Teorema de Cayley–Hamilton

Objetivos. Demostrar el teorema de Cayley–Hamilton. Conocer los conceptos de polinomios con coeficientes matriciales y de matrices con entradas polinomiales.

Requisitos. Polinomio de un operador lineal, polinomio de una matriz, matriz adjunta, definición formal del producto de polinomios.

El enunciado del teorema de Cayley–Hamilton

1. Ejemplo. Consideremos la matriz

$$A = \begin{bmatrix} 3 & -1 \\ 2 & 1 \end{bmatrix}.$$

El polinomio característico de A es $\mathrm{C}_A(\lambda) = \lambda^2 - 4\lambda + 5$. Calculemos $\mathrm{C}_A(A)$.

$$A^2 = \begin{bmatrix} 3 & -1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 3 & -1 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 9-2 & -3-1 \\ 6+2 & -2+1 \end{bmatrix} = \begin{bmatrix} 7 & -4 \\ 8 & -1 \end{bmatrix},$$

$$\mathrm{C}_A(A) = A^2 - 4A + 5I = \begin{bmatrix} 7 & -4 \\ 8 & -1 \end{bmatrix} + \begin{bmatrix} -12 & 4 \\ -8 & -4 \end{bmatrix} + \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

El teorema de Cayley–Hamilton dice que para toda matriz $A \in \mathcal{M}_n(\mathbb{F})$,

$$\mathrm{C}_A(A) = \mathbf{0}_{n,n}.$$

La demostración del teorema no es trivial. Necesitamos unas herramientas avanzadas.

Polinomios con coeficientes matriciales

2. Polinomios con coeficientes matriciales. Sean $P_0, \ldots, P_m \in \mathcal{M}_n(\mathbb{F})$. Entonces la expresión

$$P(\lambda) \coloneqq P_0 + P_1\lambda + P_2\lambda^2 + \ldots + P_m\lambda^m$$

se llama *polinomio con coeficientes matriciales.*

3. Evaluación de un polinomio con coeficientes matriciales en una matriz. Sea P un polinomio con coeficientes matriciales en $\mathcal{M}_n(\mathbb{F})$:

$$P(\lambda) = P_0 + P_1\lambda + P_2\lambda^2 + \ldots + P_m\lambda^m,$$

y sea $A \in \mathcal{M}_n(\mathbb{F})$ una matriz. Entonces se definen el *valor derecho* y el *valor izquierdo* del polinomio P en la matriz A:

$$P^{der}(A) = P_0 + P_1A + P_2A^2 + \ldots + P_mA^m,$$
$$P^{izq}(A) = P_0 + AP_1 + A^2P_2 + \ldots + A^mP_m.$$

En vez de $P^{der}(A)$ escribimos simplemente $P(A)$.

4. Ejemplo cuando $(PQ)(A) \neq P(A)Q(A)$. Polinomios con coeficientes matriciales no cumplen algunas de las propiedades de los polinomios con coeficientes numéricos. Mostremos un ejemplo de polinomios P, Q con coeficientes matriciales tales que

$$(PQ)(A) \neq P(A)Q(A).$$

Sean

$$P(\lambda) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \lambda, \qquad Q(\lambda) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \qquad A = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

Entonces

$$(PQ)(\lambda) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \lambda, \qquad (PQ)(A) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \qquad P(A)Q(A) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

5. Lema para el teorema de Cayley–Hamilton: condición suficiente para la igualdad $(PQ)(A) = P(A)Q(A)$. Sean P y Q polinomios con coeficientes matriciales:

$$P(\lambda) = \sum_{i=0}^{m} P_i \lambda^i, \qquad Q(\lambda) = \sum_{j=0}^{s} Q_j \lambda^j,$$

donde $P_i, Q_j \in \mathcal{M}_n(\mathbb{F})$, y sea $A \in \mathcal{M}_n(\mathbb{F})$ una matriz que conmuta con todos los coeficientes del polinomio Q:

$$Q_j A = A Q_j \qquad \forall j \in \{0, \ldots, s\}.$$

Entonces

$$(PQ)(A) = P(A)Q(A).$$

Demostración. Recordamos la definición del producto de polinomios:

$$(PQ)(\lambda) = \sum_{k=0}^{m+s} \left(\sum_{\substack{i,j:\\ i+j=k}} P_i Q_j \right) \lambda^k.$$

De allí

$$(PQ)(A) = \sum_{k=0}^{m+s} \left(\sum_{\substack{i,j:\\ i+j=k}} P_i Q_j \right) A^k.$$

Por otro lado,

$$P(A)Q(A) = \left(\sum_{i=0}^{m} P_i A^i \right) \left(\sum_{j=0}^{s} Q_j A^j \right) = \sum_{\substack{0 \leq i \leq m \\ 0 \leq j \leq n}} P_i A^i Q_j A^j.$$

Juntando los sumandos en grupos con $i + j = k$ podemos escribir el resultado de la siguiente manera:

$$P(A)Q(A) = \sum_{k=0}^{m+s} \left(\sum_{\substack{i,j:\\ i+j=k}} P_i A^i Q_j A^j \right).$$

Ahora usamos la condición que A conmuta con Q_j:

$$P_i A^i Q_j A^j = P_i Q_j A^{i+j}.$$

De allí obtenemos que

$$P(A)Q(A) = \sum_{k=0}^{m+s} \sum_{\substack{i,j:\\ i+j=k}} P_i Q_j A^k = (PQ)(A). \qquad \square$$

6. Ejemplo que muestra la idea de la demostración del teorema de Cayley–Hamilton. Sea

$$A = \begin{bmatrix} 3 & -1 & 2 \\ 1 & 4 & -2 \\ 3 & 2 & 3 \end{bmatrix}.$$

Definimos

$$P(\lambda) \coloneqq \operatorname{adj}(\lambda I - A) Q(\lambda) \coloneqq \lambda I - A.$$

Escriba $P(\lambda)$ y $Q(\lambda)$ en forma explícita como matrices con coeficientes polinomiales, luego escriba $P(\lambda)$ y $Q(\lambda)$ como polinomios con coeficientes matriciales y calcule el producto $P(\lambda)Q(\lambda)$.

7. Teorema (Cayley–Hamilton). Sea $A \in \mathcal{M}_n(\mathbb{F})$. Entonces

$$\mathrm{C}_A(A) = \mathbf{0}_{n,n},$$

donde $\mathrm{C}_A(\lambda) = \det(\lambda I - A)$ es el polinomio característico de A.

Demostración. Sabemos que para toda matriz cuadrada B se tiene

$$\operatorname{adj}(B)B = \det(B)I,$$

donde $\operatorname{adj}(B)$ es la matriz adjunta clásica de B o sea la matriz de cofactores transpuesta. Apliquemos este resultado a la matriz $\lambda I - A$:

$$\operatorname{adj}(\lambda I - A)(\lambda I - A) = \det(\lambda I - A) \cdot I = \mathrm{C}_A(\lambda) I.$$

Pongamos

$$P(\lambda) = \operatorname{adj}(\lambda I - A), \qquad Q(\lambda) = \lambda I - A.$$

Estas expresiones son matrices con entradas polinomiales, las vamos a tratar como polinomiales con coeficientes matriciales. Notemos que los coeficientes de Q son I y A, y estas matrices conmutan con A. Por el lema,

$$\mathrm{C}_A(A) = P(A)Q(A) = P(A) \cdot (IA - A) = \mathbf{0}_{n,n}. \qquad \square$$

8. Corolario (teorema de Cayley–Hamilton para operadores lineales). Sea V un espacio vectorial sobre un campo $\mathbb{F}$ de dimensión finita y sea $T \in \mathcal{L}(V)$. Entonces $\mathrm{C}_T(T) = \mathbf{0}$.

9. Nota. Después de estudiar la forma canónica de Jordan de una matriz veremos otra demostración del teorema de Cayley–Hamilton (para el caso $\mathbb{F} = \mathbb{C}$).

10. Corolario. Sean $A \in \mathcal{M}_n(\mathbb{F})$, $f \in \mathcal{P}(\mathbb{F})$. Entonces existe un polinomio $g \in \mathcal{P}(\mathbb{F})$ tal que $\deg(g) < n$ y $g(A) = f(A)$.

Demostración. Dividamos f entre C_A con residuo:

$$f(\lambda) = \mathrm{C}_A(\lambda)q(\lambda) + g(\lambda).$$

Aquí $q, g \in \mathcal{P}(\mathbb{F})$ y $\deg(g) < n$. Sustituyendo λ por A se obtiene que $f(A) = g(A)$. □

11. Ejemplo. Calculemos $f(A)$, donde $f(x) = x^3 - 6x^2 + x - 3$ y la matriz A es la misma que en el ejemplo anterior:

$$A = \begin{bmatrix} 3 & -1 \\ 2 & 1 \end{bmatrix}.$$

Dividamos f entre C_A:

$$f(x) = (x^3 - 4x^2 + 5x - 1) = (x^2 - 4x + 5)(x - 2) + (-12x + 7).$$

Pongamos $g(x) = -12x + 7$. Entonces

$$f(A) = g(A) = -12A + 7I = \begin{bmatrix} -29 & 12 \\ -24 & -5 \end{bmatrix}.$$