システム工学 I
第 8 回
Jordan 標準形

固有値と固有ベクトル (1)

- $\boldsymbol{A}$ を n 次の正方行列とする.
- $\boldsymbol{v} \neq \boldsymbol{0}$ で, あるスカラー λ に対して, $\boldsymbol{A}\boldsymbol{v} = \lambda\boldsymbol{v}$ となるとき, λ を $\boldsymbol{A}$ の固有値, $\boldsymbol{v}$ を固有値 λ に対応する固有ベクトルという.
- 固有値と固有ベクトルの定義はどのような係数体に対しても適用可能であるが $\cdots$

固有値と固有ベクトル (2)

- 係数体の取り方しだいで, 行列 $\boldsymbol{A}$ が固有値を持つ場合と持たない場合が出て来る.
- 応用上重要なのは実行列であるが, 実行列は実数に範囲では必ずしも固有値を持つとは限らない. これは, 実数体が代数的閉体ではないという事実に起因する.

固有値と固有ベクトル (3)

- 固有値の重要な応用のひとつは線形連立微分方程式であるが, たとえば正弦波を解として持つ 2 次元の線形微分方程式は実数の固有値を持たないため, 実行列への限定は不便
- そこで, 通常は, 取り扱いたい行列が実行列であっても, 係数体を複素数体にまで拡張して考える.

固有値と固有ベクトル (4)

- 複素数体は代数的閉体であるため, 複素行列は固有値と対応する固有ベクトルを持つ.
- 実行列は複素行列の特別な場合なので, 固有値と, 対応する固有ベクトルを持つ. ただし, 固有値や固有ベクトルの要素が実数であることは保証されない.

固有値と固有ベクトル (5)

- 以下では, 特に断らない限り, $\boldsymbol{A}$ を n 次の複素行列とする.
- λ が行列 $\boldsymbol{A}$ の固有値で, 対応する固有ベクトルが $\boldsymbol{v}$ であれば, $(\lambda \boldsymbol{I} - \boldsymbol{A})\boldsymbol{v} = \boldsymbol{0}$ であり, 定義より $\boldsymbol{v} \neq \boldsymbol{0}$ であったから, $(\lambda \boldsymbol{I} - \boldsymbol{A})$ は正則ではなく, したがって $\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = 0$ である.

固有値と固有ベクトル (6)

- s を変数とし, $\Phi_A(s) = \det(s\boldsymbol{I} - \boldsymbol{A})$ と定義する．$\Phi_A(s)$ を $\boldsymbol{A}$ の特性多項式という．なお, $\det(\boldsymbol{A} - s\boldsymbol{I})$ を $\boldsymbol{A}$ の特性多項式と定義する流儀もある (符号が違うだけ).
- 文献によっては, 固有値を「特性多項式の根」と定義しているものもある (この講義の教科書もそうである).

固有値と固有ベクトル (7)

- λ を $\boldsymbol{A}$ のひとつの固有値とし, $V_\lambda = \{\boldsymbol{v} \in \mathbb{C}^n : \boldsymbol{A}\boldsymbol{v} = \lambda\boldsymbol{v}\}$ とする. V_λ は $\mathbb{C}^n$ の部分空間である. この次元を λ に対応する固有空間の次元という.
- 相異なる固有値に対応する固有ベクトルは一次独立である (証明は次ページ).

$\lambda_1, \ldots, \lambda_k$ を $\boldsymbol{A}$ の互いに相異なる固有値, $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$ を対応する固有ベクトルとする. k に関する帰納法によって, $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$ が一次独立であることを示す. $k = 1$ のときは主張は自明. よって, 「$k-1$ について主張が正しければ k についても主張は正しい」ということを示せばよい. 「k について主張が正しくなければ $k-1$ についても主張は正しくない」という形でこれを示す. $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ が一次従属であったと仮定する. すると, $\exists j (1 \le j \le k)$, $\exists c_l (1 \le l \le k, l \ne j)$, $\boldsymbol{v}_j = \sum_{l \ne j} c_l \boldsymbol{v}_l$ である. この両辺に $\boldsymbol{A}$ を左から掛けると, $\lambda_j \boldsymbol{v}_j = \sum_{l \ne j} c_l \lambda_l \boldsymbol{v}_l$ となる. 一方, $\boldsymbol{v}_j = \sum_{l \ne j} c_l \boldsymbol{v}_l$ の両辺に λ_j を掛けると, $\lambda_j \boldsymbol{v}_j = \sum_{l \ne j} c_l \lambda_j \boldsymbol{v}_l$ となり, これらの 2 式を引くと, $\boldsymbol{0} = \sum_{l \ne j} c_l (\lambda_l - \lambda_j) \boldsymbol{v}_l$ となる. $\lambda_j \ne \lambda_l$ で, $\{c_l\}$ の中には少なくともひとつ零でないものがあるから, $k-1$ について主張が正しくないことが示された. この待遇を取ると, 先の主張が証明される.

固有値と固有ベクトル (9)

- $\boldsymbol{A}$ の固有ベクトル全体がつねに $\mathbb{C}^n$ の基底となるのであれば話は簡単なのだが $\cdots$
- 残念ながら, 必ずしもそうなるとは限らない.
- $\boldsymbol{A}$ の固有ベクトル全体が $\mathbb{C}^n$ の基底になっているときには, 固有ベクトルを集めて作った行列により $\boldsymbol{A}$ を対角化することができる.

固有値と固有ベクトル (10)

- そうでない場合には, Jordan 標準形という概念が必要になる.
- まず対角化可能な場合について見てゆく.

対角化 (1)

- $\boldsymbol{A} \in \mathbb{C}^n$ が n 個の一次独立な固有ベクトル $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ を持つものとする. $\boldsymbol{v}_i$ に対応する固有値を λ_i とする $(1 \leq i \leq n)$. 固有値の中には同一のものがあっても構わない.
- $\boldsymbol{V} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ (固有ベクトルを横に並べて作った行列) とする.

対角化 (2)

- $\boldsymbol{A}\boldsymbol{v}_i = \lambda_i \boldsymbol{v}_i$ だから, $\boldsymbol{A}\boldsymbol{V} = \boldsymbol{V}\,\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ である (順番に注意).
- したがって, $\boldsymbol{V}^{-1}\boldsymbol{A}\boldsymbol{V} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ である. このようにすることを行列の**対角化**という. なお, $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ が一次独立であるから, $\boldsymbol{V}$ は逆行列を持つ.

対角化 (3)

- 対角化可能のための十分条件を考える.
- まず, 行列 $\boldsymbol{A}$ の特性多項式が重根を持たなければ, n 個の 1 次独立な固有ベクトルを取れるので, $\boldsymbol{A}$ は対角化できる.
- これ以外で応用上重要な対角化可能行列は, 対称行列および Hermite 行列である.

対角化 (4)

- 対称行列は $\boldsymbol{A}^T = \boldsymbol{A}$ となる行列, Hermite 行列は $\boldsymbol{A}^* = \boldsymbol{A}$ となる行列であった．ただし $\boldsymbol{A}^*$ は $\boldsymbol{A}$ の共役転置 (各成分の複素共役を取ってから転置)．対称行列は成分が実数の Hermite 行列である．

- 対称行列や Hermite 行列の固有値は実数であることが示せる (次ページ).

- $\boldsymbol{A}$ を Hermite 行列とする．λ を $\boldsymbol{A}$ の固有値, $\boldsymbol{v}$ を対応する固有ベクトルとする．$\boldsymbol{v}^*(\boldsymbol{A}\boldsymbol{v}) = \lambda \boldsymbol{v}^*\boldsymbol{v}$ である．また, $\boldsymbol{v} = (v_1, \ldots, v_n)^T$ としたとき, $\boldsymbol{v}^* = (\bar{v}_1, \ldots, \bar{v}_n)$ だから ($\bar{x}$ は x の複素共役), $\boldsymbol{v}^*\boldsymbol{v} = |v_1|^2 + \cdots + |v_n|^2 > 0$ である．一方, $\boldsymbol{A}^* = \boldsymbol{A}$ より, $\boldsymbol{v}^*\boldsymbol{A}\boldsymbol{v} = (\boldsymbol{v}^*\boldsymbol{A}^*)\,\boldsymbol{v}$ であるが, $\boldsymbol{v}^*\boldsymbol{A}^* = (\boldsymbol{A}\boldsymbol{v})^* = (\lambda\boldsymbol{v})^* = \bar{\lambda}\boldsymbol{v}^*$, よって $\boldsymbol{v}^*\boldsymbol{A}\boldsymbol{v} = \bar{\lambda}\boldsymbol{v}^*\boldsymbol{v}$. $\boldsymbol{v}^*\boldsymbol{v} > 0$ だから, $\lambda = \bar{\lambda}$, よって固有値は実数である．
- $\boldsymbol{A}$ が対称行列の場合には, $\boldsymbol{A}^* = \boldsymbol{A}^T$ となるから, 上記の証明は変更なしで適用可能である．
- なお, 固有ベクトルは, $(\boldsymbol{A} - \lambda\boldsymbol{I})\boldsymbol{v} = \boldsymbol{0}$ の解だから, 対称行列の固有ベクトルは, その全要素が実数であるように選べる．この性質を直交行列による対角化のところで使う．

対角化 (6)

- $\boldsymbol{A}$ を n 次の対称行列あるいは Hermite 行列とし, λ, μ を $\boldsymbol{A}$ の相異なる固有値, $\boldsymbol{v}$, $\boldsymbol{w}$ を対応する固有ベクトルとする. $\boldsymbol{A}$ が対称行列なら $\boldsymbol{v}^T\boldsymbol{w} = 0$, Hermite 行列なら $\boldsymbol{v}^*\boldsymbol{w} = 0$ となる.

対称行列については, $\mu\boldsymbol{w}^T\boldsymbol{v} = \boldsymbol{w}^T\boldsymbol{A}^T\boldsymbol{v} = \boldsymbol{w}^T\boldsymbol{A}\boldsymbol{v} = \lambda\boldsymbol{w}^T\boldsymbol{v}$ で, $\mu \neq \lambda$ だから, $\boldsymbol{w}^T\boldsymbol{v} = 0$. Hermite 行列については, 上記の $\quad^T$ を $\quad^*$ に置き換えればよい.

対角化 (7)

- 対称行列は直交行列によって対角化できる.
- Hermite 行列はユニタリ行列によって対角化できる.

証明はどちらもほぼ同じなので, 対称行列について示す. Hermite 行列の場合の変更点は, (i) ^{T} を * に変更すること, (ii) ベクトルの内積を $\boldsymbol{v}^T\boldsymbol{w}$ から $\boldsymbol{v}^*\boldsymbol{w}$ にすること, (iii) 固有ベクトルの成分が複素数であることを許容すること, の 3 点である.

対称行列の場合を示す. n に関する帰納法による. $n=1$ の場合は明らか. $n-1$ まで主張が正しいと仮定し, $\boldsymbol{A}$ を n 次の対称行列とする. λ を $\boldsymbol{A}$ のある固有値, $\boldsymbol{v}_1$ を対応する固有ベクトルとする. $\boldsymbol{v}_1$ の全要素は実数にできる (既出). $\boldsymbol{v}_1$ を正規化し, $\boldsymbol{v}_1$ を含む正規直交基底 $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ を取る. $\boldsymbol{V} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ とする. これは直交行列である. すると, $\boldsymbol{V}^T\boldsymbol{A}\boldsymbol{V}$ は対称行列であり, さらに $\boldsymbol{V}^T\boldsymbol{A}\boldsymbol{V} = \left(\begin{array}{c|c} 1 & \mathbf{0} \\ \hline \mathbf{0} & \boldsymbol{A}_{n-1} \end{array}\right)$ となる ($\boldsymbol{A}_{n-1}$ は $n-1$ 次の対称行列). さて, W を $n-1$ 次の直交行列とすると,

$$\left(\begin{array}{c|c} 1 & \mathbf{0} \\ \hline \mathbf{0} & \boldsymbol{W} \end{array}\right)^T \left(\begin{array}{c|c} 1 & \mathbf{0} \\ \hline \mathbf{0} & \boldsymbol{A}_{n-1} \end{array}\right) \left(\begin{array}{c|c} 1 & \mathbf{0} \\ \hline \mathbf{0} & \boldsymbol{W} \end{array}\right) = \left(\begin{array}{c|c} 1 & \mathbf{0} \\ \hline \mathbf{0} & \boldsymbol{W}^T\boldsymbol{A}_{n-1}\boldsymbol{W} \end{array}\right),$$

よって帰納法の仮定により $\boldsymbol{A}$ は直交行列によって対角化可能.

Jordan 標準形 (1)

- $\boldsymbol{N} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ とする.
- $\boldsymbol{N}$ の特性多項式は $\det(s\boldsymbol{I} - \boldsymbol{N}) = s^3$ だから, 固有値は零のみ.
- 固有ベクトルは?

Jordan 標準形 (2)

- 固有ベクトルは, $(\lambda \boldsymbol{I} - \boldsymbol{N})|_{\lambda=0}\, \boldsymbol{v} = \boldsymbol{0}$ の解
- よって, $-\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$
- 上記を解くと $x_2 = 0, x_3 = 0$ が出るから, 固有ベクトルは $(1, 0, 0)^T$ の零でない定数倍のみ.

Jordan 標準形 (3)

- 「$\boldsymbol{A}$ の固有ベクトルから成る $\mathbb{C}^n$ の基底があれば $\boldsymbol{A}$ は対角化できる」のだが $\cdots$
- そのような基底は必ずしも存在しない.
- 上記の $\boldsymbol{N}$ のような行列は, 純粋な積分器を含む微分方程式系で出現するため, 応用上もこのような場合を無視するわけにはいかない.

Jordan 標準形 (4)

- 天下り的であるが, $\boldsymbol{w} = (0, 0, 1)^T$ とすると,
$$\boldsymbol{w} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \boldsymbol{N}\boldsymbol{w} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \boldsymbol{N}^2\boldsymbol{w} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$
より, この例では $\boldsymbol{w}, \boldsymbol{N}\boldsymbol{w}, \boldsymbol{N}^2\boldsymbol{w}$ が基底になっている.

Jordan 標準形 (5)

- 行列 $\boldsymbol{N}$ の例では, 特性多項式が重根を持つことに注意する. 先に述べた理由によって, その特性多項式が重根を持たない行列 $\boldsymbol{A}$ は対角化可能である. 一方, 重根を持つ場合には, 対角化はできることもできないこともある.
- 固有値 λ の特性多項式の根としての重複度を, その固有値の重複度という.

Jordan 標準形 (6)

- 行列 $\boldsymbol{A}$ が対角化不能のときにも, 固有ベクトルから (対角化よりは繁雑な手順によって) 基底を作り, 対応する表現行列が望ましい性質を持つようにできる. この典型例が, これから述べる (**Jordan 標準形**) である.
- 行列 $\boldsymbol{A}$ の Jordan 標準形を $\boldsymbol{J}$ とする. これは, 次のような形の標準形である.

Jordan 標準形 (7)

- $\boldsymbol{J}$ は, k 個のブロックが対角線上にならんだ, 次のような形のブロック対角行列である：

$$\boldsymbol{J} = \begin{pmatrix} \boldsymbol{J}_1 & & \\ & \ddots & \\ & & \boldsymbol{J}_k \end{pmatrix}$$

Jordan 標準形 (8)

- 各 $\boldsymbol{J}_i$ は m_i 次の正方行列で, $\sum_{i=1}^{k} m_i = n$ である. $\boldsymbol{J}_i$ は次のような形になっている：

$$\boldsymbol{J}_i = \lambda_i \boldsymbol{I}_{m_i} + \boldsymbol{N}_{m_i},\ \boldsymbol{N}_{m_i} = \left(\begin{array}{c|c} \mathbf{0} & \boldsymbol{I}_{m_i-1} \\ \hline 0 & \mathbf{0} \end{array}\right)$$

- λ_i は固有値である.

Jordan 標準形 (9)

- $m_i \geq 1$ であり, $m_i = 1$ のときには $\boldsymbol{N}_{m_i}$ の部分は存在しないものと見做す. すべてのブロックについて $m_i = 1$ となるのが, 対角化可能な場合である.
- 例を示す.

- Jordan 標準形の例 (空白部分は零)

$$
\left(\begin{array}{ccc|cccc}
7 & 1 & & & & & \\
 & 7 & 1 & & & & \\
 & & 7 & & & & \\
\hline
 & & & 5 & 1 & & \\
 & & & & 5 & 1 & \\
 & & & & & 5 & 1 \\
 & & & & & & 5
\end{array}\right), \quad
\left(\begin{array}{cc|cc|cc}
2 & 1 & & & & \\
 & 2 & & & & \\
\hline
 & & 2 & 1 & & \\
 & & & 2 & & \\
\hline
 & & & & 2 & 1 \\
 & & & & & 2
\end{array}\right)
$$

Jordan 標準形 (10)

- 以下では, f を線形写像とし, 変換 f の表現行列が Jordan 標準形になるような基底を構成する. 行列 $\boldsymbol{A}$ の Jordan 標準形を求めるときには, $\boldsymbol{A}$ に対応する線形写像を考えればよい.
- V を n 次元の複素ベクトル空間とし, $f : V \to V$ を線形写像とする. 示すべきことは次の事実である.

ある $k > 0$ と, k 個のスカラー $\lambda_1, \ldots, \lambda_k$ および各 λ_i に対応する m_i 個のベクトル $\boldsymbol{v}_{i,1}, \ldots, \boldsymbol{v}_{i,m_i}$ が存在し (ただし $m_i > 0$, $\sum_{i=1}^{k} m_i = n$), $\{\boldsymbol{v}_{ij} : 1 \leq i \leq k, 1 \leq j \leq m_i\}$ はV の基底で, 各 i について次式を満たす:

$$
\begin{aligned}
f(\boldsymbol{v}_{i,1}) &= \lambda_i \boldsymbol{v}_{i,1} \\
f(\boldsymbol{v}_{i,2}) &= \lambda_i \boldsymbol{v}_{i,2} + \boldsymbol{v}_{i,1} \\
&\ldots, \\
f(\boldsymbol{v}_{i,m_i}) &= \lambda_1 \boldsymbol{v}_{i,m_i} + \boldsymbol{v}_{i,m_i-1}
\end{aligned}
$$

Jordan 標準形 (12)

- $f(\boldsymbol{v}_{i,j})$ に関する式を行列を使って書き直すと，$f((\boldsymbol{v}_{i,1},\ldots,\boldsymbol{v}_{i,m_i})) = (\boldsymbol{v}_{i,1},\ldots,\boldsymbol{v}_{i,m_i})\boldsymbol{J}_i$ となる．よって，

$$f((\boldsymbol{v}_{1,1},\ldots,\boldsymbol{v}_{1,m_1},\ldots,\boldsymbol{v}_{k,1},\ldots,\boldsymbol{v}_{k,m_k})) \\ = (\boldsymbol{v}_{1,1},\ldots,\boldsymbol{v}_{1,m_1},\ldots,\boldsymbol{v}_{k,1},\ldots,\boldsymbol{v}_{k,m_k})\boldsymbol{J}$$

であり，したがってこの基底に関する f の表現行列は Jordan 標準形である．

- 証明は V の次元 n に関する帰納法による. $n=1$ の場合は, f の表現行列はスカラーだから, 証明すべきことは何もない.
- $n-1$ まで主張は正しいと仮定して, n に対しても主張が正しいことを示す.
- (定義) $f(W) \subset W$ を満たす V の部分空間 W のことを f-不変部分空間という.
- (補題) $f: V \to V$ が線形写像であれば, V の $n-1$ 次元 f-不変部分空間 W が存在する.

(補題の証明) V の基底 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ を取り, この基底に関する f の表現行列を $\boldsymbol{A}$ とする. $\boldsymbol{A}^*$ のある固有値を λ とし, 対応する固有ベクトルを $\boldsymbol{w}$ とする. $X = \{\boldsymbol{x} \in \mathbb{C}^n : \boldsymbol{w}^*\boldsymbol{x} = 0\}$ とおくと, X は $\mathbb{C}^n$ の $n-1$ 次元部分空間であり, よって $W = \{(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)\boldsymbol{x} : \boldsymbol{x} \in X\}$ は V の $n-1$ 次元線形部分空間である. ((ただし $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)\boldsymbol{x}$ は $\boldsymbol{v}_1 x_1 + \cdots + \boldsymbol{v}_n x_n$ を意味する). また, $f((\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)\boldsymbol{x}) = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)\boldsymbol{A}\boldsymbol{x}$ で, $\boldsymbol{w}^*\boldsymbol{A}\boldsymbol{x} = \overline{\boldsymbol{x}^*\boldsymbol{A}^*\boldsymbol{w}} = \overline{\lambda\boldsymbol{x}^*\boldsymbol{w}} = \overline{\lambda}\boldsymbol{w}^*\boldsymbol{x} = 0$ だから, $\boldsymbol{A}\boldsymbol{x} \in X$. よって W は f-不変である.

- 先の補題により, V には $n-1$ 次元 f-不変部分空間 W が存在する.
- 帰納法の仮定により, f を W に制限した写像については主張は正しい. すなわち, ある $k>0$ と, k 個のスカラー $\lambda_1,\ldots,\lambda_k$ および各 λ_i に対応する m_i 個のベクトル $\boldsymbol{v}_{i,1},\ldots,\boldsymbol{v}_{i,m_i}$ が存在し (ただし $m_i>0$, $\sum_{i=1}^{k} m_i = n$), $\{\boldsymbol{v}_{ij} : 1 \le i \le k, 1 \le j \le m_i\}$ は W の基底で, $f((\boldsymbol{v}_{i,1},\ldots,\boldsymbol{v}_{i,m_i})) = (\boldsymbol{v}_{i,1},\ldots,\boldsymbol{v}_{i,m_i})\boldsymbol{J}_i$ となる.
- $\Omega = \{\boldsymbol{v}_{i,j} : 1 \le i \le k, 1 \le j \le m_i\}$ とし, $\Omega \cup \{\boldsymbol{v}\}$ が V の基底となるよう $\boldsymbol{v}$ を選ぶ.
- $f(\boldsymbol{v}) = \lambda_0\boldsymbol{v} + \sum a_{i,j}\boldsymbol{v}_{i,j}$ となる (λ_0 は固有値とは無関係)
- $I(\cdot)$ を恒等写像とし, $f'(\boldsymbol{x}) = f(\boldsymbol{x}) - \lambda_0 I(\boldsymbol{x})$ とおく; すると $f'(\boldsymbol{v}) = \sum a_{i,j}\boldsymbol{w}_{i,j}$ となる

- ある座標系で f' の表現行列が Jordan 標準形になれば, f の表現行列は f' のそれに $\lambda_0 \boldsymbol{I}$ を加えたものになるので, やはり Jordan 標準形である. よって, f' について主張を示せばよい.
- f と f' の基底 $\Omega \cup \{\boldsymbol{v}\}$(この順に並べる) に関する表現行列は, それぞれ, 次のようになる. ただし, $\boldsymbol{J}'_i$ は $\boldsymbol{J}_i$ の対角要素を $\lambda_i - \lambda_1$ で置き換えたもの.

$$f \Leftrightarrow \left(\begin{array}{ccc|c} \boldsymbol{J}_1 & & & * \\ & \ddots & & \vdots \\ & & \boldsymbol{J}_k & * \\ & & & \lambda \end{array}\right), \quad f' \Leftrightarrow \left(\begin{array}{ccc|c} \boldsymbol{J}'_1 & & & * \\ & \ddots & & \vdots \\ & & \boldsymbol{J}'_k & * \\ & & & 0 \end{array}\right)$$

- $\lambda'_i = \lambda_i - \lambda_1$ と定義する.

- 基底を取り直すことで, f' の表現行列を簡単にすることを考える.
- 先のページの $*$ の部分がはじめから零であれば, すでに Jordan 標準形が得られているので, これ以上やるべきことは何もない.
- そうでない場合, $*$ の部分が零になるよう基底を取り直す.
- まず, $\forall i, \lambda_i' \neq 0$ の場合を考える. $\{p_{i,j} : 1 \leq i \leq k, 1 \leq j \leq m_i\}$ をパラメータとし, $\boldsymbol{v}' = \boldsymbol{v} + \sum_{i=1}^{k} \left(p_{i,1}\boldsymbol{v}_{i,1} + \sum_{j=2}^{m_i} p_{i,j}\boldsymbol{v}_{i,j} \right)$ とする. $\Omega \cup \{\boldsymbol{v}'\}$ は基底である. $f'(\boldsymbol{v}') = \sum_{i,j} a_{i,j}\boldsymbol{v}_{i,j} + + \sum_{i=1}^{k} \left(p_{i,1}\lambda_i\boldsymbol{v}_{i,1} + \sum_{j=2}^{m_i} p_{i,j} \left(\lambda_i\boldsymbol{v}_{i,j} + \boldsymbol{v}_{i,j-1} \right) \right)$ だから, 各 i に対し, $p_{i,m_i} = -a_{i,j_i}/\lambda_i$ とし, $j = m-1, \ldots, 1$ の順に $p_{i,j} = -(a_{i,j} - p_{i,j+1})/\lambda_i$ によってパラメータを計算すれば, $*$ の部分を零にできる.

- 次に $\exists i, \lambda'_i = 0$ の場合を考える． 基底を並べ換えて, $\forall j \leq q$, $\lambda'_j = 0$, $m_1 \geq m_2 \geq \cdots \geq m_q$, $\forall j \geq q+1$, $\lambda'_j \neq 0$ のであるようにする．

  1. $\{p_{i,j} : q+1 \leq i \leq k, 1 \leq j \leq m_i\}$ を パラメータ とし, $\boldsymbol{v}^{(1)} = \boldsymbol{v} + \sum_{i=q+1}^{k} \left(p_{i,1}\boldsymbol{v}_{i,1} + \sum_{j=2}^{m_i} p_{i,j}\boldsymbol{v}_{i,j} \right)$ とする． $\Omega \cup \{\boldsymbol{v}^{(1)}\}$ は基底である．先と同様に, 各 $i \geq q+1$ に対し, $p_{i,m_i} = -a_{i,j_i}/\lambda_i$ とし, $j = m-1, \ldots, 1$ の順に $p_{i,j} = -(a_{i,j} - p_{i,j+1})/\lambda_i$ とすれば, 第 $q+1$ ブロック以降の $*$ の部分を零にできる．
  2. $i \leq q$ であれば, $f(\boldsymbol{v}_{i,1}) = \boldsymbol{0}$ かつ $\forall j : 2 \leq j \leq m_i$, $f(\boldsymbol{v}_{i,j}) = \boldsymbol{v}_{i,j-1}$ である． $\boldsymbol{v}^{(2)} = \boldsymbol{v}^{(1)} - \sum_{i=1}^{q} \sum_{j=1}^{m_i-1} a_{i,j}\boldsymbol{v}_{i,j+1}$ とする． $\Omega \cup \{\boldsymbol{v}^{(2)}\}$ は基底である．さらに, $f(\boldsymbol{v}^{(2)}) = \sum_{i=1}^{q} a_{i,m_i}\boldsymbol{v}_{i,m_i}$ となる．

- ($\exists i, \lambda_i' = 0$ の場合, 続き (1))

3. $\forall i, a_{i,j_i} = 0$ の場合には $*$ の部分はすべて零になっているので終了．そうでない場合には, $r = \min\{j : a_{j,m_j} \neq 0\}$ とする．すると, $f(\boldsymbol{v}^{(2)}) = \sum_{i=r}^{q} a_{i,m_i}\boldsymbol{v}_{i,m_i}$ となる．
4. $\boldsymbol{v}^{(3;0)} = \boldsymbol{v}^{(2)}$ とし, $j \geq 1$ に対し, 帰納的に, $\boldsymbol{v}^{(3;j)} = f(\boldsymbol{v}^{(3;j-1)})$ と定義する．記法の 簡単のために, $l \leq 0$ のときには $\boldsymbol{v}_{i,l} = \boldsymbol{0}$ と定義すると, $\boldsymbol{v}^{(3;1)} = f(\boldsymbol{v}^{(3;0)}) = \sum_{i=r}^{q} a_{i,m_i}\boldsymbol{v}_{i,m_i}$ であり, $\boldsymbol{J}_i'$ の構造より, $\boldsymbol{v}^{(3;j)} = \sum_{i=r}^{q} a_{i,m_i}\boldsymbol{v}_{i,m_i-j+1}$ となる．したがって, $\boldsymbol{v}^{(3;m_r)} = \sum_{i=r}^{q} a_{i,m_i}\boldsymbol{v}_{i,m_i-m_r+1}$ となり, $i \geq r$ なら $m_r \geq m_i$ であったから, $f(\boldsymbol{v}^{(3;m_r)}) = \boldsymbol{0}$ である．

- $(\exists i, \lambda'_i = 0$ の場合, 続き (2))

5. 前段で構成したベクトルを逆順に並べると, $f((\boldsymbol{v}^{(3,m_i)}, \ldots, \boldsymbol{v}^{(3,0)})) = (\boldsymbol{v}^{(3,m_i)}, \ldots, \boldsymbol{v}^{(3,0)})\boldsymbol{N}_{m_r+1}$ となり, Jordan 標準形の一部が構成されている. このステップにおける新しい基底の構成に関与したベクトルは, $\boldsymbol{v}^{(2)}$ 以外には, $\{\boldsymbol{v}_{i,j} : r \leq i \leq q, 1 \leq j \leq m_i\}$ のみなので, $\mathcal{B}_1$ を $\{\boldsymbol{v}^{(3;m_r-j+1)} : 1 \leq j \leq m_r\} \cup \{\boldsymbol{v}_{i,j} : r+1 \leq i \leq q, 1 \leq j \leq m_i\}$ をこの順に並べた基底の一部, $\mathcal{B}_2$ を $\{\boldsymbol{v}_{i,j} : r \leq i \leq q, 1 \leq j \leq m_i\}$ をこの順に並べた基底の一部としたとき, $\mathcal{B}_1$ と $\mathcal{B}_2$ の張る空間が一致することを証明できれば, Jordan 標準形の構成が完了する.

- $(\exists i, \lambda'_i = 0$ の場合, 続き (3))

  6. $m_r \geq m_{r+1} \geq \cdots geqm_q$ であったことに注意する. $r \leq j \leq q$ に対し, $\boldsymbol{D}_j = a_{j,m_j}\boldsymbol{I}_{m_j}$ とし, $r+1 \leq j \leq q$ に対して $\boldsymbol{E}_j = (\boldsymbol{0}_{m_j \times (m_r - m_j)}, \boldsymbol{D}_j)$ とすると, $\mathcal{B}_1$ は $\mathcal{B}_2$ に以下の正方行列を右から乗ずることで得られる. この行列は正則なので, 我々がおこなったことが基底変換であることが示される.

$$\begin{pmatrix} \boldsymbol{D}_r & & & \\ \boldsymbol{E}_{r+1} & \boldsymbol{I}_{m_{r+1}} & & \\ \vdots & & \ddots & \\ \boldsymbol{E}_q & & & \boldsymbol{I}_{m_q} \end{pmatrix}$$

Jordan 標準形 (22)

- 以上の証明は, I. M. Gelfand (Translated by A. Shenitzer, Lectures on linear algebra, Dover, 1989 の記述に手を加えたものである. 証明が長くて難しいことに驚いたと思うが, これでも相対的には簡単な証明であり, 線形代数の (きちんとした) 教科書では, Jordan 標準形の導出に 30 ページ前後が費されているものもある.
- 以上の証明は相対的には簡単なのだが, Jordan 標準形におけるブロック $\boldsymbol{J}_i$ の大きさが行列 (あるいは対応する線型写像) から一意的に決まることが示せないという欠点がある.

線形連立微分方程式への応用 (1)

- 正方行列は, 基底の変換によって, 対角化可能できる場合もあれば, できない場合もある
- 正方行列は, つねに, 基底の変換によって Jordan 標準形に変換できる
- これらの応用のひとつは, 線形連立微分方程式の解法である

線形連立微分方程式への応用 (2)

- 微分方程式 $\dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x}$, $\boldsymbol{x}(0) = \boldsymbol{x}_0$ を考える．ただし, $\boldsymbol{x}(t) \in \mathbb{R}^n$ で, $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ とする．
- この微分方程式の解は, $\boldsymbol{x}(t) = \exp[\boldsymbol{A}t]\boldsymbol{x}_0$ であった．
- 行列が対角化されているか, Jordan 標準形になっていると, $\exp[\boldsymbol{A}t]$ を具体的に計算できる．

線形連立微分方程式への応用 (3)

- まず, ある正則行列 $\boldsymbol{P}$ によって, $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{D}$, $\boldsymbol{D} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ となっている場合を考える. このとき, $\boldsymbol{A} = \boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{-1}$, このとき, $\boldsymbol{A}^2 = (\boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{-1})(\boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{-1}) = \boldsymbol{P}\boldsymbol{D}^2\boldsymbol{P}^{-1}$ で, 以下帰納的に, $\boldsymbol{A}^k = \boldsymbol{P}\boldsymbol{D}^k\boldsymbol{P}^{-1}$.
- したがって, $\exp[\boldsymbol{A}t] = \sum_{k=0}^{\infty} \frac{1}{k!}\boldsymbol{A}^k t^k$
$= \boldsymbol{P}\left(\sum_{k=0}^{\infty} \frac{1}{k!}\boldsymbol{D}^k t^k\right)\boldsymbol{P}^{-1}$

線形連立微分方程式への応用 (4)

- $\boldsymbol{D}^k = \mathrm{diag}(\lambda_1^k, \dots, \lambda_n^k)$ であるから, 以上により, $\boldsymbol{A}$ が対角化可能な場合には,

$$\exp[\boldsymbol{A}t] = \boldsymbol{P} \begin{pmatrix} e^{\lambda_1 t} & & \\ & \ddots & \\ & & e^{\lambda_n t} \end{pmatrix} \boldsymbol{P}^{-1}$$

となる.

線形連立微分方程式への応用 (5)

- 次に, $\boldsymbol{A}$ の Jordan 標準形の特別な場合として, $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \lambda \boldsymbol{I} + \boldsymbol{N}$, $\boldsymbol{N} = \left(\begin{array}{c|c} \boldsymbol{0} & \boldsymbol{I}_{n-1} \\ \hline 0 & \boldsymbol{0} \end{array}\right)$ となっている場合を考える.
- $\boldsymbol{A}$ と $\boldsymbol{B}$ が可換な正方行列のときには, 通常の指数関数と同様に, $\exp[\boldsymbol{A}+\boldsymbol{B}] = \exp[\boldsymbol{A}]\exp[\boldsymbol{B}]$ が成り立つ.

線形連立微分方程式への応用 (6)

- $\boldsymbol{A}$ と $\boldsymbol{B}$ が非可換の場合は先の式は不成立.
- $\boldsymbol{I}$ と $\boldsymbol{N}$ は可換, よって

$$\exp[(\lambda \boldsymbol{I} + \boldsymbol{N})t] = \exp[\lambda \boldsymbol{I} t] \exp[\boldsymbol{N} t]$$
$$= e^{\lambda t} \exp[\boldsymbol{N} t].$$

- 帰納的に, $\boldsymbol{N}^k = \left(\begin{array}{c|c} \boldsymbol{0}_{n-k \times k} & \boldsymbol{I}_{n-k} \\ \hline \boldsymbol{0}_{k \times k} & \boldsymbol{0}_{k \times n-k} \end{array} \right).$

線形連立微分方程式への応用 (7)

- したがって, $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \lambda\boldsymbol{I} + \boldsymbol{N}$ の場合には

$$\exp[\boldsymbol{A}t] = \boldsymbol{P}\begin{pmatrix} e^{\lambda t} & te^{\lambda t} & \cdots & \frac{t^{n-1}}{(n-1)!}e^{\lambda t} \\ & \ddots & \ddots & \vdots \\ & & \ddots & te^{\lambda t} \\ & & & e^{\lambda t} \end{pmatrix}\boldsymbol{P}^{-1}$$

- 一般の場合は上記の組み合わせ.

線形連立微分方程式への応用 (8)

- $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \mathrm{diag}(\boldsymbol{J}_1, \ldots, \boldsymbol{J}_k)$, $\boldsymbol{J}_i = \lambda_i \boldsymbol{I}_{m_i} + \boldsymbol{N}_{m_i}$, $\boldsymbol{N}_{m_i} = \left(\begin{array}{c|c} \boldsymbol{0} & \boldsymbol{I}_{m_i-1} \\ \hline 0 & \boldsymbol{0} \end{array}\right)$ $(1 \leq i \leq k)$ とすると ⋯

$$\exp[\boldsymbol{A}t] = \boldsymbol{P}\,\mathrm{diag}(\exp[\boldsymbol{J}_1 t], \ldots, \exp[\boldsymbol{J}_k t])\boldsymbol{P}^{-1}$$

- $\exp[\boldsymbol{J}_i t]$ は ⋯

線形連立微分方程式への応用 (9)

$$\exp[\boldsymbol{J}_i t] = \begin{pmatrix} e^{\lambda_i t} & te^{\lambda_i t} & \cdots & \frac{t^{n_i-1}}{(n_i-1)!}e^{\lambda_i t} \\ & \ddots & \ddots & \vdots \\ & & \ddots & te^{\lambda_i t} \\ & & & e^{\lambda_i t} \end{pmatrix}$$

線形連立微分方程式への応用 (10)

- 以上により, 任意の正方行列 $\boldsymbol{A}$ に対し, $\exp[\boldsymbol{A}t]$ が書き下せたので, 連立線形微分方程式の解の具体的な形がわかったことになる.

Kronecker 積 (1)

- システム工学の分野では, 行列の Kronecker 積と呼ばれる演算が使われることがある．これは, $\boldsymbol{A} = (a_{ij})$ と $\boldsymbol{B} = (b_{ij})$ という任意の 2 個の行列に対して定義され (型が合っている必要はない), $\boldsymbol{A} \otimes \boldsymbol{B}$ という記号であらわされる．$\boldsymbol{A}$ を m 行 n 列の行列とすると, $\boldsymbol{A} \otimes \boldsymbol{B}$ は, 次のようなブロック行列である．

Kronecker 積 (2)

$$\boldsymbol{A} \otimes \boldsymbol{B} = \begin{pmatrix} a_{11}\boldsymbol{B} & \cdots & a_{1n}\boldsymbol{B} \\ \vdots & & \vdots \\ a_{m1}\boldsymbol{B} & \cdots & a_{mn}\boldsymbol{B} \end{pmatrix}$$

- $\boldsymbol{B}$ が p 行 q 列の行列であれば, $\boldsymbol{A} \otimes \boldsymbol{B}$ は mp 行 nq 列の行列になる.

行列値関数の微分と積分 (1)

- $\boldsymbol{A}(t)$ を各成分が時間 t の関数である行列としたとき: $\boldsymbol{A}(t) = (a_{ij}(t))_{1\le i\le m, 1\le j\le n}$;
$$\frac{d\boldsymbol{A}(t)}{dt} = \left(\frac{da_{ij}(t)}{dt}\right)_{1\le i\le m, 1\le j\le n}$$
$$\int_0^t \boldsymbol{A}(\tau)d\tau = \left(\int_0^t a_{ij}(\tau)d\tau\right)_{1\le i\le m, 1\le j\le n}$$
と定義する.

行列値関数の微分と積分 (2)

- 各時刻 t における $\boldsymbol{A}(t)$ の逆行列 $\boldsymbol{A}^{-1}(t)$ が定義され, かつ微分可能であると仮定する．このとき, $\boldsymbol{A}(t)\boldsymbol{A}^{-1}(t) = \boldsymbol{I}$ であるから, $\dfrac{d\boldsymbol{A}(t)}{dt}\boldsymbol{A}^{-1}(t)$
$\boldsymbol{A}(t)\dfrac{d\boldsymbol{A}^{-1}(t)}{dt} = \boldsymbol{0}$, よって
$\dfrac{d\boldsymbol{A}^{-1}(t)}{dt} = -\boldsymbol{A}^{-1}(t)\dfrac{d\boldsymbol{A}(t)}{dt}\boldsymbol{A}^{-1}(t)$ である．

多変数関数の微分 (1)

- システム工学では列ベクトルと行ベクトルを使い分けることが普通.
- $f : \mathbb{R}^n \ni \boldsymbol{x} \mapsto f(\boldsymbol{x}) \in \mathbb{R}$ が偏微分可能な実数値関数であるとき, $\dfrac{\partial f}{\partial \boldsymbol{x}}$ は, 通常は, $\begin{pmatrix} \dfrac{\partial f}{\partial x_1} & \cdots & \dfrac{\partial f}{\partial x_n} \end{pmatrix}$ という行ベクトルとして定義される.

多変数関数の微分 (2)

- $\boldsymbol{f} : \mathbb{R}^n \ni \boldsymbol{x} \mapsto \boldsymbol{f}(\boldsymbol{x}) \in \mathbb{R}^m$ が偏微分可能なベクトル値関数であるとき，$\dfrac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}$ は，通常は，
$$\begin{pmatrix} \dfrac{\partial f_1}{\partial x_1} & \cdots & \dfrac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \dfrac{\partial f_m}{\partial x_1} & \cdots & \dfrac{\partial f_m}{\partial x_n} \end{pmatrix}$$
という行列として定義される．

多変数関数の微分 (3)

- $\dfrac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}$ を $\boldsymbol{f}$ の Jacobi 行列あるいは Jacobian と呼ぶ.
- $\dfrac{\partial f}{\partial \boldsymbol{x}}$ と $\dfrac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}$ に定義が整合していたことに注意する: すなわち, $\dfrac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}$ は, $\dfrac{\partial f_1}{\partial \boldsymbol{x}}$ から $\dfrac{\partial f_m}{\partial \boldsymbol{x}}$ までを縦にならべたものになっている.
- ここまでは良いのだが $\cdots$

多変数関数の微分 (4)

- f の偏微分を ∇f と書く流儀もある. ∇f は, 文献によって, 行ベクトルとして定義されていることもあれば, 列ベクトルとして定義されていることもある.

多変数関数の微分 (5)

- 明示的には書かれていないが，教科書では，∇f を列ベクトルと定義しているらしい．
- 教科書では，$\dfrac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}} = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1} & \cdots & \dfrac{\partial f_m}{\partial x_1} \\ \vdots & & \vdots \\ \dfrac{\partial f_1}{\partial x_n} & \cdots & \dfrac{\partial f_m}{\partial x_n} \end{pmatrix}$ と定義している．この定義は ∇f を列ベクトルと定義する流儀と整合性があるが，少数派．

多変数関数の微分 (6)

- この講義資料の定義を使うと $\dfrac{\partial}{\partial \boldsymbol{x}}(\boldsymbol{A}\boldsymbol{x}) = \boldsymbol{A}$
- 教科書の定義を使うと $\dfrac{\partial}{\partial \boldsymbol{x}}(\boldsymbol{A}\boldsymbol{x}) = \boldsymbol{A}^T$
- 多変数関数の微分では, ベクトルの型にバリエーションがあり, 混乱が発生しやすいので, 注意すること.

参考文献

- 笠原, 線形代数と固有値問題, 増補版, 現代数学社, 2005
- 伊理, 線形代数汎論, 朝倉書店, 2009
- D. S. Bernstein, Matrix Mathematics, 2/e, Princeton University Press, 2009
- I. M. Gelfand (Translated by A. Shenitzer, Lectures on linear algebra, Dover, 1989
- 川久保, 線形代数学, 日本評論社, 1999
- 笠原, 微分方程式の基礎, 朝倉書店, 1982