線型推測論

第11回 一般化逆行列

2020/6/5

統計数理研究所

長島 健悟

計画行列$\mathbf{X}$のランク

- 今までの講義では$\mathbf{X}$がフルランク, すなわち$\mathrm{rank}(\mathbf{X}) = p \leq n$, の場合だけを扱った
- $\mathrm{rank}(\mathbf{X}) < p$の場合
  - $\mathrm{rank}(\mathbf{X}'\mathbf{X}) < p$となり, $\mathbf{X}'\mathbf{X}$は正則でなくなる; $(\mathbf{X}'\mathbf{X})^{-1}$が存在しなくなる
  - 最小二乗推定量が一意に定まらない
  - モデルによってこういうことは起こってくる

一般線型モデルのイメージ

- 一般線型モデルは以下のようにも書ける
- $\mathbf{Y} = \beta_1 \mathbf{x}_1 + \beta_2 \mathbf{x}_2 + \cdots + \beta_p \mathbf{x}_p + \boldsymbol{\epsilon}$
  - $\mathbf{x}_j$：列ベクトル
  - $\mathbf{X}$の列ベクトルが張る空間上の点に誤差ベクトルを伸ばした場所

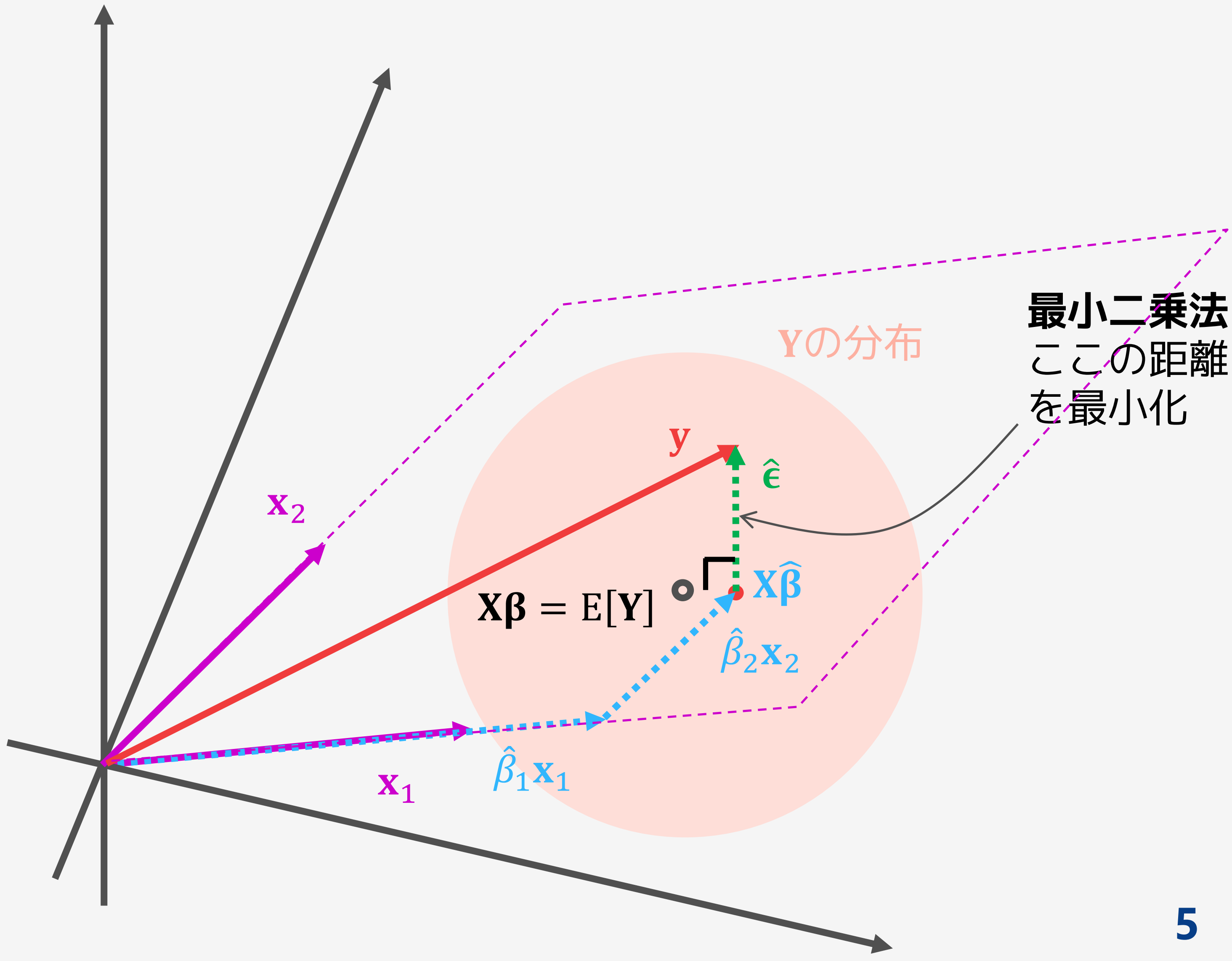
    ||

    確率変数ベクトル$\mathbf{Y}$の場所

一般線型モデルのイメージ

- 一番簡単な事例を考える
  - nは3次元
  - pは2次元
  - $\mathbf{X}$の列ベクトルが張る空間：$\beta_1 \mathbf{x}_1 + \beta_2 \mathbf{x}_2$
  - $\mathbf{y}$：確率変数$\mathbf{Y}$の観測値ベクトル

一般線型モデルのイメージ



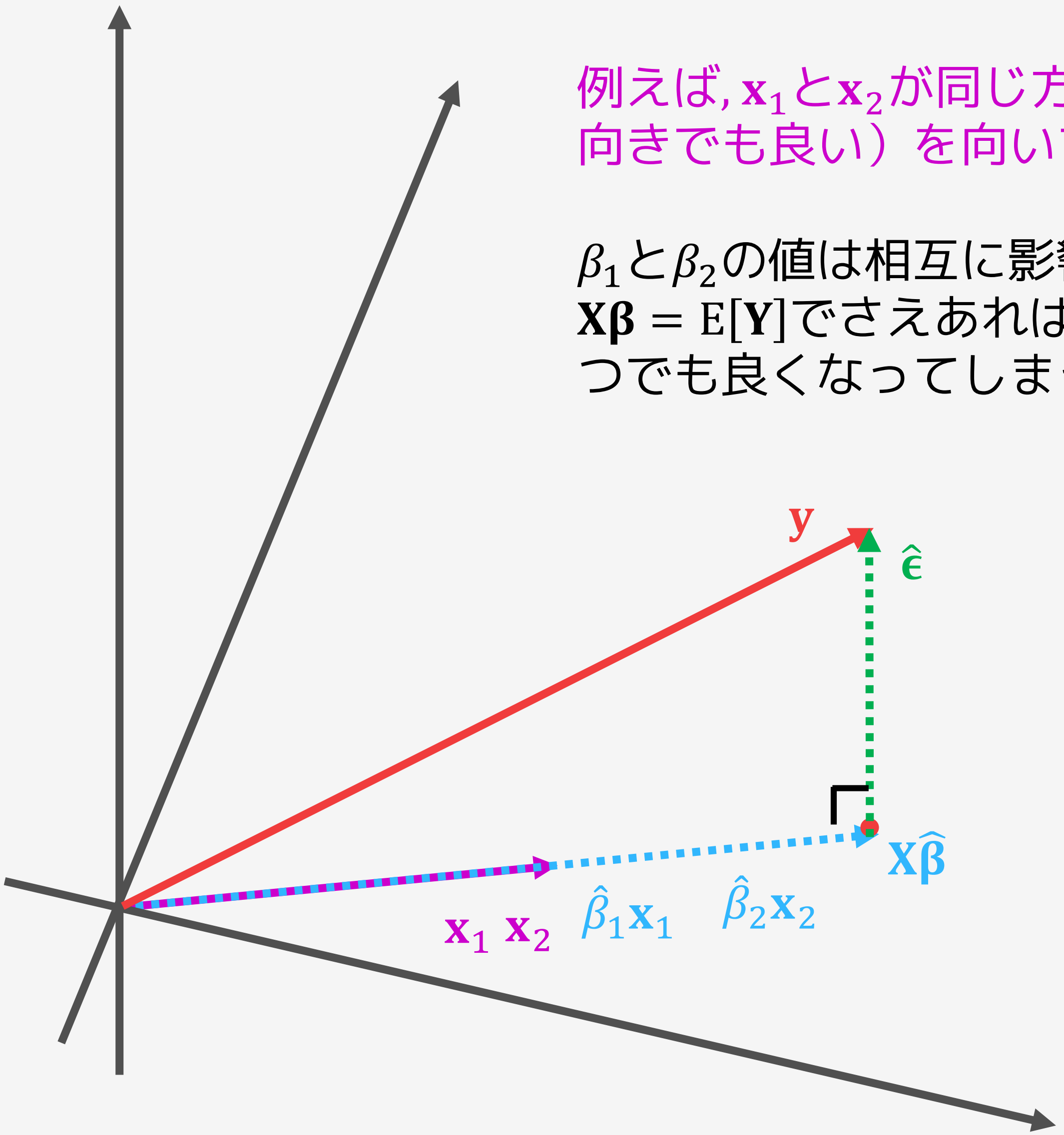
一般線型モデルのイメージ

- $\mathbf{X}$がフルランクでない場合, 一部の列ベクトル空間がつぶれてしまうので$\boldsymbol{\beta}$を一意に推定できない
  - 線形従属なβ_jの組の推定値は相互に影響しあうため$\hat{\boldsymbol{\beta}}$が一意に定まらない

ランク落ちした一般線型モデル

例えば, $\mathbf{x}_1$と$\mathbf{x}_2$が同じ方向（符号は逆向きでも良い）を向いてしまうと

β_1とβ_2の値は相互に影響し合って$\mathbf{X\beta} = \mathrm{E}[\mathbf{Y}]$でさえあれば値はいくつでも良くなってしまう・・・



計画行列Xのランク

- 次の一般線型モデル$\mathbf{Y} = \mathbf{X\beta} + \boldsymbol{\epsilon}$を考える
- $$\begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{pmatrix} = \begin{pmatrix} 6 & 2 & 2 & 2 \\ 2 & 2 & 0 & 0 \\ 2 & 0 & 2 & 0 \\ 2 & 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} \mu \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \end{pmatrix}$$
- $p = 4$
- $\mathrm{rank}(\mathbf{X}) = ?$

行列のランクの復習

- 行列$\mathbf{X}$のランク (rank; 階数) の定義
  - $\mathbf{X}$の列ベクトルの一次独立なものの最大個数を指し
  - $\mathbf{X}$の行ベクトルの一次独立なものの最大個数にも等しい

計画行列Xのランク

- 列基本変形を使う (操作は掃き出し法とほぼ同じです; 行基本変形でもよい)
- 1列目を6で割る

$$\begin{pmatrix} 1 & 2 & 2 & 2 \\ 1/3 & 2 & 0 & 0 \\ 1/3 & 0 & 2 & 0 \\ 1/3 & 0 & 0 & 2 \end{pmatrix}$$

計画行列Xのランク

- 2–4列目から1列目の2倍を引く

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 1/3 & 4/3 & -2/3 & -2/3 \\ 1/3 & -2/3 & 4/3 & -2/3 \\ 1/3 & -2/3 & -2/3 & 4/3 \end{pmatrix}$$

計画行列Xのランク

- 2列目を3/4倍
- 1列目から2列目の1/3倍を引く
- 3, 4列目に2列目の2/3倍を足す

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1/2 & -1/2 & 1 & -1 \\ 1/2 & -1/2 & -1 & 1 \end{pmatrix}$$

計画行列Xのランク

- 1列目から3列目の1/2倍を引く
- 2列目に3列目の1/2倍を足す
- 4列目に3列目を足す

$$
\begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
1 & -1 & -1 & 0
\end{pmatrix}
$$

- これ以上は変形できない
  - $\therefore \mathrm{rank}(\mathbf{X}) = 3$

計画行列Xのランク

- 因みにこの$\mathbf{X}$の場合
  - $\mathbf{x}_1 = (6,2,2,2)', \mathbf{x}_2 = (2,2,0,0)', \mathbf{x}_3 = (2,0,2,0)', \mathbf{x}_4 = (2,0,0,2)'$
- 例えば, $\beta_2\mathbf{x}_2 + \beta_3\mathbf{x}_3 + \beta_4\mathbf{x}_4$の張る空間上には$\beta_1\mathbf{x}_1$が乗っている (他の列ベクトル同士の組み合わせでもよい)

練習問題

- 次の行列$\mathbf{A}, \mathbf{B}$のランクを求めよ
- $\mathbf{A} = \begin{pmatrix} -1 & 1 \\ 2 & -2 \end{pmatrix}$
  - $\mathrm{rank}(\mathbf{A}) = 1$
- $\mathbf{B} = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & -1 \\ -1 & 0 & 2 \end{pmatrix}$
  - $\mathrm{rank}(\mathbf{B}) = 3$

ランクの性質

- 定理11-1
  - 任意の行列$\mathbf{X}$について
    $$\mathrm{rank}(\mathbf{X}'\mathbf{X}) = \mathrm{rank}(\mathbf{X}\mathbf{X}') = \mathrm{rank}(\mathbf{X}') = \mathrm{rank}(\mathbf{X})$$
  - 証明は割愛 (Gentle, 2007など; 部分行列を使う)

ランク落ちしたモデルの場合

- 正規方程式の解$\widehat{\boldsymbol{\beta}}$が無限に存在する
  - $\mathbf{X}'\mathbf{X}\widehat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{Y}$
- どうしたら良いか？
- 一応解は存在するので，ここではその中で使えそうな解を得る方針を考える
  - 一般化逆行列を用いる方法を紹介する

一般化逆行列

- $\mathbf{A}$を任意の$m \times n$行列とする
  - 以下を満たす行列$\mathbf{A}^{-}$を一般化逆行列とよぶ

$$\mathbf{A}\mathbf{A}^{-}\mathbf{A} = \mathbf{A}$$

  - $\mathbf{A}$は正方行列でなくてもよい
  - $\mathbf{A}^{-}$は一意でない
  - $\mathbf{A}$に逆行列が存在するならば, $\mathbf{A}^{-} = \mathbf{A}^{-1}$で一意に定まる

一般化逆行列の例

- $\mathbf{x} = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$のとき
  - $\mathbf{x}_1^- = (1,0,0,0)$, $\mathbf{x}_2^- = (0,1/2,0,0)$, $\mathbf{x}_3^- = (0,0,1/3,0)$, $\mathbf{x}_4^- = (0,0,0,1/4)$,
- $\mathbf{x}\mathbf{x}_i^-\mathbf{x} = \mathbf{x}$ $(i = 1,2,3,4)$を確認せよ
  - $\mathbf{x}\mathbf{x}_i^-\mathbf{x} =$

一般化逆行列の性質

- 定理11-2
  - 連立方程式$\mathbf{Ax}=\mathbf{c}$に解が存在するとき，$\mathbf{A}^{-}$を$\mathbf{A}$の一般化逆行列とすると
$$\mathbf{x}=\mathbf{A}^{-}\mathbf{c}$$
は解の一つになる

一般化逆行列の性質

- 定理11-3
  - 連立方程式$\mathbf{Ax} = \mathbf{c}$について, $\mathbf{A}^{-}$を$\mathbf{A}$の一般化逆行列とすると
$$\mathbf{AA}^{-}\mathbf{c} = \mathbf{c}$$
は解が存在する($\mathbf{x} = \mathbf{A}^{-}\mathbf{c}$が成り立つ)ための必要十分条件である

一般化逆行列による解

- 定理11-2を正規方程式に適用してみよう
  - 正規方程式：$\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{Y}$
  - $\mathbf{A} \Rightarrow \mathbf{X}'\mathbf{X}$, $\mathbf{x} \Rightarrow \hat{\boldsymbol{\beta}}$, $\mathbf{c} \Rightarrow \mathbf{X}'\mathbf{Y}$

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{Y}$$

一般化逆行列による解の性質

- 期待値$\mathrm{E}\left[\widehat{\boldsymbol{\beta}}\right]$を求めよ

$$
\begin{aligned}
\mathrm{E}\left[\widehat{\boldsymbol{\beta}}\right] &= \mathrm{E}[(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{Y}] \\
&= (\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathrm{E}[\mathbf{Y}] \\
&= (\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{X}\boldsymbol{\beta}
\end{aligned}
$$

一般化逆行列による解の性質

- $(\mathbf{X}'\mathbf{X})^- = (\mathbf{X}'\mathbf{X})^{-1}$でない限り
$$(\mathbf{X}'\mathbf{X})^-\mathbf{X}'\mathbf{X} \neq \mathbf{I}$$
  - $\hat{\boldsymbol{\beta}}$は$\boldsymbol{\beta}$の不偏推定量にはならない
  - $(\mathbf{X}'\mathbf{X})^-\mathbf{X}'\mathbf{X}\boldsymbol{\beta}$の不偏推定量になる
  - また, $(\mathbf{X}'\mathbf{X})^-$は一意でないため, $(\mathbf{X}'\mathbf{X})^-$の選び方によっても何の推定量かはかわってくる

何が推定できるのか？

- 以下の例を考える
- $$\begin{pmatrix} Y_{11} \\ Y_{12} \\ Y_{13} \\ Y_{21} \\ Y_{22} \\ Y_{23} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mu \\ \beta_1 \\ \beta_2 \end{pmatrix} + \begin{pmatrix} \epsilon_{11} \\ \epsilon_{12} \\ \epsilon_{13} \\ \epsilon_{21} \\ \epsilon_{22} \\ \epsilon_{23} \end{pmatrix}$$

何が推定できるのか？

- $\mathbf{X}'\mathbf{X}$を求めよ
  - $\mathbf{X}=\begin{pmatrix}1 & 1 & 0\\1 & 1 & 0\\1 & 1 & 0\\1 & 0 & 1\\1 & 0 & 1\\1 & 0 & 1\end{pmatrix}, \mathbf{X}'\mathbf{X}=\begin{pmatrix}6 & 3 & 3\\3 & 3 & 0\\3 & 0 & 3\end{pmatrix}$

何が推定できるのか？

- 以下に示す$(\mathbf{X}'\mathbf{X})^-$が$\mathbf{X}'\mathbf{X}$の一般化逆行列になっているかどうか確かめよ

$$(\mathbf{X}'\mathbf{X})^- = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1/3 & 0 \\ 0 & 0 & 1/3 \end{pmatrix}$$

何が推定できるのか？

- $\mathbf{X}'\mathbf{Y} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix} \mathbf{Y} = \begin{pmatrix} \bar{Y}_{\cdot\cdot} \\ \bar{Y}_{1\cdot} \\ \bar{Y}_{2\cdot} \end{pmatrix}$
  - $\widehat{\boldsymbol{\beta}} = ?$

$$\widehat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{Y}$$
$$= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1/3 & 0 \\ 0 & 0 & 1/3 \end{pmatrix} \begin{pmatrix} \bar{Y}_{\cdot\cdot} \\ \bar{Y}_{1\cdot} \\ \bar{Y}_{2\cdot} \end{pmatrix} = \begin{pmatrix} 0 \\ \bar{Y}_{1\cdot}/3 \\ \bar{Y}_{2\cdot}/3 \end{pmatrix}$$

何が推定できるのか？

- $\mathrm{E}[\hat{\boldsymbol{\beta}}]$を求めよ

$$\mathrm{E}[\hat{\boldsymbol{\beta}}] = \begin{pmatrix} 0 \\ \mathrm{E}[\bar{Y}_{1\cdot}/3] \\ \mathrm{E}[\bar{Y}_{2\cdot}/3] \end{pmatrix} = \begin{pmatrix} 0 \\ \mu + \beta_1 \\ \mu + \beta_2 \end{pmatrix}$$

$$\mathrm{E}[\bar{Y}_{i\cdot}/3] = \frac{1}{3}\mathrm{E}\left[\sum_{j=1}^{3} Y_{ij}\right] = \mu + \beta_i$$

- $\bar{Y}_{i\cdot} = \sum_{j=1}^{3} Y_{ij}$, $\mathrm{E}[Y_{ij}] = \mu + \beta_i$

何が推定できるのか？

- $\mathrm{E}[\widehat{\boldsymbol{\beta}}] = \begin{pmatrix} 0 \\ \mu + \beta_1 \\ \mu + \beta_2 \end{pmatrix}$
  - パラメータの線形結合が推定できる
  - $\widehat{\boldsymbol{\beta}}$については$(\mathbf{X}'\mathbf{X})^{-}$の選び方によって推定対象が異なる事に注意せよ

パラメータβの推定可能線型式

- パラメータ$\boldsymbol{\beta}$の推定可能線型式 (estimable functions) とは, データから(不偏に)推定することができるパラメータの線型結合$\mathbf{a}'\boldsymbol{\beta}$の事である
- 線型推定量$T = \mathbf{c}'\mathbf{Y}$について考える
  - $\mathrm{E}[T] = \mathbf{c}'\mathbf{X}\boldsymbol{\beta}$
  - つまり$\mathbf{c}'\mathbf{X}\boldsymbol{\beta} = \mathbf{a}'\boldsymbol{\beta} \Leftrightarrow \mathbf{c}'\mathbf{X} = \mathbf{a}'$を満たせば, Tは$\mathbf{a}'\boldsymbol{\beta}$の線型不偏推定量になる

パラメータβの推定可能線型式

- 推定したい$\mathbf{a}'\boldsymbol{\beta}$が推定可能かどうかは, $\mathbf{c}'\mathbf{X} = \mathbf{a}'$を満たす$\mathbf{c}'$が作れるかどうかで判定できる

パラメータβの推定可能線型式

- 以下の例を考える
- $$\begin{pmatrix} Y_{11} \\ Y_{12} \\ Y_{13} \\ Y_{21} \\ Y_{22} \\ Y_{23} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mu \\ \beta_1 \\ \beta_2 \end{pmatrix} + \begin{pmatrix} \epsilon_{11} \\ \epsilon_{12} \\ \epsilon_{13} \\ \epsilon_{21} \\ \epsilon_{22} \\ \epsilon_{23} \end{pmatrix}$$

パラメータβの推定可能線型式

- 例：$\beta_1 - \beta_2$を推定したいとする
  - $\mathbf{a}' = (0,1,-1)$
  - $\mathbf{c}'\mathbf{X} = (0,1,-1)$を満たす$\mathbf{c}'$があるかどうかを探せば良い

パラメータβの推定可能線型式

- $\mathbf{c}' = (0,0,1,-1,0,0)$のとき, $\mathbf{c}'\mathbf{X} = (0,1,-1)$となることを確認せよ

$$(0,0,1,-1,0,0)\begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix} = (\qquad\qquad)$$

- $\beta_1 - \beta_2$は推定可能

cの性質

- $\mathbf{Xd}=\mathbf{c}$を満たすベクトル$\mathbf{d}$が存在する
- 証明
  - $\mathbf{Xd}=\mathbf{c}\Leftrightarrow\mathbf{X}'\mathbf{Xd}=\mathbf{X}'\mathbf{c}$に定理11-3を適用する
    - 連立方程式は$\mathbf{X}'\mathbf{Xd}=\mathbf{X}'\mathbf{c}$ $(\mathbf{A}{:}\,\mathbf{X}'\mathbf{X},\mathbf{x}{:}\,\mathbf{d},\mathbf{c}{:}\,\mathbf{X}'\mathbf{c})$
  - $\mathbf{AA}^{-}\mathbf{c}{:}\ \mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{c}=\mathbf{X}'\mathbf{c}$ よって成立
    - $\because\ \mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'=\mathbf{X}'$

cの性質

- 定理11-4
  - $\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}' = \mathbf{X}'$
  - $\mathbf{X}(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{X} = \mathbf{X}$
    - $\{(\mathbf{X}'\mathbf{X})^{-}\}' = \{(\mathbf{X}'\mathbf{X})'\}^{-} = (\mathbf{X}'\mathbf{X})^{-}$

$\mathbf{a}'\widehat{\boldsymbol{\beta}}$の性質

- $\mathbf{a}'\widehat{\boldsymbol{\beta}} = \mathbf{c}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{Y}$
  - 先の結果から$\mathbf{c}' = \mathbf{d}'\mathbf{X}'$と書いて良い
- $\mathbf{a}'\widehat{\boldsymbol{\beta}} = \mathbf{d}'\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{Y}$
  - 定理11-4から$\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}' = \mathbf{X}'$
- $\mathbf{a}'\widehat{\boldsymbol{\beta}} = \mathbf{d}'\mathbf{X}'\mathbf{Y}$
  - $\mathbf{a}'\widehat{\boldsymbol{\beta}}$は$(\mathbf{X}'\mathbf{X})^{-}$の選び方にかかわらず一意に定まる

$\mathbf{a}'\hat{\boldsymbol{\beta}}$の性質

- 次の例を考える
  - $\mathbf{X}' = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix}$
  - $\mathbf{X}'\mathbf{X} = \begin{pmatrix} 6 & 3 & 3 \\ 3 & 3 & 0 \\ 3 & 0 & 3 \end{pmatrix}$

$\mathbf{a}'\widehat{\boldsymbol{\beta}}$の性質

- $\mathbf{a}'\widehat{\boldsymbol{\beta}} = \mathbf{a}'(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{Y}, \mathbf{a}' = (0,1,-1)$
  - $\mathbf{a}'(\mathbf{X}'\mathbf{X})_1^{-}\mathbf{X}' = ?, \mathbf{a}'(\mathbf{X}'\mathbf{X})_2^{-}\mathbf{X}' = ?$
  - $(\mathbf{X}'\mathbf{X})_1^{-} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1/3 & 0 \\ 0 & 0 & 1/3 \end{pmatrix}$
  - $(\mathbf{X}'\mathbf{X})_2^{-} = \begin{pmatrix} 1/3 & -1/3 & 0 \\ -1/3 & 2/3 & 0 \\ 0 & 0 & 0 \end{pmatrix}$

対処法のまとめ

- 一般化逆行列をそのまま使う場合は, 何を推定できる$(\mathbf{X}'\mathbf{X})^{-}$なのかを考える
  - 推定可能線型式を使ってもよい
- 他にも, パラメータを減らした新しいモデルを作る方法も考えられる
  - $y = \mu + \beta_1 x_1 + \cdots \rightarrow y = \gamma_1 x_1 + \cdots$
  - 新しいモデルのパラメータが何を表わすかはよく考える必要がある

まとめ

- ランク落ちした場合の推測について議論した
  - 一般化逆行列
  - 推定可能線型式
  - モデルの再構築

補足

- 線型独立なパラメータ$\boldsymbol{\beta}$の推定可能線型式の数は$\mathbf{X}$のランクに等しい
- ランク落ちした場合, 推定可能線型式の数までであれば線型仮説の検定が可能である
  - 分散の推定量等では$(\mathbf{X}'\mathbf{X})^{-1}$を$(\mathbf{X}'\mathbf{X})^{-}$に置き換えるだけで良い

参考文献

- Gentle JE. *Matrix Algebra Theory, Computations, and Applications in Statistics*. Springer, 2007, pp.91-92.