Theorem of Cayley and Hamilton

In the case of 2×2 matrices

Nobuyuki TOSE

October 18, 2016

Notation and a Remark

Notation $M_2(\mathbf{R})$

The set of 2×2 matrices with value in $\mathbf{R}$ is denoted by

$$M_2(\mathbf{R}) := \{A;\ 2 \times 2 \text{ matrix with value in} \mathbf{R}\}$$

Remark and Notation

We have already seen that

$$A, B \in M_2(\mathbf{R}) \ \Rightarrow AB \in M_2(\mathbf{R})$$

The nth power of $A \in M_2(\mathbf{R})$ can be defined by

$$A^n = \underbrace{A \cdots A}_{n \text{ copies}} \in M_2(\mathbf{R})$$

or recursively by

$$A^n = A^{n-1}A, \quad A^0 = I_2$$

Theorem of Cayley and Hamilton

Theorem

For $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, we have

$$A^2 - (a+d)A + |A| \cdot I_2 = O_2$$

(proof)

$$A^2 = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a^2 + bc & ab + bd \\ ca + dc & cb + d^2 \end{pmatrix} = \begin{pmatrix} a^2 + bc & b(a+d) \\ c(a+d) & cb + d^2 \end{pmatrix}$$

$$(a+d)A = \begin{pmatrix} (a+d)a & (a+d)b \\ (a+d)c & (a+d)d \end{pmatrix} = \begin{pmatrix} a^2 + ad & (a+d)b \\ (a+d)c & ad + d^2 \end{pmatrix}$$

$$A^2 - (a+d)A = \begin{pmatrix} bc - ad & 0 \\ 0 & bc - ad \end{pmatrix} = -|A| \cdot I_2$$

An Application

Let $A = \begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix}$. Then it follows from Theorem of C-H that

$$A^2 - 4A - 5I_2 = O_2$$

We associate to A the *eigen-polynomial* of A:

$$\Phi_A(\lambda) = \lambda^2 - 4\lambda - 5 = (\lambda + 1)(\lambda - 5)$$

Then the Vieta's Theorem leads us to

$$5 + (-1) = 4, \quad (-1) \cdot 5 = -5$$

Thus we get

$$A^2 - ((-1) + 5)A + (-1)5I_2 = O_2$$

An Application

We make it in the following two forms:

$$A^2 + A = 5A + 5I_2, \quad A^2 - 5A = -A + 5I_2$$

Accordingly we get

$$A(A + I_2) = 5(A + I_2), \quad A(A - 5I_2) = -(A - 5I_2)$$

We use these identities repeatedly to get

$$A^n(A + I_2) = 5^n(A + I_2), \quad A^n(A - 5I_2) = (-1)^n(A - 5I_2)$$

Namely

$$A^{n+1} + A^n = 5^n(A + I_2), \quad A^{n+1} - 5A^n = (-1)^n(A - 5I_2)$$

This leads us to

$$6A^n = 5^n(A + I_2) - (-1)^n(A - 5I_2)$$

Example

Let $A = \begin{pmatrix} 5 & 2 \\ 2 & 2 \end{pmatrix}$. Then by C-H we have

$$A^2 - 7A + 6I_2 = O_2$$

We associate the eigen-polynomial to A by

$$\Phi_A(\lambda) = \lambda^2 - 7\lambda + 6 = (\lambda - 1)(\lambda - 6).$$

The factorization above tells us that $1 + 6 = 7$ and $1 \cdot 6 = 6$. Thus the equation deduced from C-H is written by

$$A^2 - (1 + 6)A + 1 \cdot 6I_2 = O_2$$

which can be interpreted in two ways by

$$A(A - I_2) = 6(A - I_2), \quad A(A - 6I_2) = (A - 6I_2),$$

Example(2)

We try to find $A^n(A - I_2)$ and $A^n(A - 6I_2)$.

$$
\begin{aligned}
A^2(A - I_2) &= A \cdot A(A - I_2) = A \cdot 6(A - I_2) = 6 \cdot A(A - I_2) \\
&= 6 \cdot 6(A - I_2) = 6^2(A - I_2) \\
A^3(A - I_2) &= A \cdot A^2(A - I_2) = A \cdot 6^2(A - I_2) = 6^2 \cdot A(A - I_2) \\
&= 6^2 \cdot 6(A - I_2) = 6^3(A - I_2) \\
&\vdots \\
A^n(A - I_2) &= 6^n(A - I_2)
\end{aligned}
$$

We find that

$$A^n(A - I_2) = 6^n(A - I_2)$$

We can find more easily

$$A^n(A - 6I_2) = (A - 6I_2)$$

Example (3)

In the previous slides we have shown

$$A^n(A - I_2) = 6^n(A - I_2), \quad A^n(A - 6I_2) = (A - 6I_2)$$

which is views as

$$\begin{aligned} A^{n+1} - A^n &= 6^n(A - I_2) \quad (1) \\ A^{n+1} - 6A^n &= (A - 6I_2) \quad (2) \end{aligned}$$

Finally (1)−(2) is

$$5A^n = 6^n(A - I_2) - (A - 6I_2) \quad i.e. \quad A^n = \frac{6^n}{5}(A - I_2) - \frac{1}{5}(A - 6I_2)$$

Theorem

Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbf{R})$ be a 2×2 matrix and assume its eigen polynomial has two simple roots:

$$\begin{aligned} \Phi_A(\lambda) &= \lambda^2 - (a+d)\lambda + (ad - bc) \\ &= (\lambda - \alpha)(\lambda - \beta) \quad \text{with} \quad \alpha \neq \beta \end{aligned}$$

Then A^n is in the form

$$A^n = \alpha^n X_0 + \beta^n X_1$$

with $X_0, X_1 \in M_2(\mathbf{R})$ independent of n.

Polynomials of matrices

We can form polynomials in the matrix $A \in M_2(\mathbf{R})$.

Definition

For any polynomial

$$f(\lambda) = a_n\lambda^n + \cdots + a_1\lambda + a_0$$

we define

$$f(A) := a_0A^n + \cdots + a_1A + a_0I_2$$

Basic properties of this operation is given in the following theorem.

Theorem

For f, g polynomials in λ, we have

$$(f+g)(A) = f(A) + g(A), \quad (f \cdot g)(A) = f(A) \cdot g(A)$$

Eigenpolynomials

Definition and a remark

For $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, wedefine its *eigenpolynomial* by

$$\begin{aligned} \Phi_A(\lambda) = |\lambda I_2 - A| &= \begin{vmatrix} \lambda - a & -b \\ -c & \lambda - d \end{vmatrix} \\ &= \lambda^2 - (a+d)\lambda + ad - bc \\ &= \lambda^2 - (a+d)\lambda + |A| \end{aligned}$$

By Theorem of C-H we have

$$\Phi_A(A) = O_2$$

Application

Let $A = \begin{pmatrix} 1 & -1 \\ 1 & 3 \end{pmatrix}$. Then we have by Theorem of C-H

$$A^2 - 4A + 4I_2 = O_2 \quad i.e. \quad (A - 2I_2)^2 = O_2$$

We divide λ^n by $(\lambda - 2)^2$. Then there exist a polynomial $q(\lambda)$ and $a, b \in \mathbf{R}$ satisfying

$$\lambda^n = q(\lambda)(\lambda - 2)^2 + a\lambda + b \tag{1}$$

We get the derivative of this equation to get

$$n\lambda^{n-1} = q'(\lambda)(\lambda - 2)^2 + q(\lambda) \cdot 2(\lambda - 2) + a \tag{2}$$

Application

We substitute 2 for λ in (1) to get

$$2^n = 2a + b \tag{3}$$

and in (2) to get

$$n2^{n-1} = a \tag{4}$$

Thus we get

$$a = n2^{n-1}, \quad b = (1-n)2^n \tag{5}$$

Accordingly we have

$$\lambda^n = q(\lambda)(\lambda - 2)^2 + n2^{n-1}\lambda + (1-n)2^n$$

We substitute A for λ to derive

$$A^n = q(A)(A - 2I_2)^2 + n2^{n-1}A + (1-n)2^n I_2 = n2^{n-1}A + (1-n)2^n I_2$$