Solutions for Homework 5

Tong Zhou

Problems: 7.2, 7.7, 7.20, 7.25.

PROBLEM 7.2

Proof. It is obvious that

$$\frac{1}{n}(\lambda I_k) \longrightarrow \mathbf{0}, \quad \text{as } n \longrightarrow \infty$$

for any constant λ.

By law of large numbers, therefore

$$\widehat{\beta} = \left(\frac{1}{n}\sum_{i=1}^{n} X_i X_i' + \frac{1}{n}\lambda I_k\right)^{-1} \left(\frac{1}{n}\sum_{i=1}^{n} X_i y_i\right) \xrightarrow{\mathsf{P}} (\mathbb{E}[X_i X_i'])^{-1}\mathbb{E}[X_i y_i] \equiv \beta.$$

Thus, $\widehat{\beta}$ is consistent for β. □

PROBLEM 7.7

Proof. (a) Recall that in problem 4.16 of homework 4, we define a composite error

$$\nu_i = e_i + u_i$$

for each i. Then the original linear regression can be rewritten as

$$y_i = X_i'\beta + \nu_i.$$

So β is the coefficient from the linear projection of y_i on X_i.

(b) By assumptions we have

$$\mathbb{E}[X_i \nu_i] = \mathbf{0}.$$

It follows that

$$\begin{aligned}
\widehat{\beta} &= \left(\frac{1}{n}\sum_{i=1}^{n} X_i X_i'\right)^{-1} \left(\frac{1}{n}\sum_{i=1}^{n} X_i y_i\right) \\
&= \beta + \left(\frac{1}{n}\sum_{i=1}^{n} X_i X_i'\right)^{-1} \left(\frac{1}{n}\sum_{i=1}^{n} X_i \nu_i\right) \\
&\xrightarrow{\mathsf{P}} \beta + \left(\mathbb{E}[X_i X_i']\right)^{-1} \mathbb{E}[X_i \nu_i] \\
&= \beta.
\end{aligned}$$

(c) From (b), we have

$$\sqrt{n}(\widehat{\beta} - \beta) = \left(\frac{1}{n}\sum_{i=1}^{n} X_i X_i'\right)^{-1} \frac{1}{\sqrt{n}}\sum_{i}^{n} X_i \nu_i \tag{1}$$

By weak law of large numbers,

$$\frac{1}{n}\sum_{i} X_i X_i' \xrightarrow{\mathsf{P}} \mathbb{E}[X_i X_i'] \equiv \mathbf{Q}_{XX} \tag{2}$$

By the central limit theorem,

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_i \nu_i \xrightarrow{\mathsf{d}} \mathcal{N}(\mathbf{0}, \mathbf{\Omega}_\nu) \tag{3}$$

where $\mathbf{\Omega}_\nu = \mathbb{E}[X_i X_i' \nu_i^2]$.
Plugging Eqs. (2) and (3) into Eq. (1), the Slustky theorem implies

$$\sqrt{n}(\widehat{\beta} - \beta) \xrightarrow{\mathsf{d}} \mathcal{N}\left(\mathbf{0}, \mathbf{Q}_{XX}^{-1}\mathbf{\Omega}_\nu \mathbf{Q}_{XX}^{-1}\right).$$

□

PROBLEM 7.20

PROOF. (a) $S(\boldsymbol{\beta})$ can be written as:

$$S(\boldsymbol{\beta}) = (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})'\boldsymbol{W}(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})$$

where $\boldsymbol{W} = \mathrm{diag}(w_1, w_2, \dots, w_n)$, $\boldsymbol{y}, \boldsymbol{X}$ and $\boldsymbol{\beta}$ conventional notations as in the OLS model. Then, it is easy to obtain:

$$\begin{aligned}\widehat{\beta} &= (\boldsymbol{X}'\boldsymbol{W}\boldsymbol{X})^{-1}(\boldsymbol{X}'\boldsymbol{W}\boldsymbol{y}) \\ &= \left(\sum_{i=1}^{n} x_i w_i x_i'\right)^{-1}\left(\sum_{i=1}^{n} x_i w_i y_i\right)\end{aligned}$$

(b) We maintain the following assumptions:

1. $w_i > 0, \forall i$
2. $\mathbb{E}(e_i | x_i, w_i) = 0, \forall i.$

A WLS model with heteroskedasiticity, in which $w_i = w_i(x_i)$ such as Exercise 4.8, is a special case of a WLS model considered here. However, their ideas are essentially the same: $\widehat{\boldsymbol{\beta}}$ is obtained by the OLS method, while the transformed regression is:

$$\boldsymbol{y}^* = \boldsymbol{X}^*\beta + \boldsymbol{e}^*.$$

where $\boldsymbol{y}^* = \boldsymbol{W}^{\frac{1}{2}}\boldsymbol{y}$, $\boldsymbol{X}^* = \boldsymbol{W}^{\frac{1}{2}}\boldsymbol{X}$, and $\boldsymbol{e}^* = \boldsymbol{W}^{\frac{1}{2}}\boldsymbol{e}$.

(c) Under certain regularity conditions and the stated assumptions, $\widehat{\boldsymbol{\beta}} \xrightarrow{P} \boldsymbol{\beta}$.

(d)

$$\sqrt{n}\left(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right) \xrightarrow{d} N(\mathbf{0}, \boldsymbol{V}_{\beta})$$

where $\boldsymbol{V}_{\beta} = \boldsymbol{Q}_{\boldsymbol{xwx}}^{-1} \Omega \boldsymbol{Q}_{\boldsymbol{xwx}}^{-1}$, in which $\boldsymbol{Q}_{\boldsymbol{xwx}} = \mathbb{E}(x_i w_i x_i')$, $\boldsymbol{\Omega} = \mathbb{E}(x_i x_i' e_i^2 w_i^2)$

□

PROBLEM 7.25

Proof. Obviously we have

$$\begin{aligned}
\widetilde{\boldsymbol{\beta}} &= \left(\sum_i W_i X_i X_i'\right)^{-1} \left(\sum_i W_i X_i Y_i\right) \\
&= \boldsymbol{\beta} + \left(\sum_i W_i X_i X_i'\right)^{-1} \left(\sum_i W_i X_i e_i\right)
\end{aligned}$$

Thus $\widetilde{\boldsymbol{\beta}}$ is consistent for $\boldsymbol{\beta}$ if and only if

$$\mathbb{E}[W_i X_i e_i] = \mathbf{0}.$$

However, the assumption $\mathbb{E}[X_i e_i] = \mathbf{0}$ cannot imply $\mathbb{E}[W_i X_i e_i] = \mathbf{0}$. Instead, a stronger condition $\mathbb{E}[e_i \mid X_i] = 0$ ensures the consistency of $\widetilde{\boldsymbol{\beta}}$. This can be seen by

$$\begin{aligned}
\mathbb{E}\left[W_i X_i e_i\right] &= \mathbb{E}[W_i X_i \underbrace{\mathbb{E}[e_i \mid X_i]}_{=\mathbf{0}}] \\
&= \mathbf{0}
\end{aligned}$$

In the above, we use the law of iterated expectations and the fact that $W_i = w(X_i)$. □