Diagonalización y polinomio minimal

Mariano Suárez-Álvarez

21 de julio, 2020

Resumen

Damos una prueba de que un endomorfismo de un espacio vectorial de dimensión finita es diagonalizable exactamente cuando su polinomio minimal se factoriza como producto de factores mónicos lineales distintos dos a dos sin usar el *Teorema de descomposición primaria*, a diferencia de cómo lo hacemos en las notas — reemplazando ese teorema por una induccion.

Vamos a usar en la demostración varias veces el siguiente hecho:

> *Si p y q son dos polinomios de $\Bbbk[X]$ y $f : V \to V$ es un endomorfismo de un espacio vectorial V, entonces los endomorfismos $p(f)$ y $q(f)$ conmutan, de manera que $p(f) \circ q(f) = q(f) \circ p(f)$.*

Teorema. *Un endomorfismo de un espacio vectorial de dimensión finita es diagonalizable si y solamente si su polinomio minimal se factoriza como un producto de polinomios de grado 1 distintos dos a dos.*

Demostración de la suficiencia de la condición. Sea V un espacio vectorial de dimensión finita y sea $f : V \to V$ un endomorfismo de V. Supongamos primero que el polinomio minimal m_f de f tiene una factorización de la forma

$$m_f = (X - \lambda_1) \cdots (X - \lambda_r)$$

con los escalares $\lambda_1, \ldots, \lambda_r$ distintos dos a dos. Queremos probar que f es diagonalizable, y procederemos haciendo inducción con respecto al número r de factores en la factorización[1]. Si $r = 0$, entonces $m_f = 1$ y necesariamente $V = 0$: claramente en este caso no hay nada que hacer. Supongamos entonces que $r \geq 1$.

Sea $p = (X - \lambda_2) \cdots (X - \lambda_r)$. Claramente los polinomios $X - \lambda_1$ y p son coprimos, así que existen dos polinomios r y s en $\Bbbk[X]$ tales que

$$r \cdot (X - \lambda_1) + s \cdot p = 1. \tag{1}$$

[1] El argumento que vimos en clase se basaba en una inducción con respecto a la dimensión de V: lo hacemos ahora con respecto a r para variar un poco.

Consideremos las funciones lineales

$$\pi_1 \coloneqq r(f) \circ (f - \lambda_1 \mathsf{id}_V) : V \to V, \qquad \pi_2 \coloneqq s(f) \circ p(f) : V \to V$$

y mostremos que valen las siguientes igualdades:

$$\pi_1 + \pi_2 = \mathsf{id}_V, \tag{2}$$
$$\pi_1 \circ \pi_2 = 0, \qquad \pi_2 \circ \pi_1 = 0, \tag{3}$$
$$\pi_1 \circ \pi_1 = \pi_1, \qquad \pi_2 \circ \pi_2 = \pi_2, \tag{4}$$
$$\pi_1 \circ f = f \circ \pi_1, \qquad \pi_2 \circ f = f \circ \pi_2, \tag{5}$$
$$p(f) \circ \pi_1 = 0, \qquad (f - \lambda_1 \mathsf{id}_V) \circ \pi_2 = 0. \tag{6}$$

- Evaluando los dos lados de la igualdad (1) en f vemos que

$$\pi_1 + \pi_2 = r(f) \circ (f - \lambda_1 \mathsf{id}_V) + s(f) \circ p(f) = \mathsf{id}_V.$$

- Los endomorfismos $(f - \lambda_1 \mathsf{id}_V)$ y $s(f)$ conmutan, así que

$$\begin{aligned} \pi_1 \circ \pi_2 &= r(f) \circ (f - \lambda_1 \mathsf{id}_V) \circ s(f) \circ p(f) \\ &= r(f) \circ s(f) \circ (f - \lambda_1 \mathsf{id}_V) \circ p(f) \\ &= r(f) \circ s(f) \circ m_f(f) \\ &= 0. \end{aligned}$$

  De manera similar, los endomorfismos $p(f)$ y $r(f)$ conmutan, así que

$$\begin{aligned} \pi_2 \circ \pi_1 &= s(f) \circ p(f) \circ r(f) \circ (f - \lambda_1 \mathsf{id}_V) \\ &= s(f) \circ r(f) \circ p(f) \circ (f - \lambda_1 \mathsf{id}_V) \\ &= s(f) \circ r(f) \circ m_f(f) \\ &= 0. \end{aligned}$$

- Si componemos a cada lado de la igualdad (2) con π_1 por la izquierda y usamos la segunda de las igualdades de (3), vemos que

$$\pi_1 = \mathsf{id}_V \circ \pi_1 = \pi_1 \circ \pi_1 + \pi_2 \circ \pi_1 = \pi_1 \circ \pi_1.$$

  De la misma forma, a partir de (2) y la primera de las igualdades de (3) vemos que

$$\pi_2 = \mathsf{id}_V \circ \pi_2 = \pi_1 \circ \pi_2 + \pi_2 \circ \pi_2 = \pi_2 \circ \pi_2.$$

- El endomorfismo π_1 es la evaluación en f del polinomio $r \cdot (X - \lambda_1)$, así que conmuta con f, que es la evaluación en f del polinomio X. De la misma forma, el endomorfismo π_2 es la evaluación en f del polinomio $s \cdot p$, así que conmuta con f que es, como dijimos, la evaluación del polinomio X en f. Esto prueba las dos igualdades de (5).

- Finalmente, para probar (6) calculamos que

$$\begin{aligned} p(f)\circ\pi_1 &= p(f)\circ r(f)\circ(f-\lambda_1\mathsf{id}_V) = r(f)\circ p(f)\circ(f-\lambda_1\mathsf{id}_V) \\ &= r(f)\circ m_f(f) = 0 \end{aligned}$$

  porque $p(f)$ y $r(f)$ conmutan, y que

$$\begin{aligned} (f-\lambda_1\mathsf{id}_V)\circ\pi_2 &= (f-\lambda_1\mathsf{id}_V)\circ s(f)\circ p(f) = s(f)\circ(f-\lambda_1\mathsf{id}_V)\circ p(f) \\ &= s(f)\circ m_f(f) = 0, \end{aligned}$$

  porque $f-\lambda_1\mathsf{id}_V$ y $s(f)$ conmutan.

Consideremos ahora los siguientes dos subespacios de V

$$V_1 \coloneqq \pi_1(V), \qquad V_2 \coloneqq \pi_2(V)$$

y mostremos que

$$V = V_1 \oplus V_2. \tag{7}$$

Es claro que $V_1 + V_2 \subseteq V$, y si $v \in V$, entonces de acuerdo a la igualdad (2) tenemos que

$$v = \mathsf{id}_V(v) = \pi_1(v) + \pi_2(v) \in \pi_1(V) + \pi_2(V) = V_1 + V_2.$$

Esto nos dice que $V = V_1 + V_2$, y nos falta ver que esta suma es directa. Supongamos para ello que $v \in V_1 \cap V_1$. Como $v \in V_1 = \pi_1(V)$, existe $v_1 \in V$ tal que $v = \pi_1(v_1)$, y como $v \in V_2 = \pi_2(V)$, existe $v_2 \in V$ tal que $v = \pi_2(v_2)$: usando esto y las igualdades de (4), vemos que

$$\pi_1(v) = \pi_1(\pi_1(v_1)) = \pi_1(v_1) = v$$

y

$$\pi_2(v) = \pi_2(\pi_2(v_2)) = \pi_2(v_2) = v,$$

y usando esto y (3) que

$$v = \pi_1(v) = \pi_1(\pi_2(v)) = 0.$$

Con esto queda probada la afirmación (7).

Lo siguiente que queremos ver es que lo subespacios V_1 y V_2 son f-invariantes, esto es, que $f(V_1) \subseteq V_1$ y $f(V_2) \subseteq V_2$.

- Si $v \in V_1 = \pi_1(V)$, entonces existe $w \in V$ tal que $v = \pi_1(w)$ y, por lo tanto,

$$f(v) = f(\pi_1(w)) = \pi_1(f(w)) \in \pi_1(V) = V_1,$$

  por la primera igualdad de (5). De manera similar, si $v \in V_2$, entonces existe $w \in V$ tal que $v = \pi_2(w)$ y

$$f(v) = f(\pi_2(w)) = \pi_2(f(w)) \in \pi_2(V) = V_2.$$

Como consecuencia de esta invariancia, podemos considerar las funciones

$$f_1 : v \in V_1 \mapsto f(v) \in V_1, \qquad f_2 : v \in V_2 \mapsto f(v) \in V_2$$

que se obtienen restringiendo f a V_1 y a V_2.

Sean $n \coloneqq \dim V_1$ y $m \coloneqq \dim V_2$.

- Observemos que si $v \in V_1 = \pi_1(V)$, entonces existe $w \in V$ es tal que $w = \pi_1(v)$ y, gracias a la primera de las igualdades de (6), tenemos que

  $$p(f_1)(v) = p(f)(v) = p(f)(\pi_1(w)) = 0.$$

  Vemos así que $p(f_1) = 0$ y, por lo tanto, que el polinomio minimal m_{f_1} de f_1 divide a p. Como $p = (X - \lambda_2)\cdots(X - \lambda_r)$, esto implica que m_{f_1} es un producto de como mucho $r - 1$ factores mónicos lineales distintos dos a dos[2]. La hipótesis inductiva obvia nos dice entonces que el endomorfismo f_1 de V_1 es diagonalizable: podemos elegir entonces una base ordenada $\mathscr{B}_1 = (x_1, \ldots, x_n)$ de V_1 de autovalores de f_1. Como f_1 es la restricción de f a V_1, los vectores de $\mathscr{B}_1$ son también autovalores de f.

- Por otro lado, si $v \in V_2 = \pi_2(V)$, entonces existe $w \in V$ tal que $w = \pi_2(v)$ y, de acuerdo a la segunda igualdad de (6),

  $$(f_2 - \lambda_1 \mathsf{id}_V)(v) = (f - \lambda_1 \mathsf{id}_V)(v) = (f - \lambda_1 \mathsf{id}_V)(\pi_2(w)) = 0.$$

  Esto nos dice que $f(v) = f_2(v) = \lambda_1 v$ cualquiera sea $v \in V$. Elijamos una base ordenada cualquiera $\mathscr{B}_2 = (y_1, \ldots, y_m)$ de V_2.

Como $\mathscr{B}_1$ y $\mathscr{B}_2$ son bases de V_1 y V_2 y vale (7), sabemos que

$$\mathscr{B} = (x_1, \ldots, x_n, y_1, \ldots, y_m)$$

es una base ordenada de V. Como sus $n + m$ elementos son todos autovectores de f, concluimos con esto que f es diagonalizable. Esto completa la inducción y, por lo tanto, la prueba de la suficiencia de la condición de la proposición. □

Demostración de la necesidad de la condición. Supongamos que $f : V \to V$ es un endomorfismo diagonalizable de un espacio vectorial V de dimensión finita n, de manera que hay una base ordenada $\mathscr{B} = (x_1, \ldots, x_n)$ de V de autovectores de f. Sean $\lambda_1, \ldots, \lambda_n$ los autovalores correspondientes a los vectores de $\mathscr{B}$, en orden.

Para cada $i \in [\![n]\!]$ tenemos que $f(x_i) = \lambda_i x_i$, así que claramente el polinomio minimal de f en x_i es

$$m_{f,x_i} = X - \lambda_i.$$

Como $\mathscr{B}$ genera a V, es evidente que

$$\langle x_1 \rangle_f + \cdots + \langle x_n \rangle_f = V,$$

[2] De hecho, el polinomio minimal de f_1 es *exactamente* igual a p, pero no necesitamos esto para lo que sigue.

y entonces el polinomio minimal de f es

$$m_f = \operatorname{mcm}\{m_{f,x_1}, \dots, m_{f,x_n}\} = \operatorname{mcm}\{X - \lambda_1, \dots, X - \lambda_n\}.$$

Si ahora $\mu_1, \dots, \mu_r$ son los elementos del conjunto $\{\lambda_1, \dots, \lambda_n\}$ listados sin repeticiones, entonces esto nos dice que

$$m_f = (X - \mu_1) \cdots (X - \mu_r)$$

y, por lo tanto, este polinomio minimal m_f se factoriza como un producto de factores mónicos lineales distintos dos a dos. $\square$