特異値分解と一般化逆行列

佐賀大学　新井康平

特異値分解

長方形行列Aの*特異値*と対応する特異ベクトルは、スカラ σ

$$Av = \sigma u$$

$$A^T u = \sigma v$$

Σ 対角行列の対角要素に特異値を配置し、
対応する2つの直交行列の列からなる特異ベクトルをUおよびVとすると、
つぎのような関係になります。

$$AV = U\Sigma$$

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

UとVは直交であるため、これは特異値分解 になります。

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

固有値と特異値

正方行列 A による線形変換の固有値・固有ベクタは

$$Ax = \lambda x \tag{0-24}$$

のように定式化される。A が正方行列なので，A による線形変換によってベクタ空間の次元は変わらないから，上のように書くことが許されるわけだ。もし A が正方行列でなく，M 行 N 列の要素で構成されていたらどうだろうか。この場合の線形変換を例えば

$$y = Ax \tag{0-25}$$

と書いてみよう。x は N 次元ベクタであり，変換の結果である y は M 次元ベクタであって両者の次元は違うから，そもそも次式は成り立たない。

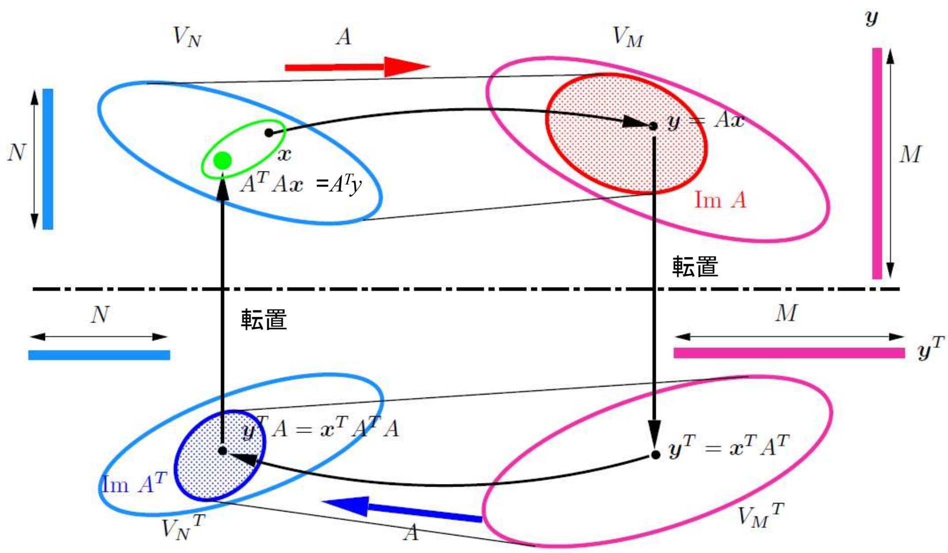
$$y(= Ax) = \lambda x \tag{0-26}$$

そこで，もう一度線形変換 A で逆に N 次元のベクタ空間に戻ったら (数学的にはおかしな表現だが) どうなるか，を考えてみる。y の転置に線形変換 A を行うと次式を得る。

$$y^T A = (Ax)^T A = x^T A^T A \tag{0-27}$$

この結果を転置しても，スカラ倍を含めて x に戻る保証はない。そこで，次のように書けるための A の構造は何かを調べよう，というのがここでのテーマだ (それが A の**特異値**分解に繋がる)。

$$(x^T A^T A)^T = A^T A x = \lambda x \tag{0-28}$$



A^TA は N 次正方行列で対称だからその固有値は実数であり，対応する固有ベクタ (N 次元) は互いに直交する。そこで，これらの固有ベクタを使って V_N の正規直交基底

$$v_n \in V_N, \quad n = 1,\ 2,\ \cdots, N \tag{0-29}$$

を作る (A^TA のランクが気になるが，それについては後述)。これらの固有ベクタに対応する固有値を λ_n, $n = 1,\ 2,\ \cdots, N$ とすると，A^TA の固有値問題は次のように書ける。

$$A^TAv_n = \lambda_n v_n \tag{0-30}$$

上式の左右両辺に左から v_n^T を作用させて N 次正方行列 A^TA の二次形式を作ると，v_n が N 次元ベクタ空間 V_N の正規直交基底であるから次式が成り立つ。

$$v_n^T(A^TAv_n) = v_n^T\lambda_n v_n, \quad \therefore (Av_n)^T(Av_n) = \lambda_n v_n^Tv_n \longrightarrow ||Av_n||^2 = \lambda_n||v_n||^2 = \lambda_n \tag{0-31}$$

上式から A^TA の固有値 λ_n は負の値にはならないことがわかるので，次式が成り立つように適切に $k_n \geq 0$ を選んでおく。

$$\lambda_n = {k_n}^2 \tag{0-32}$$

ところで，A や A^TA のランクは N とは限らない。そこで，A のランクを $R \leq N$ とし，固有値 λ_n も値が大きい順に整列して

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_R > 0, \quad \lambda_{R+1} = \lambda_{R+2} = \cdots = \lambda_N = 0 \tag{0-33}$$

としておこう。すると，値が 0 の固有値に対応する固有値 λ_{R+1}, λ_{R+2}, $\cdots\lambda_N$ に対応する正規直交基底ベクタ v_{R+1}, v_{R+2}, $\cdots, v_N$ に対しては，式 (0-31) より

$$Av_n = 0_M \in V_M, \quad n = R+1,\ R+2,\ \cdots, N \tag{0-34}$$

であり，v_{R+1}, v_{R+2}, $\cdots, v_N$ は A の核 Ker A に属する。

一方，$r \leq R$ の場合は 0_M ベクタ以外に写像されるはずだから，v_r の A による像として適当に $w_r \in V_M$ を選んで

$$Av_r = k_r w_r, \quad r = 1,\ 2,\ \cdots, R \tag{0-35}$$

と置いてみよう。ただし，k_r については上に述べたように，A^TA の固有ベクタ v_r に対応する固有値 λ_r の平方根であることに注意しておこう。N 次元ベクタ空間 V_N の正規直交基底 v_r, $r \leq R$ に対して上式のように書ける w_r の性質は何か，というのがここでのキーポイントになるわけだ。

そこで，適当に $r \leq R$ と $s \leq R$ の二つの番号を選んで式 (0-35) の内積を取り，v_r が正規直交基底をなすことを利用して整理すると次式を得る。

$$(k_rw_r)^T k_sw_s = (Av_r)^T Av_s = v_r^TA^TAv_s = v_r\lambda_sv_s = \lambda_sv_rv_s = \lambda_s\delta_{rs} = \begin{cases} \lambda_s, & (r = s) \\ 0. & (r \neq s) \end{cases} \tag{0-36}$$

上式を整理すると，w_r, w_s $(r, s \leq R)$ に関して次式が成り立つ。

$$\left.\begin{array}{ll} r = s \text{ のとき} & {k_r}^2w_r^Tw_r = \lambda_r = {k_r}^2 \\ r \neq s \text{ のとき} & k_rk_sw_r^Tw_s = 0 \end{array}\right\}, \quad \therefore w_r^Tw_s = \begin{cases} 1, & (r = s) \\ 0. & (r \neq s) \end{cases} \tag{0-37}$$

すなわち，A による v_r の像 w_r, $r \leq R$ が A の像空間 Im A を張る R 個の正規直交基底ベクタを構成することがわかった。この w_r に対しては次式が成り立つことに注意しよう。

$$A^Tw_r = A^T(k_r)^{-1}Av_r = (k_r)^{-1}A^TAv_r = (k_r)^{-1}\lambda_rv_r = k_nv_r, \quad r = 1,\ 2,\ \cdots, R \tag{0-38}$$

つまり，v_r, $r \leq R$ は A^T の像空間 Im A^T を張る正規直交基底ベクタである。

次に M 次正方行列 AA^T について調べてみよう。式 (0-30), (0-35), (0-38) などを利用して整理すると，$\boldsymbol{w}_r$, $r \leq R$ に対して次式を得る。

$$AA^T\boldsymbol{w}_r = Ak_r\boldsymbol{v}_r = k_rA\boldsymbol{v}_r = (k_r)^2\boldsymbol{w}_r = \lambda_r\boldsymbol{w}_r \tag{0-39}$$

λ_r はもともと N 次正方行列 A^TA の固有値であり，大きい順に $\lambda_1 \geq \lambda_2 \geq \cdots \lambda_R > 0$ と付番されていた。この λ_r が同時に M 次正方行列 AA^T の固有値でもあり，$\boldsymbol{w}_r$, $r \leq R$ はそれに対応する固有ベクトルで AA^T の像空間 Im AA^T を張る正規直交基底であることを上式は示している。

ところで，一般に二つの行列 A と B があってそれらの積 BA が存在する場合，行列 BA のランクは A のランクと B のランクの小さい方で決まる。したがって A^TA や AA^T のランクも A のランクと同じく R であるから，M 次元ベクタ空間 V_M で AA^T の像空間 Im AA^T 以外の空間を張る正規直交基底ベクタ $\boldsymbol{w}_m$, $m = R+1,\ R+2, \cdots, M$ に対しては次式が成り立つ。

$$AA^T\boldsymbol{w}_m = \boldsymbol{0}_M \tag{0-40}$$

上式に左から $\boldsymbol{w}_m^T$ を施せば

$$\boldsymbol{w}_m^TAA^T\boldsymbol{w}_m = ||A^T\boldsymbol{w}_m||^2 = 0 \tag{0-41}$$

が成り立つから，結局

$$A^T\boldsymbol{w}_m = \boldsymbol{0}_N \in V_N, \quad m = R+1,\ R+2, \cdots, M \tag{0-42}$$

を得る。すなわち，$\boldsymbol{w}_{R+1}$, $\boldsymbol{w}_{R+2}, \cdots, \boldsymbol{w}_M$ は A^T の核 Ker A^T に属する。

行列 A の特異値分解

M 行 N 列で構成される行列 A に対して，左から M 次単位行列 I_M を，右から N 次単位行列 I_N を，それぞれ作用させても結果は A のままである。

$$\mathrm{I}_M A\ \mathrm{I}_N = A \tag{0-43}$$

ところで，N 次元ベクタ空間 V_N の正規直交基底を $\boldsymbol{v}_n$, $n = 1,\ 2, \cdots, N$ とし，M 次元ベクタ空間の正規直交基底を $\boldsymbol{w}_m$, $m = 1,\ 2, \cdots, M$ とする。これらを使って単位行列 I_N, I_M を次のように作ることができる。

$$\mathrm{I}_N = \sum_{n=1}^{N}\boldsymbol{v}_n\boldsymbol{v}_n^T, \qquad \mathrm{I}_M = \sum_{m=1}^{M}\boldsymbol{w}_m\boldsymbol{w}_m^T \tag{0-44}$$

これより行列 A は次のように書き直せる。

$$\begin{aligned} A &= \sum_{m=1}^{M}\boldsymbol{w}_m\boldsymbol{w}_m^T \cdot A \cdot \sum_{n=1}^{N}\boldsymbol{v}_n\boldsymbol{v}_n^T = \sum_{m=1}^{M}\boldsymbol{w}_m\boldsymbol{w}_m^T\sum_{n=1}^{N}(A)\boldsymbol{v}_n\boldsymbol{v}_n^T = \sum_{m=1}^{M}\sum_{n=1}^{N}\boldsymbol{w}_m\boldsymbol{w}_m^T(A)\boldsymbol{v}_n\boldsymbol{v}_n^T \\ &= \sum_{m=1}^{M}\sum_{n=1}^{N}\boldsymbol{w}_m(\boldsymbol{w}_m^TA\boldsymbol{v}_n)\boldsymbol{v}_n^T \end{aligned} \tag{0-45}$$

ところで，式 (0-34) および (0-35) より

$$\begin{cases} A\boldsymbol{v}_r = k_r\boldsymbol{w}_r & (r \leq R) \\ A\boldsymbol{v}_r = \boldsymbol{0}_M & (r > R) \end{cases} \tag{0-46}$$

だったから，次の関係が成り立つ。

$$\boldsymbol{w}_m^TA\boldsymbol{v}_n = \boldsymbol{w}_m^Tk_n\boldsymbol{w}_m = k_n\delta_{mn} = \begin{cases} 1, & (m = n,\ m, n \leq R) \\ 0, & (m \neq n,\ \text{or } m, n > R) \end{cases} \tag{0-47}$$

ゆえに行列 A は次のように書ける。

$$A = \sum_{r=1}^{R} w_r k_r v_r^T = \sum_{r=1}^{R} k_r w_r v_r^T$$
$$= \begin{bmatrix} | & | & & | \\ w_1 & w_2 & \cdots & w_R \\ | & | & & | \end{bmatrix} \begin{bmatrix} k_1 & & & \\ & k_2 & & \\ & & \ddots & \\ & & & k_R \end{bmatrix} \begin{bmatrix} — & v_1 & — \\ — & v_2 & — \\ & \vdots & \\ — & v_R & — \end{bmatrix} \tag{0-48}$$

これを行列 A の**特異値分解**という。

$r > R$ の正規直交基底ベクタも含めて次のような直交行列 (複素線形空間ならばユニタリ行列) を考える。

$$V = \begin{bmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_N \\ | & | & & | \end{bmatrix}, \quad W = \begin{bmatrix} | & | & & | \\ w_1 & w_2 & \cdots & w_N \\ | & | & & | \end{bmatrix} \tag{0-49}$$

このとき M 行 N 列で構成される行列

ランクはR

$$\widetilde{A} = \left[\begin{array}{cccc|c} k_1 & & & & \\ & k_2 & & & \\ & & \ddots & & 0 \\ & & & k_R & \\ \hline & & 0 & & 0 \end{array}\right] = \left[\begin{array}{c|c} \Lambda_R & 0 \\ \hline 0 & 0 \end{array}\right] = W^T A V \tag{0-50}$$

を A の表現行列という。表現行列を用いると A を次のように書くこともできる。

$$A = W\widetilde{A}V^T \tag{0-51}$$

特異値分解のまとめ

1. M 行 N 列で構成される行列 A のランクを R とする。
2. $A^T A$ の固有値を大きい順にならべて上から R 個の固有値 $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_R > 0$ に対応する固有ベクタを使って正規直交基底ベクタ $v_n,\ n \leq R$ を作る。それ以外の基底ベクタ $v_n,\ n > R$ は例えば Gram-Schmidt の直交化などで作成する。
3. $\lambda_1 = k_1^2,\ \lambda_2 = k_2^2,\ \cdots \lambda_R = k_R^2$ とする。
4. $v_{R+1},\ v_{R+2},\ \cdots v_N \in \mathrm{Ker}A$
5. $Av_r = k_r w_r,\ \ r \leq R$ なる w_r は $w_r \in \mathrm{Im}\ A$
6. w_m は AA^T の固有ベクタである。
7. $A^T w_r = k_r v_r,\ \ r \leq R$ なる v_r は $v_r \in \mathrm{Im}\ A^T$
8. $w_{R+1},\ w_{R+2},\ \cdots w_M \in \mathrm{Ker}A^T$
9. $\mathrm{Im}\ A \perp \mathrm{Ker}\ A^T$
0. $\mathrm{Im}\ A^T \perp \mathrm{Ker}\ A$

Moore-Penrose 型一般化逆行列

最小ノルムかつ最小二乗型の一般化逆行列は $J=0, H=0$ であり，さらに反射型であれば $K=0$ のはずだ。この種の一般化逆行列を Moore-Penrose 型の一般化逆行列という。すなわち，行列 A の Moore-Penrose 型一般化逆行列をとくに A^+ で表すと，A^+ は次のように書ける。

$$A^+ = V\left[\begin{array}{c|c} \Lambda_R^{-1} & 0 \\ \hline 0 & 0 \end{array}\right]W^T = V_R\Lambda_R^{-1}W_R^T \tag{0-68}$$

ただし，V_R は A^TA の固有ベクトルで構成される行列であり，W_R は AA^T の固有ベクトルで構成される行列であった。

今，行列 A の Moore-Penrose 型一般化逆行列が X，Y の二種類あったとしよう。このとき，X は

$$\begin{aligned}
X &= XAX = (XA)^TX && \longleftarrow \text{最小ノルム型の直交射影} \\
&= A^TX^TX = (AYA)^TX^TX && \longleftarrow \text{一般化逆行列の定義による置き換え} \\
&= A^TY^TA^TX^TX \\
&= A^TY^T(XA)^TX = A^TY^TXAX && \longleftarrow \text{最小ノルム型の直交射影} \\
&= A^TY^TX && \longleftarrow \text{反射型} \\
&= (YA)^TX = YAX && \longleftarrow \text{最小ノルム型の直交射影} \\
&= YAYAX && \longleftarrow \text{反射型} \\
&= Y(AY)^T(AX)^Te && \longleftarrow \text{最小二乗型の直交射影} \\
&= YY^TA^TX^TA^T = YY^T(AXA)^T = YY^TA^T && \longleftarrow \text{一般化逆行列の定義による置き換え} \\
&= Y(AY)^T = YAY && \longleftarrow \text{最小二乗型の直交射影} \\
&= Y && \longleftarrow \text{反射型}
\end{aligned} \tag{0-69}$$

のように書き換えられて，結局 Moore-Penrose 型の一般化逆行列は一意であることがわかる。

Kは実数の集合　Rまたは複素数の集合　Cを表すものとします。

K上の線形空間 **V** から線形空間 **W** への写像が

任意の $\mathbf{x}, \mathbf{y} \in \mathbf{V}$ について

$$f(\mathbf{x}+\mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y})$$

任意の $\mathbf{x} \in \mathbf{V}$と $\alpha \in$ Kについて

$$f(\alpha\mathbf{x}) = \alpha f(\mathbf{x})$$

を充たすとき,この性質を線形性といい,このような性質をもつ写像 f を線形写像といいます。

例えば,n 行 m 列

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1m} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2m} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nm} \end{pmatrix}$$

により，K^mから　K^nへの写像

$$f\text{: } \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} \in \mathrm{K}^m \mapsto \mathbf{Ax} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1m} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2m} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nm} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} \in \mathrm{K}^n$$

が定義できますが,これは　K^mから　K^nへの線形写像です。

線形空間:重ね合わせの理

実際 任意の $\mathbf{x},\mathbf{y}\in \mathrm{K}^n$について

$$f(\mathbf{x}+\mathbf{y}) = \mathbf{A}(\mathbf{x}+\mathbf{y}) = \mathbf{A}\mathbf{x}+\mathbf{A}\mathbf{y} = f(\mathbf{x})+f(\mathbf{y})$$

任意の $\mathbf{x}\in \mathrm{K}^m$と $\alpha\in \mathrm{K}$について

$$f(\alpha\mathbf{x}) = \mathbf{A}\alpha\mathbf{x} = \alpha\mathbf{A}\mathbf{x} = \alpha f(\mathbf{x})$$

線形写像

行列の核空間,像空間

n行m列

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1m} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2m} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nm} \end{pmatrix}$$

により, K^mから K^nへの線形写像

$$\text{f: } \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} \in \mathrm{K}^m \mapsto \mathbf{A}\mathbf{x} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & cdots & a_{1m} \\ a_{21} & a_{22} & a_{23} & cdots & a_{2m} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & cdots & a_{nm} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} \in \mathrm{K}^n$$

が定義できます。

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} \in \mathrm{K}^m$$

について，

$$\mathrm{f}(\mathbf{x}) = \mathbf{A}\mathbf{x} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1m} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2m} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nm} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} \in \mathrm{K}^n$$

ですから，行列 $\mathbf{A}$ の各列ベクトル

$$\mathbf{a}_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix}, \mathbf{a}_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{pmatrix}, \cdots, \mathbf{a}_m = \begin{pmatrix} a_{1m} \\ a_{2m} \\ \vdots \\ a_{nm} \end{pmatrix}$$

に注目すると，

$$f(\mathbf{x}) = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_m\mathbf{a}_m$$

ゆえ，

像←Span(列ベクトルの張る空間)

$$Im(f) = span[\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}]$$

$Im(f)$ の次元 $dim(Im(f))$ は

$$\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}$$

のうちの一次独立なベクトルの個数になります。次元←列ベクトルのうち、一次独立なものの個数

$$span[\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}]$$

を $\mathbf{A}$ の像空間といいます。

ベクトルa1,a2,..で張る空間：span[a1,a2,..]

また，行列 $\mathbf{A}$ の各行ベクトル

$$\mathbf{a}_1^T = (a_{11}, a_{12}, \cdots, a_{1m})$$
$$\mathbf{a}_2^T = (a_{21}, a_{22}, \cdots, a_{2m})$$
$$\vdots$$
$$\mathbf{a}_n^T = (a_{n1}, a_{n2}, \cdots, a_{nm})$$

に注目すると，

核は行ベクトルとAの内積が0になるベクトル

$$\begin{aligned} \mathrm{Ker}(\mathrm{f}) &= \{\mathbf{x} \in \mathrm{K}^m | \mathbf{A}\mathbf{x} = \mathbf{0}\} \\ &= \left\{\mathbf{x} \in \mathrm{K}^m | \mathbf{a}_1^T\mathbf{x} = 0, \mathbf{a}_2^T\mathbf{x} = 0, \cdots, \mathbf{a}_m^T\mathbf{x} = 0\right\} \\ &= \{\mathbf{x} \in \mathrm{K}^m | < \mathbf{a}_1, \mathbf{x} >= 0, < \mathbf{a}_2, \mathbf{x} = 0 >, \cdots, < \mathbf{a}_m, \mathbf{x} >= 0\} \end{aligned}$$

で，$Ker(f)$ は，$\mathbf{A}$ の各行ベクトルとの内積が 0 になるベクトルの集合です。これを，$\mathbf{A}$ の核空間といいます。

$dim(Ker(f))$ は，$\mathbf{A}$ の各行ベクトルとの内積が 0 になるベクトルのうち，一次独立なベクトルの最大個数です。

像空間と核空間

線形空間 $\mathbf{V}$ から線形空間 $\mathbf{W}$ への線形写像

$$f:\mathbf{V}\longrightarrow\mathbf{W}$$

から像空間 $Im(f)$ と核空間 $Ker(f)$ と呼ばれる,それぞれ $\mathbf{V}$ と $\mathbf{W}$ の線形部分空間が定義されます。

$$Im(f)=f(\mathbf{V})=\{f(\mathbf{x})|\mathbf{x}\in\mathbf{V}\}$$

$$Ker(f)=f^{-1}(\mathbf{0})=\{\mathbf{x}\in\mathbf{V}|f(\mathbf{x})=\mathbf{0}\}$$

内積が0

これらが部分空間になっていることは以下のようにして確かめられます。

先ず,$Im(f)$ については,

任意の $\mathbf{z},\mathbf{w}\in Im(f)$ について

定義から,ある $\mathbf{x},\mathbf{y}\in\mathbf{V}$ が存在して,

$$\mathbf{z}=f(\mathbf{x}),\mathbf{w}=f(\mathbf{y})$$

となり,これから

$$\mathbf{x}+\mathbf{y}\in\mathbf{V},\quad \mathbf{z}+\mathbf{w}=f(\mathbf{x})+f(\mathbf{y})=f(\mathbf{x}+\mathbf{y})\in Im(f)$$

となり,これから

$$\alpha\mathbf{x}\in Ker(f)$$

任意の $\mathbf{z}Im(f)$ と $\alpha\in\mathbf{K}$ について

定義から,ある $\mathbf{x}\in\mathbf{V}$ が存在して,

$$\mathbf{z}=f(\mathbf{x})$$

となり,これから

$$\alpha\mathbf{z}=\alpha f(\mathbf{x})=f(\alpha\mathbf{x})\in Im(f)$$

次に,$Ker(f)$ については,

任意の $\mathbf{x},\mathbf{y} \in Ker(f)$ について

定義から,

$$f(\mathbf{x}) = \mathbf{0}, f(\mathbf{y}) = \mathbf{0}$$

となり,これから

$$\mathbf{x}+\mathbf{y} \in \mathbf{V},\ \ f(\mathbf{x}+\mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y}) = \mathbf{0}+\mathbf{0} = \mathbf{0}$$

従って,

$$\mathbf{x}+\mathbf{y} \in Ker(f)$$

任意の $\mathbf{x} \in Ker(f)$と $\alpha \in \mathbf{K}$について

定義から,

$$f(\mathbf{x}) = \mathbf{0}$$

であり,

$$\alpha\mathbf{x} \in \mathbf{V},\ \ f(\alpha\mathbf{x}) = \alpha f(\mathbf{x}) = \alpha\mathbf{0} = \mathbf{0}$$

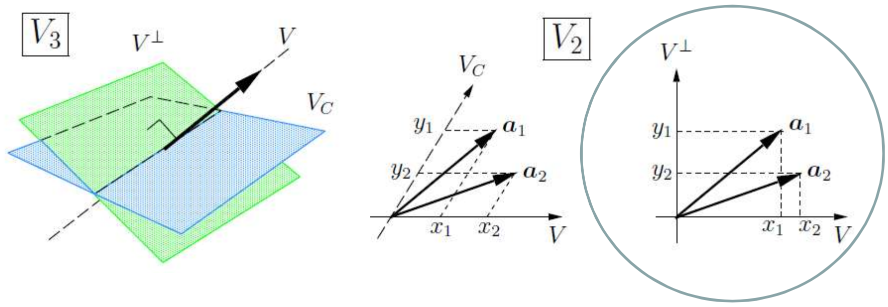
となり,これから

$$\alpha\mathbf{x} \in Ker(f)$$

直交射影、直交補空間

射影: 直交射影:

線形空間内の部分空間を一つ定めたとき，それに対する補空間は無数にある。例えば，3 次元ユークリッド空間内に図のような一次元部分空間 V を定めたとしよう。この一次元部分空間 (つまり直線) を含まないどんな二次元平面も V の補空間 V^C である。しかし，V の直交補空間 $V^{\perp}$ は一つだけ決まる。



簡単のため，二次元線形空間 V_2 のベクタ $\boldsymbol{a}_1$，$\boldsymbol{a}_2 \in V_2$ について考える。部分空間 V とその直交補空間 $V^\perp$ を取って次のように書く。

$$\boldsymbol{a}_1 = \boldsymbol{x}_1 + \boldsymbol{y}_1, \quad \boldsymbol{a}_2 = \boldsymbol{x}_2 + \boldsymbol{y}_2, \qquad \boldsymbol{x}_1, \boldsymbol{x}_2 \in V, \ \ \boldsymbol{y}_1, \boldsymbol{y}_2 \in V^\perp \tag{0-70}$$

直交射影行列，すなわち直交補空間に「沿った」射影行列を P とすると $P\boldsymbol{a}_1 = \boldsymbol{x}_1$，　$P\boldsymbol{a}_2 = \boldsymbol{x}_2$ と書ける。P が直交射影行列であれば $P\boldsymbol{a}_1$，$P\boldsymbol{a}_2 \in V$，かつ $\boldsymbol{y}_1$，$\boldsymbol{y}_2 \in V^\perp$ だから $P\boldsymbol{a}_1 \perp \boldsymbol{y}_2$，$P\boldsymbol{a}_2 \perp \boldsymbol{y}_1$ であることに注意すると，

$$(\boldsymbol{a}_1, P\boldsymbol{a}_2) = (\boldsymbol{x}_1 + \boldsymbol{y}_1, P\boldsymbol{a}_2) = (\boldsymbol{x}_1, P\boldsymbol{a}_2) = (P\boldsymbol{a}_1, P\boldsymbol{a}_2) = (P\boldsymbol{a}_1, P\boldsymbol{a}_2 + \boldsymbol{y}_2) = (P\boldsymbol{a}_1, \boldsymbol{x}_2 + \boldsymbol{y}_2) = (P\boldsymbol{a}_1, \boldsymbol{a}_2) \tag{0-71}$$

が成り立つ。したがって $P^T = P$ である。これが成り立つのは P が直交射影行列だからである。

ベクタ空間 V_N での射影行列を P とし，$\boldsymbol{x} \in V_N$ を $\boldsymbol{x} = P\boldsymbol{x} + (I_N - P)\boldsymbol{x}$ のように分解して整理すると，

$$(P\boldsymbol{x})^T (I_N - P)\boldsymbol{x} = \boldsymbol{x}^T (P^T - P^T P)\boldsymbol{x} \tag{0-72}$$

と書ける。P は射影行列であるので $PP = P$ を満たすが，さらに $P^T = P$ を満たす (つまり直交射影行列である) とき $P^T = P^T P$ と書けるから，上式は 0 となる。すなわち，$P\boldsymbol{x}$ と $(I_N - P)\boldsymbol{x}$ は互いに直交するので

$$||\boldsymbol{x}||^2 = ||P\boldsymbol{x}||^2 + ||(I_N - P)\boldsymbol{x}||^2 \geq ||P\boldsymbol{x}||^2, \quad \Longrightarrow \underline{||\boldsymbol{x}||^2 \geq ||P\boldsymbol{x}||^2} \tag{0-73}$$

が常に成り立つ。しかし，P が直交射影行列でない場合は必ずしも上式は成り立たない (どんな場合か ?)。

直交射影と最小ノルム

解くべき未知数の数に比べて方程式の数が少ない場合のように，線形方程式 $A\boldsymbol{x} = \boldsymbol{y}$ の解は一意とは限らない (例えば Ker $A \neq \{0\}$ などの場合)。このような場合は，他に何らかの知識に基づいた拘束条件を付加して，例えばノルム最小の解を求めることがある。A の一般化逆行列 $\overline{A}$ がノルム最小型であるためには，$\overline{A}A$ が直交射影行列でなければならないことを前節で述べた。

射影行列 $P = \overline{A}A$ が直交射影行列となるような $\overline{A}$ を特に A^- と書くことにし，このときに作られる直交射影行列を $P^- = A^-A$ と書くものとする。まず，次式が成り立つことを確認しておこう。

$$P^-P = A^-A \cdot \overline{A}A = A^- \cdot A\overline{A}A = A^-A = P^- \tag{0-74}$$

線形方程式 $A\boldsymbol{x} = \boldsymbol{y} \in \mathrm{Image}A$，$\boldsymbol{y} \neq \boldsymbol{0}$ が解を持つということは，適当なベクタ $\boldsymbol{v} \notin \mathrm{Ker}A$ が存在して

$$A\boldsymbol{v} = \boldsymbol{y} \tag{0-75}$$

と書けることだ。そこで，単なる反射型一般化逆行列で求めた解 $\boldsymbol{x}$ とノルム最小型一般化逆行列で求めた解 $\boldsymbol{x}^-$ のノルムを比べてみる。

$$\boldsymbol{x} = \overline{A}\boldsymbol{y} = \overline{A}A\boldsymbol{v} = P\boldsymbol{v}, \quad \boldsymbol{x}^- = A^-\boldsymbol{y} = A^-A\boldsymbol{v} = P^-\boldsymbol{v} \tag{0-76}$$

すると常に次式が成り立ち，ノルム最小型一般化逆行列で求めた解のノルムが最も最小となることがわかる。

$$||\boldsymbol{x}||^2 = ||\overline{A}A\boldsymbol{v}||^2 = ||P\boldsymbol{v}||^2 \geq ||P^-P\boldsymbol{v}||^2 = ||P^-\boldsymbol{v}||^2 = ||\boldsymbol{x}^-||^2 \Longrightarrow \underline{||\boldsymbol{x}||^2 \geq ||\boldsymbol{x}^-||^2} \tag{0-77}$$

最小二乗

$A\overline{A}$ が直交射影行列となるときに $\overline{A}$ を最小二乗型逆行列と言ったので，これをとくに A^- と書くことにしよう。このときの直交射影行列を $AA^- = P^-$ と表す。

線形方程式 $A\boldsymbol{x} = \boldsymbol{y} \in \mathrm{Image}\, A, \boldsymbol{y} \neq \boldsymbol{0}$ に対して，単なる一般化逆行列 $\overline{A}$ を用いた場合の誤差 J と最小二乗型逆行列を用いた場合との誤差 J^- の大きさを比べてみる。

$$J = ||A\boldsymbol{x} - \boldsymbol{y}||^2 = ||A\overline{A}\boldsymbol{y} - \boldsymbol{y}||^2 = ||P\boldsymbol{y} - \boldsymbol{y}||^2 = ||(P - I_M)\boldsymbol{y}||^2 \quad (0\text{-}78)$$

$$J^- = ||A\boldsymbol{x} - \boldsymbol{y}||^2 = ||AA^-\boldsymbol{y} - \boldsymbol{y}||^2 = ||P^-\boldsymbol{y} - \boldsymbol{y}||^2 = ||(P^- - I_M)\boldsymbol{y}||^2 \quad (0\text{-}79)$$

ここで，$P^-P = (AA^-)(A\overline{A}) = (AA^-A)\overline{A} = A\overline{A} = P$ であるから次式が成り立つ。

$$(P^- - I_M)(P - I_M) = P^-P - P^- - P + I_M = -(P^- - I_M) \quad (0\text{-}80)$$

ところで，P^- が (Image$A$ への) 直交射影であるから，$P^- - I_M$ は直交補空間への直交射影行列である。したがって，式 (0-73) 前後の議論や式 (0-77) と同様に次式が常に成り立つ。

$$J = ||(P - I_M)\boldsymbol{y}||^2 \geq ||(P^- - I_M)(P - I_M)\boldsymbol{y}||^2 = ||(P^- - I_M)\boldsymbol{y}||^2 = J^-, \quad \therefore \underline{J \geq J^-} \quad (0\text{-}81)$$

例: 2 次元ベクタ空間 V_2 から 3 次元ベクタ空間 V_3 への写像

$$A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \\ 1 & -1 \end{bmatrix} \quad (0\text{-}82)$$

を例にとって，その Moore-Penrose 型一般逆行列を求めてみよう。まず，行列 A^TA の固有値分解を求める。

$$A^TA = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix} \quad (0\text{-}83)$$

上式の固有方程式

$$\det \begin{bmatrix} 3-\lambda & -1 \\ -1 & 3-\lambda \end{bmatrix} = (3-\lambda)^2 - 1 = 0 \quad (0\text{-}84)$$

を解いて次式を得る (つまり，A は V_2 に関してはフルランクである)。

$$\lambda^2 - 6\lambda + 8 = (\lambda - 4)(\lambda - 2) = 0, \quad \therefore \text{固有値} = \begin{cases} \lambda_1 = 4 \\ \lambda_2 = 2 \end{cases}, \quad \text{特異値} = \begin{cases} \sqrt{\lambda_1} = k_1 = 2 \\ \sqrt{\lambda_2} = k_2 = \sqrt{2} \end{cases} \quad (0\text{-}85)$$

これらの固有値に対応する固有ベクタ $\boldsymbol{v}_1, \boldsymbol{v}_2 \in V_2$ の長さを 1 に正規化して次のようにして求める。

$$\begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 4 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad \Longrightarrow x_1 = -x_2, \quad \therefore \boldsymbol{v}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix},$$

$$\begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 2 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad \Longrightarrow x_1 = x_2, \quad \therefore \boldsymbol{v}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad (0\text{-}86)$$

$\boldsymbol{v}_1, \boldsymbol{v}_2 \in V_2$ に対応する V_3 内の正規直交基底は次のように求められる (もう一つの基底 $\boldsymbol{w}_3$ はこれらに直交する長さ 1 のベクタを選ぶ)。

$$w_1 = \frac{1}{k_1}Av_1 = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ -1 & 1 \\ 1 & -1 \end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \tag{0-87}$$

$$w_2 = \frac{1}{k_2}Av_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \\ 1 & -1 \end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \tag{0-88}$$

したがって A の特異値分解は次のように書ける。

$$\begin{aligned} A &= \sum_{r=1}^{2} k_r w_r v_r{}^T = 2\frac{1}{\sqrt{2}}\begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \end{bmatrix} + \sqrt{2}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \end{bmatrix} \\ &= 2\begin{bmatrix} 0 & 0 \\ -1/2 & 1/2 \\ 1/2 & -1/2 \end{bmatrix} + \sqrt{2}\begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \\ 1 & -1 \end{bmatrix} \end{aligned} \tag{0-89}$$

これに対して A の Moore-Penrose 型一般逆行列 A^+ は次のように表すことができる。

$$\begin{aligned} A^+ &= \sum_{r=1}^{2} k_r^{-1} v_r w_r{}^T = \frac{1}{2}\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix} 0 & -1 & 1 \end{bmatrix} + \frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \\ &= \frac{1}{2}\begin{bmatrix} 0 & -1/2 & 1/2 \\ 0 & 1/2 & -1/2 \end{bmatrix} + \frac{1}{\sqrt{2}}\begin{bmatrix} 1/\sqrt{2} & 0 & 0 \\ 1/\sqrt{2} & 0 & 0 \end{bmatrix} = \frac{1}{4}\begin{bmatrix} 2 & -1 & 1 \\ 2 & 1 & -1 \end{bmatrix} \end{aligned} \tag{0-90}$$

この例では A のランクが 2 であったので，2 行 2 列で構成される A^+A は単位行列 I_2 となるが，3 行 3 列となる AA^+ は残念ながら I_3 にはならない。ただし，$A^+AA^+ = A^+$，$AA^+A = A$ は成り立つので A^+ は**反射型**逆行列である。また，$(A^+A)^T = A^+A$ となっているから**最小ノルム型**であり，$(AA^+)^T = AA^+$ が成り立つので **最小二乗型**でもある。

まとめ

行列 A が

$$A = W\left[\begin{array}{c|c} \Lambda_R & 0 \\ \hline 0 & 0 \end{array}\right]V^T$$

と特異値分解されるとき，A の一般化逆行列 $\overline{A}$ は次式で与えられる。

$$\overline{A} = V\left[\begin{array}{c|c} \Lambda_R^{-1} & \text{任意} \\ \hline \text{任意} & \text{任意} \end{array}\right]W^T$$

上の「任意」の部分行列によって様々な一般化逆行列が存在する。