線型推測論

第09回 最尤推定量(1)

2018/6/6

統計数理研究所

長島 健悟

一般化最小二乗推定量

- $\mathrm{Cov}[\boldsymbol{\epsilon}] = \sigma^2\mathbf{I}$から$\mathrm{Cov}[\boldsymbol{\epsilon}] = \sigma^2\mathbf{V}$へ緩和した場合のBLUE等について導いた
  - $\mathbf{V}$が既知である事を仮定していた
  - 正しく特定しないとうまく推定ができない
- これに関連して, 今回は最尤推定量について議論する

二つの推定法

モデル	一般線型モデル	一般化線型モデル
基本理論	小標本理論	尤度理論
推定量	最小二乗推定量 /BLUE MINQUE/MIVQUE	最尤推定量
推定量の基準	不偏性 有効性	一致性 漸近有効性
基準の尺度	期待値 分散	確率収束 漸近分散
n	固定	∞
統計量の分布	分布を直接求める	漸近正規性

二つの推定法

- 二つの方法は全く異なる理論がベースになっている
  - 互いに関連する部分もあって, 一部については等しい推定量が得られる

SASでの実装

- SASではPROCEDUREで分かれている
  - 一般線型モデル：GLM
  - 一般化線型モデル(尤度ベース)：MIXEDやGENMOD
- 使い方は似ているが…
  - ヘルプをよく読むと, 手法の仮定や用いる事ができる統計量が全然違うことがよく分かる

SASでの実装

- GLM
  - 純粋な一般線型モデルのPROCEDURE
  - 小標本理論に基づいた各種推定方法をサポート
  - 分散成分の理論がベースとなる線型混合モデル, 繰り返し測定分散分析; MANOVAなども

SASでの実装

- MIXED
  - 尤度理論ベースの線型混合モデルのPROCEDURE
- GENMOD
  - 一般化線型モデル用のPROCEDUREに見えるが, 一般化線型モデル+GEEに対応
    - GEEは尤度をベースにしていないのに…

最尤推定量

- 最尤推定量
  - 尤度関数$L(\boldsymbol{\theta}) = f(\mathbf{Y}; \boldsymbol{\theta}) = \prod_{i=1}^{n} f(y_i; \boldsymbol{\theta})$ ($\mathbf{Y}$はi.i.d.)は同時確率関数をパラメータ$\boldsymbol{\theta} = (\boldsymbol{\beta}, \sigma^2)$の関数と見なしたもの
    - 尤度関数は確率変数$\mathbf{Y}$を含んだ関数である
    - $\boldsymbol{\theta}$に対する密度関数ではない事は注意せよ

最尤推定量

- 対数尤度関数：$l(\boldsymbol{\theta}) = \log L(\boldsymbol{\theta})$
- 最尤推定量

$$\hat{\boldsymbol{\theta}} = \mathrm{argmax}_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} l(\boldsymbol{\theta})$$

  - (対数)尤度関数を最大化する$\boldsymbol{\theta}$をパラメータの推定量とする

最尤推定量

- 一致性

$$\forall \epsilon, \lim_{n\to\infty} \Pr\left(\left|\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0\right| \geq \epsilon\right) = 0$$

  - $n \to \infty$で真値に近づく
  - $\boldsymbol{\theta}_0$はパラメータの真値

最尤推定量

- 漸近正規性：ある正則条件の下で

$$\sqrt{n}\left(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0\right) \to^d N\left(\mathbf{0}, \mathbf{I}^{-1}(\boldsymbol{\theta}_0)\right)$$

  - $\mathbf{I}(\boldsymbol{\theta}_0) = \mathrm{E}\left[\frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}\right]$：Fisher情報行列

- 本講義では扱わないが, 漸近理論の講義の収束・最尤推定量の話は非常に重要
- 数列・関数の収束は基本的素養なので一度きちんと理解しておくことが望ましい

一般線型モデルの最尤推定量

- 一般線型モデルの仮定
  - $\mathbf{Y} \sim N(\mathbf{X\beta}, \sigma^2\mathbf{I})$

- 平均ベクトル$\boldsymbol{\mu}$が, 分散共分散行列が$\boldsymbol{\Sigma}$のn変量多変量正規分布の確率密度関数は

$$\frac{1}{(2\pi)^{n/2}|\boldsymbol{\Sigma}|^{1/2}}\exp\left\{-\frac{(\mathbf{Y}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{Y}-\boldsymbol{\mu})}{2}\right\}$$

一般線型モデルの最尤推定量

- 一般線型モデル$\mathbf{Y} \sim N(\mathbf{X\beta}, \sigma^2\mathbf{I})$の尤度関数$L(\boldsymbol{\theta})$を示せ

$$L(\boldsymbol{\theta}) = \frac{1}{(2\pi)^{n/2}|\sigma^2\mathbf{I}|^{1/2}} e^{-(\mathbf{Y}-\mathbf{X\beta})'\left(\sigma^2\mathbf{I}\right)^{-1}(\mathbf{Y}-\mathbf{X\beta})/2} = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left\{-\frac{(\mathbf{Y}-\mathbf{X\beta})'(\mathbf{Y}-\mathbf{X\beta})}{2\sigma^2}\right\}$$

一般線型モデルと最尤推定量

- i.i.d.なので, iごとの正規分布の確率密度関数

$$f(Y_i;\boldsymbol{\theta}) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{\frac{(Y_i - \mathbf{x}_i'\boldsymbol{\beta})^2}{2\sigma^2}\right\}$$

から求めても良い

一般線型モデルの最尤推定量

- 一般線型モデル$\mathbf{Y} \sim N(\mathbf{X\beta}, \sigma^2 \mathbf{I})$の対数尤度関数$l(\boldsymbol{\theta})$を示せ

$$l(\boldsymbol{\theta}) = -\frac{n}{2}\log 2\pi - \frac{n}{2}\log \sigma^2 - \frac{1}{2\sigma^2}(\mathbf{Y} - \mathbf{X\beta})'(\mathbf{Y} - \mathbf{X\beta})$$

一般線型モデルの最尤推定量

- 最尤推定量を求める
  - 対数尤度関数を各パラメータで偏微分したスコア関数$\mathbf{S}(\boldsymbol{\theta})$を求め, $\mathbf{S}(\boldsymbol{\theta}) = \mathbf{0}$を満たす解を最尤推定量として求める
    - $\hat{\boldsymbol{\theta}}$における対数尤度関数のヘッセ行列が負定値であれば, $\hat{\boldsymbol{\theta}}$は極大値である
    - 最尤推定量は必ず存在するわけではない (多くの場合存在し, ヘッセ行列は負定値になっているが...)

一般線型モデルの最尤推定量

- $\frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}} = ?$

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}}$$

$$= \mathbf{0} + \frac{\partial}{\partial \boldsymbol{\beta}}\left\{-\frac{1}{2\sigma^2}(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})\right\}$$

$$= \frac{\partial}{\partial \boldsymbol{\beta}}\left\{-\frac{1}{2\sigma^2}(\mathbf{Y}'\mathbf{Y} - 2\boldsymbol{\beta}'\mathbf{X}'\mathbf{Y} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta})\right\}$$

$$= -\frac{1}{2\sigma^2}(-2\mathbf{X}'\mathbf{Y} + 2\mathbf{X}'\mathbf{X}\boldsymbol{\beta})$$

一般線型モデルの最尤推定量

- $\frac{\partial l(\boldsymbol{\theta})}{\partial \sigma^2} = ?$

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})$$

一般線型モデルの最尤推定量

- $\mathbf{S}(\boldsymbol{\theta})=\mathbf{0}$を満たす$\widehat{\boldsymbol{\theta}}$を求めよ

$$-\frac{1}{2\sigma^2}\left(-2\mathbf{X}'\mathbf{Y}+2\mathbf{X}'\mathbf{X}\widehat{\boldsymbol{\beta}}\right)=\mathbf{0}$$

$$\Leftrightarrow \widehat{\boldsymbol{\beta}}=(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$$

$$-\frac{n}{2\hat{\sigma}^2}+\frac{1}{2(\hat{\sigma}^2)^2}\left(\mathbf{Y}-\mathbf{X}\widehat{\boldsymbol{\beta}}\right)'\left(\mathbf{Y}-\mathbf{X}\widehat{\boldsymbol{\beta}}\right)=0$$

$$\Leftrightarrow \hat{\sigma}^2=\frac{1}{n}\left(\mathbf{Y}-\mathbf{X}\widehat{\boldsymbol{\beta}}\right)'\left(\mathbf{Y}-\mathbf{X}\widehat{\boldsymbol{\beta}}\right)$$

一般線型モデルの最尤推定量

- 一般線型モデルにおける最尤推定量は
  - $\widehat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$
  - $\hat{\sigma}^2 = \frac{1}{n}\left(\mathbf{Y} - \mathbf{X}\widehat{\boldsymbol{\beta}}\right)'\left(\mathbf{Y} - \mathbf{X}\widehat{\boldsymbol{\beta}}\right)$
- やや異なる結果が得られた
  - $\widehat{\boldsymbol{\beta}}$は最小二乗推定量に一致
  - $\hat{\sigma}^2$はプラグイン推定量に一致; 不偏推定量にはならない

尤度比検定

- 階層構造をなす二つのモデルを比較
  - 単純なモデル$\mathcal{M}_0$と複雑なモデル$\mathcal{M}_1$
  - $H_0: \mathcal{M}_0 = \mathcal{M}_1$の検定を考える
  - $\mathcal{M}_1$と$\mathcal{M}_0$のパラメータ数の差をν
  - $\mathcal{M}_0, \mathcal{M}_1$の各最大対数尤度$l_0(\widehat{\boldsymbol{\theta}}_0), l_1(\widehat{\boldsymbol{\theta}}_1)$
- このとき

$$T = -2\left[l_1(\widehat{\boldsymbol{\theta}}_1) - l_0(\widehat{\boldsymbol{\theta}}_0)\right] \to^d \chi^2(\nu)$$

尤度比検定

- $H_0 : \boldsymbol{\beta} = \mathbf{0}$の検定について考えてみよう
  - 単純モデルの最大対数尤度は$l_0(\mathbf{0}, \hat{\sigma}_0^2)$
  - 複雑モデルの最大対数尤度は$l_1(\hat{\boldsymbol{\beta}}, \hat{\sigma}^2)$

尤度比検定

- $l_0(\mathbf{0}, \hat{\sigma}_0^2)$を計算せよ

$$
\begin{aligned}
&l_0(\mathbf{0}, \hat{\sigma}_0^2) \\
&= -\frac{n}{2}\log 2\pi - \frac{n}{2}\log \hat{\sigma}_0^2 - \frac{1}{2\hat{\sigma}_0^2}\mathbf{Y}'\mathbf{Y} \\
&= -\frac{n}{2}\log 2\pi - \frac{n}{2}\log \frac{\mathbf{Y}'\mathbf{Y}}{n} - \frac{n}{2}
\end{aligned}
$$

- $\hat{\sigma}_0^2 = \frac{1}{n}\mathbf{Y}'\mathbf{Y}$

尤度比検定

- $l_1\left(\widehat{\boldsymbol{\beta}}, \hat{\sigma}^2\right)$を計算せよ

$$
\begin{aligned}
&l_1\left(\widehat{\boldsymbol{\beta}}, \hat{\sigma}^2\right)\\
&= -\frac{n}{2}\log 2\pi - \frac{n}{2}\log \hat{\sigma}^2 - \frac{1}{2\hat{\sigma}^2}\left(\mathbf{Y} - \mathbf{X}\widehat{\boldsymbol{\beta}}\right)'\left(\mathbf{Y} - \mathbf{X}\widehat{\boldsymbol{\beta}}\right)\\
&= -\frac{n}{2}\log 2\pi - \frac{n}{2}\log \frac{\left(\mathbf{Y} - \mathbf{X}\widehat{\boldsymbol{\beta}}\right)'\left(\mathbf{Y} - \mathbf{X}\widehat{\boldsymbol{\beta}}\right)}{n} - \frac{n}{2}
\end{aligned}
$$

尤度比検定

- $T=-2\left[l_1\left(\widehat{\boldsymbol{\theta}}_1\right)-l_0\left(\widehat{\boldsymbol{\theta}}_0\right)\right]$を求めよ

$$T=n\left(\log\frac{\left(\mathbf{Y}-\mathbf{X}\widehat{\boldsymbol{\beta}}\right)'\left(\mathbf{Y}-\mathbf{X}\widehat{\boldsymbol{\beta}}\right)}{n}-\log\frac{\mathbf{Y}'\mathbf{Y}}{n}\right)$$

尤度比検定

- ここで$\exp\left(\frac{T}{n}\right)$を求めてみよう

$$\exp\left(\frac{T}{n}\right) = \frac{\left(\mathbf{Y}-\mathbf{X}\widehat{\boldsymbol{\beta}}\right)'\left(\mathbf{Y}-\mathbf{X}\widehat{\boldsymbol{\beta}}\right)}{\mathbf{Y}'\mathbf{Y}}$$

$$= \frac{\mathbf{Y}'\mathbf{Y}-\mathbf{Y}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}}{\mathbf{Y}'\mathbf{Y}}$$

$$= \frac{1}{1+\dfrac{\mathbf{Y}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}}{\mathbf{Y}'\mathbf{Y}-\mathbf{Y}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}}}$$

尤度比検定

- Tは線型仮説のF検定と似ている
  - $\mathbf{C}=\mathbf{I}$のとき$(\mathbf{C}\widehat{\boldsymbol{\beta}})'\{\mathbf{C}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}\}^{-1}\mathbf{C}\widehat{\boldsymbol{\beta}}=\mathbf{Y}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$
  - $s^2=\frac{1}{n-p}\{\mathbf{Y}'\mathbf{Y}-\mathbf{Y}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}\}$
  - $F=\frac{\mathbf{Y}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}/p}{s^2}$
- $\exp\left(\frac{T}{n}\right)=\frac{1}{1+pF/(n-p)}$が成り立つ