作业 **9**

1. 假设 $y_1,...,y_n$ iid $\sim(\mu,\sigma^2)$, μ,σ^2 是未知参数。记样本均值和样本方差分别为 $\bar{y}$ 和 s^2。令 $\mathbf{y}=(y_1,...,y_n)^\top$, $\boldsymbol{\epsilon}=(\epsilon_1,...,\epsilon_n)^\top$, $\mathbb{1}=(1,...,1)^\top$, 则上述模型可表示为矩阵向量形式

$$\mathbf{y}=\mathbb{1}\mu+\boldsymbol{\epsilon},\ \boldsymbol{\epsilon}\sim(0,\sigma^2 I_n).$$

求 $\mathbf{y}$ 在 $C(\mathbb{1})$ 上的投影 $\hat{\mathbf{y}}$ 及 μ 的 LS 估计，以及 σ^2 的 LS 估计. （提示：$P_{\mathbb{1}}\mathbf{y}=\mathbb{1}\hat{\mu}$)

2. 假设随机样本数据 $(x_i,y_i), i=1,...,n$ 满足简单线性回归模型

$$y_i=a+bx_i+\epsilon_i,\epsilon_i,i=1,...,\text{iid}\sim N(0,\sigma^2),\text{且 } x_i \text{ 与 } \epsilon_i \text{ 独立},i=1,...,n.$$

将模型写成矩阵-向量的形式

$$\mathbf{y}=X\boldsymbol{\beta}+\boldsymbol{\epsilon}=\mathbb{1}a+\mathbf{x}b+\boldsymbol{\epsilon}$$

其中设计阵 $X=(\mathbb{1},\mathbf{x})$, $\mathbf{x}=(x_1,...,x_n)^\top$, $\boldsymbol{\beta}=(a,b)^\top$.

(a) 记 $\bar{x}=(x_1+...+x_n)/n=\mathbf{x}^\top\mathbb{1}/n$. 证明 $C(X)$ 的投影矩阵

$$P_X=\frac{\mathbb{1}\mathbb{1}^\top}{n}+\frac{(\mathbf{x}-\mathbb{1}\bar{x})(\mathbf{x}-\mathbb{1}\bar{x})^\top}{\|\mathbf{x}-\mathbb{1}\bar{x}\|^2}.$$

(提示：先将 $\mathbf{x}$ 和 $\mathbb{1}$ 正交化).

(b) 求 $\mathbf{y}$ 在 $C(X)$ 上的投影 $\hat{\mathbf{y}}=P_X\mathbf{y}$。投影表达式决定了参数 a,b 的 LS 估计: 若 $\hat{\mathbf{y}}=\mathbb{1}\xi+\mathbf{x}\eta$，则 $\mathbb{1},\mathbf{x}$ 的系数即是 LS 估计 $\hat{a}=\xi,\hat{b}=\eta$，试由 $\hat{\mathbf{y}}$ 的表达式求出 $\hat{a},\hat{b}$.

3. 假设线性模型 $y_i=\beta_0+\mathbf{x}_i^\top\mathbf{b}+\epsilon_i$, $i=1,2,...,n$, 其矩阵-向量形式为

$$\mathbf{y}_{n\times 1}=X_{n\times p}\boldsymbol{\beta}_{p\times 1}+\boldsymbol{\epsilon}=\mathbb{1}\beta_0+Z\boldsymbol{b}+\boldsymbol{\epsilon}, \tag{1}$$

其中 $\boldsymbol{\beta}=(\beta_0,\mathbf{b}^\top)^\top$, $X=(\mathbb{1},Z)$, $Z=(\mathbf{x}_1,...,\mathbf{x}_n)^\top$ 为所有自变量按行排成的 $n\times(p-1)$ 矩阵。

(a) 假设 $Z^\top\mathbb{1}=0$ （即 Z 的各个列向量都与 $\mathbb{1}$ 正交，每列之和为 0)，试利用 LS 估计公式 $\hat{\boldsymbol{\beta}}=(X^\top X)^- X^\top\mathbf{y}$ 证明 $\mathbf{b}$ 的 LS 估计

$$\hat{\mathbf{b}}=(Z^\top Z)^- Z^\top\mathbf{y}.$$

(b) 试利用投影 $P_X\mathbf{y}=P_{\mathbb{1}}\mathbf{y}+P_Z\mathbf{y}=\mathbb{1}\hat{\beta}_0+Z\hat{\mathbf{b}}$, 说明 $\hat{\mathbf{b}}=(Z^\top Z)^- Z^\top\mathbf{y}$.

(c) 当 $Z^\top\mathbb{1}\neq 0$ 时，记 Z 的中心化矩阵 $Z_c=Z-P_{\mathbb{1}}Z=Z-\mathbb{1}\bar{\mathbf{x}}^\top$，其中 $\bar{\mathbf{x}}=Z^\top\mathbb{1}/n=(\mathbf{x}_1+...+\mathbf{x}_n)/n$ 为自变量的样本均值。试利用 LS 估计的表达 $\hat{\boldsymbol{\beta}}=(\hat{\beta}_0,\hat{\mathbf{b}}^\top)^\top=(X^\top X)^{-1}X^\top\mathbf{y}$ 和分块矩阵求逆公式（第 4 讲命题 1，P12）证明

$$\hat{\mathbf{b}}=(Z_c^\top Z_c)^{-1}Z_c^\top\mathbf{y}$$

(d) 将 $Z = Z_c + \mathbb{1}\bar{\mathbf{x}}^\top$ 代入模型（1）

$$\mathbf{y} = \mathbb{1}\beta_0 + (Z_c + \mathbb{1}\bar{\mathbf{x}}^\top)\mathbf{b} + \boldsymbol{\epsilon} = \mathbb{1}\beta_0^* + Z_c\mathbf{b} + \boldsymbol{\epsilon} \tag{2}$$

其中 $\beta_0^* = \beta_0 + \bar{\mathbf{x}}^\top\mathbf{b}$, 注意 $Z_c^\top\mathbb{1} = 0$, 因此投影 $\hat{\mathbf{y}} = P_X\mathbf{y} = P_{\mathbb{1},Z_c}\mathbf{y} = P_{\mathbb{1}}\mathbf{y} + P_{Z_c}\mathbf{y}$，根据该表达说明 $\mathbf{b}$ 的最小二乘估计为 $\hat{\mathbf{b}} = (Z_c^\top Z_c)^{-1}Z_c^\top\mathbf{y}$.

(e) 记

$$S_{\mathbf{xx}} = \frac{1}{n-1}\sum_{i=1}^{n}(\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^\top, S_{\mathbf{x}y} = \frac{1}{n-1}\sum_{i=1}^{n}(\mathbf{x}_i - \bar{\mathbf{x}})y_i$$

分别为样本方差和样本协方差矩阵, 验证 (c) 或 (d) 中的 LS 估计 $\hat{\mathbf{b}} = S_{\mathbf{xx}}^{-1}S_{\mathbf{x}y}$.

4. 下面叙述的是简单回归模型中求解工具变量最小二乘估计的方法步骤：

假设随机变量 x, y 满足总体模型 $y = a + bx + \epsilon, \epsilon \sim (0, \sigma^2)$, 但其中 x 与 ϵ 不独立 (内生)。假设存在一个随机变量 z，它与 ϵ 独立 (外生)，特别地，$cov(\epsilon, z) = 0$, 即

$$0 = cov(\epsilon, z) = cov(y - a - bx, z) = \Sigma_{yz} - b\Sigma_{xz} \Rightarrow b = \Sigma_{yz}/\Sigma_{xz}$$

假设样本数据为 $(y_i, x_i, z_i), i = 1, ..., n$，为了估计 b 我们应用矩估计方法，在上述表达中代入样本协方差

$$S_{yz} = \sum_{i=1}^{n}(z_i - \bar{z})(y_i - \bar{y})/(n-1), S_{xz} = \sum_{i=1}^{n}(z_i - \bar{z})(x_i - \bar{x})/(n-1),$$

即得 b 的估计

$$\tilde{b} = S_{yz}/S_{xz} = \frac{\sum_{i=1}^{n}(z_i - \bar{z})(y_i - \bar{y})}{\sum_{i=1}^{n}(z_i - \bar{z})(x_i - \bar{x})}$$

该估计称为工具变量最小二乘估计。可以证明该估计是 b 的渐近无偏估计（此处略去）。

从 b 或 $\tilde{b}$ 的表达式来看，上述讨论中对于 z 遗漏了什么条件？