線型幾何 -スペクトル-

いろはの学習備忘録
https://168iroha.net

初回更新 2019 年 9 月 4 日
最終更新 2019 年 9 月 4 日

目次

第 1 章

正方行列のスペクトル

1.1　固有値・固有ベクトル

線型写像 $f:\mathbb{R}^n \to \mathbb{R}^n$ における変換後でベクトルの方向が不変となるベクトルを考える．このとき f の表現行列を A とすれば，$\boldsymbol{x} \in \mathbb{R}^n$ と $\lambda \in \mathbb{R}$ で

$$A\boldsymbol{x} = \lambda\boldsymbol{x}$$

が成り立つことであるが，

$$(A-\lambda I)\boldsymbol{x} = \boldsymbol{0}$$

とすれば，この連立方程式が $\boldsymbol{x} = \boldsymbol{0}$ 以外の解を持つ，すなわち $A-\lambda I$ が正則とならない λ を選べばいい．つまり，行列が正則であることの必要十分条件から

$$\det(A-\lambda I) = 0$$

の λ に関する代数方程式が得られ，λ が得られる．また，この代数方程式は明らかに n 次方程式であるため，代数学の基本定理より複素数の範囲で n 個の解 λ が存在し，λ に対応する $\boldsymbol{x}$ を求めるができる．

これより，以下の定義を与えることができる．

定義 1.1

n 次の正方行列 A に対して，$\lambda \in \mathbb{R}$ に対する代数方程式

$$p(\lambda) = \det(A-\lambda I) = 0$$

の解を A の**固有値 (eigenvalue)** といい，

$$(A-\lambda I)\boldsymbol{x} = \boldsymbol{0}$$

を満たす $\boldsymbol{0}$ ではないベクトル $\boldsymbol{x} \in \mathbb{R}^n$ を**固有ベクトル (eigenvector)** といい，これらを求める問題を**固有値問題 (eigenvalue problem)** という．また，この $p(\lambda) = 0$ のことを**固有方程式**といい，その左辺を**固有多項式 (characteristic polynomial)** という．

固有値は複素数値をとることもあるため，一般には複素数値を成分にもつ行列を考える必要がある．そこで，定義を与える．

定義 1.2

$m \times n$ 複素行列 A の全ての成分の共役をとり転置した行列を**エルミート転置 (Hermitian transpose)** もしくは**随伴行列 (adjoint matrix)** といい，A^* とあらわす．特に，$A = A^*$ である行列を**エルミート行列 (Hermitian matrix)** といい，$A^* = -A$ である行列を**歪エルミート行列 (Skew-Hermitian matrix)** もしくは**反エルミート行列 (Anti-Hermitian matrix)** という．また，n 次の複素正方行列 U で単位行列 I を用いて

$$U^*U = UU^* = I$$

となるとき，U を**ユニタリ行列 (Unitary matrix)** という．

この定義より，転置行列や交代行列における性質はエルミート行列や歪エルミート行列にただちに拡張されることがわかる．

補題 1.1　n 次の正方行列 A の相異なる固有値に対する固有ベクトルはそれぞれ線型独立である．

Proof.

A の相異なる固有値 $\lambda_1, \lambda_2 \in \mathbb{C}$ に対する固有ベクトル $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{C}^n$ で $\boldsymbol{x}_1$ と $\boldsymbol{x}_2$ が平行であると仮定すれば $k \in \mathbb{C}$ を用いて

$$\boldsymbol{x}_2 = k\boldsymbol{x}_1.$$

これより，

$$\begin{aligned} A\boldsymbol{x}_2 &= kA\boldsymbol{x}_1 \\ \Rightarrow \lambda_2\boldsymbol{x}_2 &= k\lambda_1\boldsymbol{x}_1 \\ \Rightarrow k\lambda_2\boldsymbol{x}_1 &= k\lambda_1\boldsymbol{x}_1 \\ \Rightarrow (\lambda_2 - \lambda_1)\boldsymbol{x}_1 &= \boldsymbol{0} \end{aligned}$$

固有ベクトルは零ベクトルでないことから $k \neq 0$ であるため

$$\lambda_2 - \lambda_1 = 0 \iff \lambda_1 = \lambda_2$$

となり，$\lambda_1 \neq \lambda_2$ の仮定に反する．

よって，命題は証明された．

□

補題 1.2　エルミート行列の固有値は実数である．

Proof.

エルミート行列 A の固有値を λ，λ に対応する固有ベクトルを $\boldsymbol{x} \in \mathbb{C}^n$ とすれば，固有値の定義より

$$A\boldsymbol{x} = \lambda\boldsymbol{x}, \quad \boldsymbol{x}^*A^* = \boldsymbol{x}^*A = \lambda^*\boldsymbol{x}^*.$$

この 2 式の両辺に $\boldsymbol{x}$ および $\boldsymbol{x}^*$ を作用させれば

$$\boldsymbol{x}^*A\boldsymbol{x} = \lambda\boldsymbol{x}^*\boldsymbol{x}, \quad \boldsymbol{x}^*A\boldsymbol{x} = \lambda^*\boldsymbol{x}^*\boldsymbol{x}$$

となるため，これを連立することにより

$$(\lambda - \lambda^*)\boldsymbol{x}^*\boldsymbol{x} = 0$$

となり，固有ベクトルは零ベクトルではないため

$$\lambda - \lambda^* = 0 \iff \lambda = \lambda^*.$$

よって，命題は証明された．

□

この補題により，実数における対称行列を考えれば固有値も実数となるため，実数における対称行列の固有値問題は実数のみで考えることができる．また，対称行列の固有ベクトルの著しい特性として以下の補題が与えられる．

補題 1.3　エルミート行列の相異なる固有値に対する固有ベクトルは直交する．

Proof.

補題 1.2 の証明と同様にして，エルミート行列 A の相異なる固有値 $\lambda_1, \lambda_2 \in \mathbb{R}$ に対する固有ベクトル $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{C}^n$ で

$$\boldsymbol{x}_2^*A\boldsymbol{x}_1 = \lambda_1\boldsymbol{x}_2^*\boldsymbol{x}_1, \quad \boldsymbol{x}_2^*A\boldsymbol{x}_1 = \lambda_2\boldsymbol{x}_2^*\boldsymbol{x}_1.$$

これを連立することにより

$$(\lambda_1 - \lambda_2^*)\boldsymbol{x}_2^*\boldsymbol{x}_1 = 0$$

となるが，仮定より $\lambda_1 \neq \lambda_2$ であるため $\boldsymbol{x}_2^*\boldsymbol{x}_1 = 0$ となる．

よって，命題は証明された．

□

補題 1.3 はエルミート行列に限定して考えたが，その証明中より正方行列 A と A の固有値 λ に対する固有ベクトル $\boldsymbol{x}$ で

$$\boldsymbol{x}^*A = \lambda\boldsymbol{x}^*$$

を満たす，すなわち

$$A^*\boldsymbol{x} = \lambda^*\boldsymbol{x}$$

を満たすという性質に集約される．

そこで，以下の補題を与える．

補題 1.4　正方行列 A が $AA^* = A^*A$ を満たすなら，A の固有値 λ に対する固有ベクトル $\boldsymbol{x}$ で

$$A^*\boldsymbol{x} = \lambda^*\boldsymbol{x}.$$

Proof.

ベクトル $A^*\boldsymbol{x} - \lambda^*\boldsymbol{x}$ の絶対値の 2 乗について

$$\begin{aligned}(A^*\boldsymbol{x} - \lambda^*\boldsymbol{x})^*(A^*\boldsymbol{x} - \lambda^*\boldsymbol{x}) =& \boldsymbol{x}^*(A - \lambda)(A^* - \lambda^*)\boldsymbol{x} \\ =& \boldsymbol{x}^*(AA^* - \lambda^*A - \lambda A^* + \lambda|^2)\boldsymbol{x} \\ =& \boldsymbol{x}^*(A^*A - \lambda^*A - \lambda A^* + |\lambda|^2)\boldsymbol{x} \\ =& \boldsymbol{x}^*(A^*\lambda\boldsymbol{x} - \lambda^*\lambda\boldsymbol{x} - \lambda A^*\boldsymbol{x} + \lambda|^2\boldsymbol{x})\boldsymbol{x} \\ =& 0\end{aligned}$$

となるため

$$A^*\boldsymbol{x} - \lambda^*\boldsymbol{x} = \boldsymbol{0} \;\Leftrightarrow\; A^*\boldsymbol{x} = \lambda^*\boldsymbol{x}.$$

よって，命題は証明された．

□

これより，以下の定義が与えられる．

定義 1.3

正方行列 A で $AA^* = A^*A$ を満たす行列を**正規行列 (normal matrix)** という．

正規行列はその定義よりエルミート行列を含み，補題 1.3 は正規行列について拡張されることがわかる．

また，固有値に関する性質として以下の定理が与えられる．

定理 1.1

n 次の正方行列 A の重複を含めた固有値の総和は A の対角成分の総和に等しい．

Proof.

A の固有方程式 $p(x) \in \mathbb{C}[x]$ を展開して x についての降べきの順で書き下すとき，行列式の定義より最初の 2 項に関しては A の対角成分の積にのみ依存する．つまり，$p(x)$ の最初の 2 項は A の対角成分の総和を S とすることで

$$x^n - Sx^{n-1}.$$

また，x に関する n 次の代数方程式の解の総和は x^n の係数を a_1，x^{n-1} の係数を a_2 としたとき $-\frac{a_2}{a_1}$ となるため，$p(x)$ の解の総和，すなわち重複を含めた A の固有値の総和は S となる．

よって，命題は証明された．

□

この定理より，正方行列における対格成分の総和は正方行列の重要な指標の 1 つであるといえる．これより，以下の定義が与えられる．

定義 1.4

正方行列 A の対格成分の和を行列における A の**トレース (trace)** といい，$\operatorname{tr} A$ とあらわす．

定理 1.2

正方行列 A, B のトレースの性質について以下が成り立つ．

1). $k \in \mathbb{C}$ で $\operatorname{tr}(kA) = k \cdot \operatorname{tr} A$.

2). $\operatorname{tr}(A+B) = \operatorname{tr} A + \operatorname{tr} B$.

3). $\operatorname{tr} A^T = \operatorname{tr} A$.

4). $\operatorname{tr}(AB) = \operatorname{tr}(BA)$.

Proof.

1) と 2) と 3) についてはトレースの定義より自明である．$A = (a_{i,j}), B = (b_{i,j})$ とすれば

$$AB = \left(\sum_j a_{i,j} b_{j,k}\right), \quad BA = \left(\sum_j b_{i,j} a_{j,k}\right)$$

となるため

$$\operatorname{tr}(AB) = \sum_i \sum_j a_{i,j} b_{j,i}, \quad \operatorname{tr}(BA) = \sum_i \sum_j b_{i,j} a_{j,i}$$

となるが，この総和における i, j はそれぞれ独立であるため $\operatorname{tr}(AB) = \operatorname{tr}(BA)$ となる．

よって，命題は証明された．

□

1.2　対角化

n 次の正則行列 A の重複を含んだ固有値 $\lambda_1, \lambda_2, \cdots, \lambda_n \in \mathbb{C}$ とそれに対応する固有ベクトル $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n \in \mathbb{C}^n$ を与えたとする．固有ベクトルを列ベクトルとして構成される行列を P とすれば，P は正則であるため

$$\begin{pmatrix} A\boldsymbol{x}_1 & A\boldsymbol{x}_2 & \cdots & A\boldsymbol{x}_n \end{pmatrix} = \begin{pmatrix} \lambda_1\boldsymbol{x}_1 & \lambda_2\boldsymbol{x}_2 & \cdots & \lambda_n\boldsymbol{x}_n \end{pmatrix}$$

$$A\begin{pmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \cdots & \boldsymbol{x}_n \end{pmatrix} = \begin{pmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \cdots & \boldsymbol{x}_n \end{pmatrix} \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}$$

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

これより，以下の定義が与えられる．

定義 1.5

n 次の正方行列 A と重複を含んだ固有値 $\lambda_1, \lambda_2, \cdots, \lambda_n \in \mathbb{C}$ と正則行列 P により

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

と与えることができるとき，このことを A の**対角化 (diagonalization)** といい，P を**対角化行列**という．また，A の対角化が存在するとき A は**対角化可能**であるという．

定理 1.3

n 次の正方行列 A が対角化可能な必要十分条件は A に n 本の線型独立な固有ベクトルが存在することである．

Proof.

A に n 本の線型独立な固有ベクトルが存在するときは対角化の構成により対角化可能であることは明らかである．

次に，対角化可能であるときに A に n 本の線型独立な固有ベクトルが存在することを示す．代数学の基本定理より A の固有値は重複を含んで n 個だけ存在するため，それぞれ異なる固有値を $\lambda_1, \lambda_2, \cdots, \lambda_r \in \mathbb{C}$ としてそれぞれの重解である個数を $n_1, n_2, \cdots, n_r \in \mathbb{N}$ としたときに，固有値 λ_k に対して線型独立な固有ベクトルが n_k 本生成可能であることを示せば，定理 1.1 により n 本の線型独立な固有ベクトルが存在することとなる．任意の行列は正則行列との積で階数は不変であるため，A が正則行列 P と対角行列 D により

$$P^{-1}AP = D$$

と対角化されるとすれば

$$\begin{aligned} \mathrm{rank}\,(A - \lambda_k I) =& \mathrm{rank}\,(P^{-1}(A - \lambda_k I)P) \\ =& \mathrm{rank}\,(P^{-1}AP - \lambda_k I) \\ =& \mathrm{rank}\,(D - \lambda_k I) \end{aligned}$$

となり，$D - \lambda_k I$ は n_k だけ対角成分が 0 の対角行列であるため

$$\mathrm{rank}\,(A - \lambda_k I) = n - n_k.$$

次元定理より

$$\mathrm{rank}\,(A - \lambda_k I) + \dim \mathrm{Ker}\,(A - \lambda_k I) = n \;\Leftrightarrow\; \dim \mathrm{Ker}\,(A - \lambda_k I) = n_k$$

となり，$\mathrm{Ker}\,(A - \lambda_k I)$ は λ_k に対する固有ベクトル全体の集合であるため，λ_k に対する固有ベクトルの次元は n_k，すなわち λ_k に対する線型独立な固有ベクトルは n_k 本生成可能である．

よって，命題は証明された．

□

これより，以下の系が得られる．

系 1.3-1　正方行列 A が対角化可能で正則行列 P と対角行列 D により $P^{-1}AP = D$ となるとき，D は固有値の並び順を無視すれば一意である．

Proof.

定理 1.3 の証明中より明らか．

□

系 1.3-2　正方行列 A が正則行列 P と対角行列 D により $P^{-1}AP = D$ となるとき，D の対角成分は A の重複を含めた全ての固有値であり，A は対角化可能である．

Proof.

定理 1.3 の証明中および対角化の構成から明らか.

□

なお，n 次の正方行列 A が n 本の固有ベクトルをもつ条件は A が正則であることとは関係ないことに留意する必要がある．例えば

$$\begin{pmatrix} 2 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

は正則であるが，固有値 $\lambda = 2$ の 3 重解で $\lambda = 2$ に対する固有ベクトルは定数 $c \in \mathbb{R}(c \neq 0)$ を用いて

$$c \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

と与えられるため，3 本の線型独立な固有ベクトルを選択できないため，対角化不可である．また，零行列 O は固有値 $\lambda = 0$ の 3 重解で $\lambda = 0$ に対する固有ベクトルは定数 $c_1, c_2, c_3 \in \mathbb{R}(c_1 \neq 0, c_2 \neq 0, c_3 \neq 0)$ で

$$c_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + c_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

と与えられるため，3 本の線型独立な固有ベクトルを選択することができるため，対角化可能である．

ここで，補題 1.3 よりエルミート行列の固有ベクトルが直交するということに基づいて対称行列の対角化について考える．そこで，以下の定理を与える．

定理 1.4

任意の正規行列は対角化可能である．

Proof.

n 次のエルミート行列 A の n に関する数学的帰納法により示す．$n = 1$ のときは A は対角行列であることから対角化可能であることは明らかであるため，n で成り立つことを仮定して $n+1$ で成り立つことを示す．A の固有値 $\lambda_{n+1} \in \mathbb{C}$ に対する単位ベクトルな固有ベクトルを $\boldsymbol{u}_{n+1}$ として，これを $\mathbb{C}^{n+1}$ における超平面の単位法線ベクトルとすれば，この超平面は

$$P = \left\{ \boldsymbol{v} \in \mathbb{R}^{n+1} \,\middle|\, \boldsymbol{v} \cdot \boldsymbol{u}_{n+1} = 0 \right\}$$

と与えられ，明らかに $\dim P = n$ である．$\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n \in \mathbb{C}^{n+1}$ を P の正規直交基底とすれば，補題 1.4 より $k \in \mathbb{N}(1 \leq k \leq n)$ で

$$(A\boldsymbol{u}_k) \cdot \boldsymbol{u}_{n+1} = (A\boldsymbol{u}_k)^* \boldsymbol{u}_{n+1} = \boldsymbol{u}_k^* A^* \boldsymbol{u}_{n+1} = \boldsymbol{u}_k^* \lambda_{n+1}^* \boldsymbol{u}_{n+1} = 0.$$

つまり，$A\boldsymbol{u}_k \in P$ であるため $A\boldsymbol{u}_k$ は適当な $b_{1,k}, b_{2,k}, \cdots, b_{n,k} \in \mathbb{C}$ により

$$A\boldsymbol{u}_k = b_{1,k}\boldsymbol{u}_1 + b_{2,k}\boldsymbol{u}_2 + \cdots + b_{n,k}\boldsymbol{u}_n = \begin{pmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 & \cdots & \boldsymbol{u}_n \end{pmatrix} \begin{pmatrix} b_{1,k} \\ b_{2,k} \\ \vdots \\ b_{n,k} \end{pmatrix}$$

とあらわすことができる．これより，行列 U, B を

$$A \begin{pmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 & \cdots & \boldsymbol{u}_n \end{pmatrix} = \begin{pmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 & \cdots & \boldsymbol{u}_n \end{pmatrix} \begin{pmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,n} \\ b_{2,1} & b_{2,2} & \cdots & b_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n,1} & b_{n,2} & \cdots & b_{n,n} \end{pmatrix} \Leftrightarrow AU = UB$$

とすれば n 次の U はユニタリ行列であり，

$$B^*B = (U^*AU)^*(U^*AU) = U^*A^*AU = U^*AA^*U = (U^*AU)(U^*AU)^* = BB^*$$

となることから B は n 次の正規行列である．仮定より B は対角化可能であり，正則行列 Q により B は対角化されるとする．ここで，

$$P = \begin{pmatrix} \boldsymbol{u}_{n+1} & U \end{pmatrix} \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & Q \end{pmatrix}$$

と定義すれば，$U\boldsymbol{u}_{n+1} = \mathbf{0}$ となることに留意して

$$\begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & Q^{-1} \end{pmatrix} \begin{pmatrix} \boldsymbol{u}_{n+1}^* \\ U^* \end{pmatrix} P = \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & Q^{-1} \end{pmatrix} I_n \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & Q \end{pmatrix} = I_n$$

となるため，

$$P^{-1} = \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & Q^{-1} \end{pmatrix} \begin{pmatrix} \boldsymbol{u}_{n+1}^* \\ U^* \end{pmatrix}.$$

これより，

$$\begin{aligned}
P^{-1}AP &= \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & Q^{-1} \end{pmatrix} \begin{pmatrix} \boldsymbol{u}_{n+1}^* \\ U^* \end{pmatrix} A \begin{pmatrix} \boldsymbol{u}_{n+1} & U \end{pmatrix} \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & Q \end{pmatrix} \\
&= \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & Q^{-1} \end{pmatrix} \begin{pmatrix} \boldsymbol{u}_{n+1}^* \\ U^* \end{pmatrix} \begin{pmatrix} A\boldsymbol{u}_{n+1} & AU \end{pmatrix} \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & Q \end{pmatrix} \\
&= \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & Q^{-1} \end{pmatrix} \begin{pmatrix} \boldsymbol{u}_{n+1}^* \\ U^* \end{pmatrix} \begin{pmatrix} \lambda_{n+1}\boldsymbol{u}_{n+1} & UB \end{pmatrix} \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & Q \end{pmatrix} \\
&= \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & Q^{-1} \end{pmatrix} \begin{pmatrix} \lambda_{n+1} & \mathbf{0}^T \\ \mathbf{0} & B \end{pmatrix} \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & Q \end{pmatrix} \\
&= \begin{pmatrix} \lambda_{n+1} & \mathbf{0}^T \\ \mathbf{0} & Q^{-1}BQ \end{pmatrix}
\end{aligned}$$

となり，$n+1$ でも対角化可能であるため，任意の正規行列は対角化可能である．
よって，命題は証明された．

□

これより，以下の系が得られる．

系 1.4-1　n 次の正規行列は n 本の線型独立な固有ベクトルをもつ．

Proof.
定理 1.3 と定理 1.4 から明らか．

□

ここで，正規行列に対して拡張した補題 1.3 に着目すれば正規行列の固有ベクトルで正規直交基底が構成可能である，すなわち正規行列はユニタリ行列により対角化可能であると考えられる．これは同じ固有値に対する固有ベクトルから正規直交基底が生成可能であることを示せばいい．
そこで，以下の定理を示す．

定理 1.5　グラム・シュミットの直交化法

$\mathbb{R}^n$ 上の線型独立系 $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n\}$ を与えたとき

$$\begin{aligned}\boldsymbol{u}_1 &= \boldsymbol{v}_1\\ \boldsymbol{u}_2 &= \boldsymbol{v}_2 - \frac{\boldsymbol{v}_2 \cdot \boldsymbol{u}_1}{\boldsymbol{u}_1 \cdot \boldsymbol{u}_1}\boldsymbol{u}_1\\ \boldsymbol{u}_3 &= \boldsymbol{v}_3 - \frac{\boldsymbol{v}_3 \cdot \boldsymbol{u}_1}{\boldsymbol{u}_1 \cdot \boldsymbol{u}_1}\boldsymbol{u}_1 - \frac{\boldsymbol{v}_3 \cdot \boldsymbol{u}_2}{\boldsymbol{u}_2 \cdot \boldsymbol{u}_2}\boldsymbol{u}_2\\ &\vdots\\ \boldsymbol{u}_n &= \boldsymbol{v}_n - \sum_{k=1}^{n-1}\frac{\boldsymbol{v}_n \cdot \boldsymbol{u}_k}{\boldsymbol{u}_k \cdot \boldsymbol{u}_k}\boldsymbol{u}_k\end{aligned}$$

と与えられる $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n\}$ は直交系である．このような与えられた線型独立系から直交系を得る操作を**グラム・シュミットの直交化法 (Gram-Schmidt orthogonalization)** という．

Proof.

$\{\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n\}$ の構成より明らか．

□

これより，以下の定理が得られる．

定理 1.6

行列 A が正規行列であるための必要十分条件は A がユニタリ行列で対角化可能であることである．

Proof.

A が正規行列ならば系 1.4-1 と定理 1.5 よりユニタリ行列により対角化が可能である．A がユニタリ行列 U と対角行列 D により $U^*AU = D$ と対角化可能であるとき

$$A^*A = (UDU^*)^*(UDU^*) = UD^*DU^* = UDD^*U^* = (UDU^*)(UDU^*)^* = AA^*$$

となるため A は正規行列となる．

よって，命題は証明された．

□

1.3　相似変換

任意の正方行列が対角化可能であるわけではないが，対角化と似たような操作を考えることができる．そこで，以下の定義を与える．

定義 1.6

n 次の正方行列 A に対して，適当な正則行列 P により

$$A \mapsto P^{-1}AP$$

といった写像を与えたとする．このような写像を**相似変換 (similarity transformation)** という．

相似変換に対して以下の定理が与えられる．

定理 1.7

正方行列に対する相似変換の前後により，以下の量が不変である．

1). 行列式

2). 固有値

3). 固有多項式

4). 階数

5). トレース

Proof.

n 次の正方行列 A と適当な正則行列 P を与えたとする．1) について，P の行列式は正則性から非零かつ有限であるため

$$|P^{-1}AP| = |P^{-1}||A||P| = \frac{1}{|P|}|A||P| = |A|$$

となり，A に対する行列式に等しい．2) と 3) について，$P^{-1}AP$ に対する固有多項式 $p(x) \in \mathbb{C}[x]$ は

$$p(x) = |P^{-1}AP - xI| = |P^{-1}||A - PxIP^{-1}||P| = \frac{1}{|P|}|A - xI||P| = |A - xI|$$

となり，A に対する固有多項式に等しい，つまり固有値は等しい．4) について，P は正則であることから $\mathrm{rank}\, A = \mathrm{rank}\,(P^{-1}AP)$ である．5) について，定理 1.1 よりトレースは固有値の総和であるため，2) からトレースは相似変換の前後で不変である．

よって，命題は証明された．

□

この定理より相似変換は行列における重要な指標が保存される．また，対角化は相似変換により対角行列となる変換であり，相似変換により最も単純化される変換ともいえる．

しかし，一般に全ての正方行列で対角化することはできないため，一般の場合でも考えることが可能な以下の定理を与える．

定理 1.8　三角化定理

n 次の正方行列 A の重複を含めた固有値を $\lambda_1, \lambda_2, \cdots, \lambda_n \in \mathbb{C}$ とする．このとき，適当なユニタリ行列 U によって

$$U^*AU = \begin{pmatrix} \lambda_1 & & & * \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}$$

と相似変換が可能である．このように相似変換により上三角行列に変換することを**三角化**といい，このことを**三角化定理**という．また，これを A について解いたとき，それを A の**シューア分解 (Schur decomposition)** という．

Proof.

数学的帰納法により証明をする．$n = 1$ のときは自明であるため，$n - 1$ で成り立つことを仮定して，n で成り立つことを示す．A の固有値 $\lambda_1 \in \mathbb{C}$ に対する単位固有ベクトルを $\boldsymbol{u}_1 \in \mathbb{C}^n$ とし，$\mathbb{C}^n$ の正規直交基底となるように $\boldsymbol{u}_1$ を含めた基底 $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n \in \mathbb{C}^n$ を選択し，それらによるユニタリ行列 $U_1 = (\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n)$ を与える．$\boldsymbol{e}_k$ を $\mathbb{C}^n$ における基本ベクトルとすれば

$$\begin{aligned} U_1^*AU_1 &= U_1^*(\lambda_1\boldsymbol{u}_1, A\boldsymbol{u}_2, \cdots, A\boldsymbol{u}_n) \\ &= (\lambda_1\boldsymbol{e}_1, U_1^*A\boldsymbol{u}_2, \cdots, U_1^*A\boldsymbol{u}_n) \\ &= \begin{pmatrix} \lambda_1 & * \\ \boldsymbol{0} & B \end{pmatrix} \end{aligned}$$

となり，これにより $n - 1$ 次の正方行列 B が生成される．定理 1.7 より A と $U_1^*AU_1$ は同じ固有値をもつため，B は λ_1 以外の A と同じ固有値をもつ．仮定より B はユニタリ行列により三角化可能であるため，適当なユニタリ行列 U_2

により

$$U_2^* B U_2 = \begin{pmatrix} \lambda_2 & & & * \\ & \lambda_3 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}$$

とあらわすことができる．ここで，

$$U = U_1 \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & U_2 \end{pmatrix}$$

と定義すれば

$$\begin{aligned} U^* A U &= \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & U_2 \end{pmatrix}^* U_1^* A U_1 \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & U_2 \end{pmatrix} \\ &= \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & U_2^* \end{pmatrix} \begin{pmatrix} \lambda_1 & * \\ \mathbf{0} & B \end{pmatrix} \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & U_2 \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 & * \\ \mathbf{0} & U_2^* B U_2 \end{pmatrix} = \begin{pmatrix} \lambda_1 & & & * \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}. \end{aligned}$$

よって，命題は証明された．

□

この定理は定理 1.6 を一般化したものとして考えることができる．また，三角化に関連して以下の定理を与える．

補題 1.5　n 次の三角行列 A の対角成分は A の固有値である．

Proof.

上三角行列の場合を示す．固有方程式 $p(x) \in \mathbb{C}[x]$ は A の対角成分 $\lambda_1, \lambda_2, \cdots, \lambda_n \in \mathbb{C}$ を用いて

$$\begin{aligned} p(x) =& |A - xI| \\ =& \begin{vmatrix} \lambda_1 - x & & & * \\ & \lambda_2 - x & & \\ & & \ddots & \\ 0 & & & \lambda_n - x \end{vmatrix} \\ =& (\lambda_1 - x)(\lambda_2 - x) \cdots (\lambda_n - x) \end{aligned}$$

となることから，$\lambda_1, \lambda_2, \cdots, \lambda_n$ は A の固有値である．また，A と A^T の固有値は等しいことから下三角行列の場合も同様である．

よって，命題は証明された．

□

定理 1.9

n 次の正方行列 A の行列式は固有値の積に等しい．

Proof.

定理 1.8 より A は三角化可能であり，その行列式は定理 1.7 より三角化の前後で等しい．補題 1.5 より三角行列の対角成分は固有値であり，三角行列の行列式は対角成分の積となることから，行列式は固有値の積に等しい．

よって，命題は証明された．

□

また，三角化と補題 1.5 および定理 1.2 を用いることで正方行列 A のトレースが A の対格成分の和となることがただちに証明することができる．これより，固有値の和はトレース，固有値の積は行列式となるといった対応関係が成り立つ．

1.4　2 次形式

まずは以下の定義を与える．

定義 1.7

2 次の項のみからなる多項式のことを **2 次形式 (quadratic form)** という．

変数 $x_1, x_2, \cdots, x_n \in \mathbb{C}$ による任意の 2 次形式 $q(x_1, x_2, \cdots, x_n)$ は係数 $a_{1,1}, \cdots a_{1,n}, \cdots, a_{n,1}, \cdots, a_{n,n} \in \mathbb{C}$ により

$$q(x_1, x_2, \cdots, x_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{i,j} x_i^* x_j$$

とあらわされる．これは行列を用いて表現すれば

$$\begin{aligned} q(x_1, x_2, \cdots, x_n) &= \begin{pmatrix} \sum_{i=1}^n a_{i,1} x_i^* & \sum_{i=1}^n a_{i,2} x_i^* & \cdots & \sum_{i=1}^n a_{i,n} x_i^* \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \\ &= \begin{pmatrix} x_1^* & x_2^* & \cdots & x_n^* \end{pmatrix} \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}. \end{aligned}$$

これより，$x_1, x_2, \cdots, x_n$ によるベクトルを $\boldsymbol{x}$，$a_{1,1}, \cdots a_{1,n}, \cdots, a_{n,1}, \cdots, a_{n,n}$ による行列を A とすれば，2 次形式 $q(\boldsymbol{x})$ は

$$q(\boldsymbol{x}) = \boldsymbol{x}^* A \boldsymbol{x}$$

とあらわされる．また，A がエルミート行列であれば定理 1.6 より対角行列 D とユニタリ行列により

$$A = U^* D U$$

のように対角化可能であるため，

$$q(\boldsymbol{x}) = \boldsymbol{x}^* U^* D U \boldsymbol{x} = (U\boldsymbol{x})^* D (U\boldsymbol{x}).$$

ここで，D と $U\boldsymbol{x}$ の成分を

$$D = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}, \quad U\boldsymbol{x} = \begin{pmatrix} x_1' \\ x_2' \\ \vdots \\ x_n' \end{pmatrix}$$

のようにおくことにより，$q(\boldsymbol{x})$ は $q(U\boldsymbol{x})$ へと変数変換され

$$q(U\boldsymbol{x}) = \lambda_1 |x_1'|^2 + \lambda_2 |x_2'|^2 + \cdots + \lambda_n |x_n'|^2$$

のようになり，$x_i'^* x_i'$ に関する項のみとなる．このような表現を 2 次形式における**標準形 (normal form)** という．

また，$\boldsymbol{x}$ が実ベクトルかつ A が実行列であるとすれば $i \neq j$ で

$$x_i^* x_j = x_j^* x_i, \quad a_{i,j} = a_{j,i}$$

となるため A は対称行列となるため，標準形への変換が可能である．

ここで，2 次形式の定符号に関して考えるために以下の定義を与える．

定義 1.8

ベクトル $\boldsymbol{x} \in \mathbb{C}^n$ と行列 A による 2 次形式 $q(\boldsymbol{x}) = \boldsymbol{x}^* A \boldsymbol{x}$ について $\boldsymbol{x} \neq \boldsymbol{0}$ の条件下で以下の定義を与える．

1). $q(\boldsymbol{x}) > 0$ であるとき $q(\boldsymbol{x})$ は**正定値 (positive definite)** であるといい，A を**正定値行列 (positive definite**

matrix) という．

2). $q(\boldsymbol{x}) \geq 0$ であるとき $q(\boldsymbol{x})$ は**半正定値 (positive semidefinite)** であるといい，A を**半正定値行列 (positive semidefinite matrix)** という．

3). $q(\boldsymbol{x}) < 0$ であるとき $q(\boldsymbol{x})$ は**負定値 (negative definite)** であるといい，A を**負定値行列 (negative definite matrix)** という．

4). $q(\boldsymbol{x}) \leq 0$ であるとき $q(\boldsymbol{x})$ は**半負定値 (negative semidefinite)** であるといい，A を**半負定値行列 (negative semidefinite matrix)** という．

また，これらの性質をまとめて 2 次形式 $q(\boldsymbol{x})$ および行列 A における**定値性 (definiteness)** という．

特に，エルミート行列の定値性について以下の同値な命題を与えることができる．

定理 1.10

エルミート行列 A の定値性について，以下の必要十分条件が与えられる．

1). A が正定値 $\Leftrightarrow$ A の任意の固有値は 0 より大きい．

2). A が半正定値 $\Leftrightarrow$ A の任意の固有値は 0 以上．

3). A が負定値 $\Leftrightarrow$ A の任意の固有値は 0 より小さい．

4). A が半正定値 $\Leftrightarrow$ A の任意の固有値は 0 以下．

特に，A が半正定値行列であることと適当な行列 B により $A = B^*B$ と分解可能であることが必要十分条件である．

Proof.

1) を示す．A が n 次の正定値行列であるとすれば A の任意の固有値 $\lambda \in \mathbb{C}$ に対する固有ベクトル $\boldsymbol{x}_\lambda \in \mathbb{C}^n$ で

$$\begin{aligned}\boldsymbol{x}_\lambda^* A \boldsymbol{x}_\lambda &> 0\\ \boldsymbol{x}_\lambda^* \lambda \boldsymbol{x}_\lambda &> 0\\ \lambda|\boldsymbol{x}_\lambda|^2 &> 0\end{aligned}$$

となるが，$|\boldsymbol{x}_\lambda| > 0$ であるため $\lambda > 0$ となる．補題 1.2 より A の固有値は実数でそれぞれ $\lambda_1, \lambda_2, \cdots, \lambda_n \in \mathbb{R}$ とすれば，仮定および定理 1.4 より A は対角化可能であるため，A による 2 次形式 $q(\boldsymbol{x}) = \boldsymbol{x}^* A \boldsymbol{x}$ は変数 $x'_1, x'_2, \cdots, x'_n \in \mathbb{C}$ を用いて

$$q(x'_1, x'_2, \cdots, x'_n) = \lambda_1|x'_1|^2 + \lambda_2|x'_2|^2 + \cdots + \lambda_n|x'_n|^2.$$

また，任意の k で $|x'_k|^2 \geq 0$ かつ $\boldsymbol{x} \neq \boldsymbol{0}$ ならばある i が存在して $|x'_k|^2 > 0$ である．このとき，任意の k で $\lambda_k > 0$ ならば

$$q(x'_1, x'_2, \cdots, x'_n) > 0.$$

これより，1) は成り立ち，2)，3)，4) も同様にして示される．

定理 1.6 より A はユニタリ行列 U と対角行列 D により $A = U^*DU$ と対角化されるため，D の成分の平方根をとった行列を $D^{\frac{1}{2}}$ とあらわすことで

$$A = U^* D^{\frac{1}{2}} D^{\frac{1}{2}} U = (D^{\frac{1}{2}} U)^* (D^{\frac{1}{2}} U).$$

また，$A = B^*B$ と記述可能であるとき

$$q(\boldsymbol{x}) = \boldsymbol{x}^* B^* B \boldsymbol{x} = (B\boldsymbol{x})^*(B\boldsymbol{x}) = |B\boldsymbol{x}|^2 \geq 0$$

となるため，A がエルミート行列であるとき $A = B^*B$ と分解可能であることと A が半正定値行列であることは必要十分条件である．

よって，命題は証明された．

$\square$

1.5 LU 分解

正方行列 A の逆行列を拡大行列とガウスの消去法により求めることを考える．これを拡大行列を用いずに記述すると

$$\left(\begin{array}{c|c} A & I \end{array}\right) \Leftrightarrow AA^{-1} = I$$

ということであり，左から行基本変形を示す行列をかけて A を除去することで A^{-1} について解くことができる．この過程で上三角行列 U へと変形，すなわち適当な行列 P を用いて

$$\left(\begin{array}{c|c} U & P \end{array}\right) \Leftrightarrow PAA^{-1} = UA^{-1} = P.$$

また，この操作で行の入れ替えが生じないと仮定すれば，i 行目であれば i 列目の成分を k_i として，j 行目 $(i < j \leq n)$ の i 列目の成分 k_j とすれば，i 行目を $\frac{k_j}{k_i}$ 倍して j 行目から引くことで 0 になる．つまり，この仮定の下では P は対角成分が 1 の下三角行列である．

同様にして P の逆行列を求めることを考えれば，P は対角成分が 1 の下三角行列であるため A から U への変換と同様にして対角成分が 1 の下三角行列 L を用いて $LP = I$ となることに留意すれば

$$PP^{-1} = I \Leftrightarrow LPP^{-1} = P^{-1} = L.$$

以上より，

$$UA^{-1} = P \Leftrightarrow P^{-1}U = A \Leftrightarrow A = LU$$

となり，A は下三角行列 L と上三角行列 U へと分解することができる．このような分解を LU **分解 (LU decomposition)** という．また，このとき U の対角成分が 1 となるようにすれば，これを補うように各行を定数倍する対角行列 D により

$$A = LDU$$

とすることができる．このような分解を LDU **分解 (LDU decomposition)** という．

ここで，以下の定理を示す．

定理 1.11

LDU 分解可能な行列 A について，LDU 分解は一意である．

Proof.

$A = LDU$ と LDU 分解できるとして，$A = L'D'U'$ と LDU 分解の別表現をもつとする．このとき，

$$LDU = L'D'U' \Leftrightarrow L'^{-1}LD = D'U'U^{-1}$$

となるが，$L'L$ は対角成分が 1 の下三角行列，$U'U^{-1}$ は対角成分が 1 の上三角行列であるため，それぞれは単位行列であり $D = D'$ となる．また，

$$L'^{-1}L = I \Leftrightarrow L = L', \quad U'U^{-1} = I \Leftrightarrow U = U'$$

となるため，A の LDU 分解が別表現をもつ仮定に矛盾する．

よって，命題は証明された．

□

A が LU 分解可能であるとは A の逆行列をガウスの消去法で行の入れ替えなしに求めることが可能であるということである．つまり，以下の定理が与えられる．

定理 1.12

任意の正則行列 A は適当な行を入れ替える行列 P を用いて $PA = LU$ と LU 分解することが可能である．これを LUP **分解 (LUP decomposition)** という．

Proof.

LU 分解の導出より明らかである．

□

定理 1.13

行列 A が LU 分解可能である必要十分条件は A の主小行列の行列式が非零であることである．

Proof.

A が LDU 分解可能であるとすれば，$A = LDU$ と LDU 分解するとして A の i 行 i 列目の主小行列は

$$A_{i,i} = L_{i,i}D_{i,i}U_{i,i}.$$

このとき，$L_{i,i}$ の対角成分は全て 1 の下三角行列で $U_{i,i}$ の対角成分も全て 1 の上三角行列，D は対角行列であるため，$\det A_{i,i}$ は D の対角成分の積であり，D の対角成分は非零であるため $\det A_{i,i} \neq 0$ となる．

A の主小行列の行列式が非零として，行の入れ替えなしの行基本変形により上三角行列へと変換可能であることを示す．仮定より $A_{1,1}$ は正則であり，その各行は線型独立である．このとき，

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & & & \\ \vdots & & A_{1,1} & \\ a_{n,1} & & & \end{pmatrix}$$

とあらわすとして，A の 1 行目を i 行目に対して 1 行目を $\frac{a_{i,1}}{a_{1,1}}$ 倍して引くとき，零ベクトルが生成されないならば $A_{1,1}$ の変換後の $A'_{1,1}$ の行ベクトルの線型独立性は保持される．A の j 行目で零ベクトルが生成されるならば，A の 1 行目と j 行目は等しくなり，例えば $A_{2,2}$ の行列式が非零である仮定に矛盾する．つまり，行基本変形により

$$A' = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ 0 & & & \\ \vdots & & A'_{1,1} & \\ 0 & & & \end{pmatrix}$$

としても A' の小行列 $A'_{1,1}$ の行列式は非零を保持する．また，$A'_{2,2}, A'_{3,3}, \cdots, A'_{n,n}$ も $A'_{1,1}$ のときと同様にして行列式が非零でることを保持する．この操作を繰り返すことにより A は行の入れ替えなしの行基本変形により上三角行列へと変換可能である．

よって，命題は証明された．

□

第 2 章

一般の行列のスペクトル

2.1 特異値分解

任意の行列における分解を考えるために，以下の補題を与える．

補題 2.1 任意の行列 A により

$$B = A^*A, \quad C = AA^*$$

と定義される行列 B, C は半正定値なエルミート行列かつ固有値は共通である．

Proof.

$$\begin{aligned}B^* =&(A^*A)^* = A^*(A^*)^* = A^*A\\ C^* =&(AA^*)^* = (A^*)^*A^* = AA^*\end{aligned}$$

より，B と C はそれぞれエルミート行列である．また，定理 1.10 より B と C は半正定値行列となる．B の固有値 λ とそれに対する固有ベクトル $\boldsymbol{x}$ について

$$B\boldsymbol{x} = \lambda\boldsymbol{x} \;\Leftrightarrow\; AB\boldsymbol{x} = A\lambda\boldsymbol{x} \;\Leftrightarrow\; C(A\boldsymbol{x}) = \lambda(A\boldsymbol{x})$$

となり，λ は C の固有値となる．同様にして C の固有値は B の固有値となる．

よって，命題は証明された．

□

この補題と定理 1.6 により任意の行列とそのエルミート転置の積はユニタリ行列により対角化可能である．ここで，任意の $m \times n$ 行列 A で A^*A はエルミート行列であるため，補題 1.2 より A の固有値は実数でそれぞれ重複を含めて $\lambda_1, \lambda_2, \cdots, \lambda_n \in \mathbb{R}$ とすれば，正規直交基底を構成するようにそれぞれの固有値に対応する固有ベクトル $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n \in \mathbb{C}^n$ を選択することができる．また，補題 2.1 から A^*A は半正定値行列であるため，定理 1.10 より $\lambda_1, \lambda_2, \cdots, \lambda_n$ は 0 以上であり，

$$\begin{cases}\lambda_k > 0 & (1 \leq k \leq r)\\ \lambda_k = 0 & (r < k \leq n)\end{cases}$$

としてもよい．ここで，$i, j \in \mathbb{N}(1 \leq i, j \leq r)$ で $A\boldsymbol{v}_i$ と $A\boldsymbol{v}_j$ の内積はクロネッカーのデルタ $\delta_{i,j}$ を用いて

$$\begin{aligned}(A\boldsymbol{v}_i)^*(A\boldsymbol{v}_j) =&\boldsymbol{v}_i^*A^*A\boldsymbol{v}_j\\ =&\boldsymbol{v}_i^*\lambda_j\boldsymbol{v}_j\\ =&\lambda_j\delta_{i,j}\end{aligned}$$

となることから，

$$\boldsymbol{u}_1 = \frac{1}{\sqrt{\lambda_1}}A\boldsymbol{v}_1, \quad \boldsymbol{u}_2 = \frac{1}{\sqrt{\lambda_2}}A\boldsymbol{v}_2, \quad \cdots, \quad \boldsymbol{u}_r = \frac{1}{\sqrt{\lambda_r}}A\boldsymbol{v}_r$$

とすれば $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_r$ は正規直交基底を構成する．また，$\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_r$ はそれぞれ m 次のベクトルであることから $r \leq m$ となることがわかる．

$\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_r$ に対して標準基底とグラム・シュミットの直交化法を用いて単位ベクトル $\boldsymbol{u}_{r+1}, \boldsymbol{u}_{r+2}, \cdots, \boldsymbol{u}_m \in \mathbb{C}^m$ を生成すれば，$\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_m$ は正規直交基底を構成する．ここで，対角化の構成と同様にして，$r < k \leq n$ で $(A\boldsymbol{v}_k)^*(A\boldsymbol{v}_k) = 0$ となることに留意すれば

$$\begin{pmatrix} A\boldsymbol{v}_1 & A\boldsymbol{v}_2 & \cdots & A\boldsymbol{v}_n \end{pmatrix} = \begin{pmatrix} \sqrt{\lambda_1}\boldsymbol{u}_1 & \sqrt{\lambda_2}\boldsymbol{u}_2 & \cdots & \sqrt{\lambda_r}\boldsymbol{u}_r & 0 & \cdots & 0 \end{pmatrix}$$

$$A\begin{pmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \cdots & \boldsymbol{v}_n \end{pmatrix} = \begin{pmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 & \cdots & \boldsymbol{u}_r & 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} \sqrt{\lambda_1} & & & 0 & 0 & \cdots & 0 \\ & \sqrt{\lambda_2} & & & \vdots & & \vdots \\ & & \ddots & & \vdots & & \vdots \\ 0 & & & \sqrt{\lambda_r} & 0 & \cdots & 0 \\ 0 & \cdots & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & & \vdots & \vdots & & \vdots \\ 0 & \cdots & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}$$

$$= \begin{pmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 & \cdots & \boldsymbol{u}_m \end{pmatrix} \begin{pmatrix} \sqrt{\lambda_1} & & & 0 & 0 & \cdots & 0 \\ & \sqrt{\lambda_2} & & & \vdots & & \vdots \\ & & \ddots & & \vdots & & \vdots \\ 0 & & & \sqrt{\lambda_r} & 0 & \cdots & 0 \\ 0 & \cdots & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & & \vdots & \vdots & & \vdots \\ 0 & \cdots & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}$$

となり，

$$U = \begin{pmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 & \cdots & \boldsymbol{u}_m \end{pmatrix}, \quad V = \begin{pmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \cdots & \boldsymbol{v}_n \end{pmatrix}$$

とすれば，U と V はユニタリ行列であるため

$$A = U \begin{pmatrix} \sqrt{\lambda_1} & & & 0 & 0 & \cdots & 0 \\ & \sqrt{\lambda_2} & & & \vdots & & \vdots \\ & & \ddots & & \vdots & & \vdots \\ 0 & & & \sqrt{\lambda_r} & 0 & \cdots & 0 \\ 0 & \cdots & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & & \vdots & \vdots & & \vdots \\ 0 & \cdots & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix} V^* = U\Sigma V^*$$

とすることが可能である．

これより，以下の定義を与えることができる．

定義 2.1

$m \times n$ 行列 A はユニタリ行列 U, V と対角成分が非零の r 次の正方行列 D により

$$\Sigma = \begin{pmatrix} D & O \\ O & O \end{pmatrix}$$

と定義される $m \times n$ 行列 Σ で $A = U\Sigma V^*$ と分解が可能である．これを**特異値分解 (singular value decomposition)** といい，**SVD** と略される．また，D を含む Σ の $\min m, n$ 次正方小行列の対角成分を**特異値 (singular value)** という．なお，D の対角成分を特異値とする場合もある．

特異値分解の導出より，任意の実行列でも特異値分解は可能であることがわかる．また，特異値は A^*A および AA^* の固有値の平方根として与えたが，特異値分解は $\boldsymbol{u}_{r+1}, \boldsymbol{u}_{r+2}, \cdots, \boldsymbol{u}_m$ の構成から一意でないことは明らかであり，A には Σ 内に記述される特異値以外のものが存在する可能性がある．

そこで，以下の定理を与える．

定理 2.1

$m \times n$ 行列 A が $A = U\Sigma V^*$ と特異値分解されるとき，Σ の特異値の並び順を無視すれば一意である．

Proof.

Σ が対角成分が非零の r 次の正方行列 D により

$$\Sigma = \begin{pmatrix} D & O_{n-r} \\ O_{m-r} & O_{m-r,n-r} \end{pmatrix}$$

と記述されるとき，$U^*AV = \Sigma$ となるため

$$\begin{aligned}
(U^*AV)^*(U^*AV) =& V^*(A^*A)V \\
=& \Sigma^*\Sigma \\
=& \begin{pmatrix} D & O_{m-r} \\ O_{n-r} & O_{m-r,n-r} \end{pmatrix} \begin{pmatrix} D & O_{n-r} \\ O_{m-r} & O_{m-r,n-r} \end{pmatrix} \\
=& \begin{pmatrix} D^2 & O_{n-r} \\ O_{n-r} & O_{n-r} \end{pmatrix} \\
(U^*AV)(U^*AV)^* =& U^*(AA^*)U \\
=& \Sigma\Sigma^* \\
=& \begin{pmatrix} D & O_{n-r} \\ O_{m-r} & O_{m-r,n-r} \end{pmatrix} \begin{pmatrix} D & O_{m-r} \\ O_{n-r} & O_{m-r,n-r} \end{pmatrix} \\
=& \begin{pmatrix} D^2 & O_{m-r} \\ O_{m-r} & O_{m-r} \end{pmatrix}.
\end{aligned}$$

系 1.3-1 および系 1.3-2 から $\Sigma^*\Sigma$ と $\Sigma\Sigma^*$ は A^*A および AA^* の固有値の並び順を無視すれば一意となり，r は $r \leq m$ および $r \leq n$ を満たすため D は A の非零の特異値の並び順を無視すれば一意となる．つまり，Σ は A の特異値の並び順を無視すれば一意である．

よって，命題は証明された．

□

これより，以下の系が得られる．

系 2.1-1　行列 A に対して，A^*A と AA^* の非零な固有値は重複を含め等しく，それぞれの平方根は重複を含めた特異値となる．

Proof.

定理 2.1 の証明中より明らか．

□

この定理より，A^*A および A^*A の重複を含めた非零な固有値の平方根がそのまま Σ を構成する因子となる．

そこで，以下の定理を与えることで明示的に特異値の性質を記述する．

定理 2.2

$m \times n$ 行列 A に対して非負の $\sigma \in \mathbb{R}$ と $\boldsymbol{u} \in \mathbb{C}^m, \boldsymbol{v} \in \mathbb{C}^n$ で

$$\begin{cases} A\boldsymbol{v} = \sigma\boldsymbol{u} \\ A^*\boldsymbol{u} = \sigma\boldsymbol{v} \end{cases}$$

を満たすことが σ が A の特異値であることの必要十分条件である．また，$\boldsymbol{u}$ と $\boldsymbol{v}$ をそれぞれ特異値 σ に対する**左特異ベクトル (left-singular vector)** および**右特異ベクトル (right-singular vector)** という．

Proof.

σ が A の特異値であるならば，特異値分解の構成から適当な $\boldsymbol{u} \in \mathbb{C}^m, \boldsymbol{v} \in \mathbb{C}^n$ で $A\boldsymbol{u} = \boldsymbol{v}$ かつ系 2.1-1 より

$(A^*A)\boldsymbol{v} = \sigma^2\boldsymbol{v}$ が成り立つため，σ が非零とすればこれらを連立することにより

$$\begin{aligned} A^*(A\boldsymbol{v}) =& \sigma^2\boldsymbol{v} \\ A^*\sigma\boldsymbol{u} =& \sigma^2\boldsymbol{v} \\ A^*\boldsymbol{u} =& \sigma\boldsymbol{v} \end{aligned}$$

が得られ，$A^*\boldsymbol{u} = \boldsymbol{v}$ を仮定した場合でも同様に得られる．また，$\sigma = 0$ のときは

$$\begin{aligned} (A^*A)\boldsymbol{v} =& \boldsymbol{0} \\ \boldsymbol{v}^*A^*A\boldsymbol{v} =& 0 \\ |A\boldsymbol{v}|^2 =& 0 \\ (AA^*)\boldsymbol{u} =& \boldsymbol{0} \\ \boldsymbol{u}^*AA^*\boldsymbol{v} =& 0 \\ |A^*\boldsymbol{u}|^2 =& 0 \end{aligned}$$

となるため，σ が特異値ならば

$$\begin{cases} A\boldsymbol{v} = \sigma\boldsymbol{u} \\ A^*\boldsymbol{u} = \sigma\boldsymbol{v}. \end{cases}$$

$A\boldsymbol{v} = \sigma\boldsymbol{u}$ かつ $A^*\boldsymbol{u} = \sigma\boldsymbol{v}$ ならば，A が $A = U\Sigma V^*$ と特異値分解されるとき，$\boldsymbol{u}$ は U の列ベクトルで $\boldsymbol{v}$ は V の列ベクトルとなることは特異値分解の構成より明らかであり，σ は A の特異値となる．

よって，命題は証明された．

□

また，非零な特異値について以下の定理が成り立つ．

定理 2.3

行列 A の階数は A の非零の特異値の数に等しい．

Proof.

A が $A = U\Sigma V^*$ と特異値分解されるとき，U と V は正則であるため $\operatorname{rank} A = \operatorname{rank} \Sigma$ である．つまり，A の階数は A の非零の特異値の数に等しい．

よって，命題は証明された．

□

2.2　疑似逆行列

方程式が n 本で変数が m 個の線型連立方程式は一般に解をもたない．そこで，変数 $\boldsymbol{x} \in \mathbb{C}^m$ と定数 $\boldsymbol{b} \in \mathbb{C}^n$ と $m \times n$ 係数行列 A による線型不能連立方程式 $A\boldsymbol{x} = \boldsymbol{b}$ を考えると，$\boldsymbol{b}$ と A を測定したデータ，$\boldsymbol{x}$ を推定する係数とした最小 2 乗法による線型回帰分析と考えることで最も適当な解 $\boldsymbol{x}$ を得ることができると考えられる．このとき，この方程式の残差平方和 $J(\boldsymbol{x})$ は

$$J(\boldsymbol{x}) = (A\boldsymbol{x} - \boldsymbol{b})^*(A\boldsymbol{x} - \boldsymbol{b})$$

となり，A が $A = U\Sigma V^*$ と特異値分解されるとすれば

$$J(\boldsymbol{x}) = (U\Sigma V^*\boldsymbol{x} - \boldsymbol{b})^*(U\Sigma V^*\boldsymbol{x} - \boldsymbol{b}) = (\Sigma V^*\boldsymbol{x} - U^*\boldsymbol{b})^*(\Sigma V^*\boldsymbol{x} - U^*\boldsymbol{b}).$$

ここで，

$$\boldsymbol{y} = V^*\boldsymbol{x} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad \boldsymbol{c} = U^*\boldsymbol{b} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{pmatrix}$$

として，A の非零の特異値を $\sigma_1, \sigma_2, \cdots, \sigma_r \in \mathbb{R}$ とすれば $J(\boldsymbol{x}) = J(\boldsymbol{y})$ は

$$J(\boldsymbol{y}) = \sum_{k=1}^{r} (\sigma_k y_k - c_k) - \sum_{k=r+1}^{m} c_k$$

となるため，$J(\boldsymbol{y})$ を最小にするには $\sum_{k=1}^{r} (\sigma_k y_k - c_k) = 0$，すなわち $k \in \mathbb{N}(1 \le k \le r)$ で $y_k = \frac{\sigma_k}{c_k}$ とすればいい．そのためには $\boldsymbol{c}$ を第 r 成分以降が 0 の n 次のベクトル，すなわち

$$\boldsymbol{c} \mapsto \Sigma^T \boldsymbol{c}$$

として，このベクトルの第 k 成分を $\frac{1}{\sigma_k^2}$ すればいい．このとき，$\sigma_1, \sigma_2, \cdots, \sigma_r$ による対角行列を D としたとき

$$\Sigma^T \Sigma = \begin{pmatrix} D^2 & O \\ O & O \end{pmatrix}$$

は n 次の正方行列であり，$\frac{1}{\sigma_k^2}$ という操作を達成するために

$$(\Sigma^T \Sigma)^+ = \begin{pmatrix} (D^{-1})^2 & O \\ O & O \end{pmatrix}$$

という表現を便宜的にとることで

$$\boldsymbol{y} = (\Sigma^T \Sigma)^+ \Sigma^T \boldsymbol{c} \;\Leftrightarrow\; \boldsymbol{x} = V (\Sigma^T \Sigma)^+ \Sigma^T U^* \boldsymbol{b}$$

で $J(\boldsymbol{x})$ の残差平方和は最小化する．また，Σ^+ は Σ^T の非零の成分の逆数をとったもの，すなわち Σ に対して $\sigma_1, \sigma_2, \cdots, \sigma_r$ の逆数をとって転置したものとすれば

$$(\Sigma^T \Sigma)^+ = \Sigma^+ (\Sigma^T)^+$$

といった関係式が得られ，

$$\boldsymbol{x} = V \Sigma^+ (\Sigma^T)^+ \Sigma^T U^* \boldsymbol{b} = V \Sigma^+ U^* \boldsymbol{b}.$$

以上より，以下の定義を与えることができる．

定義 2.2

行列 A が $A = U\Sigma V^*$ と特異値分解されるとき，Σ^T の非零の成分の逆数をとったものを Σ^+ としたとき，

$$A^+ = V\Sigma^+ U^*$$

を A の**擬似逆行列 (pseudo-inverse matrix)** という．

このとき，A が正則行列ならば $A^+ = A^{-1}$ となることは定義より明らかである．また，擬似逆行列について以下のような同値な定義が与えられる．

定理 2.4

行列 A が以下を満たすことと A^+ が擬似逆行列であることは必要十分条件である．

1). $AA^+A = A$.

2). $A^+AA^+ = A^+$.

3). $(AA^+)^* = AA^+$.

4). $(A^+A)^* = A^+A$.

また，このとき A の擬似逆行列は一意に定まる．

Proof.

A が $A = U\Sigma V^*$ と特異値分解されるとして，1) について，

$$AA^+A = (U\Sigma V^*)(V\Sigma^+ U^*)(U\Sigma V^*) = U\Sigma\Sigma^+\Sigma V^* = U\Sigma V^* = A.$$

2) について，

$$A^+AA^+ = (V\Sigma^+U^*)(U\Sigma V^*)(V\Sigma^+U^*) = V\Sigma^+\Sigma\Sigma^+U^* = V\Sigma^+U^* = A^+.$$

3) について，

$$\begin{aligned}(AA^+)^* =&(A^+)^*A^* = (U(\Sigma^+)^TV^*)(V\Sigma^TU^*) = U(\Sigma^+)^T\Sigma^TU^* = U(\Sigma\Sigma^+)^TU^* = U\Sigma\Sigma^+U^*\\ AA^+ =&(U\Sigma V^*)(V\Sigma^+U^*) = U\Sigma\Sigma^+U^*\end{aligned}$$

となるため，$(AA^+)^* = AA^+$ となる．4) について，

$$\begin{aligned}(A^+A)^* =&A^*(A^+)^* = (V\Sigma^TU^*)(U(\Sigma^+)^TV^*) = V\Sigma^T(\Sigma^+)^TV^* = V(\Sigma^+\Sigma)^TV^* = V\Sigma^+\Sigma V^*\\ AA^+ =&(V\Sigma^+U^*)(U\Sigma V^*) = V\Sigma^T\Sigma V^*\end{aligned}$$

となるため，$(A^+A)^* = A^+A$ となる．

また，擬似逆行列は定理 2.1 より一意であるため，1) と 2) と 3) と 4) を満たす行列 A^+ が一意であることを示せば必要十分性を示すことができる．A の相異なる 2 つの擬似逆行列 X, Y が存在するとしたとき，

$$\begin{aligned}X =&XAX = X(AX)^* = X((AYA)X)^* = X(AX)^*(AY)^* = XAXAY = X(AXA)Y = XAY\\ Y =&YAY = (YA)^*Y = (Y(AXA))^*Y = (XA)^*(YA)^*Y = XAYAY = X(AYA)Y = XAY\end{aligned}$$

となり，$X \neq Y$ の仮定に矛盾するため，1) と 2) と 3) と 4) を満たす行列 A^+ は一意である．

よって，命題は証明された．

□

また，擬似逆行列の満たす性質として以下の定理が与えられる．

定理 2.5

$m \times n$ 行列 A の疑似逆行列について以下の性質が成り立つ．

1). $(A^+)^+ = A$.

2). $(A^+)^* = (A^*)^+$.

3). $(A^*A)^+ = A^+(A^*)^+$.

4). $A^+ = (A^*A)^+A^* = A^*(AA^*)^+$.

Proof.

A が $A = U\Sigma V^*$ と特異値分解されるとする．1) について，

$$(A^+)^+ = (V\Sigma^+U^*)^+ = U(\Sigma^+)^+V^* = U\Sigma V^* = A.$$

2) について，

$$\begin{aligned}(A^+)^* =&(V\Sigma^+U^*)^* = U(\Sigma^+)^*V^*\\ (A^*)^+ =&(V\Sigma^*U^*)^+ = U(\Sigma^*)^+V^*\end{aligned}$$

で $(\Sigma^+)^* = (\Sigma^*)^+$ となるため $(A^+)^* = (A^*)^+$ となる．3) について，

$$\begin{aligned}(A^*A)^+ =&((V\Sigma^*U^*)(U\Sigma V^*))^+ = (V\Sigma^T\Sigma V^*)^+ = V(\Sigma^T\Sigma)^+V^*\\ A^+(A^*)^+ =&(V\Sigma^+U^*)(V\Sigma^TU^*)^+ = (V\Sigma^+U^*)(U(\Sigma^T)^+V^*) = V\Sigma^+(\Sigma^T)^+V^*\end{aligned}$$

で $(\Sigma^T\Sigma)^+ = \Sigma^+(\Sigma^T)^+$ となるため $(A^*A)^+ = A^+(A^*)^+$ となる．4) について，2) と 3 と定理 2.4 より

$$\begin{aligned}(A^*A)^+A^* =&A^+(A^*)^+A^* = A^+(AA^+)^* = A^+AA^+ = A^+\\ A^*(AA^*)^+ =&A^*(A^*)^+A^+ = (A^+A)^*A^+ = A^+AA^+ = A^+\end{aligned}$$

となるため $A^+ = (A^*A)^+A^* = A^*(AA^*)^+$ となる．

よって，命題は証明された．

□