5.5. EL POLINOMIO CARACTERÍSTICO Y EL POLINOMIO MÍNIMO.

Sea $p \in P_{\mathbb{K}}[x]$ un polinomio no nulo,

$$p(x) = a_0 + a_1 x + \dots + a_d x^d = \sum_{j=0}^{d} a_j x^j .$$

Para $A \in M_{n\times n}(\mathbb{K})$ se define

$$p(A) = a_0 I_n + a_1 A + \dots + a_d A^d = \sum_{j=0}^{d} a_j A^j$$

En 1858, Arthur Cayley afirmó que si $A \in M_{n\times n}(\mathbb{K})$ y $p_A(x) = |A - xI|$ es su polinomio característico, $p_A(A) = 0$. A. Cayley demostró este resultado para $A \in M_{2\times 2}(\mathbb{K})$; para $n > 2$ el resultado lo probó William Rowan Hamilton.

Lema 5.5.1

Sea $A \in \mathcal{M}_{n\times n}(\mathbb{K})$ una matriz de orden n con elementos en $\mathbb{K}$ y $p(x) \in P_{\mathbb{K}}[x]$ un polinomio de grado n con coeficientes en $\mathbb{K}$. Si $p(x)I_n = (A - xI)Q(x)$ con

$$Q(x) = Q_0 + Q_1 x + \cdots + Q_{n-1}x^{n-1},$$

donde $Q_j \in \mathcal{M}_{n\times n}(\mathbb{K})$ para $j = 0, 1, ..., n-1$, entonces

$$p(A) = 0.$$

Demostración. Sea $p(x) = a_0 + a_1 x + \cdots + a_n x^n$. La igualdad $p(x)I_n = (A - xI_n)Q(x)$ se transforma en

$$\begin{aligned} p(x)I_n &= a_0 I_n + a_1 I_n x + \cdots + a_n I_n x^n = (A - xI_n)Q(x) \\ &= AQ_0 + (AQ_1 - Q_0)x + \cdots + (AQ_{n-1} - Q_{n-2})x^{n-1} - Q_{n-1}x^n. \end{aligned}$$

Igualando los términos de igual grado obtenemos

$$\begin{aligned} a_0 I_n &= AQ_0 \\ a_1 I_n &= AQ_1 - Q_0 \\ &\vdots \\ a_{n-1} I_n &= AQ_{n-1} - Q_{n-2} \\ a_n I_n &= -Q_{n-1}. \end{aligned}$$

Multiplicar la segunda igualdad por A, la tercera por A^2, ..., y la última por A^n. Si sumamos la parte izquierda de todas las igualdades se obtiene

$$a_0 I_n + a_1 A + \cdots + a_{n-1}A^{n-1} + a_n A^n = p(A),$$

mientras que al sumar la parte derecha de todas las igualdades se obtiene

$$AQ_0 + A(AQ_1 - Q_0) + \cdots + A^{n-1}(AQ_{n-1} - Q_{n-2}) - A^n Q_{n-1} = 0,$$

de donde se deduce el resultado.

Teorema 5.5.2 (Teorema de Cayley-Hamilton)
Si $A \in \mathcal{M}_{n\times n}(\mathbb{K})$ y $p_A(\lambda) = |A - \lambda I_n|$ es su polinomio característico, se tiene que $p_A(A) = 0$.

D/ La matriz $A - xI_n$ es una matriz $n\times n$ cuyos elementos son polinomios de grado ≤ 1. Sea $\Delta(x)$ la matriz de cofactores de $A - xI_n$. Como la matriz de cofactores se hace con los menores de orden $n-1$, los elementos de $\Delta(x)$ son polinomios de grado $\leq n-1$. Por tanto

$$\Delta(x) = \Delta_0 + \Delta_1 x + \dots + \Delta_{n-1}x^{n-1}$$

con $\Delta_j \in M_{n\times n}(k)$, $j=0,1,\dots,n-1$. Pero sabemos (Teorema 4.1 del Tema 4) que

$$(A - xI_n)[\Delta(x)]^t = |A - xI_n|\, I_n = p_A(x) I_n .$$

Por el lema 5.5.1 con $Q(x) = [\Delta(x)]^t$ se deduce que $p_A(A) = 0$. ■

Ej 5.5.1. Sea $f \in \mathrm{End}(\mathbb{R}^4)$ dada por $f(x_1,x_2,x_3,x_4) = (x_1, x_2, x_3+x_4, -x_3+x_4)$. Sea $m(x) = (1-x)(x^2-2x+2)$. Prueba que $m(f) = 0$. ¿Es $m(x)$ el polinomio característico de f?

S/ $m(x)$ no es el polinomio característico de f, pero $m(x)$ divide a $p_A(x)$. Para probar que $m(f)=0$, basta probar que $m(A) = 0$ siendo A una matriz de f.

Dada $A \in M_{n\times n}(\mathbb{K})$ sea

$$I_A = \{p(x) \in P_{\mathbb{K}}[x] : p \text{ no nulo y } p(A)=0\}$$

Por el Teorema de Cayley-Hamilton, $p_A(x) \in I_A$, por lo que $I_A \neq \emptyset$. Llamamos POLINOMIO MÍNIMO de A, y escribimos $m_A(x)$, al polinomio <u>mónico</u> de menor grado entre los elementos de I_A.

Lema 5.5.3

Sea $A \in \mathcal{M}_{n\times n}(\mathbb{K})$ y $m_A(x)$ su polinomio mínimo. Las siguientes condiciones son equivalentes:

1) $p(x) \in I_A$ 2) $m_A(x)$ divide a $p(x)$

D/ 1) $\Rightarrow$ 2) Al dividir $p(x)$ entre $m_A(x)$ se tendría $p(x) = c(x)\, m_A(x) + r(x)$ con $r(x)=0$ ó grado$(r(x)) <$ grado$(m_A(x))$. Como $p(x) \in I_A$, tenemos $p(A)=0$; como $m_A(A)=0$ se tiene $0 = p(A) = c(A)\, m_A(A) + r(A) = r(A)$. Por tanto $r(A) \in I_A$. Como m_A tiene grado mínimo en I_A deducimos $r(x)=0$ y por tanto $p(x) = c(x)\, m_A(x)$.

2) $\Rightarrow$ 1) Como $m_A(x)$ divide a $p(x)$ se tiene $p(x) = c(x)\, m_A(x)$. Como $m_A(A)=0$, deducimos $p(A) = c(A) \cdot m_A(A) = 0$.

NOTA: Del lema 5.5.3 se deduce que $m_A(x)$ es único.

Ej 5.5.2. Halla el polinomio mínimo de la matriz

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

S/ $m_A(x) = (x-1)(x+1)$.

Ya sabemos que un endomorfismo es diagonalizable si y solo si se puede encontrar una base de autovectores (Proposición 5.2.4. Daremos ahora otro criterio basado en su polinomio característico y en la dimensión de sus autoespacios.

Lema 5.5.4

Sea $A: V \to V$ una aplicación lineal en un espacio vectorial V de dimensión n. Sea λ un valor propio de A. Si s es la multiplicidad de λ como raíz del polinomio $p_A(x) = |A - xI_n|$, entonces,

$$\dim(\operatorname{Ker}(A - \lambda I)) \leqslant s.$$

Demostración. Sea $r = \dim(\operatorname{Ker}(A - \lambda I))$ y $\{\vec{v}_1, ..., \vec{v}_r\}$ una base de $\operatorname{Ker}(A - \lambda I)$. Tomemos $v_{r+1}, ..., v_n$ en V tales que $B = \{v_1, ..., v_r, v_{r+1}, ..., v_n\}$ sea una base de V (~~un algoritmo para obtener esta base lo proporciona el Teorema de Steiniz, Teorema 4.5.1~~). En la base B la aplicación lineal A tiene como matriz

$$M = \left(\begin{array}{c|c} \lambda I_r & T \\ \hline 0 & N \end{array}\right)$$

donde N es una matriz de tamaño $(n - r) \times (n - r)$. Como el polinomio característico no depende de la base elegida en V,

$$p_A(x) = p_M(x) = (-1)^r (x - \lambda)^r p_N(x),$$

y como $p_N(x)$ puede tener $(x - \lambda)$ como factor, la multiplicidad de λ como raíz del polinomio $p_A(x)$ es al menos r.

Teorema 5.5.5 (Teorema de diagonalización)

Sea $A: V \to V$ una aplicación lineal en un espacio vectorial V de dimensión n sobre un cuerpo $\mathbb{K}$. Son equivalentes:

a) A es diagonalizable.

b) El polinomio característico de A se descompone en $P_{\mathbb{K}}[x]$ en factores lineales, $p_A(x) = c\prod_{i=1}^{r}(x - \lambda_i)^{n_i}$, $\lambda_i \neq \lambda_j$, y además $\dim(\operatorname{Ker}(A - \lambda_i I)) = n_i$, $i = 1, 2, ..., r$.

Demostración. a) $\Rightarrow$ b). Supongamos que A es diagonalizable. Por la Proposición 5.2.4 existe una base B formada por vectores propios de A con respecto a la cual la matriz de A es diagonal. Sea

$$D = \begin{pmatrix} \lambda_1 I_{n_1} & & 0 \\ & \ddots & \\ 0 & & \lambda_r I_{n_r} \end{pmatrix}, \qquad \lambda_i \neq \lambda_j,\ \lambda_i \in \mathbb{K},$$

esta matriz diagonal, en donde hemos reordenado la base B para poder escribir D de esta manera. Observar que $n = n_1 + \cdots + n_r$. Escribamos

$$B = B_1 \cup B_2 \cup \cdots \cup B_r$$

con $B_i = \{\vec{u}_{i1}, \vec{u}_{i2}, ..., \vec{u}_{in_i}\}$, $i = 1, 2, ..., r$, donde $A\vec{u}_{ij} = \lambda_i \vec{u}_{ij}$, $j = 1, ..., n_i$. Por tanto, $B_i \subset \operatorname{Ker}(A - \lambda_i I)$. Como B_i es parte de una base de V, B_i es un conjunto de vectores linealmente independientes. Luego

$$n_i \leqslant \dim(\operatorname{Ker}(A - \lambda_i I)), \qquad i = 1, 2, ..., r. \tag{5.5}$$

Como el polinomio característico no depende de la base elegida en V,

$$p_A(x) = p_D(x) = \prod_{j=1}^{r} (\lambda_i - x)^{n_i}, \qquad \lambda_i \neq \lambda_j,\ \lambda_i \in \mathbb{K},$$

por lo que $p_A(x)$ se descompone en $P_{\mathbb{K}}[x]$ en factores lineales. Además, la multiplicidad de λ_i es n_i. Por el Lema 5.5.4

$$\dim(\operatorname{Ker}(A - \lambda_i I)) \leqslant n_i, \qquad i = 1, 2, ..., r. \tag{5.6}$$

De (5.5) y (5.6) se deduce $\dim(\operatorname{Ker}(A - \lambda_i I)) = n_i$, $i = 1, 2, ..., r$.

b) $\Rightarrow$ a). Como $\dim(\operatorname{Ker}(A - \lambda_i I)) = n_i$, elijamos

$$B_i = \{\vec{u}_{i1}, ..., \vec{u}_{in_i}\}, \qquad i = 1, 2, ..., r,$$

una base de $\operatorname{Ker}(A - \lambda_i I)$. Sea

$$B = B_1 \cup B_2 \cup \cdots \cup B_r.$$

Como $n_1 + n_2 + \cdots + n_r = n$ porque $p_A(x) = c \prod_{i=1}^{r} (x - \lambda_i)^{n_i}$ y $p_A(x)$ es un polinomio de grado n, basta probar que B es un conjunto de vectores linealmente independiente, ya que entonces B es base de V formada por vectores propios. Por la Proposición 5.2.4 A será diagonalizable.

Probaremos que B es linealmente independiente por inducciónh en r. Si $r = 1$, el resultado es cierto porque B_1 es base de $\operatorname{Ker}(A - \lambda_1 I)$.

Supongamos que $B_1 \cup B_2 \cup \cdots \cup B_{t-1}$, $2 < t \leqslant r$ es linealmente independiente. Sea

$$\sum_{i=1}^{t} \sum_{j=1}^{n_i} \alpha_{ij}\vec{u}_{ij} = \vec{0} \tag{5.7}$$

una combinación lineal de los elementos de $B_1 \cup \cdots \cup B_t$. Aplicando A a (5.7) se deduce

$$\sum_{i=1}^{t} \sum_{j=1}^{n_i} \alpha_{ij}\lambda_i\vec{u}_{ij} = \vec{0} \tag{5.8}$$

y multiplicando (5.7) por λ_t obtenemos

$$\sum_{i=1}^{t} \sum_{j=1}^{n_i} \alpha_{ij}\lambda_t\vec{u}_{ij} = \vec{0}. \tag{5.9}$$

Restando (5.8) y (5.9) podemos escribir

$$\sum_{i=1}^{t-1} \sum_{j=1}^{n_i} \alpha_{ij}(\lambda_i - \lambda_t)\vec{u}_{ij} = \vec{0}.$$

Como $B_1 \cup \cdots \cup B_{t-1}$ es linealmente independiente por la hipótesis de inducción, de esta igualdad se deduce $\alpha_{ij}(\lambda_i - \lambda_t) = 0$. Puesto que $\lambda_i \neq \lambda_t$, $1 \leqslant i < t$, obtenemos $\alpha_{ij} = 0$, $i = 1, ..., t - 1, j = 1, ..., n_i$. Sustituyendo en (5.7) obtenemos

$$\sum_{j=1}^{n_t} \alpha_{tj}\vec{u}_{tj} = \vec{0},$$

de donde deducimos que $\alpha_{tj} = 0, j = 1, ..., n_t$, puesto que B_t es base de $\operatorname{Ker}(A - \lambda_t I)$. ▩

Ej 5.5.3. Sea $f: \mathbb{C}^4 \to \mathbb{C}^4$ el endomorfismo dado por

$$f(z_1, z_2, z_3, z_4) = (z_1, z_2, z_3 + z_4, -z_3 + z_4)$$

(ver Ej 5.5.1). Usa el Teorema 5.5.5 para probar que f es diagonalizable sobre $\mathbb{C}$, pero no como endomorfismo de $\mathbb{R}^4$.

S/ $M_f = A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -1 & 1 \end{pmatrix}$ y $P_A(x) = (1-x)^2 [(1-x)^2 + 1]$

$$= (1-x)^2 [x^2 - 2x + 2] = 0$$

$x^2 - 2x + 2 = 0 \Rightarrow x = \dfrac{2 \pm \sqrt{4-8}}{2} = 1 \pm i$

Como $P_A(x)$ no se descompone en factores simples en $\mathbb{R}$, por el Teor. 5.5.5 no es diagonalizable sobre $\mathbb{R}$.

Sobre $\mathbb{C}$

$$P_A(x) = (1-x)^2 (x - (1+i)) (x - (1-i))$$

sí se descompone en factores simples. Hay que comprobar que se cumple la condición sobre la dimensión de los autoespacios.

$\boxed{\lambda_1 = 1}$

$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}$ tiene rango 2 $\Rightarrow$ $\dim E_1(1) = 4 - 2 = 2 =$
= multiplicidad de $\lambda_1 = 1$

$\boxed{\lambda_2 = 1 + i, \lambda_3 = 1 - i}$ Siempre $\dim(E_1(\lambda_i)) \geq 1$, $i = 1, 2$

p.q. $E_1(\lambda_i) \neq \{\vec{0}\}$. Por otro lado, por el lema 5.5.4, $\dim(E_1(\lambda_i)) \geq 1$ (1 es su multiplicidad). Por tanto

$$\dim(E_1(\lambda_i)) = 1 = \text{multiplicidad de } \lambda_i \,, \; i = 1, 2.$$

Por el Teorema 5.5.5, f es diagonalizable sobre $\mathbb{C}$.