Week 3

Gauss–Markov模型、Aitken模型如何估计σ^2？σ^2的无偏估计和极大似然估计是什么？ 此外，为何Theorem 5.3 中，$\hat{\sigma}^2 = \mathrm{SSE}/(N-r)$的分母是$N-r$？（闫引桥）

Lemma：$E(z)=\mu$，$Cov(z)=\Sigma$，则 $E(z^TAz) = \mu^TA\mu + tr(A\Sigma)$

1. Gauss-Markov model

$$\begin{cases} y = Xb + e. \\ E(e) = 0, \quad Cov(e) = \sigma^2 I_N \end{cases}$$

$$y = \underset{\substack{\uparrow \\ \hat{y} = P_x y}}{Xb} + \underset{\substack{\uparrow \\ \hat{e} = (I-P_x)y}}{(y - Xb)}$$

用 $SSE = \hat{e}^T\hat{e} = \|(I-P_x)y\|^2$ 估计 σ^2

$$\begin{aligned} E(SSE) &= E(y^T(I-P_x)y) \\ &= b^TX^T(I-P_x)Xb + tr((I-P_x)\sigma^2 I_N) \\ &\qquad \downarrow Xb \in C(X) \qquad\qquad \downarrow (*) \\ &= 0 + \sigma^2(N-r) \end{aligned}$$

(*)：① For idempotent matrix P，$tr(P) = rank(P)$

（幂等矩阵的特征值为 0 或 1，$tr(P) = \Sigma\lambda_i$ = 特征值为1的个数 $= rank(P)$）

② $tr(I-P_x) = N - tr(P_x) = N - rank(P_x) = N - r$

（$rank(A^-A) = rank(A)$）

则 $\frac{SSE}{N-r}$ 是 σ^2 的无偏估计

2. Aitken model

$$\begin{cases} y = Xb + e \\ E(e) = 0, \quad Cov(e) = \sigma^2 V \quad (V \text{ 正定}) \end{cases}$$

Find R : $\underline{RVR^T = I}$ ，则有

$$\begin{cases} \underline{z} = Ry = RXb + Re =: \underline{Ub + f} \\ E(f) = 0, \quad Cov(f) = \sigma^2 I \end{cases} \longrightarrow \hat{b}_{GLS}$$

$$z = Ub + f$$

$$\hat{z} = \underline{P_U z} \qquad \hat{f} = (I - P_U) z$$

$$\underline{SSE = \hat{f}^T \hat{f}} = \| z - U\hat{b}_{GLS} \|^2 \qquad SSE = \| y - X\hat{b}_{OLS} \|^2$$

$$\begin{aligned} E(SSE) &= E(z^T (I - P_U) z) \\ &= tr((I - P_U)\sigma^2 I) \\ &= \sigma^2 [N - tr(U(U^TU)^- U^T)] \\ &= \sigma^2 [N - rank(U^TU)] \\ &= \sigma^2 [N - rank(U)] \qquad U = RX \\ &= \sigma^2 [N - rank(RX)] \\ &= \sigma^2 (N - r) \qquad (R \text{ 可逆}) \end{aligned}$$

$s^2 = \frac{1}{n-1} \Sigma (x_i - \bar{x})^2$

$\bar{x} \to$ 估计 μ ，μ 是1个

$\therefore n-1$

$\frac{SSE}{N-r}$ σ^2

$\downarrow$

自由度

$\hat{b} \to$ 估计 b (p 个)，

$\therefore$ 当 X full-column rank 时

$\frac{1}{N-p} SSE$

$\hat{\sigma}^2$ 的 MLE: Ch 6 P127

① Aitken model

$$\begin{cases} \underset{N\times 1}{y} = \underset{N\times p}{X}\,\underset{p\times 1}{b} + \underset{N\times 1}{e} \\ e \sim N(0, \sigma^2 \underset{N\times N}{V}) \end{cases}$$

则 $y \sim N(Xb, \sigma^2 V)$

则 likelihood is

$$L(b, \sigma^2) = \text{normal}(y \mid Xb, \sigma^2 V)$$

$$= (2\pi)^{-\frac{N}{2}} |\sigma^2 V|^{-\frac{1}{2}} \exp\left\{ -\frac{1}{2} (y - Xb)^T (\sigma^2 V)^{-1} (y - Xb) \right\}$$

$$\ell(b, \sigma^2) = -\frac{N}{2} \log \sigma^2 - \frac{1}{2\sigma^2} \underbrace{(y - Xb)^T V^{-1} (y - Xb)}_{= Q(b)} \quad = (z - Ub)^T (z - Ub)$$

$$\begin{cases} \dfrac{\partial \ell}{\partial b} = \dfrac{1}{2\sigma^2} (y - Xb)^T (V^{-1} + V^{-T}) X \\ \left(\dfrac{\partial u^T A v}{\partial x} = u^T A \dfrac{\partial v}{\partial x} + v^T A^T \dfrac{\partial u}{\partial x}, \quad \dfrac{\partial Ax}{\partial x} = A \right) \\ \dfrac{\partial \ell}{\partial \sigma^2} = -\dfrac{N}{2\sigma^2} + \dfrac{1}{2(\sigma^2)^2} Q(b) \end{cases}$$

令 $\dfrac{\partial \ell}{\partial b} = 0 \Rightarrow \underbrace{X^T (V^{-1} + V^{-T}) X}_{\text{rank} = r}\, b = X^T (V^{-1} + V^{-T}) y \iff U^T U b = U^T z$

$$\Rightarrow \hat{b} = [X^T (V^{-1} + V^{-T}) X]^g X^T (V^{-1} + V^{-T}) y$$

$$\left(V = I: \quad \hat{b} = (X^T X)^g X^T y = \hat{b}_{OLS} \right)$$

令 $\dfrac{\partial \ell}{\partial \sigma^2} = 0 \Rightarrow \hat{\sigma}^2 = \dfrac{1}{N} Q(\hat{b}) = \dfrac{1}{N} SSE$

② 简单情形

$$X_1, \cdots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$$

$$\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x_i-\mu)^2\right)$$

log-likelihood

$$l(\theta) = -\frac{n}{2}\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

$$\begin{cases} \dfrac{\partial l}{\partial \mu} = \dfrac{1}{\sigma^2}\displaystyle\sum_{i=1}^{n}(x_i-\mu) & \Rightarrow \quad \hat{\mu} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n}x_i = \bar{x} \\ \dfrac{\partial l}{\partial \sigma^2} = -\dfrac{n}{2\sigma^2} + \dfrac{1}{2(\sigma^2)^2}\displaystyle\sum_{i=1}^{n}(x_i-\mu)^2 & \Rightarrow \quad \hat{\sigma}^2 = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n}(x_i-\bar{x})^2 \end{cases}$$