

線型推測論

第05回 σ^2 の点推定量

2020/5/21

統計数理研究所

長島 健悟

もう一つの未知パラメータ

- 一般線型モデルのパラメータ
 - $E[Y] = X\beta$
- もう一つ未知パラメータがありました
 - $\text{Cov}[Y] = \sigma^2 I$

分散 σ^2 の推定量

- 分散 σ^2 も重要なパラメータである
 - $\text{Cov}[\hat{\beta}] = \sigma^2 (\mathbf{X}'\mathbf{X})^{-1}$
- 最小二乗法で σ^2 の推定量は構成できない
- 分散の推定は別途考えなければならない

分散 σ^2 の推定量

- 分散の推定量は線型推定量でないが、基本的な考え方は同じである
- 推定量は Y の関数である
- 推定量の満たすべき望ましい性質
 - 不偏性・最小分散性

推定量の構成方法

- 無数のタイプの推定量とその構成方法が存在する
 - 線型推定量：制約のもとで最適な線形結合を導出する
 - 最尤推定量：最尤法を使う
 - プラグイン推定量：複数パラメータある時に、一部のパラメータに推定量を代入して別のパラメータの推定量を得る
- ロバスト推定量, James–Stein推定量...

分散のプラグイン推定量

- プラグイン推定量
 - 興味あるパラメータ α が他のパラメータ β の関数になっているとき, β の推定量を代入して α の推定量を得る
 - $\alpha(\beta)$ の推定量を $\hat{\alpha}(\hat{\beta})$ として作る

一般線型モデルにおける分散

- 一般線型モデル中の分散

- $\text{Cov}[\mathbf{Y}] = \sigma^2 \mathbf{I}$

$$\begin{aligned}\text{Var}[Y_i] &= E[\{Y_i - E[Y_i]\}^2] \\ &= E[\{Y_i - \mathbf{x}'_i \boldsymbol{\beta}\}^2] \\ &= \sigma^2\end{aligned}$$

- $\mathbf{x}'_i = (x_{i1}, \dots, x_{ij}, \dots, x_{ip})$
- $\boldsymbol{\beta}$ を $\hat{\boldsymbol{\beta}}$ (BLUE)で置き換える方針を考える

分散のプラグイン推定量

- 期待値は標本平均に置き換える

- $E \left[\{Y_i - \mathbf{x}'_i \hat{\boldsymbol{\beta}}\}^2 \right] \Rightarrow \frac{1}{n} \sum_{i=1}^n (Y_i - \mathbf{x}'_i \hat{\boldsymbol{\beta}})^2$

- 分散のプラグイン推定量

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

- この推定量が良い性質を持っていれはうれしい . . .

期待値の標本平均への置き換え

- $E[T_i]$ の推定において, 標本平均による推定量 $T = \frac{1}{n} \sum_{i=1}^n T_i$ を考えた
- 各 T_i が等分散・無相関で, $E[T_i] = \theta$ が成立するとき, 以下の推定量が得られる

- $$E[T] = E \left[\frac{1}{n} \sum_{i=1}^n T_i \right] = \theta$$

- $$\text{Var}[T] = \frac{1}{n^2} \text{Var} \left[\sum_{i=1}^n T_i \right] = \frac{1}{n} \text{Var}[T_i]$$

プラグイン推定量の期待値

- $E[\hat{\sigma}^2] = \frac{1}{n} E \left[(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})' (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}) \right]$
- 練習問題（ヒント：内積, $\hat{\boldsymbol{\beta}} = ?$ ）

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

$$\begin{aligned} & (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})' (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}) \\ &= \mathbf{Y}'\mathbf{Y} - 2\hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{Y} + \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} \\ &= \mathbf{Y}'\mathbf{Y} - \mathbf{Y}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} \\ &= \mathbf{Y}'\{\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\}\mathbf{Y} \end{aligned}$$

プラグイン推定量の期待値

- 練習問題： $E \left[(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})' (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}) \right] = ?$

$$E \left[(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})' (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}) \right]$$

$$= E[\mathbf{Y}'\mathbf{A}\mathbf{Y}]$$

$$= \text{tr}(\mathbf{A}E[\mathbf{Y}\mathbf{Y}'])$$

$$= \text{tr}(\mathbf{A}\sigma^2\mathbf{I} + \mathbf{A}E[\mathbf{Y}]E[\mathbf{Y}]')$$

$$= \sigma^2\text{tr}(\mathbf{A}) + \text{tr}(E[\mathbf{Y}]'\mathbf{A}E[\mathbf{Y}])$$

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

プラグイン推定量の期待値

- 定理5-1：スカラー $Y'AY$ の期待値について以下が成立する（結構重要な性質）
- $E[Y'AY] = E[\text{tr}\{Y'AY\}] = E[\text{tr}\{AYY'\}] = \text{tr}\{E[AYY']\} = \text{tr}\{AE[YY']\}$
- $Y'AY = \text{tr}\{Y'AY\}$ (トレースの定義)
- $\text{tr}\{Y'AY\} = \text{tr}\{AYY'\}$ (定理1-6)
- $E[\text{tr}\{AYY'\}] = \text{tr}\{E[AYY']\}$ (トレースの定義)

プラグイン推定量の期待値

- 系5-2

$$E[YY'] = \text{Cov}[Y] + E[Y]E[Y]'$$

$$\therefore \text{Cov}[Y] = E[(Y - E[Y])(Y - E[Y])']$$

- 系5-3 : 誤差の三条件のもとで

$$E[YY'] = \sigma^2 \mathbf{I} + E[Y]E[Y]'$$

プラグイン推定量の期待値

- 練習問題： $\text{tr}\{\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\} = ?$

$$\begin{aligned} & \text{tr}\{\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\} \\ &= \text{tr}(\mathbf{I}) - \text{tr}\{\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\} \\ &= \text{tr}(\mathbf{I}) - \text{tr}\{(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}\} \\ &= \text{tr}(\mathbf{I}_n) - \text{tr}(\mathbf{I}_p) \\ &= n - p \end{aligned}$$

- \mathbf{X} : $n \times p$ 次元
- $\mathbf{X}'\mathbf{X}$: $(p \times n) \times (n \times p)$ 次元

プラグイン推定量の期待値

- 練習問題： $\text{tr}(\mathbf{E}[\mathbf{Y}]' \mathbf{A} \mathbf{E}[\mathbf{Y}]) = ?$

$$\text{tr}(\mathbf{E}[\mathbf{Y}]' \mathbf{A} \mathbf{E}[\mathbf{Y}])$$

$$= \mathbf{E}[\mathbf{Y}]' \mathbf{A} \mathbf{E}[\mathbf{Y}]$$

$$= \boldsymbol{\beta}' \mathbf{X}' \{ \mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}' \} \mathbf{X} \boldsymbol{\beta}$$

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

$$= 0$$

プラグイン推定量の期待値

- 以上の結果から,

$$\begin{aligned} E[\hat{\sigma}^2] &= \frac{1}{n} E \left[(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})' (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}) \right] \\ &= \frac{1}{n} (n - p) \sigma^2 \end{aligned}$$

- プラグイン推定量は不偏ではなかった . . .
- しかし, 後一步の見た目をしている

分散 σ^2 の不偏推定量 s^2

- $s^2 = \frac{(Y - X\hat{\beta})'(Y - X\hat{\beta})}{n - p}$

$$E[s^2] = \frac{1}{n - p} (n - p)\sigma^2 = \sigma^2$$

- これを用いれば $\text{Cov}[\hat{\beta}]$ の不偏推定量が得られる

$$\widehat{\text{Cov}}[\hat{\beta}] = s^2 (\mathbf{X}'\mathbf{X})^{-1}$$

分散 σ^2 の不偏推定量 s^2

- 定理5-4 : s^2 は $E[Y_i - E[Y_i]]^4 = 3\sigma^4$ のとき,
 σ^2 の最小分散不偏推定量である
- 証明は Wang & Chow (1993, pp. 161–163)
- $Y_i \sim_{\text{i.i.d.}} N(\mathbf{x}'_i \boldsymbol{\beta}, \sigma^2)$ ならば
 $E[Y_i - E[Y_i]]^4 = 3\sigma^4$ が成立 (正規分布の尖度)

分散成分の推定理論

- MINQUE (Minimum Norm Quadratic Unbiased Estimator)
- MIVQUE (Minimum Variance Quadratic Unbiased Estimator)
- Quadratic Formの中で不偏性をみだし分散がなるべく小さくなるもの
- これらは線型混合モデルの推定理論として研究されてきたが、現在はほぼ最尤法や制限付き最尤法に取って代わられた

まとめ

- 最小二乗法だけでは分散の推定量を構成できない
- 無数の推定量が存在し, その構成方法も色々である
 - 色々な推定量を勉強しておくが良い
- 一般線型モデルにおける分散の不偏推定

量は :
$$s^2 = \frac{Y' \{ I - X(X'X)^{-1}X' \} Y}{n-p}$$

引用文献

- Chow SC, Wang SG. *Advanced Linear Models: Theory and Applications*. CRC Press, 1993.