

# 计量经济学: 第三次作业

## 主讲教师: 崔文昊

### 截止日期: 2022 年 11 月 3 日

1. 请在指定日期当天课前提交作业。
2. 请用 A4 纸完成作业。如果作业为多页, 请于 A4 纸左上角装订。
3. 使用 Matlab 完成的作业, 请将代码, 连同结果一并打印提交。

## 第一题 (40 分)

考虑多元线性回归模型  $Y_i = \alpha + \beta_2 X_{2,i} + \cdots + \beta_K X_{K,i} + u_i, i = 1, 2, \dots, N$ 。用矩阵和向量可以将上述模型写作  $Y = X\beta + U$ , 其中  $X$  是一个  $N \times K$  的矩阵。

我们将变量  $X$  进行分组:  $X = [X^{(1)}, X^{(2)}]$ , 其中  $X^{(1)}$  是一个  $N \times K_1$  的矩阵,  $X^{(2)}$  是一个  $N \times K_2$  的矩阵, 并且满足  $K = K_1 + K_2$ 。同样, 我们也可以将带估计参数  $\beta$  进行分组:  $\beta = [\beta^{(1)'}, \beta^{(2)'}]'$ 。那么, 就有  $X\beta = X^{(1)}\beta^{(1)} + X^{(2)}\beta^{(2)}$ 。

### 1.1 遗漏解释变量的后果 (15 分)

假设你用  $OLS$  估计了线性回归模型  $Y = X^{(1)}\beta^{(1)} + \epsilon$ , 将对应的估计量记作  $\tilde{\beta}^{(1)}$ 。当真实的模型是  $Y = X\beta + U$  时,  $\tilde{\beta}^{(1)}$  的期望和方差是多少? 在遗漏变量  $X^{(2)}$  的情况下,  $\tilde{\beta}^{(1)}$  是否可能是  $\beta^{(1)}$  的无偏估计量?

### 1.2 添加多余解释变量 (15 分)

假设你用  $OLS$  估计了线性回归模型  $Y = X\beta + U$ , 将对应的估计量记作  $\hat{\beta} = [\hat{\beta}^{(1)'}, \hat{\beta}^{(2)'}]'$ 。当真实的模型是  $Y = X^{(1)}\beta^{(1)} + \epsilon$  时 ( $\epsilon$  满足线性回归的所有基本假设),  $\hat{\beta}$  的期望和方差是多少? 提示:  $Y = X^{(1)}\beta^{(1)} + \epsilon = X \begin{bmatrix} \beta^{(1)} \\ 0 \end{bmatrix} + \epsilon$ 。

此外, 利用分块矩阵求逆的公式

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$
$$A^{-1} = \begin{bmatrix} A_{11}^{-1} + A_{11}^{-1}A_{12}FA_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}F \\ -FA_{21}A_{11}^{-1} & F \end{bmatrix},$$

其中  $F = (A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1}$ , 计算  $\hat{\beta}^{(1)}$  的方差。

### 1.3 多余解释变量与有效性 (10 分)

假设你用  $OLS$  估计了线性回归模型  $Y = X^{(1)}\beta^{(1)} + \epsilon$ , 并且该模型就是真实模型时。根据课堂的学习, 我们知道对应的估计量具有无偏性, 并且其方差为  $\sigma^2(X^{(1)'}X^{(1)})^{-1}$ 。将其与第二问之中计算的  $\hat{\beta}^{(1)}$  方差进行对比, 添加冗余变量是否会影响估计量的有效性? 提示: 已知  $[X^{(2)'}X^{(2)} - X^{(2)'}X^{(1)}(X^{(1)'}X^{(1)})^{-1}X^{(1)'}X^{(2)}]^{-1}$  为正定矩阵。

Variable	OLS estimate	t stat
EXP	-0.0020	-1.1433
WKS	-0.0178	-5.2092
OCC	0.3181	6.8523
IND	0.0300	0.7892
SOUTH	-0.1701	-4.2745
SMSA	0.0845	2.1974
MS	0.0990	1.5515
FEM	-0.1087	-1.3714
ED	-0.0162	-1.8839
BLK	0.0502	0.7057
Const	1.1959	5.2679

## 第二题 (20 分)

我们想要对员工是否参与工会进行回归分析研究，采用 1982 年 PSID 数据集。在 1982 年观察到的 595 人之中，218 人的工资由工会确定，另外 377 人的工资则不是由工会确认。在此回归之中，我们将统一采用以下解释变量：教育年限 (ED)，工作周数 (WKS)，全职工作年限 (EXP)，职业 (OCC = 1，如果个人从事蓝领职业)，居住地 (SOUTH = 1, SMSA = 1，如果个人居住在南方，或者在大都市区域)，工业 (IND = 1，如果个人在制造业行业)，婚姻状况 (MS = 1，如果个人已婚)，性别和种族 (FEM = 1, BLK = 1，如果此人是女性或黑人)。

数据已经按顺序读入 Matlab 之中，并且命名为  $X, Y$ 。你需要在 Matlab 之中做如下回归， $Y = X\beta + U$ ，并复制出表中的结果，你需要提交你的 code 和对应的回归结果。