ベクトル空間と部分空間

部分空間: n 次元実数ベクトル全体の集合を R^n で表すとき，R^n の部分集合 W が

1. $\boldsymbol{a} \in W$, $\boldsymbol{b} \in W$ ならば $\boldsymbol{a} + \boldsymbol{b} \in W$
2. $\boldsymbol{a} \in W$ ならば $\alpha \boldsymbol{a} \in W$ (α は実数)

をみたすとき，W を**線形部分空間**と呼ぶ.

独立・直和: 2 つの部分空間 W_A, W_B に関し，$W_A \cap W_B = \{\mathbf{0}\}$ のとき，すなわち W_A と W_B の共通部分が零ベクトルのみからなるとき，その部分空間は**独立**であるという. このとき，W_A と W_B の和空間 $W_{A \cup B}$ を

$$W_{A \cup B} = W_A \oplus W_B$$

と書き，$W_{A \cup B}$ は W_A と W_B の**直和**に分解されるという.

補空間

補空間: 全空間 R^n が V と W の直和で表される，すなわち

$$R^n = V \oplus W$$

が成り立つとき，W は V の**補空間**(または V はW の補空間) といい，$W = V^C$ (または $V = W^C$) と書く．

直交補空間: W が V の補空間であり，さらに W に属する任意のベクトルが V に属する任意のベクトルと直交するとき，

$$W = V^{\perp} \text{ (または } V = W^{\perp}\text{)}$$

と書き，W は V の(または V は W の) **直交補空間**であるという．ただし，$V^{\perp}$ の定義は次式である．

$$V^{\perp} = \{\boldsymbol{a} \mid \boldsymbol{a}^T\boldsymbol{b} = 0, \quad \forall \boldsymbol{b} \in V\}$$

列空間と零空間 (核)

- $m \times n$ 行列 A を考える.
- A の**列空間**:

$$\mathrm{Im}A = \{\boldsymbol{x} \mid \boldsymbol{x} = A\boldsymbol{z}, \quad \forall \boldsymbol{z} \in R^n\} \subset R^m$$

このとき,

$$\boldsymbol{x} = A\boldsymbol{z} = \begin{bmatrix} \boldsymbol{a}_1 & \cdots & \boldsymbol{a}_n \end{bmatrix} \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} = z_1\boldsymbol{a}_1 + \cdots + z_n\boldsymbol{a}_n$$

すなわち, $\mathrm{Im}A$ は A の列ベクトルによって張られる空間である.

- A^T の**零空間 (核)**:

$$\mathrm{Ker}A^T = \{\boldsymbol{x} \mid \boldsymbol{0} = A^T\boldsymbol{x}\} \subset R^m$$

すなわち, $\mathrm{Ker}A^T$ の要素は A^T のすべての行ベクトル (A の列ベクトルに等しい) に直交する

線形代数の基本定理

定理:

$$\mathrm{Ker}A^T = (\mathrm{Im}A)^\perp$$

証明: $\boldsymbol{y}_1 \in \mathrm{Ker}A^T$ とするとき，任意の$\boldsymbol{y}_2 = A\boldsymbol{x}_2 \in \mathrm{Im}A$ に対して$\boldsymbol{y}_1^T\boldsymbol{y}_2 = \boldsymbol{y}_1^T A\boldsymbol{x}_2 = (A^T\boldsymbol{y}_1)^T\boldsymbol{x}_2 = \boldsymbol{0}$ となる．したがって $\boldsymbol{y}_1 \in \mathrm{Ker}A^T \Longrightarrow \boldsymbol{y}_1 \in (\mathrm{Im}A)^\perp$.

逆に，$\boldsymbol{y}_1 \in (\mathrm{Im}A)^\perp$とすると，任意の$n$次元ベクトル$\boldsymbol{x}_2$に対して$A\boldsymbol{x}_2 \in \mathrm{Im}A$であるから

$$\boldsymbol{0} = \boldsymbol{y}_1^T A\boldsymbol{y}_2 = (A^T\boldsymbol{y}_1)^T\boldsymbol{y}_2$$

となり，$\boldsymbol{y}_1 \in \mathrm{Ker}A^T$を得る．したがって$\boldsymbol{y}_1 \in (\mathrm{Im}A)^\perp \Longrightarrow \boldsymbol{y}_1 \in \mathrm{Ker}A^T$．(証明終わり)

- A^Tの行空間: $\mathrm{Im}A^T = \{\boldsymbol{x} \mid \boldsymbol{x} = A^T\boldsymbol{z}, \boldsymbol{z} \in R^m\} \subset R^n$
- Aの核: $\mathrm{Ker}A = \{\boldsymbol{x} \mid \boldsymbol{0} = A\boldsymbol{x}\} \subset R^n$

定理:

$$\mathrm{Ker}A = (\mathrm{Im}A^T)^\perp$$

列空間と零空間の例

- 正方行列

$$A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad \mathrm{Im}A = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}, \qquad \mathrm{Ker}A^T = \left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

$$A^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 0 \end{bmatrix}, \qquad \mathrm{Im}A^T = \left\{ \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}, \qquad \mathrm{Ker}A = \left\{ \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} \right\}$$

- 横長の場合

$$A = \begin{bmatrix} 1 & 1 & 0 & 2 \\ 2 & 0 & 1 & 4 \\ 3 & 3 & 0 & 6 \end{bmatrix}, \qquad \mathrm{Im}A = \left\{ \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}, \qquad \mathrm{Ker}A^T = \left\{ \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} \right\}$$

$$A^T = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 0 & 3 \\ 0 & 1 & 0 \\ 2 & 4 & 6 \end{bmatrix}, \qquad \mathrm{Im}A^T = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ -2 \\ 1 \\ 0 \end{bmatrix} \right\}, \qquad \mathrm{Ker}A = \left\{ \begin{bmatrix} 2 \\ 0 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ -2 \\ 0 \end{bmatrix} \right\}$$

$\boldsymbol{y} \approx A\boldsymbol{x}$ の解 (例)

- $\boldsymbol{y} \notin \mathrm{Im}A$ のとき

$$\boldsymbol{y} = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}, \qquad A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad \boldsymbol{x} = \begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix}?, \quad \begin{bmatrix} 0 \\ 3 \\ 1 \end{bmatrix}?, \quad \begin{bmatrix} 0.4 \\ 3 \\ 0.8 \end{bmatrix}?$$

- A が横長の場合

$$\boldsymbol{y} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \qquad A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \end{bmatrix}, \qquad \boldsymbol{x} = [?]$$

$$\boldsymbol{y} = \begin{bmatrix} 1 \\ 4 \\ 5 \end{bmatrix}, \qquad A = \begin{bmatrix} 1 & 1 & 0 & 2 \\ 2 & 0 & 1 & 4 \\ 3 & 3 & 0 & 6 \end{bmatrix}, \qquad \boldsymbol{x} = \begin{bmatrix} 1 \\ 0 \\ 2 \\ 0 \end{bmatrix}??$$

- A が縦長の場合

$$\boldsymbol{y} = \begin{bmatrix} 2 \\ 3 \\ 2 \end{bmatrix}, \qquad A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 2 & 0 \end{bmatrix}, \qquad \boldsymbol{x} = [?]$$

$\boldsymbol{y} \approx A\boldsymbol{x}$の解 ($\boldsymbol{y} \notin \mathrm{Im}A$ のとき)

- $\boldsymbol{y} \notin \mathrm{Im}A$ のとき，解は存在しない
  - $\boldsymbol{r} = \boldsymbol{y} - A\boldsymbol{x}$ とおき，$\|\boldsymbol{r}\|$ を最小にする $\boldsymbol{x}$ を解と定義する
  - $e = \|\boldsymbol{r}\|^2$ とおくと $e = (\boldsymbol{y} - A\boldsymbol{x})^T(\boldsymbol{y} - A\boldsymbol{x})$ であるから

$$\frac{\partial e}{\partial \boldsymbol{x}} = -2A^T(\boldsymbol{y} - A\boldsymbol{x}) = -2A^T\boldsymbol{r}$$

    となる．
  - 解 $\boldsymbol{x}_r$ は $\left.\dfrac{\partial e}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x}=\boldsymbol{x}_r} = \boldsymbol{0}$ を満たす．すなわち $\boldsymbol{x} = \boldsymbol{x}_r$ のとき $A^T\boldsymbol{r} = \boldsymbol{0}$ となり

$$\boldsymbol{y} - A\boldsymbol{x}_r \in \mathrm{Ker}A^T$$

    を得る
  - つまり $\boldsymbol{y} = \boldsymbol{y}_r + \boldsymbol{r}$，ただし $\boldsymbol{y}_r \in \mathrm{Im}A$, $\boldsymbol{r} \in \mathrm{Ker}A^T$ と直交分解したときの $\boldsymbol{y}_r$ を用いて

$$\boldsymbol{y}_r = A\boldsymbol{x}$$

    の解 $\boldsymbol{x}_r$ が最適解となる

$\boldsymbol{y} = A\boldsymbol{x}$ の解 ($\boldsymbol{y} \in \mathrm{Im}A$ のとき)

$$\boldsymbol{y} = A\boldsymbol{x} = \begin{bmatrix} \boldsymbol{a}_1 & \cdots & \boldsymbol{a}_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1\boldsymbol{a}_1 + \cdots + x_n\boldsymbol{a}_n$$

- A が正方かつフルランクの場合，解は一意に求まる

$$\boldsymbol{x} = A^{-1}\boldsymbol{y}$$

- A が縦長かつフルランクの場合，冗長な方程式が存在する
  - 冗長な方程式を除いた連立方程式を解けば良い

$$\tilde{\boldsymbol{y}} = \tilde{A}\boldsymbol{x}$$

  - 解は一意に求まる (もちろん上の $\boldsymbol{x} = \tilde{A}^{-1}\tilde{\boldsymbol{y}}$ に一致する)

$$\boldsymbol{x} = (A^TA)^{-1}A^T\boldsymbol{y}$$

- A が縦長 (正方を含む) であるがフルランクでない場合，冗長な方程式を除けば方程式の数より変数の数が多くなるので解は一意には求まらない
- A が横長の場合，方程式の数より変数の数が多いので解は一意には求まらない

$\boldsymbol{y} = A\boldsymbol{x}$ の解 ($\boldsymbol{y} \in \mathbf{Im}A$ かつ A がフルランクでない場合)

- 解が一意には求まらない場合,「$\boldsymbol{y} = A\boldsymbol{x}$ をみたすすべての解 $\boldsymbol{x}$ のうち，$\|\boldsymbol{x}\|$ を最小にする $\boldsymbol{x}$」を解と定義する
  - $\boldsymbol{y} = A\boldsymbol{x}$ をみたすひとつの解を $\boldsymbol{x}_0$ とする
  - A の行空間を $\mathrm{Im}A^T$，零空間を $\mathrm{Ker}A$ とすると任意のベクトルは行空間上への射影と零空間上への射影とに直交分解できる

$$\boldsymbol{x}_0 = \boldsymbol{x}_r + \boldsymbol{w} \qquad \boldsymbol{x}_r^T \boldsymbol{w} = 0, \qquad \boldsymbol{x}_r \in \mathrm{Im}A^T,\ \ \boldsymbol{w} \in \mathrm{Ker}A$$

  - 任意の $\boldsymbol{w} \in \mathrm{Ker}A$ に対し $A\boldsymbol{w} = \boldsymbol{0}$ が成り立つから,

$$A\boldsymbol{x}_0 = A(\boldsymbol{x}_r + \boldsymbol{w}) = A\boldsymbol{x}_r = \boldsymbol{y}$$

  - Pythagoras の定理より

$$\|\boldsymbol{x}_0\|^2 = \|\boldsymbol{x}_r + \boldsymbol{w}\|^2 = \|\boldsymbol{x}_r\|^2 + \|\boldsymbol{w}\|^2$$

  - $\boldsymbol{w} = 0$ のとき $\|\boldsymbol{x}_0\|$ が最小になるから，最適解は $\boldsymbol{x}_r \in \mathrm{Im}A^T$

$\boldsymbol{y} \approx A\boldsymbol{x}$ の解 (まとめ)

- $\boldsymbol{y} \in R^m$ の直交分解

$$\boldsymbol{y} = \boldsymbol{y}_r + \boldsymbol{r} \qquad \boldsymbol{y}_r \in \mathrm{Im}A,\ \boldsymbol{r} \in \mathrm{Ker}A^T$$

- $\boldsymbol{y}_r = A\boldsymbol{x}$ を解く
  - A がフルランクなら
    - A が縦長の時 ($A^T\boldsymbol{r} = \boldsymbol{0}$)

$$\boldsymbol{x} = (A^TA)^{-1}A^T\boldsymbol{y}_r = (A^TA)^{-1}A^T\boldsymbol{y}$$

    - A が横長の時 (A がフルランクだから $\boldsymbol{r} = \boldsymbol{0}$)
      - $\boldsymbol{x}_r \in \mathrm{Im}A^T$ ゆえ $\boldsymbol{x}_r = A^T\boldsymbol{z}$ となる $\boldsymbol{z}$ が存在する
      - $\boldsymbol{y} = AA^T\boldsymbol{z}$
      - AA^T は正則だから $\boldsymbol{z} = (AA^T)^{-1}\boldsymbol{y}$

$$\boldsymbol{x}_r = A^T(AA^T)^{-1}\boldsymbol{y}$$

  - A がフルランクでない場合,
    - $\boldsymbol{y}_r = A\boldsymbol{x}_0$ をみたす $\boldsymbol{x}_0$ を見つけ
    - $\boldsymbol{x}_0$ を直交分解

$$\boldsymbol{x}_0 = \boldsymbol{x}_r + \boldsymbol{w} \qquad \boldsymbol{x}_r \in \mathrm{Im}A^T,\ \boldsymbol{w} \in \mathrm{Ker}A$$

$\boldsymbol{y} \approx A\boldsymbol{x}$ の解 (例)

$$\boldsymbol{y} = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}, \quad A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathrm{Im}A \ni \boldsymbol{y}_r = \begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix}, \quad \mathrm{Im}A^T \ni \boldsymbol{x} = \alpha \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} + \beta \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.4 \\ 3 \\ 0.8 \end{bmatrix}$$

$$\boldsymbol{y} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \quad A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \end{bmatrix}, \quad \boldsymbol{x} = \begin{bmatrix} 0.4 \\ 3 \\ 0.8 \end{bmatrix}$$

$$\boldsymbol{y} = \begin{bmatrix} 1 \\ 4 \\ 5 \end{bmatrix}, \quad A = \begin{bmatrix} 1 & 1 & 0 & 2 \\ 2 & 0 & 1 & 4 \\ 3 & 3 & 0 & 6 \end{bmatrix}, \quad \mathrm{Ker}A^T \ni \boldsymbol{y} - \boldsymbol{y}_r = \gamma \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}, \quad \boldsymbol{y}_r = \begin{bmatrix} 1.6 \\ 4 \\ 4.8 \end{bmatrix}$$

$$\mathrm{Im}A^T \ni \boldsymbol{x} = \alpha \begin{bmatrix} 1 \\ 1 \\ 0 \\ 2 \end{bmatrix} + \beta \begin{bmatrix} 0 \\ -2 \\ 1 \\ 0 \end{bmatrix}, \quad A\boldsymbol{x} = \begin{bmatrix} 6\alpha - 2\beta \\ 10\alpha + \beta \\ 18\alpha - 6\beta \end{bmatrix}, \quad \boldsymbol{x} = \begin{bmatrix} 0.3692 \\ -0.2462 \\ 0.3077 \\ 0.7385 \end{bmatrix}$$

疑似逆行列の性質

- 任意の $\boldsymbol{y} \in R^m$ に対し
  - $\|\boldsymbol{y} - A\boldsymbol{x}\|$ を最小にし
  - $\|\boldsymbol{x}\|$ を最小にする

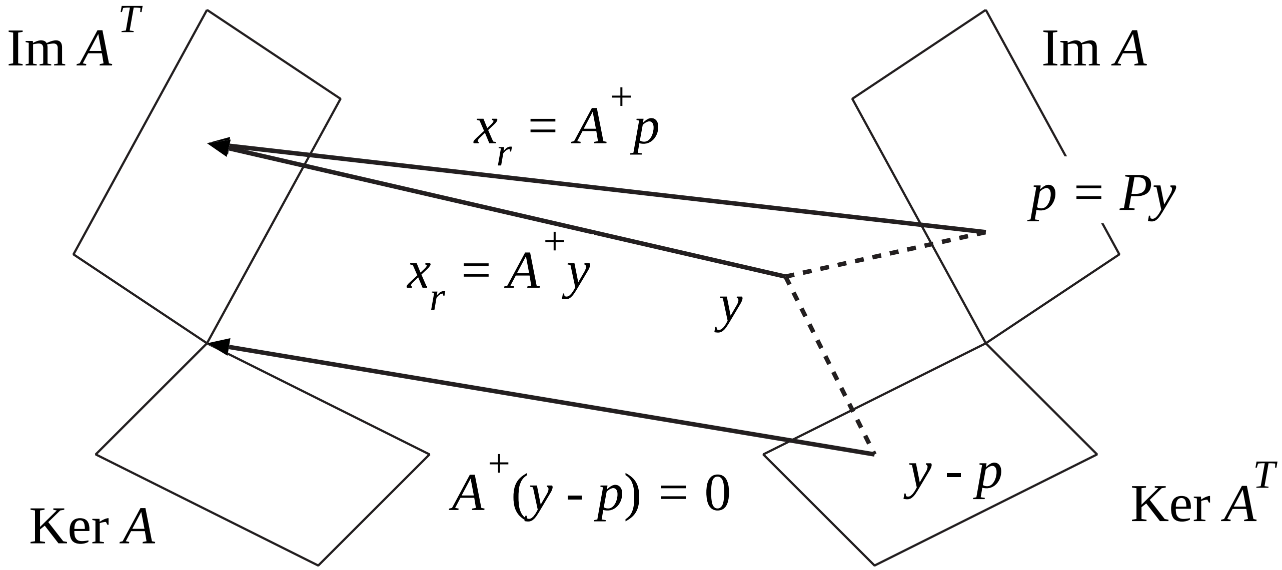
  最適解 $\boldsymbol{x}_r$ を生成するオペレータ A^+ を疑似逆行列という

$$\boldsymbol{x}_r = A^+\boldsymbol{y}$$

- A^+ の性質
  - 最適解は $\boldsymbol{x}_r \in \mathrm{Im}A^T$
  - $P = AA^+$ とおくと $P\boldsymbol{y} \in \mathrm{Im}A$
  - $\boldsymbol{y} - P\boldsymbol{y} = \boldsymbol{y} - AA^+\boldsymbol{y} = \boldsymbol{y} - A\boldsymbol{x}_r \in \mathrm{Ker}A^T$
    - ゆえに $P\boldsymbol{y}$ と $\boldsymbol{y} - P\boldsymbol{y}$ は $\boldsymbol{y}$ の直交分解
  - $P\boldsymbol{y} = A\boldsymbol{x}_r$
  - $A^+P\boldsymbol{y} = \boldsymbol{x}_r$
  - $A^+(\boldsymbol{y} - P\boldsymbol{y}) = \boldsymbol{0}$

A と A^+ の関係

- $\boldsymbol{x}_r \in \mathrm{Im} A^T$
- $P = AA^+$ とおくと $\boldsymbol{p} = P\boldsymbol{y} \in \mathrm{Im} A$
- $\boldsymbol{y} = \boldsymbol{p} + \boldsymbol{r}$ は直交分解$(\boldsymbol{p} \in \mathrm{Im} A,\ \boldsymbol{r} \in \mathrm{Ker} A^T)$
- $\boldsymbol{p} = A\boldsymbol{x}_r,\ A^+\boldsymbol{p} = \boldsymbol{x}_r,\ A^+(\boldsymbol{y} - \boldsymbol{p}) = \boldsymbol{0}$



疑似逆行列の例 (A がフルランクでない場合)

次の行列の疑似逆行列を求めたい.

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

- まず $\boldsymbol{y} = [y_1,\ y_2,\ y_3]^T$ とする.
- $P = AA^+$ のとき $\boldsymbol{p} = P\boldsymbol{y} \in \mathrm{Im}A,\ \boldsymbol{y} - \boldsymbol{p} \in \mathrm{Ker}A^T$
- ゆえに $\boldsymbol{p} = [y_1,\ y_2,\ 0]^T$
- 次に $A\boldsymbol{x} = \boldsymbol{p}$ を解くと $\boldsymbol{x} = [y_1,\ y_2,\ x_3]^T$ (x_3 は任意)
- $\boldsymbol{x}$ の長さを最小にするには $x_3 = 0$
- ゆえに

$$A^+ = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

疑似逆行列の例 (A がフルランクでない場合)

次の行列の疑似逆行列を求めたい.

$$A = \begin{bmatrix} \mu_1 & 0 & 0 & 0 \\ 0 & \mu_2 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad (\mu_1 > 0,\ \mu_2 > 0)$$

- $\boldsymbol{y} = [y_1,\ y_2,\ y_3]^T$ とすると，$\boldsymbol{p} \in \mathrm{Im}A,\ \boldsymbol{y} - \boldsymbol{p} \in \mathrm{Ker}A^T$ より $\boldsymbol{p} = P\boldsymbol{y} = [y_1,\ y_2,\ 0]^T$
- $A\boldsymbol{x} = \boldsymbol{p}$ を解くと $\boldsymbol{x} = [y_1/\mu_1,\ y_2/\mu_2,\ x_3,\ x_4]^T$
- $\boldsymbol{x}$ の長さを最小にするには $x_3 = x_4 = 0$
- ゆえに

$$A^+ = \begin{bmatrix} \mu_1^{-1} & 0 & 0 \\ 0 & \mu_2^{-1} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

特異値分解

特異値分解: 任意の $m \times n$ 行列は $A = U\Sigma V^T$ に分解することができる.

- ただし U は $m \times m$ 直交行列
- V は $n \times n$ 直交行列
- Σ は $m \times n$ 対角行列である.
- Σ の対角要素は零を含むときがある．たとえば，$m > n$ のとき次のような形になる.

$$\Sigma = \begin{bmatrix} \mu_1 & 0 & 0 & 0 & 0 \\ 0 & \mu_2 & 0 & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & 0 & \mu_r & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

疑似逆行列と特異値分解

疑似逆行列の公式: $m \times n$ 行列 A が特異値分解 $A = U\Sigma V^T$ をもつとき，A の疑似逆行列は $A^+ = V\Sigma^+ U^T$ で与えられる．

- ただし Σ^+ は次式で与えられる $n \times m$ 行列で Σ の疑似逆行列である．

$$\Sigma^+ = \begin{bmatrix} \mu_1^{-1} & 0 & 0 & 0 & 0 & 0 \\ 0 & \mu_2^{-1} & 0 & 0 & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 & 0 \\ 0 & 0 & 0 & \mu_r^{-1} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

証明: 上記の Σ^+ が Σ の疑似逆行列であることを用いる．

- まず $\|A\boldsymbol{x} - \boldsymbol{y}\| = \|U\Sigma V^T\boldsymbol{x} - \boldsymbol{y}\| = \|\Sigma V^T\boldsymbol{x} - U^T\boldsymbol{y}\|$
- 次に $\boldsymbol{z} = V^T\boldsymbol{x} = V^{-1}\boldsymbol{x}$ を導入すると，Σ^+ が Σ の疑似逆行列であるから，$\|\Sigma\boldsymbol{z} - U^T\boldsymbol{y}\|$ を最小化する解は $\boldsymbol{z} = \Sigma^+ U^T\boldsymbol{y}$
- したがって，

$$\boldsymbol{x} = V\boldsymbol{z} = V\Sigma^+ U^T\boldsymbol{y}$$

疑似逆行列

定義: 任意の $m \times n$ 実数行列 A に対し

1. $AA^+A = A$
2. $A^+AA^+ = A^+$
3. $(AA^+)^T = AA^+$
4. $(A^+A)^T = A^+A$

をみたす $n \times m$ 実数行列 A^+ がただ1つ存在する．この A^+ を A の疑似逆行列という．

- $\mathrm{rank}A = r$ とすると，A は適当な $m \times r$ 行列 B と $r \times n$ 行列 C の積として $A = BC$ と分解できる．このとき

$$D = C^T(CC^T)^{-1}(B^TB)^{-1}B^T$$

は上記の条件1から4を満足し，このような D は唯一であることが証明できる．

疑似逆行列の性質

1. $(A^+)^+ = A$
2. $(A^T)^+ = (A^+)^T, (AA^T)^+ = (A^+)^T A^+$
3. $m \times n$ 行列 A，$n \times p$ 行列 B に対して，一般には $(AB)^+ \neq B^+A^+$ であるが，$\mathrm{rank}A = \mathrm{rank}B = n$ ならば $(AB)^+ = B^+A^+$ となる．
4. $A^+ = (A^TA)^+A^T = A^T(AA^T)^+$
5. $m \times n$ 行列 A が $\mathrm{rank}A = m$ ならば $A^+ = A^T(AA^T)^{-1}$．また，$\mathrm{rank}A = n$ ならば $A^+ = (A^TA)^{-1}A^T$
6. $m \times n$ 行列 A の特異値分解を $A = U\Sigma V^T$ とする．$\mathrm{rank}A = r$ のとき，

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 & 0 & 0 & \\ 0 & \sigma_2 & 0 & 0 & \mathbf{0} \\ 0 & 0 & \ddots & 0 & \\ 0 & 0 & 0 & \sigma_r & \\ & \mathbf{0} & & & \mathbf{0} \end{bmatrix}$$

で与えられ，$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ をみたす．また，U, V は直交行列である．このとき A^+ は $A = V\Sigma^+U^T$ で与えられる．

連立方程式と疑似逆行列

連立一次方程式 $\boldsymbol{y} = A\boldsymbol{x}$ の解 $\boldsymbol{x}$ は...

- A が正則であれば $\boldsymbol{x} = A^{-1}\boldsymbol{y}$ である.
- Aが正則でない場合には,
  - $\boldsymbol{y} \in \mathrm{Im}A$ であれば解は存在するが，一意とは限らない.
  - $\boldsymbol{y} \notin \mathrm{Im}A$ であれば解は存在しない.
- ただし $\|\boldsymbol{y} - A\boldsymbol{x}\|$ を最小にする解は

$$\boldsymbol{x} = A^{+}\boldsymbol{y} + (I - A^{+}A)\boldsymbol{k}$$

  で与えられる ($\boldsymbol{k}$ は任意のベクトル).
- 解のうちで解自身のノルムを最小にするものは

$$\boldsymbol{x} = A^{+}\boldsymbol{y}$$

  である.

特異値分解 (少し詳細に)

特異値: 任意の $m \times n$ 実数行列 A が与えられたとする．このとき $A^T A$ は準正定行列 (すべての固有値は正または零) となる．$A^T A$ の固有値 λ_i $(i = 1, \ldots, n)$ を大きい順に並べて $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ とする．このとき，$\mathrm{rank} A \leq \min(m, n)$ に注意し，

$$\sigma_i = \sqrt{\lambda_i} \qquad (i = 1, \ldots, \min(m, n))$$

と定義する．$\sigma_1 \geq \cdots \geq \sigma_{\min(m,n)} \geq 0$ を A の特異値という．

特異値分解: $m \times m$ 直交行列 U と $n \times n$ 直交行列 V が存在し，

$$A = U \Sigma V^T$$

が成立する．これを A の特異値分解という．ただし Σ は

$$\Sigma = \begin{cases} \left[\begin{array}{c} \begin{matrix} \sigma_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \sigma_n \end{matrix} \\ \hline \mathbf{0} \end{array} \right] & (m \geq n) \\ \\ \left[\begin{array}{c|c} \begin{matrix} \sigma_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \sigma_m \end{matrix} & \mathbf{0} \end{array} \right] & (m < n) \end{cases}$$

で与えられる $m \times n$ 行列である．

特異値分解の性質

- U, V は直交行列であるから

$$UU^T = U^TU = I_m, \qquad VV^T = V^TV = I_n$$

- $\boldsymbol{y} = A\boldsymbol{x}$ という線形変換を考える．$\boldsymbol{y}_U = U^T\boldsymbol{y}$, $\boldsymbol{x}_V = V^T\boldsymbol{x}$ とおくと,

$$\boldsymbol{y}_U = \Sigma\boldsymbol{x}_V \qquad \text{(ただし } \Sigma \text{ は対角行列)}$$

  を得る．したがって，$\boldsymbol{x}$ から $\boldsymbol{y}$ への任意の変換は,

  – V^T による $\boldsymbol{x}$ から $\boldsymbol{x}_V$ への長さをかえない直交変換

  – Σ による $\boldsymbol{x}_V$ から $\boldsymbol{y}_U$ への $\boldsymbol{x}_v$ の第 i 要素を σ_i 倍して $\boldsymbol{y}_U$ の第 i 要素とする変換 (ただし $\boldsymbol{x}_V \in R^n$, $\boldsymbol{y}_U \in R^m$ であることに注意)

  – そして U による $\boldsymbol{y}_U$ から $\boldsymbol{y}$ への長さをかえない直交変換

  の 3 つの変換に分解できることを示している．このように，特異値分解は線形変換の 1 つの基本的な性質を表現するものである．

特異値分解の求め方

1. まず，特異値を求める．
   (a) ただし A^TA と AA^T の固有値のうち 0 でないものの個数およびそれらの値は一致するので，$m < n$ ならば AA^T の，$m > n$ ならば A^TA の固有値を求める．
   (b) 求まった固有値を λ_i とすると，特異値は $\sigma_i = \sqrt{\lambda_i}$ である．
2. U, V を求めるための準備．
   特異値のうち 0 でないものから構成した対角行列

$$\Sigma_r = \begin{bmatrix} \sigma_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \sigma_r \end{bmatrix}$$

を考える．U および V の第 i 列ベクトルをそれぞれ $\boldsymbol{u}_i$, $\boldsymbol{v}_i$ と書く．つまり

$$U_r = [\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r], \quad V_r = [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_r]$$

とする．このとき

$$A = U_r \Sigma_r V_r^T$$

となる．$U_r^T U_r = I_r$, $V_r^T V_r = I_r$ であるから，

$$A^T A V_r = V_r \Sigma_r^2$$

が成り立つ．

特異値分解の求め方 (つづき)

3. V を求める.

   (a) A^TA の固有値は σ_i^2 であり，上式 ($A^TAV_r = V_r\Sigma_r^2$) は

   $$A^TA\boldsymbol{v}_i = \sigma_i^2\boldsymbol{v}_i$$

   と書き直せることに注意すると，σ_i^2 に対応する固有ベクトルとして $\boldsymbol{v}_i$ が求められる.

   (b) V の V_r 以外の部分は直交性が成り立つように任意に定める.

4. U を求める.

   (a) U_r は $U_r = AV_r\Sigma_r^{-1}$ より求める.

   (b) U_r 以外の U は直交性が成り立つように任意に定める.

特異値分解の例

次の行列の特異値分解を求めよ.

$$A = \begin{bmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix}$$

1. $A^T A$ の固有値の平方根を求める.

$$A^T A = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 3 & -2 \\ -1 & -2 & 3 \end{bmatrix}, \quad \det(\lambda I_3 - A^T A) = \lambda(\lambda - 3)(\lambda - 5) = 0$$

ゆえに $\lambda = 5, 3, 0$, $\sigma = \sqrt{5}, \sqrt{3}, 0$.

2. $\lambda_i \neq 0$ に対する固有ベクトルを求める.

$$\boldsymbol{v}_1 = [0,\ 1/\sqrt{2},\ -1/\sqrt{2}]^T, \qquad \boldsymbol{v}_2 = [2/\sqrt{6},\ -1/\sqrt{6},\ -1/\sqrt{6}]^T$$

ゆえに

$$V = \begin{bmatrix} 0 & 2/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ -1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \end{bmatrix}$$

特異値分解の例 (つづき)

3. $U_r = AV_r\Sigma_r^{-1}$ より

$$U_r = \begin{bmatrix} 2/\sqrt{10} & -1/\sqrt{10} & -1/\sqrt{10} & 2/\sqrt{10} \\ 0 & -1/\sqrt{2} & 1/\sqrt{2} & 0 \end{bmatrix}^T$$

ゆえに

$$U = \begin{bmatrix} 2/\sqrt{10} & 0 & 1/\sqrt{2} & 1/\sqrt{10} \\ -1/\sqrt{10} & -1/\sqrt{2} & 0 & 2/\sqrt{10} \\ -1/\sqrt{10} & 1/\sqrt{2} & 0 & 2/\sqrt{10} \\ 2/\sqrt{10} & 0 & -1/\sqrt{2} & 1/\sqrt{10} \end{bmatrix}$$