ジョルダン標準形

木原浩貴

2009 年 12 月 31 日

概要

単因子論は多くの重要な応用をもっているが、その一つにジョルダン標準形の理論がある。すなわち K が代数的閉体 (例えば $\mathbb{C}$) の場合に、K の元を成分にもつ任意の n 次正方行列はジョルダン行列と呼ばれる特別な型の行列に相似であることが知られている。本書ではこの定理の単因子を用いた証明を解説する。

1　行列の単因子

■予備知識　本書は線形代数学の基礎概念を理解し、さらに「群」「群の作用」「環」「イデアル」や「体」の少なくとも定義だけは知っている読者を対象にしている (「加群」については本書では簡単に説明してある)。

■行列の相似　可換環 R の元を成分にもつ n 次正方行列を、本書では簡単に **n 次 R-行列**と呼ぶ。n 次 R-行列全体のなす環 (これは $n \geq 2$ ならば一般に非可換) を $M_n(R)$ と書く。そして $M_n(R)$ における可逆元 (単元) 全体のつくる群を $GL_n(R)$ と書く。

K を体とする。$A, B \in M_n(K)$ は $Q \in GL_n(K)$ が存在して $B = Q^{-1}AQ$ であるとき、$A \sim B$ と書いて、A と B は互いに**相似**であるという。相似は明らかに同値関係である。

本書の目標は K が代数的閉体 (例えば $K = \mathbb{C}$) である場合に、任意の $A \in M_n(K)$ がジョルダン標準形と呼ばれる特別な形をもった行列に相似なことを示すことである。その過程で、Hamilton-Cayley の定理 (A の特性多項式 $\chi_A(t)$ に対して $\chi_A(A) = 0$ が成り立つこと) も証明されるであろう。

■基本変形　このために本書では単因子の手法を用いる。単因子の手法では K-行列 A を直接あつかう代わりに A の特性行列 $tE - A$ (これは $K[t]$-行列) を扱う。

そこでまず $K[t]$-行列全体を対象とする「基本変形」について述べる。

以下の 3 つの形の行列を $M_n(K[t])$ における**基本行列**と呼ぶ：

$$
P(i,j)=\begin{array}{c}
\begin{matrix} & {\scriptstyle (i)} & & {\scriptstyle (j)} & \end{matrix}\\
\begin{bmatrix}
E_1 & \vdots & & \vdots & \\
\cdots & 0 & \cdots & 1 & \cdots \\
 & \vdots & E_2 & \vdots & \\
\cdots & 1 & \cdots & 0 & \cdots \\
 & \vdots & & \vdots & E_3
\end{bmatrix}
\end{array}
\quad (i<j), \qquad
Q(i\,;u)=\begin{array}{c}
\begin{matrix} & {\scriptstyle (i)} & \end{matrix}\\
\begin{bmatrix}
E_4 & \vdots & \\
\cdots & u & \cdots \\
 & \vdots & E_5
\end{bmatrix}
\end{array}
\quad (u\in K^\times)
$$

$$
R(i,j\,;c(t))=\begin{array}{c}
\begin{matrix} & {\scriptstyle (i)} & & {\scriptstyle (j)} & \end{matrix}\\
\begin{bmatrix}
E_6 & \vdots & & \vdots & \\
\cdots & 1 & \cdots & c(t) & \cdots \\
 & \vdots & E_7 & \vdots & \\
\cdots & 0 & \cdots & 1 & \cdots \\
 & \vdots & & \vdots & E_8
\end{bmatrix}
\end{array}
\quad (i\neq j).
$$

ここで $E_1 \ldots E_8$ は適当なサイズの単位行列である。数や行列の指定されていない場所の成分はすべて 0 とする。

$P(i,j)$ は単位行列の第 i 行と 第 j 行を置換したもの (第 i 列と 第 j 列を置換したものと言っても同じ) を、$Q(i\,;u)$ は単位行列の (i,i) 成分を単元 u に置き換えたものを意味する。$R(i,j\,;c(t))$ は単位行列の第 (i,j) 成分に $c(t)$ を加えたものを意味する (図は $i<j$ の場合を示しており、それゆえ $c(t)$ は対角線より上に来る)。

基本行列はみな可逆で、逆行列も基本行列である：

$$
P(i,j)\cdot P(i,j)=E,\quad Q(i\,;u)\cdot Q(i\,;u^{-1})=E,\quad R(i,j\,;c(t))\cdot R(i,j\,;-c(t))=E.
$$

基本行列の積全体の集合は $GL_n(K[t])$ の部分群を成す。これをしばらく $\mathcal{P}$ と書く。(実は後で分かるように $\mathcal{P}=GL_n(K[t])$ である)。

$K[t]$-行列 $A(t)$ に対し、基本行列を左から掛けることは $A(t)$ の行を扱う操作であり、基本行列を右から掛けることは $A(t)$ の列を扱う操作である。とくに

- $P(i,j)$ を左 [右] から掛ける操作は $A(t)$ の第 i 行 [列] と 第 j 行 [列] を交換する操作に等しい。
- $Q(i;u)$ を左 [右] から掛ける操作は $A(t)$ の第 i 行 [列] を u 倍する操作に等しい。
- $R(i,j;c(t))$ を左 [右] から掛ける操作は $A(t)$ の第 i 行 [第 j 列] に $A(t)$ の第 j 行 [第 i 列] の $c(t)$ 倍を加える操作に等しい。

$K[t]$-行列に対し、基本行列の積を左 [右] から掛ける操作を**左 [右] 基本変形**という。左からも右からも掛ける操作を**基本変形**という。基本変形は $M_n(K[t])$ に対する直積群 $\mathcal{P}\times\mathcal{P}$ の左作用である：$(P(t),Q(t))\,A(t)=P(t)A(t)Q(t)^{-1}$。

もし $A(t)$ に基本変形を施して $B(t)$ が得られたならば、つまり $B(t)$ が $A(t)$ の軌道上にあるならば $A(t)$ と $B(t)$ は互いに**対等**であるといい、本書では記号 $A(t)\leftrightarrow B(t)$ を用いて表す。この関係はもちろん同値関係である。

■単因子行列　ここでの目標は任意の $K[t]$-行列 $A(t)$ が次のかたちの $K[t]$-行列 $E(t)$ に対等なことを示すことである：

$$
(1)\qquad E(t)=\begin{bmatrix} e_1(t) & & \\ & \ddots & \\ & & e_n(t) \end{bmatrix}
$$

ここで $e_1(t),\ldots,e_n(t)$ は $e_i(t)\,|\,e_{i+1}(t)$ $(1\le i\le n-1)$ の関係を満たす多項式で、各 $e_i(t)$ はモニック多項式であるとする。

このような行列 $E(t)$ を本書では**単因子行列**と呼ぶ。

補題 1.1. 任意の行列 $A(t)$ は次の性質をもつ行列 $B(t)$ に対等である。すなわち $B(t)$ のすべての成分は成分 $b_{11}(t)$ の倍元である：

(2) $\quad b_{11}(t)\,|\,b_{ij}(t) \quad (\forall i,j)$

証明 $A(t)=0$ ならば自明 ($B(t)=A(t)$ でよい)。
$A(t)\ne 0$ の場合、必要なら行の置換、列の置換 (基本変形の一種) を行うことで、$(1,1)$ 成分に非 0 元をもってくることができるから、最初から $A(t)$ は $a_{11}(t)\ne 0$ を満たすと仮定しても一般性を失わない。
証明は $a_{11}(t)$ の次数 $p(\ge 0)$ に関する帰納法による。まず $p=0$ の場合は、$a_{11}(t)$ は単元だから $B(t)=A(t)$ でよい。
$p-1$ まで定理の主張が成立していると仮定する。$A(t)$ のすべての成分が $a_{11}(t)$ の倍元ならば $B(t)=A(t)$ でよい。$A(t)$ の成分のなかに $a_{11}(t)$ の倍元とならない元がある場合は次の三通りに分けられる：

1. $a_{11}(t)$ の倍元とならない元が $A(t)$ の 第 1 行 にある場合。
2. $a_{11}(t)$ の倍元とならない元が $A(t)$ の 第 1 列 にある場合
3. $A(t)$ の 第 1 行 と 第 1 列 の成分はすべて $a_{11}(t)$ の倍元で、$a_{11}(t)$ の倍元とならない元が $A(t)$ の 2 行目以降かつ 2 列目以降に現れる場合。

第一の場合、$a_{1j}(t)$ ($2\le j$) を $a_{11}(t)$ の倍元でない成分としよう。

$$a_{1j}(t)=a_{11}(t)q(t)+r(t),\quad \deg r(t)<\deg a_{11}(t)$$

とおいて、次の基本変形を行う：

1. 第 1 列 の $-q(t)$ 倍を 第 j 列 に加える。
2. 第 1 列 と 第 j 列を入れ替える。

$$\begin{bmatrix} a_{11}(t) & \ldots & a_{1j}(t) & \ldots \\ * & & * & \\ \vdots & & \vdots & \\ * & & * & \end{bmatrix} \leftrightarrow \begin{bmatrix} a_{11}(t) & \ldots & r(t) & \ldots \\ * & & * & \\ \vdots & & \vdots & \\ * & & * & \end{bmatrix}$$
$$\leftrightarrow \begin{bmatrix} r(t) & \ldots & a_{11}(t) & \ldots \\ * & & * & \\ \vdots & & \vdots & \\ * & & * & \end{bmatrix}$$

すると $(1,1)$ 成分に $r(t)\ne 0$ が現れる。この行列を $A'(t)$ と書くと、$\deg r(t)<p$ だから帰納法の仮定により $A'(t)$ は条件 (2) を満たす $B(t)$ に対等である。したがって $A(t)$ も $B(t)$ に対等である。
第二の場合も同様である。
第三の場合、$a_{11}(t)$ の倍元でない成分を $a_{ij}(t)$ $(2\le i, 2\le j)$ とする。$a_{i1}(t)=a_{11}(t)\cdot d(t)$ とおいて、次の基本変形を行う：

1. 第 1 行 の $-d(t)$ 倍を 第 i 行 に加える。
2. 第 i 行 を 第 1 行 に加える。

$$
\begin{bmatrix} a_{11}(t) & \dots & a_{1j}(t) & \dots \\ \vdots & & \vdots & \\ a_{i1}(t) & \dots & a_{ij}(t) & \dots \\ \vdots & & \vdots & \end{bmatrix} \leftrightarrow \begin{bmatrix} a_{11}(t) & \dots & a_{1j}(t) & \dots \\ \vdots & & \vdots & \\ 0 & \dots & a_{ij}(t) - d(t)a_{1j}(t) & \dots \\ \vdots & & \vdots & \end{bmatrix}
$$

$$
\leftrightarrow \begin{bmatrix} a_{11}(t) & \dots & (1 - d(t))a_{1j}(t) + a_{ij}(t) & \dots \\ \vdots & & \vdots & \\ 0 & \dots & a_{ij}(t) - d(t)a_{1j}(t) & \dots \\ \vdots & & \vdots & \end{bmatrix}
$$

この結果、$(1, j)$ 成分に $(1 - d(t))a_{1j}(t) + a_{ij}(t)$ が来る。これは $a_{11}(t)$ の倍元でない ($a_{1j}(t)$ が $a_{11}(t)$ の倍元で、$a_{ij}(t)$ が $a_{11}(t)$ の倍元でないから)。ゆえにこの行列を $A'(t)$ と書くと、$A'(t)$ は第一の場合に当てはまり、既に示したように条件 (2) を満たす行列 $B(t)$ に対等である。したがって $A(t)$ も $B(t)$ に対等である。 □

補題 1.2. $A(t) = (a_{ij}(t)) \leftrightarrow A'(t) = (a'_{ij}(t))$ とする。このとき $d(t) \mid a_{ij}(t)$ ($\forall i, j$) ならば $d(t) \mid a'_{ij}(t)$ ($\forall i, j$) である。

証明 これは基本変形の式を見れば明らかだから、詳細は述べない (後述の補題 1.5 を参照)。 □

定理 1.3. 任意の行列 $A(t)$ は単因子行列に対等である。

証明 $A(t)$ の次数 n に関する帰納法による。
まず $n = 1$ の場合は自明である。
$n - 1$ までは定理は成立と仮定する。$A(t) = 0$ ならば $A(t)$ がすでに単因子行列だから証を要さない。そこで以下、$A(t) \neq 0$ と仮定するが、補題 1.1 によると $A(t)$ について次の条件が成り立っていると仮定してよい：$A(t)$ のすべての成分は $a_{11}(t)(\neq 0)$ の倍元である。
ここで次の基本変形を行う。

1. 第 1 行 に適当な単元を掛けて、$a_{11}(t)$ をモニック多項式 $e_1(t)$ に変える。
2. 第 2 行 以降の各行に 第 1 行 の適当な倍元を加えて、$(1, 1)$ 成分を除く 第 1 列 のすべての成分を 0 にする。
3. 同様にして $(1, 1)$ 成分を除く 第 1 行 のすべての成分を 0 にする。

このようにして出来上がる行列を $A'(t) = (a'_{ij}(t))$ とおく。

$$
\begin{bmatrix} a_{11}(t) & a_{12}(t) & \dots & a_{1n}(t) \\ a_{21}(t) & a_{22}(t) & \dots & a_{2n}(t) \\ \vdots & \vdots & & \vdots \\ a_{n1}(t) & a_{n2}(t) & \dots & a_{nn}(t) \end{bmatrix} \leftrightarrow \begin{bmatrix} e_1(t) & 0 & \dots & 0 \\ 0 & a'_{22}(t) & \dots & a'_{2n}(t) \\ \vdots & \vdots & & \vdots \\ 0 & a'_{n2}(t) & \dots & a'_{nn}(t) \end{bmatrix}
$$

さて、帰納法の仮定から、行列 $(a'_{ij})_{2\leq i,2\leq j}$ は単因子行列 $(e_2(t), \dots, e_n(t))$ に対等である。これは $A'(t)$ の 2 行目以降と 2 列目以降に基本変形を施すことで $A'(t)$ を次のかたちに変形できることを意味する。

$$
\begin{bmatrix} e_1(t) & 0 & \dots & 0 \\ 0 & e_2(t) & & \\ \vdots & & \ddots & \\ 0 & & & e_n(t) \end{bmatrix}
$$

ここで補題 1.2 により $e_1(t) \mid e_i(t)$ $(2 \leq i)$ が成り立っていることに注意すると、この行列は単因子行列であると分かる。

したがって $A(t)$ は単因子行列に対等である。 □

定理 1.4. 1. 可逆な単因子行列は単位行列である。

2. 可逆行列は単位行列に対等である。
3. 可逆行列は基本行列の積のかたちに書ける。
4. 可逆行列は左基本変形だけでも右基本変形だけでも単位行列に導かれる。

証明 1. $X(t) = (x_{ij})$ を単因子行列 $E(t)$ の逆行列とすると、

$$\begin{bmatrix} e_1(t) & & \\ & \ddots & \\ & & e_n(t) \end{bmatrix} \begin{bmatrix} & & \cdot \\ & & \vdots \\ \dots & \dots & x_{nn}(t) \end{bmatrix} = \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix}$$

すなわち $e_n(t)x_{nn}(t) = 1$ 。ここから明らかに $e_n(t) = 1$。ゆえに $e_1(t) = \cdots = e_n(t) = 1$ となって $E(t)$ は単位行列。

2. $A(t)$ を可逆行列、$P(t), Q(t) \in \mathcal{P}$、$E(t)$ を単因子行列、$P(t)A(t)Q(t) = E(t)$ とする。すると $E(t)$ は可逆だから、1. により $E(t) = E$ である。故に

$$P(t)A(t)Q(t) = E. \tag{3}$$

$A(t)$ は単位行列に対等である。

3. (3) から $A(t) = P(t)^{-1}Q(t)^{-1}$ で、右辺は基本行列の積である。

4. (3) からまた $P(t)A(t) = Q(t)^{-1}$。したがって $Q(t)P(t)A(t) = E$ で、これは $A(t)$ が左基本変形だけで E に導かれることを示している。同様にして $A(t)$ は右基本変形だけでも E に導かれる。 □

■**行列式因子と単因子** 次に与えられた行列 $A(t)$ に対等な単因子行列 $E(t)$ の一意性を証明しよう。

このために行列式因子 $\{d_k(t)\}$ の概念を導入する。

$A(t)$ のすべての k 次小行列式の最大公約元を $d_k(t)$ と書く $(1 \le k \le n)$。$d_1(t), \dots, d_n(t)$ を $A(t)$ の**行列式因子**と呼ぶ。

補題 1.5. $A(t)$ に基本変形を施しても行列式因子は不変である。

証明 $A(t)$ に $P(i, j), Q(i; u), R(i, j; c(t))$ のいずれかを左から施したものを $A'(t)$ と書く。
基本行列を掛ける過程をくわしく見れば、$A'(t)$ の k 次小行列式 $\Delta'(t)$ はすべて

$$\Delta'(t) = u\Delta_1(t) + c(t)\Delta_2(t)$$

のかたちに書けることが分かる。ここで $\Delta_1(t), \Delta_2(t)$ は $A(t)$ の k 次小行列式、u は K の単元で $c(t) \in K[t]$ である。ゆえに $d_k(t) \,|\, \Delta'(t)$ で、これが任意の $\Delta'(t)$ に対して成り立つから、$d_k(t) \,|\, d'_k(t)$ である。したがって、基本変形の可逆性からまた $d'_k(t) \,|\, d_k(t)$ も導かれ、ゆえに $d_k(t) = d'_k(t)$ となる。 □

補題 1.6. 単因子行列 $E(t)$

の行列式因子を $\{\ d_k(t)\ \}$ と書くと、

$$d_1(t) = e_1(t), \quad d_2(t) = e_1(t)e_2(t), \quad \dots \quad , d_n(t) = e_1(t)e_2(t)\cdots e_n(t) \tag{4}$$

が成り立つ。

証明 これは (1) から明らかであろう。 □

ここから分かるように二つの単因子行列 $E_1(t)$ と $E_2(t)$ が対等ならばじつは $E_1(t) = E_2(t)$ が成り立つ。したがって

定理 1.7. 行列 $A(t)$ に対し、それに対等な単因子行列 $E(t)$ がただ一通りに定まる。$E(t)$ の対角成分 $\{ e_k(t) \}_k$ を $A(t)$ の**単因子 (elementary divisors)** と呼ぶ。これらと $A(t)$ の行列式因子 $\{ d_k(t) \}_k$ とは式 (4) で結ばれている。

よって $A(t)$ の単因子を求めるのに次の二つの方法がある。

1. $A(t)$ に基本変形を施して単因子行列を導き、その対角成分から求める方法。
2. $A(t)$ の行列式因子を計算して、そこから式 (4) にしたがって求める方法。

■**特性行列**　$A \in M_n(K)$ に対して $tE - A \in M_n(K[t])$ は A の**特性行列**と呼ばれる。

$A, B \in M_n(K)$ とする。$A \sim B$ ならば $tE - A \leftrightarrow tE - B$ であることは見やすい。実際 $Q \in GL_n(K)$、$B = Q^{-1}AQ$ であるとすると、$tE - B = Q^{-1}(tE - A)Q$。

逆に $tE - A \leftrightarrow tE - B$ から $A \sim B$ が導かれることを次節で示す。

2　$K[t]$-加群

■$K[t]$-加群　線型代数においてベクトル空間の元にかかる係数は、何らかの体に属する数であったが、これを一般の環にまで拡張したものが加群の理論である。ただし以下では可換環上の加群についてだけ説明する。

R を可換環とする。**R-加群**とは加法群 V に R が作用していて、その作用を $(a, x) \in R \times V \mapsto ax \in V$ と書くとき、$a(x + y) = ax + ay$, $(a + b)x = ax + bx$, $(ab)x = a(bx)$, $1x = x$ を恒等的に満たすものをいう。

前節では K-行列の世界を拡張して $K[t]$-行列の世界を考えたが、ここでは K-ベクトル空間の世界を拡張して $K[t]$-加群の世界を考えよう。

まず K^n を拡張して $(K[t])^n$ を考えるのが自然である。n 次の K-縦ベクトルを K^n の元と、n 次の K-行列を K^n の線形変換とそれぞれ同一視できたように、n 次の $K[t]$-縦ベクトルは $(K[t])^n$ の元と、n 次 の $K[t]$-行列は $K[t]^n$ の線形変換 (正しく言うと $K[t]$-自己準同型) とそれぞれ同一視できる。すなわ n 次 $K[t]$-行列 $A(t) = (a_{ij}(t))$ に対し、

$$\begin{bmatrix} y_1(t) \\ \vdots \\ y_n(t) \end{bmatrix} = \begin{bmatrix} a_{11}(t) & \dots & a_{1n}(t) \\ \vdots & & \vdots \\ a_{n1}(t) & \dots & a_{nn}(t) \end{bmatrix} \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix}$$

で定義される写像 $(x_i(t)) \in (K[t])^n \mapsto (y_i(t)) \in (K[t])^n$ は $K[t]$-自己準同型である。そして逆に $(K[t])^n$ の任意の $K[t]$-自己準同型はこのように一つのしかも唯一つの行列 $A(t)$ を用いて定義される。

次に $A \in M_n(K)$ とし、これを K^n の K-線形変換と同一視する。K^n と A の組 (K^n, A) を次の意味で $K[t]$-加群と見なすことができる。すなわち $K[t]$ の K^n への作用を

$$f(t) \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = f(A) \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

と定義すれば K^n は $K[t]$-加群になることが確かめられる。ここで $f(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_k t^k$ に対し

$$f(A) = a_0 + a_1 A + a_2 A^2 + \cdots + a_k A^k$$

である。この $K[t]$-加群の構造は K^n にある K-ベクトル空間の構造と両立している。すなわち $a \in K$ と $\mathbf{x} \in K^n$ に対し、$a\mathbf{x}$ は a を $K[t]$ の元と思っても同じ結果を与えている。

定理 2.1. $A, B \in M_n(K)$ とする。$A \sim B$ であることが (K^n, A) と (K^n, B) が $K[t]$-加群として同型であるための必要十分条件である。詳しく言うと $Q \in GL_n(K)$ に対し $B = Q^{-1}AQ$ であることと、Q が (K^n, B) から (K^n, A) への $K[t]$-同型であることとは同値である。

証明 $B = Q^{-1}AQ$ とすると、任意の $\mathbf{x} \in (K^n, B)$ に対し

$$Q(t\mathbf{x}) = Q(B\mathbf{x}) = A(Q\mathbf{x}) = t(Q\mathbf{x})$$

が成り立つ。ゆえに Q は $K[t]$-同型である。逆に Q が $K[t]$-同型ならば $QB = AQ$ であることも上の式から読み取れる。 □

したがって K-行列の相似類とは言い換えれば、K^n に入る $K[t]$-加群構造の同型類である。ただしここでいう $K[t]$-加群構造は K^n のもつ K-ベクトル空間の構造と両立するものに限る。

■ **$K[t]$-加群の短完全列** $tE - A \leftrightarrow tE - B$ ならば $A \sim B$ が成り立つことを示そう。

まず次の図式を考える：

$$0 \longrightarrow (K[t])^n \xrightarrow{\varphi_A} (K[t])^n \xrightarrow{\psi_A} (K^n, A) \longrightarrow 0 \tag{5}$$

ここで $\varphi_A : (K[t])^n \to (K[t])^n$ は特性行列 $tE - A$ の定義する $K[t]$-準同型である：

$$\varphi_A(\mathbf{p}(t)) = (tE - A)\mathbf{p}(t).$$

$\psi_A : (K[t])^n \to (K^n, A)$ は

$$\psi_A(\mathbf{q}(t)) = q_1(t)\mathbf{e}_1 + \cdots + q_n(t)\mathbf{e}_n = q_1(A)\mathbf{e}_1 + \cdots + q_n(A)\mathbf{e}_n$$

で定義される $K[t]$-準同型である。

定理 2.2. 上で定義した φ_A, ψ_A について次のことが成り立つ：

1. $\operatorname{Ker}\varphi_A = \{0\}$。つまり φ_A は単射である。
2. $\operatorname{Img}\varphi_A \subset \operatorname{Ker}\psi_A$ である。
3. ψ_A は $\psi_A(\mathbf{e}_i) = \mathbf{e}_i$ $(i = 1 \ldots n)$ を満たす $K[t]$-準同型として特徴づけられる。
4. $(K[t])^n = \operatorname{Img}\varphi_A \oplus K^n$。ただしこれは K-ベクトル空間としての直和である (K^n は $(K[t])^n$ の中では $K[t]$-部分加群とはならないことに注意)。
5. $\operatorname{Img}\varphi_A = \operatorname{Ker}\psi_A$。
6. $\operatorname{Img}\psi_A = (K^n, A)$。つまり ψ_A は全射である。

証明 1.) $\mathbf{p}(t) = (p_i(t)) \neq 0$ とする。$p_1(t), \ldots, p_n(t)$ のなかで次数最大のものを $p_i(t)$ とし、$d = \deg p_i(t)$ とおくと $d \geq 0$。$\mathbf{q}(t) = (q_i(t)) = \varphi_A(\mathbf{p}(t))$ とおくと、φ_A の定義より

$$q_i(t) = tp_i(t) - \sum_{k=1}^{n} a_{ik} p_k(t).$$

ここで第一項の次数は $d+1$ で、第二項以降の和の次数は d 以下だから、$\deg q_i(t) = d+1$ でなければならない。ゆえに $\mathbf{q}(t) \neq 0$ となり、$\mathrm{Ker}\,\varphi_A = \{0\}$ が証明された。
2.) つまり $\psi_A(\varphi_A(\mathbf{p}(t))) = 0$。これは単純な計算で出てくるから読者に任せる。
3.) $\mathbf{a} = (a_i) \in K^n$ とすると $\psi_A(\mathbf{a}) = a_1\mathbf{e}_1 + \cdots + a_n\mathbf{e}_n = \mathbf{a}$。逆に ψ_A が $\psi_A(\mathbf{e}_i) = \mathbf{e}_i$ $(i = 1\ldots n)$ を満たす $K[t]$-準同型だとすると、任意の $\mathbf{q}(t) = (q_i(t)) \in (K[t])^n$ に対し

$$\psi_A(\mathbf{q}(t)) = \psi_A\left(\sum_{i=1}^{n} q_i(t)\mathbf{e}_i\right) = \sum_{i=1}^{n} q_i(t)\psi_A(\mathbf{e}_i) = \sum_{i=1}^{n} q_i(t)\mathbf{e}_i.$$

4.) 任意の $\mathbf{q}(t) = (q_i(t)) \in (K[t])^n$ に対し

$$\begin{bmatrix} q_1(t) \\ \vdots \\ q_n(t) \end{bmatrix} = \begin{bmatrix} q_1(t) & & \\ & \ddots & \\ & & q_n(t) \end{bmatrix} \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}.$$

ところが右辺の行列 $Q(t) = (q_1(t), \ldots, q_n(t))$ は $Q(t) = (tE - A)Q_1(t) + Q$ ($Q \in M_n(K)$) のかたちに表すことができるから

$$\begin{bmatrix} q_1(t) \\ \vdots \\ q_n(t) \end{bmatrix} = (tE - A)\begin{bmatrix} p_1(t) \\ \vdots \\ p_n(t) \end{bmatrix} + \begin{bmatrix} r_1 \\ \vdots \\ r_n \end{bmatrix}.$$

故に $\mathbf{q}(t) \in \mathrm{Img}\,\varphi_A + K^n$。
次に $\mathbf{a} \in \mathrm{Img}\,\varphi_A \cap K^n$ とすると、$\mathrm{Img}\,\varphi_A \subset \mathrm{Ker}\,\psi_A$ だから $\psi_A(\mathbf{a}) = \mathbf{a} = 0$。したがって $\mathrm{Img}\,\varphi_A + K^n = \mathrm{Img}\,\varphi_A \oplus K^n$。
5.) $\mathrm{Ker}\,\psi_A \subset \mathrm{Img}\,\varphi_A$ を示せばよい。4. にしたがい $\mathbf{q}(t) \in \mathrm{Ker}\,\psi_A$ を $\mathbf{q}(t) = \mathbf{q}_1(t) + \mathbf{a}$, $\mathbf{q}_1(t) \in \mathrm{Img}\,\varphi_A$, $\mathbf{a} \in K^n$ と分解すると、2. と 3. により

$$0 = \psi_A(\mathbf{q}_1(t) + \mathbf{a}) = 0 + \mathbf{a}.$$

ゆえに $\mathbf{a} = 0$, $\mathbf{q}(t) = \mathbf{q}_1(t) \in \mathrm{Img}\,\varphi_A$ となる。
6.) これは 3. から明らか。

□

上の 1, 5, 6 を満たすとき、図式 (5) は $K[t]$-加群の**短完全列 (short exact sequence)** であるという言い方をする。

注意. ただし以下では 1. で述べたこと、つまり図式の $0 \longrightarrow$ の部分は使用されない。

さて $A, B \in M_n(K)$ とし、$tE - A$ と $tE - B$ は対等であるとする。すなわち可逆な行列 $P(t), Q(t) \in GL_n(K[t])$ が存在し、$tE - B = Q(t)^{-1}(tE - A)P(t)$ を満たす、言い換えると、下の図式の左半分を可換にするものとする。

$$\begin{array}{ccccccccc} 0 & \longrightarrow & (K[t])^n & \xrightarrow{\varphi_B} & (K[t])^n & \xrightarrow{\psi_B} & (K^n, B) & \longrightarrow & 0 \\ & & {\scriptstyle P(t)}\downarrow & & {\scriptstyle Q(t)}\downarrow & & \downarrow{\scriptstyle Q} & & \\ 0 & \longrightarrow & (K[t])^n & \xrightarrow{\varphi_A} & (K[t])^n & \xrightarrow{\psi_A} & (K^n, A) & \longrightarrow & 0 \end{array} \tag{6}$$

このとき右半分を可換にする $K[t]$-準同型 $Q : (K^n, B) \to (K^n, A)$ は一意に定まり、それは

$$Q(\psi_B(\mathbf{q}(t))) = \psi_A(Q(t)(\mathbf{q}(t)))$$

で定義される (Q が well-defined であること、$K[t]$-準同型であることを確かめよ)。$P(t), Q(t)$ が可逆だから Q も可逆である。つまり Q は (K^n, B) の (K^n, A) への $K[t]$-同型である。

このようにして我々は次の定理を得た。

定理 2.3. $A, B \in M_n(K)$ とする。$A \sim B$ であるための必要十分条件は $tE - A \leftrightarrow tE - B$ である。

■Q の求め方　上記の $Q \in M_n(K)$ は $Q(t) \in M_n(K[t])$ から次のように計算される。まず $Q(t) = (tE - A)Q_1(t) + Q'$, $Q' \in M_n(K)$ とする。このとき $Q = Q'$ である。実際、任意の $\mathbf{x} \in K^n$ に対し $Q(t)\mathbf{x} = (tE - A)Q_1(t)\mathbf{x} + Q'\mathbf{x}$. ここで第一項は $\mathrm{Img}\,\varphi_A$ の元、第二項は K^n の元。ゆえに $\psi_A\,(\,Q(t)\mathbf{x}\,) = Q'\mathbf{x}$。したがって $Q\mathbf{x} = (\,Q \circ \psi_B\,)\,\mathbf{x} = (\,\psi_A \circ Q(t)\,)\,\mathbf{x} = Q'\mathbf{x}$.

こうして次の式を得た：

(7)　$$Q(t) = (tE - A)Q_1(t) + Q.$$

Q は次のようにも計算できる。$Q(t) = t^p C_p + t^{p-1} C_{p-1} + \cdots + C_0$ と展開すると、

(8)　$$Q = A^p C_p + A^{p-1} C_{p-1} + \cdots + C_0.$$

である。これも p に関する帰納法で容易に示すことができる。詳細は読者に委ねる。

■最小多項式と特性多項式　$A \in M_n(K)$ とする。

$$I_A = \left\{\, f(t) \in K[t] \;\middle|\; f(A) = 0 \,\right\}$$

とおくと、I_A は $K[t]$ のイデアルである。ゆえに I_A は

$$I_A = K[t]m_A(t)$$

という形をしている。ここで生成元 $m_A(t)$ をモニックに限れば一意に決まり、これを A の**最小多項式**と呼ぶ。

A の最小多項式は A の特性行列 $tE - A$ の最後の単因子 $e_n(t)$ に一致することを以下で示そう。(6) に似た次の図式を考える。

(9)
$$\begin{array}{ccccccccc}
0 & \longrightarrow & (K[t])^n & \xrightarrow{\varphi_A} & (K[t])^n & \xrightarrow{\psi_A} & (K^n, A) & \longrightarrow & 0 \\
 & & \big\downarrow{\scriptstyle P(t)} & & \big\downarrow{\scriptstyle Q(t)} & & \big\downarrow{\scriptstyle Q} & & \\
0 & \longrightarrow & (K[t])^n & \xrightarrow{\rho} & (K[t])^n & \xrightarrow{\pi} & (K[t])^n/\mathrm{Img}\,\rho & \longrightarrow & 0
\end{array}$$

ここで $\rho : (K[t])^n \to (K[t])^n$ は φ_A に対等な単因子行列 $E(t) = (e_i(t))$ で定義される $K[t]$-準同型で、$P(t), Q(t)$ は基本変形をあたえる可逆行列である。φ_A が単射だから ρ も単射である。$\pi : (K[t])^n \to (K[t])^n/\mathrm{Img}\,\rho$ は商加群への自然な全射。ゆえに $\mathrm{Img}\,\rho = \mathrm{Ker}\,\pi$ は明らかで、図の下辺は $K[t]$-加群の短完全列となっている。$Q : (K^n, A) \to (K[t])^n/\mathrm{Img}\,\rho$ は図を可換にする唯一の $K[t]$-準同型として定義され、$P(t), Q(t)$ が可逆だから Q も可逆である。すなわち Q は $K[t]$-同型である。

さて $f(t) \in K[t]\ (\,f(A) = 0)$ とすると、$f \in I_A$ は「任意の $\mathbf{x} \in (K^n, A)$ に対し $f(t)\mathbf{x} = 0$」を意味する。ところが $K[t]$-同型 Q の存在により、これは任意の元 $\pi(\mathbf{q}(t)) \in (K[t])^n/\mathrm{Img}\,\rho$ に対して $f(t)\pi(\mathbf{q}(t)) = 0$ であることに同値。つまり $\pi(f(t)\mathbf{q}(t)) = 0$, あるいは $f(t)\mathbf{q}(t) \in \mathrm{Img}\,\rho$ と言い換えられる：

$$f(t) \in I_A \Longleftrightarrow f(t)\mathbf{q}(t) \in \mathrm{Img}\,\rho \quad (\forall \mathbf{q}(t) \in (K[t])^n)$$

ところが $\mathrm{Img}\,\rho$ とは

$$(e_1(t)p_1(t), \ldots, e_n(t)p_n(t)) \quad (p_i(t) \in K[t])$$

の形の元の集合。すなわち $f(t)\mathbf{q}(t) \in \mathrm{Img}\,\rho$ $(\forall \mathbf{q}(t) \in (K[t])^n)$ は

$$\begin{bmatrix} f(t)q_1(t) \\ \vdots \\ f(t)q_n(t) \end{bmatrix} = \begin{bmatrix} e_1(t)p_1(t) \\ \vdots \\ e_n(t)p_n(t) \end{bmatrix} \quad (\forall q_i(t) \in K[t], \exists p_i(t) \in K[t])$$

の意味になる。

容易に分かるようにこれが成り立つための必要十分条件は $e_n(t) \mid f(t)$ あるいは $f(t) \in K[t]e_n(t)$ であることである。

要するに $I_A = K[t]m_A(t) = K[t]e_n(t)$。ここから $m_A(t) = e_n(t)$ が結論される。こうして次の定理を得た。

定理 2.4. A の最小多項式は特性行列 $tE - A$ の最後の単因子に等しい。

この定理から Hamilton-Cayley の定理はほとんど明らかとなる。すなわち

定理 2.5 (Hamilton-Cayley の定理). $A \in M_n(K)$ の特性行列 $tE - A$ の行列式 $\chi_A(t) = |tE - A|$ は A の**特性多項式**と呼ばれる。任意の $A \in M_n(K)$ に対し

$$\chi_A(A) = 0$$

が成り立つ。

証明 行列式因子の定義から $\chi_A(t) = d_n(t)$。一方、補題 1.6 により $d_n(t) = e_1(t) \cdots e_n(t)$。したがって $\chi_A(t) = e_1(t) \cdots e_n(t)$。ここで前定理により $m_A(t) = e_n(t)$ だから、$m_A(t) \mid \chi_A(t)$ となる。したがって $\chi_A(A) = 0$。 □

3 ジョルダン標準形

以下では K は代数的閉体 (例えば $K = \mathbb{C}$) と仮定する。

■ジョルダン行列 $a \in K, k \geq 1$ とする。

$$J(a, k) = \begin{bmatrix} a & 1 & & \\ & a & & \\ & & \ddots & 1 \\ & & & a \end{bmatrix}$$

のかたちの k 次 K-行列を**ジョルダンブロック**と呼び、a をその**固有値**、k を**次数**と呼ぶ。

例 3.1.

$$J(a, 1) = \begin{bmatrix} a \end{bmatrix}, \quad J(a, 2) = \begin{bmatrix} a & 1 \\ 0 & a \end{bmatrix}, \quad J(a, 3) = \begin{bmatrix} a & 1 & 0 \\ 0 & a & 1 \\ 0 & 0 & a \end{bmatrix}, \quad \ldots$$

はジョルダンブロックである。

$$J = \begin{bmatrix} J(a_1, k_1) & & \\ & \ddots & \\ & & J(a_i, k_i) \end{bmatrix}$$

のようにジョルダンブロックを対角線上に並べてできる行列を**ジョルダン行列**と呼ぶ。

例 3.2.

$$\begin{bmatrix} a & \\ & b \end{bmatrix}, \quad \begin{bmatrix} a & & \\ & b & \\ & & c \end{bmatrix}, \quad \begin{bmatrix} a & & \\ & b & 1 \\ & & b \end{bmatrix}, \ldots$$

はジョルダン行列 (0 を省略)。

さてジョルダン行列は、そのブロックの並ぶ順序を別にすれば次のような表と一対一に対応していることが分かる (表は列の順序を変えても "同じ" 表と見なすものとする)。

(10)

a_1	a_2	$\ldots$	a_p
$k_{1,1}$	$k_{1,2}$	$\ldots$	$k_{1,p}$
$k_{2,1}$	$k_{2,2}$	$\ldots$	$k_{2,p}$
$\vdots$	$\vdots$		$\vdots$
$k_{n,1}$	$k_{n,2}$	$\ldots$	$k_{n,p}$

ここで p は $1 \leq p \leq n$ なる整数、$a_1, \ldots, a_p$ は K に属する数で、$i \neq j$ ならば $a_i \neq a_j$ であるとする。これらは J のジョルダンブロックの相異なる固有値に対応している。各 a_j の下に続く $\left\{ k_{i,j} \right\}_{1 \leq i \leq n}$ は

(11) $\quad 0 \leq k_{1,j} \leq k_{2,j} \leq \cdots \leq k_{n,j} \neq 0,$

(12) $\quad \displaystyle\sum_{i=1}^{n} \sum_{j=1}^{p} k_{i,j} = n$

を満たす整数列で、$k_{i,j}$ の記された枡目はジョルダンブロック $J(a_j, k_{i,j})$ を表す ($k_{i,j} = 0$ の場合、対応するジョルダンブロックは空行列とする)。

例 3.3. ジョルダン行列

$$\begin{bmatrix} a & 1 & & & \\ & a & & & \\ & & a & & \\ & & & a & \\ & & & & b \end{bmatrix}$$

に対応する表は

a	b
1	
1	
2	1

(0 の入る枡は空白にしてある)。

■ジョルダン行列の特性行列　はじめにジョルダンブロック $J(a, k)$ の特性行列 $tE_k - J(a, k)$ について

(13) $$tE_k - J(a, k) \leftrightarrow \begin{bmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & (t-a)^k \end{bmatrix}$$

を示そう。

これは行列式因子を計算してみれば分かる。$l < k$ とすると下のような l 次小行列式がつねにある。

$$\begin{vmatrix} -1 & & & \\ t-a & -1 & & \\ & t-a & \ddots & \\ & & & -1 \end{vmatrix} \quad (l\text{ 次})$$

この値は明らかに $(-1)^l$。ゆえに $d_l(t) = 1$。

一方 $tE_k - J(a,k)$ 全体の行列式は

$$\begin{vmatrix} t-a & -1 & & \\ & t-a & & \\ & & \ddots & -1 \\ & & & t-a \end{vmatrix}$$

この値も明らかに $(t-a)^k$。ゆえに $d_k(t) = (t-a)^k$。

したがって $tE_k - J(a,k)$ の単因子は $1, 1, \ldots, 1, (t-a)^k$ となって (13) を得る。

次にジョルダン行列の特性行列 (略して**ジョルダン特性行列**) の単因子を求めよう。次の補題が鍵になる。

補題 3.1. $a(t), b(t) \in K[t]$ とする。$d(t) = (\, a(t),\, b(t)\,)$, $m(t) = [\, a(t),\, b(t)\,]$ とおくと、

$$\begin{bmatrix} a(t) & 0 \\ 0 & b(t) \end{bmatrix} \leftrightarrow \begin{bmatrix} d(t) & 0 \\ 0 & m(t) \end{bmatrix}$$

である。とくに $a(t)$ と $b(t)$ が互いに素ならば、

$$\begin{bmatrix} a(t) & 0 \\ 0 & b(t) \end{bmatrix} \leftrightarrow \begin{bmatrix} 1 & 0 \\ 0 & a(t)b(t) \end{bmatrix}$$

である。

証明 証明では簡単のため $a = a(t), b = b(t)$ 等と略記する。$pa + qb = d$, $a = a'd$, $b = b'd$ とすると、$m = a'b$ である。ゆえに

$$\begin{aligned} &\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \leftrightarrow \begin{bmatrix} a & qb \\ 0 & b \end{bmatrix} \leftrightarrow \begin{bmatrix} a & pa+qb \\ 0 & b \end{bmatrix} = \begin{bmatrix} a & d \\ 0 & b \end{bmatrix} \\ &\leftrightarrow \begin{bmatrix} 0 & d \\ -a'b & b \end{bmatrix} \leftrightarrow \begin{bmatrix} 0 & d \\ -a'b & 0 \end{bmatrix} \leftrightarrow \begin{bmatrix} d & 0 \\ 0 & a'b \end{bmatrix} = \begin{bmatrix} d & 0 \\ 0 & m \end{bmatrix} \end{aligned}$$

□

これと前の結果からジョルダン行列の表 (10) の第 i 行にあるブロックをまとめた

$$\begin{bmatrix} J(a_1, k_{i,1}) & & & \\ & J(a_2, k_{i,2}) & & \\ & & \ddots & \\ & & & J(a_p, k_{i,p}) \end{bmatrix}$$

の特性行列は下の行列に対等と分かる。

$$\begin{bmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & \prod_{j=1}^{p} (t-a_j)^{k_{i,j}} \end{bmatrix}.$$

何故なら、$(t-a_1)^{k_{i,1}}, \ldots, (t-a_p)^{k_{i,p}}$ は互いに素だから。
　したがって J の特性行列は次の行列に対等である。

$$\begin{bmatrix} \prod_{j=1}^{p}(t-a_j)^{k_{1,j}} & & & \\ & \prod_{j=1}^{p}(t-a_j)^{k_{2,j}} & & \\ & & \ddots & \\ & & & \prod_{j=1}^{p}(t-a_j)^{k_{n,j}} \end{bmatrix}$$

ところが (11) からこれは単因子行列である。
　こうしてジョルダン特性行列 $tE-J$ の単因子 $e_1(t), \ldots, e_n(t)$ を得た。それは

$$(14) \quad e_i(t) = (t-a_1)^{k_{i,1}} \ldots (t-a_p)^{k_{i,p}} \quad (i = 1 \ldots n)$$

である。

定理 3.2. 任意の行列 $A \in M_n(K)$ はジョルダンブロックの順序の違いを別にして唯一つのジョルダン行列 $J \in M_n(K)$ に相似である。

証明 $tE-A$ の単因子を $e_1(t), \ldots, e_n(t)$ とする。K は代数的閉体だから各単因子 $e_i(t)$ は

$$e_i(t) = (t-a_1)^{k_{i,1}} \ldots (t-a_p)^{k_{i,p}}.$$

のように $K[t]$ の一次式に因数分解できる。$e_i(t) \mid e_{i+1}(t)$ ($1 \leq i \leq n-1$) だから $\left\{ k_{i,j} \right\}$ は (11) を満たす。また補題 1.6 で示したように $e_1(t) \cdots e_n(t)$ は特性行列式に一致するからその次数は n である。ゆえに $\left\{ k_{i,j} \right\}$ は (12) も満たす。すなわち $e_1(t), \ldots, e_n(t)$ はあるジョルダン行列 J の特性行列の単因子である。
ゆえに $tE-A$ と $tE-J$ は対等で、定理 2.3 から A と J は相似である。
(一意性) J_1, J_2 を n 次のジョルダン行列とする。$J_1 \sim J_2$ ならば $tE-J_1 \leftrightarrow tE-J_2$ で、$tE-J_1$ の単因子と $tE-J_2$ の単因子は等しい。ゆえに (14) から分かるように J_1 と J_2 は同じ表 (10) をもつ。つまり J_1 と J_2 は同じジョルダンブロックから成っている。 □

　A に相似なジョルダン行列 J を A の**ジョルダン標準形**という。

参考文献

[1] 斎藤正彦 「線型代数入門」東京大学出版会
[2] 堀田良之 「代数入門 − 群と加群」裳華房
[3] 堀田良之 「加群十話 − 代数学入門」朝倉書店
[4] I.M.Gel'fand "Lectures on Linear Algebra" Dover