MATH-UA. 0148 Honors Linear Algebra

Lecture 29: The Cayley - Hamilton Theorem

I/ Characteristic polynomial of an endomorphism

1) Characteristic polynomial of a square matrix

Definition: The characteristic polynomial of an $n \times n$ matrix A with real coefficients is the polynomial

$$p_A(x) = \det(A - xI)$$

Proposition 1: λ is an eigenvalue of the $n \times n$ matrix A with real coefficients if and only if λ is a root of its characteristic polynomial p_A.

We already proved this result in Lecture 21

Proposition 2: Two similar matrices A and B have the same characteristic polynomial.

Proof: Let A and B such that $B = M^{-1}AM$

$$\begin{aligned} p_B = p_{M^{-1}AM} &= \det(M^{-1}AM - xI) \\ &= \det(M^{-1}(A - xI)M) \\ &= \det(M^{-1})\, p_A \det(M) = p_A \quad \blacksquare \end{aligned}$$

<u>Proposition 3</u>: Let M be an $n \times n$ matrix with real coefficients which we decompose in the following blocks:

$$M = \begin{vmatrix} A & B \\ O & C \end{vmatrix}$$

where A and C are square matrices.

We have $p_M = p_A \, p_C$

<u>Proof</u>: $p_M = \begin{vmatrix} A - I & B \\ O & C - I \end{vmatrix} = |A - I| \, |C - I| = p_A \, p_C$

2) <u>Characteristic polynomial of an endomorphism</u>

<u>Definition</u>: An endomorphism of a vector space V is a linear transformation $L : V \longrightarrow V$

<u>Examples</u>: • Any projection P is an endomorphism, since $P(V) \subset V$

• The linear transformation

$$L : P_3 \longrightarrow P_3$$
$$p \in P_3 \longmapsto q(x) = x \frac{dp}{dx} - 2p$$

which we studied last time, is an endomorphism.

• The linear transformation

$$L : \mathbb{R}^3 \longrightarrow P_2$$
$$\vec{a} = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} \in \mathbb{R}^3 \longmapsto p(x) = a_2 x^2 + a_1 x + a_0$$

is not an endomorphism.

Since $L: V \longrightarrow V$, the matrix representation A of L in any basis $\mathcal{B}$ is a square matrix.

Now, let B be the matrix representation of L in another basis $\mathcal{B}'$. There exists an invertible matrix M such that

$$B = M^{-1} A M$$

From proposition 2 of the previous section, A and B have the same characteristic polynomial, which can be viewed as the unique characteristic polynomial of the endomorphism L

Definition: The characteristic polynomial P_L of an endomorphism L is the characteristic polynomial of the matrix representation of L in any basis of the vector space V.

λ is an eigenvalue of the endomorphism L if and only if it is a root of the characteristic polynomial P_L.

If λ is an eigenvalue of the endomorphism L, then there exists a nonzero vector $\vec{x}$ such that

$$L\vec{x} = \lambda \vec{x}$$

$$\Leftrightarrow (L - \lambda Id)\vec{x} = \vec{0}_V$$

$$\Leftrightarrow Ker(L - \lambda Id) \neq \{\vec{0}_V\}$$

$$\Leftrightarrow L - \lambda Id \text{ is not injective.}$$

II] Polynomials of matrices

1) Definition

Let A be an $n \times n$ matrix with real coefficients. For a given polynomial $p(x) = a_m x^m + a_{m-1} x^{m-1} + \dots + a_1 x + a_0$ in P_m, we define the matrix

$$P(A) = a_m A^m + a_{m-1} A^{m-1} + \dots + a_1 A^1 + a_0 I$$

Proposition 1: Let p_1 and p_2 be two polynomials of P_m. Then, for any square matrix A,

i) $(p_1 + p_2)(A) = p_1(A) + p_2(A)$
ii) $(p_1 p_2)(A) = p_1(A)\, p_2(A)$

Proof: Let $p_1 = \sum_{i=0}^{m} a_i x^i$, $p_2 = \sum_{i=0}^{m} b_i x^i$

where some of the a_i or the b_i are potentially zero. We have

$$(p_1 + p_2)(A) = \sum_{i=0}^{m} (a_i + b_i) A^i = \sum_{i=0}^{m} a_i A^i + \sum_{i=0}^{m} b_i A^i$$

$$= p_1(A) + p_2(A)$$

$$(p_1 p_2)(A) = \sum_{i=0}^{m} \sum_{j=0}^{m} a_i b_j A^{i+j} = \sum_{i=0}^{m} a_i A^i \sum_{j=0}^{m} b_j A^j$$

$$= p_1(A)\, p_2(A)$$

Observe that since any two polynomials p_1 and p_2 commute, the previous result also tell us that the matrices $p_1(A)$ and $p_2(A)$ commute.

Proposition 2: Let A be an $n\times n$ matrix with real coefficients and p_1 and p_2 be two polynomials in P_m. Then the matrices $p_1(A)$ and $p_2(A)$ commute: $p_1(A)p_2(A) = p_2(A)p_1(A)$.

Example: Let $A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$, $p_1(x) = x^2 + x + 1$ and $p_2(x) = x - 1$

$$p_1(A) = A^2 + A + 1 = \begin{bmatrix} 3 & 2 \\ 0 & 1 \end{bmatrix}, \quad p_2(A) = \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix}$$

$$p_1(A)p_2(A) = \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix} = p_2(A)p_1(A)$$

2) Matrices as roots of polynomials

Definition: Let A be an $n\times n$ matrix with real coefficients. We say that A is a root of the polynomial $p \in P_m$ if $p(A)$ is the zero matrix.

Example: Any 2×2 matrix A with real coefficients is the root of the polynomial

$$p_A(x) = x^2 - \mathrm{Trace}(A)x + \det(A)$$

Indeed, let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, $A^2 = \begin{bmatrix} a^2+bc & ab+bd \\ ac+cd & cb+d^2 \end{bmatrix}$

$$A^2 - \text{Trace}(A)A + \det A \, I = \begin{bmatrix} a^2+bc-a^2-ad+ad-bc & ab+bd-ab-bd \\ ac+cd-ac-cd & cb+d^2-ad-d^2+ad-bc \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Proposition 1: Any $n \times n$ matrix is the root of at least one polynomial in P_m with $m \in \mathbb{N}^*$

Proof: The vector space of $n \times n$ matrices with real coefficients has dimension n^2. Therefore any set of n^2+1 $n \times n$ matrices with real coefficients cannot be linearly independent. This is particularly true for $(A^{n^2}, A^{n^2-1}, \ldots, A^2, A, I)$: there exist $c_0, c_1, \ldots, c_{n^2}$ such that

$$c_{n^2} A^{n^2} + c_{n^2-1} A^{n^2-1} + \ldots + c_2 A^2 + c_1 A + c_0 I = 0$$

A is the root of the polynomial

$$p(x) = c_{n^2} x^{n^2} + \ldots + c_2 x^2 + c_1 x + c_0$$

Proposition 2: Let A be an $n\times n$ matrix with real coefficients, and suppose A is the root of the polynomial $p \in P_m$, with $m \in \mathbb{N}^*$.
Then any eigenvalue of A is a root of p.

Proof: Let $p = a_m x^m + \dots + a_2 x^2 + a_1 x + a_0$ such that $p(A) = O$. We thus have $a_n A^n + \dots + a_2 A^2 + a_1 A + a_0 = O$

For any $\vec{x} \in \mathbb{R}^n$, we can then write

$$(a_m A^m + \dots + a_2 A^2 + a_1 A + a_0)\vec{x} = \vec{0}$$

If $\vec{x}$ is an eigenvector of A with corresponding eigenvalue λ, then

$$a_1 A\vec{x} = a_1 \lambda \vec{x}$$

$$a_2 A^2 \vec{x} = a_2 \lambda A\vec{x} = a_2 \lambda^2 \vec{x}$$

$$\vdots$$

$$a_m A^m \vec{x} = a_m \lambda^m \vec{x}$$

Thus, $(a_m \lambda^m + \dots + a_2 \lambda^2 + a_1 \lambda + a_0)\vec{x} = \vec{0}$
And since $\vec{x} \neq 0$, this implies $p(\lambda) = 0$: λ is a root of p ■

⚠ The converse may not be true: p can have roots which are not eigenvalues of A.
Here is an illustration: I, the identity matrix, is a root of $p(x) = x(x-1)$, but 0 is not an eigenvalue of I.

III/ The Cayley-Hamilton theorem

Theorem (Cayley-Hamilton Theorem)

Any $n\times n$ matrix A with real coefficients is a root of its characteristic polynomial: $p_A(A)=0$.

Proof: We recall from lecture 25 that A is similar to a Jordan matrix J: there exists an invertible matrix M such that

$$J = M^{-1}AM$$

where $J = \begin{bmatrix} T_1 & & * \\ & T_2 & \ddots \\ 0 & & \ddots & * \\ & & & T_k \end{bmatrix}$ where $T_i = \begin{bmatrix} \lambda_i & & * \\ & \ddots & \\ 0 & & \lambda_i \end{bmatrix}$

We have $p_A(x) = p_J(x) = (x-\lambda_1)^{n_1}(x-\lambda_2)^{n_2}\cdots(x-\lambda_k)^{n_k}$

where the n_i are the number of rows in T_i

Now, for each i, $T_i - \lambda_i I$ is an $n_i\times n_i$ matrix with zeros on the diagonal. Hence, $(T_i-\lambda_i I)^{n_i}=0$

A upper triangular, zeros on diagonal

$A\vec{e}_1=\vec{0}$

$A e_i = \text{Span}(e_1,\dots,e_{i-1})$

continue to

$A^n\vec{x} = cA\vec{e}_1=\vec{0}$

Thus, $(J-\lambda_i I)^{n_i} = \begin{bmatrix} (T_1-\lambda_i I)^{n_i} & * & * & \cdots \\ & \ddots & & \vdots \\ & & (T_i-\lambda_i I)^{n_i} & \vdots \\ & & & * \\ & & & (T_k-\lambda_i I)^{n_k} \end{bmatrix}$

$$= \begin{bmatrix} (T_1 - \lambda_i I)^{n_1} & & & & \\ & \ddots & & & \\ & & (T_{i-1} - \lambda_{i-1} I)^{n_{i-1}} & & & \\ & & & 0 & & \\ & & & & (T_{i+1} - \lambda_{i+1} I)^{n_{i+1}} & \\ & & & & & \ddots \\ & & & & & & (T_\ell - \lambda_i I)^{n_\ell} \end{bmatrix}$$

Thus, $(J - \lambda_1 I)^{n_1} (J - \lambda_2 I)^{n_2} \cdots (J - \lambda_\ell I)^{n_\ell} = 0$

$\iff p_A(J) = 0.$

Now, $M^{-1} p_A(A) M = M^{-1} (A - \lambda_1 I)^{n_1} \cdots (A - \lambda_\ell I)^{n_\ell} M$

$= (M^{-1} A M - \lambda_1 I)^{n_1} \cdots (M^{-1} A M - \lambda_\ell I)^{n_\ell}$

$= (J - \lambda_1 I)^{n_1} \cdots (J - \lambda_\ell I)^{n_\ell}$

$= 0$

Thus, $M^{-1} p_A(A) M = 0 \Rightarrow p_A(A) = 0$ since M is invertible ■

<u>Examples:</u> • Prove the Cayley-Hamilton theorem holds for

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

$p(\lambda) = \det(A - \lambda I) = (1-\lambda)(4-\lambda) - 6 = \lambda^2 - 5\lambda - 2$

$A^2 = \begin{bmatrix} 7 & 10 \\ 15 & 22 \end{bmatrix} \Rightarrow A^2 - 5A - 2I = \begin{bmatrix} 7-5-2 & 10-10 \\ 15-15 & 22-20-2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

- Use the Cayley-Hamilton theorem to find the inverse of the matrix

$$A = \begin{bmatrix} 2 & 0 & 1 \\ -2 & 3 & 4 \\ -5 & 5 & 6 \end{bmatrix}$$

$$p(\lambda) = \begin{vmatrix} 2-\lambda & 0 & 1 \\ -2 & 3-\lambda & 4 \\ -5 & 5 & 6-\lambda \end{vmatrix} = (2-\lambda)\left[(3-\lambda)(6-\lambda)-20\right] + 5(3-\lambda) - 10$$

$$= -\lambda^3 + 11\lambda^2 - 21\lambda + 1$$

By the Cayley-Hamilton theorem,

$$A^3 - 11A^2 + 21A - I = 0$$

$$\Leftrightarrow A^{-1} = A^2 - 11A + 21I$$

$$A^2 = \begin{bmatrix} -1 & 5 & 8 \\ -30 & 29 & 34 \\ -50 & 45 & 51 \end{bmatrix}$$

$$A^{-1} = \begin{bmatrix} -1 & 5 & 8 \\ -30 & 29 & 34 \\ -50 & 45 & 51 \end{bmatrix} - \begin{bmatrix} 22 & 0 & 11 \\ -22 & 33 & 44 \\ -55 & 55 & 66 \end{bmatrix} + \begin{bmatrix} 21 & 0 & 0 \\ 0 & 21 & 0 \\ 0 & 0 & 21 \end{bmatrix}$$

$$= \begin{bmatrix} -2 & 5 & -3 \\ -8 & 17 & -10 \\ 5 & -10 & 6 \end{bmatrix}$$