

線型推測論

第10回 最尤推定量(2)

2020/6/4

統計数理研究所

長島 健悟

誤差の仮定を緩めた場合への拡張

- 分散共分散行列が $\sigma^2 \mathbf{I}$ の場合の最尤推定量を導出した
- 今度は相関がある場合の最尤推定量を考えよう

誤差の仮定を緩めた場合への拡張

- 拡張した一般線型モデルの仮定
 - $\mathbf{Y} \sim N(\mathbf{X}\boldsymbol{\beta}, \mathbf{V})$

注： σ^2 がちょっと邪魔なので、分散共分散行列の中に含まれた形で定義しなおす

一般線型モデルの最尤推定量

- 一般線型モデル $Y \sim N(X\beta, V)$ の尤度関数 $L(\theta)$ を示せ

$$L(\theta) = \frac{1}{(2\pi)^{n/2} |\mathbf{V}|^{1/2}} e^{-\frac{1}{2}(Y - X\beta)' V^{-1} (Y - X\beta)}$$

一般線型モデルの最尤推定量

- 一般線型モデル $Y \sim N(X\beta, V)$ の対数尤度関数 $l(\theta)$ を示せ

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

一般線型モデルの最尤推定量

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

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

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

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

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

一般線型モデルの最尤推定量

- $\frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}} = \mathbf{0}$ を満たす $\hat{\boldsymbol{\beta}}$ を求めよ

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

$$\Leftrightarrow \hat{\boldsymbol{\beta}} = \left(\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{X} \right)^{-1} \mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{Y}$$

- \mathbf{V} はどう推定したらよい？

Vの推定方法

- このクラスのモデルにおける推定に対しては, プロファイル尤度法を用いることができる

最尤法の改良法

- 通常的最尤法：全尤度関数に基づく
 - 全尤度関数の中には推定に無関係な局外パラメータが含まれることがある
 - 推定が複雑になる
 - 推定精度に悪影響が出ることもある
 - この問題の解決方法がいくつも提案されてきた

最尤法の改良法

- プロファイル尤度法：興味あるパラメータを与えたもとでの局外パラメータ最尤推定量を全尤度関数に代入したもの
- 条件付き尤度法：局外パラメータの十分統計量の条件付き分布を使う方法
- 周辺尤度法：局外パラメータを積分消去する方法
- 部分尤度法：Cox回帰で用いる, 実はプロファイル尤度としても解釈できる

プロファイル尤度法

- 尤度関数を $l(\theta)$
- $\theta = (\alpha', \phi)'$, 興味のあるパラメータは α , 局外パラメータは ϕ
- α を与えたもとでの ϕ の最尤推定量を $\hat{\phi}(\alpha)$ と表わす
- プロファイル尤度法

$$\hat{\alpha} = \operatorname{argmax}_{\alpha \in \Theta_{\alpha}} l\{\alpha, \hat{\phi}(\alpha)\}$$

Vのプロファイル尤度推定量

- Vの推定に興味がある
- Vを与えたもとでの最尤推定量は

$$\hat{\beta}(V) = (X'V^{-1}X)^{-1}X'V^{-1}Y$$

- これを全尤度関数に代入すると

$$l\{V, \hat{\beta}(V)\}$$

$$= -\frac{n}{2} \log 2\pi - \frac{1}{2} \log |V| - \frac{1}{2} \mathbf{r}'V^{-1}\mathbf{r}$$

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

Vのプロファイル尤度推定量

- 残念ながら解析的には解けないため、数値最適化により以下をみたす \hat{V} を計算する

$$\hat{V} = \operatorname{argmax}_{V \in \Theta_V} l\{V, \hat{\beta}(V)\}$$

- ちなみに $V = \sigma^2 I$ とおいたときは $\hat{\sigma}^2 = \frac{1}{n} (\mathbf{Y} - \mathbf{X}\hat{\beta})' (\mathbf{Y} - \mathbf{X}\hat{\beta})$ となる

制限付き最尤推定 (REML)

- Restricted maximum likelihood estimation
 - 通常的最尤推定量では, $\hat{\sigma}^2$ は不偏性を満たさなかった
 - 何とかこれの改良を考えられないだろうか？

制限付き最尤推定 (REML)

- Harville (1974)の方法
- V の推定のみに興味があるため, β に対して周辺化した周辺尤度関数

$$\int_{\beta \in \Theta_\beta} L(\theta) d\beta$$

に基づく推測としてREMLを導出する

- Θ_β : β のパラメータ空間全体

制限付き最尤推定 (REML)

- 導出では以下の式変形を用いる

$$\begin{aligned}(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})' \mathbf{V}^{-1} (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}) &= \\ &(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})' \mathbf{V}^{-1} (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}) \\ &+ (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})' (\mathbf{X}' \mathbf{V}^{-1} \mathbf{X}) (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})\end{aligned}$$

- ちょっと計算が必要だが証明は簡単
- ヒント : $\hat{\boldsymbol{\beta}}(\mathbf{V}) = (\mathbf{X}' \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}' \mathbf{V}^{-1} \mathbf{Y}$

制限付き最尤推定 (REML)

- 周辺尤度関数を求める

$$\int_{\boldsymbol{\beta} \in \Theta_{\boldsymbol{\beta}}} L(\boldsymbol{\theta}) d\boldsymbol{\beta} \quad \because a = \frac{1}{(2\pi)^{n/2} |\mathbf{V}|^{1/2}}$$
$$= a \int_{\boldsymbol{\beta} \in \Theta_{\boldsymbol{\beta}}} \exp \left\{ -\frac{1}{2} (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})' \mathbf{V}^{-1} (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}) \right\} d\boldsymbol{\beta}$$
$$= a \exp \left\{ -\frac{1}{2} (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})' \mathbf{V}^{-1} (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}) \right\}$$
$$\int_{\boldsymbol{\beta} \in \Theta_{\boldsymbol{\beta}}} \exp \left\{ -\frac{1}{2} (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})' (\mathbf{X}' \mathbf{V}^{-1} \mathbf{X}) (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}) \right\} d\boldsymbol{\beta}$$

制限付き最尤推定 (REML)

- $(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}) = \mathbf{W}^{-1}$ とおく
- $\int_{\boldsymbol{\beta} \in \Theta_{\boldsymbol{\beta}}} \exp \left\{ -\frac{1}{2} (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})' \mathbf{W}^{-1} (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}) \right\} d\boldsymbol{\beta} = ?$
$$\int_{\boldsymbol{\beta} \in \Theta_{\boldsymbol{\beta}}} \exp \left\{ -\frac{1}{2} (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})' \mathbf{W}^{-1} (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}) \right\} d\boldsymbol{\beta}$$
$$= b \int_{\boldsymbol{\beta} \in \Theta_{\boldsymbol{\beta}}} \frac{1}{b} \exp \left\{ -\frac{1}{2} (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})' \mathbf{W}^{-1} (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}) \right\} d\boldsymbol{\beta}$$
$$= b$$
- $b = (2\pi)^{p/2} |\mathbf{W}|^{1/2}$

制限付き最尤推定 (REML)

- $$\begin{aligned} & \int_{\boldsymbol{\beta} \in \Theta_{\boldsymbol{\beta}}} L(\boldsymbol{\theta}) d\boldsymbol{\beta} \\ &= ab \exp \left\{ -(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})' \mathbf{V}^{-1} (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}) / 2 \right\} \\ &= \frac{|\mathbf{X}' \mathbf{V}^{-1} \mathbf{X}|^{-1/2}}{(2\pi)^{(n-p)/2} |\mathbf{V}|^{1/2}} \\ & \quad \times \exp \left\{ -(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})' \mathbf{V}^{-1} (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}) / 2 \right\} \\ &= \frac{1}{(2\pi)^{(n-p)/2} |\mathbf{V}|^{1/2} |\mathbf{X}' \mathbf{V}^{-1} \mathbf{X}|^{1/2}} \\ & \quad \times \exp \left\{ -(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})' \mathbf{V}^{-1} (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}) / 2 \right\} \end{aligned}$$

制限付き最尤推定 (REML)

- 最後に $\hat{\beta}(V) = (X'V^{-1}X)^{-1}V^{-1}X'Y$ を代入し、整理すれば

$$\begin{aligned} l_R\{V, \hat{\beta}(V)\} &= \log \int_{\beta \in \Theta_\beta} L(\theta) d\beta = \\ &= -\frac{n-p}{2} \log 2\pi - \frac{1}{2} \log |V| \\ &\quad - \frac{1}{2} \log |X'V^{-1}X| - \frac{1}{2} r'V^{-1}r \end{aligned}$$

制限付き最尤推定 (REML)

- $V = \sigma^2 \mathbf{I}$ の場合, REMLによる分散の推定量は不偏推定量に一致することが知られている
- $$\hat{\sigma}_R^2 = \frac{1}{n-p} (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})' (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})$$

まとめ

- 一般線型モデルにおける最尤法, 制限付き最尤法を説明した
- $\hat{\beta}$ は最小二乗推定量, 一般化最小二乗推定量と一致する
- σ^2 の最尤推定量は不偏ではない
- 制限付き最尤法を使えば σ^2 の不偏な推定量を与える事ができる

引用文献

- Harville D. Bayesian inference for variance components using only error contrasts. *Biometrika* 1974;**61**(2):383–385.