

第E章 特異値分解

E.1 特異値分解

定義 E.1. $n, p \in \mathbb{N}, n \geq p$ とし, $V \subset \mathbb{R}^n$ をベクトル空間とする. $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_p \in V$ とする.

(1) $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_p$ はベクトル空間 V の基底であるとは, 任意の $\mathbf{v} \in V$ に対して, 唯一の $a_1, a_2, \dots, a_p \in \mathbb{R}$ があって,

$$\mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \dots + a_p\mathbf{v}_p$$

とかけるときをいう. この p をベクトル空間 V の次元といい, $\dim(V)$ と書く.

(2) $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_p$ は線型独立 (1 次独立) であるとは,

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \dots + a_p\mathbf{v}_p = \mathbf{0} \quad (a_1, a_2, \dots, a_p \in \mathbb{R}) \Rightarrow a_1 = a_2 = \dots = a_p = 0$$

が成り立つときをいう. ただし, $\mathbf{0} = (0, 0, \dots, 0)^\top \in \mathbb{R}^n$ である.

定義 E.2. (1) $n, p \in \mathbb{N}, n \geq p$ とする. ベクトル $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_p \in \mathbb{R}^n$ で張られた \mathbb{R}^n の部分空間 $\text{span}\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_p\}$ を

$$\text{span}\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_p\} = \{a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \dots + a_p\mathbf{x}_p; a_1, a_2, \dots, a_p \in \mathbb{R}\}$$

と定義する.

(2) $m, n \in \mathbb{N}$ とし, $\mathbf{A} \in \text{Mat}(m, n; \mathbb{R})$ とする. \mathbf{A} と \mathbf{A}^\top の像と核をそれぞれ

$$\begin{aligned} \mathbf{R}(\mathbf{A}) &= \{\mathbf{Ax}; \mathbf{x} \in \mathbb{R}^n\} \subset \mathbb{R}^m, \\ \ker(\mathbf{A}) &= \{\mathbf{x}; \mathbf{Ax} = \mathbf{0} \in \mathbb{R}^m\} \subset \mathbb{R}^n, \\ \mathbf{R}(\mathbf{A}^\top) &= \{\mathbf{A}^\top\mathbf{y}; \mathbf{y} \in \mathbb{R}^m\} \subset \mathbb{R}^n, \\ \ker(\mathbf{A}^\top) &= \{\mathbf{y}; \mathbf{A}^\top\mathbf{y} = \mathbf{0} \in \mathbb{R}^n\} \subset \mathbb{R}^m \end{aligned}$$

で定める.

定義 E.3. $n \in \mathbb{N}$ とする. 行列式は関数 $\det: \text{Mat}(n; \mathbb{R}) \rightarrow \mathbb{R}$ で以下のように定義する.

(1) 1×1 の行列 (a) の行列式を

$$\det(a) = a.$$

(2) $n \in \mathbb{N}, n \geq 2$ とき, 行列

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \in \text{Mat}(n; \mathbb{R})$$

の行列式を以下のように定める. $i, j = 1, 2, \dots, n$ に対し \mathbf{A}_{ij} を \mathbf{A} から i 行と j 列をとった $(n-1) \times (n-1)$ 行列とする. $n \times n$ 行列 \mathbf{A} の行列式を

$$\det(\mathbf{A}) = \sum_{k=1}^n (-1)^{k+1} a_{1k} \det(\mathbf{A}_{1k})$$

で定める.

命題 E.4. $n \times n$ の行列 $\mathbf{A} = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n) \in \text{Mat}(n; \mathbb{R})$ に対して, 行列式は以下の性質をみたす.

(1) $k = 1, 2, \dots, n$ と $d \in \mathbb{R}$ に対し,

$$\det(\mathbf{a}_1, \dots, \lambda \mathbf{a}_k, \dots, \mathbf{a}_n) = \lambda \det(\mathbf{a}_1, \dots, \mathbf{a}_k, \dots, \mathbf{a}_n).$$

(2) $j \neq k$ に対し,

$$\det(\mathbf{a}_1, \dots, \mathbf{a}_k + \lambda \mathbf{a}_j, \dots, \mathbf{a}_n) = \lambda \det(\mathbf{a}_1, \dots, \mathbf{a}_k, \dots, \mathbf{a}_n).$$

(3) $\det(\mathbf{I}_n) = 1$.

(4) $\mathbf{A}, \mathbf{B} \in \text{Mat}(n; \mathbb{R})$ に対し

$$\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B}).$$

定義 E.5. $\mathbf{A} \in \text{Mat}(n; \mathbb{R})$ に対し, ある $\mathbf{B} \in \text{Mat}(n; \mathbb{R})$ が存在し,

$$\mathbf{AB} = \mathbf{BA} = \mathbf{I}_n$$

が成り立つとき, \mathbf{A} は可逆であるといい, \mathbf{B} を \mathbf{A} の逆行列といい, これを \mathbf{A}^{-1} と記す.

命題 E.6. $\mathbf{A} \in \text{Mat}(n; \mathbb{R})$ とする. このとき

$$\mathbf{A} \text{ は可逆} \Leftrightarrow \det(\mathbf{A}) \neq 0.$$

定理 E.7. (特異値分解) $m, n \in \mathbb{N}$ とする. 任意の $A \in \text{Mat}(m, n; \mathbb{R})$ は次の分解をもつ.

$$A = UDV^T.$$

ただし $U \in \text{Mat}(m; \mathbb{R})$, $V \in \text{Mat}(n; \mathbb{R})$ は直交行列で, $D \in \text{Mat}(m, n; \mathbb{R})$ は対角行列で, その対角成分は $d_1 \geq d_2 \geq \cdots \geq d_{\min(m, n)} \geq 0$ である. d_j ($j = 1, 2, \dots, \min(m, n)$) を行列 A の特異値といい, $\{(d_j, \mathbf{u}_j, \mathbf{v}_j); j = 1, 2, \dots, \min(m, n)\}$ を特異値系¹という.

注意 E.8. $n \geq m$ のとき

$$D = \begin{pmatrix} d_1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & d_m & 0 & \cdots & 0 \end{pmatrix}$$

である. $m \geq n$ のとき,

$$D = \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \\ 0 & 0 & \cdots & 0 \end{pmatrix}$$

である. □

定理 E.7 の証明 $\mathbf{y} \in \mathbb{R}^m$ に対し,

$$|\mathbf{y}|_2 = \sqrt{\mathbf{y}^T \mathbf{y}}$$

とする. ただし, 次元の異なる Euclid 空間のノルムも同じ記号を流用する. 行列 A のノルム

$$\|A\| := \sup_{\mathbf{x} \in \mathbb{R}^n} \frac{|A\mathbf{x}|_2}{|\mathbf{x}|_2}$$

で定義すると

$$d_1 := \|A\| := \sup_{\mathbf{x} \in \mathbb{R}^n, |\mathbf{x}|_2=1} |A\mathbf{x}|_2$$

¹ $j = 1, 2, \dots, \min(m, n)$ に対して

$$A\mathbf{v}_j = d_j \mathbf{u}_j$$

となっている.

となる. $d_1 \neq 0$ と仮定する. そうでなければ, 定理の証明は自明となる. あるベクトル $\boldsymbol{x} \in \mathbb{R}^n$ ($|\boldsymbol{x}|_2 = 1$) が存在して

$$|\boldsymbol{Ax}|_2 = d_1$$

をみたすとする.

$$\boldsymbol{y} = \frac{1}{d_1} \boldsymbol{Ax} \in \mathbb{R}^m$$

とおけば

$$|\boldsymbol{y}|_2^2 = \boldsymbol{y}^\top \boldsymbol{y} = \frac{1}{d_1^2} \boldsymbol{x}^\top \boldsymbol{A}^\top \boldsymbol{Ax} = \frac{|\boldsymbol{Ax}|_2^2}{d_1^2} = 1$$

より $|\boldsymbol{y}|_2 = 1$ となる. $\boldsymbol{v}_2, \boldsymbol{v}_3, \dots, \boldsymbol{v}_n \in \mathbb{R}^n$ と $\boldsymbol{u}_2, \boldsymbol{u}_3, \dots, \boldsymbol{u}_m \in \mathbb{R}^m$ をうまく選んで, $\{\boldsymbol{x}, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n\}$ と $\{\boldsymbol{y}, \boldsymbol{u}_2, \dots, \boldsymbol{u}_m\}$ はそれぞれ \mathbb{R}^n と \mathbb{R}^m の正規直交基底となるようにする.

$$\boldsymbol{U}_1 = (\boldsymbol{y}, \boldsymbol{u}_2, \dots, \boldsymbol{u}_m) \in \text{Mat}(m; \mathbb{R}), \quad \boldsymbol{V}_1 = (\boldsymbol{x}, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n) \in \text{Mat}(n; \mathbb{R}),$$

$$\boldsymbol{A}_1 = \boldsymbol{U}_1^\top \boldsymbol{AV}_1$$

とおけば,

$$\begin{aligned} \boldsymbol{A}_1 = \boldsymbol{U}_1^\top \boldsymbol{AV}_1 &= \begin{bmatrix} \boldsymbol{y}^\top \\ \boldsymbol{u}_2^\top \\ \vdots \\ \boldsymbol{u}_m^\top \end{bmatrix} \boldsymbol{A}[\boldsymbol{x}, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n] = \begin{bmatrix} \boldsymbol{y}^\top \\ \boldsymbol{u}_2^\top \\ \vdots \\ \boldsymbol{u}_m^\top \end{bmatrix} [\boldsymbol{Ax}, \boldsymbol{Av}_2, \dots, \boldsymbol{Av}_n] \\ &= \begin{bmatrix} \boldsymbol{y}^\top \\ \boldsymbol{u}_2^\top \\ \vdots \\ \boldsymbol{u}_m^\top \end{bmatrix} [d_1 \boldsymbol{y}, \boldsymbol{Av}_2, \dots, \boldsymbol{Av}_n] = \begin{bmatrix} d_1 \boldsymbol{y}^\top \boldsymbol{y} & \boldsymbol{y}^\top \boldsymbol{Av}_2 & \dots & \boldsymbol{y}^\top \boldsymbol{Av}_n \\ 0 & \boldsymbol{u}_2^\top \boldsymbol{Av}_2 & \dots & \boldsymbol{u}_2^\top \boldsymbol{Av}_n \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \boldsymbol{u}_m^\top \boldsymbol{Av}_2 & \dots & \boldsymbol{u}_m^\top \boldsymbol{Av}_n \end{bmatrix} \\ &=: \begin{bmatrix} d_1 & \boldsymbol{w}^\top \\ \mathbf{0} & \boldsymbol{B} \end{bmatrix} \end{aligned}$$

と書き直せる. ただし

$$\begin{aligned} \boldsymbol{w} &= \begin{bmatrix} \boldsymbol{y}^\top \boldsymbol{Av}_2 \\ \boldsymbol{y}^\top \boldsymbol{Av}_3 \\ \vdots \\ \boldsymbol{y}^\top \boldsymbol{Av}_n \end{bmatrix} \in \mathbb{R}^{n-1}, \\ \boldsymbol{B} &= \begin{bmatrix} \boldsymbol{u}_2^\top \boldsymbol{Av}_2 & \boldsymbol{u}_2^\top \boldsymbol{Av}_3 & \dots & \boldsymbol{u}_2^\top \boldsymbol{Av}_n \\ \boldsymbol{u}_3^\top \boldsymbol{Av}_2 & \boldsymbol{u}_3^\top \boldsymbol{Av}_3 & \dots & \boldsymbol{u}_3^\top \boldsymbol{Av}_n \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{u}_m^\top \boldsymbol{Av}_2 & \boldsymbol{u}_m^\top \boldsymbol{Av}_3 & \dots & \boldsymbol{u}_m^\top \boldsymbol{Av}_n \end{bmatrix} \in \text{Mat}(m-1, n-1; \mathbb{R}) \end{aligned}$$

である. 上記の形に注意して計算すると

$$\mathbf{A}_1 \begin{bmatrix} d_1 \\ \mathbf{w} \end{bmatrix} = \begin{bmatrix} d_1 & \mathbf{w}^\top \\ \mathbf{0} & \mathbf{B} \end{bmatrix} \begin{bmatrix} d_1 \\ \mathbf{w} \end{bmatrix} = \begin{bmatrix} d_1^2 + \mathbf{w}^\top \mathbf{w} \\ \mathbf{B}\mathbf{w} \end{bmatrix}$$

となる. したがって

$$\begin{aligned} \left| \mathbf{A}_1 \begin{bmatrix} d_1 \\ \mathbf{w} \end{bmatrix} \right|_2^2 &= \left| \begin{bmatrix} d_1^2 + \mathbf{w}^\top \mathbf{w} \\ \mathbf{B}\mathbf{w} \end{bmatrix} \right|_2^2 = \{d_1^2 + \mathbf{w}^\top \mathbf{w}\}^2 + \mathbf{w}^\top \mathbf{B}^\top \mathbf{B}\mathbf{w} \\ &\geq \{d_1^2 + |\mathbf{w}|_2^2\}^2 \end{aligned}$$

から

$$\left| \mathbf{A}_1 \begin{bmatrix} d_1 \\ \mathbf{w} \end{bmatrix} \right|_2 \geq d_1^2 + |\mathbf{w}|_2^2, \quad \left| \frac{1}{\sqrt{d_1^2 + |\mathbf{w}|_2^2}} \begin{bmatrix} d_1 \\ \mathbf{w} \end{bmatrix} \right|_2 = 1$$

がわかる. よって

$$\begin{aligned} \|\mathbf{A}_1\| &= \sup_{\mathbf{x} \in \mathbb{R}^n, |\mathbf{x}|_2=1} |\mathbf{A}_1 \mathbf{x}|_2 \geq \frac{1}{\sqrt{d_1^2 + |\mathbf{w}|_2^2}} \left| \mathbf{A}_1 \begin{bmatrix} d_1 \\ \mathbf{w} \end{bmatrix} \right|_2 \\ &\geq \sqrt{d_1^2 + |\mathbf{w}|_2^2}. \end{aligned} \tag{E.1}$$

直交変換に関して $\|\cdot\|$ は不変²なので,

$$d_1 = \|\mathbf{A}\| = \|\mathbf{A}_1\|.$$

である. これと (E.1) と合わせると

$$d_1 = \|\mathbf{A}\| = \|\mathbf{A}_1\| \geq \sqrt{d_1^2 + |\mathbf{w}|_2^2} \geq d_1$$

となるので,

$$|\mathbf{w}|_2 = 0 \Leftrightarrow \mathbf{w} = \mathbf{0}.$$

よって

$$\mathbf{U}_1^\top \mathbf{A} \mathbf{V}_1 = \mathbf{A}_1 = \begin{bmatrix} d_1 & \mathbf{0} \\ \mathbf{0}^\top & \mathbf{B} \end{bmatrix}$$

² $\forall \mathbf{x} \in \mathbb{R}^n$ ($|\mathbf{x}|_2 = 1$) に対して $|\mathbf{V}_1 \mathbf{x}|_2 = 1$ に注意すれば,

$$\begin{aligned} \|\mathbf{A}\| &= \|\mathbf{U}_1 \mathbf{A}_1 \mathbf{V}_1\| = \sup_{\mathbf{x} \in \mathbb{R}^n; |\mathbf{x}|_2=1} |\mathbf{U}_1 \mathbf{A}_1 \mathbf{V}_1 \mathbf{x}|_2 = \sup_{\mathbf{x} \in \mathbb{R}^n; |\mathbf{x}|_2=1} |\mathbf{U}_1 \mathbf{A}_1 \mathbf{x}|_2 \\ &= \sup_{\mathbf{x} \in \mathbb{R}^n; |\mathbf{x}|_2=1} \sqrt{\mathbf{x}^\top \mathbf{A}_1^\top \mathbf{U}_1^\top \mathbf{U}_1 \mathbf{A}_1 \mathbf{x}} = \sup_{\mathbf{x} \in \mathbb{R}^n; |\mathbf{x}|_2=1} \sqrt{\mathbf{x}^\top \mathbf{A}_1^\top \mathbf{A}_1 \mathbf{x}} \\ &= \sup_{\mathbf{x} \in \mathbb{R}^n; |\mathbf{x}|_2=1} |\mathbf{A}_1 \mathbf{x}|_2 = \|\mathbf{A}_1\| \end{aligned}$$

からわかる.

と書けることがわかった.

次に, $d_2 = \|B\|$ とおけば,

$$\begin{aligned} d_2 &= \sup_{\mathbf{x} \in \mathbb{R}^{n-1}; |\mathbf{x}|_2=1} |\mathbf{B}\mathbf{x}|_2 = \sup_{\mathbf{x} \in \mathbb{R}^{n-1}; |\mathbf{x}|_2=1} \left\| \begin{bmatrix} d_1 & \mathbf{0}^\top \\ \mathbf{0} & \mathbf{B} \end{bmatrix} \begin{bmatrix} 0 \\ \mathbf{x} \end{bmatrix} \right\|_2 \\ &\leq \sup_{\mathbf{z} \in \mathbb{R}^n; |\mathbf{z}|_2=1} |\mathbf{A}_1\mathbf{z}|_2 = \|\mathbf{A}_1\| = \|\mathbf{A}\| = d_1. \end{aligned}$$

$d_2 = 0$ のとき, $\mathbf{B} = \mathbf{0}$ より証明はおわり. $d_2 > 0$ と仮定して議論を進める. 前と同じようにすれば,

$$d_2 = \|\mathbf{B}\|$$

とおき,

$$|\mathbf{B}\mathbf{x}|_2 = 1, \mathbf{x} \in \mathbb{R}^{n-1} \text{ s.t. } |\mathbf{x}|_2 = 1, \mathbf{y} = \frac{1}{d_2}\mathbf{B}\mathbf{x} \in \mathbb{R}^{m-1}$$

と取り, さきほどの議論を繰り返すと $\tilde{U}_2 \in \text{Mat}(m-1; \mathbb{R})$, $\tilde{V}_2 \in \text{Mat}(n-1; \mathbb{R})$ で

$$\tilde{U}_2^\top \mathbf{B} \tilde{V}_2 = \begin{bmatrix} d_2 & \mathbf{0}^\top \\ \mathbf{0} & \mathbf{C} \end{bmatrix}, \mathbf{C} \in \text{Mat}(m-2, n-2; \mathbb{R})$$

とできて,

$$\mathbf{U}_2 = \begin{bmatrix} 1 & \mathbf{0}^\top \\ \mathbf{0} & \tilde{U}_2 \end{bmatrix}, \mathbf{V}_2 = \begin{bmatrix} 1 & \mathbf{0}^\top \\ \mathbf{0} & \tilde{V}_2 \end{bmatrix}$$

とおけば, $\mathbf{U}_2, \mathbf{V}_2$ は直交行列で

$$\mathbf{U}_2^\top \mathbf{U}_1^\top \mathbf{A} \mathbf{V}_1 \mathbf{V}_2 = \begin{bmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & \mathbf{C} \end{bmatrix}$$

となる. さらに $\mathbf{U}_1 \mathbf{U}_2 \in \text{Mat}(m; \mathbb{R})$ と $\mathbf{V}_1 \mathbf{V}_2 \in \text{Mat}(n; \mathbb{R})$ も直交行列であることを注意する.

以上の操作を繰り返せば, 定理は証明される. □

定理 E.9. $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{V}^\top$ と特異値分解されたとする. ただし, $p \leq \min(m, n)$ に対し

$$d_1 \geq d_2 \geq \cdots \geq d_p > d_{p+1} = d_{p+2} = \cdots = d_{\min(m, n)} = 0$$

とする.

$$\mathbf{U} = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m), \mathbf{V} = (\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n)$$

としたとき,

$$\begin{aligned}\ker(\mathbf{A}) &= \text{span}\{\mathbf{v}_{p+1}, \mathbf{v}_{p+2}, \dots, \mathbf{v}_n\} = \mathbf{R}(\mathbf{A})^\perp, \\ \ker(\mathbf{A}^\top) &= \text{span}\{\mathbf{u}_{p+1}, \mathbf{u}_{p+2}, \dots, \mathbf{u}_m\} = \mathbf{R}(\mathbf{A}^\top)^\perp.\end{aligned}$$

Proof. \mathbf{U} の直交性より

$$\mathbf{A}\mathbf{x} = \mathbf{0} \ (x \in \mathbb{R}^n) \Leftrightarrow \mathbf{D}\mathbf{V}^\top\mathbf{x} = \begin{bmatrix} d_1\mathbf{v}_1^\top\mathbf{x} \\ d_2\mathbf{v}_2^\top\mathbf{x} \\ \vdots \\ d_p\mathbf{v}_p^\top\mathbf{x} \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \mathbf{0}.$$

よって $\ker(\mathbf{A})$ の任意の元 \mathbf{x} は $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_p$ と直交するので,

$$\ker(\mathbf{A}) = \text{span}\{\mathbf{v}_{p+1}, \mathbf{v}_{p+2}, \dots, \mathbf{v}_n\}.$$

さらに $\mathbf{A}^\top = \mathbf{V}\mathbf{D}^\top\mathbf{U}^\top$ より

$$\ker(\mathbf{A}^\top) = \text{span}\{\mathbf{u}_{p+1}, \mathbf{u}_{p+2}, \dots, \mathbf{u}_m\} = \mathbf{R}(\mathbf{A})^\perp.$$

一方, $\forall \tilde{\mathbf{x}} \in \mathbf{R}(\mathbf{A}^\top)$ とすれば, ある $\mathbf{y} \in \mathbb{R}^m$ があって, $\tilde{\mathbf{x}} = \mathbf{A}^\top\mathbf{y}$ とかける. このことより,

$$\mathbf{x} \in \ker(\mathbf{A}) \Leftrightarrow \mathbf{x}^\top\tilde{\mathbf{x}} = (\mathbf{A}\mathbf{x})^\top\mathbf{y} = 0$$

より, \mathbf{x} は $\mathbf{R}(\mathbf{A}^\top)$ と直交する. 同様に, $\forall \tilde{\mathbf{y}} \in \mathbf{R}(\mathbf{A})$ をとる. ある $\mathbf{x} \in \mathbb{R}^n$ があって, $\tilde{\mathbf{y}} = \mathbf{A}\mathbf{x}$ とかける. このことより

$$\mathbf{y} \in \ker(\mathbf{A}^\top) \Leftrightarrow \mathbf{y}^\top\tilde{\mathbf{y}} = (\mathbf{A}^\top\mathbf{y})^\top\mathbf{x} = 0$$

より, \mathbf{y} は $\mathbf{R}(\mathbf{A})$ と直交することもわかる. □

注意 E.10.

$$\mathbf{U}_1 = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_p], \mathbf{U}_2 = [\mathbf{u}_{p+1}, \mathbf{u}_{p+2}, \dots, \mathbf{u}_m],$$

$$\mathbf{V}_1 = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_p], \mathbf{V}_2 = [\mathbf{v}_{p+1}, \mathbf{v}_{p+2}, \dots, \mathbf{v}_n]$$

とおく. すなわち

$$\mathbf{U} = [\mathbf{U}_1, \mathbf{U}_2], \quad \mathbf{V} = [\mathbf{V}_1, \mathbf{V}_2]$$

である. このとき, \mathbb{R}^n から $\ker(\mathbf{A})$ への直交射影を \mathbf{P} , \mathbb{R}^m から $\ker(\mathbf{A}^\top)$ への直交射影を \mathbf{Q} とおくと

$$\mathbf{P} = \mathbf{V}_2\mathbf{V}_2^\top, \quad \mathbf{Q} = \mathbf{U}_2\mathbf{U}_2^\top$$

となる. □

E.2 Moore-Penrose の一般化逆行列

$A \in \text{Mat}(m, n; \mathbb{R})$ で $A \neq \mathbf{0}$ とする. 定理 E.7 から $1 \leq p \leq \min(m, n)$ があって

$$A = \tilde{U} \tilde{D} \tilde{V}^\top,$$

$\tilde{U} \in \text{Mat}(m; \mathbb{R}), \tilde{V} \in \text{Mat}(n; \mathbb{R})$ は直交行列,
 $\tilde{D} \in \text{Mat}(m, n; \mathbb{R})$ は対角行列 (非対角成分は 0) で
 その正の対角成分は $d_1 \geq d_2 \geq \cdots \geq d_p > 0$

と書ける. ここで $\tilde{U}, \tilde{D}, \tilde{V}$ を分解し

$$\begin{aligned} \tilde{U} &= [U, U^\perp], U \in \text{Mat}(m, p; \mathbb{R}), \\ \tilde{V} &= [V, V^\perp], V \in \text{Mat}(n, p; \mathbb{R}), \\ D &= \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_p \end{bmatrix} \in \text{Mat}(p; \mathbb{R}) \end{aligned}$$

とする. すると

$$\begin{aligned} \tilde{U}^\top \tilde{U} &= \begin{bmatrix} U^\top U & U^\top U^\perp \\ (U^\perp)^\top U & (U^\perp)^\top U^\perp \end{bmatrix} = I_m, \\ \tilde{V}^\top \tilde{V} &= \begin{bmatrix} V^\top V & V^\top V^\perp \\ (V^\perp)^\top V & (V^\perp)^\top V^\perp \end{bmatrix} = I_n \end{aligned}$$

から

$$U^\top U = I_p, U^\top U^\perp = \mathbf{0}, V^\top V = I_p, V^\top V^\perp = \mathbf{0}$$

となる. よって

$$A = \tilde{U} \tilde{D} \tilde{V}^\top = U D V^\top \tag{E.2}$$

となる. ただし

$$U^\top U = V^\top V = I_p, D = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_p \end{bmatrix}$$

となる.

(E.2) を踏まえて, $A^\dagger \in \text{Mat}(n, m; \mathbb{R})$ を

$$A^\dagger := VD^{-1}U^\top, D^{-1} = \begin{bmatrix} d_1^{-1} & 0 & \cdots & 0 \\ 0 & d_2^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_p^{-1} \end{bmatrix} \quad (\text{E.3})$$

と定める. A^\dagger を A の Moore-Penrose の一般逆行列と呼ぶことにする.
すると A^\dagger は下記の関係式をみたす $n \times m$ の行列となる.

$$AA^\dagger A = A, \quad (\text{E.4})$$

$$A^\dagger AA^\dagger = A^\dagger, \quad (\text{E.5})$$

$$AA^\dagger = (AA^\dagger)^\top, \quad (\text{E.6})$$

$$A^\dagger A = (A^\dagger A)^\top \quad (\text{E.7})$$

が成立する. 上記 (E.4) – (E.7) は (E.2) と (E.3) よりわかる. 実際

$$\begin{aligned} AA^\dagger A &= UDV^\top VD^{-1}U^\top UDV^\top = UDV^\top = A, \\ A^\dagger AA^\dagger &= VD^{-1}U^\top UDV^\top VD^{-1}U^\top = VD^{-1}U^\top = A^\dagger, \\ AA^\dagger &= UDV^\top VD^{-1}U^\top = UU^\top = (UU^\top)^\top = (AA^\dagger)^\top, \\ A^\dagger A &= VD^{-1}U^\top UDV^\top = VV^\top = (VV^\top)^\top = (A^\dagger A)^\top \end{aligned}$$

からわかる. さらに

$$\begin{aligned} P &= AA^\dagger : \mathbb{R}^m \rightarrow \text{ran}(A \subset \mathbb{R}^m), \\ Q &= A^\dagger A : \mathbb{R}^n \rightarrow \text{ran}(A^\top) \subset \mathbb{R}^n \end{aligned}$$

は射影行列になっている. すなわち $P \in \text{Mat}(m; \mathbb{R})$ と $Q \in \text{Mat}(n; \mathbb{R})$

$$P^2 = P, P^\top = P, Q^2 = Q, Q^\top = Q$$

をみたしている.

最後に A^\dagger の一意性を示す. B^\dagger も A の Moore-Penrose の一般逆行列とする. すると A^\dagger を B と替えた

$$ABA = A, \quad (\text{E.8})$$

$$BAB = B, \quad (\text{E.9})$$

$$AB = (AB)^\top, \quad (\text{E.10})$$

$$BA = (BA)^\top \quad (\text{E.11})$$

も成立する. すると

$$\begin{aligned}
 A^\dagger &= A^\dagger A A^\dagger && (\because (E.4)) \\
 &= A^\dagger (A B A) A^\dagger && (\because (E.8)) \\
 &= A^\dagger (A B)^\top (A A^\dagger)^\top && (\because (E.6) \text{ と } (E.10)) \\
 &= A^\dagger B^\top (A A^\dagger A)^\top \\
 &= A^\dagger B^\top A^\top && (\because (E.4)) \\
 &= A^\dagger (A B)^\top \\
 &= A^\dagger A B && (\because (E.10)) \\
 &= A^\dagger A B A B && (\because (E.8)) \\
 &= (A^\dagger A)^\top (B A)^\top B && (\because (E.7) \text{ と } (E.11)) \\
 &= (A A^\dagger A)^\top B^\top B \\
 &= A^\top B^\top B && (\because (E.4)) \\
 &= (B A)^\top B \\
 &= B A B && (\because (E.11)) \\
 &= B && (\because (E.8))
 \end{aligned}$$

からわかる. よって一意性を証明できた.

注意 E.11. 多くの本では, $A \in \text{Mat}(m, n; \mathbb{R})$ に対して (E.4) – (E.7) をみたす $n \times m$ の行列 A^\dagger を A の Moore-Penrose の一般逆行列と定義している. □

E.3 章末注釈と参考文献