Álgebra Lineal de Métodos de Estadística

Pedro Sancho de Salas

2018

Índice general

Capítulo 1

Espacios vectoriales

1.1. Espacio vectorial. Bases

"Un espacio vectorial es un conjunto en el que podemos sumar sus elementos y multiplicar cada elemento por un escalar, y estas operaciones cumplen propiedades muy naturales".

Sea $\mathbb{R}$ el cuerpo de los números reales (podremos considerar otros cuerpos como $\mathbb{Q}$, $\mathbb{C}$).

1. Definición: Un $\mathbb{R}$-*espacio vectorial* es un conjunto, E, dotado de dos operaciones, una llamada suma $E \times E \to E$ y se escribe $(e,e') \mapsto e+e'$, y otra llamada producto por escalares $\mathbb{R} \times E \to E$ y se escribe $(\lambda, e) \mapsto \lambda \cdot e$, verificando:

1. $(E,+)$ es un grupo abeliano, es decir,

   a) $e+(e'+e'') = (e+e')+e''$, para todo $e,e',e'' \in E$.

   b) Existe un elemento que denotamos por 0 tal que $0+e = e+0 = e$, para todo $e \in E$.

   c) Para cada $e \in E$ existe otro elemento que denotamos $-e$ tal que $e+(-e) = 0$.

   d) $e+e' = e'+e$, para todo $e,e' \in E$.

2. $\lambda \cdot (e+v) = \lambda \cdot e + \lambda \cdot v$, $\forall \lambda \in \mathbb{R}$, $e,v \in E$

3. $(\lambda+\mu) \cdot e = \lambda \cdot e + \mu \cdot e$ $\forall \lambda,\mu \in \mathbb{R}$, $e \in E$

4. $(\lambda \cdot \mu) \cdot e = \lambda \cdot (\mu \cdot e)$, $\forall \lambda,\mu \in \mathbb{R}$, $e \in E$

5. $1 \cdot e = e$, $\forall e \in E$.

Los elementos de un espacio vectorial se denominan vectores y los de $\mathbb{R}$ escalares.

2. Ejemplos: $\mathbb{R}^n$, con la suma $(\lambda_i)+(\mu_i) := (\lambda_i+\mu_i)$ y el producto por escalares $\lambda\cdot(\lambda_i) := (\lambda\cdot\lambda_i)$ es un $\mathbb{R}$-espacio vectorial. $\mathbb{R}^3$ que es el espacio en el que pensamos que vivimos es un ejemplo de $\mathbb{R}$-espacio vectorial.

Sea X un conjunto y $C(X) = \text{Aplic}(X,\mathbb{R})$. $C(X)$ con la suma estándar de funciones $(f+g)(x) := f(x)+g(x)$ y producto estándar por escalares $(\lambda\cdot f)(x) := \lambda\cdot f(x)$ es un $\mathbb{R}$-espacio vectorial.

Dados dos espacios vectoriales E,E' se define

$$E \times E' = \{(e,e'),\text{ con } e\in E, e'\in E'\}$$

$E\times E'$ es un espacio vectorial, con la definición de suma de vectores y producto por escalares como sigue

$$(e_1,e'_1)+(e_2,e'_2) := (e_1+e_2,e'_1+e'_2),\quad \lambda\cdot(e_1,e'_1) := (\lambda\cdot e_1,\lambda\cdot e'_1);\quad \forall e_1,e_2\in E,\, e'_1,e'_2\in E',\lambda\in\mathbb{R}$$

En los espacios vectoriales se cumplen las siguientes propiedades:

1. $0\cdot e = (0+0)\cdot e = 0\cdot e+0\cdot e$ y por tanto $0\cdot e = 0$.
2. Si $e+e' = 0$, sumando $-e$, obtenemos que $e' = -e$.
3. Como $0 = (1-1)\cdot e = e+(-1)\cdot e$, tenemos que $(-1)\cdot e = -e$.

3. Definición: Decimos que $F\subset E$ es un *subespacio vectorial* de E, si $f+f'\in F$ y $\lambda\cdot f\in F$, para todo $f,f'\in F$ y $\lambda\in\mathbb{R}$.

El subespacio vectorial, F, con la suma y producto por escalares es un espacio vectorial. Si F_i son subespacios vectoriales de E entonces $\cap_i F_i$ es un subespacio vectorial de E.

Sea $C = \{c_i\}_{i\in I}$, con $c_i\in E$. Llamamos subespacio vectorial de E generado por C, al mínimo subespacio vectorial de E que contiene a C, y lo denotamos $\langle C\rangle$. Es fácil probar que

$$\langle C\rangle = \{e\in E \mid e = \lambda_1\cdot c_1+\cdots+\lambda_n c_n,\ c_i\in C, \lambda_i\in\mathbb{R}, n\in\mathbb{N}\}$$

Si $C = \{e_1,\dots,e_r\}$ entonces denotamos $\langle e_1,\dots,e_r\rangle = \langle C\rangle$ y se cumple que $\langle e_1,\dots,e_r\rangle = \{\lambda_1\cdot e_1+\cdots+\lambda_r e_r,\ \lambda_i\in\mathbb{R}\}$.

Se dice que los vectores de C son un sistema generador de E si $\langle C\rangle = E$. Decimos que un espacio vectorial E es finito generado si existen $e_1,\dots,e_r\in E$ de modo que $\langle e_1,\dots,e_r\rangle = E$. Se dice que los vectores de C son linealmente independientes si $\lambda_1\cdot c_1+\cdots+\lambda_n\cdot c_n\neq 0$ si algún $\lambda_i\neq 0$, para todo n y $\{c_1,\dots,c_n\}\subset C$.

4. Definición: Se dice que los vectores $\{e_1,\dots,e_n\}$ forman una base de E si son un sistema generador de E y son linealmente independientes.

5. Ejemplo : Los vectores $(1,0,\dots,0),(0,1,\dots,0),\dots,(0,0\dots,1)$ forman una base de $\mathbb{R}^n$, denominada "base estándar" de $\mathbb{R}^n$.

Dada una base $\{e_1,\dots,e_n\}$ y un vector $e\in E$, existen $\lambda_1,\dots,\lambda_n\in\mathbb{R}$ tales que

$$e=\lambda_1 e_1+\cdots+\lambda_n e_n$$

porque los e_i generan E. Además, estos λ_i son únicos, porque si $e=\lambda_1 e_1+\cdots+\lambda_n e_n=\lambda'_1 e_1+\cdots+\lambda'_n e_n$, entonces $(\lambda_1-\lambda'_1)e_1+\cdots+(\lambda_n-\lambda'_n)e_n=0$, luego $\lambda_1-\lambda'_1=\cdots=\lambda_n-\lambda'_n=0$, porque los e_i son linealmente independientes. Por tanto, $\lambda_i=\lambda'_i$, para todo i.

Escribiremos "$e\equiv(\lambda_1,\dots,\lambda_n)$ en la base $\{e_1,\dots,e_n\}$" y diremos que los λ_i son las coordenadas de e en la base $\{e_1,\dots,e_n\}$.

6. Lema : *Sea E un espacio vectorial y $e,e_1,\dots,e_n\in E$. Entonces, $e\in\langle e_1,\dots,e_n\rangle$ si y sólo si $\langle e,e_1,\dots,e_n\rangle=\langle e_1,\dots,e_n\rangle$.*

Demostración. Si $e\in\langle e_1,\dots,e_n\rangle$, entonces existen $\lambda_i\in\mathbb{R}$ tales que $e=\sum_i\lambda_i e_i$. Dado $v=\mu e+\mu_1 e_1+\cdots+\mu_n e_n\in\langle e,e_1,\dots,e_n\rangle$, entonces $v=(\mu_1+\lambda_1)e_1+\cdots+(\mu_n+\lambda_n)e_n\in\langle e_1,\dots,e_n\rangle$. Luego, $\langle e,e_1,\dots,e_n\rangle\subseteq\langle e_1,\dots,e_n\rangle$, y como la inclusión inversa es obvia se tiene la igualdad.

Si $\langle e,e_1,\dots,e_n\rangle=\langle e_1,\dots,e_n\rangle$, entonces $e\in\langle e,e_1,\dots,e_n\rangle=\langle e_1,\dots,e_n\rangle$.

□

Los espacios vectoriales que consideremos estarán generados por un número finito de vectores, salvo que se diga lo contrario.

7. Teorema de la base: *Todo espacio vectorial $E\neq 0$ contiene alguna base. Todas las bases de E tienen el mismo número de vectores, tal número se dice que es la dimensión de E y se denota* $\dim_{\mathbb{R}}E$.

Demostración. Supongamos, pues, que $E=\langle e_1,\dots,e_n\rangle$. Sea $I\subset\{1,\dots,n\}$ un subconjunto máximo con la condición de que los vectores $\{e_i\}_{i\in I}$ sean linealmente independientes. Obviamente $I\neq\emptyset$, pues si $I=\emptyset$ entonces $e_i=0$ para todo i y $E=0$. Veamos que los vectores $\{e_i\}_{i\in I}$ forman una base de E. Tenemos que probar que $\langle e_i\rangle_{i\in I}=E$. Dado e_j, $1\leq j\leq n$, si $e_j\notin\langle e_i\rangle_{i\in I}$ entonces $\{e_j,e_i\}_{i\in I}$ serían linealmente independientes, pues si $\lambda_j\cdot e_j+\sum_i\lambda_i e_i=0$ entonces: 1. Si $\lambda_j\neq 0$ tendremos que $e_j=-\sum_i\frac{\lambda_i}{\lambda_j}\cdot e_i$ y $e_j\in\langle e_i\rangle_{i\in I}$, contradicción. 2. Si $\lambda_j=0$, entonces $\sum_i\lambda_i e_i=0$ y entonces $\lambda_i=0$, para todo $i\in I$, pues los vectores $\{e_i\}_{i\in I}$ son linealmente independientes. En conclusión, $\lambda_j=\lambda_i=0$ para todo i, luego $\{e_j,e_i\}_{i\in I}$ son linealmente independientes. Ahora bien, por la maximalidad de I, esto es contradictorio. En conclusión, $e_j\in\langle e_i\rangle_{i\in I}$, para todo $1\leq j\leq n$. Por tanto, $E=\langle e_1,\dots,e_n\rangle\subseteq\langle e_i\rangle_{i\in I}$ y $E=\langle e_i\rangle_{i\in I}$.

Veamos que todas las bases tienen el mismo número de vectores.

Sean $\{e_1,\dots,e_n\}$ l.i. y $\{v_1,\dots,v_m\}$ un sistema generador de E.

1. Veamos que $\{e_1,v_2,\dots,v_m\}$ es un sistema generador de E, reordenando si es necesario los vectores $\{v_1,\dots,v_m\}$: $e_1=\lambda_1v_1+\cdots+\lambda_mv_m$, reordenando los v_i, podemos suponer que $\lambda_1\neq 0$. Entonces, $v_1\in\langle e_1,v_2,\dots,v_m\rangle$. Por tanto, $\langle e_1,v_2,\dots,v_m\rangle = \langle e_1,v_1,v_2,\dots,v_m\rangle = E$.

2. Veamos que $\{e_1,e_2,v_3,\dots,v_m\}$ es un sistema generador de E reordenando si es necesario los vectores $\{v_2,\dots,v_m\}$: $e_2=\lambda_1e_1+\lambda_2v_2\cdots+\lambda_mv_m$. No pueden ser $\lambda_2=\cdots=\lambda_n=0$, porque e_1,e_2 son l.i. Reordenando los v_i, podemos suponer que $\lambda_2\neq 0$. Entonces, $v_2\in\langle e_1,e_2,v_3,\dots,v_m\rangle$. Por tanto, $\langle e_1,e_2,v_3,\dots,v_m\rangle=\langle e_1,e_2,v_2,\dots,v_m\rangle=E$.

3. Así sucesivamente $\{e_1,\dots,e_i,v_{i+1},\dots,v_m\}$ son un sistema generador de E.

Ahora sean $\{e_1,\dots,e_n\}$ y $\{v_1,\dots,v_m\}$ dos bases de E. Si $n>m$, entonces tomando $i=m$, tenemos que $\{e_1,\dots,e_m\}$ es un sistema generador de E, luego e_{m+1} es combinación lineal de ellos, contradicción. Luego, $n\leq m$. Cambiando el papel de los e_i por el de los v_j, tendremos igualmente que $m\leq n$, luego $n=m$.

□

En la demostración del teorema de la base hemos probado que todo subconjunto de vectores de un sistema generador, maximal con la condición de ser linealmente independientes es una base. Así pues, si $e_1,\dots,e_i$ son linealmente independientes, consideremos un sistema generador $v_1,\dots,v_m$ de E. Cualquier subconjunto de vectores de $\{e_1,\dots,e_i,v_1,\dots,v_m\}$ que contenga a $\{e_1,\dots,e_i\}$, maximal con la condición de ser linealmente independientes, formará una base, es decir, hemos obtenido una base "ampliada" que contiene a $\{e_1,\dots,e_i\}$.

Si $E'\underset{\neq}{\subset}E$ es un subespacio vectorial entonces $\dim_{\mathbb{R}}E'<\dim_{\mathbb{R}}E$, porque dada una base en E' la podemos ampliar a una base en E, luego las bases en E tienen más vectores que las de E'.

8. Proposición: *Sean E y E' dos espacios vectoriales. Entonces,*

$$\dim(E\times E')=\dim E+\dim E'$$

Demostración. Sea $\{e_1,\dots,e_n\}$ una base de E y $\{e'_1,\dots,e'_m\}$ una base de E'. Entonces, $\{(e_1,0),\dots,(e_n,0),(0,e'_1),\dots,(0,e'_m)\}$ es una base de $E\times E'$ (al lector). Luego, $\dim(E\times E')=n+m=\dim E+\dim E'$.

□

Sean E_1,E_2 dos subespacios vectoriales de E. Se denota por E_1+E_2 el mínimo subespacio vectorial de E que contiene a E_1 y E_2. Explícitamente,

$$E_1+E_2=\{e=e_1+e_2\in E,\ \text{con } e_1\in E_1, e_2\in E_2\}$$

Se dice que E_1 y E_2 están en suma directa si $E_1\cap E_2=0$ y en este caso E_1+E_2 lo denotamos $E_1\oplus E_2$. Todo vector $e\in E_1\oplus E_2$ se escribe de modo único como suma de un vector de E_1 y otro de E_2: si $e = e_1 + e_2 = e'_1 + e'_2$, con $e_1, e'_1 \in E_1$ y $e_2, e'_2 \in E_2$, entonces $0 = (e_1 - e'_1) + (e_2 - e'_2)$, luego $e_1 - e'_1 = e_2 - e'_2 \in E_1 \cap E_2 = \{0\}$ y $e_1 = e'_1$, $e_2 = e'_2$. Ahora ya, si $\{e_1, \dots, e_n\}$ una base de E_1 y $\{e'_1, \dots, e'_m\}$ una base de E_2, entonces $\{e_1, \dots, e_n, e'_1, \dots, e'_m\}$ una base de $E_1 \oplus E_2$. Por tanto,

$$\dim(E_1 \oplus E_2) = \dim E_1 + \dim E_2$$

1.2. Aplicaciones lineales. Matrices

Las aplicaciones que conservan la estructura de espacio vectorial son las aplicaciones lineales. Con precisión:

1. Definición: Sean E, E' dos $\mathbb{R}$-espacios vectoriales. Una aplicación $T : E \to E'$ es un *morfismo de $\mathbb{R}$-espacios vectoriales* (o *aplicación $\mathbb{R}$-lineal*) si

$$T(e+v) = T(e) + T(v) \quad \text{y} \quad T(\lambda \cdot e) = \lambda \cdot T(e)$$

para cualesquiera $e, v \in E$, $\lambda \in \mathbb{R}$.

Observemos que $T(0) = 0$: $T(0) = T(0+0) = T(0) + T(0)$, y sumando $-T(0)$ en ambos lados de la igualdad, obtenemos $0 = T(0)$.

Dadas dos aplicaciones lineales $T, S : E \to E'$, la aplicación $T + S : E \to E'$ definida por

$$(T+S)(e) := T(e) + S(e), \qquad \forall e \in E$$

es lineal. Dada $\lambda \in \mathbb{R}$, la aplicación $\lambda \cdot T : E \to E'$ definida por $(\lambda \cdot T)(e) := \lambda \cdot T(e)$ es lineal. El conjunto de las aplicaciones lineales de E en E', que denotamos $\mathrm{Hom}_{\mathbb{R}}(E, E')$, es un $\mathbb{R}$-espacio vectorial.

Es claro que la composición $T_1 \circ T_2 : E \to G$ de dos aplicaciones lineales $T_2 : E \to F$, $T_1 : F \to G$, es una aplicación lineal.

2. Definición: Sea $T : E \to E'$ una aplicación lineal. Se define el núcleo de T, que denotamos $\mathrm{Ker}\, T$, como sigue

$$\mathrm{Ker}\, T := \{e \in E : T(e) = 0\} = T^{-1}(0)$$

Se cumple que $\mathrm{Ker}\, T$ es un subespacio vectorial de E.

3. Proposición: *Sea $T : E \to E'$ una aplicación lineal, y $e, v \in E$. Entonces, $T(v) = T(e)$ si y sólo si existe $w \in \mathrm{Ker}\, T$ de modo que $v = e + w$.*

Demostración. Evidentemente, dado $w \in \mathrm{Ker}\, T$, $T(e+w) = T(e) + T(w) = T(e)$. Recíprocamente, si $T(v) = T(e)$, entonces $0 = T(v) - T(e) = T(v-e)$ y $w := v - e \in \mathrm{Ker}\, T$. Por tanto, $v = e + w$. □

4. Proposición: *Una aplicación* $T\colon E \to E'$ *es inyectiva si y sólo si* $\operatorname{Ker} T = \{0\}$.

Demostración. Es consecuencia inmediata de la proposición anterior. □

5. Definición: Sea $T\colon E \to E'$ una aplicación lineal. Se define la imagen de T, que denotamos $\operatorname{Im} T$, como sigue

$$\operatorname{Im} T := \{T(e) \in E', e \in E\})$$

Se cumple que $\operatorname{Im} T$ es un subespacio vectorial de E'. Si $E = \langle e_1, \dots, e_n \rangle$ entonces $\operatorname{Im} T = \langle T(e_1), \dots, T(e_n) \rangle$, pues dado $e \in E$, tenemos que $e = \sum_i \lambda_i e_i$ y $T(e) = \sum_i \lambda_i T(e_i)$.

6. Definición: Una aplicación lineal $T : E \to E'$ es un *isomorfismo* si existe otra aplicación lineal $S : E' \to E$ tal que $T \circ S = Id_{E'}$, $S \circ T = Id_E$.

La aplicación lineal T es un isomorfismo de espacios vectoriales si y sólo si es una aplicación biyectiva.

7. Proposición: *Sea E un espacio vectorial y $\{e_1, \dots, e_n\}$ una base de E. Sea $T\colon E \to E'$ una aplicación lineal. Entonces,*

1. *T es inyectiva si y sólo si $\{T(e_1), \dots, T(e_n)\}$ son linealmente independientes.*

2. *T es un epimorfismo si y sólo si $\{T(e_1), \dots, T(e_n)\}$ generan E'.*

3. *T es un isomorfismo si y sólo si $\{T(e_1), \dots, T(e_n)\}$ es una base de E'.*

Demostración. 1. Si T es inyectiva, $\operatorname{Ker} T = 0$. Si $\lambda_1 T(e_1) + \dots + \lambda_n T(e_n) = 0$ entonces $T(\lambda_1 e_1 + \dots + \lambda_n e_n) = 0$, luego $\lambda_1 e_1 + \dots + \lambda_n e_n = 0$ y $\lambda_1 = \dots = \lambda_n = 0$. Por tanto, $\{T(e_1), \dots, T(e_n)\}$ son linealmente independientes. Si $\{T(e_1), \dots, T(e_n)\}$ son linealmente independientes entonces dado $e = \sum_i \lambda_i e_i$, si $0 = T(e) = \sum_i \lambda_i T(e_i)$ entonces $\lambda_i = 0$, para todo i y $e = 0$. Luego $\operatorname{Ker} T = 0$ y T es inyectiva.

2. Inmediato, porque $\operatorname{Im} T = \langle T(e_1), \dots, T(e_n) \rangle$.

□

8. Corolario: *Si $T\colon E \to E'$ es un isomorfismo lineal, entonces* $\dim_{\mathbb{R}} E = \dim_{\mathbb{R}} E'$.

9. Corolario: *Sea E un espacio vectorial y $\{e_1, \dots, e_n\}$ una base. Entonces, la aplicación lineal $T\colon \mathbb{R}^n \to E$, $T((\lambda_1, \cdots, \lambda_n)) := \lambda_1 e_1 + \dots + \lambda_n e_n$ es un isomorfismo lineal.*

10. Definición: Se denomina rango de la aplicación lineal T a la dimensión del espacio vectorial $\operatorname{Im} T$.

11. Proposición: *Sea E un espacio vectorial y $\{e_1, \dots, e_n\}$ una base de E. Sea $T\colon E \to E'$ una aplicación lineal. Entonces,*

1. *T es inyectiva si y sólo si* $\dim E = \operatorname{rg} T$.

2. *T es un epimorfismo si y sólo si* $\operatorname{rg} T = \dim E'$.

3. *T es un isomorfismo si y sólo si* $\dim E = \operatorname{rg} T$ *y* $\operatorname{rg} T = \dim E'$.

Demostración. Es consecuencia de 1.2.7. □

Observemos que una aplicación lineal está determinada por lo que vale en los vectores de una base. Efectivamente, si $\{e_1, \dots, e_n\}$ es una base de E, y $e \in E$, tendremos que $e = \sum_i x_i e_i$, para ciertos $x_i \in \mathbb{R}$ y tendremos que $T(e) = \sum_i x_i \cdot T(e_i)$. Si $\{e_1, \dots, e_n\}$ es una base de E, dados $v_1, \dots, v_n \in E'$ la aplicación $T\colon E \to E'$, definida por $T(\sum_i \lambda_i e_i) := \sum_i \lambda_i v_i$ es lineal y cumple que $T(e_i) = v_i$.

12. Definición: Sean E y E' dos espacios vectoriales, $B = \{e_1, \dots, e_n\}$ una base de E y $B' = \{e'_1, \dots, e'_m\}$ una base de E'. Sea $T\colon E \to E'$ una aplicación lineal. Tenemos que

$$T(e_i) = \sum_{j=1}^{m} \lambda_{ji} e'_j \equiv (\lambda_{1i}, \cdots, \lambda_{mi}), \qquad \text{en la base } B',$$

para ciertas $\lambda_{ji} \in \mathbb{R}$ únicas.

Diremos que la caja de números

$$\begin{pmatrix} \lambda_{11} & \cdots & \lambda_{1n} \\ \vdots & & \vdots \\ \lambda_{m1} & \cdots & \lambda_{mn} \end{pmatrix}$$

de columnas $T(e_i) \equiv (\lambda_{1i}, \cdots, \lambda_{mi})$, es la matriz de T en las bases B y B'. Escribiremos $T \equiv (\lambda_{ji})$.

Si escribimos $e = (x_1, \dots, x_n)$ en la base B, $T(e) = (x'_1, \dots, x'_m)$ en la base B', entonces tenemos que $x'_j = \sum_{i=1}^n \lambda_{ji} x_i$, que escribiremos de modo reducido

$$\begin{pmatrix} x'_1 \\ \vdots \\ x'_m \end{pmatrix} = \begin{pmatrix} \lambda_{11} & \cdots & \lambda_{1n} \\ \vdots & & \vdots \\ \lambda_{m1} & \cdots & \lambda_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

13. Ejercicio: Sean B, B' bases de E y E'. Sea $T, S\colon E \to E'$ dos aplicaciones lineales de matrices asociadas en las bases B y B', (a_{ij}) y (b_{ij}). Calcular la matriz asociada a $T + S$ en las bases B y B'. Dada $\lambda \in \mathbb{R}$, calcular la matriz asociada a $\lambda \cdot T$.

Calculemos ahora la matriz de la composición de dos aplicaciones lineales.

Sean E, E' y E'' espacios vectoriales, $B = \{e_1, \dots, e_n\}$ una base de E y $B' = \{e'_1, \dots, e'_m\}$ una base de E' y $B'' = \{e''_1, \dots, e''_r\}$ una base de E. Sea $T\colon E \to E'$ una aplicación lineal de matriz (λ_{ji}) en las bases B, B'. Sea $S\colon E' \to E''$ una aplicación lineal de matriz (μ_{kj}) en las bases B', B. ¿Cuál es la matriz (c_{ki}) de $S \circ T\colon E \to E''$ en las bases B, B''?

$$(S\circ T)(e_i)=S(\sum_j \lambda_{ji}e'_j)=\sum_j \lambda_{ji}S(e'_j)=\sum_j\sum_k \lambda_{ji}\cdot\mu_{kj}\cdot e''_k=\sum_k\sum_j \mu_{kj}\cdot\lambda_{ji}\cdot e''_k$$

Luego $c_{ki}=\sum_{1\leq j\leq m}\mu_{kj}\cdot\lambda_{ji}$ ("la fila k de (μ_{rs}) por la columna i de (λ_{uv})"), que escribiremos de modo abreviado

$$(c_{ki})=(\mu_{kj})\cdot(\lambda_{ji})$$

1.3. Fórmulas de cambio de base

1. Definición: Sea E un espacio vectorial y $B_1=\{e_1,\dots,e_n\}$, $B_2=\{v_1,\dots,v_n\}$ dos bases de E. Sean $e_i=\sum_j a_{ji}v_j$, para cada i. Se dice que (a_{ij}) es la matriz de cambio de base de B_1 a B_2.

2. Proposición: *Sea E un espacio vectorial y $B_1=\{e_1,\dots,e_n\}$, $B_2=\{v_1,\dots,v_n\}$ dos bases de E. Sean $e_i=\sum_j a_{ji}v_j$, para cada i. Sea $e\in E$, $(\lambda_1,\dots,\lambda_n)$ las coordenadas de e en la base B_1 y $(\lambda'_1,\dots,\lambda'_n)$ las coordenadas de e en la base B_2. Entonces,*

$$\begin{pmatrix}\lambda'_1\\ \vdots\\ \lambda'_n\end{pmatrix}=\begin{pmatrix}a_{11} & \cdots & a_{1n}\\ \vdots & & \vdots\\ a_{1n} & \cdots & a_{nn}\end{pmatrix}\cdot\begin{pmatrix}\lambda_1\\ \vdots\\ \lambda_n\end{pmatrix}$$

Demostración. Sea $\mathrm{Id}\colon E\to E$, la aplicación lineal identidad. La matriz de Id en las bases B_1,B_2 es (a_{ij}), luego si $e=(\lambda_1,\dots,\lambda_n)$ en la base B_1, entonces $\mathrm{Id}(e)=e=(\lambda'_1,\dots,\lambda'_n)$ en la base B_2, de modo que

$$\begin{pmatrix}\lambda'_1\\ \vdots\\ \lambda'_n\end{pmatrix}=\begin{pmatrix}a_{11} & \cdots & a_{1n}\\ \vdots & & \vdots\\ a_{1n} & \cdots & a_{nn}\end{pmatrix}\cdot\begin{pmatrix}\lambda_1\\ \vdots\\ \lambda_n\end{pmatrix}$$

□

Sea E un espacio vectorial de base $B=\{e_1,\dots,e_n\}$ y $\mathrm{Id}\colon E\to E$ la aplicación lineal identidad, $\mathrm{Id}(e)=e$. La matriz de Id en las bases B y B es

$$\mathrm{Id}\equiv\begin{pmatrix}1 & 0 & \cdots & 0\\ 0 & 1 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & 1\end{pmatrix}$$

Sea $T: E \to E'$ es un isomorfismo lineal y (λ_{ij}) la matriz de T en las bases B y B'. Diremos que la matriz de $T^{-1}: E' \to E$ en las bases B' y B es la matriz inversa de (λ_{ij}) y la denotaremos $(\lambda_{ij})^{-1}$, que es la matriz que cumple

$$(\lambda_{ij})^{-1} \cdot (\lambda_{ij}) = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

porque $T^{-1} \circ T = \mathrm{Id}$.

3. Proposición: *Sea $T: E \to E'$ una aplicación lineal. Sean $B_1 = \{e_1, \ldots, e_n\}$ y $B_2 = \{v_1, \ldots, v_n\}$ dos bases de E y supongamos que $v_i = \sum_j a_{ji} e_j$. Sean $B'_1 = \{e'_1, \ldots, e'_m\}$ y $B'_2 = \{v'_1, \ldots, v'_m\}$ dos bases de E y supongamos que $v'_i = \sum_j a'_{ji} e'_j$.*

Si (λ_{ij}) es la matriz de T en las bases B_1, B'_1, entonces la matriz (μ_{ij}) de T en las bases B_2, B'_2 es igual a

$$(\mu_{ij}) = (a'_{ij})^{-1} \cdot (\lambda_{ij}) \cdot (a_{ij})$$

Demostración. Se deduce del diagrama conmutativo

$$\begin{tikzcd} B_1 & E \arrow[r, "(\lambda_{ij})", "T"'] & E' & B'_1 \\ B_2 & E \arrow[u, "(a_{ij})", "\mathrm{Id}"'] \arrow[r, "T", "(\mu_{ij})"'] & E' \arrow[u, "\mathrm{Id}", "(a'_{ij})"'] & B'_2 \end{tikzcd}$$

□

1.4. Aplicaciones lineales equivalentes

1. Definición: Diremos que dos aplicaciones lineales $T, S: E \to E'$ son equivalentes si existen bases B_1, B_2 de E y bases B'_1, B'_2 de E' de modo que la matriz asociada a T en las bases B_1, B'_1 sea igual a la matriz asociada a S en la bases B_2, B'_2.

2. Proposición: *Dos aplicaciones lineales $T, S: E \to E'$ son equivalentes $\iff$ existen isomorfismos lineales $\phi: E \to E$ y $\phi': E' \to E'$, tales que el diagrama*

$$\begin{tikzcd} E \arrow[r, "T"] \arrow[d, "\phi"] & E' \arrow[d, "\phi'"] \\ E \arrow[r, "S"] & E' \end{tikzcd}$$

es conmutativo.

Demostración. $\Rightarrow$) Sean $B_1=\{e_1,\dots,e_n\}$, $B_2=\{v_1,\dots,v_n\}$ bases de E y $B'_1=\{e'_1,\dots,e'_m\}$, $B'_2=\{v'_1,\dots,v'_m\}$ bases de E' de modo que la matriz (a_{ij}) asociada a T en las bases B_1,B'_1 sea igual a la matriz asociada a S en la bases B_2,B'_2. Sea $\phi\colon E\to E$, la aplicación lineal definida por $\phi(e_i):=v_i$, para todo i (de inversa obvia) y $\phi'\colon E'\to E'$, la aplicación lineal definida por $\phi'(e_i):=v'_i$, para todo i (de inversa obvia). El diagrama

$$\begin{tikzcd} B_1 & E \arrow[r, "(a_{ij})", "T"'] \arrow[d, "(\mathrm{Id})"', "\phi"] & E' \arrow[d, "\phi'"', "(\mathrm{Id})"] & B'_1 \\ B_1 & E \arrow[r, "S", "(a_{ij})"'] & E' & B'_2 \end{tikzcd}$$

es conmutativo.

$\Leftarrow$) Sean $\phi\colon E\to E$ y $\phi'\colon E'\to E'$, tales que el diagrama

$$\begin{tikzcd} E \arrow[r, "T"] \arrow[d, "\phi"] & E' \arrow[d, "\phi'"] \\ E \arrow[r, "S"] & E' \end{tikzcd}$$

es conmutativo. Sea $B_1=\{e_1,\dots,e_n\}$ una base de E, y $B'_1=\{e'_1,\dots,e'_m\}$ una base de E' y (a_{ij}) la matriz de T en las bases B_1,B'_1. Entonces, $B_2:=\{\phi(e_1),\dots,\phi(e_n)\}$ es una base de E y $B'_2:=\{\phi'(e'_1),\dots,\phi'(e'_m)\}$ es una base de E'. Del diagrama conmutativo

$$\begin{tikzcd} B_1 & E \arrow[r, "(a_{ij})", "T"'] \arrow[d, "(\mathrm{Id})"', "\phi"] & E' \arrow[d, "\phi'"', "(\mathrm{Id})"] & B'_1 \\ B_2 & E \arrow[r, "S"] & E' & B'_2 \end{tikzcd}$$

se deduce que la matriz asociada a S en las bases B_2, B'_2 es (a_{ij}).

□

3. Lema : *Sea* $\{e_1,\dots,e_s\}$ *una base de* $\operatorname{Ker} T$ *y sea* $\{e_1,\dots,e_s,\dots,e_n\}$ *una base de* E. *Entonces,* $T(e_{s+1}),\dots,T(e_n)$ *es una base de* $\operatorname{Im} T$.

Demostración. $T(e_{s+1}),\dots,T(e_n)$ generan $\operatorname{Im} T$: Dado $T(e)\in\operatorname{Im} T$, escribamos $e=\lambda_1e_1+\cdots+\lambda_se_s+\cdots+\lambda_ne_n$. Entonces, $T(e)=\lambda_{s+1}T(e_{s+1})+\cdots+\lambda_nT(e_n)$.

Son linealmente independientes: Si $\lambda_{s+1}T(e_{s+1})+\cdots+\lambda_nT(e_n)=0$, entonces $T(\lambda_{s+1}e_{s+1}+\cdots+\lambda_ne_n)=0$. Luego, $e=\lambda_{s+1}e_{s+1}+\cdots+\lambda_ne_n\in\operatorname{Ker} T$ y $e=\lambda_1e_1+\cdots+\lambda_se_s$, para ciertos $\lambda_i\in\mathbb{R}$. Por tanto,

$$0=\lambda_1e_1+\cdots+\lambda_se_s-(\lambda_{s+1}e_{s+1}+\cdots+\lambda_ne_n)$$

y $\lambda_i=0$, para todo i. □

4. Proposición: *Sea $T: E \to E'$ una aplicación lineal. Entonces,*

$$\operatorname{rg} T = \dim E - \dim \operatorname{Ker} T$$

5. Teorema: *Sea $T: E \to E'$ una aplicación lineal. Existen bases $B = \{e_1, \dots, e_n\}$ de E y $B' = \{e'_1, \dots, e'_m\}$ de modo que la matriz asociada a T es la matriz (por bloques)*

$$\left(\begin{array}{c|c} \mathrm{Id} & 0 \\ \hline 0 & 0 \end{array}\right), \quad \mathrm{Id} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & r & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & 1 \end{pmatrix}$$

Además, r ha de ser igual al rango de T. Diremos que esta matriz es la matriz reducida asociada a T.

Demostración. Sea $n = \dim E$ y $n - r = \dim \operatorname{Ker} T$. Sea $e_{r+1}, \dots, e_n$ una base de $\operatorname{Ker} T$ y ampliemos a una base $B = \{e_1, \dots, e_r, \dots, e_n\}$ de E. Sabemos por el lema 1.4.3, que $T(e_1), \dots, T(e_r)$ es una base de $\operatorname{Im} T$. Ampliemos a una base $B' = \{e'_1 = T(e_1), \dots, e'_r = T(e_r), e'_{r+1}, \dots, e'_m\}$ de E'. La matriz de T en las bases B, B' es la requerida.

Además si tenemos una matriz reducida asociada a T, es claro que la dimensión de $\operatorname{Im} T$ es r.

$\square$

6. Corolario: *Dos aplicaciones lineales $T, S: E \to E'$ son equivalentes si y sólo si tienen el mismo rango.*

Demostración. Si tienen el mismo rango, las matrices reducidas asociadas son iguales, luego T y S son equivalentes. Si son equivalentes entonces $\operatorname{Im} T \simeq \operatorname{Im} S$, luego tienen el mismo rango. $\square$

7. Teorema: *Sea $T: E \to E'$ una aplicación lineal y $B = \{e_1, \dots, e_n\}$ una base de E. Reordenando los vectores de la base de B si es necesario, existe una base B' de E' de modo que la matriz de T en las base B, B' es igual a*

$$\left(\begin{array}{c|c} \mathrm{Id} & (a_{ij}) \\ \hline 0 & 0 \end{array}\right)$$

De nuevo r ha de ser igual al rango de T. Esta matriz se dice que una matriz escalonada por filas asociada a T.

Demostración. $\operatorname{Im} T = \langle T(e_1), \dots, T(e_n)\rangle$. Cualquier subconjunto de $\{T(e_1), \dots, T(e_n)\rangle$ maximal con la condición de ser linealmente independientes es una base de $\operatorname{Im} T$. Reordenando los vectores de $B = \{e_1, \dots, e_n\}$, puedo suponer que $T(e_1), \dots, T(e_r)$ es una base de $\operatorname{Im} T$. Sea $B' = \{e'_1 = T(e_1), \dots, e'_r = T(e_r), e'_{r+1}, \dots, e'_m\}$ una base de E'. Observemos que $T(e_i) \in \operatorname{Im} T = \langle T(e_1), \dots, T(e_r)\rangle$, luego existen $a_{ij} \in \mathbb{R}$, tales que $T(e_j) = \sum_{1 \le i \le r} a_{ij} T(e_i)$, para $r < j \le n$. La matriz de T en las bases B, B' es la requerida. □

Expresemos los dos teoremas anteriores en términos meramente matriciales.

8. Definición: Diremos que dos matrices $A, B \in M_{m\times n}(\mathbb{R})$ son equivalentes cuando existen matrices cuadradas invertibles $P \in M_{n\times n}(\mathbb{R})$ y $Q \in M_{m\times n}$ de modo que $B = Q \cdot A \cdot P$.

9. Teorema: *Si $A \in M_{m\times n}(\mathbb{R})$ es una matriz de rango r, entonces es equivalente a la matriz reducida*

$$\left(\begin{array}{c|c} \mathrm{Id} & 0 \\ \hline 0 & 0 \end{array}\right)$$

10. Teorema: *Sea $A \in M_{m\times n}(\mathbb{R})$ es una matriz de rango r, entonces existe una matriz cuadrada invertible $P \in M_{m\times m}(\mathbb{R})$, de modo que*

$$P \cdot A = \left(\begin{array}{c|c} \mathrm{Id} & (a_{ij}) \\ \hline 0 & 0 \end{array}\right) \qquad (\textit{reordenando las columnas})$$

es una matriz reducida por filas.

1.5. Sistemas de ecuaciones lineales

Consideremos el sistema de m ecuaciones lineales

$$(*) \qquad \begin{array}{rcl} a_{11}x_1 + \cdots + a_{1n}x_n & = & b_1 \\ \cdots\cdots\cdots\cdots\cdots\cdots & = & \cdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n & = & b_m \end{array}$$

de variables $x_1, \dots, x_n$ y coeficientes $a_{ij}, b_i \in \mathbb{R}$. Queremos calcular las soluciones de este sistema de ecuaciones, es decir, los valores que han de tomar las $x_1, \dots, x_n$ para que se cumplan las m igualdades.

Podemos escribir el anterior sistema de modo matricial

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

Así pues, si consideramos la aplicación lineal $T\colon \mathbb{R}^n \to \mathbb{R}^m$, cuya matriz en las bases estándar de $\mathbb{R}^n$ y $\mathbb{R}^m$, es (a_{ij}), y consideramos el vector $b = (b_1,\dots,b_m) \in \mathbb{R}^m$, las soluciones del sistema $(*)$ son aquellos vectores $x = (x_1,\dots,x_n) \in \mathbb{R}^n$ tales que $T(x) = b$.

Denotemos por $v_1 = T((1,0,\cdots,0)) = (a_{11},\dots,a_{m1}),\dots, v_n = T((0,\cdots,0,1) = (a_{1n},\dots,a_{mn})$ los vectores columnas de la matriz (a_{ij}).

1. Proposición: *El sistema de ecuaciones $(*)$ tiene solución si y sólo si*

$$\dim\langle b, v_1,\dots,v_n\rangle = \dim\langle v_1,\dots,v_n\rangle.$$

Demostración. El sistema $(*)$ tiene solución $\iff b \in \operatorname{Im} T = \langle v_1,\dots,v_n\rangle \iff \langle b,v_1,\dots,v_n\rangle = \langle v_1,\dots,v_n\rangle \iff \dim\langle b,v_1,\dots,v_n\rangle = \dim\langle v_1,\dots,v_n\rangle$. □

2. Proposición: *Si el sistema de ecuaciones $(*)$ tiene una solución x', entonces*

$$\{\textit{Soluciones del sistema } (*)\} = \{x' + y,\ \forall y \in \operatorname{Ker} T\}$$

3. Definición: Se dice que un sistema de ecuaciones es compatible si existen soluciones.

4. Proposición: *Supongamos que el sistema de ecuaciones $(*)$ es compatible. Entonces tiene una única solución $\iff$ el rango de (a_{ij}) es igual a n.*

Demostración. El sistema de ecuaciones $(*)$ tiene una única solución $\iff \operatorname{Ker} T = 0 \iff \dim \operatorname{Ker} T = 0 \iff$ rango $(a_{ij}) = n - \dim \operatorname{Ker} T = n$. □

1.6. Cálculos

1.6.1. Matriz traspuesta

1. Definición: Sea $A = (a_{ij}) \in M_{m\times n}(\mathbb{R})$ una matriz. Llamaremos matriz traspuesta de A, que denotaremos $A^t = (a^t_{ij}) \in M_{n\times m}(\mathbb{R})$, a la matriz definida por $a^t_{ij} := a_{ji}$ (que es la matriz A^t cuyas filas son las columnas de A).

2. Proposición: *Se cumple que*

1. $(A\cdot B)^t = B^t \cdot A^t$, *para todo* $A \in M_{n\times m}(\mathbb{R})$ *y* $B \in M_{m\times n}(\mathbb{R})$.
2. $(A^t)^{-1} = (A^{-1})^t$, *para toda matriz cuadrada invertible* A.

Demostración. 1. Escribamos $A = (a_{ij})$, $B = (b_{ij})$, $A\cdot B = C = (c_{ij})$, donde $c_{ij} = \sum_l a_{il}\cdot b_{lj}$; $B^t\cdot A^t = C' = (c'_{ij})$ donde $c'_{ij} = \sum_l b^t_{il}\cdot a^t_{lj}$. Tenemos que ver que $c^t_{ij} = c'_{ij}$, ahora bien $c^t_{ij} = c_{ji} = \sum_l a_{jl}\cdot b_{li} = c'_{ij}$.

2. Como $A\cdot A^{-1} = \mathrm{Id}$, trasponiendo $(A^{-1})^t \cdot A^t = \mathrm{Id}^t = \mathrm{Id}$, luego $(A^t)^{-1} = (A^{-1})^t$.

□

3. Definición: Sea $A=(a_{ij})\in M_{m\times n}(\mathbb{R})$ una matriz y

$$v_1=(a_{11},\dots,a_{m1}),\dots,v_n=(a_{1n},\dots a_{mn})$$

los vectores columna de A. Llamaremos rango de A, que denotamos por $\operatorname{rg}(A)$, al número máximo de columnas linealmente independientes, es decir, a $\dim_{\mathbb{R}}\langle v_1,\dots,v_n\rangle$.

Si $T\colon\mathbb{R}^n\to\mathbb{R}^m$ es la aplicación lineal cuya matriz asociada en las bases estándar es A, entonces $\operatorname{rg}(T)=\operatorname{rg}(A)$, porque $\operatorname{rg}(T)=\dim_k\operatorname{Im}T=\dim_k\langle v_1,\dots,v_n\rangle$. Si $S\colon\mathbb{R}^n\to\mathbb{R}^n$ es un isomorfismo lineal, entonces $\operatorname{rg}(T\circ S)=\operatorname{rg}(T)$, porque $\operatorname{Im}(T\circ S)=\operatorname{Im}(T)$. Si $S'\colon\mathbb{R}^m\to\mathbb{R}^m$ es un isomorfismo lineal, entonces $\operatorname{rg}(S'\circ T)=\operatorname{rg}(T)$, porque $S'\colon\operatorname{Im}T\to\operatorname{Im}(S'\circ T)$ es un isomorfismo lineal, luego $\dim_k\operatorname{Im}T=\dim\operatorname{Im}(S'\circ T)$. En conclusión, si $B\in M_{n\times n}(\mathbb{R})$ y $B'\in M_{m\times m}(\mathbb{R})$, son matrices invertibles $\operatorname{rg}(A)=\operatorname{rg}(B'\cdot A\cdot B)$.

4. Proposición: *Se cumple que* $\operatorname{rg}(A)=\operatorname{rg}(A^t)$.

Demostración. Dada $A\in M_{m\times n}(\mathbb{R})$, por el teorema 1.4.9, existen matrices invertibles $P\in M_{m\times m}(\mathbb{R})$ y $Q\in M_{n\times n}(\mathbb{R})$, de modo que

$$A=P\cdot\left(\begin{array}{c|c}\mathrm{Id}&0\\\hline 0&0\end{array}\right)\cdot Q$$

Luego,

$$A^t=Q^t\cdot\left(\begin{array}{c|c}\mathrm{Id}&0\\\hline 0&0\end{array}\right)\cdot P^t$$

Por tanto, $\operatorname{rg}(A)=\operatorname{rg}\left(\begin{array}{c|c}\mathrm{Id}&0\\\hline 0&0\end{array}\right)=\operatorname{rg}A^t$. □

1.6.2. Transformaciones elementales

Sea E un $\mathbb{R}$-espacio vectorial de base $\{e_1,\dots,e_n\}$. Sea $E'=\langle v_1,\dots,v_r\rangle\subseteq E$ un subespacio vectorial y supongamos $v_i=\sum_j a_{ij}e_j$. Si $a_{11}\neq 0$, sean

$$v_i':=v_i-(a_{i1}/a_{11})\cdot v_1=(a_{i2}-\frac{a_{i1}a_{12}}{a_{11}})e_2+\cdots+(a_{in}-\frac{a_{i1}a_{1n}}{a_{11}})e_n\subseteq\langle e_2,\dots,e_n\rangle,\ \text{para } 1<i\le r.$$

Observemos que $\langle v_1,v_2,\dots,v_r\rangle=\langle v_1,v_2',\dots,v_r'\rangle$. Matricialmente, si

$$\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&&\vdots\\a_{r1}&a_{r2}&\cdots&a_{rn}\end{pmatrix}$$

es la matriz de filas $v_1,\dots,v_r$, entonces

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22}-\frac{a_{21}a_{12}}{a_{11}} & \cdots & a_{2n}-\frac{a_{21}a_{1n}}{a_{11}} \\ \vdots & \vdots & & \vdots \\ 0 & a_{r2}-\frac{a_{r1}a_{12}}{a_{11}} & \cdots & a_{rn}-\frac{a_{r1}a_{1n}}{a_{11}} \end{pmatrix}$$

es la matriz de filas, $v_1, v_2' = v_2-(a_{21}/a_{11})\cdot v_1,\dots,v_r' = v_r-(a_{r1}/a_{11})\cdot v_1$. Además,

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ -a_{21}/a_{11} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -a_{r1}/a_{11} & 0 & \cdots & 1 \end{pmatrix} \cdot \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{r1} & a_{r2} & \cdots & a_{rn} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22}-\frac{a_{21}a_{12}}{a_{11}} & \cdots & a_{2n}-\frac{a_{21}a_{1n}}{a_{11}} \\ \vdots & \vdots & & \vdots \\ 0 & a_{r2}-\frac{a_{r1}a_{12}}{a_{11}} & \cdots & a_{rn}-\frac{a_{r1}a_{1n}}{a_{11}} \end{pmatrix}$$

Si $a_{11}=0$ y algún $a_{i1}\neq 0$, definiremos $v_1' = v_1+v_i$ y podremos suponer que $a_{11}\neq 0$. Por último permutando los vectores de la base $\{e_1,\dots,e_n\}$, podemos suponer que algún $a_{i1}\neq 0$ (salvo que todos los a_{ij} sean nulos).

Cuando a una fila de una matriz le restemos otra fila multiplicada por un factor diremos que hemos hecho una transformación elemental de la matriz. Así sucesivamente, por transformaciones elementales de filas (salvo reordenación de las columnas) obtendremos una matriz triangular

$$\left(\begin{array}{c|c} T & B \\ \hline 0 & 0 \end{array}\right), \quad T = \begin{pmatrix} \lambda_{11} & \cdots & \lambda_{1s} \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_{ss} \end{pmatrix}, \quad B = (b_{ij})$$

con $\lambda_{11},\dots,\lambda_{ss}$ no nulos y tendremos que $\langle v_1,\dots,v_r\rangle$ esta generado por las s filas de $\left(\begin{array}{c|c} T & B \end{array}\right)$ y que $\dim_{\mathbb{R}}\langle v_1,\dots,v_r\rangle = s$.

Si seguimos operando con las filas podemos obtener una matriz

$$\left(\begin{array}{c|c} \Delta & C \\ \hline 0 & 0 \end{array}\right), \quad \Delta = \begin{pmatrix} \lambda_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_{ss} \end{pmatrix}, \quad C = (c_{ij})$$

Además habremos ido calculando la matriz P tal que $P\cdot A = \left(\begin{array}{c|c} \Delta & C \\ \hline 0 & 0 \end{array}\right)$.

Si tenemos un sistema de ecuaciones lineales $A\cdot x = b$, entonces (reordenando las variables)

$$\left(\begin{array}{c|c} \Delta & C \\ \hline 0 & 0 \end{array}\right)\cdot x = P\cdot A\cdot x = P\cdot b =: b'$$

El sistema es compatible si $b'_{s+1} = \cdots = b'_n = 0$ y en este caso las soluciones son

$$\begin{pmatrix} x_1 \\ \vdots \\ x_s \end{pmatrix} = \Delta^{-1} \cdot \begin{pmatrix} b'_1 \\ \vdots \\ b'_s \end{pmatrix} - \Delta^{-1} \cdot C \cdot \begin{pmatrix} x_{s+1} \\ \vdots \\ x_n \end{pmatrix}$$

Observemos que para $b = 0$ (luego $b' = 0$), estamos calculando $\operatorname{Ker} A$.

5. Observación: Sea $A \in M_{m\times n}(\mathbb{R})$. Mediante transformaciones elementales de A por columnas (salvo una permutación de las filas) obtenemos igualmente una matriz

$$\left(\begin{array}{c|c} \Delta & 0 \\ \hline C & 0 \end{array}\right), \quad \Delta = \begin{pmatrix} \lambda_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_{ss} \end{pmatrix}, \quad C = (c_{ij})$$

Además habremos ido calculando la matriz $P \in M_{n\times n}(\mathbb{R})$ tal que $A \cdot P = \left(\begin{array}{c|c} \Delta & 0 \\ \hline C & 0 \end{array}\right)$.

1.6.3. Determinante de una matriz. Matriz inversa

6. Notación: Sea $N_n = \{1, 2, \dots, n\}$, diremos que una biyección $\sigma \colon N_n \to N_n$ una es una permutación de N_n. El conjunto de todas las permutaciones de N_n lo denotaremos S_n.

7. Definición: Consideremos el polinomio $\Delta(x_1, \dots, x_n) = \prod_{i<j}(x_i - x_j)$. Dada una permutación $\sigma \in S_n$, tendremos que $\Delta(x_{\sigma(1)}, \dots, x_{\sigma(n)}) = \prod_{i<j}(x_{\sigma(i)} - x_{\sigma(j)}) = \pm\Delta(x_1, \dots, x_n)$. Se define $\operatorname{signo}(\sigma)$ al número real ± 1, que cumple

$$\Delta(x_{\sigma(1)}, \dots, x_{\sigma(n)}) = \operatorname{signo}(\sigma) \cdot \Delta(x_1, \dots, x_n) \qquad (*)$$

Consideremos una permutación $\tau \in S_n$, si en la igualdad $(*)$ sustituimos x_j por $x_{\tau(j)}$, para todo x_j, tendremos $\Delta(x_{\tau(\sigma(1))}, \dots, x_{\tau(\sigma(n))}) = \operatorname{signo}(\sigma) \cdot \Delta(x_{\tau(1)}, \dots, x_{\tau(n)})$, luego

$$\begin{aligned} \operatorname{signo}(\tau \circ \sigma) \cdot \Delta(x_1, \dots, x_n) &= \Delta(x_{\tau(\sigma(1))}, \dots, x_{\tau(\sigma(n))}) = \operatorname{signo}(\sigma) \cdot \Delta(x_{\tau(1)}, \dots, x_{\tau(n)}) \\ &= \operatorname{signo}(\sigma) \cdot \operatorname{signo}(\tau) \cdot \Delta(x_1, \dots, x_n) \end{aligned}$$

Por tanto, $\operatorname{signo}(\tau \circ \sigma) = \operatorname{signo}(\sigma) \cdot \operatorname{signo}(\tau)$. Como $\operatorname{signo}(\mathrm{Id}) = 1$, entonces $1 = \operatorname{signo}(\sigma \circ \sigma^{-1}) = \operatorname{signo}(\sigma) \cdot \operatorname{signo}(\sigma^{-1})$ y $\operatorname{signo}(\sigma) = \operatorname{signo}(\sigma^{-1})$.

Dados $i, j \in N_n$ distintos, denotaremos $\sigma = (i, j)$ la permutación que cumple que $\sigma(i) = j$, $\sigma(j) = i$ y $\sigma(k) = k$, para todo $k \neq i, j$.

Es fácil comprobar que $\operatorname{signo}((1,2)) = -1$. Dados i, j distintos, sea σ una permutación cualquiera tal que $\sigma(1) = i$ y $\sigma(2) = j$. Entonces, $\sigma \circ (1,2) \circ \sigma^{-1} = (i, j)$. Por tanto,

$$\operatorname{signo}((i,j)) = \operatorname{signo}(\sigma \circ (1,2) \circ \sigma^{-1}) = \operatorname{signo}(\sigma) \cdot \operatorname{signo}((1,2)) \cdot \operatorname{signo}(\sigma^{-1}) = -1$$

8. Definición: Sea $(a_{ij})_{1\leq i,j\leq n}$ una matriz "cuadrada". Llamaremos determinante de las matriz (a_{ij}), que denotaremos $\det(a_{ij})$ o $|a_{ij}|$, al número real definido por

$$\det(a_{ij}) := \sum_{\sigma\in S_n} \operatorname{signo}(\sigma)\cdot a_{\sigma(1)1}\cdots a_{\sigma(n)n}$$

9. Ejercicio: $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - cb.$

10. Ejercicio: $\begin{vmatrix} d_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & d_n \end{vmatrix} = d_1\cdots d_n.$

11. Proposición: *Sea $(a_{ij})_{1\leq i,j\leq n}$ una matriz cuadrada. Entonces,*

$$\det((a_{ij})) = \det((a_{ij})^t)$$

Demostración.

$$\begin{aligned}\det((a_{ij})^t) &= \sum_{\sigma\in S_n} \operatorname{signo}(\sigma)\cdot a^t_{\sigma(1)1}\cdots a^t_{\sigma(n)n} = \sum_{\sigma\in S_n} \operatorname{signo}(\sigma)\cdot a_{1\sigma(1)}\cdots a_{n\sigma(n)} \\ &= \sum_{\sigma\in S_n} \operatorname{signo}(\sigma)\cdot a_{\sigma^{-1}(1)1}\cdots a_{\sigma^{-1}(n)n} = \sum_{\sigma\in S_n} \operatorname{signo}(\sigma^{-1})\cdot a_{\sigma^{-1}(1)1}\cdots a_{\sigma^{-1}(n)n} \\ &= \det((a_{ij}))\end{aligned}$$

□

Dados n-vectores $v_i = (a_{1i},\dots,a_{ni})$ definiremos $\det(v_1,\dots,v_n) = \det(a_{ij})$.

12. Proposición: *Se cumple que*

1. $\det(v_1,\dots,\lambda\cdot v_i,\dots,v_n) = \lambda\cdot\det(v_1,\dots,v_i,\dots,v_n)$, *luego* $\det(v_1,\dots,0,\dots,v_n) = 0$.
2. $\det(v_1,\dots,v_i+v'_i,\dots,v_n) = \det(v_1,\dots,v_i,\dots,v_n) + \det(v_1,\dots,v'_i,\dots,v_n)$.
3. $\det(v_1,\overset{i}{\dots},v_i,\dots,v_j,\dots,v_n) = -\det(v_1,\overset{i}{\dots},v_j,\dots,v_i,\dots,v_n)$. *En particular, se cumple que* $\det(v_1,\overset{i}{\dots},v,\dots,v,\dots,v_n) = 0$.
4. *Si v_i es combinación lineal de los demás v_j, entonces* $\det(v_1,\dots,v_n) = 0$.
5. $\det(v_1,\overset{i}{\dots},v_i+\sum_{j\neq i}\lambda_j v_j,\cdots,v_n) = \det(v_1,\dots,v_n)$.

Demostración. 1. $\det(v_1,\dots,\lambda\cdot v_i,\dots,v_n) = \sum\limits_{\sigma\in S_n} \operatorname{signo}(\sigma)\cdot a_{\sigma(1)1}\cdots\lambda\cdot a_{\sigma(i)i}\cdots a_{\sigma(n)n} = \lambda\cdot$ $\sum\limits_{\sigma\in S_n} \operatorname{signo}(\sigma)\cdot a_{\sigma(1)1}\cdots a_{\sigma(i)i}\cdots a_{\sigma(n)n} = \lambda\cdot\det(v_1,\dots,v_i,\dots,v_n)$.

2. $\det(v_1,\dots,v_i+v'_i,\dots,v_n) = \sum_{\sigma\in S_n} \text{signo}(\sigma)\cdot a_{\sigma(1)1}\cdots(a_{\sigma(i)i}+a'_{\sigma(i)i})\cdots a_{\sigma(n)n}$

$$= \sum_{\sigma\in S_n} \text{signo}(\sigma)\cdot a_{\sigma(1)1}\cdots a_{\sigma(i)i}\cdots a_{\sigma(n)n} + \sum_{\sigma\in S_n} \text{signo}(\sigma)\cdot a_{\sigma(1)1}\cdots a'_{\sigma(i)i}\cdots a_{\sigma(n)n}$$
$$= \det(v_1,\dots,v_i,\dots,v_n)+\det(v_1,\dots,v'_i,\dots,v_n)$$

3. Denotemos la matriz de columnas $v_1,\overset{i}{\dots},v_j,\dots,v_i,\dots,v_n$, (b_{ij}). Para cada permutación σ sea $\sigma'=\sigma\circ(i,j)$. Observemos que $(\sigma')'=\sigma$. Además, $\text{signo}(\sigma')=-\text{signo}(\sigma)$. Por tanto,

$$\text{signo}(\sigma)\cdot a_{\sigma(1)1}\cdots a_{\sigma(i)i}\cdots a_{\sigma(j)j}\cdots a_{\sigma(n)n} = -\text{signo}(\sigma')\cdot a_{\sigma'(1)1}\cdots a_{\sigma'(j)i}\cdots a_{\sigma'(i)j}\cdots a_{\sigma'(n)n}$$
$$-\text{signo}(\sigma')\cdot b_{\sigma'(1)1}\cdots b_{\sigma'(j)j}\cdots b_{\sigma'(i)i}\cdots b_{\sigma'(n)n}$$

Luego, $\det(v_1,\overset{i}{\dots},v_i,\dots,v_j,\dots,v_n) = -\det(v_1,\overset{i}{\dots},v_j,\dots,v_i,\dots,v_n)$.

4. $\det(v_1,\overset{i}{\cdots},\sum_{j\neq i}\lambda_j v_j,\cdots,v_n) \overset{1.,2.}{=} \sum_{j\neq i}\det(v_1,\overset{i}{\cdots},v_j,\cdots,v_n) \overset{3}{=} 0.$

□

13. Definición: $A=(a_{ij})$ una matriz. Llamaremos matriz adjunta rs de A, que denotaremos $Ad(A)_{rs}$, a la matriz que se obtiene de eliminar en A la fila r y la columna s.

14. Proposición: *Sea $A=(a_{ij})_{1\leq i,j\leq n}$ una matriz cuadrada. Entonces,*

$$\det(A) = a_{11}\det(Ad(A)_{11}) - a_{21}\det(Ad(A)_{21}) + \cdots + (-1)^{n-1}a_{n1}\cdot\det(Ad(A)_{n1})$$

Demostración. Sea $B=(b_{ij})$ una matriz, tal que los coeficientes de la columna k-ésima cumplen que $b_{jk}=0$ para todo j, salvo $b_{kk}=1$. Es fácil ver que para esta matriz que $\det(b_{ij})=\det(Ad(B)_{kk})$.

Sean $v_i=(a_{1i},\dots,a_{ni})$, para todo i, es decir, en la base estándar $\mathbb{R}^n$, $v_i=\sum_j a_{ji}e_j$. Entonces, $\det(A)=\det(v_1,\dots,v_n)=\det(\sum_k a_{k1}e_k,v_2,\dots,v_n)=\sum_k a_{k1}\cdot\det(e_k,v_2,\dots,v_n)$ y permutando e_k con v_2, después con v_3 y así $k-1$ veces, obtendremos que

$$\det(A)=\sum_k(-1)^{k-1}\cdot a_{k1}\cdot\det(v_2,\overset{k}{\cdots},e_k,\cdots,v_n) = \sum_k(-1)^{k-1}\cdot a_{k1}\cdot\det(Ad(A)_{k1})$$

□

15. Observación: Hemos calculado el determinate “desarrollando a partir de la primera columna”, si queremos desarrollarlo a partir de la columna i-ésima observemos que

$$det(v_1,v_2,\dots,v_i,\dots,v_n) = (-1)^{i-1}\cdot det(v_i,v_1,v_2,\dots,v_n)$$

Luego, $\det(A)=(-1)^{i-1}\cdot(a_{1i}\det(Ad(A)_{1i})-a_{2i}\det(Ad(A)_{2i})+\cdots+(-1)^{n-1}a_{ni}\cdot\det(Ad(A)_{ni}))=\sum_j(-1)^{j+i}a_{ji}\det(Ad(A)_{ji})$.

Si queremos desarrollarlo por filas en vez de por columnas recordemos que $\det(A)=\det(A^t)$. Por tanto, $\det(A)=\sum_j(-1)^{j+i}a_{ij}\det(Ad(A)_{ij})$.

16. Ejercicio: Demostrar que $\begin{vmatrix} a & b & c\\ d & e & f\\ g & h & i\end{vmatrix}=aei+dhc+bfg-(gec+dbi+hfa)$.

Si A no es invertible, entonces alguna columna es combinación lineal de las demás, luego $\det(A)=0$. Sea A una matriz cuadrada invertible (es decir, las columnas son linealmente independientes). Mediante sucesivas transformaciones elementales por columnas podemos obtener una matriz diagonal

$$\Delta=\begin{pmatrix} d_1 & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & d_n\end{pmatrix},\quad d_i\neq 0,\forall i$$

y tendremos que $\det(A)=\det(\Delta)=d_1\cdots d_n\neq 0$. Por tanto, una matriz es invertible si y sólo si su determinante es distinto de cero.

17. Proposición: $\det(A\cdot B)=\det(A)\cdot\det(B)$.

Demostración. Si B no es invertible, entonces $A\cdot B$ no es invertible. En tal caso, $\det(A\cdot B)=0=\det(A)\cdot\det(B)$.

Supongamos que $\det(B)\neq 0$. Por transformaciones elementales de columnas

$$B\equiv\Delta=\begin{pmatrix} d_1 & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & d_n\end{pmatrix}$$

y por transformaciones elementales de columnas $A\cdot B\equiv A\cdot\Delta$ y

$$\det(A\cdot B)=\det(A\cdot\Delta)=\det(A)\cdot d_1\cdots d_n=\det(A)\cdot\det(\Delta)=\det(A)\cdot\det(B)$$

□

18. Proposición: *Si A es una matriz cuadrada invertible, entonces* $\det(A^{-1})=\det(A)^{-1}$.

Demostración. $1=\det(\mathrm{Id})=\det(A\cdot A^{-1})=\det(A)\cdot\det(A^{-1})$.

□

19. Teorema: *Sea $A=(a_{ij})\in M_n(\mathbb{R})$ una matriz. Sea $B=(b_{ij})$ la matriz definida por*

$$b_{ij}:=(-1)^{i+j}\det(Ad(A)_{ji})$$

Entonces, $A\cdot B=\det(A)\cdot\mathrm{Id}$. Por tanto, si A es invertible, $A^{-1}=B/\det(A)$.

Demostración. Escribamos $(c_{ij}) = A \cdot B$. Entonces, $c_{ij} = \sum_k a_{ik} \cdot b_{kj} = \sum_k (-1)^{k+j} \cdot a_{ik} \cdot \det(Ad(A)_{jk})$. Luego c_{ij} es el determinante de la matriz igual a A salvo que se ha sustituido la fila i de A por la fila j de A (y la fila j sigue siendo la fila j). Si una matriz tiene dos filas iguales su determinante es nulo. Por tanto, $c_{ij} = 0$ si $i \neq j$ y $c_{ii} = \det(A)$. En conclusión, $(c_{ij}) = \det(A) \cdot \mathrm{Id}$.

□

1.7. Problemas

1. Calcular las coordenadas de un vector de $\mathbb{R}^3$ respecto de la base

$$\mathscr{B}_1 = \{(1,2,3),\ (3,4,0),\ (1,1,0)\}$$

sabiendo que sus coordenadas respecto de la base $\mathscr{B}_2 = \{(1,1,0), (0,1,1), (1,0,1)\}$ son $(1,1,1)$.

2. Sean $\mathscr{B}_1 = \{\mathbf{e}_1, \mathbf{e}_2\}$, $\mathscr{B}_2 = \{\mathbf{u}_1, \mathbf{u}_2\}$ y $\mathscr{B}_3 = \{\mathbf{v}_1, \mathbf{v}_2\}$ tres bases de $\mathbb{R}^2$ tales que $\mathbf{u}_1 = \mathbf{e}_1$, $\mathbf{u}_2 = 2\mathbf{e}_1 + \mathbf{e}_2$, $\mathbf{v}_1 = \mathbf{e}_1$ y $\mathbf{v}_2 = \mathbf{e}_1 + 4\mathbf{e}_2$. Usando las matrices de cambio de bases, calcular las coordenadas del vector $\mathbf{u} = 2\mathbf{u}_1 + 5\mathbf{u}_2$ respecto de la base $\mathscr{B}_3$.

3. Dada la aplicación lineal $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ definida por $f(x,y,z) = (2x+y, y-z)$, calcular la matriz asociada a T respecto de:

   1. las bases usuales de $\mathbb{R}^3$ y $\mathbb{R}^2$;
   2. las bases $\mathscr{B} = \{(1,1,1),(0,1,2),(0,2,1)\}$ de $\mathbb{R}^3$ y $\mathscr{B}' = \{(2,1),(1,0)\}$ de $\mathbb{R}^2$.

4. Obtener la matriz asociada a la aplicación lineal $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ determinada por la igualdades $f(1,2) = (1,1,2)$, $f(2,3) = (2,10,1)$ respecto de las bases $\mathscr{B} = \{(1,1),(1,3)\}$ de $\mathbb{R}^2$ y $\mathscr{B}' = \{(1,0,1),(1,1,0),(0,0,2)\}$ de $\mathbb{R}^3$.

5. Calcular el rango de la matriz

$$\begin{pmatrix} 2 & 2 & 2 & 1 & 1 & 4 \\ -1 & -1 & -3 & 0 & 2 & -1 \\ 1 & 2 & 1 & 1 & 1 & 3 \\ 3 & 1 & 2 & -2 & -1 & -1 \\ 4 & -2 & -2 & -6 & 0 & 8 \end{pmatrix}.$$

6. Sea $E' = \langle (1,2,3,4),(2,3,1,1),(5,8,5,6),(0,0,0,1) \rangle \subseteq \mathbb{R}^4$. Calcular una base de E'.

7. Sea $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ la aplicación lineal definida como $T(x,y) = (x+y, x+y, x+y)$.

1. Hallar la matriz asociada a T en las bases usuales.
2. Calcular bases de $\ker(T)$ e $\mathrm{Im}(T)$.

8. Consideremos la aplicación lineal $T : \mathbb{R}^3 \to \mathbb{R}^4$ que respecto de las bases usuales de $\mathbb{R}^3$ y $\mathbb{R}^4$ viene dada por

$$T(x,y,z) = (x+z, y+z, x+z, y+z)$$

1. Calcular la matriz A de T respecto de las bases usuales de $\mathbb{R}^3$ y $\mathbb{R}^4$.
2. Calcular el rango r de A y determinar matrices P y Q tales que

$$Q^{-1}AP = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}.$$

3. Escribir una base de $\ker(T)$.
4. Escribir una base de $\mathrm{Im}(T)$.

9. En $\mathbb{R}^3$ consideramos una base $\mathscr{B}$ fija. Sean T y $S \in \mathrm{End}_{\mathbb{R}}(\mathbb{R}^3)$ tales que sus matrices asociadas respecto de $\mathscr{B}$ son A y B, donde

$$A = \begin{pmatrix} 1 & 1 & 2 \\ 2 & 1 & 1 \\ 1 & 2 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & 2 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 0 \end{pmatrix}.$$

Calcular las matrices asociadas a las aplicaciones $S \circ T$ y $T \circ S$ respecto de $\mathscr{B}$.

10. Se llama **determinante de Vandermonde** de unos ciertos escalares $(x_1, \dots, x_n)$ al determinante definido por la igualdad

$$V(x_1,\dots,x_n) = \begin{vmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ x_1^2 & x_2^2 & \dots & x_n^2 \\ \vdots & \vdots & & \vdots \\ x_1^{n-1} & x_2^{n-1} & \dots & x_n^{n-1} \end{vmatrix}.$$

Probar la siguiente relación de recurrencia:

$$V(x_1,\dots,x_n) = (x_n - x_1)\cdot(x_{n-1}-x_1)\cdot\ldots\cdot(x_2-x_1)\cdot V(x_2,\dots,x_n).$$

Concluir de lo anterior la siguiente igualdad: $V(x_1,\dots,x_n) = \prod_{i<j}(x_j - x_i)$. Como consecuencia, el determinante de Vandermonde de unos escalares es igual a 0 si y sólo si entre dichos escalares hay dos iguales.

Como aplicación de lo anterior probar que se satisface la igualdad

$$\begin{vmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & 2 & 2^2 & \dots & 2^{n-1} \\ 1 & 3 & 3^2 & \dots & 3^{n-1} \\ \dots & \dots & \dots & \ddots & \dots \\ 1 & n & n^2 & \dots & n^{n-1} \end{vmatrix} = 1! \cdot 2! \cdot \dots \cdot (n-1)!.$$

11. Sea $A \in \mathscr{M}_{m\times n}(\mathbb{R})$. Probar las siguientes igualdades y afirmaciones
    1. $(A^{\mathrm{t}})^{\mathrm{t}} = A$.
    2. $(A+B)^{\mathrm{t}} = A^{\mathrm{t}} + B^{\mathrm{t}}$, para cualquier matriz $B \in \mathscr{M}_{m\times n}(\mathbb{R})$.
    3. $(A\cdot B)^{\mathrm{t}} = B^{\mathrm{t}} \cdot A^{\mathrm{t}}$, para cualquier matriz $B \in \mathscr{M}_{n\times p}(\mathbb{R})$.
    4. Si A es invertible, $(A^{-1})^{\mathrm{t}} = (A^{\mathrm{t}})^{-1}$.
    5. Si A tiene coeficientes reales, entonces $A^{\mathrm{t}} \cdot A = 0$ si, sólo si, $A = 0$.

12. Sea $A \in M_n(\mathbb{R})$. Probar que
    1. $(A + A^{\mathrm{t}})$ es simétrica y $(A - A^{\mathrm{t}})$ es antisimétrica.
    2. $A = \frac{1}{2}(A + A^{\mathrm{t}}) + \frac{1}{2}(A - A^{\mathrm{t}})$
    3. A puede escribirse, de modo único,como suma de una matriz simétrica y otra antisimétrica.

13. Diremos que una matriz N cuadrada de orden n es **nilpotente** si existe un número natural $r \geq 1$ tal que $N^r = 0_n$. Probar que si N es nilpotente, entonces la matriz $I_n - N$ es invertible y, además:

$$(I-N)^{-1} = I_n + N + N^2 + \ldots + N^{r-1}.$$

Como aplicación, calcular la matriz inversa de la matriz siguiente:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

14. Probar que A y B son invertibles si, y sólo si, $A \oplus B$ es invertible. En tal caso $(A \oplus B)^{-1} = A^{-1} \oplus B^{-1}$.

15. Consideremos la matriz cuadrada

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},$$

con A_{11} y A_{22} matrices cuadradas. Probar que si A_{11} es invertible, entonces

$$|A| = |A_{11}| \cdot |A_{22} - A_{21} A_{11}^{-1} A_{12}|.$$

Capítulo 2

Endomorfismos. Matrices cuadradas

2.1. Máximo común divisor de dos polinomios

1. Definición : Llamaremos máximo común divisor de dos polinomios $p(x), q(x) \neq 0 \in \mathbb{R}[x]$, que lo denotaremos por $m.c.d.(p(x), q(x))$, al polinomio mónico de grado máximo que divide a $p(x)$ y $q(x)$.

Si $p(x) = a \cdot (x-\alpha_1)^{n_1} \cdots (x-\alpha_r)^{n_r}$ y $q(x) = b \cdot (x-\alpha_1)^{m_1} \cdots (x-\alpha_r)^{m_r}$, con $n_i, m_i \geq 0$, entonces

$$m.c.d.(p(x), q(x)) = (x-\alpha_1)^{min(n_1,m_1)} \cdots (x-\alpha_r)^{min(n_r,m_r)}$$

Vamos a desarrollar el algoritmo de Euclides para calcular el máximo común divisor de dos polinomios $p_1(x), p_2(x) \neq 0$: Sea $p_3(x)$ el resto de dividir $p_1(x)$ por $p_2(x)$. Es decir, tenemos que

$$p_1(x) = c_1(x) \cdot p_2(x) + p_3(x),$$

con $\mathrm{gr}(p_3(x)) < \mathrm{gr}(p_2(x))$ (ó $p_3(x) = 0$). Observemos que un polinomio divide a $p_1(x)$ y a $p_2(x)$ si y sólo si divide a $p_2(x)$ y $p_3(x)$. Luego, $m.c.d.(p_1(x), p_2(x)) = m.c.d.(p_2(x), p_3(x))$. Si $p_3(x) \neq 0$, repitamos la división con $p_2(x)$ y $p_3(x)$: tenemos

$$p_2(x) = c_2(x) \cdot p_3(x) + p_4(x),$$

con $\mathrm{gr}(p_4(x)) < \mathrm{gr}(p_3(x))$ (ó $p_3(x) = 0$). De nuevo, $m.c.d.(p_2(x), p_3(x)) = m.c.d.(p_3(x), p_4(x))$. Así sucesivamente, vamos obteniendo polinomios $p_1(x), p_2(x), p_3(x), \ldots, p_n(x)$. Este proceso termina cuando $p_n(x) = 0$ (recordemos que $\mathrm{gr}(p_2) > \mathrm{gr}(p_3(x) > \mathrm{gr}(p_3(x) > ...)$. Por tanto,

$$\begin{aligned} m.c.d.(p_1(x), p_2(x)) &= m.c.d.(p_2(x), p_3(x)) = \cdots = m.c.d.(p_{n-1}(x), p_n(x)) = m.c.d.(p_{n-1}(x), 0) \\ &= p_{n-1}(x). \end{aligned}$$

Observemos que

$$\begin{aligned}
m.c.d.(p_1(x),p_2(x)) &= p_{n-1}(x) = p_{n-3}(x) - c_{n-3}(x)p_{n-2}(x)\\
&= p_{n-3}(x) - c_{n-3}(x)(p_{n-4}(x) - c_{n-4}(x)p_{n-3}(x))\\
&= -c_{n-3}(x))p_{n-4}(x) + (1 + c_{n-3}(x)c_{n-4}(x))p_{n-3}(x) = \cdots = \lambda(x)p_1(x) + \mu(x)p_2(x)
\end{aligned}$$

2. Ejercicio : Calcular el máximo común divisor de $x^3 - 3x^2 + 3x + 1$ y $x^2 - 1$.

Se dice que dos polinomios de grados mayor que cero, $p(x)$ y $q(x)$, son primos entre sí si el máximo común divisor de $p(x)$ y $q(x)$ es 1. Si $p(x)$ y $q(x)$ son primos entre sí existen dos polinomios $\lambda(x), \mu(x)$ tales que $\lambda(x)\cdot p(x) + \mu(x)\cdot q(x) = 1$.

2.2. Aplicaciones

Sea $T\colon E \to E$ un endomorfismo $\mathbb{R}$-lineal.

Dado un polinomio $p(x) = a_n x^n + \cdots + a_1 x + a_0 \in \mathbb{R}[x]$, denotaremos $p(T) = a_n \cdot T^n + \cdots + a_1 \cdot T + a_0 := a_n \cdot (T \circ \overset{n}{\cdots} \circ T) + \cdots + a_1 \cdot T + a_0 \cdot \mathrm{Id}$. Observemos que dados $p(x), q(x) \in \mathbb{R}[x]$, si $h(x) = p(x)\cdot q(x)$, entonces $h(T) = p(T)\circ q(T)$. Además, como $q(x)\cdot p(x) = h(x)$, entonces $q(T)\circ p(T) = h(T) = p(T)\circ q(T)$.

1. Proposición: *Sea $T\colon E \to E$ un endomorfismo lineal. Sean $p_1(x), p_2(x)$ primos entre sí y $p(x) = p_1(x)\cdot p_2(x)$. Entonces,*

$$\operatorname{Ker} p(T) = \operatorname{Ker} p_1(T) \oplus \operatorname{Ker} p_2(T)$$

Demostración. Sean $\lambda(x), \mu(x)$ tales que $\lambda(x)\cdot p_1(x) + \mu(x)\cdot p_2(x) = 1$. Luego,

$$\mathrm{Id} = \lambda(T)\cdot p_1(T) + \mu(T)\cdot p_2(T)$$

Dado $e \in \operatorname{Ker} p(T)$, entonces

$$e = \mathrm{Id}(e) = (\lambda(T)\cdot p_1(T) + \mu(T)\cdot p_2(T))(e) = \lambda(T)(p_1(T)(e)) + \mu(T)(p_2(T)(e))$$

Ahora bien, $\lambda(T)(p_1(T)(e)) \in \operatorname{Ker} p_2(T)$, porque $p_2(T)(\lambda(T)(p_1(T)(e))) = \lambda(T)(p(T)(e)) = 0$, e igualmente $\mu(T)(p_2(T)(e)) \in \operatorname{Ker} p_1(T)$. Por tanto, $\operatorname{Ker} p(T) = \operatorname{Ker} p_1(T) + \operatorname{Ker} p_2(T)$.

Por último, $\operatorname{Ker} p_1(T) \cap \operatorname{Ker} p_2(T) = 0$, porque si $e \in \operatorname{Ker} p_1(T) \cap \operatorname{Ker} p_2(T)$, entonces

$$e = \mathrm{Id}(e) = (\lambda(T)\cdot p_1(T) + \mu(T)\cdot p_2(T))(e) = 0 + 0 = 0$$

□

2.2.1. Ecuaciones diferenciales de orden n con coeficientes constantes

Sea F el conjunto de todas las funciones de los números reales a valores complejos infinitamente diferenciables. F es un $\mathbb{C}$-espacio vectorial. Sea $D\colon F \to F$, $D(f(x)) = f'(x)$ el operador derivada. Es claro que D es un endomorfismo $\mathbb{C}$ lineal de F.

2. Fórmula de conmutación: *Sea $P(x) \in \mathbb{C}[x]$. Para toda $y \in F$ y $\alpha \in \mathbb{C}$, se cumple que*

$$P(D)(e^{\alpha x} \cdot y) = e^{\alpha x} \cdot P(D+\alpha)(y)$$

Demostración. $D(e^{\alpha x} \cdot y) = \alpha \cdot e^{\alpha x} \cdot y + e^{\alpha x} \cdot D(y) = e^{\alpha x} \cdot (D+\alpha)(y)$.

$D^2(e^{\alpha x}\cdot y) = D(D(e^{\alpha x}\cdot y)) = D(e^{\alpha x}\cdot(D+\alpha)(y)) = e^{\alpha x}\cdot(D+\alpha)((D+\alpha)(y)) = e^{\alpha x}\cdot(D+\alpha)^2(y)$.

Así sucesivamente, $D^n(e^{\alpha x} \cdot y) = e^{\alpha x} \cdot (D+\alpha)^n(y)$ y para $P(D) = \sum_i a_i D^i$ tendremos que

$$P(D)(e^{\alpha x} \cdot y) = \sum_i a_i D^i (e^{\alpha x} \cdot y) = \sum_i a_i \cdot e^{\alpha x} \cdot (D+\alpha)^i(y) = e^{\alpha x} \cdot P(D+\alpha)(y)$$

□

3. Teorema: *Se cumple que*

1. $\operatorname{Ker} D^r = \{$*Polinomios de grado menor que* $r\}$.
2. $\operatorname{Ker}(D-\alpha)^r = e^{\alpha x} \cdot \{$*Polinomios de grado menor* $r\}$.
3. *Si* $p(x) = (x-\alpha_1)^{n_1} \cdots (x-\alpha_r)^{n_r}$, *entonces*

$$\operatorname{Ker} p(D) = e^{\alpha_1 x} \cdot \{\text{Pol. de grado menor } n_1\} \oplus \overset{r}{\cdots} \oplus e^{\alpha_r x} \cdot \{\text{Pol. de grado menor } n_r\}$$

Demostración. 1. Si $p(x) = a_s x^s + a_{s-1}x^{s-1} + \cdots + a_0 \in \mathbb{C}[x]$ es un polinomio de grado s, entonces $D(p(x)) = (sa_s)x^{s-1} + (s-1)a_{s-1}x^{s-2} + \cdots + 0$ es un polinomio de grado $s-1$. Por tanto, si $p(x)$ es un polinomio de grado menor o igual que r, $D^r(p(x)) = 0$.

Veamos el recíproco. Procedamos por inducción sobre r. $\operatorname{Ker} D$ es igual a las funciones constantes, es decir, a los polinomios de grado menor que 1. Supongamos $r > 1$. Dada $f(x) \in \operatorname{Ker} D^r$, tendremos que $0 = D^r(f(x)) = D^{r-1}(D(f(x))$, luego por hipótesis de inducción $D(f(x)) = a_{r-2}x^{r-2} + \cdots + a_0$, para ciertas a_i. Luego $f(x) = \int a_{r-2}x^{r-2} + \cdots + a_0$ que es un polinomio de grado menor que r.

2. $(D-\alpha)^r f(x) = 0 \iff 0 = (D-\alpha)^r(e^{\alpha x} \cdot e^{-\alpha x} \cdot f(x)) = e^{\alpha x} \cdot D^r(e^{-\alpha x} \cdot f(x)) \iff 0 = D^r(e^{-\alpha x} \cdot f(x)) \iff e^{-\alpha x} \cdot f(x)$ es un polinomio de grado menor que $r \iff f(x)$ es $e^{\alpha x}$ multiplicado por un polinomio de grado menor que r.

3. Es consecuencia de que $\operatorname{Ker} p(D) = \operatorname{Ker}(D-\alpha_1)^{n_1} \oplus \cdots \oplus \operatorname{Ker}(D-\alpha_r)^{n_r}$ y de 2.

□

4. Ejercicio : Resolver las ecuaciones diferenciales: $y'''' - 2y''' + 2y'' = 0$, $y'' + y = 0$.

Consideremos una ecuación diferencial $P(D)(y) = z$, con $z \in F$. Sea y_0 una solución particular. Entonces, $y \in F$ cumple que $P(D)(y) = z$ si y sólo si

$$y = y_0 + y_1, \text{ con } y_1 \in \operatorname{Ker} P(D)$$

Métodos para calcular soluciones particulares.

1. Queremos resolver una ecuación $P(D)y = z$, con $z \in F$, de la que sabemos que existe un polinomio $Q(x)$ primo con $P(x)$ tal que $Q(D)z = 0$:

Sean $\lambda(x)$ y $\mu(x)$ tales que $\lambda(x)\cdot P(x) + \mu(x)\cdot Q(x) = 1$. Por tanto, $\lambda(D)\cdot P(D) + \mu(D)\cdot Q(D) = \mathrm{Id}$. Si aplicamos esta igualdad a z, obtenemos

$$z = \lambda(D)\cdot P(D)(z) = P(D)(\lambda(D)(z))$$

Por tanto, una solución particular es $y = \lambda(D)(z)$.

5. Ejercicio : Resolver la ecuación $y^{(n} - y = x^n$.

2. Vamos a denotar $\int f = \frac{1}{D} f$ y en general $\int \overset{n}{\cdots} \int f = \frac{1}{D^n} f$.

Sea $P(x) = (x-\alpha_1)^{n_1} \cdots (x-\alpha_r)^{n_r}$. Existen polinomios $Q_1(x), \dots, Q_r(x)$ tales que

$$\frac{1}{(x-\alpha_1)^{n_1}\cdots(x-\alpha_r)^{n_r}} = \frac{Q_1(x)}{(x-\alpha_1)^{n_1}} + \overset{r}{\cdots} + \frac{Q_r(x)}{(x-\alpha_r)^{n_r}}$$

Resolvamos la ecuación diferencial $P(D)y = z$.

$$y = \frac{1}{P(D)} z = \frac{Q_1(D)}{(D-\alpha_1)^{n_1}} + \overset{r}{\cdots} + \frac{Q_r(D)}{(D-\alpha_r)^{n_r}} = \sum_i \frac{Q_i(D)}{(D-\alpha_i)^{n_i}} z = \sum_i e^{\alpha_i x} \cdot \frac{Q_i(D+\alpha_i)}{(D)^{n_i}} e^{-\alpha_i x} \cdot z$$
$$\overset{2.2.2}{=} \sum_i e^{\alpha_i x} \cdot Q_i(D+\alpha_i) \cdot \int \overset{n_i}{\cdots} \int e^{-\alpha_i x} \cdot z$$

6. Ejercicio : Resolver $y'' - y = senx$.

2.2.2. Ecuaciones en diferencias finitas

Sea $Suc(\mathbb{C}) = \{(a_n)\}$ el $\mathbb{C}$-espacio vectorial de las sucesiones de números complejos. Sea el "operador siguiente" $\nabla\colon Suc(\mathbb{C}) \to Suc(\mathbb{C})$ la aplicación $\mathbb{C}$-lineal definida por $\nabla(a_n) = (a'_n)$, donde $a'_n = a_{n+1}$. Sea $\Delta = \nabla - \mathrm{Id}$, el "operador diferencia".

7. Fórmulas de conmutación : *Sea $P(x) \in \mathbb{C}[x]$. Entonces para toda sucesión $(a_n) \in Suc(\mathbb{C})$ y $\alpha \in \mathbb{C}$, se cumple que*

$$1.\ P(\nabla)((\alpha^n)\cdot(a_n)) = (\alpha^n)\cdot P(\alpha\nabla)(a_n)$$
$$2.\ P(\nabla-\alpha)((\alpha^n)\cdot(a_n)) = (\alpha^n)\cdot P(\alpha\cdot\Delta)(a_n)$$

Demostración. $\nabla((\alpha^n)\cdot(a_n)) = (\alpha^{n+1})\cdot(a_{n+1}) = \alpha\cdot(\alpha^n)\cdot\nabla(a_n) = (\alpha^n)\cdot(\alpha\nabla)(a_n)$.

$\nabla^2((\alpha^n)\cdot(a_n)) = \nabla(\nabla((\alpha^n)\cdot(a_n))) = \nabla((\alpha^n)\cdot(\alpha\nabla)(a_n)) = (\alpha^n)\cdot(\alpha\nabla)((\alpha\nabla)(a_n)) = (\alpha^n)\cdot(\alpha\nabla)^2(a_n)$.

Así sucesivamente, $\nabla^r((\alpha^n)\cdot(a_n)) = (\alpha^n)\cdot(\alpha\nabla)^r(a_n)$. Por último, si $p(x) = \sum_i c_i x^i$, entonces

$$P(\nabla)((\alpha^n)\cdot(a_n)) = \sum_i c_i\nabla^i((\alpha^n)\cdot(a_n)) = \sum_i c_i\cdot(\alpha^n)\cdot(\alpha\nabla)^i(a_n) = (\alpha^n)\cdot P(\alpha\nabla)(a_n)$$

□

8. Lema: *Sea $T\colon E\to E'$ una aplicación lineal y $V\subseteq E'$ un subespacio vectorial. Entonces,* $\dim T^{-1}(V)\leq \dim V + \dim \operatorname{Ker} T$.

Demostración. Consideremos el morfismo $T_{|T^{-1}(V)}\colon T^{-1}(V)\to V$, $T_{|T^{-1}(V)}(e) = T(e)$. Observemos que $\operatorname{Ker} T_{|T^{-1}(V)} = T^{-1}(V)\cap \operatorname{Ker} T = \operatorname{Ker} T$, y $\operatorname{Im} T_{|T^{-1}(V)} = T(T^{-1}(V))\subseteq V$. Luego,

$$\dim V + \dim\operatorname{Ker} T \geq \dim\operatorname{Im} T_{|T^{-1}(V)} + \dim\operatorname{Ker} T_{|T^{-1}(V)} = \dim T^{-1}(V)$$

□

9. Teorema: *Se cumple que*

1. *Las sucesiones* $\{(1),(n),\dots,(n^r)\}$ *son una base de* $\operatorname{Ker}\Delta^{r+1}$.
2. $\operatorname{Ker}(\nabla-\alpha)^r = (\alpha^n)\cdot\{$*(polinomios $q(n)$ de grado menor que r)*$\}$.
3. *Si $p(x) = (x-\alpha_1)^{r_1}\cdots(x-\alpha_s)^{r_s}$, entonces*

   $\operatorname{Ker} p(\nabla) = ({\alpha_1}^n)\cdot\{$*(Pol. $q(n)$ de grado menor r_1)*$\}\oplus\overset{r}{\cdots}\oplus(\alpha_s^n)\cdot\{$*Pol. $q(n)$ de grado menor r_s*$\}$

Demostración. 1. $\Delta((n^s)) = ((n+1)^s - n^s)$ que es un polinomio en en n de grado menor que s. Por tanto, si $q(n)$ es un polinomio en n de grado s entonces $\Delta(q(n))$ es un polinomio de grado menor que s. Por tanto, $\langle(1),(n),\dots,(n^r)\rangle\subseteq\operatorname{Ker}\Delta^{r+1}$.

$\operatorname{Ker}\Delta$ son las sucesiones de término general constante, es decir, $\operatorname{Ker}\Delta = \langle(1)\rangle$ y $\dim\operatorname{Ker}\Delta = 1$. Como $\operatorname{Ker}\Delta^{r+1} = \Delta^{-1}(\operatorname{Ker}\Delta^r)$, entonces

$$\dim\operatorname{Ker}\Delta^r \leq \operatorname{Ker}\Delta^r + 1 \leq \cdots \leq r+1$$

Por dimensiones, la inclusión $\langle(1),(n),\dots,(n^r)\rangle\subseteq\operatorname{Ker}\Delta^{r+1}$ es una igualdad.

2. $(s(n))\in\operatorname{Ker}(\nabla-\alpha)^r \iff 0 = (\nabla-\alpha)^r(s(n)) \iff 0 = (\nabla-\alpha)^r((\alpha^n)\cdot(\alpha^{-n})\cdot(s(n))) = (\alpha^n)\cdot(\alpha\Delta)^r((\alpha^{-n})\cdot(s(n))) \iff 0 = \Delta^r((\alpha^{-n})\cdot(s(n))) \iff$ existe un polinomio $q(n)$ de grado menor que r tal que $\alpha^{-n}\cdot s(n) = q(n)$, es decir, $s(n) = \alpha^n\cdot q(n)$.

3. Es consecuencia de que $\operatorname{Ker} p(\nabla) = \operatorname{Ker}(\nabla-\alpha_1)^{n_1}\oplus\cdots\oplus\operatorname{Ker}(\nabla-\alpha_r)^{n_r}$ y de 2.

□

10. Ejercicio: Resolver la ecuación $a_{n+2} = a_{n+1} + a_n$, con las condiciones iniciales $a_0 = 0, a_1 = 1$ (sucesión de Fibonacci).

11. Ejercicio: Calcular cuántos números de longitud n se pueden escribir con ceros y unos, de modo que nunca aparezcan dos ceros seguidos (ejemplo: los números de longitud tres cumpliendo lo dicho son $010, 011, 101, 110, 111$, que son cinco distintos).

Consideremos una ecuación en diferencias $p(\nabla)(s(n)) = z(n)$. Sea $(s_0(n))$ una solución particular. Entonces, $(s(n))$ es una solución de la ecuación en diferencias si y sólo si

$$s(n) = s_0(n) + t(n), \text{ con } t(n) \in \operatorname{Ker} p(\nabla)$$

Método para resolver una ecuación en diferencias.

1. Queremos resolver la ecuación $p(\nabla)y = z(n)$, con $z(n)$, cumpiendo que existe un polinomio $q(x)$ primo con $p(x)$ de modo que $q(\nabla)(z(n)) = 0$:

Sean $\lambda(x)$ y $\mu(x)$ tales que $\lambda(x)\cdot p(x)+\mu(x)\cdot q(x) = 1$. Por tanto, $\lambda(\nabla)\cdot p(\nabla)+\mu(\nabla)\cdot q(\nabla) = \mathrm{Id}$. Si aplicamos esta igualdad a $(z(n))$, obtenemos

$$(z(n)) = \lambda(\nabla)\cdot p(\nabla)(z(n)) = p(\nabla)(\lambda(\nabla)(z(n)))$$

Por tanto, una solución particular es $s_0(n) = \lambda(\nabla)(z(n))$.

12. Ejercicio: Resolver $a_{n+2} + 2a_{n+1} - 6a_n = 2^n$.

13. Observación: 1. Dada una sucesión de números complejos $(s(n))$, si definimos $t(n) = \sum_{i=0}^{n-1} s(i)$, entonces $\Delta((t(n)) = (t(n+1)-t(n)) = (s(n))$. Se denota $\frac{1}{\Delta}(s(n)) = (\sum_{i=0}^{n-1} s(i))$.

2. Observemos que $\langle (1), (n), \dots, (n^r)\rangle = \langle (\binom{n}{0}), (\binom{n}{1}), \dots, (\binom{n}{r})\rangle$. Además, $\Delta(\binom{n}{r}) = (\binom{n}{r-1})$. Por tanto, si escribimos $p(n) = \sum_{i=0}^r \lambda_i \binom{n}{i}$, tendremos que

$$\lambda_i = \Delta^i(p(n))_{|n=0}$$

3. $\frac{1}{\Delta}(\binom{n}{r}) = \binom{n}{r+1}$.

14. Ejercicio: Calcular $\sum_{i=0}^n (i^2 + i - 3)$.

2.3. Cambio de base en endomorfismos lineales. Determinante

1. Definición: Una aplicación lineal de un espacio vectorial en sí mismo se denomina endomorfismo lineal.

Si B es una base de E y $T: E \to E$ un endomorfismo, diremos que la matriz de T en las bases B, B es la matriz de T en la base B, por brevedad.

2. Proposición: *Sea $T\colon E \to E$ un endomorfismo lineal. Sean $B_1 = \{e_1,\dots,e_n\}$ y $B_2 = \{v_1,\dots,v_n\}$ dos bases de E y supongamos que $e_i = \sum_j a_{ji} v_j$. Si (λ_{ij}) es la matriz de T en la base B_1, entonces la matriz (μ_{ij}) de T en la base B_2, es igual a*

$$(\mu_{ij}) = (a_{ij})\cdot(\lambda_{ij})\cdot(a_{ij})^{-1}$$

3. Definición: Diremos que dos matrices $A, A' \in M_n(\mathbb{R})$ cuadradas son semejantes si existe una matriz cuadrada invertible $C \in M_n(\mathbb{R})$, tal que

$$A' = C \cdot A \cdot C^{-1}$$

4. Proposición: *Dos matrices cuadradas $A, A' \in M_n(\mathbb{R})$ son semejantes si son las matrices asociadas a un mismo endomorfismo $T\colon \mathbb{R}^n \to \mathbb{R}^n$ en ciertas bases B y B'.*

5. Proposición: *Sea $T\colon E \to E$ un endomorfismo lineal. Sean B, B' dos bases de E. Sea A la matriz asociada a T en la base B y A' la matriz asociada a T en la base B'. Entonces,*

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

(se dice que $\det(T) := \det(A)$ *es el determinante del endomorfismo lineal T).*

Demostración. Sea C la matriz de cambio de base de B a B'. Por el corolario 2.3.2

$$A = C^{-1} \cdot A' \cdot C$$

Luego, $\det(A) = \det(C^{-1} \cdot A' \cdot C) = \det(C^{-1}) \cdot \det(A') \cdot \det(C) = \det(A')$.

□

2.4. Autovectores y autovalores. Diagonalización

1. Definición: Sea $T\colon E \to E$ un endomorfismo lineal. Diremos que $0 \neq e \in E$ es un autovector de T si existe $\lambda \in \mathbb{R}$ tal que $T(e) = \lambda \cdot e$, en este caso diremos que λ es un autovalor.

2. Proposición: *Sea $T\colon E \to E$ un endomorfismo lineal. Entonces, $\lambda \in \mathbb{R}$ es un autovalor de T si y sólo si* $\operatorname{Ker}(T - \lambda \cdot \mathrm{Id}) \neq 0$, *que equivale a decir que* $\det(T - \lambda \cdot \mathrm{Id}) = 0$.

Demostración. En efecto, $\lambda \in \mathbb{R}$ es un autovalor de T si y sólo si existe un vector $0 \neq e \in E$ tal que $T(e) = \lambda \cdot e$, es decir, si y sólo si existe un vector $0 \neq e \in E$ tal que $(T - \lambda \cdot \mathrm{Id})(e) = 0$, es decir, $\operatorname{Ker}(T - \lambda \cdot \mathrm{Id}) \neq 0$. □

Notación: Diremos que $\lambda \in \mathbb{C}$ es un autovalor imaginario de T si $\det(T - \lambda \cdot \mathrm{Id}) = 0$.

3. Definición: Sea $T\colon E \to E$ un endomorfismo lineal. Llamaremos polinomio característico de T, que denotaremos $\chi_T(x)$, al polinomio definido por

$$\chi_T(x) := \det(x \cdot \mathrm{Id} - T)$$

Si $A = (a_{ij})$ es la matriz de T en una cierta base, entonces

$$\chi_T(x) = \begin{vmatrix} x - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & x - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & x - a_{nn} \end{vmatrix} =: \chi_A(x)$$

que es un polinomio mónico de grado $n = \dim E$. Si A' es una matriz semejante a A, es decir, es la matriz asociada a T en otra base, entonces $\chi_{A'}(x) = \chi_T(x) = \chi_A(x)$.

4. Proposición: *Sea $T\colon E \to E$ un endomorfismo lineal. Entonces, $\lambda \in \mathbb{R}$ es un autovalor de T si y sólo si λ es una raíz del polinomio característico de T.*

5. Definición: Se dice que un endomorfismo $T\colon E \to E$, es diagonalizable si existe una base de E, $\{e_1, \dots, e_n\}$ tal que $T(e_i) = \lambda_i \cdot e_i$, para todo i, es decir, si matriz asociada a T en la base es

$$T \equiv \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}$$

6. Proposición: *Un endomorfismo $\mathbb{R}$-lineal $T\colon E \to E$ es diagonalizable si y sólo si cumple*

1. *Todas las raíces de $\chi_T(x)$ son reales.*

2. *Si escribimos $\chi_T(x) = (x - \alpha_1)^{n_1} \cdots (x - \alpha_r)^{n_r}$, con $\alpha_i \neq \alpha_j \in \mathbb{R}$, entonces*

$$\dim \mathrm{Ker}(T - \alpha_i \cdot \mathrm{Id}) = n_i, \text{ para todo } i.$$

.

Demostración. Si T es diagonolizable existe una base en la que la matriz asociada a T es diagonal, reordenando los vectores de esta base podemos suponer que

$$T \equiv \begin{pmatrix} \alpha_1 & & & & & & \\ & \ddots^{n_1} & & & & & \\ & & \alpha_1 & & & & \\ & & & \ddots & & & \\ & & & & \alpha_r & & \\ & & & & & \ddots^{n_1} & \\ & & & & & & \alpha_r \end{pmatrix}$$

con $\alpha_i \neq \alpha_j \in \mathbb{R}$. En este caso, $\chi_T(x) = (x-\alpha_1)^{n_1}\cdots(x-\alpha_r)^{n_r}$, con $\alpha_i \neq \alpha_j \in \mathbb{R}$ y $\dim \operatorname{Ker}(T-\alpha_i \cdot \mathrm{Id}) = n_i$, para todo i.

Recíprocamente, supongamos $\chi_T(x) = (x-\alpha_1)^{n_1}\cdots(x-\alpha_r)^{n_r}$, con $\alpha_i \neq \alpha_j \in \mathbb{R}$ y $\dim \operatorname{Ker}(T-\alpha_i \cdot \mathrm{Id}) = n_i$, para todo i.. Entonces, $\sum_i n_i = \mathrm{gr}(\chi_T(x)) = \dim E$. Además, los $\operatorname{Ker}((T-\alpha_i \mathrm{Id})$ están en suma directa, pues

$$\operatorname{Ker}((T-\alpha_1 \mathrm{Id})\cdots(T-\alpha_r \mathrm{Id})) = \operatorname{Ker}(T-\alpha_1 \mathrm{Id}) \oplus \cdots \oplus \operatorname{Ker}(T-\alpha_r \mathrm{Id})$$

Por dimensiones, $E = \operatorname{Ker}(T-\alpha_1 \mathrm{Id}) \oplus \cdots \oplus \operatorname{Ker}(T-\alpha_r \mathrm{Id})$. Si consideramos una base en cada $\operatorname{Ker}((T-\alpha_i \mathrm{Id})$, obtendremos una base en la que T diagonaliza.

□

7. Proposición: *Un endomorfismo es diagonalizable si y sólo si todas las raíces del polinomio anulador son reales y de multiplicidad* 1.

2.5. Teorema de Hamilton-Cayley

1. Definición: Diremos que un polinomio $p(x)$ anula a un endomorfismo $T: E \to E$, si $p(T) = 0$.

2. Teorema de Hamilton-Cayley: *Sea $T: E \to E$ un endomorfismo. Entonces, el polinomio característico de T anula a T.*

Demostración. Si $A \in M_n(\mathbb{R})$ es una matriz cuadrada, entonces $\chi_A(x) = \det(x \cdot \mathrm{Id} - A)$ y tenemos que probar que $\chi_A(A) = 0$. Podemos pensar que A es la matriz del correspondiente endomorfismo $\mathbb{R}$-lineal $\mathbb{R}^n \to \mathbb{R}^n$ en la base estándar, o igualmente, que es la matriz del correspondiente endomorfismo $\mathbb{C}$-lineal $\mathbb{C}^n \to \mathbb{C}^n$ en la base estándar, obviamente $\chi_A(x)$, ni $\chi_A(A)$ dependen de ello. En conclusión, podemos suponer que E es un $\mathbb{C}$-espacio vectorial.

Vamos a proceder por inducción sobre $\dim E$. Si $\dim E = 1$, sea $\{e_1\}$ una base de E y tendremos que $T(e) = \lambda \cdot e$, $\chi_T(x) = (x-\lambda)$ y $\chi_T(T) = 0$. Suponemos el teorema cierto para $\dim E < n$ y $n > 1$.

Sea $\lambda \in \mathbb{C}$ una raíz del polinomio característico de T. Sea pues $e_1 \in E$ un vector propio de valor propio λ y $\{e_1, \ldots, e_n\}$ una base de E.

Sea

$$T = \begin{pmatrix} \lambda & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

la matriz asociada a T. Sea $S = \begin{pmatrix} a_{22} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{n2} & \cdots & a_{nn} \end{pmatrix}$, entonces $T = \left(\begin{array}{c|c} \lambda & C \\ \hline 0 & S \end{array}\right)$. Dado un polinomio $q(x)$ se tiene que

$$q(T) = \left(\begin{array}{c|c} q(\lambda) & D \\ \hline 0 & q(S) \end{array}\right)$$

Observemos que $\chi_T(x) = (x-\lambda)\cdot\chi_S(x)$. Por tanto,

$$\chi_T(T) = (T-\lambda)\cdot\chi_S(T) = (T-\lambda)\cdot\left(\begin{array}{c|c} \chi_S(\lambda) & D \\ \hline 0 & \chi_S(S) \end{array}\right)$$
$$= \left(\begin{array}{c|c} 0 & D' \\ \hline 0 & D'' \end{array}\right)\cdot\left(\begin{array}{c|c} \chi_S(\lambda) & D \\ \hline 0 & 0 \end{array}\right) = 0$$

□

3. Definición: Sea $T\colon E \to E$ un endomorfismo lineal. El conjunto de los polinomios que anulan a T es un ideal de $\mathbb{R}[x]$. Llamaremos polinomio anulador de T, que denotaremos $p_{an(T)}(x)$, al polinomio mónico de grado más pequeño que anula a T.

Todo polinomio que anule a T es múltiplo del polinomio anulador de T.

El teorema de Hamilton-Cayley nos dice que el polinomio característico es múltiplo del anulador.

4. Proposición: *Sea $T\colon E \to E$ un endomorfismo lineal. Sea $\chi_T(x) = p_1(x)^{n_1}\cdots p_r(x)^{n_r}$ la descomposición de $\chi_T(x)$ en producto de polinomios irreducibles (con $p_i(x)$ primo con $p_j(x)$, para todo $i \neq j$). Sea $p_{an(T)}(x)$ el polinomio mínimo anulador de T. Entonces,*

1. $E = \operatorname{Ker} p_1(T)^{n_1} \oplus \cdots \oplus \operatorname{Ker} p_r(T)^{n_r}$.
2. $\dim \operatorname{Ker} p_i(T)^{n_i} = \operatorname{gr}(p_i(x)^{n_i}) = n_i \cdot \operatorname{gr}(p_i(x))$.
3. $p_{an(T)}(x) = p_1(x)^{m_1}\cdots p_r(x)^{m_r}$, *con* $0 < m_i \leq n_i$, *para todo* i. *"El polinomio anulador tiene las mismas raíces que el característico, aunque con multiplicidad menor o igual".*
4. $\operatorname{Ker} p_i(x)^{m_i} = \operatorname{Ker} p_i(T)^{n_i}$, *para todo* i.

Demostración. Veamos que si $\operatorname{gr}(q(x)) > 0$ y $q(x)$ divide a $\chi_T(x)$, entonces $\operatorname{Ker} q(T) \neq 0$: Sea $\alpha \in \mathbb{C}$ una raíz de $q(x)$ y escribamos $q(x) = (x-\alpha)\cdot q_2(x)$. Observemos que α es una raíz del polinomio característico, luego $0 = \chi_T(\alpha) = \det(\alpha \operatorname{Id} - T)$. Entonces,

$$\det(q(T)) = \det(T - \alpha \operatorname{Id}) \cdot \det(q_2(T)) = 0$$

y $\operatorname{Ker} q(T) \neq 0$.

1. $\chi_T(T) = 0$, luego $E = \operatorname{Ker}\chi_T(T) = \operatorname{Ker} p_1(T)^{n_1} \oplus \cdots \oplus \operatorname{Ker} p_r(T)^{n_r}$.

2. Escribamos $E_i = \operatorname{Ker} p_i(T)^{n_i}$. Si $e \in E_i$, entonces $T(e) \in E_i$, porque

$$p_i(T)^{n_i}(T(e)) = T(p_i(T)^{n_i}(e)) = 0$$

Denotemos $T_i \colon E_i \to E_i$ a la restricción de T en E_i. Observemos que $p_i(T_i)^{n_i} = 0$, porque $p_i(T_i)^{n_i}(e) = p_i(T)^{n_i}(e) = 0$, para todo $e \in E_i$. Si $\chi_{T_i}(x) = p_i(x)^{m_i} \cdot q(x)$, con $q(x)$ primo con $p_i(x)$, como $E_i = \operatorname{Ker} p_i(T)^{n_i}$ está en suma directa con $\operatorname{Ker} q(T_i)$, entonces $\operatorname{Ker} q(T_i) = 0$ y $q(x)$ no puede aparecer en la descomposición de $\chi_{T_i}(x)$. Como $\chi_T(x) = \chi_{T_1}(x) \cdots \chi_{T_r}(x)$, tenemos que $\chi_{T_i}(x) = p_i(x)^{n_i}$, para todo i. Además, $\dim E_i = \operatorname{gr}(\chi_{T_i}(x)) = \operatorname{gr}(p_i(x)^{n_i})$.

3. y 4. $\chi_T(x)$ es múltiplo de $p_{an(T)}$, luego $m_i \leq n_i$. Luego, $\operatorname{Ker} p_i(T)^{m_i} \subseteq \operatorname{Ker} p_i(T)^{n_i}$. Como $p_{an(T)}(T) = 0$, entonces $E = \operatorname{Ker} p_1(T)^{m_1} \oplus \cdots \oplus \operatorname{Ker} p_r(T)^{m_r}$, luego $\operatorname{Ker} p_i(T)^{m_i} = \operatorname{Ker} p_i(T)^{n_i}$ y $m_i > 0$.

□

2.6. Base de Jordan

En esta sección $T \colon E \to E$ es un endomorfismo lineal cuyo polinomio característico es $\chi_T(x) = (x-\alpha_1)^{n_1} \cdots (x-\alpha_r)^{n_r}$, con α_i en el cuerpo considerado.

Sabemos que $E = \operatorname{Ker}(T-\alpha_1)^{n_1} \oplus \cdots \operatorname{Ker}(T-\alpha_r)^{n_r}$, que $T = T_1 \oplus \cdots \oplus T_r$, donde T_i es la restricción de T a $\operatorname{Ker}(T-\alpha_i)^{n_i}$ y $\chi_{T_i}(x) = (x-\alpha_1)^{n_1}$.

Queremos dar una base de cada $E_i = \operatorname{Ker}(T-\alpha_i)^{n_i}$ en la que la matriz de T_i sea sencilla. Así obtendremos una base de E en la que la matriz de T es sencilla, base que se llamará base de Jordan.

1. Definición: Sea $E_1 \subseteq E_2$ un subespacio vectorial. Diremos que $\{e_1, \dots, e_r\}$ es una base suplementaria de E_1 en E_2 si existe una base $\{e_{r+1}, \dots, e_n\}$ de E_1 de modo que $\{e_1, \dots, e_n\}$ es una base de E_2.

Si $\{e_1, \dots, e_r\}$ es una base suplementaria de E_1 en E_2 entonces

$$E_2 = \langle e_1, \dots, e_r \rangle \oplus \langle e_{r+1}, \dots, e_n \rangle = \langle e_1, \dots, e_r \rangle \oplus E_1$$

y toda base de E_1, junto con $\{e_1, \dots, e_r\}$ es una base de E_2.

Si $\{v_1, \dots, v_s\}$ es una base de E_1 y la completamos a una base $\{v_1, \dots, v_n\}$ de E_2, entonces $\{v_{s+1}, \dots, v_n\}$ es una base suplementaria de E_1 en E_2.

2. Proposición: *Sea $0 = E_0 \subseteq E_1 \subseteq E_2 \subseteq \cdots \subseteq E_n$ inclusiones de espacios vectoriales. Si para cada i, $\{e_{ij}\}_j$ es una base suplementaria de E_i en E_{i+1}, entonces $\{e_{ij}\}_{ij}$ es una base de E_n.*

3. Proposición : *Sea $S\colon E \to E$ un endomorfismo lineal. Sea $r \geq 2$. Si $e_1,\dots,e_s$ es una base suplementaria de* $\operatorname{Ker} S^{r-1}$ *en* $\operatorname{Ker} S^r$*, entonces existen vectores $e'_1,\dots,e'_{s'} \in$* $\operatorname{Ker} S^{r-1}$ *tales que $\{S(e_1),\dots,S(e_s),e'_1,\dots,e'_{s'}\}$ es una base suplementaria de* $\operatorname{Ker} S^{r-2}$ *en* $\operatorname{Ker} S^{r-1}$.

Demostración. 1. $\{S(e_1),\dots,S(e_s)\}$ son linealmente independientes: Si $0 = \sum_i \lambda_i S(e_i) = S(\sum_i \lambda_i e_i)$, entonces $\sum_i \lambda_i e_i \in \operatorname{Ker} S \subseteq \operatorname{Ker} S^{r-1}$, luego $\sum_i \lambda_i e_i = 0$ y $\lambda_i = 0$, para todo i.

2. $\langle S(e_1),\dots,S(e_s)\rangle \cap \operatorname{Ker} S^{r-2} = 0$: Si $\sum_i \lambda_i S(e_i) \in \operatorname{Ker} S^{r-2}$, entonces $S(\sum_i \lambda_i e_i) \in \operatorname{Ker} S^{r-2}$ y $\sum_i \lambda_i e_i \in \operatorname{Ker} S^{r-1}$, luego $\lambda_i = 0$, para todo i.

3. Tenemos que $\langle S(e_1),\dots,S(e_s)\rangle + \operatorname{Ker} S^{r-2} \subset \operatorname{Ker} S^{r-1}$. Sea $v_1,\dots,v_m$ una base de $\operatorname{Ker} S^{r-2}$. Entonces, $S(e_1),\dots,S(e_s),v_1,\dots,v_m$ es una base de

$$\langle S(e_1),\dots,S(e_s)\rangle + \operatorname{Ker} S^{r-2}$$

Sean $e'_1,\dots,e'_{s'} \in \operatorname{Ker} S^{r-1}$ tales que $S(e_1),\dots,S(e_s),v_1,\dots,v_m,e'_1,\dots,e'_{s'}$ sean una base de $\operatorname{Ker} S^{r-1}$. Entonces, $\{S(e_1),\dots,S(e_s),e'_1,\dots,e'_{s'}\}$ es una base suplementaria de $\operatorname{Ker} S^{r-2}$ en $\operatorname{Ker} S^{r-1}$.

□

Sea $S\colon E \to E$ un endomorfismo tal que $S^n = 0$. Consideremos la cadena de subespacios vectoriales $0 = \operatorname{Ker} S^0 \subseteq \operatorname{Ker} S^1 \subseteq \cdots \subseteq \operatorname{Ker} S^n = E$. Sea

$$\{e_{11},\dots,e_{1r_1}\}$$

una base suplementaria de $\operatorname{Ker} S^{n-1}$ en $\operatorname{Ker} S^n$. Sea

$$\{e_{21} := S(e_{11}),\dots,e_{2r_1} = S(e_{1r_1}),e_{2,r_1+1},\dots,e_{2,r_1+r_2}\}$$

una base suplementaria de $\operatorname{Ker} S^{n-2}$ en $\operatorname{Ker} S^{n-1}$. Sea

$$\{e_{31} := S(e_{21}),\dots,e_{3,r_1+r_2} = S(e_{2,r_1+r_2}),\dots,e_{3,r_1+r_2+r_3}\}$$

una base suplementaria de $\operatorname{Ker} S^{n-3}$ en $\operatorname{Ker} S^{n-2}$. Procediendo así sucesivamente, por la proposición 2.6.2, $\{e_{ij}\}$ es una base de E. Ordenemosla por columnas como sigue

$\operatorname{Ker} S^n \downarrow$	e_{11}	$\cdots$	e_{1r_1}							
$\operatorname{Ker} S^{n-1} \downarrow$	e_{21}	$\cdots$	e_{2r_1}	e_{2,r_1+1}	$\cdots$	e_{2,r_1+r_2}				
$\cup$	$\vdots$		$\vdots$	$\vdots$		$\vdots$	$\cdots$			
$\operatorname{Ker} S$	e_{n1}	$\cdots$	e_{nr_1}	e_{n,r_1+1}	$\cdots$	e_{n,r_1+r_2}	$\cdots$	$e_{n,r_1+\cdots+r_{n-1}+1}$	$\cdots$	$e_{n,r_1+\cdots+r_n}$

Observemos que cada columna j (con $0 < j - r_1 - \cdots - r_{i-1} \leq r_i$) $\{e_{i,j}, e_{i+1,j}, \ldots, e_{n,j}\}$ es estable por S y la matriz de S sobre los vectores de esta columna es

$$B_{ii} = \begin{array}{c} n-i+1 \\ \begin{pmatrix} 0 & \cdots & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 \end{pmatrix} \end{array}$$

En la base $\{e_{ij}\}$, ordenada como sigue: $e_{ij} < e_{i'j'}$ si $j < j'$ ó $j = j'$ y $i < i'$, la matriz de S es

$$S \equiv \left(\begin{array}{c|c|c} B_1 & 0 & 0 \\ \hline 0 & \ddots & 0 \\ \hline 0 & 0 & B_n \end{array}\right), \quad B_i = \left(\begin{array}{c|c|c} B_{ii} & 0 & 0 \\ \hline 0 & \ddots^{\,r_i} & 0 \\ \hline 0 & 0 & B_{ii} \end{array}\right), \quad B_{ii} = \begin{array}{c} n-i+1 \\ \begin{pmatrix} 0 & \cdots & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 \end{pmatrix} \end{array}$$

Observemos que los números r_i (hay r_i columnas de altura $n - i + 1$), determinan S. Si denotamos $s_j := \dim \operatorname{Ker} S^{n-j}$, para $0 < j \leq n$, el lector puede comprobar que $s_{j-1} - s_j = r_1 + \cdots + r_j$, luego $r_j = (s_{j-1} - s_j) - (s_{j-2} - s_{j-1}) = 2 \cdot s_{j-1} - (s_j + s_{j-2})$.

Por último, si $T: E \to E$ es un endomorfismo tal que $(T - \lambda \cdot \mathrm{Id})^n = 0$, entonces $S = T - \lambda \cdot \mathrm{Id}$ cumple que $S^n = 0$ y $T = S + \lambda \cdot \mathrm{Id}$. En la base escogida para S, la matriz de T es

$$T \equiv \left(\begin{array}{c|c|c} B_1 & 0 & 0 \\ \hline 0 & \ddots & 0 \\ \hline 0 & 0 & B_n \end{array}\right), \quad B_i = \left(\begin{array}{c|c|c} B_{ii} & 0 & 0 \\ \hline 0 & \ddots^{\,r_i} & 0 \\ \hline 0 & 0 & B_{ii} \end{array}\right), \quad B_{ii} = \begin{array}{c} n-i+1 \\ \begin{pmatrix} \lambda & \cdots & \cdots & 0 \\ 1 & \lambda & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 1 & \lambda \end{pmatrix} \end{array}$$

4. Teorema: *Dos matrices cuadradas $A, A' \in M_n(\mathbb{C})$ son semejantes si y sólo si*

1. $\chi_A(x) = \chi_{A'}(x)$.
2. *Si escribimos $\chi_A(x) = (x - \alpha_1)^{n_1} \cdots (x - \alpha_r)^{n_r}$, con $\alpha_i \neq \alpha_j$, entonces se cumple que*

$$\operatorname{rg}((A - \alpha_i \cdot \mathrm{Id})^j) = \operatorname{rg}((A' - \alpha_i \cdot \mathrm{Id})^j)$$

para todo i y todo $j \leq n_i$.

2.7. Problemas

1. Un banco nos presta un capital K, a devolver en N años, a un tipo de interés anual I. ¿Cuánto dinero D deberemos pagar al año, de modo que todos los años paguemos la misma cantidad y en los N años hayamos saldado nuestra deuda con el banco?

   Resolución: Sea i_n el dinero que pagamos en el año n por los intereses del capital que tenemos prestado durante el año n y a_n el dinero que amortizamos en el año n por el capital prestado. Entonces $D = a_n + i_n$. Además, $i_n = I\cdot(K - \sum_{r=1}^{n-1} a_r)$. Por tanto, $D = a_n + I\cdot(K-\sum_{r=1}^{n-1} a_r)$. Si aplicamos el operador diferencia Δ entonces $0 = \Delta(a_n) - I\cdot a_n = (\nabla - (1+I))(a_n) = 0$. Por tanto, $a_n = (1+I)^n \cdot \lambda$. Si $n = 1$, entonces

   $$D = a_1 + IK = (1+I)\lambda + IK$$

   Tenemos que calcular λ. Nos falta decir que amortizamos la hipoteca en N años, es decir

   $$K = \sum_{r=1}^{N} a_r$$

   Calculemos $b_n = \sum_{r=1}^{n-1} a_r$. Tenemos que $(\nabla - 1)(b_n) = \Delta(b_n) = a_n$ y $\nabla - (1+I)$ anula a (a_n). Como $\frac{\nabla-1}{I} - \frac{\nabla-(1+I)}{I} = 1$ tenemos

   $$(\nabla - 1)(b_n) = (a_n) = (\frac{\nabla - 1}{I} - \frac{\nabla-(1+I)}{I})(a_n) = (\frac{\nabla-1}{I})(a_n)$$

   Luego $b_n = \frac{a_n}{I} + \mu$. Como $b_2 = a_1$ tenemos que $a_1 = \frac{a_2}{I} + \mu$, es decir, $(1+I)\lambda = \frac{(1+I)^2\lambda}{I} + \mu$ y $\mu = \frac{1+I}{-I}\cdot\lambda$. Luego $K = b_{N+1} = \frac{a_{N+1}}{I} + \mu = \frac{(1+I)^{N+1}\lambda}{I} - \frac{(1+I)\lambda}{I}$ y despejando obtenemos $\lambda = \frac{IK}{(1+I)^{N+1}-(1+I)}$. Por tanto,

   $$D = (1+I)\lambda + IK = \cdots = \frac{IK}{1 - \frac{1}{(1+I)^N}}$$

2. Un préstamo de $K = 10^5$ euros se quiere devolver durante $N = 20$ años, pagando cada año n una anualidad d_n de modo que $d_n = d_{n-1} + 10^3$. Se suponen que nos prestan el dinero a un tipo de interés anual del $I = 5\,\%$. Determinar d_1.

   Resolución: Sea i_n es el dinero que pagamos en el año n como pago de los intereses del capital que tenemos prestado durante el año n y a_n el dinero que amortizamos en el año n. Entonces, $d_n = i_n + a_n$. Tenemos que $i_n = I\cdot(K - \sum_{r=1}^{n-1} a_r)$. Por tanto,

   $$d_n = a_n + I\cdot(K - \sum_{r=1}^{n-1} a_r).$$

Si aplicamos el operador diferencia Δ entonces $10^3 = \Delta(a_n) - I \cdot a_n = (\nabla - (1 + I))(a_n) = 0$. Por tanto, $a_n = (1+I)^n \cdot \lambda + \frac{-10^3}{I}$. Si $n = 1$, entonces

$$d_1 = a_1 + IK = (1+I)\lambda + \frac{-10^3}{I} + IK$$

Tenemos que calcular λ. Nos falta decir que amortizamos la hipoteca en N años, es decir,

$$K = \sum_{r=1}^{N} a_r$$

Tenemos que $b_n = \sum_{r=1}^{n-1} a_r = \frac{(1+I)^n\lambda}{I} - \frac{10^3\cdot(n-1)}{I} + \mu$. Como $b_2 = a_1$ tenemos que $\frac{(1+I)^2\lambda}{I} - \frac{10^3}{I} + \mu = (1+I)\cdot\lambda - \frac{10^3}{I}$ y $\mu = \frac{-\lambda(1+I)}{I}$. Luego $K = b_{N+1} = \frac{(1+I)^{N+1}\lambda}{I} - \frac{10^3\cdot N}{I} + \frac{-\lambda(1+I)}{I}$ y despejando obtenemos que $\lambda = \frac{IK+10^3N}{(1+I)^{N+1}-(1+I)}$ Por tanto,

$$d_1 = (1+I)\lambda + \frac{-10^3}{I} + IK = \frac{IK + \frac{10^3N}{(1+I)^N}}{1 - \frac{1}{(1+I)^N}} - \frac{10^3}{I}$$

3. Dado el endomorfismo $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ definido por $T(x, y) = (x + y, x - y)$, obtener su matriz respecto de la base usual de $\mathbb{R}^2$. Obtener también las matrices de los endomorfismos $T^2 - \mathrm{Id}_{\mathbb{R}^2}$ y $T^3 = T \circ T \circ T$.

4. Sea V un espacio vectorial de dimensión 2 y sea T un endomorfismo de V no nulo y **nilpotente** (se dice que un endomorfismo es nilpotente si existe un número natural $p > 1$ tal que $T^p = 0$, donde T^p es $T \circ \cdots \circ T$ p veces). Probar que existe una base de V respecto de la cual la matriz asociada a T es $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Aplicar lo anterior al endomorfismo del $\mathbb{C}$-espacio vectorial $\mathbb{C}^2$ cuya matriz asociada respecto cierta base es $\begin{pmatrix} \mathrm{i} & 1 \\ 1 & -\mathrm{i} \end{pmatrix}$.

5. Sean $A_1, \ldots, A_r$ matrices tales que $A_i \in \mathscr{M}_{m_i}(\mathbb{R})$, $i = 1, \ldots, r$. Probar que si los autovalores de A_i son $\lambda_{i,1}, \ldots, \lambda_{i,s_i}$, $i = 1, \ldots, r$, entonces los autovalores de $A_1 \oplus \ldots \oplus A_r$ son $\{\lambda_{ij} \mid i = 1, \ldots, r;\ j = 1, \ldots, s_i\}$.

6. Sea $\chi_T(x) = a_0 + a_1x + \ldots + a_{n-1}x^{n-1} + x^n$ el polinomio característico de un endomorfismo T de un $\mathbb{R}$ -espacio vectorial V de dimensión finita $n > 0$. Probar que el determinante de T es igual a $(-1)^n a_0$.

7. Sea V un $\mathbb{R}$ -espacio vectorial de dimensión finita $n > 0$ y $T \in \mathrm{End}_{\mathbb{R}}(V)$ tal que la suma de las entradas de cada una de las filas de su matriz asociada $A \in \mathcal{M}_n(\mathbb{R})$ respecto de alguna base de V es igual 1 (es decir, A es una **matriz estocástica**). Probar que 1 es un autovalor de T.

8. Sean V un $\mathbb{R}$ -espacio vectorial de dimensión finita y $T \in \mathrm{End}_{\mathbb{R}}(V)$ biyectivo, es decir, T es un automorfismo de V. Probar que λ es un autovalor de T si y sólo si $\lambda \neq 0$ y λ^{-1} es autovalor de T^{-1}.

9. Comprobar que si $\{\lambda_1, \dots, \lambda_r\}$ son los autovalores de una matriz A, entonces

    1. Los autovalores de αA (siendo $\alpha \neq 0$) son $\{\alpha\lambda_1, \dots, \alpha\lambda_r\}$. Un vector $\mathbf{v}$ es autovector de A asociado a λ_i si, y sólo si $\mathbf{v}$ es autovector de αA asociado a $\alpha\lambda_i$.
    2. A es invertible si, y solo si, $0 \notin \{\lambda_1, \dots, \lambda_r\}$ y en este caso, los autovalores de A^{-1} son $\{\lambda_1^{-1}, \dots, \lambda_r^{-1}\}$. Un vector $\mathbf{v}$ es autovector de A asociado a λ_i si, y sólo si $\mathbf{v}$ es autovector de A^{-1} asociado a λ_i^{-1}.

10. Probar que si $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ son todos los autovalores (no necesariamente distintos) de una matriz $A \in \mathcal{M}_n(\mathbb{C})$, entonces

    1. $|A| = \lambda_1 \cdots \lambda_n$.
    2. $\mathrm{tr}(A) = \lambda_1 + \ldots + \lambda_n$.

11. Sean $V = \mathbb{R}^4$ y $T \in \mathrm{End}_{\mathbb{R}}(V)$ tal que su matriz asociada respecto de la base usual de $\mathbb{R}^4$ es
$$(a) \begin{pmatrix} -1 & -2 & 3 & 2 \\ 0 & 1 & 1 & 0 \\ -2 & -2 & 4 & 2 \\ 0 & 0 & 0 & 2 \end{pmatrix}, \qquad (b) \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 3 & 5 \end{pmatrix}.$$
Estudiar si T es diagonalizable.

12. Sean $V = \mathbb{R}^3$ y $T \in \mathrm{End}_{\mathbb{R}}(V)$ tal que su matriz asociada respecto de alguna base de V es
$$(a) \begin{pmatrix} a & -1 & 1 \\ 0 & 1 & 3 \\ 0 & 2 & 2 \end{pmatrix}, \qquad (b) \begin{pmatrix} 1 & a & b \\ 0 & 2 & c \\ 0 & 0 & 1 \end{pmatrix}, \qquad (c) \begin{pmatrix} 5 & 0 & 0 \\ 0 & -1 & b \\ 3 & 0 & a \end{pmatrix}$$
con a, b y $c \in \mathbb{R}$. Estudiar (según los valores de a, b y $c \in \mathbb{R}$) si T es diagonalizable. Resolver el mismo problema pero con $V = \mathbb{C}^3$.

13. Sean $V = \mathbb{R}^4$ y $T \in \operatorname{End}_{\mathbb{R}}(V)$ tal que su matriz asociada respecto de alguna base de V es
$$\begin{pmatrix} 1 & -1 & 0 & 0 \\ -4 & 1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 0 & a & 1 & 3 \end{pmatrix}.$$
Estudiar, según el valor de $a \in \mathbb{R}$, si T es diagonalizable, y calcular, cuando sea posible, una base de V respecto de cual la matriz de T sea diagonal.

14. Sean $V = \mathbb{R}^3$ y T y $T' \in \operatorname{End}_{\mathbb{R}}(V)$ tales que $T(v_1, v_2, v_3) = (v_1 + v_2 + v_3, 2v_1 + 5v_2 + 2v_3, -2v_1 - 5v_2 - 2v_3)$ y $T'(v_1, v_2, v_3) = (-2v_2 - 2v_3, 0, 2v_2 + 2v_3)$, para cada $\mathbf{v} = (v_1, v_2, v_3) \in \mathbb{R}^3$. Hallar, si es posible, sendas bases de V respecto de las cuales las matrices de T y T' sean diagonales.

15. Sea V un $\mathbb{R}$ -espacio vectorial de dimensión finita $n > 0$ y $T \in \operatorname{End}_{\mathbb{R}}(V)$ tal que $\mathrm{Id} + T^2 = 0$. Probar que T no tiene autovalores reales.

16. Sean T y T' dos endomorfismos de un $\mathbb{C}$-espacio vectorial V de dimensión finita. Probar que si T y T' conmutan, entonces T y T' tienen autovectores comunes.

17. Sea V un $\mathbb{R}$ -espacio vectorial de dimensión n y $T \in \operatorname{End}_{\mathbb{R}}(V)$ nilpotente. Probar que $\chi_T(x) = x^n$. Concluir que los autovalores de un endomorfismo nilpotente son todos nulos. ¿Es cierto el recíproco?

18. Sean V un espacio vectorial de dimensión finita sobre un cuerpo K y T un endomorfismo de V. Probar que

    1. Si $K = \mathbb{C}$ y V no tiene subespacios invariantes por T distintos del cero y el total, entonces la dimensión de V es 1.
    2. Si $K = \mathbb{R}$ y V no tiene subespacios invariantes por T distintos del cero y el total, entonces la dimensión de V es menor o igual que dos.

19. Sean T y S dos endomorfismos de un $\mathbb{R}$ -espacio vectorial V de dimensión finita. Probar:

    (a) Si T es diagonalizable, entonces para todo subespacio L de V que es invariante por T el endomorfismo $T|_L$ también es diagonalizable.

    (b) Los endomorfismos T y S son simultáneamente diagonalizables (esto es, existe una base de V formada por autovectores de los dos endomorfismos) si y sólo si T y S son diagonalizables y conmutan.

20. Clasificar los endomorfismos de un $\mathbb{C}$-espacio vectorial de dimensión 4.

21. Dadas las matrices

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

¿representan todas al mismo endomorfismo?

22. Calcular la forma canónica y la base de Jordan de los siguientes endomorfismos cuyas matrices respecto de la base canónica del correspondiente $\mathbb{C}$-espacio vectorial son:

$$\text{(a)} \begin{pmatrix} 3 & -2 & 0 \\ -2 & 3 & 0 \\ 0 & 0 & 5 \end{pmatrix}, \quad \text{(b)} \begin{pmatrix} -14 & 1 & 12 \\ -13 & 0 & 12 \\ -17 & 1 & 15 \end{pmatrix}, \quad \text{(c)} \begin{pmatrix} -1 & 2 & -1 \\ -2 & 3 & -2 \\ -2 & 2 & -1 \end{pmatrix}$$

$$\text{(d)} \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 3 & 5 \end{pmatrix}, \quad \text{(e)} \begin{pmatrix} 3 & 45 & 37 & -9 \\ 2 & 12 & 8 & -5 \\ -2 & -4 & -1 & 4 \\ 3 & 33 & 26 & -8 \end{pmatrix},$$

$$\text{(f)} \begin{pmatrix} 3 & 67 & 59 & -9 \\ 2 & -16 & -20 & -5 \\ -2 & 28 & 31 & 4 \\ 3 & 31 & 24 & -8 \end{pmatrix}, \quad \text{(g)} \begin{pmatrix} 3 & 17 & 9 & -9 \\ 2 & 16 & 12 & -5 \\ -2 & -12 & -9 & 4 \\ 3 & 17 & 10 & -8 \end{pmatrix},$$

$$\text{(h)} \begin{pmatrix} 3 & 45 & 37 & -9 \\ 2 & 10 & 6 & -5 \\ -2 & -2 & 1 & 4 \\ 3 & 32 & 25 & -8 \end{pmatrix}, \quad \text{(i)} \begin{pmatrix} 3 & 31 & 23 & -9 \\ 2 & 7 & 3 & -5 \\ -2 & -1 & 2 & 4 \\ 3 & 21 & 14 & -8 \end{pmatrix},$$

$$\text{(j)} \begin{pmatrix} 3 & 42 & 34 & -9 \\ 2 & -29 & -33 & -5 \\ -2 & 38 & 41 & 4 \\ 3 & 7 & 0 & -8 \end{pmatrix},$$

Capítulo 3

Potencias de matrices

3.1. Potencias de matrices cuadradas

Dado una matriz cuadrada A, queremos calcular A^n del modo más sencillo y rápido posible.

Si la matriz A fuese diagonal $A = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_r \end{pmatrix}$ es claro que $A^n = \begin{pmatrix} \lambda_1^n & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_r^n \end{pmatrix}$.

Si A es la matriz asociada a un endomorfismo diagonalizable, sabemos que A es semejante a una matriz diagonal, es decir, existe una matriz B (la matriz formada por los vectores en los que el endomorfismo diagonaliza) tal que $A = B \cdot D \cdot B^{-1}$, con $D = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_r \end{pmatrix}$ y por tanto

$$A^n = (B \cdot D \cdot B^{-1})^n = B \cdot D^n \cdot B^{-1} = B \cdot \begin{pmatrix} \lambda_1^n & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_r^n \end{pmatrix} \cdot B^{-1}$$

En general, $A \in M_r(\mathbb{C})$ no es diagonalizable. Si $B = \{v_1 = (b_{11}, \dots, b_{r1}), \dots, v_r = (b_{1r}, \dots, b_{rr})\}$ es una base de Jordan del endomorfismo asociado a A y J es la matriz asociada al endomorfismo en la base de Jordan, entonces $A = B \cdot J \cdot B^{-1}$ y

$$A^n = B \cdot J^n \cdot B^{-1}$$

El problema es calcular J^n, que se reduce al problema de calcular T^n cuando

$$T = \begin{array}{c} s \\ \begin{pmatrix} \lambda & \cdots & \cdots & 0 \\ 1 & \lambda & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 1 & \lambda \end{pmatrix} \end{array}$$

Ahora bien, $T = \lambda \cdot \mathrm{Id} + N$, con

$$N = \begin{array}{c} s \\ \begin{pmatrix} 0 & \cdots & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 \end{pmatrix} \end{array}$$

Como $\lambda \cdot \mathrm{Id}$ y N conmutan y $N^s = 0$ entonces, $T^n = (\lambda \cdot \mathrm{Id} + N)^n = \sum_{0 \leq i \leq n,s} \binom{n}{i} \lambda^{n-i} \cdot N^i$, es decir,

$$T^n = \begin{pmatrix} \lambda^n & 0 & 0 & \cdots & 0 \\ \binom{n}{1}\lambda^{n-1} & \lambda^n & 0 & \cdots & 0 \\ \binom{n}{2}\lambda^{n-2} & \binom{n}{1}\lambda^{n-1} & \lambda^n & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ & \cdots & \binom{n}{2}\lambda^{n-2} & \binom{n}{1}\lambda^{n-1} & \lambda^n \end{pmatrix}$$

3.1.1. Exponencial de una matriz

1. Definición: Dada una matriz cuadrada $A \in M_r(\mathbb{C})$, llamaremos exponencial de A, que denotaremos e^A, a la matriz cuadrada

$$e^A := \mathrm{Id} + A + A^2/2! + \cdots + A^n/n! + \cdots = \sum_{i=0}^{\infty} A^i/i!$$

Se cumple que $B \cdot e^A \cdot B^{-1} = e^{B \cdot A \cdot B^{-1}}$ y si A y A' conmutan $e^{A+A'} = e^A \cdot e^{A'}$.

Sigamos notaciones anteriores. $A = B \cdot J \cdot B^{-1}$, entonces $e^A = B \cdot e^J \cdot B^{-1}$. El problema de calcular e^J se reduce al problema de calcular e^C, cuando

$$C = \begin{array}{c} s \\ \begin{pmatrix} \lambda & \cdots & \cdots & 0 \\ 1 & \lambda & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 1 & \lambda \end{pmatrix} \end{array}$$

Ahora bien, $C = \lambda \cdot \mathrm{Id} + N$, con

$$N = \begin{array}{c} \qquad\quad s \\ \begin{pmatrix} 0 & \cdots & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 \end{pmatrix} \end{array}$$

Como $\lambda \cdot \mathrm{Id}$ y N conmutan y $N^s = 0$ entonces, $e^C = e^\lambda \cdot e^N = e^\lambda \cdot (\mathrm{Id} + N + N^2/2 + \cdots + N^{s-1}/(s-1)!)$, es decir,

$$e^C = e^\lambda \cdot \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ 1/2 & 1 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 1/(s-1)! & \cdots & 1/2 & 1 & 1 \end{pmatrix}$$

Calculemos $e^{C\cdot x} = e^{\lambda \cdot x} \cdot e^{N \cdot x} = e^\lambda \cdot (\mathrm{Id} + N \cdot x + N^2 \cdot x^2/2 + \cdots + N^{s-1} \cdot x^{s-1}/(s-1)!)$, es decir,

$$e^{C\cdot x} = e^{\lambda\cdot x} \cdot \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ x & 1 & 0 & \cdots & 0 \\ x^2/2 & x & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ x^{s-1}/(s-1)! & \cdots & x/2 & x & 1 \end{pmatrix}$$

Sea $B = \{f : \mathbb{R} \to \mathbb{C}$, infinitamente derivables$\}$. Buscamos $y_1(x), \dots, y_n(x) \in B$ tales que

$$\begin{array}{rcl} y_1' & = & a_{11} \cdot y_1 + \cdots + a_{1n} \cdot y_n \\ \cdots & = & \cdots \\ y_n' & = & a_{n1} \cdot y_1 + \cdots + a_{nn} \cdot y_n \end{array} \qquad\qquad a_{ij} \in \mathbb{C}$$

Es decir,

$$\begin{pmatrix} y_1' \\ \vdots \\ y_n' \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

Que podemos denotar de modo reducido $Y' = A \cdot Y$, o $DY = A \cdot Y$. Es decir, $(D - A)(Y) = 0$.

Las soluciones de este sistema de ecuaciones diferenciales lineal con coeficientes constantes son $\{Y = e^{Ax} \cdot C$, con $C \in \mathbb{C}^n\}$: Observemos que $D(e^{Ax}W) = e^{Ax}(D + A)(W)$. $(D-A)(Y) = 0$ si y sólo si $0 = (D-A)(e^{Ax} \cdot e^{-Ax} \cdot Y) = e^{Ax} D(e^{-Ax} \cdot Y)$, si y sólo si $e^{-Ax} \cdot Y = C$, con $C \in \mathbb{C}^n$, es decir, $Y = e^{Ax} \cdot C$.

Sea $B = \{s\colon \mathbb{N} \to \mathbb{C}\}$ el $\mathbb{C}$-espacio vectorial de las sucesiones de números complejos. Buscamos $s_1(n),\dots,s_r(n) \in B$ tales que

$$\begin{array}{rcl} \nabla(s_1) & = & a_{11}\cdot s_1 + \cdots + a_{1r}\cdot s_r \\ \cdots & = & \cdots \qquad\qquad\qquad\qquad a_{ij}\in\mathbb{C} \\ \nabla(s_r) & = & a_{r1}\cdot s_1 + \cdots + a_{rr}\cdot s_r \end{array}$$

Es decir,

$$\begin{pmatrix} \nabla(s_1) \\ \vdots \\ \nabla(s_r) \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & & \vdots \\ a_{r1} & \cdots & a_{rr} \end{pmatrix} \cdot \begin{pmatrix} s_1 \\ \vdots \\ s_r \end{pmatrix}$$

Que podemos denotar de modo reducido $\nabla S = A\cdot S$. Es decir, $(\nabla - A)(S) = 0$. Las soluciones de este sistema son $S = A^n\cdot C$, con $C \in \mathbb{C}^r$.

3.2. Comportamiento asintótico de las potencias de una matriz

Supongamos que $A \in M_n(\mathbb{R})$ tiene un autovalor real $\rho > 0$ de multiplicidad 1 y de modo que todo otro autovalor $\lambda \in \mathbb{C}$ cumple que $|\lambda| < \rho$. Sea $\chi_A(x) = (x-\rho)\cdot q(x)$ el polinomio anulador de A y $E' = \operatorname{Ker} q(A)$ y $\langle v\rangle = \operatorname{Ker}(A-\rho)$. Se cumple que

$$\lim_{m\to\infty} (A^m/\rho^m)(e) = \begin{cases} 0, \text{ si } e \in E' \\ v^t, \text{ si } e = v^t \end{cases}$$

Efectivamente, existe una base $B = \{v, e_2, \dots, e_e\}$, donde $\{e_2,\dots,e_n\}$ es una base de E', de modo que si P es la matriz de cambio de base de B a la usual entonces tenemos la descomposición de Jordan,

$$A = P\cdot \begin{pmatrix} \rho & & & \\ & B_1 & & \\ & & \ddots & \\ & & & B_n \end{pmatrix} \cdot P^{-1}, \quad B_i = \begin{pmatrix} B_{i1} & & \\ & \ddots & \\ & & B_{im} \end{pmatrix}, \quad B_{ij} = \begin{pmatrix} \lambda_i & 1 & & \\ & \ddots & \ddots & \\ & & \lambda_i & 1 \\ & & & \lambda_i \end{pmatrix}$$

Es fácil probar que $\lim\limits_{m\to\infty} (B_{ij}^m/\rho^m) = 0$ (pues $|\lambda_i| < \rho$) y que

$$\lim_{m\to\infty} (A^m/\rho^m) = P\cdot \begin{pmatrix} 1 & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{pmatrix} \cdot P^{-1}$$

Luego, $\lim_{m\to\infty}(A^m/\rho^m)$ se anula sobre E' y $(\lim_{m\to\infty}(A^m/\rho^m))(v^t)=v^t$. Con lo que se concluye.

Por tanto, $\operatorname{Im}\lim_{m\to\infty}(A^m/\rho^m)=\langle v^t\rangle$. Sea $w\in\mathbb{R}^n$ tal que $w\cdot e^t=0$ para todo $e\in E'$ y $w\cdot v^t=1$, entonces

$$\lim_{m\to\infty}(A^m/\rho^m)=v^t\cdot w.$$

3.3. Matrices no negativas irreducibles

Buscamos condiciones suficientes para que podamos afirmar que una matriz $A\in M_n(\mathbb{R})$ tiene un autovalor real $\rho>0$ de multiplicidad 1, de modo que todo otro autovalor $\lambda\in\mathbb{C}$ cumple que $|\lambda|<\rho$.

1. Definición: $A=(a_{ij})\in M_n(\mathbb{R})$ diremos que es no negativa si $a_{ij}\geq 0$ para todo i,j. Lo denotaremos $A\geq 0$. Diremos que A es positiva si $a_{ij}>0$, para todo i,j. Se denotará $A>0$.

2. Definición: Diremos que A es irreducible si no existe una matriz de permutación (de vectores de la base) $P\in M_n(\mathbb{R})$ tal que

$$PAP^{-1}=\begin{pmatrix}A_{11} & A_{12}\\ 0 & A_{22}\end{pmatrix},\qquad (P^{-1}=P^t)$$

Es decir, si A es la matriz de un endomorfismo T en cierta base $\{e_1,\dots,e_n\}$, entonces A es irreducible si los subespacios vectoriales generados por unos cuantos vectores de la base no son invariantes por T, salvo que los cojamos todos.

3.3.1. Teorema de Perron-Fröbenius

3. Lema: *Sea $A\in M_n(\mathbb{R})$ no negativa e irreducible. Para todo $v=(v_1,\dots,v_n)\geq 0$ (es decir, $v_i\geq 0$, para todo i) no nulo se cumple que:*

$$(\mathrm{Id}+A)^{n-1}v^t>0$$

Demostración. Observemos que

$$(\mathrm{Id}+A)v^t=v^t+Av^t\geq v^t\geq 0$$

Supongamos que v tiene r entradas positivas y las demás son nulas. Veamos que $(\mathrm{Id}+A)v^t$ tiene al menos $r+1$ entradas positivas: Permutando los vectores de la base, puedo suponer que

$$v^t=\begin{pmatrix}u^t\\ 0\end{pmatrix}$$

con $0 < u \in \mathbb{R}^r$. Escribamos

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

con $A_{11} \in M_r(\mathbb{R})$, no negativa y A_{21} no nula (pues A es irreducible) y no negativa. Entonces,

$$(\mathrm{Id} + A)v^t = v^t + Av^t = \begin{pmatrix} u^t \\ 0 \end{pmatrix} + \begin{pmatrix} A_{11}u^t \\ A_{21}u^t \end{pmatrix} = \begin{pmatrix} u^t + A_{11}u^t \\ A_{21}u^t \end{pmatrix}$$

Como $u^t + A_{11}u^t > 0$ y $0 \neq A_{21}u^t \geq 0$, entonces $(\mathrm{Id} + A)v^t$ tiene al menos $r+1$ entradas positivas. Por tanto, $(\mathrm{Id} + A)^2 v^t$ tiene al menos $r+2$ entradas positivas. En conclusión, $(\mathrm{Id} + A)^{n-1} v^t$ tiene todas las entradas positivas. □

4. Definición: Sea $A \in M_n(\mathbb{R})$ no negativa. Sea $N = \{x \in \mathbb{R}^n : x \geq 0, x \neq 0\}$. Dado $x \in N$, definimos $\rho(x) := \text{máx}\{\lambda \in \mathbb{R} : (A - \rho\,\mathrm{Id})x^t \geq 0\}$.

Obviamente $\rho(x) > 0$.

5. Lema: *Sea $A \in M_n(\mathbb{R})$ no negativa y $x \in N$. Entonces,*

$$\rho(x) = mín\{\frac{\sum_j^n a_{ij}x_j}{x_i} : x_i \neq 0,\ i = 1, \dots, n\}$$

Demostración. Es inmediata. □

6. Proposición: *Sea $A \in M_n(\mathbb{R})$ no negativa e irreducible. Existe $v > 0$ tal que*

$$\rho(v) = máx\{\rho(x) : x \in N\}$$

Demostración. Sea $M = \{x \in N : x_1^2 + \cdots + x_n^2 = 1\}$. Observemos que $\rho(\lambda \cdot x) = \rho(x)$, para todo $\lambda > 0$ y $x \in N$. Por tanto,

$$\text{máx}\{\rho(x) : x \in N\} = \text{máx}\{\rho(x) : x \in M\}$$

Si ρ fuese continua en M, que es compacto, tendríamos que existe $w \in M$ tal que $\rho(w)$ es máximo. Obviamente, ρ es continua en $N \setminus \cup_i \{x_i = 0\}$. Por el lema anterior, $(\mathrm{Id} + A)^{n-1}M \subset N \setminus \cup_i \{x_i = 0\}$, y es un compacto, pues $(\mathrm{Id} + A)^{n-1}$ es continua. En conclusión, ρ es continua en $(\mathrm{Id} + A)^{n-1}M$ y existe $v \in (\mathrm{Id} + A)^{n-1}M$ donde ρ alcanza un máximo.

Observemos que si $x \in M$ e $y^t = (\mathrm{Id} + A)^{n-1}x^t$ entonces $\rho(x) \leq \rho(y)$:

$$(A - \rho(x))y^t = (A - \rho(x))(\mathrm{Id} + A)^{n-1}x^t = (\mathrm{Id} + A)^{n-1}(A - \rho(x))x^t \geq 0$$

por el Lema 3.3.3, ya que $(A - \rho(x))x^t \geq 0$. Por tanto, $\rho(y) \geq \rho(x)$.

En conclusión, "Máximo de ρ en M ≤ Máximo de ρ en $(\mathrm{Id} + A)^{n-1}M$ ≤ Máximo de ρ en N". Como "Máximo de ρ en M = Máximo de ρ en N", entonces "Máximo de ρ en $(\mathrm{Id} + A)^{n-1}M$ = Máximo de ρ en N". Luego, ρ alcanza el máximo en $v > 0$.

□

7. Definición: Sea $A \in M_n(\mathbb{R})$ no negativa e irreducible. Denotaremos $\rho = \max\{\rho(x), x \in N\}$ y diremos que $v \geq 0$ es extremal si $\rho(v) = \rho$.

8. Lema: *Sea $A \in M_n(\mathbb{R})$ no negativa e irreducible.*

1. *Si u es extremal entonces es un autovector de autovalor ρ y $u > 0$.*

2. *Si $\lambda \in \mathbb{C}$ es un autovalor de A entonces $|\lambda| \leq \rho$, si $|\lambda| = \rho$ y $u = (u_1, \dots, u_n) \in \mathbb{C}^n$ es un autovector de A de autovalor λ entonces $|u| := (|u_1|, \dots, |u_n|)$ es extremal.*

Demostración. 1. Si $(A-\rho)u^t \neq 0$, entonces $0 < (\mathrm{Id}+A)^{n-1}((A-\rho)u^t) = (A-\rho)((\mathrm{Id}+A)^{n-1}u^t)$. Si denotamos $x = (\mathrm{Id}+A)^{n-1}u^t$, $(A-\rho)x > 0$ y $\rho < \rho(x)$, contradicción. Como $0 < (\mathrm{Id}+A)^{n-1}u^t = (1+\rho)^{n-1}u^t$ se tiene que $u > 0$.

2. Sea $u \in \mathbb{C}^n$ un autovector de autovalor λ. Tomando valores absolutos $A|u^t| \geq |Au^t| = |\lambda \cdot u^t| = |\lambda||u^t|$, luego $|\lambda| \leq \rho$. Si $\lambda = \rho$, entonces $|u|$ es extremal.

□

9. Teorema de Perron-Fröbenius: *Sea $A \in M_n(\mathbb{R})$ no negativa e irreducible. Entonces,*

1. *ρ es un autovalor de A de multiplicidad* 1.

2. *Si $\lambda \in \mathbb{C}$ es un autovalor de A, entonces $|\lambda| \leq \rho$.*

Demostración. Por el lema 3.3.8 sólo nos falta ver que la multiplicidad del autovalor ρ es 1.

Supongamos que existen dos autovectores linealmente independientes de autovalor ρ, v_1, v_2. Sean λ y μ tales que $u = \lambda v_1 + \mu v_2$ tengan alguna entrada nula. Por Lema 3.3.8, $|u|$ es extremal, lo cual es contradictorio porque tiene entradas nulas. Por tanto, en la matriz de Jordan asociada a A sólo aparece un bloque asociado a ρ.

Para concluir que la multiplicidad de ρ es 1, tenemos que probar que no existe w tal que $(A-\rho)w^t = v^t$ es extremal: Si A es no negativa e irreducible entonces A^t es no negativa e irreducible. Como los autovalores de A y A^t son iguales, por 2. el ρ asociado a A coincide con el asociado a A^t. Sea $z > 0$ extremal de A^t, entonces $0 = (z \cdot (A-\rho))w^t = z \cdot (A-\rho)w^t = z \cdot v > 0$, contradicción.

□

10. Definición: Sea $A \in M_n(\mathbb{R})$ no negativa e irreducible. Al único $v = (v_1, \dots, v_n) \in \mathbb{R}^n$ tal que $(A-\rho)(v^t) = 0$ y tal que $v_1 + \cdots + v_n = 1$ se le denomina autovector de Perron. A ρ se le denomina autovalor de Perron.

11. Proposición: *Sea $A \in M_n(\mathbb{R})$ no negativa e irreducible. Si $A = (a_{ij})$ tiene una fila j de entradas no nulas y λ es un autovalor de A distinto de ρ, entonces $|\lambda| < \rho$.*

Demostración. Sea u un autovector de autovalor λ. Si $|\lambda| = \rho$ entonces, por Lema 3.3.8, $A|u^t| = \rho|u^t| = |\lambda u^t| = |Au^t|$. Entonces, $\sum_i a_{ji}|u_i| = |\sum_i a_{ji}u_i|$, luego $|u_i| = u_i$ (o $|u_i| = -u_i$) para todo i y $\rho = \lambda$. □

12. Definición: Diremos que una matriz cuadrada $A = (a_{ij}) \geq 0$ es primitiva si existe $m \in \mathbb{N}$ tal que $A^m > 0$.

Si A es primitiva entonces es irreducible: Si

$$A = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix} \quad \Rightarrow A^m = \begin{pmatrix} A_{11}^m & A'_{12} \\ 0 & A_{22}^m \end{pmatrix}$$

que no es positiva.

13. Teorema: *Sea A primitiva. Existe un único autovalor real $\rho > 0$ de multiplicidad 1 y de modo que todo otro autovalor λ cumple que $|\lambda| < \rho$.*

Demostración. Si $\lambda_1, \dots, \lambda_n$ son las raíces del polinomio característico de A entonces $\lambda_1^m, \dots, \lambda_n^m$ son las raíces del polinomio característico de A^m. Por el teorema de Perron-Frobenious y 3.3.11, el autovalor de Perron de A^m, ρ_{A^m}, que podemos decir, que es λ_1^m cumple que $|\lambda_2^m|, |\dots, |\lambda_n^m| < \lambda_1^m$. Por tanto, $|\lambda_2|, \dots, |\lambda_n| < |\lambda_1|$ y además λ_1 es el autovalor de Perron de A. □

14. Proposición: *Sea $A \in M_n(\mathbb{R})$ primitiva y $v > 0$ el autovector de Perron de A. Para todo vector no negativo $q \in \mathbb{R}^n$, se cumple que*

$$\lim_{m\to\infty} \frac{A^m q^t}{\|A^m q^t\|_1} = v^t$$

Demostración. Si A es primitiva entonces A^t también lo es. El autovalor de Perron de A coincide con el de A^t. Sabemos que $\lim_{m\to\infty}(A^m/\rho^m) = v^t \cdot w$, donde $v > 0$ es el autovector de Perron de A y por simetría $w > 0$ es el autovector de Perron de A^t (salvo por un factor multiplicativo).

$$\lim_{m\to\infty} \frac{A^m q^t}{\|A^m q^t\|_1} = \lim_{m\to\infty} \frac{A^m/\rho^m(q^t)}{\|A^m/\rho^m(q^t)\|_1} = \frac{v^t \cdot w \cdot q^t}{(1,\cdots,1)\cdot v^t \cdot w \cdot q^t} = \frac{v^t}{(1,\cdots,1)\cdot v^t} = v^t$$

□

3.3.2. Matrices de Leslie

Dividamos la población de hembras de una misma especie en distintos grupos de edad $G_1, G_2, \ldots, G_n$, donde cada grupo tiene la misma amplitud. Así, si la vida más larga se estima en L años, la amplitud de cada grupo de edades es de L/n años. El grupo G_1 está formado por los individuos cuya edad está en el intervalo $[0, L/n)$ es decir, que tienen menos de L/n años. El siguiente grupo por edades G_1, lo forman los individuos cuya edad está en el intervalo $[L/n, 2L/n)$. El siguiente grupo lo forman los individuos con edad en $[2L/n, 3L/n)$, y así, hasta llegar al último grupo formado por los individuos cuya edad está comprendida en el intervalo $[(n-1)L/n, L]$.

Supongamos que los censos de población se realizan en intervalos de tiempo iguales a la amplitud de los grupos de edades, y consideremos las tasas de fecundidad y supervivencia: denotamos por f_i el número promedio de hijas de cada hembra del grupo G_i (esto es la tasa de fecundidad específica del grupo G_i). Llamamos s_i a la fracción de individuos del grupo G_i que sobreviven al intervalo entre censos y pasan a formar parte del grupo G_{i+1}.

Si $p_i(m)$ es el número de hembras de G_i en el instante m, entonces se sigue que

$$\begin{array}{rcl} p_1(m+1) & = & p_1(m)f_1 + p_2(m)f_1 + \ldots + p_n(m)f_n \\ p_i(m+1) & = & p_{i-1}(m)s_{i-1}; \quad \text{para } i = 2, \ldots, n. \end{array} \tag{3.3.1}$$

Además,

$$P_i(m) = \frac{p_i(m)}{p_0(m) + p_1(m) + \ldots + p_n(m)}$$

es la proporción de población en G_i en el instante m.

El vector $\mathbf{P}(m) = (P_1(m), P_2(m), \ldots, P_n(m))^t$ representa a la *distribución de edades de la población* en el instante m, y, suponiendo que existe, $\mathbf{P}^* = \lim_{m\to\infty} \mathbf{P}(m)$ es la *distribución de edades de la población a largo plazo*.

Las ecuaciones (3.3.1) constituyen un sistema de ecuaciones lineales en diferencias homogéneo que se puede escribir en forma matricial como

$$\mathbf{p}(m) = A\,\mathbf{p}(m-1), \quad \text{donde} \quad A = \begin{pmatrix} f_1 & f_2 & \ldots & f_{n-1} & f_n \\ s_1 & 0 & \ldots & 0 & 0 \\ 0 & s_2 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & s_{n-1} & 0 \end{pmatrix} \in \mathcal{M}_n(\mathbb{R}) \tag{3.3.2}$$

y $\mathbf{p}(m) = (p_1(m), \ldots, p_n(m))^t$, para todo $m \geq 0$. De modo que $\mathbf{p}(m) = A^m \mathbf{p}(0)$ para todo $m > 0$.

La matriz A se llama **Matriz de Leslie** en honor de P.H. Leslie que introdujo este modelo en 1945.

La matriz A es una matriz no negativa, pues $s_i \geq 0,\ i = 1,\ldots,n-1$ y $f_i \geq 0,\ i = 1,\ldots,n$. Además, si $n > 2$ y $f_{n-1} \cdot f_n$ y $s_1 \cdots s_{n-1}$ son no nulos, entonces A es primitiva (ejercicio 20). En este último caso, por 3.3.14, $\mathbf{P}^*$ es el autovector de Perron de A, es decir, el autovector de Perron es la distribución de edades de la población a largo plazo.

3.3.3. Cadenas de Markov

15. Definición: Sea $P = (p_{ij}) \in \mathscr{M}_n(\mathbb{R})$ tal que $p_{ij} \in [0,1],\ i,j = 1,\ldots,n$. Se dice que P es una **matriz estocástica** cuando sus columnas suman 1. Diremos que es **doblemente estocástica** cuando sus columnas y filas suman 1.

Nos centraremos en el caso en que las columnas suman 1. No es raro encontrar textos donde esta condición se supone sobre las filas, pero los resultados son semejantes.

De esta forma una matriz estocástica tiene como columnas a vectores de probabilidad. Nótese que las matrices estocásticas son no negativas.

16. Definición: Un **vector** no negativo $\mathbf{p} = (p_1,\ldots,p_n)^{\mathrm{t}} \in \mathbb{R}^m$ se dice que es **de probabilidad** si $\|\mathbf{p}\|_1 := \sum_{i=1}^n p_i = 1$.

17. Proposición: *Si $A \in M_n(\mathbb{R})$ es una matriz estocástica y $\mathbf{p} \in \mathbb{R}^{\mathbf{n}}$ es un vector de probabilidad, entonces $A \cdot \mathbf{p^t}$ es un vector de probabilidad.*

18. Proposición: *Si $A \in M_n(\mathbb{R})$ es una matriz estocástica, entonces 1 es un autovalor de A y cualquier otro autovalor $\lambda \in \mathbb{C}$, cumple que $|\lambda| \leq 1$.*

Demostración. El vector $(1,\cdots,1)$ es un autovector de A^t, de autovalor 1. Por tanto, 1 es un autovalor de A.

Sea $\lambda \in \mathbb{C}$ un autovalor de A, tal que $|\lambda| > 1$. Para todo vector de probabilidad p' se cumple que $0 \leq p'/|\lambda|^m \leq (1/|\lambda|^m,\ldots,1/|\lambda|^m)$. Por tanto,

$$\lim_{m\to\infty}((A^m/\lambda^m)(p)) = 0$$

para todo vector de probabilidad p. Como la base estándar está formada por vectores de probabilidad, tenemos que $\lim\limits_{m\to\infty}(A^m/\lambda^m) = 0$. Por otra parte si v es un autovector de autovalor λ, tenemos que $\lim\limits_{m\to\infty}((A^m/\lambda^m)v) = v$ y hemos llegado a contradicción.

□

19. Teorema: *Sea $A \in M_n(\mathbb{R})$ una matriz estocástica primitiva. Entonces $\rho = 1$ es el autovalor de Perron de A. Si v es el autovector de Perron, para todo vector p de* *probabilidad* $\lim_{m\to\infty}(A^m p^t) = v^t$, *es decir,*

$$\lim_{m\to\infty} A^m = (v^t, \overset{n}{\ldots}, v^t).$$

Supongamos que estamos observando algún fenómeno aleatorio a lo largo del tiempo, y que en cualquier punto concreto del tiempo nuestra observación puede tomar uno de los valores, a veces llamados estados, $a_1, a_2, \ldots, a_s$. En otras palabras, tenemos una sucesión de variables aleatorias X_m; para periodos de tiempo $m = 0, 1, \ldots, n$ donde cada variable puede ser igual a $a_1, a_2, \ldots, a_s$. Si la probabilidad de que X_m se encuentre en el estado a_i sólo depende del estado en que se hallase X_{m-1} y no en los estados de periodos anteriores de tiempo, entonces el proceso se dice que es una cadena de Markov. Si la probabilidad tampoco depende del valor de m; entonces la cadenas de Markov se dice que es homogénea, y si el número de estados es finito, como es nuestro caso, la cadena de Markov se dice finita. En el caso de las cadenas de Markov homogéneas y finitas, la probabilidades de cualquier periodo de tiempo se pueden calcular a partir de la probabilidades iniciales

En el caso de las cadenas de Markov homogéneas y finitas, la probabilidades de cualquier periodo de tiempo se pueden calcular a partir de la probabilidades iniciales de los estados y lo que se conoce como probabilidades de transición. Denotaremos

$$\mathbf{p}_0 = \begin{pmatrix} p_{a_1}^{(0)} \\ \vdots \\ p_{a_s}^{(0)} \end{pmatrix}$$

al vector de probabilidades iniciales, donde $p_i^{(0)}$ es la probabilidad de que el proceso comience en el estado a_i. La **matriz de transición de probabilidades** es la matriz $P \in \mathcal{M}_s(\mathbb{R})$ cuya entrada (i,j)-ésima, p_{ij}, da la probabilidad de que X_m se halle en el estado a_i supuesto que X_{m-1} se hallaba en el estado a_j. Por consiguiente, si

$$\mathbf{p}_m = \begin{pmatrix} p_{a_1}^{(m)} \\ \vdots \\ p_{a_s}^{(m)} \end{pmatrix}$$

siendo $p_{a_i}^{(m)}$ la probabilidad de que el sistema se encuentre en el estado a_i en el instante m, entonces, por el teorema de la probabilidad total se tiene que

$$\begin{aligned} \mathbf{p}_1 &= P\,\mathbf{p}_0, \\ \mathbf{p}_2 &= P\,\mathbf{p}_1 = P\,P\,\mathbf{p}_0 = P^2\mathbf{p}_0, \end{aligned}$$

y en general,

$$\mathbf{p}_m = P^m \mathbf{p}_0.$$

Nótese que P es una matriz estocástica pues su columna j-ésima nos indica la probabilidad de los posibles estados en un determinado instante cuando en el instante inmediatamente anterior el estado sea a_j.

Si tenemos una población considerable de individuos sujetos a este proceso aleatorio, entonces $p_{a_i}^{(m)}$ se puede describir como la proporción de individuos en el estado a_i al instante m, mientras que $p_i^{(0)}$ sería la proporción de individuos que comienzan en el estado a_i. De modo natural nos podemos hacer las siguientes preguntas ¿qué ocurre con estas proporciones cuando m aumenta? Es decir, ¿podemos determinar el comportamiento límite de $\mathbf{p}_m$? Nótese que la respuesta depende del comportamiento asintótico de P^m, y que P es una matriz no negativa ya que cada una de sus entradas es una probabilidad. Por consiguiente, si P es primitiva, sabemos que el autovalor de Perron de P es 1 y que si $\mathbf{p}$ es el autovector de Perron de P, entonces $\lim_{m\to\infty} \mathbf{p}_m = \mathbf{p}$. Por tanto, el sistema se aproxima a un punto de equilibrio en que las proporciones de los distintos estados vienen dadas por las entradas de $\mathbf{p}$. Además, el comportamiento límite no depende de las proporciones iniciales.

3.3.4. Sistemas de seguridad

Consideremos un sistema que tiene dos controles independientes, A y B, que previene que el sistema sea destruido. El sistema se activa en momentos discretos $t_1, t_2, t_3, \ldots$, y el sistema se considera bajo control si alguno de los controles A o B funciona en el momento de la activación. El sistema se destruye si A y B fallan simultáneamente. Por ejemplo, un automóvil tiene dos sistemas de frenado independientes, el freno de pedal y el freno de mano. El automóvil está bajo control si al menos uno de los sistemas de frenado está operativo cuando intentamos parar, pero choca si ambos sistemas fallan simultáneamente.

Si uno de los controles falla en un punto de activación pero el otro control funciona, entonces el control defectuoso es reemplazado antes de la siguiente activación. Si un control funciona en el momento t entonces se considera fiable en un 90% para la activación $t+1$. Sin embargo, si un control falla en el instante t, entonces su recambio no probado se considera fiable en un 60% para $t+1$.

La pregunta que nos planteamos es: ¿Puede el sistema funcionar indefinidamente sin ser destruido? Si no, ¿cuánto tiempo se espera que el sistema funcione antes de la destrucción?

Este problema se puede modelizar con una cadena de Markov con cuatro estados, definidos por los controles que funcionen en un momento de activación. Podemos poner entonces que el espacio de estados es el conjunto de pares (a,b) tales que

$$a = \begin{cases} 1 & \text{si } A \text{ funciona,} \\ 0 & \text{si } A \text{ falla,} \end{cases} \text{ y } b = \begin{cases} 1 & \text{si } B \text{ funciona,} \\ 0 & \text{si } B \text{ falla.} \end{cases}$$

El estado $(0,0)$ es absorbente, es decir, si se llega a él no se puede salir.

Por simplicidad, escribiremos $1,2,3$ y 4 en vez de $(1,1)$, $(1,0)(0,1)$ y $(0,0)$, respectivamente. De este modo la matriz de transición es

$$P = \begin{pmatrix} 0{,}81 & 0{,}54 & 0{,}54 & 0 \\ 0{,}09 & 0{,}36 & 0{,}06 & 0 \\ 0{,}09 & 0{,}06 & 0{,}36 & 0 \\ 0{,}01 & 0{,}04 & 0{,}04 & 1 \end{pmatrix}$$

La cuarta fila de matriz P es positiva. El autovalor de Perron es 1, que es de multiplicidad 1, el autovector de Perron es $v = (0,0,0,1)$ y cualquier otro autovalor $\lambda \in \mathbb{C}$, cumple que $|\lambda| < 1$. Entonces, existe el límite $\lim_{n\to\infty} P^n$, y es igual a

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

Esto significa que el estado absorbente se alcanza siempre, partamos de donde partamos. Así que tenemos respondida a la primera pregunta: el sistema se destruirá, a la larga, con probabilidad 1. La segunda cuestión que planteábamos es en cuántos procesos de activación llegaremos al desastre.

Partamos de un estado concreto, por ejemplo $(1,0,0,0)$. La probabilidad m_n de que en el momento del proceso n no estemos en estado de desastre es

$$m_n = (1,1,1,0) \cdot P^n \cdot (1,0,0,0)^t$$

Sea P_{11} la submatriz de P formada por las tres primeras filas y columnas, $P = \begin{pmatrix} P_{11} & \mathbf{0} \\ \mathbf{p}_{12} & 1 \end{pmatrix}$. Escribamos $u = (1,1,1)$. Entonces, $m_n = u \cdot P_{11}^n \cdot (1,0,0)^t$.

Por tanto, nuestra esperanza de vida es

$$\begin{aligned} m &= 1 + u \cdot P_{11} \cdot (1,0,0)^t + \cdots + u \cdot P_{11}^n \cdot (1,0,0)^t + \cdots = u \cdot (\mathrm{Id} + P_{11} + \cdots + P_{11}^n + \cdots) \cdot (1,0,0)^t \\ &= u \cdot (\mathrm{Id} - P_{11})^{-1} \cdot (1,0,0)^t \end{aligned}$$

En nuestro caso,

$$(\mathrm{Id} - P_{11})^{-1} = \begin{pmatrix} 44{,}615 & 41{,}538 & 41{,}538 \\ 6{,}9231 & 8{,}022 & 6{,}5934 \\ 6{,}9231 & 6{,}5934 & 8{,}022 \end{pmatrix},$$

y

$$u \cdot (\mathrm{Id} - P_{11})^{-1} = \begin{pmatrix} 58{,}462 & 56{,}154 & 56{,}154 \end{pmatrix}.$$

Interpretemos los resultados. El tiempo medio para fallo si partimos con los dos controles probados es algo más de 58 pasos, mientras que el tiempo medio para fallo si partimos con uno de los controles no probado está alrededor de los 56 pasos. La diferencia no parece significativa, pero vamos a considerar qué ocurre usamos solamente un control en el sistema. En este caso, solamente hay dos estados en la cadena de Markov: 1 (control que funciona) y 2 (control que no funciona). La matriz de transición queda

$$P = \begin{pmatrix} 0{,}9 & 0 \\ 0{,}1 & 1 \end{pmatrix}$$

por lo que el tiempo medio de fallo es únicamente de $u(\mathrm{Id} - P_{11})^{-1} = 10$ pasos ¿Qué ocurrirá si usamos tres controles independientes?

3.4. Problemas

1. *a*) Sea $X' = AX$ un sistema homogéneo de ecuaciones diferenciales, siendo A una matriz cuadrada de coeficientes constantes. Probar que $e^{At} \cdot C$ son las soluciones del sistema, siendo C una matriz columna de constantes.

   b) Sea $X' = AX + B(t)$ un sistema lineal de ecuaciones diferenciales. Calcular la matriz columna $C(t)$ tal que $e^{At} \cdot C(t)$ sea una solución del sistema.

2. Resuélvanse los siguientes sistemas de ecuaciones diferenciales

$$\begin{array}{lll} \frac{dx}{dt} = x - 3y + 3z & \frac{dx}{dt} = 3x - y & \frac{dx}{dt} = -11x - 4y \\ \frac{dy}{dt} = -2x - 6y + 13z & \frac{dy}{dt} = x + y & \frac{dy}{dt} = 15x + 6y \\ \frac{dz}{dt} = -x - 4y + 8z & \frac{dy}{dt} = 3x + 5z - 3u & \\ & \frac{du}{dt} = 4x - y + 3z - u & \end{array}$$

3. Sea $P(x) \in \mathbb{R}[x]$ un polinomio de grado n. Probar que la ecuación diferencial $P(D)y = f(x)$ es equivalente a un sistema de ecuaciones diferenciales lineales con coeficientes constantes de primer orden de n variables.

4. *a*) Sea $P(x) \in \mathbb{R}[x]$ un polinomio mónico de grado n. Sean $s_1(x),\dots,s_n(x)$ soluciones, linealmente independientes, de la ecuación diferencial $P(D)y = 0$. Probar que si $c_1(x),\dots,c_n(x)$ cumplen las ecuaciones

$$\begin{aligned}
&c_1(x)'s_1(x) + \ldots + c_n(x)'s_n(x) = 0\\
&\ldots\\
&c_1(x)'s_1(x)^{n-2)} + \ldots + c_n(x)'s_n(x)^{n-2)} = 0\\
&c_1(x)'s_1(x)^{n-1)} + \ldots + c_n(x)'s_n(x)^{n-1)} = f(x)
\end{aligned}$$

entonces $c_1(x)s_1(x) + \ldots + c_n(x)s_n(x)$ es una solución particular de $P(D)y = f(x)$.

b) Pruébese este resultado como caso particular de 1 (b).

5. Sea A una matriz con coeficientes en $k[D]$. Probar que mediante las transformaciones elementales, el problema de resolver los sistemas $A(X(t)) = Y(t)$, se reduce al problema de resolver ecuaciones $P(D)f(t) = h(t)$.

6. Resolver el sistema de ecuaciones diferenciales

$$\begin{aligned}
&x'' - x + y' = e^t\\
&x'' + 2x' + x + y'' = e^t
\end{aligned}$$

7. Sea

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

Los autovalores de esta matriz son $\lambda_1 = 1+\sqrt{3}$, $\lambda_2 = 1-\sqrt{3}$ y $\lambda_3 = 0$. El autovalor de mayor módulo es λ_1. Asociado a este autovalor tenemos el autovector $\mathbf{v} = (\sqrt{3}-1, 1, 1)$ de componentes estrictamente positivas.

Para un vector cualquiera $\mathbf{b}$, comprobar que el producto $B^m\mathbf{b}$ se aproxima, para valores grandes de m a $c\lambda_1^m\mathbf{v}_1$, donde c es una cierta constante y $\mathbf{v_1}$ es un autovector asociado a λ_1.

8. Sean V un $\mathbb{R}$-espacio vectorial de dimensión $n > 0$ y $T \in \operatorname{End}_{\mathbb{R}}(V)$ diagonalizable. Dado $r \in \mathbb{Z}_+$, diremos que $S \in \operatorname{End}_{\mathbb{R}}(V)$ es una **raíz r-ésima de T** si $S^r = T$. Encontrar condiciones necesarias y suficientes para que existan raíces r-ésimas de T.

Sean $V = \mathbb{R}^3$ y $T \in \operatorname{End}_{\mathbb{R}}(V)$ tal que su matriz asociada respecto de la base usual de $\mathbb{R}^3$ es

$$\begin{pmatrix} 8 & -6 & 4\\ -6 & 9 & -2\\ 4 & -2 & 4\end{pmatrix}.$$

Hallar, si es posible, la matriz asociada a la raíz cuadrada de T respecto de la base usual de $\mathbb{R}^3$.

9. Sean $V = \mathbb{R}^3$ y $T \in \mathrm{End}_{\mathbb{R}}(V)$ tal que su matriz asociada respecto de la base usual de $\mathbb{R}^3$ es
$$(a)\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \qquad (b)\begin{pmatrix} 5 & 0 & 0 \\ 0 & -1 & 1 \\ 3 & 0 & 2 \end{pmatrix}.$$
Hallar la matriz asociada T^m respecto de la base usual de $\mathbb{R}^3$

10. Dado el sistema de ecuaciones en diferencias $\mathbf{u}_n = A\mathbf{u}_{n-1}$, siendo
$$A = \begin{pmatrix} 0 & a^2 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & a^2 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$
   1. Obtener la expresión general de $\mathbf{u}_n$.
   2. Calcular $\mathbf{u}_{10}$, dado el vector inicial $\mathbf{u}_0 = (0,2,0,2)$.

11. Dadas $(a_{ij}),(b_{ij}) \in \mathcal{M}_n(\mathbb{R})$ diremos que $(a_{ij}) \geq (b_{ij})$ si $a_{ij} \geq b_{ij}$, para todo i,j. Supongamos que $(a_{ij}) \geq (b_{ij})$. Probar que si (b_{ij}) es no negativa entonces (a_{ij}) es no negativa. Probar que si (b_{ij}) es no negativae irreducible entonces (a_{ij}) es no negativa e irreducible. Probar que si (b_{ij}) es primitiva entonces (a_{ij}) es primitiva.

12. Sean $A \in \mathcal{M}_n(\mathbb{R})$ y $\varepsilon > 0$. Probar que si A es no negativa e irreducible, entonces $(\varepsilon I_n + A)^{n-1} > 0$.

13. Sea $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{R})$ una matriz no negativa e irreducible. Si $a_{ii} \neq 0$, para todo $i = 1,\ldots,n$, entonces A es primitiva. [Tómese $\varepsilon = \min\{a_{ii} \mid i = 1,\ldots,n\}$, compruébese que $B = A - \varepsilon I_n$ es no negativa e irreducible, y úsese el ejercicio 12 para concluir que $A = \varepsilon I_n + B$ es primitiva.

14. Verificar que si P es una matriz de primitiva de orden n, cuyas filas suman todas uno, entonces su vector de Perron tiene todas sus componentes iguales a $1/n$.

15. Sea $A \in \mathcal{M}_n(\mathbb{R})$ una matriz positiva e irreducible. Probar que si la suma de las entradas de toda fila (o columna) es ρ, entonces el autovalor de Perron de A es ρ.

16. Comprobar el teorema de Perron-Fröbenius calculando los autovalores y autovectores de la matriz
$$A = \begin{pmatrix} 7 & 2 & 3 \\ 1 & 8 & 3 \\ 1 & 2 & 9 \end{pmatrix}.$$
Encontrar el autovalor y el autovector de Perron de A.

17. Calcular el autovalor y el autovector de Perron de la matriz
$$A = \begin{pmatrix} 1-\alpha & \beta \\ \alpha & 1-\beta \end{pmatrix},$$
donde $\alpha + \beta = 1$ con α y $\beta > 0$.

18. Sea
$$A = \begin{pmatrix} 0 & 1 & 0 \\ 3 & 0 & 3 \\ 0 & 2 & 0 \end{pmatrix}.$$
    1. Probar que A es irreducible.
    2. Hallar el autovalor y el autovector de Perron de A.

19. Demuestre que el polinomio característico de la matriz
$$A = \begin{pmatrix} f_1 & f_2 & f_3 \\ s_1 & 0 & 0 \\ 0 & s_2 & 0 \end{pmatrix}$$
es igual a
$$\chi_A(x) = \det(xI - A) = x^3 - f_1x^2 - f_2s_1x - f_3s_1s_2.$$
Demuestre que el polinomio característico de la matriz
$$A = \begin{pmatrix} f_1 & f_2 & f_3 & f_4 \\ s_1 & 0 & 0 & 0 \\ 0 & s_2 & 0 & 0 \\ 0 & 0 & s_3 & 0 \end{pmatrix}$$
es igual a
$$\chi_A(x) = \det(xI - A) = x^4 - f_1x^3 - f_2s_1x^2 - f_3s_1s_2x - f_4s_1s_2s_3.$$

Dada la matriz de Leslie

$$A = \begin{pmatrix} f_1 & f_2 & \dots & f_{n-1} & f_n \\ s_1 & 0 & \dots & 0 & 0 \\ 0 & s_2 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & s_{n-1} & 0 \end{pmatrix}$$

intente deducir una fórmula para su polinomio característico.

20. Sea $A \in \mathcal{M}_n(\mathbb{R})$ una matriz de Leslie tal que $f_{n-1} \cdot f_n \neq 0$ y $s_1 \cdots s_{n-1} \neq 0$. Probar que

    1. A es irreducible.
    2. Si $f_1 = \ldots = f_{n-2} = 0$, entonces

    $$A^n = s_1 \cdots s_{n-2} f_{n-1} A + s_1 \cdots s_{n-1} f_n I_n.$$

    Usando esta igualdad concluir que es no negativa e irreducible y, por el ejercicio 13, que es primitiva.

    3. En general $A^n = s_1 \cdots s_{n-2} f_{n-1} A + s_1 \cdots s_{n-1} f_n I_n + B$ para cierta matriz B no negativa. Usando esta igualdad concluir que es no negativa e irreducible y, por el ejercicio 13, que es primitiva.

21. Un estudio ha determinado que el sector de ocupación de un niño, cuando sea adulto, depende del sector en que trabaje su padre, y está dada por la siguiente matriz de transición, con los sectores de producción P = sector primario, S = sector secundario, T = sector terciario.

$$\begin{array}{cc} & \text{Sector del padre} \\ & \begin{array}{ccc} \text{T} & \text{S} & \text{P} \end{array} \\ \text{Sector del hijo} \begin{array}{c} \text{T} \\ \text{S} \\ \text{P} \end{array} & \begin{pmatrix} 0{,}8 & 0{,}3 & 0{,}2 \\ 0{,}1 & 0{,}5 & 0{,}2 \\ 0{,}1 & 0{,}2 & 0{,}6 \end{pmatrix} \end{array}$$

Así, la probabilidad de que el hijo de alguien que trabaja en el sector terciario también lo haga en ese sector es 0,8.

1. ¿Cuál es la probabilidad de que el nieto de un trabajador del sector terciario trabaje en ese sector?
2. A largo plazo, ¿qué proporción de la población trabajará en el sector secundario?

22. Para la matriz de transición
$$P = \begin{pmatrix} 0{,}4 & 0{,}5 \\ 0{,}6 & 0{,}5 \end{pmatrix},$$
   1. calcular $\mathbf{x}^{(m)}$ para $n = 1,2,3,4,5$, si $\mathbf{x}^{(0)} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$;
   2. probar que P es una matriz primitiva y calcular el vector de estado estacionario.

23. Consideremos la matriz de transición
$$P = \begin{pmatrix} 0{,}5 & 0 \\ 0{,}5 & 1 \end{pmatrix}.$$
   1. Probar que P no es primitiva.
   2. Probar que cuando $m \to \infty$, $P^m\mathbf{x}^{(0)}$ se aproxima a $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$, para cualquier vector de probabilidad inicial $\mathbf{x}^{(0)}$.

24. Probar que la matriz de transición
$$P = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}$$
es primitiva, y aplicar el ejercicio 23 para calcular su vector de estado estacionario.

25. Consideremos la sucesión de matrices de transición $\{P_2, P_3, P_4, \ldots\}$, con
$$P_2 = \begin{pmatrix} 0 & \frac{1}{2} \\ 1 & \frac{1}{2} \end{pmatrix}, \qquad P_3 = \begin{pmatrix} 0 & 0 & \frac{1}{3} \\ 0 & \frac{1}{2} & \frac{1}{3} \\ 1 & \frac{1}{2} & \frac{1}{3} \end{pmatrix},$$
$$P_4 = \begin{pmatrix} 0 & 0 & 0 & \frac{1}{4} \\ 0 & 0 & \frac{1}{3} & \frac{1}{4} \\ 0 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \end{pmatrix}, \quad P_5 = \begin{pmatrix} 0 & 0 & 0 & 0 & \frac{1}{5} \\ 0 & 0 & 0 & \frac{1}{4} & \frac{1}{5} \\ 0 & 0 & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ 0 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{pmatrix}.$$
y sucesivamente. Probar que estas matrices de transición son regulares, y determinar los vectores de estado estacionarios $\mathbf{x}_m$ tales que $P_m\mathbf{x}_m = \mathbf{x}_m$, para $m = 2,3,\ldots,n$.

Capítulo 4

Espacio vectorial euclídeo

4.1. Aplicaciones bilineales

Sea V un $\mathbb{R}$-espacio vectorial.

1. Definición: Una aplicación $T_2\colon V \times V \to \mathbb{R}$ diremos que es una aplicación bilineal sobre V su cumple:

1. $T_2(\lambda v_1 + v_2, v_3) = \lambda T_2(v_1, v_3) + T_2(v_1, v_3)$, para todo $v_1, v_2, v_3 \in V$ y $\lambda \in \mathbb{R}$.

2. $T_2(v_3, \lambda v_1 + v_2) = \lambda T_2(v_3, v_1) + T_2(v_3, v_2)$, para todo $v_1, v_2, v_3 \in V$ y $\lambda \in \mathbb{R}$.

Diremos, además, que T_2 es simétrica si $T_2(v, v') = T_2(v', v)$, para todo $v, v' \in V$. Diremos que es antisimétrica si $T_2(v, v') = -T_2(v', v)$, para todo $v, v' \in V$.

2. Definición: Sea $B = \{v_1, \ldots, v_n\}$ una base de V. Se llama matriz asociada a T_2 respecto de la base B a la matriz $A = (a_{ij}) \in M_n(\mathbb{R})$, definida por

$$a_{ij} := T_2(v_i, v_j), \quad \forall i, j$$

3. Proposición: *Sea V un $\mathbb{R}$-espacio vectorial, $B = \{v_1, \ldots, v_n\}$ una base de V, T_2 una aplicación bilineal en V y $A = (a_{ij})$ la matriz asociada a T_2 en la base B. Dados $v, v' \in V$ de coordenadas $(x_1, \ldots, x_n)$ e $(y_1, \ldots, y_n)$ respecto de la base B, entonces*

$$T_2(v, v') = (x_1, \ldots, x_n) \cdot (a_{ij}) \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

Demostración.

$$T_2(v, v') = T_2(\sum_i x_i v_i, \sum_j y_j v_j) = \sum_{ij} x_i y_j T_2(v_i, v_j) = \sum_{ij} x_i a_{ij} y_j = (x_1, \ldots, x_n) \cdot (a_{ij}) \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

□

Por tanto, la aplicación bilineal queda determinada por su matriz asociada.

Recíprocamente, sea $B = \{v_1,\dots,v_n\}$ una base de V. Dada una matriz (a_{ij}) si definimos

$$T_2(v,v') := (x_1,\dots,x_n)\cdot(a_{ij})\cdot\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

donde $(x_1,\dots,x_n)$ e $(y_1,\dots,y_n)$ son las coordenadas de v,v' respecto de la base B, entonces T_2 es una aplicación bilineal de matriz asociada (a_{ij}) en la base B.

4. Ejemplo: Sea $V = \mathbb{R}^n$.

$$\begin{array}{rccl} T_2\colon & \mathbb{R}^n\times\mathbb{R}^n & \to & \mathbb{R} \\ & ((x_1,\dots,x_n),(y_1,\dots,y_n)) & \mapsto & x_1y_1+\dots+x_ny_n \end{array}$$

es una aplicación bilineal. En la base usual de $\mathbb{R}^n$, la matriz asociada a T_2 es

$$T_2 \equiv \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix}$$

Si definimos $T_2((x_1,\dots,x_n),(y_1,\dots,y_n)) := \sum_{ij} a_{ij}x_iy_j$, entonces la matriz asociada a T_2 en la base usual de $\mathbb{R}^n$ es (a_{ij}).

5. Proposición: *Sea V un $\mathbb{R}$-espacio vectorial, $B = \{v_1,\dots,v_n\}$ una base de V, T_2 una aplicación bilineal en V y $A = (a_{ij})$ la matriz asociada a T_2 en la base B. Entonces, T_2 es simétrica si y sólo si (a_{ij}) es simétrica.*

Demostración. Si T_2 es simétrica entonces $a_{ij} = T_2(v_i,v_j) = T_2(v_j,v_i) = a_{ij}$ y la matriz (a_{ij}) es simétrica.

Supongamos ahora que (a_{ij}) es simétrica. Sea $v = (x_1,\dots,x_n)$ y $v' = (y_1,\dots,y_n)$, entonces

$$T_2(v,v') = \sum_{ij} x_iy_ja_{ij} = \sum_{ij} x_iy_ja_{ji} = T_2(v',v)$$

Luego T_2 es simétrica. □

6. Proposición: *Sea V un $\mathbb{R}$-espacio vectorial, $B = \{v_1,\dots,v_n\}$ una base de V, T_2 una aplicación bilineal en V y $A = (a_{ij})$ la matriz asociada a T_2 en la base B. Sea $B' = \{v'_1,\dots,v'_n\}$ otra base de V y $A' = (a'_{ij})$ la matriz asociada a T_2 en la base B'. Si $P = (p_{ij})$ es la matriz de cambio de base de B' a B entonces*

$$A' = P^t\cdot A\cdot P$$

Demostración. Se tiene que

$$\begin{aligned} a'_{ij} &= T_2(v'_i, v'_j) = T_2(\sum_l p_{li} v_l, \sum_{l'} p_{l'j} v_{l'}) = \sum_{l,l'} p_{li} p_{l'j} T_2(v_l, v_{l'}) = \sum_{l,l'} p_{li} a_{ll'} p_{l'j} \\ &= \sum_{l'} (P^t A)_{il'} p_{l'j} = (P^t A P)_{ij} \end{aligned}$$

Por tanto, $A' = P^t \cdot A \cdot P$. □

7. Definición: Dadas $A, A' \in M_n(\mathbb{R})$ se dice que A' es congruente con A si existe una matriz invertible $P \in M_n(\mathbb{R})$ tal que

$$A' = P^t A P$$

Es decir, A y A' son congruentes si son las matrices asociadas a una aplicación bilineal T_2 en dos bases B y B' (si P es la matriz de cambio de base de B' a B, entonces $A' = P^t A P$).

4.2. Producto escalar

1. Definición: Diremos que una aplicación bilineal $T_2 \colon V \times V \to \mathbb{R}$ es definida positiva si $T_2(v,v) > 0$, para todo $v \neq 0$. Diremos que es semidefinida positiva si $T_2(v,v) \geq 0$, para todo $v \neq 0$.

2. Definición: Si T_2 es una aplicación bilineal, simétrica, definida positiva diremos que T_2 es un producto escalar y diremos que la pareja (V, T_2) es un espacio vectorial euclídeo.

3. Ejemplo: Sea $V = \mathbb{R}^n$. La aplicación bilineal

$$\begin{array}{rccc} T_2\colon & \mathbb{R}^n \times \mathbb{R}^n & \to & \mathbb{R} \\ & ((x_1, \dots, x_n), (y_1, \dots, y_n)) & \mapsto & x_1 y_1 + \ldots + x_n y_n \end{array}$$

es simétrica, pues la matriz asociada en la base usual de $\mathbb{R}^n$ es la identidad, que es simétrica, y es definida positiva, ya que

$$T_2((x_1, \dots, x_n), (x_1, \dots, x_n)) = \sum_i x_i^2 > 0$$

si $(x_1, \dots, x_n) \neq 0$.

Se dice que T_2 es el producto escalar usual del espacio vectorial euclídeo $\mathbb{R}^n$.

4. Notación: Si (V, T_2) es un espacio vectorial euclídeo también se suele denotar $(V, T_2) = (V, \cdot)$ y $T_2(v, v') = v \cdot v'$.

5. Definición: Sea $(E,\cdot)$ un espacio vectorial euclídeo. Diremos que los vectores no nulos $\{e_1,\dots,e_r\} \in E$ son ortogonales si $e_i \cdot e_j = 0$, para todo $i \neq j$. Diremos que son ortonormales si $e_i \cdot e_j = \delta_{ij}$, para todo i, j.

Probaremos que en espacios vectoriales euclídeos siempre existen bases ortonormales de vectores. Una base $\{e_1,\dots,e_r\} \in E$ es ortonormal si y sólo si la matriz asociada al producto escalar, T_2, en esta base es la matriz identidad.

6. Definición: Sea $(E,\cdot)$ un espacio vectorial euclídeo y $e \in E$ un vector. Llamaremos módulo o norma de e y lo denotaremos $||e||$ a:

$$||e|| = \sqrt{e \cdot e}$$

Llamaremos distancia de e a e', que denotaremos $d(e,e')$ a:

$$d(e,e') := ||e - e'||$$

Observemos que $||\lambda \cdot e|| = \sqrt{(\lambda \cdot e)\cdot(\lambda \cdot e)} = \sqrt{\lambda^2 \cdot (e \cdot e)} = |\lambda| \cdot ||e||$, para todo $\lambda \in \mathbb{R}$ y $e \in E$. Por tanto, dado un vector no nulo e, se cumple que $v = e/||e||$ es un vector de norma 1 y $v \cdot v = 1$.

4.3. Método de ortonormalización de Gram-Schmidt

Sea $\{e_1,\dots,e_n\}$ un conjunto de vectores linealmente independientes de un espacio vectorial euclídeo $(E,\cdot)$. Vamos a definir n vectores ortonormales $\{v_1,\dots,v_n\}$ de modo que $\langle e_1,\dots,e_i\rangle = \langle v_1,\dots,v_i\rangle$, para todo $i \leq n$, mediante el proceso de ortonormalización de Gram-Schmidt:

Sea $v_1 = \frac{e_1}{||e_1||}$. Observemos que $v_1 \cdot v_1 = 1$. Además, $\langle e_1\rangle = \langle v_1\rangle$.

Sea $v_2' = e_2 - (v_1 \cdot e_2)\cdot v_1$, se cumple que $v_1 \cdot v_2' = 0$ y además $\langle e_1, e_2\rangle = \langle v_1, v_2'\rangle$. Por tanto, si definimos $v_2 = v_2'/||v_2'||$ tendremos que $\{v_1, v_2\}$ son ortonormales y $\langle e_1, e_2\rangle = \langle v_1, v_2\rangle$.

Sea $v_3' = e_3 - (v_1 \cdot e_3)\cdot v_1 - (v_2 \cdot e_3)\cdot v_2$, se cumple que $v_1 \cdot v_3' = 0$ y $v_2 \cdot v_3' = 0$ y además $\langle e_1, e_2, e_3\rangle = \langle v_1, v_2, v_3'\rangle$. Por tanto, si definimos $v_3 = v_3'/||v_3'||$ tendremos que $\{v_1, v_2, v_3\}$ son ortonormales y $\langle e_1, e_2, e_3\rangle = \langle v_1, v_2, v_3\rangle$.

Así sucesivamente vamos construyendo $\{v_1,\dots,v_n\}$.

1. Teorema: *En todo espacio vectorial euclídeo existen bases ortonormales.*

Demostración. Sea $\{e_1 \dots, e_n\}$ una base del espacio vectorial euclídeo E. Por el proceso de ortonormalización de Gram-Schmidt obtenemos una base ortonormal $\{v_1,\dots,v_n\}$. □

2. Proposición: *Sea T_2 una aplicación bilineal simétrica de un $\mathbb{R}$-espacio vectorial E y $A \in M_n(\mathbb{R})$ la matriz asociada en una base. T_2 es un producto escalar $\iff$ todos los menores principales de A son positivos.*

Demostración. $\Rightarrow$) Sabemos que existe una base ortonormal. Si P es la matriz de cambio de base, entonces

$$A = P \cdot \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix} \cdot P^t$$

y $|A| = |P|^2 > 0$. Por tanto, el determinante de toda matriz asociada a un producto escalar es positivo.

La aplicación bilineal T_2 restringida a $\langle e_1, \dots, e_i \rangle$ es un producto escalar y su matriz asociada es la submatriz de $A = (a_{rs})$, $(a_{rs})_{r,s \le i}$. Por tanto, $|(a_{rs})_{r,s \le i}| > 0$. Es decir, todos los menores principales de A son positivos.

$\Leftarrow$) Procedemos por inducción sobre n. Si $n = 1$ entonces $A = (\lambda)$ y $|A) = \lambda > 0$, luego T_2 es un producto escalar.

Supongamos que el teorema es cierto para matrices de orden $\le n-1$.

Si $A = (a_{ij})$ es la matriz asociada a T_2 (en la base usual $B = \{e_1, \dots, e_n\}$ de $\mathbb{R}^n$) entonces la matriz asociada a la restricción de T_2 a $\langle e_1, \dots, e_{n-1} \rangle$ es $(a_{rs})_{r,s \le n-1}$. Todos los menores principales de $(a_{rs})_{r,s \le n-1}$ son positivos. Por hipótesis de inducción la restricción de T_2 a $\{e_1, \dots, e_{n-1}\}$ es definida positiva. Por tanto, existe una base $\{e'_1, \dots, e'_{n-1}\}$ de $\langle e_1, \dots, e_{n-1} \rangle$ tal que

$$T_{2|\langle e_1, \dots, e_{n-1} \rangle} \equiv \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_{n-1} \end{pmatrix}$$

con $\lambda_1, \dots, \lambda_{n-1} > 0$. Sea $e'_n = e_n - \sum_i \frac{T_2(e_n, e'_i)}{\lambda_i} e'_i$. La matriz de T_2 en la base $B' = \{e'_1, \dots, e'_n\}$ es

$$T_2 \equiv \begin{pmatrix} \lambda_1 & & & \\ & \ddots & & \\ & & \lambda_{n-1} & \\ & & & T_2(e'_n, e'_n) \end{pmatrix}$$

Si Q es la matriz de cambio de base de B a B' tenemos que

$$Q^t \begin{pmatrix} \lambda_1 & & & \\ & \ddots & & \\ & & \lambda_{n-1} & \\ & & & T_2(e'_n, e'_n) \end{pmatrix} Q = A$$

y $0 < |A| = \lambda_1 \cdots \lambda_{n-1} \cdot T_2(e'_n, e'_n) \cdot |Q|^2$. Luego, $T_2(e'_n, e'_n) > 0$ y T_2 es un producto escalar. □

3. Descomposición de Cholesky: *Siguiendo las notaciones observemos que* $e_i \in \langle e_1, \dots, e_i \rangle = \langle v_1, \dots, v_i \rangle$. *Luego la matriz de cambio de base* R *de* $\{e_1, \dots, e_n\}$ *a* $\{v_1, \dots, v_n\}$ *es triangular. La matriz asociada al producto escalar en* $\{v_1, \dots, v_n\}$ *es* Id. *Por tanto, si* A *es la matriz de asociada al producto escalar en la base* $\{e_1, \dots, e_n\}$ *tendremos que*

$$A = R^t \cdot \mathrm{Id} \cdot R = R^t \cdot R$$

(con R *triangular).*

Luego toda matriz simétrica definida positiva se puede escribir como el producto de una matriz triangular por su transpuesta.

Supongamos ahora que tenemos dos bases ortonormales $B = \{e_1, \dots, e_n\}$ y $B' = \{e'_1, \dots, e'_n\}$ y sea $Q = (q_{ij})$ la matriz de cambio de base de la base B' a B. Entonces

$$\mathrm{Id} = Q^t \cdot \mathrm{Id} \cdot Q$$

por tanto, $Q^t \cdot Q = \mathrm{Id}$ (las matrices cuadradas que cumplen esta igualdad se denominan matrices ortogonales).

Recíprocamente, sea Q una matriz ortogonal que sea la matriz de cambio de base de una base ortonormal $B = \{e_1, \dots, e_n\}$ en otra base $C = \{v_1, \dots, v_n\}$. Entonces la matriz asociada al producto escalar en la base C es $C = (Q^{-1})^t \, \mathrm{Id} \cdot Q^{-1} = Q \cdot Q^{-1} = \mathrm{Id}$. Luego, la base $\{v_1, \dots, v_n\}$ es ortonormal.

En conclusión, las matrices ortogonales están caracterizadas por aplicar bases ortonormales en bases ortonormales.

4. Factorización QR: *Sea* $B = \{e_1, \dots, e_n\}$ *la base ortonormal usual de* $\mathbb{R}^n$. *Sea* $A \in M_n(\mathbb{R})$ *una matriz invertible. Sea* $C = \{v_1, \dots, v_n\}$ *las columnas de* A *(*A *es la matriz de cambio de base de* C *a* B*). A partir de* C*, por el proceso de ortonormalización de Gram-Schmidt, definamos la base ortonormal* $C' = \{v'_1, \dots, v'_n\}$. *Sea* R *la matriz triangular de cambio de base de* C *a* C' *y* Q *la matriz ortogonal de cambio de base de* C' *a* B *entonces*

$$A = Q \cdot R$$

5. Observación: Observemos que si $A = QR$ entonces la descomposición de Cholesky de $A^t A$ es

$$A^t A = (R^t Q^t) Q R = R^t R$$

4.4. Subespacio ortogonal y proyección ortogonal

1. Definición: Sea $(E,\cdot)$ un espacio vectorial euclídeo. Dado un subespacio vectorial $L \subseteq E$ se define el ortogonal de L, que denotaremos $L^\perp$ como

$$L^\perp := \{e \in E : e \cdot l = 0, \forall l \in L\}$$

$L^\perp$ es un subespacio vectorial de E: Dados $v_1, v_2 \in L^\perp$, entonces $(v_1+v_2)\cdot l = v_1 \cdot l + v_2 \cdot l = 0$, para todo $l \in L$, luego $v_1 + v_2 \in L^\perp$. Dados $\lambda \in \mathbb{R}$ y $v \in L^\perp$, entonces $(\lambda \cdot v)\cdot l = \lambda \cdot (v \cdot l) = 0$ para todo $l \in L$, luego $\lambda \cdot v \in L^\perp$.

2. Proposición: *Se cumplen las siguientes propiedades:*

1. $E^\perp = \{0\}$ *y* $\{0\}^\perp = E$.
2. *Si* $L \subseteq L'$ *entonces* $L^\perp \supseteq L'^\perp$.
3. $(L^\perp)^\perp = L$.
4. $E = L \oplus L^\perp$ *(en particular,* $L \cap L^\perp = \{0\}$*).*
5. $\dim E = \dim L + \dim L^\perp$.
6. $(L+L')^\perp = L^\perp \cap L'^\perp$ *y* $(L \cap L')^\perp = L^\perp + L'^\perp$.

Demostración. 1. Si $e \in E^\perp$ entonces $e \cdot e = 0$, luego $e = 0$ y $E^\perp = 0$. Obviamente, $\{0\}^\perp = E$.

2. Obvio.

3. 4. y 5. Sea $\langle e_1, \dots, e_n \rangle$ una base ortonormal de L. Ampliando, sea $\langle e_1, \dots, e_n, e'_1, \dots, e'_m \rangle$ una base de E. Por el proceso de ortonormalización de Gram-Schmidt, obtenemos una base ortonormal

$$\{e_1, \dots, e_n, e''_1, \dots, e''_m\}$$

de E. Ahora ya es claro que $L^\perp = \langle e''_1, \dots, e''_m \rangle$ y $(L^\perp)^\perp = \langle e_1, \dots, e_n \rangle = L$.

Además, $E = L \oplus L^\perp$ y tomando dimensiones

$$\dim E = \dim L + \dim L^\perp$$

6. Obviamente, $(L+L')^\perp = L^\perp \cap L'^\perp$. Por 3., $(L^\perp \cap L'^\perp)^\perp = L + L'$. Denotemos, $N = L^\perp$ y $N' = L'^\perp$. Entonces, tenemos que $(N \cap N')^\perp = N^\perp + N'^\perp$. Hemos terminado, porque todo subespacio es ortogonal de un subespacio (efectivamente, de su ortogonal).

□

Dado un espacio vectorial euclídeo $(E,\cdot)$ y un subespacio $L \subseteq E$ sabemos que $E = L \oplus L^{\perp}$. Por tanto, dado $e \in E$ tenemos que $e = e_1 + e_2$ de modo único con $e_1 \in L$ y $e_2 \in L^{\perp}$.

La aplicación lineal

$$\begin{array}{rcl} E = L \oplus L^{\perp} & \xrightarrow{\pi} & L \\ e = e_1 + e_2 & \mapsto & e_1 \end{array}$$

se dice que es la proyección ortogonal de E sobre L. Es claro que $\operatorname{Ker} \pi = L^{\perp}$ y $\pi_{|L} = \mathrm{Id}$ y que estas dos igualdades determinan π.

Sea $\{l_1, \dots, l_m\}$ una base ortonormal de L. El morfismo

$$\phi \colon E \to L,\ \phi(e) = \sum_i (e \cdot l_i) l_i$$

coincide con π: Sea $v = \sum_i (e \cdot l_i) l_i$. como $(\phi(e) - v) \cdot l_i = e \cdot l_i - v \cdot l_i = 0$, para todo i, entonces $\phi(e) - v = 0$ y $\phi(e) = v$.

3. Proposición: *Sea $\pi : E \to L$ la protección ortogonal de E en L. Dado $e \in E$ se cumple que $\pi(e) \in L$ es el vector de L más cercano a e.*

Demostración. Tenemos que $e = e_1 + e_2$, $e_1 \in L$, $e_2 \in L^{\perp}$ y $\pi(e) = e_1$. Podemos escribir $L = \langle e_1 \rangle \oplus \langle e_1 \rangle^{\perp}$ (denotamos por $\langle e_1 \rangle^{\perp}$ el subespacio de L ortogonal a $\langle e_1 \rangle$). Dado $l \in L$ tendremos que $l = \lambda e_1 + l_2$, con $l_2 \in \langle e_1 \rangle^{\perp}$. Entonces,

$$\begin{aligned} d(l,e)^2 &= (l - e) \cdot (l - e) = ((\lambda - 1) e_1 + l_2 - e_2) \cdot ((\lambda - 1) e_1 + l_2 - e_2) \\ &= (\lambda - 1)^2 \cdot (e_1 \cdot e_1) + l_2 \cdot l_2 + e_2 \cdot e_2 \geq e_2 \cdot e_2 = d(e_1, e)^2 \end{aligned}$$

y $d(l,e) = d(e_1,e)$, si y sólo si $\lambda = 1$ y $l_2 = 0$, es decir, $l = e_1$. □

Sea $B = \{e_1, \dots, e_n\}$ una base de E y A la matriz asociada al producto escalar en la base B. Sea $B' = \{l_1, \cdots, l_m\}$ una base ortonormal de L y Q la matriz cuyas columnas son las coordenadas de los vectores $\{l_i\}$ respecto de B. Calculemos la matriz de la proyección ortogonal a L, en las bases B, B': Dado $e \in E$, $\pi(e) = \sum_i (e \cdot l_i) \cdot l_i$. Si $e = (x_1, \dots, x_n)$ en la base B,

$$\pi((x_1, \dots, x_n)) = (x_1, \dots, x_n) \cdot A \cdot Q = (y_1, \dots, y_m)$$

en la base B'. Es decir, la matriz de π en las bases B, B' es $Q^t \cdot A$.

Por tanto, la matriz P de la composición,

$$E \xrightarrow{\pi} L \overset{i}{\hookrightarrow} E$$

en la base B, es $P = Q \cdot Q^t \cdot A$.

Sea ahora $B' = \{e'_1, \ldots, e'_m\}$ una base de L y $Q \in M_{n\times m}(\mathbb{R})$ la matriz cuyas columnas son las coordenadas de los vectores e'_i en la base B, es decir, Q es la matriz asociada asociada a la inclusión $L \hookrightarrow E$ en las bases B' y B.

Calculemos la matriz P asociada a la aplicación lineal $\pi\colon E \to L$ en las bases B y B':

Dado $v \in E$ sabemos que $v - \pi(v) \in L^\perp$, o dicho de otro modo,

$$v \cdot l = \pi(v) \cdot l,\ \forall l \in L$$

Sea $v = (x_1, \ldots, x_n)$ en la base B y $l = (y_1, \ldots, y_m)$ en la base B'. Entonces

$$v \cdot l = (x_1, \ldots, x_n) \cdot A \cdot Q \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

Por otra parte, $\pi(v)^t = Q \cdot P \cdot (x_1, \ldots, x_n)^t$ en la base B. Por tanto,

$$\pi(v) \cdot l = (x_1, \ldots, x_n) P^t Q^t A Q \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

Luego, $AQ = P^tQ^tAQ$, luego $P^t = AQ(Q^tAQ)^{-1}$ (Q^tAQ es la matriz asociada al producto escalar de L en la base B') y la matriz de la composición $E \xrightarrow{\pi} L \xrightarrow{i} E$ en la base B es

$$P = (Q^tAQ)^{-1}(Q^tA).$$

La matriz de la composición $E \xrightarrow{\pi} L \xhookrightarrow{i} E$ en la base B es

$$Q(Q^tAQ)^{-1}(Q^tA).$$

4.5. Inversas generalizadas

Sea $T\colon \mathbb{R}^n \to \mathbb{R}^m$ una aplicación lineal (entre espacios vectoriales euclídeos). $\mathbb{R}^n = (\operatorname{Ker} T)^\perp \oplus \operatorname{Ker} T$ y $\mathbb{R}^m = \operatorname{Im} T \oplus (\operatorname{Im} T)^\perp$ y con esta descomposición

$$T = \begin{pmatrix} T_1 & 0 \\ 0 & 0 \end{pmatrix}$$

donde $T_1\colon (\operatorname{Ker} T)^\perp \simeq \operatorname{Im} T$, es la restricción de T a $(\operatorname{Ker} T)^\perp$.

1. Definición: Se dice que $T^-\colon \mathbb{R}^m \to \mathbb{R}^n$ es una inversa generalizada de T si y sólo si TT^- es la identidad sobre $\operatorname{Im} T$, es decir, $TT^-T = T$.

Siguiendo notaciones, si $T = \begin{pmatrix} T_1 & 0 \\ 0 & 0 \end{pmatrix}$ entonces $T^- = \begin{pmatrix} T_1^{-1} & B \\ C & D \end{pmatrix}$, con B, C, D cualesquiera.

2. Definición: Se dice que $T^\square \colon \mathbb{R}^m \to \mathbb{R}^n$ es una inversa de mínimos cuadrados de T si $TT^\square$ es la identidad sobre $\operatorname{Im} T$ y es nulo sobre $(\operatorname{Im} T)^\perp$, es decir, sólo si $TT^\square T = T$ y $T^\square((\operatorname{Im} T)^\perp) \subseteq \operatorname{Ker} T$.

Con la descomposición, $\mathbb{R}^n = (\operatorname{Ker} T)^\perp \oplus \operatorname{Ker} T$ y $\mathbb{R}^m = \operatorname{Im} T \oplus (\operatorname{Im} T)^\perp$,

$$T^\square = \begin{pmatrix} T_1^{-1} & 0 \\ C & D \end{pmatrix}$$

3. Definición: Diremos que $T^+ \colon \mathbb{R}^m \to \mathbb{R}^n$ es la inversa de Moore-Penrose de T si $T^+_{|(\operatorname{Im} T)} = T_1^{-1}$ y $T^+_{|(\operatorname{Im} T)^\perp} = 0$.

Con la descomposición, $\mathbb{R}^n = (\operatorname{Ker} T)^\perp \oplus \operatorname{Ker} T$ y $\mathbb{R}^m = \operatorname{Im} T \oplus (\operatorname{Im} T)^\perp$,

$$T^+ = \begin{pmatrix} T_1^{-1} & 0 \\ 0 & 0 \end{pmatrix}$$

4.6. Aplicación a sistemas de ecuaciones lineales

Dados $A \in M_{m\times n}(\mathbb{R})$ y $b \in \mathbb{R}^m$ consideremos el sistema de ecuaciones lineales

$$Ax = b$$

1. Proposición: *Sea A^- una inversa generalizada de A. El sistema de ecuaciones $Ax = b$ es compatible si y sólo si $AA^-b = b$, en cuyo caso A^-b es una solución particular del sistema de ecuaciones.*

Demostración. Obviamente si $AA^-b = b$ entonces A^-b es una solución particular del sistema de ecuaciones y el sistema es compatible.

Si el sistema es compatible entonces $b = Ac$ para cierto c. Luego, $AA^-b = AA^-Ac = Ac = b$.

□

2. Lema: *Sea $A \in M_{m\times n}(\mathbb{R})$ una matriz y $A^- \in M_{n\times m}(\mathbb{R})$ una inversa generalizada. Entonces,*

$$\operatorname{Ker} A = \{(\operatorname{Id} - A^-A)y,\ \forall\, y \in \mathbb{R}^n\}$$

Demostración. $\operatorname{Ker} A \subseteq \{(\operatorname{Id} - A^-A)y \colon y \in \mathbb{R}^n\}$: Dado $y' \in \operatorname{Ker} A$ tenemos que $(\operatorname{Id} - A^-A)y' = y'$, luego $y' \in \{(\operatorname{Id} - A^-A)y \colon y \in \mathbb{R}^n\}$. $\{(\operatorname{Id} - A^-A)y \colon y \in \mathbb{R}^n\} \subseteq \operatorname{Ker} A$: Dado $y' = (\operatorname{Id} - A^-A)y$, tenemos que $Ay' = (A - AA^-A)y = (A - A)y = 0$, luego $y' \in \operatorname{Ker} A$. □

3. Teorema: *Si el sistema $Ax = b$ es compatible y A^- es una inversa generalizada de A, entonces*

$$\{\textit{Soluciones de } Ax = b\} = \{A^-b + (\mathrm{Id} - A^-A)y,\ \forall\, y \in \mathbb{R}^n\}$$

Demostración. {Soluciones de $Ax = b$} = {Sol. particular+Ker$A$}. Como A^-b es una solución particular y $\operatorname{Ker} A = \{(\mathrm{Id} - A^-A)y\colon y \in \mathbb{R}^n\}$, hemos concluido. □

4. Definición: Se dice que $\tilde{x} \in \mathbb{R}^n$ es una solución aproximada mínimo cuadrática del sistema de ecuaciones $Ax = b$ si $d(A\tilde{x}, b) \leq d(Ay, b)$, para todo $y \in \mathbb{R}^n$.

5. Proposición: *Sea $A^\square \in M_{n\times m}(\mathbb{R})$ una inversa mínimo cuadrática de $A \in M_{m\times n}(\mathbb{R})$ y $b \in \mathbb{R}^m$. Entonces $A^\square b$ es una solución aproximada mínimo cuadrática del sistema $Ax = b$.*

Demostración. $\mathbb{R}^m = \operatorname{Im} A \oplus (\operatorname{Im} A)^\perp$ y $b = b_1 + b_2$, con $b_1 = Ac \in \operatorname{Im} A$ y $b_2 \in (\operatorname{Im} A)^\perp$. Observemos que $b_1 \in \operatorname{Im} A$, es el vector de $\operatorname{Im} A$ más cercano a b. Por tanto, $d(b_1, b) \leq d(Ay, b)$, para todo $y \in \mathbb{R}^n$. Sabemos que $A^\square(b) = A^\square(b_1) + A^\square(b_2)$ y $A^\square(b_2) \in \operatorname{Ker} A$. Por tanto, $A(A^\square(b)) = AA^\square(b_1) = AA^\square A(c) = Ac = b_1$. Por tanto, $A^\square(b)$ es una solución aproximada mínimo cuadrática de $Ax = b$. □

6. Teorema: *Si el sistema $Ax = b$ es compatible y $A^\square$ es una inversa mínimo cuadrática de A, entonces*

$$\{\textit{Soluciones aproximadas mínimo cuadráticas de } Ax = b\} = \{A^\square b + (\mathrm{Id} - A^\square A)y,\ \forall\, y \in \mathbb{R}^n\}$$

Demostración. $A^\square(b)$ es una solución mínimo cuadrática, luego $A(A^\square(b)) =: b_1$ es el vector de $\operatorname{Im} A$ más cercano a b. El vector $e \in \mathbb{R}^n$ es otra solución mínimo cuadrática si y sólo si $A(e) = b_1 = A(A^\square(b))$, es decir, si y sólo si $A(e - A^\square(b)) = 0$. Por tanto, el vector $e \in \mathbb{R}^n$ es otra solución mínimo cuadrática si y sólo si $e - A^\square(b) \in \operatorname{Ker} A$, es decir, $e \in A^\square(b) + \operatorname{Ker} A$. Hemos concluido por 4.6.2. □

7. Definición: Se dice que $x^+ \in \mathbb{R}^n$ es la solución óptima mínimo cuadrática del sistema de ecuaciones $Ax = b$ si es la solución aproximada mínimo cuadrática de norma (euclidea) mínima.

8. Proposición: *Si el sistema $Ax = b$ es compatible, entonces $x^+ = A^+(b)$ es la solución aproximada óptima mínimo cuadrática del sistema $Ax = b$.*

Demostración. $A^+(b)$ es una solución mínimo cuadrática. Por el teorema anterior y 4.6.2, cualquier otra solución mínimo cuadrática es de la forma $e = A^+(b) + d$, con $d \in \operatorname{Ker} A$. Como $A^+(b) \in (\operatorname{Ker} A)^\perp$ tenemos que $||e||^2 = ||A^+(b)||^2 + ||d||^2$. Luego $A^+(b)$ es la solución aproximada óptima mínimo cuadrática del sistema $Ax = b$. □

4.7. Cálculo de la inversa de Moore-Penrose

Sea $T\colon \mathbb{R}^n \to \mathbb{R}^m$ una aplicación lineal (entre los espacio vectoriales euclídeo usuales) de matriz asociada A. Queremos calcular la matriz A^+ asociada a T^+.

Observemos que $(\operatorname{Ker} T)^\perp$ está generado por las filas de A: El vector $(x_1,\dots,x_n) \in \mathbb{R}^n$ pertenece a $\operatorname{Ker} T$ si y sólo si

$$A \cdot (x_1,\dots,x_n)^t = 0$$

es decir, si y sólo si $(x_1,\dots,x_n)$ es ortogonal a las filas de A. En conclusión,

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

Es decir, $\operatorname{Ker} T$ es el espacio ortogonal al espacio generado por las filas de A.

Sea r el rango de T y $\{a_1,\dots,a_r\}$ r filas de A linealmente independientes. Sea Ar la matriz de columnas los vectores $\{a_1,\dots,a_r\}$. Una base de $\operatorname{Im} T$ es $\{a'_1 = T(a_1),\dots,a'_r = T(a_r)\}$. Sea $\{l_1,\dots,l_r\}$ una base ortonormal de $\operatorname{Im} T$ (obtenida, por ejemplo por Gram-Schmidt). Sea R la matriz de cambio de base de $\{l_1,\dots,l_r\}$ a $\{a'_1,\dots,a'_r\}$ y $Q = A \cdot Ar \cdot R$ la matriz de columnas las coordenadas de los vectores l_i en la base usual de $\mathbb{R}^m$.

La matriz de la proyección de $\mathbb{R}^m$ en $\operatorname{Im} T$, $\mathbb{R}^m \to \operatorname{Im} T$, $v \mapsto \sum_i (v\cdot l_i)\cdot l_i$, en las bases la usual de $\mathbb{R}^m$ y $\{l_1,\dots,l_r\}$ de $\operatorname{Im} T$ es Q^t, en la base $\{T(a_1) = a'_1,\dots,T(a_r) = a'_r\}$ es $R\cdot Q^t$. Por tanto, la matriz de T^+ en las bases usuales de $\mathbb{R}^m$ y $\mathbb{R}^n$ es

$$A^+ = Ar \cdot R \cdot Q^t = (Ar\cdot R)\cdot (Ar\cdot R)^t \cdot A^t$$

4.8. Matrices hermíticas

Sea E un $\mathbb{C}$-espacio vectorial (por ejemplo, $E = \mathbb{C}^n$).

1. Definición: Diremos que una aplicación $T_2\colon E\times E \to \mathbb{C}$ es una forma hermítica si para todo $u,u',v,v' \in E$ y $\lambda \in \mathbb{C}$, cumple que

1. $T_2(u+u',v) = T_2(u,v) + T_2(u',v)$.
2. $T_2(u,v+v') = T_2(u,v) + T_2(u,v')$.
3. $T_2(\lambda u,v) = \bar\lambda T_2(u,v)$ ($\bar\lambda$ es el número complejo conjugado de λ).
4. $T_2(u,\lambda v) = \lambda T_2(u,v)$.
5. $T_2(u,v) = \overline{T_2(v,u)}$.

(que es el concepto equivalente en espacios vectoriales complejos al concepto de aplicación bilineal simétrica que dábamos en espacios vectoriales reales).

Observemos que $T_2(u,u) = \overline{T_2(u,u)}$, luego $T_2(u,u) \in \mathbb{R}$.

Sea $B = \{e_1,\dots,e_n\}$ una base de E y $A = (a_{ij})$ la matriz asociada a T_2 en esta base

$$a_{ij} := T_2(e_i,e_j)$$

Si T_2 es hermítica, entonces $a_{ij} = T_2(e_j,e_i) = \overline{T_2(e_i,e_j)} = \bar{a}_{ji}$. Es decir, $(a_{ij}) = \overline{(a_{ij})^t} =: (a_{ij})^*$. Es decir, (a_{ij}) es una matriz hermítica.

Dados $e = \sum_i x_i e_i \equiv (x_1,\dots,x_n)$ y $e' = \sum_i y_i e_i \equiv (y_1,\dots,y_n)$, entonces

$$T_2(e,e') = T_2(\sum_i x_i e_i, \sum_j y_j e_j) = \sum_{ij} \bar{x}_i y_j T_2(e_i,e_j) = \sum_{ij} \bar{x}_i a_{ij} y_j = \overline{(x_1,\dots,x_n)}(a_{ij}) \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

Sea $B' = \{e'_1,\dots,e'_n\}$ otra base de E y $P = (p_{ij})$ la matriz de cambio de base de B' a B. Sea (a'_{ij}) la matriz asociada a T_2 en la base B'. Entonces

$$a'_{ij} = T_2(e'_i,e'_j) = \overline{(0,\overset{i}{\dots},1,\dots,0)(p_{ij})^t}(a_{ij})(p_{ij}) \begin{pmatrix} 0 \\ \vdots j \\ 1 \\ \vdots \\ 0 \end{pmatrix} = (P^*AP)_{ij}$$

Luego, $A' = P^*AP$.

2. Definición : Diremos que la forma hermítica T_2 es un producto escalar si $T_2(u,u) \in \mathbb{R}^+$, para todo $u \neq 0$. De nuevo denotaremos T_2 por $\cdot$ y diremos que $(E,\cdot)$ es un espacio vectorial complejo unitario.

$\mathbb{C}^n$ con el producto escalar usual $(x_1,\dots,x_n)\cdot(y_1,\dots,y_n) = \sum_i \bar{x}_i y_i$ es un espacio vectorial complejo unitario. La base usual de $\mathbb{C}^n$ es una base ortonormal.

Del mismo modo que en espacios vectoriales euclídeos, tenemos el proceso de ortonormalización de Gram-Schmidt en espacios vectoriales complejos unitarios. Por tanto, en todo espacio vectorial unitario existen bases ortonormales.

3. Definición : Diremos que $(a_{ij}) \in M_n(\mathbb{C})$ es una matriz unitaria si

$$(a_{ij})^*(a_{ij}) = \mathrm{Id}$$

Sea $L\colon \mathbb{C}^n \to \mathbb{C}^n$ una aplicación lineal de matriz asociada (a_{ij}). De nuevo, (a_{ij}) es unitaria si y sólo si T aplica bases ortonormales en base ortonormales (es decir, conserva distancias).

Del mismo modo que en espacios vectoriales euclideos, toda matriz (a_{ij}) tiene su descomposición UR, es decir, es el producto de una matriz unitaria por una triangular.

4. Factorización de Schur: *Sea* $A \in M_n(\mathbb{C})$*. Entonces, existe una matriz unitaria* $U \in M_n(\mathbb{C})$ *tal que*

$$U^*AU = T, \quad (T \text{ triangular}).$$

Demostración. Existe $P \in M_n(\mathbb{C})$ tal que

$$P^{-1}AP = J, \quad (J \text{ matriz de Jordan}).$$

Sea $P = UR$ una descomposición UR. Entonces,

$$J = (UR)^{-1}A(UR) = R^{-1}(U^{-1}AU)R$$

Luego, $U^{-1}AU = T$, (con $T = RJR^{-1}$ triangular). □

4.9. Problemas

1. Sobre $\mathbb{R}^3$ consideramos una aplicación bilineal $T_2 : \mathbb{R}^3 \times \mathbb{R}^3 \longrightarrow \mathbb{R}$ cuya matriz asociada respecto de la base usual de $\mathbb{R}$ es $A \in \mathscr{M}_3(\mathbb{R})$. Determinar si T_2 es simétrica cuando A es:

$$(a)\begin{pmatrix} -1 & 2 & 3 \\ 2 & 4 & 1 \\ 3 & 1 & 5 \end{pmatrix}, \quad (b)\begin{pmatrix} 1 & 1 & -1 \\ 1 & 2 & 1 \\ -1 & 1 & 6 \end{pmatrix}, \quad (c)\begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 6 \\ 1 & 1 & 4 \end{pmatrix}.$$

2. Hallar la matriz respecto de la base usual $\mathbb{R}^3$ de la métrica simétrica $T_2 : \mathbb{R}^3 \times \mathbb{R}^3 \longrightarrow \mathbb{R}$ definida por

$$\begin{array}{lllllllll} T_2(\mathbf{u}_1,\mathbf{u}_1) & = & 5; & T_2(\mathbf{u}_1,\mathbf{u}_2) & = & 0; & T_2(\mathbf{u}_1,\mathbf{u}_3) & = & -1; \\ & & & T_2(\mathbf{u}_2,\mathbf{u}_2) & = & 1; & T_2(\mathbf{u}_2,\mathbf{u}_3) & = & 4; \\ & & & & & & T_2(\mathbf{u}_3,\mathbf{u}_3) & = & 0; \end{array}$$

donde $\mathbf{u}_1 = (1,2,1)$, $\mathbf{u}_2 = (-1,2,0)$ y $\mathbf{u}_3 = (1,0,1)$.

3. Sean V un $\mathbb{R}$-espacio vectorial de dimensión $n > 0$ y T_2 una aplicación bilineal sobre V. Probar que si $\mathscr{B} = \{\mathbf{v}_1,\dots,\mathbf{v}_n\}$ es una base de V tal que $T_2(\mathbf{v}_i,\mathbf{v}_j) = \delta_{ij}$, donde δ_{ij} es la función Delta de Kronecker, entonces T_2 es un producto escalar sobre V.

4. Sean V un espacio vectorial euclídeo y $L \subseteq V$ un subespacio de V. Probar que la restricción del producto escalar de V a L dota a este último de estructura de espacio vectorial euclídeo; es decir, *todo subespacio vectorial de un espacio vectorial euclídeo hereda una estructura natural de espacio vectorial euclídeo.*

5. Aplicar el método de Gram-Schmidt para calcular bases ortonormales, a partir de la bases dadas en los siguientes espacios vectoriales euclídeos:

    1. $\{(1,1,1),(0,1-1),(0,2,0)\}$ en $\mathbb{R}^3$, con el producto escalar usual.
    2. $\{1,x,x^2\}$ en el espacio V de los polinomios de $\mathbb{R}[x]$ con grado menor o igual que 2, y el producto escalar $T_2(P,Q) = P(0)Q(0) + P(1)Q(1) + P(2)Q(2)$.
    3. $\{1,x,x^2\}$ en el espacio V de los polinomios de $\mathbb{R}[x]$ con grado menor o igual que 2, y el producto escalar $\bar{T}_2(P,Q) = \int_0^1 P(x)Q(x)dx$.

6. En el $\mathbb{R}$-espacio vectorial $\mathbb{R}^2$ consideramos la aplicación

$$\begin{array}{rcl} T_2 : \mathbb{R}^2 \times \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ ((x_1,y_1),(x_2,y_2)) & \longmapsto & T_2((x_1,y_1),(x_2,y_2)) = x_1x_2 + x_1y_2 + x_2y_1 + 2y_1y_2. \end{array}$$

    1. Probar que T_2 es un producto escalar.
    2. Obtener una base ortonormal del espacio vectorial euclídeo $(\mathbb{R}^2, T_2)$.

7. Sobre $\mathbb{R}^3$ se considera la aplicación bilineal T_2 cuya matriz en la base usual es

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

    1. Probar que T_2 es un producto escalar.
    2. Hallar una base ortonormal del espacio vectorial euclídeo $(\mathbb{R}^3, T_2)$.
    3. Calcular el módulo de $\mathbf{v} = (1,3,-2)$ y el ángulo que forman los vectores $\mathbf{u}_1 = (1,-2-2)$ y $\mathbf{u}_2 = (2,1,0)$ para el producto escalar T_2.

8. Sobre $\mathbb{R}^3$ consideramos la aplicación bilineal T_2 cuya matriz en la base $\mathscr{B}$ usual es

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

    1. Probar que T_2 es producto escalar.
    2. Hallar una base ortonormal del espacio vectorial euclídeo $(\mathbb{R}^3, T_2)$.
    3. Calcular el módulo del vector $\mathbf{v} \in \mathbb{R}^3$ de coordenadas $(1,0,2)$ respecto de $\mathscr{B}$. Calcular el ángulo que forman el vector $\mathbf{v}$ con el vector $\mathbf{u}$ de coordenadas $(1,0,0)$ respecto de $\mathscr{B}$.

9. Sea V el espacio vectorial de las matrices simétricas reales de orden dos.

1. Hallar una base de V.
2. Probar que la aplicación

$$\begin{array}{ccc} V \times V & \longrightarrow & \mathbb{R} \\ (A,B) & \longmapsto & A \cdot B = \operatorname{tr}(AB) \end{array}$$

es un producto escalar sobre V y obtener su matriz en la base hallada en el apartado (a).

3. Calcular una base ortonormal de V para el producto escalar anterior.

10. Consideremos en $\mathbb{R}^3$ el producto escalar T_2 que en la base $B = \{\mathbf{v}_1 = (1,1,0), \mathbf{v}_2 = (1,0,1), \mathbf{v}_3 = (0,1,1)\}$ tiene matriz

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

1. Calcular una base ortonormal para T_2.
2. Escribir la matriz de T_2 en la base usual de $\mathbb{R}^3$.
3. Calcular las ecuaciones, en la base B, del subespacio ortogonal al plano π que en la base usual tiene ecuación $z = 0$.
4. Calcular la distancia de $\mathbf{v}_1$ a π.

11. Consideremos en $\mathbb{R}^4$ el producto escalar euclídeo

$$T_2(x, y) = 4x_1y_1 + x_1y_2 + x_2y_1 + 2x_2y_2 + x_3y_3.$$

Calcular la proyección ortogonal del vector $\mathbf{v} = (0,1,0)$ sobre el subespacio $L = \langle(1,0,0), (0,0,1)\rangle$ y determinar la distancia de $\mathbf{v}$ a L.

12. Sean V un $\mathbb{R}$-espacio vectorial de dimensión 4, $\mathscr{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ una base de V. Consideramos el producto escalar definido por

$$\begin{array}{llllllllllll} \mathbf{v}_1\cdot\mathbf{v}_1 & = & 7; & \mathbf{v}_1\cdot\mathbf{v}_2 & = & 3; & \mathbf{v}_1\cdot\mathbf{v}_3 & = & 3; & \mathbf{v}_1\cdot\mathbf{v}_4 & = & -1; \\ & & & \mathbf{v}_2\cdot\mathbf{v}_2 & = & 2; & \mathbf{v}_2\cdot\mathbf{v}_3 & = & 1; & \mathbf{v}_2\cdot\mathbf{v}_4 & = & 0; \\ & & & & & & \mathbf{v}_3\cdot\mathbf{v}_3 & = & 2; & \mathbf{v}_3\cdot\mathbf{v}_4 & = & -1; \\ & & & & & & & & & \mathbf{v}_4\cdot\mathbf{v}_4 & = & 1, \end{array}$$

que dota a V de estructura de espacio vectorial euclídeo. Dado el subespacio L de V generado por los vectores $\mathbf{u}_1 = \mathbf{v}_2 + \mathbf{v}_4$, $\mathbf{u}_2 = 2\mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3$ y $\mathbf{u}_3 = \mathbf{v}_3 - 2\mathbf{v}_4 - \mathbf{v}_5$, obtener una base de $L^\perp$.

13. Sean $V = \mathbb{R}^4$, $\mathscr{B}$ la base usual de $\mathbb{R}^4$ y T_2 la aplicación bilineal simétrica cuya matriz respecto de $\mathscr{B}$ es
$$\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}.$$
Si L es un subespacio de V, definimos
$$L^{\perp} = \{\mathbf{v} \in V \mid T_2(\mathbf{v},\mathbf{u}) = 0, \ \forall \mathbf{u} \in L\}.$$
   1. Probar $L^{\perp}$ es un subespacio de V.
   2. Hallar una base de $V^{\perp}$.
   3. Sea L el subespacio de V definido por $\{(x_1,x_2,x_3,x_4) \mid x_1 - x_4 = x_2 - x_3 = 0\}$. Comprobar que $(L^{\perp})^{\perp} \neq L$.
   4. ¿Contradicen los apartados anteriores a las propiedades vistas para el subespacio ortogonal de un subespacio de un espacio vectorial euclídeo? Justificar la respuesta.

14. Sea $\mathscr{B} = \{\mathbf{v}_1 = (1,1,0), \mathbf{v}_2 = (1,0,1), \mathbf{v}_3 = (0,1,1)\}$ una base de $\mathbb{R}^3$. Sobre $\mathbb{R}^3$ consideramos el producto escalar cuya matriz respecto de $\mathscr{B}$ es
$$A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & \sqrt{2} \\ 0 & \sqrt{2} & 3 \end{pmatrix}$$
que lo dota de estructura de espacio vectorial euclídeo.
   1. Calcular una base ortonormal de $\mathbb{R}^3$.
   2. Calcular la matriz del producto escalar respecto de la base usual de $\mathbb{R}^3$.
   3. Dado el subespacio $L = \{(x,y,z) \in \mathbb{R}^3 \mid z = 0\}$, calcular $L^{\perp}$.

15. Sobre $V = \mathscr{M}_2(\mathbb{R})$, esto es, el espacio vectorial de las matrices reales de orden 2, se considera el producto escalar dado por la igualdad $A \cdot B := \mathrm{tr}(A^{t}B)$, para cada A y $B \in \mathscr{M}_2(\mathbb{R})$.
   1. Calcular el ortogonal del subespacio L formado por las matrices diagonales de $\mathscr{M}_2(\mathbb{R})$.
   2. Determinar la proyección ortogonal de cada matriz $C \in \mathscr{M}_2(\mathbb{R})$ sobre L.

16. Sean $\mathscr{B} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ una base ortonormal de un espacio vectorial euclídeo V, L un subespacio de V, $\{\mathbf{v}_1, \dots, \mathbf{v}_r\}$ una base de L y $A \in \mathscr{M}_{n\times r}(\mathbb{R})$ la matriz cuyas columnas son las coordenadas de $\mathbf{v}_1, \dots, \mathbf{v}_r$ respecto de $\mathscr{B}$.

    1. Probar que la matriz $A^{\mathrm{t}}A$ es invertible.
    2. Dado un vector $\mathbf{v} = \lambda_1\mathbf{u}_1 + \dots + \lambda_n\mathbf{u}_n$, demostrar que las coordenadas de la proyección ortogonal de $\mathbf{v}$ sobre L respecto de $\mathscr{B}$ son
    $$A(A^{\mathrm{t}}A)^{-1}A^{\mathrm{t}}\begin{pmatrix}\lambda_1\\ \vdots \\ \lambda_n\end{pmatrix}.$$
    3. Aplicar lo anterior para calcular, en $\mathbb{R}^4$ con su producto escalar usual, la proyección ortogonal de $(-1,2,-3,-1)$ sobre $L = \langle(1,3,-2,0),(3,2,0,0)\rangle$.

17. Sea $A \in M_{n\times m}(\mathbb{R})$ una matriz de rango r. Probar que existen matrices $B \in M_r(\mathbb{R})$ invertible, y $P \in M_n(\mathbb{R})$, $Q \in M_m(\mathbb{R})$ ortogonales de modo que
$$A = Q\cdot\left(\begin{array}{c|c} B & 0 \\ \hline 0 & 0\end{array}\right)\cdot P^t$$
Probar que las r primeras columnas de P son una base (ortonormal) de $(\mathrm{Ker}A)^\perp$, las $n-r$ últimas columnas son una base (ortonormal) de $\mathrm{Ker}A$, las r primeras columnas de Q es una base (ortonormal) de $\mathrm{Im}A$ y las $m-r$ columnas de Q es una base (ortonormal) de $(\mathrm{Im}A)^\perp$. Probar que
$$A^+ = P^t\cdot\left(\begin{array}{c|c} B^{-1} & 0 \\ \hline 0 & 0\end{array}\right)\cdot Q$$

18. Sea $A \in M_{n\times m}(\mathbb{R})$. Probar

    a) $(\lambda\cdot A)^+ = \lambda^{-1}\cdot A^+$.
    b) $(A^+)^t = (A^t)^+$.
    c) $(A^+)^+ = A$.
    d) $A^+ = A^{-1}$ si A es invertible (y cuadrada).
    e) $(AA^+)^+ = AA^+$ y $(A^+A)^+ = A^+A$.
    f) $A^+A = \mathrm{Id}$ si y sólo si $\mathrm{rg}(A) = n$.
    g) $A^+ = (A^tA)^{-1}A^t$ si $\mathrm{rg}(A) = n$.

19. Sea $A \in M_{n\times m}(\mathbb{R})$. Probar que

    a) $A^+ = (A^tA)^+A^t$.

    b) La proyección ortogonal de $\mathbb{R}^m$ en $\operatorname{Im} A$ es igual a $AA^+ = A(A^tA)^+A^t$.

    c) $A^{\square} = (A^tA)^-A^t$.

20. Usar el ejercicio 18(g) para calcular la inversa de Moore-Penrose de

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

21. Sea $A = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. Probar que $(A^2)^+ \neq (A^+)^2$.

22. Consideremos la matriz

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

    1. Calcular la inversa generalizada de Moore-Penrose de AA^{t}, y usar la proposición 18(g) para hallar A^+.
    2. Usar A^+ para calcular la matriz de proyección ortogonal de $\mathbb{R}^n$ sobre $\operatorname{im}(A)$ y de $\mathbb{R}^m$ sobre $\operatorname{im}(A^{\mathrm{t}})$.

23. Sea $A \in \mathcal{M}_n(\mathbb{R})$. Probar que si A es simétrica, entonces

    1. A^+ es simétrica.
    2. $AA^+ = A^+A$.

24. Sea $A \in \mathcal{M}_{m\times n}(\mathbb{R})$. Probar que si $\operatorname{rg}(A) = 1$, entonces $A^+ = \alpha^{-1}A^{\mathrm{t}}$, donde $\alpha = \operatorname{tr}(A^+A)$.

25. Sea $A \in \mathcal{M}_{m\times n}(\mathbb{R})$. Probar que

    1. $AB = \mathbf{0}$ si, y sólo si, $B^+A^+ = \mathbf{0}$, con $B \in \mathcal{M}_{n\times p}(\mathbb{R})$.
    2. $A^+B = \mathbf{0}$ si, y sólo si, $A^{\mathrm{t}}B = \mathbf{0}$, con $B \in \mathcal{M}_{m\times p}(\mathbb{R})$.

26. Sea $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ simétrica y de rango r. Probar que si A tiene un autovalor λ no nulo de multiplicidad r, entonces $A^+ = \lambda^{-2}A$.

27. Sean $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ y $B \in \mathcal{M}_{n\times p}(\mathbb{R})$. Probar que si B tiene rango pleno por filas (es decir, $\mathrm{rg}(B) = n$), entonces
$$AB(AB)^+ = AA^+.$$

28. Sean $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ y $B \in \mathcal{M}_{n\times m}(\mathbb{R})$. Probar que $(AB)^+ = B^+A^+$ si $A^tABB^t = BB^tA^tA$.

29. Calcular la inversa de Moore-Penrose de
$$\begin{pmatrix} 2 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 4 \end{pmatrix}.$$

30. Consideremos la matriz diagonal $A = \mathrm{diag}(0,2,3)$.
    1. Hallar una inversa generalizada de A de rango 2.
    2. Hallar una inversa generalizada de A de rango 3 y que sea diagonal.
    3. Hallar una inversa generalizada de A que no sea diagonal.

31. Sea $A \in \mathcal{M}_n(\mathbb{R})$ una matriz divida por bloques de la siguiente manera
$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},$$
con $A_{11} \in \mathcal{M}_r(\mathbb{R})$. Probar que si $\mathrm{rg}(A_{11}) = \mathrm{rg}(A) = r$, entonces
$$\begin{pmatrix} A_{11}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$$
es una inversa generalizada de A.

32. Sean $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ y A^- una inversa generalizada de A. Probar que:
    1. AA^-, A^-A, $I_n - A^-A$ e $I_m - AA^-$ son idempotentes.
    2. $\mathrm{rg}(I_n - A^-A) = n - \mathrm{rg}(A)$ y $\mathrm{rg}(I_m - AA^-) = m - \mathrm{rg}(A)$.

33. Sean $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ y $B \in \mathcal{M}_{n\times p}(\mathbb{R})$. Probar que B^-A^- será una inversa generalizada de AB para cualquier elección de A^- y B^- si $\mathrm{rg}(B) = n$.

34. Sean $A \in \mathscr{M}_{m\times n}(\mathbb{R})$ y $B \in \mathscr{M}_{n\times p}(\mathbb{R})$. Probar que para cualquier elección de A^- y B^-, B^-A^- es una inversa generalizada de AB si, y sólo si, A^-BB^- es idempotente.

35. Probar que la matriz B es una inversa generalizada de A si, y sólo si, AB es idempotente y $\operatorname{rg}(A) = \operatorname{rg}(AB)$.

36. Sean $A \in \mathscr{M}_{m\times n}(\mathbb{R})$ y $B \in \mathscr{M}_{n\times m}(\mathbb{R})$. Probar que B es la inversa de Moore-Penrose de A si, y sólo si, B es una inversa mínimo cuadrática de A y A es una inversa mínimo cuadrática de B.

37. Sea $A\mathbf{x} = \mathbf{b}$ un sistema de ecuaciones compatible. Probar que si B es una inversa generalizada de A, entonces $\mathbf{x} = B\mathbf{b}$ es una solución, y para cualquier solución $\hat{\mathbf{x}}$, existe una inversa generalizada B de A, tal que $\hat{\mathbf{x}} = B\mathbf{b}$.

38. Consideremos el sistema de ecuaciones $A\mathbf{x} = \mathbf{b}$, donde $A \in \mathscr{M}_{4\times 3}(\mathbb{R})$ es la matriz de ejercicio 20 y

$$\mathbf{b} = \begin{pmatrix} 1 \\ 3 \\ -1 \\ 0 \end{pmatrix}.$$

    1. Probar que el sistema es compatible.
    2. Hallar una solución de este sistema de ecuaciones.
    3. ¿Cuántas soluciones linealmente independientes hay?

39. Consideremos el sistema de ecuaciones $AXC = B$, donde $X \in \mathscr{M}_3(\mathbb{R})$ es una matriz de incógnitas y

$$A = \begin{pmatrix} 1 & 3 & 1 \\ 3 & 2 & 1 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & -1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{y} \quad B = \begin{pmatrix} 4 & 2 \\ 2 & 1 \end{pmatrix}.$$

    1. Probar que el sistema de ecuaciones es compatible.
    2. Hallar la expresión de la solución general de este sistema.

40. Sea $A \in \mathscr{M}_{m\times n}(\mathbb{R})$ y $\mathbf{b} \in \mathbb{R}^m$. Probar que $\hat{\mathbf{x}}$ es una solución aproximada mínimo cuadrática del sistema $A\mathbf{x} = \mathbf{b}$ si, y sólo si, $\hat{\mathbf{x}}$ forma parte de una solución del sistema ampliado

$$\begin{pmatrix} I_m & A \\ A^t & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{y} \\ \hat{\mathbf{x}} \end{pmatrix} = \begin{pmatrix} \mathbf{b} \\ \mathbf{0} \end{pmatrix}$$

No es extraño encontrar problemas de mínimos cuadrados en los que la matriz A es muy grande pero contiene muchos ceros. Para esta situación, el anterior sistema ampliado contendrá menos entradas no nulas que el sistema de ecuaciones normales, y evitará los problemas de memoria que suelen dar los algoritmos de resolución. Además, se evita el cálculo de $A^{t}A$ que puede producir problemas de mal condicionamiento.

41. Consideremos el problema de calcular la solución de norma mínima del problema de mínimos cuadrados mín $\|A\mathbf{x}-b\|^2$, donde

$$A=\begin{pmatrix}1&0\\0&0\end{pmatrix}\quad \text{y}\quad \mathbf{b}=\begin{pmatrix}1\\0\end{pmatrix}.$$

Probar que

1. la solución $\hat{\mathbf{x}}=(1,0)^{t}$.
2. Consideremos la perturbación de A

$$E_1=\begin{pmatrix}0&\delta\\0&0\end{pmatrix}$$

donde δ es un número positivo pequeño. Resolver la versión perturbada del problema anterior mín $\|A_1\mathbf{y}-b\|^2$, donde $A_1=A+E_1$. ¿Qué le ocurre a $\|\hat{\mathbf{x}}-\mathbf{y}\|$ cuando δ se aproxima a cero?

3. Ahora consideremos la perturbación de A

$$E_2=\begin{pmatrix}0&0\\0&\delta\end{pmatrix}$$

donde δ es un número positivo pequeño. Resolver la versión perturbada del problema anterior mín $\|A_2\mathbf{z}-b\|^2$, donde $A_2=A+E_2$. ¿Qué le ocurre a $\|\hat{\mathbf{x}}-\mathbf{z}\|$ cuando δ se aproxima a cero?

Capítulo 5

Clasificación euclídea de aplicaciones lineales y formas cuadráticas

5.1. Matrices simétricas reales

1. Definición: Sea $(\mathbb{R}^n,\cdot)$ el espacio euclídeo usual. Diremos que un endomorfismo lineal $T\colon \mathbb{R}^n \to \mathbb{R}^n$ es simétrico si $T(x)\cdot y = x\cdot T(y)$ para todo $x,y\in\mathbb{R}^n$.

2. Proposición: $T\colon \mathbb{R}^n \to \mathbb{R}^n$ *es simétrico si y sólo si la matriz asociada a T en la base usual de $\mathbb{R}^n$ es simétrica.*

Demostración. Sean $x \equiv (x_1,\dots,x_n)$ y $y \equiv (y_1,\dots,y_n)$ en la base usual de $\mathbb{R}^n$ y A la matriz asociada a T en la base usual. Entonces

$$(Tx)\cdot y = (A\cdot x^t)^t\cdot \mathrm{Id}\cdot y^t = (x_1,\dots,x_n)\cdot A^t\cdot \begin{pmatrix} y_1\\ \vdots \\ y_n\end{pmatrix}$$
$$x\cdot(Ty) = (x_1,\dots,x_n)\cdot Id\cdot A\cdot \begin{pmatrix} y_1\\ \vdots \\ y_n\end{pmatrix} = (x_1,\dots,x_n)\cdot A\cdot \begin{pmatrix} y_1\\ \vdots \\ y_n\end{pmatrix}$$

Luego $T(x)\cdot y = x\cdot T(y)$, para todo $x,y\in\mathbb{R}^n$ si y sólo si A es simétrica.

□

3. Lema: *Sea $(\mathbb{R}^n,\cdot)$ el espacio euclídeo usual y $T\colon\mathbb{R}^n\to\mathbb{R}^n$ un endomorfismo lineal simétrico. Entonces, para todo $x,y\in\mathbb{R}^n$ se cumple que*

1. $(T^m x)\cdot y = x\cdot(T^m y)$, para todo $m\in\mathbb{N}$.

2. $(p(T)x)\cdot y = x\cdot(p(T)y)$, para todo polinomio $p(z)\in\mathbb{R}[z]$.

Demostración. 1. Usando la proposición anterior, tenemos que $(T^2x)y=(Tx)\cdot(Ty)=x\cdot(T^2y)$, etc.

2. $((\sum_i \lambda_i T^i)x)\cdot y = \sum_i \lambda_i((T^ix)\cdot y) = \sum_i \lambda_i(x\cdot(T^iy)) = x\cdot((\sum_i \lambda_i T^i)y)$.

□

4. Teorema: *Sea $(\mathbb{R}^n,\cdot)$ el espacio euclídeo usual y $T:\mathbb{R}^n\to\mathbb{R}^n$ un endomorfismo lineal tal que su matriz asociada $A\in M_n(\mathbb{R})$ en la base usual es una matriz simétrica. Entonces, existe una base ortonormal en la que T diagonaliza. Es decir, existen una matriz ortogonal P y una matriz diagonal D tales que*

$$A = P\cdot D\cdot P^t$$

Demostración. a) Los autovalores de T son reales. Tenemos que ver que las raíces del polinomio característico de T, $\chi_T(x)$ son reales. Sea $\alpha = a+bi$, con $b\neq 0$ una raíz compleja de $\chi_T(x)$. Entonces $\chi_T(x)=(x-\alpha)\cdot(x-\bar{\alpha})\cdot q(x) = ((x-a)^2+b^2)\cdot q(x)$. Entonces existe $v\in\mathbb{R}^n$, no nulo, de modo que $0=((T-a)^2+b^2)v=(T-a)^2v+b^2v$, de donde

$$0=((T-a)^2v+b^2v)\cdot v = (T-a)v\cdot(T-a)v+(bv)\cdot(bv)>0$$

Hemos llegado a contradicción.

b) T es diagonalizable. Tenemos que ver que si $\lambda\in\mathbb{R}$ es un autovalor de T entonces $\operatorname{Ker}(T-\lambda)=\operatorname{Ker}(T-\lambda)^2$ (es decir, todos los bloques de Jordan son de rango 1). Sea $v\in\operatorname{Ker}(T-\lambda)^2$. Entonces

$$0=((T-\lambda)^2v)\cdot v = ((T-\lambda)v)\cdot((T-\lambda)v)$$

luego $(T-\lambda)v=0$ y $v\in\operatorname{Ker}(T-\lambda)$.

c) Si u,v son dos autovectores de autovalores $\lambda\neq\mu$, entonces son ortogonales:

$$(\lambda u)\cdot v = (Tu)\cdot v = u\cdot(Tv) = u\cdot(\mu v) = \mu u\cdot v$$

Luego, $u\cdot v = 0$.

Entonces $\mathbb{R}^n = \operatorname{Ker}(T-\lambda_1)\oplus\cdots\oplus\operatorname{Ker}(T-\lambda_r)$, donde $\lambda_1,\dots,\lambda_r$ son los autovalores de T. Consideremos una base ortonormal en cada $\operatorname{Ker}(T-\lambda_i)$. Todas juntas forman una base B ortonormal de $\mathbb{R}^n$ (por c)), en la que T diagonaliza. Sea P la matriz ortogonal de cambio de base de B a la base ortonormal usual de $\mathbb{R}^n$. Se cumple que

$$A = P\cdot\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_r\end{pmatrix}\cdot P^t$$

□

Del mismo modo que el caso real, en el caso complejo tenemos:

5. Proposición: *Si $A \in M_n(\mathbb{C})$ es hermítica entonces todos sus autovalores son reales y existe una matriz unitaria U de modo que $U^* \cdot A \cdot U$ es una matriz diagonal real D, luego*

$$A = U \cdot D \cdot U^*$$

Demostración. En primer lugar hay que demostrar que todos los coeficientes del polinomio característico son reales y después se continúa como en el caso real. □

6. Teorema: *Sea $(\mathbb{R}^n, \cdot)$ el espacio euclídeo usual y T_2 una forma bilineal simétrica de $\mathbb{R}^n$. Entonces, existe una base ortonormal para $\cdot$, ortogonal para T_2.*

Demostración. Sea A la matriz (simétrica) asociada a T_2 en la base usual. Por 5.1.4, existen una matriz ortogonal P y una matriz digonal D tal que

$$A = P \cdot D \cdot P^t$$

Es decir, si B es la base (ortonormal para $\cdot$) formada por las columnas de P, entonces la matriz asociada a T_2 en B es D. Es decir, B es una base ortogonal para T_2.

□

5.2. Descomposición en valores singulares

1. Teorema: *Sean $(E, \cdot)$ y $(E', \cdot)$ dos espacios vectoriales euclídeos. Sea $T: E \to E'$ una aplicación lineal. Entonces, existen bases ortonormales B, B' de E y E' en las que T "diagonaliza".*

Demostración. Sea T_2 la forma bilineal simétrica de E definida por

$$T_2(e, e') := T(e) \cdot T(e'), \ \ \forall e, e' \in E$$

Por 5.1.6, existe una base ortonormal de E, $B = \{e_1, \ldots, e_n\}$, ortogonal para T_2. Por tanto, $T(e_i) \cdot T(e_j) = T_2(e_i, e_j) = \lambda_i \cdot \delta_{ij}$. Reordenando la base, podemos suponer que $\lambda_1, \ldots, \lambda_r$ son positivos y que $\lambda_{r+1} = \cdots = \lambda_n = 0$. Por tanto, $T(e_1), \ldots, T(e_r)$ son vectores no nulos ortogonales entre sí y $T(e_{r+1}) = \cdots = T(e_n) = 0$. Sea $B' = \{e'_1 = T(e_1)/\sqrt{\lambda_1}, \cdots, e'_r =$

$T(e_r)/\sqrt{\lambda_r}, e'_{r+1}, \cdots, e'_m\}$ una base ortonormal de E'. La matriz asociada a T en las bases B, B' es

$$T \equiv \begin{pmatrix} \sqrt{\lambda_1} & & & & & \\ & \ddots & & & & \\ & & \sqrt{\lambda_r} & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \\ & & & & & \end{pmatrix}$$

□

2. Observación: Dados $e = (x_1, \cdots, x_n), e' = (y_1, \ldots, y_n) \in \mathbb{R}^n$,

$$T_2(e, e') = T(e) \cdot T(e') = (x_1, \ldots, x_n) \cdot A^t \cdot A \cdot (y_1, \ldots, y_n)^t$$

Luego la matriz asociada a $T_2 = A^t A$. Recordemos que $\lambda_1, \ldots, \lambda_r$ son los autovalores no nulos de la matriz $A^t A$ y que $B = \{e_i\}$ es una base ortonormal de autovectores de $A^t A$. Observemos además que $e_{r+1}, \ldots, e_n$ es una base de $\operatorname{Ker} T$, luego $e_1, \ldots, e_r$ es una base de $(\operatorname{Ker} T)^\perp$. Por Ãºltimo, $\{e'_1 = T(e_1)/||T(e_1)||, \ldots, e'_r = T(e_r)/||T(e_r)||\}$ es una base ortonormal de $\operatorname{Im} T$ y $\{e'_{r+1}, \ldots, e'_m\}$ es una base ortonormal de $(\operatorname{Im} T)^\perp$.

3. Descomposición en valores singulares: *Sea $A \in M_{m\times n}(\mathbb{R})$ de rango r. Existen matrices ortogonales $P \in M_m(\mathbb{R})$ y $Q \in M_n(\mathbb{R})$, tales que*

$$A = P \cdot \begin{pmatrix} \sqrt{\lambda_1} & & & & & \\ & \ddots & & & & \\ & & \sqrt{\lambda_r} & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \\ & & & & & \end{pmatrix} \cdot Q^t \qquad (*)$$

Demostración. Es consecuencia directa del teorema anterior, donde $T\colon \mathbb{R}^n \to \mathbb{R}^m$ es una aplicación lineal de matriz asociada A, en las bases usuales de $\mathbb{R}^n$ y $\mathbb{R}^m$, Q es la matriz de cambio de base de la base B a la base usual de $\mathbb{R}^n$ y P es la matriz de cambio de base de la base B' a la base usual de $\mathbb{R}^m$. □

4. Definición: Se dice que $\sqrt{\lambda_1}, \cdots, \sqrt{\lambda_r}$ son los valores singulares de A y la descomposición $(*)$ se llama la descomposición en valores singulares de A.

Los valores singulares de A no dependen de la descomposición en valores singulares de A: Si $A = P\cdot D\cdot Q^t$, con P,Q matrices ortogonales y D una matriz diagonal con coeficientes no negativos entonces $A^tA = Q\cdot D^2\cdot Q^t$. Por tanto, los coeficientes de la diagonal de D son las raíces cuadradas de los autovalores de A^tA.

5. Descomposición en valores singulares corta: *Sigamos las notaciones de 5.2.3. Sea $Q_1 \in M_{n\times r}(\mathbb{R})$ la matriz formada por las r-primeras columnas de Q y $P_1 \in M_{m\times r}$ la matriz formada por las r primeras columnas de P y $\Delta \in M_r(\mathbb{R})$ la matriz diagonal de coeficientes en la diagonal $\sqrt{\lambda_i}$. Entonces*

$$A = P_1\cdot\Delta\cdot Q_1^t \qquad (**)$$

Además, por ser las columnas de P y Q vectores ortonormales se tiene que $P_1^t\cdot P_1 = \mathrm{Id}$ y $Q_1^t\cdot Q = \mathrm{Id}$.

Demostración. Con las notaciones obvias $P = (P_1,P_2)$, $Q = (Q_1,Q_2)$ y

$$A = (P_1,P_2)\cdot\begin{pmatrix}\Delta & 0\\ 0 & 0\end{pmatrix}\cdot\begin{pmatrix}Q_1^t\\ Q_2^t\end{pmatrix} = (P_1,P_2)\cdot\begin{pmatrix}\Delta\cdot Q_1^t\\ 0\end{pmatrix} = P_1\cdot\Delta\cdot Q_1^t$$

□

6. Definición: Se dice que la descomposición $(**)$ es la descomposición en valores singulares corta de A.

Calculemos la descomposición en valores singulares de T^+.

Existen unas bases ortonormales $\{e_1,\dots,e_n\}$, ($\{e_1,\dots,e_r\}$ base de $(\mathrm{Ker}\,T)^\perp$ y $\{e_{r+1},\dots,e_n\}$ base de $(\mathrm{Ker}\,T)$) y $\{e'_1,\dots,e'_m\}$ ($\{e'_1,\dots,e'_r\}$ base de $\mathrm{Im}\,T$) y $\{e'_{r+1},\dots,e'_m\}$ base de $(\mathrm{Im}\,T)^\perp$) de modo que $T(e_i) = \sqrt{\lambda_i}\cdot e'_i$, si $i\le r$ y $T(e_i) = 0$ si $i > r$. La matriz asociada a T en estas bases es la diagonal, D

$$D = \begin{pmatrix}\Delta & 0\\ 0 & 0\end{pmatrix}, \qquad \Delta = \begin{pmatrix}\sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_r}\end{pmatrix}$$

La inversa de Moore-Penrose T^+ está definida por $T(e'_i) = (1/\sqrt{\lambda_i})\cdot e_i$, si $i\le r$ y $T(e'_i) = 0$ si $i > r$.

Si A es la matriz asociada a T en la bases usuales de $\mathbb{R}^n$ y $\mathbb{R}^m$ teníamos la descomposición en valores singulares

$$A = P\cdot\begin{pmatrix}\Delta & 0\\ 0 & 0\end{pmatrix}\cdot Q^t$$

Por tanto, la matriz asociada a T^+ en las bases usuales es

$$A^+ = Q \cdot \begin{pmatrix} \Delta^{-1} & 0 \\ 0 & 0 \end{pmatrix} \cdot P^t$$

Si consideramos la descomposición en valores singulares corta tenemos que

$$A = P_1 \cdot \Delta \cdot Q_1^t, \quad A^+ = Q_1 \cdot \Delta^{-1} \cdot P_1^t$$

Por último,

$$A^- = Q \cdot \begin{pmatrix} \Delta^{-1} & B \\ C & D \end{pmatrix} \cdot P^t \quad \text{y} \quad A^\square = Q \cdot \begin{pmatrix} \Delta^{-1} & 0 \\ C & D \end{pmatrix} \cdot P^t$$

5.3. Formas cuadráticas

1. Definición : Una forma cuadrática en $\mathbb{R}^n$ es un polinomio homogéneo de grado dos en n-variables,

$$\begin{array}{ccc} \mathbb{R}^n & \xrightarrow{q} & \mathbb{R} \\ (x_1, \dots, x_n) & \mapsto & q(x) = \sum_{i \le j} a_{ij} x_i x_j \end{array}$$

2. Proposición : *El conjunto de formas cuadráticas de $\mathbb{R}^n$ está en correspondencia biunívoca con el conjunto de matrices simétricas de orden n (o equivalentemente con el conjunto de formas bilineales simétricas).*

Demostración. Asignemos a la matriz simétrica $A = (a_{ij})$, la forma cuadrática

$$q(x_1, \dots, x_n) := (x_1, \dots, x_n) \cdot A \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \sum_{ij} a_{ij} x_i x_j = \sum_i a_{ii} x_i^2 + \sum_{i<j} 2a_{ij} x_i x_j$$

Asignemos a la forma cuadrática, $q(x_1, \dots, x_n) = \sum_{i \le j} b_{ij} x_i x_j$ la matriz $A = (a_{ij})$ donde $a_{ij} = a_{ji} = b_{ij}/2$ si $i < j$ y $a_{ii} = b_i$.

Ambas asignaciones son inversas entre sí. □

Dada una forma cuadrática

$$q(x_1, \dots, x_n) = \sum_{i \le j} b_{ij} x_i x_j = (x_1, \dots, x_n) \cdot (a_{ij}) \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

y un cambio de coordenadas

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (a'_{ij}) \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$$

se tiene que

$$q(x_1,\ldots,x_n) = (x_1,\ldots,x_n)(a_{ij}) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (x'_1,\ldots,x'_n)(a'_{ij})^t \cdot (a_{ij}) \cdot (a'_{ij}) \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$$

Sabemos que existe un cambio de coordenadas (a'_{ij}) ortogonal (de columnas una base ortonormal donde la matriz (a_{ij}) diagonaliza) tal que $(a'_{ij})^t \cdot (a_{ij}) \cdot (a'_{ij}) = D$ es una matriz diagonal de coeficientes $\lambda_1,\ldots,\lambda_n$. Luego,

$$q(x_1,\ldots,x_n) = \sum_{i\leq j} b_{ij} x_i x_j = \sum_i \lambda_i {x'_i}^2$$

5.3.1. Clasificación afín euclídea de cuádricas

Consideremos un polinomio de grado dos $p(x_1,\ldots,x_n) = \sum_{i\leq j} a_{ij} x_i \cdot x_j + \sum_i a_i x_i + a$ (estamos pensando en la cuádrica $\sum_{i\leq j} a_{ij} x_i \cdot x_j + \sum_i a_i x_i + a = 0$). Por un cambio de coordenadas (a'_{ij}) ortogonal

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (a'_{ij}) \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

podemos conseguir que $\sum_{i\leq j} a_{ij} x_i x_j = b_1 y_1^2 + \cdots + b_r y_r^2$ (con $b_i \neq 0$). Por tanto,

$$p(x_1,\ldots,x_n) = b_1 y_1^2 + \cdots + b_r y_r^2 + \sum_i b'_i y_i + b'$$

Por la traslación $y_i = (z_i - \frac{b'_i}{2\cdot b_i})$, para todo $i \leq r$, $y_j = z_j$ para todo $j > r$, tenemos que

$$p(x_1,\ldots,x_n) = b_1 z_1^2 + \cdots + b_r z_r^2 + b_{r+1} z_{r+1} + \cdots + b_n z_n + b''$$

Por otro cambio de coordenadas ortogonal ($u_1 = z_1, \ldots, u_r = z_r$, $u_{r+1} = \frac{b_{r+1} z_{r+1} + \cdots + b_n z_n}{\sqrt{\sum_i b_{r+i}^2}}$, etc.) obtenemos

$$p(x_1,\ldots,x_n) = b_1 u_1^2 + \cdots + b_r u_r^2 + c \cdot u_{r+1} + d$$

Si $c = 0$, entonces

$$p(x_1,\ldots,x_n) = b_1 u_1^2 + \cdots + b_r u_r^2 + d$$

(aquí hay que distinguir dos casos $d=0$ y $d\neq 0$).

Si $c\neq 0$, por traslación podemos suponer que

$$p(x_1,\dots,x_n)=b_1u_1^2+\cdots+b_ru_r^2+c\cdot u_{r+1}$$

Vamos a considerar $p(x_1,\dots,x_n)$ equivalente a $\lambda\cdot p(x_1,\dots,x_n)$, $\lambda\in\mathbb{R}$. Módulo un factor multiplicativo podemos suponer que c o d son iguales a -1. Reordenando las r primeras variables, si es necesario podemos suponer que $b_i>0$, para $i\leq s$ y $b_j<0$ para $s<j\leq r$. En conclusión, por cambio de coordenadas afín euclídeo

$$\boxed{p(x_1,\dots,x_n)=\begin{cases} \text{1. } a_1x_1^2+\cdots+a_sx_s^2-a_{s+1}\cdot x_{s+1}^2-\cdots-a_r\cdot x_r^2-1\\ \text{ó}\\ \text{2. } a_1x_1^2+\cdots+a_sx_s^2-a_{s+1}\cdot x_{s+1}^2-\cdots-a_r\cdot x_r^2\\ \text{ó}\\ \text{3. } a_1x_1^2+\cdots+a_sx_s^2-a_{s+1}\cdot x_{s+1}^2-\cdots-a_r\cdot x_r^2-x_{r+1}\end{cases}\\ \text{(con } a_1,\dots,a_r>0)}$$

3. Observación: El caso 3. $p(x_1,\dots,x_n)=a_1x_1^2+\cdots+a_sx_s^2-a_{s+1}\cdot x_{s+1}^2-\cdots-a_r\cdot x_r^2-x_{r+1}$, $a_i>0$ es equivalente (multiplicando $p(x)$ por -1, haciendo el cambiando x_{r+1} por $-x_{r+1}$ y reordenando las r primeras coordenadas) a $b_1x_1^2+\cdots+b_{r-s}x_{r-s}^2-b_{r-s+1}\cdot x_{r-s+1}^2-\cdots-b_r\cdot x_r^2-x_{r+1}$, $b_i>0$ (con $b_i=a_{s+i}$, para $i\leq r-s$ y $b_{r-s+i}=a_i$, para $i\leq s$).

Igualmente, el caso 2. $p(x_1,\dots,x_n)=a_1x_1^2+\cdots+a_sx_s^2-a_{s+1}\cdot x_{s+1}^2-\cdots-a_r\cdot x_r^2$ es equivalente a $b_1x_1^2+\cdots+b_{r-s}x_{r-s}^2-b_{r-s+1}\cdot x_{r-s+1}^2-\cdots-b_r\cdot x_r^2$.

En conclusión, en los casos 2. y 3., podemos suponer que s es menor o igual que la parte entera de $r/2$ (pues ó $s=r/2$ y $r-s=r/2$, ó sólo uno de los dos es menor que la parte entera de $r/2$).

4. Ejemplo: Clasificación afín euclídea de las cónicas. Es decir, clasificación afín euclídea de los polinomios en dos variables de grado dos (salvo un factor multiplicativo). Sigamos notaciones anteriores ($a,b>0$),

1. $r=2, s=0$: $ax^2+by^2-1=0$, (elipse); $ax^2+by^2=0$ (dos rectas imaginarias no paralelas).
2. $r=2, s=1: ax^2-by^2-1=0$ (hipérbola); $ax^2-by^2=0$ (par de rectas, no paralelas).
3. $r=2, s=2$: $-ax^2-by^2-1=0$ (elipse imaginaria).
4. $r=1, s=0$: $ax^2-1=0$ (dos rectas paralelas), $ax^2=0$ (recta doble), $ax^2-y=0$ (parábola).
5. $r=1, s=1$: $-ax^2-1=0$ (dos rectas imaginarias paralelas).

6. $r=0$: $-1=0$ (el vacío), $0=0$ (todo el plano), $-x=0$ (una recta).

5. Ejemplo : Clasificación euclídea afín de las cuádricas de $\mathbb{R}^3$. Es decir, clasificación afín euclídea de los polinomios en tres variables de grado dos (salvo un factor multiplicativo). Sigamos notaciones anteriores $(a,b,c>0)$,

1. $r=3, s=0$: $ax^2+by^2+cz^2-1=0$ (elipsoide), $ax^2+by^2+cz^2=0$ (cono imaginario).
2. $r=3, s=1$: $ax^2+by^2-cz^2-1=0$ (hiperboloide reglado), $ax^2+by^2-cz^2=0$ (cono).
3. $r=3, s=2$: $ax^2-by^2-cz^2-1=0$ (hiperboloide no reglado).
4. $r=3, s=3$: $-ax^2-by^2-cz^2-1=0$ (elipsoide imaginario).
5. $r=2, s=0$: $ax^2+by^2-1=0$ (cilindro elíptico), $ax^2+by^2=0$ (dos hiperplanos imaginarios no paralelos), $ax^2+by^2-z=0$ (paraboloide).
6. $r=2, s=1$: $ax^2-by^2-1=0$ (cilindro hiperbólico), $ax^2-by^2=0$ (dos hiperplanos no paralelos), $ax^2-by^2-z=0$ (paraboloide reglado).
7. $r=2, s=2$: $-ax^2-by^2-1=0$ (cilindro imaginario).
8. $r=1, s=0$: $ax^2-1=0$ (par de planos paralelos), $ax^2=0$ (plano doble), $ax^2-y=0$ (cilindro parabólico).
9. $r=1, s=1$: $-ax^2-1=0$ (par de planos imaginarios paralelos).
10. $r=0$: $-1=0$ (vacío), $0=0$ (todo $\mathbb{R}^3$), $-x=0$ (un plano).

5.4. Problemas

1. Calcular la descomposición en valores singulares (larga y corta) de la matriz
$$A=\begin{pmatrix} 1 & 2 & 2 & 1 \\ 1 & 1 & 1 & -1 \end{pmatrix}.$$
2. Sea $A \in \mathcal{M}_{m\times n}(\mathbb{R})$.
    1. Probar que los valores singulares de A son los mismos que los de A^{t}.
    2. Probar que los valores singulares de A son los mismos que los de UAV, si $U \in \mathcal{M}_m(\mathbb{R})$ y $V \in \mathcal{M}_n(\mathbb{R})$ son matrices ortogonales.

3. Si $\alpha \neq 0$ es un escalar, ¿cómo son los valores singulares de αA en comparación con los de A?

3. Sea $A \in \mathscr{M}_{m\times n}(\mathbb{R})$. Si A tiene rango r y su descomposición en valores singulares (larga) es
$$A = P\begin{pmatrix} \Delta & 0 \\ 0 & 0 \end{pmatrix} Q^{\mathrm{t}},$$
probar que, si $\mathbf{v}_i$ y $\mathbf{u}_i$ denotan, respectivamente, la columna i-ésima de P y Q, entonces, $\mathbf{v}_i = (1/\sigma_i)A^{\mathrm{t}}\mathbf{u}_i,\ i = 1,\ldots,r$.

4. Sean V un $\mathbb{R}$-espacio vectorial de dimensión $3, \mathscr{B}$ una base de V y T_2 la forma bilineal sobre V cuya matriz respecto de $\mathscr{B}$ es
$$\begin{pmatrix} 2 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 6 \end{pmatrix}.$$
Hallar una base de V respecto de la cual la matriz de T_2 sea diagonal.

5. Probar las siguientes afirmaciones:
   1. Si $A \in \mathscr{M}_n(\mathbb{R})$ es simétrica y P es una matriz invertible, entonces A es (semi)definida positiva si, y sólo si, lo es $P^{\mathrm{t}}AP$.
   2. Si Si $A \in \mathscr{M}_n(\mathbb{R})$ es simétrica, entonces A es definida positiva si, y sólo si, existe una matriz P invertible tal que $P^{\mathrm{t}}AP = I_n$.
   3. Si $A \in \mathscr{M}_n(\mathbb{R})$ es simétrica, entonces A es definida positiva si, y sólo si, existe una matriz Q invertible tal que $A = Q^{\mathrm{t}}Q$.
   4. Si $A \in \mathscr{M}_{m\times n}(\mathbb{R})$, las matrices $A^{\mathrm{t}}A$ y AA^{t} son semidefinidas positivas.
   5. Si $A \in \mathscr{M}_{m\times n}(\mathbb{R})$, entonces el rango de A es m si, y sólo si, la matriz AA^{t} es definida positiva.
   6. Si $A \in \mathscr{M}_{m\times n}(\mathbb{R})$, entonces el rango de A es n si, y sólo si, la matriz $A^{\mathrm{t}}A$ es definida positiva.
   7. Si $A \in \mathscr{M}_n(\mathbb{R})$ es simétrica de rango r, entonces existe una matriz $B \in \mathscr{M}_{n\times r}(\mathbb{C})$ de rango r tal que $A = BB^*$. Además, si A es semidefinida positiva, entonces B puede tomarse real.

6. Aplicar los distintos criterios para determinar si las siguientes formas cuadráticas son definidas positivas (negativas) o semidefinidas positivas o negativas. Escribir también la forma reducida de cada una de ellas.

1. $q_1(x,y,z) = 3x^2 + 16y^2 + 139z^2 + 12xy + 30xz + 92yz$.
2. $q_2(x,y,z) = -4x^2 - 5y^2 - 2z^2 + 4xz$.
3. $q_3(x,y,z) = x^2 + 4y^2 - 4xy$.
4. $q_4(x,y,z,t) = -4x^2 + 4xy - y^2 - 9z^2 + 6zt - t^2$.
5. $q_5(x,y) = xy$.
6. $q_6(x,y,z,t) = 2xt + 2yz$.

7. Dada la matriz $A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$,

   a. Escribir una matriz ortogonal P tal que $P^{-1}AP$ sea una matriz diagonal D.

   b. Escribir una matriz Q, que pueda expresarse como producto de matrices de transformaciones elementales del tipo permutar dos filas y sumar a una fila λ-veces otra fila, tal que $Q^{t}AQ$ sea una matriz diagonal $\bar{D}$

   c. Escribir, si es posible, una matriz R, que pueda expresarse como producto de matrices de transformaciones elementales, tal que $R^{t}AR = I_3$.

   Sea T_2 la forma bilineal simétrica que, en la base usual de $\mathbb{R}^3$ tiene matriz A y sea q la forma cuadrática asociada a T_2.

   d. Comprobar que T_2 es un producto escalar.

   e. Las columnas de P forman una base ortonormal para el producto escalar usual de $\mathbb{R}^3$. Comprobar que dichas columnas forman una base ortogonal para T_2.

   f. Comprobar que las columnas de Q forman una base ortogonal para T_2 y que las de R forman una base ortonormal para T_2.

   g. Escribir la expresión de q en coordenadas para las bases dadas por las columnas de P, de Q y de R.

8. Sean $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ y $B \in \mathcal{M}_{m\times n}(\mathbb{R})$ simétricas y semidefinidas positivas tales que $A - B$ también es semidefinida positiva. Probar que $B^+ - A^+$ es semidefinida positiva si, y sólo si, $\mathrm{rg}(A) = \mathrm{rg}(B)$.