

行列の特異値分解について

Note on singular value decomposition

2017年11月11日改訂版

1 実対称行列の固有値と固有ベクトル

一般に $M \times M$ の正方行列 A において,

$$A\mathbf{u}_j = \lambda_j \mathbf{u}_j \quad (1)$$

を満たす M 個の列ベクトル \mathbf{u}_j と定数 λ_j ($j = 1, \dots, M$) が存在するとき, λ_j を固有値, \mathbf{u}_j を固有ベクトルと呼ぶ. ここで, A を実対称行列とする. すなわち, A の全ての要素が実数で $A^T = A$ が成り立つ. 実対称行列では固有値と固有ベクトルは必ず存在し, 以下の性質がある.

1. 固有値は必ず実数となる. 証明は以下の通りである. 上付きの $*$ を複素共役を表すものとして, 式 (1) において両辺に左から \mathbf{u}_j^{T*} を乗じれば,

$$\mathbf{u}_j^{T*} A \mathbf{u}_j = \lambda_j \mathbf{u}_j^{T*} \mathbf{u}_j = \lambda_j \|\mathbf{u}_j\|^2 \quad (2)$$

を得る. 一方, 式 (1) においての転置と複素共役を取れば,

$$\mathbf{u}_j^{T*} A^{T*} = \lambda_j^* \mathbf{u}_j^{T*}$$

であり, さらに, 上式に右から \mathbf{u}_j を乗じれば,

$$\mathbf{u}_j^{T*} A^{T*} \mathbf{u}_j = \lambda_j^* \mathbf{u}_j^{T*} \mathbf{u}_j = \lambda_j^* \|\mathbf{u}_j\|^2$$

ここで, 実対称行列であれば $A^{T*} = A$ が成り立つので,

$$\mathbf{u}_j^{T*} A \mathbf{u}_j = \lambda_j^* \mathbf{u}_j^{T*} \mathbf{u}_j = \lambda_j^* \|\mathbf{u}_j\|^2 \quad (3)$$

である, 式 (2) と式 (3) の左辺は全く同じであるので,

$$\lambda_j \|\mathbf{u}_j\|^2 = \lambda_j^* \|\mathbf{u}_j\|^2$$

すなわち,

$$\lambda_j = \lambda_j^* \quad (4)$$

が成立する. 式 (4) は固有値 λ_j が実数であることを示している¹.

2. 実対称行列であれば固有ベクトルも実数である. 固有値 λ が与えられたとして, 固有ベクトル $\mathbf{u} = [u_1, \dots, u_M]^T$ の要素 u_1, \dots, u_M は M 行 M 列の線型方程式:

$$(A - \lambda I)\mathbf{u} = 0 \quad (5)$$

を解いて求める. この線型方程式は係数が全て実数であるので, 得られる解 (すなわち固有ベクトルの要素 u_1, \dots, u_M) も実数であり, したがって固有ベクトル \mathbf{u} も実数となる².

¹実際, ここの部分の証明は行列 A が $A^{T*} = A$ の性質を有することを用いている. しかし複素数値行列でも $A^{T*} = A$ の性質を有する行列は存在し, この性質を持つ行列はエルミート行列と呼ばれる. したがって, 固有値が実数であるのは, 実対称行列でなくてもエルミート行列なら成り立つ性質である.

²当然ながら複素数のエルミート行列では固有ベクトルは実数とはならない.

3. $M \times N$ の実数行列を F として, $A = F^T F$ とすれば, 行列 F は (実対称行列でありしかも) 半正定値行列である. 証明は以下の通りである. 任意の $M \times 1$ 列ベクトル x に対して,

$$x^T A x = x^T (F^T F) x = (F x)^T (F x) = \|F x\|^2 \geq 0 \quad (6)$$

である. x を固有ベクトル u_j とすれば,

$$u_j^T A u_j = x^T (F^T F) x = (F x)^T (F x) = \lambda_j^2 \|x\|^2 \geq 0 \quad (7)$$

したがって, $\lambda_j^2 \geq 0$ が成り立ち, A は半正定値行列である.

4. $M \times M$ の実対称行列を $A : A = F^T F$ とすれば, A の固有値は正またはゼロであり, 固有ベクトルは実数である (半正定値行列の性質)

2 行列の特異値分解

2.1 定義と諸性質

$M \times N$ の実数値行列 F に対して, $M \leq N$ の場合,

$$F = U \begin{bmatrix} \gamma_1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \gamma_2 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \gamma_M & 0 & \cdots & 0 \end{bmatrix} V^T \quad (8)$$

を与える $M \times M$ の直交行列 U と $N \times N$ の直交行列 V が必ず存在する. 上式を行列 F の特異値分解 (singular-value decomposition) と呼び, 略して SVD と呼ばれる. ここで, γ_j を行列 F の特異値, 直交行列 U の列ベクトル u_j ($j = 1, \dots, M$) を左特異値ベクトル, V の列ベクトルを v_j ($j = 1, \dots, N$) 右特異値ベクトルと呼ぶ. 非正方行列 F の次元が $M \geq N$ である場合には, 特異値分解は

$$F = U \begin{bmatrix} \gamma_1 & 0 & \cdots & 0 \\ 0 & \gamma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \gamma_N \\ 0 & \cdot & \cdot & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \cdot & \cdot & 0 \end{bmatrix} V^T \quad (9)$$

となる. 式 (8) および式 (9) を満たす直交行列 U と V が必ず存在することの証明は第 2.3 節に述べる.

式 (8) において右辺の行列計算を行うと

$$F = [u_1, u_2, \dots, u_M] \begin{bmatrix} \gamma_1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \gamma_2 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \gamma_M & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_N^T \end{bmatrix} = \sum_{j=1}^M \gamma_j u_j v_j^T \quad (10)$$

となる．式 (9) においても

$$\mathbf{F} = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_M] \begin{bmatrix} \gamma_1 & 0 & \cdots & 0 \\ 0 & \gamma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \gamma_N \\ 0 & \cdot & \cdot & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \cdot & \cdot & 0 \end{bmatrix} \begin{bmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \\ \vdots \\ \mathbf{v}_N^T \end{bmatrix} = \sum_{j=1}^N \gamma_j \mathbf{u}_j \mathbf{v}_j^T \quad (11)$$

であるので，結局， $R = \min\{M, N\}$ として，

$$\mathbf{F} = \sum_{j=1}^R \gamma_j \mathbf{u}_j \mathbf{v}_j^T \quad (12)$$

が成り立つ．行列 \mathbf{F} のランクが \mathcal{R} ($\mathcal{R} \leq R$) の場合には，特異値分解は

$$\begin{aligned} \mathbf{F} &= [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_M] \begin{bmatrix} \gamma_1 & 0 & \cdots & \cdot & \cdot & \cdots & 0 \\ 0 & \gamma_2 & \cdots & \cdot & \cdot & \cdot & \cdot \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdot & \cdots & \gamma_{\mathcal{R}} & 0 & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdot & \cdots & \cdot & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \\ \vdots \\ \mathbf{v}_N^T \end{bmatrix} \\ &= [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{\mathcal{R}}] \begin{bmatrix} \gamma_1 & 0 & \cdots & 0 \\ 0 & \gamma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \gamma_{\mathcal{R}} \end{bmatrix} \begin{bmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \\ \vdots \\ \mathbf{v}_{\mathcal{R}}^T \end{bmatrix} = \sum_{j=1}^{\mathcal{R}} \gamma_j \mathbf{u}_j \mathbf{v}_j^T \quad (13) \end{aligned}$$

となる．

式 (13) から特異値分解は

$$\mathbf{F} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T \quad (14)$$

と書くことができる．ここで，

$$\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{\mathcal{R}}] \quad (15)$$

$$\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{\mathcal{R}}] \quad (16)$$

は直交行列であり，対角行列 $\mathbf{\Sigma}$ は

$$\mathbf{\Sigma} = \begin{bmatrix} \gamma_1 & 0 & \cdots & 0 \\ 0 & \gamma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \gamma_{\mathcal{R}} \end{bmatrix}$$

である．したがって，

$$\mathbf{F} \mathbf{F}^T = (\mathbf{U} \mathbf{\Sigma} \mathbf{V}^T) (\mathbf{U} \mathbf{\Sigma} \mathbf{V}^T)^T = \mathbf{U} \mathbf{\Sigma}^2 \mathbf{U}^T \quad (17)$$

であり，さらに

$$\mathbf{F}^T \mathbf{F} = (\mathbf{U} \mathbf{\Sigma} \mathbf{V}^T)^T (\mathbf{U} \mathbf{\Sigma} \mathbf{V}^T) = \mathbf{V} \mathbf{\Sigma}^2 \mathbf{V}^T \quad (18)$$

も成り立つ．式 (17) によれば U は行列 FF^T の固有ベクトルから，式 (18) によれば V は行列 $F^T F$ の固有ベクトルから求めることができる．また，特異値は FF^T あるいは $F^T F$ の固有値の平方根を計算することにより求めることができる．

2.2 よく使う性質 (Rayleigh-Ritz formula の拡張版)

$M \times 1$ の単位ベクトル x ($\|x\| = 1$) と $N \times 1$ の単位ベクトル y ($\|y\| = 1$) を考えれば，最適化問題

$$\operatorname{argmax}_{x,y} x^T A y \quad (19)$$

を満たす x と y は行列 A の最大特異値 γ_1 に対応した特異値ベクトル u_1 と v_1 であり， $x^T A y$ の最大値は特異値 γ_1 に等しい．さらに，

$$\operatorname{argmax}_{x,y} x^T A y, \quad \text{subject to } x^T u_1 = 0, \quad \text{and } y^T v_1 = 0 \quad (20)$$

を満たす解は特異値 γ_2 に対応した特異値ベクトル u_2 と v_2 であり， $x^T A y$ の値は γ_2 となる．全く同様に，

$$\operatorname{argmax}_x x^T A y, \quad \text{subject to } x^T u_1 = 0, \quad x^T u_2 = 0, \quad \text{and } x^T v_1 = 0, \quad x^T v_2 = 0 \quad (21)$$

の解は特異値ベクトル u_3 と v_3 であり， $x^T A y$ の値は λ_3 となる．4 次以上の特異値，特異値ベクトルについても全く同様の関係が成り立つ．

2.3 特異値分解：式 (8) および式 (9) の成立の証明

ここでは，任意の実数行列 F に対して，式 (8) あるいは式 (9) を満たす直交行列 U と V が必ず存在することを示す．本節の証明は [1] p 400 を参考にした．まず，全ての $j = 1, \dots, M$ に対して， $\gamma_j > 0$ を仮定する．このとき， F は full-row-rank であると言われる． $F^T F$ は実対称行列であり，正定値行列 (positive definite matrix) であるので，この行列は正の固有値を持つ．行列 $F^T F$ の固有値の平方根を行列 F の特異値と呼ぶ． $F^T F$ の固有ベクトルは実ベクトルであり， \mathbb{R}^N の正規直交基底をなす．これら固有ベクトル v_1, \dots, v_N を列ベクトルとする $N \times N$ の行列 V を

$$V = [v_1, \dots, v_M, v_{M+1}, \dots, v_N] \quad (22)$$

とする．行列 $F^T F$ はランク M であるので， v_1, \dots, v_M は M 個のノンゼロ固有値に対応した固有ベクトルである．残りの v_{M+1}, \dots, v_N はゼロ固有値に対応した固有ベクトルである．すなわち，

$$F^T F v_j = \gamma_j^2 v_j \quad \text{for } j = 1, \dots, M \quad (23)$$

$$F^T F v_j = 0 v_j = \mathbf{0} \quad \text{for } j = M + 1, \dots, N \quad (24)$$

である．ここで，ベクトル u_j を，

$$u_j = \frac{1}{\gamma_j} F v_j \quad (25)$$

と定義する．すると，

$$u_i^T u_j = \frac{1}{\gamma_i \gamma_j} (F v_i)^T (F v_j) = \frac{1}{\gamma_i \gamma_j} v_i^T F^T F v_j = \frac{1}{\gamma_i \gamma_j} v_i^T \gamma_j^2 v_j = \frac{\gamma_j}{\gamma_i} v_i^T v_j = \delta_{i,j} \quad (26)$$

であり，ベクトル u_1, \dots, u_M も正規直交系をなす．この u_1, \dots, u_M を列ベクトルとして持つ行列 U を

$$U = [u_1, \dots, u_M] \quad (27)$$

と定義する．式 (22) の V と上式の U を用いて，

$$\begin{aligned}
U^T FV &= \\
U^T F[v_1, \dots, v_M, v_{M+1}, \dots, v_N] &= U^T [Fv_1, \dots, Fv_M, Fv_{M+1}, \dots, Fv_N] = \\
U^T [\gamma_1 u_1, \dots, \gamma_M u_M, \mathbf{0}, \dots, \mathbf{0}] &= \\
[\gamma_1 U^T u_1, \dots, \gamma_M U^T u_M, \mathbf{0}, \dots, \mathbf{0}] &= \begin{bmatrix} \gamma_1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \gamma_2 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \gamma_M & 0 & \cdots & 0 \end{bmatrix} \quad (28)
\end{aligned}$$

を得，したがって，特異値分解の式 (8) をを得る．

次に式 (9) に示す $M \geq N$ の場合を証明する．ここで，全ての $j = 1, \dots, N$ に対して， $\gamma_j > 0$ を仮定する．この場合は full-column-rank とよばれ， $F^T F$ の固有値はすべて正の値を持つ． $F^T F$ の固有ベクトル v_1, \dots, v_N を列ベクトルとする $N \times N$ の行列 V を

$$V = [v_1, \dots, v_N] \quad (29)$$

と定義する．式 (25) を用いて u_j を定義すれば，

$$FF^T u_j = \frac{1}{\gamma_j} F(F^T F)v_j = \frac{1}{\gamma_j} F(\gamma_j^2 v_j) = \gamma_j Fv_j = \gamma_j^2 u_j \quad (30)$$

であるので， u_j ($j = 1, \dots, N$) は $M \times M$ の行列 $F^T F$ の非ゼロ固有値に対応した固有ベクトルである．行列 FF^T の固有値ゼロに対応した固有ベクトルを u_{N+1}, \dots, u_M とすれば，ベクトル $u_1, \dots, u_N, u_{N+1}, \dots, u_M$ は正規直交系をなす．これら固有ベクトルを列にもつ行列 U を

$$U = [u_1, \dots, u_N, u_{N+1}, \dots, u_M] \quad (31)$$

と定義する．式 (29) の V と，式 (31) の U を用いれば，

$$\begin{aligned}
U^T FV &= \\
U^T F[v_1, \dots, v_N] &= U^T [Fv_1, \dots, Fv_N] = \\
U^T [\gamma_1 u_1, \dots, \gamma_N u_N] &= \\
[\gamma_1 U^T u_1, \dots, \gamma_N U^T u_N] &= \begin{bmatrix} \gamma_1 & 0 & \cdots & 0 \\ 0 & \gamma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \gamma_N \\ 0 & \cdot & \cdot & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \cdot & \cdot & 0 \end{bmatrix} \quad (32)
\end{aligned}$$

References

- [1] J. L. Goldberg, *Matrix Theory with Applications*. New York, NY: McGraw-Hill, Inc., 1991.
- [2] I. C. Ipsen, *Numerical matrix analysis: Linear systems and least squares*. SIAM, 2009.