1 Cayley-Hamilton Theorem

The goal of this section is to prove the famous Cayley-Hamilton Theorem, which asserts that if $p(x)$ is the characteristic polynomial of an n by n matrix A, then $p(A) = 0$.

Definition 1 *Let $p(x) = a_0 + a_1 x + ... + a_k x^k \in \mathbb{K}[X]$, and let $A \in \mathcal{M}_n(\mathbb{K})$. Define the matrix $p(A)$ by*

$$p(A) = a_0 I_n + a_1 A + ... + a_k A^k.$$

In other words, $p(A)$ is the matrix obtained by replacing x^i by A^i, for each $i = 0, 1, ..., k$, in the expression of p, with the convention $A^0 = I_n$.

Remark 2 *If we replace x by A in the formula of the characteristic polynomial $p_A(x)$, which gives*

$$p_A(A) = \det(A - A) = \det(0) = 0.$$

This is impossible since $p_A(A) \in \mathcal{M}_n(\mathbb{K})$ and $\det(A - A) = \det(0) \in \mathbb{K}$.

Let us recall the statement of one of the very classical theorem.

Theorem 3 (Cayley-Hamilton Theorem) *Let $A \in \mathcal{M}_n(\mathbb{R})$ and let $p_A(x)$ be its characteristic polynomial. Then $p_A(A) = 0$.*

In the proof, we need to use the following lemma.

Lemma 4 *For each $A \in \mathcal{M}_n(\mathbb{R})$, we have*

$$A\,(com\,(A))^t = (com\,(A))^t\,A = \det A I_n. \tag{1}$$

In particular, if A is invertible, its inverse is given by

$$A^{-1} = \frac{1}{\det(A)}\,(com\,(A))^t\,.$$

For example, if $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_2(\mathbb{R})$, we have

$$\begin{aligned} A.\,(com\,(A))^t &= \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} ad - bc & 0 \\ 0 & ad - bc \end{pmatrix} \\ &= (ad - bc) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \det(A)\, I_2. \end{aligned}$$

Proof of Cayley-Hamilton Theorem. Let

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \in M_n(\mathbb{R})\,.$$

Assume further that $p_A(x) = x^n + c_{n-1}x^{n-1} + c_{n-2}x^{n-2} + ... + c_1 x + c_0$. Applying Lemma 4 using the matrix $xI_n - A$, we obtain

$$(xI_n - A)\, com\,(xI - A)^t = \det(xI_n - A)\, I_n,$$

where

$$xI - A = \begin{pmatrix} x - a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & x - a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots \\ a_{n1} & a_{n2} & \dots & x - a_{nn} \end{pmatrix}.$$

Hence

$$com\,(xI - A) = \begin{pmatrix} p_{n-1}^{(1,1)}(x) & p_{n-1}^{(1,2)}(x) & \dots & p_{n-1}^{(1,n)}(x) \\ p_{n-1}^{(2,1)}(x) & p_{n-1}^{(2,2)}(x) & \dots & p_{n-1}^{(2,n)}(x) \\ \vdots & \vdots & \vdots & \vdots \\ p_{n-1}^{(n,1)}(x) & p_{n-1}^{(n,2)}(x) & \dots & p_{n-1}^{(n,n)}(x) \end{pmatrix},$$

where $p_{n-1}^{(i,j)}$ are polynomials of degree $n-1$. Setting

$$com\,(xI - A)^t = B_0 + xB_1 + x^2 B_2 + ... + x^{n-1}B_{n-1}, \text{ where } (B_i)_{i=0,1,...,n-1} \in M_n(\mathbb{R}).$$

We deduce that

$$\begin{aligned} (xI - A)\left(B_0 + xB_1 + x^2B_2 + ... + x^{n-1}B_{n-1}\right) &= \det(xI_n - A).I_n \\ &= x^n I_n + c_{n-1}x^{n-1}I_n + ... + c_1 x I_n + c_0 I_n. \end{aligned}$$

It follows that

$$\begin{aligned} & x^n B_{n-1} + x^{n-1}(B_{n-2} - AB_{n-1}) + ... + x(B_0 - AB_1) - AB_0 \\ = \; & x^n I_n + c_{n-1}x^{n-1}I_n + ... + c_1 x I_n + c_0 I_n. \end{aligned}$$

Then

$$\begin{cases} B_{n-1} = I_n \\ B_{n-2} - AB_{n-1} = c_{n-1}x^{n-1}I_n \\ \qquad \vdots \\ B_0 - AB_1 = c_1 I_n \\ -AB_0 = c_0 I_n. \end{cases}$$

Which gives

$$\begin{aligned} p_A(A) &= c_0 I_n + c_1 A + ... + c_{n-1}A^{n-1} + A^n \\ &= -AB_0 + A(B_0 - AB_1) + ... + A^{n-1}(B_{n-2} - AB_{n-1}) + A^n B_{n-1} \\ &= 0. \end{aligned}$$

This completes the proof. ■

Example 5 *Let* $A = \begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix}$. *Find a polynomial* $p(x)$ *of degree* 2 *such that* $p(A) = 0$.
Ans. $p(x) = x^2 - 3x - 2$.

Corollary 6 *Let $A \in \mathcal{M}_n(\mathbb{R})$ with*

$$p_A(x) = x^n + c_{n-1}x^{n-1} + c_{n-2}x^{n-2} + ... + c_1 x + c_0,$$

where $c_0 \in \mathbb{R}^*$ and $c_1, c_2, ..., c_{n-1} \in \mathbb{R}$. Then

$$A^{-1} = \frac{-1}{c_0}\left(\sum_{i=1}^{n-1} c_i A^{i-1} + A^{n-1}\right).$$

Proof. Since

$$p_A(A) = c_0 I + c_1 A + c_2 A^2 + ... + c_{n-1}A^{n-1} + A^n = 0,$$

it follows that

$$\left(c_1 I + c_2 A + ... + c_{n-1}A^{n-2} + A^{n-1}\right) A = -c_0 I,$$

and so

$$A^{-1} = \frac{-1}{c_0}\left(c_1 I + c_2 A + ... + c_{n-1}A^{n-2} + A^{n-1}\right).$$

This completes the proof. ■

Example 7 *Using Cayley-Hamilton Theorem, calculate the inverse of the matrix*

$$A = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 0 & 0 \\ 2 & 0 & -1 \end{pmatrix}.$$

Solution. First, let us calculate $p_A(x)$:

$$\begin{aligned} p_A(x) &= \begin{vmatrix} \overset{+}{x-1} & \overset{-}{1} & \overset{+}{0} \\ -1 & x & 0 \\ 2 & 0 & x+1 \end{vmatrix} \\ &= (x-1)\left[x(x+1)\right] + (x+1) \\ &= (x-1)\left(x^2 - x + 1\right) \\ &= x^3 + 1. \end{aligned}$$

Therefore, $p_A(x) = x^3 + 1$, and hence

$$\begin{aligned} p_A(A) &= 0 \Rightarrow A^3 + I_3 = 0 \\ &\Rightarrow A^{-1} = -A^2. \end{aligned}$$

Finally, we get

$$A^{-1} = -\begin{pmatrix} 1 & 1 & 0 \\ -1 & 0 & 0 \\ 2 & 0 & -1 \end{pmatrix}\begin{pmatrix} 1 & 1 & 0 \\ -1 & 0 & 0 \\ 2 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & -2 & -1 \end{pmatrix}.$$