2.6 Polinomios anuladores y los Teoremas A y B

En esta sección vamos a suponer que todos los espacios vectoriales son sobre $\mathbb{K} = \mathbb{C}$. Esta sección está basada en la Sección 14.6 de Volumen 2 del Libro de *Álgebra Lineal* de Fernando, Gamboa, Ruiz.

Polinomios anuladores de un endomorfismo

Sea $f : V \to V$ un endomorfismo de un $\mathbb{C}$-espacio vectorial. Denotamos como es habitual

$$f^0 = \mathrm{I}, \quad \text{y} \quad f^k = f \circ \overset{k)}{\cdots} \circ f \quad \text{para } k \geq 1.$$

Dado un polinomio

$$P(T) = \sum_{k=0}^{p} a_k T^k \in \mathbb{C}[T]$$

queremos darle sentido a sustituir $T = f$ en $P(T)$. Simplemente interpretamos T^k como la composición f^k, de forma que queda

$$P(f) = \sum_{k=0}^{p} a_k f^k$$

el cual también es un endomorfismo de V bien definido.

Observación 2.2 1) Si f respecto de una base B tiene matriz A, entonces la matriz del endomorfismo $P(f)$ respecto de B es evidentemente

$$P(A) = \sum_{k=0}^{p} a_k A^k.$$

2) La sustitución anterior respeta el producto de polinomio, es decir, dados dos polinomios $P(T), Q(T) \in \mathbb{C}[T]$, se cumple

$$(PQ)(f) = P(f) \circ Q(f).$$

Por ejemplo, si $P(T) = T^2 - 5$ y $Q(T) = T^3 - T^2$ se tiene que

$$\begin{aligned} P(T)Q(T) =& (T^2 - 5)(T^3 - T^2) = T^2(T^3 - T^2) - 5(T^3 - T^2) = \\ =& T^{2+3} - T^{2+2} - 5T^3 + 5T^2 = T^5 - T^4 - 5T^3 + 5T^2 \\ \Rightarrow (PQ)(f) =& f^5 - f^4 - 5f^3 + 5f^2. \end{aligned}$$

Por otro lado $P(f) = f^2 - 5I$ y $Q(f) = f^3 - f^2$ y se tiene que

$$\begin{aligned} P(f) \circ Q(f) =& (f^2 - 5I) \circ (f^3 - f^2) = f^2 \circ (f^3 - f^2) - 5I \circ (f^3 - f^2) = \\ =& f^{2+3} - f^{2+2} - 5f^3 + 5f^2 = f^5 - f^4 - 5f^3 + 5f^2. \end{aligned}$$

Vemos que la demostración en general se reduce a la igualdad evidente $f^k \circ f^\ell = f^{k+\ell}$, después de hacer cuidadosamente todas las operaciones. En particular, como el producto entre polinomios es conmutativo, deducimos que

$$P(f) \circ Q(f) = (PQ)(f) = (QP)(f) = Q(f) \circ P(f).$$

Obsérvese que esto nos dice que si $P(T)$ es un producto de polinomios $Q_1, \dots, Q_\ell$ y alguno de ellos es anulador de f, entonces P también es anulador de f. Claro, siempre podemos reordenarlos y podemos suponer que Q_ℓ es el anulador de f, de forma que $P(f) = Q_1(f) \circ \cdots Q_{\ell-1}(f) \circ Q_\ell(f) = Q_1(f) \circ \cdots Q_{\ell-1} \circ 0 = 0$.

¿Y al revés? ¿Si P es un polinomio anulador de f y $P = Q_1, \dots, Q_\ell$, entonces algún Q_i es un anulador de f? No, eso no es cierto. Por ejemplo, tomemos $Q_1(T) = T$ y $Q_2(T) = T$, y consideremos el producto $P(T) = Q_1(T)Q_2(T) = T^2$. Se tiene entonces que el endomorfismo $f : \mathbb{R}^2 \to \mathbb{R}^2 : (x, y) \mapsto (x - y, x - y)$ se anula en $P(T)$ pero no se anula en Q_1 ni en Q_2.

3) Supongamos que tenemos polinomios $P_1(T), \dots, P_\ell(T)$ y que sabemos que para cierto vector $v \in V$ se tiene que $P_i(f)(v) = 0$. Entonces para el polinomio $P(T) = P_1(T) \cdots P_\ell(T)$ también se tiene que $P(f)(v) = 0$. En efecto,

$$\begin{aligned}
P(f)(v) &= (P_1 \cdots P_\ell)(f)(v) = \\
&= (P_1 \cdots P_{i-1} P_{i+1} P_\ell P_i)(f)(v) = \\
&= (P_1 \cdots P_{i-1} P_{i+1} P_\ell) \circ P_i(f)(v) = \\
&= (P_1 \cdots P_{i-1} P_{i+1} P_\ell)(0) = 0.
\end{aligned}$$

4) Si $\lambda \in \mathbb{C}$ es un autovalor de f, y u un autovector asociado, es decir, $f(u) = \lambda u$, entonces se tiene que

$$P(f)(u) = P(\lambda)u.$$

En efecto, observamos primero que

$$f^2(u) = f(f(u)) = f(\lambda u) = \lambda f(u) = \lambda(\lambda u) = \lambda^2 u.$$

y repitiendo el cálculo se ve que $f^k(u) = \lambda^k u$ para cualquier k. En particular,

$$P(f)(u) = \sum_{k=0}^{p} a_k f^k(u) = \sum_{k=0}^{p} a_k \lambda^k u = (\sum_{k=0}^{p} a_k \lambda^k)u = P(\lambda)u.$$

■ **Definición** Dado un endomorfismo $f : V \to V$, decimos que un polinomio $P(T) \in \mathbb{C}[T]$ es un *polinomio anulador* de f si $P(f) = 0$.

Observación 2.3 (1) Siempre hay un polinomio anulador. En efecto, supongamos que hemos fijado una base B de V y que la matriz de f es A respecto de esta base. Ya sabemos que el espacio de matrices $\mathcal{M}_n(\mathbb{C})$ tiene dimensión n^2, así que entre las siguientes $n^2 + 1$ matrices,

$$\mathrm{I}, A, A^2, \dots, A^{n^2},$$

debe haber una dependencia lineal, es decir, existen $a_0, \dots, a_{n^2}$ tales que

$$a_0 I + a_1 A + \dots + a_{n^2} A^{n^2} = 0.$$

En particular estamos diciendo que $P(T) = \sum_{k=0}^{n^2} a_k T^k$ es una polinomio anulador de f.

(2) Los autovalores de f son raíces de cualquier polinomio anulador $P(T)$. En efecto, si $f(u) = \lambda u$ con $u \neq 0$, entonces por (2)

$$0 = P(f)(u) = P(\lambda)u,$$

y como $u \neq 0$, es $P(\lambda) = 0$.

(3) Además, siempre podemos encontrar un polinomio anulador que no tenga otras raíces. En efecto, sea $P(T)$ un polinomio anulador y μ una raíz suya que no es autovalor de f. Entonces $P(T) = (T-\mu)Q(T)$, y para $u \in V$ se tiene:

$$0 = P(f)(u) = (f - \mu \mathrm{I}) \circ Q(f)(u) = f\big(Q(f)(u)\big) - \mu Q(f)(u),$$

luego $f(v) = \mu v$ con $v = Q(f)(u)$. Como μ no es autovalor, el vector $v = Q(f)(u)$ debe ser nulo. Por tanto, $Q(T)$ es también un polinomio anulador. De esta manera, podemos eliminar todas las raíces que no son autovalores.

■ **Proposición 2.4** *Sea $f : V \to V$ un endomorfismo, y sean $\lambda_1, \dots, \lambda_r \in \mathbb{C}$ todos sus autovalores. Supongamos que el polinomio $P(T) = (T-\lambda_1)^{m_1} \cdots (T-\lambda_r)^{m_r} = \prod_i (T-\lambda_i)^{m_i}$ es un polinomio anulador de f. Entonces*

1) $E_{m_j}(\lambda_j) \cap \sum_{i \neq j} E_{m_i}(\lambda_i) = \{0\}$,

2) $V = E_{m_1}(\lambda_1) + \cdots + E_{m_r}(\lambda_r)$,

3) $M(\lambda_i) = E_{m_i}(\lambda_i)$ *y por tanto* $\mathrm{nil}(\lambda_i) \leq m_i$.

Demostración. 1) Para cada $j = 1, \dots, r$ consideramos los polinomios

$$P_j(T) = (T - \lambda_j)^{m_j} \quad \text{y} \quad Q_j(T) = \prod_{i \neq j} P_i(T) = \prod_{i \neq j} (T - \lambda_i)^{m_i}.$$

Observamos que estos polinomios tiene la siguiente propiedad: dado $u \in E_{m_j}(\lambda_j)$,

$$P_j(f)(u) = (f - \lambda_j I)^{m_j}(u) = 0,$$

y por tanto dado $v \in E_{m_\ell}(\lambda_\ell)$ con $\ell \neq j$, como $P_\ell(f)(v) = 0$, se tiene por la Observación 2.2.(3) que

$$Q_j(f)(v) = 0.$$

Observamos también que $P_j(T)$ y $Q_j(T)$ no tiene raíces comunes y por tanto podemos encontrar ciertos polinomios $A(T)$ y $B(T)$ de forma que

$$1 = A(T)P_j(T) + B(T)Q_j(T)$$

y sustituyendo $T = f$,

$$\mathrm{I} = A(f) \circ P_j(f) + B(f) \circ Q_j(f).$$

Por tanto, para cada $u \in V$ resulta

$$u = A(f)(P_j(f)(u)) + B(f)(Q_j(f)(u)) \tag{$\star$}$$

Finalmente, tomemos

$$u \in E_{m_j}(\lambda_j) \cap \sum_{i \neq j} E_{m_i}(\lambda_i)$$

y veamos que $u = 0$. Por una parte, como $u \in E_{m_j}(\lambda_j)$ ya sabemos que $P_j(f)(u) = 0$, y por otra parte, como $u = \sum_{i \neq j} u_i$ para ciertos $u_i \in E_{m_i}(\lambda_i)$, se tiene que

$$Q_j(f)(u) = Q_j(f)\big(\sum_{i \neq j} u_i\big) = \sum_{i \neq j} Q_j(f)(u_i) = 0.$$

En consecuencia, aplicando esto a la ecuación ($\star$),

$$u = A(f)(P_j(f)(u)) + B(f)(Q_j(f)(u)) = A(f)(0) + B(f)(0) = 0.$$

es decir, $u = 0$ que es lo que queríamos probar.

2) Como los polinomios $Q_1(T), \ldots, Q_r(T)$ no tienen ninguna raíz en común, existen polinomios $A_1(T), \ldots, A_r(T)$ tales que

$$1 = A_1(T)Q_1(T) + \cdots + A_r(T)Q_r(T),$$

y substituyendo $T = f$,

$$\mathrm{I} = A_1(f) \circ Q_1(f) + \cdots + A_r(f) \circ Q_r(f).$$

Si aplicamos los endomorfismos de la izquierda y de la derecha a un vector $u \in V$ arbitrario deducimos que

$$u = A_1(f) \circ Q_1(f)(u) + \cdots + A_r(f) \circ Q_r(f)(u).$$

Denotemos $u_i = A_i(f) \circ Q_i(f)(u)$ para cada $i = 1, \ldots, r$, es decir,

$$u = \underbrace{A_1(f) \circ Q_1(f)(u)}_{u_1} + \cdots + \underbrace{A_r(f) \circ Q_r(f)(u)}_{u_r}.$$

Evidentemente, si probamos que

$$u_i \in E_{m_i}(\lambda_i) = \ker(f - \lambda_i \mathrm{I})^{m_i} = \ker(P_i(f)),$$

habremos terminado. Pero lo anterior es inmediato ya que

$$\begin{aligned} P_i(f)(u_i) &= P_i(f) \circ A_i(f) \circ Q_i(f)(u) = \\ &= A_i(f) \circ (Q_i P_i)(f)(u) = A_i(f) \circ P(f)(u) = A_i(f)(0) = 0. \end{aligned}$$

3) Los dos puntos anteriores nos dicen que tenemos la descomposición en suma directa

$$(*) \qquad V = E_{m_1}(\lambda_1) \oplus \cdots \oplus E_{m_r}(\lambda_r).$$

Evidentemente el polinomio

$$Q(T) = P(T)(T - \lambda_1) \cdots (T - \lambda_r) = (T - \lambda_1)^{m_1+1} \cdots (T - \lambda_r)^{m_r+1}$$

también es un polinomio anulador de f. Luego también

$$(**) \qquad V = E_{m_1+1}(\lambda_1) \oplus \cdots \oplus E_{m_r+1}(\lambda_r).$$

Como $(*)$ y $(**)$ son dos sumas directas, y cada sumando de la primera está contenido en el correspondiente de la segunda, tiene que ser

$$E_{m_i}(\lambda_i) = E_{m_i+1}(\lambda_i).$$

En otras palabras la sucesión de espacios invariantes de λ_i se estabiliza al menos a partir de $E_{m_i}(\lambda_i)$, así que éste es el subespacio invariante maximal $M(\lambda_i)$. $\square$

La demostración de los Teoremas A y B es ahora trivial.

Demostración del Teorema B. Por la Observación 2.3 sabemos que f tiene al menos un polinomio anulador cuyas raíces son los autovalores de f. Por la proposición anterior deducimos directamente que $V = M(\lambda_1) \bigoplus \cdots \bigoplus M(\lambda_r)$. $\square$

Demostración del Teorema A. Ya hemos probado que $V = M(\lambda_1) \bigoplus \cdots \bigoplus M(\lambda_r)$. Por el Teorema de Jordan con un único autovalor (véase la Observacíon 2.1) tenemos que cada

$$f|_{M(\lambda_i)} : M(\lambda_i) \to M(\lambda_i)$$

tiene una base respecto de la cual su matriz es de Jordan J_i y cuyo orden es evidentemente $\dim(M(\lambda_i))$. Denotemos $d_i = \dim(M(\lambda_i))$. Si juntamos todas las bases obtendremos una base de V respecto de la cual f tiene matriz

$$\left(\begin{array}{c|c|c} J_1 & & \\ \hline & \ddots & \\ \hline & & J_r \end{array}\right)$$

Así pues, el polinomio característico de f es

$$\begin{aligned} P(T) &= \det(J - T\,I_n) = \det(J_1 - T\,I_{d_1}) \cdots \det(J_r - T\,I_{d_r}) \\ &= (\lambda_1 - T)^{d_1} \cdots (\lambda_r - T)^{d_r}, \end{aligned}$$

En particular, vemos que d_i debe ser exactamente la multiplicidad algebraica de λ_i, es decir, se tiene que $\mathrm{mult}_a(\lambda_i) = \dim(M(\lambda_i))$. □

■ **Definición** Sean $\lambda_1, \ldots, \lambda_r \in \mathbb{C}$ todos los autovalores de $f : V \to V$. Entonces llamamos al polinomio

$$P_{\min}(T) = \prod_i (T - \lambda_i)^{\mathrm{nil}(\lambda_i)}$$

el *polinomio mínimo* de f.

■ **Proposición** *El polinomio mínimo es un polinomio anulador de f.*

Demostración. Sea $v \in M(\lambda_i) = E_{\mathrm{nil}(\lambda_i)}(\lambda_i)$. Evidentemente

$$(f - \lambda_i I)^{\mathrm{nil}(\lambda_i)}(v) = 0,$$

y por la Observación 2.2.(3) se tiene entonces que $P_{\min}(f)(v) = 0$. Por tanto dado cualquier vector $v \in V$, como lo podemos escribir

$$v = v_1 + \cdots + v_r$$

con $v_i \in M(\lambda_i)$, deducimos que

$$P_{\min}(f)(v) = P_{\min}(f)(v_1) + \cdots + P_{\min}(f)(v_r) = 0$$

y por tanto $P_{\min}(f) = 0$. □

■ **Proposición** *Si $Q(T)$ es un polinomio anulador de f, entonces $P_{min}(T)$ divide a $Q(T)$.*

Demostración. Ya hemos visto que los autovalores deben ser raíces de $Q(T)$, por tanto

$$Q(T) = (T - \lambda_1)^{m_1} \cdots (T - \lambda_r)^{m_r} R(T)$$

para algún polinomio $R(T)$ que no se anula en $T = \lambda_i$ con $i = 1, \ldots, r$. Ya vimos en la Observación 2.3.(3) que si μ es una raíz de $P(T)$ que no es un autovalor (es por tanto una raíz de $R(T)$), entonces si dividimos $P(T)$ entre $T - \mu$ nos sigue quedando un polinomio anulador de f. Así que si hacemos esto repetidamente llegamos a que $(T - \lambda_1)^{m_1} \cdots (T - \lambda_r)^{m_r}$ es un polinomio anulador de f. Y por Proposición 2.4.(3) deducimos que $m_i \geq \mathrm{nil}(\lambda_i)$. En particular,

$$Q(T) = P_{\min}(T)(T - \lambda_1)^{m_1 - \mathrm{nil}(\lambda_1)} \cdots (T - \lambda_r)^{m_r - \mathrm{nil}(\lambda_r)} R(T),$$

es decir, $P_{\min}(T)$ divide a $Q(T)$. □

■ **Corolario** (Teorema de Cayley-Hamilton) *El polinomio característico* $P(T)$ *de* f *es divisible entre* $P_{min}(T)$. *En particular, se tiene que* $P(f) = 0$.

Demostración. Si $\lambda_1, \dots, \lambda_r$ son todos los autovalores de f, lo que queremos probar es que

$$P(T) = (T - \lambda_1)^{\mathrm{mult}_a(\lambda_1)} \cdots (T - \lambda_r)^{\mathrm{mult}_a(\lambda_r)}$$

es divisible entre

$$P_{\min}(T) = (T - \lambda_1)^{\mathrm{nil}(\lambda_1)} \cdots (T - \lambda_r)^{\mathrm{nil}(\lambda_r)},$$

es decir, necesitamos probar que $\mathrm{nil}(\lambda_i) \leq \mathrm{mult}_a(\lambda_i)$ para cualquier $i = 1, \dots, r$. Por definición

$$E_1(\lambda_i) \subsetneq E_2(\lambda_i) \subsetneq \cdots \subsetneq E_{\mathrm{nil}(\lambda_i)-1}(\lambda_i) \subsetneq E_{\mathrm{nil}(\lambda_i)}(\lambda_i) = M(\lambda_i),$$

y como los contenidos son estrictos, cada vez que pasamos de un subespacio propio generalizado al siguiente aumentamos al menos en 1 la dimensión, por tanto

$$\dim(M(\lambda_i)) =$$

$$= \dim(E_1(\lambda_i)) + \big(\dim(E_2(\lambda_i)) - \dim(E_1(\lambda_i))\big) + \cdots + \big(\dim(E_{\mathrm{nil}(\lambda_i)}(\lambda_i)) - \dim(E_{\mathrm{nil}(\lambda_i)-1}(\lambda_i))\big) \geq$$

$$\geq \dim(E_1(\lambda_i)) + \big(\mathrm{nil}(\lambda_i) - 1\big) \geq 1 + \big(\mathrm{nil}(\lambda_i) - 1\big) = \mathrm{nil}(\lambda_i),$$

y puesto que $\dim(M(\lambda_i)) = \mathrm{mult}_a(\lambda_i)$, deducimos que $\mathrm{nil}(\lambda_i) \leq \mathrm{mult}_a(\lambda_i)$. □