n(k)$.

Prop: $A \cdot {}^t\operatorname{com}(A) = {}^t\operatorname{com}(A)\cdot A = \det(A)\cdot I_n$. En part, $A^{-1} = \frac{1}{\det(A)}{}^t\operatorname{com}(A)$ si $A \in GL_n(k)$.

Dem: El término (i,j) de $A\cdot{}^t\operatorname{com}(A)$ es $\sum_{k=1}^{n} a_{ik}C_{jk}$. Si $i=j$, el Teorema anterior dice que esto vale $\det(A)$ ✓

Si $i \neq j$, el teorema anterior dice que esto vale el desarrollo de la j-ésima fila del determinante de la matriz M_{ij} que se obtiene a partir de A copiando la i-ésima fila en la j-ésima fila (**Ejercicio**). Como M_{ij} tiene dos filas iguales, $\det(M_{ij}) = 0$, y luego $A\cdot{}^t\operatorname{com}(A) = \det(A)\cdot I_n$.

La igualdad ${}^t\operatorname{com}(A)\cdot A = \det(A)\cdot I_n$ se obtiene al desarrollar columnas. $\blacksquare$

§14. Valores y vectores propios

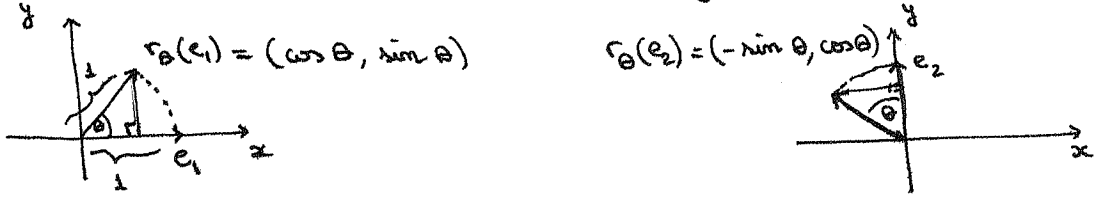
Motivación: Sea $u: V \to V$ endomorfismo. Nos gustaría hallar una base $\mathcal{B}$ de V tq $A = Mat_{\mathcal{B}}(u) \in M_n(k)$ sea lo más "simple" posible.

Por ejemplo, si $\exists\, \mathcal{B} = (v_1, \dots, v_n)$ tq $A = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ es diagonal, entonces $u(v_i) = \lambda_i v_i$ para todo $i \in \{1, \dots, n\}$, ie, v_i y $u(v_i)$ son colineales.

Notar que en tal caso podemos calcular fácilmente $\det(u) = \lambda_1 \cdots \lambda_n$ y $A^m = \begin{pmatrix} \lambda_1^m & & 0 \\ & \ddots & \\ 0 & & \lambda_n^m \end{pmatrix}$ $\forall m \in \mathbb{N}^{\geq 1}$, por ejemplo.

Lamentablemente, esto no siempre es posible:

Ejemplo: Sea $V = \mathbb{R}^2$ con base canónica $\mathcal{B} = (e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix})$, y sea $r_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ la rotación (anti-horaria) en ángulo θ:



$$\Rightarrow A_\theta = Mat_{\mathcal{B}}(r_\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Veremos que r_θ no posee vectores propios a menos que $\theta \equiv 0 \pmod{\pi}$.

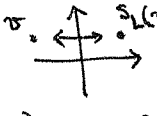
Def: Sea $u: V \to V$ endomorfismo. Si existe un escalar $\lambda \in k$ y un vector $v \in V$, $v \neq 0$, tal que $u(v) = \lambda v$, decimos que λ es un <u>valor propio</u> (o eigenvalor) y que v es un <u>vector propio</u> (o eigenvector) asociado a λ.

Obs: Todo multiplo $\neq 0$ de un vector propio es de nuevo un vector propio, asociado al mismo valor propio. En part, los vectores propios no son únicos!

Ejemplos:

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

2) $V = \mathbb{R}^2$, $S_L : \mathbb{R}^2 \to \mathbb{R}^2$, $(x,y) \mapsto (-x,y)$ simetría resp. a la recta $L = Vect_{\mathbb{R}}(e_2)$



e_1 v.p. asociado a $\lambda_1 = -1$
e_2 v.p. asociado a $\lambda_2 = 1$.

3) $V = \mathcal{C}^\infty(\mathbb{R},\mathbb{R}) = \{f: \mathbb{R} \to \mathbb{R} \text{ clase } \mathcal{C}^\infty\}$. y $u: V \to V$, $f \mapsto f'$.
La función $f(t) = e^{\lambda t}$ es un v.p. asociada al valor propio λ: $f' = \lambda f$.

4) $V = \mathcal{C}^\infty(\mathbb{R},\mathbb{R})$ 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 v.p. asociado a $\lambda = 0 \Rightarrow f_{a,b}(t) \in \ker(u)$.
MAT023: Verán que $\ker(u) = Vect_{\mathbb{R}}\langle \cos(t), \sin(t) \rangle \cong \mathbb{R}^2$.

Terminología: Sea $u: V \to V$ endomorfismo. El conjunto (eventualmente vacío) de valores propios de u es denotado $\sigma(u) \subseteq k$ y es llamado el espectro de u: $\sigma(u) = \{\lambda \in k \text{ tq } \exists v \neq 0 \text{ con } u(v) = \lambda v\}$.

Ejercicio Sea $u \in End_k(V)$ y $v \in V$ un vector propio asociado al valor propio $\lambda \in k$. Probar que $\forall Q \in k[X]$ se tiene que $Q(u)(v) = Q(\lambda)v$.

Def: Sea $u: V \to V$ endomorfismo y $\lambda \in k$ un valor propio de u. El sub-e.v.
$$V_\lambda := \ker(\lambda id_V - u) \subseteq V$$
es llamado el sub-espacio propio asociado a λ.

Obs: Notar que $u(v) = \lambda v = \lambda id_V(v) \iff (\lambda id_V - u)(v) = 0 \iff v \in V_\lambda$. Más aún, por definición de valor propio, $V_\lambda \neq \{0\}$.

⚠ En todo lo que sigue supondremos que V es de dimensión finita. La teoría espectral de dimensión infinita será estudiada en Análisis Funcional.

Prop: Sea $u: V \to V$ endomorfismo y $\lambda \in k$ escalar. Entonces, λ es valor propio de u (ie, $\lambda \in \sigma(u)$) $\iff \det(\lambda id_V - u) = 0$.

Dem: ($\Rightarrow$) Si $\lambda \in \sigma(u)$, existe $v \neq 0$ vector propio $\Rightarrow \lambda id_V - u$ no es inyectiva (pues $v \in V_\lambda = \ker(\lambda id_V - u)$). En part, $\lambda id_V - u$ no es un automorfismo $\iff \det(\lambda id_V - u) = 0$.
($\Leftarrow$) Dado que V de dim. finita: $\lambda id_V - u$ automorfismo $\iff$ inyectiva. Luego, $\det(\lambda id_V - u) = 0 \Rightarrow \lambda id_V - u$ no es inyectivo $\Rightarrow \exists v \neq 0$ tq $u(v) = \lambda v$ ∎

Terminología: Si $V_1, V_2 \subseteq V$ sub-e.v. Decimos que V es suma directa de V_1 y V_2 si todo elemento $v \in V$ se escribe de manera única como
$$v = v_1 + v_2, \quad \text{con } v_1 \in V_1,\ v_2 \in V_2.$$
En tal caso escribimos $V = V_1 \oplus V_2$. En otras palabras, la aplicación lineal $\varphi: V_1 \times V_2 \to V$, $(v_1, v_2) \mapsto \varphi(v_1, v_2) := v_1 + v_2$ es biyectiva.

Ejercicio $V = V_1 \oplus V_2 \iff \begin{cases} a) & V = Vect_k(V_1, V_2). \\ b) & V_1 \cap V_2 = \{0\}. \end{cases}$

Ejercicio* Definamos $Im(\varphi) := V_1 + V_2$. Probar que $V_1 + V_2 = Vect_k(V_1, V_2)$ y que $\ker(\varphi) = V_1 \cap V_2$. Deducir (eg. usando el Teorema del rango) que
$$\dim_k(V_1 + V_2) = \dim_k(V_1) + \dim_k(V_2) - \dim_k(V_1 \cap V_2).$$

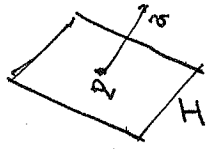
Def: Sean $V_1, \dots, V_p \subseteq V$ sub-e.v. Diremos que V es suma directa de $V_1, \dots, V_p$ si la aplicación lineal
$$\varphi: V_1 \times \dots \times V_p \to V, \quad (v_1, \dots, v_p) \mapsto v_1 + \dots + v_p$$
es biyectiva, en cuyo caso escribimos $V = V_1 \oplus \dots \oplus V_p$ o $V = \bigoplus_{j=1}^{p} V_j$.
Si φ es inyectiva, decimos que los $V_1, \dots, V_p$ están en suma directa.

Def: Sean $V_1, V_2 \subseteq V$ sub-e.v. Decimos que V_1 y V_2 son *suplementarios* si $V = V_1 \oplus V_2$.

Lema: Todo sub-e.v. de V admite un suplementario.

Dem: Sup. que $\dim_k(V) = n$ y sea $W \subseteq V$ sub-e.v. Si $W = \{0\}$ entonces W' es un suplementario. Si $\dim_k(W) = m \geq 1$, consideramos $(e_1, \dots, e_m)$ base de W. Completamos $(e_1, \dots, e_m)$ en una base $\mathcal{B} = (e_1, \dots, e_m, e_{m+1}, \dots, e_n)$ de V, y definimos $W' = \mathrm{Vect}_k(e_{m+1}, \dots, e_n)$. Luego, $V = \mathrm{Vect}_k(W, W')$ y $W \cap W' = \{0\}$ ■

Ejemplo: Sea $f : V \to k$ una forma lineal en V. Sup. que $f \neq 0$ y sea $H := \ker(f)$. Como $f \neq 0$, $\exists v \in V$ tq $f(v) \neq 0$ en k.
Reemplazando v por $f(v)^{-1} v$, podemos suponer que $f(v) = 1$.
$\Rightarrow \forall w \in V$, $f(w - f(w)v) = f(w) - f(w)\underbrace{f(v)}_{=1} = 0 \Rightarrow w - f(w)v \in \ker(f) = H$.
Así, $w = h + f(w)v$, con $h := w - f(w)v \in H \Rightarrow V = \mathrm{Vect}_k(H, v)$.
Por otro lado, si $\lambda v \in H \Rightarrow 0 = f(\lambda v) = \lambda \underbrace{f(v)}_{=1} \Rightarrow H \cap \mathrm{Vect}_k(v) = \{0\}$.
Conclusión: $V = H \oplus \mathrm{Vect}_k(v)$, ie, H y la recta $\mathrm{Vect}_k(v)$ son suplementarios.



Decimos que H es un *hiperplano* en V.
$\Rightarrow \dim_k(H) = n - 1$.
Lema

Prop: Sean $V_1, \dots, V_p \subseteq V$ sub-e.v. tq $V = V_1 \oplus \dots \oplus V_p$. Entonces:
① $\dim_k(V) = \dim_k(V_1) + \dots + \dim_k(V_p)$
② Si $\mathcal{B}_j$ es una base de V_j, para $j \in \{1, \dots, p\}$, entonces $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_p)$ es una base de V.

Dem: Dado que $\varphi : V_1 \times \dots \times V_p \to V$ es un isomorfismo, tenemos que $\dim_k(V) = \dim_k(V_1 \times \dots \times V_p) = \dim_k(V_1) + \dots + \dim_k(V_p)$, de donde deducimos ① ✓
Del mismo modo, $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_p)$ es generadora y, por ①, tiene el mismo cardinal, $\dim_k(V)$, de una base de V. Luego $\mathcal{B}$ es base de V ② ✓ ■

Ejercicio importante Sean $V_1, \dots, V_p \subseteq V$ sub-e.v. tq $V = V_1 \oplus \dots \oplus V_p$ y sea $\mathcal{B}_j$ una base de V_j para cada $j \in \{1, \dots, p\}$. Sea $u : V \to V$ endomorfismo tq $u(V_j) \subseteq V_j$ $\forall j \in \{1, \dots, p\}$. Probar que

$$A = \mathrm{Mat}_{\mathcal{B}}(u) = \left(\begin{array}{c|c|c} A_1 & 0 & 0 \\ \hline 0 & \ddots & 0 \\ \hline 0 & 0 & A_p \end{array}\right)$$

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

Prop: Los sub-espacios propios de un endomorfismo asociados a valores propios distintos están en suma directa.

Dem: Por inducción en el número p de valores propios: $p=1$ ✓

Sean $\lambda_1,\dots,\lambda_p$ valores propios de $u: V\to V$ tq $\lambda_i\neq\lambda_j$ si $i\neq j$.

Queremos probar que $\varphi: V_{\lambda_1}\times\dots\times V_{\lambda_p}\to V$, $(v_1,\dots,v_p)\mapsto v_1+\dots+v_p$ es inyectiva, ie, si $v_1+\dots+v_p=0 \Rightarrow v_1=\dots=v_p=0$.

Al aplicar u a la ecuación $v_1+\dots+v_p=0$ obtenemos $0=\lambda_1 v_1+\dots+\lambda_p v_p$.

Si restamos $\lambda_1(v_1+\dots+v_p)=0$, obtenemos $(\lambda_2-\lambda_1)v_2+\dots+(\lambda_p-\lambda_1)v_p=0$.

$\overset{\text{Hip. Inducción}}{\Longrightarrow} V_{\lambda_2},\dots,V_{\lambda_p}$ están en suma directa y luego $(\lambda_j-\lambda_1)v_j=0$ para $j=2,\dots,p$. Como $\lambda_1\neq\lambda_j \Rightarrow v_2=\dots=v_p=0$ y luego $v_1=0$ ■

Obs: Sean $\lambda_1,\dots,\lambda_p$ los valores propios (distintos) de $u: V\to V$. Entonces, $V_{\lambda_1}+\dots+V_{\lambda_p}:=\mathrm{Vect}_k(V_{\lambda_1},\dots,V_{\lambda_p})=V_{\lambda_1}\oplus\dots\oplus V_{\lambda_p}$, pero no necesariamente $V=\bigoplus_{j=1}^{p}V_{\lambda_j}$. Más adelante: $V=\bigoplus_{j=1}^{p}V_{\lambda_j}\iff u$ es "diagonalizable".

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

Dem: Sean $\lambda_1,\dots,\lambda_p$ los valores propios de u. Luego: $\bigoplus_{j=1}^{p}V_{\lambda_j}\subseteq V$

$\Rightarrow \dim_k(V_{\lambda_1})+\dots+\dim_k(V_{\lambda_p})=\dim_k\left(\bigoplus_{j=1}^{p}V_{\lambda_j}\right)\leq\dim_k(V)$.

Dado que $\dim_k(V_{\lambda_j})\geq 1$ (pues $\exists\, 0\neq v\in V_{\lambda_j}$ vector propio) $\Rightarrow p\leq\dim_k(V)$ ■

§15. Endomorfismos diagonalizables

Def: Un endomorfismo $u: V\to V$ es diagonalizable si existe una base de V formada por vectores propios de u.

Obs: Si $\mathcal{B}=(v_1,\dots,v_m)$ es una base de V forma por vectores propios de u.

$\Rightarrow \mathrm{Mat}_{\mathcal{B}}(u)=\begin{pmatrix}\lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_m\end{pmatrix}$, donde $u(v_j)=\lambda_j v_j$.

Recíprocamente, si $\exists\,\mathcal{B}$ base de V tq $\mathrm{Mat}_{\mathcal{B}}(u)$ es diagonal, entonces necesariamente los vectores de $\mathcal{B}$ son vectores propios de u.

Prop: Sea $u: V\to V$ endomorfismo y sean $\lambda_1,\dots,\lambda_p$ los valores propios distintos de u. Son equivalentes:

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

Dem: 3 ⇒ 2: Vimos que $\varphi: V_{\lambda_1}\times\dots\times V_{\lambda_p}\to V$ inyectiva. Si 3 se verifica entonces φ biyectiva ⇒ 2 ✓

2 ⇒ 1: Sea $\mathcal{B}_j$ base de V_{λ_j} $\overset{2}{\Rightarrow}\mathcal{B}=(\mathcal{B}_1,\dots,\mathcal{B}_p)$ base de V formada por valores propios ⇒ 1 ✓

1 ⇒ 3: Sea $\mathcal{B}$ base de vectores propios y sea n_j el número de vectores de $\mathcal{B}$ asociados a $\lambda_j \Rightarrow \dim_k(V_{\lambda_j})\geq n_j \Rightarrow \dim_k(V_{\lambda_1})+\dots+\dim_k(V_{\lambda_p})\geq n_1+\dots+n_p=\dim_k(V)$ ■

Corolario: Sup. que $\dim_k(V) = n$ y sea $u: V \to V$ endomorfismos. Si u posee n valores propios distintos $\Rightarrow u$ es diagonalizable.

Dem: Con la notación anterior: $p \geq n$. Además, $\dim(V_{\lambda_j}) \geq 1 \Rightarrow \sum_{j=1}^{p} \dim_k(V_{\lambda_j}) \geq n$ ■

Obs: El recíproco es falso (e.g. $u = \lambda id_V \Rightarrow V_\lambda = V$, para $p = 1$).

Ejercicio Sea $u: V \to V$ diagonalizable. Probar que $\lambda \in k$ es el único valor propios de $u \iff u = \lambda id_V$ es una homotecia.

§16. Polinomio característico de una matriz cuadrada

Recuerdo: Si $A = (a_{ij}) \in M_n(k)$, 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, donde $\mathcal{B} = (e_1, \dots, e_n)$ base canónica.

Def: Sea $A \in M_n(k)$. Los valores y vectores propios de A son aquellos del endomorfismo $u: k^n \to k^n$ asociado. En otras palabras, $\lambda \in k$ es un valor propio de A si $\exists\, v \in k^n$ con $v \neq 0$ y tal que $Av = \lambda v$; el vector v es un vector propio de A asociado a λ.

Ejemplo: Sea $A = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix}$ y sea $v = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \Rightarrow Av = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 \\ 6 \end{pmatrix} = 3\begin{pmatrix} 1 \\ 2 \end{pmatrix}$
Luego, $v \neq 0$ es un vector propio asociado a $\lambda = 3$.

Def: Decimos que $A \in M_n(k)$ es diagonalizable si $u_A: k^n \to k^n$ lo es.
Es decir, si $\exists\, P \in Gl_n(k)$ tq $P^{-1}AP = D$ es una matriz diagonal.

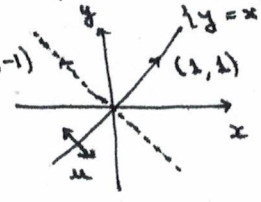
Luego, todas las propiedades discutidas en §15. valen para matrices, por ejemplo: "Toda matriz $A \in M_n(k)$ con n valores propios distintos es diagonalizable".
O bien y otro ejemplo:

Prop: Sea $A \in M_n(k)$. Un escalar $\lambda \in k$ es un valor propio de A si y sólo si $\det(\lambda I_n - A) = 0$.

Ejemplo: Sea $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \in M_2(\mathbb{R})$, y sea $\lambda \in \mathbb{R}$. Entonces:

$$\det(\lambda I_2 - A) = \begin{vmatrix} \lambda & -1 \\ -1 & \lambda \end{vmatrix} = \lambda^2 - 1 = (\lambda+1)(\lambda-1) = 0 \iff \lambda_1 = 1,\ \lambda_2 = -1.$$

Luego, A es diagonalizable sobre $\mathbb{R}$, pues posee dos valores propios distintos.
Alternativamente: $u_A(x,y) = (y,x)$, pues $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} y \\ x \end{pmatrix}$.
$\Rightarrow$ Geométricamente, A (o bien, siendo estrictos, u_A) es la reflexión en $\mathbb{R}^2$ respecto a la recta $y = x$:



$\{y = x\} = \text{Vect}_{\mathbb{R}}\langle(1,1)\rangle$

Observamos (geométricamente), que $v_1 = (1,1)$ y $v_2 = (1,-1)$ son vectores propios de A, asociados a $\lambda_1 = 1$ y $\lambda_2 = -1$, resp.
$\mathcal{B} = (v_1, v_2)$ es una base de $\mathbb{R}^2$, pues $\begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix} = -2 \neq 0$.

En el ejemplo anterior se puede apreciar que para hallar los valores propios de $A \in M_n(k)$ hay que resolver una ecuación polinomial. Esto se formaliza de la manera siguiente:

A partir del anillo de polinomios $k[X]$ podemos formar un "cuerpo de fracciones" (cf. $\mathbb{Q}$ a partir de $\mathbb{Z}$) cuyos elementos son de la forma $\frac{P(X)}{Q(X)}$ con $P, Q \in k[X]$ y donde $\frac{P(X)}{Q(X)} = \frac{R(X)}{S(X)} \overset{\text{def}}{\Leftrightarrow} PS = RS$ en $k[X]$.

Dicho cuerpo se llama el <u>cuerpo de funciones racionales</u> en la variable X y se denota $k(X)$.

Todo lo que hemos discutido hasta ahora es válido sobre <u>cualquier cuerpo</u> (!), en particular sobre el cuerpo $K = k(X)$. Notar que $k \subseteq k(X)$ <u>sub-cuerpo</u>.

Sea $A \in M_n(k)$ y consideremos la matriz en $M_n(k(X))$ dada por:

$$XI_n - A = \begin{pmatrix} X - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & X - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ -a_{n1} & -a_{n2} & \cdots & X - a_{nn} \end{pmatrix}$$

i.e., una matriz con coeficientes dados por polinomios en $k[X]$.

<u>Lema útil</u>: Sea $P \in k[X]$ dado por

$$P(X) = (X - \lambda_1) \cdots (X - \lambda_n) = X^n + a_{n-1}X^{n-1} + \ldots + a_2X^2 + a_1X + a_0.$$

Entonces: $a_{n-1} = -\lambda_1 - \lambda_2 - \ldots - \lambda_n$

$$a_{n-2} = \lambda_1\lambda_2 + \lambda_1\lambda_3 + \ldots + \lambda_2\lambda_3 + \ldots + \lambda_{n-1}\lambda_n = \sum_{1 \le i < j \le n} \lambda_i\lambda_j.$$

$$\vdots$$

$$a_1 = (-1)^{n-1}(\lambda_2\lambda_3 \cdots \lambda_n + \lambda_1\lambda_3\lambda_4 \cdots \lambda_n + \ldots + \lambda_1\lambda_2 \cdots \lambda_{n-1}) = (-1)^{n-1} \sum_{i=1}^{n} \prod_{j \neq i} \lambda_j$$

$$a_0 = (-1)^n \lambda_1\lambda_2 \cdots \lambda_n.$$

<u>Dem</u>: Comparar los coeficientes ■

<u>Teorema</u>: Sea $A \in M_n(k)$. El determinante $\det(XI_n - A)$ es un polinomio en $k[X]$ que es unitario (i.e., coeficiente principal 1) y de grado n. Dicho polinomio es llamado el <u>polinomio característico</u> de A y lo denotamos P_A. Más aún, $P_A(X) = X^n - \operatorname{tr}(A)X^{n-1} + \ldots + (-1)^n \det(A)$, donde

$$\operatorname{tr}(A) = a_{11} + \ldots + a_{nn}$$

es la <u>traza</u> de A, i.e., la suma de los términos diagonales de A.

<u>Dem</u>: Escribamos $XI_n - A = (b_{ij}) \in M_n(k(X))$, i.e., $b_{ij} = \begin{cases} -a_{ij} & \text{si } i \neq j \\ X - a_{ii} & \text{si } i = j \end{cases}$.

$$\Rightarrow P_A(X) = \sum_{\sigma \in S_n} \varepsilon(\sigma)\, b_{1\sigma(1)} \cdots b_{n\sigma(n)}$$

Dado que cada $b_{ij} \in k[X]$ es un polinomio $\Rightarrow P_A(X)$ es un polinomio ✓
El término de grado maximal en X es aquel que corresponde al producto de los términos de la diagonal, ie, $\sigma = id$:

$$\underbrace{\varepsilon(id)}_{=1} b_{11} \cdots b_{nn} = (X - a_{11}) \cdots (X - a_{nn}) = X^n + \dots$$

$\Rightarrow P_A$ es unitario de grado n.

Del mismo modo, para tener un término con X^{n-1} hay que tomar $n-1$ términos diagonales : $\varepsilon(\sigma) b_{11} \cdots b_{i-1,i-1} b_{i\sigma(i)} b_{i+1,i+1} \cdots b_{nn}$.
Dado que $\sigma(j) = j$ para $j \neq i$ y $\sigma \in S_n$ biyección, necesariamente $\sigma(i) = i$, ie,

$$\sigma = id \Rightarrow b_{11} \cdots b_{nn} = (X - a_{11})(X - a_{22}) \cdots (X - a_{nn})$$
$$\underset{\text{lema útil}}{=} X^n - \underbrace{(a_{11} + \dots + a_{nn})}_{tr(A)} X^{n-1} + \dots + (-1)^n a_{11} \cdots a_{nn}.$$

Finalmente, si escribimos $P_A(X) = X^n - tr(A)X + \dots + a_1 X + a_0$
$\Rightarrow P_A(0) = \det(-A) = (-1)^n \det(A) = a_0$. ∎

Ejemplos:

① Sea $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \in M_n(k)$, reflexión respecto a la recta $y = x$ en k^2.
$\Rightarrow P_A(X) = \det(XI_2 - A) = \begin{vmatrix} X & -1 \\ -1 & X \end{vmatrix} = X^2 - 1 = (X+1)(X-1)$ tiene raíces $\lambda_1 = 1$, $\lambda_2 = -1$ sobre todo cuerpo k.
Si $car(k) = 2$: $\lambda_1 = \lambda_2$; pero si $car(k) \neq 2$: $\lambda_1 \neq \lambda_2$ y luego A es diagonalizable en tal caso.

② Sea $r_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ la rotación en ángulo $\theta \in \mathbb{R}$, cuya matriz en la base canónica está dada por $A_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \in M_2(\mathbb{R})$.

$$\Rightarrow P_{A_\theta}(X) = \begin{vmatrix} X - \cos\theta & \sin\theta \\ -\sin\theta & X - \cos\theta \end{vmatrix} = (X^2 - \cos\theta)^2 - (-\sin^2\theta)$$
$$= X^2 - 2\cos\theta X + \underbrace{\cos^2\theta + \sin^2\theta}_{=1} = X^2 - 2\cos\theta X + 1.$$

$\Delta = b^2 - 4ac = 4\cos^2\theta - 4 = -4\sin^2\theta \leq 0$
Luego, P_{A_θ} posee raíces en $\mathbb{R} \Leftrightarrow \sin\theta = 0 \Leftrightarrow \theta \equiv 0 \pmod{\pi}$.
⚠ Si pensamos $A_\theta \in M_2(\mathbb{C})$ como matriz con coeficientes complejos, entonces las raíces de P_{A_θ} están dadas por $\frac{-b \pm \sqrt{\Delta}}{2a} = \frac{2\cos\theta \pm 2i\sin\theta}{2} = e^{\pm i\theta}$
$\Rightarrow \lambda_1 = e^{i\theta}$ y $\lambda_2 = e^{-i\theta}$ valores propios.
Ejercicio $v_1 = \begin{pmatrix} 1 \\ i \end{pmatrix}$ y $v_2 = \begin{pmatrix} 1 \\ -i \end{pmatrix}$ son vectores propios y luego A_θ es diagonalizable sobre $\mathbb{C}$.

③ Si $A = (a_{ij}) \in M_n(k)$ es una matriz triangular (superior o inferior), entonces $XI_n - A$ lo es también $\Rightarrow P_A(X) = (X - a_{11}) \cdots (X - a_{nn})$ y luego los valores propios de A son sus términos diagonales. En part, si los coeficientes a_{ii} son todos distintos $\Rightarrow A$ diagonalizable.

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: Toda matriz compleja $A \in M_n(\mathbb{C})$ tiene un valor propio en $\mathbb{C}$.

Dem: $P_A(X) = \det(XI_n - A) \in \mathbb{C}[X]$ es de grado $n \geq 1$ y luego $\exists \lambda \in \mathbb{C}$ tq $P_A(\lambda) = 0 \Leftrightarrow \lambda$ valor propio de A. ■

⚠ Si k es un cuerpo arbitrario, el teorema de Steinitz (1910), ajirma que existe un cuerpo $\overline{k}$ tq $k \subset \overline{k}$ y tal que todo $P \in \overline{k}[X]$ no constante posee una raíz en $\overline{k}$; $\overline{k}$ es la clausura algebraica de k.

§17. Polinomio característico de un endomorfismo

Sea $u : V \to V$ un endomorfismo y sea $\mathcal{B}$ una base de V. Si denotamos $A = \mathrm{Mat}_{\mathcal{B}}(u) \in M_n(k)$, entonces $\det(\lambda \mathrm{id}_V - u) = \det(\lambda I_n - A)$ para todo $\lambda \in k$. En part, los valores propios de u son los valores propios de su matriz asociada respecto a cualquier base.

Más aún, dado que A y B son semejantes (ie, $\exists P \in Gl_n(k)$ tq $P^{-1}AP = B$) si y sólo si representan un mismo endomorfismo (ie, $\exists u : V \to V$ y bases $\mathcal{B}$ y $\mathcal{E}$ de V tq $A = \mathrm{Mat}_{\mathcal{B}}(u)$ y $B = \mathrm{Mat}_{\mathcal{E}}(u)$), se tiene que matrices semejantes tienen los mismos valores propios.

El resultado siguiente generaliza la discusión anterior:

Teorema: Matrices semejantes tienen el mismo polinomio característico.

Dem: Basta probar que si $A, B \in M_n(k)$ son semejantes, entonces $XI_n - A$ y $XI_n - B$ son semejantes en $M_n(k(X))$.

Sup. que $\exists P \in Gl_n(k)$ tq $B = P^{-1}AP$, entonces

$$P^{-1}(XI_n - A)P = P^{-1}(XI_n)P - P^{-1}AP = XI_n - B,$$

y luego $XI_n - A$ y $XI_n - B$ tienen el mismo determinante ■

Corolario: Sea $u : V \to V$ endomorfismo, $\mathcal{B}$ una base de V y $A = (a_{ij}) = \mathrm{Mat}_{\mathcal{B}}(u)$. Entonces, el polinomio característico de u, $P_u \in k[X]$, dado por la expresión

$$P_u(X) = \det(X \mathrm{id}_V - u) = \det(XI_n - A) = X^n - \mathrm{tr}(u)X^{n-1} + \ldots + (-1)^n \det(u)$$

está bien definido, donde $n = \dim_k(V)$ y el coeficiente $\mathrm{tr}(u)$ es la traza de u, que coincide con $\mathrm{tr}(A) = \sum_{i=1}^{n} a_{ii}$ (y no depende de la base $\mathcal{B}$). ■

Ejercicio Sean $A, B \in M_n(k)$. Probar (por definición) que $\mathrm{tr}(AB) = \mathrm{tr}(BA)$.

Recuerdo: Sean $A, B \in k[X]$ polinomios. Decimos que B **divide** a A (o que A es un **múltiplo** de B) si $\exists\, Q \in k[X]$ tq $A = BQ$.

Teorema (división euclídeana): Sean $A, B \in k[X]$. Si $B \neq 0$, existen únicos polinomios Q (cociente) y R (resto) tq
$$A = BQ + R \quad \text{y} \quad gr(R) < gr(B).$$

Dem: Unicidad: Si $A = BQ + R = BQ' + R' \Rightarrow R - R' = B(Q' - Q)$ y luego B divide $R - R'$. $\underset{gr(R-R') < gr(B)}{\Rightarrow} R' - R = 0$, ie, $R = R' \underset{B \neq 0}{\Rightarrow} Q = Q'$ ✓

Existencia: Por inducción en $n = gr(A)$: Si $A = 0$, consideramos $Q = R = 0$ ✓
Sup. que $A \neq 0$. Si $n < gr(B)$ consideramos $Q = 0$ y $R = A$, sino: $n \geq gr(B) =: m$ y escribimos $A = \sum_{i=0}^{n} a_i X^i$ y $B = \sum_{i=0}^{m} b_i X^i$, con $a_n \neq 0$ y $b_m \neq 0$.
$\Rightarrow A' := A - \frac{a_n}{b_m} X^{n-m} B = (a_n X^n + \ldots) - \frac{a_n}{b_m} X^{n-m} (b_m X^m + \ldots)$ tiene grado $< n$.
Hipótesis de inducción $\Rightarrow A' = BQ' + R'$ con $gr(R') < gr(B)$
$\Rightarrow A = B \cdot \left(\frac{a_n}{b_m} X^{n-m} + Q' \right) + R'$ y luego $Q = \frac{a_n}{b_m} X^{n-m} + Q'$ y $R = R'$ ■

Ejemplo: La división euclídeana de $A(X) = X^3 + X + 1$ por $B(X) = X + 1$ es:

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

$$\Rightarrow A(X) = B(X)\underbrace{(X^2 - X + 2)}_{Q(X)} \underbrace{- 1}_{R(X)}.$$

Corolario: Sea $P \in k[X]$. Entonces, $a \in k$ es una raíz de P (ie, $P(a) = 0$) $\Leftrightarrow$ P es divisible por $X - a$.

Dem: La división euclídeana de P por $X - a$ nos da $P = (X-a)Q + R$, donde $gr(R) < 1$ y luego $R \in k$ es constante $\Rightarrow P(a) = 0 + R$, de donde concluimos ■

Recuerdo: Sea $P \in k[X]$, y sea $a \in k$ una raíz de P. La **multiplicidad** de la raíz $a \in k$ es el máximo $m \in \mathbb{N}^{\geq 1}$ tal que $(X-a)^m$ divide a $P(X)$. En caso que $m = 1$ (resp $m = 2$, resp. $m = 3$, etc) decimos que a es una **raíz simple** (resp. **raíz doble**, resp. **raíz triple**, etc).

Ejemplo: Sea $P \in \mathbb{C}[X]$ dado por $P(X) = X^n - 1$, para cierto $n \in \mathbb{N}^{\geq 1}$.
$\Rightarrow P(X) = (X-1)(X-\xi)(X-\xi^2)\cdots(X-\xi^{n-1})$ con $\xi = e^{2\pi i/n}$ (raíces simples).

Ejercicio Sea $P \in k[X]$ y $a \in k$. Probar que a es una raíz **múltiple** de P (ie, de multiplicidad ≥ 2) $\Leftrightarrow P(a) = P'(a) = 0$.

Def: Sea $I \subseteq k[X]$ sub-conjunto. Decimos que I es un *ideal* de $k[X]$ si:
(i) I es un sub-a.v. de $k[X]$.
(ii) $\forall P \in I$ y $\forall Q \in k[X]$ se tiene que $PQ \in I$.

Ejemplo: Sean $P_1, \dots, P_r \in k[X]$ y sea $I = \{P_1 Q_1 + \dots + P_r Q_r$ donde $Q_i \in k[X]\}$. Entonces I es el ideal más pequeño que contiene los $P_1, \dots, P_r$ y es llamado el *ideal generado* por $P_1, \dots, P_r$, que denotamos $\langle P_1, \dots, P_r \rangle$.

Terminología: Decimos que un ideal $I \subseteq k[X]$ es un *ideal principal* si $\exists P \in k[X]$ tq $I = \langle P \rangle$, ie, si I está formado por todos los múltiplos PQ de P.

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

Dem: Si $I = \{0\}$ entonces $I = \langle 0 \rangle$ está generado por el polinomio nulo ✓
Sup. 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 $\mathbb{N}$. $\Rightarrow$ Admite un elemento mínimal d. Sea $P \in I - \{0\}$ tq $gr(P) = d$. Reemplazando P por $a^{-1}P$ si fuese necesario, donde $a \in k - \{0\}$ es el coeficiente principal de P, podemos suponer que P es unitario. Veamos 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: $R = \underbrace{F}_{\in I} - \underbrace{PQ}_{\in I} \in I$.
Luego, por minimalidad de d, necesariamente $R = 0$.
$\Rightarrow F = PQ$ y luego $I = \langle P \rangle$ ✓

Veamos que P es único:
Si $I = \langle P \rangle = \langle P' \rangle$ con P' unitario, entonces $\exists Q, Q' \in k[X]$ tq $P' = PQ$ (pues $I = \langle P \rangle$) y $P = P'Q'$ (pues $I = \langle P' \rangle$) $\Rightarrow gr(Q) = gr(Q') = 0$, ie, que $Q = 1$ y luego $P' = P$. ■

Def: Sean $A, B \in k[X] \setminus \{0\}$ no-nulos. Decimos que $D \in k[X]$ es un *máximo común divisor* (mcd) de A y B si D divide A y B, y si todo divisor común de A y B divide a D. Diremos que A y B son "*primos entre sí*" si 1 es un mcd de A y B.

⚠ 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 (en caso de existir) se obtienen de este modo y en particular $\exists!$ $\mathrm{mcd}(A,B)$ *unitario*.

Prop: Sean $A, B \in k[X] - \{0\}$ no nulos. Entonces, existe $D \in k[X]$ un mcd de A y B. Dicho D se escribe como $D = AU + BV$ para ciertos $U, V \in k[X]$. En part, $\exists!$ $\mathrm{mcd}(A,B) \in k[X]$ unitario.

Dem: El conjunto $I = \langle A, B\rangle = \{AP + BQ \text{ donde } P, Q \in k[X]\}$ es un ideal $\neq \{0\}$.
$\Rightarrow I = \langle D\rangle$ para cierto $D \in k[X]$ (único si lo suponemos unitario).
En part, $D = AU + BV$ para ciertos $U, V \in k[X]$. Como $A, B \in I = \langle D\rangle$, ellos son múltiplos de D. Más aún, todo polinomio que divide tanto A y B también divide $AU + BV = D \Rightarrow D$ es un mcd de A y B. ■

Teorema de Bézout: Sean $A, B \in k[X] - \{0\}$ no nulos. Entonces, A y B son primos entre sí $\Leftrightarrow \exists\, U, V \in k[X]$ tq $AU + BV = 1$.

Dem: A y B son primos entre sí $\overset{\text{dy}}{\Leftrightarrow}$ 1 es un mcd de A y B
$\Leftrightarrow \mathrm{mcd}(A,B) = 1 \underset{\text{Prop}}{\Rightarrow} \exists\, U, V \in k[X]$ tq $AU + BV = 1$ ✓
Recíprocamente, si $AU + BV = 1$ entonces todo mcd de A y B divide 1 y es por lo tanto constante $\Rightarrow \mathrm{mcd}(A,B) = 1$ ✓ ■

Lema de Gauss: Sean $A, B, C \in k[X] - \{0\}$ polinomios no nulos. Entonces:
① Si $\mathrm{mcd}(A,B) = 1$ y A divide $BC \Rightarrow A$ divide C.
② Si $\mathrm{mcd}(A,B) = 1$ y $\mathrm{mcd}(A,C) = 1 \Rightarrow \mathrm{mcd}(A, BC) = 1$.

Dem: ① Como $\mathrm{mcd}(A,B) = 1 \Rightarrow \exists\, U, V \in k[X]$ tq $AU + BV = 1$.
Como A divide BC, $\exists P \in k[X]$ tq $BC = AP$.
$\Rightarrow C = C \cdot 1 = C(AU + BV) = ACU + \underbrace{BCV}_{=AP} = ACU + APV = A(\underbrace{CU + PV}_{Q \in k[X]})$
$\Rightarrow A$ divide C ✓

② Sea $D = \mathrm{mcd}(A, BC)$. Notar que $\mathrm{mcd}(B, D) = 1$, pues todo divisor común a D y B divide A y B, y estos últimos son primos entre sí.
Como $\mathrm{mcd}(B,D) = 1$ y D divide $BC \overset{①}{\Rightarrow} D$ divide C
Luego, D divide A y D divide $C \Rightarrow D = 1$ ✓ ■

Corolario: Sean $P, P_1, \ldots, P_m \in k[X] - \{0\}$ polinomios no nulos. Entonces:
① Si $\mathrm{mcd}(P_i, P_j) = 1$ para todo $i \neq j$ y si P_i divide P para todo i
$\Rightarrow$ El producto $P_1 \cdots P_m$ divide P.
② Si $\mathrm{mcd}(P, P_i) = 1$ para todo $i \Rightarrow \mathrm{mcd}(P, P_1 \cdots P_m) = 1$.

Dem: ① Por inducción en m: si $P = P_1 Q$ y si P_j con $j \geq 2$ divide P y $\mathrm{mcd}(P_j, P_1) = 1 \Rightarrow P_j$ divide Q (Lema de Gauss)
Por hipótesis de inducción, Q es divisible por $P_2 \cdots P_m$ ✓
② Por inducción en m: el caso base ($m=2$) dado por el Lema de Gauss ✓ ■

Ejemplo importante: Sea $P \in k[X]$ un polinomio. Vimos que $a \in k$ es una raíz de P (ie, $P(a)=0$) $\Leftrightarrow X-a$ divide a P.

En part, si $a \in k$ no es una raíz de P ($\Rightarrow P \neq 0$ es no-nulo) entonces $\mathrm{mcd}(P, X-a) = 1$ (este último es necesariamente de grado 0).

El Corolario anterior ② $\Rightarrow P$ y $(X-a)^n$ son primos entre sí $\forall n \in \mathbb{N}^{\geq 1}$.

Por otra parte, si $\lambda_1, \dots, \lambda_p \in k$ son raíces distintas de P, de multiplicidades $m_1, \dots, m_p \in \mathbb{N}^{\geq 1}$ entonces el Corolario anterior ① implica que P es divisible por el producto $\prod_{j=1}^{p} (X-\lambda_j)^{m_j} = (X-\lambda_1)^{m_1}(X-\lambda_2)^{m_2} \cdots (X-\lambda_p)^{m_p}$.

Def: Sea $P \in k[X]-\{0\}$ polinomio no nulo. Decimos que P escinde sobre k si se factoriza (sobre k) como un producto de polinomios de grado 1:

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

donde $a \in k$ y $\lambda_1, \dots, \lambda_p \in k$. En otras palabras, P escinde sobre k si y sólo si la suma $s(P) = m_1 + \dots + m_p$ de las multiplicidades de sus raíces en k es igual al grado de P: $s(P) = \mathrm{gr}(P)$.

Lema útil: Sean $P, Q \in k[X]-\{0\}$ no nulos. Si P escinde sobre k y si Q divide a $P \Rightarrow Q$ escinde sobre k.

Dem: Si $P = AQ \Rightarrow \mathrm{gr}(P) = s(P) = s(A) + s(Q)$.

Como $\mathrm{gr}(P) = \mathrm{gr}(A) + \mathrm{gr}(Q)$, $s(A) \leq \mathrm{gr}(A)$ y $s(Q) \leq \mathrm{gr}(Q)$, necesariamente $s(Q) = \mathrm{gr}(Q)$ y luego Q escinde sobre k. ■

Recuerdo: El Teorema Fundamental del Álgebra asegura que todo polinomio complejo $P \in \mathbb{C}[X]$ de grado ≥ 1 posee al menos una raíz en $\mathbb{C}$.

Como consecuencia de la división euclideana, y usando la terminología anterior, esto último equivale a:

Teorema Fundamental del Álgebra: Todo polinomio $P \in \mathbb{C}[X]-\{0\}$ no nulo con coeficientes complejos escinde sobre $\mathbb{C}$.

Ejemplos: Esto último no es cierto sobre otros cuerpos:

① $P \in \mathbb{R}[X]$ dado por $P(X) = X^3 - 1 = (X-1)(X^2+X+1)$ no escinde sobre $\mathbb{R}$, pues X^2+X+1 no posee raíces reales.

② Pequeño Teorema de Fermat (1636): Sea $a \in \mathbb{N}$ y p un número primo, entonces $a^p \equiv a \pmod p$. En part, $a^{-1} \equiv a^{p-2}$ para todo $a \not\equiv 0 \pmod p$.

Idea (inducción): $a=0$ ✓ y $(a+1)^p - (a+1) = \underbrace{a^p - a}_{\equiv 0 \pmod p} + \sum_{k=0}^{p-1} \underbrace{\binom{p}{k}}_{\equiv 0 \pmod p} a^{p-k}$ ■

Consecuencia: El polinomio $P \in \mathbb{F}_p[X]$ dado por $P(X) = X^p - X + 1$ no posee raíces en $\mathbb{F}_p$ (pues $P(a) = 1 \ \forall a \in \mathbb{F}_p$) y luego no escinde sobre $\mathbb{F}_p$.

§19. Trigonalización y Teorema de Cayley-Hamilton

Def: Sea $u: V \to V$ un endomorfismo, donde $\dim_k(V) = n$. Decimos que u es trigonalizable si existe una base $\mathcal{B}$ de V en la cual la matriz $A = \mathrm{Mat}_{\mathcal{B}}(u) \in M_n(k)$ 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, ie, si $\exists P \in Gl_n(k)$ tq $P^{-1}AP$ es triangular superior.

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

En part, podríamos reemplazar "superior" por "inferior" en la definición anterior.

Ejemplo: Sea $u: V \to V$ endomorfismo trigonalizable y sea $\mathcal{B}$ una base de V tq $A = \mathrm{Mat}_{\mathcal{B}}(u)$ es triangular superior, ie,

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

$\Rightarrow P_A(X) = \det(XI_n - A) = (X - a_{11})(X - a_{22}) \cdots (X - a_{nn})$ y luego los valores propios de A (o equivalentemente, de u) son los términos diagonales a_{ii}. En part, $P_u \in k[X]$ escinde sobre k.

Teorema: Sea $u: V \to V$ un endomorfismo, donde $\dim_k(V) = n$. Entonces:
u es trigonalizable $\Leftrightarrow$ su polinomio característico $P_u \in k[X]$ escinde sobre k.

Dem: Inducción en n: Toda matriz 1×1 es triangular ($n = 1$) ✓

Sup. $n \geq 2$ y sea $P_u(X) = \prod_{j=1}^{n} (X - \lambda_j)$, donde $\lambda_1, \dots, \lambda_n \in k$ son los valores propios de u (no necesariamente distintos).

Dado que $\lambda_n \in k$ es un valor propio de u, el endomorfismo $u - \lambda_n \mathrm{id}_V$ no es inyectivo, y luego no es sobreyectivo (pues $\dim_k(V)$ es finita).

Sea $W := \mathrm{Im}(u - \lambda_n \mathrm{id}_V) \subsetneq V \Rightarrow r := \dim_k(W) = \mathrm{rg}(u - \lambda_n \mathrm{id}_V) < n$.

Si $(e_1', \dots, e_r')$ es una base de W, la completamos en una base $\mathcal{B}' = (e_1', \dots, e_n')$ de V, y definimos $\mathcal{U} := \mathrm{Vect}_k(e_1', \dots, e_{n-1}')$.

Notar que $W \subseteq \mathcal{U}$ y que para todo $j \in \{1, \dots, n-1\}$ tenemos que:

$$u(e_j') = u(e_j') - \lambda_n e_j' + \lambda_n e_j' = \underbrace{(u - \lambda_n \mathrm{id}_V)(e_j')}_{\in W} + \underbrace{\lambda_n e_j'}_{\in \mathcal{U}} \Rightarrow u(e_j') \in \mathcal{U}.$$

Más aún, $u(e_n') = \underbrace{(u - \lambda_n \mathrm{id}_V)(e_n')}_{\in \mathcal{U}} + \lambda_n e_n' \Rightarrow \mathrm{Mat}_{\mathcal{B}'}(u) = \begin{pmatrix} A & C \\ 0 & \lambda_n \end{pmatrix}$,

donde $A \in M_{n-1}(k)$ y $C \in k^{n-1}$ es una columna $\Rightarrow P_u(X) = P_A(X) \cdot (X - \lambda_n)$.

Luego, $P_A(X) = \prod_{j=1}^{n-1} (X - \lambda_j)$ escinde sobre $k \underset{\text{Hip. Inducción}}{\Rightarrow} u|_{\mathcal{U}}: \mathcal{U} \to \mathcal{U}$ es trigonalizable, ie, $\exists (e_1, \dots, e_{n-1})$ base de $\mathcal{U}$ tq $A' = \mathrm{Mat}_{(e_1, \dots, e_{n-1})}(u|_{\mathcal{U}})$ es triangular superior. Luego, si $\mathcal{B} := (e_1, \dots, e_{n-1}, e_n') \Rightarrow \mathrm{Mat}_{\mathcal{B}}(u)$ triang. superior. ∎

Corolario: Sean $A \in M_n(\mathbb{C})$ y $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ los valores propios de A (no necesariamente distintos). Entonces, $\det(A) = \prod_{j=1}^{n} \lambda_m$ y $\operatorname{tr}(A) = \sum_{j=1}^{n} \lambda_m$.

Dem: El polinomio $P_A(X) = (X-\lambda_1)\cdots(X-\lambda_n)$ escinde sobre $\mathbb{C}$ (Teo. Fund. del Álgebra)
$\Rightarrow A$ es trigonalizable, ie, $\exists P \in GL_n(\mathbb{C})$ tq $B := P^{-1}AP$ triangular superior.
Vista: $P_B(X) = P_A(X)$ y luego $\operatorname{tr}(A) = \operatorname{tr}(B)$ y $\det(A) = \det(B)$.
Por otro lado, los términos diagonales de B son exactamente las raíces de $P_B = P_A$

$$\Rightarrow B = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \Rightarrow \operatorname{tr}(A) = \operatorname{tr}(B) = \lambda_1 + \dots + \lambda_n$$
$$\text{y } \det(A) = \det(B) = \lambda_1 \cdots \lambda_n. \quad \blacksquare$$

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

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

Recuerdo: Si $u \in \operatorname{End}_k(V)$ y $P \in k[X]$ dado por $P(X) = a_0 + a_1 X + \dots + a_d X^d$
$\Rightarrow P(u) \in \operatorname{End}_k(V)$ está dado por $P(u) = a_0 \operatorname{id}_V + a_1 u + \dots + a_d u^d$.
Más aún, si V es de dimensión finita: $\forall u \in \operatorname{End}_k(V)$, $\exists P \in k[X] - \{0\}$ tq $P(u) = 0$.

Teorema de Cayley-Hamilton: Sea $u: V \to V$ endomorfismo, entonces $P_u(u) = 0$.

Dem: Sea $\mathcal{B}$ una base de V y sea $A = \operatorname{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \in M_n(k)$.
Entonces, $P_u(u) = 0 \iff P_A(A) = 0$. Para esto último, consideremos:
$B := XI_n - A \in M_n(k(X))$, ie, $B = (b_{ij})$ con $b_{ij} = \begin{cases} X - a_{ii} & \text{si } i = j \\ -a_{ij} & \text{si } i \neq j \end{cases}$
y sea $C := {}^t\operatorname{com}(B)$ comatriz traspuesta. Dado que cada cofactor de B es un polinomio de grado $\leq n-1$, podemos escribir

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

donde $C_j \in M_n(k)$. Por otro lado, $B \cdot \underbrace{{}^t\operatorname{com}(B)}_{=C} = \underbrace{\det(B)}_{P_A(X)} I_n$.
Luego, si $P_A(X) = X^n + a_{n-1}X^{n-1} + \dots + a_1 X + a_0$ entonces:

$$B \cdot C = (XI_n - A)(C_0 + XC_1 + \dots + X^{n-1}C_{n-1}) = X^n C_{n-1} + X^{n-1}(C_{n-2} - AC_{n-1}) + \dots + X(C_0 - AC_1) + (-AC_0)$$
$$\| \qquad\qquad \| \qquad\qquad\qquad\

Ejemplo: Sea $A = \begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix} \in M_2(\mathbb{R})$. Entonces $P_A(X) = \begin{vmatrix} X & -1 \\ -2 & X-3 \end{vmatrix} = X^2 - 3X - 2$.

Cayley-Hamilton: $A^2 - 3A - 2I_2 = 0$. En efecto, $A^2 = \begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 6 & 11 \end{pmatrix}$

$\Rightarrow A^2 - 3A - 2I_2 = \begin{pmatrix} 2 & 3 \\ 6 & 11 \end{pmatrix} - 3\begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix} - 2\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ ✓

Usos típicos de Cayley-Hamilton: Como $A^2 - 3A - 2I_2 = 0$ (*)

$\Rightarrow A^2 = 3A + 2I_2$, $A^3 = 3A^2 + 2A \overset{(*)}{=} 3(3A + 2I_2) + 2A = 11A + 6I_2$, etc.

Notar que $\det(A) = \begin{vmatrix} 0 & 1 \\ 2 & 3 \end{vmatrix} = 0 - 2 \neq 0$, ie, $A \in GL_2(\mathbb{R})$.

Multiplicando (*) por A^{-1}: $A - 3I_2 - 2A^{-1} = 0 \Rightarrow A^{-1} = \dfrac{A - 3I_2}{2} = \begin{pmatrix} -3/2 & 1/2 \\ 1 & 0 \end{pmatrix}$.

§20. Polinomio minimal

Sea V un k-e.v. de dimensión finita y $u : V \to V$ un endomorfismo.

Sabemos que siempre $\exists P \in k[X]$ no-nulo tq $P(u) = 0$ en $\mathrm{End}_k(V)$.

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

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

Diremos que m_u es el polinomio minimal de u.

Dem: El sub-conjunto $I = \{P \in k[X] \text{ tq } P(u) = 0\} \subseteq k[X]$ es un ideal:

(i) Si $\lambda \in k$ y $P, Q \in I \Rightarrow \lambda P + Q \in I$ pues $(\lambda P + Q)(u) = \lambda P(u) + Q(u) = 0$

Luego, I es un sub-e.v. de $k[X]$.

(ii) Si $P \in I$ y $Q \in k[X] \Rightarrow PQ \in I$ pues $(PQ)(u) = \underbrace{P(u)}_{=0} Q(u) = 0$.

Más aún, $I \neq \{0\}$ (ver §7). $\overset{(\S 18)}{\Rightarrow} \exists!\, m_u \in k[X]$ polinomio unitario tq $I = \langle m_u \rangle$, ie, $\forall P \in k[X]$ se tiene que $P(u) = 0 \iff P \in I = \langle m_u \rangle$

$\iff P = m_u Q$ para cierto $Q \in k[X]$ ■

Terminología: Si $A \in M_n(k)$, definimos el polinomio minimal de A como el polinomio minimal $m_A \in k[X]$ del endomorfismo $u_A : k^n \to k^n$, $x \mapsto Ax$ asociado.

Ejercicio importante: Sea $u : V \to V$ un endomorfismo. Probar que:

$\mathrm{gr}(m_u) = 1$, ie, $m_u(X) = X - \lambda$ para cierto $\lambda \in k \iff u = \lambda \mathrm{id}_V$ (homotecia).

En part, $m_u(X) = X \iff u = 0$.

Teorema de Cayley-Hamilton (reformulación): Sea $u : V \to V$ un endomorfismo, entonces el polinomio minimal m_u divide al polinomio característico P_u.

Ejemplos: (1) $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in M_2(k)$ verifica $A^2 = 0$. Luego m_A divide $P(X) = X^2$.

Dado que $A \neq 0$, $m_A(X) \neq X$ y luego $m_A(X) = X^2$.

(2) $A = \begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix} \in M_2(\mathbb{R})$ verifica $P_A(X) = X^2 - 3X - 2$. Como A no es una homotecia, $\mathrm{gr}(m_A) \geq 2 \Rightarrow m_A(X) = P_A(X)$.

Cayley-Hamilton

Dado que m_u divide P_u, y las raíces de P_u son los valores propios de u, tenemos que toda raíz de m_u es un valor propio de u. El recíproco es cierto:

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

Dem: Sea $\lambda \in k$ un valor propio de u y sea $v \neq 0$ un vector propio asociado, ie, $u(v) = \lambda v$. La división euclídeana de m_u por $(X-\lambda)$ nos da:

$$m_u(X) = Q(X)(X-\lambda) + c \text{, donde } c \text{ polinomio de grado } < 1, \text{ ie, } c \in k$$

$$\Rightarrow 0 = m_u(u) = Q(u)(u - \lambda Id_V) + c\,Id_V \text{ en } End_k(V). \text{ Aplicándolo a } v:$$

$$0 = \underbrace{Q(u)((u-\lambda Id_V)(v))}_{=0} + c\,\underbrace{Id_V(v)}_{=v} = \underbrace{Q(u)(0)}_{=0} + cv = cv \underset{v \neq 0}{\Rightarrow} c = 0$$

Luego, $m_u(X) = Q(X)(X-\lambda)$, ie, $\lambda \in k$ es una raíz de m_u. ∎

Ejemplo: Sup. que $\dim_k(V) = n$, entonces para todo endomorfismo $u: V \to V$ se tiene $gr(P_u) = n$. Cayley-Hamilton: m_u divide $P_u \Rightarrow gr(m_u) \leq n$. En part, si $u: V \to V$ tiene n valores propios distintos entonces el Teorema anterior implica que todos ellos son raíces de $m_u \Rightarrow gr(m_u) \geq n$. Luego, en ese caso tenemos que $m_u = P_u$.

Ejemplo: Sea $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \in M_3(k)$ y $B = A^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \in M_3(k)$.

$\Rightarrow P_A(X) = P_B(X) = X^3$. Como A y B no son homotecias, $gr(m_A) \geq 2$ y $gr(m_B) \geq 2$. Como $B^2 = A^3 = 0$ tenemos que $m_A(X) = P_A(X) = X^3$ y $m_B(X) = X^2$ $(\neq P_B(X))$.

§21. Reducción de endomorfismos y Descomposición de Dunford:

Comencemos por probar el siguiente resultado útil:

Lema de los kernel: Sea $u: V \to V$ un endomorfismo y $P \in k[X]$ un polinomio. Entonces:

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

Dem: 1. Notar que u y $P(u)$ conmutan: $u \circ P(u) = P(u) \circ u$.
Si $v \in \ker(P(u))$, ie, $P(u)(v) = 0$, entonces $P(u)(u(v)) = u(P(u))(v)$
$\Rightarrow u(v) \in \ker P(u)$ ✓ $= u(\underbrace{P(u)(v)}_{=0}) = 0$

2. Teorema de Bézout $\Rightarrow \exists U, V \in k[X]$ tq $AU + BV = 1$.
$\Rightarrow A(u)U(u) + B(u)V(u) = id_V$. Si $v \in \ker A(u)$ y $v \in \ker B(u)$
$\Rightarrow v = id_V(v) = (A(u)U(u))(v) + (B(u)V(u))(v) = U(u)(\underbrace{A(u)(v)}_{=0}) + V(u)(\underbrace{B(u)(v)}_{=0}) = 0$

Luego, $\ker A(u)$ y $\ker B(u)$ están en suma directa. Veamos que generan $\ker P(u)$:

Primero que todo, si $v \in \ker A(u)$, ie, $A(u)(v) = 0$, entonces:

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

De manera similar $\ker B(u) \subseteq \ker P(u)$ y luego $\ker P(u) \supseteq \ker A(u) \oplus \ker B(u)$.

Sea $v \in \ker P(u)$, ie, $P(u)(v) = 0$. Si escribimos

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

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

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

De manera similar, calculamos que $A(u)(B(u)V(u)(v)) = 0$ ✓ ■

Corolario: Sea $u: V \to V$ un endomorfismo y sean $P_1, \dots, P_m \in k[X]$ polinomios tq $\text{mcd}(P_i, P_j) = 1$ si $i \neq j$. Si denotamos $P = P_1 \cdots P_m$, entonces:

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

Dem: Inducción en m: si $m = 2$ OK ✓ (Lema de los kernel).

Si $P_1, \dots, P_m$ son dos a dos primos entre sí (ie, $\text{mcd}(P_i, P_j) = 1$ si $i \neq j$), entonces el (corolario de) Lema de Gauss implica que P_1 y $P_2 \cdots P_m$ son primos entre sí. Luego, el Lema de los kernel implica:

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

Por hipótesis de inducción tenemos $\ker (P_2 \cdots P_m)(u) = \ker P_2(u) \oplus \dots \oplus \ker P_m(u)$. ■

Obs: El Teorema de Cayley-Hamilton afirma que $m_u \mid P_u$ para todo $u: V \to V$ endomorfismo. En part, si P_u escinde sobre $k \Rightarrow m_u$ escinde sobre k.

Ejercicio: Construir $A \in M_2(\mathbb{R})$ tq $m_A(X) = P_A(X) = X^2 + 1$.

Recuerdo ① Las raíces del polinomio mínimal m_u del endomorfismo $u: V \to V$ son los valores propios $\lambda_1, \dots, \lambda_p \in k$ (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 $j \in \{1, \dots, p\}$.

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.

Terminología: Sea $u: V \to V$ endomorfismo tq P_u escinde sobre k, ie,

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

i) El entero $\text{mult}_{alg}(\lambda_j) := m_j \geq 1$ es la multiplicidad algebraica de λ_j.

ii) El entero $\text{mult}_{geom}(\lambda_j) := \dim_k(V_\lambda)$ es la multiplicidad geométrica de λ_j.

Obs: Cayley-Hamilton $\Rightarrow m_j \leq n_j = \text{mult}_{alg}(\lambda_j)$.

El siguiente resultado permite comparar las multiplicidades:

Prop: Sea $u: V \to V$ un endomorfismo de un espacio vectorial de $\dim_k(V) = n$ y sean $\lambda_1, \dots, \lambda_p \in k$ (distintos) sus valores propios. Sup. que P_u escinde sobre k. Entonces:

① $\text{mult}_{geom}(\lambda_i) \leq \text{mult}_{alg}(\lambda_i)$ para todo $i \in \{1, \dots, p\}$.

② u es diagonalizable $\iff \text{mult}_{geom}(\lambda_i) = \text{mult}_{alg}(\lambda_i)$ para todo $i \in \{1, \dots, p\}$.

Dem: ① Escribamos $P_u(X) = \prod_{j=1}^{p} (X - \lambda_j)^{n_j}$ y sea $d_j = \dim_k(V_{\lambda_j})$ para $j \in \{1, \dots, p\}$. Queremos probar que $d_j \leq n_j$: Sea $\mathcal{E}_j$ una base de V_{λ_j}. Dado que los V_{λ_j} están en suma directa (ver §14), tenemos que $\mathcal{E} := (\mathcal{E}_1, \dots, \mathcal{E}_p)$ es una familia l.i. y luego se puede completar en una base $\mathcal{B}$ de V. Luego,

$$A = \text{Mat}_{\mathcal{B}}(u) = \left[\begin{array}{c|c|c|c} \lambda_1 I_{d_1} & 0 & 0 & * \\ \hline 0 & \ddots & 0 & * \\ \hline 0 & 0 & \lambda_p I_{d_p} & * \\ \hline 0 & 0 & 0 & B \end{array}\right]$$

donde $B \in M_\ell(k)$ con $\ell = n - (d_1 + \dots + d_p)$. En part, A es triangular superior por bloques y por ende tenemos: $P_u(X) = \det(XI_n - A) = P_B(X) \prod_{j=1}^{p} (X - \lambda_j)^{d_j}$
$\Rightarrow d_j \leq n_j$ para todo $j \in \{1, \dots, p\}$ ✓

② u diagonalizable $\iff \ell = 0$ $(\iff P_B(X) = 1) \iff d_j = n_j$ para todo $j \in \{1, \dots, p\}$ ∎

Para completar el círculo de ideas, falta ver cómo interviene m_j (ie, la multiplicidad de λ_j en el polinomio minimal) en la discusión:

Def: Sea $u: V \to V$ endomorfismo y sean $\lambda_1, \dots, \lambda_p \in k$ sus valores propios (distintos). Si suponemos que $m_u \in k[X]$ escinde sobre k, ie,

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

entonces el sub-ev.

$$V_{(\lambda_j)} := \ker\left[(u - \lambda_j \text{Id}_V)^{m_j}\right] \subseteq V$$

es llamado el (sub-) espacio característico (o espacio propio generalizado) asociado al valor propio $\lambda_j \in k$.

Obs: ① Si $P_j \in k[X]$ es $P_j(X) := (X - \lambda_j)^{m_j} \Rightarrow V_{(\lambda_j)} = \ker P_j(u)$.

② El Lema de los kernels $\Rightarrow u(V_{(\lambda)}) \subseteq V_{(\lambda)}$ es estable por u.

③ $V_\lambda \subseteq V_{(\lambda)}$ para todo valor propio λ.

④ Un vector no-nulo $v \in V \setminus \{0\}$ pertenece a $V_{(\lambda)} \iff \exists k \in \{1, \dots, m_\lambda\}$ tq $(u - \lambda \text{Id}_V)^k(v) = 0$, 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: Sea $A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \in M_3(\mathbb{R})$ triangular superior $\Rightarrow P_A(X) = (X-2)^2(X-3)$.

Cayley-Hamilton $\Rightarrow m_A(X) = (X-2)^m(X-3)$ con $m = 1$ o 2.

Si $m = 1$, entonces $(A-2I_3)(A-3I_3) = 0$, pero:

$$(A-2I_3)(A-3I_3) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$ no es nula $\Rightarrow \boxed{m=2}$ ✓

Espacios propios:

① $\lambda = 2$: $V_2 = \ker(A-2I_3) = ?$ $\quad \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} y \\ 0 \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow y = z = 0$

Luego, $V_2 = \text{Vect}_{\mathbb{R}}\left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\rangle = \text{Vect}_{\mathbb{R}}\langle e_1 \rangle$ "eje x"

② $\lambda = 3$: $V_3 = \ker(A-3I_3) = ?$ $\quad \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -x+y \\ -y \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow x = y = 0$

Luego, $V_3 = \text{Vect}_{\mathbb{R}}\langle e_3 \rangle$ "eje z"

Notar que $V_2 \oplus V_3$ no genera $\mathbb{R}^3$, i.e., A no es diagonalizable.

Espacios característicos:

① $\lambda = 2$: $V_{(2)} = \ker\left[(A-2I_3)^2\right] = ?$ $\quad \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

$\Rightarrow V_{(2)} = \text{Vect}_{\mathbb{R}}\langle e_1, e_2 \rangle$ "plano xy" (Notar que $m = 2 = \dim_{\mathbb{R}} V_{(2)}$).

② $\lambda = 3$: $V_{(3)} = V_3$ pues $\lambda = 3$ es raíz simple de m_A.

Notar que en este ejemplo $V_{(2)}$ y $V_{(3)}$ están en suma directa y además $\mathbb{R}^3 = V_{(2)} \oplus V_{(3)}$. ❗

Teorema: Sea $u : V \to V$ un endomorfismo y sean $\lambda_1, \dots, \lambda_p \in k$ sus valores propios (distintos). Si el polinomio característico $P_u \in k[X]$ escinde sobre k, i.e.,

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

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

① V es suma directa de los espacios característicos de u, i.e.,

$$V = \bigoplus_{j=1}^{p} V_{(\lambda_j)} = V_{(\lambda_1)} \oplus \dots \oplus V_{(\lambda_p)}, \text{ donde } V_{(\lambda_j)} = \ker\left[(u-\lambda_j \text{Id}_V)^{m_j}\right].$$

② $\dim_k(V_{(\lambda_j)}) = \text{mult}_{alg}(\lambda_j)$ es la multiplicidad algebraica n_j del valor propio λ_j para todo $j \in \{1, \dots, p\}$. En part, dado que $V_\lambda \subseteq V_{(\lambda)}$ se tiene que $\text{mult}_{geom}(\lambda) = \dim_k(V_\lambda) \leq \dim_k(V_{(\lambda)}) = \text{mult}_{alg}(\lambda)$.

Dem: ① Sea $P_j(X) := (X-\lambda_j)^{m_j}$ para $j \in \{1, \dots, p\}$. Entonces $\text{mcd}(P_i, P_j) = 1$ si $i \neq j$ y $m_u = P_1 \cdots P_p$. Luego, el (corolario de) Lema de los kernels implica que

$$\underbrace{\ker(m_u(u))}_{=0} = \underbrace{\ker P_1(u)}_{V_{(\lambda_1)}} \oplus \dots \oplus \underbrace{\ker P_p(u)}_{V_{(\lambda_p)}} \qquad \checkmark$$

$\ker(0) = V$

② Los espacios característicos son estables por u, ie, $u(V_{(\lambda)}) \subseteq V_{(\lambda)}$ para todo valor propio $\lambda \in k$. Dado que $V = V_{(\lambda_1)} \oplus \dots \oplus V_{(\lambda_p)}$, escogiendo una base $\mathcal{B}_j$ de cada $V_{(\lambda_j)}$ formamos una base $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_p)$ de V tq $\mathrm{Mat}_{\mathcal{B}}(u)$ es una matriz diagonal por bloques (cf. Ejercicio importante en §14):

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

donde $A_j \in M_{s_j}(k)$, con $s_j = \dim_k(V_{(\lambda_j)})$. Queremos probar que $s_j = n_j$:
$P_u(X) = \det(X I_n - A) = P_{A_1}(X) \cdots P_{A_p}(X)$. En part, cada P_{A_j} escinde sobre k.
Finalmente, por definición de $V_{(\lambda_j)}$ se tiene que $(A_j - \lambda_j I_{s_j})^{n_j} = 0$
$\Rightarrow \det(A_j - \lambda_j I_{s_j})^{n_j} = 0 \Rightarrow \lambda_j$ es valor propio de A_j.
Más aún, como $V_\lambda \subseteq V_{(\lambda)}$ para todo valor propio λ, y los $V_{(\lambda_1)}, \dots, V_{(\lambda_p)}$ están en suma directa, λ_j es el <u>único</u> valor propio de A_j.
Como P_{A_j} escinde sobre k y $\mathrm{gr}(P_{A_j}) = s_j \Rightarrow P_{A_j}(X) = (X - \lambda_j)^{s_j}$.
Así, $P_u(X) = P_{A_1}(X) \cdots P_{A_p}(X) = \prod_{j=1}^{p} (X - \lambda_j)^{s_j} \Rightarrow s_j = n_j$ para todo j. ∎

<u>Obs</u>: La demostración anterior prueba que <u>toda</u> matriz cuyo polinomio característico escinde 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 sólo a $\lambda_j \in k$ como valor propio y verifica $(A_j - \lambda_j I_{n_j})^{n_j} = 0$.

[<u>Corolario</u>: Sea $u: V \to V$ un endomorfismo tq P_u escinde sobre k, entonces:
u diagonalizable $\Leftrightarrow \mathrm{mult}_{geom}(\lambda) = \mathrm{mult}_{alg}(\lambda)$ para todo valor propio λ.
$\Leftrightarrow V_\lambda = V_{(\lambda)}$ para todo valor propio λ.

<u>Dem</u>: La primera equivalencia ya fue probada, y la segunda equivalencia se obtiene de $\mathrm{mult}_{geom}(\lambda) = \dim_k(V_\lambda)$ y $\mathrm{mult}_{alg}(\lambda) = \dim_k(V_{(\lambda)})$. ∎

El resultado siguiente generaliza el Corolario anterior:

[<u>Teorema</u>: Sea $u: V \to V$ un endomorfismo. Son equivalentes:
① u es diagonalizable.
② El polinomio característico P_u escinde sobre k y $\dim_k(V_\lambda) = \mathrm{mult}_{alg}(\lambda)$ para todo valor propio λ.
③ El polinomio mínimal m_u escinde sobre k y sólo tiene raíces simples.

Dem: Sup. ①, ie, u diagonalizable. Sea $\mathcal{B}$ base de V tq $A = Mat_{\mathcal{B}}(u)$ es diagonal: $A = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_p \end{pmatrix}$ donde $\lambda_1, \dots, \lambda_p \in k$ son los valores propios (distintos) de A. Luego, $P_u(X) = \prod_{j=1}^{n} (X-\lambda_j)^{n_j}$ donde $n_j = \dim_k(V_{\lambda_j})$, ie, ② ✓

Más aún, si $P(X) := \prod_{j=1}^{n} (X-\lambda_j) \Rightarrow P(A) = 0$ y luego m_u divide P. Dado que las raíces de m_u son $\lambda_1, \dots, \lambda_p \Rightarrow m_u(X) = P(X) = \prod_{j=1}^{n} (X-\lambda_j)$, ie, ① ⇒ ③ ✓

Sup ②, ie, $P_u(X) = \prod_{j=1}^{n} (X-\lambda_j)^{n_j}$ con $n_j \overset{\text{def}}{=} \text{mult}_{alg}(\lambda_j) \overset{②}{=} \dim_k(V_{\lambda_j}) \overset{\text{def}}{=} \text{mult}_{geom}(\lambda_j)$

⇒ u es diagonalizable, ie, ② ⇒ ① ✓

Sup ③, ie, $m_u(X) = \prod_{j=1}^{n} (X-\lambda_j)$. Luego $V_{(\lambda_j)} = \ker((u-\lambda \text{Id}_V)^1) = V_{(\lambda_j)}$ para todo $j \in \{1, \dots, p\}$ ⇒ u diagonalizable, ie, ③ ⇒ ① ✓ ∎

Finalizamos esta sección con una consecuencia importante de los resultados previos: la "descomposición de Dunford" expresa un endomorfismo u como la suma $u = u_s + u_n$, donde u_s es diagonalizable y u_n "nilpotente". Además, u_s y u_n conmutan: $u_s u_n = u_n u_s$. Debido a lo anterior, comenzamos por presentar las siguientes nociones y resultados previos:

Def: Sea $u: V \to V$ endomorfismo. Decimos que u es nilpotente si $\exists r \in \mathbb{N}^{\geq 1}$ tq $u^r = 0$. En tal caso, el menor r con esta propiedad es llamado el índice de nilpotencia de u.

Ejemplo: ① Si $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \in M_3(k) \Rightarrow A^2 = 0$ y luego A es nilpotente de índice de nilpotencia 2.

② Si $B = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \in M_3(k) \Rightarrow B^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ y $B^3 = 0$, y luego B es nilpotente de índice de nilpotencia 3.

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

① u nilpotente $\iff m_u(X) = X^r$ para cierto $r \in \mathbb{N}^{\geq 1}$.
Más aún, r es el índice de nilpotencia de u. En part, $u^n = 0$.

② Si u es nilpotente, 0 es el único valor propio de u.

③ Si u es nilpotente y diagonalizable $\Rightarrow u = 0$.

Dem: ① ($\Rightarrow$) Sup $u^r = 0$ con $r \in \mathbb{N}^{\geq 1}$ índice de nilpotencia. Sea $P(X) = X^r$.
$\Rightarrow P(u) = 0$ y luego $m_u(X)$ divide $P(X) = X^r$ (por definición de m_u).
$\Rightarrow m_u(X) = X^m$ con $m \leq r$. Dado que $m_u(u) = u^m = 0$, tenemos $m = r$ por minimalidad de r ✓

① ($\Leftarrow$) Si $m_u(X) = X^r \Rightarrow m_u(u) = u^r = 0$ y luego u nilpotente ✓.
Dado que m_u divide P_u (Cayley-Hamilton) $\Rightarrow r = gr(m_u) \leq gr(P_u) = n$
y luego $u^n = 0$.

② Los valores propios de u son las raíces de $m_u(X) = X^r$, ie, $\lambda = 0$ ✓

③ u nilpotente $\Rightarrow m_u(X) = X^r$. En part, u diagonalizable $\Leftrightarrow r = 1$ (raíz simple)
$\Rightarrow u = 0$. ■

Ejemplo: Sea $A \in M_n(k)$ matriz triangular superior estricta, ie, de la forma

$$A = \begin{pmatrix} 0 & * & & * \\ & 0 & * & \\ & & \ddots & * \\ 0 & & & 0 \end{pmatrix}$$

$\Rightarrow P_A(X) = X^n$ y luego $P_A(A) = A^n = 0$. Más aún, A diagonalizable $\Leftrightarrow A = 0$.

Def: Sean $u: V \to V$ y $v: V \to V$ endomorfismos. Decimos que u y v conmutan si $u \circ v = v \circ u$.

Obs: Si u y v conmutan, el Teorema del binomio de Newton implica que para todo $n \in \mathbb{N}^{\geq 1}$ se tiene que

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

⚠ Si u y v no conmutan esto es falso (eg. $(u+v)^2 \neq u^2 + 2uv + v^2$ para $u = E_{12}$ y $v = E_{21}$ en $M_2(k)$)!

Lema: Sean $u: V \to V$ y $v: V \to V$ endomorfismos que conmutan. Entonces:

① Sea $\lambda \in k$ un valor propio de u. Entonces, V_λ y $V_{(\lambda)}$ son estables por v, ie, $v(V_\lambda) \subseteq V_\lambda$ y $v(V_{(\lambda)}) \subseteq V_{(\lambda)}$.

② Si u y v son diagonalizables, $\exists \mathcal{B}$ base de V formada por vectores que son vectores propios de u y de v simultaneamente. En part, $u+v$ y uv son diagonalizables.

③ Si u y v son nilpotentes, entonces $u+v$ y uv son nilpotentes.

Dem: ① Sea $x \in V_\lambda$. entonces $u(v(x)) = v(\underbrace{u(x)}_{\lambda x}) = \lambda v(x) \Rightarrow v(x) \in V_\lambda$.
Del mismo modo, v y $P(u) = (u - \lambda Id_V)^{m_\lambda}$ conmutan, por lo que si
$x \in V_{(\lambda)}$, ie, $x \in \ker P(u) \Rightarrow P(u)(v(x)) = v(\underbrace{P(u)(x)}_{=0}) = 0 \Rightarrow v(x) \in V_{(\lambda)}$ ✓

② Sean $\lambda_1, \ldots, \lambda_p \in k$ los valores propios (distintos) de u y escribamos $V = V_{\lambda_1} \oplus \ldots \oplus V_{\lambda_p}$. Sea $\lambda = \lambda_j$ con $j \in \{1, \ldots, p\}$ fijo. Queremos ver que V_λ admite una base $\mathcal{B}_\lambda$ formada por vectores propios de v.

Sean $\mu_1, \ldots, \mu_r \in k$ los valores propios (distintos) de v. Dado que v es diagonalizable, todo $x \in V_\lambda$ se escribe como $x = x_1 + \ldots + x_r$, con $x_i \in V_{\mu_i}$.

Inducción en r (nº de valores propios de v): Si $r=1 \Rightarrow x = x_1 \in V_{\mu_1}$ y $x \in V_\lambda$ ✓

Sup. $r \geq 2$ y escribamos $x = x_1 + \ldots + x_r$ (de manera única) con $x_i \in V_{\mu_i}$.

$$\Rightarrow \tilde{x} := (v - \mu_r \mathrm{Id}_V)(x) = (v(x_1) - \mu_r x_1) + \ldots + (v(x_{r-1}) - \mu_r x_{r-1}) + \underbrace{(v(x_r) - \mu_r x_r)}_{=0}$$

$$= \underbrace{(\mu_1 - \mu_r)}_{\neq 0} x_1 + \ldots + \underbrace{(\mu_{r-1} - \mu_r)}_{\neq 0} x_{r-1}$$

Por ④, $v(V_\lambda) \subseteq V_\lambda \Rightarrow \tilde{x} \in V_\lambda$. Hip. de Inducción $\Rightarrow x_1, \ldots, x_{r-1} \in V_\lambda$.

Luego, $x_r = \underbrace{x}_{\in V_\lambda} - (\underbrace{x_1 + \ldots + x_{r-1}}_{\in V_\lambda}) \in V_\lambda \Rightarrow$ Cada $x_i \in V_{\mu_i}$ pertenece a V_λ.

Así, V_λ admite una base $\mathcal{B}_\lambda$ formada por vectores propios x_i de v, que son todos vectores propios de u asociados a λ, ie, $x_i \in V_\lambda$.

$\Rightarrow \mathcal{B} = (\mathcal{B}_{\lambda_1}, \ldots, \mathcal{B}_{\lambda_p})$ es una base de V formada por vectores propios comunes de u y v. En part,

$$\mathrm{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} \lambda_1 \cdots \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_p \cdots \lambda_p \end{pmatrix} \quad \text{y} \quad \mathrm{Mat}_{\mathcal{B}}(v) = \begin{pmatrix} \mu_1 \cdots \mu_1 & & 0 \\ & \ddots & \\ 0 & & \mu_r \cdots \mu_r \end{pmatrix}$$

$\Rightarrow$ En la base $\mathcal{B}$, $u+v$ y uv son diagonales ✓

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

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

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, para que $u^k v^{n-k}$ sea $\neq 0$ es necesario que $k \leq r-1$ y $n-k \leq s-1$

$\Rightarrow n \leq r+s-2$. En part, $(u+v)^{r+s-1} = 0$ ✓. ■

Ejercicio ① Sean $A = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$ y $B = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}$ en $M_2(\mathbb{R})$. Probar que $AB \neq BA$, que A y B son diagonalizables, pero $A+B$ y AB no lo son.

② Sean $E_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ y $E_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ en $M_2(k)$. Probar que $E_{12}^2 = E_{21}^2 = 0$, pero $E_{12}E_{21} \neq E_{21}E_{12}$ y además $E_{12}E_{21} = E_{11}$ no es nilpotente, y $S = E_{21} + E_{12}$ tampoco es nilpotente.

(En part, si u y v no conmutan, las conclusiones del Lema son falsas).

Teorema (Descomposición de Dunford): Sea $u : V \to V$ un endomorfismo tq el 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, ie, $u_S u_n = u_n u_S$.

↳ ⚠ También conocida como "Descomposición de Jordan-Chevalley"

Dem: Existencia: Vimos (pag. 47) que todo endomorfismo $u: V \to V$ tq P_u escinde sobre k admite una base $\mathcal{B}$ de V tq $A = Mat_{\mathcal{B}}(u) \in M_n(k)$ es diagonal por bloques:

$$A = \begin{pmatrix} A_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & A_p \end{pmatrix}$$

donde cada bloque A_j posee sólo a $\lambda_j \in k$ como valor propio y verifica $(A_j - \lambda_j I_{m_j})^{m_j} = 0$ para $m_j \in \mathbb{N}^{\geq 1}$ (la mult. de λ_j como raíz de m_u).

Sea $u_s : V \to V$ el endomorfismo con $S = Mat_{\mathcal{B}}(u_s)$ dada por la matriz diagonal

$$S = \begin{pmatrix} \lambda_1 I_{m_1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_p I_{m_p} \end{pmatrix}.$$

En part, u_s es diagonalizable. Sea $u_n := u - u_s$, cuya matriz $N = Mat_{\mathcal{B}}(u_n)$ está dada por

$$N = \begin{pmatrix} A_1 - \lambda_1 I_{m_1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & A_p - \lambda_p I_{m_p} \end{pmatrix}.$$

Dado que $(A_j - \lambda_j I_{m_j})^{m_j} = 0$, tenemos que $u_n^r = 0$ para $r \geq \max(m_1, \dots, m_p)$.

Finalmente, notar que $u_s|_{V(\lambda_j)} = \lambda_j \, Id_{V(\lambda_j)}$ y que si $v \in V(\lambda_j) \Rightarrow u(v) \in V(\lambda_j)$

Luego, para $v \in V(\lambda_j)$ tenemos que $u(u_s(v)) = u(\lambda_j v) = \lambda_j u(v)$ y que $u_s(\underbrace{u(v)}_{\in V(\lambda_j)}) = \lambda_j u(v)$, ie, $u u_s = u_s u$ en $V(\lambda_j)$. Dado que $V = \bigoplus_{j=1}^{p} V(\lambda_j)$

$\Rightarrow u u_s = u_s u$ en V. Luego, $u_n u_s = (u - u_s) u_s = u u_s - u_s^2 = u_s u - u_s^2$

$= u_s (u - u_s) = u_s u_n$ ✓

Unicidad: Sea $u = u_s' + u_n'$ otra descomposición con las mismas propiedades.

Como u_s' y u_n' conmutan, $u_s' u = u u_s'$ y $u_n' u = u u_n'$. En part, por las propiedades de endomorfismos que conmutan (Lema en pag 49), tenemos que $u_s'(V(\lambda_j)) \subseteq V(\lambda_j)$ y $u_n'(V(\lambda_j)) \subseteq V(\lambda_j)$, ie, el espacio característico $V(\lambda_j)$ de u es estable por u_s' y u_n'. Dado que $u_s|_{V(\lambda_j)} = \lambda_j \, Id_{V(\lambda_j)}$ es una homotecia, y $V = \bigoplus_{j=1}^{p} V(\lambda_j)$, $u_s u_s' = u_s' u_s$ y $u_s u_n' = u_n' u_s$.

En part, u_n' y u_s' conmutan con $u_n = u - u_s$.

Finalmente, como u_s y u_s' (resp. u_n y u_n') conmutan y son diagonalizables (resp. nilpotentes) $\Rightarrow u_s - u_s'$ es diagonalizable y $u_n' - u_n$ es nilpotente.

Así, $u = u_s + u_n = u_s' + u_n' \Rightarrow u_s - u_s' = u_n' - u_n$ es un endomorfismo diagonalizable y nilpotente, luego debe ser nulo, ie, $u_s = u_s'$ y $u_n = u_n'$ ■

Caso particular importante: $k = \mathbb{R}$.

Sea $A \in M_n(\mathbb{R})$ matriz real, 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 característicos $V_{(\mu_j)}$ 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:

i) mínimal real $m_A \in \mathbb{R}[X]$, ie, el polinomio real unitario de grado más pequeño tal que $m_A(A) = 0$ en $M_n(\mathbb{R})$.

ii) mínimal complejo $M_A \in \mathbb{C}[X]$, ie, el polinomio complejo unitario de grado más pequeño tq $M_A(A) = 0$ en $M_n(\mathbb{C})$.

[Prop: $m_A = M_A \in \mathbb{R}[X]$. (no necesariamente escinde sobre $\mathbb{R}$).

Dem: Sabemos que M_A divide a m_A. Sea $M_A(X) = X^m + a_{m-1}X^{m+1} + \dots + a_1 X + a_0$ con $a_i \in \mathbb{C}$. Dado que $M_A(A) = 0$, tenemos que $\mathrm{Re}(M_A)(A) = 0$, ie,

$$\mathrm{Re}(M_A)(A) = A^m + \mathrm{Re}(a_{m-1})A^{m-1} + \dots + \mathrm{Re}(a_1)A + \mathrm{Re}(a_0)I_m = 0$$

$\Rightarrow m_A$ divide $\mathrm{Re}(M_A)$, y luego:

$$\left.\begin{array}{l} m = gr(M_A) \underset{M_A \text{ divide } m_A}{\leq} gr(m_A) \underset{m_A \text{ divide } \mathrm{Re}(M_A)}{\leq} gr(\mathrm{Re}(M_A)) = m \end{array}\right\} \Rightarrow m_A = M_A \quad \blacksquare$$

Ejercicio Sea $A \in M_n(\mathbb{R})$ matriz real. Sup. que $\exists P \in GL_n(\mathbb{C})$ matriz invertible compleja tq $P^{-1}AP$ es diagonal con coeficientes reales. Probar que $\exists Q \in GL_n(\mathbb{R})$ matriz invertible real tq $Q^{-1}AQ$ es diagonal.

Ejemplo: Sea $A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix} \in M_4(\mathbb{C})$ (matriz de permutación).

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

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

Para calcular los espacios propios:

$\lambda_1 = 1$: $A - \lambda_1 I_4 = \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 1 & 0 & 0 & -1 \end{pmatrix}$ $\rightsquigarrow$ Para encontrar una base de $V_1 = \ker[A - I_4]$ usamos operaciones columna (cf. §6, pág 12):

$$\begin{pmatrix} \boxed{-1} & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 1 & 0 & 0 & -1 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow{C_2 \mapsto C_2 + C_1} \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & \boxed{-1} & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 1 & 1 & 0 & -1 \\ \hline 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow{C_3 \mapsto C_3 + C_2} \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & \boxed{-1} & 1 \\ 1 & 1 & 1 & -1 \\ \hline 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow{C_4 \mapsto C_4 + C_3} \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 1 & 1 & 1 & \boxed{0} \\ \hline 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \Rightarrow V_1 = \text{Vect}_{\mathbb{C}} \left\langle \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \right\rangle$$

Otra forma: $\begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 1 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} y-x \\ z-y \\ t-z \\ x-t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow x = y = z = t \Rightarrow V_1 = \text{Vect}_{\mathbb{C}} \left\langle \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \right\rangle$.

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

Importante: Gracias a la discusión anterior, basta calcular V_i para obtener $\overline{V_i} = V_{-i}$ (conjugados)! (Pues $A \in M_4(\mathbb{C})$ tiene coeficientes reales).

Finalmente, dado que A es diagonalizable, su descomposición de Dunford es $A = S + N$, con $S = A$ (diagonalizable) y $N = 0$ (nilpotente).

Ejemplo: Sea $A = \begin{pmatrix} 2 & -1 & 2 \\ 10 & -5 & 7 \\ 4 & -2 & 2 \end{pmatrix} \in M_3(\mathbb{R})$.

$$\Rightarrow P_A(X) = \det \begin{pmatrix} X-2 & \boxed{1} & -2 \\ -10 & X+5 & -7 \\ -4 & 2 & X-2 \end{pmatrix} = \det \begin{pmatrix} X & 1 & -2 \\ 2X & X+5 & -7 \\ 0 & 2 & X-2 \end{pmatrix} = X \cdot \det \begin{pmatrix} \boxed{1} & 1 & -2 \\ 2 & X+5 & -7 \\ 0 & 2 & X-2 \end{pmatrix}$$

$$= X \det \begin{pmatrix} 1 & 1 & -2 \\ 0 & X+3 & -3 \\ 0 & 2 & X-2 \end{pmatrix} \overset{C_2 \mapsto C_2 + 2C_1}{=} X \det \begin{pmatrix} 1 & -1 & -2 \\ 0 & X & -3 \\ 0 & X & X-2 \end{pmatrix} \overset{C_2 \mapsto C_2 + C_3}{=} X \left(X^2 - 2X + 3X\right)$$

$F_2 \mapsto F_2 - 2F_1$

$\Rightarrow P_A(X) = X^2(X+1)$. Luego, los valores propios de A son $\lambda_1 = -1$ (raíz simple) y $\lambda_2 = 0$ (raíz doble): $\text{mult}_{alg}(\lambda_1) = 1$, $\text{mult}_{alg}(\lambda_2) = 2$.

Obs: Dado que $\text{mult}_{geom}(\lambda) \leq \text{mult}_{alg}(\lambda)$ $\Rightarrow$ $\dim(V_{-1}) = 1$

$\overset{\text{dg}}{=} \dim_{\mathbb{R}}(V_\lambda) \geq 1$ $\qquad$ $\dim(V_0) = 1$ ó 2.

$\lambda_1 = -1$: Dado que λ_1 es raíz simple, $V_{-1} = V_{(-1)} = \text{Vect}_{\mathbb{R}}(v_1)$ para cierto $v_1 \in \mathbb{R}^3$. en $\ker[A - (-1)I_3] = \ker[A + I_3]$. Basta resolver:

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

$x = 1$: $\begin{cases} -y + 2z = -3 \ /\cdot(-4) \\ -4y + 7z = -10 \end{cases} \Rightarrow -z = 2 \Rightarrow z = -2$; $y = +3 + 2z = -1$.

$\Rightarrow v_1 = \begin{pmatrix} 1 \\ -1 \\ -2 \end{pmatrix}$ genera la recta $V_{\lambda_1} = V_{-1}$.

$\lambda_2 = 0$: $V_{\lambda_2} = V_0 = \ker[A - 0\cdot I_3] = \ker[A]$. Operaciones columna:

$$\begin{pmatrix} 2 & -1 & 2 \\ 10 & -5 & 7 \\ 4 & -2 & 2 \end{pmatrix} \xrightarrow[C_1 \leftrightarrow C_2]{C_2 \mapsto -C_2} \begin{pmatrix} 1 & 2 & 2 \\ 5 & 10 & 7 \\ 2 & 4 & 2 \end{pmatrix} \xrightarrow[C_3 \mapsto C_3 - 2C_1]{C_2 \mapsto C_2 - 2C_1} \begin{pmatrix} 1 & 0 & 0 \\ 5 & 0 & -3 \\ 2 & 0 & -2 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} 0 & 1 & 0 \\ -1 & 2 & 2 \\ 0 & 0 & 1 \end{pmatrix}$$

$\Rightarrow \ker(A) = \mathrm{Vect}_{\mathbb{R}}(v_2)$ con $v_2 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} \in \mathbb{R}^3$

Obs: En part, $\underbrace{\mathrm{mult}_{geom}(\lambda_2)}_{= \dim_{\mathbb{R}}(V_{\lambda_2})} = 1 < \underbrace{\mathrm{mult}_{alg}(\lambda_2)}_{= \dim_{\mathbb{R}}(V_{(\lambda_2)})} = 2$ y A no es diagonalizable.

Más aún, $m_A(X) = X^2(X+1)$.

Espacio característico: $V_{(\lambda_2)} = V_{(0)} = \ker(A^2)$. Calculamos $A^2 = \begin{pmatrix} 2 & -1 & 1 \\ -2 & 1 & -1 \\ -4 & 2 & -2 \end{pmatrix}$

$$\begin{pmatrix} 2 & -1 & 1 \\ -2 & 1 & -1 \\ -4 & 2 & -2 \end{pmatrix} \xrightarrow[C_3 \mapsto C_3 + C_2]{C_1 \mapsto C_1 + 2C_2} \begin{pmatrix} 0 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 0 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

$\Rightarrow \ker(A^2) = \mathrm{Vect}_{\mathbb{R}}(v_2, v_3)$ con $\underbrace{v_2 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}}_{\text{vector propio}}$ y $\underbrace{v_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}}_{\text{vector propio generalizado}}$.

Sea $\mathcal{B} = (v_1, v_2, v_3)$ y $P = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 2 & 1 \\ -2 & 0 & 1 \end{pmatrix}$

Vimos (pág. 47): $P^{-1}AP = \left(\begin{array}{c|c} A_1 & 0 \\ \hline 0 & A_2 \end{array}\right)$ donde $\begin{cases} \bullet\ A_1 \in M_1(\mathbb{R}) \text{ posee sólo } \lambda_1 = -1 \text{ como valor propio} \\ \bullet\ A_2 \in M_2(\mathbb{R}) \text{ posee sólo } \lambda_2 = 0 \text{ como valor propio} \end{cases}$

Notar que $\begin{cases} A(v_1) = (-1)v_1 \Rightarrow A_1 = (-1) \in M_1(\mathbb{R}) \\ A(v_2) = 0 \cdot v_2 \end{cases}$ y $A(v_3) = \begin{pmatrix} 2 & -1 & 2 \\ 10 & -5 & 7 \\ 4 & -2 & 2 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} = v_2 \Rightarrow A_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in M_2(\mathbb{R})$

$\Rightarrow P^{-1}AP = \left(\begin{array}{c|cc} -1 & 0 & 0 \\ \hline 0 & 0 & 1 \\ 0 & 0 & 0 \end{array}\right)$. Más aún, $P^{-1} = \begin{pmatrix} 2 & -1 & 1 \\ -1 & 1 & -1 \\ 4 & -2 & 3 \end{pmatrix}$ (Ejercicio)

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

Finalmente, para encontrar la descomposición de Dunford $A = S + N$ procedemos como en la demostración del Teorema (ver pág. 51):

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

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

Luego, $S = PS'P^{-1} = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 2 & 1 \\ -2 & 0 & 1 \end{pmatrix}\begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 2 & -1 & 1 \\ -1 & 1 & -1 \\ 4 & -2 & 3 \end{pmatrix} = \begin{pmatrix} -2 & 1 & -1 \\ 2 & -1 & 1 \\ 4 & -2 & 2 \end{pmatrix}$ (diagonalizable)

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

Ejercicio Verificar que $A = S + N$, $SN = NS$, S diagonalizable y N nilpotente.

Ejemplo importante: Una Ecuación Diferencial lineal homogenea de orden n con coeficientes constantes es una ecuación de la forma

$$(E) \quad x^{(n)}(t) + a_{n-1}x^{(n-1)}(t) + \ldots + a_1 x'(t) + a_0 x(t) = 0,$$

donde los $a_j \in \mathbb{C}$ son constantes y $x : \mathbb{R} \to \mathbb{C}$, $t \mapsto x(t)$ es una función a valores complejos n veces derivable.

Si definimos $x_0(t) := x(t)$, $x_1(t) := x'(t)$, $\ldots$, $x_{n-1}(t) := x^{(n-1)}(t)$ y consideramos los vectores $X(t) = \begin{pmatrix} x_0(t) \\ \vdots \\ x_{n-1}(t) \end{pmatrix}$ y $X'(t) = \begin{pmatrix} x_0'(t) \\ \vdots \\ x_{n-1}'(t) \end{pmatrix} = \begin{pmatrix} x_1(t) \\ \vdots \\ x_{n-1}(t) \\ x^{(n)}(t) \end{pmatrix}$

Entonces, $(E) \Leftrightarrow X'(t) = AX(t)$ donde

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & & & & 1 \\ -a_0 & -a_1 & & & -a_{n-1} \end{pmatrix}$$

Veremos más adelante que para resolver (E) hay que analizar $P_A \in \mathbb{C}[X]$:

[**Obs**: Sea $A \in M_n(k)$ y $(XI_n - A) \in M_n(k(X))$. Entonces, ${}^t(XI_n - A) = XI_n - {}^tA$ en $M_n(k(X)) \Rightarrow P_A = P_{{}^tA}$ para toda matriz $A \in M_n(k)$! **Similar**: $m_A = m_{{}^tA}$

Luego, basta calcular $P_B \in \mathbb{C}[X]$, donde $B = \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & & 0 & -a_1 \\ 0 & 1 & \ddots & 0 & \vdots \\ \vdots & & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & 1 & -a_{n-1} \end{pmatrix}$

Contrariamente a lo hecho hasta ahora, calcularemos m_B para deducir P_B:
Sea $u = u_B : \mathbb{C}^n \to \mathbb{C}^n$, $v \mapsto Bv$ el endomorfismo asociado y $\mathcal{B} = (e_1, \ldots, e_n)$ la base canónica de $\mathbb{C}^n$. Entonces:

$u(e_1) = e_2$, $u(e_2) = e_3$, $\ldots$, $u(e_{n-1}) = e_n$, $u(e_n) = -a_0 e_1 - \ldots - a_{n-1} e_n$.

Sea $m_B(X) = b_0 + b_1 X + \ldots + b_{m-1} X^{m-1} + X^m$ polinomio mínimal. Si $m < n$:

$(m_B(u))(e_1) = b_0 e_1 + b_1 e_2 + \ldots + b_{m-1} e_m + e_{m+1} \neq 0$ ↯ (pues $m_B(u) = 0$)

Luego, $m = n$ y en particular $P_B = m_B$. Más aún:

$$\begin{aligned}(m_B(u))(e_1) &= b_0 e_1 + b_1 e_2 + \ldots + b_{n-1} e_n - a_0 e_1 - \ldots - a_{n-1} e_n \\ &= (b_0 - a_0) e_1 + \ldots + (b_{n-1} - a_{n-1}) e_n = 0\end{aligned}$$

$\Rightarrow a_0 = b_0$, $a_1 = b_1$, $\ldots$, $a_{n-1} = b_{n-1}$, i.e.,

$$P_B(X) = m_B(X) = X^n + a_{n-1}X^{n-1} + \ldots + a_1 X + a_0$$

En conclusión: $P_A(X) = m_A(X) = X^n + a_{n-1}X^{n-1} + \ldots + a_1 X + a_0$. En part,

A diagonalizable $\Leftrightarrow$ Todas las raíces de $P_A = m_A$ son simples

$\underset{P_A = m_A}{\Leftrightarrow}$ A posee n valores propios distintos.

§23. Forma canónica de Jordan

Motivación: Incluso sobre $\mathbb{C}$, no toda matriz es diagonalizable (eg. $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in M_2(\mathbb{C})$). Sin embargo, vimos (cf. Observación en pág. 47) que al considerar bases de los espacios característicos $V(\lambda)$ podemos probar que toda matriz en $M_n(k)$ cuyo polinomio característico escinde sobre k (eg. $k = \mathbb{C}$) 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 A_j posee sólo $\lambda_j \in k$ como valor propio y $A_j - \lambda_j I_{n_j}$ nilpotente.
Una consecuencia importante de lo anterior es que todo endomorfismo $u: V \to V$ tq P_u escinde sobre k se descompone como

$$u = u_s + u_n \quad \text{(descomposición de Dunford)}$$

donde u_s es diagonalizable y u_n es nilpotente.
En esta sección, veremos un resultado importante de Camille Jordan (1838-1922) que permite describir de manera más precisa las matrices A_j antes mencionadas.

Def: Sea $n \in \mathbb{N}^{\geq 1}$ y $\lambda \in k$. Un **bloque de Jordan** de tamaño n asociado a λ es una matriz en $M_n(k)$ dada por:

$$J_n(\lambda) := \begin{pmatrix} \lambda & 1 & & 0 \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix}.$$

Un **bloque de Jordan nilpotente** de tamaño n es $J_n := J_n(0)$, ie,

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

Notar que el índice de nilpotencia de J_n es n (ie, $J_n^{n-1} \neq 0$ y $J_n^n = 0$) y que $J_n(\lambda) = \lambda I_n + J_n$ es la descomposición de Dunford de $J_n(\lambda)$.

Ejemplo: $J_1(\lambda) = (\lambda)$, $J_2(\lambda) = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$, $J_3(\lambda) = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}$, etc.

Def: Sea $n \in \mathbb{N}^{\geq 1}$ un entero. Una **partición de n** es una sucesión finita y decreciente de naturales $\underline{n} = (n_1, \ldots, n_s) \in (\mathbb{N}^{\geq 1})^s$ para cierto $s \geq 1$ tal que

$$|\underline{n}| := n_1 + \ldots + n_s = n.$$

Es decir, es una escritura de n como suma de enteros positivos.

Ejemplo: Las posibles particiones de $n = 4$ son: (4), $(3,1)$, $(2,2)$, $(2,1,1)$ y $(1,1,1,1)$.

Obs: Las particiones son muy importantes en matemáticas y en física (eg. Bohr y Kalckar, 1937), y tienen muchas propiedades interesantes (eg. Ramanujan: "si el último dígito de n es 4 o 9, el nº de particiones de n es múltiplo de 5").

Def: Sea $\lambda \in k$, $m \in \mathbb{N}^{\geq 1}$ y $\underline{m} = (m_1, \dots, m_s) \in (\mathbb{N}^{\geq 1})^s$ una partición de m. La matriz de Jordan de tamaño m asociada a $\lambda \in k$ y a $\underline{m} = (m_1, \dots, m_s)$ es la matriz por bloques dada por:

$$J_{\underline{m}}(\lambda) = J_{(m_1,\dots,m_s)}(\lambda) = \begin{pmatrix} J_{m_1}(\lambda) & & 0 \\ & \ddots & \\ 0 & & J_{m_s}(\lambda) \end{pmatrix} \in M_m(k)$$

Ejemplos: Para $m = 4$ y $\lambda \in k$ tenemos:

$$J_{(4)}(\lambda) = J_4(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{pmatrix}, \quad J_{(3,1)}(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 0 \\ 0 & 0 & 0 & \lambda \end{pmatrix}, \quad J_{(2,2)}(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 0 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{pmatrix}$$

$$J_{(2,1,1)}(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 0 & 0 \\ 0 & 0 & \lambda & 0 \\ 0 & 0 & 0 & \lambda \end{pmatrix}, \quad J_{(1,1,1,1)}(\lambda) = \begin{pmatrix} \lambda & 0 & 0 & 0 \\ 0 & \lambda & 0 & 0 \\ 0 & 0 & \lambda & 0 \\ 0 & 0 & 0 & \lambda \end{pmatrix}$$

Teorema: Sea V un k-e.v. de $\dim_k(V) = n$ y sea $u : V \to V$ nilpotente. Entonces, $\exists\, \mathcal{B}$ base de V tal que $\mathrm{Mat}_{\mathcal{B}}(u) = J_{\underline{m}} := J_{\underline{m}}(0)$ para cierta partición $\underline{m}$ de n.

Más aún, la partición $\underline{m}$ está completamente determinada por u. En part, toda matriz nilpotente $A \in M_n(k)$ es semejante a una única matriz $J_{\underline{m}}$, y luego hay una biyección entre clases de semejanza de matrices nilpotentes de $M_n(k)$ y el conjunto $\mathcal{P}(n)$ de particiones de n.

Dem: Sea $e_0 \in V \setminus \{0\}$ vector no-nulo y sea $r \in \mathbb{N}^{\geq 1}$ el menor entero positivo tq $u^r(e_0) = 0$. En part, $r \leq n$.

Paso 1 Sea $e_i := u^i(e_0)$ para $i \in \{0, \dots, r-1\}$. Entonces, $\mathcal{F} = (e_0, \dots, e_{r-1})$ es l.i.

En efecto, por inducción en r tenemos que: OK si $r = 1$: $\mathcal{F} = (e_0)$ ✓

Si $r \geq 2$ y $\exists\, a_0, \dots, a_{r-1} \in k$ tq $a_0 e_0 + \dots + a_{r-1} e_{r-1} = 0$ $(\star)$, entonces u aplicado a $(\star)$ nos da: $a_0 e_1 + \dots + a_{r-2} e_{r-1} = 0$.

Dado que el menor entero positivo $r' \in \mathbb{N}^{\geq 1}$ tq $u^{r'}(u(e_0)) = u^{r'}(e_1) = 0$ es $r' = r-1$, la hipótesis de inducción implica que $(e_1, \dots, e_{r-1})$ es l.i.

$\Rightarrow a_0 = \dots = a_{r-2} = 0$ y $(\star)$ se reduce a $a_{r-1} e_{r-1} = 0 \underset{e_{r-1} \neq 0}{\Longrightarrow} a_{r-1} = 0$ ✓

Caso particular: Si $r = n$ entonces $\mathcal{B} = (e_0, \dots, e_{n-1})$ es una base de V. En part, la relación $u(e_0) = e_1, \dots, u(e_{n-2}) = e_{n-1}$ y $u(e_{n-1}) = 0$ implica $\mathrm{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & & 0 & \vdots \\ 0 & 1 & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & 0 & 0 \\ 0 & 0 & & 1 & 0 \end{pmatrix}$. Luego, si $\mathcal{B} = (e_{n-1}, \dots, e_1, e_0)$ entonces

$$\mathrm{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & 0 & 1 & \ddots & 0 \\ \vdots & \vdots & 0 & \ddots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix} = J_n \quad \text{matriz de Jordan.}$$

Caso general: Sea $p \in \mathbb{N}^{\geq 1}$ el índice de nilpotencia de u (ie, $u^{p-1} \neq 0$ y $u^p = 0 \Leftrightarrow m_u(X) = X^p$). Dejamos $N_i := \ker(u^i)$ para $i \in \{0, \dots, p\}$. Notar que $u^i(x) = 0 \Rightarrow u^{i+1}(x) = u(u^i(x)) = 0$. Luego, tenemos inclusiones:

$$\{0\} = N_0 \subseteq N_1 \subseteq \dots \subseteq N_p = V = V_{(0)}$$

Paso 2 La inclusión $N_{i-1} \subseteq N_i$ es estricta, ie, $N_{i-1} \neq N_i$

En efecto, si $N_{i-1} = N_i$ para cierto $i \in \{1, \dots, p\}$ entonces para todo $x \in V$ se tiene que: $u^p(x) = 0 = u^i(u^{p-i}(x)) \Rightarrow u^{p-i}(x) \in N_i = N_{i-1}$
$\Rightarrow u^{i-1}(u^{p-i}(x)) = u^{p-1}(x) = 0$ ↯ (p es mínimal !!). ✓

Luego, $\{0\} = N_0 \subsetneq N_1 = \ker(u) \subsetneq N_2 = \ker(u^2) \subsetneq \dots \subsetneq N_p = V$.
En part, $\dim_k(N_i) \geq i$ para todo $i \in \{1, \dots, p\}$ y $p \leq \dim_k(N_p) = n$.

Paso 3 $u(N_{i+1}) \subseteq N_i$ para todo $i \in \{0, \dots, p-1\}$:

Si $x \in N_{i+1}$ entonces $u^{i+1}(x) = 0 = u^i(u(x))$. En part, $u(x) \in N_i$ ✓

Paso 4 Usando inducción descendente sobre $i \in \{0, \dots, p\}$, construiremos un sub-e.v. $M_i \subseteq V$ tq $N_i = N_{i-1} \oplus M_i$ y $u(M_i) \subseteq M_{i-1}$.

[Recuerdo (inducción descendente): Sea $n_0 \in \mathbb{N}$ y $P(n)$ proposición lógica.
Si ① $P(n_0)$ es verdad; ② $\forall n \in \{0, \dots, n_0\}$, $P(n)$ verdad $\Rightarrow P(n-1)$ verdad.
$\Rightarrow P(n)$ es verdad $\forall n \in \{0, \dots, n_0\}$.]

$n_0 = p$: Sea M_p cualquier espacio complementario de N_{p-1} en V.
$\Rightarrow V = N_p = N_{p-1} \oplus M_p$.

Sup que M_i ($i \geq 2$) está construido, y notar que:

i) $u(M_i) \subseteq N_{i-1}$: pues $M_i \subseteq N_i \Rightarrow u(M_i) \subseteq u(N_i) \subseteq N_{i-1}$ ✓

ii) $u(M_i) \cap N_{i-2} = \{0\}$: pues $M_i \cap N_{i-1} = \{0\}$. En efecto, si existiera $x \in (u(M_i) \cap N_{i-2}) \setminus \{0\}$ no-nulo, entonces $x = u(m)$ con $m \in M_i \subseteq N_i \setminus N_{i-1}$ (ie, $u^{i-1}(m) \neq 0$ y $u^i(m) = 0$). Luego, $u^{i-2}(x) \neq 0$ y $u^{i-1}(x) = 0$, es decir, $x \in N_{i-1} \setminus N_{i-2}$ ↯

Así, basta tomar un sub-e.v. M_{i-1} en N_{i-1} complementario de N_{i-2} y que contenga $u(M_i)$
$\Rightarrow N_{i-1} = N_{i-2} \oplus M_{i-1}$ y $u(M_i) \subseteq M_{i-1}$ ✓

Obs: Notar que
$$\begin{array}{ccccc} N_i & = & N_{i-1} & \oplus & M_i \\ u\downarrow & & u\downarrow & & u\downarrow \\ N_{i-1} & = & N_{i-2} & \oplus & M_{i-1} \end{array}$$
pues $u(N_i) \subseteq N_{i-1}$ y $u(M_i) \subseteq M_{i-1}$
"u respeta la descomposición"

$$\Rightarrow V = N_p = N_{p-1} \oplus M_p = N_{p-2} \oplus M_{p-1} \oplus M_p = \dots = \underbrace{M_1 \oplus \dots \oplus M_i}_{= N_i} \oplus M_{i+1} \oplus \dots \oplus M_p$$

Más aún, dados que $M_1 = N_1 = \ker(u)$ y $M_i \cap M_1 = \{0\}$ para $i \geq 2$, se tiene que $u|_{M_i} : M_i \hookrightarrow M_{i-1}$ es inyectiva para $i \geq 2$. En part, $\dim_k(M_i) \leq \dim_k(M_{i-1})$.

Importante: Gracias a la discusión anterior, tenemos que

$$\underline{m} = (m_1, \dots, m_p) \in (\mathbb{N}^{\geq 1})^p$$

con $m_j = \dim_k(M_j)$, es una partición de $n = \dim_k(V)$ completamente determinada por u. En efecto, $V = \bigoplus_{j=1}^{p} M_j$ implica $m_1 + \dots + m_p = n$ y además tenemos que $m_1 \geq m_2 \geq \dots \geq m_p$. Finalmente, dado que $N_i = \bigoplus_{j=1}^{i} M_j$ se tiene que $m_i = \dim_k(N_i) - \dim_k(N_{i-1})$ está determinado por u, pues $N_i = \ker(u^i)$ ✓

Paso 5 Construcción de una base $\mathcal{B}$ tq $\mathrm{Mat}_{\mathcal{B}}(u) = J_{\underline{m}}$ para cierta partición $\underline{m}$:

Consideremos las inclusiones:

$$(\dim_k M_p = m_p) \leq (\dim_k M_{p-1} = m_{p-1}) \leq \dots \leq (\dim_k M_1 = m_1)$$

$$M_p \xhookrightarrow{u} M_{p-1} \xhookrightarrow{u} M_{p-2} \xhookrightarrow{u} \dots \xhookrightarrow{u} M_1 = N_1 = \ker(u)$$

M_p	M_{p-1}	M_{p-2}	$M_1 = N_1 = \ker(u)$
$e_{p,1}$	$u(e_{p,1})$	$u^2(e_{p,1})$	$u^{p-1}(e_{p,1})$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
e_{p,m_p}	$u(e_{p,m_p})$	$u^2(e_{p,m_p})$	$u^{p-1}(e_{p,m_p})$
(base)	e_{p-1,m_p+1}	$u(e_{p-1,m_p+1})$	$u^{p-2}(e_{p-1,m_p+1})$
	$\vdots$	$\vdots$	$\vdots$
	$e_{p-1,m_{p-1}}$	$u(e_{p-1,m_{p-1}})$	$u^{p-2}(e_{p-1,m_{p-1}})$
	(base)	$e_{p-2,m_{p-1}+1}$	$u^{p-3}(e_{p-2,m_{p-1}+1})$
		$\vdots$	$\vdots$
		$e_{p-2,m_{p-2}}$	$u^{p-3}(e_{p-2,m_{p-2}})$
		(base)	$\vdots$
			$u(e_{2,m_2})$
			e_{1,m_2+1}
			$\vdots$
			e_{1,m_1}
			(base)

Los vectores que aparecen en este diagrama forman una base de V, que ordenamos de la manera siguiente:

Comenzamos en la esquina superior derecha con $u^{p-1}(e_{p,1})$ y leemos la primera fila de derecha a izquierda: $\mathcal{B}_1 := (u^{p-1}(e_{p,1}), \dots, u(e_{p,1}), e_{p,1})$

Notar que $u(\mathcal{B}_1) \subseteq \mathcal{B}_1$ y, gracias al Paso 1 (Caso Particular), obtenemos un bloque de Jordan nilpotente J_p.

La segunda línea de derecha a izquierda: $\mathcal{B}_2 := (u^{p-1}(e_{p,2}), \dots, u(e_{p,2}), e_{p,2}) \rightsquigarrow J_p$

Continuamos así hasta la última línea: $\mathcal{B}_{m_1} = (e_{1,m_1})$ con $u(e_{1,m_1}) = 0$ y luego obtenemos el bloque $J_1 = (0)$.

Así, si consideramos $\mathcal{B} = (\mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_{m_1})$ base de V, obtenemos:

$$\mathrm{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} J_p & & & & \\ & J_{p-1} & & & \\ & & J_{p-1} & & \\ & & & \ddots & \\ & & & & J_1 \end{pmatrix}$$

En otras palabras, la matriz de u en la base $\mathcal{B}$ es diagonal por bloques de tipo J_m, con m_p bloques de tipo J_p, $m_{p-1} - m_p$ bloques de tipo J_{p-1}, y así sucesivamente hasta tener $m_1 = \dim_k(\ker(u))$ bloques de tipo $J_1 = (0)$.

Luego, si $\underline{m} = (\underbrace{p, \dots, p}_{m_p \text{ veces}}, \underbrace{p-1, \dots, p-1}_{m_{p-1}-m_p \text{ veces}}, \dots, \underbrace{1, \dots, 1}_{m_1 \text{ veces}}) \Rightarrow \mathrm{Mat}_{\mathcal{B}}(u) = J_{\underline{m}}$ matriz de Jordan.

Paso 6 Unicidad de $\underline{m}$ (partición) tq $\exists \mathcal{B}$ base con $\mathrm{Mat}_{\mathcal{B}}(u) = J_{\underline{m}}$:
Si existe otra base $\mathcal{C}$ de V tq $\mathrm{Mat}_{\mathcal{C}}(u) = J_{\underline{m}'}$, entonces reagrupando bloques de Jordan del mismo tamaño podemos construir M_j' tq $N_i = \bigoplus_{j=1}^{i} M_j'$. En part, tendríamos como antes que $m_j' = \dim_k(M_j')$ están determinados por $u \Rightarrow m_j = m_j'$ para todo $j \Rightarrow \underline{m}' = \underline{m}$. ✓ ■

Ejercicio Calcular el número de clases de semejanza de matrices nilpotentes en $M_6(k)$. Dar un representante de cada clase.

Teorema (Jordan, 1870): Sea V un k-e.v. de $\dim_k(V) = n$ y sea $u: V \to V$ un endomorfismo tal que su polinomio característico escinde sobre k, ie,

$$P_u(X) = \prod_{j=1}^{s} (X - \lambda_j)^{n_j} \qquad (n_j \geq 1 \text{ y } \lambda_i \neq \lambda_j \text{ si } i \neq j).$$

Entonces, $\exists \mathcal{B}$ base de V tal que $J := \mathrm{Mat}_{\mathcal{B}}(u)$ es una matriz diagonal por bloques de Jordan:

$$J = \begin{pmatrix} J_{\underline{n}_1}(\lambda_1) & & \\ & \ddots & \\ & & J_{\underline{n}_s}(\lambda_s) \end{pmatrix}$$

"forma canónica de Jordan de u"

donde $J_{\underline{n}_j}(\lambda_j) \in M_{n_j}(k)$ es la matriz de Jordan de tamaño n_j asociada a $\lambda_j \in k$ y a la partición $\underline{n}_j$ del entero n_j. Más aún, J está determinada de manera única por u "módulo permutación", ie, si $\exists \mathcal{B}'$ base de V tal que $J' = \mathrm{Mat}_{\mathcal{B}'}(u)$ es una matriz diagonal por bloques de Jordan $J_{\underline{n}_1'}(\mu_1), \dots, J_{\underline{n}_s'}(\mu_s)$ con $\mu_i \neq \mu_j$ si $i \neq j$, entonces $s = p$ y $\exists \sigma \in S_p$ permutación tal que $\mu_i = \lambda_{\sigma(i)}$ y $\underline{n}_i' = \underline{n}_{\sigma(i)}$.

Dem: Para cada $j \in \{1,\dots,p\}$ denotamos $u_j := u|_{V_{(\lambda_j)}} : V_{(\lambda_j)} \to V_{(\lambda_j)}$. Así, dado que $V_{(\lambda_j)} = \ker((u-\lambda_j \mathrm{Id}_V)^{m_j})$, tenemos que $u_j - \lambda_j \mathrm{Id}_{V_{(\lambda_j)}}$ es un endomorfismo nilpotente de $V_{(\lambda_j)} \Rightarrow \exists\, \mathcal{B}_j$ base de $V_{(\lambda_j)}$ tq $\mathrm{Mat}_{\mathcal{B}_j}(u_j - \lambda_j \mathrm{Id}_{V_{(\lambda_j)}}) = J_{\underline{m_j}}$ para una única partición $\underline{m_j}$ de $m_j = \dim_k V_{(\lambda_j)} = \mathrm{mult}_{alg}(\lambda_j)$.

$\Rightarrow \mathrm{Mat}_{\mathcal{B}_j}(u_j) = J_{\underline{m_j}} + \lambda_j I_{m_j} = J_{\underline{m_j}}(\lambda_j)$.

Dado que $V = \bigoplus_{j=1}^{p} V_{(\lambda_j)}$, si consideramos la base $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_p)$ tenemos

$$J := \mathrm{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} J_{\underline{m_1}}(\lambda_1) & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & J_{\underline{m_p}}(\lambda_p) \end{pmatrix}$$ matriz diagonal por bloques de Jordan ✓ ("forma canónica de Jordan").

Veamos la unicidad módulo permutación: Sup que $\exists\, \mathcal{B}'$ base de V tq

$$\mathrm{Mat}_{\mathcal{B}'}(u) = \begin{pmatrix} J_{\underline{m_1'}}(\mu_1) & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & J_{\underline{m_s'}}(\mu_s) \end{pmatrix} \quad \text{con } \mu_i \neq \mu_j \text{ si } i \neq j$$

Dado que esta matriz es triangular superior, obtenemos $P_u(X) = \prod_{j=1}^{s} (X-\mu_j)^{m_j'}$

$\Rightarrow p = s$ y $\exists\, \sigma \in S_p$ permutación tq $\mu_j = \lambda_{\sigma(j)}$ y $m_j' = m_{\sigma(j)}$ para todo $j \in \{1,\dots,p\}$.

Resta probar que las particiones $\underline{m_j'}$ y $\underline{m_{\sigma(j)}}$ son iguales: reordenando los μ_j, podemos suponer que $\mu_j = \lambda_j$ (ie, $\sigma = \mathrm{Id}$ en S_p) y probar que las particiones $\underline{m_j}$ y $\underline{m_j'}$ de $m_j \in \mathbb{N}^{\geq 1}$ son iguales:

Sea $\mathcal{B}_j'$ el conjunto de vectores de $\mathcal{B}'$ que corresponden al bloque de Jordan $J_{\underline{m_j'}}(\lambda_j) \in M_{m_j}(k) \Rightarrow W_j := \mathrm{Vect}_k(\mathcal{B}_j')$ verifica $u(W_j) \subseteq W_j$ y $\dim_k(W_j) = m_j$. Además, si $u_{W_j} := u|_{W_j} : W_j \to W_j$ entonces tenemos que $\mathrm{Mat}_{\mathcal{B}_j'}(u_{W_j}) = J_{\underline{m_j'}}(\lambda_j)$ por construcción.

Como $J_{\underline{m_j'}}$ es triangular superior estricta de tamaño $m_j \times m_j$, tenemos

$$J_{\underline{m_j'}}^{m_j} = (J_{\underline{m_j'}}(\lambda_j) - \lambda_j I_{m_j})^{m_j} = 0 \Rightarrow (u_{W_j} - \lambda_j \mathrm{Id}_{W_j})^{m_j} = 0.$$

$$\Rightarrow W_j \subseteq V_{(\lambda_j)} \underset{\dim_k W_j = \dim V_{(\lambda_j)}}{\Longrightarrow} W_j = V_{(\lambda_j)}$$

Luego, $\mathcal{B}_j$ y $\mathcal{B}_j'$ son bases del mismo sub-e.v. $V_{(\lambda_j)}$ y verifican

$\mathrm{Mat}_{\mathcal{B}_j}(u_j - \lambda_j \mathrm{Id}_{V_{(\lambda_j)}}) = J_{\underline{m_j}}$ y $\mathrm{Mat}_{\mathcal{B}_j'}(u_j - \lambda_j \mathrm{Id}_{V_{(\lambda_j)}}) = J_{\underline{m_j'}}$.

La unicidad de la partición en el caso nilpotente implica que $\underline{m_j} = \underline{m_j'}$. ■

[Obs: En términos matriciales: Sea $A \in M_n(k)$ tq $P_A \in k[X]$ escinde sobre k.
$\Rightarrow \exists P \in Gl_n(k)$ tq

$$P^{-1}AP = \begin{pmatrix} J_{\underline{m_1}}(\lambda_1) & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & J_{\underline{m_p}}(\lambda_p) \end{pmatrix} =: J \quad \text{"forma canónica de Jordan"}$$

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

Del Teorema anterior y de la demostración de unicidad en el caso nilpotente, obtenemos:

Corolario: Sean $A, B \in M_n(k)$ tq P_A y P_B escinden sobre k. Son equivalentes:

① A y B son semejantes (ie, $\exists P \in GL_n(k)$ tq $P^{-1}AP = B$).

② A y B tienen el mismo 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 $\lambda \in k$ y todo $m \in \mathbb{N}^{\geq 1}$.

③ A y B tienen la misma forma canónica de Jordan (módulo permutación).

Ejercicio Sean $\lambda_1, \dots, \lambda_p \in \mathbb{C}$ complejos tq $\lambda_i \neq \lambda_j$ si $i \neq j$. Sea $P(X) = \prod_{j=1}^{p} (X-\lambda_j)^{n_j}$ con $n_j \geq 1$ y $n_1 + \dots + n_p = n$. El Corolario anterior implica que el número de clases de semejanza de matrices $A \in M_n(\mathbb{C})$ tq $P_A = P$ es igual a $p(n_1) \cdots p(n_p)$, donde $p(n_j)$ es el número de particiones de $n_j \in \mathbb{N}^{\geq 1}$.

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 $C_P(A) := \{A \in M_n(\mathbb{C}) \text{ tq } P_A(X) = P(X)\}$. ¿Bajo qué condición sobre P el conjunto $C_P(A)$ está formado por una sola clase de semejanza?

⚠ **Importante**: A partir de las demostraciones anteriores se deduce el siguiente cálculo algorítmico de la forma canónica 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:

1º Calculamos $d_m := \dim_k \ker[(A-\lambda_j I_n)^m]$ para $m = 1, 2, \dots$ hasta obtener $d_{r_j} = \dim_k \ker[(A-\lambda_j I_n)^{r_j}] = n_j = \dim_k V(\lambda_j)$ para cierto r_j.

(Obs ① Si conocemos $m_A(X) = \prod_{j=1}^{p} (X-\lambda_j)^{m_j}$ entonces $r_j = m_j$
② Teorema del rango: $d_m = n - \operatorname{rg}[(A-\lambda_j I_n)^m]$)

Output: Obtenemos una secuencia $0 := d_0 < d_1 < \dots < d_{r_j} = n_j$

2º La secuencia $d_0, d_1, \dots, d_{r_j}$ determina cuantos bloques de Jordan $J_m(\lambda_j)$ hay en la matriz de Jordan $J_{n_j}(\lambda_j)$: $d_1 = d_1 - d_0 = \dim_k \ker(A-\lambda_j I_n)$ es el número de bloques de Jordan (de tamaño ≥ 1) que aparecen en $J_{n_j}(\lambda_j)$.
(Obs: Si $d_1 = 1 \Rightarrow J_{n_j}(\lambda_j) = J_{n_j}(\lambda_j) = \begin{pmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix} \in M_{n_j}(k)$).

$d_2 - d_1$ es el número de bloques de Jordan de tamaño ≥ 2 que aparecen en $J_{n_j}(\lambda_j)$.
$d_3 - d_2$ es el número de bloques de Jordan de tamaño ≥ 3 que aparecen en $J_{n_j}(\lambda_j)$, etc.

Output: Matriz de Jordan $J_{n_j}(\lambda_j) \in M_{n_j}(k)$ asociada a λ_j.

Finalmente, $J = \left(\begin{array}{c|c|c|c} J_{m_1}(\lambda_1) & 0 & \cdots & 0 \\ \hline 0 & \ddots & & \vdots \\ \hline \vdots & & \ddots & 0 \\ \hline 0 & \cdots & 0 & J_{m_p}(\lambda_p) \end{array}\right) \in M_n(k)$.

(Obs: Veremos en ejemplos, cómo obtener $P \in GL_n(k)$ a partir de A y J.)

Ejemplo: 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 $P_A(X)$:

$$\begin{vmatrix} X & -1 & 1 & 0 \\ 3 & X+1 & -3 & -1 \\ 2 & -1 & X-1 & -1 \\ 0 & 0 & 0 & X-2 \end{vmatrix} = (X-2)\begin{vmatrix} X & -1 & 1 \\ 3 & X+1 & -3 \\ 2 & -1 & X-1 \end{vmatrix} \underbrace{=}_{C_3 \mapsto C_3+C_2} (X-2)\begin{vmatrix} X & -1 & 0 \\ 3 & X+1 & X-2 \\ 2 & -1 & X-2 \end{vmatrix} = (X-2)^2\begin{vmatrix} X & -1 & 0 \\ 3 & X+1 & 1 \\ 2 & -1 & 1 \end{vmatrix}$$

$$\underbrace{=}_{F_2 \mapsto F_2-F_3} (X-2)^2\begin{vmatrix} X & -1 & 0 \\ 1 & X+2 & 0 \\ 2 & -1 & 1 \end{vmatrix} = (X-2)^2\begin{vmatrix} X & -1 \\ 1 & X+2 \end{vmatrix} = (X-2)^2(X^2+2X+1) = (X-2)^2(X+1)^2.$$

Sea $\lambda_1 = -1$ y $\lambda_2 = 2$, y sea $n_1 = \text{mult}_{alg}(-1) = 2$, $n_2 = \text{mult}_{alg}(2) = 2$.
Queremos determinar las matrices de Jordan $J_{n_1}(-1)$, $J_{n_2}(2) \in M_2(\mathbb{R})$.

$\lambda_1 = -1$: Sea $B = A - \lambda_1 I_4 = A + I_4 = \begin{pmatrix} 1 & 1 & -1 & 0 \\ -3 & 0 & 3 & 1 \\ -2 & 1 & 2 & 1 \\ 0 & 0 & 0 & 3 \end{pmatrix}$ ¿$\ker(B)$, $\ker(B^2)$?

Para calcular una base de $\ker(B) = V_{\lambda_1}$ podemos hacer operaciones columna o bien:

$$B\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} x_1+x_2-x_3 \\ -3x_1+3x_3+x_4 \\ -2x_1+x_2+2x_3+x_4 \\ 3x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \overset{F_4}{\Rightarrow} x_4 = 0 \overset{F_2}{\Rightarrow} x_1 = x_3 \overset{F_1 \text{ o } F_3}{\Rightarrow} x_2 = 0$$

$\Rightarrow \ker(B) = \text{Vect}_{\mathbb{R}}(v_1)$ con $v_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$. (En part, $\text{mult}_{geom}(-1) = 1$ y A no es diagonalizable)

⚠ Dado que $\dim_{\mathbb{R}} \ker(B) = 1$, $J_{n_1}(-1)$ posee 1 bloque de Jordan

$\Rightarrow J_{n_1}(-1) = J_2(-1) = \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix}$. Más aún, $J_{n_1}(-1) = \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix}$ nos dice cómo escoger una "buena" base de $V_{(-1)}$: $\mathcal{B}_1 = (v_1, v_2)$ con $Av_1 = (-1)v_1$, ie, $(A+I_4)v_1 = 0$ y $Av_2 = v_1 + (-1)v_2 \iff (A+I_4)v_2 = v_1$.

$\Rightarrow$ Tomamos $v_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$ y v_2 tq $Bv_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$. eg; (resolviendo el sistema): $v_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$

(Otra forma: Calcular B^2 y una base de $\ker(B^2)$ dada por $\mathcal{B}_1 = (v_1, v_2)$ con $\ker(B) = \text{Vect}_{\mathbb{R}}(v_1)$ y $\ker(B^2) = \ker(B) \oplus \text{Vect}_{\mathbb{R}}(v_2)$ Ejercicio)

$\lambda_2 = 2$: Sea $C = A - 2I_4 = A - 2I_4 = \begin{pmatrix} -2 & 1 & -1 & 0 \\ -3 & -3 & 3 & 1 \\ -2 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ ¿$\ker(C)$, $\ker(C^2)$?

Verifiquemos mediante operaciones columna que $\ker(C) = \text{Vect}_{\mathbb{R}}(v_3)$ con $v_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}$.

Dado que $\dim_{\mathbb{R}} \ker(C) = 1$, $J_{n_2}(2) = J_2(2) = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$ y luego, al calcular una solución de $(A-2I_4)v_4 = v_3 \iff Cv_4 = v_3$ (eg. $v_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$), obtenemos una base $\mathcal{B}_2 = (v_3, v_4)$ de $V_{(2)}$.

Sea $\mathcal{B} = (v_1, v_2, v_3, v_4)$ y $P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \Rightarrow P^{-1}AP = J = \left(\begin{array}{cc:cc} -1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ \hdashline 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{array}\right)$

forma canónica de Jordan de A.

Ejercicio Importante Sea $u: V \to V$ endomorfismo tq P_u escinde sobre $\mathbb{R}$, y sea $u = u_s + u_n$ su descomposición de Dunford. Sea $\mathcal{B}$ una base de V tq

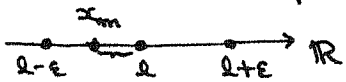
$$J = \mathrm{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} J_{m_1}(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & J_{m_p}(\lambda_p) \end{pmatrix}$$

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

$$S = \begin{pmatrix} \lambda_1 I_{m_1} & & 0 \\ & \ddots & \\ 0 & & \lambda_p I_{m_p} \end{pmatrix} \quad \text{y} \quad N = \begin{pmatrix} J_{m_1} & & 0 \\ & \ddots & \\ 0 & & J_{m_p} \end{pmatrix}.$$

§24. Espacios vectoriales normados

Motivación: Recordemos que una sucesión $(x_m)_{m\in\mathbb{N}}$ de números reales converge a un límite $l \in \mathbb{R}$ si: $\forall \varepsilon > 0$ existe $m_0 \in \mathbb{N}$ tq si $m \geq m_0$ entonces $|x_m - l| < \varepsilon$.



Esta última desigualdad dice que la "distancia" entre x_m y l es menor que ε. En esta sección extenderemos estas ideas a sucesiones $(\mathbb{x}_m)_{m\in\mathbb{N}}$ donde $\mathbb{x}_m \in V$ es un vector: necesitamos una noción de "distancia".

A modo de ejemplo consideremos $V = \mathbb{R}^2$: si $v_1 = (x_1, y_1)$ y $v_2 = (x_2, y_2)$ definimos la distancia euclideana entre v_1 y v_2 por $\|v_1 - v_2\|_2$, donde

$$\|(x,y)\|_2 := \sqrt{x^2 + y^2} \quad \text{para todo } (x,y) \in \mathbb{R}^2.$$

Es natural entonces decir que una sucesión $(v_m)_{m\in\mathbb{N}}$ de vectores en $\mathbb{R}^2$ converge a $l \in \mathbb{R}^2$ si $\forall \varepsilon > 0$, $\exists m_0 \in \mathbb{N}^*$ tq si $m \geq m_0$ entonces $\|v_m - l\|_2 < \varepsilon$.

Por otro lado, si escribimos coordenadas: $v_m = (x_m, y_m) \in \mathbb{R}^2$ y $l = (a,b) \in \mathbb{R}^2$, es natural también decir que $(v_m)_{m\in\mathbb{N}}$ converge a l si (x_m) converge a $a \in \mathbb{R}$ y si (y_m) converge a $b \in \mathbb{R}$. No es difícil notar que esto último ocurre si $\forall \varepsilon > 0$, $\exists m_0 \in \mathbb{N}$ tq si $m \geq m_0$ entonces $\|v_m - l\|_\infty < \varepsilon$, donde

$$\|(x,y)\|_\infty := \max(|x|, |y|) \quad \text{para todo } (x,y) \in \mathbb{R}^2.$$

¿Cómo comparar ambas nociones de convergencia?

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

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

Deducir que las nociones de convergencia anteriores coinciden, ie, si $l \in \mathbb{R}^2$ y $(v_m)_{m\in\mathbb{N}}$ sucesión en $\mathbb{R}^2$ entonces: "$v_m \xrightarrow{\|\cdot\|_2} l$" $\Leftrightarrow$ "$v_m \xrightarrow{\|\cdot\|_\infty} l$".

Conclusión: Necesitamos una noción de "distancia". Aunque no hay una única forma de medir distancia, la noción de convergencia es la misma si podemos comparar las distancias como en $(\star)$.

En esta sección $k = \mathbb{R}$ o $\mathbb{C}$ y $|\cdot|$ denota el valor absoluto o módulo en k, i.e., $|x| = \max(x, -x)$ si $k = \mathbb{R}$ y $|z| = \sqrt{z\bar{z}} = \sqrt{a^2+b^2}$ si $z = a+ib \in \mathbb{C}$.

Def: Sea V un k-e.v. Una norma de V es una función $\|\cdot\| : V \to \mathbb{R}^{\geq 0}$ tq para todos $x, y \in V$ y $\lambda \in k$ se cumple:

① $\|x\| = 0 \iff x = 0$

② $\|\lambda x\| = |\lambda| \cdot \|x\|$.

③ $\|x+y\| \leq \|x\| + \|y\|$ ("desigualdad triangular").

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

Ejemplo: Sea $V = \mathbb{R}^n$ o $\mathbb{C}^n$, definimos las normas:

$$\|(x_1, \ldots, x_n)\|_2 = \sqrt{|x_1|^2 + \ldots + |x_n|^2}$$

$$\|(x_1, \ldots, x_n)\|_1 = |x_1| + \ldots + |x_n|$$

$$\|(x_1, \ldots, x_n)\|_\infty = \max(|x_1|, \ldots, |x_n|).$$

Las propiedades ①, ② y ③ no son difíciles de probar para $\|\cdot\|_1$ y $\|\cdot\|_\infty$, pero la desigualdad triangular ③ para $\|\cdot\|_2$ es más sutil y la admitiremos por el momento (Más adelante: Es consecuencia de la "desigualdad de Cauchy-Schwarz").

Ejercicio Probar que $\|x\|_\infty \leq \|x\|_2 \leq \|x\|_1 \leq n\|x\|_\infty$ para todo $x \in V = \mathbb{R}^n$ o $\mathbb{C}^n$.

Def: Sea V un k-e.v. y sean $\|\cdot\|$ y $\|\cdot\|'$ dos normas en V. Decimos que las normas $\|\cdot\|$ y $\|\cdot\|'$ son equivalentes si existen constantes $a, b \in \mathbb{R}^{>0}$ tq

$$a\|x\| \leq \|x\|' \leq b\|x\|$$

para todo $x \in V$.

Hecho (sin prueba): Si $\dim_k(V)$ es finita, todas las normas de V son equivalentes ! (Análisis I).

En todo lo que sigue, supondremos que V es de dimensión finita. En particular, la definición siguiente no depende de la norma $\|\cdot\|$ escogida:

Def: Sea V un k-e.v. y sea $(v_m)_{m \in \mathbb{N}}$ una sucesión de elementos de V. Decimos que $(v_m)_{m \in \mathbb{N}}$ converge a un vector $l \in V$ si:

"Para todo $\varepsilon > 0$, existe $m_0 \in \mathbb{N}^*$ tq para todo $m \geq m_0$ se tiene $\|v_m - l\| < \varepsilon$."

Obs importante: Si $\mathcal{B}$ es una base de V, todo $v \in V$ se escribe de forma única $v = x_1 v_1 + \ldots + x_n v_n$, donde $(x_1, \ldots, x_n) \in k^n$ son las coordenadas de v respecto a $\mathcal{B} = (v_1, \ldots, v_n)$. En otras palabras:

$\varphi_{\mathcal{B}} : \mathbb{K}^n \to V$, $(x_1, \dots, x_n) \mapsto x_1 v_1 + \dots + x_n v_n$ es un isomorfismo. Luego, toda norma $\|\cdot\|$ en V define una norma en $\mathbb{K}^n$, y viceversa.

En part, si escogemos la norma $\|\cdot\|_\infty$ en $\mathbb{K}^n$ tenemos que una sucesión $(v_m)_{m\in\mathbb{N}} = ((x_{m,1}, \dots, x_{m,n}))_{m\in\mathbb{N}}$ de vectores en $\mathbb{K}^n$ converge al vector $l = (l_1, \dots, l_n) \in \mathbb{K}^n$ si y sólo si para todo $j \in \{1, \dots, n\}$ la sucesión real o compleja $(x_{m,j})_{m\in\mathbb{N}}$ "j-ésima componente" converge a $l_j \in \mathbb{K}$.

⚠ La observación anterior implica que la mayoría de los resultados sobre sucesiones reales o complejas son verdad para sucesiones vectoriales. Por ejemplo: una sucesión convergente es acotada, el límite de la suma es la suma de los límites (pero no el producto, pues no hemos definido el producto de vectores), etc.

Es importante notar que la definición de convergencia tiene el inconveniente de requerir conocer *a priori* el valor del límite $l \in V$. Es útil tener un criterio de convergencia que pueda verificarse sólo a partir de $(v_m)_{m\in\mathbb{N}}$:

Def: Sea $(V, \|\cdot\|)$ un e.v. normado. Decimos que una sucesión $(v_m)_{m\in\mathbb{N}}$ de elementos de V es una **sucesión de Cauchy** si:
"Para todo $\varepsilon > 0$, existe $m_0 \in \mathbb{N}$ tq para todos $p, q \geq m_0$ se tiene $\|v_p - v_q\| < \varepsilon$".

Ejemplo: Si $(v_m)_{m\in\mathbb{N}}$ converge a $l \in V$, entonces $(v_m)_{m\in\mathbb{N}}$ es de Cauchy. En efecto, dado $\varepsilon > 0$ tenemos que $\frac{\varepsilon}{2} > 0$ y luego, dado que $\lim_{m\to\infty} v_m = l$, $\exists m_0 \in \mathbb{N}$ tq si $m \geq m_0$ entonces $\|v_m - l\| < \frac{\varepsilon}{2}$. En part, si $p, q \geq m_0$ entonces $\|v_p - v_q\| = \|(v_p - l) + (l - v_q)\| \underbrace{\leq}_{\text{desig. triangular}} \|v_p - l\| + \|v_q - l\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$ ✓

Conclusión: Toda sucesión convergente es de Cauchy. Más aún (Análisis I):

Hecho (sin prueba): Si $\dim_{\mathbb{K}}(V)$ es finita, toda sucesión de Cauchy es convergente.

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

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

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 $\sum_{k=0}^{\infty} v_k$ converge si y sólo si $\forall \varepsilon > 0$, existe $m_0 \in \mathbb{N}$ tq para todos $p > q \geq m_0$ se tiene

$$\|S_p - S_q\| = \|v_{q+1} + v_{q+2} + \dots + v_p\| < \varepsilon.$$

Def: Sea V un k-e.v. ($\text{con } k = \mathbb{R} \text{ o } \mathbb{C}$). Decimos que una serie vectorial $\sum_{k=0}^{\infty} v_k$ de elementos de V converge absolutamente si la serie real $\sum_{k=0}^{\infty} \|v_k\|$ converge.

Prop: Toda serie absolutamente convergente es convergente.

Dem: Sea $\sum_{k=0}^{\infty} v_k$ absolutamente convergente y sea $\varepsilon > 0$. Usaremos el criterio de Cauchy: Dado que $\sum_{k=0}^{\infty} \|v_k\|$ converge, es de Cauchy

$\Rightarrow \exists m_0 \in \mathbb{N}$ tq para todos $p > q \geq m_0$ se tiene $\sum_{k=q+1}^{p} \|v_k\| < \varepsilon$. Por otro lado:

$$\left\| \sum_{k=q+1}^{p} v_k \right\| \underbrace{\leq}_{\text{desig. triangular}} \sum_{k=q+1}^{p} \|v_k\| < \varepsilon \quad \Rightarrow \left(\sum_{k=0}^{m} v_k \right)_{m \in \mathbb{N}} \text{ es de Cauchy} \Rightarrow \text{convergente} \quad \blacksquare$$

Obs: La desigualdad triangular implica que si $\sum v_k$ es absolutamente convergente:

$$\left\| \sum_{k=0}^{m} v_k \right\| \leq \sum_{k=0}^{m} \|v_k\| \leq \sum_{k=0}^{\infty} \|v_k\|$$

Tomando $m \to \infty$, obtenemos en este caso que $\left\| \sum_{k=0}^{\infty} v_k \right\| \leq \sum_{k=0}^{\infty} \|v_k\|$.

Caso particular importante: $V = M_n(k)$ ($\text{con } k = \mathbb{R} \text{ o } \mathbb{C}$).

Podemos considerar normas sobre $M_n(k) \cong k^{n^2}$, sucesiones de matrices $(A_m)_{m \in \mathbb{N}}$, series de matrices $\sum_{k=0}^{\infty} A_k$, etc.

Ejercicio Sea $(A_m)_{m \in \mathbb{N}^{\geq 1}}$ sucesión en $M_2(\mathbb{R})$ dada por $A_m = \begin{pmatrix} m \operatorname{sin}(1/m) & \frac{m}{m^2+1} \\ \frac{2m+3}{m} & 4 \end{pmatrix}$

Calcular (si existe) $\lim_{m \to \infty} A_m$.

⚠ Notar que el producto en $M_n(k)$ no es conmutativo. Luego, vale la pena redemostrar propiedades del límite del producto en este contexto:

Lema: Para todas $A, B \in M_n(k)$ se tiene $\|AB\|_1 \leq \|A\|_1 \|B\|_1$.

Dem: Si $A = (a_{ij})$ y $B = (b_{ij})$ entonces

$$\|AB\|_1 = \sum_{i,j} \left| \sum_k a_{ik} b_{kj} \right| \underbrace{\leq}_{\substack{\text{des.} \\ \text{triang.}}} \sum_{i,j,k} |a_{ik}| |b_{kj}| \leq \left(\sum_{i,k} |a_{ik}| \right) \left(\sum_{j,j} |b_{ij}| \right) = \|A\|_1 \|B\|_1 \quad \blacksquare$$

Importante: Dado que $M_n(k)$ es de dimensión finita, todas las normas son equivalentes. Luego, en todos los resultados sobre convergencia podemos considerar (sin pérdida de generalidad) la norma $\|\cdot\|_1$ (útil gracias al Lema anterior).

Prop: Sea $(A_m)_{m \in \mathbb{N}}$ una sucesión de matrices que converge a $A \in M_n(k)$ y sea $(B_m)_{m \in \mathbb{N}}$ una sucesión de matrices que converge a $B \in M_n(k)$. Entonces, la sucesión $(A_m B_m)_{m \in \mathbb{N}}$ converge a la matriz AB.

Dem: Sea $\varepsilon > 0$ y consideremos $\varepsilon_1 := \dfrac{\varepsilon}{2(\varepsilon + \|B\|_1)} > 0$ y $\varepsilon_2 := \dfrac{\varepsilon}{2(\|A\|_1 + 1)} > 0$.

Sea $m_0 \in \mathbb{N}$ tq para todo $m \geq m_0$ se tiene que

$$\|A_m - A\|_1 < \varepsilon_1 = \frac{\varepsilon}{2(\varepsilon + \|B\|_1)} \quad \text{y} \quad \|B_m - B\|_1 < \varepsilon_2 = \frac{\varepsilon}{2(\|A\|_1 + 1)} < \varepsilon .$$

En part, $\|B_m\|_1 = \|(B_m - B) + B\|_1 \leq \|B_m - B\|_1 + \|B\|_1 < \varepsilon + \|B\|_1$. Luego, si $m \geq m_0$:

$$\begin{aligned}
\|A_m B_m - AB\|_1 &= \|(A_m B_m - AB_m) + (AB_m - AB)\|_1 \\
&\underset{\text{des. tricang.}}{\leq} \|(A_m - A)B_m\|_1 + \|A(B_m - B)\|_1 \underset{\text{Lema}}{\leq} \|A_m - A\|_1 \|B_m\|_1 + \|A\|_1 \|B_m - B\|_1 \\
&< \|A_m - A\|_1 (\varepsilon + \|B\|_1) + \|A\|_1 \|B_m - B\|_1 \\
&< \frac{\varepsilon}{2(\varepsilon + \|B\|_1)} \cdot (\varepsilon + \|B\|_1) + \|A\|_1 \cdot \frac{\varepsilon}{2(\|A\|_1 + 1)} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon
\end{aligned}$$

■

Ejercicio Sea $(A_m)_{m \in \mathbb{N}}$ sucesión de matrices tal que $A_m \in Gl_m(k)$ para todo $m \in \mathbb{N}$. Sup. que $(A_m)_{m \in \mathbb{N}}$ converge a $A \in M_m(k)$ y $(A_m^{-1})_{m \in \mathbb{N}}$ converge a $B \in M_m(k)$. Probar que $A \in Gl_m(k)$ y $A^{-1} = B$.

El producto de series convergentes de matrices es más delicado:

Prop: Sean $\sum A_k$ y $\sum B_k$ series de matrices <u>absolutamente convergentes</u> con límites respectivos A y B. La serie $\sum C_k$ de término general

$$C_k := A_0 B_k + A_1 B_{k-1} + \ldots + A_k B_0$$

es absolutamente convergente con límite AB.

Dem: Notar que para $m \in \mathbb{N}$ se tiene:

$$\sum_{k=0}^{m} \|C_k\|_1 \leq \sum_{\substack{i,j \geq 0 \\ i+j \leq m}} \|A_i B_j\|_1 \leq \sum_{\substack{i,j \geq 0 \\ i+j \leq m}} \|A_i\|_1 \|V_j\|_1 \leq \underbrace{\left(\sum_{i=0}^{m} \|A_i\|_1\right)}\left(\underbrace{\sum_{j=0}^{m} \|B_j\|_1}\right)$$

$$\leq \sum_{i=0}^{\infty} \|A_i\|_1 \qquad \leq \sum_{j=0}^{\infty} \|B_j\|_1$$

$\Rightarrow$ Las sumas parciales de la serie real $\sum_{k=0}^{\infty} \|C_k\|_1$ de términos positivos son <u>acotadas</u> $\Rightarrow \sum_{k=0}^{\infty} \|C_k\|_1$ converge $\Rightarrow \sum_{k=0}^{\infty} C_k$ absolutamente convergente ✓

Finalmente, observamos que

$$\begin{aligned}
\left\| \sum_{k=0}^{2m} C_k - \left(\sum_{i=0}^{m} A_i\right)\left(\sum_{j=0}^{m} B_j\right) \right\|_1 &= \left\| \sum_{\substack{i>m \text{ o } j>m \\ i+j \leq 2m}} A_i B_j \right\|_1 \leq \sum_{\substack{i>m \text{ o } j>m \\ i+j \leq 2m}} \|A_i\|_1 \|B_j\|_1 \\
&\leq \left(\sum_{i=0}^{2m} \|A_i\|_1\right)\left(\sum_{j=0}^{2m} \|B_j\|_1\right) - \left(\sum_{i=0}^{m} \|A_i\|_1\right)\left(\sum_{j=0}^{m} \|B_j\|_1\right)
\end{aligned}$$

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, deducimos que $\sum_{k=0}^{\infty} C_k = AB$. ■

§25. Exponencial de matrices:

En esta sección $k = \mathbb{R}$ o $\mathbb{C}$.

Def: Sea $A \in M_n(k)$. Definimos la exponencial de A mediante la serie

$$\exp(A) = e^A := \sum_{k=0}^{\infty} \frac{A^k}{k!} \in M_n(k).$$

⚠ La exponencial e^A está bien definida pues para todo $k \in \mathbb{N}$ se tiene

$$\left\| \frac{A^k}{k!} \right\|_1 \leq \frac{\|A\|_1^k}{k!} \Rightarrow \sum_{k=0}^{\infty} \left\| \frac{A^k}{k!} \right\|_1 \leq \sum_{k=0}^{\infty} \frac{\|A\|_1^k}{k!} = e^{\|A\|_1} \Rightarrow \sum_{k=0}^{\infty} \frac{A^k}{k!} \text{ absolutamente convergente } \checkmark$$

Ejemplo: Si $A = \begin{pmatrix} A_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & A_p \end{pmatrix}$ matriz diagonal por bloques, entonces

$$A^k = \begin{pmatrix} A_1^k & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & A_p^k \end{pmatrix} \text{ para todo } k \in \mathbb{N} \Rightarrow e^A = \begin{pmatrix} e^{A_1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & e^{A_p} \end{pmatrix}$$

En part, si $A = \begin{pmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \ddots & \lambda_m \end{pmatrix}$ matriz diagonal $\Rightarrow e^A = \begin{pmatrix} e^{\lambda_1} & & 0 \\ & e^{\lambda_2} & \\ 0 & & \ddots & e^{\lambda_m} \end{pmatrix}$.

Ejemplo: Si $A \in M_n(k)$ es nilpotente de índice de nilpotencia $p \in \mathbb{N}^{\geq 1}$, ie, $A^p = 0$ pero $A^{p-1} \neq 0$, entonces $e^A = \sum_{k=0}^{p-1} \frac{A^k}{k!} = I_n + A + \frac{1}{2}A^2 + \ldots + \frac{1}{(p-1)!}A^{p-1}$.

Por ejemplo, $A = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ cumple $A^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ y $A^3 = 0$

$$\Rightarrow e^A = I_3 + A + \frac{A^2}{2} = \begin{pmatrix} 1 & 1 & 3/2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

Obs: 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:

$$\exp\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \text{ y } \exp\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \text{ no conmutan.}$$

Prop: Sean $A, B \in M_n(k)$ dos matrices que conmutan, ie, $AB = BA$. Entonces:

$$e^{A+B} = e^A e^B = e^B e^A.$$

Dem: Dado que A y B conmutan, el Teorema del Binomio de Newton afirma que

$$(A+B)^k = \sum_{j=0}^{k} \binom{k}{j} A^{k-j} B^j = \sum_{j=0}^{k} \frac{k!}{(k-j)!\, j!} A^{k-j} B^j$$

$$\Rightarrow e^{A+B} = \sum_{k=0}^{\infty} \frac{(A+B)^k}{k!} = \sum_{k=0}^{\infty} \underbrace{\left(\sum_{j=0}^{k} \frac{A^{k-j}}{(k-j)!} \cdot \frac{B^j}{j!} \right)}_{=: C_k} = \sum_{k=0}^{\infty} C_k$$

Visto: $\sum_{k=0}^{\infty} C_k$ converge al producto $\left(\sum_{i=0}^{\infty} \frac{A^i}{i!} \right)\left(\sum_{j=0}^{\infty} \frac{B^j}{j!} \right) = e^A \cdot e^B$ ∎

Corolario: Sea $A \in M_n(k)$. Entonces $e^A \in GL_n(k)$ y $(e^A)^{-1} = e^{-A}$.

Dem: Aplicar la proposición a A y $B := -A$ (que conmutan), y notar que $e^{A+(-A)} = e^A \cdot e^{-A} = e^{0_n} = I_n$ ■

Prop: Sean $A, B \in M_n(k)$ dos matrices semejantes, ie, $\exists P \in GL_n(k)$ tal que $B = P^{-1}AP$. Entonces e^A y e^B son semejantes. Más precisamente:

$$e^B = P^{-1} e^A P.$$

Dem: Notar que para todo $K \in \mathbb{N}$ se tiene

$$(P^{-1}AP)^K = (P^{-1}AP)(P^{-1}AP)\cdots(P^{-1}AP) = P^{-1}A^K P$$

$$\Rightarrow \sum_{K=0}^{m} \frac{(P^{-1}AP)^K}{K!} = P^{-1}\left(\sum_{K=0}^{m} \frac{A^K}{K!}\right)P \xrightarrow[m\to\infty]{} e^B = P^{-1}e^A P. \quad ■$$

Obs: La proposición anterior implica que para todo endomorfismo $u: V \to V$ de un k-e.v. V de dimensión finita (con $k = \mathbb{R}$ o $\mathbb{C}$) podemos definir $\exp(u) = e^u \in \mathrm{End}_k(V)$: es el endomorfismo asociado a $\exp(\mathrm{Mat}_{\mathcal{B}}(u))$, donde $\mathcal{B}$ es cualquier base de V.

Ejercicio Sea $A \in M_n(k)$.

① Probar que $\exp({}^tA) = {}^t\exp(A)$.

② Probar que $\det(e^A) = e^{\mathrm{tr}(A)}$.

Ejemplo: Sea $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \Rightarrow P_A(X) = X^2 - 1 = (X-1)(X+1)$.

Luego, A es diagonalizable con valores propios $\lambda_1 = -1$ y $\lambda_2 = 1$, y con vectores propios $v_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ y $v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, respectivamente.

$$\Rightarrow P = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \quad P^{-1} = \begin{pmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \end{pmatrix} \text{ y luego } A = P\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}P^{-1}$$

$$\Rightarrow e^A = P\begin{pmatrix} e^{-1} & 0 \\ 0 & e \end{pmatrix}P^{-1} = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} e^{-1} & 0 \\ 0 & e \end{pmatrix}\begin{pmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} e+e^{-1} & e-e^{-1} \\ e-e^{-1} & e+e^{-1} \end{pmatrix} \checkmark$$

⚠ **Importante**: Incluso si $A \in M_n(\mathbb{C})$ no es diagonalizable, el polinomio $P_A \in \mathbb{C}[X]$ escinde sobre $\mathbb{C}$ y luego A admite una única descomposición de Dunford $A = S + N$ con S diagonalizable, N nilpotente y donde $SN = NS$ (¡conmutan!). Luego, $e^A = e^S e^N = e^N e^S$ donde e^S se calcula como en el ejemplo anterior y donde $e^N = I_n + N + \frac{N^2}{2} + \ldots + \frac{N^{n-1}}{(n-1)!}$.

Caso particular: Si A posee sólo un valor propio λ, entonces su descomposición de Dunford es $A = \underbrace{\lambda I_n}_{S} + \underbrace{(A - \lambda I_n)}_{N}$, de donde obtenemos:

$$e^A = e^{\lambda I_m} \cdot e^{(A-\lambda I_m)} = e^\lambda \left(I_m + (A-\lambda I_m) + \frac{(A-\lambda I_m)^2}{2} + \ldots + \frac{(A-\lambda I_m)^{m-1}}{(m-1)!} \right).$$

Caso general: A es semejante a una matriz diagonal por bloques, donde cada bloque posee sólo un valor propio. Basta aplicar el cálculo anterior a cada bloque.

Ejemplo: Sea $A = \begin{pmatrix} -4 & 1 & 1 \\ 1 & -1 & -2 \\ -2 & 1 & -1 \end{pmatrix} \in M_3(\mathbb{R})$.

$$\Rightarrow P_A(X) = \begin{vmatrix} X+4 & -1 & -1 \\ -1 & X+1 & 2 \\ 2 & -1 & X+1 \end{vmatrix} \overset{\substack{C_1 \mapsto C_1+2C_2 \\ = \\ C_3 \mapsto C_3 - C_2}}{\begin{vmatrix} X+2 & -1 & 0 \\ 2X+1 & X+1 & -X+1 \\ 0 & -1 & X+2 \end{vmatrix}} \overset{\substack{F_1 \mapsto F_1 - F_3 \\ = \\ F_2 \mapsto F_2 + F_3}}{\begin{vmatrix} X+2 & 0 & -(X+2) \\ 2X+1 & X & 3 \\ 0 & -1 & X+2 \end{vmatrix}}$$

$$= (X+2) \begin{vmatrix} 1 & 0 & -1 \\ 2X+1 & X & 3 \\ 0 & -1 & X+2 \end{vmatrix} = (X+2) \underbrace{\left[1 \cdot (X^2+2X+3) + (-1) \cdot (-2X-1) \right]}_{(X^2+4X+4)} = (X+2)^3.$$

$\Rightarrow \lambda = -2$ es el único valor propio de A.

$\Rightarrow e^A = e^{-2I_3} e^{(A+2I_3)} = e^{-2} \left(I_3 + (A+2I_3) + \frac{1}{2}(A+2I_3)^2 \right)$.

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 Sea $A = \begin{pmatrix} 2 & -1 & 2 \\ 10 & -5 & 7 \\ 4 & -2 & 2 \end{pmatrix} \in M_3(\mathbb{R})$. Calcular e^A (cf. §22, pág. 53).

Lema útil: Sea $J_m = \begin{pmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix} \in M_m(k)$ bloque de Jordan nilpotente de tamaño $m \times m$ y sea $t \in k$. Entonces:

$$\exp(tJ_m) = \begin{pmatrix} 1 & t & t^2/2 & \cdots & t^{m-1}/(m-1)! \\ 0 & 1 & t & \ddots & \vdots \\ 0 & 0 & 1 & \ddots & t^2/2 \\ \vdots & \vdots & \ddots & \ddots & t \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

Dem: Sea $\mathcal{B} = (e_1, \ldots, e_m)$ la base canónica de k^m. Entonces, $Je_1 = 0$ y $Je_i = e_{i-1}$ para $i \in \{2, \ldots, m\}$. Luego, para todo $K \in \{1, \ldots, m-1\}$:

$$J_m^K e_i = \begin{cases} 0 & \text{si } i = 1, \ldots, K \\ e_{i-K} & \text{si } i = K+1, \ldots, d. \end{cases}$$

En otras palabras, $J_m^2 = \begin{pmatrix} 0 & 0 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & 1 \\ & & & \ddots & 0 \\ 0 & & & & 0 \end{pmatrix}$, $\ldots$, $J_m^{m-1} = \begin{pmatrix} 0 & \cdots & 0 & 1 \\ & \ddots & & 0 \\ & & \ddots & \vdots \\ 0 & & & 0 \end{pmatrix}$, $J_m^m = 0$.

$$\Rightarrow \exp(tJ_m) = I_m + (tJ_m) + \frac{1}{2}(tJ_m)^2 + \ldots + \frac{1}{(m-1)!}(tJ_m)^{m-1}$$

$$= \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix} + \begin{pmatrix} 0 & t & & \\ & \ddots & \ddots & \\ & & \ddots & t \\ & & & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & t^2/2 & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & t^2/2 \\ & & & \ddots & 0 \\ & & & & 0 \end{pmatrix} + \ldots + \begin{pmatrix} 0 & \cdots & 0 & t^{m-1}/(m-1)! \\ & \ddots & & 0 \\ & & \ddots & \vdots \\ 0 & & & 0 \end{pmatrix} \quad \blacksquare$$

Prop: Sean $m \in \mathbb{N}^{\geq 1}$ y $\lambda, t \in \mathbb{K}$. Si consideramos el bloque de Jordan de tamaño m asociado a λ, i.e., $J_m(\lambda) = \begin{pmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix} \in M_m(\mathbb{K})$. Entonces:

$$\exp(tJ_m(\lambda)) = e^{\lambda t}\left(I_m + tJ_m + \frac{t^2}{2}J_m^2 + \ldots + \frac{t^{m-1}}{(m-1)!}J_m^{m-1}\right) = e^{\lambda t}\begin{pmatrix} 1 & t & t^2/2 & \cdots & t^{m-1}/(m-1)! \\ 0 & 1 & t & \ddots & \vdots \\ 0 & 0 & 1 & \ddots & t^2/2 \\ \vdots & \vdots & \ddots & \ddots & t \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

En part, para toda matriz de Jordan

$$J = \left(\begin{array}{c|c|c} J_{m_1}(\lambda_1) & 0 & 0 \\ \hline 0 & \ddots & 0 \\ \hline 0 & 0 & J_{m_p}(\lambda_p) \end{array}\right) \quad \text{se tiene} \quad \exp(tJ) = \left(\begin{array}{c|c|c} e^{\lambda_1 t}\exp(tJ_{m_1}) & 0 & 0 \\ \hline 0 & \ddots & 0 \\ \hline 0 & 0 & e^{\lambda_p t}\exp(tJ_{m_p}) \end{array}\right).$$

Dem: La descomposición de Dunford de $J_m(\lambda)$ es $J_m(\lambda) = \lambda I_m + J_m$.
$\Rightarrow \exp(tJ_m(\lambda)) = \exp(\lambda t I_m)\cdot\exp(tJ_m) = e^{\lambda t}\exp(tJ_m)$, por lo que la primera parte sigue gracias al Lema útil. La última parte se obtiene al considerar cada bloque de Jordan por separado, y usar la primera parte. ■

Terminamos esta sección estudiando la continuidad y derivabilidad de la exponencial (esto será útil en aplicaciones).

Recuerdo: Sea $\Omega \subseteq \mathbb{R}$ subconjunto no-vacío y $f: \Omega \to \mathbb{R}$, $t \mapsto f(t)$ función real. Decimos que f es continua en $t_0 \in \Omega$ si
"Para todo $\varepsilon > 0$, existe $\delta > 0$ tq: si $t \in \Omega$ y $|t - t_0| < \delta$ entonces $|f(t) - f(t_0)| < \varepsilon$."

Si consideramos $\Omega \subseteq \mathbb{R}$ y $f: \Omega \to V$ función con valores en el espacio vectorial normado $(V, \|\cdot\|)$, entonces la definición anterior se adapta fácilmente:

Def: Sea $\Omega \subseteq \mathbb{R}$ subconj no-vacío y $(V, \|\cdot\|)$ e.v.n. Una función $f: \Omega \to V$ es continua en $t_0 \in \Omega$ si $\forall \varepsilon > 0$, $\exists \delta > 0$ tq si $t \in \Omega$ y $|t - t_0| < \delta$ entonces $\|f(t) - f(t_0)\| < \varepsilon$.

Obs: Como siempre, si $\dim_{\mathbb{K}}(V)$ es finita entonces todas las normas de V son equivalentes y luego la continuidad de $f: \Omega \to V$ es independiente de la norma de V que escojamos (Ejercicio).

⚠ Si $V = \mathbb{R}^n$ o $\mathbb{C}^n$, entonces $f: \Omega \to \mathbb{K}^n$ está determinada por sus n funciones componentes $f_1, \ldots, f_n: \Omega \to \mathbb{K}$ tales que

$$f(t) = (f_1(t), \ldots, f_n(t)) \quad , t \in \Omega.$$

Si escogemos la norma $\|\cdot\|_\infty$ notamos que f es continua en $t_0 \in \Omega$ si y sólo si cada componente $f_1, \ldots, f_n$ es continua en $t_0 \in \Omega$.

Ejemplos: ① $V = \mathbb{R}^3$: $f(t) = (t^2, t^3, \sin(t))$ es continua en todo $\mathbb{R}$.

② $V = M_2(\mathbb{R}) \cong \mathbb{R}^4$: $f(t) = \begin{pmatrix} t & t^2+1 \\ \sqrt{t} & e^t \end{pmatrix}$ es continua en $\Omega = [0, +\infty[$.

Obs: La noción de límite es completamente análoga: Sea $\Omega \subseteq \mathbb{R}$ no vacío, $f: \Omega \to V$ y $t_0 \in \Omega$. Decimos que $\lim_{t \to t_0} f(t) = l$, para cierto $l \in V$, si: "$\forall \varepsilon > 0, \exists \delta > 0$ tq si $t \in \Omega$ y $|t - t_0| < \delta$ entonces $\|f(t) - l\| < \varepsilon$".

Igual que antes, si $V = \mathbb{R}^n$ o $\mathbb{C}^n$ y escogemos la norma $\|\cdot\|_\infty$, entonces para $f: \Omega \to \mathbb{K}^n$, $t \mapsto (f_1(t), \dots, f_n(t))$ y $l = (l_1, \dots, l_n) \in \mathbb{K}^n$ se tiene:

$$\lim_{t \to t_0} f(t) = l \iff \lim_{t \to t_0} f_1(t) = l_1, \dots, \lim_{t \to t_0} f_n(t) = l_n.$$

Este último permite probar directamente, por ejemplo, que el límite de una suma es la suma de los límites.

Prop: Sea $\Omega \subseteq \mathbb{R}$ no-vacío, y sean $F: \Omega \to M_{m \times n}(\mathbb{C})$ y $G: \Omega \to M_{n \times p}(\mathbb{C})$ funciones con valores matriciales tales que para $t_0 \in \Omega$ los límites $\lim_{t \to t_0} F(t)$ y $\lim_{t \to t_0} G(t)$ existen. Entonces, la función $FG: \Omega \to M_{m \times p}(\mathbb{C})$, $t \mapsto F(t)G(t)$ tiene un límite en $t_0 \in \Omega$ y además $\lim_{t \to t_0} (F(t)G(t)) = \left(\lim_{t \to t_0} F(t)\right) \cdot \left(\lim_{t \to t_0} G(t)\right)$.

Dem: Análoga a la demostración en el caso de funciones a valores reales (cf. pág 68). ∎

Def: Sea $\Omega \subseteq \mathbb{R}$ no-vacío y $(V, \|\cdot\|)$ e.v.n. Decimos que f es <u>derivable</u> 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 es denotado $f'(t_0) \in V$: la derivada de f en t_0.

⚠ Si $V = \mathbb{R}^n$ o $\mathbb{C}^n$ y escogemos la norma $\|\cdot\|_\infty$, entonces se prueba que para $f: \Omega \to \mathbb{K}^n$, $t \mapsto (f_1(t), \dots, f_n(t))$ y $t_0 \in \Omega$ se tiene:

f derivable en $t_0 \in \Omega \iff f_1, \dots, f_n$ son derivables en t_0.

Más aún, en tal caso se tiene que $f'(t_0) = (f_1'(t_0), \dots, f_n'(t_0))$.

Ejemplo: ① $V = \mathbb{R}^3$: $f(t) = (t^2, t^3, \sin(t))$ es derivable en todo $\mathbb{R}$ y $f'(t_0) = (2t_0, 3t_0^2, \cos(t_0))$.

② $V = M_2(\mathbb{R}) \cong \mathbb{R}^4$: $f(t) = \begin{pmatrix} t & t^2+1 \\ \sqrt{t} & e^t \end{pmatrix}$ es derivable en $]0, +\infty[$ y $f'(t_0) = \begin{pmatrix} 1 & 2t_0 \\ 1/(2\sqrt{t_0}) & e^{t_0} \end{pmatrix}$.

Obs: Gracias a las propiedades de límites, se puede verificar directamente que la derivada de una suma es la suma de las derivadas.

Prop: Sea $\Omega \subseteq \mathbb{R}$ no-vacío, y sean $F: \Omega \to M_{m\times n}(\mathbb{C})$ y $G: \Omega \to M_{n\times p}(\mathbb{C})$ funciones con valores matriciales, definidas en un intervalo abierto en torno a $t_0 \in \Omega$ y derivables en t_0. Entonces, FG es derivable en $t_0 \in \Omega$ y

$$(FG)'(t_0) = F'(t_0)G(t_0) + F(t_0)G'(t_0).$$

Dem: Sea $\delta(t) = \dfrac{F(t)G(t) - F(t_0)G(t_0)}{t-t_0}$. Basta calcular $\lim_{t\to t_0} \delta(t)$:

Escribimos $\delta(t) = \dfrac{F(t)G(t) - F(t_0)G(t) + F(t_0)G(t) - F(t_0)G(t_0)}{t-t_0}$

$$= \left(\frac{F(t)-F(t_0)}{t-t_0}\right) G(t) + F(t_0)\left(\frac{G(t)-G(t_0)}{t-t_0}\right),$$

de donde obtenemos $\lim_{t\to t_0} \delta(t) = F'(t_0)G(t_0) + F(t_0)G'(t_0)$ pues es el límite de un producto de matrices (cf. pág 73). ∎

⚠ Muy importante: Dado que el producto de matrices no es conmutativo, es importante prestar atención al orden de las matrices en la fórmula anterior!

Para estudiar la función exponencial usaremos la siguiente desigualdad:

Lema: Sea $A \in M_n(k)$, entonces $\|e^A - I_n - A\|_1 \leq \|A\|_1^2 \, e^{\|A\|_1}$.

Dem: Tenemos que $e^A - I_n - A = \sum_{k=2}^{\infty} \frac{A^k}{k!}$.

$$\Rightarrow \|e^A - I_n - A\|_1 \leq \sum_{k=2}^{\infty} \frac{\|A^k\|_1}{k!} = \sum_{k=0}^{\infty} \frac{\|A^{k+2}\|_1}{(k+2)!} \leq \|A\|_1^2 \sum_{k=0}^{\infty} \frac{\|A\|_1^k}{(k+2)!} \leq \|A\|_1^2 \underbrace{\sum_{k=0}^{\infty} \frac{\|A\|_1^k}{k!}}_{e^{\|A\|_1}}$$ ∎

Teorema: Sea $A \in M_n(k)$. La función $f: \mathbb{R} \to M_n(k)$, $t \mapsto f(t) = e^{tA}$ es derivable en todo $\mathbb{R}$, y además $f'(t_0) = A e^{t_0 A} = e^{t_0 A} A$. En particular, $f(t) = e^{tA}$ es de clase C^∞ y su derivada m-ésima es $f^{(m)}(t_0) = A^m e^{t_0 A} = e^{t_0 A} A^m$.

Dem: Queremos probar que $\dfrac{e^{tA} - e^{t_0 A}}{t-t_0} - A e^{t_0 A} = \dfrac{e^{tA} - e^{t_0 A} - (t-t_0) A e^{t_0 A}}{t-t_0} =: \Delta(t)$

tiende a 0 cuando $t \to t_0$. Dado que $t_0 A$ y $(t-t_0)A$ conmutan, tenemos que $e^{tA} = e^{(t-t_0)A + t_0 A} = e^{(t-t_0)A} \cdot e^{t_0 A}$. Sea $h := t - t_0$, entonces:

$$e^{tA} - e^{t_0 A} - (t-t_0) A e^{t_0 A} = e^{(t-t_0)A} e^{t_0 A} - e^{t_0 A} - (t-t_0) A e^{t_0 A} = (e^{hA} - I_n - hA) e^{t_0 A}.$$

El lema anterior aplicado a hA implica que:

$$\|e^{tA} - e^{t_0 A} - (t-t_0) A e^{t_0 A}\|_1 \leq \|e^{hA} - I_n - hA\|_1 \, \|e^{t_0 A}\|_1 \leq h^2 \|A\|_1^2 \, e^{|h|\|A\|_1} \, \|e^{t_0 A}\|_1$$

$$\Rightarrow \|\Delta(t)\|_1 \leq h \|A\|_1^2 \, e^{|h|\|A\|_1} \, \|e^{t_0 A}\|_1 \longrightarrow 0 \quad \text{cuando } h \to 0, \text{ i.e., } f'(t_0) = A e^{t_0 A}.$$

Por inducción, obtenemos $f^{(m)}(t_0) = A^m e^{t_0 A}$ para todo $m \in \mathbb{N}$. Finalmente, notamos que $A^m \left(\sum_{k=0}^{N} \frac{(tA)^k}{k!}\right) = \left(\sum_{k=0}^{N} \frac{(tA)^k}{k!}\right) A^m \xrightarrow[N\to\infty]{} A^m e^{tA} = e^{tA} A^m$ ✓ ∎

§26. Sistemas y EDO lineales con coeficientes constantes

Motivación: Sea $a \in \mathbb{C}$. Resolver la Ecuación Diferencial Ordinaria (EDO) $x' = ax$ significa encontrar una función derivable $f: \mathbb{R} \to \mathbb{C}$ tq para todo $t \in \mathbb{R}$ se cumpla $f'(t) = a f(t)$.

Por ejemplo, $f(t) = C e^{at}$ donde $C \in \mathbb{C}$ constante, verifica $f'(t) = aC e^{at} = a f(t)$ ✓

Por otra parte, si $g(t)$ es una solución de $x' = ax$ y definimos la función $F(t) := g(t) e^{-at}$ entonces: $F'(t) = g'(t) e^{-at} - a g(t) e^{-at} = (g'(t) - a g(t)) e^{-at} = 0$ para todo $t \in \mathbb{R} \Rightarrow F(t) = C$ es constante, ie, $g(t) = C e^{at}$.

El objetivo de esta sección es generalizar el cálculo anterior:

Def: Un **sistema diferencial lineal** (homogéneo) con coeficientes constantes es un sistema de la forma

$$(S) \begin{cases} x_1'(t) = a_{11} x_1(t) + a_{12} x_2(t) + \dots + a_{1n} x_n(t) \\ x_2'(t) = a_{21} x_1(t) + a_{22} x_2(t) + \dots + a_{2n} x_n(t) \\ \vdots \\ x_n'(t) = a_{n1} x_1(t) + a_{n2} x_2(t) + \dots + a_{nn} x_n(t) \end{cases}$$

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 $A = (a_{ij}) \in M_n(\mathbb{C})$ y $X(t) = \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix}$.

Teorema: 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 $\mathbb{C}$-s.v. de dimensión n.

Dem: 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

$$X'(t) = A e^{tA} X_0 + e^{tA} \cdot 0 = A e^{tA} X_0 = AX'(t) \quad ✓$$

Por otra parte, si $X(t)$ es una solución de (S), definimos la función

$$\begin{aligned} F(t) := e^{-tA} X(t) \Rightarrow F'(t) &= -A e^{-tA} X(t) + e^{-tA} X'(t) \\ &= -e^{-tA} AX(t) + e^{-tA} X'(t) \\ &= e^{-tA} (AX(t) - X'(t)) = 0 \end{aligned}$$

Luego, $F(t) = C \in \mathbb{C}^n$ constante. En part, $C = F(0) = X(0)$.

$\Rightarrow e^{-tA} X(t) = X(0)$. Dado que $e^{-tA} \in Gl_n(\mathbb{C})$ y $(e^{-tA})^{-1} = e^{tA}$, obtenemos $X(t) = e^{tA} X(0)$ ✓

Sea $V = \{X: \mathbb{R} \to \mathbb{C}^n$ derivable tq $X' = AX\}$ conj. de soluciones de (S). Entonces, la aplicación $\varphi: \mathbb{C}^n \to V$ dada por $\varphi(X_0) := e^{tA} X_0$ es lineal e inyectiva, con inversa dada por $X \mapsto X(0) \in \mathbb{C}^n$. Luego, $V \cong \mathbb{C}^n$. ■

Corolario: 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 <u>única</u> solución de (S) que verifica $X(t_0) = X_0$.

Dem: La función $X(t) = e^{tA}(e^{-t_0 A} X_0)$ es solución de (S) gracias al Teorema anterior, y ella verifica $X(t_0) = e^{0_m} X_0 = I_m X_0 = X_0$ ✓
Por otro lado, si $Y(t)$ es una solución de (S) tq $Y(t_0) = X_0$, entonces la función $Y(t+t_0)$ es solución de (S) y vale X_0 en $t=0$. La demostración del Teorema anterior implica que $Y(t+t_0) = e^{tA} X_0$, ie, $Y(t) = e^{(t-t_0)A} X_0$ ✓ ■

Nos gustaría tener una descripción más explícita de las soluciones de (S). Veamos primero el caso más simple: A diagonalizable.

Teorema: Supongamos que existe una base $(v_1, \dots, v_n)$ de $\mathbb{C}^n$ formada por vectores propios de A asociados a valores propios $\lambda_1, \dots, \lambda_n$ (no necesariamente distintos), ie, A es diagonalizable. Entonces, las soluciones de (S) son de la forma

$$X(t) = c_1 e^{\lambda_1 t} v_1 + c_2 e^{\lambda_2 t} v_2 + \dots + c_n e^{\lambda_n t} v_n .$$

donde $c_1, \dots, c_n \in \mathbb{C}$ son constantes.

Dem: Toda "condición inicial" $X_0 \in \mathbb{C}^n$ se escribe como $X_0 = \sum_{j=1}^{n} c_j v_j$. Luego, si $P \in GL_n(\mathbb{C})$ es la matriz de cambio de base de la base canónica a la base $(v_1, \dots, v_n)$, entonces $X_0 = P \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$.
Por otro lado, $P^{-1} A P = D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ matriz diagonal.

$$\Rightarrow e^{tA} X_0 = P e^{tD} P^{-1} X_0 = P e^{tD} \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = P \begin{pmatrix} c_1 e^{\lambda_1 t} \\ \vdots \\ c_n e^{\lambda_n t} \end{pmatrix} = \sum_{j=1}^{n} c_j e^{t\lambda_j} v_j \quad \checkmark$$

(Otra forma: En la base $(v_1, \dots, v_n)$ el sistema (S) se reduce a $y_j' = \lambda_j y_j$; $j = 1, \dots, n$) ■

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 $v_1 = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$, $v_2 = \begin{pmatrix} 4 \\ -1+3i \\ -1-3i \end{pmatrix}$ 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

$$\begin{cases} x_1(t) = c_1 + 4c_2 e^{3it} + 4c_3 e^{-3it} \\ x_2(t) = 2c_1 + (-1+3i)c_2 e^{3it} + (-1-3i)c_3 e^{-3it} \\ x_3(t) = 2c_1 + (-1-3i)c_2 e^{3it} + (-1+3i)c_3 e^{-3it} \end{cases}$$

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 $X_1(t)$, $(X_2(t)+X_3(t))/2 = \mathrm{Re}(X_2)(t)$ y $(X_2(t)-X_3(t))/2i = \mathrm{Im}(X_2)(t)$ son reales y linealmente independientes. Luego, toda solución real es combinación lineal con coeficientes reales de

$$X_1(t) = \begin{pmatrix}1\\2\\2\end{pmatrix}, \quad \mathrm{Re}(X_2)(t) = \begin{pmatrix}4\\-1\\-1\end{pmatrix}\cos(3t), \quad \mathrm{Im}(X_2)(t) = \begin{pmatrix}4\\3\\-3\end{pmatrix}\sin(3t).$$

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 mínimal m_A, es decir,

$$m_A(X) = \prod_{j=1}^{p} (X-\lambda_j)^{m_j} \quad \text{y} \quad P_A(X) = \prod_{j=1}^{p} (X-\lambda_j)^{n_j}.$$

Más aún, $0 < m_j \leq n_j$ (Cayley-Hamilton).

Teorema: Sean $\lambda_1, \ldots, \lambda_p$ los valores propios (distintos) de A y sean $m_1, \ldots, m_p$ sus multiplicidades como raíces del polinomio mínimal m_A. Entonces, las soluciones de (S) son de la forma

$$X(t) = e^{\lambda_1 t} P_1(t) + \ldots + e^{\lambda_p t} P_p(t),$$

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.

Dem: Recordemos que $V_{(\lambda_j)} = \ker\left[(A-\lambda_j I_n)^{m_j}\right]$ y que $\mathbb{C}^n = V_{(\lambda_1)} \oplus \ldots \oplus V_{(\lambda_p)}$. Por otra parte, toda solución $X(t)$ de (S) es de la forma $X(t) = e^{tA} X_0$ para cierta $X_0 \in \mathbb{C}^n$. Escribamos $X_0 = v_1 + \ldots + v_p$ 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 todo vector $v_j \in V_{(\lambda_j)}$, deducimos que:

$$e^{tA} v_j = e^{\lambda_j t} \sum_{k=0}^{m_j-1} \frac{((A-\lambda_j I_n)t)^k}{k!} v_j$$

$$\Rightarrow X(t) = \sum_{j=1}^{p} e^{\lambda_j t} P_j(t), \quad \text{donde } P_j(t) = \sum_{k=0}^{m_j-1} \frac{t^k}{k!} (A-\lambda_j I_n)^k v_j. \quad \blacksquare$$

Obs: A diagonalizable $\Leftrightarrow m_j = 1$ 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. Teorema en pág. 76).

Ejemplo: Consideremos el sistema

$$\begin{cases} x_1' = 2x_1 - x_2 + 2x_3 \\ x_2' = 10x_1 - 5x_2 + 7x_3 \\ x_3' = 4x_1 - 2x_2 + 2x_3 \end{cases} \Leftrightarrow \begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix} = \underbrace{\begin{pmatrix} 2 & -1 & 2 \\ 10 & -5 & 7 \\ 4 & -2 & 2 \end{pmatrix}}_{=A} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

Vimos (Ejemplo en §22, pág 53) 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 $v_1 = \begin{pmatrix}1\\-1\\-2\end{pmatrix}$

y $V_{\lambda_2} = \ker(A) = \text{Vect}_{\mathbb{R}}(v_2)$ y $V_{(\lambda_2)} = \ker(A^2) = \text{Vect}_{\mathbb{R}}(v_2, v_3)$ con $v_2 = \begin{pmatrix}1\\2\\0\end{pmatrix}$ y $v_3 = \begin{pmatrix}0\\1\\1\end{pmatrix}$.

Método 1: Gracias a (la demostración del) Teorema anterior, tenemos que una base para el espacio de soluciones del sistema está dado por:

$$X_1(t) = \underbrace{e^{\lambda_1 t} v_1}_{= e^{tA} v_1} = \begin{pmatrix}1\\-1\\-2\end{pmatrix} e^{-t}, \quad X_2(t) = \underbrace{e^{\lambda_2 t} v_2}_{= e^{tA} v_2} = \begin{pmatrix}1\\2\\0\end{pmatrix} \quad \text{y} \quad X_3(t) = e^{tA} v_3 = \underbrace{e^{\lambda_2 t}}_{=1} (v_3 + t \underbrace{A v_3}_{v_2})$$

$$= \begin{pmatrix}t\\2t+1\\1\end{pmatrix}$$

Luego, la solución general del sistema es

$$\begin{cases} x_1(t) = c_1 e^{-t} + c_2 + c_3 t \\ x_2(t) = -c_1 e^{-t} + 2c_2 + c_3(2t+1) \\ x_3(t) = -2c_1 e^{-t} + c_3 \end{cases}$$

Método 2: Calcular explícitamente $X(t) = e^{tA} X_0$: Dado que $\mathcal{B} = (v_1, v_2, v_3)$ es una base, escribimos $X_0 = c_1 v_1 + c_2 v_2 + c_3 v_3 \Rightarrow X_0 = P \begin{pmatrix}c_1\\c_2\\c_3\end{pmatrix}$ con $P = \begin{pmatrix}1 & 1 & 0\\-1 & 2 & 1\\-2 & 0 & 1\end{pmatrix}$.

Más aún, $P^{-1} A P = \left(\begin{array}{c|cc} -1 & 0 & 0 \\ \hline 0 & 0 & 1 \\ 0 & 0 & 0 \end{array}\right) = J$ forma canónica de Jordan.

$\Rightarrow e^{tA} = P e^{tJ} P^{-1}$, donde $e^{tJ} = \left(\begin{array}{c|cc} e^{(-1)t} & & \\ \hline & e^{0t} & e^{0t} \cdot t \\ & 0 & e^{0t} \end{array}\right) = \left(\begin{array}{c|cc} e^{-t} & & \\ \hline & 1 & t \\ & 0 & 1 \end{array}\right)$ (Cf. Lema útil en §25, pág. 71)

Luego, $e^{tA} X_0 = P e^{tJ} P^{-1} X_0 = P e^{tJ} \begin{pmatrix}c_1\\c_2\\c_3\end{pmatrix} = P \begin{pmatrix}c_1 e^{-t}\\c_2 + c_3 t\\c_3\end{pmatrix}$

$$\Rightarrow \begin{cases} x_1(t) = c_1 e^{-t} + c_2 + c_3 t \\ x_2(t) = -c_1 e^{-t} + 2c_2 + 2c_3 t + c_3 = -c_1 e^{-t} + 2c_2 + c_3(2t+1). \\ x_3(t) = -2c_1 e^{-t} + c_3 \end{cases}$$

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 (Ejemplo en §25, pág 71) 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:

$$e^{tA} \underset{\text{Dunford}}{=} e^{-2I_3 t} e^{(A+2I_3)t} = e^{-2t} \left(I_3 + (A+2I_3)t + \frac{1}{2} t^2 (A+2I_3)^2\right)$$

Dado que $A + 2I_3 = \begin{pmatrix}-2 & 1 & 1\\1 & 1 & -2\\-2 & 1 & 1\end{pmatrix}$ y $(A+2I_3)^2 = \begin{pmatrix}3 & 0 & -3\\3 & 0 & -3\\3 & 0 & -3\end{pmatrix}$ obtenemos

$$\begin{cases} x_1(t) = e^{-2t}\left[\left(1 - 2t + \frac{3t^2}{2}\right)c_1 + t c_2 + \left(t - \frac{3t^2}{2}\right)c_3\right] \\ x_2(t) = e^{-2t}\left[\left(t + \frac{3t^2}{2}\right)c_1 + (1+t)c_2 - \left(2t + \frac{3t^2}{2}\right)c_3\right] \\ x_3(t) = e^{-2t}\left[\left(-2t + \frac{3t^2}{2}\right)c_1 + t c_2 + \left(1 + t - \frac{3t^2}{2}\right)c_3\right] \end{cases}$$

⚠ **Consejos prácticos**: Para resolver el sistema diferencial $X'(t) = AX(t)$:

(1º) Primero que todo, calculamos los valores y vectores propios de A. Si los vectores propios forman una base (ie, A es diagonalizable) entonces el Teorema en pág 76 nos da la solución.

(2º) Si A no es diagonalizable pero sólo posee un valor propio λ, entonces

$$e^{tA} \underset{\text{Dunford}}{=} e^{\lambda t}\, e^{t(A-\lambda I_n)} = e^{\lambda t} \sum_{k=0}^{n-1} (A-\lambda I_n)^k \frac{t^k}{k!}$$

(3º) Si A no es diagonalizable y posee varios valores propios $\lambda_1, \dots, \lambda_p$ con multiplicidades algebraicas $n_1, \dots, n_p$ 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).

(Obs: De hecho, el Teorema en pág 77 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 $(v_j^1, v_j^2, \dots)$ de $V_{(\lambda_j)} = \ker\left[(A-\lambda_j I_n)^{m_j}\right]$ de tal suerte que $v_j^l \in \ker\left[(A-\lambda_j I_n)^l\right]$ y luego calcular $e^{tA} v_j^l$ como en (2º).

(Obs: 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 $\ker\left[(A-\lambda_j I_n)^2\right]$ tq v_j y v_j' sean l.i. En este caso obtenemos soluciones

$$X_j(t) = e^{\lambda_j t} v_j \quad \text{y} \quad X_j'(t) = e^{\lambda_j t}\left[v_j' + t(A-\lambda_j I_n) v_j'\right].$$

(c) Podemos calcular la forma canónica de Jordan de $A = PJP^{-1}$ y obtener e^{tJ} usando la Proposición en pág. 72.

(Obs: Vale la pena señalar que no es necesario calcular la inversa P^{-1} para encontrar la solución general (cf. "Método 2" en pág. 78).)

(4º) Si la matriz $A \in M_n(\mathbb{R})$ tiene coeficientes reales sus valores propios pueden agruparse en valores propios reales $\lambda_1, \dots, \lambda_p \in \mathbb{R}$ y en pares conjugados de valores propios complejos $\mu_1, \overline{\mu_1}, \dots, \mu_q, \overline{\mu_q} \in \mathbb{C}$. 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 $X_1(t), \dots, X_p(t), Y_1(t), \overline{Y_1(t)}, \dots, Y_q(t), \overline{Y_q(t)}$, con $X_j(t)$ real. Así, una base de soluciones reales es:

$$X_1(t), \dots, X_p(t), (\mathrm{Re}\, Y_1)(t), (\mathrm{Im}\, Y_1)(t), \dots, (\mathrm{Re}\, Y_q)(t), (\mathrm{Im}\, Y_q)(t).$$

Ejercicio Resolver el sistema diferencial

$$\begin{cases} x' = x + y \\ y' = -4x - 3y \end{cases}$$

Def: Una EDO homogenea lineal de orden n con coeficientes constantes es

$$(E)\quad x^{(n)}(t) + a_{n-1}x^{(n-1)}(t) + \ldots + a_1 x'(t) + a_0 x(t) = 0\ ,$$

donde los $a_j \in \mathbb{C}$ son constantes y $x : \mathbb{R} \to \mathbb{C}$ es una función n-veces derivable.

Si definimos $x_0 := x$, $x_1 = x'$, $x_2 = x''$, $\ldots$, $x_{n-1} = x^{(n-1)}$ entonces $(E) \Leftrightarrow X'(t) = AX(t)$

donde $X(t) = \begin{pmatrix} x_0(t) \\ \vdots \\ x_{n-1}(t) \end{pmatrix}$ y $A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & & & & 1 \\ -a_0 & -a_1 & & \cdots & -a_{n-1} \end{pmatrix}$.

¡Podemos aplicar los resultados anteriores!

Luego, el conjunto de soluciones de (E) es un $\mathbb{C}$-e.v. de dimensión n.
Más aún, vimos (cf. §22, pág. 55) que $P_A(X) = m_A(X) = X^n + a_{n-1}X^{n-1} + \ldots + a_1 X + a_0$, por lo que P_A es el "polinomio característico de (E)" y "$P_A(\lambda) = 0$" es la ecuación característica de (E).

Teorema: Sean $\lambda_1, \ldots, \lambda_p$ las raíces (distintas) de la ecuación característica y sean $n_1, \ldots, n_p$ sus multiplicidades. Entonces, las soluciones de (E) son todas las funciones de la forma

$$x(t) = Q_1(t)\, e^{\lambda_1 t} + \ldots + Q_p(t)\, e^{\lambda_p t},$$

donde $Q_j \in \mathbb{C}[t]$ es un polinomio de grado $< n_j$.

Dem: Toda solución de $X' = AX$ es de la forma

$$X(t) = e^{\lambda_1 t} P_1(t) + \ldots + e^{\lambda_p t} P_p(t)$$

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 $x(t) = e^{\lambda_1 t} P_{11}(t) + e^{\lambda_2 t} P_{21}(t) + \ldots + e^{\lambda_p t} P_{p1}(t)$, 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 las $n = n_1 + \ldots + n_p$ funciones $t \mapsto t^k e^{\lambda_j t}$, para $0 \le k < n_j$ y $j \in \{1, \ldots, p\}$. Como el espacio de soluciones es de dimensión n, ambos espacios son iguales. ∎

Ejemplo: ① $x''' + 3x'' + 3x' + x = 0 \Rightarrow P(\lambda) = (\lambda+1)^3 = 0$ y luego las soluciones son todas las funciones $x(t) = (at^2 + bt + c)\, e^{-t}$, con $a, b, c \in \mathbb{C}$.

② $x^{(4)} + 2x'' + x = 0 \Rightarrow P(\lambda) = (\lambda^2+1)^2 = 0$ 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 part, 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 reales de la EDO $x^{(n)} = x$.
(Obs: Tratar por separado el caso n par y el caso n impar).

§27. Espacios vectoriales cocientes

Motivación: Sabemos que si $f: V \to W$ es una aplicación lineal entre dos k-e.v. entonces $\ker(f) \subseteq V$ es un sub-e.v. de V. ¿Será cierto que todo sub-e.v. de V es el núcleo de alguna aplicación lineal? ¿Unicidad?

Sea $K := \ker(f) \subseteq V$ y consideremos $w \in \mathrm{Im}(f) \subseteq W$, ie, $w \in W$ tq existe $v \in V$ con $f(v) = w$. Si $v' \in V$ también cumple $f(v') = w$ entonces tenemos $f(v) = f(v') \iff v - v' \in K$. Luego, podemos definir la relación de equivalencia

$$v \sim v' \overset{\text{def}}{\iff} v - v' \in K \quad (\iff f(v) = f(v')) \quad \text{en } V.$$

Así, cada elemento $w \in \mathrm{Im}(f) \subseteq W$ se identifica a una única clase de equivalencia

$$[v] = \{v' \in V \text{ tq } v \sim v'\} = \{v' \in V \text{ tq } v - v' \in K\} =: v \bmod K$$
$$= \{v + x \,;\, x \in K\} =: v + K,$$

donde $v \in V$ es cualquier vector tq $f(v) = w$.

Generalizaremos esta construcción a todo sub-e.v. $\mathcal{U} \subseteq V$ y construiremos de manera "canónica" un k-e.v. denotado $V/\mathcal{U}$, y llamado el cociente de V por $\mathcal{U}$, junto a una aplicación lineal $\pi: V \twoheadrightarrow V/\mathcal{U}$ tq $\ker(\pi) = \mathcal{U}$, llamada proyección canónica.

Intuitivamente, trabajar sobre $V/\mathcal{U}$ nos permite "ignorar" los elementos de $\mathcal{U}$ y por ende simplificar muchas situaciones (veremos que esto es útil al hacer demostraciones por inducción en la dimensión, por ejemplo). Esto último es análogo al hecho que es más fácil hacer cálculos "módulo n" que en $\mathbb{Z}$ (eg. módulo 7 se tiene $3^{2010} \equiv 9^{1005} \equiv 2^{1005} \equiv 8^{335} \equiv 1 \bmod 7$ sin tener que calcular 3^{2010} explícitamente!).

Def: Sea $(A,+)$ un grupo abeliano (ie, conmutativo) y $B \subseteq A$ sub-conj. Decimos que B es un subgrupo de A si:
a) $0 \in B$.
b) Para todos $b_1, b_2 \in B$ se tiene $b_1 + b_2 \in B$.
c) Para todo $b \in B$ se tiene $-b \in B$.

Ejemplos: ① El conjunto $n\mathbb{Z} = \{kn,\ k \in \mathbb{Z}\}$ de múltiplos de $n \in \mathbb{N}$ es un subgrupo de $(\mathbb{Z},+)$.
② Sea V un k-e.v. Entonces todo sub-e.v. $W \subseteq V$ es un subgrupo de $(V,+)$. (que además cumple con ser estable por multiplicación por escalares).

Sea A un grupo abeliano y $B \subseteq A$ un sub-grupo. Definimos en A la relación de equivalencia $\sim_B$ mediante: $x \sim_B y \overset{\text{def}}{\iff} x - y \in B$.

Para todo $a \in A$ denotamos por

$$[a] := \{b \in A \text{ tq } b - a \in B\} = \{a + x\}_{x \in B} =: a \pmod{B} =: a + B$$

su clase de equivalencia, y por A/B al conjunto cociente (ie, el conjunto cuyos elementos son todas las clases de equivalencia). Denotamos por

$$\pi : A \longrightarrow A/B \qquad (\text{"proyección canónica"})$$
$$a \longmapsto [a] = a + B$$

la función que a todo $a \in A$ le asocia su clase $[a]$.

Prop: Sea A grupo abeliano y $B \subseteq A$ sub-grupo. Entonces, el cociente A/B admite una única estructura de grupo abeliano tq la proyección canónica $\pi : A \to A/B$ es un morfismo de grupos, ie,

$$\pi(a+b) = \pi(a) + \pi(b) \quad \text{para todos } a, b \in A.$$

Dem: La condición sobre π implica que en A/B se tiene que

$$[a] + [b] = (a+B) + (b+B) = \pi(a) + \pi(b) = \pi(a+b) = a + b + B = [a+b].$$

Luego, nos gustaría definir $[a] + [b] := [a+b]$. Sin embargo, la clase de equivalencia no determina únicamente el elemento $a \in A$: en efecto, si $a \sim_B a'$ entonces $[a] = [a']$. Luego, tenemos que probar que la expresión $[a+b] =: [a] + [b]$ no depende en los "representantes" $a, b \in A$:

Sean $a', b' \in A$ tales que $a \sim_B a'$ y $b \sim_B b'$, ie, $\exists\, x, y \in B$ tales que $a' = a + x$ y $b' = b + y$. Como $(A, +)$ es un grupo abeliano, tenemos:

$$a' + b' = (a+x) + (b+y) = a + b + \underbrace{(x+y)}_{\in B} \overset{\text{def}}{\Longrightarrow} (a+b) \sim_B (a'+b').$$

Así, $[a+b] =: [a] + [b] = [a'] + [b'] := [a'+b']$ bien definida ✓

En part, la suma en A/B es asociativa y conmutativa, y admite un elemento neutro dado por $[0] = 0 + B = B$. Finalmente, para todo $a \in A$, la clase $[-a] = -a + B$ es el opuesto de $[a] = a + B$. ∎

Ejemplo: Si $A = (\mathbb{Z}, +)$ y $B = n\mathbb{Z}$ entonces $x \sim_B y \overset{\text{def}}{\iff} x \equiv y \pmod{n}$ y $A/B = \mathbb{Z}/n\mathbb{Z}$.

Obs: Dado que para todo $a \in A$ se tiene que $a \in [a]$, tenemos que la proyección canónica $\pi : A \twoheadrightarrow A/B$ es siempre sobreyectiva.

Ejercicio Determinar cuándo $\pi : A \to A/B$ es inyectiva.

Teorema: Sea V un k-e.v. y $\mathcal{U} \subseteq V$ un sub-e.v. Entonces:

① El grupo abeliano $V/\mathcal{U}$ puede ser dotado de una única estructura de k-e.v. tq la proyección $\pi : V \to V/\mathcal{U}$ es lineal: para todos $\lambda \in k$ y $v \in V$ se tiene $\pi(\lambda v) = \lambda \pi(v)$. Decimos que $V/\mathcal{U}$ es el espacio vectorial cociente de V por $\mathcal{U}$.

② Si V es de dimensión finita, $\dim_k(V/\mathcal{U}) = \dim_k(V) - \dim_k(\mathcal{U})$.

③ Si V es de dimensión finita y $W \subseteq V$ es un suplementario de $\mathcal{U}$ (ie, $V = \mathcal{U} \oplus W$) entonces $\pi|_W : W \xrightarrow{\sim} V/\mathcal{U}$ es un isomorfismo. En part, si $\mathcal{C} = (w_1, \dots, w_d)$ es una base de W, con $d = \dim_k(V) - \dim_k(\mathcal{U})$, entonces $\pi(\mathcal{C}) = (\pi(w_1), \dots, \pi(w_d))$ es una base de $V/\mathcal{U}$.

Dem: ① Ya sabemos que $\pi(v+v') = \pi(v) + \pi(v') \iff v + v' + \mathcal{U} = (v+\mathcal{U}) + (v'+\mathcal{U})$. Por lo que basta verificar que para todo $\lambda \in k$ y $v \in V$ la expresión

$$\lambda \pi(v) = \lambda (v + \mathcal{U}) := \pi(\lambda v) = \lambda v + \mathcal{U}$$

está bien definida: sea $v' \in V$ tq $v \sim_{\mathcal{U}} v'$ (ie, $v - v' \in \mathcal{U}$) y sea $x \in \mathcal{U}$ tq $v' = v + x \Rightarrow \lambda v' = \lambda v + \lambda x$. Dado que $\mathcal{U}$ es un sub-e.v. $\lambda x \in \mathcal{U}$ y luego $\lambda v \sim_{\mathcal{U}} \lambda v'$. Así, $\lambda \cdot [v] := [\lambda v] = [\lambda v'] =: \lambda [v']$ bien definida ✓

② La aplicación lineal $\pi : V \twoheadrightarrow V/\mathcal{U}$ es sobreyectiva, y además

$$\ker(\pi) = \pi^{-1}([0]) = \pi^{-1}(0 + \mathcal{U}) = \{ v \in V \text{ tq } v - 0 \in \mathcal{U} \} = \mathcal{U}.$$

Luego, si V es de dimensión finita, el teorema del rango implica:

$$\dim_k(V) = \dim_k \ker(\pi) + \operatorname{rg}(\pi) = \dim_k(\mathcal{U}) + \dim_k(V/\mathcal{U}) \text{ ✓}$$

③ Sea $\pi_W = \pi|_W : W \to V/\mathcal{U}$ la restricción de π a $W \subseteq V$.

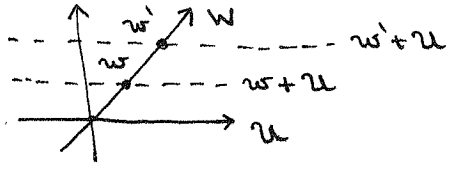
$\Rightarrow \ker(\pi_W) = \ker(\pi) \cap W = \mathcal{U} \cap W = \{0\}$, pues $V = \mathcal{U} \oplus W$.

$\Rightarrow \pi_W : W \hookrightarrow V/\mathcal{U}$ inyectiva. Por otra parte, todo $x \in V/\mathcal{U}$ es igual a $\pi(v)$ para cierto $v \in V$. Dado que $V = \mathcal{U} \oplus W$ existen únicos $u \in \mathcal{U}$ y $w \in W$ tq $v = u + w \Rightarrow \pi(v) = \underbrace{\pi(u)}_{=0} + \pi(w) = \pi(w)$, ie, $x = \pi_W(w)$

$\Rightarrow \pi_W : W \twoheadrightarrow V/\mathcal{U}$ sobreyectiva y luego un isomorfismo. En part, la imagen por π de cualquier base de W es una base de $V/\mathcal{U}$. ■

Ejercicio Sea $V = \mathbb{R}^2$ y $\mathcal{U} = \operatorname{Vect}_{\mathbb{R}}(e_1)$ eje x, describir el cociente $V/\mathcal{U}$.

Obs: El suplementario $W \subseteq V$ no es único (ie, el isomorfismo $W \cong V/\mathcal{U}$ no es "canónico").



El siguiente resultado, llamado la propiedad universal del cociente, permite caracterizar al cociente y a la proyección canónica:

Teorema (propiedad universal): Sea V un k-e.v., $\mathcal{U} \subseteq V$ un sub-e.v. y sea $f: V \to W$ una aplicación lineal tal que $\mathcal{U} \subseteq \ker(f)$. Entonces, existe una única aplicación lineal $\hat{f}: V/\mathcal{U} \to W$ tal que $f = \hat{f} \circ \pi$, donde $\pi: V \to V/\mathcal{U}$ es la proyección canónica. En otras palabras, el diagrama

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ {\scriptstyle\pi}\downarrow & \circlearrowleft & \nearrow_{\exists!\,\hat{f}} \\ V/\mathcal{U} & & \end{array}$$

es conmutativo.

Dem: Si existiera $\hat{f}: V/\mathcal{U} \to W$ tq $f = \hat{f} \circ \pi$ entonces la única forma de definir $\hat{f}$ es mediante $f(v) = \hat{f}(\pi(v)) = \hat{f}([v])$. Luego, basta ver que esto último está bien definido: Sea $v' \in V$ tq $\pi(v) = \pi(v')$, ie, $v - v' \in \mathcal{U}$. Sea $x \in \mathcal{U}$ tq $v' = v + x \Rightarrow f(v') = f(v) + f(x) = f(v)$ pues $x \in \mathcal{U} \subseteq \ker(f)$. Luego, $\hat{f}: V/\mathcal{U} \to W$, $[v] \mapsto f(v)$ bien definida ✓
Más aún, para todos $v, v' \in V$ y $\lambda \in k$ se tiene
$\hat{f}(\lambda[v] + [v']) = \hat{f}([\lambda v + v']) = f(\lambda v + v') = \lambda f(v) + v' = \lambda \hat{f}([v]) + \hat{f}([v'])$
$\Rightarrow \hat{f}$ es una aplicación lineal. ∎

Una consecuencia de lo anterior es el llamado "Teorema del Isomorfismo" de Emmy Noether (1882-1935):

Teorema (Noether, 1927): Sean V, W k-e.v. y $f: V \to W$ lineal. Entonces f induce un isomorfismo de espacios vectoriales

$$\hat{f}: V/\ker(f) \xrightarrow{\sim} \mathrm{Im}(f).$$

Dem: Podemos reemplazar $f: V \to W$ por la aplicación lineal sobreyectiva $f: V \to \mathrm{Im}(f)$. Luego, la propiedad universal del cociente (aplicada a $\mathcal{U} = \ker(f)$) implica que $\exists!\ \hat{f}: V/\ker(f) \to \mathrm{Im}(f)$ tal que $\hat{f}([v]) = f(v)$ para todo $v \in V$. En part, f sobreyectiva $\Rightarrow \hat{f}$ sobreyectiva ✓
Veamos que $\hat{f}$ es inyectiva: sea $x \in \ker(\hat{f})$ y sea $v \in V$ tq $\pi(v) = [v] = x$.
$\Rightarrow 0 = \hat{f}([v]) = f(v) \Rightarrow v \in \ker(f)$, ie, $x = 0$ en el cociente $V/\ker(f)$ (!) ✓ ∎

El siguiente resultado establece una conexión entre los espacios cocientes y las matrices triangulares por bloques:

Prop: Sea V un k-e.v de $\dim_k(V) = n$. Sea $u: V \to V$ endomorfismo y sea $\mathcal{U} \subseteq V$ un sub-e.v estable por u (ie, $u(\mathcal{U}) \subseteq \mathcal{U}$). Denotemos por $u_{\mathcal{U}}: \mathcal{U} \to \mathcal{U}$ la restricción $u|_{\mathcal{U}}$ y por $\pi: V \to V/\mathcal{U}$ la proyección canónica. Entonces:

① u induce un endomorfismo $\hat{u} := u_{V/\mathcal{U}}: V/\mathcal{U} \to V/\mathcal{U}$ tal que $u_{V/\mathcal{U}}(\pi(v)) = \pi(u(v))$ para todo $v \in V$.

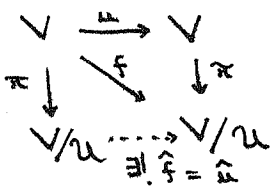
② Sea $\mathcal{B} = (e_1, \dots, e_n)$ base de V tq $\mathcal{C} = (e_1, \dots, e_r)$ es una base de $\mathcal{U}$. Entonces:

$$\mathrm{Mat}_{\mathcal{B}}(u) = \left(\begin{array}{c|c} A & B \\ \hline 0 & D \end{array}\right) \qquad \text{(triangular superior por bloques)}.$$

donde $A = \mathrm{Mat}_{\mathcal{C}}(u_{\mathcal{U}})$, $B \in M_{r,n-r}(k)$ y $D \in M_{n-r}(k)$. Más aún, D es la matriz de $u_{V/\mathcal{U}}$ respecto a la base $\mathcal{D} = (\pi(e_{r+1}), \dots, \pi(e_n))$ de $V/\mathcal{U}$.

③ Tenemos que $P_u(X) = P_{u_{\mathcal{U}}}(X)\, P_{u_{V/\mathcal{U}}}(X)$ en $k[X]$.

Dem: ① Como $u(\mathcal{U}) \subseteq \mathcal{U}$, tenemos que $(\pi \circ u)(\mathcal{U}) = \{0\}$ y luego la propiedad universal asegura la existencia de $\hat{u}$ (ie, consideramos $f = \pi \circ u: V \to V/\mathcal{U}$):



$(f = \pi \circ u$ cumple $\mathcal{U} \subseteq \ker(f))$

$\Rightarrow \hat{u}([v]) = f(v) \iff \hat{u}(\pi(v)) = \pi(u(v))$

② Sabemos que si $\mathcal{U} = \mathrm{Vect}_k(e_1, \dots, e_r)$ es estable por u, entonces $M = \mathrm{Mat}_{\mathcal{B}}(u)$ tiene la forma indicada. Más aún, si escribimos $B = (b_{ij})$ y $D = (d_{ij})$ donde $i = 1, \dots, r$ y $j, l = 1, \dots, n-r$, entonces:

$$u(e_{r+j}) = \sum_{i=1}^{r} b_{ij} e_i + \sum_{l=1}^{r} d_{lj} e_{r+l}$$

$$\Rightarrow \hat{u}(\pi(e_{r+j})) = \pi(u(e_{r+j})) = \sum_{l=1}^{r} d_{lj}\, \pi(e_{r+l}) \quad \text{para } j = 1, \dots, n-r,$$

ie, D es la matriz de $\hat{u}$ resp. a la base $\mathcal{D} = (\pi(e_{r+1}), \dots, \pi(e_n))$ de $V/\mathcal{U}$. ✓

③ Finalmente, $P_u(X) = P_M(X) = P_A(X) \cdot P_D(X) = P_{u_{\mathcal{U}}}(X) \cdot P_{u_{V/\mathcal{U}}}(X)$ ✓ ∎

Obs: Si consideramos V de dimensión finita y $\mathcal{U} \subseteq V$ sub-e.v. no-nulo $\Rightarrow \dim_k(V/\mathcal{U}) = \dim_k(V) - \dim_k(\mathcal{U}) < \dim_k(V)$. Luego, considerar cocientes es muchas veces útil para realizar demostraciones por inducción en la dimensión.

Ejercicio Sea $f: V \to W$ lineal y $\mathcal{U} \subseteq V$ sub-e.v. Sea $T := \mathrm{Im}(u|_{\mathcal{U}}) = u(\mathcal{U})$ sub-e.v de W. Probar que f induce una aplicación lineal

$$\hat{g}: V/\mathcal{U} \to W/T.$$

y describirla en términos de $f: V \to W$, $\pi_{\mathcal{U}}: V \to V/\mathcal{U}$ y $\pi_T: W \to W/T$.

(**Obs**: De hecho, basta que $T \subseteq W$ sub-e.v cumpla $u(\mathcal{U}) \subseteq T$).

§28. Formas lineales y espacio dual.

Recuerdo: Sea V un k-e.v. Una **forma lineal** en V es $f: V \to k$ lineal.

Ejemplo: Si $V = k^n$ entonces $pr_i : k^n \to k$, $(x_1, \dots, x_n) \mapsto x_i$ es una forma lineal.

Def: Sea V un k-e.v. Definimos el **espacio dual** (o simplemente **dual**) de V como
$V^* := \mathrm{Hom}_k(V, k) = \{ f : V \to k \text{ lineal} \}$.

Sup. que $\dim_k(V) = n$ y sea $\mathcal{B} = (e_1, \dots, e_n)$ una base de V. Para $j = 1, \dots, n$ definimos la forma $e_j^* = e^j \in V^*$ mediante

$$e_j^*(e_k) := \delta_{jk} = \begin{cases} 1 & \text{si } j = k \\ 0 & \text{si } j \neq k \end{cases}$$

Lema: La familia $\mathcal{B}^* = (e_1^*, \dots, e_n^*)$ es una base de V^*, llamada la **base dual** de la base $\mathcal{B}$. En part, $\dim_k(V) = \dim_k(V^*)$ y luego $V \cong V^*$.

Dem: Si $\sum_{j=1}^n \lambda_j e_j^* = 0$ en $V^* \Rightarrow 0 = \left(\sum_{j=1}^n \lambda_j e_j^*\right)(e_k) = \sum_{j=1}^n \lambda_j \underbrace{e_j^*(e_k)}_{\delta_{jk}} = \lambda_j \Rightarrow \mathcal{B}^*$ l.i. ✓

Sea $f \in V^*$ entonces $f = \sum_{j=1}^n f(e_j) e_j^*$ pues tienen el mismo valor en cada e_k ✓ ■

⚠ La notación e^* sólo tiene sentido si e es parte de la base!

Def: Definimos el **bidual** de V por $V^{**} := (V^*)^* = \mathrm{Hom}_k(V^*, k)$, y definimos la **evaluación canónica** $\psi : V \to V^{**}$ por $\psi(x)(f) := f(x)$ para todos $x \in V$ y $f \in V^*$.

Teorema (reflexividad): La evaluación canónica $\psi : V \to V^{**}$ es lineal y biyectiva.

Dem: Veamos que ψ es lineal: sean $x, y \in V$ y $\lambda \in k$, entonces para $f \in V^*$:
$\psi(\lambda x + y)(f) = f(\lambda x + y) = \lambda f(x) + f(y) = \lambda \psi(x)(f) + \psi(y)(f) = (\lambda \psi(x) + \psi(y))(f)$ ✓

Dado que $\dim_k(V) = \dim_k(V^*) = \dim_k(V^{**})$, basta probar que ψ es inyectiva:
Sea $x \in V \setminus \{0\} \Rightarrow \exists \mathcal{B} = (e_1, \dots, e_n)$ base de V tq $e_1 = x$ (teo. de base incompleta)
Sea $\mathcal{B}^* = (e_1^*, \dots, e_n^*)$ base dual $\Rightarrow \psi(x)(e_1^*) = e_1^*(x) = 1 \Rightarrow \psi(x) \neq 0$ ✓ ■

Def: Sea V un k-e.v. y $\mathcal{U} \subseteq V$ un sub-conj no-vacío de V. Definimos el **ortogonal** (o el **anulador**) **de** $\mathcal{U}$ en V^* como el conjunto

$$\mathcal{U}^\circ := \{ f \in V^* \text{ tq } f|_{\mathcal{U}} = 0 \} = \{ f \in V^* \text{ tq } f(x) = 0 \ \forall x \in \mathcal{U} \}.$$

Obs: ① Notar que $\mathcal{U}^\circ \subseteq V^*$ es un **sub-ev**: si $f, g \in \mathcal{U}^\circ$ y $\lambda \in k$ entonces para todo $x \in \mathcal{U}$ se tiene $(\lambda f + g)(x) = \lambda f(x) + g(x) = 0 \Rightarrow \lambda f + g \in \mathcal{U}^\circ$.

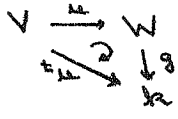
② **Importante**: Si $W \subseteq V^*$ sub-conj no-vacío $\Rightarrow W^\circ \subseteq V^{**}$.
Si $x \in V$ entonces la imagen $\psi(x) \in V^{**}$ por la evaluación canónica está en W°
$\Leftrightarrow \forall f \in W$ se tiene $0 = \psi(x)(f) = f(x)$. Luego, si **identificamos** V y V^{**} vía ψ:

$$W^\circ = \{ x \in V \text{ tq } : \forall f \in W,\ f(x) = 0 \} \subseteq V.$$

Teorema: Sea V un k-ev de dimensión finita y sea $\mathcal{U} \subseteq V$ sub-ev. Entonces:
$$\dim_k(\mathcal{U}) + \dim_k(\mathcal{U}^\circ) = \dim_k(V).$$

Dem: Si $\mathcal{U} = \{0\} \Rightarrow \mathcal{U}^\circ = V^*$ ✓ Si $\mathcal{U} \neq \{0\}$, consideramos $(e_1, \dots, e_r)$ base de $\mathcal{U}$, donde $r = \dim_k(\mathcal{U}) \geq 1$, y la completamos en una base $\mathcal{B} = (e_1, \dots, e_n)$ de V. Sea $\mathcal{B}^* = (e_1^*, \dots, e_n^*)$ la base dual de V^*. Notar que una forma lineal $f \in V^*$ pertenece a $\mathcal{U}^\circ \Leftrightarrow f(e_1) = \dots = f(e_r) = 0$.
Luego, si escribimos $f = \sum_{j=1}^{n} \lambda_j e_j^*$ con $\lambda_j = f(e_j)$, entonces esta última condición equivale a $\lambda_1 = \dots = \lambda_r = 0$, ie, $\mathcal{U}^\circ = \text{Vect}_k(e_{r+1}^*, \dots, e_n^*) \Rightarrow \dim_k(\mathcal{U}^\circ) = n - r$. ■

Def: Sea $u: V \to W$ lineal. La aplicación **transpuesta** (o **dual**) de u es la aplicación ${}^t u: W^* \to V^*$ dada por
$${}^t u(g) := g \circ u$$
para todo $g \in W^*$.



Obs: ① Notar que ${}^t u: W^* \to V^*$ es **lineal** Ejercicio.
② Si consideramos ${}^t({}^t u): V^{**} \to W^{**}$ y si identificamos $V^{**} \cong V$, $W^{**} \cong W$ usando las evaluaciones canónicas, entonces ${}^t({}^t u) = u$:
En efecto, $\forall x \in V$ y todo $g \in W^*$ se tiene que
$${}^t({}^t u)(\psi_V(x))(g) \underset{\text{def } {}^t({}^t u)}{=} \psi_V(x)({}^t u(g)) \underset{\text{def } {}^t u}{=} \psi_V(x)(g \circ u) \underset{\text{def } \psi_V}{=} (g \circ u)(x) = g(u(x)) = \psi_W(u(x))(g)$$
ie, ${}^t({}^t u)(\psi_V(x)) = \psi_W(u(x))$ ✓

Teorema: Sea $u: V \to W$ lineal entre k-ev. de dimensión finita. Entonces:
$$\ker({}^t u) = (\text{Im}(u))^\circ.$$
En part, u y ${}^t u$ son del mismo rango: $\text{rg}(u) = \text{rg}({}^t u)$.

Dem: Por def, $g \in W^*$ pertenece a $\ker({}^t u) \Leftrightarrow {}^t u(g) = 0$ en V^*
$\Leftrightarrow ({}^t u(g))(x) = g(u(x)) = 0 \quad \forall x \in V$
Esto último equivale a que g es ortogonal a todos los $u(x)$, ie, $g \in \text{Im}(u)^\circ$ ✓
Por otro lado, tenemos que: $\dim_k(W^*) = \dim_k \ker({}^t u) + \dim_k \text{Im}({}^t u)$ (teo. del rango)
y $\dim_k(W) = \dim_k \text{Im}(u) + \dim_k \text{Im}(u)^\circ$ (teo. anterior). Luego, dado que $\dim_k(W) = \dim_k(W^*)$ y $\dim_k \ker({}^t u) = \dim_k \text{Im}(u)^\circ$, tenemos $\dim_k \text{Im}(u) = \dim_k \text{Im}({}^t u)$ ■

Corolario: Sea $u: V \to W$ lineal entre k-ev. de dimensión finita. Entonces:
a) ${}^t u$ inyectiva $\Leftrightarrow$ u sobreyectiva.
b) ${}^t u$ sobreyectiva $\Leftrightarrow$ u inyectiva.

Dem: (a) es consecuencia del Teorema anterior, y (b) es consecuencia de (a) aplicado a ${}^t u$ (usando que ${}^t({}^t u) = u$). ■

Prop: Sean V y W k-ev, $\mathcal{B} = (e_1, \dots, e_n)$ base de V y $\mathcal{C} = (f_1, \dots, f_m)$ base de W. Sea $u : V \to W$ aplicación lineal. Entonces, la matriz de ${}^t u : W^* \to V^*$ respecto a las bases duales $\mathcal{C}^*$ y $\mathcal{B}^*$ es:

$$\mathrm{Mat}_{\mathcal{B}^*, \mathcal{C}^*}({}^t u) = {}^t \mathrm{Mat}_{\mathcal{C}, \mathcal{B}}(u) \qquad \text{(matriz traspuesta)}$$

Dem: Sea $A = \mathrm{Mat}_{\mathcal{C}, \mathcal{B}}(u) \in M_{m \times n}(k)$ definida por $u(e_j) = \sum_{i=1}^{m} a_{ij} f_i \rightsquigarrow A = (a_{ij})$

Luego, ${}^t u(f_i^*)(e_j) = f_i^*(u(e_j)) = \sum_{l=1}^{m} a_{lj} f_i^*(f_l) = a_{ij}$ y así para cada $j = 1, \dots, n$ se tiene ${}^t u(f_i^*) = \sum_{j=1}^{n} {}^t u(f_i^*)(e_j)\, e_j^* = \sum_{j=1}^{n} a_{ij} e_j^*$ en V^*.

Finalmente, la matriz $B = \mathrm{Mat}_{\mathcal{B}^*, \mathcal{C}^*}({}^t u) \in M_{n \times m}(k)$ está definida por $B = (b_{ij})$ tq ${}^t u(f_i^*) = \sum_{j=1}^{n} b_{ji} e_j^*$ (cf. §4) $\Rightarrow b_{ji} = a_{ij} \Rightarrow B = {}^t A$ ✓ ■

Corolario: Sea $A \in M_{m \times n}(k)$, entonces $\mathrm{rg}(A) = \mathrm{rg}({}^t A)$.

Dem: Consideran $u = u_A : k^n \to k^m$, $x \mapsto Ax$. Si $\mathcal{B}$ (resp. $\mathcal{C}$) es la base canónica de k^n (resp. k^m), entonces $\mathrm{Mat}_{\mathcal{C}, \mathcal{B}}(u_A) = A$ y $\mathrm{Mat}_{\mathcal{B}^*, \mathcal{C}^*}({}^t u_A) = {}^t A$. Más aún, sabemos que $\mathrm{rg}(u_A) = \mathrm{rg}({}^t u_A) \iff \mathrm{rg}(A) = \mathrm{rg}({}^t A)$ ■

Ejercicio Sea V un k-ev. de dimensión n, $\mathcal{B}$ una base de V, y $\mathcal{B}^*$ la base dual de V^*:

① Sea $\mathcal{C}$ otra base de V, $\mathcal{C}^*$ la base dual de V^* y sea $P = \mathrm{Mat}_{\mathcal{B}}(\mathcal{C})$ la matriz de cambio de base (cf. §5). Determinar la matriz $\mathrm{Mat}_{\mathcal{B}^*}(\mathcal{C}^*)$ en términos de P.

② Probar que para toda base $\mathcal{D}$ de V^*, existe una única base $\mathcal{B}$ de V tal que $\mathcal{B}^* = \mathcal{D}$. (Se dice que $\mathcal{B}$ es la base predual de $\mathcal{D}$).

§29. Formas bilineales y formas cuadráticas

⚠ En todo lo que sigue, supondremos que k es un cuerpo con $\mathrm{car}(k) \neq 2$ (eg. $k = \mathbb{Q}, \mathbb{R}$ o $\mathbb{C}$).

Def: Sea V un k-ev. Una forma bilineal sobre V es una forma 2-multilineal (ver §9), ie, es una aplicación $B : V \times V \to k$ que cumple:

(a) Para todo $y \in V$ fijo: La aplicación $B^y : V \to k$, $x \mapsto B(x, y)$ es una forma lineal.

(b) Para todo $x \in V$ fijo: La aplicación ${}^x B : V \to k$, $y \mapsto B(x, y)$ es una forma lineal.

Obs: Explícitamente, $B : V \times V \to k$ es bilineal si $\forall x, y, z \in V$ y $\forall \lambda \in k$ se tiene $B(\lambda x, y) = B(x, \lambda y) = \lambda B(x, y)$; $B(x+y, z) = B(x, z) + B(y, z)$; $B(x, y+z) = B(x, y) + B(x, z)$

Ejemplo: $B((x_1, \dots, x_n), (y_1, \dots, y_n)) = x_1 y_1 + \dots + x_n y_n$ es bilineal en k^n.

⚠ Si $W = V \times V$, entonces $B : W \to k$ no es necesariamente lineal (sólo es lineal en cada variable!).

Sea $\mathcal{B} = (e_1, \dots, e_n)$ una base de V, y escribamos $x = \sum_{j=1}^n x_j e_j$ e $y = \sum_{j=1}^n y_j e_j$ en V. Entonces,

$$B(x,y) = B\left(\sum_{j=1}^n x_j e_j, \sum_{j=1}^n y_j e_j\right) = \sum_{1 \le i,j \le n} B(e_i, e_j)\, x_i\, y_j$$

por lo que la forma bilineal B está determinada por la matriz $(B(e_i,e_j))_{i,j}$

Def: Sea $\mathcal{B} = (e_1, \dots, e_n)$ una base del k-ev V, y sea $B : V \times V \to k$ bilineal. Definimos la matriz de la forma bilineal B resp. a la base $\mathcal{B}$ mediante

$$\mathrm{Mat}_{\mathcal{B}}(B) := (B(e_i,e_j))_{i,j} \in M_n(k).$$

Si la forma bilineal es clara en el contexto, escribimos $A_{\mathcal{B}} := \mathrm{Mat}_{\mathcal{B}}(B)$ simplemente.

⚠ **Importante**: Si asociamos a $x, y \in V$ sus vectores coordenadas resp. a $\mathcal{B}$, ie, $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in k^n$ e $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in k^n$ entonces:

El escalar $B(x,y) = \sum_{1 \le i,j \le n} B(e_i,e_j) x_i y_j = \underbrace{(x_1\, x_2 \dots x_n)}_{\text{vector fila}} A_{\mathcal{B}} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = {}^t X A_{\mathcal{B}} Y$.

Luego, si $\mathcal{B}'$ es otra base de V y $P = \mathrm{Mat}_{\mathcal{B}}(\mathcal{B}')$ la matriz de cambio de base entonces los vectores coord. $X', Y' \in k^n$ de $x, y \in V$ resp. a $\mathcal{B}'$ verifican (cf. §5): $X = PX'$ e $Y = PY'$. Luego:

$$B(x,y) = {}^t X A_{\mathcal{B}} Y = {}^t(PX') A_{\mathcal{B}} (PY') = {}^t X' {}^t P A_{\mathcal{B}} P Y' = {}^t X' A_{\mathcal{B}'} Y'$$

$$\Rightarrow \boxed{A_{\mathcal{B}'} = {}^t P A_{\mathcal{B}} P} \quad (\text{"fórmula de cambio de base"}).$$

¡**Atención**!: $\det(A_{\mathcal{B}'}) = \det(P)^2 \det(A_{\mathcal{B}})$, por lo que $\det(A_{\mathcal{B}})$ depende de la base!

Def: Sea V un k-ev de dimensión finita y $B : V \times V \to k$ forma bilineal. Definimos las aplicaciones lineales:

$$\hat{B} : V \to V^* \qquad \text{y} \qquad \check{B} : V \to V^*$$
$$y \mapsto B^y = B(\cdot, y) \qquad\qquad x \mapsto {}^x B = B(x, \cdot)$$

Lema: Si identificamos V y V^{**} mediante la evaluación canónica $\gamma : V \to V^{**}$, entonces la aplicación transpuesta ${}^t\check{B} : V^{**} \to V^*$ cumple ${}^t\check{B} = \hat{B}$. En part, $\mathrm{rg}(\hat{B}) = \mathrm{rg}(\check{B})$.

Dem: Sean $x, y \in V$, entonces:

$${}^t\check{B}(\gamma(y))(x) \underset{\text{def de transp.}}{=} \gamma(y)(\check{B}(x)) \underset{\text{def de } \gamma}{=} \check{B}(x)(y) \underset{\text{def de } \check{B}}{=} B(x,y) \underset{\text{def de } \hat{B}}{=} \hat{B}(y)(x) \quad \Rightarrow {}^t\check{B} \circ \gamma = \hat{B} \quad \blacksquare$$

Def: Sea $B : V \times V \to k$ bilineal. Definimos el rango de B, $\mathrm{rg}(B)$, como el rango $\mathrm{rg}(\hat{B}) = \mathrm{rg}(\check{B})$. Decimos que B es no-degenerada si $\mathrm{rg}(B) = \dim_k(V)$, ie, si $\hat{B}$ y $\check{B}$ son biyectivas.

Terminología: Si B es no-degenerada, también se dice que $B : V \times V \to k$ es un emparejamiento perfecto: en tal caso B induce un isomorfismo $V \cong V^*$.

En términos matriciales:

Prop: Sea $B: V \times V \to k$ bilineal, $\mathcal{B}$ una base de V y $\mathcal{B}^*$ la base dual de V^*. Sea $A_{\mathcal{B}} \in M_n(k)$ la matriz de B resp. a $\mathcal{B}$, y sea $\hat{B}: V \to V^*$ como antes. Entonces:

$$Mat_{\mathcal{B}^*,\mathcal{B}}(\hat{B}) = A_{\mathcal{B}}.$$

Dem: Sea $\mathcal{B} = (e_1, \dots, e_n)$ y $\mathcal{B}^* = (e_1^*, \dots, e_n^*)$. Sea $M = Mat_{\mathcal{B}^*,\mathcal{B}}(\hat{B}) = (b_{ij}) \in M_n(k)$ y veamos que $b_{ij} = B(e_i, e_j)$:

$$B(e_i, e_j) = \hat{B}(e_j)(e_i) = \Big(\sum_{l=1}^{n} b_{lj}\, e_l^*\Big)(e_i) = \sum_{l=1}^{n} b_{lj}\, \underbrace{e_l^*(e_i)}_{\delta_{li}} = b_{ij} \quad \checkmark \quad \blacksquare$$

Obs importante: En part, $rg(B) = rg(A_{\mathcal{B}})$ para <u>cualquier</u> base $\mathcal{B}$ de V. Luego,

B <u>no-degenerada</u> $\iff$ $A_{\mathcal{B}} \in GL_n(k)$ para cualquier base $\mathcal{B}$

$\iff$ $\det(A_{\mathcal{B}}) \neq 0$ para cualquier base $\mathcal{B}$.

$\iff$ Para todo $y \in V \setminus \{0\}$ no-nulo, $\exists\, x \in V$ tq $B(x,y) \neq 0$.

En muchos casos, estudiaremos formas bilineales que además son "simétricas":

Def: Una forma bilineal $B: V \times V \to k$ es <u>simétrica</u> si para todos $x, y \in V$ se tiene:

$$B(x,y) = B(y,x).$$

Una matriz $A \in M_n(k)$ es <u>simétrica</u> si $A = {}^tA$.

Ejemplo: La forma $B: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $((x_1,\dots,x_n),(y_1,\dots,y_n)) \mapsto x_1y_1 + \dots + x_ny_n$ es <u>simétrica</u>.

Obs: Si $B: V \times V \to k$ es simétrica, entonces $\hat{B} = \check{B}: V \to V^*$ coinciden. En tal caso definimos el <u>kernel de B</u> como $\ker(B) := \ker(\hat{B}) = \ker(\check{B}) \leq V$, ie,

$$\ker(B) = \{x \in V \text{ tq } B(x,y) = 0 \text{ para } \underline{\text{todo}}\ y \in V\}.$$

Prop: Sea $B: V \times V \to k$ forma bilineal. Entonces:

B es una forma bilineal <u>simétrica</u> $\iff$ $A_{\mathcal{B}} = Mat_{\mathcal{B}}(B)$ es simétrica para <u>toda</u> base $\mathcal{B}$ de V.

Dem: ($\Rightarrow$) Si B es simétrica entonces $A_{\mathcal{B}} = (B(e_i,e_j))$ cumple ${}^tA_{\mathcal{B}} = (B(e_j,e_i))$ $= (B(e_i,e_j)) = A_{\mathcal{B}}$, ie, $A_{\mathcal{B}}$ es simétrica ✓

($\Leftarrow$) Sea $\mathcal{B} = (e_1,\dots,e_n)$ una base de V tq $A_{\mathcal{B}} = (B(e_i,e_j))$ es simétrica $\Rightarrow B(e_i,e_j) = B(e_j,e_i)$ para todos $i,j \in \{1,\dots,n\}$. Luego, para todo $x, y \in V$:

$$B(x,y) = \sum_{1 \leq i,j \leq n} B(e_i,e_j)\, x_i y_j = \sum_{1 \leq i,j \leq n} B(e_j,e_i)\, y_j x_i = B(y,x),$$ ie, B es simétrica ✓ $\blacksquare$

Def: Una <u>forma cuadrática</u> en un k-ev V es una aplicación $Q: V \to k$ tal que $\exists\, B: V \times V \to k$ <u>bilineal</u> con $Q(x) = B(x,x)$ para todo $x \in V$.

Obs: Si $\mathcal{B} = (e_1,\dots,e_n)$ base de V y $x = \sum_{j=1}^{n} x_j e_j$ entonces

$$Q(x) = \sum_{1 \leq i,j \leq n} B(e_i,e_j)\, x_i x_j$$

En part, las formas cuadráticas son lo mismo que las funciones (en varias variables) polinomiales homogéneas de grado 2 (ie, verifican $Q(\lambda x) = \lambda^2 Q(x)$ para todo $x \in V$ y $\lambda \in k$). ✱ Luego, las formas cuadráticas no son lineales!

Teorema: Sea $Q : V \to k$ una forma cuadrática. Entonces, existe una única forma bilineal simétrica $B : V \times V \to k$ tq $B(x,x) = Q(x)$. Más aún, B está dada por:

$$B(x,y) = \frac{1}{2}\left(Q(x+y) - Q(x) - Q(y)\right) \quad (\text{"fórmula de polarización"})$$

Diremos que B es la forma bilineal simétrica asociada a Q.

Dem: Sea $B : V \times V \to k$ una forma bilineal tq $Q(x) = B(x,x)$, entonces:

$$\begin{aligned} Q(x+y) = B(x+y, x+y) &= B(x,x) + B(x,y) + B(y,x) + B(y,y) \\ &= Q(x) + B(x,y) + B(y,x) + Q(y) \end{aligned}$$

Si B es simétrica entonces $B(x,y) = B(y,x)$, de donde obtenemos la fórmula ✓ ■

Ejemplos: ① Dada una forma cuadrática $Q : V \to k$, en general existen infinitas formas bilineales B tq $Q(x) = B(x,x)$. Por ejemplo: En $V = \mathbb{R}^2$ consideramos $Q(x_1, x_2) = 2x_1x_2$ forma cuadrática: su forma bilineal simétrica asociada es

$$B((x_1,x_2),(y_1,y_2)) = \frac{1}{2}\left(2(x_1+y_1)(x_2+y_2) - 2x_1x_2 - 2y_1y_2\right) = x_1y_2 + x_2y_1.$$

Sin embargo: Para todo $\lambda \in \mathbb{R}$ la forma bilineal $B_\lambda((x_1,x_2),(y_1,y_2)) = \lambda x_1y_2 + (2-\lambda)x_2y_1$ verifica $B_\lambda(x,x) = Q(x)$. Notar que, si $\mathcal{B} = (e_1, e_2)$ base canónica de $\mathbb{R}^2$, entonces:

$$\mathrm{Mat}_{\mathcal{B}}(B_\lambda) = \begin{pmatrix} B_\lambda(e_1,e_1) & B_\lambda(e_1,e_2) \\ B_\lambda(e_2,e_1) & B_\lambda(e_2,e_2) \end{pmatrix} = \begin{pmatrix} 0 & \lambda \\ 2-\lambda & 0 \end{pmatrix} \text{ simétrica} \iff \lambda = 1.$$

② Sea $Q : k^n \to k$ la forma cuadrática dada por:

$$Q(x_1, \dots, x_n) = \sum_{1 \le i \le j \le n} a_{ij} x_i x_j$$

Luego, la matriz (b_{ij}) en la base canónica de la forma bilineal simétrica asociada a Q está dada por:

$$b_{ii} = a_{ii}, \quad b_{ij} = \frac{1}{2}a_{ij} \text{ si } i < j, \quad b_{ij} = \frac{1}{2}a_{ji} \text{ si } i > j.$$

③ Sea $f \in V^* \setminus \{0\}$ forma lineal no-nula. Entonces $Q(x) := f(x)^2$ es una forma cuadrática en V, cuya forma bilineal simétrica es $B(x,y) = f(x)f(y)$. En part, $\widehat{B}(y) = f(y)f \in V^*$ y luego $\mathrm{Im}(\widehat{B}) = \mathrm{Vect}_k(f) \subseteq V^*$ es la "recta" generada por f, ie, $\mathrm{rg}(Q) := \mathrm{rg}(B) = 1$.

Def: Sea $Q : V \to k$ forma cuadrática, y $B : V \times V \to k$ su forma bilineal simétrica asociada. Definimos el rango de Q como $\mathrm{rg}(Q) := \mathrm{rg}(B)$.

Ejercicio Probar que toda forma cuadrática de rango 1 es proporcional a una forma cuadrática como en el Ejemplo ③.

§30. Ortogonalidad respecto a una forma cuadrática:

Def: 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 $\mathcal{U} \subseteq V$ sub-conj no-vacío, al conjunto

$$\mathcal{U}^{\perp} := \{ y \in V \text{ tq } : \forall x \in \mathcal{U},\ B(x,y) = 0 \}$$

es el ortogonal de $\mathcal{U}$.

③ Un vector $x \in V$ es isótropo si $Q(x) = B(x,x) = 0$ (ie, si es ortogonal a sí mismo).

Ejemplos: ① Sea $V = \mathbb{R}^2$ y $Q(x_1, x_2) = x_1^2 - x_2^2$ forma cuadrática. La 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 $L = \text{Vect}_{\mathbb{R}}(e_1)$ entonces $L^{\perp} = \{ (y_1,y_2) \in \mathbb{R}^2 \text{ tq } y_1 = 0 \} = \text{Vect}_{\mathbb{R}}(e_2)$.

(c) El vector $v = (1,1)$ es isótropo.

② Sea $B: V \times V \to k$ bilineal simétrica. Recordemos que

$$\ker(B) = \{ x \in V \text{ tq } B(x,y) = 0 \text{ para todo } y \in V \},$$

ie, $\ker(B) = V^{\perp}$.

Lema: Sea $B: V \times V \to k$ forma bilineal simétrica, $\widehat{B}: V \to V^*$, $y \mapsto B^y = B(\cdot, y)$ definida anteriormente, y sea $\mathcal{U} \subseteq V$ sub-conj no-vacío. Entonces,

$$\mathcal{U}^{\perp} = \widehat{B}^{-1}(\mathcal{U}^{\circ}).$$

En part, $\mathcal{U}^{\perp} \subseteq V$ es un sub-ev. Además, $\mathcal{U}^{\perp} \cong \mathcal{U}^{\circ}$ si B es no-degenerada.

Dem: Notar que

$$\mathcal{U}^{\perp} = \{ y \in V \text{ tq } \forall x \in \mathcal{U},\ B(x,y) = 0 \} = \{ y \in V \text{ tq } \forall x \in \mathcal{U},\ \widehat{B}(y)(x) = 0 \}$$
$$= \{ y \in V \text{ tq } \widehat{B}(y) \in \mathcal{U}^{\circ} \} = \widehat{B}^{-1}(\mathcal{U}^{\circ}) \ \checkmark \ \blacksquare$$

Teorema: Sea $B: V \times V \to k$ forma bilineal simétrica y $\mathcal{U} \subseteq V$ sub-ev. Entonces:

$$\dim_k(\mathcal{U}) + \dim_k(\mathcal{U}^{\perp}) \geq \dim_k(V) \quad \text{y} \quad \mathcal{U} \subseteq (\mathcal{U}^{\perp})^{\perp}.$$

Más aún, si B es no-degenerada entonces

$$\dim_k(\mathcal{U}) + \dim_k(\mathcal{U}^{\perp}) = \dim_k(V) \quad \text{y} \quad \mathcal{U} = (\mathcal{U}^{\perp})^{\perp}.$$

Dem: Sea $(e_1, \dots, e_r)$ una base de $\mathcal{U}$. El espacio vectorial $\mathcal{U}^{\perp} \subseteq V$ está dejinido por las r ecuaciones lineales $B(e_1,y) = \dots = B(e_r,y) = 0$. Luego, se tiene que $\dim_k(\mathcal{U}^{\perp}) \geq \dim_k(V) - \#r$ (podrían haber ecuaciones l.d.).

Por otro lado, sabemos que $\mathcal{U}^{\perp} = \widehat{B}^{-1}(\mathcal{U}^{\circ})$, por lo que si B es no-degenerada entonces $\dim_k(\mathcal{U}^{\perp}) = \dim_k(\mathcal{U}^{\circ}) = \dim_k(V) - \dim_k(\mathcal{U})$ (cf. §28) ✓

Si $x \in \mathcal{U}$ entonces $B(x,y) = 0 \ \forall y \in \mathcal{U}^{\perp}$ (por dy) $\Leftrightarrow B(y,x) = 0$ pues B simétrica. $\Rightarrow \mathcal{U} \subseteq (\mathcal{U}^{\perp})^{\perp}$. Por otro lado, si B es no degenerada entonces el cálculo anterior prueba que $\dim_k(\mathcal{U}) = \dim_k(\mathcal{U}^{\perp})^{\perp}$ y luego $\mathcal{U} = (\mathcal{U}^{\perp})^{\perp}$ ✓ $\blacksquare$

⚠ ¡Atención! Incluso si $B: V \times V \to k$ es no-degenerada, no necesariamente tenemos que $V = \mathcal{U} \oplus \mathcal{U}^\perp$! Por ejemplo, si $V = \mathbb{R}^2$ y $B((x_1,x_2),(y_1,y_2)) = x_1y_1 - x_2y_2$, entonces la recta $\mathcal{U} = \{(x_1,x_2) \in \mathbb{R}^2 \text{ tq } x_1 = x_2\}$ verifica $\mathcal{U} = \mathcal{U}^\perp$ (vectores isótropos). Luego, $\dim_{\mathbb{R}}(\mathcal{U}) + \dim_{\mathbb{R}}(\mathcal{U}^\perp) = 2 = \dim_{\mathbb{R}}(\mathbb{R}^2)$ pero $\mathcal{U} \cap \mathcal{U}^\perp \neq \{0\}$.

Def: Sea $B: V \times V \to k$ una forma bilineal simétrica. Decimos que una base $\mathcal{B} = (e_1, \dots, e_n)$ de V es ortogonal (resp. a B) si $B(e_i, e_j) = 0$ para $i \neq j$.

Obs: Matricialmente, $\mathcal{B}$ es ortogonal resp. a B si la matriz $A_{\mathcal{B}} = \mathrm{Mat}_{\mathcal{B}}(B)$ es diagonal: $A_{\mathcal{B}} = \begin{pmatrix} B(e_1,e_1) & & 0 \\ & \ddots & \\ 0 & & B(e_n,e_n) \end{pmatrix}$

Teorema: Toda forma bilineal simétrica posee una base ortogonal.

Dem: Sea $B: V \times V \to k$ bilineal simétrica. Por inducción en $n = \dim_k(V)$:
Si $n = 1$ toda base es ortogonal ✓ Sup. $n \geq 2$:
Si $B = 0$ toda base es ortogonal ✓ Sup. que $B \neq 0$, luego la fórmula de polarización

$$B(x,y) = \frac{1}{2}(Q(x+y) - Q(x) - Q(y))$$

implica que $\exists\, e_1 \in V \setminus \{0\}$ vector no-isótropo ($i.e.$, $Q(e_1) \neq 0$). Sea $L = \mathrm{Vect}_k(e_1)$ la recta generada por $e_1 \Rightarrow L \cap L^\perp = \{0\}$. Más aún, sabemos que: $\dim_k(L) + \dim_k(L^\perp) \geq n$ y luego necesariamente $L \oplus L^\perp = V$.

Por hipótesis de inducción, existe una base $(e_2, \dots, e_n)$ de $L^\perp$ que es ortogonal respecto a la restricción $B|_{L^\perp}: L^\perp \times L^\perp \to k$ de B a $L^\perp$. Dado que además $B(e_1, e_j) = 0$ para $j \in \{2, \dots, n\}$ (por def. de $L^\perp$) tenemos que $\mathcal{B} = (e_1, e_2, \dots, e_n)$ es una base de V que es ortogonal respecto a B. ■

Corolario: Toda forma cuadrática en un k-ev de dimensión n es combinación lineal de n cuadrados de formas lineales.

Dem: Sea $Q: V \to k$ forma cuadrática, y $B: V \times V \to k$ su forma bilineal simétrica asociada. Sea $\mathcal{B} = (e_1, \dots, e_n)$ una base de V ortogonal respecto a B. Entonces, dado que todo $x \in V$ se escribe como $x = \sum_{j=1}^n x_j e_j$, tenemos que:

$$Q(x) = B(x,x) = \sum_{1 \leq i,j \leq n} B(e_i,e_j) x_i x_j = \sum_{j=1}^n B(e_j,e_j) x_j^2.$$

Por otro lado, $x_j = e_j^*(x)$, donde $\mathcal{B}^* = (e_1^*, \dots, e_n^*)$ es la base dual de $\mathcal{B}$.
Así, tenemos que: $Q = \sum_{j=1}^n B(e_j,e_j)(e_j^*)^2$ ✓ ■

Ejercicio Sea $B: V \times V \to k$ forma bilineal. Supongamos que para todos $x, y \in V$ la condición "$B(x,y) = 0$" implica "$B(y,x) = 0$". Probar que B es necesariamente simétrica o alternada (cf. §9).

(Indicación: Considerar la expresión $B(x, B(x,y)z - B(x,z)y)$.)

§31. Clasificación de formas cuadráticas reales y complejas

En esta sección nos enfocaremos en el caso $k = \mathbb{R}$ y $\mathbb{C}$. Las formas cuadráticas definidas sobre $\mathbb{Q}$ o $\mathbb{Z}$ son muy importantes también, pero su clasificación es más delicada (Ver Jean-Pierre Serre "A course in arithmetic").

Def: Sean $Q: V \to k$ y $Q': V \to k$ dos formas cuadráticas. Decimos que Q y Q' son <u>equivalentes</u>, y escribimos $Q \sim Q'$, si existe un automorfismo $u \in GL(V)$ tq:
Para todo $x \in V$, $Q(x) = Q'(u(x))$.

En términos de las:
a) Forma bilineal simétrica asociada: $B(x,y) = B'(u(x), u(y))$ para todos $x, y \in V$. (gracias a la fórmula de polarización).
b) Matrices asociadas: Existen bases $\mathcal{B}$ y $\mathcal{B}'$ de V tales que $\mathrm{Mat}_{\mathcal{B}}(B) = \mathrm{Mat}_{\mathcal{B}'}(B')$
En otras palabras, $\exists P \in GL_m(k)$ invertible tq $\mathrm{Mat}_{\mathcal{B}}(B') = {}^tP \,\mathrm{Mat}_{\mathcal{B}}(B) P$.

Obs: En part, formas cuadráticas equivalentes tienen el <u>mismo rango</u>.

El resultado siguiente es el recíproco de la observación anterior: sobre $\mathbb{C}$ las formas cuadráticas están determinadas por su rango:

Teorema: <u>Toda</u> forma de rango r sobre un $\mathbb{C}$-e.v. V se escribe como
$$Q = f_1^2 + \dots + f_r^2$$
donde $f_1, \dots, f_r \in V^*$ son formas lineales en V linealmente independientes.

Dem: Sea $Q: V \to \mathbb{C}$ forma cuadrática compleja, donde $V \cong \mathbb{C}^n$, y sea $\mathcal{B} = (e_1, \dots, e_n)$ una base ortogonal (respecto a Q) de V. Luego, existen $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ tales que $Q\left(\sum_{j=1}^n x_j e_j\right) = \sum_{j=1}^n \lambda_j x_j^2$.
La matriz de Q en la base $\mathcal{B}$ está dada por $A_{\mathcal{B}} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \in M_n(\mathbb{C})$.
Por ende, el rango r de Q es el número de $\lambda_j \neq 0$: Reordenando la base $\mathcal{B}$ si fuese necesario, podemos suponer que $\lambda_1, \dots, \lambda_r \neq 0$ y $\lambda_j = 0$ para $j > r$.
Sobre $\mathbb{C}$, cada $\lambda_1, \dots, \lambda_r \neq 0$ puede ser escrito como $\lambda_j = \mu_j^2$ para ciertos $\mu_1, \dots, \mu_r \neq 0$ complejos no-nulos (esto último <u>no</u> siempre es posible en $\mathbb{R}$).
Si $(f_1, \dots, f_n)$ es la base dual de la base $\left(\frac{e_1}{\mu_1}, \dots, \frac{e_r}{\mu_r}, e_{r+1}, \dots, e_n\right)$ de V, entonces $Q = f_1^2 + \dots + f_r^2$. ∎

⚠ Las formas $f_1, \dots, f_r \in V^*$ linealmente independientes <u>no</u> son únicas: Sea $Q: \mathbb{C}^2 \to \mathbb{C}$ dada por $Q(x,y) = 2x^2 + 2y^2$, entonces:
$$Q(x,y) = (x\sqrt{2})^2 + (y\sqrt{2})^2 = (x+y)^2 + (x-y)^2.$$

Obs: La conclusión del teorema <u>no</u> es cierta en $\mathbb{R}$, pues se tendría $Q(x) \geq 0$ para todo $x \in V \cong \mathbb{R}^n$ (lo cual <u>no</u> es necesariamente cierto!).

Corolario: Sean $Q: V \to \mathbb{C}$ y $Q': V \to \mathbb{C}$ dos formas cuadráticas en un $\mathbb{C}$-ev V. Entonces, $Q \sim Q' \Leftrightarrow rg(Q) = rg(Q')$.

Dem: Ya vimos que $Q \sim Q' \Rightarrow rg(Q) = rg(Q')$. Por otro lado, si $rg(Q) = rg(Q')$ tienen el mismo rango r, el teorema anterior implica que existen bases de V tq Q y Q' tienen la misma matriz $\left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0_{m-r} \end{array}\right) \Rightarrow Q \sim Q'$ ✓ ■

Veamos ahora lo que ocurre sobre $\mathbb{R}$:

Def: Sea $Q: V \to \mathbb{R}$ forma cuadrática real sobre un $\mathbb{R}$-ev $V \cong \mathbb{R}^m$. Decimos que la forma cuadrática Q es:

(a) Definida positiva si $Q(x) > 0$ para todo $x \neq 0$ en V.

(a)' Semi-definida positiva si $Q(x) \geq 0$ para todo $x \in V$.

(b) Definida negativa si $Q(x) < 0$ para todo $x \neq 0$ en V.

(b)' Semi-definida negativa si $Q(x) \leq 0$ para todo $x \in V$.

Terminología: Decimos que una forma bilineal simétrica $B: V \times V \to \mathbb{R}$ (o bien, una matriz real simétrica A) es (semi-) definida positiva o negativa si la forma cuadrática asociada $Q: V \to \mathbb{R}$ lo es.

El siguiente resultado, también conocido como "Ley de inercia de Sylvester", permite descomponer las formas cuadráticas reales en partes positivas y negativas:

Teorema (Sylvester, 1852): Toda forma cuadrática de rango r sobre un $\mathbb{R}$-ev V se escribe 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 formas lineales en V linealmente independientes, y donde $p \in \mathbb{N}$ es un entero que sólo depende de Q. Escribimos $q := r - p$ y decimos que el par (p, q) es la signatura de la forma cuadrática Q.

Dem: Sea $Q: V \to \mathbb{R}$ forma cuadrática real, donde $V \cong \mathbb{R}^m$, y sea $\mathcal{B} = (e_1, \ldots, e_m)$ una base ortogonal (resp. a Q) de V. Luego, existen $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$ tales que $Q\left(\sum_{j=1}^m x_j e_j\right) = \sum_{j=1}^m \lambda_j x_j^2$. Tal como en el caso complejo, podemos suponer que los $\lambda_1, \ldots, \lambda_r \neq 0$ son no-nulos y $\lambda_j = 0$ para $j > r$. Más aún, reordenando $\mathcal{B}$ si fuese necesario, podemos suponer que los primeros p valores $\lambda_1, \ldots, \lambda_p > 0$ son positivos y los q siguientes $\lambda_{p+1}, \ldots, \lambda_r < 0$ son negativos. Sobre $\mathbb{R}$, consideramos $\mu_j := \sqrt{|\lambda_j|}$ para $j = 1, \ldots, r$ (esto último no siempre es posible en $\mathbb{Q}$). Si $(f_1, \ldots, f_m)$ es la base dual de la base $\left(\frac{e_1}{\mu_1}, \ldots, \frac{e_r}{\mu_r}, e_{r+1}, \ldots, e_m\right)$ de V, entonces $Q = f_1^2 + \ldots + f_p^2 - f_{p+1}^2 - \ldots - f_r^2$.

Veamos ahora que p no depende de la base $\mathcal{B}$ (ie, está determinado de manera geométrica por la forma cuadrática Q):

Sean $\mathcal{U} = \mathrm{Vect}_{\mathbb{R}}(e_1, \dots, e_p)$ y $W = \mathrm{Vect}_{\mathbb{R}}(e_{p+1}, \dots, e_n)$ sub-ev de V. Entonces, la restricción $Q|_{\mathcal{U}} : \mathcal{U} \to \mathbb{R}$ es definida positiva y la restricción $Q|_W$ es semi-definida negativa. En part, si $\mathcal{P} \subseteq V$ es cualquier sub-ev tq $Q|_{\mathcal{P}} : \mathcal{P} \to \mathbb{R}$ es definida positiva entonces necesariamente $\mathcal{P} \cap W = \{0\}$ y luego $\dim_{\mathbb{R}}(\mathcal{P}) + \underbrace{\dim_{\mathbb{R}}(W)}_{= n-p} \leq n$, es decir, $\dim_{\mathbb{R}}(\mathcal{P}) \leq p$. En otras palabras:

$$p = \max\{\dim_{\mathbb{R}}(\mathcal{P}), \text{ donde } Q|_{\mathcal{P}} \text{ definida positiva}\}$$

está completamente determinado por la forma cuadrática Q ✓ ■

Corolario: Sean $Q : V \to \mathbb{R}$ y $Q' : V \to \mathbb{R}$ dos formas cuadráticas en un $\mathbb{R}$-ev V, y sean (p,q) y (p',q') las signaturas respectivas (donde $\mathrm{rg}(Q) = p+q$ y $\mathrm{rg}(Q') = p'+q'$). Entonces, $Q \sim Q' \iff (p,q) = (p',q')$.

Dem: Ya vimos que $Q \sim Q'$ implica que $\mathrm{rg}(Q) = \mathrm{rg}(Q')$. Más aún, el entero p (resp. p') está caracterizado geométricamente en términos de Q (resp. Q'). En part, si $Q \sim Q'$ entonces $p = p'$ ($\Rightarrow q = q'$).
Por otro lado, si $(p,q) = (p',q')$, entonces (gracias al teorema anterior) existen bases de V tq Q y Q' tienen la misma matriz $\begin{pmatrix} I_p & 0 & 0 \\ 0 & -I_q & 0 \\ 0 & 0 & 0_{n-r} \end{pmatrix} \Rightarrow Q \sim Q'$ ✓ ■

Ejemplos: ① Sea $V \cong \mathbb{R}^n$ y $Q : V \to \mathbb{R}$ forma cuadrática de signatura (p,q), entonces:

i) Q semi-definida positiva (resp. negativa) $\iff q = 0$ (resp. $p = 0$).

ii) Q definida positiva (resp. negativa) $\iff (p,q) = (n,0)$ (resp. $(p,q) = (0,n)$).

② Sea $c \in \mathbb{R}^{>0}$ constante real positiva. 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 $(3,1)$.

③ Si $Q : V \to \mathbb{R}$ tiene signatura (p,q), entonces la forma cuadrática $-Q : V \to \mathbb{R}$, $x \mapsto -Q(x)$ tiene signatura (q,p).

Ejercicio Sea $V \cong \mathbb{R}^n$ y $Q : V \to \mathbb{R}$ forma cuadrática no-degenerada de signatura $(1, n-1)$. Sea $x \in V$ tal que $Q(x) > 0$.

① Probar que la restricción de Q a $H := L^{\perp}$ es definida negativa, donde $L = \mathrm{Vect}_{\mathbb{R}}(x)$ es la recta generada por $x \in V$.

② Sea $B : V \times V \to \mathbb{R}$ la forma bilineal simétrica asociada y sea $y \in V$ tal que $B(x,y) = 0$. Probar que $Q(y) \leq 0$, y que $Q(y) = 0 \iff y = 0$.

§32. 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 2 casos a considerar

Caso 1 Si Q "contiene un cuadrado", ie, un término de la forma x_j^2:
Reordenando si fuese necesario, podemos suponer que Q se escribe como

$$Q(x_1,\dots,x_n) = \lambda_1 x_1^2 + x_1 f(x_2,\dots,x_n) + R(x_2,\dots,x_n),$$

donde $\lambda_1 \neq 0$, f forma lineal y R forma cuadrática. Formamos un cuadrado de binomio al escribir

$$Q(x_1,\dots,x_n) = \lambda_1\left(x_1 + \frac{f}{2\lambda_1}\right)^2 - \frac{f^2}{4\lambda_1} + R \quad ; \text{ y definimos } f_1 := x_1 + \frac{f}{2\lambda_1}.$$

Por inducción en el número de variables, podemos escribir la forma cuadrática $-\frac{f^2}{4\lambda_1} + R$ como $\lambda_2 f_2^2 + \dots + \lambda_r f_r^2$, donde las $f_2,\dots,f_r$ son formas lineales en las variables $(x_2,\dots,x_n)$ que son l.i. Dado que f_1 contiene un término no-nulo en la variable x_1, f_1 no es combinación lineal de $f_2,\dots,f_r$
$\Rightarrow Q = \lambda_1 f_1^2 + \dots + \lambda_r f_r^2$ con $f_1,\dots,f_r$ formas l.i. ■

Caso 2 Si Q "no contiene cuadrados":
Reordenando si fuese necesario, podemos suponer que Q se escribe como

$$Q(x_1,\dots,x_n) = a x_1 x_2 + x_1 g_1(x_3,\dots,x_n) + x_2 g_2(x_3,\dots,x_n) + R(x_3,\dots,x_n),$$

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:

$$\begin{aligned} Q(x_1,\dots,x_n) &= a\left(x_1 + \frac{g_2}{a}\right)\left(x_2 + \frac{g_1}{a}\right) - \frac{g_1 g_2}{a} + R \\ &= \frac{a}{4}\left(x_1 + x_2 + \frac{g_2+g_1}{a}\right)^2 - \frac{a}{4}\left(x_1 - x_2 + \frac{g_2-g_1}{a}\right)^2 - \frac{g_1 g_2}{a} + R \end{aligned}$$

y definimos $f_1 := x_1 + x_2 + \frac{g_2+g_1}{a}$, $f_2 := x_1 - x_2 + \frac{g_2-g_1}{a}$. Tal como antes, aplicamos la hipótesis de inducción a la forma cuadrática dada por $-\frac{1}{a} g_1 g_2 + R$ y la escribimos como $\lambda_3 f_3^2 + \dots + \lambda_r f_r^2$, donde las $f_3,\dots,f_r$ son formas lineales en las variables $(x_3,\dots,x_n)$ que son l.i. Igual que antes:
$\Rightarrow Q = \lambda_1 f_1^2 + \dots + \lambda_r f_r^2$ con $f_1,\dots,f_r$ formas l.i. ■

Ejemplos: ① Sea $Q: \mathbb{R}^3 \to \mathbb{R}$ dada por $Q(x,y,z) = xy + yz + zx$ ("sin cuadrados")
Luego: $Q(x,y,z) = xy + \underbrace{xz}_{g_1} + \underbrace{yz}_{g_2} = (x+z)(y+z) - z^2$

$$= \frac{1}{4}(x+y+2z)^2 - \frac{1}{4}(x-y)^2 - z^2 \quad \rightsquigarrow \text{signatura } (1,2)$$
$$\text{y } \operatorname{rg}(Q) = 3 \text{ (no-degenerada)}.$$

② Sea $Q:\mathbb{R}^4 \to \mathbb{R}$ dada por $Q(x,y,z,t) = xy + yz + zt + tx$ ("sin cuadrados").

Luego: $Q(x,y,z,t) = xy + x\underbrace{t}_{g_1} + y\underbrace{z}_{g_2} + \underbrace{zt}_{R} = (x+z)(y+t) - \cancel{tz} + \cancel{zt}$

$$= \frac{1}{4}(x+y+z+t)^2 - \frac{1}{4}(x-y+z-t)^2$$

$\Rightarrow$ Signatura de Q es $(1,1)$ y $rg(Q) = 2$ (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:

$$Q(x,y,z,t) = 3z^2 - 2zt + \underbrace{(x^2 + 2y^2 + tx + 3xy - yt)}_{R_1(x,y,t)} = 3\left(z - \frac{t}{3}\right)^2 - \frac{t^2}{3} + R_2(x,y,t)$$

En la expresión $-\frac{t^2}{3} + R_2 = x^2 + 2y^2 + tx + 3xy - yt - \frac{t^2}{3}$ escogemos la variable x:

$$x^2 + x(t+3y) + \left(2y^2 - yt - \frac{t^2}{3}\right) = \left(x + \frac{1}{2}(t+3y)\right)^2 - \frac{1}{4}(t+3y)^2 + 2y^2 - yt - \frac{t^2}{3}$$

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:

$$-\frac{1}{4}y^2 - \frac{7}{12}t^2 - \frac{5}{2}yt = -\frac{1}{4}\left(y^2 + 10yt\right) - \frac{7}{12}t^2 = -\frac{1}{4}(y+5t)^2 + \frac{17}{3}t^2$$

$$\Rightarrow Q(x,y,z,t) = 3\left(z - \frac{t}{3}\right)^2 + \left(x + \frac{t}{2} + \frac{3}{2}y\right)^2 - \frac{1}{4}(y+5t)^2 + \frac{17}{3}t^2$$

Luego, la signatura de Q es $(3,1)$ y $rg(Q) = 4$ (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:

$$\begin{aligned} Q(x,y,z) &= x^2 + x(4y+2z) + (y^2 + 3z^2 + 2yz) \\ &= (x+2y+z)^2 - (2y+z)^2 + y^2 + 3z^2 + 2yz \\ &= (x+2y+z)^2 - 3y^2 + 2z^2 - 2yz. \end{aligned}$$

En la expresión $-3y^2 + 2z^2 - 2yz$ escogemos y:

$$-3y^2 + 2z^2 - 2yz = -3\left(y^2 + \frac{2}{3}yz\right) + 2z^2 = -3\left(y + \frac{z}{3}\right)^2 + \frac{7}{3}z^2$$

$$\Rightarrow Q(x,y,z) = (x+2y+z)^2 - 3\left(y + \frac{z}{3}\right)^2 + \frac{7}{3}z^2$$

Luego, la signatura de Q es $(2,1)$ y $rg(Q) = 3$ (no-degenerada).

Ejercicio 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 $rg(Q)$.

② Determinar una base de $\mathbb{R}^3$ ortogonal respecto a Q.

§33. Grupo ortogonal

Sea V un k-ev de $\dim_k(V) = n$ y sea $Q : V \to k$ forma cuadrática.
Denotamos por $B : V \times V \to k$ la (única) forma bilineal simétrica asociada a Q.

Lema: Sea $u : V \to V$ endomorfismo. Son equivalentes:
① u preserva la forma bilineal B: $B(u(x), u(y)) = B(x,y)$ $\forall x, y \in V$.
② u preserva la forma cuadrática Q: $Q(u(x)) = Q(x)$ $\forall x \in V$.

Dem: ① $\Rightarrow$ ② pues por definición $Q(x) = B(x,x)$ ✓ Por otro lado, si $u : V \to V$ preserva Q entonces la fórmula de polarización implica

$$B(u(x), u(y)) = \frac{1}{2}\big(Q(\underbrace{u(x)+u(y)}_{u(x+y)}) - Q(u(x)) - Q(u(y))$$
$$= \frac{1}{2}(Q(x+y) - Q(x) - Q(y)) = B(x,y) \quad \checkmark \quad \blacksquare$$

Def: Sea V un k-ev y $Q : V \to k$ forma cuadrática. Decimos que un automorfismo $u : V \xrightarrow{\sim} V$ (ie, $u \in GL(V)$) es ortogonal respecto a Q si preserva Q ($\circ B$).

Notación: $O(Q) := \{u \in GL(V) \text{ tal que } u \text{ es ortogonal respecto a } Q\} \subseteq GL(V)$.

Prop: Sea $Q : V \to k$ forma cuadrática. El conjunto $O(Q)$ es un sub-grupo de $GL(V)$, llamado el grupo ortogonal de Q.

Dem: (i) La identidad $Id_V \in GL(V)$ preserva Q.
(ii) Si $u, v \in O(Q)$ preservan Q, entonces $u \circ v$ también, pues

$$Q((u\circ v)(x)) = Q(u(v(x)) \underset{\uparrow u \in O(Q)}{=} Q(v(x)) \underset{\uparrow v \in O(Q)}{=} Q(x) \quad \checkmark \quad \Rightarrow u \circ v \in O(Q).$$

(iii) Si $u \in O(Q)$ preserva Q, entonces u^{-1} preserva Q: Sea $x \in V$ y sea $x' := u^{-1}(x)$ entonces $Q(u^{-1}(x)) = Q(x') \underset{\uparrow u \in O(Q)}{=} Q(u(x')) \underset{\uparrow u(x') = x}{=} Q(x)$ ✓ $\Rightarrow u^{-1} \in O(Q)$ $\blacksquare$

Importante: Si $Q : V \to k$ es no-degenerada, entonces todo endomorfismo $u : V \to V$ que preserva Q es necesariamente biyectivo (ie, $u \in GL(V)$). En efecto, si $u(x) = 0$ para $x \in V$ entonces (dado que u preserva la forma bilineal simétrica B):

$$B(x,y) = B(u(x), u(y)) = 0 \quad \text{para todo } y \in V$$

$\Rightarrow x \in V^{\perp} \underset{B \text{ no-deg}}{\Rightarrow} x = 0$. Luego, u es inyectivo y por lo tanto biyectivo.

Recuerdo: Sea $Q : V \to k$ forma cuadrática, y $B : V \times V \to k$ forma bilineal simétrica asociada. Si $\mathcal{B}$ es una base de V, decimos que la matriz de Q en la base $\mathcal{B}$ es la matriz de B en la base $\mathcal{B}$:

$$Mat_{\mathcal{B}}(Q) := Mat_{\mathcal{B}}(B) = A_{\mathcal{B}} \in M_n(k).$$

Prop: Sea $Q: V \to k$ forma cuadrática, $\mathcal{B}$ una base de V y sea $A_{\mathcal{B}} = Mat_{\mathcal{B}}(Q)$ la matriz de Q en la base $\mathcal{B}$. Sea $u: V \to V$ endomorfismo y sea $M = Mat_{\mathcal{B}}(u)$. Entonces:

① u es ortogonal respecto a $Q \iff {}^tMA_{\mathcal{B}}M = A_{\mathcal{B}}$.

② Si Q es no-degenerada y $u \in O(Q)$, entonces $\det(u) = \pm 1$. Más aún, el conjunto $SO(Q) := \{u \in O(Q) \text{ tal que } \det(u) = 1\}$ es un sub-grupo de $O(Q)$, llamado el grupo especial ortogonal de Q.

Dem: ① Sean $x, y \in V$. Si representamos x e y por los vectores columna X e Y en k^n dados por sus coordenadas respecto a la base $\mathcal{B}$, entonces sabemos que:

$$B(x,y) = {}^tXA_{\mathcal{B}}Y \Rightarrow B(u(x),u(y)) = {}^t(MX)A_{\mathcal{B}}(MY) = {}^tX\,{}^tMA_{\mathcal{B}}MY.$$

Es decir, la matriz de la forma bilineal $(x,y) \mapsto B(u(x),u(y))$ en la base $\mathcal{B}$ es ${}^tMA_{\mathcal{B}}M$. Luego, u es ortogonal resp. a $\mathcal{B}$ (ie, resp. a Q) si y sólo si las formas bilineales $B(\cdot,\cdot) = B(u(\cdot),u(\cdot))$ son iguales $\iff {}^tMA_{\mathcal{B}}M = A_{\mathcal{B}}$ ✓

② Q es no-degenerada $\iff A_{\mathcal{B}} \in GL_n(k) \iff \det(A_{\mathcal{B}}) \neq 0$. Luego, si $u \in O(Q)$ entonces ${}^tMA_{\mathcal{B}}M = A_{\mathcal{B}} \Rightarrow \det(M)^2 \det(A_{\mathcal{B}}) = \det(A_{\mathcal{B}}) \Rightarrow \det(M)^2 = 1$. Dado que $\text{car}(k) \neq 2$, $\det(M) = 1$ o -1. Finalmente, notar que:

(i) $Id_V \in O(Q)$ cumple $\det(Id_V) = 1$, ie, $Id_V \in SO(Q)$.

(ii) Si $u, v \in SO(Q)$ entonces $\det(u \circ v) = \underbrace{\det(u)}_{=1}\underbrace{\det(v)}_{=1} = 1 \Rightarrow u \circ v \in SO(Q)$

(iii) Si $u \in SO(Q)$ entonces $\det(u^{-1}) = 1/\det(u) = 1 \Rightarrow u^{-1} \in SO(Q)$ ■

Prop: Sean $Q: V \to k$ y $Q': V \to k$ dos formas cuadráticas en V. Entonces: Si $Q \sim Q'$ son equivalentes $\Rightarrow$ Los grupos $O(Q) \cong O(Q')$ son isomorfos.

Dem: Si $Q \sim Q'$ entonces existe un automorfismo $v \in GL(V)$ tq para todo $x \in V$ se tiene $Q(x) = Q'(v(x))$. Sea $u: V \to V$ endomorfismo tq $u \in O(Q)$ (en part, $u \in GL(V)$) y sea $x \in V$, entonces:

$$Q'(x) = Q'(v \circ v^{-1}(x)) \underset{Q \sim Q'}{=} Q(v^{-1}(x)) \underset{u \in O(Q)}{=} Q(u \circ v^{-1}(x)) \underset{Q \sim Q'}{=} Q'(v \circ u \circ v^{-1}(x)).$$

Luego, $v \circ u \circ v^{-1} \in O(Q')$. Sea $\varphi: O(Q) \to O(Q')$, $u \mapsto \varphi(u) := v \circ u \circ v^{-1}$. $\Rightarrow \varphi$ es biyectiva (pues su inversa es $\varphi^{-1}: O(Q') \to O(Q)$, $u' \mapsto v^{-1} \circ u' \circ v$) y es un morfismo de grupos (ie, $\varphi(u_1 \circ u_2) = \varphi(u_1) \circ \varphi(u_2)$) $\Rightarrow O(Q) \cong O(Q')$. ■

Notación: El estudio de $O(Q)$ para formas $Q: V \to k$ no-degeneradas se reduce a:

① Sobre $k = \mathbb{C}$: $Q(x) = x_1^2 + \ldots + x_n^2 \rightsquigarrow O(n,\mathbb{C}) = O_n(\mathbb{C})$

② Sobre $k = \mathbb{R}$: $Q(x) = x_1^2 + \ldots + x_p^2 - x_{p+1}^2 - \ldots - x_{p+q}^2$, $n = p+q \rightsquigarrow O(p,q,\mathbb{R}) = O(p,q)$.

Si $(p,q) = (n,0)$ escribimos $O(n,0) = O(n,\mathbb{R}) = O_n(\mathbb{R})$.

Ejercicio ① Probar que $O(p,q) \cong O(q,p)$. En part, $O(Q) \cong O(Q')$ no implica $Q \sim Q'$.

② Probar que $O_n(k) = \{A \in M_n(k) \text{ tq } {}^tAA = I_n\}$, donde $k = \mathbb{R}$ o $\mathbb{C}$.

§34. Productos escalar y espacios euclideanos

En toda la sección, $V \cong \mathbb{R}^n$ será un $\mathbb{R}$-ev de dimensión finita.

Def: Sea V un $\mathbb{R}$-ev de dimensión finita. Un **producto escalar** (o "producto interno", o "producto punto") en V es una forma bilineal simétrica 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 denota usualmente como $\langle x, y \rangle$. Así, la forma **bilineal**

$$\begin{aligned} \langle \cdot, \cdot \rangle : V \times V &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto \langle x, y \rangle \end{aligned}$$

es (i) **simétrica**: $\langle x, y \rangle = \langle y, x \rangle$ para todos $x, y \in V$.

(ii) **definida positiva**: $\langle x, x \rangle \in \mathbb{R}^{>0}$ para $x \neq 0$ ($\Rightarrow$ el único vector isótropo es el vector nulo).

Obs: En part, $\langle \cdot, \cdot \rangle$ es **no-degenerada**. Más aún, dado que la signatura de $\langle \cdot, \cdot \rangle$ es $(n, 0)$, donde $n = \dim_{\mathbb{R}}(V)$, siempre existe una base ortogonal $\mathcal{B} = (e_1, \dots, e_n)$ de $V \cong \mathbb{R}^n$ tal que si $x = \sum_{j=1}^{n} x_j e_j$, entonces

$$\langle x, x \rangle = x_1^2 + \dots + x_n^2 .$$

Ejemplo principal: Sea $V = \mathbb{R}^n$, la relación

$$\langle (x_1, \dots, x_n), (y_1, \dots, y_n) \rangle := x_1 y_1 + \dots + x_n y_n$$

define un producto escalar "**canónico**" en $\mathbb{R}^n$, dotándolo así de estructura de espacio euclideano.

Ejercicio Sea $V = \mathbb{R}_d[X]$ el $\mathbb{R}$-ev de polinomios con coeficientes reales en la variable X de grado $\leq d$. Probar que la relación

$$\langle P, Q \rangle = \int_{-1}^{1} P(t)\, Q(t)\, dt$$

define un producto escalar en V.

Notación: Sea V espacio euclideano. Para todo $x \in V$ escribimos

$$\|x\| := \sqrt{\langle x, x \rangle} .$$

En part, $\|x\|^2 = \langle x, x \rangle$ es la forma cuadrática asociada a $\langle \cdot, \cdot \rangle$. Luego, la fórmula de **polarización** se escribe como

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2\langle x, y \rangle .$$

Así, si x e y son **ortogonales** (ie, $\langle x, y \rangle = 0$) entonces se tiene

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2 \quad \text{(Pitágoras)}$$

Lema (desigualdad de Cauchy-Schwarz): Sea V un espacio euclideano. Para todos $x, y \in V$ se tiene $|\langle x,y\rangle| \leq \|x\|\|y\|$. Más aún, $|\langle x,y\rangle| = \|x\|\|y\|$ si y sólo si x e y son linealmente dependientes (ie, colineales).

Dem: Si $x = 0$ el resultado es automático. Sup. que $x \neq 0$ y consideremos para $t \in \mathbb{R}$ la función $f(t) = \|tx + y\|^2 \geq 0$. Notamos que

$$\|tx+y\|^2 = \langle tx+y, tx+y\rangle = t^2\langle x,x\rangle + 2t\langle x,y\rangle + \langle y,y\rangle$$
$$= t^2\|x\|^2 + 2t\langle x,y\rangle + \|y\|^2 \geq 0 \quad \text{para todo } t \in \mathbb{R}.$$

En part, el discriminante Δ del polinomio de grado 2 $f(t)$ es ≤ 0.

$\Rightarrow \Delta = b^2 - 4ac = 4\langle x,y\rangle^2 - 4\|x\|^2\|y\|^2 \leq 0 \Rightarrow |\langle x,y\rangle| \leq \|x\|\|y\|$ ✓

Más aún, la igualdad se alcanza cuando $\Delta = 0$, ie, cuando $f(t)$ tiene una raíz doble $t_0 \in \mathbb{R}$: $f(t_0) = \|t_0 x + y\|^2 = 0 \Rightarrow t_0 x + y = 0$ ✓ ■

La desigualdad de Cauchy-Schwarz implica que si $x, y \in V \setminus \{0\}$ son no-nulos entonces $-1 \leq \dfrac{\langle x,y\rangle}{\|x\|\|y\|} \leq 1$. Luego, dijimos:

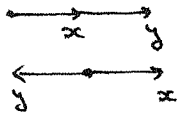
Def: Sea V un espacio euclideano. Sean $x, y \in V \setminus \{0\}$ vectores no-nulos, entonces definimos el ángulo no-orientado en x e y como el único real $\theta = \theta(x,y)$ en $[0,\pi]$ tal que

$$\cos(\theta) = \frac{\langle x,y\rangle}{\|x\|\|y\|}.$$

Obs: Decimos que $\theta(x,y)$ es no-orientado pues $\theta(x,y) = \theta(y,x)$. Notar que, gracias a la desigualdad de Cauchy-Schwarz, tenemos que:

① $\theta(x,y) = 0 \iff y = \lambda x$ con $\lambda > 0$:

② $\theta(x,y) = \pi \iff y = \lambda x$ con $\lambda < 0$:



③ $\theta(x,y) = \dfrac{\pi}{2} \iff x$ e y son ortogonales (ie, $\langle x,y\rangle = 0$):



Teorema: Sea V un espacio euclideano. Entonces, la aplicación $\|\cdot\| : V \to \mathbb{R}^{\geq 0}$ es una norma. Es decir, verifica:

① $\|x\| = 0 \iff x = 0$

② $\|\lambda x\| = |\lambda| \cdot \|x\|$

③ $\|x+y\| \leq \|x\| + \|y\|$ ("desigualdad triangular")

Dem: ① y ② se obtienen de la definición $\|x\| = \sqrt{\langle x,x\rangle}$. Para ③ notamos que

$$\|x+y\|^2 = \langle x+y, x+y\rangle = \|x\|^2 + 2\langle x,y\rangle + \|y\|^2 \underset{CS}{\leq} \underbrace{\|x\|^2 + 2\|x\|\|y\| + \|y\|^2}_{(\|x\| + \|y\|)^2}$$

$\Rightarrow \|x+y\| \leq \|x\| + \|y\|$ ✓ (con igualdad $\iff x$ e y son colineales). ■

§35. Bases ortonormadas y ortogonalización de Gram-Schmidt

Recordemos que una forma bilineal simétrica definida positiva no posee vectores isótropos no-nulos. En particular, tenemos:

Teorema: Sea V un espacio euclideano. Entonces, para todo $\mathcal{U} \subseteq V$ sub-esp se tiene que $V = \mathcal{U} \oplus \mathcal{U}^\perp$.

Dem: Sabemos que $\dim_{\mathbb{R}}(\mathcal{U}) + \dim_{\mathbb{R}}(\mathcal{U}^\perp) = \dim_{\mathbb{R}}(V)$. Por otro lado, todo vector en $\mathcal{U} \cap \mathcal{U}^\perp$ es isótropo, y luego es nulo. ∎

Obs: Dado que $\langle \cdot, \cdot \rangle$ es definida positiva, i.e., de signatura $(n, 0)$ donde $n = \dim_{\mathbb{R}}(V)$, tenemos que V admite una base ortogonal $\mathcal{B} = (\tilde{e}_1, \dots, \tilde{e}_n)$ respecto a $\langle \cdot, \cdot \rangle$ tal que $\langle \tilde{e}_i, \tilde{e}_i \rangle = \|\tilde{e}_i\|^2 > 0$ para todo $i = 1, \dots, n$ (ver §31). En part, si definimos $e_i := \dfrac{\tilde{e}_i}{\|\tilde{e}_i\|} \in V$ se tiene que $\|e_i\| = 1$.

Definición: Sea V un espacio euclideano, y sea $\mathcal{B} = (e_1, \dots, e_n)$ una base de V. Decimos que $\mathcal{B}$ es una base ortonormal (o base ortonormada) si $\langle e_i, e_j \rangle = \delta_{ij}$ para todos $i, j \in \{1, \dots, n\}$. De manera más general, decimos que un conjunto $x_1, \dots, x_m$ de vectores es una familia ortonormal si $\langle x_i, x_j \rangle = \delta_{ij}$ para todos $i, j \in \{1, \dots, m\}$.

Obs: ① Vimos que todo espacio euclideano posee una base ortonormal.

② La matriz $A_{\mathcal{B}}$ de la forma bilineal $\langle \cdot, \cdot \rangle$ resp. a una base ortonormal $\mathcal{B}$ es $A_{\mathcal{B}} = I_n$ (identidad).

③ Toda familia ortonormal $x_1, \dots, x_m$ es linealmente independiente (en part, $m \leq n$): si $\lambda_1 x_1 + \dots + \lambda_m x_m = 0 \Rightarrow 0 = \langle \lambda_1 x_1 + \dots + \lambda_m x_m, x_j \rangle = \lambda_j \underbrace{\langle x_j, x_j \rangle}_{=1} \Rightarrow \lambda_j = 0 \ \forall j$

Prop: Sea V un espacio euclideano y sea $\mathcal{B} = (e_1, \dots, e_n)$ una base ortonormal de V. Entonces, para todo $x \in V$ se tiene:

$$x = \sum_{j=1}^{n} \langle x, e_j \rangle e_j \qquad \text{y} \qquad \|x\|^2 = \sum_{j=1}^{n} \langle x, e_j \rangle^2$$

Dem: Escribamos $x = \sum_{j=1}^{n} x_j e_j \Rightarrow \langle x, e_i \rangle = \langle x_1 e_1 + \dots + x_n e_n, e_i \rangle = x_i \underbrace{\langle e_i, e_i \rangle}_{=1}$

$\Rightarrow x_i = \langle x, e_i \rangle$ para todo $i \in \{1, \dots, n\}$. En particular,

$$\|x\|^2 = \langle x, x \rangle = \left\langle \sum_{j=1}^{n} \langle x, e_j \rangle e_j, \sum_{j=1}^{n} \langle x, e_j \rangle e_j \right\rangle = \langle x, e_1 \rangle^2 \underbrace{\|e_1\|^2}_{=1} + \dots + \langle x, e_n \rangle^2 \underbrace{\|e_n\|^2}_{=1}$$ ∎

Recuerdo (ver §33): Si $Q : V \to \mathbb{R}$ forma cuadrática, y $B : V \times V \to \mathbb{R}$ forma bilineal simétrica asociada. Entonces un endomorfismo $u : V \to V$ es ortogonal respecto a Q si: $Q(u(x)) = Q(x) \ \forall x \in V \iff B(u(x), u(y)) = B(x, y) \ \forall x, y \in V$. Más aún, si B es no-degenerada $\Rightarrow u \in GL(V)$. En part, $u \in O(Q)$.

Prop/Def: Sea V un espacio euclídeano, y sea $u: V \to V$ 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 ortonormal $\mathcal{B}$ de V tal que la imagen $u(\mathcal{B})$ es una base ortonormal.
4. Para Toda base ortonormal $\mathcal{B}$ de V, la imagen $u(\mathcal{B})$ es ortonormal.
5. La matriz $A = \mathrm{Mat}_{\mathcal{B}}(u)$ de u respecto a una base ortonormal $\mathcal{B}$ de V es una matriz ortogonal, ie, ${}^tAA = I_n$.

Un endomorfismo $u: V \to V$ que cumple estas propiedades es una isometría de V.

Dem: Vimos que (1) $\Leftrightarrow$ (2) ✓ Por otro lado, (1) $\Rightarrow$ (4) $\Rightarrow$ (3) ✓ Veamos que (3) $\Rightarrow$ (2): Sea $\mathcal{B} = (e_1, \dots, e_n)$ base ortonormal tq $u(\mathcal{B}) = (u(e_1), \dots, u(e_n))$ es una base ortonormal. Sea $x \in V$ tal que $x = \sum_{j=1}^n x_j e_j \Rightarrow u(x) = \sum_{j=1}^n x_j u(e_j)$ (*)

Visto: $x = \sum_{j=1}^n \langle x, e_j\rangle e_j$ (ie, $x_j = \langle x, e_j\rangle$), y además (considerando la base $u(\mathcal{B})$): $u(x) = \sum_{j=1}^n \langle u(x), u(e_j)\rangle u(e_j)$ (**)

$\overset{(*)\,\&\,(**)}{\Longrightarrow} \langle u(x), u(e_j)\rangle = \langle x, e_j\rangle$ para todo $j \in \{1, \dots, n\}$. En part,

$$\|x\|^2 = \sum_{j=1}^n \langle x, e_j\rangle^2 = \sum_{j=1}^n \langle u(x), u(e_j)\rangle^2 = \|u(x)\|^2$$

de donde obtenemos (2) ✓

Finalmente, sabemos (ver §33, pág. 100) que un endomorfismo $u: V \to V$ preserva la forma bilineal simétrica $\langle \cdot, \cdot \rangle \Leftrightarrow {}^tA A_{\mathcal{B}} A = A_{\mathcal{B}}$, donde $A_{\mathcal{B}} = (\langle e_i, e_j\rangle)_{i,j}$ es I_n si $\mathcal{B} = (e_1, \dots, e_n)$ es ortonormal, ie, (1) $\Leftrightarrow$ (5) ✓ ■

Terminología: Si V es un espacio euclídeano, podemos definir la distancia entre $x, y \in V$ como:

$$d(x, y) = \|x - y\|.$$

Si $u: V \to V$ es un endomorfismo que preserva $\langle \cdot, \cdot \rangle$ (y luego, u es necesariamente biyectivo) 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 esta razón que se prefiere el término isometría (en lugar de endomorfismo ortogonal) en el contexto de espacios euclídeanos. Finalmente, recordemos que una isometría $u: V \to V$ verifica $\det(u) = \pm 1$. En este contexto, decimos que u es una:

1. isometría directa si $\det(u) = 1$ (ie, $u \in SO(Q)$, con $Q = \|\cdot\|^2$)
2. isometría indirecta si $\det(u) = -1$.

Corolario: Sea V un espacio euclídeano y sea $\mathcal{B}$ una base ortonormal. Entonces, una base $\mathcal{B}'$ de V es ortonormal $\Leftrightarrow$ La matriz $P = Mat_{\mathcal{B}}(\mathcal{B}')$ es ortogonal.

Dem: Sea $\mathcal{B} = (e_1, \dots, e_n)$ y $\mathcal{B}' = (e_1', \dots, e_n')$. Consideremos el endomorfismo $u: V \to V$ definido por $u(e_j) := e_j'$ para todo $j \in \{1, \dots, n\}$. Entonces, $Mat_{\mathcal{B}}(u) = P$ y luego: $\mathcal{B}'$ ortonormal $\Leftrightarrow {}^tPP = I_n$ (gracias a (4) $\Leftrightarrow$ (5)) ■

Prop: Sea V un espacio euclídeano y sea $u: V \to V$ una isometría. Entonces, para todo $\mathcal{U} \subseteq V$ subespacio se tiene:
(a) $u(\mathcal{U}^\perp) = u(\mathcal{U})^\perp$
(b) Si $u(\mathcal{U}) \subseteq \mathcal{U}$ entonces $u(\mathcal{U}^\perp) \subseteq \mathcal{U}^\perp$.

Dem: (a) Sea $y \in \mathcal{U}^\perp$, entonces para todo $x \in \mathcal{U}$ se tiene $\langle x, y\rangle = \langle u(x), u(y)\rangle = 0$, i.e. $u(y)$ es ortogonal a todos los elementos de $u(\mathcal{U}) \Leftrightarrow u(y) \in u(\mathcal{U})^\perp$.
$\Rightarrow u(\mathcal{U}^\perp) \subseteq u(\mathcal{U})^\perp$. Por otro lado, dado que $u \in GL(V)$ es un automorfismo:

$$\dim_{\mathbb{R}}(u(\mathcal{U}^\perp)) = \dim_{\mathbb{R}}(\mathcal{U}^\perp) = \dim_{\mathbb{R}}(V) - \dim_{\mathbb{R}}(\mathcal{U}) = \dim_{\mathbb{R}}(V) - \dim_{\mathbb{R}}(u(\mathcal{U})) = \dim_{\mathbb{R}}(u(\mathcal{U})^\perp)$$

$\Rightarrow u(\mathcal{U}^\perp) = u(\mathcal{U})^\perp$ ✓

(b) Dado que $u \in GL(V)$, $\dim_{\mathbb{R}}(\mathcal{U}) = \dim_{\mathbb{R}}(u(\mathcal{U}))$. Luego, si $u(\mathcal{U}) \subseteq \mathcal{U}$ entonces $u(\mathcal{U}) = \mathcal{U} \underset{(a)}{\Rightarrow} u(\mathcal{U})^\perp = u(\mathcal{U}^\perp) = \mathcal{U}^\perp$ ✓ ■

A continuación describiremos un algoritmo, llamado usualmente "proceso de ortogonalización de Gram-Schmidt", que permite construir una base ortonormal a partir de una base cualquiera. Este método fue descubierto por Laplace en 1816, pero usualmente se le atribuye (erroneamente) a Gram (1883) y a Schmidt (1907).

Teorema: Sea V un espacio euclídeano. Dada $(e_1, \dots, e_n)$ base de V, existe una única base ortonormal $(e_1', \dots, e_n')$ de V tal que:
(1) Para todo $j \in \{1, \dots, n\}$ se tiene $Vect_{\mathbb{R}}(e_1, \dots, e_j) = Vect_{\mathbb{R}}(e_1', \dots, e_j')$, y
(2) $\langle e_j, e_j'\rangle > 0$.

Dem: En la práctica, primero construimos una base $(e_1'', \dots, e_n'')$ que es sólo ortogonal y que cumple (1) y (2), y luego se "normaliza":
(1°) Definimos $e_1'' := e_1$
(2°) Suponemos $e_1'', \dots, e_{j-1}''$ ya construidos y buscamos e_j'' de la forma

$$e_j'' = e_j + \lambda_1 e_1'' + \dots + \lambda_{j-1} e_{j-1}''.$$

Para $k \in \{1, \dots, j-1\}$ se tiene $\langle e_j'', e_k''\rangle = 0 \Leftrightarrow \langle e_j, e_k''\rangle + \lambda_k \|e_k''\|^2 = 0$, lo cual determina $\lambda_1, \dots, \lambda_{j-1}$ y luego determina e_j''.

③ Por inducción, construimos $(e_1'', \dots, e_m'')$ base ortogonal, y notamos que: Verifica ① por construcción, y $\langle e_j, e_j'' \rangle = \langle e_j'', e_j'' \rangle = \|e_j''\|^2 > 0 \rightsquigarrow$ ② ✓

④ Definimos $e_j' := \frac{e_j''}{\|e_j''\|}$ para $j \in \{1, \dots, n\}$ y luego $(e_1', \dots, e_m')$ es una base ortonormal que verifica ① y ②, cuya unicidad se obtiene por construcción ■

[Corolario (Factorización QR): Sea $A \in Gl_m(\mathbb{R})$ matriz invertible. Entonces, existe un único par (Q, R) de matrices, donde:
① Q es ortogonal (ie, ${}^tQQ = I_m$),
② R es triangular superior con coeficientes diagonales > 0,
③ $A = QR$.

Dem: Los vectores columna $e_1, \dots, e_m$ de A forman una base $\mathcal{B}$ de $\mathbb{R}^n$, a la cual aplicamos Gram-Schmidt para obtener una base ortonormal $\mathcal{B}'$ de $\mathbb{R}^n$. Sea ξ la base canónica de $\mathbb{R}^n$, entonces (por definición): $A = Mat_\xi(\mathcal{B})$.
Además, la demostración del proceso de Gram-Schmidt nos dice precisamente que $Mat_{\mathcal{B}}(\mathcal{B}')$ es triangular superior con coeficientes diagonales estrictamente positivos. Por otro lado, sabemos que $\mathcal{B}'$ ortonormal $\Leftrightarrow Q := Mat_\xi(\mathcal{B}')$ es ortogonal.
Finalmente, $A = Mat_\xi(\mathcal{B}) = Mat_\xi(\mathcal{B}')\, Mat_{\mathcal{B}'}(\mathcal{B}) = Q\, Mat_{\mathcal{B}'}(\mathcal{B}) = Q\, Mat_{\mathcal{B}}(\mathcal{B}')^{-1}$, donde $R := Mat_{\mathcal{B}}(\mathcal{B}')^{-1}$ es triangular superior con coeficientes diagonales > 0 ■

Ejemplo: Sea $(e_1, e_2, e_3) = ((1,1,1), (1,1,0), (1,0,0))$ base de $\mathbb{R}^3$. Gram-Schmidt:
$e_1'' := e_1 = (1,1,1) \rightsquigarrow e_2'' := e_2 - \frac{\langle e_2, e_1'' \rangle}{\|e_1''\|^2} e_1'' = (1,1,0) - \frac{2}{3}(1,1,1) = \left(\frac{1}{3}, \frac{1}{3}, \frac{-2}{3}\right)$.
$\rightsquigarrow e_3'' := e_3 - \frac{\langle e_3, e_1'' \rangle}{\|e_1''\|^2} e_1'' - \frac{\langle e_3, e_2'' \rangle}{\|e_2''\|^2} e_2'' = (1,0,0) - \frac{1}{3}(1,1,1) - \frac{1}{2}\left(\frac{1}{3}, \frac{1}{3}, \frac{-2}{3}\right) = \left(\frac{1}{2}, \frac{-1}{2}, 0\right)$
$\Rightarrow e_1' = \frac{e_1''}{\|e_1''\|} = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)$, $e_2' = \frac{e_2''}{\|e_2''\|} = \left(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{-2}{\sqrt{6}}\right)$, $e_3' = \frac{e_3''}{\|e_3''\|} = \left(\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}, 0\right)$ ✓

⚠ 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 $e_1' = (1,0,0)$, $e_2' = (0,1,0)$, $e_3' = (0,0,1)$!
En términos de factorización QR: Si $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$ entonces la matriz

$Q = \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{6} & 1/\sqrt{2} \\ 1/\sqrt{3} & 1/\sqrt{6} & -1/\sqrt{2} \\ 1/\sqrt{3} & -2/\sqrt{6} & 0 \end{pmatrix}$ es ortogonal y $R = Q^{-1}A = {}^tQA = \begin{pmatrix} \sqrt{3} & 2/\sqrt{3} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{6} \\ 0 & 0 & 1/\sqrt{2} \end{pmatrix}$

Luego, $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} = QR = \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{6} & 1/\sqrt{2} \\ 1/\sqrt{3} & 1/\sqrt{6} & -1/\sqrt{2} \\ 1/\sqrt{3} & -2/\sqrt{6} & 0 \end{pmatrix} \begin{pmatrix} \sqrt{3} & 2/\sqrt{3} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{6} \\ 0 & 0 & 1/\sqrt{2} \end{pmatrix}$

[Obs: Notar que el proceso de ortogonalización permite asociar a una familia linealmente independiente $(e_1, \dots, e_j)$ (no nec. una base), una familia ortonormal.

§36. Endomorfismo adjunto y endomorfismos simétricos (o auto-adjuntos)

Prop: Sea V un espacio euclídeano. Para todo endomorfismo $u: V \to V$ existe un único endomorfismo $u^*: V \to V$ que cumple:

$$\langle u(x), y\rangle = \langle x, u^*(y)\rangle \quad \text{para todos } x, y \in V.$$

Además, $rg(u) = rg(u^*)$. Decimos que u^* es el adjunto de u.

Dem: Si escribimos $B = \langle\cdot,\cdot\rangle : V \times V \to \mathbb{R}$ producto escalar, entonces definimos $\hat{B}: V \to V^*$, $y \mapsto \langle \cdot, y\rangle = \langle y, \cdot\rangle$. Luego, la identidad buscada se escribe como $\hat{B}(y)(u(x)) = \hat{B}(u^*(y))(x)$ o bien (ver §28) ${}^t u \circ \hat{B} = \hat{B} \circ u^*$.

Dado que B es no-degenerada, $\hat{B}$ es un isomorfismo y luego definimos $u^* := \hat{B}^{-1} \circ {}^t u \circ \hat{B}$ como el adjunto de u. En part, $rg(u^*) = rg({}^t u) = rg(u)$ ■

Ejemplo: Sea $u: V \xrightarrow{\sim} V$ automorfismo, ie, $u \in GL(V)$. Entonces, u es una isometría (ie, $u \in O(n)$) $\Leftrightarrow \langle u(x), u(y)\rangle = \langle x, y\rangle \ \forall x, y \in V$

$$\Leftrightarrow \langle x, (u^* \circ u)(y)\rangle = \langle x, y\rangle \ \forall x, y \in V$$

$$\Leftrightarrow u^* \circ u = u \circ u^* = Id_V \Leftrightarrow u^* = u^{-1}.$$

Prop: Sea V un espacio euclídeano y sea $\mathcal{B}$ una base ortonormal de V. Entonces, para todo endomorfismo $u: V \to V$ se tiene que:

$$Mat_{\mathcal{B}}(u^*) = {}^t Mat_{\mathcal{B}}(u).$$

Dem: Si $\mathcal{B} = (e_1, \dots, e_n)$ y escribimos $u(e_j) = \sum_{i=1}^n a_{ij} e_i$ para $j \in \{1, \dots, n\}$, entonces dado que $\mathcal{B}$ es ortonormal tenemos (cf. §35) que: $a_{ij} = \langle u(e_j), e_i\rangle = \langle e_j, u^*(e_i)\rangle$. Del mismo modo, $u^*(e_i) = \sum_{j=1}^n \langle u^*(e_i), e_j\rangle e_j = \sum_{j=1}^n a_{ij} e_j$ y por ende: $u^*(e_j) = \sum_{i=1}^n a_{ji} e_i$ ✓ ■

Ejercicio Sean u y v endomorfismos de un espacio euclídeano V. Probar que $(u \circ v)^* = v^* \circ u^*$ y que $(u^*)^* = u$.

Def: Sea V un espacio euclídeano, y sea $u: V \to V$ un endomorfismo. Decimos que u es simétrico (o auto-adjunto) si $u = u^*$. Es decir,

$$\langle u(x), y\rangle = \langle x, u(y)\rangle \quad \text{para todos } x, y \in V.$$

Obs: Sea $\mathcal{B}$ una base ortonormal de V. Entonces, un endomorfismo $u: V \to V$ es simétrico $\Leftrightarrow A := Mat_{\mathcal{B}}(u)$ es simétrica (ie, $A = {}^t A$).

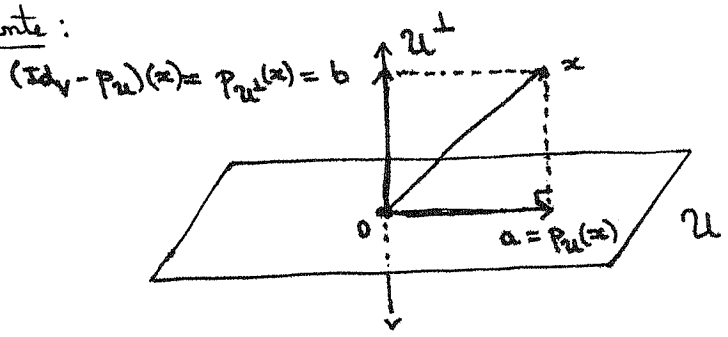
Ejercicio Sea $V \cong \mathbb{R}^n$ espacio euclídeano. Probar que el conjunto $S \subseteq End(V)$ dado por $S = \{u: V \to V \text{ endomorfismo simétrico}\} \subseteq End(V) \cong M_n(\mathbb{R})$ es un sub-ev y que $\dim_{\mathbb{R}}(S) = \frac{n(n+1)}{2}$.

Recuerdo: Sea V un espacio euclideano. Entonces, para todo $\mathcal{U} \subseteq V$ sub-esp se cumple que $V = \mathcal{U} \oplus \mathcal{U}^\perp$. En part, todo vector $x \in V$ se escribe de manera única como $x = a + b$, con $a \in \mathcal{U}$ y $b \in \mathcal{U}^\perp$. Luego, podemos definir

$$P_{\mathcal{U}} : V \to V, \quad x \mapsto a \qquad \text{y} \qquad P_{\mathcal{U}^\perp} : V \to V, \quad x \mapsto b$$

En particular, $P_{\mathcal{U}} + P_{\mathcal{U}^\perp} = Id_V$.

Geométricamente:



Notación: Sea $\mathcal{U} \subseteq V$ sub-esp de un espacio euclideano (o simplemente un esp) V, y sea $x \in V$ un vector. La distancia entre x y $\mathcal{U}$ está definida por:

$$d(x, \mathcal{U}) := \inf_{y \in \mathcal{U}} d(x, y) = \inf_{y \in \mathcal{U}} \|x - y\|.$$

Es decir, $d(x, \mathcal{U})$ es la distancia más pequeña entre $x \in V$ y un vector en $\mathcal{U}$.

Teorema: Sea V un espacio euclideano y $\mathcal{U} \subseteq V$ un sub-esp. Entonces, la aplicación $P_{\mathcal{U}} : V \to V$ definida anteriormente es lineal y cumple $P_{\mathcal{U}}^2 = P_{\mathcal{U}}$, su kernel es $\ker(P_{\mathcal{U}}) = \mathcal{U}^\perp$ y su imagen es $\mathrm{Im}(P_{\mathcal{U}}) = \mathcal{U}$. Más aún, $P_{\mathcal{U}}$ es un endomorfismo simétrico (ie, auto-adjunto: $P_{\mathcal{U}}^* = P_{\mathcal{U}}$) que llamamos la proyección ortogonal sobre $\mathcal{U}$. Además, para todo $x \in V$ y todo $y \in \mathcal{U}$ se tiene:

$$\|x - P_{\mathcal{U}}(x)\| \leq \|x - y\| \iff d(x, \mathcal{U}) = \|x - P_{\mathcal{U}}(x)\|.$$

Dem: Sean $x, x' \in V$ y $\lambda \in \mathbb{R}$. Escribimos $x = a + b$ y $x' = a' + b'$, donde $a, a' \in \mathcal{U}$ y $b, b' \in \mathcal{U}^\perp$. Luego, $x + \lambda x' = \underbrace{(a + \lambda a')}_{\in \mathcal{U}} + \underbrace{(b + \lambda b')}_{\in \mathcal{U}^\perp}$ por lo que:

$P_{\mathcal{U}}(x + \lambda x') = a + \lambda a' = P_{\mathcal{U}}(x) + \lambda P_{\mathcal{U}}(x')$, ie, $P_{\mathcal{U}} : V \to V$ es lineal ✓

Más aún, dado que $V = \mathcal{U} \oplus \mathcal{U}^\perp$, se tiene (por definición de $P_{\mathcal{U}}$) que $\ker(P_{\mathcal{U}}) = \mathcal{U}^\perp$ y que $P_{\mathcal{U}}(a) = a$ si $a \in \mathcal{U}$. En part, $\mathrm{Im}(P_{\mathcal{U}}) = \mathcal{U}$ y

$P_{\mathcal{U}}^2(x) = P_{\mathcal{U}}(P_{\mathcal{U}}(x)) = P_{\mathcal{U}}(a) = a = P_{\mathcal{U}}(x)$, ie, $P_{\mathcal{U}}^2 = P_{\mathcal{U}}$ ✓

Más aún: $\langle P_{\mathcal{U}}(x), x' \rangle = \langle a, x' \rangle = \langle a, a' + b' \rangle = \langle a, a' \rangle$ y, del mismo modo, $\langle x, P_{\mathcal{U}}(x') \rangle = \langle x, a' \rangle = \langle a, a' \rangle \Rightarrow P_{\mathcal{U}}^* = P_{\mathcal{U}}$ (ie, $P_{\mathcal{U}}$ es simétrico).

Finalmente, si $y \in \mathcal{U}$ entonces:

$$\|x - y\|^2 = \|\underbrace{(a - y)}_{\in \mathcal{U}} + \underbrace{b}_{\in \mathcal{U}^\perp}\|^2 \overset{\text{Pitágoras}}{=} \|a - y\|^2 + \|b\|^2 \geq \|b\|^2 \overset{b = x - a}{=} \|x - P_{\mathcal{U}}(x)\|^2 \quad \blacksquare$$

⚠ Observación importante: Sea V un espacio euclideano y $\mathcal{U} \subseteq V$ sub-esv. Si $\dim_{\mathbb{R}}(\mathcal{U}) = r$ y $(e_1, \dots, e_r)$ es una base ortogonal de $\mathcal{U}$, entonces para todo vector $x \in V$ se tiene que:

$$P_{\mathcal{U}}(x) = \langle x, e_1 \rangle \frac{e_1}{\|e_1\|^2} + \dots + \langle x, e_r \rangle \frac{e_r}{\|e_r\|^2}.$$

En efecto, sea $a := \langle x, e_1 \rangle \frac{e_1}{\|e_1\|^2} + \dots + \langle x, e_r \rangle \frac{e_r}{\|e_r\|^2} \in \mathcal{U}$ y sea $b := x - a$. Luego, para todo $j \in \{1, \dots, r\}$ se tiene que:

$$\langle b, e_j \rangle = \langle x - a, e_j \rangle = \langle x, e_j \rangle - \langle a, e_j \rangle = \langle x, e_j \rangle - \frac{\langle x, e_j \rangle}{\|e_j\|^2} \langle e_j, e_j \rangle = 0$$

$\Rightarrow b \in \mathcal{U}^{\perp}$ y luego $P_{\mathcal{U}}(x) = a$ ✓

En particular, si $(e_1, \dots, e_r)$ es una base ortonormal de $\mathcal{U}$, entonces

$$P_{\mathcal{U}}(x) = \langle x, e_1 \rangle e_1 + \dots + \langle x, e_r \rangle e_r.$$

¡Atención!: La fórmula de $P_{\mathcal{U}}(x)$ apareció en el proceso de ortogonalización de Gram-Schmidt. En efecto, usando la misma notación que en §35, teníamos que:

$$e_j'' = e_j - \langle e_j, e_1'' \rangle \frac{e_1''}{\|e_1''\|^2} - \dots - \langle e_j, e_{j-1}'' \rangle \frac{e_{j-1}''}{\|e_{j-1}''\|^2} = e_j - P_{\mathcal{U}_{j-1}}(e_j),$$

donde $\mathcal{U}_{j-1} = \mathrm{Vect}_{\mathbb{R}}(e_1, \dots, e_{j-1}) = \mathrm{Vect}_{\mathbb{R}}(e_1'', \dots, e_{j-1}'')$. Dado que $P_{\mathcal{U}} + P_{\mathcal{U}^{\perp}} = \mathrm{Id}_V$, tenemos que $e_j'' = e_j - P_{\mathcal{U}_{j-1}}(e_j) = (\mathrm{Id}_V - P_{\mathcal{U}_{j-1}})(e_j) = P_{\mathcal{U}_{j-1}^{\perp}}(e_j)$, por lo que e_j'' es la proyección ortogonal de e_j a $\mathcal{U}_{j-1}^{\perp}$ (y en particular e_j'' es ortogonal a $\mathcal{U}_{j-1}$, tal como discutimos en §35).

Def: Sea V un espacio euclideano y $\mathcal{U} \subseteq V$ un sub-esv. Definimos el endomorfismo $S_{\mathcal{U}}: V \to V$ mediante:

$$S_{\mathcal{U}}(x) := 2P_{\mathcal{U}}(x) - x \text{ para todo } x \in V \iff S_{\mathcal{U}} = 2P_{\mathcal{U}} - \mathrm{Id}_V \iff P_{\mathcal{U}} = \frac{S_{\mathcal{U}} + \mathrm{Id}_V}{2},$$

que llamaremos la simetría ortogonal respecto a $\mathcal{U}$. Si $\mathcal{U} = H$ es un hiperplano (ie, $\dim_{\mathbb{R}}(H) = n-1$), decimos que S_H es una reflexión respecto a H.

La terminología anterior está justificada por las siguientes propiedades:

Prop: Sea V un espacio euclideano y $\mathcal{U} \subseteq V$ un sub-esv. Entonces, la simetría ortogonal $S_{\mathcal{U}}: V \to V$ es un automorfismo (ie, $S_{\mathcal{U}} \in GL(V)$) que es ortogonal (ie, $S_{\mathcal{U}} \in O(n)$) y simétrico (ie, $S_{\mathcal{U}}^* = S_{\mathcal{U}}$), y que verifica $S_{\mathcal{U}}^2 = \mathrm{Id}_V$.

Dem: La simetría $S_{\mathcal{U}} := 2P_{\mathcal{U}} - Id_V$ es combinación lineal de los endomorfismos simétricos $P_{\mathcal{U}}$ y Id_V, por lo que $S_{\mathcal{U}}$ es simétrico: $S_{\mathcal{U}}^* = 2P_{\mathcal{U}}^* - Id_V^* = 2P_{\mathcal{U}} - Id_V = S_{\mathcal{U}}$ ✓

Más aún, $S_{\mathcal{U}}^2 = S_{\mathcal{U}} \circ S_{\mathcal{U}} = (2P_{\mathcal{U}} - Id_V)\circ(2P_{\mathcal{U}} - Id_V) = 4\underbrace{P_{\mathcal{U}}^2}_{=P_{\mathcal{U}}} - 4P_{\mathcal{U}} + Id_V = Id_V$ ✓

En part, $\det(S_{\mathcal{U}})^2 = 1 \neq 0$ y luego $S_{\mathcal{U}} \in GL(V)$.

Finalmente, dado que $S_{\mathcal{U}}$ es simétrico: $S_{\mathcal{U}} \circ S_{\mathcal{U}}^* = S_{\mathcal{U}} \circ S_{\mathcal{U}} = Id_V \Rightarrow S_{\mathcal{U}}^* = S_{\mathcal{U}}^{-1}$ y luego $S_{\mathcal{U}}$ es una isometría (ie, $S_{\mathcal{U}} \in O(n)$): ver Ejemplo en pág. 107 ✓ ■

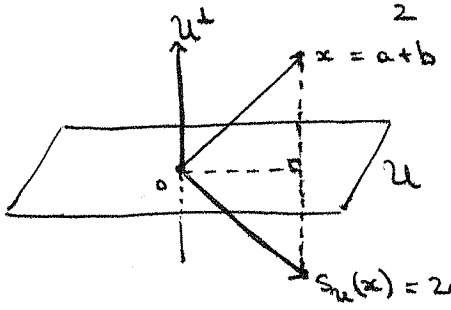
Obs: ① Notar que $P_{\mathcal{U}} + P_{\mathcal{U}^\perp} = Id_V$ implica que:

$$S_{\mathcal{U}^\perp} = 2P_{\mathcal{U}^\perp} - Id_V = 2(Id_V - P_{\mathcal{U}}) - Id_V = Id_V - 2P_{\mathcal{U}} = -S_{\mathcal{U}} \iff \boxed{S_{\mathcal{U}^\perp} = -S_{\mathcal{U}}}$$

② Geométricamente: Dados $x \in V$, el vector $S_{\mathcal{U}}(x)$ está caracterizado por:

(i) $x - S_{\mathcal{U}}(x)$ es ortogonal a $\mathcal{U}$

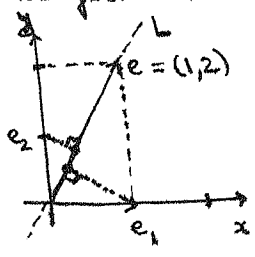
(ii) El "punto medio" $\frac{x + S_{\mathcal{U}}(x)}{2} = P_{\mathcal{U}}(x)$ pertenece a $\mathcal{U}$



$$S_{\mathcal{U}} : V \longrightarrow V$$
$$x = a+b \longmapsto S_{\mathcal{U}}(x) = a-b$$

(cf. conjugación $z \mapsto \bar{z}$ en $\mathbb{C}$).

Ejemplos: ① Sea $L = Vect_{\mathbb{R}}\langle(1,2)\rangle$ recta en $\mathbb{R}^2$. Entonces, gracias a la fórmula de P_L (ver pág 109):



$$P_L((1,0)) = \langle(1,0), e\rangle \frac{e}{\|e\|^2} = \frac{1}{5}e \quad \text{y} \quad P_L((0,1)) = \langle(0,1), e\rangle \frac{e}{\|e\|^2} = \frac{2}{5}e$$

Luego, la matriz de P_L resp. a la base canónica de $\mathbb{R}^2$ es:

$$A = \begin{pmatrix} 1/5 & 2/5 \\ 2/5 & 4/5 \end{pmatrix}$$ y la matriz de S_L es $2A - I_2 = \begin{pmatrix} -3/5 & 4/5 \\ 4/5 & 3/5 \end{pmatrix}$. (simétrica)

Notar que las columnas de $2A - I_2$ son ortogonales (tal como predice la teoría).

Observar que $\det(2A - I_2) = \frac{-9}{25} - \frac{16}{25} = -1$, y luego $S_L \in O(2)$ pero $S_L \notin SO(2)$ (cf. §33).

② Sea $\Pi \subseteq \mathbb{R}^3$ el plano de ecuación $3x - y + 2z = 0$. Para encontrar la matriz de P_Π resp. a la base canónica es más fácil deducirla a partir de $P_{\Pi^\perp}$:

La recta $L := \Pi^\perp$ está generada por $e = (3,-1,2)$ y luego:

$$P_L((1,0,0)) = \frac{3}{14}e,\ P_L((0,1,0)) = \frac{-1}{14}e,\ P_L((0,0,1)) = \frac{2}{14}e \rightsquigarrow \begin{pmatrix} 9/14 & -3/14 & 6/14 \\ -3/14 & 1/14 & -2/14 \\ 6/14 & -2/14 & 4/14 \end{pmatrix} = A$$

Luego, la matriz de $P_\Pi = Id_{\mathbb{R}^3} - P_L$ es $B := I_3 - A = \begin{pmatrix} 5/14 & 3/14 & -3/7 \\ 3/14 & 13/14 & 1/7 \\ -3/7 & 1/7 & 5/7 \end{pmatrix}$

y la matriz de S_Π es $2B - I_3 = \begin{pmatrix} -2/7 & 3/7 & -6/7 \\ 3/7 & 6/7 & 2/7 \\ -6/7 & 2/7 & 3/7 \end{pmatrix}$.

Obs importante: Sea V un espacio euclideano y $\mathcal{U} \subseteq V$ un sub-esv. Entonces:

① $p_{\mathcal{U}}$ es diagonalizable: En efecto, $p_{\mathcal{U}}^2 - p_{\mathcal{U}} = 0$ implica que el polinomio mínimal $m_{p_{\mathcal{U}}}$ divide $P(X) = X^2 - X = X(X-1)$ que escinde sobre $\mathbb{R}$ con raíces simples ✓

Más aún, $V = V_0 \oplus V_1$ donde $V_0 = \ker(p_{\mathcal{U}}) = \mathcal{U}^\perp$ y $V_1 = \ker(p_{\mathcal{U}} - Id_V) = \ker(p_{\mathcal{U}^\perp}) = \mathcal{U}$ $= \mathrm{Im}(p_{\mathcal{U}})$

② $s_{\mathcal{U}}$ es diagonalizable: Similar: $s_{\mathcal{U}}^2 = Id_V$ implica que $m_{s_{\mathcal{U}}}$ divide a $P(X) = X^2 - 1 = (X-1)(X+1)$ que escinde sobre $\mathbb{R}$ con raíces simples ✓

Más aún, $V = V_{-1} \oplus V_1$ donde $V_{-1} = \ker(s_{\mathcal{U}} + Id_V) = \ker(p_{\mathcal{U}}) = \mathcal{U}^\perp$ y donde $V_1 = \ker(s_{\mathcal{U}} - Id_V) = \ker(p_{\mathcal{U}^\perp}) = \mathcal{U}$.

Teorema: Sea V un espacio euclideano. Entonces, todo endomorfismo $u: V \to V$ simétrico (ie, auto-adjunto: $u^* = u$) es diagonalizable en una base ortonormal.

Dem: Por inducción en $n = \dim_{\mathbb{R}}(V)$: Sea $u: V \to V$ endomorfismo simétrico y $\mathcal{B}$ una base ortonormal de V. Entonces, $A = \mathrm{Mat}_{\mathcal{B}}(u)$ es una matriz simétrica. Veamos que u posee sólo valores propios reales:

Sea $\lambda \in \mathbb{C}$ valor propio de A y sea $X \in \mathbb{C}^n \setminus \{0\}$ vector columna tq $AX = \lambda X$.

Dado que $A \in M_n(\mathbb{R})$ es real, tenemos que $\overline{AX} = A\overline{X} = \overline{\lambda}\,\overline{X}$ y luego:

Si escribimos $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ entonces ${}^t X \overline{X} = {}^t\overline{X} X = (\overline{x_1} \cdots \overline{x_n}) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1\overline{x_1} + \ldots + x_n\overline{x_n}$

$= |x_1|^2 + \ldots + |x_n|^2 = \sum_{j=1}^{n} |x_j|^2 > 0$ en $\mathbb{R}$!

$\Rightarrow \overline{\lambda}\, {}^t X \overline{X} = {}^t X(\overline{\lambda}\overline{X}) = {}^t X A \overline{X} \overset{\text{Matriz } 1\times 1}{=} {}^t({}^t X A \overline{X}) = {}^t\overline{X}\, {}^t A X \overset{{}^tA = A}{=} {}^t\overline{X} A X = \lambda\, {}^t\overline{X} X$

($ {}^t X \overline{X} \in \mathbb{R}^{>0}$, $ {}^t\overline{X} X \in \mathbb{R}^{>0}$)

por lo 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. Para todo $y \in H$ tenemos que:

$$\langle u(y), x \rangle = \langle y, u^*(x) \rangle = \langle y, u(x) \rangle = \lambda \langle x, y \rangle = 0.$$

por lo que $u(y) \in L^\perp = H$, ie, $u(H) \subseteq H$. ("H es estable por 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 de inducción prueba que $u|_H$ es diagonalizable en cierta base ortonormal $(e_2, \ldots, e_n)$ de $H = L^\perp$.

$\Rightarrow u$ es diagonalizable en la base ortonormal $\left(\frac{x}{\|x\|}, e_2, \ldots, e_n\right)$ de V. ■

⚠ 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"

(Esto también se puede probar "a la mano": si $\lambda \neq \mu$ y si $x, y \in V \setminus \{0\}$ cumplen $u(x) = \lambda x$ y $u(y) = \mu y \Rightarrow \lambda \langle x,y \rangle = \langle u(x), y \rangle = \langle x, u(y) \rangle = \mu \langle x,y \rangle \underset{\lambda \neq \mu}{\Rightarrow} \langle x,y \rangle = 0$).

Comunmente se dice que V es la <u>suma directa ortogonal</u> de los espacios propios $V_{\lambda_1}, \dots, V_{\lambda_p}$ de u, y escribimos $V = V_{\lambda_1} \overset{\perp}{\oplus} V_{\lambda_2} \overset{\perp}{\oplus} \dots \overset{\perp}{\oplus} V_{\lambda_p} = \overset{p}{\underset{j=1}{\overset{\perp}{\bigoplus}}} V_{\lambda_j}$.

<u>Corolario</u>: Sea $A \in M_n(\mathbb{R})$ matriz <u>simétrica</u> real. Entonces, existe una matriz $P \in Gl_n(\mathbb{R})$ <u>ortogonal</u> (ie, ${}^tPP = I_n \Leftrightarrow P^{-1} = {}^tP$) y una matriz $D \in M_n(\mathbb{R})$ <u>diagonal real</u> tal que $A = PDP^{-1} = PD{}^tP$.

<u>Dem</u>: Sea $u = u_A : \mathbb{R}^n \to \mathbb{R}^n$, $x \mapsto Ax$ el endomorfismo asociado a A. Si $\mathcal{E}$ es la base canónica de $\mathbb{R}^n$, entonces $Mat_{\mathcal{E}}(u) = A$ matriz simétrica $\Rightarrow u^* = u$ ✓

Gracias al teorema anterior, existe $\mathcal{B}$ base ortonormal de $\mathbb{R}^n$ tal que $D = Mat_{\mathcal{B}}(u)$ es diagonal real. Por otra parte, la matriz de cambio de base $P = Mat_{\mathcal{E}}(\mathcal{B})$ es ortogonal y $Mat_{\mathcal{B}}(u) = Mat_{\mathcal{B}}(\mathcal{E})\, Mat_{\mathcal{E}}(u)\, Mat_{\mathcal{E}}(\mathcal{B}) \Leftrightarrow D = P^{-1}AP$. ■

<u>Ejemplo</u>: Sea $A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 3 & 0 \\ 2 & 0 & 1 \end{pmatrix} \in M_3(\mathbb{R})$ matriz simétrica real. Su polinomio característico es: $P_A(X) = \det(XI_3 - A) = \begin{vmatrix} X-1 & 0 & -2 \\ 0 & X-3 & 0 \\ -2 & 0 & X-1 \end{vmatrix} = (X-3)\begin{vmatrix} X-1 & -2 \\ -2 & X-1 \end{vmatrix} = (X-3)(X^2-2X+1-4)$

$= (X-3)(X^2-2X-3) = (X+1)(X-3)^2 \Rightarrow \lambda = -1$ y $\mu = 3$ valores propios.

$\boxed{\lambda = -1}$: $A + I_3 = \begin{pmatrix} 2 & 0 & 2 \\ 0 & 4 & 0 \\ 2 & 0 & 2 \end{pmatrix} \Rightarrow \begin{pmatrix} 2 & 0 & 2 \\ 0 & 4 & 0 \\ 2 & 0 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x+2z \\ 4y \\ 2x+2z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow y = 0,\ x = -z$

Luego $\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$ vector propio. <u>Normalizamos</u>: $v_1 = \begin{pmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{pmatrix}$ genera V_λ y $\|v_1\| = 1$.

$\boxed{\mu = 3}$: $A - 3I_3 = \begin{pmatrix} -2 & 0 & 2 \\ 0 & 0 & 0 \\ 2 & 0 & -2 \end{pmatrix} \Rightarrow \begin{pmatrix} -2 & 0 & 2 \\ 0 & 0 & 0 \\ 2 & 0 & -2 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2z-2x \\ 0 \\ 2x-2z \end{pmatrix} \Rightarrow x - z = 0$.

⚠ Notar que el plano $x - z = 0$ es precisamente el ortogonal de V_λ, ie, $V_\mu = V_\lambda^\perp$! (se podría haber calculado directamente de esta forma).

Luego, V_μ está generado por $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ y $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$. <u>Normalizamos</u>: $v_2 = \begin{pmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{pmatrix}$ y $v_3 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$.

↖ ↑ <u>ortogonales</u>.

¡Atención! En general, se busca una base cualquiera del espacio propio, y se le aplica el proceso de Gram-Schmidt para ortonormalizarla.

Luego, obtenemos que:

$$\underbrace{\begin{pmatrix} 1 & 0 & 2 \\ 0 & 3 & 0 \\ 2 & 0 & 1 \end{pmatrix}}_{A} = \underbrace{\begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \\ -1/\sqrt{2} & 1/\sqrt{2} & 0 \end{pmatrix}}_{P} \underbrace{\begin{pmatrix} -1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}}_{D} \underbrace{\begin{pmatrix} 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 1 & 0 \end{pmatrix}}_{{}^tP = P^{-1}}.$$

§37. Isometrías del plano y orientación

En toda esta sección, a menos que se diga lo contrario, $V \cong \mathbb{R}^2$ será un espacio euclideano de dimensión 2, ie, un plano. Estudiaremos las isometrías de V (ie, elementos de $O(2)$) mirando sus matrices respecto a una base ortonormal:

Sea $V \cong \mathbb{R}^2$ un plano euclideano y sea $\mathcal{B}$ una base ortonormal de V. Consideremos una isometría $u: V \xrightarrow{\sim} V$ (ie, $u \in O(2)$) y sea $A = \mathrm{Mat}_{\mathcal{B}}(u)$. Si escribimos

$$A = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \in GL_2(\mathbb{R})$$

entonces sabemos que $\delta := \det(A) = \pm 1$ y que $A\,{}^tA = I_2$, o bien:

$$A = {}^tA^{-1} = \frac{1}{\delta}\,{}^t\begin{pmatrix} d & -c \\ -b & a \end{pmatrix} = \frac{1}{\delta}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

lo cual equivale a: $d = \delta a$, $c = -\delta b$, por lo que $\delta = ad - bc = \delta a^2 + \delta b^2$

Prop: Sea $A \in GL_2(\mathbb{R})$ la matriz de una isometría indirecta (ie, $\delta = -1$) $u: V \to V$ resp. a una base ortonormal. Entonces, A es de la forma

$$A = \begin{pmatrix} a & b \\ b & -a \end{pmatrix} \quad \text{donde } a^2 + b^2 = 1.$$

Más aún, existe una única recta $L \subseteq V$ (ie, $\dim_{\mathbb{R}}(L) = 1$) tal que $u = s_L$ es la reflexión respecto a L.

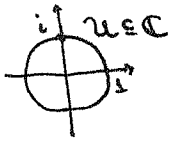
Dem: Si $\delta = -1$ entonces $d = -a$ y $c = +b$, donde $a^2 + b^2 = 1$. Notar que ${}^tA = A$ en este caso y luego $A\,{}^tA = I_2 \iff A^2 = I_2$. Luego, los valores propios de u son 1 y -1, y los vectores propios asociados e_1 y e_{-1} son ortogonales (cf. pág 111). Así, $u = s_L$ donde $L = V_1 = \mathrm{Vect}_{\mathbb{R}}(e_1)$. ■

Corolario: Toda isometría de un plano euclideano es el producto de a lo más dos reflexiones.

Dem: Sea $V \cong \mathbb{R}^2$ plano euclideano y $u: V \xrightarrow{\sim} V$ isometría. Si $\det(u) = -1$ entonces $u = s$ es una reflexión ✓ Si $\det(u) = 1$ y $s: V \to V$ es una reflexión arbitraria, entonces $\det(s \circ u) = \underbrace{\det(s)}_{=-1}\underbrace{\det(u)}_{=1} = -1$ y luego $s \circ u$ es una reflexión. Finalmente, $u = s \circ (s \circ u)$, pues $s^2 = \mathrm{Id}_V$ ✓ ■

Notación: Sea $\mathcal{U} \subseteq \mathbb{C}$ el círculo de radio 1. Explícitamente:

$$\mathcal{U} = \{z \in \mathbb{C} \text{ tq } |z| = 1\} = \{a + ib \in \mathbb{C} \text{ tq } a, b \in \mathbb{R} \text{ y } a^2 + b^2 = 1\}$$



Entonces, $\mathcal{U}$ es un grupo abeliano respecto a la multiplicación en $\mathbb{C}$.

Prop: Sea $A \in GL_2(\mathbb{R})$ la matriz de una isometría directa (ie, $\delta = 1$) $u: V \to V$ resp. a una base ortonormal. Entonces, A es de la forma

$$A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \quad \text{donde } a^2 + b^2 = 1.$$

Más aún, $\varphi: \mathcal{U} \xrightarrow{\sim} SO(2)$, $a + ib \mapsto \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ es un isomorfismo de grupos. En part, $SO(2)$ es un grupo abeliano.

Dem: Si $\delta = 1$ entonces $d = +a$ y $c = -b$, donde $a^2 + b^2 = 1$. Luego, $\varphi: \mathcal{U} \to SO(2)$, $a+ib \mapsto \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ es biyectiva. Veamos que respeta las estructuras de grupo (ie, $\varphi(zz') = \varphi(z)\varphi(z')$): si $z = a+ib$ y $z' = a'+ib'$ entonces $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}\begin{pmatrix} a' & -b' \\ b' & a' \end{pmatrix} = \begin{pmatrix} aa'-bb' & -ab'-a'b \\ a'b+ab' & aa'-bb' \end{pmatrix}$, mientras que

$zz' = (a+ib)(a'+ib') = (aa'-bb') + i(ab'+a'b)$ ✓ ■

Def: Sea V un plano euclidiano. Los elementos de $SO(2)$ (ie, isometrías de determinante 1) serán llamados **rotaciones**. En part, dado que $SO(2) \cong \mathcal{U}$ es un grupo abeliano, si $r_1, r_2 \in SO(2)$ son rotaciones entonces $r_1 \circ r_2 = r_2 \circ r_1$.

Obs importante: Sea $r: V \xrightarrow{\sim} V$ una rotación de $V \cong \mathbb{R}^2$ y sea $\mathcal{B}$ una base ortonormal. Entonces, $R = Mat_{\mathcal{B}}(r)$ es de la forma

$$R = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \quad \text{donde } a^2 + b^2 = 1.$$

Notamos la similitud con la matriz de rotación (cf. pág 28) $R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$. Nos gustaría asociar a r un "ángulo" $\theta \in \mathbb{R}$ tq $R = R_\theta$, ie, $a = \cos\theta$ y $b = \sin\theta$:

Notamos que $2a = tr(R)$ y luego $a = \frac{1}{2}tr(r)$ es **independiente** de $\mathcal{B}$ (ie, está definido **intrínsecamente** por $r: V \to V$). **Sin embargo**, el coeficiente b está determinado por $a^2 + b^2 = 1$ y luego se requiere **escoger** un signo!

(Idea: $b = \sin\theta$ si percibe el "**signo**" del ángulo θ). Explícitamente: Sea $\mathcal{B}'$ otra base ortonormal de V y sea $P = Mat_{\mathcal{B}}(\mathcal{B}')$. ($\Rightarrow {}^tPP = I_2$).

Caso 1: $\det(P) = 1$, ie, $P \in SO(2)$:

Dado que $SO(2)$ es conmutativo, $PR = RP$ en este caso. Luego:

$R' := Mat_{\mathcal{B}'}(r) = P^{-1}RP = P^{-1}PR = R$.

Caso 2: $\det(P) = -1$, ie, P es una **reflexión**:

Notar que $\det(PR) = -1$, y luego P y PR son reflexiones $\Rightarrow P^2 = (PR)^2 = I_2$.

Luego, $R' = P^{-1}RP = PRP = \underbrace{(PRPR)}_{=I_2}R^{-1} = R^{-1} = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$.

"El cambio de base **indirecto** invierte el "ángulo" de rotación"

Conclusión: Una vez **fijada** una base ortonormal, hay 2 tipos de cambio de base: los que preservan el sentido de rotación (caso 1) y los que lo invierten (caso 2).

Def/Prop (orientación): Sea $V \cong \mathbb{R}^n$ un $\mathbb{R}$-ev. Definimos la relación de equivalencia siguiente en el conjunto de bases de V:

$$\mathcal{B} \sim \mathcal{B}' \overset{\text{def}}{\iff} \det_{\mathcal{B}}(\mathcal{B}') > 0.$$

Hay exactamente **dos** clases de equivalencia, cada una llamada una **orientación** de V.

Dem: Sabemos que $\det_{\mathcal{B}}(\mathcal{B}) = 1$ y luego $\mathcal{B} \sim \mathcal{B}$ (reflexiva). Sean $\mathcal{B}, \mathcal{B}', \mathcal{B}''$ bases de $V \Rightarrow \mathrm{Mat}_{\mathcal{B}}(\mathcal{B}'') = \mathrm{Mat}_{\mathcal{B}}(\mathcal{B}')\,\mathrm{Mat}_{\mathcal{B}'}(\mathcal{B}'') \Rightarrow \det_{\mathcal{B}}(\mathcal{B}'') = \det_{\mathcal{B}}(\mathcal{B}')\det_{\mathcal{B}'}(\mathcal{B}'')$ y luego $\det_{\mathcal{B}}(\mathcal{B}') > 0$ y $\det_{\mathcal{B}'}(\mathcal{B}'') > 0$ implica $\det_{\mathcal{B}}(\mathcal{B}'') > 0$ (transitiva). En part, si $\mathcal{B}'' = \mathcal{B}$ entonces $\det_{\mathcal{B}}(\mathcal{B}')\det_{\mathcal{B}'}(\mathcal{B}) = 1$ por lo que $\det_{\mathcal{B}}(\mathcal{B}') > 0 \Leftrightarrow \det_{\mathcal{B}'}(\mathcal{B}) > 0$ (simétrica). Luego, $\sim$ es una relación de equivalencia ✓

Sea $\mathcal{B}_0 = (e_1, \dots, e_n)$ una base de V y sea $\mathcal{E}_0 := (-e_1, e_2, \dots, e_n)$ base de V. $\Rightarrow \det_{\mathcal{B}_0}(\mathcal{E}_0) = -1$ y luego $\mathcal{B}_0 \not\sim \mathcal{E}_0$ no están relacionadas. Veamos que $[\mathcal{B}_0]$ y $[\mathcal{E}_0]$ son las únicas clases: Sea $\mathcal{B}$ una base cualquiera de V. Si $\det_{\mathcal{B}_0}(\mathcal{B}) > 0$ entonces $\mathcal{B} \in [\mathcal{B}_0]$, y si $\det_{\mathcal{B}_0}(\mathcal{B}) < 0$ entonces:

$$\det_{\mathcal{E}_0}(\mathcal{B}) = \underbrace{\det_{\mathcal{E}_0}(\mathcal{B}_0)}_{=-1}\underbrace{\det_{\mathcal{B}_0}(\mathcal{B})}_{<0} > 0 \Rightarrow \mathcal{B} \in [\mathcal{E}_0] \quad ✓ \quad \blacksquare$$

Obs: Sea $V \cong \mathbb{R}^n$ un $\mathbb{R}$-ev, entonces "elegir una orientación de V" es escoger una base $\mathcal{B}_0$ de V y la orientación que le corresponde, ie, todas las bases $\mathcal{B}$ de V tal que $\det_{\mathcal{B}_0}(\mathcal{B}) > 0$; las cuales llamaremos bases directas, mientras que las otras serían bases indirectas. Diremos que V es un espacio orientado.

⚠ Por convención, orientaremos canonicamente a $\mathbb{R}^n$ declarando que la base canónica es directa: es la orientación canónica de $\mathbb{R}^n$.

Def: Sea $V \cong \mathbb{R}^2$ un plano euclideano orientado y sea $r: V \xrightarrow{\sim} V$ una rotación de V (ie, $r \in SO(2)$). Entonces (por la discusión anterior), para toda base directa ortonormal $\mathcal{B}$ de V existe $\theta \in \mathbb{R}$ (único módulo $2\pi\mathbb{Z}$) tal que

$$R = \mathrm{Mat}_{\mathcal{B}}(r) = R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Llamamos a θ el ángulo de rotación de r y escribiremos $r = r_\theta$.

Ejercicio: Sea V un plano euclideano orientado y sean $\theta, \theta' \in \mathbb{R}$. Verifican que las rotaciones r_θ y $r_{\theta'}$ cumplen $r_\theta \circ r_{\theta'} = r_{\theta+\theta'}$, $r_0 = \mathrm{Id}_V$ y $r_\pi = -\mathrm{Id}_V$.

Lema: Sea $V \cong \mathbb{R}^2$ un plano euclideano, y sean $x, y \in V \setminus \{0\}$ no-nulos tal que $\|x\| = \|y\|$. Entonces, existe una única rotación $r: V \to V$ tq $r(x) = y$.

Dem: Sea $e_1 = \frac{x}{\|x\|}$ vector unitario. Sea $e_2 \in V$ tq $\mathcal{B} = (e_1, e_2)$ es una base ortonormal (eg. completar e_1 en una base y usar Gram-Schmidt). Si escribimos $\frac{y}{\|y\|} = a e_1 + b e_2$ entonces $a^2 + b^2 = 1$. Sea r la rotación de matriz

$$R = \mathrm{Mat}_{\mathcal{B}}(r) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}.$$

$\Rightarrow r(e_1) = a e_2 + b e_2 \Leftrightarrow r\left(\frac{x}{\|x\|}\right) = \frac{y}{\|y\|} \underset{\|x\|=\|y\|}{\Leftrightarrow} r(x) = y$ ✓ Si r' es otra rotación tal que $r'(x) = y \Rightarrow (r'^{-1} \circ r)(x) = x$ y luego $\lambda_1 = 1$ es valor propio de $r'' = r'^{-1} \circ r$. Como $\det(r'') = 1 = \underbrace{\lambda_1}_{=1}\lambda_2 \Rightarrow \lambda_2 = 1$ (2^{do} valor propio) $\Rightarrow r'' = \mathrm{Id}_V \Leftrightarrow r = r'$ $\blacksquare$

Def: Sea $V \cong \mathbb{R}^2$ un plano euclideano orientado y sean $x, y \in V \setminus \{0\}$ no-nulos. Sea $r = r_\theta$ la única rotación que envía $\frac{x}{\|x\|}$ en $\frac{y}{\|y\|}$. Definimos el ángulo orientado en x, y como

$$\measuredangle(x,y) := \theta ,$$

donde $\theta \in \mathbb{R}$ es único módulo $2\pi\mathbb{Z}$. En part, único si suponemos $\theta \in]-\pi, \pi]$.

Observación: ① Gracias a la demostración del lema anterior:

$$\frac{y}{\|y\|} = a e_1 + b e_2 \Rightarrow a = \cos \measuredangle(x,y) = \left\langle e_1, \frac{y}{\|y\|} \right\rangle = \frac{\langle x, y\rangle}{\|x\|\|y\|} \Rightarrow \boxed{\cos(\measuredangle(x,y)) = \frac{\langle x,y\rangle}{\|x\|\|y\|}}.$$

$$b = \det \begin{vmatrix} 1 & a \\ 0 & b \end{vmatrix} = \det_{\mathcal{B}}\left(e_1, \frac{y}{\|y\|}\right) = \frac{\det_{\mathcal{B}}(x,y)}{\|x\|\|y\|} = \underbrace{\det_{\mathcal{B}}(\mathcal{B}')}_{=1} \frac{\det_{\mathcal{B}'}(x,y)}{\|x\|\|y\|}$$ donde $\mathcal{B}'$ es cualquier base ortonormal directa $\Rightarrow$ $\boxed{\operatorname{sin}(\measuredangle(x,y)) = \frac{\det_{\mathcal{B}}(x,y)}{\|x\|\|y\|}}$.

② Si $-\pi < \measuredangle(x,y) \leq \pi$, el ángulo no-orientado (cf §34) entre x e y es $\theta(x,y) = |\measuredangle(x,y)|$.

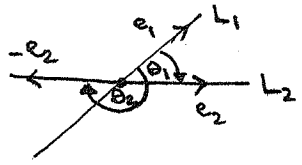
③ Si $r \in SO(2)$ es una rotación, entonces $(r \circ r_\theta)(x) = r(y)$ y como $r_\theta \circ r = r_\theta \circ r$ (pues $SO(2) \cong \mathcal{U}$ conmutativo) se tiene $(r \circ r_\theta)(x) = r_\theta(r(x))$ para todo $x, y \in V$ con $r_\theta(x) = y$. En part, $\measuredangle(r(x), r(y)) = \measuredangle(x,y) \mod 2\pi\mathbb{Z}$.

Ejercicio ① Sea $S: V \to V$ una reflexión. Probar que $\measuredangle(S(x), S(y)) = -\measuredangle(x,y) \mod 2\pi\mathbb{Z}$

② Sean $x, y, z \in V \setminus \{0\}$ no-nulos. Probar que

$\measuredangle(x,y) + \measuredangle(y,z) = \measuredangle(x,z) \mod 2\pi\mathbb{Z}$, $\measuredangle(y,x) = \measuredangle(x,-y) \mod 2\pi\mathbb{Z}$, y que $\measuredangle(x,-y) = \measuredangle(x,y) + \pi \mod 2\pi\mathbb{Z}$. [Indicación: Usar Ejercicio en pág. 115].

Obs: La última relación implica que si L_1 y L_2 son rectas en V, generadas por vectores unitarios e_1 y e_2, respectivamente, entonces módulo $2\pi\mathbb{Z}$ se tiene $\measuredangle(e_1, e_2) = \measuredangle(e_1, -e_2) + \pi$, y luego $\measuredangle(e_1, e_2) = \measuredangle(e_1, -e_2)$ módulo $\pi\mathbb{Z}$.



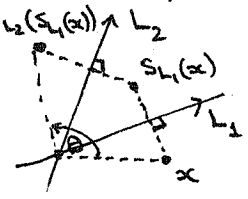
$\theta_1 \equiv \theta_2 \pmod{\pi\mathbb{Z}}$. En part, $\measuredangle(e_1, e_2) \mod \pi\mathbb{Z}$ sólo depende de L_1 y L_2 !

Def: Sean $L_1, L_2 \subseteq V$ dos rectas en el plano euclideano orientado V, y sean e_1 y e_2 vectores directores unitarios. Definimos el ángulo entre L_1 y L_2 como $\measuredangle(L_1, L_2) := \measuredangle(e_1, e_2) \mod \pi\mathbb{Z}$. En part, $\measuredangle(L_1, L_2)$ es único si lo suponemos en $[0, \pi[$.

Prop: Sean $L_1, L_2 \subseteq V$ dos rectas en un plano euclideano orientado V. Entonces,

$$S_{L_2} \circ S_{L_1} = r_{2\measuredangle(L_1, L_2)}.$$

Dem: Sean e_1 y e_2 vectores directores unitarios. Dado que $\det(s_{L_2} \circ s_{L_1}) = (-1)^2 = 1$, sabemos que $s_{L_2} \circ s_{L_1} = r_\theta \in SO(2)$ es una rotación. Basta calcular θ:



Si $x = e_1$: $\theta = \measuredangle(e_1, s_{L_2}(s_{L_1}(e_1))) = \measuredangle(e_1, s_{L_2}(e_1))$

$= \measuredangle(e_1, e_2) + \measuredangle(e_2, s_{L_2}(e_1)) = \measuredangle(e_1, e_2) + \measuredangle(s_{L_2}(e_2), s_{L_2}(e_1))$

(Ejercicio anterior)

$= \measuredangle(e_1, e_2) - \measuredangle(e_2, e_1) = 2\measuredangle(e_1, e_2)$ ■

§38. Isometrías del espacio

En toda esta sección $V \cong \mathbb{R}^3$ será un espacio euclídeano de dimensión 3. Para describir los elementos de $O(3)$ usaremos argumentos más geométricos.

Teorema: Sea $V \cong \mathbb{R}^3$ un espacio euclídeano y sea $u: V \xrightarrow{\sim} V$ una isometría de V (ie, $u \in O(3)$). Entonces, existe una base ortonormal $\mathcal{B}$ de V tal que $A = \mathrm{Mat}_{\mathcal{B}}(u)$ es de la forma

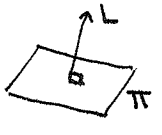
$$A = \begin{pmatrix} \delta & 0 & 0 \\ 0 & a & -b \\ 0 & b & a \end{pmatrix}$$

donde $a^2 + b^2 = 1$ y $\delta = \pm 1$ es el determinante de u.

Dem: El polinomio característico $P_u(X) \in \mathbb{R}[X]$ es de grado 3, por lo que posee necesariamente una raíz real $\lambda \in \mathbb{R}$ (por continuidad). Sea x un vector propio asociado al valor propio λ, entonces:

$$\|x\| \underset{u \in O(3)}{=} \|u(x)\| = \|\lambda x\| = |\lambda| \, \|x\| \underset{x \neq 0}{\Longrightarrow} |\lambda| = 1, \text{ ie, } \lambda = \pm 1.$$

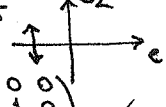
Caso 1 ($\lambda = 1$): En este caso, existe $x \neq 0$ (que podemos suponer unitario, ie, $\|x\| = 1$) tal que $u(x) = x$. Sea $L = \mathrm{Vect}_{\mathbb{R}}(x)$ recta, estable por u: $u(L) = L$. Sea $\Pi := L^\perp$ plano ortogonal a L, estable por u pues $u(L^\perp) = u(L)^\perp$ ✓



$\Pi \cong \mathbb{R}^2$ es un plano euclídeano y podemos aplicar los resultados de la sección anterior. Sea $u_\Pi := u|_\Pi : \Pi \xrightarrow{\sim} \Pi$ isometría. Notar que $\det(u) = \lambda \det(u_\Pi) = \det(u_\Pi)$ en este caso!

Subcaso (a): Si $\delta = \det(u) = \det(u_\Pi) = -1$ entonces $u_\Pi \in O(2)$ es una reflexión. En part, $\exists \mathcal{B}'$ base ortonormal de $\Pi \cong \mathbb{R}^2$ formada por vectores propios de u_Π

$\Rightarrow B = \mathrm{Mat}_{\mathcal{B}'}(u_\Pi) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, con $\mathcal{B}' = (e_1, e_2)$



Si $\mathcal{B} := (e_2, e_1, x)$ base ortonormal $\Rightarrow A = \mathrm{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ ✓

Subcaso (b): Si $\delta = +1$ entonces $u_\Pi \in SO(2)$ es una rotación. En part, $\exists \mathcal{B}'$ base ortonormal de $\Pi \cong \mathbb{R}^2$ tq $B = \mathrm{Mat}_{\mathcal{B}'}(u_\Pi) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ con $a^2 + b^2 = 1$.

Si $\mathcal{B} := (x, e_1, e_2)$ base ortonormal $\Rightarrow A = \mathrm{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & a & -b \\ 0 & b & a \end{pmatrix}$ con $a^2 + b^2 = 1$ ✓

Caso 2 ($\lambda = -1$): Considerar $-u \in O(3)$ que posee a 1 como valor propio y cumple $\det(-u) = (-1)^3 \det(u) = -\delta$. Luego, basta aplicar el Caso 1 a $-u$ ✓ ■

Obs: Notamos que $\mathrm{tr}(u) = 2a + \delta$, por lo que $\delta = \det(u)$ y $a = \frac{1}{2}(\mathrm{tr}(u) - \det(u))$ están definidos intrínsecamente por $u: V \to V$. Sin embargo, b está determinado por $a^2 + b^2 = 1$ (hay que escoger un signo).

Prop: Sea $V \cong \mathbb{R}^3$ espacio euclídeano y sea $u \in SO(3)$ isometría directa (ie, $\det(u) = 1$). Entonces, $-1 \leq \mathrm{tr}(u) \leq 3$ y además $\mathrm{tr}(u) = 3 \Leftrightarrow u = \mathrm{Id}_V$. Si $\mathrm{tr}(u) < 3$ entonces existe una única recta $L \subseteq V$ invariante por u (ie, $u(L) = L$) tq la restricción de u al plano ortogonal $\Pi := L^\perp$ es una rotación (ie, $u|_\Pi \in SO(2)$). Decimos en ese caso que u es una rotación de eje L.

Dem: Sea $\mathcal{B}$ base ortonormal de V tq $A = \mathrm{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} \delta & 0 & 0 \\ 0 & a & -b \\ 0 & b & a \end{pmatrix}$, con $\delta = 1$ y $a^2 + b^2 = 1$ ($\Rightarrow -1 \leq a \leq 1$). Luego:

$\mathrm{tr}(u) = 1 + 2a \Rightarrow 1 - 2 = -1 \leq \mathrm{tr}(u) \leq 3 = 1 + 2$ y $\mathrm{tr}(u) = 3 \Leftrightarrow a = 1$ ($\Rightarrow b = 0$) $\Leftrightarrow u = \mathrm{Id}_V$ ✓

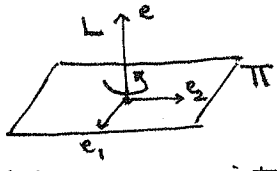
Notamos que $P_u(X) = (X - \delta)(X^2 - 2aX + 1) = (X-1)(X^2 - (\mathrm{tr}(u) - 1)X + 1)$, y que $\lambda = 1$ es raíz de $X^2 - (\mathrm{tr}(u) - 1)X + 1 \Leftrightarrow 1 - \mathrm{tr}(u) + 1 + 1 = 3 - \mathrm{tr}(u) = 0 \Leftrightarrow \mathrm{tr}(u) = 3$.

Luego, si $\mathrm{tr}(u) < 3$, entonces $V_1 = \ker(u - \mathrm{Id}_V) =: L$ es una tq $u|_L = \mathrm{Id}_L$.

Sea $\Pi := L^\perp$ plano ortogonal. Entonces $u(\Pi) = \Pi$ y $\det(u) = \det(u|_\Pi) = 1$, por lo que $u|_\Pi : \Pi \xrightarrow{\sim} \Pi$ es una rotación del plano euclídeano $\Pi \cong \mathbb{R}^2$ ■

Obs: Si orientamos $\Pi \cong \mathbb{R}^2$ entonces $u|_\Pi = r_\theta$ para un único $\theta \in \mathbb{R}$ (mod $2\pi\mathbb{Z}$) que verifica $\mathrm{tr}(u) = 2\cos(\theta) + 1$.

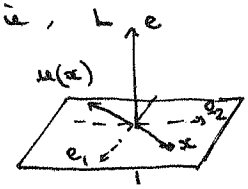
⚠ Atención: No basta con orientar $V \cong \mathbb{R}^3$ para determinar θ módulo $2\pi\mathbb{Z}$. Además, hay que orientar el plano $\Pi = L^\perp$.



Convención: Una vez fijada una orientación de $V \cong \mathbb{R}^3$, diremos que una base (e_1, e_2) de Π es directa si la base $\mathcal{B} = (e, e_1, e_2)$ de V es directa, donde $e \neq 0$ es un vector director de L. Así, orientar L eligiendo un vector no-nulo ("arriba" o "abajo") induce una orientación natural de Π ("anti-horario" u "horario"): esto último es la "regla de la mano derecha".

Ejemplo importante: Sea $u \in SO(3)$ rotación de ángulo π, ie, L, e

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

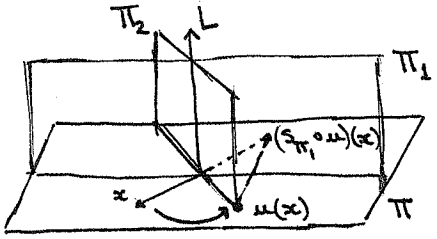


y luego u es una simetría respecto a una recta L, y está caracterizada por las igualdades $\det(u) = 1$ y $\mathrm{tr}(u) = -1$.

Ejercicio Probar que $SO(3)$ no es conmutativo. [Indicación: Considerar $V = \mathbb{R}^3$ y r_1 (resp. r_2) la rotación del plano XY (resp. YZ) en $\frac{\pi}{2}$ y probar que $r_1 r_2 \neq r_2 r_1$].

Corolario: Sea $V \cong \mathbb{R}^3$ un espacio euclideano y sea $u \in SO(3)$ una isometría directa. Entonces, u es el producto de dos reflexiones.

Dem: Si $u = Id_V$ entonces $Id_V = S \circ S = S^2$ para toda reflexión $S: V \to V$ ✓ Si $u \neq Id_V$ entonces u es una rotación de eje L, y la restricción a $\Pi = L^\perp$ es una rotación del plano euclideano $\Pi \cong \mathbb{R}^2$, por lo que $u|_\Pi$ es producto de dos reflexiones (ver pág. 113). Explícitamente:



Sea Π_1 un plano conteniendo L, entonces $S_{\Pi_1} \circ u$ es una reflexión respecto a un plano Π_2 conteniendo L.

$\Rightarrow S_{\Pi_1} \circ u = S_{\Pi_2} \Rightarrow u = S_{\Pi_1}^{-1} \circ S_{\Pi_2} = S_{\Pi_1} \circ S_{\Pi_2}$ ■

Prop: Sea $V \cong \mathbb{R}^3$ un espacio euclideano y sea $u \in O(3)$ una isometría indirecta (u, $\det(u) = \delta = -1$). Entonces, existe una recta $L \subseteq V$ invariante por u (u, $u(L) = L$) tal 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 $\Pi = L^\perp$. En particular, u es producto de tres reflexiones.

Dem: Sea $\mathcal{B} = (e_1, e_2, e_3)$ base ortonormal de V tal que $A = Mat_{\mathcal{B}}(u) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & a & -b \\ 0 & b & a \end{pmatrix}$ con $a^2 + b^2 = 1$. Notar que:

$$\underbrace{\begin{pmatrix} -1 & 0 & 0 \\ 0 & a & -b \\ 0 & b & a \end{pmatrix}}_{A} = \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & a & -b \\ 0 & b & a \end{pmatrix}}_{R} \underbrace{\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{S} = \underbrace{\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{S} \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & a & -b \\ 0 & b & a \end{pmatrix}}_{R}.$$

Si $L := Vect_{\mathbb{R}}(e_1)$ entonces $\Pi = Vect_{\mathbb{R}}(e_2, e_3)$ y $R = Mat_{\mathcal{B}}(r_L)$ es una rotación de eje L, mientras que $S = Mat_{\mathcal{B}}(S_\Pi)$ es la reflexión respecto a Π ✓

Vimos en el corolario anterior que r_L es el producto de dos reflexiones ■

⚠ Vimos en §37 que todo elemento de $O(2)$ es producto de a lo más 2 reflexiones, y acabamos de probar que todo elemento de $O(3)$ es producto de a lo más 3 reflexiones. Veremos que todo elemento de $O(n)$ es producto de a lo más n reflexiones!

Terminaremos esta sección dando una definición formal del "producto cruz" (o "producto vectorial") estudiado en 1er año, y que será útil para determinar ángulos de rotación de elementos de $SO(3)$:

Recuerdo: Sea $V \cong \mathbb{R}^3$ un espacio euclideano orientado y sea $\mathcal{B}$ una base ortonormal directa de V. La aplicación lineal $f: V \to \mathbb{R}$ dada por

$$z \mapsto f(z) = \det_{\mathcal{B}}(x, y, z),$$

donde $x, y \in V$ son dos vectores fijos, no depende de la base ortonormal directa, pues si $\mathcal{B}'$ es otra base ortonormal directa, entonces $\det_{\mathcal{B}'}(\mathcal{B}) = 1$.

Luego, si escribimos $B = \langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ producto escalar, entonces $\widehat{B} : V \xrightarrow{\sim} V^*$ dada por $\widehat{B}(y) = \langle \cdot, y \rangle = \langle y, \cdot \rangle$ es un isomorfismo. En part, dados $x, y \in V$, existe un único vector en V asociado a $f : \widehat{B}^{-1}(f)$.

Def: Sea $V \cong \mathbb{R}^3$ un espacio euclideano orientado, y sean $x, y \in V$ dos vectores. El producto cruz entre x e y es el único vector $x \times y \in V$ tal que

$$\det_B(x, y, z) = \langle x \times y, z \rangle \quad \text{para todo } z \in V,$$

donde B es cualquier base ortonormal directa de V.

Prop: Sea $V \cong \mathbb{R}^3$ un espacio euclideano orientado. Entonces:

① La aplicación $V \times V \to V$, $(x, y) \mapsto x \times y$ es bilineal alternada.

② El vector $x \times y$ es ortogonal a x e y; además $x \times y = 0 \Leftrightarrow x$ e y son l.d. (i.e. colineales). 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

$$\|x \times y\| = \|x\| \|y\| |\sin \theta|.$$

Dem: ① Para todo $z \in V$ se tiene que

$$\langle y \times x, z \rangle \underset{\text{def}}{=} \det_B(y, x, z) = -\det_B(x, y, z) = -\langle x \times y, z \rangle \Rightarrow y \times x = -x \times y.$$

② El cálculo anterior implica que $\langle x \times y, x \rangle = \det_B(x, y, x) = 0$ y que $\langle x \times y, y \rangle = \det_B(x, y, y) = 0$, por lo que $x \times y$ es ortogonal a x e y. Si x e y no son colineales, podemos completar (x, y) en una base (x, y, z) de V y luego $\langle x \times y, z \rangle = \det_B(x, y, z) \neq 0 \Rightarrow x \times y \neq 0$. Por otro lado,

$$\|x \times y\|^2 = \langle x \times y, x \times y \rangle \underset{\text{def}}{=} \det_B(x, y, x \times y)$$

por lo que si x e y son colineales entonces $\|x \times y\|^2 = 0 \Rightarrow x \times y = 0$, y además si x e y no son colineales $\det_B(x, y, x \times y) > 0$ y luego la base $(x, y, x \times y)$ es directa.

③ Supongamos primero que x e y son unitarios y ortogonales (i.e. $\|x\| = \|y\| = 1$ y $\langle x, y \rangle = 0$) y completemos (x, y) en una base ortonormal directa $B = (x, y, e)$. Luego, para todo $z \in V$ se tiene $z = \langle z, x \rangle x + \langle z, y \rangle y + \langle z, e \rangle e$

$\Rightarrow \det_B(x, y, z) = \langle z, e \rangle \underbrace{\det_B(x, y, e)}_{=1} = \langle e, z \rangle \Rightarrow e = x \times y$. y además $\|e\|^2 = \|x \times y\|^2 = 1$ ✓ Notar que $\langle x, y \rangle = 0 \Rightarrow \theta(x, y) = \frac{\pi}{2} \Rightarrow |\sin(\theta)| = 1$, por lo que $1 = \|x \times y\| = \|x\| \|y\| |\sin \theta|$ en este caso ✓

Si x e y son sólamente ortogonales, la bilinealidad del producto cruz implica $1 = \left\| \frac{x}{\|x\|} \times \frac{y}{\|y\|} \right\| \Leftrightarrow \|x \times y\| = \|x\| \|y\|$ ✓

Si x e y son sólamente unitarios, entonces x y $x\langle x, y \rangle - y$ son ortogonales.

Entonces, $\|x\times y\| \underset{x\times x=0}{=} \|x\times(x\langle x,y\rangle - y)\| \underset{\text{caso anterior}}{=} \|x\|\,\|x\langle x,y\rangle - y\| = \|x\langle x,y\rangle - y\|$, mientras que $\|x\langle x,y\rangle - y\|^2 = \langle x,y\rangle^2 - 2\langle x,y\rangle^2 + \underbrace{\|y\|^2}_{=1} = 1-\langle x,y\rangle^2 = 1-\cos^2(\theta) = \sin^2(\theta)$

$\Rightarrow \|x\times y\| = |\sin\theta|$ en este caso ✓ Finalmente, el caso general se deduce por bilinealidad del producto cruz: $\left\|\frac{x}{\|x\|}\times\frac{y}{\|y\|}\right\| = |\sin\theta| \iff \|x\times y\| = \|x\|\|y\||\sin\theta|$ ■

Obs importante: ① Sea $V\cong\mathbb{R}^3$ espacio euclideano orientado y sea $\mathcal{B}$ una base ortonormal directa. Si $(x_1,x_2,x_3)\in\mathbb{R}^3$ (resp. $(y_1,y_2,y_3)\in\mathbb{R}^3$) son las coordenadas de $x\in V$ (resp. $y\in V$) respecto a $\mathcal{B}$, entonces:

Las coordenadas de $x\times y\in V$ son dadas por los determinantes

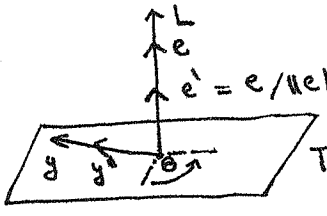
$$\left(\begin{vmatrix} x_2 & y_2\\ x_3 & y_3\end{vmatrix}, -\begin{vmatrix} x_1 & y_1\\ x_3 & y_3\end{vmatrix}, \begin{vmatrix} x_1 & y_1\\ x_2 & y_2\end{vmatrix}\right)\in\mathbb{R}^3$$

En efecto, ellos son los coeficientes de la forma lineal

$$\det_{\mathcal{B}}(x,y,z) = \begin{vmatrix} x_1 & y_1 & z_1\\ x_2 & y_2 & z_2\\ x_3 & y_3 & z_3\end{vmatrix}$$

obtenida al desarrollar la última columna.

② Sea $V\cong\mathbb{R}^3$ espacio euclideano orientado y sea $u = r_L\in SO(3)$ una rotación de eje L. Orientamos la recta L al escoger un vector director $e\in L$ no-nulo



Luego, el plano $\Pi = L^\perp$ posee una orientación "inducida" y $u|_\Pi = r_\theta\in SO(2)$ para un <u>único</u> ángulo $\theta\in\,]-\pi,\pi]$

Sea $y\in V\setminus\{0\}$ vector no nulo ortogonal a L, y sean $e' := \frac{e}{\|e\|}$ e $y' := \frac{y}{\|y\|}$.

$\Rightarrow \mathcal{B}' = (e',y',e'\times y')$ base ortonormal directa de V. Más aún,

$$\mathrm{Mat}_{\mathcal{B}'}(u) = \begin{pmatrix} 1 & 0 & 0\\ 0 & \langle u(y'),y'\rangle & \langle u(e'\times y'),y'\rangle\\ 0 & \langle u(y'),e'\times y'\rangle & \langle u(e'\times y'),e'\times y'\rangle\end{pmatrix} = \begin{pmatrix} 1 & 0 & 0\\ 0 & \cos\theta & -\sin\theta\\ 0 & \sin\theta & \cos\theta\end{pmatrix}$$

$$\Rightarrow \sin(\theta) = \langle u(y'),e'\times y'\rangle = \langle e'\times y',u(y')\rangle \underset{\text{dy}}{=} \det_{\mathcal{B}}(e',y',u(y')) = \frac{1}{\|y\|^2\|e\|}\det_{\mathcal{B}}(e,y,u(y))$$

Ahora, si $x\in V\setminus\{0\}$ es <u>cualquier</u> vector <u>no</u> colineal a e, entonces $y := x-\langle x,e'\rangle e'$ es no-nulo y ortogonal a e. Luego, $\sin(\theta)$ tiene el mismo signo que

$$\begin{aligned}\det_{\mathcal{B}}(e,y,u(y)) &= \det_{\mathcal{B}}(e,x-\langle x,e'\rangle e', u(x-\langle x,e'\rangle e'))\\ &= \det_{\mathcal{B}}(e,x,u(x)) - \langle x,e'\rangle \underbrace{u(e')}_{=e'}) = \det_{\mathcal{B}}(e,x,u(x)).\end{aligned}$$

<u>Conclusión</u>: Si $u = r_{L,\theta}\in SO(3)$ es la rotación de eje L, orientado por un vector director e, de ángulo $\theta\in\,]-\pi,\pi]$. Entonces:

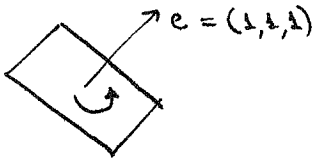
ⓐ $\mathrm{tr}(u) = 2\cos(\theta)+1$

ⓑ $\sin(\theta)$ tiene el <u>mismo</u> signo de $\det_{\mathcal{B}}(e,x,u(x))$, donde x es cualquier vector fuera del eje L y $\mathcal{B}$ es cualquier base directa.

Ejemplo: ① Sea $A = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \in M_3(\mathbb{R})$. Notamos que A es una matriz ortogonal pues ${}^tAA = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I_3$ y además $\det(A) = 1$. Luego, $A \in SO(3)$ y es la matriz (resp. a la base canónica) de una rotación $r_{L,\theta}$ cuyo eje L está generado por un vector propio asociado al valor propio $\lambda = 1$.

Calculamos $L = V_1 = \ker(A - I_3)$: $\begin{pmatrix} -1 & 0 & 1 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} z-x \\ x-y \\ y-z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow x = y = z$.

Consideramos $e = (1,1,1)$ como generador, que usamos para orientar L.



Para calcular θ, notamos que:

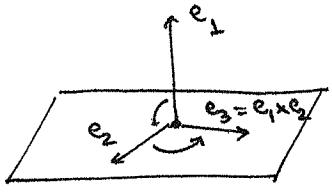
(a) $\mathrm{tr}(A) = 0 = 2\cos(\theta) + 1 \Rightarrow \cos(\theta) = -\frac{1}{2}$, y luego $\theta = \frac{2\pi}{3}$ o $\theta = -\frac{2\pi}{3}$.

(b) El vector $x = (1,0,0)$ no es colineal a e y cumple $r_L(x) = (0,1,0) \Rightarrow \det_{\mathcal{B}}(e, x, r_L(x)) = \begin{vmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{vmatrix} = 1 > 0$

$\Rightarrow \sin(\theta) > 0$ y luego $\theta = \frac{2\pi}{3}$.

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 part, el vector $e_2 = (4,3,0) \in \Pi$. Sea $e_3 := e_1 \times e_2 = (-75, 100, 300)$ producto cruz.

$\Rightarrow (e_1, e_2, e_3)$ base directa.



"Regla de la mano derecha"

Al normalizar, obtenemos:

$e_1' = \frac{e_1}{\|e_1\|} = \left(\frac{36}{65}, \frac{-48}{65}, \frac{25}{65}\right)$; $e_2' = \frac{e_2}{\|e_2\|} = \left(\frac{4}{5}, \frac{3}{5}, 0\right)$; $e_3' = \frac{e_3}{\|e_3\|} = \left(\frac{-3}{13}, \frac{4}{13}, \frac{12}{13}\right)$

y luego $\mathcal{B} = (e_1', e_2', e_3')$ base ortonormal directa. En esta base notamos que:

$A = \mathrm{Mat}_{\mathcal{B}}(r) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$. La matriz de cambio de base a la base canónica es $P = \begin{pmatrix} 36/65 & 4/5 & -3/13 \\ -48/65 & 3/5 & 4/13 \\ 25/65 & 0 & 12/13 \end{pmatrix} = \frac{1}{65}\begin{pmatrix} 36 & 52 & -15 \\ -48 & 39 & 20 \\ 25 & 0 & 60 \end{pmatrix}$

$\Rightarrow PAP^{-1} = PA{}^tP = \frac{1}{4225}\begin{pmatrix} 1296 & -3353 & -2220 \\ -103 & 2309 & -3540 \\ 4020 & 1140 & 625 \end{pmatrix}$ matriz de r en la base canónica de $\mathbb{R}^3$.

Notar que la traza es $\frac{1}{4225}(1296 + 2304 + 625) = 1 = 1 + 2\cos\left(\frac{\pi}{2}\right)$ ✓

§39. Isometrías en dimensión arbitraria y Teorema de Cartan-Dieudonné

Vimos que toda isometría de $V \cong \mathbb{R}^2$ (resp. $V \cong \mathbb{R}^3$) es el producto de a lo más 2 (resp. 3) reflexiones (ie, simetrías respecto a un hiperplano). Un resultado de Élie Cartan (1869-1951) y Jean Dieudonné (1906-1992) generaliza lo anterior a dimensión (finita) arbitraria.

Lema: Sea $V \cong \mathbb{R}^n$ espacio euclídeano y sean $x, y \in V \setminus \{0\}$ no-nulos tales que $x \neq y$ con $\|x\| = \|y\|$. Sea $H := (x-y)^\perp$ hiperplano ortogonal a la recta generada por $x-y$. Entonces, $s_H(x) = y$.

Dem: Buscamos $a \in H$ y $b \in H^\perp = \mathrm{Vect}_{\mathbb{R}}(x-y)$ tal que $x = a+b$. Dado que $\langle x+y, x-y\rangle = \|x\|^2 - \|y\|^2 = 0$, tenemos que $a := \frac{x+y}{2} \in H \Rightarrow b = x-a = \frac{x-y}{2}$.

Así, $x = \left(\frac{x+y}{2}\right) + \left(\frac{x-y}{2}\right)$ con $\frac{x+y}{2} \in H$, $\frac{x-y}{2} \in H^\perp \Rightarrow s_H(x) = \frac{x+y}{2} - \left(\frac{x-y}{2}\right) = y$ ∎

Teorema (Cartan-Dieudonné): Toda isometría de un espacio euclídeano de dimensión $n \geq 2$ es producto de a lo más n reflexiones.

Dem: Por inducción en $n \geq 2$: $n=2$ OK ✓

Sea $u \in O(n)$ una isometría de $V \cong \mathbb{R}^n$. Si $u = \mathrm{Id}_V$ entonces $\mathrm{Id}_V = s^2 = s \circ s$ para toda reflexión $s \in O(n)$ ✓ Si $u \neq \mathrm{Id}_V$ entonces $\exists\, x \in V \setminus \{0\}$ tq $u(x) \neq x$. Dado que $u \in O(n)$, $\|x\| = \|u(x)\|$ tienen la misma norma $\overset{\text{Lema}}{\Rightarrow} \exists\, s$ reflexión tq $s(x) = u(x) \Rightarrow s^2(x) = x = (s \circ u)(x)$. En part, $s \circ u \in O(n)$ deja invariante la recta generada por x y el hiperplano ortogonal $H := x^\perp$.

Hipótesis de inducción $\Rightarrow (s \circ u)|_H$ se escribe como $s_{H_1'} \circ \cdots \circ s_{H_m'}$, donde $m \leq n-1$ y $H_1', \ldots, H_m'$ son hiperplanos en H (ie, $\dim_{\mathbb{R}} H_j' = n-2$). Para $j \in \{1, \ldots, m\}$ definimos $H_j := \mathrm{Vect}_{\mathbb{R}}(H_j', x)$ hiperplano en V generado por H_j' y por x.

$\Rightarrow$ El producto $s_{H_1} \circ \cdots \circ s_{H_m}$ coincide con $s \circ u$ al restringirse a H, y al restringirse a $L = \mathrm{Vect}_{\mathbb{R}}(x)$ (pues $(s_{H_1} \circ \cdots \circ s_{H_m})|_L = \mathrm{Id}_L$).

Dado que $V = H \oplus L$ concluimos que $s_{H_1} \circ \cdots \circ s_{H_m} = s \circ u$ en todo V

$\Rightarrow u = s^{-1} \circ s_{H_1} \circ \cdots \circ s_{H_m} = s \circ s_{H_1} \circ \cdots \circ s_{H_m}$ ∎

Def: Sea $V \cong \mathbb{R}^n$ un espacio euclídeano. Una función (no necesariamente lineal) $\varphi : V \to V$ es una función isométrica si para todos $x, y \in V$ se tiene

$$d(\varphi(x), \varphi(y)) = d(x,y) \overset{\text{def}}{\Leftrightarrow} \|\varphi(x) - \varphi(y)\| = \|x - y\|.$$

Ejemplo: ① Toda isometría (lineal) $u \in O(n)$ es una función isométrica.

② Sea $a \in V$ vector fijo. La traslación en a, dada por

$$t_a : V \to V, \quad x \mapsto t_a(x) = x + a$$

es una función isométrica (no lineal si $a \neq 0$).

Teorema: Sea $V \cong \mathbb{R}^n$ un espacio euclídeano y sea $\varphi: V \to V$ función isométrica. Entonces, existe un único vector $a \in V$ y una única isometría (lineal) $u \in O(n)$ tal que

$$\varphi = t_a \circ u \iff \varphi(x) = u(x) + a \text{ para todo } x \in V.$$

Dem: Para la unicidad, notamos que si $\varphi = t_a \circ u = t_b \circ v$ con $a, b \in V$, $u, v \in O(n)$

$$\Rightarrow \varphi(0) = \underbrace{u(0)}_{=0} + a = \underbrace{v(0)}_{=0} + b \Rightarrow a = b \Rightarrow u = t_{-a} \circ \varphi = v \checkmark$$

Para la existencia: Sea $a := \varphi(0) \in V$ y consideremos $\psi := t_{-a} \circ \varphi : V \to V$ que cumple $\psi(0) = 0$ y $d(\psi(x), \psi(0)) = d(x, 0) \iff \|\psi(x)\| = \|x\|$ para todo $x \in V$.

Veamos que $\psi : V \to V$ es lineal (y luego, $\psi = u \in O(n)$): Sea $(e_1, \dots, e_n)$ base ortonormal de V, entonces para $i \neq j$ tenemos que

$$\|\psi(e_i) - \psi(e_j)\| = d(\psi(e_i), \psi(e_j)) = d(e_i, e_j) = \|e_i - e_j\| = \sqrt{2}.$$

Luego, $\|\underbrace{\psi(e_i) - \psi(e_j)}_{=2}\|^2 = \underbrace{\|\psi(e_i)\|^2}_{=1} + \underbrace{\|\psi(e_j)\|^2}_{=1} - 2\langle \psi(e_i), \psi(e_j)\rangle \Rightarrow \langle \psi(e_i), \psi(e_j)\rangle = 0.$

$\Rightarrow (\psi(e_1), \dots, \psi(e_n))$ es una base ortonormal de V. Sea $x \in V$ arbitrario y escribamos ① $x = \sum_{j=1}^n x_j e_j$ usando la base $(e_1, \dots, e_n)$.

② $\psi(x) = \sum_{j=1}^n y_j \psi(e_j)$ usando la base $(\psi(e_1), \dots, \psi(e_n))$.

$\Rightarrow d(x, e_k)^2 = \|x - e_k\|^2 = \|x\|^2 - 2\langle x, e_k\rangle + \|e_k\|^2 = \|x\|^2 - 2x_k + 1$, y también

$$d(\psi(x), \psi(e_k))^2 = \|\psi(x) - \psi(e_k)\|^2 = \|\psi(x)\|^2 - 2y_k + 1 = \|x\|^2 - 2y_k + 1$$

Dado que $d(\psi(x), \psi(e_k)) = d(x, e_k)$, concluimos que $x_k = y_k$ para todo k

$\Rightarrow \psi\left(\sum_{j=1}^n x_j e_j\right) = \sum_{j=1}^n x_j \psi(e_j)$ y luego ψ es lineal ✓ ■

§40. Reducción simultánea de formas cuadráticas e hipersuperficies cuádricas

Sea $V \cong \mathbb{R}^n$ un espacio euclídeano. Gracias al Teorema de Sylvester (ver §31) sabemos que existe una base ortogonal (de hecho, ortonormal) tal que la forma cuadrática dada por la norma $\|\cdot\|^2$ se reduce a $\|x\|^2 = x_1^2 + \dots + x_n^2$. De manera más general, toda forma cuadrática $Q: V \to \mathbb{R}$ admite una escritura de la forma

$$Q(x) = \lambda_1 x_1^2 + \dots + \lambda_p x_p^2 - \lambda_{p+1} x_{p+1}^2 - \dots - \lambda_{p+q} x_{p+q}^2$$

donde (p, q) es la signatura de Q y $(x_1, \dots, x_n)$ son las coord. de x resp. a una base ortogonal. El resultado siguiente afirma que podemos reducir simultáneamente $\|\cdot\|^2$ y Q:

Teorema: Sea $V \cong \mathbb{R}^n$ un espacio euclídeano y sea $Q: V \to \mathbb{R}$ una forma cuadrática. Entonces, existe una base ortonormal de V que es ortogonal para Q.

Dem: Denotemos momentaneamente $E := \langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ el producto escalar en V, y por $B: V \times V \to \mathbb{R}$ la forma bilineal simétrica asociada a Q.

Construiremos un endomorfismo $u: V \to V$ tal que $B(x, y) = \langle x, u(y)\rangle \ \forall x, y \in V$:

Si escribimos $B(x,y) = \widehat{B}(y)(x)$ y $\langle x, u(y)\rangle = E(x,u(y)) = \widehat{E}(u(y))(x)$, la igualdad buscada se reescribe como $\widehat{B} = \widehat{E} \circ u$. Dado que $E = \langle\cdot,\cdot\rangle$ es no-degenerada, $\widehat{E}: V \xrightarrow{\sim} V^*$ es un isomorfismo y podemos definir $u := \widehat{E}^{-1} \circ \widehat{B} : V \to V$.

$\Rightarrow \langle x, u(y)\rangle = B(x,y) = B(y,x) = \langle y, u(x)\rangle$ y luego $u = u^*$ es simétrico. En part, u es diagonalizable en una base ortonormal $\mathcal{B} = (e_1, \dots, e_m)$ de V, ie, $u(e_j) = \lambda_j e_j$ para ciertos $\lambda_1, \dots, \lambda_m \in \mathbb{R}$.

$\Rightarrow B(e_i, e_j) = \langle e_i, u(e_j)\rangle = \lambda_j \langle e_i, e_j\rangle = \lambda_j \delta_{ij} = \begin{cases} 0 & \text{si } i \neq j \\ \lambda_j & \text{si } i = j \end{cases}$ ✓ ■

Una de las aplicaciones del resultado anterior es el estudio de hipersuperficies cuádricas (o simplemente "cuádricas").

Def: Sea $V \cong \mathbb{R}^m$ espacio euclideano. Sean $Q: V \to \mathbb{R}$ una forma cuadrática, $f: V \to \mathbb{R}$ forma lineal (ie, $f \in V^*$) y $c \in \mathbb{R}$ una constante. Decimos que el conjunto

$$X = \{x \in V \text{ tal que } Q(x) + f(x) + c = 0\} \subseteq V$$

es una cuádrica en V. Más aún, diremos que la cuádrica $X \subseteq V$ es suave si la forma cuadrática $\widetilde{Q}: V \times \mathbb{R} \to \mathbb{R}$ dada por

$$\widetilde{Q}(x,t) = Q(x) + f(x)t + ct^2$$

en $V \times \mathbb{R} \cong \mathbb{R}^{m+1}$ es no-degenerada; en caso contrario, diremos que $X \subseteq V$ es singular.

Obs: Sea $p \in V$ y consideremos la traslación $t_{-p}: V \to V$, $x \mapsto x - p$. Entonces, el conjunto $t_{-p}(X)$ es el 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. Luego, $t_{-p}(X)$ es la cuádrica asociada a la forma cuadrática Q, a la forma lineal $(2\widehat{B}(p) + f) \in V^*$ y a la constante $c' := Q(p) + f(p) + c$.

Si la forma cuadrática Q es no-degenerada entonces $\widehat{B}: V \xrightarrow{\sim} V^*$ es un isomorfismo.

$\Rightarrow$ Si tomamos $p := -\widehat{B}^{-1}(f/2)$ entonces la ecuación de $t_{-p}(X)$ es de la forma $Q(x) + c' = 0$ (no hay término lineal), y luego $t_{-p}(X)$ es simétrica respecto al origen (ie, $x \in t_{-p}(X) \iff -x \in t_{-p}(X)$).

Def: Sea $V \cong \mathbb{R}^m$ espacio euclideano y sea $X = \{x \in V \text{ tq } Q(x) + f(x) + c = 0\} \subseteq V$ una cuádrica. Diremos que X es una cuádrica centrada si la forma cuadrática $Q: V \to \mathbb{R}$ es no-degenerada. Más aún, diremos que X es una cuádrica centrada en el origen si además se tiene $f = 0$.

Obs: Por la discusión anterior, toda cuádrica centrada puede ser trasladada a una cuádrica centrada en el origen. El punto $p = -\widehat{B}^{-1}(f/2)$ es llamado el centro de la cuádrica centrada $Q(x) + f(x) + c = 0$.

Prop: Sea $V \cong \mathbb{R}^n$ espacio euclídeano y sea $X = \{x \in V \text{ tq } Q(x) + c = 0\} \subseteq V$ cuádrica centrada en el origen. Entonces, existe una base ortonormal $\mathcal{B} = (e_1, \dots, e_n)$ tal que

$$x = x_1 e_1 + \dots + x_n e_n \in X \iff \lambda_1 x_1^2 + \dots + \lambda_n x_n^2 + c = 0 \quad (\text{"Ecuación reducida de } X\text{"}),$$

donde $\lambda_1, \dots, \lambda_n \in \mathbb{R} \setminus \{0\}$ no-nulos. En part, X suave $\iff c \neq 0$.

Dem: Vimos que existe una base ortonormal $\mathcal{B} = (e_1, \dots, e_n)$ de V que es ortogonal respecto a la forma cuadrática Q. En part, la matriz $A_{\mathcal{B}}$ de Q está dada por

$$A_{\mathcal{B}} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

donde $\det(A_{\mathcal{B}}) = \lambda_1 \cdots \lambda_n \neq 0$, pues Q es no-degenerada. Así, si $(x_1, \dots, x_n) \in \mathbb{R}^n$ son las coordenadas de $x \in V$ resp. a $\mathcal{B}$, entonces $Q(x) = Q(x_1, \dots, x_n) = \sum_{j=1}^{n} \lambda_j x_j^2$ ✓

Finalmente, X es una cuádrica suave $\overset{\text{dy}}{\iff} \widetilde{Q}(x,t) = Q(x) + ct^2$ es no-degenerada $\iff c \neq 0$ (pues Q es no-degenerada). ■

Terminología: Con la notación anterior:

① Un eje de la cuádrica $X \subseteq V \cong \mathbb{R}^n$ es toda recta $L \subseteq V$ generada por un vector propio asociado a alguno de los valores propios $\lambda_1, \dots, \lambda_n \in \mathbb{R}$. Un vértice de X es un punto $p \in X$ que está sobre un eje (ie, $p \in X \cap L$, para algún eje L de X).

② Una cuádrica en el plano $V \cong \mathbb{R}^2$ es llamada una cónica.

③ Una cuádrica en el espacio $V \cong \mathbb{R}^3$ es llamada una superficie cuádrica.

Ejemplos: ① La cónica en $V = \mathbb{R}^2$ de ecuación $y = x^2 + 1$ ("parábola") no es una cónica centrada, pues $Q(x,y) = x^2$ es degenerada. Sin embargo, la forma cuadrática $\widetilde{Q}(x,y,t) = x^2 - yt + t^2$ es no-degenerada (Ejercicio, cf. §32) y luego $C = \{(x,y) \in \mathbb{R}^2 \text{ tq } y = x^2 + 1\} \subseteq \mathbb{R}^2$ es una cónica suave.

② Consideremos la cónica C en $\mathbb{R}^2$ dada por la ecuación $3x^2 + 2xy + 3y^2 + 4x - 4y = 0$. La matriz de $Q(x,y) = 3x^2 + 2xy + 3y^2$ es $A = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} \Rightarrow \det(A) = 8 \neq 0$. Luego, $C \subseteq \mathbb{R}^2$ es una cónica centrada; sea $p = (a,b) \in \mathbb{R}^2$ su centro y sea $f(x,y) = 4x - 4y$. Para determinar $p = (a,b)$ impondremos la anulación del término lineal de

$$3(x+a)^2 + 2(x+a)(y+b) + 3(y+b)^2 + 4(x+a) - 4(y+b) = 0$$

$$\iff 6ax + 2ay + 2bx + 6by + 4x - 4y = 0 \ \forall x, y \iff \begin{cases} 6a + 2b + 4 = 0 \\ 2a + 6b - 4 = 0 \end{cases} \Rightarrow \boxed{(a,b) = (-1,1)}$$

Sea $C' = t_{-p}(C)$ la traslación de C por $(1,-1)$.

$$\Rightarrow C' = \{(x,y) \in \mathbb{R}^2 \text{ tq } 3x^2 + 2xy + 3y^2 + \underbrace{Q(-1,1)}_{=4} + \underbrace{f(-1,1)}_{=-8} = 0\} \ *$$

$$= \{(x,y) \in \mathbb{R}^2 \text{ tq } 3x^2 + 2xy + 3y^2 = 4\}.$$

Calculamos $P_A(X) = \begin{vmatrix} X-3 & -1 \\ -1 & X-3 \end{vmatrix} = (X-3)^2 - 1 = X^2 - 6X + 8 = (X-2)(X-4)$

$\Rightarrow \lambda_1 = 2$ y $\lambda_2 = 4$ valores propios (reales, tal como predice la teoría!).

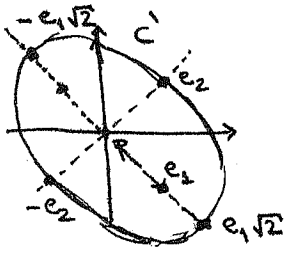
Vectores propios: $V_2 = \ker(A-2I_2)$: $\begin{pmatrix}1&1\\1&1\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}x+y\\x+y\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix} \Rightarrow x = -y$.

$v_1 = (1,-1) \rightsquigarrow e_1 := \frac{v_1}{\|v_1\|} = \left(\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}\right)$ vector propio unitario.

$V_1 = V_2^\perp$: $\langle (x,y), (1,-1)\rangle = x-y = 0 \Rightarrow x = y \rightsquigarrow v_2 = (1,1) \Rightarrow e_2 := \frac{v_2}{\|v_2\|} = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$ vector propio unitario.

Luego, la ecuación reducida de C' en la base $\mathcal{B} = (e_1, e_2)$ es:

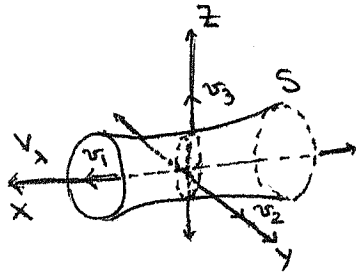
$$2X^2 + 4Y^2 = 4 \quad (\text{"elipse"})$$



Los ejes de C' son las rectas (pasando por $p = (-1,1)$) generadas por e_1 y e_2. Los vértices son $\pm e_1\sqrt{2}$ y $\pm e_2$.

③ Consideramos la superficie cuádrica $S \subseteq \mathbb{R}^3$ dada por $x^2+3y^2+z^2+4xz = 3$. La matriz de $Q(x,y,z) = x^2+3y^2+z^2+4xz$ es $A = \begin{pmatrix}1&0&2\\0&3&0\\2&0&1\end{pmatrix}$, que fue analizada en §36 (pág 112). Sus valores propios son $\lambda = -1$ y $\mu = 3$ (doble). Una base ortonormal de vectores propios es $v_1 = \begin{pmatrix}1/\sqrt{2}\\0\\-1/\sqrt{2}\end{pmatrix} \in V_\lambda$, $v_2 = \begin{pmatrix}1/\sqrt{2}\\0\\1/\sqrt{2}\end{pmatrix} \in V_\mu$, $v_3 = \begin{pmatrix}0\\1\\0\end{pmatrix} \in V_\mu$. Luego, la ecuación reducida de S en la base $\mathcal{B} = (v_1, v_2, v_3)$ es:

$$-X^2 + 3Y^2 + 3Z^2 = 3 \quad (\text{"hiperboloide de una hoja"}).$$



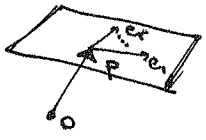
Los ejes de S son las rectas generadas por v_1 (ie, V_λ) y todas las rectas contenidas en $V_\mu = V_\lambda^\perp$ (ie, todas las rectas ortogonales a V_λ).

Def: Sea $V \cong \mathbb{R}^n$ espacio vectorial real y $\mathcal{U} \subseteq V$ un sub-ev. Para todo vector $p \in V$, definimos el sub-espacio afín de V dirigido por $\mathcal{U}$ pasando por p como:

$$p + \mathcal{U} = \{x \in V \text{ tal que } x - p \in \mathcal{U}\} = \{p + u,\ u \in \mathcal{U}\},$$

ie, $p+\mathcal{U}$ es la preimagen de $[p]$ en el cociente $V/\mathcal{U}$ (ver §27). Explícitamente, si $\mathcal{B} = (e_1, \dots, e_r)$ es una base de $\mathcal{U}$ entonces

$$p + \mathcal{U} = \{p + \lambda_1 e_1 + \dots + \lambda_r e_r, \text{ donde } \lambda_1, \dots, \lambda_r \in \mathbb{R}\}.$$



Ejemplo: ① Sea $\mathcal{B} = (e_1, \dots, e_n)$ base de $V \cong \mathbb{R}^n$ y sea

$$H = \{x = \sum_{j=1}^n x_j e_j \in V \text{ tal que } a_1x_1 + \dots + a_nx_n = 0\} \text{ hiperplano}.$$

Si $p = \sum_{j=1}^n p_j e_j$ vector de coord $(p_1, \dots, p_n)$, el hiperplano afín dirigido por H pasando por p es $\{x \in V \text{ tq } x - p \in H\} = \{x = \sum_{j=1}^n x_j e_j \in V \text{ tq } a_1(x_1-p_1) + \dots + a_n(x_n-p_n) = 0\}$.

② Sea $V \cong \mathbb{R}^n$ espacio euclídeano y sea $X = \{x \in V \text{ tq } Q(x) + f(x) + c = 0\}$ una cuádrica. Si $\mathcal{L} = p + \mathrm{Vect}_{\mathbb{R}}(e) = \{p + te,\ t \in \mathbb{R}\}$ es la recta afín dirigida por e pasando por p, entonces la intersección $X \cap \mathcal{L}$ está dada por:

$$0 = Q(p+te) + f(p+te) + c = Q(e)t^2 + (2B(e,p) + f(e))t + (Q(p) + f(p) + c).$$

La ecuación de grado ≤ 2 tiene discriminante $\Delta = (2B(e,p)+f(e))^2 - 4Q(e)(Q(p)+f(p)+c)$.
Luego, ya sea $\ell \subseteq X$ o bien $X \cap \ell$ consiste en a lo más dos puntos. Si $Q(e) \neq 0$:

① $\Delta < 0$ y $X \cap \ell = \emptyset$
② $\Delta > 0$ y $X \cap \ell = \{p_1, p_2\}$
③ $\Delta = 0$ y $X \cap \ell = \{p\}$



En general, incluso si $Q(e)=0$, diremos que ℓ es <u>tangente</u> a X si $\Delta = 0$.

Prop: Sea X una cuádrica en el espacio euclideano $V \cong \mathbb{R}^n$, y sea $p \in X$ un punto de la cuádrica (ie, $Q(p)+f(p)+c=0$). Si consideramos el <u>espacio tangente de X en p</u> dado por

$$T_pX := \{\ell \subseteq V \text{ recta afín pasando por } p \text{ tal que } \ell \text{ tangente a } X\},$$

entonces: ① $T_pX = V$ y X es una cuádrica singular; o bien
② T_pX es un hiperplano afín. ■

En part, X es una cuádrica suave $\Leftrightarrow T_pX$ es un hiperplano afín $\forall p \in X$.

<u>Dem</u>: Como $p \in X$, la condición $\Delta = 0$ se reduce a $0 = 2B(e,p) + f(e)$, ie, $(2\hat{B}(p) + f)(e) = 0$. Luego, la unión de las rectas tangentes a X en p es el sub-espacio afín de V dirigido por $H := \ker(2\hat{B}(p)+f)$ y pasando por p. H es efectivamente un hiperplano a menos que $2\hat{B}(p)+f = 0$ en V^*. Si escribimos $\tilde{Q}(x,t) = Q(x) + f(x)t + ct^2$ en $V \times \mathbb{R} \cong \mathbb{R}^{n+1}$, entonces (<u>Ejercicio</u>):

$\tilde{B}((x,t),(y,s)) = B(x,y) + \frac{1}{2}f(x)s + \frac{1}{2}f(y)t + cts$ forma bilineal simétrica asociada.

$\Rightarrow \tilde{B}((p,1),(y,0)) = B(p,y) + \frac{1}{2}f(y) = \frac{1}{2}(2\hat{B}(p)+f)(y)$.

En otras palabras, la condición $2\hat{B}(p)+f$ equivale a decir que el vector $(p,1) \in V \times \mathbb{R}$ es ortogonal a $V \times \{0\}$ respecto a $\tilde{Q}$. Dado que $\tilde{Q}(p,1) = 0$ (pues $p \in X$), esto implica que $(p,1) \in \ker(\tilde{Q})$ y esto sólo es posible si $\tilde{Q}$ es degenerada, ie, X es una cuádrica singular. ■

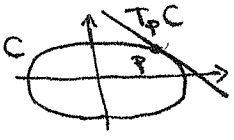
<u>Ejemplo</u>: Si X está centrada en el origen (ie, Q no degenerada y $f=0$) entonces T_pX está dirigido por $H = \ker(\hat{B}(p)) = \{x \in V \text{ tq } B(x,p)=0\} =: p^{\perp_Q}$, ie, el ortogonal de p respecto a Q.

<u>Ejemplo</u>: Sea $C \subseteq \mathbb{R}^2$ elipse dada por $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$, centrada en el origen.

$\Rightarrow Q(x,y) = \frac{x^2}{a^2} + \frac{y^2}{b^2} \Rightarrow B((x_1,y_1),(x_2,y_2)) = \frac{x_1x_2}{a^2} + \frac{y_1y_2}{b^2}$. Luego, si $p = (x_0,y_0) \in C$

$\Rightarrow L = p^{\perp_Q} = \{(x,y) \in \mathbb{R}^2 \text{ tq } \frac{xx_0}{a^2} + \frac{yy_0}{b^2} = 0\}$ y por ende:

$$T_pC = \{(x,y) \in \mathbb{R}^2 \text{ tq } \frac{(x-x_0)x_0}{a^2} + \frac{(y-y_0)y_0}{b^2} = 0\} = \{(x,y) \in \mathbb{R}^2 \text{ tq } \frac{xx_0}{a^2} + \frac{yy_0}{b^2} = 1\}$$



Si $S \subseteq \mathbb{R}^3$ es una superficie cuadrática suave y centrada en el origen, entonces su ecuación puede escribirse (en una base conveniente) como:

(1º) Elipsoide: $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1 \rightsquigarrow$ signatura $(3,0)$

(2º) Hiperboloide de una hoja:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1 \rightsquigarrow \text{signatura } (2,1)$$

(3º) Hiperboloide de dos hojas:

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1 \rightsquigarrow \text{signatura } (1,2)$$



Teorema: Sea $V \cong \mathbb{R}^3$ espacio euclídeo y sea $S \subseteq V$ hiperboloide de una hoja. Entonces, S es una *superficie reglada*: por cada punto $p \in S$ pasan dos rectas (afines) contenidas en S, dadas por la intersección de S y el plano tangente T_pS.

Dem: Como S está centrada en el origen, tenemos que T_pS es un plano afín dirigido por $p^{\perp_Q}$, el ortogonal de p respecto a Q. Escribamos

$$Q(x,y,z) = \frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2}$$



Sea $p' \in p^{\perp_Q}$ y consideremos $p+p'$ en T_pS. Entonces, $p+p' \in S$ si y sólo si:

$$0 = Q(p+p') - 1 = \underbrace{Q(p)}_{=1} + 2\underbrace{B(p,p')}_{=0} + Q(p') - 1 = Q(p').$$

$\Rightarrow T_pS \cap S$ es la traslación por p del conjunto de vectores *isótropos* respecto a $Q' := Q|_{p^{\perp_Q}}$. Dado que $Q(p) = 1 > 0$ y la signatura de Q es $(2,1)$, la signatura de Q' en $p^{\perp_Q} \cong \mathbb{R}^2$ es $(1,1)$.

$\underset{\text{Sylvester}}{\Rightarrow} Q'(u,v) = u^2 - v^2 = (u+v)(u-v)$ en una base conveniente de $p^{\perp_Q} \cong \mathbb{R}^2$.

Luego, los vectores isótropos de Q' son la unión de las rectas $u = v$ y $u = -v$ ✓ ■

Ejercicio Sea $V \cong \mathbb{R}^3$ espacio euclídeo y sea $S \subseteq V$ un elipsoide o un hiperboloide de dos hojas. Probar que para todo $p \in S$, la intersección $T_pS \cap S$ es exactamente el punto p. [Indicación: Analizar la signatura de Q'].

Ejercicio Considerar la superficie cuádrica $S \subseteq \mathbb{R}^3$ dada por:

$$xy + xz + yz + 2y + 1 = 0$$

Determinar una base ortonormal en la cual S admite una ecuación reducida. Determinar dicha ecuación y la naturaleza geométrica de S.

Motivación: Nos gustaría extender la noción de espacio euclideano a espacios vectoriales complejos $V \cong \mathbb{C}^n$. Si $n=1$, (ie, $V \cong \mathbb{C}$) podemos considerar el módulo de $z \in \mathbb{C}$ dado por $|z|^2 = z\overline{z} = h(z,z)$, donde

$$h : \mathbb{C}\times\mathbb{C} \to \mathbb{C} \quad , \quad h(z_1,z_2) = z_1\overline{z_2} .$$

Sin embargo, h no es una forma bilineal compleja: h es $\mathbb{C}$-lineal en la 1ª variable, pero h es "anti-lineal" (o lineal conjugada) resp. a la 2ª variable:

$$h(z_1, \lambda z_2 + z_3) = \overline{\lambda}\, h(z_1,z_2) + h(z_1,z_3) .$$

Def: Sea $V \cong \mathbb{C}^n$ espacio vectorial complejo. Decimos que una función $h : V\times V \to \mathbb{C}$ es una forma hermitiana si h es:

① lineal en la primera variable: $h(\lambda x + y, z) = \lambda h(x,z) + h(y,z)$
② anti-lineal en la segunda variable: $h(x, \lambda y + z) = \overline{\lambda}\, h(x,y) + h(x,z)$
} "sesqui lineal"

③ simétrica hermitiana: $h(x,y) = \overline{h(y,x)}$.

En part. $h(x,x) \in \mathbb{R}$ $\forall x \in V$ y además $h(x,y) = 0 \iff h(y,x) = 0$.

Ejemplo: En $\mathbb{C}^n$, la forma $h : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ dada por

$$h((z_1,\dots,z_n),(w_1,\dots,w_n)) = z_1\overline{w_1} + z_2\overline{w_2} + \dots + z_n\overline{w_n} \text{ es hermitiana.}$$

Obs: Tal como en el caso de formas bilineales reales, podemos describir las formas hermitianas matricialmente: Sea $\mathcal{B} = (e_1,\dots,e_n)$ base de $V \cong \mathbb{C}^n$ y escribamos $x = x_1e_1 + \dots + x_ne_n$, $y = y_1e_1 + \dots + y_ne_n \in V$. Entonces:

$$h(x,y) = h\left(x_1e_1+\dots+x_ne_n, y\right) = \sum_{j=1}^n x_j\, h(e_j,y) = \sum_{j=1}^n x_j\, h\Big(e_j, \sum_{k=1}^n y_k e_k\Big) = \sum_{j=1}^n\sum_{k=1}^n x_j\overline{y_k}\, h(e_j,e_k).$$

Luego, si definimos $h_{jk} := h(e_j,e_k) \in \mathbb{C}$ entonces

$$h(x,y) = \sum_{1\le j,k\le n} h_{jk}\, x_j\, \overline{y_k} .$$

Más aún, la simetría hermitiana $h(x,y) = \overline{h(y,x)}$ equivale a $h_{jk} = \overline{h_{kj}}$.

Def: Sea $\mathcal{B} = (e_1,\dots,e_n)$ una base de $V \cong \mathbb{C}^n$, y sea $h : V\times V \to \mathbb{C}$ forma hermitiana. Definimos la matriz de h resp. a $\mathcal{B}$ mediante

$$H = \mathrm{Mat}_{\mathcal{B}}(h) := (h_{jk})_{1\le j,k\le n} \text{, donde } h_{jk} := h(e_j,e_k).$$

Más aún, si $A = (a_{jk}) \in M_{n\times p}(\mathbb{C})$ definimos la matriz adjunta de A por:

$$A^* := {}^t\overline{A} = (\overline{a_{kj}}) \in M_{p\times n}(\mathbb{C}).$$

En part, $H^* = H$ y diremos que H es una matriz hermitiana.

Obs: Si $A \in M_{n\times p}(\mathbb{C})$ y $B \in M_{p\times r}(\mathbb{C})$ entonces $(AB)^* = B^*A^*$.

⚠ Importante: Sea $\mathcal{B}=(e_1,\dots,e_m)$ base de $V\cong\mathbb{C}^m$ y sea $h:V\times V\to\mathbb{C}$ forma hermitiana, con $H=\mathrm{Mat}_{\mathcal{B}}(h)=(h_{jk})$. Entonces, para $x=\sum_{j=1}^m x_j e_j$ e $y=\sum_{k=1}^m y_k e_k$:

$$h(x,y)=\sum_{j=1}^m\sum_{k=1}^m x_j\,\overline{y_k}\,h(e_j,e_k)=\sum_{j=1}^m\sum_{k=1}^m x_j\,h_{jk}\,\overline{y_k}={}^tX H\overline{Y},$$

donde $X=\begin{pmatrix}x_1\\ \vdots\\ x_m\end{pmatrix}$, $Y=\begin{pmatrix}y_1\\ \vdots\\ y_m\end{pmatrix}\in\mathbb{C}^m$ vectores coordenadas resp. a $\mathcal{B}$.

Ejercicio Probar que si $\mathcal{B}'$ es otra base de V y $P=\mathrm{Mat}_{\mathcal{B}}(\mathcal{B}')$ es la matriz de cambio de base, entonces $H'=\mathrm{Mat}_{\mathcal{B}'}(h)={}^tP H\overline{P}$.

Def: Sea $V\cong\mathbb{C}^m$ e.v. y sea $h:V\times V\to\mathbb{C}$ forma hermitiana. La forma cuadrática hermitiana asociada a h es $Q:V\to\mathbb{R}$ dada por

$$Q(x)=h(x,x)\quad\text{para todo } x\in V.$$

En part, para todos $x\in V$ y $\lambda\in\mathbb{C}$ se cumple

$$Q(\lambda x)=h(\lambda x,\lambda x)=\lambda\overline{\lambda}\,h(x,x)=|\lambda|^2 Q(x).$$

Ejemplo: En $\mathbb{C}^m$, la forma cuadrática hermitiana

$$Q(z_1,\dots,z_m)=|z_1|^2+\dots+|z_m|^2$$

está asociada a la forma hermitiana h del ejemplo anterior (pág 130).

Obs: Dado que $Q(x)=h(x,x)\in\mathbb{R}$ para todo $x\in V\cong\mathbb{C}^m$, tenemos que Q está también asociada a $B:V\times V\to\mathbb{R}$ dada por $B(x,y):=\mathrm{Re}\,(h(x,y))$, la cual es bilineal real simétrica pues

$$B(y,x)=\mathrm{Re}\,h(y,x)=\mathrm{Re}\,\overline{h(x,y)}=\mathrm{Re}\,h(x,y)=B(x,y).$$

En part, Q es una forma cuadrática real en $V\cong\mathbb{R}^{2m}$. Más aún, ella verifica $Q(\lambda x)=|\lambda|^2 Q(x)\ \ \forall x\in V,\ \lambda\in\mathbb{C}$.

Lema (fórmula de polarización): Sea $V\cong\mathbb{C}^m$ y $h:V\times V\to\mathbb{C}$ forma hermitiana. Si Q es la forma cuadrática hermitiana asociada a h, entonces:

① $\mathrm{Re}\,h(x,y)=\frac{1}{2}\big(Q(x+y)-Q(x)-Q(y)\big)=\frac{1}{4}\big(Q(x+y)-Q(x-y)\big)$

② $\mathrm{Im}\,h(x,y)=\frac{1}{2}\big(Q(x+iy)-Q(x)-Q(y)\big)=\frac{1}{4}\big(Q(x+iy)-Q(x-iy)\big)$

Luego, $h(x,y)=\frac{1}{4}\big(Q(x+y)-Q(x-y)+iQ(x+iy)-iQ(x-iy)\big)$

está completamente determinada por Q.

Dem: La fórmula de polarización real aplicada a $B(x,y)=\mathrm{Re}\,h(x,y)$ da ① ✓

Dado que $\forall z\in\mathbb{C}$ se tiene $\mathrm{Im}(z)=\mathrm{Re}(-iz)$ tenemos que

$$\mathrm{Im}\,h(x,y)=\mathrm{Re}(-ih(x,y))=\mathrm{Re}\,h(x,iy),$$

por lo que ② se obtiene de ① reemplazando y por iy, y usando que $Q(iy)=|i|^2Q(y)=Q(y)$ ✓ Finalmente, $h(x,y)=\mathrm{Re}\,h(x,y)+i\,\mathrm{Im}\,h(x,y)$. ∎

Def: Sea $V \cong \mathbb{C}^n$ ev y $h: V \times V \to \mathbb{C}$ forma hermitiana. Dada $\mathcal{B}$ una base de V, definimos el <u>rango de h</u> (o de Q) como $rg(h) := rg(H)$, donde H es la matriz de h resp. a $\mathcal{B}$. Decimos que h (o que Q) es <u>no-degenerada</u> si $rg(h) = \dim_{\mathbb{C}}(V) = n$, ie, $H \in Gl_n(\mathbb{C})$ es invertible.

Obs: La fórmula de cambio de base $H' = {}^t\overline{P} H P$ implica $rg(H) = rg(H')$.

Ejercicio Determinar la forma hermitiana de $\mathbb{C}^n$ asociada a

$$Q(z_1, \dots, z_n) = |z_1|^2 + \dots + |z_p|^2 - |z_{p+1}|^2 - \dots - |z_{p+q}|^2 \quad , (p+q \leq n)$$

y determinar su rango.

El hecho que $h(x,y) = 0 \iff h(y,x) = 0$ para una forma hermitiana h nos permite definir la noción de ortogonalidad:

Def: Sea $V \cong \mathbb{C}^n$ ev y $h: V \times V \to \mathbb{C}$ forma hermitiana. Diremos que:

1. Dos vectores $x, y \in V$ son <u>ortogonales</u> respecto a h (o Q) si $h(x,y) = 0$.
2. Si $\mathcal{U} \subseteq V$ sub-conjunto no-vacío, el <u>sub-espacio</u>
$$\mathcal{U}^{\perp} = \mathcal{U}^{\perp_h} = \mathcal{U}^{\perp_Q} := \{y \in V \text{ tq } \forall x \in \mathcal{U},\ h(x,y) = 0\}$$
es el <u>ortogonal de $\mathcal{U}$</u>.
3. El <u>kernel</u> de h (o de Q) es
$$\ker(h) = \ker(Q) := \{x \in V \text{ tq } \forall y \in V,\ h(x,y) = 0\} = V^{\perp}$$
4. Un vector $x \in V$ es <u>isótropo</u> si $Q(x) = h(x,x) = 0$.

Ejemplo: En $\mathbb{C}^2$, la forma cuadrática hermitiana $Q(z_1, z_2) = |z_1|^2 - |z_2|^2$ está asociada a $h((z_1, z_2), (w_1, w_2)) = z_1 \overline{w_1} - z_2 \overline{w_2}$, cuya matriz resp. a la base canónica de $\mathbb{C}^2$ es $H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ $\Rightarrow$ h es no-degenerada. Notar que $\ker(h) = \{0\} \neq \{(z_1, z_2) \in \mathbb{C}^2 \text{ tq } |z_1| = |z_2|\}$ (vectores isótropos).

Teorema: Sea $V \cong \mathbb{C}^n$ ev. y $h: V \times V \to \mathbb{C}$ forma hermitiana. Entonces, para todo $\mathcal{U} \subseteq V$ sub-ev se tiene que:

1. $\mathcal{U} \subseteq (\mathcal{U}^{\perp})^{\perp}$ y $\dim_{\mathbb{C}}(\mathcal{U}^{\perp}) \geq \dim_{\mathbb{C}}(V) - \dim_{\mathbb{C}}(\mathcal{U})$
2. $\ker(h) = \{0\} \iff h$ es no-degenerada.
3. Si h es no-degenerada, entonces
$$\mathcal{U} = (\mathcal{U}^{\perp})^{\perp} \quad \text{y} \quad \dim_{\mathbb{C}}(\mathcal{U}^{\perp}) = \dim_{\mathbb{C}}(V) - \dim_{\mathbb{C}}(\mathcal{U}).$$
4. Si $\mathcal{U} \cap \mathcal{U}^{\perp} = \{0\}$ entonces $V = \mathcal{U} \oplus \mathcal{U}^{\perp}$ (sin suponer necesariamente que h es no-degenerada; en cuyo caso la condición $\mathcal{U} \cap \mathcal{U}^{\perp} = \{0\}$ siempre se cumple).

Dem: La prueba es idéntica al caso de formas bilineales simétricas (Ejercicio). ∎

§42. Teorema de Sylvester en el caso hermitiano y grupo unitario

Def: Sea $V \cong \mathbb{C}^n$ ev y $h: V \times V \to \mathbb{C}$ forma hermitiana, y sea $Q(x) = h(x,x)$. Decimos que una base $\mathcal{B} = (e_1, \dots, e_n)$ de V es <u>ortogonal</u> respecto a h (o Q) si:

$h(e_j, e_k) = 0$ para todos $j \neq k \iff H = Mat_{\mathcal{B}}(h) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \in M_n(\mathbb{R})$ diagonal <u>real</u>

$\iff Q(z_1, \dots, z_n) = \lambda_1 |z_1|^2 + \dots + \lambda_n |z_n|^2$, donde $(z_1, \dots, z_n)$ son las coord. resp. a $\mathcal{B}$.

Lema: Toda forma hermitiana admite una base ortogonal.

Dem: Por inducción en $n = \dim_{\mathbb{C}}(V)$: OK si $n = 1$ o si $h = 0$. Sup $h \neq 0$:
$\Rightarrow \exists\, e_1 \in V$ tq $Q(e_1) \neq 0$. Sea $L = Vect_{\mathbb{C}}(e_1)$ y $H = L^{\perp}$. Dado que $h(e_1, e_1) \neq 0$ tenemos que $L \cap L^{\perp} = L \cap H = \{0\} \Rightarrow V = L \oplus H$.
$\overset{\text{Hip. Ind}}{\Rightarrow} \exists$ base $(e_2, \dots, e_n)$ de H tq $h(e_j, e_k) = 0$ para todos $j \neq k$ en $\{2, \dots, n\}$.
Luego, $\mathcal{B} = (e_1, e_2, \dots, e_n)$ base de V ortogonal resp. a h ✓ ■

Teorema (Sylvester): Sea $V \cong \mathbb{C}^n$ ev y sea $Q: V \to \mathbb{R}$ forma cuadrática hermitiana asociada a la forma hermitiana $h: V \times V \to \mathbb{C}$. Sea $\mathcal{B} = (e_1, \dots, e_n)$ una base de V ortogonal resp. a h, y sea $H = Mat_{\mathcal{B}}(h)$ matriz diagonal <u>real</u> de coef. diagonales $\lambda_1, \dots, \lambda_n \in \mathbb{R}$. Entonces:

① Módulo permutación de índices, podemos suponer que $\lambda_1, \dots, \lambda_r \neq 0$ y que $\lambda_j = 0$ si $j > r$. En part, si $(z_1, \dots, z_n)$ son las coordenadas respecto a $\mathcal{B}$ entonces:

$$Q(z_1, \dots, z_n) = \lambda_1 |z_1|^2 + \dots + \lambda_r |z_r|^2.$$

② Sea $p \in \mathbb{N}$ (resp. $q \in \mathbb{N}$) el número de índices $j \in \{1, \dots, n\}$ tales que $Q(e_j) > 0$ (resp. $Q(e_j) < 0$). Entonces, el par (p,q) no depende de la base ortogonal escogida y se llama la <u>signatura</u> de Q (o de h). Más aún, $p + q = r = rg(h)$.

③ $\ker(h) = Vect_{\mathbb{C}}(e_{r+1}, \dots, e_n)$ está dado por las ecuaciones $z_1 = \dots = z_r = 0$.

④ Más aún, podemos escoger $\mathcal{B}$ tal que

$$H = \left(\begin{array}{c|c|c} I_p & 0 & 0 \\ \hline 0 & -I_q & 0 \\ \hline 0 & 0 & 0_{n-r} \end{array}\right)$$

Dem: La demostración es completamente análoga al caso de formas cuádricas reales (ver §31) y queda como Ejercicio ■

⚠ <u>Obs importante</u>: Sea $V \cong \mathbb{C}^n$ ev y $h: V \times V \to \mathbb{C}$ forma hermitiana. Sea $u: V \to V$ endomorfismo. Entonces, la fórmula de polarización implica que:

① u preserva h, ie, $h(u(x), u(y)) = h(x,y)$ $\forall x, y \in V$.

② u preserva Q, ie, $Q(u(x)) = Q(x)$ $\forall x \in V$

son propiedades equivalentes.

Def: Sea $V \cong \mathbb{C}^n$ ev y sea $Q: V \to \mathbb{R}$ forma cuadrática hermitiana. Diremos que un automorfismo $u: V \xrightarrow{\sim} V$ es <u>unitario</u> resp. a Q si u preserva Q ($\circ$ h). El sub-grupo de $GL(V)$ dado por

$$\mathcal{U}(Q) := \{ u \in GL(V) \text{ tq } u \text{ es unitario resp. a } Q \} \subseteq GL(V)$$

es llamado el <u>grupo unitario de Q</u>. Más aún, el sub-grupo de $\mathcal{U}(Q)$ dado por

$$SU(Q) = \{ u \in \mathcal{U}(Q) \text{ tq } \det(u) = 1 \} \subseteq \mathcal{U}(Q)$$

es llamado el <u>grupo especial unitario de Q</u>.

Ejercicio Probar que si Q es <u>no-degenerada</u> y u preserva Q ($\circ$ h), entonces $u \in GL(V)$ es un automorfismo.

Prop: Sea $V \cong \mathbb{C}^n$ ev y sea $Q: V \to \mathbb{R}$ forma cuadrática hermitiana asociada a h. Sea $\mathcal{B}$ una base de V y sea $H = Mat_{\mathcal{B}}(h)$. Dado un automorfismo $u: V \xrightarrow{\sim} V$ con $A = Mat_{\mathcal{B}}(u) \in GL_n(\mathbb{C})$, tenemos que:

u es <u>unitario</u> (i.e., $u \in \mathcal{U}(Q)$) $\iff {}^tA H \overline{A} = H$ $(\iff A^* \overline{H} A = \overline{H})$.

En part, si Q es no-degenerada y $u \in \mathcal{U}(Q)$, entonces $|\det(u)| = 1$.

Dem: Si $X, Y \in \mathbb{C}^n$ son los vectores coord. de $x, y \in V$ resp. a la base $\mathcal{B}$, entonces $h(x,y) = {}^tX H \overline{Y} \Rightarrow h(u(x), u(y)) = {}^t(AX) H (\overline{AY}) = {}^tX ({}^tA H \overline{A}) Y$, por lo que $u \in \mathcal{U}(Q)$ si y sólo si $H = {}^tA H \overline{A}$ ✓

En part, si h no-degenerada $\Rightarrow \det(H) \neq 0$ y luego la relación $H = {}^tA H \overline{A}$ implica $\det(H) = \det(A)\det(H)\det(\overline{A}) \Rightarrow \det(A)\det(\overline{A}) = |\det(A)|^2 = 1$ ■

<u>Caso particular importante</u>: El Teorema de Sylvester reduce el estudio de formas cuadráticas hermitianas de $V \cong \mathbb{C}^n$ <u>no-degeneradas</u> de signatura (p,q) a estudiar

$$Q(z_1, \dots, z_n) = |z_1|^2 + \dots + |z_p|^2 - |z_{p+1}|^2 - \dots - |z_{p+q}|^2 \qquad (p+q=n).$$

El grupo unitario (resp. especial unitario) correspondiente es denotado $\mathcal{U}(p,q)$ (resp. $SU(p,q)$). Explícitamente, si consideramos

$$H_{p,q} := \begin{pmatrix} I_p & 0 \\ 0 & -I_q \end{pmatrix} \in GL_n(\mathbb{C})$$

entonces $\mathcal{U}(p,q) \cong \{ A \in GL_n(\mathbb{C}) \text{ tq } {}^tA H_{p,q} \overline{A} = H_{p,q} \}$ y $SU(p,q) \cong \{ A \in \mathcal{U}(p,q) \text{ tq } \det(A) = 1 \}$

Si $(p,q) = (n,0)$ decimos que Q es una forma cuadrática hermitiana <u>definida positiva</u>, y escribimos $\mathcal{U}(n)$ (resp. $SU(n)$) en lugar de $\mathcal{U}(n,0)$ (resp. $SU(n,0)$).

Explícitamente:

$$\mathcal{U}(n) \cong \{ A \in GL_n(\mathbb{C}) \text{ tq } {}^tA\overline{A} = I_n \} = \{ A \in GL_n(\mathbb{C}) \text{ tq } A^*A = I_n \}.$$

§43. Productos escalares complejos y espacios hermitianos.

Def: Sea $V \cong \mathbb{C}^n$ ev. Un *producto escalar complejo* es una forma hermitiana $h: V \times V \to \mathbb{C}$ definida positiva. Un *espacio hermitiano* es un espacio vectorial complejo de dimensión finita dotado de un producto escalar complejo.

Notación: El producto escalar entre dos vectores $x, y \in V$ se denota $\langle x, y \rangle$. Luego, la forma hermitiana $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}$ verifica: Para todos $x, y, z \in V$ y $\lambda \in \mathbb{C}$

1. Lineal en 1ra variable: $\langle \lambda x + y, z \rangle = \lambda \langle x, z \rangle + \langle y, z \rangle$.
2. Anti-lineal en 2da variable: $\langle x, \lambda y + z \rangle = \overline{\lambda} \langle x, y \rangle + \langle x, z \rangle$.
3. Simetría hermitiana: $\langle x, y \rangle = \overline{\langle y, x \rangle}$
4. Definida positiva: $\langle x, x \rangle \in \mathbb{R}^{>0}$ para todo $x \neq 0$.

Ejemplo principal (cf. Teorema de Sylvester): Para $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n)$ en $\mathbb{C}^n$

$$\langle x, y \rangle := x_1 \overline{y_1} + \dots + x_n \overline{y_n} \quad \text{"producto escalar canónico"}.$$

Def: Sea $V \cong \mathbb{C}^n$ espacio hermitiano, y sea $\mathcal{B} = (e_1, \dots, e_n)$ una base de V. Decimos que $\mathcal{B}$ es una *base ortonormal* si $\langle e_j, e_k \rangle = \delta_{jk} = \begin{cases} 1 & \text{si } j = k \\ 0 & \text{si } j \neq k \end{cases}$.

Obs: El Teorema de Sylvester implica que todo espacio hermitiano posee al menos una base ortonormal.

Teorema (proyección ortogonal): Sea $V \cong \mathbb{C}^n$ espacio hermitiano y sea $\mathcal{U}^\perp$ el ortogonal (resp. al producto escalar) de un sub-ev $\mathcal{U} \subseteq V$. Entonces:

1. $V = \mathcal{U} \oplus \mathcal{U}^\perp$ y la aplicación $p_{\mathcal{U}}: V \to V$ inducida es lineal y verifica $\operatorname{Im}(p_{\mathcal{U}}) = \mathcal{U}$ y $\ker(p_{\mathcal{U}}) = \mathcal{U}^\perp$, es la *proyección ortogonal* sobre $\mathcal{U}$.
2. Si $\dim_{\mathbb{C}}(\mathcal{U}) = r$ y $(e_1, \dots, e_r)$ base ortonormal de $\mathcal{U}$, entonces:
$$p_{\mathcal{U}}(x) = \langle x, e_1 \rangle e_1 + \dots + \langle x, e_r \rangle e_r \quad \text{para todo } x \in V.$$
3. $(\mathcal{U}^\perp)^\perp$ y en particular $p_{\mathcal{U}} + p_{\mathcal{U}^\perp} = \mathrm{Id}_V$.

Dem: Idéntica al caso euclideano; ver §35 y §36. ∎

Notación: Sea $V \cong \mathbb{C}^n$ espacio hermitiano. Para todo $x \in V$ definimos la *norma hermitiana* de x por $\|x\| := \sqrt{Q(x)}$, donde $Q(x) = \langle x, x \rangle \in \mathbb{R}^{\geq 0}$.

Notar que $\|x\| = 0 \Leftrightarrow x = 0$ (pues Q definida positiva) y para todo $\lambda \in \mathbb{C}$ se tiene: $\|\lambda x\| = \sqrt{Q(\lambda x)} = \sqrt{|\lambda|^2 Q(x)} = |\lambda| \sqrt{Q(x)} = |\lambda| \|x\|$. Más aún:

Teorema (desigualdad de Cauchy-Schwarz): Sea $V \cong \mathbb{C}^n$ espacio hermitiano. Entonces:
$|\langle x, y \rangle| \leq \|x\| \|y\|$ $\forall x, y \in V$; con igualdad $\Leftrightarrow$ x e y son colineales.
En part, $\|x + y\| \leq \|x\| + \|y\|$ para todos $x, y \in V$ (desigualdad triangular).

Dem: Sabemos que $Q(x) = \langle x, x \rangle = \operatorname{Re} \langle x, x \rangle$ está asociada a la forma bilineal real $B(x, y) := \operatorname{Re} \langle x, y \rangle$ en $V \cong \mathbb{R}^{2n}$. En part, Cauchy-Schwarz real implica:

$|\mathrm{Re}\langle x,y\rangle| \le \sqrt{Q(x)}\cdot\sqrt{Q(y)}$. Consideremos $z = \langle x,y\rangle \in \mathbb{C}$ y escribamos $\langle x,y\rangle = r e^{i\theta}$, con $r = |\langle x,y\rangle|$ y $\theta \in \mathbb{R}$. Entonces, dado que $\overline{e^{i\theta}} = e^{-i\theta}$:

$\Rightarrow r = e^{-i\theta}\langle x,y\rangle = \langle x, e^{i\theta}y\rangle = \mathrm{Re}\langle x, e^{i\theta}y\rangle$. Por lo tanto:

$$|\langle x,y\rangle| = |\mathrm{Re}\langle x,e^{i\theta}y\rangle| \le \sqrt{Q(x)}\sqrt{Q(e^{i\theta}y)} = \sqrt{Q(x)}\sqrt{\underbrace{|e^{i\theta}|^2}_{=1}Q(y)} = \sqrt{Q(x)}\cdot\sqrt{Q(y)}.$$

Más aún, Cauchy-Schwarz real implica que la igualdad ocurre $\Leftrightarrow$ x e $e^{i\theta}y$ son $\mathbb{R}$-lin. dep, y en particular serían $\mathbb{C}$-lin. dep. Recíprocamente, si x e y son $\mathbb{C}$-lin. dep. entonces se calcula directamente que la igualdad ocurre. ✓

Finalmente, dado que para todo $z = a+ib \in \mathbb{C}$ se tiene $\mathrm{Re}(z) = a \le \sqrt{a^2+b^2} = |z|$:

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2\underbrace{\mathrm{Re}\langle x,y\rangle}_{\le |\langle x,y\rangle|} \overset{CS}{\le} \|x\|^2 + \|y\|^2 + 2\|x\|\|y\| = (\|x\|+\|y\|)^2 \quad ✓ \ \blacksquare$$

Ejemplo principal: La norma hermitiana (canónica) en $\mathbb{C}^n$ está dada por

$$\|(z_1,\dots,z_n)\| = \sqrt{|z_1|^2 + \dots + |z_n|^2}.$$

⚠ Importante: En un espacio hermitiano, las fórmulas de polarización son:

① $\mathrm{Re}\langle x,y\rangle = \frac{1}{2}(\|x+y\|^2 - \|x\|^2 - \|y\|^2) = \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2)$.

② $\mathrm{Im}\langle x,y\rangle = \frac{1}{2}(\|x+iy\|^2 - \|x\|^2 - \|y\|^2) = \frac{1}{4}(\|x+iy\|^2 - \|x-iy\|^2)$.

$$\Rightarrow \langle x,y\rangle = \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2) \text{ para todos } x,y \in V \cong \mathbb{C}^n$$

Ejercicio Sea $V \cong \mathbb{C}^n$ espacio hermitiano y sean $x_1,\dots,x_m \in V$ ortogonales. Probar que:

$$\|x_1+\dots+x_m\|^2 = \|x_1\|^2 + \dots + \|x_m\|^2 \quad (\text{"Pitágoras"})$$

Prop: Sea $V \cong \mathbb{C}^n$ espacio hermitiano y sea $u: V \to V$ endomorfismo. Son equivalentes:

① u preserva la norma: $\|u(x)\| = \|x\| \ \forall x \in V$.

② u preserva el producto escalar: $\langle u(x), u(y)\rangle = \langle x,y\rangle \ \forall x,y \in V$.

③ Para toda base ortonormal $\mathcal{B}$ de V, la imagen $u(\mathcal{B})$ es una base ortonormal.

④ Existe una base ortonormal $\mathcal{B}$ de V tq $u(\mathcal{B})$ es una base ortonormal.

⑤ La matriz $A = \mathrm{Mat}_{\mathcal{B}}(u)$ resp. a una base ortonormal $\mathcal{B}$ de V es una matriz unitaria, i.e., $A^*A = I_n$ $(\Leftrightarrow {}^tA\bar{A} = I_n)$.

En tal caso, u es necesariamente biyectivo (i.e., $u \in GL(V)$), y diremos que u es una isometría de V. En part, el conjunto de isometrías de $V \cong \mathbb{C}^n$ se identifica con el grupo unitario $\mathcal{U}(n)$.

Dem: Idéntica al caso de isometrías de un espacio euclideano: ver §35. ■

Ejercicio Sea $A \in M_n(\mathbb{C})$ matriz unitaria. Probar que las columnas de A forman una base ortonormal de $\mathbb{C}^n$.

Teorema (adjunción): Sea $V \cong \mathbb{C}^n$ espacio hermitiano y sea $u: V \to V$ endomorfismo. Entonces, existe un único **endomorfismo adjunto** $u^*: V \to V$ verificando

$$\langle u(x), y\rangle = \langle x, u^*(y)\rangle \quad \text{para todos } x, y \in V.$$

Más aún, para toda base **ortonormal** $\mathcal{B}$ de V se tiene que:

$$\mathrm{Mat}_{\mathcal{B}}(u^*) = \mathrm{Mat}_{\mathcal{B}}(u)^* \underset{\text{dy}}{=} {}^t\overline{\mathrm{Mat}_{\mathcal{B}}(u)}.$$

En particular, $(u^*)^* = u$.

Dem: Supongamos que existe $u^*: V \to V$ verificando $\langle u(x), y\rangle = \langle x, u^*(y)\rangle$ $\forall x, y \in V$, y sea $\mathcal{B}$ una base ortonormal de V. Entonces, $H = \mathrm{Mat}_{\mathcal{B}}(\langle\cdot,\cdot\rangle) = I_n$. Sean $A = \mathrm{Mat}_{\mathcal{B}}(u)$ y $B = \mathrm{Mat}_{\mathcal{B}}(u^*)$. Si $X, Y \in \mathbb{C}^n$ son las coord. de $x, y \in V$ respecto a $\mathcal{B}$, entonces: $\langle x, y\rangle = {}^tXH\overline{Y} = {}^tX\overline{Y}$ y luego

$$\langle u(x), y\rangle = {}^t(AX)\overline{Y} = {}^tX\,{}^tA\overline{Y} \underset{\text{Hipótesis}}{=} \langle x, u^*(y)\rangle = {}^tX\,\overline{(BY)} = {}^tX\overline{B}\,\overline{Y} \Rightarrow {}^tA = \overline{B}$$

En otras palabras, $B = {}^t\overline{A} = A^*$. Así, si u^* existe entonces necesariamente se cumple $\mathrm{Mat}_{\mathcal{B}}(u^*) = \mathrm{Mat}_{\mathcal{B}}(u)^*$ y luego está determinado únicamente por u. Recíprocamente: sea $u^*: V \to V$ el endomorfismo tq $\mathrm{Mat}_{\mathcal{B}}(u^*) = A^*$, entonces $\forall x, y \in V$:

$$\langle x, u^*(y)\rangle = {}^tX\,\overline{(A^*Y)} = {}^tX\,\overline{A^*}\,\overline{Y} = {}^tX\,{}^tA\overline{Y} = {}^t(AX)\overline{Y} = \langle u(x), y\rangle \quad \checkmark$$

Finalmente, el hecho que $(A^*)^* = A$ equivale a $(u^*)^* = u$. ■

El siguiente resultado será útil para estudiar endomorfismos diagonalizables:

Lema: Sea $V \cong \mathbb{C}^n$ espacio hermitiano y sea $u: V \to V$ endomorfismo. Entonces, para todo sub-esp $\mathcal{U} \subseteq V$ se tiene que:
$u(\mathcal{U}) \subseteq \mathcal{U}$ (ie, $\mathcal{U}$ es estable por u) $\iff$ $u^*(\mathcal{U}^\perp) \subseteq \mathcal{U}^\perp$ (ie, $\mathcal{U}^\perp$ es estable por u^*).

Dem: Sup. que $u(\mathcal{U}) \subseteq \mathcal{U}$ y sea $y \in \mathcal{U}^\perp$. Veamos que $u^*(y) \in \mathcal{U}^\perp$: Para todo $x \in \mathcal{U}$

$$\langle x, u^*(y)\rangle \underset{\text{dy}}{=} \langle \underbrace{u(x)}_{\in\mathcal{U}}, \underbrace{y}_{\in\mathcal{U}^\perp}\rangle = 0 \Rightarrow u^*(y) \in \mathcal{U}^\perp.$$

Recíprocamente, si $u^*(\mathcal{U}^\perp) \subseteq \mathcal{U}^\perp$ entonces el argumento anterior (reemplazando u por u^* y $\mathcal{U}$ por $\mathcal{U}^\perp$) implica $(u^*)^*((\mathcal{U}^\perp)^\perp) \subseteq (\mathcal{U}^\perp)^\perp \iff u(\mathcal{U}) \subseteq \mathcal{U}$. ■

Def: Sea $V \cong \mathbb{C}^n$ espacio hermitiano y sea $u: V \to V$ endomorfismo. Diremos que u es:

1. **auto-adjunto** (o **hermitiano**) si $u^* = u$.
2. **anti-hermitiano** si $u^* = -u$.
3. **unitario** si $uu^* = \mathrm{Id}_V$ (ie, $u \in GL(V)$ y $u^{-1} = u^*$).

Más generalmente, diremos que u es un endomorfismo **normal** si conmuta con su adjunto u^*, ie, $uu^* = u^*u$. En part, los endomorfismos normales engloban los tres casos anteriores.

Ejercicio Dar un ejemplo de una matriz $A \in M_2(\mathbb{C})$ que sea normal (ie, $AA^* = A^*A$), pero que no sea hermitiana, ni anti-hermitiana, ni unitaria.

Teorema espectral: Sea $V \cong \mathbb{C}^n$ espacio hermitiano y sea $u: V \to V$ endomorfismo **normal**. Entonces, u es diagonalizable y los espacios propios asociados a valores propios distintos son ortogonales. En part, existe una base ortonormal $\mathcal{B}$ de V formada por vectores propios de u. Más aún:

① Si u es **hermitiano**, sus valores propios $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ son **reales**:
$$Mat_{\mathcal{B}}(u) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_m \end{pmatrix} \in M_m(\mathbb{R}).$$

② Si u es **anti-hermitiano**, sus valores propios $i\lambda_1, \dots, i\lambda_m \in i\mathbb{R}$ son **imaginarios puros**:
$$Mat_{\mathcal{B}}(u) = \begin{pmatrix} i\lambda_1 & & 0 \\ & \ddots & \\ 0 & & i\lambda_m \end{pmatrix} \in M_m(\mathbb{R}), \text{ con } \lambda_1, \dots, \lambda_m \in \mathbb{R}.$$

③ Si u es **unitario**, sus valores propios $\lambda_1, \dots, \lambda_m \in \mathbb{C}$ son de módulo 1:
$$Mat_{\mathcal{B}}(u) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_m \end{pmatrix} \in GL_m(\mathbb{C}), \text{ con } |\lambda_1| = \dots = |\lambda_m| = 1.$$

Dem: Por inducción en $n = \dim_{\mathbb{C}}(V)$: OK si $n = 1$ ✓ Sup. $n \geq 2$: Dado que $\mathbb{C}$ es algebraicamente cerrado, $P_u(X)$ posee al menos una raíz $\lambda \in \mathbb{C}$.

Sea V_λ el espacio propio asociado a λ, entonces $u^*(V_\lambda) \subseteq V_\lambda$. En efecto:

Para todo $x \in V_\lambda$: $u(u^*(x)) \overset{\text{Hip.}}{=} u^*(u(x)) \overset{x \in V_\lambda}{=} u^*(\lambda x) = \lambda u^*(x)$, ie, $u^*(x) \in V_\lambda$.

Luego, el Lema anterior implica que $V_\lambda^\perp$ es estable por u y por u^*. Por otro lado, $V = V_\lambda \oplus V_\lambda^\perp$ y $\dim_{\mathbb{C}}(V_\lambda) > 0 \Rightarrow \dim_{\mathbb{C}}(V_\lambda^\perp) < \dim_{\mathbb{C}}(V)$.

Denotemos $v := u|_{V_\lambda^\perp}: V_\lambda^\perp \to V_\lambda^\perp$ y $v^* := u^*|_{V_\lambda^\perp}: V_\lambda^\perp \to V_\lambda^\perp$. Así, para todos $x, y \in V_\lambda^\perp$ se tiene que $v(x) = u(x)$ y $v^*(x) = u^*(x)$

$\Rightarrow \langle v(x), y \rangle = \langle u(x), y \rangle = \langle x, u^*(y) \rangle = \langle x, v^*(y) \rangle$, ie, v^* adjunto de v.

Dado que u y u^* conmutan, v y v^* también, ie, 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. Como $V = V_\lambda \oplus V_\lambda^\perp$, esto último es cierto también para u ✓

En part, considerando bases ortonormales para cada espacio propio de u, obtenemos que existe una base ortonormal $\mathcal{B}$ de V formada por vectores propios de u.

Finalmente, sea $D = Mat_{\mathcal{B}}(u) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_m \end{pmatrix}$ matriz diagonal. Entonces si u es:

① hermitiano $\Rightarrow D^* = \overline{D} = D$, ie, $\lambda_j \in \mathbb{R}$ $\forall j$.

② anti-hermitiano $\Rightarrow D^* = \overline{D} = -D$, ie, $\lambda_j \in i\mathbb{R}$ imaginario puro $\forall j$.

③ unitario $\Rightarrow D^* D = \overline{D} D = I_m$, ie, $|\lambda_j| = 1$ $\forall j$. ■

⚠ **Importante**: El Teorema espectral también cubre el caso de matrices **reales** simétricas, anti-simétricas y ortogonales. En efecto, si $A \in M_m(\mathbb{R})$ 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).

§44. Producto tensorial de espacios vectoriales:

Motivación: Sea k un cuerpo arbitrario. El objetivo de los "tensores" es describir formas multilineales (ie, lineales en cada variable) $T: V_1 \times \dots \times V_p \longrightarrow W$, donde $V_1, \dots, V_p, W$ son k-ev.

Los tensores son muy útiles en matemática y en física, y fueron introducidos en 1846 por Hamilton. Comencemos por el caso "2-multilineal", ie, bilineal:

Def: Sean V y W dos k-ev. Un producto tensorial de V y W es un par (T, t), donde T es un k-ev. y una aplicación bilineal $t: V \times W \to T$ verificando la siguiente "propiedad universal":

"Para todo k-ev $\mathcal{U}$ y toda aplicación bilineal $B: V \times W \to \mathcal{U}$, existe una única aplicación lineal $\hat{B}: T \to \mathcal{U}$ tal que $B = \hat{B} \circ t$."

En otras palabras, el diagrama

$$\begin{array}{ccc} V \times W & \xrightarrow{B} & \mathcal{U} \\ & {\scriptstyle t}\searrow \quad \nearrow {\scriptstyle \exists! \hat{B}} & \\ & T & \end{array}$$

es conmutativo.

Teorema: Sean V y W dos k-ev. Entonces, existe un producto tensorial (T, t) de V y W, el cual es único módulo un único isomorfismo. Denotaremos $T := V \otimes_k W$ (o simplemente $V \otimes W$) y para $v \in V$ y $w \in W$, $t(v, w) := v \otimes w$.

Dem: Veamos la existencia: Sea $k^{V \times W}$ el k-ev de base (no necesariamente finita!) $\{e_{(v,w)}\}_{(v,w) \in V \times W}$, ie, un elemento de $k^{V \times W}$ es de la forma:

$$\sum_{\text{finita}} \lambda_{(v,w)} e_{(v,w)}, \quad \text{donde } \lambda_{(v,w)} \in k.$$

Importante: la aplicación $V \times W \to k^{V \times W}$, $(v, w) \mapsto e_{(v,w)}$ no es bilineal! Sin embargo, si $S \subseteq k^{V \times W}$ es el k-sub-ev generado por todos los vectores de la forma:

$e_{(v+v', w)} - e_{(v,w)} - e_{(v',w)}$; $e_{(v, w+w')} - e_{(v,w)} - e_{(v,w')}$; $e_{(\lambda v, w)} - \lambda e_{(v,w)}$ y $e_{(v, \lambda w)} - \lambda e_{(v,w)}$, donde $v, v' \in V$, $w, w' \in W$ y $\lambda \in k$.

$\Rightarrow$ El k-ev. cociente $T := k^{V \times W}/S$ dotado de la aplicación bilineal

$$t: V \times W \to T, \quad (v, w) \mapsto [e_{(v,w)}]$$

es un producto tensorial de V y W. En efecto: si denotamos $V \otimes W := T$ y $v \otimes w := t(v, w) = [e_{(v,w)}]$, entonces para toda aplicación $B: V \times W \to \mathcal{U}$ bilineal se tiene que la aplicación inducida $\tilde{B}: k^{V \times W} \to \mathcal{U}$, $e_{(v,w)} \mapsto B(v, w)$ es lineal y, por construcción (!), se anula en S: $\tilde{B}(S) = \{0_{\mathcal{U}}\}$.

Propiedad universal del cociente $\Rightarrow \exists! \hat{B}: k^{V \times W}/S = T \to \mathcal{U}$ tal que $B = \hat{B} \circ t$.

Explícitamente: si $v \otimes w \in T$, entonces $\hat{B}(v \otimes w) = B(v, w)$.

Unicidad (módulo únicos isomorfismo): Usamos la propiedad universal de (T,t): si (T',t') es otro producto tensorial de V y W, entonces:

$$\begin{array}{ccc} V\times W & \xrightarrow{t'} & T' \\ & \searrow^{t} \quad \nearrow_{\exists!\,\alpha} & \\ & T & \end{array} \qquad \begin{array}{ccc} V\times W & \xrightarrow{t} & T \\ & \searrow^{t'} \quad \nearrow_{\exists!\,\beta} & \\ & T' & \end{array}$$

donde $t' = \alpha\circ t$ y $t = \beta\circ t' \Rightarrow t' = \alpha\circ t = (\alpha\circ\beta)\circ t'$ y $t = \beta\circ t' = (\beta\circ\alpha)\circ t$. Como α y β son únicos, entonces $\alpha\circ\beta$ y $\beta\circ\alpha$ son únicos. Luego, necesariamente $\alpha\circ\beta = Id_{V'}$ y $\beta\circ\alpha = Id_V$, ie, α y β son isomorfismos ✓ ∎

Recuerdo (notación): Sean $\mathcal{U}, V, W$ tres k-evs. Recordemos que:

$Hom_k(V,W) = \{f: V\to W \text{ lineal}\}$, $V^* = Hom_k(V,k)$ (espacio dual) y

$Bil_k(V\times W, \mathcal{U}) := \{B: V\times W \to \mathcal{U} \text{ aplicación } k\text{-bilineal}\}$ son k-evs.

Corolario: Sean $\mathcal{U}, V, W$ tres k-evs. Entonces, hay un isomorfismo:

$$Bil_k(V\times W, \mathcal{U}) \cong Hom_k(V\otimes W, \mathcal{U}).$$

En part, $Bil_k(V\times W, k) = \{B: V\times W \to k \text{ forma bilineal}\} \cong (V\otimes W)^*$.

Dem: La aplicación $B \mapsto \hat{B}$ es un isomorfismo. ∎

Ejemplo: Sea $V = \mathbb{R}^2$ y $\mathcal{B} = (e_1, e_2)$ base canónica. La forma bilineal $B: V\times V\to\mathbb{R}$ dada por $B((x_1,y_1),(x_2,y_2)) = ax_1y_1 + bx_1y_2 + cx_2y_1 + dx_2y_2$ está determinada por $a = B(e_1,e_1)$, $b = B(e_1,e_2)$, $c = B(e_2,e_1)$ y $d = B(e_2,e_2)$. Por otro lado: $V\otimes V = Vect_{\mathbb{R}}\langle e_1\otimes e_1, e_1\otimes e_2, e_2\otimes e_1, e_2\otimes e_2\rangle \cong \mathbb{R}^4$ donde por definición se cumple: $\lambda(e_i\otimes e_j) = (\lambda e_i)\otimes e_j = e_i\otimes(\lambda e_j)$ $\forall \lambda\in\mathbb{R}$.

En part, $(x_i e_i \otimes y_j e_j) = x_i y_j\,(e_i\otimes e_j)$ para $x_i, y_j\in\mathbb{R}$ y luego

$(x_1e_1 + y_1e_2)\otimes(x_2e_1 + y_2e_2) = x_1x_2(e_1\otimes e_1) + x_1y_2(e_1\otimes e_2) + x_2y_1(e_2\otimes e_1) + x_2y_2(e_2\otimes e_2)$

Finalmente, $f: V\otimes V\to\mathbb{R}$ en $(V\otimes V)^*$ está determinada por los valores $a' = f(e_1\otimes e_1)$, $b' = f(e_1\otimes e_2)$, $c' = f(e_2\otimes e_1)$ y $d' = f(e_2\otimes e_2)$.

$\Rightarrow f\in(V\otimes V)^*$ y $B\in Bil(V\times V,\mathbb{R})$ son "lo mismo".

⚠ Importante: En general, un vector en $V\otimes_k W$ no es de la forma $v\otimes w$ con $v\in V$ y $w\in W$, sino que es una suma finita de la forma

$$\lambda_1(v_1\otimes w_1) + \lambda_2(v_2\otimes w_2) + \ldots + \lambda_r(v_r\otimes w_r) \quad \text{con } v_i\in V,\ w_i\in W,\ \lambda_i\in k.$$

Los elementos de $V\otimes W$ serán 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; ellos generan el k-ev. $V\otimes_k W$ por construcción.

Ejercicio: 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 $T = ae_1\otimes f_1 + be_1\otimes f_2 + ce_2\otimes f_1 + de_2\otimes f_2 \in V\otimes W$ es un tensor simple si y sólo si $ad - bc = 0$.

Lema: Sean V y W dos k-ev. 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 part, si V y W son de dimensión finita, entonces:

$$\dim_k(V\otimes W) = \dim_k(V)\cdot \dim_k(W).$$

Dem: Por construcción de $V\otimes W$, $\{v_i\otimes w_j\}_{(i,j)\in I\times J}$ es una familia generadora. Veamos que el 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 $f_i(v_k) = \delta_{ik} = \begin{cases} 1 & \text{si } i=k \\ 0 & \text{si } i\neq k \end{cases}$ y por $g_j(w_k) = \delta_{jk} = \begin{cases} 1 & \text{si } j=k \\ 0 & \text{si } j\neq k \end{cases}$. Ahora, consideremos $\varphi_{ij}: V\times W \to k$ forma bilineal dada por $\varphi_{ij}(v,w) = f_i(v)\,g_j(w) \overset{\text{Prop. universal}}{\Longrightarrow} \exists!\ \widehat{\varphi}_{ij}: V\otimes W \to k$ lineal tal que

$$\widehat{\varphi}_{ij}(v_k\otimes w_l) = \varphi_{ij}(v_k,w_l) = f_i(v_k)\,g_j(w_l) = \delta_{ik}\,\delta_{jl} = \begin{cases} 1 & \text{si } i=k \text{ y } j=l \\ 0 & \text{sino.} \end{cases}$$

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 estos son linealmente independientes ✓ ■

El siguiente resultados muestra que podemos tensorizar aplicaciones lineales:

Prop (functorialidad): Sean $f: V\to V'$ y $g: W\to W'$ aplicaciones lineales entre k-ev. Entonces, existe una única aplicación lineal

$$f\otimes g: V\otimes W \to V'\otimes W'$$

que para todo tensor simple $v\otimes w$ verifica $(f\otimes g)(v\otimes w) = f(v)\otimes g(w)$. Más aún, el producto tensorial de aplicaciones lineales verifica:

$$(f'\otimes g')\circ(f\otimes g) = (f'\circ f)\otimes(g\circ g').$$

Dem: Basta completar el diagrama conmutativo

$$\begin{array}{ccc} V\times W & \xrightarrow{(f,g)} & V'\times W' \\ {\scriptstyle t}\downarrow & \overset{\text{bilineal}}{\searrow} & \downarrow{\scriptstyle t'} \\ V\otimes W & \underset{\exists!\ f\otimes g}{\dashrightarrow} & V'\otimes W' \end{array}$$

Es decir, $f\otimes g$ se obtiene al aplicar la propiedad universal a la aplicación bilineal $t'\circ(f,g)$. La última parte se deduce del mismo modo (Ejercicio). ■

A continuación describimos las principales propiedades del producto tensorial:

Teorema: Sean U, V y W tres k-ev. Entonces, hay isomorfismos "canónicos":

1. $k\otimes V \xrightarrow{\sim} V$, dado por $\lambda\otimes v \mapsto \lambda v$.
2. $(U\oplus V)\otimes W \xrightarrow{\sim} (U\otimes W)\oplus(V\otimes W)$, dado por $(u+v)\otimes w \mapsto u\otimes w + v\otimes w$.
3. $U\otimes V \xrightarrow{\sim} V\otimes U$, dado por $u\otimes v \mapsto v\otimes u$.
4. $U\otimes(V\otimes W) \xrightarrow{\sim} (U\otimes V)\otimes W$, dado por $u\otimes(v\otimes w)\mapsto(u\otimes v)\otimes w$.

Aquí, la fórmula de cada isomorfismo está descrita para tensores simples, y se extiende linealmente para tensores arbitrarios (ie, suma finita de tensores simples).

Dem: La aplicación $k\times V\to V$, $(\lambda, v)\mapsto \lambda v$ es bilineal $\Rightarrow \exists!$ aplicación lineal inducida $k\otimes V\to V$, $\lambda\otimes v\mapsto \lambda v$. Su inversa es $v\mapsto 1\otimes v$. $\rightsquigarrow$ (1) ✓

Para probar ② consideramos la aplicación bilineal $(U\oplus V)\times W \longrightarrow (U\otimes W)\oplus(V\otimes W)$
$(u+v, w) \longmapsto (u\otimes w + v\otimes w)$

$\Rightarrow \exists!\ \varphi : (U\oplus V)\otimes W \to (U\otimes W)\oplus(V\otimes W)$ tal que $\varphi((u+v)\otimes w) = u\otimes w + v\otimes w$.

Por otro lado, las inclusiones $i_1 : U \hookrightarrow U\oplus V$, $u \mapsto u+0$ y $i_2 : V \hookrightarrow U\oplus V$, $v \mapsto 0+v$ inducen, al tensorizar con $\mathrm{Id}_W : W \to W$ aplicaciones lineales

$$i_1\otimes \mathrm{Id}_W : U\otimes W \longrightarrow (U\oplus V)\otimes W \quad \text{y} \quad i_2\otimes \mathrm{Id}_W : V\otimes W \to (U\oplus V)\otimes W.$$

$\Rightarrow \psi := (i_1\otimes \mathrm{Id}_W, i_2\otimes \mathrm{Id}_W) : (U\otimes W)\oplus(V\otimes W) \to (U\oplus V)\otimes W$ es la inversa de φ.

En efecto, $\psi(\varphi((u+v)\otimes w)) = \psi(u\otimes w + v\otimes w) = (i_1\otimes \mathrm{Id}_W)(u\otimes w) + (i_2\otimes \mathrm{Id}_W)(v\otimes w)$

$$= (u+0)\otimes w + (0+v)\otimes w = (u+v)\otimes w.$$

i.e., $\psi\circ\varphi = \mathrm{Id}_{(U\oplus V)\otimes W}$, similar para $\varphi\circ\psi$ ✓ Los isomorfismos ③ y ④ son similares. ∎

Ejemplo importante (producto de Kronecker): Sean V_1, V_2, W_1 y W_2 k-ev de dimensión finita. Sean $(v_{1,j})_{j\in I_1}$ base de V_1; $(v_{2,i})_{i\in I_2}$ base de V_2; $(w_{1,\ell})_{\ell\in J_1}$ base de W_1 y sea $(w_{2,\kappa})_{\kappa\in J_2}$ base de W_2.

Si $f : V_1 \to V_2$ y $g : W_1 \to W_2$ son aplicaciones lineales cuyas matrices en las bases respectivas son $A = (a_{ij})_{i\in I_2,\, j\in I_1}$ y $B = (b_{\kappa\ell})_{\kappa\in J_2,\, \ell\in J_1}$ entonces:

$$(f\otimes g)(v_{1,j}\otimes w_{1,\ell}) = f(v_{1,j})\otimes g(w_{1,\ell}) = \sum_{i\in I_2,\, \kappa\in J_2} a_{ij}\, b_{\kappa\ell}\, (v_{2,i}\otimes w_{2,\kappa})$$

Luego, la matriz de $f\otimes g$ respecto a las bases $(v_{1,j}\otimes w_{1,\ell})_{(j,\ell)\in I_1\times J_1}$ de $V_1\otimes W_1$ y $(v_{2,i}\otimes w_{2,\kappa})_{(i,\kappa)\in I_2\times J_2}$ de $V_2\otimes W_2$ es:

$$A\otimes B := (a_{ij}\, b_{\kappa\ell})_{(i,\kappa)\in I_2\times I_2,\ (j,\ell)\in I_1\times J_1}. \quad (\text{"producto de Kronecker"})$$

Concretamente, si por ejemplo todos los espacios V_1, V_2, W_1 y W_2 son de dimensión 2:

$$A\otimes B = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \otimes \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11}B & a_{12}B \\ a_{21}B & a_{22}B \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} & a_{11}b_{12} & a_{12}b_{11} & a_{12}b_{12} \\ a_{11}b_{21} & a_{11}b_{22} & a_{12}b_{21} & a_{12}b_{22} \\ a_{12}b_{11} & a_{12}b_{12} & a_{22}b_{11} & a_{22}b_{12} \\ a_{12}b_{21} & a_{12}b_{22} & a_{22}b_{21} & a_{22}b_{22} \end{pmatrix}$$

En general, $A\otimes B \neq B\otimes A$.

Ejercicio Probar que $\mathrm{rg}(A\otimes B) = \mathrm{rg}(A)\,\mathrm{rg}(B)$. Más aún, probar que si $A\in M_n(k)$ y $B\in M_m(k)$ son matrices cuadradas, entonces $\mathrm{tr}(A\otimes B) = \mathrm{tr}(A)\,\mathrm{tr}(B)$ y $\det(A\otimes B) = \det(A)^m \det(B)^n$.

Ejemplo importante (las matrices son tensores): Veremos que toda matriz puede ser pensada como un tensor. Más generalmente, consideremos la aplicación lineal

$$\varphi : V^*\otimes W \longrightarrow \mathrm{Hom}_k(V, W)$$
$$f\otimes w \longmapsto (v \mapsto f(v)\,w)$$

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 $\sum_{\text{finita}\ j\in J} f_j\otimes w_j$, donde cada $f_j\in V^*$ forma lineal en V.

Si su imagen por φ es nula en $\mathrm{Hom}_k(V,W)$, entonces $\sum_{\text{finita}} f_j(v)w_j = 0 \ \forall v \in V$.
$\Rightarrow f_j(v) = 0 \ \forall v \in V$ para todo j, pues $\{w_j\}_{j\in J}$ es una base de $W \Rightarrow f_j = 0 \ \forall j \in J$ ✓

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 part, si V o W es de dimensión finita $\Rightarrow \varphi: V^* \otimes W \xrightarrow{\sim} \mathrm{Hom}_k(V,W)$ isomorfismo.

Explícitamente, si $\dim_k(V) = n$ y $(v_1, \dots, v_n)$ base de V, de base dual $(v_1^*, \dots, v_n^*)$ 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 ¿Cuál es la imagen por φ del conjunto de tensores simples de $V^* \otimes W$?

Una aplicación típica de los tensores es usarlos para formalizar la intuición de pasar de un k-ev a un K-ev, donde $k \subseteq K$ (eg. $\mathbb{R} \subseteq \mathbb{C}$):

Ejemplo (extensión de escalares): Sean k y K cuerpos tales que $k \subseteq K$. Si V es un k-ev, entonces dado K también es un k-ev podemos considerar

$$V_K := K \otimes_k V \cong V \otimes_k K.$$

$\Rightarrow V_K$ es un k-ev, pero también un K-ev: si $\lambda \in K$, la multiplicación por λ es $m_\lambda : K \to K$, $x \mapsto \lambda x$ 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-ev son fáciles de verificar.
Decimos que V_K se obtiene a partir de V por extensión de escalares de k a K.
Más aún, $\dim_K(V_K) = \dim_k(V)$ pues 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-ev.
Por otro lado, 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 $A \in M_n(k)$ resp. a una base $\{v_i\}$ de V, entonces u_K posee la misma matriz $A \in M_n(K)$ resp. a la base $\{1 \otimes v_i\}$ de V_K.
Caso particular importante: si $k = \mathbb{R}$ y $K = \mathbb{C}$, entonces $V_{\mathbb{C}} = \mathbb{C} \otimes_{\mathbb{R}} V$ es la complejificación del $\mathbb{R}$-ev V. En part, $\mathbb{C} \otimes_{\mathbb{R}} \mathbb{R}^n \cong \mathbb{C}^n$.

Ejercicio Sean V_1 y V_2 dos espacios vectoriales complejos de dimensión finita > 0. Probar que los espacios $V_1 \otimes_{\mathbb{C}} V_2$ y $V_1 \otimes_{\mathbb{R}} V_2$ no son isomorfos como espacios vectoriales reales.

Ejercicio Importante Sean V y W dos k-ev. de dimensión finita. Probar que
$\psi : V^* \otimes W^* \xrightarrow{\sim} (V \otimes W)^*$, $f \otimes g \mapsto (v \otimes w \mapsto f(v)g(w))$ es un isomorfismo.

§45. Álgebra tensorial

Comencemos por una generalización natural del concepto de m-forma multilineal (ver §9):

Def: Sean $V_1, \dots, V_d, \mathcal{U}$ k-evs. Una <u>aplicación d-lineal</u> (o d-multilineal) es una función $f: V_1 \times \dots \times V_d \to \mathcal{U}$ que es lineal en cada variable. Denotaremos $\mathrm{Mult}^d(V_1 \times \dots \times V_d, \mathcal{U}) = \{f: V_1 \times \dots \times V_d \to \mathcal{U} \ d\text{-lineal}\}$ al k-ev correspondiente.

⚠ Siguiendo exactamente el mismo procedimiento que en el caso bilineal (ver §44), dados $V_1, \dots, V_d$ k-evs construimos un k-ev $V_1 \otimes \dots \otimes V_d$ producto tensorial junto con una aplicación d-lineal universal

$$t : V_1 \times \dots \times V_d \to V_1 \otimes \dots \otimes V_d,$$

de tal suerte que $\mathrm{Mult}^d(V_1 \times \dots \times V_d, \mathcal{U}) \cong \mathrm{Hom}_k(V_1 \otimes \dots \otimes V_d, \mathcal{U})$ para todo k-ev $\mathcal{U}$. Más aún, si tenemos d aplicaciones lineales $f_i : V_i \to W_i$ $(i = 1, \dots, d)$, la propiedad universal permite "tensorizarlas" y obtener una única aplicación lineal

$$f_1 \otimes \dots \otimes f_d : V_1 \otimes \dots \otimes V_d \to W_1 \otimes \dots \otimes W_d$$

que para todo tensor simple $v_1 \otimes \dots \otimes v_d$ verifica $(f_1 \otimes \dots \otimes f_d)(v_1 \otimes \dots \otimes v_d) = f_1(v_1) \otimes \dots \otimes f_d(v_d)$.

Obs: Vimos en §44 que $(V_1 \otimes V_2) \otimes V_3$ y $V_1 \otimes (V_2 \otimes V_3)$ son <u>canónicamente</u> isomorfos (ie, existe un isomorfismo construido sin escoger bases). No es difícil verificar que ambos son canónicamente isomorfos a $V_1 \otimes V_2 \otimes V_3$.

Recuerdo (k-álgebra): Recordemos que (ver §4, pág 8) una <u>k-álgebra</u> es un k-ev A dotado de un producto $A \times A \to A$, $(a,b) \mapsto ab$ que es una aplicación k-bilineal y que dota a A de estructura de <u>anillo</u>. Así, la multiplicación es asociativa pero <u>no</u> necesariamente conmutativa.

⚠ <u>Todas</u> las álgebras que consideraremos poseen una <u>unidad</u>, ie, $\exists! \ 1 \in A$ tal que $a \cdot 1 = 1 \cdot a = a \ \ \forall a \in A$. Notar que $1 = 0 \iff A = 0$.

Def: Sea A una k-álgebra. Decimos que A es una <u>álgebra graduada</u> si existe una descomposición (eventualmente infinita)

$$A = \bigoplus_{d \in \mathbb{N}} A_d = A_0 \oplus A_1 \oplus A_2 \oplus \dots$$

en suma directa de espacios vectoriales tal que $\forall d, e \in \mathbb{N} : A_d \cdot A_e \subseteq A_{d+e}$. En part, $1 \in A_0$. $(\Rightarrow k := k \cdot 1 \subseteq A_0)$. Más aún, un <u>morfismo de álgebras graduadas</u> es un morfismo de álgebras $f : A \to B$ (ie, $f(a_1 a_2) = f(a_1) f(a_2) \ \forall a_1, a_2 \in A$ y f lineal) que además preserva la graduación: $f(A_d) \subseteq B_d$ para todo $d \in \mathbb{N}$.

Ejemplo: El álgebra (conmutativa) $k[X]$ de polinomios en una variable está graduada (por el grado): $k[X] = \bigoplus_{d \geq 0} \mathrm{Vect}_k\langle X^d \rangle$. Más generalmente, el álgebra $A = k[X_1, \dots, X_r]$ de polinomios en r variables está graduada si definimos a A_d como el sub-ev generado por polinomios <u>homogéneos</u> de grado total d, ie, generados por los monomios $X_1^{i_1} \cdots X_r^{i_r}$ con $i_1 + \dots + i_r = d$.

Terminología: Sea $A = \bigoplus_{d\geq 0} A_d$ k-álgebra graduada. Un elemento no-nulo $x \in A_{\neq 0}$ se dice homogéneo si existe $d \in \mathbb{N}$ (necesariamente único) tq $x \in A_d$. Diremos que x es de grado d.

Def: Sea V un k-ev. Definimos las potencias tensoriales de V como $T^0V := k$ y, para $d \geq 1$:

$$T^d V := \underbrace{V \otimes \dots \otimes V}_{d \text{ veces}} =: V^{\otimes d}.$$

En part (y por definición), $(T^d V)^* \cong \mathrm{Mult}^d(V^d, k)$ es el k-ev de formas d-lineales en V. Más aún, definimos el álgebra tensorial de V por

$$TV := \bigoplus_{d\geq 0} T^d V = k \otimes V \oplus T^2V \oplus T^3V \oplus \dots$$

donde el producto en TV está dado por

$$T^d V \times T^e V \to T^{d+e} V$$
$$(v_1 \otimes \dots \otimes v_d, w_1 \otimes \dots \otimes w_e) \mapsto v_1 \otimes \dots \otimes v_d \otimes w_1 \otimes \dots \otimes w_e,$$

y cuya unidad es $1 \in k = T^0V$. Así, TV es una k-álgebra graduada.

Obs: 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 $i : V \hookrightarrow TV$.

③ Si $\{e_i\}_{i\in I}$ es una base de V, entonces TV tiene por base todos los tensores $e_{i_1} \otimes \dots \otimes e_{i_d}$ para $d \in \mathbb{N}$ y $i_1, \dots, i_d \in I$. En part, incluso si V es de dimensión finita, el k-ev TV es siempre de dimensión infinita si $V \neq 0$.

Teorema (propiedad universal): El álgebra tensorial TV satisface la propiedad universal siguiente:

"Para toda aplicación lineal $f : V \to A$ a una k-álgebra con unidad A, existe un único morfismo de álgebras $\hat{f} : TV \to A$ tal que $f = \hat{f} \circ i$, con $i : V \hookrightarrow TV$."

En otras palabras, el diagrama

$$\begin{array}{ccc} V & \xrightarrow{f} & A \\ \searrow & & \nearrow \exists! \hat{f} \\ & TV & \end{array}$$

es conmutativo.

Dem: Como la aplicación $V^d \to A$, $(v_1, \dots, v_d) \mapsto f(v_1) \cdots f(v_d)$ es d-lineal, la propiedad universal de T^dV permite definir la aplicación lineal $\hat{f}_d : T^dV \to A$ para todo $d \in \mathbb{N}$, y en particular $\hat{f} = \bigoplus_{d\geq 0} \hat{f}_d : TV \to A$. Basta ver que $\hat{f}$ es un morfismo de álgebras: es suficiente verificarlo para tensores simples (que generan TV)

$$\hat{f}((v_1 \otimes \dots \otimes v_d) \otimes (w_1 \otimes \dots \otimes w_e)) = f(v_1) \cdots f(v_d) f(w_1) \cdots f(w_e)$$
$$= \hat{f}(v_1 \otimes \dots \otimes v_d)\, \hat{f}(w_1 \otimes \dots \otimes w_e) \quad \checkmark \quad \blacksquare$$

Obs: Como en todos los casos anteriores, la propiedad universal implica que la construcción es "functorial": si $f : V \to W$ aplicación lineal, existe un único morfismo de álgebras $Tf : TV \to TW$ tal que el diagrama

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ \downarrow i_V & & \downarrow i_W \\ TV & \xrightarrow[Tf]{} & TW \end{array} \quad \text{es conmutativo.}$$

En efecto, Tf no es nada más que $\bigoplus_{d\geq 0} T^d f$, donde $T^d f = f^{\otimes d} = f \otimes \overset{d \text{ veces}}{\dots} \otimes f$.
Más aún, $T(f\circ g) = Tf \circ Tg$ (En términos más sofisticados, T es un "<u>functor</u>" entre las "<u>categorías</u>" de k-espacios vectoriales y la de k-álgebras).

<u>Caso particular importante</u>: En Física (y en Geometría Diferencial) se consideran ciertos tensores particulares: Sea V un k-ev de $\dim_k(V) = n$ y sea $V^* = \mathrm{Hom}_k(V,k)$ su espacio <u>dual</u>. Durante este parrafo (y para simplificar la notación), si $\mathcal{B} = (e_1,\dots,e_n)$ es una base de V, entonces denotaremos por $\mathcal{B}^* = (e^1,\dots,e^n)$ la respectiva base dual de V^* (i.e. $e^i(e_j) = \delta_{ij}$ o bien $e^i(e_j) = \delta^i_j$).

<u>Def</u>: Sea $V \cong k^n$ un k-ev y sea V^* el espacio dual. Un tensor <u>p-covariante y q-contravariante</u> es un elemento del producto tensorial

$$\underbrace{V^*\otimes\dots\otimes V^*}_{p \text{ veces}} \otimes \underbrace{V\otimes\dots\otimes V}_{q \text{ veces}} = T^pV^* \otimes T^qV.$$

También diremos que son <u>tensores de tipo (p,q)</u>. 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:

$$T = \sum_{\substack{i_1,\dots,i_q \\ j_1,\dots,j_p}} T^{i_1\dots i_q}_{j_1\dots j_p}\, e^{j_1}\otimes\dots\otimes e^{j_p}\otimes e_{i_1}\otimes\dots\otimes e_{i_q}$$

pues los vectores de la forma $e^{j_1}\otimes\dots\otimes e^{j_p}\otimes e_{i_1}\otimes\dots\otimes e_{i_q}$ forman una base de $T^pV^*\otimes T^qV$.

<u>Convención de Einstein</u>: La "notación de Einstein" (aunque en realidad es un caso particular del "cálculo de Ricci" desarrollado 30 años antes...) consiste en omitir los símbolos de sumatorias. Así, un tensor de tipo (p,q) se escribe como:

$$T = T^{i_1\dots i_q}_{j_1\dots j_p}\, e^{j_1}\otimes\dots\otimes e^{j_p}\otimes e_{i_1}\otimes\dots\otimes e_{i_q}$$

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, $y = \sum_{i=1}^{3} x^i e_i = x^1e_1 + x^2e_2 + x^3e_3$ se reduce a escribir $y = x^ie_i$.

<u>Ejemplo</u>: Sea $v = v^ie_i$ un vector de V. Vimos que (§44, pág 142) todo endomorfismo es un tensor de tipo $(1,1)$, pues $\mathrm{End}_k(V) \underset{\text{def}}{=} \mathrm{Hom}_k(V,V) \cong V^*\otimes V$. Así, si $(e'_1,\dots,e'_n)$ es otra base de V y $P = (P^i_j)$ es la matriz de cambio de base $\Rightarrow e'_j = P^i_j e_i$. Luego, si escribimos $v = v'^j e'_j \Rightarrow v = v'^j e'_j = v'^j P^i_j e_i$ y luego $v^i = P^i_j v'^j$, o bien: $v'^j = (P^{-1})^j_i v^i$.

Observamos que las coordenadas de v se transforman en sentido inverso a los vectores de las bases (de aquí el término "contravariante"). Recíprocamente, para una forma lineal $f = f_j e^j = f'_i e'^i$ se tiene que:

$f_{i'} = f(e_{i'}) = f(P^i_{i'} e_j) = P^i_{i'} f_j$, ie, las coordenadas de $f \in V^*$ se transforman en el mismo sentido que las bases (de aquí el término "covariante").

Ejercicio Para un tensor de tipo (p,q) de coordenadas $(T^{i_1 \dots i_q}_{j_1 \dots j_p})$, probar que las coordenadas en la base $(e_{1'}, \dots, e_{n'})$ son

$$(T')^{i_1' \dots i_q'}_{j_1' \dots j_p'} = (P^{-1})^{i_1'}_{i_1} \cdots (P^{-1})^{i_q'}_{i_q} P^{j_1}_{j_1'} \cdots P^{j_p}_{j_p'} T^{i_1 \dots i_q}_{j_1 \dots j_p}.$$

Obs: Generalmente, los cálculos anteriores se hacen en presencia de una "métrica", ie, una forma bilineal $B : V \times V \to k$ no-degenerada (eg, producto escalar o forma de Lorentz) y decimos que B es el tensor métrico:
Dado que $\mathrm{Bil}(V \times V, k) \cong (V \otimes V)^* \cong V^* \otimes V^*$, pensamos B como un tensor de tipo $(2,0)$ (o 2-covariante) y denotamos su matriz por $g_{ij} = B(e_i, e_j)$ en la base $(e_1, \dots, e_n)$ de V.

Como B es no-degenerada, $\hat{B} : V \xrightarrow{\sim} V^*$, $x \mapsto B(x, \cdot)$ es un isomorfismo que identifica vectores (contravariantes) y formas lineales (covariantes). En coordenadas:
$\hat{B}(e_j)(e_i) = B(e_j, e_i) = g_{ji} \Rightarrow \hat{B}(v^i e_j) = v^i g_{ji} e^i$

Así, pasamos de coord. ~~covariantes~~ contravariantes (v^i) a coord. covariantes (v_i) mediante la fórmula $v_i = v^j g_{ji}$.

Finalmente, observamos que el isomorfismo $\hat{B}$ permite transportar la métrica B a una métrica B^* en V^*. Explícitamente, la matriz $g^{ij} := B^*(e^i, e^j)$ de B^* resp. a la base dual $(e^1, \dots, e^n)$ de V^* es la inversa de la matriz (g_{ij}), ie, $g_{ij} g^{jk} = \delta_i^k$. Este "tensor métrico dual" permite pasar de coordenadas covariantes a contravariantes mediante la fórmula $v^i = v_i g^{ij}$, y en particular:

$$B(u,v) = u^i v_i = u_i v^i = g_{ij} u^i v^j = g^{ij} u_i v_j.$$

§46. Álgebra exterior y formas multilineales alternadas

Por un lado, en la construcción del álgebra tensorial $TV = \bigoplus_{d \geq 0} T^d V$ tenemos que $(T^d V)^* \cong \mathrm{Mult}^d(\overbrace{V \times \dots \times V}^{d \text{ veces}}, k)$ son canónicamente isomorfos. Por otro lado, estudiamos en §9 las formas multilineales alternadas: una forma d-lineal $\omega : \overbrace{V \times \dots \times V}^{d \text{ veces}} \to k$ es alternada si $\omega(v_1, \dots, v_d) = 0$ cuando dos de los $v_1, \dots, v_d$ son iguales. Más aún, si denotamos por

$$\mathrm{Alt}^d(V) = \{ \omega : \overbrace{V \times \dots \times V}^{d \text{ veces}} \to k \ d\text{-lineal alternada} \}$$

al k-ev de formas d-lineales alternadas, entonces vimos en §10 que si $n = \dim_k(V)$ entonces toda forma n-lineal alternada es proporcional al determinante, ie, $\mathrm{Alt}^n(V) = \mathrm{Vect}_k(\det)$.

El objetivo de esta sección es realizar una construcción análoga a $\mathrm{Alt}^d(V)$ en TV (que será dual a $\mathrm{Alt}^d(V)$ en un sentido preciso).

Concretamente: Nos gustaría construir a partir de TV una nueva álgebra ΛV tal que para cada tensor $v_1 \otimes \dots \otimes v_d$ en TV asociamos un "símbolo" $v_1 \wedge \dots \wedge v_d$ (ie, vector en ΛV) tal que $v_1 \wedge \dots \wedge v_d = 0$ si dos de los $v_1, \dots, v_d$ son iguales.

Como siempre, los cocientes son útiles para construir objetos a partir de otros: basta cocientar por las "relaciones" que deseamos que se verifiquen!

Def: Sea V un k-ev y TV su álgebra tensorial. Sea $I \subseteq TV$ el sub-ev generado por todos los tensores de la forma

$$a \otimes v \otimes v \otimes b \quad \text{con } v \in V \text{ y } a, b \in TV.$$

Definimos el <u>álgebra exterior</u> de V como $\Lambda V := TV/I$, y denotamos por $v_1 \wedge \dots \wedge v_d$ la imagen de $v_1 \otimes \dots \otimes v_d$ en el cociente. Más aún, la composición $V \hookrightarrow TV \twoheadrightarrow \Lambda V$ será denotada $i : V \to \Lambda V$.

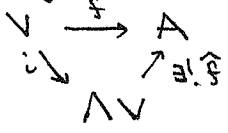
Hecho (MAT214): Se puede probar que I es un <u>ideal</u> del álgebra TV (cf. §18, pág 37). Esto último implica exactamente que el cociente $\Lambda V = TV/I$ también es una k-álgebra, ie, la multiplicación (no conmutativa) en TV define una multiplicación en el álgebra exterior ΛV que llamamos <u>producto exterior</u>:

$$(v_1 \wedge \dots \wedge v_d, w_1 \wedge \dots \wedge w_e) \longmapsto v_1 \wedge \dots \wedge v_d \wedge w_1 \wedge \dots \wedge w_e .$$

Teorema (propiedad universal): El álgebra exterior ΛV satisface la propiedad universal siguiente:

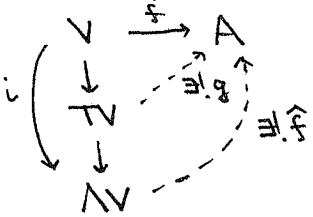
"Para toda aplicación lineal $f : V \to A$ a una k-álgebra con unidad A tal que $f(v)^2 = 0$ para todo $v \in V$, existe un <u>único</u> morfismo de álgebras $\hat{f} : \Lambda V \to A$ tal que $f = \hat{f} \circ i$, con $i : V \to \Lambda V$."

En otras palabras, el diagrama



es conmutativo.

Dem: Por la propiedad universal de TV, $\exists!\, g : TV \to A$ inducido por f. Dado que $g(v \otimes v) = f(v) f(v) = f(v)^2 = 0$, la aplicación lineal g se anula en el ideal I. Luego, la propiedad universal del cociente implica que $\exists!\, \hat{f} : \Lambda V = TV/I \to A$ 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:



■

Describamos ahora la "funtorialidad" asociada a la propiedad universal anterior, cuya demostración se deja como ejercicio:

Prop (functorialidad): Si $f: V \to W$ es una aplicación lineal entre k-ev. Entonces, existe un único morfismo de álgebras $\Lambda f: \Lambda V \to \Lambda W$ tq $i_W \circ f = \Lambda f \circ i_V$, ie, el diagrama

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ {\scriptstyle i_V}\downarrow & & \downarrow{\scriptstyle i_W} \\ \Lambda V & \xrightarrow[\Lambda f]{} & \Lambda W \end{array}$$

es conmutativo.

Más aún, $\Lambda(f \circ g) = \Lambda f \circ \Lambda g$.

Veamos ahora cómo describir el álgebra exterior ΛV concretamente:

Recordemos que $I \subseteq TV$ es el sub-ev generado por tensores de la forma

$$a \otimes v \otimes v \otimes b \quad \text{con } v \in V \text{ y } a, b \in TV.$$

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

$$\Lambda^d V := T^d V / I_d$$

llamado la <u>d-ésima potencia exterior de V</u>. Así, el álgebra exterior también es una k-álgebra graduada:

$$\Lambda V = \bigoplus_{d \in \mathbb{N}} T^d V / I_d = \bigoplus_{d \in \mathbb{N}} \Lambda^d V.$$

<u>Importante</u>: Dado que los vectores no-nulos de I contienen los elementos de grado 2 de la forma $v \otimes v$, tenemos que $T^0 V \cap I = \{0\}$ y $T^1 V \cap I = \{0\}$ $\Rightarrow \Lambda^0 V \cong k$ y $\Lambda^1 V \cong T^1 V \cong V$. En part, $i: V \hookrightarrow \Lambda V$ es <u>inyectivo</u>.

Así: $\Lambda V = \underset{\lambda}{k} \oplus \underset{v}{V} \oplus \underset{v_1 \wedge v_2}{\Lambda^2 V} \oplus \underset{v_1 \wedge v_2 \wedge v_3}{\Lambda^3 V} \oplus \dots$

Luego, el k-ev ΛV está generado por los vectores de la forma $v_1 \wedge \dots \wedge v_d$ con $d \in \mathbb{N}$ y donde $v_1 \wedge \dots \wedge v_d = 0$ si dos de los $v_1, \dots, v_d \in V$ son iguales. Además, el hecho que ΛV sea una álgebra graduada se escribe como

$$(\Lambda^d V) \wedge (\Lambda^e V) \subseteq \Lambda^{d+e} V.$$

Más aún, si $f: V \to W$ es una aplicación lineal, entonces $\Lambda f = \bigoplus_{d \geq 0} \Lambda^d f$, donde $\Lambda^d f: \Lambda^d V \to \Lambda^d W$ está dada por $(\Lambda^d f)(v_1 \wedge \dots \wedge v_d) = f(v_1) \wedge \dots \wedge f(v_d)$.

Prop: Sea V un k-ev. Entonces, para todo $d \geq 2$ la aplicación d-lineal

$$V^d = \underbrace{V \times \dots \times V}_{d \text{ veces}} \longrightarrow \Lambda^d V$$
$$(v_1, \dots, v_d) \mapsto v_1 \wedge \dots \wedge v_d$$

es <u>alternada</u>. En particular, para toda permutación $\sigma \in S_d$ y $v_1, \dots, v_d \in V$ se tiene:

$$v_{\sigma(1)} \wedge \dots \wedge v_{\sigma(d)} = \varepsilon(\sigma)\, v_1 \wedge \dots \wedge v_d.$$

Dem: Por definición de I y de ΛV, la expresión $v_1 \wedge \dots \wedge v_d$ es nula si dos vectores consecutivos son iguales, el caso general se deduce al expresar una permutación arbitraria como producto de trasposiciones (cf. §9, pág 18). ■

Corolario: El álgebra graduada $\bigwedge V$ es anti-conmutativa: si $\alpha \in \bigwedge^d V$ y $\beta \in \bigwedge^e V$ entonces se tiene que $\beta \wedge \alpha = (-1)^{de} \alpha \wedge \beta$.

Dem: Basta probarlo en el caso que $\alpha = v_1 \wedge \dots \wedge v_d$ y $\beta = w_1 \wedge \dots \wedge w_e$ (el caso general es una combinación lineal 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: Sea V un k-ev y sea $d \in \mathbb{N}$. Entonces:

① Para todo $d \geq 1$, la aplicación d-lineal alternada $V^d \to \bigwedge^d V$, $(v_1, \dots, v_d) \mapsto v_1 \wedge \dots \wedge v_d$ satisface la propiedad universal siguiente:

"Para todo k-ev $\mathcal{U}$ y toda aplicación d-lineal alternada $\omega : V^d \to \mathcal{U}$, existe una única aplicación lineal $\widehat{\omega} : \bigwedge^d V \to \mathcal{U}$ tal que el diagrama

$$\begin{array}{ccc} V^d & \xrightarrow{\omega} & \mathcal{U} \\ \downarrow & \nearrow_{\exists!\, \widehat{\omega}} & \\ \bigwedge^d V & & \end{array}$$

es conmutativo."

En particular, $\mathrm{Alt}^d(V) \cong (\bigwedge^d V)^*$.

② Si V es de dimensión finita n y $(e_1, \dots, e_n)$ es una base de V, entonces $(e_{i_1} \wedge \dots \wedge e_{i_d})_{1 \leq i_1 < \dots < i_d \leq n}$ 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

$$\begin{array}{ccc} \bigwedge^d V \times \bigwedge^{n-d} V & \longrightarrow & \bigwedge^n V \cong k \\ (\alpha, \beta) & \longmapsto & \alpha \wedge \beta \end{array}$$

es no-degenerada. En part, induce un isomorfismo canónico $\bigwedge^n V \otimes (\bigwedge^d V)^* \cong \bigwedge^{n-d} V$ y un isomorfismo no-canónico (ie, depende de la elección de un isomorfismo $\bigwedge^n V \cong k$) $(\bigwedge^d V)^* \cong \bigwedge^{n-d} V$.

④ La aplicación $\bigwedge^d (V^*) \times \bigwedge^d V \to k$, $(f_1 \wedge \dots \wedge f_d, v_1 \wedge \dots \wedge v_d) \mapsto \det((f_i(v_j))_{1 \leq i,j \leq d})$ es bilineal y no-degenerada. En part, si V es de dimensión finita entonces hay un isomorfismo canónico $\bigwedge^d V^* \cong (\bigwedge^d V)^*$.

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, $(\bigwedge^n u)(e_1 \wedge \dots \wedge e_n) = \det(u)\, e_1 \wedge \dots \wedge e_n$.

Dem: ① La propiedad universal de TV implica que la aplicación d-lineal (alternada) $\omega : V^d \to \mathcal{U}$ induce una única $g : T^d V \to \mathcal{U}$ lineal tq $\omega = g \circ i$, con $i : V^d \to T^d V$. Pero, dado que ω es alternada, g se anula en $I_d = I \cap T^d V$ y luego la propiedad universal del cociente $\Rightarrow \exists!\, \widehat{\omega} : \bigwedge^d V = T^d V / I_d \to \mathcal{U}$ ✓ En part, si $\mathcal{U} = k$ obtenemos $\mathrm{Alt}^d(V) \cong (\bigwedge^d V)^*$.

② y ③: Dado que $v_{\sigma(1)} \wedge \dots \wedge v_{\sigma(d)} = \varepsilon(\sigma)\, v_1 \wedge \dots \wedge v_d$, tenemos que los vectores $(e_{i_1} \wedge \dots \wedge e_{i_d})_{1 \leq i_1 < \dots < i_d \leq n}$ generan $\bigwedge^d V$ (podemos suponer los índices ordenados).

Veamos que son l.i: Si $d=n$ entonces sabemos (ver §10) que $\mathrm{Alt}^n(V) = \mathrm{Vect}_k(\det)$ es de dimensión 1, y luego $\Lambda^n V \cong (\mathrm{Alt}^n(V))^* \cong k$ es de dimensión 1, por lo que el vector no-nulo $e_1 \wedge e_2 \wedge \ldots \wedge e_n$ es una base de $\Lambda^n V$. Supongamos $d<n$ y fijemos índices $1 \le j_1 < \ldots < j_{n-d} \le n$, entonces el único caso en que

$$(e_{i_1} \wedge \ldots \wedge e_{i_d}) \wedge (e_{j_1} \wedge \ldots \wedge e_{j_{n-d}}) \neq 0 \quad \text{en } \Lambda^n V \cong k$$

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, concluímos que los $(e_{i_1} \wedge \ldots \wedge e_{i_d})_{1 \le i_1 < \ldots < i_d \le n}$ son l.i. (basta tomar productos exteriores de cualquier combinación lineal con $e_{j_1} \wedge \ldots \wedge e_{j_{n-d}}$) y luego una base de $\Lambda^d V$. Más aún, esto también prueba que $\Lambda^d V \times \Lambda^{n-d} V \to \Lambda^n V$, $(\alpha, \beta) \mapsto \alpha \wedge \beta$ es bilineal no-degenerada ✓

En part, "fijar la segunda variable" induce un isomorfismo canónico entre $\Lambda^{n-d} V$ y $\mathrm{Hom}_k(\Lambda^d V, \Lambda^n V) \cong (\Lambda^d V)^* \otimes \Lambda^n V$ (ver §44, pág 142) ✓

④ Las propiedades del determinante vistas en §10 implican que la aplicación $V^* \times \overset{d \text{ veces}}{\ldots} \times V^* \times V \times \overset{d \text{ veces}}{\ldots} \times V \to k$, $((f_1, \ldots, f_d), (v_1, \ldots, v_d)) \mapsto \det((f_i(v_j))_{1 \le i,j \le d})$ es alternada en los f_j y es alternada en los v_i, por lo que induce una aplicación bilineal $B: \Lambda^d V^* \times \Lambda^d V \to k$.

Si $(e_1, \ldots, e_m)$ es una base de V y $(e_1^*, \ldots, e_m^*)$ es la base dual de V^*, y si $1 \le i_1 < \ldots < i_d \le m$ y $1 \le j_1 < \ldots < j_d \le m$ entonces (Ejercicio):

$$B(e_{i_1}^* \wedge \ldots \wedge e_{i_d}^*, e_{j_1} \wedge \ldots \wedge e_{j_d}) = \begin{cases} 1 & \text{si } i_\kappa = j_\kappa \text{ para todo } \kappa \\ 0 & \text{sino.} \end{cases}$$

Así, la forma B es no-degenerada e induce una dualidad en la cual $(e_{i_1}^* \wedge \ldots \wedge e_{i_d}^*)_{1 \le i_1 < \ldots < i_d \le n}$ es la base dual de $(e_{i_1} \wedge \ldots \wedge e_{i_d})_{1 \le i_1 < \ldots < i_d \le n}$ ✓

⑤ Sea $A = (a_{ij}) = \mathrm{Mat}_{\mathcal{B}}(u)$ resp. a una base $\mathcal{B} = (e_1, \ldots, e_m)$ de V. Entonces:

$$\begin{aligned}(\Lambda^m u)(e_1 \wedge \ldots \wedge e_m) &\underset{\text{def}}{=} u(e_1) \wedge \ldots \wedge u(e_m) = \Big(\sum_i a_{i1} e_i\Big) \wedge \ldots \wedge \Big(\sum_i a_{im} e_i\Big) \\ &= \Big(\sum_{i_1=1}^m a_{i_1 1} e_{i_1}\Big) \wedge \ldots \wedge \Big(\sum_{i_m=1}^m a_{i_m m} e_{i_m}\Big) = \sum_{i_1, \ldots, i_m = 1}^m a_{i_1 1} \cdots a_{i_m m}\, e_{i_1} \wedge \ldots \wedge e_{i_m} \\ &= \sum_{\sigma \in S_m} a_{\sigma(1)1} \cdots a_{\sigma(m)m} \underbrace{(e_{\sigma(1)} \wedge \ldots \wedge e_{\sigma(m)})}_{\varepsilon(\sigma)\, e_1 \wedge \ldots \wedge e_m} = \det(u)\, e_1 \wedge \ldots \wedge e_m \quad ✓\end{aligned}$$

El siguiente resultado es una aplicación, muy útil en la práctica, del Teorema anterior:

Corolario: Sea V un k-ev de dimensión finita. Entonces, los vectores $v_1, \ldots, v_d \in V$ son l.i. $\Leftrightarrow v_1 \wedge \ldots \wedge v_d \neq 0$ en $\Lambda^d V$. En part, si los vectores $v_1, \ldots, v_d \in V$ son li y $v \in V$ es otro vector, entonces $v \in \mathrm{Vect}_k(v_1, \ldots, v_d) \Leftrightarrow v_1 \wedge \ldots \wedge v_d \wedge v = 0$ en $\Lambda^{d+1} V$.

Dem: Si los $v_1, \ldots, v_d$ son li, entonces se pueden completar en una base de V y luego $v_1 \wedge \ldots \wedge v_d$ es parte de una base de $\Lambda^d V$, en part. es no-nulo. Recíprocamente, si $v_1, \ldots, v_d$ son ld y, por ejemplo, $v_d = \lambda_1 v_1 + \ldots + \lambda_{d-1} v_{d-1}$ $\Rightarrow v_1 \wedge \ldots \wedge v_{d-1} \wedge v_d = v_1 \wedge \ldots \wedge v_{d-1} \wedge (\lambda_1 v_1 + \ldots + \lambda_{d-1} v_{d-1}) = 0$ ✓ Finalmente, si los $v_1, \ldots, v_d$ son li y $v \in V$ entonces $v \in \mathrm{Vect}_k(v_1, \ldots, v_d) \Leftrightarrow \{v_1, \ldots, v_d, v\}$ es ld. ∎

Ejemplo: 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 li pues: $v_1 \wedge v_2 = (e_1 + e_3) \wedge (e_2 + e_4) = e_1 \wedge e_2 + e_1 \wedge e_4 + e_3 \wedge e_2 + e_3 \wedge e_4$
$= e_1 \wedge e_2 + e_1 \wedge e_4 - e_2 \wedge e_3 + e_3 \wedge e_4 \neq 0$ en $\Lambda^2 \mathbb{R}^4 \cong \mathbb{R}^{\binom{4}{2}} = \mathbb{R}^6$

Por otro lado, un vector $v \in \mathbb{R}^4$ pertenece al plano $\Pi = \mathrm{Vect}_{\mathbb{R}}(v_1, v_2) \simeq \mathbb{R}^2$ si y sólo si $v_1 \wedge v_2 \wedge v = 0$.

Ejemplo (producto cruz): Recordemos que si $V \cong \mathbb{R}^3$ es un espacio euclideano orientado, entonces el producto cruz $u \times v \in V$ de $u, v \in V$ está dejinido por la relación

$$\det_{\mathcal{B}}(u, v, w) = \langle u \times v, w \rangle \quad \forall w \in V,$$

donde $\mathcal{B} = (e_1, e_2, e_3)$ es cualquier base ortonormal directa de V (ver §38).
Por otro lado, el producto exterior $e_1 \wedge e_2 \wedge e_3 \in \Lambda^3 V$ provee un generador canónico (gracias a la orientación!) de $\Lambda^3 V \cong \mathbb{R}$. Así, en este caso el isomorfismo $\Lambda^{n-d} V \cong (\Lambda^d V)^* \otimes \Lambda^n V$ se reduce (donde $n=3$ y $d=1$) a $\Lambda^2 V \cong V^*$. Más aún, si consideramos el isomorfismo $V^* \cong V$ dado por el producto escalar, obtenemos un isomorfismo $\Lambda^2 V \cong V$ y por ende $u \wedge v \in \Lambda^2 V$ corresponde a un único vector en V. Ejercicio: Probar que dicho vector es $u \times v \in V$.

Caso particular importante: Supongamos que $\mathrm{car}(k) = 0$ (eg. $k = \mathbb{Q}, \mathbb{R}$ o $\mathbb{C}$).
En este caso, podemos pensar a $\Lambda^d V$ como un sub-espacio vectorial de $T^d V$ de la manera siguiente:
Para cada $\sigma \in S_d$ permutación, entonces dejinimos el endomorfismo $\tilde{\sigma}$ de $T^d V$ mediante la fórmula

$$\tilde{\sigma}(v_1 \otimes \dots \otimes v_d) = v_{\sigma(1)} \otimes \dots \otimes v_{\sigma(d)}.$$

Def: Un tensor $T \in T^d V$ se dice anti-simétrico si $\tilde{\sigma}(T) = \varepsilon(\sigma) T$ para todo $\sigma \in S_d$.
Denotamos por $A^d V \subseteq T^d V$ al sub-ev de tensores anti-simétricos. Más aún, dejinimos la aplicación de anti-simetrización mediante:

$$p : T^d V \longrightarrow T^d V$$
$$T \longmapsto \frac{1}{d!} \sum_{\sigma \in S_d} \varepsilon(\sigma) \tilde{\sigma}(T)$$

Ejemplo ($d=2$): $p(v_1 \otimes v_2) = \frac{1}{2}(v_1 \otimes v_2 - v_2 \otimes v_1)$.

Prop: La aplicación lineal p es un proyector (ie, $p^2 = p$), de kernel $I_d = I \cap T^d V$ y de imagen $A^d V$. En part, $A^d V \cong \Lambda^d V$.

Dem: Para todo $\tau \in S_d$ se tiene:

$$p(\tilde{\tau}(T)) = \frac{1}{d!} \sum_{\sigma \in S_d} \varepsilon(\sigma) \tilde{\sigma} \tilde{\tau}(T) \underset{\sigma' = \sigma\tau}{=} \frac{1}{d!} \sum_{\sigma' \in S_d} \varepsilon(\sigma'\tau^{-1}) \tilde{\sigma}'(T) = \varepsilon(\tau^{-1}) p(T) = \varepsilon(\tau) p(T).$$

Del mismo modo:

$$\tilde{\tau}(p(T)) = \frac{1}{d!} \sum_{\sigma \in S_d} \varepsilon(\sigma) \tilde{\tau} \tilde{\sigma}(T) = \frac{1}{d!} \sum_{\sigma \in S_d} \varepsilon(\tau^{-1}\sigma) \tilde{\sigma}'(T) = \varepsilon(\tau) p(T), \text{ ie, } p(T) \in A^d V.$$

Además: $p^2(T) = p(p(T)) = \frac{1}{d!} \sum_{\tau \in S_d} \varepsilon(\tau) \tilde{\tau}(\underbrace{p(T)}_{= \varepsilon(\tau) p(T)}) = \frac{1}{d!} \sum_{\tau \in S_d} \underbrace{\varepsilon(\tau)^2}_{=1} p(T) = \frac{d!}{d!} p(T)$

$\Rightarrow p^2 = p$ y $\mathrm{Im}(p) \subseteq A^d V$. Veamos que $\mathrm{Im}(p) = A^d V$:

Sea $T \in A^d V \Rightarrow P(T) = \frac{1}{d!}\sum_{\sigma\in S_d} \varepsilon(\sigma)\underbrace{\widetilde{\sigma}(T)}_{=\varepsilon(\sigma)T} = \frac{1}{d!}\sum_{\sigma\in S_d}\underbrace{\varepsilon(\sigma)^2}_{=1} T = \frac{d!}{d!}T$. Así, $P^2 = P$ y $P|_{A^dV} = Id_{A^dV}$ (i.e., $\mathrm{Im}(P) = A^dV$) $\rightsquigarrow$ proyector de imagen A^dV.

Veamos que $I_d \subseteq \ker(P)$: Sabemos que I_d está generado por los $T = v_1 \otimes \dots \otimes v_d$ con $v_i = v_{i+1}$ para cierto $i \in \{1,\dots,d-1\}$. Sea $\tau = (i, i+1) \in S_d$ transposición, entonces $\widetilde{\tau}(T) = T$ y luego $P(T) = P(\widetilde{\tau}(T)) = \varepsilon(\tau)P(T) = -P(T) \underset{\mathrm{car}(k)=0}{\Rightarrow} P(T) = 0$ ✓

Luego, la propiedad universal del cociente implica que $\exists!\ \widehat{P} : \Lambda^d V = T^dV/I_d \to A^dV$. Si $\pi : T^dV \twoheadrightarrow \Lambda^d V$ es la proyección canónica al cociente, entonces:

$$(\pi\circ\widehat{P})(v_1\wedge\dots\wedge v_d) = (\pi\circ P)(v_1\otimes\dots\otimes v_d) = \pi\left(\frac{1}{d!}\sum_{\sigma\in S_d}\varepsilon(\sigma)\, v_{\sigma(1)}\otimes\dots\otimes v_{\sigma(d)}\right)$$

$$= \frac{1}{d!}\sum_{\sigma\in S_d}\varepsilon(\sigma)\, v_{\sigma(1)}\wedge\dots\wedge v_{\sigma(d)} = \frac{1}{d!}\sum_{\sigma\in S_d}\underbrace{\varepsilon(\sigma)^2}_{=1} v_1\wedge\dots\wedge v_d$$

$$= v_1\wedge\dots\wedge v_d.$$

$\Rightarrow \pi\circ\widehat{P} = Id_{\Lambda^dV}$ y luego $\widehat{P}$ es inyectiva, i.e., $I_d = \ker(P)$. ■

§47. Pfaffiano y matrices antisimétricas

El objetivo de esta sección es aplicar el teorema espectral (ver §43) y las propiedades del álgebra exterior (ver §46) para estudiar matrices anti-simétricas, i.e., que cumplen ${}^tA = -A$.

¡Atención! Durante toda esta sección asumiremos $\mathrm{car}(k) = 0$ (eg. $k = \mathbb{R} = \mathbb{C}$). En particular, $\mathbb{Q} \subseteq k$.

Motivación: Toda matriz anti-simétrica 2×2 se escribe como

$$A = \begin{pmatrix} 0 & \lambda \\ -\lambda & 0 \end{pmatrix}, \quad \lambda \in k.$$

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.

Obs: Si $A \in M_m(k)$ anti-simétrica es invertible (i.e., $A \in Gl_m(k)$) entonces $m = 2n$ es par: ${}^tA = -A \Rightarrow \det(A) = \det({}^tA) = \det(-A) = (-1)^m\det(A) = 0$ si m impar.

Teorema espectral (caso anti-simétrico): Sea $V \cong \mathbb{R}^n$ espacio euclídeano y $u : V \to V$ endomorfismo anti-simétrico (i.e., $u^* = -u$). Entonces, existe una base ortonormal $\mathcal{B}$ de V y reales $\lambda_1,\dots,\lambda_s \in \mathbb{R}^{>0}$ positivos tales que

$$\mathrm{Mat}_{\mathcal{B}}(u) = \left(\begin{array}{cc|c|c|c|c} 0 & \lambda_1 & 0 & \cdots & 0 \\ -\lambda_1 & 0 & 0 & \cdots & 0 \\ \hline & & \ddots & & \\ \hline & & 0 & 0 & \lambda_s \\ \hline & & 0 & -\lambda_s & 0 \\ \hline 0 & & 0 & 0 & 0_{n-r} \end{array}\right) \quad \text{donde } r = \mathrm{rg}(u) = 2s$$

y donde $\pm i\lambda_1,\dots,\pm i\lambda_s \in i\mathbb{R}$ son los valores propios no-nulos de u.

Dem: Recordemos que si $\mathcal{U} \subseteq V$ es un sub-esv, entonces $u(\mathcal{U}) \subseteq \mathcal{U} \Leftrightarrow u^*(\mathcal{U}^\perp) \subseteq \mathcal{U}^\perp$ (ver §43, pág 137). En part, $K = \ker(u)$ estable por u 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 $\Rightarrow \dim_{\mathbb{R}}(K^\perp) = \operatorname{rg}(u) = r = 2s$ es par. Luego, el Teorema espectral (ver §43) 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 $K = V$ entonces $u = 0$ y el resultado 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 $X, Y \in \mathbb{R}^n$
$\Rightarrow AZ = i\lambda Z \Leftrightarrow AX + iAY = -\lambda Y + i\lambda X \Leftrightarrow AX = -\lambda Y$ y $AY = \lambda X$.
Así, $\Pi := \operatorname{Vect}_{\mathbb{R}}(X, Y) \cong \mathbb{R}^2$ es un plano invariante por A y $\langle X, Y\rangle = 0$.

El cálculo anterior implica que si $\mathcal{B}_1 = (a_1, b_1)$ 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 $\operatorname{Mat}_{\mathcal{B}_1}(u|_{\Pi_1}) = \begin{pmatrix} 0 & \lambda_1 \\ -\lambda_1 & 0 \end{pmatrix}$. Más aún, intercambiando $i\lambda_1$ con $-i\lambda_1$ si fuera necesario, podemos suponer $\lambda_1 > 0$.
Finalmente, por inducción en $n = \dim_{\mathbb{R}}(V)$ y considerando $V = \Pi_1 \oplus \Pi_1^\perp$, tenemos que necesariamente $K \subseteq \Pi_1^\perp$ y que (por hipótesis de inducción) $\Pi_1^\perp = \Pi_2 \oplus \dots \oplus \Pi_s \oplus K$, con $\operatorname{Mat}_{\mathcal{B}_j}(u|_{\Pi_j}) = \begin{pmatrix} 0 & \lambda_j \\ -\lambda_j & 0 \end{pmatrix}$ para $j = 2, \dots, s$. ■

Corolario: Sea $M \in M_n(\mathbb{R})$ matriz anti-simétrica real tal que $r = \operatorname{rg}(A) = 2s$ y $\pm i\lambda_1, \dots, \pm i\lambda_s \in i\mathbb{R}$ son los valores propios no-nulos de A. Entonces, existe $P \in O(n)$ matriz ortogonal tal que $M = PAP^{-1} = PA^tP$, donde

$$A = \begin{pmatrix} \begin{matrix} 0 & \lambda_1 \\ -\lambda_1 & 0 \end{matrix} & & & \\ & \ddots & & \\ & & \begin{matrix} 0 & \lambda_s \\ -\lambda_s & 0 \end{matrix} & \\ & & & 0_{n-r} \end{pmatrix} \quad \text{con } \lambda_1, \dots, \lambda_s > 0.$$

El resto de la sección nos concentraremos en matrices cuadradas anti-simétricas de tamaño par $2n \times 2n$.

Def: 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

$$\omega(A) := \sum_{1 \le i < j \le 2n} a_{ij}\, e_i \wedge e_j \quad \text{en } \Lambda^2 V.$$

Si para todo $k \in \mathbb{N}^{\ge 1}$ denotamos $\omega(A)^k := \omega(A)^{\wedge k} = \underbrace{\omega(A) \wedge \dots \wedge \omega(A)}_{k \text{ veces}} \in \Lambda^{2k} V$.
Entonces, definimos el pfaffiano de A como el único escalar $\operatorname{Pf}(A) \in k$ tal que:

$$\omega(A)^n = n!\, \operatorname{Pf}(A)\, e_1 \wedge \dots \wedge e_{2n} \Leftrightarrow \frac{1}{n!}\omega(A)^n = \operatorname{Pf}(A)\, e_1 \wedge \dots \wedge e_{2n} \text{ en } \Lambda^{2n} V \cong k.$$

Ejemplo: ① La matriz $A = \begin{pmatrix} 0 & a_{12} \\ -a_{12} & 0 \end{pmatrix}$ cumple $\omega(A) = a_{12}\, e_1 \wedge e_2$. Aquí $2 = 2n \Rightarrow n = 1$ y luego $Pf(A) = \frac{a_{12}}{1!} = a_{12}$.

② ($n=2$) La matriz anti-simétrica 4×4

$$A = \begin{pmatrix} 0 & a_{12} & a_{13} & a_{14} \\ -a_{12} & 0 & a_{23} & a_{24} \\ -a_{13} & -a_{23} & 0 & a_{34} \\ -a_{14} & -a_{24} & -a_{34} & 0 \end{pmatrix}$$

cumple $\omega(A) = a_{12}\, e_1 \wedge e_2 + a_{13}\, e_1 \wedge e_3 + a_{14}\, e_1 \wedge e_4 + a_{23}\, e_2 \wedge e_3 + a_{24}\, e_2 \wedge e_4 + a_{34}\, e_3 \wedge e_4$.

Luego, el pfaffiano de A se calcula mediante $\omega(A) \wedge \omega(A) = 2!\, Pf(A)\, e_1 \wedge e_2 \wedge e_3 \wedge e_4$:

$$\begin{aligned} \omega(A) \wedge \omega(A) = {} & (a_{12} e_1 \wedge e_2) \wedge (a_{34} e_3 \wedge e_4) + (a_{13} e_1 \wedge e_3) \wedge (a_{24} e_2 \wedge e_4) + \\ & (a_{14} e_1 \wedge e_4) \wedge (a_{23} e_2 \wedge e_3) + (a_{23} e_2 \wedge e_3) \wedge (a_{14} e_1 \wedge e_4) + \\ & (a_{24} e_2 \wedge e_4) \wedge (a_{13} e_1 \wedge e_3) + (a_{34} e_3 \wedge e_4) \wedge (a_{12} e_1 \wedge e_2) \end{aligned}$$

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

$\omega(A) \wedge \omega(A) = 2(a_{12}a_{34} - a_{13}a_{24} + a_{14}a_{23})\, e_1 \wedge e_2 \wedge e_3 \wedge e_4$ y luego obtenemos que $Pf(A) = a_{12}a_{34} - a_{13}a_{24} + a_{14}a_{23}$ es el pfaffiano de A en este caso.

Ejercicio Calcular $Pf(A)$ para una matriz anti-simétrica 6×6.

Ejemplo importante: Consideremos la matriz anti-simétrica

$$A = \left(\begin{array}{cc|c|c} 0 & \lambda_1 & & \\ -\lambda_1 & 0 & & \\ \hline & & \ddots & \\ \hline & & & \begin{matrix} 0 & \lambda_n \\ -\lambda_n & 0 \end{matrix} \end{array}\right) \in M_{2n}(k)$$

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 $Pf(A) = \lambda_1 \cdots \lambda_n$, ie, que

$$\omega(A)^n = n!\, \lambda_1 \cdots \lambda_n\, e_1 \wedge e_2 \wedge \ldots \wedge e_{2n-1} \wedge e_{2n}.$$

Sea $\eta = \lambda_1 e_1 \wedge e_2 + \ldots + \lambda_{n-1} e_{2n-3} \wedge e_{2n-2} \underset{\text{Hip.}}{\Rightarrow} \eta^{n-1} = (n-1)!\, \lambda_1 \cdots \lambda_{n-1}\, e_1 \wedge \ldots \wedge e_{2n-2}$.

y además $\eta^n = \eta^{n-1} \wedge \eta = 0$.

> **Recuerdo** (ver §46, pág 150): Si $\alpha \in \Lambda^d V$ y $\beta \in \Lambda^e V \Rightarrow \beta \wedge \alpha = (-1)^{de}\, \alpha \wedge \beta$.
> En part, si $\alpha := \lambda_n e_{2n-1} \wedge e_{2n} \in \Lambda^2 V$ entonces $\beta \wedge \alpha = \alpha \wedge \beta$ para todo β !
> $\Rightarrow$ 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 $\omega = \eta + \alpha \Rightarrow \omega^n = (\eta+\alpha)^n = \underbrace{\eta^n}_{=0} + n\, \eta^{n-1} \wedge \alpha + \underbrace{\binom{n}{2} \eta^{n-2} \wedge \alpha^2 + \ldots}_{=0 \text{ para } \alpha^k \text{ con } k \geq 2}$

$$\Rightarrow \omega^n = n\, \eta^{n-1} \wedge \alpha = \underbrace{n \cdot (n-1)!}_{=n!}\, \lambda_1 \cdots \lambda_{n-1}\, e_1 \wedge \ldots \wedge e_{2n-2} \wedge (\lambda_n e_{2n-1} \wedge e_{2n})$$

$$= n!\, \lambda_1 \cdots \lambda_n\, e_1 \wedge \ldots \wedge e_{2n}, \text{ y luego } Pf(A) = \lambda_1 \cdots \lambda_n.$$

⚠ Notar además que $\det(A) = \lambda_1^2 \cdots \lambda_n^2 = Pf(A)^2$.

Ejercicio* Sea $A=(a_{ij})\in M_{2n}(k)$ matriz anti-simétrica. Probar (eg. por inducción) la fórmula general

$$Pf(A)=\frac{1}{2^n n!}\sum_{\sigma\in S_{2n}}\varepsilon(\sigma)\,a_{\sigma(1),\sigma(2)}\cdots a_{\sigma(2n-1),\sigma(2n)}=\sum_{\sigma\in F_{2n}}\varepsilon(\sigma)\,a_{\sigma(1),\sigma(2)}\cdots a_{\sigma(2n-1),\sigma(2n)}$$

donde $F_{2n}:=\{\sigma\in S_{2n}$ tal que $\sigma(1)<\sigma(3)<\ldots<\sigma(2n-1)$ y $\sigma(1)<\sigma(2),\ \sigma(3)<\sigma(4),\ldots,\sigma(2n-1)<\sigma(2n)\}$

Recuerdo: Si $u:V\to V$ es un endomorfismo, entonces (por funtorialidad) existe $\Lambda^2 u:\Lambda^2 V\to\Lambda^2 V$ tal que $(\Lambda^2 u)(e_i\wedge e_j)=u(e_i)\wedge u(e_j)$. Matricialmente, si P es una matriz entonces denotamos de igual modo $(\Lambda^2 P)(e_i\wedge e_j)=P(e_i)\wedge P(e_j)$.

Lema: Sea $A\in M_{2n}(k)$ matriz anti-simétrica y sea $P\in M_{2n}(k)$ una matriz. Entonces,

$$\omega(PA^tP)=(\Lambda^2 P)(\omega(A))\ \text{ en }\Lambda^2 k^{2n}.$$

Dem: Si $P=(p_{ij})$ y $A=(a_{kl})$ entonces $PA^tP=(c_{ij})$ con $c_{ij}=\sum_{k,l}p_{ik}a_{kl}p_{jl}$

$$\Rightarrow\omega(PA^tP)=\sum_{i<j}c_{ij}\,e_i\wedge e_j=\sum_{i<j}\sum_{k,l}p_{ik}a_{kl}p_{jl}\,e_i\wedge e_j=\frac{1}{2}\sum_{i,j,k,l}p_{ik}a_{kl}p_{jl}\,e_i\wedge e_j$$

$$=\frac{1}{2}\sum_{i,j,k,l}a_{kl}(p_{ik}e_i)\wedge(p_{jl}e_j)=\frac{1}{2}\sum_{k,l}a_{kl}P(e_k)\wedge P(e_l)$$

$$=\sum_{k<l}a_{kl}\underbrace{P(e_k)\wedge P(e_l)}_{(\Lambda^2P)(e_k\wedge e_l)}=(\Lambda^2P)(\omega(A)).\ \blacksquare$$

El lema anterior permite probar la siguiente identidad útil.

Prop: Sea $A\in M_{2n}(k)$ matriz anti-simétrica y sea $P\in M_{2n}(k)$ una matriz. Entonces,

$$Pf(PA^tP)=\det(P)\,Pf(A).$$

Dem: La identidad $\omega(PA^tP)=(\Lambda^2P)(\omega(A))$ implica que

$$n!\,Pf(PA^tP)\,e_1\wedge\ldots\wedge e_{2n}\overset{\text{def}}{=}\omega(PA^tP)^n=((\Lambda^2P)(\omega(A)))^n.$$

Por otro lado, si $V=k^{2n}$, entonces en $\Lambda V=k\oplus V\oplus\Lambda^2V\oplus\ldots\oplus\Lambda^{2n}V$ se tiene $\Lambda P=1\oplus P\oplus\Lambda^2P\oplus\ldots\oplus\Lambda^{2n}P$. Así, dado que $\omega(A)\in\Lambda^2V$ (homogéneo de grado 2) tenemos $(\Lambda^2P)(\omega(A))=(\Lambda P)(\omega(A))$. Sin embargo, ΛP es un morfismo de k-álgebras (¡respeta el producto!) $\Rightarrow(\Lambda P)(\underbrace{\omega(A)^n}_{\in\Lambda^{2n}V})=(\Lambda^{2n}P)(\omega(A)^n)$.

Finalmente, dado que $\Lambda^{2n}P$ es la homotecia de factor $\det(P)$ en $\Lambda^{2n}V\cong k$, obtenemos que:

$$((\Lambda^2P)(\omega(A)))^n=(\Lambda^{2n}P)(\omega(A)^n)=(\det P)\cdot\omega(A)^n=n!\,\det(P)\,Pf(A)\,e_1\wedge\ldots\wedge e_{2n}\ \checkmark\ \blacksquare$$

Teorema (Cayley, 1847): Sea $A\in M_{2n}(k)$ matriz anti-simétrica. Entonces,

$$\det(A)=Pf(A)^2.$$

Dem: Basta probarlo para $k=\mathbb{Q}$ y, en part, $k\subseteq\mathbb{R}$. El teorema espectral (pág 153) implica que $A=P\tilde{A}^tP$ con $P\in O(n)$ y $\tilde{A}$ de la forma del "Ejemplo Importante" (pág 155). Finalmente, $Pf(\tilde{A})^2=\det(\tilde{A})=\det(A)$ y $Pf(A)=\det(P)\,Pf(\tilde{A})=\pm Pf(\tilde{A})$ $\checkmark\ \blacksquare$

§48. Álgebra simétrica y anillos de polinomios

Motivación: El álgebra exterior ΛV 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 $\hat{B}: V \otimes V \to k$ es la forma lineal asociada entonces $\hat{B}(x \otimes y) = \hat{B}(y \otimes x) \iff \hat{B}(x \otimes y - y \otimes x) = 0 \ \forall x, y \in V$.

En particular, si denotamos por $J_2 \subseteq T^2 V = V \otimes V$ al sub-ev generado por todos los tensores de la forma $x \otimes y - y \otimes x$ con $x, y \in V$, entonces $J_2 \subseteq \ker(\hat{B})$.

Concretamente: Construiremos a partir de TV una nueva álgebra SV tal que para cada tensor $v_1 \otimes \dots \otimes v_d$ en TV asociamos un "símbolo" $v_1 v_2 \cdots v_d$ (ie, 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 part, $v_1 v_2 = v_2 v_1$.

Dy: Sea V un k-ev y TV su álgebra tensorial. Sea $J \subseteq TV$ el sub-espacio generado por todos los tensores de la forma

$$a \otimes (v \otimes w - w \otimes v) \otimes b \quad \text{con } v, w \in V \text{ y } a, b \in TV.$$

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 $v_1 \otimes \dots \otimes v_d$ en el cociente. Más aún, la composición $V \hookrightarrow TV \twoheadrightarrow SV$ será denotada $j: V \to SV$.

Obs: Tal cómo en el caso de ΛV, el cociente $SV = TV/J$ es una k-álgebra. Más aún, la definición de J implica que la multiplicación

$$(v_1 \cdots v_d, w_1 \cdots w_e) \longmapsto v_1 \cdots v_d w_1 \cdots w_e$$

es **conmutativa**!

Teorema (propiedad universal): El álgebra simétrica SV satisface la propiedad universal siguiente:

"Para toda aplicación lineal $f: V \to A$ a una k-álgebra conmutativa con unidad A, existe un único morfismo de álgebras $\hat{f}: SV \to A$ tal que el diagrama

$$\begin{array}{ccc} V & \xrightarrow{f} & A \\ {\scriptstyle j}\downarrow & \nearrow {\scriptstyle \exists! \hat{f}} & \\ SV & & \end{array}$$

es conmutativo (ie, $f = \hat{f} \circ j$)".

Dem: Análoga a la propiedad universal de ΛV, reemplazando $I \subseteq TV$ por $J \subseteq TV$. ∎

Corolario (funtorialidad): Si $f: V \to W$ es una aplicación lineal entre k-ev. Entonces, existe un único morfismo de k-algebras $Sf: SV \to SW$ tal que

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ {\scriptstyle j_V}\downarrow & & \downarrow{\scriptstyle j_W} \\ SV & \xrightarrow{Sf} & SW \end{array}$$

es conmutativo (ie, $j_W \circ f = Sf \circ j_V$).

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-esv generado por los tensores de la forma

$$a \otimes (v \otimes w - w \otimes v) \otimes b \quad \text{con } v, w \in V \text{ y } a, b \in TV.$$

Para $d \in \mathbb{N}$, definimos $J_d := J \cap T^d V$ como el sub-esv de J generado por elementos homogéneos de grado d. Luego, $J = \bigoplus_{d \geq 0} J_d$ y en particular

$$SV = \bigoplus_{d \in \mathbb{N}} S^d V, \quad \text{donde } S^d V := T^d V / J_d$$

es la <u>d-ésima potencia simétrica de V</u>. En part, SV es una k-álgebra graduada.

<u>Importante</u>: Dado que los vectores no-nulos de J contienen los elementos de grado 2 de la forma $v \otimes w - w \otimes v$, tenemos que $T^0 V \cap J = \{0\}$ y $T^1 V \cap J = \{0\}$ $\Rightarrow S^0 V \cong k$ y $S^1 V \cong T^1 V \cong V$. En part, $j : V \hookrightarrow SV$ es <u>inyectivo</u>.

Explícitamente: $SV = \underset{\wedge}{k} \oplus \underset{\vee}{V} \oplus \underset{\vee v_1 v_2}{S^2 V} \oplus \underset{\vee v_1 v_2 v_3}{S^3 V} \oplus \dots$

Más aún, si $f : V \to W$ aplicación lineal, entonces $Sf = \bigoplus_{d \geq 0} S^d f$, donde $S^d f : S^d V \to S^d W$ está dada por $(S^d f)(v_1 \cdots v_d) = f(v_1) \cdots f(v_d)$.

El siguiente resultado resume las principales propiedades de las potencias simétricas $S^d V$:

<u>Teorema</u>: Sea V un k-esv y sea $d \in \mathbb{N}$. Entonces:

① La aplicación d-lineal $V^d \to S^d V$ dada por $(v_1, \dots, v_d) \mapsto v_1 \cdots v_d$ es una aplicación d-lineal <u>simétrica</u> (ie, $f(v_{\sigma(1)}, \dots, v_{\sigma(d)}) = f(v_1 \cdots v_d)$ para toda $\sigma \in S_d$). En part, $(S^d V)^*$ es el k-esv de formas d-lineales simétricas en V.

② Si V es de dimensión finita n, entonces $SV \cong k[X_1, \dots, X_n]$ es isomorfa al álgebra de polinomios en n variables. Más aún, si $(e_1, \dots, e_n)$ es una base de V, entonces $SV \xrightarrow{\sim} k[X_1 \dots, X_n]$, $e_i \mapsto X_i$ es un isomorfismo.

③ Si V es de dimensión finita n y $(e_1, \dots, e_n)$ es una base de V, entonces una base de $S^d V$ está dada por los $e_{i_1}^{k_1} \cdots e_{i_r}^{k_r}$ para toda r-tupla con $1 \leq i_1 < \dots < i_r \leq n$ y enteros k_i tales que $k_1 + \dots + k_r = d$ (cf. polinomios homogéneos de grado d). Luego,

$$\dim_k(S^d V) = \binom{n+d-1}{n-1} = \binom{n+d-1}{d}.$$

Así, si $V \neq 0$, el álgebra simétrica SV es de dimensión infinita.

④ Si $\operatorname{car}(k) = 0$, podemos pensar a $S^d V$ como un <u>sub-esv</u> de $T^d V$ generado por <u>tensores simétricos</u> (ie, $T \in T^d V$ tal que $\widetilde{\sigma}(T) = T$ para toda $\sigma \in S_d$). En efecto, la <u>aplicación de simetrización</u> dada por

$$q : T^d V \to T^d V$$

$$v_1 \otimes \dots \otimes v_d \mapsto \frac{1}{d!} \sum_{\sigma \in S_d} v_{\sigma(1)} \otimes \dots \otimes v_{\sigma(d)}$$

tiene imagen $\Lambda^d(V) \subseteq T^d V$ el sub-esv de tensores simétricos y $\ker(q) = J_d$. En part, $\Lambda^d(V) \cong S^d V = T^d V / J_d$.

<u>Dem</u>: La demostración es análoga al caso de $\Lambda^d V$ y queda como ejercicio. ∎

Observación final: 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:

$$M = S + A, \text{ con } S = \frac{M + {}^tM}{2} \text{ y } A = \frac{M - {}^tM}{2}$$

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, como suma de un tensor anti-simétrico y un tensor simétrico, ie,

$$T^2V = A^2V \oplus \Lambda^2_S V,$$

con $A^2V \simeq \Lambda^2 V$ y $\Lambda^2_S V \simeq S^2V$ (Obs: Basta que $\mathrm{car}(k) \neq 2$ en el caso $d=2$).

En general, para $d \geq 3$ tenemos una inclusión $T^dV \supsetneq A^dV \oplus \Lambda^d_S V$ pero ella es estricta: si $d=3$ calculamos

$$\dim_k(T^3V) = \dim_k(V \otimes V \otimes V) = n^3 > \dim_k(A^3V) + \dim_k(\Lambda^3_S V) = \binom{n}{3} + \binom{n+2}{3}$$
$$= \frac{n(n-1)(n-2)}{6} + \frac{n(n+1)(n+2)}{6}.$$

La parte faltante es un sub-S_n 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

$$T^dV = V^{\otimes d} = \bigoplus_{\substack{\lambda_1 \geq \dots \geq \lambda_m \geq 0 \\ \lambda_1 + \dots + \lambda_m = d}} S_{(\lambda_1, \dots, \lambda_m)}(V)^{\oplus m_\lambda}$$

donde $m_\lambda \in \mathbb{N}^{\geq 1}$ y los $S_{(\lambda_1, \dots, \lambda_m)}$ son los llamados functores de Schur, que verifican (no escribimos los $\lambda_j = 0$):

① $S_{(1,\dots,1)}(V) \simeq \Lambda^d V$ (d veces)

② $S_{(d)}(V) \simeq S^dV$

③ $\dim_k S_{(\lambda_1, \dots, \lambda_m)}(V) = \prod_{1 \leq i < j \leq n} \left(\frac{\lambda_i - \lambda_j + j - i}{j - i} \right)$.

Aquí, la functorialidad significa que para todo $\lambda = (\lambda_1, \dots, \lambda_m)$ y toda aplicación lineal $f : V \to W$, existe una única aplicación lineal inducida

$$S_\lambda(f) : S_\lambda(V) \to S_\lambda(W)$$

con $S_\lambda(\mathrm{Id}) = \mathrm{Id}$ y $S_\lambda(f \circ g) = S_\lambda(f) \circ S_\lambda(g)$.

La construcción de los functores de Schur está relacionada a la "teoría de representaciones" y a objetos combinatorios llamados "tablas de Young", pero eso es otra historia...