

こんにちは。加田修です。今回は Cayley-Hamilton の定理を証明します。この定理は重要な定理でいろいろな応用があります。

定理 (Cayley-Hamilton)

A を実数を成分とする n -次正方行列とし、(実は任意の可換環でもいいのですが) そして $P_A(t) := \det(tI - A)$ を A の特性多項式とします。すると t に A を代入すると零行列になります: すなわち、 $P_A(A) = O$ 。

ここで、多項式 $f(t) = \sum_{k=0}^m a_k t^k$ に対して、 $f(A) = \sum_{k=0}^m a_k A^k$ ($A^0 = I$) と定義します。例えば、 t に A を代入することによって、 $t^2 + 2$ は $A^2 + 2I$ に、 t は A に、 a は aI に、 0 は O 、零行列に変換されます。

$$A = (a_{ij}) = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \text{ に対して}$$

$$\begin{aligned} P_A(t) &= \left| \begin{pmatrix} t & & \\ & \ddots & \\ & & t \end{pmatrix} - \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \right| = \left| \begin{matrix} t - a_{11} & \cdots & -a_{1n} \\ \vdots & & \vdots \\ -a_{n1} & \cdots & t - a_{nn} \end{matrix} \right| \\ &= t^n - \operatorname{tr}(A)t^{n-1} + \cdots + (-1)^n \det A, \end{aligned}$$

$$\begin{aligned} \text{そして} \quad P_A(A) &= A^n - \operatorname{tr}(A)A^{n-1} + \cdots + (-1)^n \det(A)I \\ &= O. \end{aligned}$$

$n = 2$ のときを考えてみましょう。 $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ とすると、

$\det P_A(t) = \det(tI - A) = t^2 - (a+d)t + ad - bc$ であり、 t に A を代入して $\det P_A(A) = \det(AI - A) = \det O = 0$? いえ、これは間違っています。

t に A を代入すると、 $t, a, 0$ は A, aI, O 変換されることを思い出しましょう。従って、 $\begin{pmatrix} t & 0 \\ 0 & t \end{pmatrix}$

は $\begin{pmatrix} A & O \\ O & A \end{pmatrix}$ に、そして $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ は $\begin{pmatrix} aI & bI \\ cI & dI \end{pmatrix}$ に代わります。ここで変換された行列は 2×2 で成分が行列 A の多項式とみています。

例えば $\begin{pmatrix} I & O \\ O & I \end{pmatrix}$ の行列式は、 2×2 で成分が行列 A の多項式の行列の行列式とみると $I^2 - O^2 = I$ です。けれども 4×4 で成分が実数の行列とみるとその行列式は 1 になります。

同様に $\begin{pmatrix} A & O \\ O & A \end{pmatrix}$ の行列式は 2×2 で成分が A の多項式とみると $A^2 - O^2 = A^2$ で、 4×4 の行列で成分が実数の行列とみると $(\det A)^2$ です。

$$P_A(A) = A^2 - (a+d)A + (ad-bc)I \text{ で、共通の } A \text{ をくくりだすと } P_A(A) = \begin{pmatrix} -d & b \\ c & -a \end{pmatrix} A + (ad-bc)I.$$

ここで行列 $\begin{pmatrix} -d & b \\ c & -a \end{pmatrix}$ は A の余因子行列の -1 倍ですから、
 $P_A(A) = -\widetilde{A}A + (ad - bc)I = -\det(A)I + \det(A)I = O$ となります。

さて一般の n のときの Cayley-Hamilton の定理を証明します。

(定理の証明)

$B(t) := tI - A^T = \begin{pmatrix} t & & \\ & \ddots & \\ & & t \end{pmatrix} - \begin{pmatrix} a_{11} & \cdots & a_{n1} \\ \vdots & & \vdots \\ a_{1n} & \cdots & a_{nn} \end{pmatrix}$ とおきます。ここで A^T は A の転置行列です。

そうすると、 $B(A) = \begin{pmatrix} A & & \\ & \ddots & \\ & & A \end{pmatrix} - \begin{pmatrix} a_{11}I & \cdots & a_{n1}I \\ \vdots & & \vdots \\ a_{1n}I & \cdots & a_{nn}I \end{pmatrix} \in M_n(\mathbb{R}[A])$

になります。ここで $B(A)$ は $n \times n$ の成分が A の多項式の行列とみています。

$\mathbb{R}[A]$ は 非可換環 $M_n(\mathbb{R})$ の可換な部分環ですから、それを成分とする行列には、行列、行列式の定理が使えます。

$\widetilde{B(A)}$ を $M_n(\mathbb{R}[A])$ の中での $(M_{n^2}(\mathbb{R}))$ の中ではない $B(A)$ の余因子行列とします。

例えば $A \in M_2(\mathbb{R})$ に対して $C := \begin{pmatrix} A & O \\ O & O \end{pmatrix} \in M_2(\mathbb{R}[A])$ と置いたとき、 C の余因子行列は C を $M_2(\mathbb{R}[A])$ の要素と見たときは $\begin{pmatrix} O & O \\ O & A \end{pmatrix}$ ですが、 $M_4(\mathbb{R})$ の要素と見ると $\begin{pmatrix} O & O \\ O & O \end{pmatrix}$ となります。

そうすると

$\widetilde{B(A)}B(A) = \begin{pmatrix} \det B(A) & & \\ & \ddots & \\ & & \det B(A) \end{pmatrix} = \begin{pmatrix} \det P_A(A) & & \\ & \ddots & \\ & & \det P_A(A) \end{pmatrix}$ となります。

$B(A)$ の成分は行列ですから、右から $\begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}$ をかけることができます。ここで e_1, \dots, e_n は \mathbb{R}^n の

標準基底です。そして

$$B(A) \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = \begin{pmatrix} Ae_1 \\ \vdots \\ Ae_n \end{pmatrix} - \begin{pmatrix} a_{11}e_1 + \cdots + a_{n1}e_n \\ \vdots \\ a_{1n}e_1 + \cdots + a_{nn}e_n \end{pmatrix} = \begin{pmatrix} \vec{0} \\ \vdots \\ \vec{0} \end{pmatrix} \text{ となります。}$$

従って、 $\begin{pmatrix} \vec{0} \\ \vdots \\ \vec{0} \end{pmatrix} = \widetilde{B(A)}B(A) \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = \begin{pmatrix} P_A(A)e_1 \\ \vdots \\ P_A(A)e_n \end{pmatrix}$ が成り立ち、 $P_A(A) = O$ となります。

これで Cayley-Hamilton の定理が証明できたことになります。□

感じをつかむために、次の例を考えてみましょう。

$$A = \begin{pmatrix} & -1 \\ 0 & \\ 2 & \end{pmatrix} = (2e_3, 0, -e_1) \text{ とおきます。そうすると、}$$

$$B(t) := tI - A^T = \begin{pmatrix} t & & \\ & t & \\ & & t \end{pmatrix} - \begin{pmatrix} & 2 \\ & 0 \\ -1 & \end{pmatrix} = \begin{pmatrix} t & & -2 \\ & t & \\ 1 & & t \end{pmatrix},$$

$$B(A) = \begin{pmatrix} A & & \\ & A & \\ & & A \end{pmatrix} - \begin{pmatrix} & 2I \\ & 0 \\ -I & \end{pmatrix} = \begin{pmatrix} A & & -2I \\ & A & \\ I & & A \end{pmatrix},$$

$$\widetilde{B(A)} = \begin{pmatrix} A^2 & 0 & 2A \\ 0 & A^2 + 2I & 0 \\ -A & 0 & A^2 \end{pmatrix}, P_A(t) = \det B(t) = t^3 + 2t,$$

$$P_A(A) = \det \begin{pmatrix} A & & -2I \\ & A & \\ I & & A \end{pmatrix} = A^3 + 2A,$$

$$\widetilde{B(A)}B(A) = \begin{pmatrix} A^2 & 0 & 2A \\ 0 & A^2 + 2I & 0 \\ -A & 0 & A^2 \end{pmatrix} \begin{pmatrix} A & & -2I \\ & A & \\ I & & A \end{pmatrix} = \begin{pmatrix} A^3 + 2A & & \\ & A^3 + 2A & \\ & & A^3 + 2A \end{pmatrix}$$

$$\begin{aligned} B(A) \begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix} &= \begin{pmatrix} A & & \\ & A & \\ & & A \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix} - \begin{pmatrix} & 2I \\ & 0 \\ -I & \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix} \\ &= \begin{pmatrix} Ae_1 \\ Ae_2 \\ Ae_3 \end{pmatrix} - \begin{pmatrix} 2e_1 \\ \vec{0} \\ -e_1 \end{pmatrix} = \begin{pmatrix} 2e_1 \\ \vec{0} \\ -e_1 \end{pmatrix} - \begin{pmatrix} 2e_1 \\ \vec{0} \\ -e_1 \end{pmatrix} = \begin{pmatrix} \vec{0} \\ \vec{0} \\ \vec{0} \end{pmatrix}. \end{aligned}$$

$$\text{従って、} \begin{pmatrix} A^3 + 2A & & \\ & A^3 + 2A & \\ & & A^3 + 2A \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix} = \begin{pmatrix} (A^3 + 2A)e_1 \\ (A^3 + 2A)e_2 \\ (A^3 + 2A)e_3 \end{pmatrix} = \begin{pmatrix} \vec{0} \\ \vec{0} \\ \vec{0} \end{pmatrix} \text{ となり、}$$

$P_A(A) = A^3 + 2A = \mathbf{O}$ であることがわかります。

ですから、 $B(A) = tI - A^T$ というのはゼロではないわけですが、 $\begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix}$ を右からかける

とゼロになる、 $P_A(A)$ がゼロになることをいうためにはそれで十分なわけです。

$P_A(A)$ というのはこの場合 3×3 の行列なわけですが、 $B(A)$ というのは 3×3 の成分も 3×3 の行列で、結局 9×9 の実数を成分とする行列なわけで、 3×3 の行列がゼロ行列であること

をいうために、 9×9 の行列を考えたということになります。

又、 $B(A)$ はゼロでないけれども余因子行列をかけるとゼロになるなるわけです。