updated: 11/23/00

Hayashi *Econometrics*: Answers to Selected Review Questions

Chapter 7

Section 7.1

1. $m(\mathbf{w}_t;\boldsymbol{\theta}) = -[y_t - \boldsymbol{\Phi}(\mathbf{x}_t'\boldsymbol{\theta})]^2$.

2. Since $\mathrm{E}(y_t \mid \mathbf{x}_t) = \boldsymbol{\Phi}(\mathbf{x}_t'\boldsymbol{\theta}_0)$, we have:

$$\mathrm{E}[\mathbf{x}_t \cdot (y_t - \boldsymbol{\Phi}(\mathbf{x}_t'\boldsymbol{\theta}_0)) \mid \mathbf{x}_t] = \mathbf{x}_t\, \mathrm{E}[y_t - \boldsymbol{\Phi}(\mathbf{x}_t'\boldsymbol{\theta}_0) \mid \mathbf{x}_t] = \mathbf{0}.$$

Use the Law of Total Expectations. $g(\mathbf{w}_t;\boldsymbol{\theta}) = \mathbf{x}_t \cdot (y_t - \boldsymbol{\Phi}(\mathbf{x}_t'\boldsymbol{\theta}))$.

5. Q_n is (7.1.3) with $g(\mathbf{w}_t;\boldsymbol{\theta}) = \mathbf{x}_t \cdot (y_t - \theta_0 z_t)$. $\widetilde{Q}_n$ is (7.1.3) with $g(\mathbf{w}_t;\boldsymbol{\theta}) = \mathbf{x}_t \cdot (z_t - \lambda_0 y_t)$.

Section 7.2

2. Sufficiency is proved in the text. To show necessity, suppose (7.2.10) were false. Then there exists a $\boldsymbol{\theta}_1$ in Θ such that $\phi(\mathbf{x}_t;\boldsymbol{\theta}_1) = \phi(\mathbf{x}_t;\boldsymbol{\theta}_0)$. Then from (7.2.9), $\mathrm{E}[\{y_t - \phi(\mathbf{x}_t;\boldsymbol{\theta}_1)\}^2] = \mathrm{E}[\{y_t - \phi(\mathbf{x}_t;\boldsymbol{\theta}_0)\}^2]$. This is a contradiction because $\boldsymbol{\theta}_0$ is the only maximizer.

3. What needs to be proved is: "$\mathrm{E}(\mathbf{x}_t\mathbf{x}_t')$ nonsingular" $\Rightarrow$ "$\mathbf{x}_t'\boldsymbol{\theta} \neq \mathbf{x}_t'\boldsymbol{\theta}_0$ for $\boldsymbol{\theta} \neq \boldsymbol{\theta}_0$". Use the argument developed in Example 7.8.

4. What needs to be proved is: "$\mathrm{E}(\mathbf{x}_t\mathbf{x}_t')$ nonsingular" $\Rightarrow$ "$\boldsymbol{\Phi}(\mathbf{x}_t'\boldsymbol{\theta}) \neq \boldsymbol{\Phi}(\mathbf{x}_t'\boldsymbol{\theta}_0)$ for $\boldsymbol{\theta} \neq \boldsymbol{\theta}_0$". It was shown in the previous review question that the nonsingularity condition implies $\mathbf{x}_t'\boldsymbol{\theta} \neq \mathbf{x}_t'\boldsymbol{\theta}_0$ for $\boldsymbol{\theta} \neq \boldsymbol{\theta}_0$.

7. The Hessian matrix for linear GMM is negative definite. So the objective function is strictly concave.

8. So the identification condition is $\mathrm{E}[\mathbf{g}(\mathbf{w}_t;\boldsymbol{\theta}_0)] = \mathbf{0}$ and $\mathbf{W}\,\mathrm{E}[\mathbf{g}(\mathbf{w}_t;\boldsymbol{\theta})] \neq \mathbf{0}$ for $\boldsymbol{\theta} \neq \boldsymbol{\theta}_0$.

Section 7.3

1. A better question would be as follows.

> Consider a random sample $(\mathbf{w}_1,\dots,\mathbf{w}_n)$. Let $f(\mathbf{w}_t;\boldsymbol{\theta}_0)$ be the density of $\mathbf{w}_t$, where $\boldsymbol{\theta}_0$ is the p-dimensional true parameter vector. The log likelihood of the sample is
>
> $$L(\mathbf{w}_1,\dots,\mathbf{w}_n;\boldsymbol{\theta}) = \sum_{t=1}^{n} \log f(\mathbf{w}_t;\boldsymbol{\theta}).$$
>
> Let $\mathbf{r}_n(\boldsymbol{\theta})$ be the score vector of this log likelihood function. That is, $\mathbf{r}_n(\boldsymbol{\theta})$ is the p-dimensional gradient of L. In Chapter 1, we defined the **Cramer-Rao bound**

to be the inverse of $\mathrm{E}[\mathbf{r}_n(\boldsymbol{\theta}_0)\mathbf{r}_n(\boldsymbol{\theta}_0)']$. Define the **asymptotic Cramer-Rao bound** as the inverse of

$$\mathbf{J} \equiv \lim_{n\to\infty} \frac{1}{n}\,\mathrm{E}[\mathbf{r}_n(\boldsymbol{\theta}_0)\mathbf{r}_n(\boldsymbol{\theta}_0)'].$$

Assume that all the conditions for the consistency and asymptotic normality of the (unconditional) maximum likelihood estimator are satisfied. Show that the asymptotic variance matrix of the ML estimator equals the asymptotic Cramer-Rao bound.

The answer is as follows.

Define $\mathbf{s}(\mathbf{w}_t;\boldsymbol{\theta})$ as the gradient of $\log f(\mathbf{w}_t;\boldsymbol{\theta})$. Then

$$\mathbf{r}_n(\boldsymbol{\theta}) = \sum_{t=1}^{n} \mathbf{s}(\mathbf{w}_t;\boldsymbol{\theta}).$$

Since $\mathrm{E}[\mathbf{s}(\mathbf{w}_t;\boldsymbol{\theta}_0)] = \mathbf{0}$ and $\{\mathbf{s}(\mathbf{w}_t;\boldsymbol{\theta}_0)\}$ is i.i.d., we have

$$\mathrm{E}[\mathbf{r}_n(\boldsymbol{\theta}_0)\mathbf{r}_n(\boldsymbol{\theta}_0)'] = \mathrm{Var}(\mathbf{r}_n(\boldsymbol{\theta}_0)) = \sum_{t=1}^{n} \mathrm{Var}(\mathbf{s}(\mathbf{w}_t;\boldsymbol{\theta}_0)) = n \cdot \mathrm{E}[\mathbf{s}(\mathbf{w}_t;\boldsymbol{\theta}_0)\mathbf{s}(\mathbf{w}_t;\boldsymbol{\theta}_0)'].$$

By the information matrix equality, it follows that

$$\frac{1}{n}\,\mathrm{E}[\mathbf{r}_n(\boldsymbol{\theta}_0)\mathbf{r}_n(\boldsymbol{\theta}_0)'] = -\,\mathrm{E}[\mathbf{H}(\mathbf{w}_t;\boldsymbol{\theta}_0)],$$

where $\mathbf{H}(\mathbf{w}_t;\boldsymbol{\theta})$ is the hessian of the log likelihood for observation t. Therefore, trivially, the limit as $n \to \infty$ of $\frac{1}{n}\,\mathrm{E}[\mathbf{r}_n(\boldsymbol{\theta}_0)\mathbf{r}_n(\boldsymbol{\theta}_0)']$ is $-\,\mathrm{E}[\mathbf{H}(\mathbf{w}_t;\boldsymbol{\theta}_0)]$, which is the inverse of the asymptotic variance matrix.