ロボット工学第二２００４　補遺

倉林大輔

8　特異値分解・擬似逆行列

8.1　特異値分解とは

いま、A を $m \times n$ の任意の実行列とする。$A^T A$ は準正定であり、その固有値（$det(\lambda I_n - A^T A) = 0$ の解）は非負の実数 $\lambda_1, \lambda_2, \cdots, \lambda_n$ であるとする。ただし、$\lambda_1 \geq \lambda_2 \geq \cdots \geq 0$ とする。また、$\sigma_i = \sqrt{\lambda_i}$ とおく（$i = 1, 2, \cdots, min(m, n)$）。

このとき、ある $m \times m$ 直交行列 U と $n \times n$ 直交行列 V が存在して、

$$
\begin{aligned}
A &= U \Sigma V^T \qquad (159) \\
\Sigma &= \begin{cases} \left[\begin{array}{ccc} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \\ \hline & 0 & \end{array}\right] & (m \geq n) \\ \left[\begin{array}{ccc|c} \sigma_1 & & & \\ & \ddots & & 0 \\ & & \sigma_n & \end{array}\right] & (m < n) \end{cases}
\end{aligned}
$$

と書ける。(159) を A の特異値分解、σ_i を A の特異値という。0 でない特異値の個数 r は $r = rank\, A$ で与えられる。

ここで、U,V は直交行列なので、

$$
\begin{aligned}
UU^T &= U^T U = I_m \\
VV^T &= V^T V = I_n
\end{aligned} \qquad (160)
$$

を満たす。

このことは、$\boldsymbol{y} = A\boldsymbol{x}$ という線形変換が、(i)V^T による長さを変えない直交変換、(ii)Σ による第 i 要素を σ_i 倍する方向を変えない変換、(iii)U による長さを変えない直交変換、の 3 つに分割されることを示している。

8.2　特異値分解の求めかた

$A^T A$ と AA^T の固有値のうち、0 でないものの個数と値は同じである。そのため、$m < n$ であれば AA^T、$m > n$ なら $A^T A$ の固有値を求めれば良い。固有値から特異値が求まる。

ついで、特異値のうち 0 でないもの（r 個ある）から、

$$
\Sigma_r = \begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_r \end{bmatrix} \qquad (161)
$$

という対角行列を作る。また、

$$
\begin{aligned}
U_r &= [\,\boldsymbol{u}_1,\, \boldsymbol{u}_2,\, \cdots,\, \boldsymbol{u}_r\,] \\
V_r &= [\,\boldsymbol{v}_1,\, \boldsymbol{v}_2,\, \cdots,\, \boldsymbol{v}_r\,]
\end{aligned}
$$

と書く。ただし、$\boldsymbol{u}_i$ および $\boldsymbol{v}_i$ はそれぞれ U, V の第 i 列ベクトルとする。すると、

$$
\begin{cases}
A = U_r \Sigma_r V_r^T \\
U_r^T U_r = I_r \\
V_r^T V_r = I_r
\end{cases}
\tag{162}
$$

なので、

$$
A^T A V_r = V_r \Sigma_r^2 \tag{163}
$$

$$
U_r = A V_r \Sigma_r^{-1} \tag{164}
$$

となる。

ここで、(163) を $A^T A \boldsymbol{v}_i = \boldsymbol{v}_i \sigma_i^2$ と要素で考えると、$\boldsymbol{v}_i$ が $A^T A$ の固有値 λ_i に対する、単位長さ固有ベクトルであることが分かる。

よって、$A^T A$ の固有ベクトルを求めれば、V_r が求まる。V の V_r 以外の部分 $\boldsymbol{v}_{r+1}, \boldsymbol{v}_{r+2}, \cdots, \boldsymbol{v}_n$ は、$V^T V = V V^T = I_n$ を満たすように適当に決めれば良い。

さらに、(164) より U_r が求まる。U の U_r 以外の部分は、V と同様に $U U^T = U^T U = I_m$ を満たすよう、適当に決めればよい。なお、(163)(164) の代わりに、

$$
\begin{aligned}
A A^T U_r &= U_r \Sigma_r^2 \\
V_r &= A^T U_r \Sigma_r^{-1}
\end{aligned}
$$

としても同じである。

計算例

$$
A = \begin{bmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix}
\tag{165}
$$

を特異値分解する。

$$
A^T A = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 3 & -2 \\ -1 & -2 & 3 \end{bmatrix}
\tag{166}
$$

であるので、

$$
\begin{aligned}
\lambda_1 &= 5 & \lambda_2 &= 3 & \lambda_3 &= 0 \\
\sigma_1 &= \sqrt{5} & \sigma_2 &= \sqrt{3} & \sigma_3 &= 0
\end{aligned}
\tag{167}
$$

である。また、$A^T A$ の固有ベクトルは、

$$
\begin{aligned}
\boldsymbol{v}_1 &= [\,0\ \frac{1}{\sqrt{2}}\ -\frac{1}{\sqrt{2}}\,]^T \\
\boldsymbol{v}_2 &= [\,\frac{2}{\sqrt{6}}\ -\frac{1}{\sqrt{6}}\ -\frac{1}{\sqrt{6}}\,]^T
\end{aligned}
$$

である。$V_r = [\boldsymbol{v}_1,\ \boldsymbol{v}_2]$ に対して、

$$\begin{aligned} U_r &= AV_r\Sigma_r^{-1} \\ &= \begin{bmatrix} \frac{2}{\sqrt{10}} & -\frac{1}{\sqrt{10}} & -\frac{1}{\sqrt{10}} & \frac{2}{\sqrt{10}} \\ 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{bmatrix}^T \end{aligned}$$

を得る。

以上から、A についての特異値分解は、

$$\Sigma = \left[\begin{array}{ccc} \sqrt{5} & & \\ & \sqrt{3} & \\ & & 0 \\ \hline & 0 & \end{array}\right] \qquad U = \begin{bmatrix} \frac{2}{\sqrt{10}} & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{10}} \\ -\frac{1}{\sqrt{10}} & -\frac{1}{\sqrt{2}} & 0 & \frac{2}{\sqrt{10}} \\ -\frac{1}{\sqrt{10}} & -\frac{1}{\sqrt{2}} & 0 & \frac{2}{\sqrt{10}} \\ \frac{2}{\sqrt{10}} & 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{10}} \end{bmatrix} \qquad V = \begin{bmatrix} 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix}$$

と得られる。

8.3 擬似逆行列

擬似逆行列 A^+ は、正方でない行列 A に対して逆行列的な性質を持つものとして定義する。

ある $m \times n$ 行列 A に対し、

$$AA^+A = A \tag{168}$$

$$A^+AA^+ = A^+ \tag{169}$$

$$(AA^+)^T = AA^+ \tag{170}$$

$$(A^+A)^T = A^+A \tag{171}$$

を全て満たす $n \times m$ 行列 A^+ がただ 1 つ存在する。これを擬似逆行列という。

証明

1. $A = 0$ のとき、$A^+ = 0$
2. $A \neq 0$ で $rank\ A = r$ とする。このとき、A は適当な $B(m \times r)$ と $C(r \times n)$ の積で $A = BC$ と書ける。この B,C で $n \times m$ 行列 D を $D = C^T(CC^T)^{-1}(B^TB)^{-1}B^T$ と表すと、D は A^+ の性質を全て満たしている。つまり、A に対して A^+ が必ず存在する。
3. A^+ の一意性を示す。いま、A に対して擬似逆行列 A_1^+,A_2^+ が存在すると仮定する。このとき、

$$\begin{aligned} A_1^+ - A_2^+ &= A_1^+AA_1^+ - A_2^+AA_2^+ \qquad (172) \\ &= A^T{A_1^+}^T A_1^+ - A_2^+{A_2^+}^T A_2^+ \\ &= (AA_2^+A)^T{A_1^+}^T A_1^+ - A_2^+{A_2^+}^T(AA_1^+A)^T \\ &= A^T{A_2^+}^T A^T{A_1^+}^T A_1^+ - A_2^+{A_2^+}^T A^T{A_1^+}^T A^T \\ &= A^T{A_2^+}^T A_1^+ - A_2^+{A_2^+}^T A^T \\ &= A_2^+AA_1^+ - A_2^+AA_1^+ \\ &= 0 \end{aligned}$$

なので、$A_1^+ = A_2^+$。つまり、A^+ は一意に定まる。■

さらに、A^+ には以下の性質がある。

1. $(A^+)^+ = A$
2. $(A^T)^+ = (A^+)^T$, $(AA^T)^+ = (A^+)^T A^+$
3. 一般には $m \times n$ 行列 A と $n \times p$ 行列 B について $(AB)^+ = B^+ A^+$ は成立しないが、$rank\, A = rank\, B = n$ であれば成立する。
4. $A^+ = (A^T A)^+ A^T = A^T (AA^T)^+$
5. $m \times n$ 行列 A について $rank\, A = m$ ならば $A^+ = A^T (AA^T)^{-1}$、また $rank\, A = n$ ならば $A^+ = (A^T A)^{-1} A^T$
6. $m \times n$ 行列 A の特異値分解が $A = U \Sigma V^T$（U,V は直交行列、$rank\, A = r$）ならば、$m \times n$ の Σ は

$$\Sigma = \left[\begin{array}{ccc|c} \sigma_1 & & & \\ & \ddots & & O_{r\times(n-r)} \\ & & \sigma_r & \\ \hline & O_{(m-r)\times r} & & O_{(m-r)\times(n-r)} \end{array}\right] \tag{173}$$

で与えられ、$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r \geq 0$ となる。このとき、

$$A^+ = V \Sigma^+ U^T \tag{174}$$

であり、Σ^+（$n \times m$）は

$$\Sigma^+ = \left[\begin{array}{ccc|c} \sigma_1^{-1} & & & \\ & \ddots & & O_{r\times(m-r)} \\ & & \sigma_r^{-1} & \\ \hline & O_{(n-r)\times r} & & O_{(n-r)\times(m-r)} \end{array}\right] \tag{175}$$

7. $A\boldsymbol{x} = \boldsymbol{b}$（$A$ は $m \times n$、$\boldsymbol{x}$,$\boldsymbol{b}$ は n 次元）について、A が正則であれば当然 $\boldsymbol{x} = A^{-1}\boldsymbol{b}$。
A が正則でないとき、$\boldsymbol{b} \in R(A)$ ならば解はあるが一意でない。そうでなければ解はない。
このような場合、$||A\boldsymbol{x} - \boldsymbol{b}||$ を最小化する $\boldsymbol{x}$ を次善の解とすると、それは

$$\boldsymbol{x} = A^+ \boldsymbol{b} + (I - A^+ A)\boldsymbol{k} \tag{176}$$

で与えられる。ここで $\boldsymbol{k}$ は任意の n 次元定数ベクトルである。解が一意でないとき、解自身のノルムを最小化する $\boldsymbol{x}$ は $\boldsymbol{x} = A^+ \boldsymbol{b}$ で与えられる。