CMT-2

quarta-feira, 12 de abril de 2023 17:14

$E[Y_i] =: \mu_i(\beta) = \tilde{x}_i^T \beta$

$\tilde{x}_i$: i-ésima linha de $\tilde{X}$

β : $p+1 \times 1$

1. Escreva a log-verossimilhança e deduza seu gradiente e a sua derivada segunda (hessiana);

$$f(y_i \mid \beta, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(\tilde{x}_i\beta - y_i)^2\right)$$

$$f_X(Y \mid \beta, \sigma^2) = \prod^{m} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(\tilde{x}_i\beta - Y)^2\right)$$

$$L(\beta, \sigma^2 \mid Y) = \frac{1}{(2\pi\sigma^2)^{m/2}} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{m}(\tilde{x}_i\beta - Y)^2\right) = \frac{1}{(2\pi\sigma^2)^{m/2}} \exp\left(-\frac{1}{2\sigma^2}\|\tilde{X}\beta - Y\|^2\right)$$

$$\log\left(f_X(Y \mid \beta, \sigma^2)\right) = -\frac{\|\tilde{X}\beta - Y\|^2}{2\sigma^2} - \frac{m}{2}\log(2\pi\sigma^2)$$

$$= (\tilde{X}\beta - Y)^T(\tilde{X}\beta - Y)$$

Gradiente:

$$\frac{\partial l}{\partial \beta} = \frac{-2\tilde{X}^T \cdot (\tilde{X}\beta - Y)}{2\sigma^2} = \frac{-\tilde{X}^T(\tilde{X}\beta - Y)}{\sigma^2}$$

Hessiana:

$$H = \frac{\partial^2 l}{\partial \beta} = \frac{-\tilde{X}^T X}{\sigma^2}$$

2. Com base nos cálculos do item anterior, mostre a forma do estimador de máxima verossimilhança para $\boldsymbol{\beta}$, $\hat{\boldsymbol{\beta}}$;

$$\frac{\partial l}{\partial \beta} = 0$$

$$\frac{-\tilde{X}^T(\tilde{X}\beta - Y)}{\sigma^2} = 0$$

$$-\tilde{X}^T\tilde{X}\beta + \tilde{X}^T Y = 0$$

$$\tilde{X}^T X \beta = \tilde{X}^T Y$$

$$\hat{\beta} = (\tilde{X}^T X)^{-1}\tilde{X}^T Y$$

3. Mostre que $\hat{\boldsymbol{\beta}}$ é não-viesado;

queremos $E[\hat{\beta}] = \beta$, $E[Y] = E[\tilde{X}\beta + \epsilon] = \tilde{X}\beta$

$$E\left[(\tilde{X}^T\tilde{X})^{-1}\tilde{X}^T Y\right] = \underbrace{(\tilde{X}^T\tilde{X})^{-1}\tilde{X}^T\tilde{X}}_{I}\beta = \beta$$

4. Considere um outro estimador não-viesado de $\boldsymbol{\beta}$: $\tilde{\boldsymbol{\beta}} = \boldsymbol{M}\boldsymbol{y}$, onde

$$\boldsymbol{M} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T + \boldsymbol{D},$$

e $\boldsymbol{D}$ é uma matriz $P \times n$ cujas entradas são não-zero.

Mostre que $\boldsymbol{R} := \mathrm{Var}(\tilde{\boldsymbol{\beta}}) - \mathrm{Var}(\hat{\boldsymbol{\beta}})$ é positiva-definida.

Dica: Compute $E[\tilde{\boldsymbol{\beta}}]$ e considere o que deve valer para $\boldsymbol{D}$ sob a premissa de que $\tilde{\boldsymbol{\beta}}$ é não-viesado.

$\tilde{\beta}$ não-viesado: $E[\tilde{\beta}] = \beta$, $E[Y] = E[\tilde{X}\beta + \epsilon] = \tilde{X}\beta$

$$E\left[\left((X^TX)^{-1}X^T + D\right)y\right] \quad E\left[(X^TX)^{-1}X^Ty + Dy\right] = \beta$$

$$\underbrace{(X^TX)^{-1}X^T\tilde{X}}_{I}\beta + D\tilde{X}\beta = \beta$$

$$\beta + D\tilde{X}\beta = \beta$$

$$D\tilde{X}\beta = 0$$

$$D\tilde{X} = 0 \qquad \mathrm{Var}(Ay) = A\,\mathrm{Var}(y)\,A^T$$

$$\mathrm{Var}(\tilde{\beta}) = \mathrm{Var}\left(\left((X^TX)^{-1}X^T + D\right)y\right) = \left((X^TX)^{-1}X^T + D\right)\underbrace{\mathrm{var}(y)}_{\sigma^2}\left((X^TX)^{-1}X^T + D\right)^T$$

$$= \sigma^2\left((X^TX)^{-1}X^T + D\right)\left((X^TX)^{-1}X^T + D\right)^T$$

$$= \sigma^2\left((X^TX)^{-1}X^T + D\right)\left(\left((X^TX)^{-1}X^T\right)^T + D^T\right)$$

$$= \sigma^2\left((X^TX)^{-1}X^T + D\right)\left(X(X^TX)^{-1} + D^T\right)$$

$DX = 0$ $\qquad$ $DX = 0$

$$= \sigma^2\left(\underbrace{(X^TX)^{-1}X^TX}_{I}(X^TX)^{-1} + \cancel{(X^TX)^{-1}\underbrace{X^TD^T}_{(DX)^T}} + \cancel{DX(X^TX)^{-1}} + DD^T\right)$$

$$= \sigma^2(X^TX)^{-1} + \sigma^2(DD^T)$$

$$\mathrm{Var}(\hat{\beta}) = \mathrm{Var}\left((\tilde{X}^T\tilde{X})^{-1}X^TY\right) = \left((\tilde{X}^T\tilde{X})^{-1}X^T\right)\mathrm{var}(y)\left((\tilde{X}^T\tilde{X})^{-1}\tilde{X}^T\right)^T$$

$$= \sigma^2\left(\underbrace{(\tilde{X}^T\tilde{X})^{-1}X^T)(\tilde{X}}_{I}(\tilde{X}^T\tilde{X})^{-1}\right)$$

$$= \sigma^2(X^T\tilde{X})^{-1}$$

$$R = \mathrm{Var}(\tilde{\beta}) - \mathrm{Var}(\hat{\beta}) = \cancel{\sigma^2(X^TX)^{-1}} + \sigma^2(DD^T) - \cancel{\sigma^2(X^T\tilde{X})^{-1}}$$

$$= \sigma^2(\underbrace{DD^T}_{\text{positiva definida}}) \qquad XX^T \text{ positiva definida}$$