Observación. No se puede probar este teorema sustituyendo $\lambda = A$ para obtener $P(A) = Q(A)(A - A) = 0$. Esto se debe a que es posible encontrar polinomios $P(\lambda)$ y $Q(\lambda)$ con coeficientes matriciales tales que $F(\lambda) = P(\lambda)Q(\lambda)$, pero $F(A) \neq P(A)Q(A)$. (Vea el problema 8.8.17.)

Ahora se puede establecer el teorema principal.

Teorema 8.8.2 Teorema de Cayley-Hamilton

Toda matriz cuadrada satisface su propia ecuación característica. Es decir, si $p(\lambda) = 0$ es la ecuación característica de A, entonces $p(A) = 0$.

Nota

El teorema recibe el nombre en honor de Sir William Rowan Hamilton y Arthur Cayley (1821-1895). Cayley publicó el primer análisis de este famoso teorema en 1858. Por su parte, Hamilton descubrió (pero no demostró) el resultado en su trabajo sobre cuaterniones.

Demostración

Se tiene

$$p(\lambda) = \det(A - \lambda I) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix}$$

Es claro que cualquier cofactor de $(A - \lambda I)$ es un polinomio en λ. Así, la adjunta de $A - \lambda I$ (vea la definición 3.3.1) es una matriz de $n \times n$ en la que cada componente es un polinomio en λ. Es decir,

$$\operatorname{adj}(A - \lambda I) = \begin{pmatrix} p_{11}(\lambda) & p_{12}(\lambda) & \cdots & p_{1n}(\lambda) \\ p_{21}(\lambda) & p_{22}(\lambda) & \cdots & p_{2n}(\lambda) \\ \vdots & \vdots & & \vdots \\ p_{n1}(\lambda) & p_{n2}(\lambda) & \cdots & p_{nn}(\lambda) \end{pmatrix}$$

Esto significa que se puede pensar en adj $(A - \lambda I)$ como en un polinomio, $Q(\lambda)$, en λ cuyos coeficientes son matrices de $n \times n$. Para entender esto, véase lo siguiente:

$$\begin{pmatrix} \lambda^2 - 2\lambda + 1 & 2\lambda^2 - 7\lambda - 4 \\ 4\lambda^2 + 5\lambda - 2 & -3\lambda^2 - \lambda + 3 \end{pmatrix} = \begin{pmatrix} -1 & 2 \\ 4 & -3 \end{pmatrix}\lambda^2 + \begin{pmatrix} -2 & -7 \\ 5 & -1 \end{pmatrix}\lambda + \begin{pmatrix} 1 & -4 \\ -2 & 3 \end{pmatrix}$$

Del teorema 3.3.2,

$$[\det(A - \lambda I)]I = [\operatorname{adj}(A - \lambda I)][A - \lambda I] = Q(\lambda)(A - \lambda I) \tag{8.8.5}$$

Pero $\det(A - \lambda I)I = p(\lambda)I$. Si

$$P(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0$$

entonces se define

$$p(\lambda) = p(\lambda)I = \lambda^n I + a_{n-1}\lambda^{n-1}I + \cdots + a_1\lambda I + a_0 I$$

Por lo tanto, de (8.8.5) se tiene $P(\lambda) = Q(\lambda)(A - \lambda I)$. Por último, del teorema 8.8.1, $P(A) = 0$. Esto completa la prueba.