システム工学 I
第 9 回
一般化逆行列と
特異値分解

はじめに (1)

- 連立一次方程式 $\boldsymbol{Ax} = \boldsymbol{b}$ の解は, $\boldsymbol{A}$ が正則であれば, $\boldsymbol{A}$ の逆行列 $\boldsymbol{A}^{-1}$ を用いることにより, $\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}$ と書ける. $\boldsymbol{A}$ が正則でない (あるいは正方行列でない) 場合にこの考え方を拡張したものが一般化逆行列である.
- 固有値を正方でない行列に適用できるように拡張したものが特異値である.

はじめに (2)

- 工学では連立一次方程式は応用上解くべき問題を数式で表現することで得られるが, 数式表現を作る際に, 変数の数と方程式の数が合わないということがよく起こる. したがって, 正方行列を前提とした数学的手法は不自由であって, 一般化逆行列や特異値といった考え方が必要になる.

一般化逆行列 (1)

- 一般化逆行列の考え方を理解するために, $\boldsymbol{A}$ を m 行 n 列の行列とし, 連立一次方程式 $\boldsymbol{Ax} = \boldsymbol{b}$ を解く問題を考える.
- 行列 $\boldsymbol{A}$ のランクを r とする ($r \leq \min\{m, n\}$).
- $\boldsymbol{A}$ を基本変形によって階段行列に変形すると (空白の部分は零, $*$ の部分は任意) $\cdots$

$$
\begin{array}{rc}
 & \begin{array}{ccccccc} i_1 & & i_2 & & & i_r & \end{array} \\
\begin{array}{r} 1) \\ \vdots \\ \vdots \\ r) \\ r+1) \\ \vdots \\ n) \end{array} &
\left(\begin{array}{ccccccc}
1 & * & \cdots & & & & \\
 & & 1 & * & \cdots & & \\
 & & & & \cdots & & \\
 & & & & & 1 & * \\
\hline
 & & & & & & \\
 & & & & & & \\
 & & & & & &
\end{array}\right)
\end{array}
$$

一般化逆行列 (3)

- $\boldsymbol{A}$ の行基本変形は, 基本行列を $\boldsymbol{A}$ に左から掛けることに相当する. $\boldsymbol{U}_L \cdots \boldsymbol{U}_1 \boldsymbol{A}$ が上記の階段行列になったものとし, $\boldsymbol{U} = \boldsymbol{U}_L \cdots \boldsymbol{U}_1$ と定義する.
- $\boldsymbol{\beta} = \boldsymbol{U}\boldsymbol{b}$ と定義する.

一般化逆行列 (4)

- $\{j_1, \ldots, j_{n-r}\} = \{1, \ldots, n\} \setminus \{i_1, \ldots, i_r\}$ とする. ただし, $j_1 < \cdots < j_r$ とする.
- $\boldsymbol{x}$ の添字を $(i_1, \ldots, i_r, j_1, \ldots, j_{n-r})$ の順に並べかえると, 次ページのようになる.

$$
\left(\begin{array}{ccc|ccc}
1 & * & \cdots & \cdots & \cdots & * \\
 & \ddots & \ddots & & & \vdots \\
 & & 1 & * & \cdots & * \\
\hline
 & & & & & \\
 & & & & & \\
 & & & & &
\end{array}\right)
\left(\begin{array}{c}
x_1 \\ \vdots \\ x_{i_r} \\ \hline x_{j_1} \\ \vdots \\ x_{j_{n-r}}
\end{array}\right)
=
\left(\begin{array}{c}
\beta_1 \\ \vdots \\ \beta_r \\ \hline \beta_{r+1} \\ \vdots \\ \beta_m
\end{array}\right)
$$

一般化逆行列 (6)

- $\boldsymbol{x}_B = (x_{i_1}, \ldots, x_{i_r})^T$, $\boldsymbol{x}_N = (x_{j_1}, \ldots, x_{j_{n-r}})^T$, $\boldsymbol{\beta}_1 = (\beta_1, \ldots, \beta_r)^T$, $\boldsymbol{\beta}_2 = (\beta_{r+1}, \ldots, \beta_m)^T$ とおく．また, $\boldsymbol{A}_B = \begin{pmatrix} 1 & & * \\ & \ddots & \\ & & 1 \end{pmatrix}$ を前ページの行列の左上のブロックとし, その右のブロックを $\boldsymbol{A}_N$ とおく．

一般化逆行列 (7)

- 以上の定義を使うと, 先ほどの連立一次方程式は, 次のように書ける.

$$\begin{aligned} \boldsymbol{A}_B \boldsymbol{x}_B + \boldsymbol{A}_N \boldsymbol{x}_N &= \boldsymbol{\beta}_1 \\ \boldsymbol{0} &= \boldsymbol{\beta}_2 \end{aligned}$$

- 数学的には, 解が存在するための必要十分条件は $\boldsymbol{\beta}_2 = \boldsymbol{0}$ である. 当面はこれを仮定する.

一般化逆行列 (8)

- $\boldsymbol{A}_B$ は正則だから, 連立一次方程式 $\boldsymbol{A}_B\boldsymbol{x}_B + \boldsymbol{A}_N\boldsymbol{x}_N = \boldsymbol{\beta}_1$ は解を持ち, その解は $\boldsymbol{x}_B = -\boldsymbol{A}_B^{-1}\boldsymbol{A}_N\boldsymbol{x}_N + \boldsymbol{A}_B^{-1}\boldsymbol{\beta}_1$ である. $\boldsymbol{x}_N$ はフリーパラメータである.

- 解をまとめて書くと ⋯

$$\begin{pmatrix} \boldsymbol{x}_B \\ \boldsymbol{x}_N \end{pmatrix} = \begin{pmatrix} -\boldsymbol{A}_B^{-1}\boldsymbol{A}_N \\ \boldsymbol{I}_{n-r} \end{pmatrix} \boldsymbol{x}_N + \begin{pmatrix} \boldsymbol{A}_B^{-1} \begin{pmatrix} \boldsymbol{I}_r & \boldsymbol{0}_{r\times(m-r)} \end{pmatrix} \boldsymbol{U} \\ \boldsymbol{0}_{(n-r)\times m} \end{pmatrix} \boldsymbol{b}$$

一般化逆行列 (9)

- 次に, $\boldsymbol{x}_N$ を $\boldsymbol{b}$ から決めることを考える. $\boldsymbol{x}_N$ は任意だから, $\boldsymbol{K}$ を任意の $n-r$ 行 m 列の行列とし, $\boldsymbol{x}_N = \boldsymbol{K}\boldsymbol{b}$ とすればよい. すると

$$\begin{pmatrix} \boldsymbol{x}_B \\ \boldsymbol{x}_N \end{pmatrix} = \begin{pmatrix} -\boldsymbol{A}_B^{-1}\boldsymbol{A}_N\boldsymbol{K} + \boldsymbol{A}_B^{-1}\begin{pmatrix} \boldsymbol{I}_r & \boldsymbol{0}_{r\times(m-r)} \end{pmatrix}\boldsymbol{U} \\ \boldsymbol{K} \end{pmatrix} \boldsymbol{b}$$

- これは, $\boldsymbol{A}$ が正則な場合の解の表現 $\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}$ と似た形になっている.

一般化逆行列 (10)

- $\boldsymbol{x}_B = \boldsymbol{A}_B^{-1}\boldsymbol{\beta}_1$ 以外の解を探す理由は $\cdots$
- たとえば, $\boldsymbol{x}_B$ の各成分に消費エネルギーという物理的な意味がある場合には, フリーパラメータをうまく調整して, 消費エネルギー最小の解を求めることが望ましい.
- 次ページの例を考える.

一般化逆行列 (11)

- $$\left(\begin{array}{cc|c} 10^{-1} & 0 & 1 \\ 0 & 1 & 0 \end{array}\right)\begin{pmatrix} x_1 \\ x_2 \\ \hline x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$
- $x_1 + x_2 + x_3$ が小さい解が望ましいものとする
- $x_2 = 1$ は一意的. x_1 と x_3 は一意的ではなく, $10^{-3}x_1 + x_3 = 1$ を満たせばよい.

一般化逆行列 (12)

- $(x_1, x_3) = (10^3, 0)$ は解だが, $(x_1, x_3) = (0, 1)$ も解. 後者の方が $x_1 + x_3$ が小さい (この例では $(x_1, x_3) = (0, 1)$ が $x_1 + x_2 + x_3$ を最小にする解になっている)
- このように, 何らかの目的でフリーパラメータを調整する可能性があるため, その調整の余地を残しておきたい

一般化逆行列 (13)

- $\boldsymbol{A}$ が正則なら連立一次方程式 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ の解が $\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}$ であることを踏まえ, 上記の例のように, うまく「逆行列のような行列」$\boldsymbol{A}^{\dagger}$ を取って, どのような $\boldsymbol{b}$ に対しても, $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ が解を持つのであれば, $\boldsymbol{x} = \boldsymbol{A}^{\dagger}\boldsymbol{b}$ が解を与えるようにしたい ($\boldsymbol{A}^{\dagger}$ には調整の余地がある)

一般化逆行列 (14)

- このような条件を満たす行列 $\boldsymbol{A}^{\dagger}$ を, $\boldsymbol{A}$ の一般化逆行列 (あるいは一般逆行列) という.
- この定義でまず問題となるのは, 一般化逆行列が存在するか否かであるが, 我々はすでにある種の一般化逆行列を構成している.

一般化逆行列 (15)

- $\boldsymbol{x} = \boldsymbol{V}(\boldsymbol{x}_B^T, \boldsymbol{x}_N^T)^T$ とすると ($\boldsymbol{V}$ は成分の並べかえに対応する行列), 先に述べた結果から,
$$\boldsymbol{V}\begin{pmatrix} -\boldsymbol{A}_B^{-1}\boldsymbol{A}_N\boldsymbol{K} + \boldsymbol{A}_B^{-1}\begin{pmatrix}\boldsymbol{I}_r & \boldsymbol{0}_{r\times(m-r)}\end{pmatrix}\boldsymbol{U} \\ \boldsymbol{K} \end{pmatrix}$$
が一般化逆行列になる. $\boldsymbol{K}$ はフリーパラメータから成る行列である.

一般化逆行列 (16)

- $\boldsymbol{A}$ を m 行 n 列の行列とする．このとき, n 行 m 列の行列 $\boldsymbol{A}^{\dagger}$ が $\boldsymbol{A}$ の一般化逆行列であるための必要十分条件は, $\boldsymbol{A}\boldsymbol{A}^{\dagger}\boldsymbol{A} = \boldsymbol{A}$ となることである (証明は次ページ)
- したがって, 一般化逆行列をひとつ決めるということは, $\boldsymbol{A}\boldsymbol{A}^{\dagger}\boldsymbol{A} = \boldsymbol{A}$ を満たす $\boldsymbol{A}^{\dagger}$ をひとつ決める, ということである.

- $\mathrm{rank}(\boldsymbol{A}, \boldsymbol{b}) = \mathrm{rank}\boldsymbol{A}$ であるということは, ある n 次のベクトル $\boldsymbol{z}$ が存在し, $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{z}$ となるということである. この性質を使って先に述べた等価性を示す.
- $\boldsymbol{A}^{\dagger}\boldsymbol{b}$ が任意の $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{z}$ に対して $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ の解になっているということは, $\forall\boldsymbol{z}, \boldsymbol{A}\boldsymbol{A}^{\dagger}\boldsymbol{A}\boldsymbol{z} = \boldsymbol{A}\zeta$, すなわち $\boldsymbol{A}\boldsymbol{A}^{\dagger}\boldsymbol{A} = \boldsymbol{A}$ を意味する.
- 逆に, $\boldsymbol{A}\boldsymbol{A}^{\dagger}\boldsymbol{A} = \boldsymbol{A}$ とすると, $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{z}$ に対し, $\boldsymbol{x} = \boldsymbol{A}^{\dagger}\boldsymbol{b} = \boldsymbol{A}^{\dagger}\boldsymbol{A}\boldsymbol{z}$ とおくと, $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}\boldsymbol{A}^{\dagger}\boldsymbol{A}\boldsymbol{z} = \boldsymbol{A}\boldsymbol{z} = \boldsymbol{b}$ だから, $\boldsymbol{x} = \boldsymbol{A}^{\dagger}\boldsymbol{b}$ は $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ の解である. よって, $\boldsymbol{A}^{\dagger}$ は $\boldsymbol{A}$ の一般化逆行列である.

一般化逆行列 (18)

- 次に，一般化逆行列を完全に特徴付ける．まず第一に，$\boldsymbol{A}^{\dagger}\boldsymbol{b}$ は $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ の解だから，$\boldsymbol{A}^{\dagger}$ は n 行 m 列の行列でなければならない．
- $\boldsymbol{U}\boldsymbol{A}\boldsymbol{V}=\begin{pmatrix}\boldsymbol{A}_B & \boldsymbol{A}_N\\ \boldsymbol{0} & \boldsymbol{0}\end{pmatrix}$ に列基本変形を施して $\boldsymbol{A}_B$ を単位行列に変え，$\boldsymbol{A}_N$ を消去する．対応する基本行列を $\boldsymbol{V}'$ とすると $\cdots$

一般化逆行列 (19)

- $\boldsymbol{UAVV'} = \begin{pmatrix} \boldsymbol{I}_r & \boldsymbol{0}_{r\times(n-r)} \\ \boldsymbol{0}_{(m-r)\times r} & \boldsymbol{0}_{(m-r)\times(n-r)} \end{pmatrix}$
- $\boldsymbol{VV'} = \boldsymbol{W}$ と定義し, 上式に左から $\boldsymbol{U}^{-1}$, 右から $\boldsymbol{W}^{-1}$ を掛けると ⋯

$$\boldsymbol{A} = \boldsymbol{U}^{-1} \begin{pmatrix} \boldsymbol{I}_r & \boldsymbol{0}_{r\times(n-r)} \\ \boldsymbol{0}_{(m-r)\times r} & \boldsymbol{0}_{(m-r)\times(n-r)} \end{pmatrix} \boldsymbol{W}^{-1}$$

一般化逆行列 (20)

- $\boldsymbol{A}^{\dagger} = \boldsymbol{W} \begin{pmatrix} \boldsymbol{D}_1 & \boldsymbol{D}_2 \\ \boldsymbol{D}_3 & \boldsymbol{D}_4 \end{pmatrix} \boldsymbol{U}$ とおき, $\boldsymbol{D}_1$ から $\boldsymbol{D}_4$ までが満たすべき条件を求める. ただし $\boldsymbol{D}_1$ は r 次の正方行列である. $\boldsymbol{D}_2$ 等の行と列の大きさは, ここから自動的に決まる.
- $\boldsymbol{A}\boldsymbol{A}^{\dagger}\boldsymbol{A} = \boldsymbol{A}$ にこれらを代入すると $\cdots$

一般化逆行列 (21)

$$\boldsymbol{U}^{-1}\begin{pmatrix} \boldsymbol{D}_1 & \boldsymbol{0}_{r\times(n-r)} \\ \boldsymbol{0}_{(m-r)\times r} & \boldsymbol{0}_{(m-r)\times(n-r)} \end{pmatrix}\boldsymbol{W}^{-1} = \boldsymbol{U}^{-1}\begin{pmatrix} \boldsymbol{I}_r & \boldsymbol{0}_{r\times(n-r)} \\ \boldsymbol{0}_{(m-r)\times r} & \boldsymbol{0}_{(m-r)\times(n-r)} \end{pmatrix}\boldsymbol{W}^{-1}$$

したがって, $\boldsymbol{D}_1 = \boldsymbol{I}_r$ で, $\boldsymbol{D}_2$, $\boldsymbol{D}_3$, $\boldsymbol{D}_4$ は任意. $\boldsymbol{U}$ と $\boldsymbol{W}$ は正則行列だが, 一意的ではない.

一般化逆行列 (22)

- 一般化逆行列が持つ自由度を利用して, 何らかの意味 (後述) で都合が良い一般化逆行列を構成する, ということがおこなわれる.
- 上記に関する議論に先立って, まず QR 分解について説明する.

QR分解 (1)

- $\boldsymbol{A}$ を m 行 n 列で階数 r の実あるいは複素行列とする.
- $\boldsymbol{A}$ の列をならべかえることで, $\boldsymbol{A}$ の左側の r 個の列ベクトルが線形独立であるようにできる. このような列の並べかえに対応する行列を $\boldsymbol{P}$ とし, $\boldsymbol{AP} = (\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r, \boldsymbol{\alpha}_{r+1}, \ldots, \boldsymbol{\alpha}_n)$ とする. $(\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r)$ は線形独立である.

QR 分解 (2)

- $(\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r)$ から Gram-Schmidt の直交化法によって作った正規直交系を $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_r)$ とし (実あるいは複素内積を使う), これらを含む正規直交基底 $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_r, \boldsymbol{v}_{r+1}, \ldots, \boldsymbol{v}_m)$ を構成する. このとき, $\boldsymbol{\alpha}_i$ $(1 \leq i \leq r)$ は $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_r)$ の線形結合である.

QR 分解 (3)

- $(\boldsymbol{\alpha}_{r+1}, \ldots, \boldsymbol{\alpha}_n)$ は $(\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r)$ に線形従属だから，これらは $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_r)$ の線形結合となる．
- 以下しばらく, 標準基底を固定し, ベクトルを標準基底に関する数ベクトルと同一視する．
- $\boldsymbol{Q} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m)$ とする．

QR 分解 (4)

- $\boldsymbol{A}$ が実行列の場合には, $\boldsymbol{Q}$ は直交行列で, 次式が成り立つ.

$$
\boldsymbol{Q}^T(\boldsymbol{AP}) = \begin{array}{cc}
 & \begin{array}{cccccc} \underbrace{1} & \cdots & \underbrace{r} & \underbrace{r+1} & \cdots & \underbrace{n} \end{array} \\
\begin{array}{r} 1) \\ \vdots \\ r) \\ r+1) \\ \vdots \\ m) \end{array} &
\begin{pmatrix}
* & \cdots & * & \cdots & \cdots & * \\
 & \ddots & \vdots & & & \vdots \\
 & & * & \cdots & \cdots & * \\
 & & & & & \\
 & & & & & \\
 & & & & &
\end{pmatrix}
\end{array}
$$

QR 分解 (5)

- $\boldsymbol{A}$ が複素行列の場合には, $\boldsymbol{Q}$ はユニタリ行列で, 次式が成り立つ.

$$
\boldsymbol{Q}^*(\boldsymbol{AP}) = \begin{array}{c} \\ 1) \\ \vdots \\ r) \\ r+1) \\ \vdots \\ m) \end{array}
\begin{array}{c}
\begin{array}{cccccc} \underbrace{}_{1} & \cdots & \underbrace{}_{r} & \underbrace{}_{r+1} & \cdots & \underbrace{}_{n} \end{array} \\
\begin{pmatrix}
* & \cdots & * & \cdots & \cdots & * \\
 & \ddots & \vdots & & & \vdots \\
 & & * & \cdots & \cdots & * \\
 & & & & & \\
 & & & & & \\
 & & & & &
\end{pmatrix}
\end{array}
$$

QR 分解 (6)

- いずれの場合も, 右辺の行列を $\boldsymbol{R}$ とすると, $\boldsymbol{AP} = \boldsymbol{QR}$ となる．このような行列の分解表現を QR 分解という.
- 以上で見たように, 任意の実あるいは複素行列は, 適切に列をならべかえることで, QR 分解できる.

QR 分解 (7)

- QR 分解で $\boldsymbol{R} = \left(\begin{array}{c|c} \boldsymbol{R}_1 & \boldsymbol{R}_2 \\ \hline \boldsymbol{0}_{(m-r)\times r} & \boldsymbol{0}_{(m-r)\times(n-r)} \end{array}\right)$ とおくと, $\boldsymbol{R}$ の右下がりの対角線より下の要素はすべて零で (このような行列を上三角行列などという), $\boldsymbol{Q}$ が $\boldsymbol{AP}$ の左側の r 列から Gram-Schmidt の直交化法により構成されていることから, $\boldsymbol{R}_1$ は正則である.

最小ノルム型一般化逆行列 (1)

- $\boldsymbol{Ax} = \boldsymbol{b}$ を満たす解が複数あるとき, その解の中でノルムが最小のものを与える一般化逆行列を最小ノルム型一般化逆行列といい, $\boldsymbol{A}^{\vee}$ であらわす.
- 以下の議論では, 重複を避けるため, $\boldsymbol{A}$ が複素行列の場合のみを考える. 実行列では, $\quad^*$ を $\quad^T$ に置き換えればよい.

最小ノルム型一般化逆行列 (2)

- $\boldsymbol{A}^*\boldsymbol{P}$ が QR 分解できるような置換行列 $\boldsymbol{P}$ を取り, $\boldsymbol{AP} = \boldsymbol{QR}$ とする. $\boldsymbol{Q}^*\boldsymbol{AP} = \boldsymbol{R}$ だから, $\boldsymbol{P}^*\boldsymbol{AQ} = \boldsymbol{R}^* = \begin{pmatrix} \boldsymbol{R}_1^* & \boldsymbol{0} \\ \boldsymbol{R}_2^* & \boldsymbol{0} \end{pmatrix}$ である.
- 上記の右辺が階段行列となるように行基本変形を施す.

最小ノルム型一般化逆行列 (3)

- $\boldsymbol{T}$ をこれに対応する基本行列の積とし, $\boldsymbol{S} = \boldsymbol{T}\boldsymbol{P}^*$ と定義すると, $\boldsymbol{S}\boldsymbol{A}\boldsymbol{Q} = \begin{pmatrix} \boldsymbol{I}_r & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}$
- 上記が一般化逆行列の「一般形」と異なるのは, $\boldsymbol{Q}$ がユニタリ行列に限定されていることである. このようにするのは, ベクトルの直交性を議論に組み込むため.

最小ノルム型一般化逆行列 (4)

- $\boldsymbol{z} = \boldsymbol{Q}\boldsymbol{x}$ とおくと, 解くべき連立一次方程式は $\boldsymbol{S}\boldsymbol{A}\boldsymbol{Q}\boldsymbol{z} = \boldsymbol{S}\boldsymbol{b}$ となる. $\boldsymbol{S}\boldsymbol{b}$ の第 1 要素から第 r 要素までをならべたベクトルを $\boldsymbol{\beta}_1$, 残りを $\boldsymbol{\beta}_2$ とする.
- $\boldsymbol{\beta}_2 = \boldsymbol{0}$ のとき, $\begin{pmatrix} \boldsymbol{I}_r & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix} \boldsymbol{z} = \begin{pmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \end{pmatrix}$ のノルム最小の解は $\cdots$

最小ノルム型一般化逆行列 (5)

- $\boldsymbol{z} = \begin{pmatrix} \boldsymbol{I}_r & \boldsymbol{D}_2 \\ \boldsymbol{0} & \boldsymbol{D}_4 \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \end{pmatrix}$ がノルム最小の解で, $\boldsymbol{\beta}_2 = \boldsymbol{0}$ を仮定するから, $\boldsymbol{D}_2$ と $\boldsymbol{D}_4$ は任意.
- もとの座標系では $\boldsymbol{x} = \boldsymbol{Q} \begin{pmatrix} \boldsymbol{I}_r & \boldsymbol{D}_2 \\ \boldsymbol{0} & \boldsymbol{D}_4 \end{pmatrix} \boldsymbol{S}\boldsymbol{b}$

最小ノルム型一般化逆行列 (6)

- 以上により, 最小ノルム型一般化逆行列の一般形は次の通り：
  $\boldsymbol{A}^{\vee} = \boldsymbol{Q}\begin{pmatrix} \boldsymbol{I}_r & \boldsymbol{D}_2 \\ \boldsymbol{0} & \boldsymbol{D}_4 \end{pmatrix}\boldsymbol{S}$. これも一意的ではなく, $\boldsymbol{D}_2$ と $\boldsymbol{D}_4$ は任意である.

最小二乗型一般化逆行列 (1)

- 最小ノルム型一般化逆行列では, $\boldsymbol{Ax} = \boldsymbol{b}$ の解の中から, ノルムが最小のものを求めた.
- 工学的な応用問題では, 誤差などのために, $\boldsymbol{Ax} = \boldsymbol{b}$ が厳密解を持たないことがあり得る. このような状況で, $\|\boldsymbol{Ax} - \boldsymbol{b}\|$ を最小にする近似解を求める問題を考える.

最小二乗型一般化逆行列 (2)

- 最小ノルム型一般化逆行列で $\boldsymbol{A}$ に左からユニタリ行列 (直交行列) を掛けたことと比較すると, $\boldsymbol{A}$ に右からユニタリ行列 (直交行列) を掛け, 先と類似した計算をおこなうことで, 近似解が構成できると考えられる. これに対応する一般化逆行列を最小二乗型一般化逆行列といい, $\boldsymbol{A}^{\wedge}$ であらわす.

最小二乗型一般化逆行列 (3)

- $\boldsymbol{A}$ を複素行列とする．実行列では, * を ^{T} に置き換えればよい．
- $\boldsymbol{A}$ を m 行 n 列とし, $\boldsymbol{AP}$ が QR 分解できるような置換行列 $\boldsymbol{P}$ を取る． $\boldsymbol{AP} = \boldsymbol{QR}$, $\boldsymbol{R} = \begin{pmatrix} \boldsymbol{R}_1 & \boldsymbol{R}_2 \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}$ とする． $\boldsymbol{R}_1$ は正則な上三角行列である．

最小二乗型一般化逆行列 (4)

- $\boldsymbol{Q}^*\boldsymbol{A}\boldsymbol{P} = \boldsymbol{R}$ であるが, この両辺に列基本変形を施して, $\boldsymbol{R}_1$ を単位行列に変形する. これに対応する基本行列の積を $\boldsymbol{V}$ とし, $\boldsymbol{P}\boldsymbol{V} = \boldsymbol{T}$ とおくと, $\boldsymbol{Q}^*\boldsymbol{A}\boldsymbol{T} = \begin{pmatrix} \boldsymbol{I}_r & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}$ である.
- $\boldsymbol{x} = \boldsymbol{T}\boldsymbol{z}$ とおき, $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ の両辺に $\boldsymbol{Q}^*$ を左から掛けると, $\boldsymbol{Q}^*\boldsymbol{A}\boldsymbol{T}\boldsymbol{z} = \boldsymbol{Q}^*\boldsymbol{b}$.

最小二乗型一般化逆行列 (5)

- $\boldsymbol{z}_1 = (z_1, \ldots, z_r)^T$, $\boldsymbol{z}_2 = (z_{r+1}, \ldots, z_n)^T$ とすると, 解を最小二乗近似するには, $\boldsymbol{z}_1 = (\boldsymbol{I}_r, \boldsymbol{0})\boldsymbol{Q}^*\boldsymbol{b}$ とすればよく, $\boldsymbol{z}_2$ は任意である. よって, 解の一般形は, $\boldsymbol{z} = \begin{pmatrix} \boldsymbol{I}_r & \boldsymbol{0} \\ \boldsymbol{D}_3 & \boldsymbol{D}_4 \end{pmatrix} \boldsymbol{Q}^*\boldsymbol{b}$.

最小二乗型一般化逆行列 (6)

- もとの座標系では, $\boldsymbol{x} = \boldsymbol{T}\begin{pmatrix} \boldsymbol{I}_r & \boldsymbol{0} \\ \boldsymbol{D}_3 & \boldsymbol{D}_4 \end{pmatrix}\boldsymbol{Q}^*\boldsymbol{b}$,
  よって $\boldsymbol{A}^{\wedge} = \boldsymbol{T}\begin{pmatrix} \boldsymbol{I}_r & \boldsymbol{0} \\ \boldsymbol{D}_3 & \boldsymbol{D}_4 \end{pmatrix}\boldsymbol{Q}^*$ であり, $\boldsymbol{D}_3$ と $\boldsymbol{D}_4$ は任意である.

Moore-Penrose の一般化逆行列 (1)

- Moore-Penrose の一般化逆行列 ($\boldsymbol{A}^+$ と書く) は, 最小ノルム型と最小二乗型の一般化逆行列の特徴を併せ持つ一般化逆行列である. その導出には, 一般化逆行列のフリーパラメータをすべて使う.
- $\boldsymbol{A}$ を複素行列とする. 実行列では, * を ^{T} に置き換えればよい.

Moore-Penrose の一般化逆行列 (2)

- $\boldsymbol{A}^*\boldsymbol{P}$ が QR 分解できるような置換行列 $\boldsymbol{P}$ を取り，対応する QR 分解を $\boldsymbol{A}^*\boldsymbol{P} = \boldsymbol{Q}_1\boldsymbol{R}$ とする．$\boldsymbol{R} = \begin{pmatrix} \boldsymbol{R}_1 & \boldsymbol{R}_2 \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}$ で，$\boldsymbol{R}_1$ は上三角な正方行列，$\boldsymbol{Q}_1$ はユニタリ行列である．$\boldsymbol{P}$ は置換行列だからユニタリ行列であることに注意する．

Moore-Penrose の一般化逆行列 (3)

- $\boldsymbol{P}^* \boldsymbol{A} \boldsymbol{Q}_1 = \boldsymbol{R}^*$ であるが, 更に $\boldsymbol{R}^*$ を QR 分解する. $\boldsymbol{R}$ の最初の r 個の列ベクトルは線形独立なので, $\boldsymbol{R}^*$ は置換なしで QR 分解でき, $\boldsymbol{R}^* = \boldsymbol{Q}_2 \begin{pmatrix} \boldsymbol{R}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$ という形になる (後半の零のみから成る列は QR 分解で形を変えないことに注意).

Moore-Penrose の一般化逆行列 (4)

- $\boldsymbol{Q}_3 = \boldsymbol{P}\boldsymbol{Q}_2$ とおくと, $\boldsymbol{Q}_3^* \boldsymbol{A} \boldsymbol{Q} = \begin{pmatrix} \boldsymbol{R}_r & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}$ である. $\boldsymbol{x} = \boldsymbol{Q}_1 \boldsymbol{z}$ とおくと, $\boldsymbol{Q}_3^* \boldsymbol{A} \boldsymbol{Q}_1 \boldsymbol{z} = \boldsymbol{Q}_3^* \boldsymbol{b}$ であり, したがって, $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ のノルムが最小の最小二乗解は $\boldsymbol{z} = \begin{pmatrix} \boldsymbol{R}_r^{-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix} \boldsymbol{Q}_3^* \boldsymbol{b}$ である.

Moore-Penrose の一般化逆行列 (5)

- もとの座標系では, $\boldsymbol{x} = \boldsymbol{Q}_1 \begin{pmatrix} \boldsymbol{R}_r^{-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix} \boldsymbol{Q}_3^* \boldsymbol{b}$,

  よって $\boldsymbol{A}^+ = \boldsymbol{Q}_1 \begin{pmatrix} \boldsymbol{R}_r^{-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix} \boldsymbol{Q}_3^*$ が $\boldsymbol{A}$ の Moore-Penrose の一般化逆行列である. これはフリーパラメータを持たない.

計算例

$\boldsymbol{A} = \left(\begin{array}{cc|ccc} 2 & 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \end{array}\right)$ とおき，$\boldsymbol{A}^+$ を求める．$\boldsymbol{Q}_3^* \boldsymbol{A} \boldsymbol{Q}_1 = \begin{pmatrix} \boldsymbol{R}_r & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}$ となる $\boldsymbol{Q}_3^*$ と $\boldsymbol{Q}_1$ を求める必要があるが，$\boldsymbol{A}$ ははじめから求める形になっているので，$\boldsymbol{Q}_3^* = \boldsymbol{I}$，$\boldsymbol{Q}_1 = \boldsymbol{I}$ としてよい． よって，$\boldsymbol{A}^+ = \left(\begin{array}{cc|c} \frac{1}{2} & -\frac{1}{6} & 0 \\ 0 & \frac{1}{3} & 0 \\ \hline 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right)$． $\boldsymbol{A}^+$ の左上のブロックが $\begin{pmatrix} 2 & 1 \\ 0 & 3 \end{pmatrix}$ の逆行列になっていることに注意.

特異値分解 (1)

- 固有値は, 行列に対応する線型写像の作用の「倍率」を測る尺度のひとつであるが, 正方行列に対してしか定義できない.
- 正方でない行列に対して固有値と似た性質を持つ量が定義できると便利.
- この役割を果たすのが特異値である.

特異値分解 (2)

- 特異値は, 正方とは限らない実行列あるいは複素行列に対して定義される (関数行列に拡張されることもある)
- 以下では, 議論の重複を避けるために, $\boldsymbol{A}$ を m 行 n 列の複素行列とする. 実行列については, $\quad^*$ を $\quad^T$ で読み換えればよい.

特異値分解 (3)

- A^*A の正の固有値の平方根を A の特異値という (零固有値を無視することに注意).
- A^*A は Hermite 行列であるから, ユニタリ行列によって対角化され, また A^*A の固有値はすべて非負なので, その特異値は曖昧さなく定まる.

特異値分解 (4)

- A^*A が r 個の正の固有値を持つものとし, これらを大きい順に並べたものを $\lambda_1, \ldots, \lambda_r$ とする. 特異値は $(\sigma_1, \ldots, \sigma_r) = (\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_r})$ である.

- $(\lambda_1, \ldots, \lambda_r)$ に対応する固有ベクトルを $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_r\}$ とする. また, $\{\boldsymbol{v}_{r+1}, \ldots, \boldsymbol{v}_n\}$ を, 零固有値に対応する固有ベクトルとする.

特異値分解 (5)

- $\boldsymbol{A}^*\boldsymbol{A}$ は Hermite 行列なので, $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_r\}$ を正規直交系にすることができる.
- 標準基底を取り, $\boldsymbol{v}_i$ を, それを標準基底により成分表示した列ベクトルと同一視する.
- $1 \le i \le r$ に対し, $\boldsymbol{u}_i = \dfrac{1}{\sigma_i}\boldsymbol{A}\boldsymbol{v}_i$ と定義する.

特異値分解 (6)

- $\boldsymbol{u}_i^* \boldsymbol{u}_j = \frac{1}{\sigma_i \sigma_j} \boldsymbol{v}_i^* \boldsymbol{A}^* \boldsymbol{A} \boldsymbol{v}_j = \delta_{ij}$ である (δ_{ij} は Kronecker のデルタ). よって, $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r\}$ は正規直交系である. これを含む正規直交基底 $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r, \boldsymbol{u}_{r+1}, \ldots, \boldsymbol{u}_m\}$ を作る.
- $1 \leq i, j \leq r$ に対し, $\boldsymbol{u}_i^* \boldsymbol{A} \boldsymbol{v}_j = \frac{1}{\sigma_i} \boldsymbol{v}_i^* \boldsymbol{A}^* \boldsymbol{A} \boldsymbol{v}_j = \frac{\lambda_j}{\sigma_i} \boldsymbol{v}_i^* \boldsymbol{v}_j = \sigma_i \delta_{ij}$ であることに注意

特異値分解 (7)

- $\ker \boldsymbol{A} = \{\boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}\}$ とおく．$\ker \boldsymbol{A} = \ker \boldsymbol{A}^*\boldsymbol{A}$ であることを示す．$\boldsymbol{A}\boldsymbol{x} = 0$ なら $\boldsymbol{A}^*\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ だから，$\ker \boldsymbol{A} \subset \ker \boldsymbol{A}^*\boldsymbol{A}$．逆に，$\boldsymbol{x} \in \ker \boldsymbol{A}^*\boldsymbol{A}$ であれば，$\boldsymbol{A}^*\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ より，$\boldsymbol{x}^*\boldsymbol{A}^*\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$．よって，$\boldsymbol{A}\boldsymbol{x} = 0$．したがって，$\boldsymbol{x} \in \ker \boldsymbol{A}$．ゆえに $\ker \boldsymbol{A}^*\boldsymbol{A} \subset \ker \boldsymbol{A}$．

特異値分解 (8)

- $\boldsymbol{v}_{r+1}, \ldots, \boldsymbol{v}_n$ は $\boldsymbol{A}^*\boldsymbol{A}$ の零固有値に対応する固有ベクトルだから, $\boldsymbol{A}^*\boldsymbol{A}\boldsymbol{v}_j = 0 \ (j \geq r+1)$. したがって, 先に述べた結果より, $\boldsymbol{A}\boldsymbol{v}_j = \boldsymbol{0}$ $(j \geq r+1)$ である.
- $r+1 \leq i \leq m$, $1 \leq j \leq r$ に対し, $\boldsymbol{u}_i^*\boldsymbol{A}\boldsymbol{v}_j = \boldsymbol{u}_i^*\sigma_j\boldsymbol{u}_j = 0$ である (正規直交基底だから).

特異値分解 (9)

- $\boldsymbol{V} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_r, \boldsymbol{v}_{r+1}, \ldots, \boldsymbol{v}_n)$, $\boldsymbol{U} = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r, \boldsymbol{u}_{r+1}, \ldots, \boldsymbol{u}_m)$ とおき, 以上の結果をまとめると,
$$\boldsymbol{U}^*\boldsymbol{A}\boldsymbol{V} = \begin{pmatrix} \mathrm{diag}(\sigma_1, \ldots, \sigma_r) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$$
となる.

特異値分解 (10)

- $\boldsymbol{A} = \boldsymbol{U} \begin{pmatrix} \mathrm{diag}(\sigma_1, \ldots, \sigma_r) & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix} \boldsymbol{V}^*$ を, $\boldsymbol{A}$ の特異値分解という.
- 特異値分解は, 最小二乗問題の解法, 伝達関数のノルムの定義などに使われる. Moore-Penrose の一般化逆行列もここから得られる.

計算例 (1)

1. $\boldsymbol{A} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 0 \end{pmatrix}$ とする．$\boldsymbol{A}^T\boldsymbol{A} = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}$, $\det \begin{pmatrix} \lambda - 2 & -2 \\ -2 & \lambda - 2 \end{pmatrix} = \lambda^2 - 4\lambda$ より, $\boldsymbol{A}^T\boldsymbol{A}$ の固有値は 0 と 4 である (実行列なので　* が ^{T} に変わっていることに注意).
2. 標準基底を固定し, ベクトルを数ベクトルと同一視する.
3. $\boldsymbol{A}^T\boldsymbol{A}$ の固有値 4 と 0 に対応する (正規化された) 固有ベクトルはそれぞれ $\boldsymbol{v}_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\boldsymbol{v}_2$ う $= \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ である．$\boldsymbol{A}$ の特異値は唯一で, その値は $\sigma_1 = \sqrt{4} = 2$ である.

計算例 (2)

4. $\boldsymbol{u}_1 = \frac{1}{\sigma_1}\boldsymbol{A}\boldsymbol{v}_1 = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\\0\end{pmatrix}$ である．$\boldsymbol{u}_2 = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\\0\end{pmatrix}$, $\boldsymbol{u}_3 = \begin{pmatrix}0\\0\\1\end{pmatrix}$ とおくと, $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$ は $\boldsymbol{u}_1$ を含む正規直交基底である．
5. $\boldsymbol{U} = (\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3)$, $\boldsymbol{V} = (\boldsymbol{v}_1, \boldsymbol{v}_2)$ とおく．
6. $\boldsymbol{U}^T\boldsymbol{A}\boldsymbol{V} = \begin{pmatrix}2 & 0\\0 & 0\\0 & 0\end{pmatrix}$ となる．よって, $\boldsymbol{A} = \boldsymbol{U}\begin{pmatrix}2 & 0\\0 & 0\\0 & 0\end{pmatrix}\boldsymbol{V}^T$ が $\boldsymbol{A}$ の特異値分解である．

参考文献

- 伊理, 韓, 線形代数, 教育出版, 1977
- 伊理, 線形代数汎論, 朝倉書店, 2009
- D. S. Bernstein, Matrix Mathematics, 2/e, Princeton University Press, 2009
- 室田, 杉原, 線形代数 I, 丸善, 2015
- 室田, 杉原, 線形代数 II, 丸善, 2015
- 太田, システム制御のための数学 (1) 線形代数編, コロナ社, 2000