Quantitative Analysis
Chung-Ming Kuan
Jui-Chung Yang

Problem Set 6: Solution

Part One: Hand-Written Exercise

1. For the probit model, its NLS estimator is obtained by solving the F.O.C.:

$$\frac{\partial}{\partial \boldsymbol{\theta}} \sum_{i=1}^{n} \left[y_i - \boldsymbol{\Phi}(\mathbf{x_i'}\boldsymbol{\theta}) \right]^2 = -2 \sum_{i=1}^{n} \left[y_i - \boldsymbol{\Phi}(\mathbf{x_i'}\boldsymbol{\theta}) \right] \phi(\mathbf{x_i'}\boldsymbol{\theta}) \mathbf{x_i} = \mathbf{0},$$

which is not the same as the F.O.C of the ML method:

$$\sum_{i=1}^{n} \frac{y_i - \boldsymbol{\Phi}(x_i'\theta)}{\boldsymbol{\Phi}(x_i'\theta)\left[1 - \Phi(x_i'\theta)\right]} \phi(\mathbf{x_i'}\boldsymbol{\theta}) \mathbf{x_i} = \mathbf{0}.$$

For the logit model, its NLS estimator is obtained by solving the F.O.C.:

$$\frac{\partial}{\partial \boldsymbol{\theta}} \sum_{i=1}^{n} \left[y_i - G(\mathbf{x_i'}\boldsymbol{\theta}) \right]^2 = -2 \sum_{i=1}^{n} \left[y_i - G(\mathbf{x_i'}\boldsymbol{\theta}) \right] G'(\mathbf{x_i'}\boldsymbol{\theta}) \mathbf{x_i} = \mathbf{0},$$

which is not the same as the F.O.C of the ML method:

$$\sum_{i=1}^{n} \left[y_i - G(\mathbf{x_i'}\boldsymbol{\theta}) \right] \mathbf{x_i} = \mathbf{0}. \;\blacksquare$$

2. First, we have

$$\begin{aligned}
\mathbf{H}(\boldsymbol{\theta}_0) &= \mathbb{E}[\nabla^2 \ln \ell(\boldsymbol{\theta}_0)] \\
&= \mathbb{E}\left(\nabla \left[\frac{\left[y_i - \Phi(\mathbf{x}_i'\boldsymbol{\theta}_0)\right] \phi(\mathbf{x}_i'\boldsymbol{\theta}_0)}{\Phi(\mathbf{x}_i'\boldsymbol{\theta}_0)\left[1 - \Phi(\mathbf{x}_i'\boldsymbol{\theta}_0)\right]} \mathbf{x}_i \right] \right) \\
&= \mathbb{E}\left(\frac{\phi'(\mathbf{x}_i'\boldsymbol{\theta}_0)\Phi(\mathbf{x}_i'\boldsymbol{\theta}_0) y_i - \phi'(\mathbf{x}_i'\boldsymbol{\theta}_0)\Phi^2(\mathbf{x}_i'\boldsymbol{\theta}_0) - \phi'(\mathbf{x}_i'\boldsymbol{\theta}_0)\Phi^2(\mathbf{x}_i'\boldsymbol{\theta}_0) y_i + \phi'(\mathbf{x}_i'\boldsymbol{\theta}_0)\Phi^3(\mathbf{x}_i'\boldsymbol{\theta}_0)}{\left(\Phi(\mathbf{x}_i'\boldsymbol{\theta}_0)\left[1 - \Phi(\mathbf{x}_i'\boldsymbol{\theta}_0)\right]\right)^2} \mathbf{x}_i \mathbf{x}_i' \right) \\
&\quad - \mathbb{E}\left(\frac{\left[y_i - 2y_i\Phi(\mathbf{x}_i'\boldsymbol{\theta}_0) + \Phi^2(\mathbf{x}_i'\boldsymbol{\theta}_0)\right] \phi^2(\mathbf{x}_i'\boldsymbol{\theta}_0)}{\left(\Phi(\mathbf{x}_i'\boldsymbol{\theta}_0)\left[1 - \Phi(\mathbf{x}_i'\boldsymbol{\theta}_0)\right]\right)^2} \mathbf{x}_i \mathbf{x}_i' \right) \\
&= 0 - \mathbb{E}\left(\frac{\left[y_i - 2y_i\Phi(\mathbf{x}_i'\boldsymbol{\theta}_0) + \Phi^2(\mathbf{x}_i'\boldsymbol{\theta}_0)\right] \phi^2(\mathbf{x}_i'\boldsymbol{\theta}_0)}{\left(\Phi(\mathbf{x}_i'\boldsymbol{\theta}_0)\left[1 - \Phi(\mathbf{x}_i'\boldsymbol{\theta}_0)\right]\right)^2} \mathbf{x}_i \mathbf{x}_i' \right) \\
&= -\mathbb{E}\left(\frac{\phi^2(\mathbf{x}_i'\boldsymbol{\theta}_0)}{\Phi(\mathbf{x}_i'\boldsymbol{\theta}_0)\left[1 - \Phi(\mathbf{x}_i'\boldsymbol{\theta}_0)\right]} \mathbf{x}_i \mathbf{x}_i' \right),
\end{aligned}$$

where that last two equations are by the law of iterated expectation and the fact that $\mathbb{E}\left(y_i|\mathbf{x}_i\right) = \Phi(\mathbf{x}_i'\boldsymbol{\theta}_0)$.

Now, for the information matrix $\mathbf{B}(\boldsymbol{\theta}_0)$, we have:

$$
\begin{aligned}
\mathbf{B}(\boldsymbol{\theta}_0) &= \mathbb{E}\left(\frac{1}{N}(\nabla L_n(\boldsymbol{\theta}_0))(\nabla L_n(\boldsymbol{\theta}_0))'\right) \\
&= \mathbb{E}\left(\frac{\left[y_i - \Phi(\mathbf{x}_i'\boldsymbol{\theta}_0)\right]^2 \phi^2(\mathbf{x}_i'\boldsymbol{\theta}_0)}{\left(\Phi(\mathbf{x}_i'\boldsymbol{\theta}_0)\left[1-\Phi(\mathbf{x}_i'\boldsymbol{\theta}_0)\right]\right)^2}\mathbf{x}_i\mathbf{x}_i'\right) \quad \text{when } (y_i, \mathbf{x}_i')' \text{ are iid data.} \\
&= \mathbb{E}\left(\frac{\left[y_i^2 - 2y_i\Phi(\mathbf{x}_i'\boldsymbol{\theta}_0) + \Phi^2(\mathbf{x}_i'\boldsymbol{\theta}_0)\right] \phi^2(\mathbf{x}_i'\boldsymbol{\theta}_0)}{\left(\Phi(\mathbf{x}_i'\boldsymbol{\theta}_0)\left[1-\Phi(\mathbf{x}_i'\boldsymbol{\theta}_0)\right]\right)^2}\mathbf{x}_i\mathbf{x}_i'\right) \\
&= \mathbb{E}\left(\frac{\phi^2(\mathbf{x}_i'\boldsymbol{\theta}_0)}{\Phi(\mathbf{x}_i'\boldsymbol{\theta}_0)\left[1-\Phi(\mathbf{x}_i'\boldsymbol{\theta}_0)\right]}\mathbf{x}_i\mathbf{x}_i'\right),
\end{aligned}
$$

where the last equation is due to the fact that:

$$
\mathrm{Var}\left(y_i|\mathbf{x}_i\right) = \Phi(\mathbf{x}_i'\boldsymbol{\theta}_0)\left(1-\Phi(\mathbf{x}_i'\boldsymbol{\theta}_0)\right),
$$

and

$$
\begin{aligned}
\mathrm{Var}\left(y_i|\mathbf{x}_i\right) &= \mathbb{E}\left(y_i^2|\mathbf{x}_i\right) - \mathbb{E}^2\left(y_i|\mathbf{x}_i\right) \\
&= \mathbb{E}\left(y_i^2|\mathbf{x}_i\right) - \Phi^2(\mathbf{x}_i'\boldsymbol{\theta}_0),
\end{aligned}
$$

which leads to the fact

$$
\mathbb{E}\left(y_i^2|\mathbf{x}_i\right) = \Phi(\mathbf{x}_i'\boldsymbol{\theta}_0) = \mathbb{E}\left(y_i|\mathbf{x}_i\right).
$$

It follows that the information equality holds: $\mathbf{H}(\boldsymbol{\theta}_0) + \mathbf{B}(\boldsymbol{\theta}_0) = 0$. ■

3. (a)

$$
\begin{aligned}
&\because \frac{(n-k-1)\hat{\sigma}^2}{\sigma_o^2}|\mathbf{X} \sim \chi^2(n-k-1) \\
&\therefore \mathrm{var}\left(\frac{(n-k-1)\hat{\sigma}^2}{\sigma_o^2}|\mathbf{X}\right) = 2(n-k-1) \\
&\Rightarrow \mathrm{var}\left(\hat{\sigma}^2|\mathbf{X}\right) = \left(\frac{1}{n-k-1}\right)2\sigma_o^4 \\
&\Rightarrow \mathrm{var}\left(\sqrt{n}\hat{\sigma}^2|\mathbf{X}\right) = \left(\frac{n}{n-k-1}\right)2\sigma_o^4 > 2\sigma_o^4.
\end{aligned}
$$

(b)

$$
\begin{aligned}
\because\ \tilde{\sigma}^2 &= \frac{n-k-1}{n}\hat{\sigma}_o^2 \\
\therefore\ \operatorname{var}(\tilde{\sigma}^2|\mathbf{X}) &= \left(\frac{n-k-1}{n}\right)\frac{2\sigma_o^4}{n} \\
\Rightarrow \operatorname{var}(\sqrt{n}\tilde{\sigma}^2|\mathbf{X}) &= \left(\frac{n-k-1}{n}\right)2\sigma_o^4 < 2\sigma_o^4.
\end{aligned}
$$

(c) Yes. In general, $\operatorname{var}(\sqrt{n}\tilde{\sigma}^2|\mathbf{X}) < 2\sigma_o^4 < \operatorname{var}\left(\sqrt{n}\hat{\sigma}^2|\mathbf{X}\right)$.

(d) No. Since $\tilde{\sigma}^2$ is biased, the Cramér-Rao lower bound doesn't apply to $\tilde{\sigma}^2$. Also, note that $\hat{\sigma}^2$ doesn't reach the Cramér-Rao lower bound.