ÁLGEBRA LINEAL (R211 - CE9) 2024

4.4 El teorema de Cayley-Hamilton

El teorema de Cayley-Hamilton nos dice que dada una matriz $A \in F^{n\times n}$, el polinomio minimal divide al característico, y sigue inmediatamente que entonces el característico anula a A. Es un Very Important Theorem, tiene nombre propio.

Aquí es importante observar lo siguiente: hemos trabajado con la expresión $\chi_A(X) = det(XI - A)$, de modo que parecería ser trivial la afirmación del teorema de Cayley-Hamilton: $\chi_A(A) = det(AI - A) = det(0) = 0 \in F$. Pero esto es incorrecto, pues al aplicar un polinomio a una matriz debe resultar en una matriz: si el polinomio es $\chi_A(X) = det(XI - A) = p(X) = \sum_{i=0}^{n} a_i X^i = a_0 + a_1 X + \cdots + a_n X^n$, entonces $\chi_A(A) = a_0 I + a_1 A + \cdots + a_n A^n \in F^{n\times n}$. Hay que tener, entonces, mucho cuidado al aplicar la expresión para el polinomio característico a las matrices.

Observemos previamente lo siguiente: consideremos $A \in F^{n\times n}$ y $T : F^n \to F^n$ definida por $T(x) = Ax$. Sean $v \in F^n$ y $k \in \mathbb{N}$ tales que $B = \{v, Tv, \ldots, T^k v\}$ es li y $T^{k+1}v = -a_0 v - a_1 Tv - \cdots - a_k T^k v$ es ld. Para calcular $[T]_B$, puesto que sus columnas son las componentes en la base B de los transformados de la base B, tenemos que:

- el primer elemento v de B se transforma en Tv, que es el segundo elemento de B, luego $[Tv]_B = e_2$;

- el segundo elemento Tv de B se transforma en $T^2 v$, que es el tercer elemento de B, luego $[T(Tv)]_B = e_3$;

- así sucesivamente,

- el elemento k-ésimo de B es $T^{k-1}v$ y se transforma en $T^k v$, que es el elemento $k+1$-ésimo de B, luego $[T(T^{k-1}v)]_B = e_{k+1}$;

- el elemento $k+1$-ésimo de B es $T^k v$ y se transforma en $T^{k+1}v = -a_0 v - a_1 Tv - \cdots - a_k T^k v$, luego $[T(T^k v)]_B = (-a_0, -a_1, \ldots, -a_{k-1}, -a_k)$.

Entonces $[T]_B$ es la matriz $k+1 \times k+1$ dada por

$$[T]_B = \begin{pmatrix} 0 & 0 & 0 & \ldots & 0 & -a_0 \\ 1 & 0 & 0 & \ldots & 0 & -a_1 \\ 0 & 1 & 0 & \ldots & 0 & -a_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 0 & -a_{k-1} \\ 0 & 0 & 0 & \ldots & 1 & -a_k \end{pmatrix}.$$

Además, si queremos calcular su polinomio característico $\chi_T(X) = \chi_{[T]_B}(X) = det(XI - [T]_B)$, donde

$$XI - [T]_B = \begin{pmatrix} X & 0 & 0 & \dots & 0 & a_0 \\ -1 & X & 0 & \dots & 0 & a_1 \\ 0 & -1 & X & \dots & 0 & a_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & X & a_{k-1} \\ 0 & 0 & 0 & \dots & -1 & X + a_k \end{pmatrix},$$

tenemos que $\chi_T(X) = X^{k+1} + a_k X^n + \dots + a_1 X + a_0$. Nos convencemos fácilmente de este resultado si desarrollamos por la última fila.

Desafío 1 *Convencerse.*

Very important: *convencerse* es sólo un eufemismo para *demostrar*. Hay que probarlo, como todo en la ciencia. (*Ciencia* siendo un eufemismo para *Álgebra Lineal*, como pretendo convencerlos este cuatrimestre, o quizás debería decir demostrarles, je.)

Todos estos cálculos nos permitirán desarrollar la prueba del Teorema de Cayley-Hamilton.

Very Important Theorem 1 ***El teorema de Cayley-Hamilton.*** *$A \in F^{n\times n}$. Ents. $m_A | \chi_A$. Equivalentemente, $\chi_A(A) = 0 \in F^{n\times n}$.*

Demostración: Consideremos $T : F^n \to F^n$ definida por $T(x) = Ax$. Sean $v \in F^n$ y $k \in F$ tq $B_1 = \{v, Tv, \dots, T^k v\}$ es li y $T^{k+1}v = -a_0 v - a_1 Tv - \dots - a_k T^k v$. Ents. $m_v(X) = X^{k+1} + a_k X^k + \dots + a_1 X + a_0$. Extendemos el conj. li a una base de F^n: $B = \{v, Tv, \dots, T^k v, w_{k+2}, \dots, w_n\}$. Así, gracias a la observación que hicimos previamente, tenemos que

$$[T]_B = \left(\begin{array}{cccccc|c} 0 & 0 & 0 & \dots & 0 & -a_0 & \\ 1 & 0 & 0 & \dots & 0 & -a_1 & \\ 0 & 1 & 0 & \dots & 0 & -a_2 & \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & M \\ 0 & 0 & 0 & \dots & 0 & -a_{k-1} & \\ 0 & 0 & 0 & \dots & 1 & -a_k & \\ \hline & & & \mathbf{0} & & & N \end{array}\right).$$

Observemos esta matriz por bloques: el bloque superior izquierdo es $k + 1 \times k + 1$, la matriz de la observación previa. Debajo de ese bloque, hay un bloque $n - (k + 1) \times k + 1$ cuyas entradas son todas 0. El resto de las entradas quedan en dos bloques que llamamos M y N, con N cuadrada.

Tenemos que $\chi_A(X) = \chi_T(X) = \chi_{[T]_B}(X) = det(XI - [T]_B)$. Luego,

$$XI - [T]_B = \left(\begin{array}{cccccc|c} X & 0 & 0 & \dots & 0 & a_0 & \\ -1 & X & 0 & \dots & 0 & a_1 & \\ 0 & -1 & X & \dots & 0 & a_2 & \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & -M \\ 0 & 0 & 0 & \dots & X & a_{k-1} & \\ 0 & 0 & 0 & \dots & -1 & X + a_k & \\ \hline & & & \mathbf{0} & & & XI_{n-(k+1)} - N \end{array}\right).$$

Así, de acuerdo a lo observado anteriormente, $\chi_A(X) = (X^{k+1} + a_k X^n + \dots + a_1 X + a_0) det(XI_{n-(k+1)} - N) = m_v(X)Q(X)$. Sigue que $m_v | \chi_A$. Como v es arbitrario, en particular para la base canónica $B = \{e_1, \dots, e_n\}$ se tiene que $m_{e_i} | \chi_A$ para todo $i = 1, \dots, n$, de donde $m_A = m.c.m.\{m_{e_1}, \dots, m_{e_n}\} | \chi_A$.

□

Observaciones 1 *Como consecuencias inmediatas sigue que, si $A \in F^{n \times n}$,*

1. *$gr(m_A) \leq n$,*
2. *si $gr(m_A) = n$ ents. $m_A = \chi_A$,*
3. *si existe $v \in F$ tq $gr(m_v) = n$ ents. $m_v = m_A = \chi_A$.*

Ejemplo 1 $A = \begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix} \in \mathbb{R}^{n \times n}$. *Daremos una fórmula general para A^n, y más. Calculemos $det(XI - A) = ... = (X - 1)^2$ (EJERCICIO). Luego $\chi_A(X) = X^2 - 2X + I$. Por el teorema de Cayley-Hamilton es $\chi_A(A) = A^2 - 2A + I = 0$. Sigue que $A^2 = 2A - I$. Y podemos inducir $A^3 = A^2 A = (2A - I)A = 2A^2 - A = 2(2A - I) - A = 3A - 2I$, $A^4 = 4A - 3I$, $\dots$, $A^n = nA - (n - 1)I$. (EJERCICIO). Más aún, observemos lo siguiente: nuevamente, de $\chi_A(A) = A^2 - 2A + I = 0$ sigue que $I = 2A - A^2 = A(2I - A)$ y por lo tanto $A^{-1} = 2I - A$. Es claro que en este ejemplo A es invertible.*

Desafío 2 *Justificar las siguientes observaciones:*

- *Para dos autovalores distintos, digamos $\lambda_1 \neq \lambda_2$, y dos autovalores v_1, v_2 correspondientes a λ_1, λ_2, respectivamente, se tiene que $\{v_1, v_2\}$ es li.*
- *Si A es diagonalizable y $\lambda = 0$ es autovalor de A, entonces A no es invertible.*

En general tenemos entonces que:

Proposición 1 *$A \in F^{n \times n}$ invertible. Luego $A^{-1} \in span\{I, A, \dots, A^{n-1}\}$.*

Demostración: Por el teorema de Cayley-Hamilton sabemos que $\chi_A(X) = X^n + a_{n-1}X^{n-1} + \dots + a_1 X + a_0$ es tal que $\chi_A(A) = 0 \in F^{n \times n}$. De donde

$$A^n + a_{n-1}A^{n-1} + \dots + a_1 A + a_0 I = 0.$$

Ahora bien, $\chi_A(0) = a_0 \in F$, luego $\chi_A(0) = a_0 = det(0I_n - A) = det(-A) = (-1)^n det(A) \neq 0$. De estas igualdades sigue que

$$I_n = -\frac{1}{a_0}(A^n + a_{n-1}A^{n-1} + \dots + a_1 A) = -\frac{1}{a_0}(A^{n-1} + a_{n-1}A^{n-2} + \dots + a_1 I)A,$$

luego

$$A^{-1} = -\frac{1}{a_0}(A^{n-1} + a_{n-1}A^{n-2} + \dots + a_1 I).$$

$\square$

Observemos que si $\chi_A(X)$ se factoriza linealmente como $\chi_A(X) = (X - \lambda_1)\dots(X - \lambda_n)$ con $\lambda_i \neq \lambda_j$ para $i, j = 1, \dots, n$, entonces $F^n = \oplus_{i=1}^n E_{\lambda_i}$, de donde F^n tiene una base de autovectores y por lo tanto A resulta diagonalizable. La recíproca no es cierta (la identidad es un ejemplo simple de esto, por qué?). Es decir, no podemos caracterizar diagonalizabilidad de esta forma mediante el polinomio característico, pero sí podemos hacerlo con el polinomio minimal:

Proposición 2 *$A \in F^{n\times n}$. Entonces A es diagonalizable en $F^{n\times n}$ si y sólo si m_A tiene todas sus raíces en F y son simples.*

Demostración:

$\Rightarrow$) A diagonalizable. Sean $\lambda_1, \dots, \lambda_r$ autovalores de A distintos y sea $B = \{v_1, \dots, v_n\}$ base de autovectores de A. Para cada $v \in V$, su polinomio minimal respecto de A es $m_A(X) = m.c.m.\{m_{v_1}, \dots, m_{v_n}\} = m.c.m.\{X - \lambda_{j_1}, \dots, X - \lambda_{j_n}\}$, puesto que cada v_i es autovector correspondiente a un autovalor λ_{j_i} y luego $m_{v_i}(X) = X - \lambda_{v_j} \in F[X]$ (recordar ejemplo anterior cuando definimos polinomio minimal, en el caso de un autovector). Resulta entonces $m_A(X) = (X - \lambda_1)\dots(X - \lambda_r)$.

$\Leftarrow$) $m_A(X) = (X - \lambda_1)\dots(X - \lambda_r)$ con $\lambda_i \neq \lambda_j$ y $\lambda_i \in F$ para $i, j = 1, \dots, r$, todos los autovalores de A en F.

Vamos a ver que $F^n = \oplus_{i=1}^r E_{\lambda_i}$. En efecto, $E_{\lambda_i} = \{x \in F^n : Ax = \lambda_i x\}$.

Consideremos $v \in F^n$, y sea $k \in \mathbb{N}$ tal que $B_1 = \{v, Av, \dots, A^k v\}$ es li y $\{v, Av, \dots, A^k v, A^{k+1} v\}$ es ld. Sea $U = span(B_1)$. Tenemos así que $gr(m_v) = k+1$ y $m_v(X) = a_0 + a_1 X + \dots + a_k X^k + X^{k+1}$. Además, si definimos $T : U \to U$ por $T(u) = Au$, como ya hemos observado, tenemos que

$$[T]_{B_1} = \begin{pmatrix} 0 & 0 & \dots & 0 & -a_0 \\ 1 & 0 & \dots & 0 & -a_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & -a_{k-1} \\ 0 & 0 & \dots & 1 & -a_k \end{pmatrix}.$$

No perder de vista que $T \in L(U)$, es decir, es un endomorfismo del subespacio U. Sigue que $\chi_{T_A}(X) = det(XI_{k+1} - [T]_{B_1}) = a_0 + a_1 X + \dots + a_k X^k + X^{k+1} = m_v(X)$. Como $\chi_T = m_v | m_A$ (¿porqué?) sigue que todas las raíces de χ_T son autovalores de A, luego son simples y están en el cuerpo F. Por la observación que hicimos antes, resulta que T es diagonalizable en U, de donde $\chi_T(X) = (X - \lambda_{i_1})\dots(X - \lambda_{i_{k+1}})$. Como $v \in U$ existen únicos $v_{i_1}, \dots, v_{i_{k+1}}$ autovectores de T correspondientes $\lambda_{i_1}, \dots, \lambda_{i_{k+1}}$, respectivamente, tales que $v = v_{i_1} + \dots + v_{i_{k+1}}$. Vale decir, $v \in \oplus_{j=1}^{k+1} E_{\lambda_{i_j}}$.

Como $v \in F^n$ es arbitrario, sigue que $F^n \subset \oplus_{i=1}^r E_{\lambda_i}$. Concluimos que A es diagonalizable.

$\square$

Ejemplos 1 1. *Sea $A \in \mathbb{C}^{n\times n}$ tq $A^k = I$ para algún $k \in \mathbb{N}$. Veamos que A es diagonalizable. En efecto, $A^k = I \Rightarrow A^k - I = 0 \Rightarrow p(X) = X^k - 1$ anula a A. $\therefore m_A | p$. Como p tiene todas sus raíces en $\mathbb{C}$ y son simples, resulta que m_A también, de donde sigue que A es diagonalizable.*

2. *Sea $A = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \in \mathbb{R}^{n\times n}$. A es tal que $A^3 = I$, pero A no es diagonalizable: $m_A(X) = (X-1)(X^2+X+1)$ (EJERCICIO) no tiene todas sus raíces en $\mathbb{R}$.*