

定理 11.2 : $\lambda_1, \lambda_2, \dots, \lambda_n$ を n 次 Hermite 行列の固有値とする.

- (a) $\lambda_1, \lambda_2, \dots, \lambda_n$ は実数であるが, 対応する固有ベクトルは複素ベクトルになる.
 (b) 固有値 λ_i に対応する固有ベクトル $\mathbf{p}_i$ をお互いに直交し, かつ, その長さを 1 にとることができる. すなわち,

$$(\mathbf{p}_i)^T \bar{\mathbf{p}}_j = 0 \quad (i \neq j), \quad \|\mathbf{p}_i\| = 1. \quad (i = 1, 2, \dots, n)$$

したがって, $\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_n$ は線形独立になる.

証明 (略):

- (a) λ を B の固有値とする. すなわち,

$$B\mathbf{y} = \lambda\mathbf{y}, \quad \mathbf{y} \neq \mathbf{0}, \quad \mathbf{y} \in \mathbb{C}^n$$

を仮定する. 複素共役を $\bar{}$ で表すとする. $(B\mathbf{y})^T = \lambda\mathbf{y}^T$ の複素共役をとると, $\bar{\mathbf{y}}^T \bar{B}^T = \bar{\mathbf{y}}^T B^* = \bar{\lambda}\bar{\mathbf{y}}^T$ となる. したがって,

$$\lambda\bar{\mathbf{y}}^T \mathbf{y} = \bar{\mathbf{y}}^T (\lambda\mathbf{y}) = \bar{\mathbf{y}}^T B\mathbf{y} = \bar{\mathbf{y}}^T B^* \mathbf{y} = \bar{\lambda}\bar{\mathbf{y}}^T \mathbf{y}$$

となる. ここで, $\bar{\mathbf{y}}^T \mathbf{y} = \|\mathbf{y}\|^2$ はゼロでない実数であるので $\lambda = \bar{\lambda}$ を得る.

注: $\mathbf{y} \in \mathbb{C}^n$, $\mathbf{y} = (a_1 + ib_1, a_2 + ib_2, \dots, a_n + ib_n)^T$ の場合, $\|\mathbf{y}\|^2 = \mathbf{y}^T \bar{\mathbf{y}} = \bar{\mathbf{y}}^T \mathbf{y} = \sum_{k=1}^n (a_k - ib_k)(a_k + ib_k) = \sum_{k=1}^n (a_k^2 + b_k^2)$ となる.

- (b) $\mathbf{y}, \mathbf{z}$ を λ, μ の相異なる固有値に対応する固有ベクトルとする. すなわち,

$$B\mathbf{y} = \lambda\mathbf{y} \quad \mathbf{y} \neq \mathbf{0}, \quad \bar{B}\bar{\mathbf{z}} = \mu\bar{\mathbf{z}}, \quad \mathbf{z} \neq \mathbf{0}.$$

このとき,

$$\begin{aligned} \bar{\mathbf{z}}^T B\mathbf{y} &= \bar{\mathbf{z}}^T (\lambda\mathbf{y}) = \lambda\bar{\mathbf{z}}^T \mathbf{y}, \\ \bar{\mathbf{z}}^T B\mathbf{y} &= \bar{\mathbf{z}}^T B^* \mathbf{y} = \mu\bar{\mathbf{z}}^T \mathbf{y} \end{aligned}$$

となる. よって,

$$0 = \bar{\mathbf{z}}^T B\mathbf{y} - \bar{\mathbf{z}}^T B\mathbf{y} = (\lambda - \mu)\bar{\mathbf{z}}^T \mathbf{y}.$$

仮定より, $\lambda \neq \mu$ であるから, $\bar{\mathbf{z}}^T \mathbf{y} = 0$ を得る.

注: ベクトル空間 $\mathbb{C}^n$ において, $\mathbf{y} \in \mathbb{C}^n$ と $\mathbf{z} \in \mathbb{C}^n$ の内積は $\mathbf{y}^T \bar{\mathbf{z}} = \bar{\mathbf{z}}^T \mathbf{y} = \sum_{k=1}^n (a_k^y + ib_k^y)(a_k^z - ib_k^z)$ として定義される.

以上のことから異なる固有値に対する固有ベクトルは直交している. さらに, $B\mathbf{y} = \lambda\mathbf{y}$ であれば, 両側を $\|\mathbf{y}\| = \sqrt{\mathbf{y}^T \bar{\mathbf{y}}}$ で割ることによって

$$B \frac{\mathbf{y}}{\|\mathbf{y}\|} = \lambda \frac{\mathbf{y}}{\|\mathbf{y}\|}$$

ノルムが 1 の固有ベクトル $\mathbf{p} = \frac{\mathbf{y}}{\|\mathbf{y}\|}$ を得ることができる. その他の事実についても実対称行列の場合と同じなので省略する.

□

11.1 Gersgorin の定理 (弱い形)

定理 11.3 : n 次複素正方行列 $\mathbf{H} = (h_{ij})$ の任意の固有値 λ は

$$G_k := \{z \in \mathbb{C} \mid |z - h_{kk}| \leq r_k\}, \quad r_k = \sum_{j=1, j \neq k}^n |h_{kj}|$$

で定義される, 複素平面上の n 個の円 G_k ($k = 1, 2, \dots, n$) の和集合 $\bigcup_{k=1}^n G_k$ の中にある.

証明 : $\mathbf{H}$ の固有値 λ と対応する固有ベクトルを $\mathbf{x} \neq \mathbf{0}$ としたとき, $\mathbf{H}\mathbf{x} = \lambda\mathbf{x}$ より,

$$\sum_{j=1}^n h_{ij}x_j = \lambda x_i \quad (i = 1, 2, \dots, n)$$

となる. ここで $k \in \{1, 2, \dots, n\}$ を $|x_k| = \max_j |x_j| \neq 0$ とおくと,

$$\sum_{j=1, j \neq k}^n h_{kj}x_j + h_{kk}x_k = \lambda x_k$$

になるので,

$$|\lambda - h_{kk}| \leq \sum_{j=1, j \neq k}^n |h_{kj}| \frac{|x_j|}{|x_k|} \leq \sum_{j=1, j \neq k}^n |h_{kj}| = r_k$$

を得る.

例 :

$\mathbf{H} = \begin{pmatrix} -2 & -i \\ i & 5 \end{pmatrix}$ の固有値は $\frac{3 \pm \sqrt{53}}{2}$ であり, $r_1 = 1$, $r_2 = 1$ となるので,

$$G_1 = \{z \in \mathbb{C} \mid |z + 2| \leq 1\}, \quad G_2 = \{z \in \mathbb{C} \mid |z - 5| \leq 1\}$$

の Gersgorin の円が得られる. 図 24 を参照.

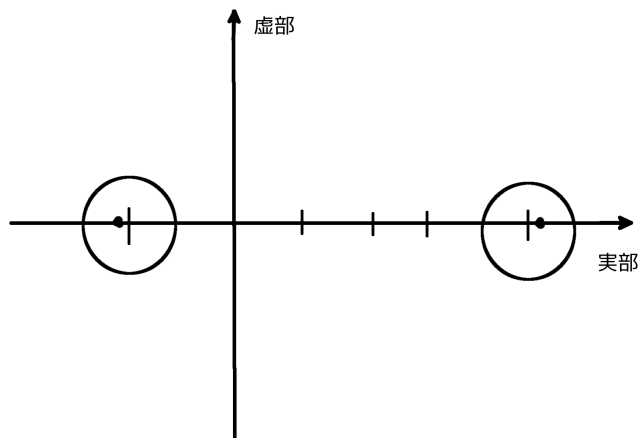


図 24: $\mathbf{H}$ に対する Gersgorin の円.

12 特異値分解

近年，データマイニング，機械学習，ビックデータなどのキーワードの元で行列の低ランク近似や行列の要素の補間などが取り上げられている．その中，行列の特異値を用いた問題がしばしば登場する．正方行列に対する固有値と異なり，正方でない任意の行列に対して特異値は定義されている．

12.1 定義と基本的な性質

定義 12.1 : $\mathbf{A} \in \mathbb{C}^{m \times n}$ に対し，その特異値分解とは

$$\mathbf{A} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^*$$

となるものである．ただし， $\mathbf{U} \in \mathbb{C}^{m \times m}$ ， $\mathbf{V} \in \mathbb{C}^{n \times n}$ はユニタリ行列， $\mathbf{\Sigma} \in \mathbb{R}^{m \times n}$ は対角行列である．それに加え， $\mathbf{\Sigma}$ の対角要素 σ_i は非負で非増加順に並んでいると仮定する．つまり $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_p \geq 0$ ， $p = \min(m, n)$ である．

例 1 :

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

例 2 :

$$\begin{pmatrix} -\frac{2}{5} & -\frac{2\sqrt{3}}{5} \\ \frac{40\sqrt{3}-9}{50} & -\frac{40+9\sqrt{3}}{50} \\ \frac{15\sqrt{3}+6}{25} & \frac{6\sqrt{3}-15}{25} \end{pmatrix} = \begin{pmatrix} 0 & -\frac{4}{5} & -\frac{3}{5} \\ \frac{4}{5} & -\frac{9}{25} & \frac{12}{25} \\ \frac{3}{5} & \frac{12}{25} & \frac{16}{25} \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix}$$

定理 12.2 : 行列 $\mathbf{A} \in \mathbb{C}^{m \times n}$ に対して必ず特異値分解が存在する．さらに，定義 12.1 の特異値 $\{\sigma_j\}$ は一意に定まる．

定理 12.3 : 行列 $\mathbf{A}$ のランク（階数） r は $\mathbf{A}$ の正の特異値の数と一致する．

証明 :

$\mathbf{A}$ の特異値分解 $\mathbf{A} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^*$ を考え， $\mathbf{U}, \mathbf{V}$ がユニタリ行列であるので， $\mathbf{A}$ の左から行に関する基本変形を $\mathbf{U}^*$ として， $\mathbf{A}$ の右からその列に対する基本形を $\mathbf{V}$ によって行くと， $\mathbf{U}^* \mathbf{A} \mathbf{V} = (\mathbf{U}^* \mathbf{U}) \mathbf{\Sigma} (\mathbf{V}^* \mathbf{V}) = \mathbf{\Sigma}$ となる．結局， $\mathbf{A}$ のランクは $\mathbf{\Sigma}$ の対角の非ゼロ要素の数と一致する． $\square$

定理 12.4 : $\|\mathbf{A}\|_2 = \sigma_1$ であり， $\|\mathbf{A}\|_F = \sqrt{\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_r^2}$ である．ただし， r は $\mathbf{A}$ のランクである．

証明 :

$\|\mathbf{A}\|_F := \sqrt{\text{tr}(\mathbf{A}^* \mathbf{A})}$ より， $\|\mathbf{A}\|_F = \sqrt{\text{tr}(\mathbf{V} \mathbf{\Sigma}^T \mathbf{U}^* \mathbf{U} \mathbf{\Sigma} \mathbf{V}^*)} = \sqrt{\text{tr}(\mathbf{V} \mathbf{\Sigma}^T \mathbf{\Sigma} \mathbf{V}^*)} = \sqrt{\text{tr}(\mathbf{\Sigma}^T \mathbf{\Sigma} \mathbf{V}^* \mathbf{V})} = \sqrt{\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_r^2}$ ．

$$\|\mathbf{A}\|_2 := \sup_{\mathbf{x} \in \mathbb{C}^n, \mathbf{x} \neq \mathbf{0}} \frac{\|\mathbf{A}\mathbf{x}\|_2}{\|\mathbf{x}\|_2} = \sup_{\mathbf{x} \in \mathbb{C}^n, \mathbf{x} \neq \mathbf{0}} \frac{\sqrt{\mathbf{x}^* \mathbf{V} \mathbf{\Sigma}^T \mathbf{U}^* \mathbf{U} \mathbf{\Sigma} \mathbf{V}^* \mathbf{x}}}{\sqrt{\mathbf{x}^* \mathbf{V} \mathbf{V}^* \mathbf{x}}} = \sup_{\mathbf{y} \in \mathbb{C}^n, \mathbf{y} \neq \mathbf{0}} \frac{\sqrt{\mathbf{y}^* \mathbf{\Sigma}^T \mathbf{\Sigma} \mathbf{y}}}{\sqrt{\mathbf{y}^* \mathbf{y}}} = \sigma_1$$

ただし, $\mathbf{y} = \mathbf{V}^* \mathbf{x}$ という変数変換を行っている. $\mathbf{V}$ はユニタリ行列であるので, 正則行列であることに注意. $\square$

定理 12.5 : $\mathbf{A} \in \mathbb{C}^{m \times n}$ は r 個のランク 1 の和である.

$$\mathbf{A} = \sum_{j=1}^r \sigma_j \mathbf{u}_j \mathbf{v}_j^*$$

ただし, $\mathbf{u}_j$ は $\mathbf{U} \in \mathbb{C}^{m \times m}$ の j 列であり, $\mathbf{v}_j$ は $\mathbf{V} \in \mathbb{C}^{n \times n}$ の j 列である.

12.2 特異値分解の計算

定理 12.6 : $\mathbf{A}$ の非ゼロ特異値は $\mathbf{A}^* \mathbf{A}$ もしくは $\mathbf{A} \mathbf{A}^*$ の固有値の平方根と一致する.

証明 :

$\mathbf{A}^* \mathbf{A} = \mathbf{V} \boldsymbol{\Sigma}^T \mathbf{U}^* \mathbf{U} \boldsymbol{\Sigma} \mathbf{V}^* = (\mathbf{V}) \boldsymbol{\Sigma}^T \boldsymbol{\Sigma} (\mathbf{V}^*)$ となるので, $\mathbf{A}^* \mathbf{A}$ の固有値は $\sigma_1^2, \sigma_2^2, \dots, \sigma_r^2, 0, \dots, 0$ となる. ただし, r は $\mathbf{A}^* \mathbf{A}$ のランクである. よって, 定理の主張が示せた. $\mathbf{A} \mathbf{A}^*$ についても同様である. $\square$

1. 上記の定理から, 行列 $\mathbf{A} \in \mathbb{C}^{m \times n}$ の特異値分解を得るためにはまず $\mathbf{A}^* \mathbf{A} \in \mathbb{C}^{n \times n}$ の固有値分解を行い, $\mathbf{A}^* \mathbf{A} = \mathbf{P} \boldsymbol{\Lambda} \mathbf{P}^*$ を計算する. $\mathbf{A}^* \mathbf{A}$ は対称行列になっていることに注意.
2. つぎに, $\boldsymbol{\Lambda} \in \mathbb{R}^{n \times n}$ の対角要素 ($\mathbf{A}^* \mathbf{A}$ の固有値) の平方根を $\sigma_1, \sigma_2, \dots, \sigma_n$ として, 正のものだけをその順番に $\boldsymbol{\Sigma} \in \mathbb{R}^{m \times n}$ の対角におく.
3. $\mathbf{P} = \mathbf{V} \in \mathbb{C}^{n \times n}$ とおき, 最後に $\mathbf{A} = \mathbf{U} \boldsymbol{\Sigma} \mathbf{V}^*$ となるように, ユニタリ行列 $\mathbf{U} \in \mathbb{C}^{m \times m}$ を求める. 定理 12.1 からわかるように, 行列 $\mathbf{A}$ の特異値は一意であるが, その特異値分解は一意ではないので $\mathbf{U}$ を計算するとき自由度がある.

[問題 12-01] 例 1, 例 2 に出てくる行列の特異値分解を上記の方法を用い確かめよ.

12.3 Moore-Penrose 形一般逆行列

定理 12.7 : $\mathbf{A} \in \mathbb{C}^{m \times n}$ の特異値分解を $\mathbf{U} \boldsymbol{\Sigma} \mathbf{V}^*$ とすると $\mathbf{A}$ の Moore-Penrose 一般逆行列を $\mathbf{A}^\dagger = \mathbf{V} \boldsymbol{\Sigma}^\dagger \mathbf{U}^*$ と定義する. ただし, $\boldsymbol{\Sigma}^\dagger \in \mathbb{R}^{n \times m}$ の対角行列でその要素は

$$\begin{cases} \frac{1}{\sigma_j} & , \text{ if } \sigma_j > 0 \\ 0 & , \text{ if } \sigma_j = 0 \end{cases}$$

として定められている.

上記の $\mathbf{A}^\dagger \in \mathbb{C}^{n \times m}$ は $\mathbf{A}$ の疑似逆行列とよばれることもあり, 以下の性質を満す唯一の行列 $\mathbf{X} \in \mathbb{C}^{n \times m}$ である.

1. $\mathbf{A} \mathbf{X} \mathbf{A} = \mathbf{A}$
2. $\mathbf{X} \mathbf{A} \mathbf{X} = \mathbf{X}$
3. $(\mathbf{A} \mathbf{X})^* = \mathbf{A} \mathbf{X}$

4. $(XA)^* = XA$

当然ながら, $\text{rank}(\mathbf{A}) = n (\leq m)$ のとき $\mathbf{A}^\dagger = (\mathbf{A}^* \mathbf{A})^{-1} \mathbf{A}^*$ であり, $\text{rank}(\mathbf{A}) = n = m$ であるとき $\mathbf{A}^\dagger = \mathbf{A}^{-1}$ である.

このことから $\mathbf{A}$ の Moore-Penrose 一般逆行列は $\|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2$ の最小ノルムを満すベクトル $\mathbf{x}$ のなかでさらに最小ノルムのものを実現するベクトル $\bar{\mathbf{x}} = \mathbf{A}^\dagger \mathbf{b}$ に関連していることがわかる.