MEI LT TS Problem Set 1[1]

Timothee Carayol

March 15, 2010

[1] Available on http://personal.lse.ac.uk/carayolt/ec402.htm

Question 1

Question

- $y_t = x_t'\beta + u_t$ with $u_t = \varepsilon_t + \theta\varepsilon_{t-1}$, for $t = 1, \ldots, T$ and where $\varepsilon_t \sim i.i.d.(0, \sigma^2)$; $\varepsilon_0 = 0$; x_t non-stochastic.
- i.e. linear regression model with $MA(1)$ error term.
- 1. Find $Var(u) = Var\left[(u_1, \ldots, u_T)'\right]$

Answer

- For $|t - s| > 1$, $cov(u_t, u_{t-s}) = 0$ (because the ε's are i.i.d.)
- $cov(u_t, u_{t-1}) = cov(\varepsilon_t + \theta\varepsilon_{t-1}, \varepsilon_{t-1} + \theta\varepsilon_{t-2}) = cov(\theta\varepsilon_{t-1}, \varepsilon_{t-1}) = \theta\sigma^2$.
- $Var(u_t) = Var(\varepsilon_t - \theta\varepsilon_{t-1}) = Var(\varepsilon_t) + \theta^2 Var(\varepsilon_{t-1})$ because ε i.i.d.; so $Var(u_t) = (1 + \theta^2)\sigma^2$.

Question 1

Answer

- In the end:

$$Var(u) = \sigma^2\Omega = \sigma^2 \begin{pmatrix} 1 & \theta & 0 & \cdots & 0 \\ \theta & 1+\theta^2 & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \theta \\ 0 & \cdots & 0 & \theta & 1+\theta^2 \end{pmatrix}.$$

Question 1

Question

2. Find matrix L such that $\varepsilon = Lu$

Answer

- $\varepsilon_t = u_t - \theta\varepsilon_{t-1} = \sum_{s=0}^{t-1}(-\theta)^s u_{t-s}$.
- This means: $\varepsilon = \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 \\ -\theta & 1 & \ddots & \ddots & \vdots \\ \theta^2 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 0 \\ -\theta^{T-1} & \cdots & \theta^2 & -\theta & 1 \end{pmatrix} u.$

Question 1

Question

3. Assuming θ is known, describe a way to compute the *BLUE* for β without inverting Ω.

Answer

- We've done the hard work in 2.: now we just have to notice that the model $Ly = LX\beta + Lu$ satisfies $A4GM$; so we just have to run OLS on this transformed model (which gives us GLS of the original model) to obtain the *BLUE*.

Question 2

Question

- $y_t = x_t'\beta + u_t$ with $u_t = \varphi u_{t-1} + \varepsilon_t$, for $t = 1, \dots, T$ and where $\varepsilon_t \sim i.i.d.(0, \sigma^2)$; $|\varphi| < 1$; x_t' and ε_t are process independent; and $plim \left(\frac{1}{T} \sum_{t=1}^{T} x_t x_t' \right) = \Sigma_{xx}$ non-singular; $plim \left(\frac{1}{T} \sum_{t=1}^{T} x_t x_{t-1}' \right) = \Sigma_{xx-}$
- i.e. linear regression model with $AR(1)$ error term (and well-behaved regressors).
- 1. Run OLS and obtain $\hat{u}_t$. Show that $\hat{\varphi} = \frac{\sum_{t=2}^{T} \hat{u}_t \hat{u}_{t-1}}{\sum_{t=2}^{T} \hat{u}_{t-1}^2}$ is consistent for φ.

Question 2

Answer

- First note that this came up in Michaelmas Term, PS8, question 1 d) (except we had skipped that bit of the proof).
- Rewrite $\hat{u}_t = y_t - \hat{y}_t = x_t'\beta + u_t - x_t'\hat{\beta} = u_t - x_t'(\hat{\beta} - \beta)$.
- The numerator of $\hat{\varphi}$ (divided by T) is then $\frac{1}{T}\sum \hat{u}_t\hat{u}_{t-1} = \frac{1}{T}\sum u_t u_{t-1} + \frac{1}{T}\sum A_t$ where $A_t = -u_t x_{t-1}'(\hat{\beta} - \beta) - u_{t-1}x_t'(\hat{\beta} - \beta) + x_t'(\hat{\beta} - \beta)x_{t-1}'(\hat{\beta} - \beta)$.
- Hence, $plim\left[\frac{\sum A_t}{T}\right] = -plim(\frac{\sum u_t x_{t-1}'}{T} + \frac{\sum u_{t-1}x_t'}{T})plim(\hat{\beta} - \beta) + plim(\frac{(\hat{\beta}-\beta)'(\sum x_t x_{t-1}')(\hat{\beta}-\beta)}{T})$.

Question 2

Answer

- Now: as x_t and u_t are process independent, and using $LLN2$, $plim(\frac{\sum u_t x'_{t-1}}{T}) = plim(\frac{\sum u_t x'_t}{T}) = 0$; likewise, $\hat{\beta}$ is consistent so $plim(\hat{\beta} - \beta) = 0$; and finally, by assumption $plim\left(\frac{\sum x_t x_{t-1}}{T}\right) = \Sigma_{xx-}$. So using Slutzky, $plim(\frac{\sum A_t}{T}) = -(0+0).0 + (0.\Sigma_{xx-}.0) = 0$.
- Likewise, we could show that the *plim* of the denominator of $\hat{\varphi}$ is $plim(\frac{\sum u_{t-1}^2}{T})$.
- $plim(\hat{\varphi}) = \frac{plim\left(\frac{1}{T}\sum u_t u_{t-1}\right)}{plim\left(\frac{1}{T}\sum u_{t-1}^2\right)} = \frac{cov(u_t, u_{t-1})}{Var(u_{t-1})}$. Then two (equivalent) ways to conclude $plim(\hat{\varphi}) = \varphi$.

Question 2

Answer

- First way: compute those covariances manually and simplify. $plim(\hat{\varphi}) = \frac{\frac{\varphi\sigma^2}{1-\varphi^2}}{\frac{\sigma^2}{1-\varphi^2}} = \varphi$.
- Second way: Look at $u_t = \varphi u_{t-1} + \varepsilon$, and call $\tilde{\varphi}$ the OLS estimator from the regression of u_t on u_{t-1} (we can't compute it as we do not know the true u_t, but we can still talk about it theoretically). Then easy to see that $\tilde{\varphi} = \frac{\sum u_t u_{t-1}}{\sum u_{t-1}^2}$, hence from the previous slide $plim(\hat{\varphi}) = plim(\tilde{\varphi})$; but we know, since *A3fi* is satisfied, that $\tilde{\varphi}$ is consistent, implying $plim(\hat{\varphi}) = \varphi$.

Question 3

Question

- $y_t = \varphi y_{t-1} + \varepsilon_t + \theta\varepsilon_{t-1}$ for $t = 1, \dots, T$, with $|\varphi| < 1$ and $\varepsilon_t \sim i.i.d.\mathbf{N}(0, \sigma^2)$.
- i.e. y_t follows an $ARMA(1,1)$ process and we'd like to estimate φ and θ.
- 1. Assuming $y_1 = \varepsilon_1 = 0$, write down the likelihood.

Answer

- Since $E(\varepsilon) = 0$ we know that $y_t | I_{t-1} \sim \mathbf{N}(\varphi y_{t-1} + \theta\varepsilon_{t-1}, \sigma^2)$ (where I_{t-1} is all the information available at $t-1$, i.e. the series of past y_s and past ε_s, $s \leq t-1$.
- Hence we know that

$$f(y_t | I_{t-1}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp - \frac{(\overbrace{y_t - \varphi y_{t-1} - \theta\varepsilon_{t-1}}^{\varepsilon_t})^2}{2\sigma^2}$$

Question 3

Answer

- So the log likelihood is simply:
- $$\log L(\theta, \varphi, \sigma; y_1, y_2, .., y_T) = -\frac{T-1}{2}\log(2\pi) - \frac{T-1}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{t=2}^{T}(y_t - \varphi y_{t-1} - \theta\varepsilon_{t-1})^2 + \log f(y_1)$$
- We know that $y_1 = \varepsilon_1 = 0$ so we can simply condition on y_1, i.e. we can take it as fixed. In this case we can just drop it from the likelihood. This is most usually done in large samples where losing one observation is not much of an issue. (Otherwise we would have to include its distribution in the likelihood).

Question 3

Question

2. Obtain the FOCs with respect to φ and θ (and σ^2, though not in the question).

Answer

- $\frac{\partial \log L}{\partial \varphi} = -\frac{1}{\sigma^2} \sum \varepsilon_t \frac{\partial \varepsilon_t}{\partial \varphi} = \frac{1}{\sigma^2} \sum \varepsilon_t (y_{t-1} + \theta \frac{\partial \varepsilon_{t-1}}{\partial \varphi})$
- $\frac{\partial \log L}{\partial \theta} = -\frac{1}{\sigma^2} \sum \varepsilon_t \frac{\partial \varepsilon_t}{\partial \theta} = \frac{1}{\sigma^2} \sum \varepsilon_t (\varepsilon_{t-1} + \theta \frac{\partial \varepsilon_{t-1}}{\partial \theta})$
- $\frac{\partial \log L}{\partial \sigma^2} = -\frac{T-1}{2\sigma^2} + \frac{1}{(\sigma^2)^2} \sum \varepsilon_t^2$

Question 3

Question

3. Obtain $\frac{1}{T} I(\psi)$

Answer

- As suggested, let us define $\psi = \begin{pmatrix} \varphi \\ \theta \\ \sigma^2 \end{pmatrix}$
- Remember that
$$I(\psi) = -E[H(\psi)]$$

.

Question 3

Answer

Next we need to compute second order derivatives. The Hessian matrix looks like this:

$$H(\psi) = \begin{pmatrix} \frac{\partial^2 \log L}{\partial \varphi^2} & \frac{\partial^2 \log L}{\partial \varphi \partial \theta} & \frac{\partial^2 \log L}{\partial \varphi \partial \sigma^2} \\ \frac{\partial^2 \log L}{\partial \theta \partial \varphi} & \frac{\partial^2 \log L}{\partial \theta^2} & \frac{\partial^2 \log L}{\partial \theta \partial \sigma^2} \\ \frac{\partial^2 \log L}{\partial \sigma^2 \partial \varphi} & \frac{\partial^2 \log L}{\partial \sigma^2 \partial \theta} & \frac{\partial^2 \log L}{\partial (\sigma^2)^2} \end{pmatrix}$$

Question 3

Answer

- Makes more sense to compute directly the information matrix (as suggested in the question), as taking the expectation leads to many simplifications..
- After some (somewhat tedious) algebra, should find:

$I(\beta) = \frac{1}{\sigma^2}$

$$\begin{pmatrix} \sum E(y_{t-1} + \theta\frac{\partial\varepsilon_{t-1}}{\partial\varphi})^2 & \sum E(\varepsilon_{t-1} + \theta\frac{\partial\varepsilon_{t-1}}{\partial\theta})(y_{t-1} + \theta\frac{\partial\varepsilon_{t-1}}{\partial\varphi}) & 0 \\ \sum E(\varepsilon_{t-1} + \theta\frac{\partial\varepsilon_{t-1}}{\partial\theta})(y_{t-1} + \theta\frac{\partial\varepsilon_{t-1}}{\partial\varphi}) & \sum E(\varepsilon_{t-1} + \theta\frac{\partial\varepsilon_{t-1}}{\partial\theta})^2 & 0 \\ 0 & 0 & \frac{T-1}{2\sigma^2} \end{pmatrix}$$

Question 3

Question

How would you estimate the covariance matrix of $\hat{\psi}$?

Answer

- Can show that the limit distribution of Maximum Likelihood is given by $\sqrt{T}(\hat{\psi} - \psi) \sim \mathbf{N}(0, IA(\psi)^{-1})$, where $IA(\psi) = \lim_{T\to\infty} \frac{I(\psi)}{T}$. So with large enough sample (asymptotic approximation), $(\hat{\psi} - \psi) \sim \mathbf{N}(0, \frac{IA(\psi)^{-1}}{T})$.
- Different ways to estimate the inverse of $IA(\psi)$.

Question 3

Answer

- Based on empirical information matrix: $\frac{I(\hat{\psi})}{T}$. (Substitute $\hat{\psi}$ for ψ in the matrix found previously). Problem: not easy to compute the expected values in the formulae.
- Instead, we can go back to the Hessian and just look at $\frac{-H(\hat{\psi})}{T}$. Then no problem with the expected values; but LLN still guarantees convergence.
- Finally, an alternative is $\frac{s(\hat{\psi})s'(\hat{\psi})}{T}$, where s is the score vector (gradient of the log-likelihood). Indeed, can be shown that this outer product has the same expectation as the hessian.