Solutions

Instructions

Problem 1

Suppose the data is characterized by the following process

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

(a) **[12 points]** Establish the consistency and asymptotic normality of the OLS estimator for ρ is a sample of T observations, conditional on y_1. 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 $\frac{1}{T}\sum \varepsilon_t^2 y_{t-k}^2 \xrightarrow{p} \sigma^2 \gamma_0$ where $\gamma_0 = E(y_t^2)$.

Solution

$$\hat{\rho} = \frac{\sum_{t=2}^{T} y_{t-1} y_t}{\sum_{t=2}^{T} y_{t-1}^2} = \rho + \frac{\sum_{t=2}^{T} y_{t-1} \varepsilon_t}{\sum_{t=2}^{T} y_{t-1}^2} \qquad \textbf{(1)}$$

Notice that:

$$\frac{1}{T}\sum y_{t-1}^2 \xrightarrow{p} \gamma_0 < \infty \qquad \gamma_0 = \frac{\sigma^2}{1-\rho^2} \text{ and } \frac{1}{T}\sum y_{t-1}\varepsilon_t \xrightarrow{p} 0$$

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

$$\frac{1}{\sqrt{T}}\sum \varepsilon_t y_{t-1} \xrightarrow{d} N\left(0, \frac{\sigma^4}{(1-\rho^2)}\right)$$

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

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

(b) **[7 points]** Given the OLS estimate whose asymptotic distribution you have just derived, derive the distribution of the impulse response function parameters ρ^j.
Solution

Using the delta method and realizing that

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

then

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

(c) **[5 points]** Given the DGP, derive the expression for y_{t+j} as a function of y_{t-1} 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) **[7 points]** 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 ρ^{j+1}
Solution

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

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) **[7 points]** Derive the asymptotic distribution of the least squares estimator in part (d)
Solution

Using the results in (a), notice that

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

where

$$z_j \sim N\left(0, \rho^{2j}(1-\rho^2)\right)$$

Hence,

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

(f) **[7 points]** Discuss why the usual estimates of the variance of the OLS estimator in part (d) are not efficient. Briefly, how could you estimate ρ^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 model

$$y_t = x_t'\beta_0 + u_t$$
$$u_t = \rho u_{t-1} + \varepsilon_t$$

where $\beta_0 \in \Re^k$ and ε_t is a martingale difference sequence with mean zero and variance σ^2. In addition, you may assume that:

(i) $\left\{ \begin{pmatrix} x_t' & \varepsilon_t \end{pmatrix} \right\}$ is *i.i.d.*

(ii) (a) $E(x_t \varepsilon_t) = 0$

(b) $E|x_{it}\varepsilon_t| < \infty$, $i = 1, \ldots, k$

(iii) (a) $E|x_{it}|^2 < \infty$

(b) $E(x_t x_t') = M < \infty$; M positive definite.

(a) **[5 points]** Derive the asymptotic distribution of $\bar{u}_T = \frac{1}{T}\sum^T u_t$.

Solution

Here we can apply directly the central limit theorem for dependent processes. Notice that $C(1)^2 = 1/(1 - \rho^2)$ and hence

$$\sqrt{T}\bar{u}_T \xrightarrow{d} N\left(0, \frac{\sigma^2}{1-\rho^2}\right)$$

(b) **[12 points]** Let $g_t(\beta) = (y_t - x_t'\beta)$. Set-up the GMM problem and derive the asymptotic distribution of $\hat{\beta}_{GMM}$. Comment on how you would obtain the optimal weighting matrix in practice.

Solution

$$\min_\beta Q(\beta) = \left(\frac{1}{T}\sum^T (y_t - x_t'\beta)'\right)\hat{W}\left(\frac{1}{T}\sum^T (y_t - x_t'\beta)\right)$$

Notice that

$$G = -\frac{1}{T}\sum^T x_t'$$
$$E[(y_t - x_t\beta_0)'(y_t - x_t\beta_0)] = E(u_t^2) = \frac{\sigma^2}{1-\rho^2}$$

Hence, in the easy case

$$\sqrt{T}\left(\hat{\beta}_T - \beta_0\right) \xrightarrow{d} N\left(0, \frac{\sigma^2}{1-\rho^2} E(x_t x_t')^{-1}\right)$$

setting $W = \frac{1-\rho^2}{\sigma^2}$. *Notice that in practice,* $\hat{E}(x_t x_t')^{-1} = \left\{\frac{1}{T}\sum^T x_t x_t'\right\}^{-1}$. *If we do not assume homoskedasticity, the correct answer is really (in truth I should use the population values but I thought this might be clearer)*

$$\sqrt{T}\left(\hat{\beta}_T - \beta_0\right) \xrightarrow{d} N\left(0, \frac{\sigma^2}{1-\rho^2}\left\{\frac{1}{T}\sum^T x_t x_t'\right\}^{-1}\left\{\frac{1}{T}\sum^T x_t u_t u_t' x_t'\right\}\left\{\frac{1}{T}\sum^T x_t x_t'\right\}^{-1}\right)$$

for the optimal weighting matrix W. In practice, to obtain W we would have to use a two-step GMM procedure to estimate $\frac{\sigma^2}{1-\rho^2}$ *in the first step and adjust the estimate of the variance-covariance matrix appropriately.*

(c) **[7 points]** Set-up a least-squares formulation of the problem that allows you to estimate the parameter β directly and indicate what its distribution would look like.

Solution:

$$y_t = \rho y_{t-1} + x_t'\beta + x_{t-1}'\gamma + \varepsilon_t$$

Under assumptions of the model, the usual LS formula would give the correct distribution (notice γ= -$\beta\rho$). Effectively, this formulation of the problem parametrizes the nature of the residual serial correlation directly so that it is estimated along with the parameters of interest.

Problem 3

If

$$y_t = \mu + x_t + z_t \qquad \text{with}$$
$$x_t = \varepsilon_t + \theta\varepsilon_{t-1} \qquad \varepsilon_t \overset{i.i.d.}{\sim} D(0,\sigma^2)$$
$$z_t = \beta z_{t-1} + u_t \qquad u_t \overset{i.i.d.}{\sim} D(0,\phi^2)$$

with ε and η independent of each other. Then:

(a) **[5 points]** What is the process for y_t?
Solution

$$y_t = (1-\beta)\mu + \beta y_{t-1} + \varepsilon_t + (\theta-\beta)\varepsilon_{t-1} - \theta\beta\varepsilon_{t-2} + u_t \quad \text{ARMA(1,2)}$$

(b) **[7 points]** Give conditions to ensure y_t is covariance stationary and invertible.
Solution
Covariance stationarity: $|\beta| < 1$
Invertibility: if there is no u_t in this process, then we can solve for the roots of
$1 + (\theta-\beta)z - \theta\beta z^2 = 0$, outside the unit circle. Solving this quadratic form yields as solutions:

$$-\frac{1}{\theta} \text{ and } \frac{1}{\beta}$$

therefore, the conditions are $|\theta| < 1$ and $|\beta| < 1$

However, this process has two disturbance terms, ε_t and u_t. A better way to check conditions for stationarity and invertibility is to write this process as VARMA(1,2):

$$\begin{bmatrix} y_t \\ y_{t-1} \end{bmatrix} = \begin{bmatrix} \beta & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ y_{t-2} \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \varepsilon_t \\ u_t \end{bmatrix} + \begin{bmatrix} (\theta-\beta) & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \varepsilon_{t-1} \\ u_{t-1} \end{bmatrix} + \begin{bmatrix} -\theta\beta & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \varepsilon_{t-2} \\ u_{t-2} \end{bmatrix}$$

or $W_t = AW_{t-1} + B\eta_t + C\eta_{t-1} + D\eta_{t-2}$, where $W_t = \begin{bmatrix} y_t \\ y_{t-1} \end{bmatrix}$, $\eta_t = \begin{bmatrix} \varepsilon_t \\ u_t \end{bmatrix}$. (Constant term is ignored for clean expression)

Since $(I - AL)W_t = (B + CL + DL^2)\eta_t$, it is easy to verify that $\det(B + Cz + Dz^2) = 0$. In other words, this process is NEVER invertible.

(c) **[10 points]** Find the long-horizon forecast for y_t and its variance.

Solution

For horizons beyond two periods only the AR terms will matter and hence,

$$\hat{y}_{t+s|t} = \beta^s y_t + \sum_{j=0}^{s-1} \beta^j (1-\beta)\mu$$

As $s \to \infty$ if $|\beta| < 1$ then $\hat{y}_{t+s|t} \to \mu$. The variance of the long run forecast will converge to the unconditional variance whenever the model is stationary. The unconditional variance is

$$V(y_t) = E\left(\varepsilon_t (1+\theta L) + \frac{u_t}{1-\beta L} \right) = (1+\theta^2)\sigma^2 + \frac{\phi^2}{1-\beta^2}$$