線型推測論

第05回 σ^2の点推定量

2022/5/19

慶應義塾大学病院

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
  - $\mathrm{E}[\mathbf{Y}] = \mathbf{X}\boldsymbol{\beta}$
- もう一つ未知パラメータがありました
  - $\mathrm{Cov}[\mathbf{Y}] = \sigma^2\mathbf{I}$

分散σ^2の推定量

- 分散σ^2も重要なパラメータである
  - $\mathrm{Cov}\left[\hat{\boldsymbol{\beta}}\right] = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$
- 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$\mathbf{Y}$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
  - $\mathrm{Cov}[\mathbf{Y}] = \sigma^2 \mathbf{I}$

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

  - $\mathbf{x}_i' = \left(x_{i1}, \dots, x_{ij}, \dots, x_{ip}\right)$
- $\boldsymbol{\beta}$を$\widehat{\boldsymbol{\beta}}$ (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
  - $\mathrm{E}\left[\left\{Y_i - \mathbf{x}_i'\widehat{\boldsymbol{\beta}}\right\}^2\right] \Rightarrow \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \mathbf{x}_i'\widehat{\boldsymbol{\beta}}\right)^2$
- 分散のプラグイン推定量
  - $\hat{\sigma}^2 = \frac{1}{n}\left(\mathbf{Y} - \mathbf{X}\widehat{\boldsymbol{\beta}}\right)'\left(\mathbf{Y} - \mathbf{X}\widehat{\boldsymbol{\beta}}\right)$
  - この推定量が良い性質を持っていればうれしい・・・

期待値の標本平均への置き換え

- $\mathrm{E}[T_i]$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, $\mathrm{E}[T_i] = \theta$が成立するとき, 以下の推定量が得られる
  - $\mathrm{E}[T] = \mathrm{E}\left[\frac{1}{n}\sum_{i=1}^{n} T_i\right] = \theta$
  - $\mathrm{Var}[T] = \frac{1}{n^2}\mathrm{Var}[\sum_{i=1}^{n} T_i] = \frac{1}{n}\mathrm{Var}[T_i]$

プラグイン推定量の期待値

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

$$
\begin{aligned}
&\left(\mathbf{Y} - \mathbf{X}\widehat{\boldsymbol{\beta}}\right)'\left(\mathbf{Y} - \mathbf{X}\widehat{\boldsymbol{\beta}}\right) \\
&= \\
&= \\
&= \mathbf{Y}'\{\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\}\mathbf{Y}
\end{aligned}
$$

プラグイン推定量の期待値

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

$$
\begin{aligned}
&\mathrm{E}\left[\left(\mathbf{Y}-\mathbf{X}\widehat{\boldsymbol{\beta}}\right)'\left(\mathbf{Y}-\mathbf{X}\widehat{\boldsymbol{\beta}}\right)\right]\\
&=\mathrm{E}[\mathbf{Y}'\mathbf{A}\mathbf{Y}]\\
&=\mathrm{tr}(\mathbf{A}\mathrm{E}[\mathbf{Y}\mathbf{Y}'])\\
&=\mathrm{tr}(\mathbf{A}\sigma^2\mathbf{I}+\mathbf{A}\mathbf{E}[\mathbf{Y}]\mathbf{E}[\mathbf{Y}]')\\
&=\sigma^2\mathrm{tr}(\mathbf{A})+\mathrm{tr}(\mathbf{E}[\mathbf{Y}]'\mathbf{A}\mathbf{E}[\mathbf{Y}])
\end{aligned}
$$

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

プラグイン推定量の期待値

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

プラグイン推定量の期待値

- 系5-2

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

$$\because\ \mathrm{Cov}[\mathbf{Y}] = \mathrm{E}[(\mathbf{Y} - \mathrm{E}[\mathbf{Y}])(\mathbf{Y} - \mathrm{E}[\mathbf{Y}])']$$

- 系5-3：誤差の三条件のもとで

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

プラグイン推定量の期待値

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

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

- $\mathbf{X}$： $n\times p$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：$\mathrm{tr}(\mathbf{E[Y]'AE[Y]}) = ?$

$$
\begin{aligned}
&\mathrm{tr}(\mathbf{E[Y]'AE[Y]}) \\
&= \mathbf{E[Y]'AE[Y]} \\
&= \boldsymbol{\beta}'\mathbf{X}'\{\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\}\mathbf{X}\boldsymbol{\beta} \\
&= \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta} - \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}\boldsymbol{\beta} \\
&= 0
\end{aligned}
$$

プラグイン推定量の期待値

- 以上の結果から,

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

$$= \frac{1}{n}(n-p)\sigma^2$$

- プラグイン推定量は不偏ではなかった・・・
- しかし, 後一歩の見た目をしている

分散σ^2の不偏推定量s^2

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

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

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

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

分散σ^2の不偏推定量s^2

- 定理5-4：s^2は$\mathrm{E}\left[Y_i - \mathrm{E}[Y_i]\right]^4 = 3\sigma^4$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 $\mathrm{E}\left[Y_i - \mathrm{E}[Y_i]\right]^4 = 3\sigma^4$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の中で不偏性をみた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と良い
- 一般線型モデルにおける分散の不偏推定量は：$s^2 = \dfrac{\mathbf{Y}'\left\{\mathbf{I}-\mathbf{X}\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\right\}\mathbf{Y}}{n-p}$

引用文献

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