Econometrics I: Solutions of Homework #13

Hiroki Kato *

July 15, 2020

Contents

1 Solutions: Question 1

1.1 (1)

OLS estimator of β is given by

$$\hat{\beta} = (X'X)^{-1}X'y = \beta + (X'X)^{-1}X'u = \beta + \left(\frac{1}{T}\sum_{t=1}^{T} X_t X_t'\right)^{-1}\left(\frac{1}{T}\sum_{t=1}^{T} X_t u_t\right), \tag{1}$$

*e-mail: vge008kh@student.econ.osaka-u.ac.jp. Room 503. If you find any errors in handouts and materials, please contact me via e-mail.

where X_t is $k \times 1$ vector whose elements are explanatory variables for observetion t. Taking the probability limit to both sides yields

$$\operatorname*{plim}_{T\to\infty} \hat{\beta} = \beta + \lim_{T\to\infty} \left(\frac{1}{T} \sum_{t=1}^{T} X_t X_t' \right)^{-1} \operatorname*{plim}_{T\to\infty} \left(\frac{1}{T} \sum_{t=1}^{T} X_t u_t \right) \tag{2}$$

First, we assume the following stationarity condition:

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} X_t X_t' = M_{xx}. \tag{3}$$

Then, by $g(X_n) \xrightarrow{p} g(X)$ if $X_n \xrightarrow{p} X$, we have

$$\operatorname*{plim}_{T\to\infty} \left(\frac{1}{T} \sum_{t=1}^{T} X_t X_t' \right)^{-1} = M_{xx}^{-1}. \tag{4}$$

We further assume zero covariance between X and u, that is,

$$\operatorname*{plim}_{T\to\infty} \left(\frac{1}{T} \sum_{t=1}^{T} X_t u_t \right) = 0. \tag{5}$$

By these two assumptions, we finally obtain

$$\operatorname*{plim}_{T\to\infty} \hat{\beta} = \beta + M_{xx}^{-1} \cdot 0 = \beta. \tag{6}$$

We can show the consistency of OLSE.

1.2 (2)

Greenberg and Webster (1983) states the central limit theorem as follows:

> $Z_1, \ldots, Z_n$ are mutually independent. Z_i is distributed with mean μ and variance Σ_i for $i = 1, \ldots, n$. Then, we have the follwing result:
>
> $$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} (Z_i - \mu) \xrightarrow{d} N(0, \Sigma),$$
>
> where
>
> $$\Sigma = \lim_{n\to\infty} \left(\frac{1}{n} \sum_{i=1}^{n} \Sigma_i \right)$$

Using this theorem, we derive the asymptotic distribution of $\frac{1}{\sqrt{n}}X'u$. Define $Z_i = X_t u_t$. By the assumption of u_t, $\mu = E(Z_i) = 0$ and the variance of Z_i is given by

$$\Sigma_i = V(X_t u_t) = E[(X_t u_t)(X_t u_t)'] = \sigma^2 X_t X_t'. \tag{7}$$

Then, we obtain Σ as follows:

$$\Sigma = \lim_{T\to\infty}\left(\frac{1}{n}\sum_{t=1}^{T}\sigma^2 X_t X_t'\right) = \sigma^2 \lim_{T\to\infty}\left(\frac{1}{n}\sum_{t=1}^{T} X_t X_t'\right) = \sigma^2 M_{xx}. \tag{8}$$

By the central limit theorm, we have

$$\frac{1}{\sqrt{n}}X'u = \frac{1}{\sqrt{n}}\sum_{t=1}^{T}(X_t u_t - 0) \xrightarrow{d} N(0, \sigma^2 M_{xx}). \tag{9}$$

1.3 (3)

We rewrite the equation (1) as follows:

$$\sqrt{T}(\hat{\beta} - \beta) = \left(\frac{1}{T}\sum_{t=1}^{T} X_t X_t'\right)^{-1}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T} X_t u_t\right). \tag{10}$$

We use the following proerty: $z_n = H_n y_n \xrightarrow{d} N(H\mu, H\Omega H')$ where H_n is an $r \times k$ matrix with $\text{plim}_{n\to\infty} H_n = H$ and y_n is a $\times 1$ vector with $y_n \xrightarrow{d} N(\mu, \Omega)$. Because the equation (10) is seen as the form z_n, $\sqrt{T}(\hat{\beta} - \beta)$ has a limiting normal distribution with zero mean and variance given by

$$\sigma^2 M_{xx}^{-1} M_{xx} M_{xx}^{-1} = \sigma^2 M_{xx}^{-1}. \tag{11}$$

Thus, we can write

$$\sqrt{T}(\hat{\beta} - \beta) \xrightarrow{d} N(0, \sigma^2 M_{xx}^{-1}). \tag{12}$$

2 Solutions: Question 2

2.1 (1)

The log-likelihood function is

$$\log L_T(\theta) = \log \prod_{i=1}^{T} f(X_i;\theta) = \sum_{i=1}^{T} \log f(X_i;\theta). \tag{13}$$

The maximum likelihood estimator is defined by

$$\hat{\theta} \in \arg\max_{\theta} \frac{1}{T} \sum_{i=1}^{T} \log f(X_i;\theta). \tag{14}$$

Assuming the domain of X_i does not depend on the parameter θ, we define the following expectation of $\log f(X_i;\theta)$ as

$$\log L(\theta) = E[\log f(X_i;\theta)] = \int (\log f(X_i;\theta)) f(X_i|\theta_0) dx, \tag{15}$$

where θ_0 is a unknown true parameter. The function $\log L(\theta)$ is maximized at $\theta = \theta_0$, that is, $\log L(\theta) \leq \log L(\theta_0)$ for any θ bacuase

$$\begin{aligned}
& E[\log f(X_i;\theta) - \log f(X_i;\theta_0)] \\
=& E\left[\log \frac{f(X_i;\theta)}{f(X_i;\theta_0)}\right] \\
\leq& E\left[\frac{f(X_i;\theta)}{f(X_i;\theta_0)} - 1\right] \\
=& \int \left(\frac{f(X_i;\theta)}{f(X_;\theta_0)} - 1\right) f(X_i;\theta_0) dx \\
=& \int f(X_i;\theta) dx - \int f(X_i;\theta_0) dx = 0.
\end{aligned}$$

Note that $\log x \leq x - 1$.

By the weak law of large numbers, for any θ, $T^{-1} \log L_T(\theta) \to \log L(\theta)$. By defenition, $\hat{\theta}$ is the maximizer of $T^{-1} \log L_T(\theta)$. Thus, the probability limit of $T^{-1} \log L_T(\theta)$ also maximizes $\log L(\theta)$. This implies that $\hat{\theta} \xrightarrow{p} \theta_0$.

2.2 (2)

Applying the central limit theorem with unequal variance yields

$$\sqrt{T}\left(\frac{1}{T}\frac{\partial \log L(\theta)}{\partial \theta} - \mu\right) = \frac{1}{\sqrt{T}}\sum_{i=1}^{T}\left(\frac{\partial \log f(X_i;\theta)}{\partial \theta} - \mu\right) \xrightarrow{d} N(0,\Sigma), \tag{16}$$

where $\frac{\partial}{\partial \theta}\log f(X_i;\theta)$ is distributed with mean μ and variance Σ_i, and $\Sigma = \lim_{T\to\infty}(T^{-1}\sum_{i=1}^{T}\Sigma_i)$. From (16), we have

$$\lim_{T\to\infty} E\left[\frac{1}{T}\frac{\partial \log L(\theta)}{\partial \theta}\right] = \mu, \tag{17}$$

$$\lim_{T\to\infty} T \cdot \mathrm{Var}\left[\frac{1}{T}\frac{\partial \log L(\theta)}{\partial \theta}\right] = \Sigma. \tag{18}$$

We need the expectation and variance of $\frac{\partial}{\partial \theta}\log L(\theta)$. Because $L(\theta)$ is a joint distribution, $\int L(\theta)dx = 1$. Taking the first-order derivative with respect to θ on both sides yields

$$\begin{aligned}\int \frac{\partial L(\theta)}{\partial \theta}dx &= 0\\ \int \frac{\partial \log L(\theta)}{\partial \theta}L(\theta)dx &= 0\\ E\left[\frac{\partial \log L(\theta)}{\partial \theta}\right] &= 0.\end{aligned}$$

Thus, we obtain

$$\lim_{T\to\infty} E\left[\frac{1}{T}\frac{\partial \log L(\theta)}{\partial \theta}\right] = \lim_{T\to\infty}\frac{1}{T}E\left[\frac{\partial \log L(\theta)}{\partial \theta}\right] = 0 = \mu. \tag{19}$$

To obtain the variance, taking the second-order derivative of $\int L(\theta)dx = 1$ with respect to θ,

$$\begin{aligned}\int \frac{\partial^2 \log L(\theta)}{\partial \theta \partial \theta'}L(\theta)dx + \int \frac{\partial \log L(\theta)}{\partial \theta}\frac{\partial \log L(\theta)}{\partial \theta'}L(\theta)dx &= 0\\ -E\left[\frac{\partial^2 \log L(\theta)}{\partial \theta \partial \theta'}\right] &= E\left[\frac{\partial \log L(\theta)}{\partial \theta}\frac{\partial \log L(\theta)}{\partial \theta'}\right]\end{aligned}$$

Thus, we obtain the variance as follows:

$$
\begin{aligned}
&\mathrm{Var}\left[\frac{\partial \log L(\theta)}{\partial \theta}\right] \\
=&E\left[\frac{\partial \log L(\theta)}{\partial \theta}\frac{\partial \log L(\theta)}{\partial \theta'}\right] - E\left[\frac{\partial \log L(\theta)}{\partial \theta}\right]^2 \\
=&E\left[\frac{\partial \log L(\theta)}{\partial \theta}\frac{\partial \log L(\theta)}{\partial \theta'}\right] \\
=&-E\left[\frac{\partial^2 \log L(\theta)}{\partial \theta \partial \theta'}\right] \\
=&I(\theta),
\end{aligned} \tag{20}
$$

where $I(\theta)$ is the information matrix. This leads to

$$
\lim_{T\to\infty} T \cdot \mathrm{Var}\left[\frac{1}{T}\frac{\partial \log L(\theta)}{\partial \theta}\right] = \lim_{T\to\infty} \frac{1}{T}\mathrm{Var}\left[\frac{1}{T}\frac{\partial \log L(\theta)}{\partial \theta}\right] = \lim_{T\to\infty} \frac{1}{T} I(\theta) = \Sigma. \tag{21}
$$

Hence, the asymptotic distribution is

$$
\frac{1}{\sqrt{T}} \sum_{i=1}^{T} \frac{\partial \log f(X_i;\theta)}{\partial \theta} \xrightarrow{d} N(0, \Sigma) \tag{22}
$$

where $\Sigma = \lim_{T\to\infty} T^{-1} I(\theta)$.

2.3 (3)

Taking the first-order approximation of $\frac{\partial}{\partial \theta} \log L(\hat{\theta}) = 0$ around $\hat{\theta} = \theta$ yields

$$
\frac{\partial \log L(\theta)}{\partial \theta} + \frac{\partial^2 \log L(\theta)}{\partial \theta \partial \theta'}(\hat{\theta} - \theta) = 0.
$$

We rewrite it as follows:

$$
\begin{aligned}
\hat{\theta} - \theta &= -\left(\frac{\partial^2 \log L(\theta)}{\partial \theta \partial \theta'}\right)^{-1} \frac{\partial \log L(\theta)}{\partial \theta} \\
\sqrt{T}(\hat{\theta} - \theta) &= \left(-\frac{1}{T}\frac{\partial^2 \log L(\theta)}{\partial \theta \partial \theta'}\right)^{-1} \left(\frac{1}{\sqrt{T}}\frac{\partial \log L(\theta)}{\partial \theta}\right).
\end{aligned} \tag{23}
$$

Note that

$$
-\frac{1}{T}\frac{\partial \log L(\theta)}{\partial \theta \partial \theta'} \to \lim_{T\to\infty} \frac{1}{T} E\left(-\frac{\partial \log L(\theta)}{\partial \theta \partial \theta'}\right) = \lim_{T\to\infty} \frac{1}{T} I(\theta) = \Sigma.
$$

Since Σ is symmetric, using the Slutsky's theorem, we have the asymptotic distribution of $\sqrt{T}(\hat{\theta}-\theta)$ as follows:

$$\sqrt{T}(\hat{\theta}-\theta) \xrightarrow{d} N(0, \Sigma^{-1}\Sigma\Sigma^{-1}) \tag{24}$$

$$\xrightarrow{d} N(0, \Sigma^{-1}) \tag{25}$$