多変量解析
第 3 回: 回帰分析 (2) 最尤法

藤本 衡

2018 年 4 月 25 日

本日の内容

- 最小 2 乗パラメータの性質
- より一般的な推定方法：最尤法
- 最尤法と最小 2 乗法の関係

最小 2 乗法の解 (再掲)

$$\beta^* = \frac{s_{xy}}{s_x}, \quad \alpha^* = \bar{y} - \frac{s_{xy}}{s_x}\bar{x} \tag{1}$$

ただし
$$\begin{cases} s_x = \dfrac{1}{n}\sum(x_i - \bar{x}) = \dfrac{1}{n}\sum x_i - \bar{x} \\ \quad (x \text{ の分散}) \\ s_{xy} = \dfrac{1}{n}\sum(x_i - \bar{x})(y_i - \bar{y}) = \dfrac{1}{n}\sum x_i y_i - \bar{x}\bar{y} \\ \quad (x \text{ と } y \text{ の共分散}) \end{cases}$$

最小 2 乗法の解の性質

- **最小 2 乗法の解＝最小 2 乗推定量**
- **推定量の性質って何があったっけ？**
  - **一致推定量** (consistent estimator)
  - **不偏推定量** (unbiased estimator)
  - **有効推定量** (efficient estimator)
  - **最尤推定量** (maximum likelihood estimator)
  - etc.

一致性と不偏性

一致性 (consistency)

$$\forall \delta > 0, \quad \lim_{n \to \infty} \mathbb{P}\left(|\beta^* - \beta| < \delta\right) = 1 \quad \textbf{(確率収束)}$$

が成り立つとき β^* は一致推定量であるという

不偏性 (unbiasedness)

$$\mathbb{E}\left[\beta^*\right] = \beta$$

が成り立つとき β^* は不偏推定量であるという

α^*, β^* の分布

$$\beta^* = \frac{s_{xy}}{s_x} = \sum_{i=1}^{n} \frac{x_i - \bar{x}}{n s_x} y_i = \sum_{i=1}^{n} B_i y_i$$

$$\alpha^* = \bar{y} - \frac{s_{xy}}{s_x}\bar{x} = \sum_{i=1}^{n} \left\{ \frac{1}{n} - \frac{\bar{x}(x_i - \bar{x})}{n s_x} \right\} y_i = \sum_{i=1}^{n} A_i y_i$$

$$\sum B_i = 0, \quad \sum B_i x_i = 1, \quad \sum B_i^2 = \frac{1}{n s_x}$$

$$\sum A_i = \boxed{\text{計算して!}}, \quad \sum A_i x_i = \boxed{\text{計算して!}}, \quad \sum A_i^2 = \boxed{\text{計算して!}}$$

α^*, β^* の分布 (2)

y_i を回帰モデルによって置き換えると

$$\beta^* = \sum_{i=1}^{n} B_i(\alpha + \beta x_i + \varepsilon_i) = \beta + \sum_{i=1}^{n} B_i \varepsilon_i$$

$$\alpha^* = \sum_{i=1}^{n} A_i(\alpha + \beta x_i + \varepsilon_i) = \alpha + \sum_{i=1}^{n} A_i \varepsilon_i$$

期待値の線形性より、$\mathbb{E}[\varepsilon_i] = 0$ ならば以下が成り立つ

$$\mathbb{E}[\beta^*] = \beta, \quad \mathbb{E}[\alpha^*] = \alpha$$ (不偏性の証明)

$\varepsilon \sim N(0, \sigma^2)$（正規分布に従う）と仮定すると、
α^*, β^* はともに正規分布に従う

α^*, β^* の分布 (3)

残差の分散が一定であり ($\mathbb{V}[\varepsilon_i] = \sigma^2$) かつ残差が無相関であるという仮定が成り立つとき、

$$\mathbb{V}[\beta^*] = \mathbb{V}[\varepsilon] \sum_i B_i^2 = \frac{\sigma^2}{n s_x}$$

$$\mathbb{V}[\alpha^*] = \mathbb{V}[\varepsilon] \sum_i A_i^2 = \sigma^2 \left\{ \frac{1}{n} + \frac{\bar{x}^2}{n s_x} \right\}$$

ともに $n \to \infty$ で 0 に収束するので、チェビシェフの不等式より

$$\beta^* \xrightarrow{p} \beta, \quad \alpha^* \xrightarrow{p} \alpha$$ (一致性の証明)

BLUE (Best Linear Unbiased Estimator)

- **有効推定量**: **不偏推定量のうち分散最小のもの**
- **線形推定量**: **確率変数** (ε_i) **の線形関数として与えられる推定量**
- **最小** 2 **乗推定量は線形不偏推定量の中で分散最小である（有効推定量）**
  $\Longrightarrow$ **最良線形不偏推定量** (BLUE)
- **(証明は面倒なので省略)**

y_i の条件付き密度関数

$\varepsilon_i \sim N(0, \sigma^2)$ **（正規分布に従う）と仮定すると、**
$x_1, \dots, x_n$ **が与えられたときの** $y_1, \dots, y_n$ **の条件付き同時密度関数は**

$$
\begin{aligned}
& f(y_1, \dots, y_n | x_1, \dots, x_n;\ \sigma^2, \alpha, \beta) \\
&= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{ -\frac{1}{2} \left(\frac{\varepsilon_i}{\sigma} \right)^2 \right\} \\
&= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{ -\frac{1}{2\sigma^2} (y_i - \alpha - \beta x_i)^2 \right\}
\end{aligned}
$$

最尤推定法 (maximum likelihood estimation: MSE)

y_i の条件付き同時密度関数は、未知パラメータ σ^2, α, β の関数と解釈することもできる

$$
\begin{aligned}
&L(\sigma^2, \alpha, \beta | y_1, \dots, y_n, x_1, \dots, x_n) \\
&= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{ -\frac{1}{2\sigma^2} (y_i - \alpha - \beta x_i)^2 \right\}
\end{aligned}
$$

- 与えられたデータで尤度関数を最大にするような σ^2, α, β を求めよう
  $\Longrightarrow$ 最尤推定法

対数尤度関数 (log-likelihood function)

- 尤度関数が積表現→微分がめんどい
- 対数をとっても極大点は変わらない
- 最尤法では対数尤度関数を使うことが多い

対数尤度関数

$$\log L(\sigma^2, \alpha, \beta | y_1, \ldots, y_n, x_1, \ldots, x_n) = \boxed{\text{計算してみて!}}$$

最尤法と最小 2 乗法

- 尤度関数の最大化 $\Longleftrightarrow$ 残差平方和の最小化
- あれ？最小 2 乗法と同じ？
- そうです、残差が正規分布に従う場合は最小 2 乗法は最尤推定をしていることになります
- 残差の分布が正規分布でない場合は結果は異なる
  - 最尤法はより一般的（広く適用できる）と言える