4.9 HW 9

作业 9 链接

练习 4.1 假设 $y_1,\ldots,y_n$iid $\sim\left(\mu,\sigma^2\right),\mu,\sigma^2$ 是未知参数。记样本均值和样本方差分别为 $\bar{y}$ 和 s^2 。令 $\mathbf{y}=\left(y_1,\ldots,y_n\right)^{\top},\boldsymbol{\epsilon}=\left(\epsilon_1,\ldots,\epsilon_n\right)^{\top},\mathbb{1}=(1,\ldots,1)^{\top}$, 则上述模型可表示为矩阵向量形式

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

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: $P_{\mathbb{1}}\mathbf{y}=\mathbb{1}\hat{\mu}$)

解 利用投影矩阵:

$$\hat{y}=P_{\mathbb{1}}(y)=\mathbb{1}\left(\mathbb{1}^{\mathbf{T}}\mathbb{1}\right)^{-}\mathbb{1}^{\mathbf{T}}y=\mathbb{1}\bar{y}.$$

已知 $P_{\mathbb{1}}\boldsymbol{y}=\mathbb{1}\hat{\mu}$, 因此 μ 的 *LS* 估计为 $\bar{y}$ 。方差的 *LS* 估计为:

$$\hat{\sigma^2}=\frac{RSS}{n-p}=\frac{\sum_{i=1}^n(y-\bar{y})^2}{n-1}=s^2.$$

练习 4.2 假设随机样本数据 $\left(x_i,y_i\right),i=1,\ldots,n$ 满足简单线性回归模型

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

将模型写成矩阵-向量的形式

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

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

(a) 记 $\bar{x}=\left(x_1+\ldots+x_n\right)/n=\mathbf{x}^{\top}\mathbb{1}/n$. 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 $\hat{a},\hat{b}$.

解

1. 先将 x 正交化:

$$\boldsymbol{x}^{\perp}=x-P_{\mathbb{1}}x=x-\mathbb{1}\bar{x}$$

已知 $P_x=P_{\mathbb{1}}+P_{x^{\perp}}$, 因此:

$$P_x=P_{\mathbb{1}}+P_{x^{\perp}}=\frac{\mathbf{1}\mathbf{1}^{\boldsymbol{T}}}{n}+\frac{(\boldsymbol{x}-\mathbb{1}\bar{x})(\boldsymbol{x}-\mathbb{1}\bar{x})^T}{\|\boldsymbol{x}-\mathbb{1}\bar{x}\|^2}$$

2.

$$\begin{aligned}\hat{y}=P_x(y)&=\left(\frac{\mathbf{1}\mathbf{1}^{\boldsymbol{T}}}{n}+\frac{(\boldsymbol{x}-\mathbb{1}\bar{x})(\boldsymbol{x}-\mathbb{1}\bar{x})^T}{\|\boldsymbol{x}-\mathbb{1}\bar{x}\|^2}\right)y\\&=\mathbb{1}\left(\frac{\mathbb{1}^{\boldsymbol{T}}y}{n}-\bar{x}\frac{(\boldsymbol{x}-\mathbb{1}\bar{x})^Ty}{\|\boldsymbol{x}-\mathbb{1}\bar{x}\|^2}\right)+\boldsymbol{x}\frac{(\boldsymbol{x}-\mathbb{1}\bar{x})^Ty}{\|\boldsymbol{x}-\mathbb{1}\bar{x}\|^2}\end{aligned}$$

若 $\hat{\boldsymbol{y}}=\mathbb{1}\xi+\boldsymbol{x}\eta$, 则:

$$\hat{a} = \xi = \frac{\mathbb{1}^T y}{n} - \bar{x}\frac{(\boldsymbol{x} - \mathbb{1}\bar{x})^T y}{\|\boldsymbol{x} - \mathbb{1}\bar{x}\|^2} = \bar{y} - \bar{x}\eta$$
$$\hat{b} = \eta = \frac{(\boldsymbol{x} - \mathbb{1}\bar{x})^T y}{\|\boldsymbol{x} - \mathbb{1}\bar{x}\|^2} = \frac{S_{yy}}{S_{xx}}$$

练习 4.3 假设线性模型 $y_i = \beta_0 + \mathbf{x}_i^\top \mathbf{b} + \epsilon_i, i = 1, 2, \ldots, n$, 其矩阵-向量形式为

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

其中 $\beta = \left(\beta_0, \mathbf{b}^\top\right)^\top, X = (\mathbb{1}, Z), Z = (\mathbf{x}_1, \ldots, \mathbf{x}_n)^\top$ 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 $\hat{\beta} = \left(X^\top X\right)^- X^\top \mathbf{y}$ 证明 $\mathbf{b}$ 的 LS 估计

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

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

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

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

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

$$\mathbf{y} = \mathbb{1}\beta_0 + \left(Z_c + \mathbb{1}\overline{\mathbf{x}}^\top\right)\mathbf{b} + \boldsymbol{\epsilon} = \mathbb{1}\beta_0^* + Z_c\mathbf{b} + \boldsymbol{\epsilon}$$

其中 $\beta_0^* = \beta_0 + \overline{\mathbf{x}}^\top \mathbf{b}$, 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 $\hat{\mathbf{b}} = \left(Z_c^\top Z_c\right)^{-1} Z_c^\top \mathbf{y}$.

(e) 记

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

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 $\hat{\mathbf{b}} = S_{\mathbf{xx}}^{-1} S_{\mathbf{x}y}$.

证明

1. 将 $\hat{\beta} = \left(X^T X\right)^- X^T y$ 展开,

$$\begin{pmatrix} \hat{\beta}_0 \\ \hat{b} \end{pmatrix} = \begin{pmatrix} \mathbb{1}^T\mathbb{1} & \mathbb{1}^T Z \\ Z^T\mathbb{1} & Z^T Z \end{pmatrix}^- \begin{pmatrix} \mathbb{1}^T \\ Z^T \end{pmatrix} y$$

且我们知道 $Z^T\mathbb{1} = 0$, 上式化简为:

$$\begin{pmatrix} \hat{\beta}_0 \\ \hat{b} \end{pmatrix} = \begin{pmatrix} \mathbb{1}^T\mathbb{1} & 0 \\ 0 & Z^T Z \end{pmatrix}^- \begin{pmatrix} \mathbb{1}^T \\ Z^T \end{pmatrix} y = \begin{pmatrix} \frac{\mathbb{1}^T y}{n} \\ \left(Z^T Z\right)^- Z^T y \end{pmatrix}$$

即得 $\hat{b}$ 的 LS 估计。

2. 已知 $Z^T\mathbb{1} = 0$, 所以

$$P_x = P_{\mathbb{1}} + P_Z = \frac{\mathbf{1}\mathbf{1}^T}{n} + Z\left(Z^T Z\right)^- Z^T$$

因此:

$$P_X\boldsymbol{y}=P_{\mathbb{1}}+P_Z(y)=\mathbb{1}\bar{y}+Z\left(Z^TZ\right)^{-}Z^Ty=\mathbb{1}\hat{\beta}_0+Z\hat{b}$$

于是 $Z\left(\left(Z^TZ\right)^{-}Z^Ty-\hat{b}\right)=0$, 由于样本彼此应不同, 矩阵列满秩, 故: $\hat{b}=\left(Z^TZ\right)^{-}Z^Ty$.

3.

$$\begin{aligned}\begin{pmatrix}\hat{\beta}_0\\ \hat{b}\end{pmatrix}&=\begin{pmatrix}\mathbb{1}^{\boldsymbol{T}}\mathbb{1} & \mathbb{1}^{\boldsymbol{T}}Z\\ Z^T\mathbb{1} & Z^TZ\end{pmatrix}^{-1}\begin{pmatrix}\mathbb{1}^{\boldsymbol{T}}\\ Z^T\end{pmatrix}y\\ &=\begin{pmatrix} * & * \\ -\left(Z^TZ-n\bar{x}\bar{x}^T\right)^{-1}\bar{x} & \left(Z^TZ-n\bar{x}\bar{x}^T\right)^{-1}\end{pmatrix}\begin{pmatrix}\mathbb{1}^{\boldsymbol{T}}\\ Z^T\end{pmatrix}y\\ &=\begin{pmatrix} * \\ \left(Z^TZ-n\bar{x}\bar{x}^T\right)^{-1}\left(Z^T-\bar{x}\mathbb{1}^T\right)y\end{pmatrix}\end{aligned}$$

同时有:

$$\left(Z_c^TZ_c\right)^{-1}Z_c^T=\left(Z^TZ-n\bar{x}\bar{x}^T\right)^{-1}\left(Z^T-\bar{x}\mathbb{1}^T\right)$$

因此: $\hat{b}=\left(Z_c^TZ_c\right)^{-1}Z_c^Ty$.

4. 利用中心化, 我们得到 $Z_c^T\mathbb{1}=0$. 与第 2 问相同, 即得:

$$\hat{b}=\left(Z_c^TZ_c\right)^{-1}Z_c^Ty$$

5. 事实上, $Z_c^TZ_c=\sum\left(x_i-\bar{x}\right)\left(x_i-\bar{x}\right)^T=(n-1)S_{xx}, Z_c^Ty=(n-1)S_{xy}$. 故: $\hat{b}=\left(Z_c^TZ_c\right)^{-1}Z_c^Ty=S_{xx}^{-1}S_{xy}$.

注 分块矩阵求逆公式，参见 Week04 第 12 页。

练习 4.4 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 $y=a+bx+\epsilon,\epsilon\sim\left(0,\sigma^2\right)$, 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, $\operatorname{cov}(\epsilon,z)=0$, 即

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

假设样本数据为 $\left(y_i,x_i,z_i\right),i=1,\ldots,n$, 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

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

即得 b 的估计

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

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?

解 还需要 z 与 x 相关。