　みなさん、こんにちは。加田修です。前回はケーリー・ハミルトンの定理を成分が行列の多項式であるような行列を考えて証明しました。今回は、R-module, 環上の加群を考える方法で証明します。ここで環 R は多項式環 $\mathbb{R}[t]$ です。定理は次です。

定理 (Cayley-Hamilton)
A を 実数を成分とする n-次正方行列とし (実は実数でなくても任意の可換環の要素でいいのですが)、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, $P_A(A) = \mathrm{O}$.

　ここで、$P_A(t) = \sum_{i=0}^n a_i t^i,\ a_n = 1, a_{n-1} = -\mathrm{tr}(A), \ldots, a_0 = (1)^n \det A$ に対して, $P_A(t) = \sum_{k=0}^n a_k A^k$ $(A^0 = I)$ と定義します。つまり任意の正方行列は自分自身の特性方程式の零点になるということです。

それでは定理を証明します。

Proof. $B(t) := tI - A^T \in M_n(\mathbb{R}[t])$ とおきます。ここで A^T は A の転置行列です。そうするとその行列式は特性多項式になります。$\det B(t) = \det(tI - A^T) = \det(tI - A) = P_A(t)$.
$\widetilde{B(t)} \in M_n(\mathbb{R}[t])$ を $B(t)$ の余因子行列とします。そうすると

$$B(t)\widetilde{B(t)} = \widetilde{B(t)}B(t) = (\det B(t))I = P_A(t)I.$$

となります。

　$R := \mathbb{R}[t],\ V := \mathbb{R}^n$ とおき, n-次元ベクトル空間を次の掛け算によって R-加群とみます：すなわち $f(t) = \sum_{i=0}^n a_i t^i \in \mathbb{R}[t]$ と $x \in V$ に対して,

$$f(t)\cdot x := f(A)x = \sum_{i=0}^n a_i A^i x \in V \text{ と定義します。}$$

ここで R-加群の定義をざっとおさらいします。環 R と アーベル群 M に対して M が R-加群であるとは次が成り立つときをいいます：
$r \in R$ と $x \in M$ に対して, $r\cdot x \in M$ が定義されていて、 $r, s \in R, x, y \in M$ に対して次が成り立つ：

$$\begin{aligned}
&\text{(i) } (rs)\cdot x = r\cdot(s\cdot x),\\
&\text{(ii) } 1\cdot x = x\\
&\text{(iii) } r\cdot(x + y) = r\cdot x + r\cdot y,\\
&\text{(iv) } (r + s)\cdot x = r\cdot x + s\cdot x.
\end{aligned}$$

今の場合、R は多項式環 $\mathbb{R}[t]$ で, $M = V = \mathbb{R}^n$, そして R の M の作用は上で定義したものです。条件 (i) が満たされているのは次からわかります：

$$\begin{aligned}
\text{(i)} \qquad (a_l t^l b_m t^m)\cdot x &= (a_l b_m t^{l+m})\cdot x\\
&= a_l b_m A^{l+m} x,\\
(a_l t^l) \cdot (b_m t^m \cdot x) &= (a_l t^l)\cdot(b_m A^m x)\\
&= a_l A^l (b_m A^m x)\\
&= a_l b_m A^{l+m} x, \text{ and}\\
\text{(ii)} \qquad 1 \cdot x &= A^0 x = Ix = x.
\end{aligned}$$

$R' := M_n(R)$ とおき、$V^n = (\mathbb{R}^n)^n$ を行列かけるベクトルによって R'-加群とみます。

$B(t) = tI - A^T \in R' = M_n(\mathbb{R}[t])$ は多項式が成分の行列で、$\begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} \in V^n = (\mathbb{R}^n)^n$

（ここで $e_1,\dots,e_n$ は $V=\mathbb{R}^n$ の標準基底）ですから

$$
\begin{aligned}
B(t)\begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} &= (tI-A^T)\begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = \left(\begin{pmatrix} t & & \\ & \ddots & \\ & & t \end{pmatrix} - \begin{pmatrix} a_{11} & \dots & a_{n1} \\ \vdots & & \vdots \\ a_{1n} & \dots & a_{nn} \end{pmatrix}\right)\begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} \\
&= \begin{pmatrix} t{\cdot}e_1 \\ \vdots \\ t{\cdot}e_n \end{pmatrix} - \begin{pmatrix} a_{11}{\cdot}e_1+\dots+a_{n1}{\cdot}e_n \\ \vdots \\ a_{1n}{\cdot}e_1+\dots+a_{nn}{\cdot}e_n \end{pmatrix} = \begin{pmatrix} Ae_1 \\ \vdots \\ Ae_n \end{pmatrix} - \begin{pmatrix} \begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix} \\ \vdots \\ \begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix} \end{pmatrix} = \begin{pmatrix} \vec{0} \\ \vdots \\ \vec{0} \end{pmatrix}
\end{aligned}
$$

となります。ですから次がわかります：

$$
\begin{aligned}
\begin{pmatrix} \vec{0} \\ \vdots \\ \vec{0} \end{pmatrix} &= \widetilde{B(t)}B(t)\begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = \begin{pmatrix} P_A(t) & & \\ & \ddots & \\ & & P_A(t) \end{pmatrix}\begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} \\
&= \begin{pmatrix} P_A(t){\cdot}e_1 \\ \vdots \\ P_A(t){\cdot}e_n \end{pmatrix} = \begin{pmatrix} P_A(A)e_1 \\ \vdots \\ P_A(t)e_n \end{pmatrix},
\end{aligned}
$$

ですから $P_A(A)=\mathrm{O}$. □

次の例を考えてみましょう。$A=\begin{pmatrix} & & -1 \\ & 0 & \\ 2 & & \end{pmatrix}=(2e_3,\ \vec{0},\ -e_1)$. すると次のようになります。

$$
\begin{aligned}
B(t) &:= tI-A^T = \begin{pmatrix} t & & \\ & t & \\ & & t \end{pmatrix} - \begin{pmatrix} & & 2 \\ & 0 & \\ -1 & & \end{pmatrix} = \begin{pmatrix} t & & -2 \\ & t & \\ 1 & & t \end{pmatrix}, \\
\widetilde{B(t)} &= \begin{pmatrix} t^2 & & 2t \\ & t^2+2I & \\ -t & & t^2 \end{pmatrix}, P_A(t)=\det B(t)=t^3+2t, \\
P_A(A) &= A^3+2A, \\
\widetilde{B(t)}B(t) &= \begin{pmatrix} t^2 & & 2t \\ & t^2+2 & \\ -t & & t^2 \end{pmatrix}\begin{pmatrix} t & & -2 \\ & t & \\ 1 & & t \end{pmatrix} = \begin{pmatrix} t^3+2t & & \\ & t^3+2t & \\ & & t^3+2t \end{pmatrix}, \\
B(t)\begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix} &= \begin{pmatrix} t & & \\ & t & \\ & & t \end{pmatrix}\begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix} - \begin{pmatrix} & & 2 \\ & 0 & \\ -1 & & \end{pmatrix}\begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix} \\
&= \begin{pmatrix} t{\cdot}e_1 \\ t{\cdot}e_2 \\ t{\cdot}e_3 \end{pmatrix} - \begin{pmatrix} 2e_3 \\ 0 \\ -e_1 \end{pmatrix} = \begin{pmatrix} Ae_1 \\ Ae_2 \\ Ae_3 \end{pmatrix} - \begin{pmatrix} 2e_3 \\ 0 \\ -e_1 \end{pmatrix} = \begin{pmatrix} \vec{0} \\ \vec{0} \\ \vec{0} \end{pmatrix}.
\end{aligned}
$$

従って

$$
\begin{aligned}
\begin{pmatrix} \vec{0} \\ \vec{0} \\ \vec{0} \end{pmatrix} &= \widetilde{B(t)} B(t) \begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix} = \begin{pmatrix} t^3+2t & & \\ & t^3+2t & \\ & & t^3+2t \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix} \\
&= \begin{pmatrix} (t^3+2t)\cdot e_1 \\ (t^3+2t)\cdot e_2 \\ (t^3+2t)\cdot e_3 \end{pmatrix} = \begin{pmatrix} (A^3+2A)e_1 \\ (A^3+2A)e_2 \\ (A^3+2A)e_3 \end{pmatrix},
\end{aligned}
$$

よって $A^3 + 2A = P_A(A) = \mathrm{O}$.