Capítulo 6

Diagonalización

En este capítulo empezaremos a estudiar la estructura de los endomorfismos de un espacio vectorial de dimensión finita.

6.1 Nociones básicas

Dada una transformación lineal $f : K^n \to K^n$, y dos bases B_1 y B_2 de K^n se tiene que

$$|f|_{B_1} = C(B_2, B_1)|f|_{B_2}C(B_1, B_2) = C(B_2, B_1)|f|_{B_2}C(B_2, B_1)^{-1},$$

y por lo tanto, existe una matriz $C \in GL(n, K)$ tal que $|f|_{B_1} = C. |f|_{B_2}. C^{-1}$. Recíprocamente, si $A, B \in K^{n\times n}$ son tales que existe una matriz $C \in GL(n, K)$ tal que $A = C.B.C^{-1}$, definiendo $f : K^n \to K^n$ como $f(x) = A.x$ y considerando $B_1 = E$ la base canónica de K^n y B_2 la base de K^n formada por las columnas de C, resulta que

$$A = |f|_{B_1} \quad \text{y} \quad B = C^{-1}.A.C = C(E, B_2)|f|_E\, C(B_2, E) = |f|_{B_2}.$$

Esto da lugar a la siguiente definición (ya introducida en el Ejercicio 35 de la Sección 3.8):

Definición 6.1 Sean $A, B \in K^{n\times n}$. Se dice que A y B son *semejantes*, y se nota $A \sim B$, si existe una matriz $C \in GL(n, K)$ tal que $A = C.B.C^{-1}$.

Por lo tanto, se demostró la siguiente propiedad (que es lo propuesto por el Ejercicio 36 de la Sección 3.8):

Proposición 6.2 Sean $A, B \in K^{n\times n}$. Entonces $A \sim B$ si y sólo si existen una transformación lineal $f : K^n \to K^n$ y bases B_1 y B_2 de K^n tales que $|f|_{B_1} = A$ y $|f|_{B_2} = B$.

Por lo que vimos, una misma transformación lineal da lugar a matrices semejantes si calculamos sus matrices en distintas bases. Es por eso que, en lo que sigue, estudiaremos la semejanza de matrices. El primer problema que consideraremos es el de determinar si una matriz es semejante a una matriz diagonal.

Definición 6.3 Una matriz $A \in K^{n\times n}$ se dice *diagonalizable* si existe una matriz $C \in GL(n,K)$ tal que $C.A.C^{-1}$ es una matriz diagonal.

En otras palabras, una matriz diagonalizable es una matriz que es semejante a una matriz diagonal. La noción correspondiente para transformaciones lineales es la siguiente:

Definición 6.4 Sea V un K-espacio vectorial de dimensión finita, y sea $f : V \to V$ una transformación lineal. Se dice que f es *diagonalizable* o *diagonal* si existe una base B de V tal que $|f|_B$ es diagonal.

Teniendo en cuenta que la semejanza de matrices es una relación de equivalencia (ver Ejercicio 35 de la Sección 3.8) deducimos que:

Observación 6.5 Sea V un K-espacio vectorial de dimensión finita y sea $f : V \to V$ una transformación lineal. Entonces f es diagonalizable si y sólo si $|f|_B$ es diagonalizable para toda base B de V.

6.1.1 Autovalores y autovectores

Sea V un K-espacio vectorial de dimensión n y sea $f : V \to V$ una transformación lineal diagonalizable. Luego, existe una base $B = \{v_1, \dots, v_n\}$ de V tal que $|f|_B$ es diagonal:

$$|f|_B = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix}.$$

Entonces, para cada $1 \le i \le n$, $f(v_i) = \lambda_i v_i$.

Recíprocamente, si para una base $B = \{v_1, \dots, v_n\}$ de V y $\lambda_1, \dots, \lambda_n \in K$ se cumple que $f(v_i) = \lambda_i v_i$ para cada $1 \le i \le n$, la matriz $|f|_B$ es diagonal y, en consecuencia, f es diagonalizable.

Esto nos lleva a la siguiente definición:

Definición 6.6 Sea V un K-espacio vectorial, y sea $f : V \to V$ una transformación lineal. Se dice que $v \in V$, $v \ne 0$, es un *autovector* de f si existe $\lambda \in K$ tal que $f(v) = \lambda.v$. El elemento $\lambda \in K$ se llama un *autovalor* de f.

Usando estas definiciones, el razonamiento anterior se puede reescribir de esta forma:

Proposición 6.7 *Sea V un K-espacio vectorial de dimensión n y sea $f : V \to V$ una transformación lineal. Entonces f es diagonalizable si y sólo si existe una base B de V formada por autovectores de f.*

La mismas nociones se pueden definir para matrices: Dada $A \in K^{n\times n}$, se le puede asociar una transformación lineal $f_A : K^n \to K^n$ definida por $f_A(x) = A.x$. Notar que $|f_A|_E = A$, donde E es la base canónica de K^n. Entonces $v \in K^n$, $v \neq 0$, es un autovector de f_A de autovalor λ si y sólo si $A.v = \lambda.v$.

Definición 6.8 Sea $A \in K^{n\times n}$. Se dice que $v \in K^n$, $v \neq 0$, es un *autovector* de A si existe $\lambda \in K$ tal que $A.v = \lambda.v$. El elemento $\lambda \in K$ que verifica la condición anterior se llama un *autovalor* de A.

Podemos dar también un resultado análogo a la Proposición 6.7 para matrices:

Proposición 6.9 *Sea $A \in K^{n\times n}$. Entonces A es diagonalizable si y sólo si existe una base B de K^n formada por autovectores de A.*

Ejemplos.

1. Decidir si $A = \begin{pmatrix} 2 & 3 \\ 2 & 1 \end{pmatrix} \in \mathbb{R}^{2\times 2}$ es diagonalizable.

   En virtud de la proposición anterior, basta buscar los autovectores de A, es decir, los vectores $x = (x_1, x_2) \in \mathbb{R}^2$ tales que $(x_1, x_2) \neq (0, 0)$ y $A.\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \lambda.x_1 \\ \lambda.x_2 \end{pmatrix}$ para algún $\lambda \in \mathbb{R}$.

   Para esto, buscaremos en primer término los elementos $\lambda \in \mathbb{R}$ para los cuales el sistema $A.x = \lambda.x$ tiene solución no trivial (autovalores de A) y después, para cada uno de los valores hallados, los vectores $(x_1, x_2) \in \mathbb{R}^2$ que son soluciones del sistema lineal correspondiente.

   Observamos que

   $$A.x = \lambda.x \iff (\lambda I_2 - A).x = 0 \iff \begin{pmatrix} \lambda-2 & -3 \\ -2 & \lambda-1 \end{pmatrix} . \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

   Este sistema homogéneo tiene solución no trivial si y sólo si el determinante de su matriz asociada es 0, o sea, si y sólo si $\lambda^2 - 3\lambda - 4 = 0$. Luego, los autovalores de A son $\lambda = -1$ y $\lambda = 4$.

   Busquemos ahora los autovectores correspondientes a cada autovalor:

   Para $\lambda = -1$, queda el sistema

   $$\begin{pmatrix} -3 & -3 \\ -2 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

   cuyo conjunto de soluciones es $< (1, -1) >$. Luego el conjunto de los autovectores asociados a $\lambda = -1$ es $< (1, -1) > - \{(0, 0)\}$.

   Para $\lambda = 4$, el sistema es

   $$\begin{pmatrix} 2 & -3 \\ -2 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

cuyo conjunto de soluciones es $< (3,2) >$. Luego el conjunto de los autovectores asociados a $\lambda = 4$ es $< (3,2) > - \{(0,0)\}$.

En consecuencia, A es diagonalizable, puesto que $B = \{(1,-1),(3,2)\}$ es una base de $\mathbb{R}^2$ formada por autovectores de A. Más aún, si $C = C(E, B)$ se tiene que

$$C.\,A.\,C^{-1} = \begin{pmatrix} -1 & 0 \\ 0 & 4 \end{pmatrix}.$$

2. Decidir si $A = \begin{pmatrix} 3 & 0 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}$ es diagonalizable en $\mathbb{R}^{3\times 3}$.

Busquemos los autovalores de A, es decir, los valores de $\lambda \in \mathbb{R}$ para los cuales el sistema $A.x = \lambda.x$ tiene solución no trivial o, equivalentemente, el sistema $(\lambda.I_3 - A).x = 0$ tiene solución no trivial. Pero esto vale si y sólo si $\det(\lambda.I_3 - A) = 0$, es decir $(\lambda - 3)^3 = 0$. Luego, $\lambda = 3$ es el único autovalor de A.

Si A fuese diagonalizable, existiría $C \in GL(n, K)$ tal que

$$C.A.C^{-1} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix} \iff A = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

Luego, A no es diagonalizable.

6.1.2 Polinomio característico

Como vimos en la sección anterior, un método para determinar si una matriz es diagonalizable consiste en buscar sus autovalores y luego ver si se puede armar una base de autovectores.

Sea $A \in K^{n\times n}$ y sea $\lambda \in K$. Se tiene que:

$$\begin{aligned} \lambda \text{ es autovalor de } A \quad &\iff \quad \exists\, x \in K^n - \{0\} \text{ tal que } A.x = \lambda.x. \\ &\iff \quad \text{El sistema } A.x = \lambda.x \text{ tiene solución no trivial.} \\ &\iff \quad \text{El sistema } (\lambda.I_n - A).x = 0 \text{ tiene solución no trivial.} \\ &\iff \quad \det(\lambda.I_n - A) = 0. \end{aligned}$$

(Comparar con el Ejercicio 13 de la Sección 5.7.)

Definición 6.10 Sea $A \in K^{n\times n}$. Se llama *polinomio característico de* A, y se nota $\mathcal{X}_A$, al polinomio $\mathcal{X}_A = \det(X.I_n - A) \in K[X]$.

Si $A \in K^{n\times n}$, $\mathcal{X}_A$ resulta ser un polinomio mónico de grado n (notar que en la matriz $X.I_n - A$, sólo aparece n veces X y que el signo del término $(X - a_{11}) \dots (X - a_{nn})$ en el determinante es 1). Por lo anterior, tenemos:

Proposición 6.11 *Sea $A \in K^{n\times n}$ y sea $\lambda \in K$. Entonces λ es autovalor de A si y sólo si λ es raíz del polinomio característico de A.*

Ejemplo. Decidir si $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ es diagonalizable en $\mathbb{Q}^{2\times 2}, \mathbb{R}^{2\times 2}$ y $\mathbb{C}^{2\times 2}$.

Los autovalores de A son las raíces del polinomio

$$\mathcal{X}_A = \det \begin{pmatrix} X & -1 \\ 1 & X \end{pmatrix} = X^2 + 1.$$

Como este polinomio no tiene raíces en $\mathbb{Q}$ ni en $\mathbb{R}$, resulta que A no es diagonalizable en $\mathbb{Q}^{2\times 2}$ ni en $\mathbb{R}^{2\times 2}$.

Considerada como matriz en $\mathbb{C}^{2\times 2}$, los autovalores de A son i y $-i$, y los autovectores asociados son $< (1, i) > - \{(0, 0)\}$ y $< (-1, i) > - \{(0, 0)\}$. Como $B = \{(1, i), (-1, i)\}$ es una base de $\mathbb{C}^2$ formada por autovectores de A, entonces A es diagonalizable en $\mathbb{C}^{2\times 2}$.

Queremos definir el polinomio característico asociado a un endomorfismo de un espacio vectorial de dimensión finita. Para eso, veremos que dos matrices semejantes tienen el mismo polinomio característico.

Proposición 6.12 *Sea $A \in K^{n\times n}$ y sea $C \in GL(n, K)$. Entonces $\mathcal{X}_{C.A.C^{-1}} = \mathcal{X}_A$.*

Demostración. Se tiene que

$$\begin{aligned} \mathcal{X}_{C.A.C^{-1}} &= \det(X.I_n - C.A.C^{-1}) \;=\; \det(C.X.I_n.C^{-1} - C.A.C^{-1}) \\ &= \det(C.(X.I_n - A).C^{-1}) \;=\; \det(X.I_n - A) \;=\; \mathcal{X}_A. \end{aligned}$$

□

Definición 6.13 Sea V un K-espacio vectorial de dimensión finita, y sea $f : V \to V$ una transformación lineal. Se define el *polinomio característico* de f como $\mathcal{X}_f = \mathcal{X}_{|f|_B}$, donde B es una base cualquiera de V.

Como en el caso de matrices, se tiene que:

Observación 6.14 Sea V un K-espacio vectorial de dimensión finita. Sea $f : V \to V$ una transformación lineal y sea $\lambda \in K$. Entonces λ es autovalor de f si y sólo si λ es raíz de $\mathcal{X}_f$.

6.2 Una caracterización de matrices diagonalizables

6.2.1 Suma directa de subespacios

Para lo que sigue, vamos a necesitar generalizar la noción de suma directa de dos subespacios de un K-espacio vectorial (ver Sección 1.4.2) al caso de cualquier cantidad finita de subespacios.

Definición 6.15 Sea V un K-espacio vectorial y sean $S_1, S_2, \dots, S_r$ subespacios de V. Se define la *suma* de $S_1, S_2, \dots, S_r$ como

$$W = S_1 + S_2 + \cdots + S_r = \{s_1 + \cdots + s_r \,/\, s_i \in S_i,\ 1 \le i \le r\}.$$

Es fácil ver que W es un subespacio de V.

Definición 6.16 Sea V un K-espacio vectorial y sean $S_1, S_2, \dots, S_r$ subespacios de V. Se dice que $S_1, S_2, \dots, S_r$ están en *suma directa* si, para cada $w \in W = S_1 + S_2 + \cdots + S_r$ existen *únicos* $s_i \in S_i$, $1 \le i \le r$, tales que $w = s_1 + \ldots + s_r$. En este caso se dice que W es la *suma directa de los subespacios* $S_1, \dots, S_r$, y se nota

$$W = S_1 \oplus S_2 \oplus \cdots \oplus S_r = \bigoplus_{i=1}^{r} S_i.$$

Vamos a dar una definición equivalente de la suma directa de varios subespacios:

Proposición 6.17 *Sea V un K-espacio vectorial y sean $S_1, S_2, \dots, S_r$ subespacios de V. Son equivalentes:*

i) $W = \bigoplus_{i=1}^{r} S_i$.

ii) $W = S_1 + \ldots + S_r$ *y para cada* $1 \le j \le r$, *vale*

$$S_j \ \cap\ (S_1 + S_2 + \cdots + S_{j-1} + S_{j+1} + \cdots + S_r) = \{0\}.$$

Demostración.

i) $\Rightarrow$ *ii)* Sea $1 \le j \le r$. Sea $x \in S_j \cap (S_1 + S_2 + \cdots + S_{j-1} + S_{j+1} + \cdots + S_r)$. Entonces

$$\begin{array}{rcl} x & = & 0 + \cdots + \ \ 0 \ \ + x + \ \ 0 \ \ + \cdots + 0, \\ x & = & s_1 + \cdots + s_{j-1} + 0 + s_{j+1} + \cdots + s_r. \end{array}$$

Por la unicidad de la escritura en la suma directa, resulta que $x = 0$.

ii) $\Rightarrow$ *i)* Por hipótesis, existen $s_1, \dots, s_r$ con $s_i \in S_i$ para cada $1 \le i \le r$ tales que $w = \sum_{i=1}^{r} s_i$.

Supongamos que $w = \sum_{i=1}^{r} s_i = \sum_{i=1}^{r} s_i'$ con $s_i, s_i' \in S_i$ para cada $1 \le i \le r$.

Entonces, para cada $1 \le j \le r$, se tiene que

$$s_j - s_j' = \sum_{\substack{i=1 \\ i \ne j}}^{r} (s_i' - s_i).$$

Como $s_j - s_j' \in S_j$ y $\sum_{\substack{i=1 \\ i \ne j}}^{r} (s_i' - s_i)$ es una suma donde cada $s_i' - s_i \in S_i$, de la hipótesis se deduce que $s_j - s_j' = 0$. Luego, $s_j = s_j'$. □

Como en la suma directa de dos subespacios, en este caso también es posible obtener una base del espacio suma uniendo bases de cada uno de los sumandos. La demostración de este resultado es análoga a la de la Proposición 1.46.

Proposición 6.18 *Sea V un K-espacio vectorial y sean $S_1, S_2, \dots, S_r$ subespacios de V. Para cada $1 \le i \le r$, sea B_i una base de S_i. Son equivalentes:*

i) $W = \bigoplus\limits_{i=1}^{r} S_i$.

ii) $B = B_1 \cup B_2 \cup \dots \cup B_r$ *es una base de* W.

Observamos que en la condición ii), B es la *familia* obtenida mediante la unión de las familias $B_1, B_2, \dots, B_r$.

6.2.2 Espacios de autovectores y diagonalización

Dado un autovalor λ de una matriz $A \in K^{n\times n}$, el conjunto de los autovectores de autovalor λ no es un subespacio de K^n, puesto que, por definición, 0 no es un autovector de A. Sin embargo, podemos considerar el siguiente subespacio que consiste en agregar el vector 0 a ese conjunto:

Definición 6.19 Sea $A \in K^{n\times n}$ y sea λ un autovalor de A. Se define

$$E_\lambda = \{v \in K^n \,/\, A.v = \lambda.v\} = \{v \in K^n \,/\, (\lambda\, I_n - A).v = 0\}.$$

Observamos que E_λ es un subespacio de K^n, puesto que es el conjunto de soluciones de un sistema lineal homogéneo.

Proposición 6.20 *Sea $A \in K^{n\times n}$ y sean $\lambda_1, \dots, \lambda_r$ autovalores distintos de A. Entonces $E_{\lambda_1}, \dots, E_{\lambda_r}$ están en suma directa.*

Demostración. Lo probaremos por inducción en la cantidad r de autovalores considerados.

Para $r = 2$: Sean λ_1 y λ_2 autovalores distintos de A. Si $v \in E_{\lambda_1} \cap E_{\lambda_2}$, se tiene que $A.v = \lambda_1.v$ y $A.v = \lambda_2.v$, de donde $(\lambda_1 - \lambda_2).v = 0$. Como $\lambda_1 - \lambda_2 \neq 0$, resulta que $v = 0$. Luego, $E_{\lambda_1} \cap E_{\lambda_2} = \{0\}$ y la suma es directa.

Supongamos ahora que el resultado vale para el caso de r autovalores distintos, y sean $\lambda_1, \dots, \lambda_r, \lambda_{r+1}$ autovalores distintos de A.

Debemos probar que para cada $1 \le i \le r+1$, $E_{\lambda_i} \cap \bigoplus\limits_{\substack{j=1\\ j\neq i}}^{r+1} E_{\lambda_j} = \{0\}$. Sin pérdida de generalidad, supongamos que $i = r+1$.

Sea $v \in E_{\lambda_{r+1}} \cap \bigoplus\limits_{j=1}^{r} E_{\lambda_j}$. Entonces, existen $v_j \in E_{\lambda_j} (1 \le j \le r)$ tales que $v = v_1 + \dots + v_r$. Multiplicando esta igualdad por la matriz A, tenemos

$$\lambda_{r+1} v = \lambda_1 v_1 + \dots + \lambda_r v_r,$$

pero si la multiplicamos por λ_{r+1}, se tiene

$$\lambda_{r+1}v = \lambda_{r+1}v_1 + \ldots + \lambda_{r+1}v_r.$$

Restando miembro a miembro,

$$0 = (\lambda_1 - \lambda_{r+1})v_1 + \ldots + (\lambda_r - \lambda_{r+1})v_r.$$

Como por hipótesis inductiva, los subespacios $E_{\lambda_j}(1 \le j \le r)$ están en suma directa, el vector cero se escribe de forma única como suma de ceros. Luego, $(\lambda_j - \lambda_{r+1})v_j = 0$ para cada $1 \le j \le r$ y, por lo tanto, $v_j = 0$ para cada $1 \le j \le r$, con lo cual $v = 0$. $\square$

Ya vimos que todo autovalor λ de una matriz A es raíz de su polinomio característico. Ahora veremos que siempre existe una relación entre la dimensión de E_λ y la multiplicidad de λ como raíz de $\mathcal{X}_A$.

Proposición 6.21 *Sea $A \in K^{n\times n}$ y sea $\lambda \in K$. Sea r la multiplicidad de λ como raíz de $\mathcal{X}_A$ (es decir, $\mathcal{X}_A = (X-\lambda)^r P$ con $P(\lambda) \ne 0$) y sea $E_\lambda = \{x \in K^n \,/\, A.x = \lambda x\}$. Entonces* $\dim(E_\lambda) \le r$.

Demostración. Sea $f_A : K^n \to K^n$ la transformación lineal definida por $f_A(x) = A.x$. Sea $s = \dim(E_\lambda)$ y sea $\{v_1, \ldots, v_s\}$ una base de E_λ. Sean $v_{s+1}, \ldots, v_n \in K^n$ tales que $B = \{v_1, \ldots, v_s, v_{s+1}, \ldots, v_n\}$ es una base de K^n. Se tiene que

$$|f_A|_B = \left(\begin{array}{cc} \overbrace{\begin{matrix} \lambda & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda \end{matrix}}^{s\times s} & N \\ \mathbf{0} & M \end{array}\right),$$

de donde

$$\begin{aligned} \mathcal{X}_{f_A} &= \det \left(\begin{array}{cc} \overbrace{\begin{matrix} X-\lambda & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & X-\lambda \end{matrix}}^{s\times s} & -N \\ \mathbf{0} & XI_{n-s} - M \end{array}\right) \\ &= (X-\lambda)^s \det(XI_{n-s} - M) \\ &= (X-\lambda)^s Q. \end{aligned}$$

Por hipótesis, $\mathcal{X}_A = (X-\lambda)^r P$ con $P \in K[X]$ tal que $P(\lambda) \neq 0$. Entonces

$$(X-\lambda)^s Q = \mathcal{X}_{f_A} = \mathcal{X}_A = (X-\lambda)^r P \qquad \text{con } P(\lambda) \neq 0,$$

de donde $s \le r$, es decir, $\dim(E_\lambda) \le \text{mult}(\lambda, \mathcal{X}_A)$. □

El siguiente teorema establece condiciones necesarias y suficientes sobre los subespacios E_λ asociados a los autovalores de una matriz para que ésta sea diagonalizable.

Teorema 6.22 *Sea $A \in K^{n\times n}$ y sean $\lambda_1, \dots, \lambda_r$ todos los autovalores de A en K, con $\lambda_i \neq \lambda_j$ si $i \neq j$. Son equivalentes:*

i) *A es diagonalizable en $K^{n\times n}$.*

ii) $\bigoplus_{i=1}^{r} E_{\lambda_i} = K^n$.

iii) $\mathcal{X}_A = (X-\lambda_1)^{a_1} \dots (X-\lambda_r)^{a_r}$ *y* $a_i = \dim E_{\lambda_i}$ *para cada* $1 \le i \le r$.

Demostración.

i) $\Rightarrow$ *ii)* Si A es diagonalizable en $K^{n\times n}$, existe una base $B = \{v_1, \dots, v_n\}$ de K^n formada por autovectores de A. Para cada $v_j \in B$, existe i, $1 \le i \le r$, tal que v_j es autovector de A de autovalor λ_i (porque $\lambda_1, \dots, \lambda_r$ son *todos* los autovalores de A), es decir $v_j \in E_{\lambda_i}$ para algún $1 \le i \le r$, lo que implica que $v_j \in \bigoplus_{i=1}^{r} E_{\lambda_i}$.

En consecuencia, $K^n = < v_1, \dots, v_n > = \bigoplus_{i=1}^{r} E_{\lambda_i}$.

ii) $\Rightarrow$ *iii)* Por la proposición anterior, para cada $1 \le i \le r$, $\dim E_{\lambda_i} \le \text{mult}(\lambda_i, \mathcal{X}_A)$. Si $K^n = \bigoplus_{i=1}^{r} E_{\lambda_i}$ se tiene que

$$n = \dim K^n = \sum_{i=1}^{r} \dim E_{\lambda_i} \le \sum_{i=1}^{r} \text{mult}(\lambda_i, \mathcal{X}_A) \le \text{gr}(\mathcal{X}_A) = n.$$

Luego, en la cadena anterior, son todas igualdades.

En particular:

- La igualdad $\sum_{i=1}^{r} \text{mult}(\lambda_i, \mathcal{X}_A) = \text{gr}(\mathcal{X}_A)$ implica que $\mathcal{X}_A$ se puede factorizar como producto de polinomios de grado 1 en $K[X]$: si $a_i = \text{mult}(\lambda_i, \mathcal{X}_A)$ $(1 \le i \le r)$, entonces $\mathcal{X}_A = (X-\lambda_1)^{a_1} \dots (X-\lambda_r)^{a_r}$.
- Como $\dim E_{\lambda_i} \le \text{mult}(\lambda_i, \mathcal{X}_A)$ para cada $1 \le i \le r$, de la igualdad $\sum_{i=1}^{r} \dim E_{\lambda_i} = \sum_{i=1}^{r} \text{mult}(\lambda_i, \mathcal{X}_A)$ se deduce que $\dim E_{\lambda_i} = \text{mult}(\lambda_i, \mathcal{X}_A)$ para cada $1 \le i \le r$.

iii) $\Rightarrow$ *i)* Para cada $1 \le i \le r$ sea B_i una base de E_{λ_i}. Por la Proposición 6.18, $B = \bigcup_{i=1}^{r} B_i$ es una base de $\bigoplus_{i=1}^{r} E_{\lambda_i} \subseteq K^n$. Ahora,

$$\#B = \sum_{i=1}^{r} \#B_i = \sum_{i=1}^{r} \dim E_{\lambda_i} = \sum_{i=1}^{r} a_i = \mathrm{gr}(\mathcal{X}_A) = n,$$

de donde $\dim \left(\bigoplus_{i=1}^{r} E_{\lambda_i} \right) = \dim K^n$. Luego $\bigoplus_{i=1}^{r} E_{\lambda_i} = K^n$ y entonces B es una base (formada por autovectores de A) de K^n. En consecuencia, A es diagonalizable. $\square$

Ejemplo. Decidir si $A = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix} \in \mathbb{R}^{4\times 4}$ es diagonalizable.

Calculamos $\mathcal{X}_A = (X-1)^3(X-2)$. Para el autovalor 1 se tiene que

$$E_1 = \{x \in \mathbb{R}^4 \,/\, (I_4 - A).x = 0\} = < (0,1,0,0), (0,0,1,0) >,$$

de donde $\dim E_1 = 2 < 3 = \mathrm{mult}(1, \mathcal{X}_A)$. El teorema anterior implica entonces que A no es diagonalizable.

6.3 Polinomios minimales

En lo que sigue, a cada matriz le asociaremos un polinomio. Este polinomio, entre otras propiedades, nos dará un nuevo criterio para decidir si la matriz es diagonalizable.

6.3.1 Polinomio minimal de una matriz

Sea $P \in K[X]$, $P = a_0 + a_1 X + \cdots + a_r X^r$. Dada $A \in K^{n\times n}$ definimos

$$P(A) = a_0 I_n + a_1 A + \cdots + a_r A^r \ \in \ K^{n\times n}.$$

Observemos que si $P, Q \in K[X]$ y $A \in K^{n\times n}$, entonces $(P+Q)(A) = P(A) + Q(A)$ y $(P.Q)(A) = P(A).Q(A)$.

Análogamente, si V es un K-espacio vectorial de dimensión finita y $f : V \to V$ es una transformación lineal, definimos

$$P(f) = a_0 id_V + a_1 f + \cdots + a_r f^r \ \in \ \mathrm{Hom}_K(V,V),$$

donde, para $k \in \mathbb{N}$, $f^k = \underbrace{f \circ f \circ \ldots \circ f}_{k \text{ veces}}$ es la composición.

Dada una matriz $A \in K^{n\times n}$ nos interesa considerar polinomios $P \in K[X]$ que anulen a A, es decir, tales que $P(A) = 0$. El siguiente resultado asegura que para cualquier matriz existe un polinomio no nulo con esta propiedad.

Lema 6.23 *Sea $A \in K^{n\times n}$. Existe un polinomio $P \in K[X]$, $P \neq 0$, tal que $P(A) = 0$.*

Demostración. Consideremos el conjunto $\{I_n, A, A^2, \dots, A^{n^2}\} \subseteq K^{n\times n}$. Este conjunto es linealmente dependiente, puesto que tiene $n^2 + 1$ elementos y $\dim(K^{n\times n}) = n^2$. Luego, existen $a_0, a_1, \dots, a_{n^2} \in K$ no todos nulos, tales que $\sum\limits_{i=0}^{n^2} a_i A^i = 0$. Sea $P \in K[X]$ el polinomio $P = \sum\limits_{i=0}^{n^2} a_i X^i$. Entonces $P \neq 0$ y $P(A) = 0$. □

Para cada matriz, distinguimos un polinomio particular entre todos los polinomios que la anulan: el de grado mínimo y mónico. Veamos que para toda matriz existe un polinomio con estas propiedades y que es único:

Existencia. Sea $H = \{\mathrm{gr}(P) : P \in K[X],\ P \neq 0,\ P(A) = 0\} \subseteq \mathbb{N}$. Por el Lema 6.23, $H \neq \emptyset$. Luego, H tiene primer elemento r. Entonces existe un polinomio $Q \neq 0$ de grado r tal que $Q(A) = 0$ y Q es mónico (si no lo fuera, bastaría dividir por su coeficiente principal).

Unicidad. Supongamos que existe un polinomio $Q' \in K[X]$ mónico, $Q' \neq Q$, tal que $\mathrm{gr}(Q') = r$ y $Q'(A) = 0$. Entonces $(Q' - Q)(A) = 0$ y $\mathrm{gr}(Q' - Q) < r$, puesto que Q y Q' son ambos mónicos, lo que contradice que r es el primer elemento de H.

Definición 6.24 Sea $A \in K^{n\times n}$. Llamaremos *polinomio minimal* de A, y lo notaremos m_A, al polinomio mónico de grado mínimo que anula a A.

Ejemplo. Sea $A = \begin{pmatrix} -1 & 0 \\ 1 & -1 \end{pmatrix}$. Calcular m_A.

Puesto que $\{I_2, A\}$ es linealmente independiente, no existe $P \in \mathbb{R}[X]$ tal que $\mathrm{gr}(P) = 1$ y $P(A) = 0$.

Buscamos entonces $P = X^2 + aX + b$ que anule a A, para lo cual estudiamos la independecia lineal de $\{I_2, A, A^2\}$.

$$\begin{aligned} A^2 + aA + bI_2 = 0 \quad &\Longleftrightarrow \quad \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix} + a \begin{pmatrix} -1 & 0 \\ 1 & -1 \end{pmatrix} + b \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 0 \\ &\Longleftrightarrow \quad \begin{cases} 1 - a + b = 0 \\ -2 + a = 0 \end{cases} \quad \Longleftrightarrow \quad a = 2,\ b = 1. \end{aligned}$$

Luego, $m_A = X^2 + 2X + 1$. (Observar que, en este caso, m_A coincide con $\mathcal{X}_A$.)

Dada una matriz $A \in K^{n\times n}$, el conjunto de todos los polinomios que anulan a A puede caracterizarse a partir de su polinomio minimal m_A.

Proposición 6.25 *Sea $A \in K^{n\times n}$ y sea $P \in K[X]$. Entonces $P(A) = 0$ si y sólo si m_A divide a P.*

Demostración.

($\Leftarrow$) Si $m_A \mid P$, existe $Q \in K[X]$ tal que $P = Q.\, m_A$. Luego

$$P(A) = (Q.\, m_A)(A) = Q(A).\, m_A(A) = Q(A).\, 0 = 0.$$

($\Rightarrow$) Por el algoritmo de división en $K[X]$, existen polinomios $Q, R \in K[X]$ tales que $P = Q.\, m_A + R$ con $R = 0$ o $\mathrm{gr}(R) < \mathrm{gr}(m_A)$. Se tiene que

$$0 = P(A) = Q(A).\, m_A(A) + R(A) = Q(A).\, 0 + R(A) = R(A).$$

Como m_A es de grado mínimo entre los polinomios que anulan a A, no puede ser $\mathrm{gr}(R) < \mathrm{gr}(m_A)$ y, por lo tanto, $R = 0$. □

Como sucede en el caso del polinomio característico, vemos que dos matrices semejantes tienen el mismo polinomio minimal. Para eso, analizamos en primer lugar la relación entre los polinomios que anulan a dos matrices semejantes. Utilizaremos el siguiente hecho:

Observación 6.26 Si $A, B \in K^{n\times n}$ y $C \in GL(n, K)$ son matrices tales que $A = C.\, B.\, C^{-1}$ entonces $A^k = C.\, B^k.\, C^{-1}$ para todo $k \in \mathbb{N}$.

En efecto, es claro que la igualdad vale para $k = 1$ y, que si vale para $k \in \mathbb{N}$, entonces $A^{k+1} = A.\, A^k = C.\, B.\, C^{-1}.\, C.\, B^k.\, C^{-1} = C.\, B^{k+1}.\, C^{-1}$.

Lema 6.27 *Sean $A, B \in K^{n\times n}$, $A \sim B$. Entonces $P(A) \sim P(B)$ para todo $P \in K[X]$. En particular, $P(A) = 0$ si y sólo si $P(B) = 0$.*

Demostración. Sea $P \in K[X]$, $P = \sum_{i=0}^{r} a_i X^i$. Si $A \sim B$, existe $C \in GL(n, K)$ tal que $A = C.B.C^{-1}$, y entonces

$$\begin{aligned} P(A) &= P(C.\, B.\, C^{-1}) = \sum_{i=0}^{r} a_i.(C.\, B.\, C^{-1})^i \\ &= \sum_{i=0}^{r} a_i.\, C.\, B^i.\, C^{-1} = C.\Big(\sum_{i=0}^{r} a_i B^i\Big).\, C^{-1} = C.\, P(B).\, C^{-1}. \end{aligned}$$

Luego, $P(A) \sim P(B)$. □

Proposición 6.28 *Sean $A, B \in K^{n\times n}$, $A \sim B$. Entonces $m_A = m_B$.*

Demostración. Por el lema anterior y la Proposición 6.25,

- $m_A(A) = 0 \Rightarrow m_A(B) = 0 \Rightarrow m_B \mid m_A$.
- $m_B(B) = 0 \Rightarrow m_B(A) = 0 \Rightarrow m_A \mid m_B$.

Puesto que m_A y m_B son ambos mónicos, resulta que $m_A = m_B$. □

Este resultado implica que si $f : V \to V$ es una transformación lineal definida en un K-espacio vectorial V de dimensión finita, y consideramos la matriz de f en dos bases de V distintas, los polinomios minimales de estas dos matrices coinciden. Esto nos permite dar la siguiente definición de polinomio minimal para transformaciones lineales:

Definición 6.29 Sea V un K-espacio vectorial de dimensión finita, y sea B una base de V. Sea $f : V \to V$ una transformación lineal. Se define el *polinomio minimal de* f como $m_f = m_{|f|_B}$.

En la Proposición 6.11 vimos que las raíces del polinomio característico de una matriz son sus autovalores. El siguiente resultado muestra que lo mismo vale para el polinomio minimal.

Proposición 6.30 *Sea $A \in K^{n\times n}$, y sea m_A el polinomio minimal de A. Sea $\lambda \in K$. Entonces λ es autovalor de A si y sólo si λ es raíz de m_A.*

Demostración.

($\Rightarrow$) Sea λ un autovalor de A. Por el algoritmo de división en $K[X]$ existen $Q \in K[X]$ y $R \in K$ tales que $m_A = Q.\,(X - \lambda) + R$. Entonces

$$0 = m_A(A) = Q(A).(A - \lambda I_n) + R.\,I_n.$$

Como λ es autovalor de A, existe $v \in K^n$, $v \neq 0$, tal que $A.v = \lambda.v$. Se tiene que

$$0 = Q(A).(A - \lambda I_n).v + R.v = Q(A).(Av - \lambda\, v) + R.v = Q(A).\,0 + R.v = R.v,$$

es decir, $R.v = 0$. Como $v \neq 0$, debe ser $R = 0$.

En consecuencia, $m_A = Q.\,(X - \lambda)$, y entonces λ es raíz de m_A.

($\Leftarrow$) Sea $\lambda \in K$ una raíz de m_A. Entonces $m_A = (X - \lambda).Q$ y, por lo tanto, $0 = m_A(A) = (A - \lambda I_n).Q(A)$.

Observamos que $Q(A) \neq 0$, puesto que $\operatorname{gr}(Q) = \operatorname{gr}(m_A) - 1$. En consecuencia, existe $w \in K^n$ tal que $Q(A).w \neq 0$. Sea $v = Q(A).w$. Entonces

$$(A - \lambda I_n).v = (A - \lambda I_n).Q(A).w = 0.w = 0,$$

de donde λ es un autovalor de A. □

6.3.2 Polinomio minimal de un vector

Sea $A \in K^{n\times n}$ y sea $v \in K^n$. Dado $P \in K[X]$, definimos $P(v) = P(A).v$. Diremos que P anula a v si $P(v) = 0$.

Observación 6.31 Sea $A \in K^{n\times n}$ y sea m_A el polinomio minimal de A. Entonces, para cada $v \in K^n$, se tiene que

$$m_A(v) = m_A(A).v = 0.v = 0.$$

Luego, para cada $v \in K^n$ existe un polinomio mónico que anula a v.

Definición 6.32 Sea $A \in K^{n\times n}$ y sea $v \in K^n$. El *polinomio minimal de* v, que notaremos m_v, es el polinomio mónico de grado mínimo que anula a v.

La existencia y unicidad de un polinomio que cumple las condiciones de la definición se prueban, a partir de la Observación 6.31, en forma análoga a lo hecho para el polinomio minimal de una matriz.

Ejemplos.

1. Sea $A \in K^{n\times n}$ y sea $v \in K^n$ un autovector de A de autovalor λ. Entonces $m_v = X - \lambda$.

2. Sea $A = \begin{pmatrix} -1 & 0 \\ 1 & -1 \end{pmatrix} \in \mathbb{R}^{2\times 2}$, y sea $e_1 = (1, 0)$. Calcular m_{e_1}.

   Comenzamos buscando un polinomio mónico $P = X + b \in \mathbb{R}[X]$ de grado 1 tal que $P(e_1) = 0$: Observamos que

   $$P(e_1) = 0 \iff (A + bI_2).e_1 = 0 \iff A.e_1 + b.e_1 = 0 \iff (-1, 1) + (b, 0) = (0, 0),$$

   pero esto no vale para ningún valor de $b \in \mathbb{R}$. Luego, $\text{gr}(m_{e_1}) \geq 2$.

   Buscamos entonces un polinomio $P \in \mathbb{R}[X]$ mónico de grado 2: $P = X^2 + aX + b$. En este caso

   $$\begin{aligned} P(e_1) = 0 &\iff (A^2 + aA + bI_2).e_1 = 0 \\ &\iff A^2 e_1 + a.Ae_1 + b.e_1 = 0 \\ &\iff (1, -2) + a(-1, 1) + b(1, 0) = (0, 0) \\ &\iff \begin{cases} 1 - a + b = 0 \\ -2 + a = 0 \end{cases} \\ &\iff a = 2,\ b = 1. \end{aligned}$$

   Luego, el polinomio minimal de e_1 es $P = X^2 + 2X + 1$.

Cómo hallar el polinomio minimal de un vector:

Sea $A \in K^{n\times n}$ y sea $v \in K^n$. Si $m_v = X^m + a_{m-1}X^{m-1} + \cdots + a_2X^2 + a_1X + a_0$ es el polinomio minimal de v, entonces

$$A^m.v + a_{m-1}A^{m-1}.v + \cdots + a_2A^2.v + a_1A.v + a_0.v = 0.$$

Además, como m_v es de grado mínimo entre los polinomios que satisfacen $P(v) = 0$, se tiene que $\{v, A.v, \ldots, A^{m-1}.v\}$ es un conjunto linealmente independiente.

Luego, para hallar el polinomio minimal de v, un primer paso consiste en buscar el mínimo $m \in \mathbb{N}$ tal que $\{v, A.v, \ldots, A^m.v\}$ es linealmente dependiente. Si $A^m.v = -a_0.v - a_1.A.v - \cdots - a_{m-1}.A^{m-1}.v$, entonces $m_v = X^m + a_{m-1}X^{m-1} + \cdots + a_1X + a_0$.

Al igual que para el polinomio minimal de una matriz (ver Proposición 6.25), fijada una matriz A, todo polinomio que anula a un vector dado resulta ser múltiplo de su polinomio minimal.

Proposición 6.33 *Sean $A \in K^{n\times n}$, $v \in K^n$ y $P \in K[X]$. Entonces $P(v) = 0$ si y sólo si m_v divide a P. En particular, m_v divide a m_A.*

Demostración. Dado $P \in K[X]$ se tiene que $P = Q.\, m_v + R$ con $Q, R \in K[X]$, $R = 0$ o $\text{gr}(R) < \text{gr}(m_v)$. En consecuencia,

$$P(v) = P(A).\, v = Q(A).\, m_v(A).\, v + R(A).\, v = Q(A).\, 0 + R(v) = R(v),$$

de donde $P(v) = 0$ si y sólo si $R(v) = 0$.

Como m_v es de grado mínimo entre los polinomios que anulan a v, resulta que $R(v) = 0$ si y sólo si $R = 0$, es decir, si y sólo si $m_v \mid P$. □

La siguiente proposición muestra cómo calcular el polinomio minimal de una matriz $A \in K^{n\times n}$ a partir de los polinomios minimales de los vectores de una base de K^n.

Proposición 6.34 *Sea $A \in K^{n\times n}$ y sea $B = \{v_1, \ldots, v_n\}$ una base de K^n. Entonces $m_A = \text{mcm}\{m_{v_i} : i = 1, \ldots, n\}$ (*mcm *denota el mínimo común múltiplo).*

Demostración. Sea $P = \text{mcm}\{m_{v_i} : i = 1, \ldots, n\}$.

Por la proposición anterior, $m_{v_i} \mid m_A$ para cada $1 \le i \le n$. Luego, $P \mid m_A$.

Por otro lado, como $m_{v_i} \mid P$ para cada $1 \le i \le n$, se tiene que $P(A).v_i = 0$ para cada $1 \le i \le n$. Sea $v \in K^n$ y sean $\alpha_1, \ldots, \alpha_n \in K$ tales que $v = \sum_{i=1}^{n} \alpha_i v_i$. Entonces

$$P(A).v = P(A)\Big(\sum_{i=1}^{n} \alpha_i.v_i\Big) = \sum_{i=1}^{n} \alpha_i P(A).v_i = 0.$$

En consecuencia, se tiene que $P(A) \in K^{n\times n}$ satisface $P(A).v = 0 \ \forall v \in K^n$ y, por lo tanto, $P(A) = 0$. Luego, $m_A \mid P$.

Como m_A y P son dos polinomios mónicos que se dividen mutuamente, $m_A = P$. $\square$

Ejemplo. Calcular m_A para $A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$.

Consideremos $E = \{e_1, e_2, e_3\}$ la base canónica de $\mathbb{R}^3$. Por la proposición anterior, $m_A = \text{mcm}\{m_{e_1}, m_{e_2}, m_{e_3}\}$.

Busquemos entonces los polinomios minimales de cada uno de los vectores de E:

m_{e_1}: Vemos que $\{e_1, A.e_1\} = \{e_1, e_1 + e_2\}$ es un conjunto linealmente independiente, pero $\{e_1, A.e_1, A^2.e_1\} = \{e_1,\ e_1 + e_2,\ e_1 + 2e_2\}$ no lo es.

Además, $0 = (e_1 + 2e_2) - 2.(e_1 + e_2) + e_1 = A^2.e_1 - 2.A.e_1 + e_1$, de donde $m_{e_1} = X^2 - 2X + 1 = (X - 1)^2$.

m_{e_2}: Como $A.e_2 = e_2$, entonces $m_{e_2} = X - 1$.

m_{e_3}: Como $A.e_3 = 2.e_3$, entonces $m_{e_3} = X - 2$.

Luego, $m_A = \text{mcm}\{(X - 1)^2, X - 1, X - 2\} = (X - 1)^2(X - 2)$.

6.3.3 Teorema de Hamilton-Cayley

El teorema de Hamilton-Cayley, cuya demostración damos a continuación, establece que para cualquier matriz A, el polinomio característico de A anula a A.

Teorema 6.35 (Teorema de Hamilton-Cayley) *Sea $A \in K^{n\times n}$ y sea $\mathcal{X}_A$ el polinomio característico de A. Entonces $m_A \mid \mathcal{X}_A$ (lo que es equivalente a que $\mathcal{X}_A(A) = 0$).*

Demostración. Sea $f_A : K^n \to K^n$ la transformación lineal definida por $f_A(x) = A.x$.

Sea $v \in K^n$. Supongamos que $\{v, f_A(v), \dots, f_A^k(v)\}$ es un conjunto linealmente independiente y que $f_A^{k+1}(v) = (-a_k)f_A^k(v) + \cdots + (-a_1)f_A(v) + (-a_0)$. Entonces $m_v = X^{k+1} + a_kX^k + \cdots + a_1X + a_0$.

Extendemos el conjunto $\{v, f_A(v), \dots, f_A^k(v)\}$ a una base de K^n: sean $w_{k+2}, \dots, w_n \in K^n$ tales que

$$B = \{v, f_A(v), \dots, f_A^k(v), w_{k+2}, \dots, w_n\}$$

es una base de K^n. Se tiene que

$$|f_A|_B = \left(\begin{array}{ccccc|c} 0 & 0 & \dots & 0 & -a_0 & \\ 1 & 0 & & 0 & -a_1 & \\ 0 & 1 & & \vdots & \vdots & M \\ \vdots & & \ddots & 0 & -a_{k-1} & \\ 0 & \dots & & 1 & -a_k & \\ & & 0 & & & N \end{array} \right).$$

Sabemos que $\mathcal{X}_A = \mathcal{X}_{f_A} = \mathcal{X}_{|f_A|_B}$, con lo que

$$\begin{aligned} \mathcal{X}_A &= \det \left(\begin{array}{ccccc|c} X & 0 & \dots & 0 & a_0 & \\ -1 & X & & 0 & a_1 & \\ 0 & -1 & & \vdots & \vdots & -M \\ \vdots & & \ddots & X & a_{k-1} & \\ 0 & \dots & & -1 & X + a_k & \\ & & 0 & & & X.\, I_{n-k-1} - N \end{array} \right) \\ &= \det \begin{pmatrix} X & 0 & \dots & 0 & a_0 \\ -1 & X & & 0 & a_1 \\ 0 & -1 & & \vdots & \vdots \\ \vdots & & \ddots & X & a_{k-1} \\ 0 & \dots & & -1 & X + a_k \end{pmatrix} \det(X.\, I_{n-k-1} - N). \end{aligned}$$

Por lo calculado en el primer ejemplo de la Sección 5.4, obtenemos que

$$\begin{aligned} \mathcal{X}_A &= (X^{k+1} + a_k X^k + \dots + a_1 X + a_0).\, \det(X.\, I_{n-k-1} - N) \\ &= m_v.\, \det(X.\, I_{n-k-1} - N). \end{aligned}$$

Por lo tanto, $m_v \mid \mathcal{X}_A$ para $v \in K^n$ arbitrario.

Sea $E = \{e_1, \dots, e_n\}$ la base canónica de K^n. Por lo anterior, $m_{e_i} \mid \mathcal{X}_A$ para cada $1 \le i \le n$. En consecuencia, $m_A = \mathrm{mcm}\{m_{e_1}, \dots, m_{e_n}\}$ divide a $\mathcal{X}_A$. $\square$

A partir del Teorema de Hamilton-Cayley podemos deducir algunos resultados relacionados con el polinomio minimal de una matriz.

Observación 6.36 Sea $A \in K^{n\times n}$. Entonces:

1. $\mathrm{gr}(m_A) \le n$.

2. Si $\text{gr}(m_A) = n$, entonces $m_A = \mathcal{X}_A$.
3. Si existe $v \in K^n$ tal que $\text{gr}(m_v) = n$, entonces $m_v = m_A = \mathcal{X}_A$.

En el ejemplo que sigue, mostramos cómo puede utilizarse el hecho de conocer un polinomio que anula a una matriz (en este caso, su polinomio característico) para calcular las potencias de la matriz.

Ejemplo. Sea $A = \begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix} \in \mathbb{R}^{2\times 2}$. Calcular A^n para cada $n \in \mathbb{N}$.

Calculamos el polinomio característico de A: $\mathcal{X}_A = (X-1)^2$. Por el Teorema de Hamilton-Cayley, $\mathcal{X}_A(A) = 0$, es decir, $A^2 - 2.A + I_2 = 0$.

La idea para calcular A^n para cada $n \in \mathbb{N}$ es ver que cada potencia de A es combinación lineal de I_2 y A (observar que la igualdad dada por el teorema de Hamilton-Cayley dice que $A^2 = 2A - I_2$), encontrar esta combinación lineal y usarla para dar una fórmula cerrada para los coeficientes de A^n en función de n.

Sea $n \in \mathbb{N}$. Por el algoritmo de división, existen $P(X) \in \mathbb{R}[X]$ y $a_n, b_n \in \mathbb{R}$ tales que

$$X^n = (X-1)^2 P(X) + a_n.X + b_n. \tag{6.1}$$

Evaluando esta igualdad en $X = 1$, se obtiene

$$1 = 0.\,P(1) + a_n + b_n \iff a_n + b_n = 1.$$

Derivando la identidad (6.1) resulta que

$$n.X^{n-1} = 2(X-1)P + (X-1)^2 P' + a_n,$$

y evaluando esta igualdad en $X = 1$, se deduce que $a_n = n$. En consecuencia, $b_n = 1 - n$.

Luego, para cada $n \in \mathbb{N}$ se tiene que

$$X^n = (X-1)^2 P(X) + nX + 1 - n,$$

de donde

$$A^n = n.A + (1-n)I_2 = n.\begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix} + (1-n)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1-n & -n \\ n & n+1 \end{pmatrix}.$$

Aplicando el Teorema de Hamilton-Cayley, también podemos calcular A^{-1} como combinación lineal de I_2 y A. Basta observar que

$$A^2 - 2A + I_2 = 0 \iff I_2 = 2A - A^2 \iff I_2 = A(2I_2 - A) = (2I_2 - A)A,$$

lo que implica que $A^{-1} = 2I_2 - A$.

El procedimiento para calcular A^{-1} del ejemplo anterior puede generalizarse para hallar la inversa de cualquier matriz en $GL(n, K)$:

Observación 6.37 Sea $A \in GL(n, K)$. Entonces $A^{-1} \in < I_n, A, A^2, \dots, A^{n-1} >$.

Supongamos que $\mathcal{X}_A = X^n + a_{n-1}X^{n-1} + \cdots + a_1X + a_0$. Entonces, por el Teorema de Hamilton-Cayley,

$$A^n + a_{n-1}A^{n-1} + \cdots + a_1A + a_0I_n = 0.$$

Observemos que $a_0 = \mathcal{X}_A(0) = \det(0.I_n - A) = \det(-A) = (-1)^n \det(A) \neq 0$, puesto que A es inversible. En consecuencia,

$$\begin{aligned} I_n &= \frac{-1}{a_0} . (A^n + a_{n-1}A^{n-1} + \cdots + a_2A^2 + a_1A) \\ &= A. \Big(\frac{-1}{a_0} . (A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_2A + a_1I_n) \Big) \\ &= \Big(\frac{-1}{a_0} . (A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_2A + a_1I_n) \Big) . A, \end{aligned}$$

de donde $A^{-1} = \dfrac{-1}{a_0} . (A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_2A + a_1I_n)$.

6.3.4 Un criterio de diagonalización usando el polinomio minimal

En primer lugar, notemos que vale lo siguiente:

Observación 6.38 Sea $A \in K^{n\times n}$ y sea $\mathcal{X}_A \in K[X]$ el polinomio característico de A. Si $\mathcal{X}_A$ se factoriza linealmente en $K[X]$ y tiene todas sus raíces simples, entonces A es diagonalizable.

En efecto, si $\mathcal{X}_A = (X-\lambda_1) \dots (X-\lambda_n)$, con $\lambda_i \neq \lambda_j$ para cada $i \neq j$, A tiene n autovalores distintos. Además, para cada $1 \leq i \leq n$, existe $v_i \in K^n$ autovector de A de autovalor λ_i. Puesto que $E_{\lambda_1}, \dots, E_{\lambda_n}$ están en suma directa, $\{v_1, \dots, v_n\}$ resulta una base de K^n formada por autovectores de A, y por lo tanto, A es diagonalizable.

La recíproca de este resultado no es cierta, es decir, $\mathcal{X}_A$ puede tener raíces múltiples y, de todas formas, A ser diagonalizable. Basta considerar, por ejemplo, la matriz $A = I_n$.

Sin embargo, es posible dar una condición necesaria y suficiente para la diagonalización, considerando la factorización del polinomio minimal de la matriz.

Proposición 6.39 *Sea $A \in K^{n\times n}$. Entonces A es diagonalizable en $K^{n\times n}$ si y sólo si m_A tiene todas sus raíces en K y son simples.*

Demostración.

($\Rightarrow$) Supongamos que A es diagonalizable. Sean $\lambda_1, \dots, \lambda_r$ los autovalores de A, con $\lambda_i \neq \lambda_j$ si $i \neq j$ y sea $\{v_1, \dots, v_n\}$ una base de autovectores de A. Si m_v denota el minimal del vector v para la matriz A, se tiene que:

$$\begin{aligned} m_A &= \text{mcm}\{m_{v_1}, \dots, m_{v_n}\} \\ &= \text{mcm}\{X - \lambda_1, \dots, X - \lambda_1, \dots, X - \lambda_r, \dots, X - \lambda_r\} \\ &= (X - \lambda_1)(X - \lambda_2) \dots (X - \lambda_r) \end{aligned}$$

En consecuencia, m_A tiene todas las raíces en K y son simples.

($\Leftarrow$) Supongamos que $m_A = (X-\lambda_1)\dots(X-\lambda_r)$ con $\lambda_i \neq \lambda_j$ si $i \neq j$. Entonces $\lambda_1,\dots,\lambda_r$ son todos los autovalores de A en K.

Veamos que $K^n = \bigoplus_{i=1}^{r} E_{\lambda_i}$, donde $E_{\lambda_i} = \{x \in K^n : A.x = \lambda_i.x\}$.

Sea $v \in K^n - \{0\}$. Consideremos el subespacio

$$S = < v, A.v, A^2.v, \dots, A^m.v, \dots > \subseteq K^n.$$

Supongamos que $\{v, A.v, \dots, A^k.v\}$ es un conjunto linealmente independiente, pero $\{v, A.v, \dots, A^k.v, A^{k+1}.v\}$ es linealmente dependiente.

Luego $A^j.v \in < v, A.v, \dots, A^k.v >$ para todo $j \in \mathbb{N}$ y $B_S = \{v, A.v, A^2.v, \dots, A^k.v\}$ resulta ser una base de S. Además, $\mathrm{gr}(m_v) = k+1$, o sea, m_v es de la forma

$$m_v = X^{k+1} + a_kX^k + \dots + a_1X + a_0.$$

Por la construcción de S, si $x \in S$ resulta que $A.x \in S$. Sea $f_A : S \to S$ la transformación lineal definida por $f_A(x) = A.x$. Se tiene que

$$|f_A|_{B_S} = \begin{pmatrix} 0 & 0 & \dots & 0 & -a_0 \\ 1 & 0 & & \vdots & -a_1 \\ 0 & 1 & \ddots & & \vdots \\ \vdots & \vdots & \ddots & 0 & -a_{k-1} \\ 0 & 0 & \dots & 1 & -a_k \end{pmatrix}.$$

Observamos que $\mathcal{X}_{f_A} = \det(XI_{k+1} - |f_A|_{B_S}) = X^{k+1} + a_kX^k + \dots + a_1X + a_0 = m_v$.

Puesto que $m_v \mid m_A$ y, por hipótesis, m_A tiene todas sus raíces en K y son simples, resulta que $\mathcal{X}_{f_A} = m_v$ tiene todas sus raíces en K, son simples y son algunos de los autovalores de A. Por la Observación 6.38 concluimos que f_A es diagonalizable sobre S. Además, si $\mathcal{X}_{f_A} = (X-\lambda_{i_1})\dots(X-\lambda_{i_{k+1}})$, como $v \in S$, existen $v_{i_1},\dots,v_{i_{k+1}}$ autovectores de f_A (donde v_{i_j} es un autovector de autovalor λ_{i_j}) tales que $v = v_{i_1} + \dots + v_{i_{k+1}}$. Pero si v_{i_j} es un autovector de f_A de autovalor λ_{i_j}, entonces es un autovector de A con el mismo autovalor. En consecuencia $v \in \bigoplus_{i=1}^{r} E_{\lambda_i}$.

Como $v \in K^n$ era arbitrario, resulta que $K^n = \bigoplus_{i=1}^{r} E_{\lambda_i}$. La equivalencia dada por el Teorema 6.22 dice entonces que A es diagonalizable. $\square$

A continuación presentamos un ejemplo en el que mostramos cómo puede aplicarse este resultado para determinar si una matriz es diagonalizable.

Ejemplo. Sea $A \in \mathbb{C}^{n\times n}$ tal que $A^k = I_n$ para algún $k \in \mathbb{N}$. Entonces A es diagonalizable.

Por hipótesis, $A^k - I_n = 0$, con lo que el polinomio $X^k - 1$ anula a A. En consecuencia, $m_A \mid X^k - 1$. Como $X^k - 1$ tiene todas sus raíces en $\mathbb{C}$ y son simples, entonces m_A también. Luego, A es diagonalizable.

Notar que si $A \in \mathbb{R}^{n\times n}$ satisface $A^k = I_n$, A no es necesariamente diagonalizable sobre $\mathbb{R}$. Por ejemplo, $A = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ cumple $A^3 = I_3$, pero A no es diagonalizable, puesto que $m_A = (X-1)(X^2+X+1)$, que no tiene todas sus raíces en $\mathbb{R}$.

6.4 Subespacios invariantes

Dada una transformación lineal $f : V \to V$ donde V es un K-espacio vectorial de dimensión finita, una posible manera de estudiar f consiste en descomponer el espacio V como suma directa de subespacios $V = \bigoplus_{i=1}^{r} S_i$ y analizar la restricción de f a cada uno de estos subespacios S_i. Ahora, para que esto sea posible, es necesario que la restricción de f a cada S_i sea una transformación lineal de S_i en S_i, es decir, que la imagen por f del subespacio esté incluida en el mismo.

Definición 6.40 Sea V un K-espacio vectorial y sea $f : V \to V$ una transformación lineal. Un subespacio $S \subseteq V$ se dice *invariante por f* (o *f-invariante*) si $f(S) \subseteq S$.

Ejemplos.

1. Sea $A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix} \in \mathbb{R}^{4\times 4}$.

   Si $E = \{e_1, e_2, e_3, e_4\}$ es la base canónica de $\mathbb{R}^4$, algunos subespacios invariantes por f_A son: $< e_1, e_2 >$, $< e_2 >$, $< e_3 >$, $< e_4 >$, $< e_3, e_4 >$, $< e_1, e_2, e_4 >$.

2. Hallar todos los subespacios invariantes por f_A siendo $A = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} \in \mathbb{R}^{2\times 2}$.

   (a) Subespacios invariantes de dimensión 0: $S = \{0\}$.

   (b) Subespacios invariantes de dimensión 2: $S = \mathbb{R}^2$.

   (c) Subespacios invariantes de dimensión 1: Observamos que $S = < v >$ es un subespacio de dimensión 1 invariante por f_A si y sólo si $v \neq 0$ y $A.v \in < v >$ o, equivalentemente, $v \neq 0$ y $A.v = \lambda.v$ para algún $\lambda \in \mathbb{R}$.

   Luego, un subespacio $S = < v >$ de dimensión 1 es f_A-invariante si y sólo si v es autovector de A.

   Es fácil ver entonces que el único subespacio f_A-invariante de dimensión 1 es $S = < (0,1) >$.

En la siguiente proposición se prueban algunas propiedades sobre subespacios invariantes.

Proposición 6.41 *Sea V un K-espacio vectorial. Sea $f : V \to V$ una transformación lineal. Entonces:*

i) $\mathrm{Nu}(f)$ *e* $\mathrm{Im}(f)$ *son subespacios f-invariantes de V.*

ii) *S es un subespacio f-invariante de V de dimensión 1 si y sólo si $S = < v >$ con $v \in V$ un autovector de f.*

iii) *Si S y T son subespacios f-invariantes de V, entonces $S \cap T$ y $S + T$ son subespacios f-invariantes de V.*

Demostración.

i) Se tiene que $f(\mathrm{Nu}(f)) = \{0\} \subseteq \mathrm{Nu}(f)$, con lo que $\mathrm{Nu}(f)$ es invariante por f.

Además, es claro que $f(\mathrm{Im}(f)) \subseteq \mathrm{Im}(f)$, de donde $\mathrm{Im}(f)$ es invariante por f.

ii) Sea S un subespacio f-invariante de V de dimensión 1. Entonces $S = < v >$ para algún $v \in V$, $v \neq 0$, tal que $f(v) \in < v >$. Esta última condición implica que $f(v) = \lambda. v$ para algún $\lambda \in K$, y siendo $v \neq 0$, resulta que v es un autovector de f.

Recíprocamente, si $S = < v >$ con v un autovector de f, como $v \neq 0$, entonces $\dim S = 1$, y como $f(v) = \lambda. v$ para algún $\lambda \in K$, resulta que $f(S) \subseteq S$. Luego, S es un subespacio f-invariante de V de dimensión 1.

iii) Sean S y T subespacios de V invariantes por f. Entonces $f(S \cap T) \subseteq f(S) \subseteq S$, puesto que S es f-invariante. Análogamente, $f(S \cap T) \subseteq T$. En consecuencia, $f(S \cap T) \subseteq S \cap T$, de donde $S \cap T$ es invariante por f.

Para $S + T$, teniendo en cuenta que f es una transformación lineal y que S y T son invariantes por f, se tiene que $f(S + T) \subseteq f(S) + f(T) \subseteq S + T$ y, por lo tanto, $S + T$ es invariante por f. □

Dada una transformación lineal $f : V \to V$ y un subespacio S de V invariante por f, la restricción de f a S, que notaremos $f_{|S}$ resulta ser una transformación lineal de S en S. En lo que sigue, analizamos la relación entre los polinomios minimales y característicos de esta restricción $f_{|S}$ y los de f.

Proposición 6.42 *Sea V un K-espacio vectorial de dimensión finita, sea $f : V \to V$ una transformación lineal y sea $S \subseteq V$ un subespacio invariante por f. Sea $f_{|S} : S \to S$ la restricción. Entonces:*

i) $m_{f_{|S}} \mid m_f$.

ii) $\mathcal{X}_{f_{|S}} \mid \mathcal{X}_f$.

Demostración. Sean $n = \dim V$ y $s = \dim S$. Sea $B_S = \{v_1, \dots, v_s\}$ una base de S y sean $v_{s+1}, \dots, v_n \in V$ tales que $B = \{v_1, \dots, v_s, v_{s+1}, \dots, v_n\}$ es una base de V.

i) Sabemos que

$$\begin{aligned} m_{f_{|S}} &= \operatorname{mcm}\{m_{v_1}, \dots, m_{v_s}\} \\ m_f &= \operatorname{mcm}\{m_{v_1}, \dots, m_{v_s}, m_{v_{s+1}}, \dots, m_{v_n}\}. \end{aligned}$$

Ahora, como $m_{v_i} \mid m_f$ para cada $1 \le i \le s$, el mcm de estos polinomios también lo divide, es decir, $m_{f_{|S}} \mid m_f$.

ii) Para la base B considerada, se tiene que

$$|f|_B = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix} \in K^{n\times n} \qquad \text{con } A = |f_{|S}|_{B_S} \in K^{s\times s}.$$

Luego,

$$\begin{aligned} \mathcal{X}_f = \mathcal{X}_{|f|_B} &= \det \begin{pmatrix} X.I_s - A & -B \\ 0 & X.I_{n-s} - C \end{pmatrix} \\ &= \det(X.I_s - A).\det(X.I_{n-s} - C) \;=\; \mathcal{X}_{f_{|S}}.Q, \end{aligned}$$

con lo que $\mathcal{X}_{f_{|S}} \mid \mathcal{X}_f$. □

Si f es una transformación lineal definida en un espacio V de dimensión finita y el espacio V puede descomponerse como suma directa de subespacios f-invariantes, entonces existe una base de V en la que la matriz de f tiene forma diagonal por bloques. Más precisamente:

Observación 6.43 Sea V un K-espacio vectorial de dimensión finita y sea $f : V \to V$ una transformación lineal. Sean S y T subespacios de V invariantes por f tales que $S \oplus T = V$.

Supongamos que $\dim(S) = s > 0$, $\dim(T) = t > 0$. Sean $B_S = \{v_1, \dots, v_s\}$ y $B_T = \{w_1, \dots, w_t\}$ bases de S y T respectivamente. Entonces

$$B = B_S \cup B_T = \{v_1, \dots, v_s, w_1, \dots, w_t\}$$

es una base de V y

$$|f|_B = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix},$$

donde $A_1 \in K^{s\times s}$ y $A_2 \in K^{t\times t}$. Más aún, si $f_{|S} : S \to S$ y $f_{|T} : T \to T$ son las restricciones, se tiene que $A_1 = |f_{|S}|_{B_S}$ y $A_2 = |f_{|T}|_{B_T}$.

Esto nos lleva a la siguiente definición:

Definición 6.44 Sea V un K-espacio vectorial y sea $f : V \to V$ una transformación lineal. Sea $S \subseteq V$ un subespacio f-invariante. Un *complemento invariante para S* es un subespacio T de V tal que T es f-invariante y $S \oplus T = V$.

Observamos que dada una transformación lineal $f : V \to V$ y un subespacio S de V invariante por f no siempre existe un complemento invariante para S.

Por ejemplo, sea $f : \mathbb{R}^2 \to \mathbb{R}^2$ definida por $f(x, y) = (0, x)$. Entonces $S = < (0, 1) >$ es f-invariante ($S = \text{Nu}(f)$), pero no admite un complemento invariante. En efecto, si $S \oplus T = \mathbb{R}^2$, entonces $\dim T = 1$. Luego, si T es f-invariante, $T = < v >$ con $v \in \mathbb{R}^2$ un autovector de f. Ahora, el conjunto de autovectores de f es $S - \{0\}$, con lo que $v \in S$, contradiciendo que T es un complemento para S.

En el caso en que V sea suma directa de subespacios f-invariantes, podemos relacionar los polinomios característico y minimal de f con los de sus restricciones a estos subespacios:

Proposición 6.45 *Sea V un K-espacio vectorial de dimensión finita y sea $f : V \to V$ una transformación lineal. Sean S y T subespacios f-invariantes de V tales que $S \oplus T = V$. Entonces:*

i) $\mathcal{X}_f = \mathcal{X}_{f_{|S}} . \mathcal{X}_{f_{|T}}$

ii) $m_f = \text{mcm}(m_{f_{|S}}, m_{f_{|T}})$

Demostración.

i) Se deduce inmediatamente de la Observación 6.43.

ii) Sea $P = \text{mcm}(m_{f_{|S}}, m_{f_{|T}})$. Puesto que S y T son subespacios f-invariantes de V, por la Proposición 6.42, se tiene que $m_{f_{|S}} \mid m_f$ y $m_{f_{|T}} \mid m_f$. Luego $P \mid m_f$.

Por otro lado, por la Observación 6.43, si B_S y B_T son bases de S y T respectivamente, y $B = B_S \cup B_T$, entonces
$$|f|_B = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}$$
con $A_1 = |f_{|S}|_{B_S}$ y $A_2 = |f_{|T}|_{B_T}$.

Como $m_{f_{|S}} \mid P$ y $m_{f_{|T}} \mid P$, resulta que $P(A_1) = 0$ y $P(A_2) = 0$. Entonces, operando por bloques,
$$P(|f|_B) = \begin{pmatrix} P(A_1) & 0 \\ 0 & P(A_2) \end{pmatrix} = 0,$$
de donde $m_f \mid P$.

Luego $P = m_f$, puesto que P y m_f son dos polinomios mónicos que se dividen mutuamente. □

6.5 Ejercicios

Ejercicio 1. Calcular el polinomio característico, los autovalores y los autovectores de la matriz A en cada uno de los siguientes casos ($a \in \mathbb{R}$):

(Analizar por separado los casos $K = \mathbb{R}$ y $K = \mathbb{C}$)

i) $A = \begin{pmatrix} 1 & 3 \\ -3 & -1 \end{pmatrix}$ ii) $A = \begin{pmatrix} 1 & 3 \\ 3 & -1 \end{pmatrix}$ iii) $A = \begin{pmatrix} 0 & a \\ -a & 0 \end{pmatrix}$

iv) $A = \begin{pmatrix} 0 & 2 & 1 \\ -2 & 0 & 3 \\ -1 & -3 & 0 \end{pmatrix}$ v) $A = \begin{pmatrix} 3 & 1 & 0 \\ -4 & -1 & 0 \\ 4 & -8 & -2 \end{pmatrix}$ vi) $A = \begin{pmatrix} a & 1 & 1 \\ 1 & a & 1 \\ 1 & 1 & a \end{pmatrix}$

vii) $A = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}$ viii) $A = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ ix) $A = \begin{pmatrix} 1 & 0 & 0 \\ a & 1 & 0 \\ 0 & a & 1 \end{pmatrix}$

Ejercicio 2. Para cada una de las matrices A del ejercicio anterior, sea U una base de K^n y sea $f : K^n \to K^n$ la tranformación lineal tal que $|f|_U = A$. Decidir si es posible encontrar una base B de K^n tal que $|f|_B$ sea diagonal. En caso afirmativo, calcular $C(U, B)$.

Ejercicio 3. Sea $f : \mathbb{R}^3 \to \mathbb{R}^3$ la transformación lineal definida por:

$$f(x, y, z) = (-x - 2.y + 2.z, -y, -x - 3.y - 4.z)$$

i) Encontrar una base B de $\mathbb{R}^3$ tal que $|f|_B$ sea diagonal.

ii) Calcular $\begin{pmatrix} -1 & -2 & 2 \\ 0 & -1 & 0 \\ -1 & -3 & -4 \end{pmatrix}^n$, $\forall n \in \mathbb{N}$. (Sugerencia: ver Observación 6.26.)

iii) ¿Existe una matriz $P \in \mathbb{R}^{3\times 3}$ tal que $P^2 = \begin{pmatrix} -1 & -2 & 2 \\ 0 & -1 & 0 \\ -1 & -3 & -4 \end{pmatrix}$?

Ejercicio 4.

i) Sea $A = \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \in K^{2\times 2}$. Determinar todos los a, b y $c \in K$ para los que A es diagonalizable.

ii) Probar que toda matriz $A \in \mathbb{C}^{2\times 2}$ es diagonalizable o bien es semejante a una matriz del tipo $\begin{pmatrix} \alpha & 0 \\ 1 & \alpha \end{pmatrix}$.

Ejercicio 5. Diagonalizar las matrices $A \in \mathbb{R}^{n\times n}$ y $B \in \mathbb{R}^{6\times 6}$ encontrando sus autovectores:

$$A = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & 1 & 1 & \dots & 1 \\ 1 & 1 & 1 & \dots & 1 \\ \dots & \dots & \dots & \dots & \dots \\ 1 & 1 & 1 & \dots & 1 \end{pmatrix} \qquad \text{y} \qquad B = \begin{pmatrix} 2 & 1 & 2 & 1 & 2 & 1 \\ 1 & 2 & 1 & 2 & 1 & 2 \\ 2 & 1 & 2 & 1 & 2 & 1 \\ 1 & 2 & 1 & 2 & 1 & 2 \\ 2 & 1 & 2 & 1 & 2 & 1 \\ 1 & 2 & 1 & 2 & 1 & 2 \end{pmatrix}$$

Ejercicio 6. Se sabe que la matriz $A \in \mathbb{R}^{2\times 2}$ tiene a $(1, -1)$ como autovector de autovalor $\sqrt{2}$ y, además, $\mathcal{X}_A \in \mathbb{Q}[X]$. Decidir si A es diagonalizable en $\mathbb{R}^{2\times 2}$. ¿Es A única?

Ejercicio 7.

i) Sea $A \in \mathbb{R}^{3\times 3}$ diagonalizable con $tr(A) = -4$. Calcular los autovalores de A, sabiendo que los autovalores de $A^2 + 2.A$ son -1, 3 y 8.

ii) Sea $A \in \mathbb{R}^{4\times 4}$ tal que $\det(A) = 6$; 1 y -2 son autovalores de A y -4 es autovalor de la matriz $A - 3.I_4$. Hallar los restantes autovalores de A.

Ejercicio 8. Sea $A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \in \mathbb{R}^{2\times 2}$.

i) Probar que, para todo $n \in \mathbb{N}$, $A^n.\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} F_n \\ F_{n+1} \end{pmatrix}$ donde F_i es el i-ésimo término de la sucesión de Fibonacci (es decir, $F_0 = 0$, $F_1 = 1$ y $F_{i+1} = F_i + F_{i-1}$).

ii) Encontrar una matriz $P \in GL(2, \mathbb{R})$ tal que $P.A.P^{-1}$ sea diagonal.

iii) Hallar la fórmula general para el término F_n, $\forall\, n \in \mathbb{N}_0$ (comparar con el Ejercicio 50 de la Sección 1.5).

iv) Se define la sucesión $\{a_n\}_{n\in\mathbb{N}_0}$ de la siguiente manera:

$$\begin{cases} a_0 = 0\,,\ a_1 = 1\,,\ a_2 = 1 \\ a_{n+3} = 6.a_{n+2} - 11.a_{n+1} + 6.a_n \quad \forall\, n \in \mathbb{N}_0 \end{cases}$$

Hallar una fórmula general para el término a_n, $\forall\, n \in \mathbb{N}_0$.

Ejercicio 9. Resolver el siguiente sistema de ecuaciones diferenciales

$$\begin{cases} x'(t) = 6x(t) + 2y(t) \\ y'(t) = 2x(t) + 3y(t) \end{cases}$$

con condiciones iniciales $x(0) = 3$, $y(0) = -1$.

Sugerencia: Hallar una matriz $C \in GL(2, \mathbb{R})$ tal que $C^{-1}\begin{pmatrix} 6 & 2 \\ 2 & 3 \end{pmatrix} C$ sea diagonal y hacer el cambio de variables $\begin{pmatrix} u(t) \\ v(t) \end{pmatrix} = C^{-1}.\begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$.

Ejercicio 10. Sea $A \in K^{n\times n}$. Probar que A y A^t tienen los mismos autovalores. Dar un ejemplo en el que los autovectores sean distintos.

Ejercicio 11. Sea $\delta : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$ la transformación lineal derivación. Mostrar que todo número real es un autovalor de δ y exhibir un autovector correspondiente.

Ejercicio 12. Analizar la validez de las siguientes afirmaciones:

i) $A \in \mathbb{R}^{n\times n}$ inversible $\Rightarrow 0$ no es autovalor de A.

ii) $A \in \mathbb{R}^{n\times n}$ inversible, x autovector de $A \Rightarrow x$ autovector de A^{-1}.

iii) $A \in \mathbb{R}^{n\times n}$ con n impar $\Rightarrow A$ admite un autovalor real.

Ejercicio 13.

i) Sea $f : K^n \to K^n$ un proyector con $\dim(\mathrm{Im}(f)) = s$. Calcular $\mathcal{X}_f$. ¿Es f diagonalizable?

ii) Sea K un cuerpo incluido en $\mathbb{C}$ y sea $f : K^n \to K^n$ un morfismo nilpotente no nulo. Calcular $\mathcal{X}_f$. ¿Es f diagonalizable?

Ejercicio 14. Sea $A \in \mathbb{R}^{n\times n}$ que verifica $A^2 + I_n = 0$. Probar que A es inversible, que no tiene autovalores reales y que n debe ser par.

Ejercicio 15. Sea V un K-espacio vectorial de dimensión finita y sea $f : V \to V$ una transformación lineal tal que $\dim(\mathrm{Im}(f)) = 1$. Probar que f es diagonalizable si y sólo si $\mathrm{Nu}(f) \cap \mathrm{Im}(f) = \{0\}$.

Ejercicio 16. Sea $D \in K^{n\times n}$ una matriz inversible y diagonal. Sea $f : K^{n\times n} \to K^{n\times n}$ la transformación lineal definida por $f(A) = D^{-1}.A.D$. Hallar los autovalores y los autovectores de f y probar que es diagonalizable.

Ejercicio 17. Sea $f : \mathbb{C}^n \to \mathbb{C}^n$ una transformación lineal. Probar que existe una base B de $\mathbb{C}^n$ tal que $|f|_B$ es triangular superior.

Ejercicio 18. Sea $A \in \mathbb{C}^{n\times n}$ y sean $\lambda_1, \dots, \lambda_n$ las raíces de $\mathcal{X}_A$ contadas con multiplicidad.

i) Probar que $\det(A) = \prod_{i=1}^{n} \lambda_i$.

ii) Probar que $tr(A) = \sum_{i=1}^{n} \lambda_i$.

Ejercicio 19. Sean $A \in K^{m\times n}$ y $B \in K^{n\times m}$.

i) Probar que las matrices $\begin{pmatrix} A.B & 0 \\ B & 0 \end{pmatrix}$ y $\begin{pmatrix} 0 & 0 \\ B & B.A \end{pmatrix}$ en $K^{(m+n)\times(m+n)}$ son semejantes.

ii) Deducir que, si $n = m$, $\mathcal{X}_{A.B} = \mathcal{X}_{B.A}$.

Ejercicio 20. Dadas las matrices $A \in \mathbb{C}^{2\times 2}$ y los polinomios $P \in \mathbb{C}[X]$, calcular $P(A)$ para:

i) $A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$, a) $P = X - 1$, b) $P = X^2 - 1$, c) $P = (X-1)^2$

ii) $A = \begin{pmatrix} i & 0 \\ 1 & -i \end{pmatrix}$, $P = X^3 - i.X^2 + 1 + i$

Ejercicio 21. Sea $A \in \mathbb{R}^{n\times n}$. Probar que el minimal de A como matriz real y el minimal de A como matriz compleja coinciden.

Ejercicio 22. Hallar el polinomio minimal de las siguientes matrices (comparar con el polinomio característico):

i) $\begin{pmatrix} 4 & 1 \\ -2 & 1 \end{pmatrix}$ ii) $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ iii) $\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$ iv) $\begin{pmatrix} i & 0 \\ 1 & i \end{pmatrix}$

v) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ vi) $\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$ vii) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ viii) $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$

ix) $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$ x) $\begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 2 & -1 & 0 \\ 3 & 4 & 0 & -1 \end{pmatrix}$ xi) $\begin{pmatrix} 1 & i & 1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$

xii) $\begin{pmatrix} a & 0 & 0 & 0 \\ 0 & a & 0 & 0 \\ 0 & 0 & a & 0 \\ 0 & 0 & 0 & a \end{pmatrix}$ xiii) $\begin{pmatrix} a & 0 & 0 & 0 \\ 1 & a & 0 & 0 \\ 0 & 0 & a & 0 \\ 0 & 0 & 1 & a \end{pmatrix}$ xiv) $\begin{pmatrix} a & 0 & 0 & 0 \\ 1 & a & 0 & 0 \\ 0 & 1 & a & 0 \\ 0 & 0 & 0 & a \end{pmatrix}$ xv) $\begin{pmatrix} a & 0 & 0 & 0 \\ 1 & a & 0 & 0 \\ 0 & 1 & a & 0 \\ 0 & 0 & 1 & a \end{pmatrix}$

Ejercicio 23. Calcular el polinomio minimal para cada una de las siguientes transformaciones lineales:

i) $f : \mathbb{R}_2[X] \to \mathbb{R}_2[X]$, $f(P) = P' + 2.P$

ii) $f : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$, $f(A) = A^t$

Ejercicio 24. Sea $A \in K^{n\times n}$ la matriz

$$A = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & -a_0 \\ 1 & 0 & 0 & \dots & 0 & -a_1 \\ 0 & 1 & 0 & \dots & 0 & -a_2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & -a_{n-2} \\ 0 & 0 & 0 & \dots & 1 & -a_{n-1} \end{pmatrix}$$

Calcular su polinomio minimal y su polinomio característico. Comparar con lo calculado en la Sección 5.4.

Ejercicio 25. Sea $\delta : \mathbb{R}[X] \to \mathbb{R}[X]$ la transformación lineal derivada. Probar que δ no admite ningún polinomio minimal.

Ejercicio 26. Utilizando el Teorema de Hamilton-Cayley:

i) Calcular $A^4 - 4.A^3 - A^2 + 2.A - 5.I_2$ para $A = \begin{pmatrix} 2 & -1 \\ 1 & 3 \end{pmatrix}$.

ii) Calcular A^{1000} para $A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$.

iii) Dada $A = \begin{pmatrix} 1 & 3 \\ -1 & 4 \end{pmatrix}$, expresar a A^{-1} como combinación lineal de A y de I_2.

iv) Dada $A = \begin{pmatrix} 1 & -1 \\ 2 & 5 \end{pmatrix}$, expresar a $(2.A^4 - 12.A^3 + 19.A^2 - 29.A - 37.I_2)^{-1}$ como combinación lineal de A y de I_2.

v) Dada $A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$, calcular A^{-1}, A^3 y A^{-3}.

vi) Calcular $\begin{pmatrix} 2 & 2 & 0 \\ 1 & 2 & 1 \\ 1 & 2 & 1 \end{pmatrix}^n$ $\forall n \in \mathbb{N}$.

Ejercicio 27. Sea V un K-espacio vectorial de dimensión finita y sea $f : V \to V$ una transformación lineal. Probar que f es un isomorfismo si y sólo si el término constante de $\mathcal{X}_f$ es no nulo. En dicho caso, hallar la expresión general de f^{-1} como polinomio en f.

Ejercicio 28. Exhibir una matriz $A \in \mathbb{C}^{n\times n}$ tal que $A^2 + I_n = 0$. Comparar con el Ejercicio 14.

Ejercicio 29.

i) Sea $f : \mathbb{R}^2 \to \mathbb{R}^2$ la transformación lineal definida por $f(x, y) = (x + 3.y, 3.x - 2.y)$. Hallar todos los subespacios de $\mathbb{R}^2$ que sean f-invariantes.

ii) Sea $f_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ la rotación de ángulo θ. Probar que, para todo $\theta \neq k.\pi$ ($k \in \mathbb{Z}$), f_θ no es diagonalizable. Hallar todos los subespacios de $\mathbb{R}^2$ que sean f_θ-invariantes.

iii) Sea $\theta \in \mathbb{R}$ y $g_\theta : \mathbb{C}^2 \to \mathbb{C}^2$ la transformación $\mathbb{C}$-lineal cuya matriz en la base canónica es

$$|g_\theta|_E = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}$$

¿Es g_θ diagonalizable? Hallar todos los subespacios de $\mathbb{C}^2$ que sean g_θ-invariantes.

Ejercicio 30. Sea $f : \mathbb{R}^n \to \mathbb{R}^n$ una tranformación lineal nilpotente tal que $f^n = 0$ y $f^{n-1} \neq 0$. Probar que existe un hiperplano de $\mathbb{R}^n$ que es f-invariante pero que no admite un complemento f-invariante (comparar con el Ejercicio 23. ii) de la Sección 3.8).

Ejercicio 31.

i) Sea V un K-espacio vectorial de dimensión finita y sea $f : V \to V$ una transformación lineal diagonalizable. Si S es un subespacio de V f-invariante, probar que $f : S \to S$ es diagonalizable.

ii) Sean $A, B \in K^{n\times n}$ tales que $A.B = B.A$ y sea $E_\lambda = \{x \in K^n \,/\, A.x = \lambda.x\}$. Probar que E_λ es B-invariante.

iii) Sean $A, B \in K^{n\times n}$ dos matrices diagonalizables tales que $A.B = B.A$. Probar que existe $C \in GL(n, K)$ tal que $C.A.C^{-1}$ y $C.B.C^{-1}$ son diagonales. (Es decir, A y B se pueden diagonalizar simultáneamente.)

Ejercicio 32.

i) Hallar una matriz $A \in \mathbb{C}^{3\times 3}$ tal que $m_A(X) = X^3 - 5X^2 + 6X + 8$. Decidir si A es diagonalizable.

ii) Hallar una matriz $A \in \mathbb{C}^{4\times 4}$ tal que $m_A(X) = X^4 + 4X^3 + 8X^2 + 8X + 4$. Decidir si A es diagonalizable.

Ejercicio 33. Sea $A \in K^{n\times n}$.

i) Probar que si A es nilpotente, entonces existe $k \in \mathbb{N}$ tal que $m_A(X) = X^k$. Calcular todos los autovalores de A.

ii) Si $K = \mathbb{C}$ y el único autovalor de A es el 0, probar que A es nilpotente. ¿Qué pasa si $K = \mathbb{R}$?

Ejercicio 34. Sea $A \in \mathbb{C}^{n\times n}$ una matriz de traza nula. Probar que A es semejante a una matriz que tiene toda la diagonal nula.