

1 ベクトルと行列

1.1 行列

n, p を 1 以上の整数とする．数を n 行 p 列に並べたものを行列という． $\mathbf{A}$ が $n \times p$ の行列であるとするとき，

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{bmatrix}$$

である．

1.2 ベクトル

行列のうち， $p = 1$ であるものを列ベクトル， $n = 1$ であるものを行ベクトルという． $\mathbf{a}$ が n 次元の列ベクトルであるとするとき，

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

である．また， $\mathbf{b}'$ が p 次元の行ベクトルであるとするとき，

$$\mathbf{b}' = \begin{bmatrix} b_1 & b_2 & \cdots & b_p \end{bmatrix}$$

である．

1.3 転置行列

行列の行と列を入れ替えたものを転置行列という． $\mathbf{A}$ の転置行列は $\mathbf{A}'$ ($p \times n$ 行列) と表され，

$$\mathbf{A}' = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & \cdots & \vdots \\ a_{1p} & a_{2p} & \cdots & a_{np} \end{bmatrix}$$

である．転置行列について以下の関係が成り立つ．

$$(\mathbf{A}')' = \mathbf{A}$$

1.4 部分行列と分割行列

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \cdots & \mathbf{A}_{1p} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \cdots & \mathbf{A}_{2p} \\ \vdots & \vdots & \cdots & \vdots \\ \mathbf{A}_{n1} & \mathbf{A}_{n2} & \cdots & \mathbf{A}_{np} \end{bmatrix}$$

ただし, A_{ij} は, 同じ i に対して行数が等しく, 同じ j に対して列数が等しいとする. このとき, A_i を部分行列, A を分割行列という.

1.5 行列のベクトル表示

p 次元ベクトル a'_i ($i = 1, 2, \dots, n$) が $n \times p$ 行列 A の各行を要素に持つベクトルであるとする,

$$A = \begin{bmatrix} a'_1 \\ a'_2 \\ \vdots \\ a'_n \end{bmatrix}$$

と書ける.

また, n 次元ベクトル a_j ($j = 1, 2, \dots, p$) が $n \times p$ 行列 A の各列を要素に持つベクトルであるとする,

$$A = \begin{bmatrix} a_1 & a_2 & \cdots & a_p \end{bmatrix}$$

と書ける.

1.6 零行列

すべての要素が 0 である行列を零行列といい, O と表す.

1.7 正方行列

行数と列数が等しい ($n = p$) 行列を (n 次) 正方行列という.

1.8 対称行列

n 次正方行列 A において, $a_{ij} = a_{ji}$ つまり $A = A'$ であるとき, A は対称行列であるという.

1.9 対角行列

n 次対称行列 A において, $i \neq j$ のとき $a_{ij} = 0$ (非対角要素がすべて 0) である行列を対角行列という.

1.10 単位行列

n 次対角行列 A のうち, すべての対角要素 a_{ii} が 1 である行列を単位行列といい, I または I_n で表す.

1.11 三角行列

n 次正方行列 A において, $i > j$ のとき $a_{ij} = 0$ である行列を上三角行列, $i < j$ のとき $a_{ij} = 0$ となる行列を下三角行列という.

1.12 2次形式と正定置行列

n 次元ベクトル x と対称行列 A があるとき, $x'Ax$ を x に関する 2 次形式という.

0 でない任意のベクトル x に対して 2 次形式が正 $x'Ax > 0$ となるとき, A は正定置行列であるといい, $A > 0$ と表す.

また, 0 でない任意のベクトル x に対して 2 次形式が非負 $x'Ax \geq 0$ となるとき, A は半正定置行列であるといい, $A \geq 0$ と表す.

2 行列計算

2.1 行列の和と差

行列の和と差は, サイズが等しい行列において, それぞれの要素の和 (差) で定義される.

$$A \pm B = \begin{bmatrix} a_{11} \pm b_{11} & a_{12} \pm b_{12} & \cdots & a_{1p} \pm b_{1n} \\ a_{21} \pm b_{21} & a_{22} \pm b_{22} & \cdots & a_{2p} \pm b_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} \pm b_{n1} & a_{n2} \pm b_{n2} & \cdots & a_{np} \pm b_{np} \end{bmatrix}$$

行列の和 (差) について以下の関係が成り立つ.

$$A + B = B + A$$

$$(A \pm B)' = A' \pm B'$$

2.2 行列の定数倍

c を定数 (スカラー) とすると,

$$cA = c \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{bmatrix} = \begin{bmatrix} c a_{11} & c a_{12} & \cdots & c a_{1p} \\ c a_{21} & c a_{22} & \cdots & c a_{2p} \\ \vdots & \vdots & \cdots & \vdots \\ c a_{n1} & c a_{n2} & \cdots & c a_{np} \end{bmatrix}$$

である.

2.3 ベクトルの内積

2 つのベクトル a, b が同じ次元を持つとき, この 2 つのベクトルの内積は,

$$a'b = \sum_i a_i b_i$$

で定義される.

2.4 行列の積

行列

$$\mathbf{A} = \begin{bmatrix} a'_1 \\ a'_2 \\ \vdots \\ a'_n \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} b_1 & b_2 & \cdots & b_p \end{bmatrix}$$

の積 $\mathbf{AB}$ は, $\mathbf{A}$ の列数と $\mathbf{B}$ の行数が等しいときに定義され,

$$\begin{aligned} \mathbf{AB} &= \begin{bmatrix} \mathbf{A}b_1 & \mathbf{A}b_2 & \cdots & \mathbf{A}b_p \end{bmatrix} \\ &= \begin{bmatrix} a'_1\mathbf{B} \\ a'_2\mathbf{B} \\ \vdots \\ a'_n\mathbf{B} \end{bmatrix} \\ &= \begin{bmatrix} a'_1b_1 & a'_1b_2 & \cdots & a'_1b_p \\ a'_2b_1 & a'_2b_2 & \cdots & a'_2b_p \\ \vdots & \vdots & \cdots & \vdots \\ a'_nb_1 & a'_nb_2 & \cdots & a'_nb_p \end{bmatrix} \end{aligned}$$

である.

同様に,

$$\mathbf{A} = \begin{bmatrix} a_1 & a_2 & \cdots & a_p \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} b'_1 \\ b'_2 \\ \vdots \\ b'_n \end{bmatrix}$$

とすると,

$$\mathbf{AB} = a_1b'_1 + a_2b'_2 + \cdots + a_pb'_p = \sum_{i=1}^p a_ib'_i$$

である.

行列の積に関して以下の関係が成り立つ.

$$(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$$

$$\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$$

$$\mathbf{AI} = \mathbf{IA} = \mathbf{A}$$

$$(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'$$

すべての要素が 1 である n 次正方行列を $\mathbf{J}_n$ とし, $\mathbf{A} = (a\mathbf{I}_n + b\mathbf{J}_n)$, $\mathbf{B} = (c\mathbf{I}_n + d\mathbf{J}_n)$ とするとき,

$$\mathbf{AB} = \mathbf{BA} = ac\mathbf{I}_n + (ad + bc + nbd)\mathbf{J}_n$$

2.5 分割行列の積

2つの分割行列 A, B において, 部分行列の積 $A_{ij}B_{jk}$ がすべての i, j, k について定義されるとき, 分割行列の積 AB を定義することができる.

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

とすると,

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

である.

2.6 クロネッカー積

$m \times n$ 行列 A と $p \times q$ 行列 B のクロネッカー積は $A \otimes B$ と表され, 以下のように定義される.

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{bmatrix}$$

これは $mp \times nq$ 行列である.

クロネッカー積に関して以下の性質が成り立つ.

$$c(A \otimes B) = (cA) \otimes B = A \otimes (cB) = cA \otimes B$$

$$A \otimes (B \otimes C) = (A \otimes B) \otimes C = A \otimes B \otimes C$$

$$(A \otimes B)' = A' \otimes B'$$

$$(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$$

$$(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$$

$$(A + B) \otimes C = A \otimes C + B \otimes C$$

$$A \otimes (B + C) = A \otimes B + A \otimes C$$

3 トレース

正方行列 A の対角要素の和とトレースといい,

$$\text{tr}A = \sum_i a_{ii}$$

で表す.

トレースに関して以下の性質が成り立つ.

c をスカラー, A, B を同じサイズの正方行列, C, D' を同じサイズの行列 ($n \times p$), x_i を A の行 (列)

と次元が同じベクトルであるとするとき，

$$\operatorname{tr} c = c$$

$$\operatorname{tr} a\mathbf{A} = a \operatorname{tr} \mathbf{A}$$

$$\operatorname{tr}(\mathbf{A} \pm \mathbf{B}) = \operatorname{tr} \mathbf{A} \pm \operatorname{tr} \mathbf{B}$$

$$\operatorname{tr}(\mathbf{CD}) = \operatorname{tr}(\mathbf{DC}) = \sum_{i,j} c_{ij} d_{ji}$$

$$\operatorname{tr}(\mathbf{AB}) = \operatorname{tr}(\mathbf{BA})$$

$$\operatorname{tr}(\mathbf{AA}') = \operatorname{tr}(\mathbf{A}'\mathbf{A}) = \sum_{i,j} a_{ij}^2$$

$$\sum_i \mathbf{x}'_i \mathbf{A} \mathbf{x}_i = \operatorname{tr} \mathbf{AT}, \quad \text{ただし } \mathbf{T} = \sum_i \mathbf{x}_i \mathbf{x}'_i$$

4 逆行列

n 次正方行列 $\mathbf{A}$ に関して，

$$\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}_n$$

となる行列 $\mathbf{A}^{-1}$ が存在するとき， $\mathbf{A}^{-1}$ を $\mathbf{A}$ の逆行列という．

$n = 2$ のとき，行列 $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ の逆行列は，

$$\mathbf{A}^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}, \quad \text{ただし } a_{11}a_{22} - a_{12}a_{21} \neq 0$$

である．

逆行列に関して以下の性質が成り立つ．ただし， $\mathbf{A}$, $\mathbf{B}$ は逆行列を持つ n 次正方行列， $\mathbf{C}$, $\mathbf{D}$ は $n \times p$ 行列， $\mathbf{E}$ は逆行列を持つ p 次正方行列とする．

$$(\mathbf{A}^{-1})^{-1} = \mathbf{A}$$

$$(\mathbf{A}')^{-1} = (\mathbf{A}^{-1})'$$

$$(c\mathbf{A})^{-1} = c^{-1}\mathbf{A}^{-1}$$

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$$

$$(\mathbf{A} + \mathbf{CED})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{C}(\mathbf{E}^{-1} + \mathbf{DA}^{-1}\mathbf{C})^{-1}\mathbf{DA}^{-1}$$

$$(\mathbf{A} + \mathbf{B})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}(\mathbf{B}^{-1} + \mathbf{A}^{-1})^{-1}\mathbf{A}^{-1}$$

$$= \mathbf{B}^{-1} - \mathbf{B}^{-1}(\mathbf{A}^{-1} + \mathbf{B}^{-1})^{-1}\mathbf{B}^{-1}$$

$$= \mathbf{A}^{-1}(\mathbf{A}^{-1} + \mathbf{B}^{-1})^{-1}\mathbf{B}^{-1}$$

$$= \mathbf{B}^{-1}(\mathbf{A}^{-1} + \mathbf{B}^{-1})^{-1}\mathbf{A}^{-1}$$

$$(\mathbf{y} - \mathbf{c})'\mathbf{A}^{-1}(\mathbf{y} - \mathbf{c}) + (\mathbf{y} - \mathbf{d})'\mathbf{B}^{-1}(\mathbf{y} - \mathbf{d})$$

$$= (\mathbf{y} - [\mathbf{A}(\mathbf{A} + \mathbf{B})^{-1}\mathbf{B}][\mathbf{A}^{-1}\mathbf{c} + \mathbf{B}^{-1}\mathbf{d}])' [\mathbf{A}(\mathbf{A} + \mathbf{B})^{-1}\mathbf{B}]^{-1} (\mathbf{y} - [\mathbf{A}(\mathbf{A} + \mathbf{B})^{-1}\mathbf{B}][\mathbf{A}^{-1}\mathbf{c} + \mathbf{B}^{-1}\mathbf{d}]) \\ + (\mathbf{c} - \mathbf{d})'(\mathbf{A} + \mathbf{B})^{-1}(\mathbf{c} - \mathbf{d})$$

$$= (\mathbf{y} - [\mathbf{B}(\mathbf{A} + \mathbf{B})^{-1}\mathbf{c} + \mathbf{A}(\mathbf{A} + \mathbf{B})^{-1}\mathbf{d}])' [\mathbf{A}(\mathbf{A} + \mathbf{B})^{-1}\mathbf{B}]^{-1} (\mathbf{y} - [\mathbf{B}(\mathbf{A} + \mathbf{B})^{-1}\mathbf{c} + \mathbf{A}(\mathbf{A} + \mathbf{B})^{-1}\mathbf{d}])$$

$$\begin{aligned}
& + (c-d)'(A+B)^{-1}(c-d) \\
= & (y - [B(A+B)^{-1}c + (I - B(A+B)^{-1})d])' [A(A+B)^{-1}B]^{-1} \\
& \times (y - [B(A+B)^{-1}c + (I - B(A+B)^{-1})d]) + (c-d)'(A+B)^{-1}(c-d) \\
\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}^{-1} = & \begin{bmatrix} A^{11} & A^{12} \\ A^{21} & A^{22} \end{bmatrix} \\
A^{11} = & (A_{11} - A_{12}A_{22}^{-1}A_{21})^{-1} \\
A^{22} = & (A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1} \\
A^{12} = & -A_{11}^{-1}A_{12}A^{22} = -A^{11}A_{12}A_{22}^{-1} \\
A^{21} = & -A_{22}^{-1}A_{21}A^{11} = -A^{22}A_{21}A_{11}^{-1} \\
a \neq 0, a+nb \neq 0 \text{ であれば,} \\
(aI_n + bJ_n)^{-1} = & \frac{1}{a} \left(I_n - \frac{b}{a+nb} J_n \right) \\
p \neq 1, -(p-1)^{-1} \text{ であれば,} \\
((1-p)I_n + pJ_n)^{-1} = & \frac{1}{1-p} \left(I_n - \frac{p}{a+(n-1)p} J_n \right)
\end{aligned}$$

5 行列式

n 次の正方行列 A の行ベクトルによって作られる超平面体の超体積を行列 A の行列式とよび, $|A|$ で表す.
 $|A| \neq 0$ であるとき, A は正則であるといい, 逆行列 A^{-1} が存在する (必要十分条件).

$n = 2$ のとき, 行列 $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ の行列式は,

$$|A| = a_{11}a_{22} - a_{12}a_{21}$$

である.

また $n = 3$ のとき, 行列 $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$ の行列式は,

$$|A| = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}$$

である.

行列式に関して以下の性質が成り立つ. ただし, c はスカラー, A, B は逆行列を持つ n 次正方行列, C, D は $n \times p$ 行列とする.

$$\begin{aligned}
|A'| &= |A| \\
|cA| &= c^n |A| \\
|A^{-1}| &= |A|^{-1} \\
|AB| &= |A||B| = |B||A| = |BA| \\
|(A+B)^{-1}| &= |A^{-1}(A^{-1} + B^{-1})^{-1}B^{-1}| = |A|^{-1}|A^{-1} + B^{-1}|^{-1}|B|^{-1} \\
|(A^{-1} + B^{-1})^{-1}| &= |A(A+B)^{-1}B| = |A||A+B|^{-1}|B| \\
|I_n + CD| &= |I_p + DC|
\end{aligned}$$

$$\begin{aligned}
|\mathbf{A} + \mathbf{CD}| &= |\mathbf{A}||\mathbf{I}_n + \mathbf{A}^{-1}\mathbf{CD}| = |\mathbf{A}||\mathbf{I}_p + \mathbf{DA}^{-1}\mathbf{C}| \\
\begin{vmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{vmatrix} &= |\mathbf{A}_{11}||\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}| = |\mathbf{A}_{22}||\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}| \\
\begin{vmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{O} & \mathbf{B} \end{vmatrix} &= |\mathbf{A}||\mathbf{B}|
\end{aligned}$$

$\mathbf{M}$ が m 次正定置行列であるとき, $|\mathbf{A} \otimes \mathbf{M}| = |\mathbf{A}|^m |\mathbf{M}|^n$

$\mathbf{A}$ が三角行列または対角行列のとき, $|\mathbf{A}| = \sum_i a_{ii}$

$\mathbf{A}$ が正定置行列ならば $|\mathbf{A}| > 0$

$\mathbf{A}$ が半正定置行列ならば $|\mathbf{A}| \geq 0$

6 線形独立と階数

p 個の n 次元ベクトル $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_p$ に対して,

$$x_1 \mathbf{a}_1 + x_2 \mathbf{a}_2 + \dots + x_p \mathbf{a}_p = \mathbf{0}$$

を満たす p 次元ベクトル $\mathbf{x} = (x_1, x_2, \dots, x_p)'$ が $\mathbf{x} = \mathbf{0}$ 以外に存在しないとき, $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_p$ は線形独立 (1 次独立) であるという.

$n \times p$ の行列 $\mathbf{A}$ に関して, 線形独立な行 (列) ベクトルの数を, 行列 $\mathbf{A}$ の階数 (ランク) といい $r(\mathbf{A})$ または $\text{rank } \mathbf{A}$ などと表す.

階数に関して以下の性質が成り立つ.

$$0 \leq r(\mathbf{A}) \leq \min(n, p)$$

$$r(\mathbf{A}) = r(\mathbf{A}') = r(\mathbf{A}^{-1})$$

$$r(\mathbf{A} + \mathbf{B}) \leq r(\mathbf{A}) + r(\mathbf{B})$$

$$r(\mathbf{AB}) \leq \min(r(\mathbf{A}), r(\mathbf{B}))$$

$$r(\mathbf{AA}') = r(\mathbf{A}'\mathbf{A}) = r(\mathbf{A})$$

$$n \text{ 次の正則行列を } \mathbf{B}, p \text{ 次の正則行列を } \mathbf{C} \text{ とするとき, } r(\mathbf{BAC}) = r(\mathbf{A})$$

$$\mathbf{A} \text{ が } n \text{ 次の正則行列であるとき, } r(\mathbf{A}) = n$$

7 ベクトル微分

関数 f が n 次元ベクトル x の関数 $f = f(x)$ であるとき, f の x による微分は,

$$\frac{\partial f}{\partial x} = \left[\frac{\partial f}{\partial x_i} \right] = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix}$$

と定義される. 同様にして, f の x' による微分は,

$$\frac{\partial f}{\partial x'} = \left[\frac{\partial f}{\partial x_i} \right] = \left[\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n} \right]$$

となる.

2 階微分は次のように定義される.

$$\frac{\partial^2 f}{\partial x \partial x'} = \left[\frac{\partial^2 f}{\partial x_i \partial x'_j} \right] = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n} \end{bmatrix}$$

ベクトル微分に関して以下の性質が成り立つ.

$$\begin{aligned} \frac{\partial a'x}{\partial x} &= \frac{\partial x'a}{\partial x} = a \\ \frac{\partial a'Ax}{\partial x} &= \frac{\partial x'A'a}{\partial x} = A'a \\ \frac{\partial x'x}{\partial x} &= 2x \\ \frac{\partial x'Ax}{\partial x} &= (A + A')x \\ \frac{\partial x'Ax}{\partial x \partial x} &= (A + A') \\ \frac{\partial (x-a)'A(x-a)}{\partial x} &= (A + A')(x-a) \\ \frac{\partial^2 (x-a)'A(x-a)}{\partial x \partial x'} &= (A + A') \\ \frac{\partial (a-Ax)'(a-Ax)}{\partial x} &= 2(A'Ax - A'a) \\ \frac{\partial^2 (a-Ax)'(a-Ax)}{\partial x \partial x'} &= 2A'A \end{aligned}$$

$A = A'$ とすると,

$$\begin{aligned} \frac{\partial x'Ax}{\partial x} &= 2Ax \\ \frac{\partial x'Ax}{\partial x \partial x'} &= 2A \\ \frac{\partial (x-a)'A(x-a)}{\partial x} &= 2A(x-a) \end{aligned}$$

$$\begin{aligned}\frac{\partial^2(\mathbf{x} - \mathbf{a})' \mathbf{A}(\mathbf{x} - \mathbf{a})}{\partial \mathbf{x} \partial \mathbf{x}'} &= 2\mathbf{A} \\ \frac{\partial(\mathbf{a} - \mathbf{A}\mathbf{x})'(\mathbf{a} - \mathbf{A}\mathbf{x})}{\partial \mathbf{x}} &= 2(\mathbf{A}^2\mathbf{x} - \mathbf{A}\mathbf{a}) \\ \frac{\partial^2(\mathbf{a} - \mathbf{A}\mathbf{x})'(\mathbf{a} - \mathbf{A}\mathbf{x})}{\partial \mathbf{x} \partial \mathbf{x}'} &= 2\mathbf{A}^2\end{aligned}$$

8 重みなし最小2乗法 (OLS)

モデル

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{e}$$

誤差ベクトル $\mathbf{e}$ の長さの2乗を $f(\beta)$ とすると,

$$\begin{aligned}f(\beta) &= \|\mathbf{e}\|^2 \\ &= \mathbf{e}'\mathbf{e} \\ &= (\mathbf{y} - \mathbf{X}\beta)'(\mathbf{y} - \mathbf{X}\beta) \\ &= \mathbf{y}'\mathbf{y} - \mathbf{y}'\mathbf{X}\beta - \beta'\mathbf{X}'\mathbf{y} + \beta'\mathbf{X}'\mathbf{X}\beta \\ &= \mathbf{y}'\mathbf{y} - 2\mathbf{y}'\mathbf{X}\beta + \beta'\mathbf{X}'\mathbf{X}\beta\end{aligned}$$

$f(\beta)$ を β で偏微分して $= 0$ を解くと,

$$\frac{\partial f}{\partial \beta} = 0 - 2\mathbf{X}'\mathbf{y} + 2\mathbf{X}'\mathbf{X}\beta$$

より,

$$\begin{aligned}-2\mathbf{X}'\mathbf{y} + 2\mathbf{X}'\mathbf{X}\beta &= 0 \\ \mathbf{X}'\mathbf{X}\beta &= \mathbf{X}'\mathbf{y} \\ \beta &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}\end{aligned}$$

9 重み付き最小2乗法 (WLS) ・ 一般化最小2乗法 (GLS)

モデル

$$\begin{aligned}\mathbf{y} &= \mathbf{X}\beta + \mathbf{e} \\ \mathbf{e} \text{ の共分散行列 } \text{Cov}(\mathbf{e}) &= \mathbf{G}\end{aligned}$$

$\mathbf{G}$ は対称行列であるから, $\mathbf{G}$ を $\mathbf{G} = (\mathbf{T}^{1/2})'\mathbf{T}^{1/2}$ と分解し, $\mathbf{T}^{-1/2}$ で重み付けたモデル,

$$\mathbf{T}^{-1/2}\mathbf{y} = \mathbf{T}^{-1/2}\mathbf{X}\beta + \mathbf{T}^{-1/2}\mathbf{e}$$

を考える. もし, $\mathbf{G}$ が対角行列で, 対角要素が $\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2$ であるとしたら, 重み付け行列 $\mathbf{T}^{-1/2}$ も対角行列で, 対角要素は $1/\sigma_1, 1/\sigma_2, \dots, 1/\sigma_n$ となる.

G を対角行列とした場合を重み付き最小 2 乗法, G を対称行列とした場合を一般化最小 2 乗法という. 重み付けた誤差ベクトル $T^{-1/2}e$ の長さの 2 乗を $f(\beta)$ とすると,

$$\begin{aligned}
 f(\beta) &= \|T^{-1/2}e\|^2 \\
 &= e'(T^{-1/2})'T^{-1/2}e \\
 &= e'G^{-1}e \\
 &= (y - X\beta)'G^{-1}(y - X\beta) \\
 &= y'G^{-1}y - y'G^{-1}X\beta - \beta'X'G^{-1}y + \beta'X'G^{-1}X\beta \\
 &= y'G^{-1}y - 2y'G^{-1}X\beta + \beta'X'G^{-1}X\beta
 \end{aligned}$$

$f(\beta)$ を β で偏微分して $= 0$ を解くと,

$$\frac{\partial f}{\partial \beta} = 0 - 2X'G^{-1}y + 2X'G^{-1}X\beta$$

より,

$$\begin{aligned}
 -2X'G^{-1}y + 2X'G^{-1}X\beta &= 0 \\
 X'G^{-1}X\beta &= X'G^{-1}y \\
 \beta &= (X'G^{-1}X)^{-1}X'G^{-1}y
 \end{aligned}$$

y の予測ベクトル $\hat{y}$ は,

$$\begin{aligned}
 \hat{y} &= X\beta \\
 &= X(X'G^{-1}X)^{-1}X'G^{-1}y
 \end{aligned}$$

となる. $P = X(X'G^{-1}X)^{-1}X'G^{-1}$ は, 擬ノルム G^{-1} ($\|x\|_{G^{-1}}^2 = x'G^{-1}x$) に基づく $S(X)$ への直交射影行列である.

β の一般化最小 2 乗推定量は最良線形不偏推定量 (best linear unbiased estimator, BLUE) であり, 線形不偏推定量の中で分散が最小である. 重み付けのない最小 2 乗推定量は, 不偏推定量ではあるが BLUE ではない. よって, 一般化最小 2 乗推定量のほうが精度が良い.

不安定な標本ほど重みを軽くして (誤差が大きくなったってあまり構わないとして) 推定したほうが, 重みをつけないで (すべての観測値を同等に扱って) 推定するよりも推定量が安定する. 不安定な標本はどうせ誤差が大きいんだから, 予測が他の安定した標本ほど正確でなくてもよい. 全体の誤差を最小化するより, 安定した標本ほど誤差が小さくなるような予測をしたほうがよい.

例えば, 集団の平均値をデータとする場合 (クラス平均, 学校平均をデータとする研究など), 集団の大きさの違いなどから各データの誤差分散が異なってくるので, 一般化最小 2 乗法を用いた方がよい.

10 因子分析

10.1 モデル

平均偏差データ行列を $\mathbf{X}$, 因子スコア行列を $\mathbf{F}$, 因子負荷行列を $\mathbf{\Lambda}$, 誤差行列を $\mathbf{E}$, また , 観測変数の共分散行列 $\frac{1}{n}\mathbf{X}'\mathbf{X} = \mathbf{\Sigma}$, 因子スコアの共分散行列を $\frac{1}{n}\mathbf{F}'\mathbf{F} = \mathbf{\Phi}$, 誤差スコアの共分散行列を $\frac{1}{n}\mathbf{E}'\mathbf{E} = \mathbf{\Psi}$ とすると , 因子分析モデルは , 次式で表される .

$$\mathbf{X} = \mathbf{F}\mathbf{\Lambda} + \mathbf{E}$$

$\mathbf{\Sigma}$ を $\mathbf{\Phi}, \mathbf{\Psi}$ を用いて表す . 因子スコアと誤差スコアは直交する ($\mathbf{F}'\mathbf{E} = \mathbf{E}'\mathbf{F} = \mathbf{0}$) とすると ,

$$\mathbf{\Sigma} = \mathbf{\Lambda}'\mathbf{\Phi}\mathbf{\Lambda} + \mathbf{\Psi} \quad (\text{斜交モデル})$$

$$\mathbf{\Sigma} = \mathbf{\Lambda}'\mathbf{\Lambda} + \mathbf{\Psi} \quad (\text{直交モデル, } \mathbf{\Phi} = \mathbf{I}, \text{ 因子分散}=1, \text{ 因子間相関}=0)$$

10.2 重みなし最小 2 乗法 (OLS)

観測変数の標本共分散行列を $\mathbf{S}$ とする . モデルをよくあてはめるためには , $\mathbf{\Sigma}$ が $\mathbf{S}$ をよく再現すればよい . 即ち , $\mathbf{S}$ と $\mathbf{\Sigma}$ のズレが最小になればよい .

$$\text{tr}(\mathbf{A}'\mathbf{A}) = \sum \sum a_{ij}^2$$

であるので ,

$$L = \text{tr}[(\mathbf{S} - \mathbf{\Sigma})^2] = \text{tr}[(\mathbf{S} - \mathbf{\Lambda}'\mathbf{\Phi}\mathbf{\Lambda} - \mathbf{\Psi})^2]$$

を最小化すればよい . それには , L を $\mathbf{\Lambda}, \mathbf{\Psi}$ で偏微分して $= 0$ とした式を解けばよい .

10.3 重み付き最小 2 乗法 (WLS) ・ 一般化最小 2 乗法 (GLS)

$\mathbf{S}$ の下三角部分を要素に持つベクトルを $\mathbf{s}$, $\mathbf{\Sigma}$ の下三角部分を要素に持つベクトルを $\mathbf{\sigma}$ とする .

$$\begin{aligned} \mathbf{s} &= (s_{11}, s_{21}, \dots, s_{p1}, s_{22}, s_{32}, \dots, s_{p2}, \dots, s_{pp})' \\ \mathbf{\sigma} &= (\sigma_{11}, \sigma_{21}, \dots, \sigma_{p1}, \sigma_{22}, \sigma_{32}, \dots, \sigma_{p2}, \dots, \sigma_{pp})' \end{aligned}$$

$\mathbf{s} = \mathbf{\sigma} + \mathbf{e}$ として , $\mathbf{e}$ の共分散行列を $\mathbf{G} = (\mathbf{T}^{1/2})'\mathbf{T}^{1/2}$ とすると , 重み付き誤差 $\mathbf{T}^{-1/2}\mathbf{e}$ の長さの 2 乗 L を最小化する最小 2 乗法の適用が考えられる . ある条件の下では ,

$$L = \|\mathbf{T}^{-1/2}\mathbf{e}\|^2 = \mathbf{e}'(\mathbf{T}^{-1/2})'\mathbf{T}^{-1/2}\mathbf{e} = (\mathbf{s} - \mathbf{\sigma})'\mathbf{G}^{-1}(\mathbf{s} - \mathbf{\sigma}) = \text{tr}[(\mathbf{S} - \mathbf{\Sigma})\mathbf{\Sigma}^{-1}]^2]$$

を最小化すればよい . しかし , このままの式で最小化するのは困難なので , 実際には , 最後の $\mathbf{\Sigma}$ を近似値の $\mathbf{S}$ で置き換えて ,

$$L^* = \text{tr}[(\mathbf{S} - \mathbf{\Sigma})\mathbf{S}^{-1}]^2 = \text{tr}[(\mathbf{S} - \mathbf{\Lambda}'\mathbf{\Phi}\mathbf{\Lambda} - \mathbf{\Psi})\mathbf{S}^{-1}]^2]$$

を最小化する .

標本共分散行列とモデルから構成される共分散構造行列とのズレを全体として最小にするよりも , 不安定な共分散のところはズレが少々あまくてもよい (どうせ不安定なんだから , 他よりも再現性が低くてもいい) ということにしておいたほうが , 推定量としての性質が安定する .

11 母共分散行列が Σ となる標本の生成

任意の値を要素とする $p \times p$ の母共分散行列を Σ , 標準正規乱数を要素とする $n \times p$ の行列を F とする . Σ を ,

$$\Sigma = A'A$$

とコレスキー分解して (A は上三角行列のコレスキー因子)

$$Z = FA$$

とする . $Z'Z/(n-1)$ の期待値を考えると ,

$$\begin{aligned} E[Z'Z/(n-1)] &= E[(FA)'FA/(n-1)] \\ &= E[A'F'FA/(n-1)] \\ &= A'E[F'F/(n-1)]A \end{aligned}$$

となるが , F の要素は標準正規乱数 (平均 0 , 分散 1) なので ,

$$E[F'F/(n-1)] = I$$

となり ,

$$\begin{aligned} E[Z'Z/(n-1)] &= A'E[F'F/(n-1)]A \\ &= A'IA \\ &= A'A \\ &= \Sigma \end{aligned}$$

となる . よって , $Z'Z/(n-1)$ の期待値は Σ となり , Z は Σ を母共分散行列とする標本データとなる .

R プログラム

```
#-----  
# 母平均を M , 母共分散行列を S とする標本の生成  
  
covmatrix.b <- function(n, p, M=rep(c(0),p), S=diag(n)){  
  F <- matrix(rnorm(n*p),n,p)  
  A <- chol(S)  
  C <- matrix(c(1),n,1)  
  M <- matrix(M,1,p)  
  Z <- F %*% A + C %*% M  
  return(Z)  
}  
#-----  
  
# オブザベーション数  
n <- 100
```

```

# 変数の数
p <- 5

# 母平均ベクトル
M <- c(1, 2, 3, 4, 5)

# 母標準偏差ベクトル
s <- c(6, 7, 8, 9, 10)

# 母相関係数行列
R <- matrix(c(1, 0, 0, 0, 0,
              0, 1, 0, 0, 0,
              0, 0, 1, 0, 0,
              0, 0, 0, 1, 0,
              0, 0, 0, 0, 1),
            p,p)

# 母共分散行列の計算
D <- diag(s)
S <- D %*% R %*% D

# オブザベーション数, 変数の数, 母平均ベクトル, 母共分散行列を指定して, 標本を作成
Z <- covmatrix.b(n, p, M, S)
Z

round(apply(Z,2,mean), 3)
round(sd(Z)*sqrt((n-1)/n), 3)
round(cor(Z),3)

write.table(Z, "d:\\p_val_data_b.csv", row.names=FALSE, sep=",")

```

12 標本共分散行列が S となる標本の生成

任意の値を要素とする $n \times p$ のデータ行列を Y とする．ただし, Y の各列の平均は 0 となるように, 各列を平均偏差化しておく． Y の標本共分散行列を $Y'Y/n = V$ とする．また, 目的とする値を要素とする $n \times p$ の標本共分散行列を S とする． V と S を,

$$\begin{aligned} V &= U'U \\ S &= A'A \end{aligned}$$

とコレスキー分解して

$$Z = YU^{-1}A$$

とする．データ行列 Z の標本共分散行列を考えると，

$$\begin{aligned}
 Z'Z/n &= (YU^{-1}A)'YU^{-1}A/n \\
 &= A'(U^{-1})'Y'YU^{-1}A/n \\
 &= A'(U^{-1})'(Y'Y/n)U^{-1}A \\
 &= A'(U^{-1})'VU^{-1}A \\
 &= A'(U^{-1})'U'UU^{-1}A \\
 &= A'(UU^{-1})'(UU^{-1})A \\
 &= A'I'A \\
 &= A'A \\
 &= S
 \end{aligned}$$

となる．よって，データ行列 Z は， S を標本共分散行列とする標本データである．

R プログラム

```

#-----
# 標本平均を M，標本共分散行列を S とする標本の生成

covmatrix.h <- function(n, p, M=rep(c(0),p), S=diag(n)){
  F <- matrix(rnorm(n*p),n,p)
  C <- matrix(c(1),n,1)
  MF <- apply(F, 2, mean)
  Y <- F - C %*% MF
  V <- 1/n * t(Y) %*% Y
  U <- chol(V)
  A <- chol(S)
  C <- matrix(c(1),n,1)
  M <- matrix(M,1,p)
  Z <- Y %*% solve(U) %*% A + C %*% M
  return(Z)
}
#-----

# オブザベーション数
n <- 50

# 変数の数
p <- 4

# 標本平均ベクトル
M <- c(0, 0, 0, 0)

```

```

# 標本標準偏差ベクトル
s <- c(1, 1, 1, 1)

# 標本相関係数行列
R <- matrix(c(1, 0.8, 0, 0,
              0.8, 1, 0, 0,
              0, 0, 1, -0.1,
              0, 0, -0.1, 1),
            p,p)

# 標本共分散行列の計算
D <- diag(s)
S <- D %*% R %*% D

# オブザベーション数, 変数の数, 標本平均ベクトル, 標本共分散行列を指定して, 標本を作成
Z <- covmatrix.h(n, p, M, S)
Z

round(apply(Z,2,mean), 3)
round(sd(Z)*sqrt((n-1)/n), 3)
round(cor(Z),3)

write.table(Z, "d:\\p_val_data_h.csv", row.names=FALSE, sep=",")

```