

No.6 特異値分解 (発展)

この回のゴール

- ${}^tAA$ が半正定値の対称行列であることを示し、特異値を定義できる.
- 特異値分解 $A = U\Sigma V$ の存在を、対称行列の対角化から導ける.
- 2 次正方行列や小さい長方形行列の特異値分解を求められる.
- 特異値分解の図形的な意味と、低ランク近似の考え方を説明できる.

この回は発展

対角化は正方行列にしか使えず、しかも対角化できない正方行列もあった. 特異値分解は、**どんな行列 (長方形でもよい) にも使える「対角化の一般化」**である. No.5 の対称行列の対角化を土台にする.

1. 特異値

A を $m \times n$ 行列とする. ${}^tAA$ は n 次正方行列で、 ${}^t({}^tAA) = {}^tAA$ なので対称行列である.

定理 1

${}^tAA$ の固有値はすべて 0 以上である.

【証明】 ${}^tAA\boldsymbol{v} = \lambda\boldsymbol{v}$ ($\boldsymbol{v} \neq \mathbf{0}$) とすると $\lambda\|\boldsymbol{v}\|^2 = ({}^tAA\boldsymbol{v}, \boldsymbol{v}) = {}^t\boldsymbol{v}{}^tAA\boldsymbol{v} = \|A\boldsymbol{v}\|^2 \geq 0$. よって $\lambda \geq 0$. $\square$

定義 1 特異値

${}^tAA$ の固有値を大きい順に $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ とするとき、 $\sigma_i = \sqrt{\lambda_i}$ を A の**特異値**という.

例 1

$A = \begin{pmatrix} 3 & 0 \\ 4 & 5 \end{pmatrix}$ のとき ${}^tAA = \begin{pmatrix} 25 & 20 \\ 20 & 25 \end{pmatrix}$ で、固有値は 45, 5 (固有ベクトル ${}^t(1, 1)$, ${}^t(1, -1)$). 特異値は $\sigma_1 = 3\sqrt{5}$, $\sigma_2 = \sqrt{5}$.

注意

特異値は A の固有値とは一般に異なる. 例 1 の A の固有値は (下三角行列なので) 3, 5 である. ただし A が対称行列なら、特異値は固有値の絶対値になる (練習問題 4).

■ 2. 特異値分解

定理 2 特異値分解

A を $m \times n$ 行列, $r = \text{rank } A$ とする. m 次直交行列 U と n 次直交行列 V をうまく選ぶと

$$A = U \Sigma {}^t V, \quad \Sigma = \begin{pmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & O \end{pmatrix} \quad (m \times n \text{ 行列})$$

と書ける. ここで $\sigma_1 \geq \dots \geq \sigma_r > 0$ は A の正の特異値である.

【証明】 V を作る: ${}^t A A$ は対称行列なので, No.5 定理 4 より固有ベクトルからなる正規直交基底 $\mathbf{v}_1, \dots, \mathbf{v}_n$ がとれる (固有値 $\sigma_1^2 \geq \dots \geq \sigma_n^2$ の順に並べる). $V = (\mathbf{v}_1 \cdots \mathbf{v}_n)$ とする.

U の最初の r 列を作る:

$$(A\mathbf{v}_i, A\mathbf{v}_j) = (\mathbf{v}_i, {}^t A A \mathbf{v}_j) = \sigma_j^2 (\mathbf{v}_i, \mathbf{v}_j) = \sigma_j^2 \delta_{ij}$$

なので, $A\mathbf{v}_1, \dots, A\mathbf{v}_n$ は互いに直交し, $\|A\mathbf{v}_i\| = \sigma_i$. $\text{Im } f_A = \langle A\mathbf{v}_1, \dots, A\mathbf{v}_n \rangle$ の次元は r なので, $\sigma_i > 0$ となるのはちょうど $i \leq r$ のとき. $\mathbf{u}_i = \frac{1}{\sigma_i} A\mathbf{v}_i$ ($i \leq r$) は正規直交系である.

U を完成させる: $\mathbf{u}_1, \dots, \mathbf{u}_r$ を $\mathbb{R}^m$ の正規直交基底 $\mathbf{u}_1, \dots, \mathbf{u}_m$ に延長し, $U = (\mathbf{u}_1 \cdots \mathbf{u}_m)$ とする.

確認: $A\mathbf{v}_i = \sigma_i \mathbf{u}_i$ ($i \leq r$), $A\mathbf{v}_i = \mathbf{0}$ ($i > r$) を並べると $AV = U\Sigma$. $V^{-1} = {}^t V$ なので $A = U\Sigma {}^t V$. $\square$

手順

1. ${}^t A A$ の固有値 (大きい順) と, 正規直交な固有ベクトル $\mathbf{v}_i$ を求める. $\sigma_i = \sqrt{\lambda_i}$.
2. $\sigma_i > 0$ の i について $\mathbf{u}_i = \frac{1}{\sigma_i} A\mathbf{v}_i$.
3. $\mathbf{u}_i$ が m 本に足りなければ, 正規直交基底に延長する.

例題 1

$A = \begin{pmatrix} 3 & 0 \\ 4 & 5 \end{pmatrix}$ の特異値分解を求めよ.

【解答】 例 1 より $\sigma_1 = 3\sqrt{5}$, $\sigma_2 = \sqrt{5}$, $\mathbf{v}_1 = \frac{1}{\sqrt{2}} {}^t(1, 1)$, $\mathbf{v}_2 = \frac{1}{\sqrt{2}} {}^t(1, -1)$.

$$\mathbf{u}_1 = \frac{1}{3\sqrt{5}} A\mathbf{v}_1 = \frac{1}{3\sqrt{10}} \begin{pmatrix} 3 \\ 9 \end{pmatrix} = \frac{1}{\sqrt{10}} \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \quad \mathbf{u}_2 = \frac{1}{\sqrt{5}} A\mathbf{v}_2 = \frac{1}{\sqrt{10}} \begin{pmatrix} 3 \\ -1 \end{pmatrix}.$$

よって

$$A = \underbrace{\frac{1}{\sqrt{10}} \begin{pmatrix} 1 & 3 \\ 3 & -1 \end{pmatrix}}_U \underbrace{\begin{pmatrix} 3\sqrt{5} & 0 \\ 0 & \sqrt{5} \end{pmatrix}}_\Sigma \underbrace{\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}}_{{}^t V}.$$

($\mathbf{u}_1 \perp \mathbf{u}_2$ は $3 - 3 = 0$ で確かめられる.)

例題 2 長方形の行列

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 0 \end{pmatrix} \text{ の特異値分解を求めよ.}$$

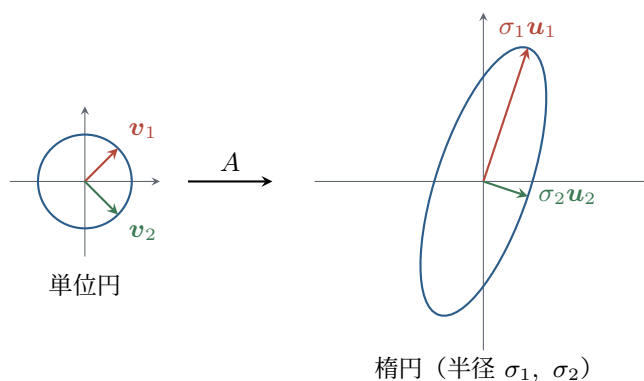
【解答】 ${}^tAA = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}$ の固有値は 4, 0, 固有ベクトル $\mathbf{v}_1 = \frac{1}{\sqrt{2}}{}^t(1, 1)$, $\mathbf{v}_2 = \frac{1}{\sqrt{2}}{}^t(1, -1)$. $\sigma_1 = 2$, $\sigma_2 = 0$ なので $\text{rank } A = 1$. $\mathbf{u}_1 = \frac{1}{2}A\mathbf{v}_1 = \frac{1}{\sqrt{2}}{}^t(1, 1, 0)$. これを $\mathbb{R}^3$ の正規直交基底に延長して $\mathbf{u}_2 = \frac{1}{\sqrt{2}}{}^t(1, -1, 0)$, $\mathbf{u}_3 = {}^t(0, 0, 1)$.

$$A = \underbrace{\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix}}_U \underbrace{\begin{pmatrix} 2 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}}_\Sigma \underbrace{\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}}_{}^tV.$$

■ 3. 図形的な意味

$A = U\Sigma V$ は, 線形写像 $\mathbf{x} \mapsto A\mathbf{x}$ を 3 段階に分解している.

- (1) ${}^tV$: 回転 (または折り返し). $\mathbf{v}_i$ を $\mathbf{e}_i$ に移す.
- (2) Σ : 各座標軸の方向に σ_i 倍する (伸縮).
- (3) U : 回転 (または折り返し). $\mathbf{e}_i$ を $\mathbf{u}_i$ に移す.



(図は例題 1 の A を, 縮尺 $1/\sqrt{5}$ で描いたもの.) 単位円は A によって楕円に移り, その長軸・短軸の半分の長さが特異値 σ_1, σ_2 , 軸の方向が $\mathbf{u}_1, \mathbf{u}_2$ である. どんな線形写像も「回転して, 軸方向に伸縮して, 回転する」だけで表せる, というのが特異値分解の主張である.

■ 4. 低ランク近似

特異値分解を列ごとに書くと

$$A = \sigma_1 \mathbf{u}_1 {}^t \mathbf{v}_1 + \sigma_2 \mathbf{u}_2 {}^t \mathbf{v}_2 + \cdots + \sigma_r \mathbf{u}_r {}^t \mathbf{v}_r$$

となる (各項 $\mathbf{u}_i {}^t \mathbf{v}_i$ はランク 1 の $m \times n$ 行列). 大きい特異値の項ほど A への寄与が大きい.

定理 3 低ランク近似 (エッカート・ヤングの定理)

$k < r$ とし, $A_k = \sigma_1 \mathbf{u}_1 {}^t \mathbf{v}_1 + \cdots + \sigma_k \mathbf{u}_k {}^t \mathbf{v}_k$ とおく. ランク k 以下の $m \times n$ 行列 B の中で, $\|A - B\|$ (行列の内積 $(X, Y) = \text{tr}({}^tXY)$ によるノルム) の最小値は $B = A_k$ で達成され, そのとき

$$\|A - A_k\|^2 = \sigma_{k+1}^2 + \cdots + \sigma_r^2.$$

($\sigma_k = \sigma_{k+1}$ のように特異値が重複する場合は, 最小値を与える B は 1 つとは限らない.)

(証明は省略する.)

例 2

例題 1 の A で $k = 1$ とすると

$$A_1 = 3\sqrt{5} \mathbf{u}_1 {}^t \mathbf{v}_1 = \begin{pmatrix} 1.5 & 1.5 \\ 4.5 & 4.5 \end{pmatrix}, \quad A - A_1 = \begin{pmatrix} 1.5 & -1.5 \\ -0.5 & 0.5 \end{pmatrix}.$$

$\|A - A_1\|^2 = 2.25 + 2.25 + 0.25 + 0.25 = 5 = \sigma_2^2$ となっている.

コラム: 画像圧縮とデータ解析

白黒画像は, 各画素の明るさを並べた大きな行列と見なせる. 特異値分解して大きい特異値の項だけを残すと, 少ないデータで元の画像をよく近似できる (画像圧縮). 同じ考え方は, データの主要な傾向を取り出す**主成分分析**や, 推薦システム, 機械学習にも使われる (分野「数学探究」第5章 No.1・No.2, 分野「プログラミング (数学のための)」第2章 No.2 で実際に計算できる).

■ 練習問題

1. $A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$ の特異値分解を求めよ.
2. $A = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ (2×1 行列) の特異値分解を求めよ.
3. 直行列の特異値はすべて 1 であることを示せ.
4. A が対称行列なら, A の特異値は A の固有値の絶対値であることを示せ. (ヒント: No.5 定理 4 で $A = QD{}^tQ$ とすると ${}^tAA = QD^2{}^tQ$.)
5. (考えてみよう) $m \times n$ 行列 A と ${}^tA$ の正の特異値は一致することを示せ. (ヒント: $A = U\Sigma{}^tV$ の転置をとる.)

■ 練習問題の解答

1. ${}^tAA = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ の固有値は 2, 0, $\mathbf{v}_1 = \frac{1}{\sqrt{2}}{}^t(1, 1)$, $\mathbf{v}_2 = \frac{1}{\sqrt{2}}{}^t(1, -1)$. $\sigma_1 = \sqrt{2}$, $\mathbf{u}_1 = \frac{1}{\sqrt{2}}A\mathbf{v}_1 = {}^t(1, 0)$,
延長して $\mathbf{u}_2 = {}^t(0, 1)$.

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{2} & 0 \\ 0 & 0 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

2. ${}^tAA = (5)$ なので $\sigma_1 = \sqrt{5}$, $V = (1)$. $\mathbf{u}_1 = \frac{1}{\sqrt{5}}{}^t(1, 2)$, 延長して $\mathbf{u}_2 = \frac{1}{\sqrt{5}}{}^t(2, -1)$.

$$A = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} \sqrt{5} \\ 0 \end{pmatrix} (1).$$

3. 直交行列 Q なら ${}^tQQ = E$ の固有値はすべて 1 なので, 特異値はすべて $\sqrt{1} = 1$. (長さを保つ変換なので, 単位円は単位円に移る.)

4. ${}^tAA = A^2 = QD^2{}^tQ$ なので, ${}^tAA$ の固有値は λ_i^2 (λ_i は A の固有値). 特異値は $\sqrt{\lambda_i^2} = |\lambda_i|$.

5. ${}^tA = V\Sigma{}^tU$ は ${}^tA$ の特異値分解の形をしており, Σ の対角成分は Σ と同じ $\sigma_1, \dots, \sigma_r$. (または: $A{}^tA$ と ${}^tAA$ の 0 でない固有値が一致することを示してもよい.)