第 32 章

特異値と特異値分解

正定値行列と半正定値行列

固有値がすべて 0 以上になる対称行列は、応用上さまざまな場面で現れる

- 半正定値行列：すべての固有値が非負（正または零）である対称行列
- 正定値行列：すべての固有値が正である対称行列

def – 正定値行列

A をエルミート行列（対称行列）とし、任意のベクトル $\boldsymbol{x} \in \mathbb{C}^n$（$\boldsymbol{x} \in \mathbb{R}^n$）に対して、

$$(A\boldsymbol{x}, \boldsymbol{x}) > 0$$

が成り立つとき、A は正定値行列であるという

theorem – 正定値性と固有値の正実性

エルミート行列 A が正定値行列であることと、A のすべての固有値が正の実数であることは同値である

証明

正定値行列 $\Longrightarrow$ 固有値が正

A の固有値を λ、対応する固有ベクトルを $\boldsymbol{x}$ とすると、

$$A\boldsymbol{x} = \lambda\boldsymbol{x}$$

両辺で $\boldsymbol{x}$ との内積をとると、

$$(A\boldsymbol{x}, \boldsymbol{x}) = \lambda(\boldsymbol{x}, \boldsymbol{x}) = \lambda\|\boldsymbol{x}\|^2$$

A が正定値行列であることから、$(A\boldsymbol{x}, \boldsymbol{x}) > 0$ が成り立ち、

$$\lambda\|\boldsymbol{x}\|^2 > 0$$

ここで、固有ベクトルは零ベクトルではないので、$\|\boldsymbol{x}\|^2 > 0$ である

よって、$\lambda\|\boldsymbol{x}\|^2 > 0$ の両辺を $\|\boldsymbol{x}\|^2$ で割ることにより、

$$\lambda > 0$$

が得られる ■

固有値が正 $\Longrightarrow$ 正定値行列

A の固有値を $\lambda_1, \ldots, \lambda_n > 0$ とする

A はエルミート行列であることから、ユニタリ行列 U を用いて次のように対角化できる

$$A = UDU^{-1} = UDU^* = U \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} U^*$$

theorem 20.3「随伴による標準内積の表現」より、

$$(A\boldsymbol{x}, \boldsymbol{x}) = \boldsymbol{x}^* A\boldsymbol{x} = \boldsymbol{x}^* UDU^*\boldsymbol{x}$$

ここで、$\boldsymbol{y} = U^*\boldsymbol{x}$ とおくと、

$$\boldsymbol{y}^* = (U^*\boldsymbol{x})^* = \boldsymbol{x}^* U$$

となるので、次のように書き換えられる

$$(A\boldsymbol{x}, \boldsymbol{x}) = \boldsymbol{y}^* D\boldsymbol{y} = (D\boldsymbol{y}, \boldsymbol{y})$$

左辺の内積を計算すると、

$$
\begin{aligned}
(D\boldsymbol{y}, \boldsymbol{y}) &= \begin{pmatrix} \lambda_1 y_1 & & \\ & \ddots & \\ & & \lambda_n y_n \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \\
&= \lambda_1 |y_1|^2 + \cdots + \lambda_n |y_n|^2
\end{aligned}
$$

$\lambda_1, \ldots, \lambda_n > 0$ であることから、すべての項が正になるので、

$$(A\boldsymbol{x}, \boldsymbol{x}) = (D\boldsymbol{y}, \boldsymbol{y}) > 0$$

よって、A は正定値行列である ■

半正定値行列は、正定値行列の条件に等号を含むようにしたものである

def - 半正定値行列

A をエルミート行列（対称行列）とし、任意のベクトル $\boldsymbol{x} \in \mathbb{C}^n$（$\boldsymbol{x} \in \mathbb{R}^n$）に対して、

$$(A\boldsymbol{x}, \boldsymbol{x}) \geq 0$$

が成り立つとき、A は**半正定値行列**であるという

theorem - 半正定値性と固有値の非負実性

エルミート行列 A が半正定値行列であることと、A のすべての固有値が非負の実数であることは同値である

対称行列を構成する行列積

theorem 31.1「対称行列のスペクトル分解」を任意の長方行列に拡張したものが**特異値分解**である

対称行列から任意の行列へ議論を拡張するにあたって、次の定理が重要となる

theorem - 自身の随伴行列との積で構成されるエルミート行列

A を任意の複素行列（長方行列）とするとき、A^*A および AA^* はエルミート行列である

証明

theorem 20.2「積に対するエルミート共役の順序反転性」より、積をエルミート行列にすると順序が入れ替わることに注意して、

$$(A^*A)^* = A^*(A^*)^* = A^*A$$

よって、A^*A はエルミート行列である

同様に、

$$(AA^*)^* = (A^*)^*A^* = AA^*$$

よって、AA^* もエルミート行列である ■

A を実行列とすれば、次が成り立つ

theorem 32.1 - 自身の転置行列との積で構成される対称行列

A を任意の実行列（長方行列）とするとき、$A^\top A$ および $AA^\top$ は対称行列である

A^*A および AA^* という形の行列には、さらに固有値に関する重要な性質がある

theorem - 自身の随伴行列との積で構成される半正値行列

任意の行列 A に対して、AA^* および A^*A はともに半正値行列である

証明

エルミート行列 AA^* の固有ベクトルを $\boldsymbol{u}$ とし、その固有値を $\lambda \in \mathbb{C}$ とすると、

$$AA^*\boldsymbol{u} = \lambda\boldsymbol{u}$$

両辺で $\boldsymbol{u}$ との内積をとると、

$$(\boldsymbol{u}, AA^*\boldsymbol{u}) = \lambda(\boldsymbol{u}, \boldsymbol{u}) = \lambda\|\boldsymbol{u}\|^2$$

この左辺は、theorem 20.4「随伴公式」を用いて、

$$\begin{aligned}(\boldsymbol{u}, AA^*\boldsymbol{u}) &= (\boldsymbol{u}, A(A^*\boldsymbol{u})) \\ &= (A^*\boldsymbol{u}, A^*\boldsymbol{u}) \\ &= \|A^*\boldsymbol{u}\|^2 \geq 0\end{aligned}$$

外側の A に
随伴公式を適用

となるので、

$$\|A^*\boldsymbol{u}\|^2 = \lambda\|\boldsymbol{u}\|^2 \geq 0$$

ここで、固有ベクトルは零ベクトルではないので、$\|\boldsymbol{u}\|^2 > 0$ である

よって、$\lambda\|\boldsymbol{u}\|^2 \geq 0$ の両辺を $\|\boldsymbol{u}\|^2$ で割ることにより、

$$\lambda \geq 0$$

が得られる

A^*A についても同様に、

$$(\boldsymbol{u}, A^*A\boldsymbol{u}) = (A\boldsymbol{u}, A\boldsymbol{u}) = \|A\boldsymbol{u}\|^2 \geq 0$$

から、$\lambda \geq 0$ が得られる ■

theorem 32.2 – 特異値と左右特異ベクトルの対応関係

A を O でない任意の行列とするとき、$A^\top A$ と $AA^\top$ は共通の正の固有値 σ^2 を持ち、それぞれの固有ベクトル $\boldsymbol{u}, \boldsymbol{v}$ は次の関係を満たす

$$A\boldsymbol{v} = \sigma\boldsymbol{u}, \quad A^\top\boldsymbol{u} = \sigma\boldsymbol{v}$$

証明

$AA^\top$ の固有値が σ^2 と仮定した場合

$AA^\top$ の固有値が非負の固有値 σ^2 を持ち、対応する固有ベクトルが $\boldsymbol{u}$ であるとすると、

$$AA^\top \boldsymbol{u} = \sigma^2 \boldsymbol{u}$$

この両辺に左から $A^\top$ をかけて、

$$A^\top AA^\top \boldsymbol{u} = A^\top \sigma^2 \boldsymbol{u}$$

ここで、$\boldsymbol{v} = \dfrac{A^\top \boldsymbol{u}}{\sigma}$ とおくと、$A^\top \boldsymbol{u} = \sigma \boldsymbol{v}$ となるので、

$$\begin{aligned} A^\top A \sigma \boldsymbol{v} &= \sigma^3 \boldsymbol{v} \\ A^\top A \boldsymbol{v} &= \sigma^2 \boldsymbol{v} \end{aligned}$$

よって、σ^2 は $A^\top A$ の固有値でもあり、対応する固有ベクトル $\boldsymbol{v}$ は

$$A^\top \boldsymbol{u} = \sigma \boldsymbol{v}$$

を満たす ■

$A^\top A$ の固有値が σ^2 と仮定した場合

$A^\top A$ の固有値が非負の固有値 σ^2 を持ち、対応する固有ベクトルが $\boldsymbol{v}$ であるとすると、

$$A^\top A \boldsymbol{v} = \sigma^2 \boldsymbol{v}$$

この両辺に左から A をかけて、

$$AA^\top A \boldsymbol{v} = A \sigma^2 \boldsymbol{v}$$

ここで、$\boldsymbol{u} = \dfrac{A \boldsymbol{v}}{\sigma}$ とおくと、$A \boldsymbol{v} = \sigma \boldsymbol{u}$ となるので、

$$\begin{aligned} AA^\top \sigma \boldsymbol{u} &= \sigma^3 \boldsymbol{u} \\ AA^\top \boldsymbol{u} &= \sigma^2 \boldsymbol{u} \end{aligned}$$

よって、σ^2 は $AA^\top$ の固有値でもあり、対応する固有ベクトル $\boldsymbol{u}$ は

$$A \boldsymbol{v} = \sigma \boldsymbol{u}$$

を満たす ■

特異値と特異ベクトル

スペクトル分解の拡張である特異値分解では、任意の行列がその特異値と特異ベクトルによって表せる

def – 特異値と特異ベクトル

零行列ではない任意の $m \times n$ 行列 A に対して、

$$A\boldsymbol{v} = \sigma \boldsymbol{u}, \quad A^\top \boldsymbol{u} = \sigma \boldsymbol{v}$$

となる正の数 σ を特異値と呼び、

- 左特異ベクトル：m 次元ベクトル $\boldsymbol{u}\,(\neq \boldsymbol{0})$
- 右特異ベクトル：n 次元ベクトル $\boldsymbol{v}\,(\neq \boldsymbol{0})$

を合わせて特異ベクトルと呼ぶ

特異ベクトルと固有ベクトルの関係

特異値と特異ベクトルの関係式

$$A\boldsymbol{v} = \sigma \boldsymbol{u}, \quad A^\top \boldsymbol{u} = \sigma \boldsymbol{v}$$

において、第 1 式の両辺に $A^\top$ を左からかけると、

$$\begin{aligned} A^\top A\boldsymbol{v} &= \sigma A^\top \boldsymbol{u} \\ &= \sigma^2 \boldsymbol{v} \end{aligned} \quad \text{第 2 式を代入}$$

また、第 2 式の両辺に A を左からかけると、

$$\begin{aligned} AA^\top \boldsymbol{u} &= \sigma A\boldsymbol{v} \\ &= \sigma^2 \boldsymbol{u} \end{aligned} \quad \text{第 1 式を代入}$$

得られた結果をまとめると、

$$AA^\top \boldsymbol{u} = \sigma^2 \boldsymbol{u}, \quad A^\top A\boldsymbol{v} = \sigma^2 \boldsymbol{v}$$

ここで、A は任意の長方行列だが、theorem 32.1「自身の転置行列との積で構成される対称行列」より、$AA^\top$ と $A^\top A$ は対称行列となる

すなわち、

- 左特異ベクトル $\boldsymbol{u}$ は m 次対称行列 $AA^\top$ の固有ベクトル
- 右特異ベクトル $\boldsymbol{v}$ は n 次対称行列 $A^\top A$ の固有ベクトル

であり、theorem 32.2「特異値と左右特異ベクトルの対応関係」より、特異値の 2 乗 σ^2 は $AA^\top, A^\top A$ 共通の固有値である

特異ベクトルの正規直交化

A の特異値を $\sigma_1 \geq \cdots \geq \sigma_r > 0$ とする
ここで、重複があってもよい

対応する r 本の左特異ベクトル $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r$ と r 本の右特異ベクトル $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_r$ は、どちらも対称行列の固有ベクトルであるから、対称行列の固有ベクトルの正規直交化［第 31 章］で述べたように、それぞれを正規直交系に選ぶことができる

特異値分解

k 本の左特異ベクトルの正規直交系 $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k$ を拡張して、$\mathbb{R}^m$ の正規直交基底 $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k, \boldsymbol{u}_{k+1}, \ldots, \boldsymbol{u}_m$ が定義できる。
同様に、k 本の右特異ベクトルの正規直交系 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$ を拡張して、$\mathbb{R}^n$ の正規直交基底 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k, \boldsymbol{v}_{k+1}, \ldots, \boldsymbol{v}_n$ が定義できる。

$\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ と $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ はそれぞれ $AA^\top$ と $A^\top A$ の固有ベクトルであり、これらに対応する共通の固有値を $\lambda_1, \ldots, \lambda_n$ とおく。

$AA^\top$ および $A^\top A$ は半正定値行列であるので、その固有値はすべて零か正の数である。
また、$AA^\top$ および $A^\top A$ は対称行列であり、対称行列の階数 r は非零の固有値の個数に等しい。
n 個の固有値のうち、r 個ある正の固有値は特異値の条件を満たすので、

- $\lambda_1, \ldots, \lambda_r$ は特異値（正の固有値）$\sigma_1, \ldots, \sigma_r$
- $\lambda_{r+1}, \ldots, \lambda_n$ は零の固有値

とする。

特異値が r 個あることから、左特異ベクトルと特異値の組の個数、右特異ベクトルと特異値の組の個数は、どちらも r であることがいえる。

$$k = r$$

以上の議論をまとめると、

$$AA^\top \boldsymbol{u}_i = \begin{cases} \sigma_i \boldsymbol{u}_i & (i = 1, \ldots, r) \\ \boldsymbol{o} & (i = r+1, \ldots, m) \end{cases}$$

$$A^\top A \boldsymbol{v}_i = \begin{cases} \sigma_i \boldsymbol{v} & (i = 1, \ldots, r) \\ \boldsymbol{o} & (i = r+1, \ldots, n) \end{cases}$$

ここで、$i = 1, \ldots, r$ の範囲に限っては、特異値と特異ベクトルの関係より、

$$A^\top \boldsymbol{u}_i = \sigma_i \boldsymbol{v}_i$$
$$A \boldsymbol{v}_i = \sigma_i \boldsymbol{u}_i$$

という形で書ける。

$i > r$ の場合についても同じ形で書くために、次の定理を示す。

theorem - 行列積による零化

i. $AA^\top \boldsymbol{u} = \boldsymbol{o} \Longrightarrow A^\top \boldsymbol{u} = \boldsymbol{o}$

ii. $A^\top A \boldsymbol{v} = \boldsymbol{o} \Longrightarrow A \boldsymbol{v} = \boldsymbol{o}$

証明

(i) $AA^\top \boldsymbol{u} = \boldsymbol{o}$ について

$AA^\top \boldsymbol{u} = \boldsymbol{o}$ の両辺で $\boldsymbol{u}$ との内積をとって、

$$(\boldsymbol{u}, AA^\top \boldsymbol{u}) = 0$$

このとき、左辺は、

$$\begin{aligned}(\boldsymbol{u}, AA^\top\boldsymbol{u}) &= (\boldsymbol{u}, A(A^\top\boldsymbol{u})) \\ &= (A^\top\boldsymbol{u}, A^\top\boldsymbol{u}) \\ &= \|A^\top\boldsymbol{u}\|^2\end{aligned}$$

外側の A に
随伴公式を適用

と変形できるので、

$$\|A^\top\boldsymbol{u}\|^2 = 0$$

が成り立つ。

ここで、内積の正値性

$$\|A^\top\boldsymbol{u}\|^2 = (A^\top\boldsymbol{u}, A^\top\boldsymbol{u}) \geq 0$$

において、等号が成立するのは、

$$A^\top\boldsymbol{u} = \boldsymbol{o}$$

の場合のみである。 ■

(ii) $A^\top A\boldsymbol{v} = \boldsymbol{o}$ について

$A^\top A\boldsymbol{v} = \boldsymbol{o}$ の両辺で $\boldsymbol{v}$ との内積をとって、

$$(\boldsymbol{v}, A^\top A\boldsymbol{v}) = 0$$

このとき、左辺は、

$$(\boldsymbol{v}, A^\top A\boldsymbol{v}) = (A\boldsymbol{v}, A\boldsymbol{v}) = \|A\boldsymbol{v}\|^2$$

と変形できるので、

$$\|A\boldsymbol{v}\|^2 = 0$$

が成り立つ。

ここで、内積の正値性

$$\|A\boldsymbol{v}\|^2 = (A\boldsymbol{v}, A\boldsymbol{v}) \geq 0$$

において、等号が成立するのは、

$$A\boldsymbol{v} = \boldsymbol{o}$$

の場合のみである。 ■

この定理を用いると、

$$A\boldsymbol{v}_i = \begin{cases} \sigma_i \boldsymbol{u}_i & (i = 1, \ldots, r) \\ \boldsymbol{o} & (i = r+1, \ldots, m) \end{cases}$$

$$A^\top \boldsymbol{u}_i = \begin{cases} \sigma_i \boldsymbol{v}_i & (i = 1, \ldots, r) \\ \boldsymbol{o} & (i = r+1, \ldots, n) \end{cases}$$

とまとめられる。

このことから、次に示すように、任意の行列は、その特異値と特異ベクトルによって表すことができる。これを特異値分解と呼ぶ。

A の特異値分解

A は $\mathbb{R}^n$ の正規直交基底 $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ をそれぞれ

$$\sigma_1 \boldsymbol{u}_1, \ldots, \sigma_r \boldsymbol{u}_r, \boldsymbol{o}, \ldots, \boldsymbol{o}$$

に写像するから、theorem 14.2「正規直交基底による表現行列の展開」より、A は

$$A = \sigma_1 \boldsymbol{u}_1 \boldsymbol{v}_1^\top + \cdots + \sigma_r \boldsymbol{u}_r \boldsymbol{v}_r^\top \quad (\sigma_1 \geq \cdots \geq \sigma_r > 0)$$

と表すことができる。

$A^\top$ の特異値分解

同様に、$A^\top$ は $\mathbb{R}^m$ の正規直交基底 $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m\}$ をそれぞれ

$$\sigma_1 \boldsymbol{v}_1, \ldots, \sigma_r \boldsymbol{v}_r, \boldsymbol{o}, \ldots, \boldsymbol{o}$$

に写像するから、$A^\top$ は

$$A^\top = \sigma_1 \boldsymbol{v}_1 \boldsymbol{u}_1^\top + \cdots + \sigma_r \boldsymbol{v}_r \boldsymbol{u}_r^\top \quad (\sigma_1 \geq \cdots \geq \sigma_r > 0)$$

と表すことができる。

列空間と行空間の正規直交基底

ここでは、A の特異値の個数は A の階数に等しく、特異ベクトルは行空間と列空間の正規直交基底を成すことを示す。

def - 列空間

行列 A の n 本の列の張る $\mathbb{R}^m$ の部分空間を A の列空間という。

def - 行空間

行列 A の m 本の行の張る $\mathbb{R}^n$ の部分空間を A の行空間という。

列空間を $\mathcal{U}$、行空間を $\mathcal{V}$ と表記することにする。

列空間の正規直交基底

A の列 $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ の任意の線形結合を考える。

$$c_1\boldsymbol{a}_1 + \cdots + c_n\boldsymbol{a}_n = \begin{pmatrix} \boldsymbol{a}_1 & \cdots & \boldsymbol{a}_n \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = A\boldsymbol{c}$$

ここで、A の特異値分解［第 32 章］の式

$$A = \sigma_1 \boldsymbol{u}_1 \boldsymbol{v}_1^\top + \cdots + \sigma_r \boldsymbol{u}_r \boldsymbol{v}_r^\top$$

の両辺に $\boldsymbol{c}$ を右からかけると、

$$\begin{aligned} A\boldsymbol{c} &= \sigma_1 \boldsymbol{u}_1 \boldsymbol{v}_1^\top \boldsymbol{c} + \cdots + \sigma_r \boldsymbol{u}_r \boldsymbol{v}_r^\top \boldsymbol{c} \\ &= \sigma_1 (\boldsymbol{v}_1, \boldsymbol{c}) \boldsymbol{u}_1 + \cdots + \sigma_r (\boldsymbol{v}_r, \boldsymbol{c}) \boldsymbol{u}_r \end{aligned}$$

すなわち、A の列の任意の線形結合 $A\boldsymbol{c}$ は、互いに直交する左特異ベクトル $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r$ の線形結合で書ける。

よって、$\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ の張る列空間 $\mathcal{U}$ は、左特異ベクトル $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r$ を正規直交基底とする r 次元部分空間である。

このことから、

r 本の列のみが線型独立である

ということもいえる。

行空間の正規直交基底

A の行は $A^\top$ の列であるので、同様の議論を $A^\top$ に対して行う。

$A^\top$ の特異値分解［第 32 章］の式

$$A^\top = \sigma_1 \boldsymbol{v}_1 \boldsymbol{u}_1^\top + \cdots + \sigma_r \boldsymbol{v}_r \boldsymbol{u}_r^\top$$

の両辺に $\boldsymbol{c}$ を右からかけることで、

$$\begin{aligned} A^\top \boldsymbol{c} &= \sigma_1 \boldsymbol{v}_1 \boldsymbol{u}_1^\top \boldsymbol{c} + \cdots + \sigma_r \boldsymbol{v}_r \boldsymbol{u}_r^\top \boldsymbol{c} \\ &= \sigma_1 (\boldsymbol{u}_1, \boldsymbol{c}) \boldsymbol{v}_1 + \cdots + \sigma_r (\boldsymbol{u}_r, \boldsymbol{c}) \boldsymbol{v}_r \end{aligned}$$

となり、A の行の任意の線形結合 $A^\top \boldsymbol{c}$ は、互いに直交する右特異ベクトル $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_r$ の線形結合で書けることがわかる。

よって、A の行の張る行空間 $\mathcal{V}$ は、右特異ベクトル $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_r$ を正規直交基底とする r 次元部分空間である。

このことから、

r 本の行のみが線型独立である

ということもいえる。

特異値の個数と特異ベクトルによる基底

以上の議論から、次のことがわかる。

theorem 32.3 – 特異値の個数と階数

行列 A の階数 r は、A の特異値の個数に等しい。

⚓ theorem – 特異ベクトルと列空間・行空間の正規直交基底

左特異ベクトル $\{\boldsymbol{u}_i\}_{i=1}^r$ と右特異ベクトル $\{\boldsymbol{v}_i\}_{i=1}^r$ は、それぞれ列空間 $\mathcal{U}$ と行空間 $\mathcal{V}$ の正規直交基底を成す。

列空間と行空間への射影

特異ベクトルが列空間・行空間の正規直交基底をなすことから、これらを用いて射影行列の展開［第 23 章］を考えることができる。

$\mathbb{R}^m$ の列空間 $\mathcal{U}$ への射影行列を $P_{\mathcal{U}}$、$\mathbb{R}^n$ の行空間 $\mathcal{V}$ への射影行列を $P_{\mathcal{V}}$ とすると、

$$P_{\mathcal{U}} = \sum_{i=1}^{r} \boldsymbol{u}_i \boldsymbol{u}_i^\top$$

$$P_{\mathcal{V}} = \sum_{i=1}^{r} \boldsymbol{v}_i \boldsymbol{v}_i^\top$$

ここで、$\boldsymbol{u}_i$ は列空間 $\mathcal{U}$ の正規直交基底であることから、$\boldsymbol{u}_i \in \mathcal{U}$ である。

よって、部分空間への射影［第 23 章］で議論したように、$P_{\mathcal{U}}$ は列空間 $\mathcal{U}$ の元をそのまま写すので、

$$P_{\mathcal{U}} \boldsymbol{u}_i = \boldsymbol{u}_i$$

このことから、A の特異値分解［第 32 章］の式

$$A = \sigma_1 \boldsymbol{u}_1 \boldsymbol{v}_1^\top + \cdots + \sigma_r \boldsymbol{u}_r \boldsymbol{v}_r^\top$$

の両辺に左から $P_{\mathcal{U}}$ をかけても変化しないことが次のように導かれる。

$$\begin{aligned} P_{\mathcal{U}} A &= P_{\mathcal{U}} \left(\sigma_1 \boldsymbol{u}_1 \boldsymbol{v}_1^\top + \cdots + \sigma_r \boldsymbol{u}_r \boldsymbol{v}_r^\top \right) \\ &= \sigma_1 P_{\mathcal{U}} \boldsymbol{u}_1 \boldsymbol{v}_1^\top + \cdots + \sigma_r P_{\mathcal{U}} \boldsymbol{u}_r \boldsymbol{v}_r^\top \\ &= \sigma_1 \boldsymbol{u}_1 \boldsymbol{v}_1^\top + \cdots + \sigma_r \boldsymbol{u}_r \boldsymbol{v}_r^\top \\ &= A \end{aligned}$$

行についても同様に、$P_{\mathcal{V}} \boldsymbol{v}_i = \boldsymbol{v}_i$ であるから、$A^\top$ の特異値分解［第 32 章］の式

$$A^\top = \sigma_1 \boldsymbol{v}_1 \boldsymbol{u}_1^\top + \cdots + \sigma_r \boldsymbol{v}_r \boldsymbol{u}_r^\top$$

の両辺に右から $P_{\mathcal{V}}$ をかけても変化しない。

$$
\begin{aligned}
AP_{\mathcal{V}} &= \left(\sigma_1 \boldsymbol{v}_1 \boldsymbol{u}_1^\top + \cdots + \sigma_r \boldsymbol{v}_r \boldsymbol{u}_r^\top\right) P_{\mathcal{V}} \\
&= \sigma_1 \boldsymbol{v}_1 P_{\mathcal{V}} \boldsymbol{u}_1^\top + \cdots + \sigma_r \boldsymbol{v}_r P_{\mathcal{V}} \boldsymbol{u}_r^\top \\
&= \sigma_1 \boldsymbol{v}_1 \boldsymbol{u}_1^\top + \cdots + \sigma_r \boldsymbol{v}_r \boldsymbol{u}_r^\top \\
&= A^\top
\end{aligned}
$$

以上をまとめると、次の式が成り立つ。

$$
\begin{aligned}
P_{\mathcal{U}} A &= A \\
AP_{\mathcal{V}} &= A^\top
\end{aligned}
$$

特異値分解の行列表記

特異値分解の式

$$
A = \sigma_1 \boldsymbol{u}_1 \boldsymbol{v}_1^\top + \cdots + \sigma_r \boldsymbol{u}_r \boldsymbol{v}_r^\top
$$

は、次のように行列で表すこともできる。

$$
\begin{aligned}
A &= \begin{pmatrix} \sigma_1 \boldsymbol{u}_1 & \cdots & \sigma_r \boldsymbol{u}_r \end{pmatrix} \begin{pmatrix} \boldsymbol{v}_1^\top \\ \vdots \\ \boldsymbol{v}_r^\top \end{pmatrix} \\
&= \begin{pmatrix} \boldsymbol{u}_1 & \cdots & \boldsymbol{u}_r \end{pmatrix} \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{pmatrix} \begin{pmatrix} \boldsymbol{v}_1^\top \\ \vdots \\ \boldsymbol{v}_r^\top \end{pmatrix}
\end{aligned}
$$

ここで、

$$
U_r = \begin{pmatrix} \boldsymbol{u}_1 & \cdots & \boldsymbol{u}_r \end{pmatrix}, \quad V_r = \begin{pmatrix} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_r \end{pmatrix},
$$

$$
\Sigma_r = \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{pmatrix}
$$

とおくと、

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

という、行列 A を 3 つの行列を用いて分解した式として表すことができる。

この式は、A の簡約された特異値分解と呼ばれる。

特異値分解のより一般的な形

先ほどの式が「簡約された」特異値分解と呼ばれるということは、簡約する前のより一般的な形も考えられるということである。

元々、特異値分解の式は、A が $\mathbb{R}^n$ の正規直交基底 $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ をそれぞれ

$$\sigma_1 \boldsymbol{u}_1, \ldots, \sigma_r \boldsymbol{u}_r, \boldsymbol{o}, \ldots, \boldsymbol{o}$$

に写像することから導かれた。

そこで、$r+1$ 番以降の項も省略せずに書くと、

$$\begin{aligned} A &= \sigma_1 \boldsymbol{u}_1 \boldsymbol{v}_1^\top + \cdots + \sigma_r \boldsymbol{u}_r \boldsymbol{v}_r^\top + \boldsymbol{o} \boldsymbol{v}_{r+1}^\top + \cdots + \boldsymbol{o} \boldsymbol{v}_n^\top \\ &= \begin{pmatrix} \sigma_1 \boldsymbol{u}_1 & \cdots & \sigma_r \boldsymbol{u}_r & \boldsymbol{o} & \cdots & \boldsymbol{o} \end{pmatrix} \begin{pmatrix} \boldsymbol{v}_1^\top \\ \vdots \\ \boldsymbol{v}_r^\top \\ \boldsymbol{v}_{r+1}^\top \\ \vdots \\ \boldsymbol{v}_n^\top \end{pmatrix} \\ &= \begin{pmatrix} \boldsymbol{u}_1 & \cdots & \boldsymbol{u}_m \end{pmatrix} \left(\begin{array}{ccc|c} \sigma_1 & & & \\ & \ddots & & O \\ & & \sigma_r & \\ \hline & O & & O \end{array}\right) \begin{pmatrix} \boldsymbol{v}_1^\top \\ \vdots \\ \boldsymbol{v}_n^\top \end{pmatrix} \end{aligned}$$

補足

この式変形は、ブロック行列の積の計算に基づいている。

たとえば、

$$A = \begin{pmatrix} A_{11} & A_{12} \end{pmatrix}, \quad B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$$

とおけば、

$$AB = \begin{pmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \end{pmatrix}$$

のように計算できる。

そこで、

$$U_1 = \begin{pmatrix} \boldsymbol{u}_1 & \cdots & \boldsymbol{u}_r \end{pmatrix}, \quad U_2 = \begin{pmatrix} \boldsymbol{u}_{r+1} & \cdots & \boldsymbol{u}_m \end{pmatrix},$$
$$D = \mathrm{diag}(\sigma_1, \ldots, \sigma_r)$$

とおくと、

$$\begin{aligned} \begin{pmatrix} U_1 & U_2 \end{pmatrix} \begin{pmatrix} D & O \\ O & O \end{pmatrix} &= \begin{pmatrix} U_1 D + U_2 O & U_1 O + U_2 O \end{pmatrix} \\ &= \begin{pmatrix} U_1 D & O \end{pmatrix} \\ &= \begin{pmatrix} \sigma_1 \boldsymbol{u}_1 & \cdots & \sigma_r \boldsymbol{u}_r & \boldsymbol{o} & \cdots & \boldsymbol{o} \end{pmatrix} \end{aligned}$$

という式変形が確かめられる。

ここで、

$$U = \begin{pmatrix} \boldsymbol{u}_1 & \cdots & \boldsymbol{u}_m \end{pmatrix}, \quad V = \begin{pmatrix} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_n \end{pmatrix},$$

$$\Sigma = \begin{array}{c} \begin{array}{cc} \overleftrightarrow{\quad r \quad} & \overleftrightarrow{\quad n-r \quad} \end{array} \\ \left(\begin{array}{ccc|c} \sigma_1 & & & \\ & \ddots & & O \\ & & \sigma_r & \\ \hline & O & & O \end{array}\right) \begin{array}{l} \updownarrow r \\ \\ \updownarrow m-r \end{array} \end{array}$$

とおくと、

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

と表せる。この式を A の特異値分解と呼ぶ。

簡約された特異値分解は、特異値分解において U の余計な列と Σ の零行を省いたものだといえる。