MIDTERM – FEBRUARY 8, 2004

Instructions

You should attempt all questions in this exam. If you are unable to give analytical solutions to the problems, credit will be given if you manage to suggest an adequate numerical solution technique or at least some intuition about the expected result. Be specific in the assumptions and steps that you take. This will earn you partial credit if your answer is incorrect further down the line. However, you need not prove every single result as long as it is a well established result or it is in your class notes. I will demand a high standard of precision and clarity. Sloppy presentation will almost surely result in a poor score. Please invest time in presenting your work in an intelligible and neat manner. Good luck!

Please answer in the space provided for each question as neatly as possible. To ensure neatness, please first work out the solution in scratch paper and then enter the answer in the appropriate space. The exam is for a total score of 100 points allocated as indicated for each question.

Problem 1

Suppose the data is characterized by the following process

\[ y_t = \rho y_{t-1} + \varepsilon_t , \quad \varepsilon_t \sim N(0, \sigma^2) \quad |\rho| < 1, \quad \sigma < \infty \]

(a) Establish the consistency and asymptotic normality of the OLS estimator for \( \rho \) is a sample of \( T \) observations, conditional on \( y_t \). Be sure to express this distribution as a function of the parameters of the model only.

Hints: \( \{\varepsilon_t, y_{t-k}\} \) is a martingale difference sequence for \( k \geq 1 \) and \( \frac{1}{T} \sum_{t=1}^{T} \varepsilon_t^2 y_{t-k} \to \sigma^2 \gamma_0 \) where \( \gamma_0 = E(y_t^2) \).

Solution

\[
\hat{\rho} = \frac{\sum_{t=2}^{T} y_{t-1} y_t}{\sum_{t=2}^{T} y_{t-1}^2} = \rho + \frac{\sum_{t=2}^{T} y_{t-1} \varepsilon_t}{\sum_{t=2}^{T} y_{t-1}^2}
\]

Notice that:

\[
\frac{1}{T} \sum_{t=1}^{T} y_{t-1}^2 \to \gamma_0 < \infty \quad \gamma_0 = \frac{\sigma^2}{1 - \rho^2} \quad \text{and} \quad \frac{1}{T} \sum_{t=1}^{T} y_{t-1} \varepsilon_t \to 0
\]

Hence, the OLS estimator is consistent.
To show asymptotic normality, we need to establish the convergence in distribution of the numerator in expression (1). By the hint, this is a martingale difference sequence so we can use the central limit theorem for MDS. The denominator converges in probability to a constant and by Slutsky’s theorem, we can combine the two easily to obtain the desired distribution.

Hence

\[
\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \varepsilon_t y_{t-1} \xrightarrow{d} N \left(0, \frac{\sigma^2}{(1 - \rho^2)} \right)
\]

Combining the numerator and denominator in expression (1), the asymptotic distribution is:

\[
\sqrt{T} (\hat{\rho} - \rho) \xrightarrow{d} N \left(0, (1 - \rho^2) \right)
\]

(b) Given the OLS estimate whose asymptotic distribution you have just derived, derive the distribution of the impulse response function parameters \( \rho^j \).

**Solution**

Using the delta method and realizing that

\[
g(\rho) = \rho^j; \quad g'(\rho) = j \rho^{j-1}
\]

then

\[
\sqrt{T} (\hat{\rho}^j - \rho^j) \xrightarrow{d} N \left(0, j^2 \rho^{2(j-1)} \left(1 - \rho^2 \right) \right)
\]

(c) Given the DGP, derive the expression for \( y_{t+j} \) as a function of \( y_{t+j} \) and the residuals. This can be done easily by recursive substitution.

**Solution**

By recursive substitution, this is trivially

\[
y_{t+j} = \rho^{j+1} y_{t+1} + \varepsilon_{t+j} + \rho \varepsilon_{t+j-1} + \ldots + \rho^j \varepsilon_t
\]

(d) Consider estimating the coefficients of the impulse response with the least squares regression

\[
y_{t+j} = \beta_{j+1} y_{t+1} + u_t
\]

Show that this estimator is consistent for \( \rho^{j+1} \)
Solution

\[ \hat{\beta}_{j+1} = \frac{\sum_{t=j+1}^{T} y_{t+j} y_{t-1}}{\sum_{t=j+1}^{T} y_{t-1}^2} = \rho^{j+1} + \frac{\sum_{t=j+1}^{T} \varepsilon_{t+j} y_{t-1}}{\sum_{t=j+1}^{T} y_{t-1}^2} + \rho^2 + \frac{\sum_{t=j+1}^{T} \varepsilon_{t+j-1} y_{t-1}}{\sum_{t=j+1}^{T} y_{t-1}^2} + \ldots + \rho^j + \frac{\sum_{t=j+1}^{T} \varepsilon_t y_{t-1}}{\sum_{t=j+1}^{T} y_{t-1}^2} \]

Hence, although we have more terms to worry about, they are all similar to how we derived the answer in (a) and by the same mechanisms we get consistency.

(e) Derive the asymptotic distribution of the least squares estimator in part (d)

Solution

Using the results in (a), notice that

\[ \sqrt{T}(\hat{\beta}_{j+1} - \rho^{j+1}) \xrightarrow{d} z_1 + z_2 + \ldots + z_j \]

where

\[ z_j \sim N(0, \rho^{2j}(1 - \rho^2)) \]

Hence,

\[ \sqrt{T}(\hat{\beta}_{j+1} - \rho^{j+1}) \xrightarrow{d} N\left(0, (1 + \rho^2 + \ldots + \rho^{2j})(1 - \rho^2)\right) \]

(f) Discuss why the usual estimates of the variance of the OLS estimator in part (d) are not efficient. Briefly, how could you estimate \( \rho^j \) more efficiently?

Solution

The OLS expression in part (d) is not fully efficient since the residuals have a moving average component. This does not affect consistency as we have shown. To obtain more efficient estimates, it would be natural to estimate the model by maximum likelihood and imposing the moving average structure on the residuals from part (c). This is the full information maximum likelihood estimator which is the most efficient given that in this case, we know the DGP.
Problem 2

Consider the following two processes:

\[ y_t = c + \epsilon_t + \psi \epsilon_{t-1} \]
\[ x_t = c + \psi \epsilon_t + \epsilon_{t-1} \]

where \( \epsilon_t \) is a martingale difference sequence and \( |\psi| > 1 \).

(a) Calculate \( E(y_t) \) and \( E(x_t) \) and verify that they are the same.

**Solution**

*Trivially they are both equal to \( c \)*

(b) Calculate \( V(y_t) \) and \( V(x_t) \) and verify that they are the same.

**Solution**

\[ V(y_t) = (1 + \psi^2)\sigma_\epsilon^2 = V(x_t) \]

(c) Is \( y_t \) invertible?

**Solution:** no

(d) Is \( x_t \) invertible?

**Solution**

*Check the lag polynomial, which is simply (\( \psi + z \)) = 0. Hence \( |z| = |\psi| > 1 \) and therefore stationary.*

*This is an unusual property of MA(1) processes – for every non-invertible MA(1) there exists an invertible counterpart that shares the same moments.*
Problem 3

You assume that a sample of $T$ observations on $y_t$ is distributed as an exponential, such that

$$y_t = \frac{\varepsilon_t}{\lambda_t}; \quad \lambda_t = \beta y_{t-1}; \text{ such that } f(y_t; \lambda_t) = \lambda_t \exp(-\lambda_t y_t)$$

Consider estimating the parameter $\beta$.

(a) Derive the maximum likelihood estimator for $\beta$ conditional on $y_0 = 1$.

**Solution**

The log-likelihood function is:

$$L(\theta) = \sum_{t=1}^{T} \log(\lambda_t) - \sum_{t=1}^{T} \lambda_t y_t$$

The score is:

$$\frac{\partial L(\theta)}{\partial \theta} = \sum_{t=1}^{T} \frac{y_{t-1}}{\beta y_{t-1}} - \sum_{t=1}^{T} y_{t-1} y_t$$

hence,

$$\hat{\beta} = \frac{T}{\sum_{t=1}^{T} y_{t-1} y_t}$$

(b) Derive the variance of your estimator from the second derivative of the likelihood

**Solution**

$$\frac{\partial^2 L(\theta)}{\partial \theta^2} = -\frac{T}{\beta^2}$$

hence

$$\hat{\nu} (\hat{\beta}) = \frac{\hat{\beta}^2}{T}$$

(c) For this problem, use the mean value theorem to derive the asymptotic distribution of the estimator you derived in (a).

*Hint:* at some point in the derivations you may want to use that $y_t = \frac{\varepsilon_t}{\lambda_t}$

**Solution**
The mean value theorem uses a Taylor series expansion of the score around the true parameter. In this case:

\[
\frac{\partial L}{\partial \beta} = \frac{\partial L}{\partial \beta_0} + \frac{\partial^2 L}{\partial \beta^2} (\hat{\beta} - \beta_0)
\]

\[
0 = \left( \frac{1}{\sum \beta_0} - \sum y_{t-1} y_t \right) - \frac{1}{\sum \beta^2} (\hat{\beta} - \beta)
\]

Notice that:

\[
\left( \frac{1}{\sum \beta_0} - \sum y_{t-1} y_t \right) = \left( \frac{1}{\sum \beta_0} - \sum \frac{\varepsilon_t}{\beta_0} \right)
\]

since \(\varepsilon_t\) is a MDS, then

\[
\frac{1}{\sqrt{T}} \left( \frac{1}{\sum \beta_0} - \sum \frac{\varepsilon_t}{\beta_0} \right) \xrightarrow{d} \mathcal{N} \left( 0, T \frac{\beta_0^2}{\beta_0^2} \right)
\]

\[
\bar{\beta} \xrightarrow{p} \beta_0
\]

hence

\[
\sqrt{T}(\hat{\beta} - \beta_0) \xrightarrow{d} \mathcal{N} \left( 0, \frac{\beta^2}{T} \frac{T \beta_0^2 \beta_0^2}{T} \right) = \mathcal{N} \left( 0, \frac{\beta_0^2}{T} \right)
\]