

EL TEOREMA DE CAYLEY-HAMILTON

Francisco Gómez García
Universidad de Murcia

RESUMEN

El teorema de Cayley-Hamilton establece que cada matriz cuadrada A satisface su ecuación característica: Si $p(\lambda) = \det(A - \lambda I)$ es el polinomio característico de A , entonces $p(A)$ es la matriz nula.

Entre las diversas demostraciones del teorema hemos encontrado en R. Bellman (1965) una puramente algebraica, que es la que detallamos, con algún matiz, en nuestro trabajo.

El interés de la demostración radica en la utilidad que puede tener para nuestros alumnos de primer curso, la exposición de un desarrollo lógico basado en sus conocimientos básicos de cálculo matricial. También es inmediato y puede ser igualmente útil calcular, a partir del teorema, la inversa de A , cuando A sea no singular.

1. TEOREMA DE CAYLEY-HAMILTON

Sea $p(\lambda) = (-1)^n \lambda^n + c_{n-1} \lambda^{n-1} + c_{n-2} \lambda^{n-2} + \dots + c_2 \lambda^2 + c_1 \lambda + c_0$ el polinomio característico de una matriz A de orden n . Entonces $p(A) = (-1)^n A^n + c_{n-1} A^{n-1} + c_{n-2} A^{n-2} + \dots + c_1 A + c_0 I$ es la matriz nula. Es decir, cada matriz cuadrada A satisface su ecuación característica $p(A) = 0$.

Nota: A es una matriz de orden n con elementos en un cuerpo K ; por tanto, los coeficientes c_i del polinomio característico $\det(A - \lambda I)$ pertenecen a dicho cuerpo K .

2. DEMOSTRACIÓN

Por las propiedades de las matrices se cumple que:

$$(A - \lambda I) \text{Adj}(A - \lambda I)^t = p(\lambda)I$$

donde $\text{Adj}(A - \lambda I)^t$ es la matriz traspuesta de la matriz de los adjuntos de los elementos respectivos de la matriz $A - \lambda I$ y $p(\lambda) = \det(A - \lambda I)$ es el polinomio característico de la matriz A .

Si denotamos $B(\lambda) = \text{Adj}(A - \lambda I)^t$, entonces $B(\lambda)$ es una matriz polinómica en λ , de grado $n-1$, que se puede escribir como:

$$B(\lambda) = B_{n-1} \lambda^{n-1} + B_{n-2} \lambda^{n-2} + \dots + B_2 \lambda^2 + B_1 \lambda + B_0$$

donde cada B_i es una matriz de orden n , con elementos en el cuerpo K . Entonces el producto $(A - \lambda I) B(\lambda)$ vale:

$$\begin{aligned} (A - \lambda I) B(\lambda) &= (A - \lambda I)(B_{n-1} \lambda^{n-1} + B_{n-2} \lambda^{n-2} + \dots + B_2 \lambda^2 + B_1 \lambda + B_0) = -B_{n-1} \lambda^n + \\ & (AB_{n-1} - B_{n-2}) \lambda^{n-1} + (AB_{n-2} - B_{n-3}) \lambda^{n-2} + \dots + (AB_2 - B_1) \lambda^2 + (AB_1 - B_0) \lambda + AB_0 \end{aligned}$$

Por otro lado $p(\lambda)I$ es la matriz polinómica:

$$p(\lambda)I = (-1)^n I \lambda^n + c_{n-1} I \lambda^{n-1} + c_{n-2} I \lambda^{n-2} + \dots + c_2 I \lambda^2 + c_1 I \lambda + c_0 I$$

Luego, igualando las matrices polinómicas, con elementos en el dominio $K(\lambda)$, $(A - \lambda I) B(\lambda) = p(\lambda)I$, se deduce que:

$$\begin{aligned} -B_{n-1} &= (-1)^n I \\ AB_{n-1} - B_{n-2} &= c_{n-1} I \\ AB_{n-2} - B_{n-3} &= c_{n-2} I \\ &\vdots \\ AB_2 - B_1 &= c_2 I \\ AB_1 - B_0 &= c_1 I \\ AB_0 &= c_0 I \end{aligned}$$

Si vamos sustituyendo cada matriz B_i en la siguiente ecuación hasta llegar a la penúltima resulta:

$$\begin{aligned}
 -B_{n-1} &= (-1)^n I \\
 -B_{n-2} &= (-1)^n A + c_{n-1} I \\
 -B_{n-3} &= (-1)^n A^2 + c_{n-1} A + c_{n-2} I \\
 -B_{n-4} &= (-1)^n A^3 + c_{n-1} A^2 + c_{n-2} A + c_{n-3} I \\
 &\vdots \\
 -B_2 &= (-1)^n A^{n-3} + c_{n-1} A^{n-4} + c_{n-2} A^{n-5} + \dots + c_4 A + c_3 I \\
 -B_1 &= (-1)^n A^{n-2} + c_{n-1} A^{n-3} + c_{n-2} A^{n-4} + \dots + c_3 A + c_2 I \\
 B_0 &= (-1)^n A^{n-1} + c_{n-1} A^{n-2} + c_{n-2} A^{n-3} + \dots + c_2 A + c_1 I
 \end{aligned}$$

Entonces sustituyendo B_0 en la última ecuación $A B_0 = c_0 I$ se obtiene:

$$-A B_0 = (-1)^n A^n + c_{n-1} A^{n-1} + c_{n-2} A^{n-2} + \dots + c_2 A^2 + c_1 A = -c_0 I$$

Por tanto, $(-1)^n A^n + c_{n-1} A^{n-1} + c_{n-2} A^{n-2} + \dots + c_2 A^2 + c_1 A + c_0 I = 0$. Es decir, $p(A) = 0$ c.q.d.

3. COMENTARIOS

R. Bellman supone en su razonamiento que λ no es raíz característica, lo cual no es necesario. Después prueba que A y B_i conmutan (en nuestra demostración es inmediato), para a continuación afirmar que $p(\lambda) I = (A - \lambda I) B(\lambda)$ es válida no sólo para escalares λ , sino para cualquier matriz que conmute con B_i , y en particular lo es para $\lambda = A$ y entonces $p(A) = 0$. Conclusión que resulta de esa igualdad $p(\lambda) I = (A - \lambda I) B(\lambda) = -B_{n-1} \lambda^n + (AB_{n-1} - B_{n-2}) \lambda^{n-1} + (AB_{n-2} - B_{n-3}) \lambda^{n-2} + \dots + (AB_2 - B_1) \lambda^2 + (AB_1 - B_0) \lambda + A B_0$ tomando $\lambda = A$:

$$\begin{aligned}
 p(A) I &= -B_{n-1} A^n + (AB_{n-1} - B_{n-2}) A^{n-1} + (AB_{n-2} - B_{n-3}) A^{n-2} + \\
 &\quad + \dots + (AB_2 - B_1) A^2 + (AB_1 - B_0) A + A B_0 = 0
 \end{aligned}$$

ya que A y B_i conmutan y cada sumando se anula con el siguiente.

4. REFERENCIAS BIBLIOGRÁFICAS

BELLMAN, R (1965). “Introducción al Análisis Matricial”. Ed. Reverté, S.A., p. 223.

DUBREIL, P. y DUBREIL-JACOTIN, M.L. (1971). “Lecciones de álgebra moderna”. Ed. Reverté, S.A., p. 363.