# 線形代数 II 第 6 回講義資料

担当：大矢 浩徳 (OYA Hironori)*

前回までの講義資料で一般の $n$ 次正方行列の対角化可能性の判定および対角化の手順について，理論的な側面も含め一通り解説を行った．次回以降，実対称行列，およびエルミート行列と呼ばれる特別な行列の対角化について，さらに詳細なことを解説する．今回はそこへ向かう前の (大事な) 寄り道である．まず，前回の証明において行った考察の応用として，任意の(複素)$n$ 次正方行列 $A$ に対して，複素 $n$ 次正方行列 $P$ が存在し，$P^{-1}AP$ が上三角行列になるという事実を証明する (**三角化可能性**)．対角化は可能な場合と不可能な場合があるが，三角化はいつでも可能なのである．さらにこの事実の応用として，**ケイリー・ハミルトンの定理 (Cayley–Hamilton theorem)** と呼ばれる定理を証明する．

## 6.1 正方行列の三角化とケイリー・ハミルトンの定理

本節では正方行列の三角化とその応用としてのケイリー・ハミルトンの定理について述べよう．まず，三角化とは以下のようなものである．

**定義 6.1**

$A$ を $n$ 次正方行列とする．ある正則な $n$ 次正方行列 $P$ が存在して

$$P^{-1}AP$$

が上三角行列となるとき，**$A$ は $P$ によって三角化される**という．ここで，上三角行列とは

$$\begin{pmatrix} c_1 & * & * & \cdots & * \\ 0 & c_2 & * & \cdots & * \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & * \\ 0 & 0 & \cdots & 0 & c_n \end{pmatrix}$$

という形の行列 ($*$ には 0 とは限らない $\mathbb{K}$ の元が入っている) で，特に対角成分 $c_1, c_2, \ldots, c_n$ は 0 とは限らないものを考える．($c_1 = c_2 = \cdots = c_n = 0$ となる上三角行列は**狭義三角行列**と呼ばれる．)

以降本節の中では $\mathbb{K} = \mathbb{C}$ とする (理由は後に述べる)．すでに見てきたように，$n$ 次正方行列 $A$ には対角化可能なものと対角化不可能なものが存在したが，実は**三角化は (複素行列を用いれば) 常に可能**である．

* *e-mail*：hoya@shibaura-it.ac.jp

**定理 6.2**

$A$ を $n$ 次正方行列とする．このとき，$A$ はある複素正方行列 $P$ によって三角化される．さらに，三角化の結果が

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & * & * & \cdots & * \\ 0 & \lambda_2 & * & \cdots & * \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & * \\ 0 & 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

という形になるならば，$A$ の固有多項式 $\Phi_A(t)$ は

$$\Phi_A(t) = (t-\lambda_1)\cdots(t-\lambda_n)$$

と因数分解される．ただし，$\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ には重複もあり得る．すなわち，対角成分には必ず $A$ の各固有値が重複度の数だけ現れる．

**証明.** 行列のサイズ $n$ に関する帰納法で三角化可能性を示す．まず，$n=1$ のときは任意の $1\times 1$ 行列 $(a)$ は上の定義で上三角行列とみなせるので，主張は自明である．次に任意の $n$ 次正方行列が三角化可能であったと仮定して，$n+1$ 次正方行列 $A$ の三角化可能性を示す $(n \geq 1)$．このとき，$A$ の固有方程式 $\Phi_A(t) = 0$ は未知数 $t$ に関する $n+1$ 次方程式なので，必ずある複素数解 $\lambda$ をもつ．(代数学の基本定理．ここで $\mathbb{K} = \mathbb{C}$ を用いた！．実数の範囲では一般には解を持たないかもしれない．) すると，$A$ は固有値 $\lambda$ の固有ベクトルを持つのでそれを $\boldsymbol{p}_1$ とする．このとき，定理 5.1 より，ある $\boldsymbol{p}_2, \ldots, \boldsymbol{p}_{n+1} \in \mathbb{K}^{n+1}$ が存在して，$\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_{n+1}$ が $\mathbb{K}^{n+1}$ の基底となるようにできる ($\boldsymbol{p}_2, \ldots, \boldsymbol{p}_{n+1} \in \mathbb{K}^{n+1}$ は $A$ の固有ベクトルとは限らない)．すると系 4.5 より，

$$\tilde{P} = (\boldsymbol{p}_1\ \boldsymbol{p}_2\ \cdots\ \boldsymbol{p}_{n+1})$$

は正則である．行列の積の定義より，$\tilde{P}\boldsymbol{e}_1 = \boldsymbol{p}_1$ なので，両辺に $\tilde{P}^{-1}$ を左から掛けて $\boldsymbol{e}_1 = \tilde{P}^{-1}\boldsymbol{p}_1$ となることに注意すると，

$$\begin{aligned} \tilde{P}^{-1}A\tilde{P} &= \tilde{P}^{-1}(A\boldsymbol{p}_1\ A\boldsymbol{p}_2\ \cdots\ A\boldsymbol{p}_{n+1}) \\ &= \tilde{P}^{-1}(\lambda\boldsymbol{p}_1\ A\boldsymbol{p}_2\ \cdots\ A\boldsymbol{p}_{n+1}) \quad (\boldsymbol{p}_1 \text{は } A \text{ の固有値} \lambda \text{の固有ベクトルなので}) \\ &= (\lambda\tilde{P}^{-1}\boldsymbol{p}_1\ \tilde{P}^{-1}A\boldsymbol{p}_2\ \cdots\ \tilde{P}^{-1}A\boldsymbol{p}_{n+1}) \\ &= (\lambda\boldsymbol{e}_1\ \tilde{P}^{-1}A\boldsymbol{p}_2\ \cdots\ \tilde{P}^{-1}A\boldsymbol{p}_{n+1}) \\ &= \left(\begin{array}{c|c} \lambda & \boldsymbol{b} \\ \hline \boldsymbol{0} & B \end{array}\right) \end{aligned}$$

という形になる．ここで，$\boldsymbol{b}$ はある $1 \times n$ 行列，$B$ はある $n$ 次正方行列，$\boldsymbol{0}$ は $n$ 次ゼロベクトルである[*1]．ここで，帰納法の仮定より，ある $n$ 次正則行列 $Q$ が存在して，

$$Q^{-1}BQ = \begin{pmatrix} \lambda_1 & * & * & \cdots & * \\ 0 & \lambda_2 & * & \cdots & * \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & * \\ 0 & 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

とできる．この $Q$ を用いて，$n+1$ 次正方行列を

$$\tilde{Q} \coloneqq \left(\begin{array}{c|c} 1 & {}^t\boldsymbol{0} \\ \hline \boldsymbol{0} & Q \end{array}\right)$$

[*1] ここでの計算は第 5 回講義資料 5.2.2 節の事実 4.1 (1) の証明で用いた手法と同様である．

とすると，$\tilde{Q}$ は正則で，

$$\tilde{Q}^{-1} = \left(\begin{array}{c|c} 1 & {}^t\mathbf{0} \\ \hline \mathbf{0} & Q^{-1} \end{array}\right).$$

さらに，

$$P \coloneqq \tilde{P}\tilde{Q}$$

とすると，$P$ も正則な $n+1$ 次正方行列である．これを用いると，

$$\begin{aligned}
P^{-1}AP &= \tilde{Q}^{-1}\tilde{P}^{-1}A\tilde{P}\tilde{Q} \\
&= \tilde{Q}^{-1}\left(\begin{array}{c|c} \lambda & \boldsymbol{b} \\ \hline \mathbf{0} & B \end{array}\right)\tilde{Q} \\
&= \left(\begin{array}{c|c} 1 & {}^t\mathbf{0} \\ \hline \mathbf{0} & Q^{-1} \end{array}\right)\left(\begin{array}{c|c} \lambda & \boldsymbol{b} \\ \hline \mathbf{0} & B \end{array}\right)\left(\begin{array}{c|c} 1 & {}^t\mathbf{0} \\ \hline \mathbf{0} & Q \end{array}\right) \\
&= \left(\begin{array}{c|c} \lambda & \boldsymbol{b}Q \\ \hline \mathbf{0} & Q^{-1}BQ \end{array}\right) \\
&= \left(\begin{array}{c|ccccc} \lambda & & & \boldsymbol{b}Q & & \\ \hline 0 & \lambda_1 & * & * & \cdots & * \\ 0 & 0 & \lambda_2 & * & \cdots & * \\ 0 & 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & * \\ 0 & 0 & 0 & \cdots & 0 & \lambda_n \end{array}\right)
\end{aligned}$$

と三角化できる．以上より，三角化可能性は示された．

後半の主張については，

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & * & * & \cdots & * \\ 0 & \lambda_2 & * & \cdots & * \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & * \\ 0 & 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

となるとき，命題 3.3 (4) より，

$$\Phi_A(t) = \Phi_{P^{-1}AP}(t) = |tI_n - P^{-1}AP| = \left|\begin{pmatrix} t-\lambda_1 & *' & *' & \cdots & *' \\ 0 & t-\lambda_2 & *' & \cdots & *' \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & *' \\ 0 & 0 & \cdots & 0 & t-\lambda_n \end{pmatrix}\right| = (t-\lambda_1)\cdots(t-\lambda_n)$$

となることからわかる． □

三角化可能性からケイリー・ハミルトンの定理と呼ばれる重要な定理が証明される．

**定理 6.3(ケイリー・ハミルトンの定理 (Cayley–Hamilton theorem))**

$A$ を $n$ 次正方行列とし，$\Phi_A(t)$ をその固有多項式とする．このとき，

$$\Phi_A(A) = O.$$

ここで，$\Phi_A(A)$ は固有多項式の変数 $t$ に行列 $A$ を代入して計算して得られる $n$ 次正方行列 (ただし定数項 $(-1)^n|A|$ は $(-1)^n|A|I_n$ に置き換える)，$O$ は $n$ 次正方ゼロ行列を表す．

**証明.** 定理 6.2 より，$A$ はある複素正方行列 $P$ によって，

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & * & * & \cdots & * \\ 0 & \lambda_2 & * & \cdots & * \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & * \\ 0 & 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

という形に三角化される．このとき，$\Phi_A(t) = (t-\lambda_1)\cdots(t-\lambda_n)$ となるので，示すべきことは

$$(A-\lambda_1 I_n)(A-\lambda_2 I_n)\cdots(A-\lambda_n I_n) = O$$

である．ここで，

$$\begin{aligned} &(A-\lambda_1 I_n)(A-\lambda_2 I_n)\cdots(A-\lambda_n I_n) = O \\ \Leftrightarrow\ & P^{-1}(A-\lambda_1 I_n)(A-\lambda_2 I_n)\cdots(A-\lambda_n I_n)P = O \\ \Leftrightarrow\ & P^{-1}(A-\lambda_1 I_n)PP^{-1}(A-\lambda_2 I_n)P\cdots P^{-1}(A-\lambda_n I_n)P = O \\ \Leftrightarrow\ & (P^{-1}AP-\lambda_1 I_n)\cdots(P^{-1}AP-\lambda_n I_n) = O \end{aligned}$$

という同値関係が成り立つので，最後の式を示せば良い．ここで，最後の式の左辺の積

$$\begin{pmatrix} 0 & * & * & \cdots & * \\ 0 & \lambda_2-\lambda_1 & * & \cdots & * \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & * \\ 0 & 0 & \cdots & 0 & \lambda_n-\lambda_1 \end{pmatrix} \begin{pmatrix} \lambda_1-\lambda_2 & * & * & \cdots & * \\ 0 & 0 & * & \cdots & * \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & * \\ 0 & 0 & \cdots & 0 & \lambda_n-\lambda_2 \end{pmatrix} \cdots \begin{pmatrix} \lambda_1-\lambda_n & * & * & \cdots & * \\ 0 & \lambda_2-\lambda_n & * & \cdots & * \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & * \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix}$$

を左から順に具体的に計算すると $O$ となることが確かめられる．[*2] これより，示すべきことは示された． □

注意 1. この節では $\mathbb{K} = \mathbb{C}$ とすると仮定したが，実数は複素数の一部なので，ケイリー・ハミルトンの定理はもちろん実行列に対しても成立する．

**例 1.** 2 次正方行列 $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ の固有多項式は，

$$\Phi_A(t) = |tI_2 - A| = \left|\begin{pmatrix} t-a & -b \\ -c & t-d \end{pmatrix}\right| = t^2 - (a+d)t + (ad-bc)$$

となるので，ケイリー・ハミルトンの定理より，

$$A^2 - (a+d)A + (ad-bc)I_2 = O.^{*3}$$

---

[*2] 状況のわかりやすさのため，$n=6$ の場合に様子を見てここの計算を “納得” してみよう．

$$\begin{pmatrix} 0 & * & * & * & * & * \\ 0 & \lambda_2-\lambda_1 & * & * & * & * \\ 0 & 0 & \lambda_3-\lambda_1 & * & * & * \\ 0 & 0 & 0 & \lambda_4-\lambda_1 & * & * \\ 0 & 0 & 0 & 0 & \lambda_5-\lambda_1 & * \\ 0 & 0 & 0 & 0 & 0 & \lambda_6-\lambda_1 \end{pmatrix} \begin{pmatrix} \lambda_1-\lambda_2 & * & * & * & * & * \\ 0 & 0 & * & * & * & * \\ 0 & 0 & \lambda_3-\lambda_2 & * & * & * \\ 0 & 0 & 0 & \lambda_4-\lambda_2 & * & * \\ 0 & 0 & 0 & 0 & \lambda_5-\lambda_2 & * \\ 0 & 0 & 0 & 0 & 0 & \lambda_6-\lambda_2 \end{pmatrix} = \begin{pmatrix} 0 & 0 & * & * & * & * \\ 0 & 0 & * & * & * & * \\ 0 & 0 & * & * & * & * \\ 0 & 0 & 0 & * & * & * \\ 0 & 0 & 0 & 0 & * & * \\ 0 & 0 & 0 & 0 & 0 & * \end{pmatrix}$$

$$\begin{pmatrix} 0 & 0 & * & * & * & * \\ 0 & 0 & * & * & * & * \\ 0 & 0 & * & * & * & * \\ 0 & 0 & 0 & * & * & * \\ 0 & 0 & 0 & 0 & * & * \\ 0 & 0 & 0 & 0 & 0 & * \end{pmatrix} \begin{pmatrix} \lambda_1-\lambda_3 & * & * & * & * & * \\ 0 & \lambda_2-\lambda_3 & * & * & * & * \\ 0 & 0 & 0 & * & * & * \\ 0 & 0 & 0 & \lambda_4-\lambda_3 & * & * \\ 0 & 0 & 0 & 0 & \lambda_5-\lambda_3 & * \\ 0 & 0 & 0 & 0 & 0 & \lambda_6-\lambda_3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & * & * & * \\ 0 & 0 & 0 & * & * & * \\ 0 & 0 & 0 & * & * & * \\ 0 & 0 & 0 & * & * & * \\ 0 & 0 & 0 & 0 & * & * \\ 0 & 0 & 0 & 0 & 0 & * \end{pmatrix}$$

$$\begin{pmatrix} 0 & 0 & 0 & * & * & * \\ 0 & 0 & 0 & * & * & * \\ 0 & 0 & 0 & * & * & * \\ 0 & 0 & 0 & * & * & * \\ 0 & 0 & 0 & 0 & * & * \\ 0 & 0 & 0 & 0 & 0 & * \end{pmatrix} \begin{pmatrix} \lambda_1-\lambda_4 & * & * & * & * & * \\ 0 & \lambda_2-\lambda_4 & * & * & * & * \\ 0 & 0 & \lambda_3-\lambda_4 & * & * & * \\ 0 & 0 & 0 & 0 & * & * \\ 0 & 0 & 0 & 0 & \lambda_5-\lambda_4 & * \\ 0 & 0 & 0 & 0 & 0 & \lambda_6-\lambda_4 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 & * & * \\ 0 & 0 & 0 & 0 & * & * \\ 0 & 0 & 0 & 0 & * & * \\ 0 & 0 & 0 & 0 & * & * \\ 0 & 0 & 0 & 0 & * & * \\ 0 & 0 & 0 & 0 & 0 & * \end{pmatrix}$$

$$\begin{pmatrix} 0 & 0 & 0 & 0 & * & * \\ 0 & 0 & 0 & 0 & * & * \\ 0 & 0 & 0 & 0 & * & * \\ 0 & 0 & 0 & 0 & * & * \\ 0 & 0 & 0 & 0 & * & * \\ 0 & 0 & 0 & 0 & 0 & * \end{pmatrix} \begin{pmatrix} \lambda_1-\lambda_5 & * & * & * & * & * \\ 0 & \lambda_2-\lambda_5 & * & * & * & * \\ 0 & 0 & \lambda_3-\lambda_5 & * & * & * \\ 0 & 0 & 0 & \lambda_4-\lambda_5 & * & * \\ 0 & 0 & 0 & 0 & 0 & * \\ 0 & 0 & 0 & 0 & 0 & \lambda_6-\lambda_5 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & * \\ 0 & 0 & 0 & 0 & 0 & * \\ 0 & 0 & 0 & 0 & 0 & * \\ 0 & 0 & 0 & 0 & 0 & * \\ 0 & 0 & 0 & 0 & 0 & * \\ 0 & 0 & 0 & 0 & 0 & * \end{pmatrix}$$

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & * \\ 0 & 0 & 0 & 0 & 0 & * \\ 0 & 0 & 0 & 0 & 0 & * \\ 0 & 0 & 0 & 0 & 0 & * \\ 0 & 0 & 0 & 0 & 0 & * \\ 0 & 0 & 0 & 0 & 0 & * \end{pmatrix} \begin{pmatrix} \lambda_1-\lambda_6 & * & * & * & * & * \\ 0 & \lambda_2-\lambda_6 & * & * & * & * \\ 0 & 0 & \lambda_3-\lambda_6 & * & * & * \\ 0 & 0 & 0 & \lambda_4-\lambda_6 & * & * \\ 0 & 0 & 0 & 0 & \lambda_5-\lambda_6 & * \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} = O.$$

このように “一つ掛けるごとにゼロベクトルの列が一つ増えていく” という形になり，最終的に $O$ になる．一般的な形の証明を書くことにも是非チャレンジしてもらいたい．

[*3] 直接検算してみると確かに正しいことがわかるだろう．

例えば，

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

のとき，

$$A^2 - 5A - 2I_2 = O,\ \text{すなわち，} A^2 = 5A + 2I_2 = \begin{pmatrix} 7 & 10 \\ 15 & 22 \end{pmatrix}$$

が確かに成立している．

**例 2.** 3 次正方行列 $A = \begin{pmatrix} -3 & 4 & 1 \\ -2 & 2 & 2 \\ 1 & 0 & 1 \end{pmatrix}$ の固有多項式は，

$$\begin{aligned}
\Phi_A(t) = |tI_3 - A| &= \left| \begin{pmatrix} t+3 & -4 & -1 \\ 2 & t-2 & -2 \\ -1 & 0 & t-1 \end{pmatrix} \right| \\
&= \left| \begin{pmatrix} t+3 & -4 & -1 \\ 0 & t-2 & 2t-4 \\ -1 & 0 & t-1 \end{pmatrix} \right| \quad (\text{第 2 行に第 3 列の 2 倍を加えた}) \\
&= \left| \begin{pmatrix} t+3 & -4 & 7 \\ 0 & t-2 & 0 \\ -1 & 0 & t-1 \end{pmatrix} \right| \quad (\text{第 3 列に第 2 列の} -2 \text{倍を加えた}) \\
&= (t-2) \left| \begin{pmatrix} t+3 & 7 \\ -1 & t-1 \end{pmatrix} \right| = (t-2)(t^2 + 2t + 4) = t^3 - 8
\end{aligned}$$

となるので，ケイリー・ハミルトンの定理より，

$$A^3 - 8I_3 = O,\ \text{すなわち，} A^3 = \begin{pmatrix} 8 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 8 \end{pmatrix}$$

となることがわかる．これを用いれば，例えば

$$A^{100} = (A^3)^{33} A = (8^{33} I_n) A = 2^{99} \cdot \begin{pmatrix} -3 & 4 & 1 \\ -2 & 2 & 2 \\ 1 & 0 & 1 \end{pmatrix}$$

というような計算ができる．