第 1 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

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 $\mathbf{x} \in \mathbb{C}^n$ ($\mathbf{x} \in \mathbb{R}^n$) に対して、

が成り立つとき、A は正定値行列であるという

証明 証明

正定値行列 = 固有値が正

A の固有値を λ 、対応する固有ベクトルを x とすると、

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

両辺で 変 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、(Ax, 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 = 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

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

随伴による内積の表現より、

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

zzv, $\boldsymbol{y} = U^*\boldsymbol{x}$ zzv

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

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、

$$(Doldsymbol{y},oldsymbol{y}) = egin{pmatrix} \lambda_1 y_1 & & & & \\ & \ddots & & \\ & & \lambda_n y_n \end{pmatrix} egin{pmatrix} y_1 \ dots \ y_n \end{pmatrix} \ & = \lambda_1 |y_1|^2 + \dots + \lambda_n |y_n|^2$$

 $\lambda_1, \ldots, \lambda_n > 0$ 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

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 $\mathbf{x} \in \mathbb{C}^n$ ($\mathbf{x} \in \mathbb{R}^n$) 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

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

スペクトル分解は対称行列に対するものだったが、これ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

♣ 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

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

 $oldsymbol{\$}$ 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、 $A^{\mathsf{T}}A$ および AA^{T} 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

証明 証明

エルミート行列 AA^* の固有ベクトルを 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 **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、随伴公式を用いて、

$$(\boldsymbol{u}, AA^*\boldsymbol{u}) = (\boldsymbol{u}, A(A^*\boldsymbol{u}))$$
 外側の A に $= (A^*\boldsymbol{u}, A^*\boldsymbol{u})$ 随伴公式を適用 $= \|A^*\boldsymbol{u}\|^2 \geq 0$

となるので、

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

ここで、固有ベクトルは零ベクトルではないので、 $\| {m u} \|^2 > 0$ であるよって、 $\lambda \| {m u} \|^2 \ge 0$ の両辺を $\| {m u} \|^2$ で割ることにより、

$$\lambda > 0$$

が得られる

A*A についても同様に、

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

から、 $\lambda \geq 0$ が得られる

・ 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、 $A^{\mathsf{T}}A$ と AA^{T} 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

$$A\mathbf{v} = \sigma \mathbf{u}, \quad A^{\mathsf{T}}\mathbf{u} = \sigma \mathbf{v}$$

証明 証明

AA^{\top} の固有値が σ^2 と仮定した場合

 AA^{T} 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 $m{u}$ であるとすると、

$$AA^{\mathsf{T}}\boldsymbol{u} = \sigma^2\boldsymbol{u}$$

この両辺に左から A^{T} をかけて、

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

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

$$A^{\top} A \sigma \boldsymbol{v} = \sigma^3 \boldsymbol{v}$$
$$A^{\top} A \boldsymbol{v} = \sigma^2 \boldsymbol{v}$$

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

$$A^{\mathsf{T}}\boldsymbol{u} = \sigma \boldsymbol{v}$$

を満たす

$A^{T}A$ の固有値が σ^{2} と仮定した場合

 $A^{\mathsf{T}}A$ 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、

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

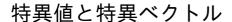
ここで、 $\boldsymbol{u} = \frac{A\boldsymbol{v}}{\sigma}$ とおくと、 $A\boldsymbol{v} = \sigma\boldsymbol{u}$ となるので、

$$AA^{\mathsf{T}}\sigma\boldsymbol{u} = \sigma^{3}\boldsymbol{u}$$
$$AA^{\mathsf{T}}\boldsymbol{u} = \sigma^{2}\boldsymbol{u}$$

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

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

を満たす



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<mark>特異値と特異ベクトル</mark>によって表せる

| 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、

$$A\mathbf{v} = \sigma \mathbf{u}, \quad A^{\mathsf{T}}\mathbf{u} = \sigma \mathbf{v}$$

となる正の数 σ を<mark>特異値</mark>と呼び、

- 左特異ベクトル: m 次元ベクトル $u \neq 0$
- 右特異ベクトル: n 次元ベクトル v (≠ 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

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

において、第 1 式の両辺に A^{T} を左からかけると、

$$A^{\mathsf{T}}A\boldsymbol{v} = \sigma A^{\mathsf{T}}\boldsymbol{u}$$

= $\sigma^2 \boldsymbol{v}$ 第 2 式を代入

また、第2式の両辺に A を左からかけると、

$$AA^{\mathsf{T}}\boldsymbol{u} = \sigma A \boldsymbol{v}$$

= $\sigma^2 \boldsymbol{u}$ 第 1 式を代入

得られた結果をまとめると、

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

ここで、A は任意の長方行列だが、 AA^{T} と $A^{\mathsf{T}}A$ は対称行列となる

すなわち、

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

であり、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 $\sigma_1 \ge \cdots \ge \sigma_r > 0$ 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 u_1, \ldots, u_r と r 本の右特異ベクトル v_1, \ldots, v_r は、ど ちらも対称行列の固有ベクトルであるから、それぞれを正規直交系に選ぶことができる



特異值分解

k 本の左特異ベクトルの正規直交系 u_1, \ldots, u_k を拡張して、 \mathbb{R}^m の正規直交基底 $u_1, \ldots, u_k, u_{k+1}, \ldots, u_m$ 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 $oldsymbol{v}_1,\ldots,oldsymbol{v}_k$ を拡張して、 \mathbb{R}^n の正規直交基底 $oldsymbol{v}_1,\ldots,oldsymbol{v}_k,oldsymbol{v}_{k+1},\ldots,oldsymbol{v}_n$ が定義できる。

 $m{u}_1,\ldots,m{u}_n$ と $m{v}_1,\ldots,m{v}_n$ 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^{T} および $A^{\mathsf{T}}A$ 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^{T} および $A^{\mathsf{T}}A$ 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、

$$AA^{ op}oldsymbol{u}_i = egin{cases} \sigma_ioldsymbol{u}_i & (i=1,\ldots,r) \ oldsymbol{o} & (i=r+1,\ldots,m) \end{cases}$$
 $A^{ op}Aoldsymbol{v}_i = egin{cases} \sigma_ioldsymbol{v} & (i=1,\ldots,r) \ oldsymbol{o} & (i=r+1,\ldots,n) \end{cases}$

ここで、i = 1, ..., 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。

→ 行列積による零化

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

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

▲ 証明

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

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

このとき、左辺は、

$$(\boldsymbol{u}, AA^{\top}\boldsymbol{u}) = (\boldsymbol{u}, A(A^{\top}\boldsymbol{u}))$$
 外側の A に
$$= (A^{\top}\boldsymbol{u}, A^{\top}\boldsymbol{u})$$
 随伴公式を適用
$$= \|A^{\top}\boldsymbol{u}\|^{2}$$

と変形できるので、

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

が成り立つ。

ここで、内積の正値性

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

において、等号が成立するのは、

$$A^{\mathsf{T}}\boldsymbol{u} = \boldsymbol{o}$$

の場合のみである。

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

 $A^{\mathsf{T}}A\boldsymbol{v}=\boldsymbol{o}$ 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

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

において、等号が成立するのは、

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

の場合のみである。

この定理を用いると、

$$egin{aligned} egin{aligned} egin{aligned} eta oldsymbol{v}_i &= \left\{egin{aligned} \sigma_i oldsymbol{u}_i & (i = 1, \ldots, r) \ oldsymbol{o} & (i = r+1, \ldots, r) \ oldsymbol{o} & (i = r+1, \ldots, n) \end{aligned} \end{aligned}$$

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、正規直交基底による表現行列の展開より、A は

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

と表すことができる。

A^T の特異値分解

同様に、 A^{T} 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^{T} は

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

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。

▶ 列空間 行列 A の n 本の列の張る ℝ の部分空間を A の列空間という。

ightharpoonup 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U、行空間を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 c を右からかけると、

$$Ac = \sigma_1 \boldsymbol{u}_1 \boldsymbol{v}_1^{\top} \boldsymbol{c} + \dots + \sigma_r \boldsymbol{u}_r \boldsymbol{v}_r^{\top} \boldsymbol{c}$$
$$= \sigma_1(\boldsymbol{v}_1, \boldsymbol{c}) \boldsymbol{u}_1 + \dots + \sigma_r(\boldsymbol{v}_r, \boldsymbol{c}) \boldsymbol{u}_r$$

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

よって、 $oldsymbol{a}_1,\ldots,oldsymbol{a}_n$ の張る列空間 $oldsymbol{U}$ は、左特異ベクトル $oldsymbol{u}_1,\ldots,oldsymbol{u}_r$ を正規直交基底とする $oldsymbol{r}$ 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^{T} の列であるので、同様の議論を A^{T} に対して行う。

A^T の特異値分解の式

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

の両辺にcを右からかけることで、

$$A^{\top} \boldsymbol{c} = \sigma_1 \boldsymbol{v}_1 \boldsymbol{u}_1^{\top} \boldsymbol{c} + \dots + \sigma_r \boldsymbol{v}_r \boldsymbol{u}_r^{\top} \boldsymbol{c}$$
$$= \sigma_1(\boldsymbol{u}_1, \boldsymbol{c}) \boldsymbol{v}_1 + \dots + \sigma_r(\boldsymbol{u}_r, \boldsymbol{c}) \boldsymbol{v}_r$$

となり、A の行の任意の線形結合 $A^{\mathsf{T}}\mathbf{c}$ は、互いに直交する右特異ベクトル $\mathbf{v}_1,\ldots,\mathbf{v}_r$ 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 $\pmb v_1,\ldots,\pmb v_r$ を正規直交基底とする $\pmb r$ 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。

- ・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。
- 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 $\{m{u}_i\}_{i=1}^r$ と右特異ベクトル $\{m{v}_i\}_{i=1}^r$ は、それぞれ列空間 $m{U}$ と行空間 $m{V}$ 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 展開式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、

$$P_{\mathcal{U}} = \sum_{i=1}^r oldsymbol{u}_i oldsymbol{u}_i^ op$$

$$P_{\mathcal{V}} = \sum_{i=1}^r oldsymbol{v}_i oldsymbol{v}_i^ op$$

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

よって、部分空間への射影で議論したように、 $P_{\mathcal{U}}$ 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 Pu をかけても変化しないことが次のように導かれる。

$$P_{\mathcal{U}}A = P_{\mathcal{U}} \left(\sigma_{1} \boldsymbol{u}_{1} \boldsymbol{v}_{1}^{\top} + \cdots + \sigma_{r} \boldsymbol{u}_{r} \boldsymbol{v}_{r}^{\top} \right)$$

$$= \sigma_{1} P_{\mathcal{U}} \boldsymbol{u}_{1} \boldsymbol{v}_{1}^{\top} + \cdots + \sigma_{r} P_{\mathcal{U}} \boldsymbol{u}_{r} \boldsymbol{v}_{r}^{\top}$$

$$= \sigma_{1} \boldsymbol{u}_{1} \boldsymbol{v}_{1}^{\top} + \cdots + \sigma_{r} \boldsymbol{u}_{r} \boldsymbol{v}_{r}^{\top}$$

$$= A$$

行についても同様に、 $P_{\nu} \boldsymbol{v}_{i} = \boldsymbol{v}_{i}$ であるから、 A^{\top} の特異値分解の式

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

の両辺に右から Py をかけても変化しない。

$$AP_{\mathcal{V}} = (\sigma_{1}\boldsymbol{v}_{1}\boldsymbol{u}_{1}^{\top} + \cdots + \sigma_{r}\boldsymbol{v}_{r}\boldsymbol{u}_{r}^{\top}) P_{\mathcal{V}}$$

$$= \sigma_{1}\boldsymbol{v}_{1}P_{\mathcal{V}}\boldsymbol{u}_{1}^{\top} + \cdots + \sigma_{r}\boldsymbol{v}_{r}P_{\mathcal{V}}\boldsymbol{u}_{r}^{\top}$$

$$= \sigma_{1}\boldsymbol{v}_{1}\boldsymbol{u}_{1}^{\top} + \cdots + \sigma_{r}\boldsymbol{v}_{r}\boldsymbol{u}_{r}^{\top}$$

$$= A^{\top}$$

以上をまとめると、次の式が成り立つ。

$$P_{\mathcal{U}}A = A$$
$$AP_{\mathcal{V}} = A^{\top}$$



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。

$$egin{aligned} A &= egin{pmatrix} \sigma_1 oldsymbol{u}_1 & \cdots & \sigma_r oldsymbol{u}_r \end{pmatrix} egin{pmatrix} oldsymbol{v}_1^{ op} \ oldsymbol{v}_r^{ op} \end{pmatrix} \ &= egin{pmatrix} oldsymbol{u}_1 & \cdots & oldsymbol{u}_r \end{pmatrix} egin{pmatrix} oldsymbol{\sigma}_1 & & & \ & \ddots & & \ & & \sigma_r \end{pmatrix} egin{pmatrix} oldsymbol{v}_1^{ op} \ oldsymbol{v}_1^{ op} \ oldsymbol{v}_r^{ op} \end{pmatrix} \end{aligned}$$

ここで、

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 $\{m{v}_1,\ldots,m{v}_n\}$ 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、

$$A = \sigma_{1} \boldsymbol{u}_{1} \boldsymbol{v}_{1}^{\top} + \cdots + \sigma_{r} \boldsymbol{u}_{r} \boldsymbol{v}_{r}^{\top} + o \boldsymbol{v}_{r+1}^{\top} + \cdots + o \boldsymbol{v}_{n}^{\top}$$

$$= \left(\sigma_{1} \boldsymbol{u}_{1} \cdots \sigma_{r} \boldsymbol{u}_{r} \quad o \cdots o\right) \begin{pmatrix} \boldsymbol{v}_{1}^{\top} \\ \vdots \\ \boldsymbol{v}_{r}^{\top} \\ \boldsymbol{v}_{r+1}^{\top} \\ \vdots \\ \boldsymbol{v}_{n}^{\top} \end{pmatrix}$$

$$= \left(\boldsymbol{u}_{1} \cdots \boldsymbol{u}_{m}\right) \begin{pmatrix} \sigma_{1} & & & & \\ & \ddots & & \\ & & \sigma_{r} & & \\ & & & O \end{pmatrix} \begin{pmatrix} \boldsymbol{v}_{1}^{\top} \\ \vdots \\ \boldsymbol{v}_{n}^{\top} \end{pmatrix}$$

● 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、

$$AB = (A_{11}B_{11} + A_{12}B_{21} \quad A_{11}B_{12} + A_{12}B_{22})$$

のように計算できる。

そこで、

$$U_1 = \begin{pmatrix} \boldsymbol{u}_1 & \cdots & \boldsymbol{u}_r \end{pmatrix}, \quad U_2 = \begin{pmatrix} \boldsymbol{u}_{r+1} & \cdots & \boldsymbol{u}_m \end{pmatrix},$$

$$D = \operatorname{diag}(\sigma_1, \ldots, \sigma_r)$$

とおくと、

$$\begin{pmatrix} U_1 & U_2 \end{pmatrix} \begin{pmatrix} D & O \\ O & O \end{pmatrix} = \begin{pmatrix} U_1D + U_2O & U_1O + U_2O \end{pmatrix}$$

$$= \begin{pmatrix} U_1D & O \end{pmatrix}$$

$$= \begin{pmatrix} \sigma_1 \boldsymbol{u}_1 & \cdots & \sigma_r \boldsymbol{u}_r & \boldsymbol{o} & \cdots & \boldsymbol{o} \end{pmatrix}$$

という式変形が確かめられる。

ここで、

$$U = \begin{pmatrix} \boldsymbol{u}_1 & \cdots & \boldsymbol{u}_m \end{pmatrix}, \quad V = \begin{pmatrix} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_n \end{pmatrix},$$
 $\Sigma = \begin{pmatrix} \sigma_1 & & & \\ & \ddots & & O \\ & & \sigma_r & & \\ \hline & O & & O \end{pmatrix} \uparrow_{m-r}$

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。