vrije Universiteit amsterdam

Studentnumber:
Name:

School of Business and Economics

Exam:	Multivariate Econometrics
Code:	E_EORM_MVE
Examinator:	dr. H. Karabiyik
Co-reader:	dr. P. Gorgi
Date:	March 17, 2020
Time:	18:45
Duration:	2 hours and 45 minutes
Calculator allowed:	**Yes**
Graphical calculator allowed:	**No**
Scrap paper	**Yes**
Number of questions:	4 (where each question consists of multiple parts)
Type of questions:	Open
Answer in:	English

Remarks:
- Motivate all your answers.

Credit score:	A score of 100 points counts for a grade 10 for the exam. Each question is 25 points.
Grades:	The grades will be made public within 10 working days after the exam.
Inspection:	TBA
Number of pages:	9 (Including front page)

Good luck!

Instructions

(i) All questions should be answered to get full points.

(ii) Each question is worth 25 points.

(iii) Read the instructions in the questions carefully.

(iv) Answer the questions as detailed as possible. Use mathematical expressions when necessary. You can use words when you cannot provide a formal mathematical answer to the questions.

(v) If a question is not clear to you, make your own assumptions to clarify the meaning of the question and then answer the question based on your assumptions.

(vi) See the back of this page for some standard results that you may make use of while answering the questions.

Some standard results

Suppose that the scalar process $\{z_t\}$ follows the following data generating process:

$$z_t = z_{t-1} + u_t,$$

where $z_0 = 0$ and u_t has the following properties:

(a) $u_t = \psi(L)\epsilon_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}$ where $\sum_{j=0}^{\infty} j \cdot |\psi_j| < \infty$ and $\{\epsilon_t\}$ is an $i.i.d$ sequence with mean zero and variance σ_ϵ^2, and finite fourth moment;

(b) σ^2 denotes the long run variance of $\{u_t\}$ and σ_u^2 denotes the contemporaneous variance of $\{u_t\}$.

Note that under these assumptions z_t can be written as a partial sum as

$$z_t = \sum_{s=1}^{t} u_s.$$

Let $W(r)$ be a standard Brownian motion process associated with u_t. Then the following results hold:

(1) $T^{-1/2} \sum_{t=1}^{T} u_t \xrightarrow{d} \sigma W(1)$;

(2) $T^{-1} \sum_{t=1}^{T} u_t^2 \xrightarrow{p} \sigma_u^2$;

(3) $T^{-1} \sum_{t=1}^{T} z_{t-1} u_t \xrightarrow{d} \frac{1}{2}\sigma^2 \left[W(1)^2 - \frac{\sigma_u^2}{\sigma^2} \right]$;

(4) $T^{-3/2} \sum_{t=1}^{T} t u_{t-j} \xrightarrow{d} \sigma \left\{ W(1) - \int_0^1 W(r) dr \right\}$ for $j = 0, 1, \ldots$;

(5) $T^{-3/2} \sum_{t=1}^{T} z_{t-1} \xrightarrow{d} \sigma \int_0^1 W(r) dr$;

(6) $T^{-2} \sum_{t=1}^{T} z_{t-1}^2 \xrightarrow{d} \sigma^2 \int_0^1 W(r)^2 dr$;

(7) $T^{-5/2} \sum_{t=1}^{T} t z_{t-1} \xrightarrow{d} \sigma \int_0^1 r W(r) dr$;

(8) $T^{-3} \sum_{t=1}^{T} t z_{t-1}^2 \xrightarrow{d} \sigma^2 \int_0^1 r W(r)^2 dr$;

(9) $T^{-(v+1)} \sum_{t=1}^{T} t^v \to 1/(v+1)$ for $v = 0, 1, \ldots$;

(10) Suppose that the DGP of another time series process y_t follows the model

$$y_t = y_{t-1} + e_t,$$

where $y_t = 0$ and e_t satisfy the same assumptions as **(a)** and has a long run variance σ_e^2 and $W_e(r)$ is a standard Brownian motion processs associated with e_t, then

$$T^{-2} \sum_{t=1}^{T} z_t y_t \xrightarrow{d} \sigma \sigma_e \int_0^1 W(r) W_e(r) dr.$$

Question 1: Conceptual Questions (25 points out of 100 points)

Below you will find **3** statements. All these are related to the concepts/techqniques that have been discussed during the lectures. Some of these statements are correct, some are wrong, some need further clarification. You need to provide a brief, to the point answer that would contain **(i) short explanations/definitions of the concepts mentioned in the statement, (ii) your judgement about the statement about whether it is correct/wrong/unclear/incomplete, and an explanation of your judgement (iii) a correction of the statement.** The concepts that you need to explain and define are written in *italics*. A formal answer using mathematics is possible, sometimes very useful but not always necessary.

(a) (5 points) Suppose that y_t and x_t are two time series processes. If there is a real causal relation between x_t and y_t, we say that x_t *Granger causes* y_t.

(b) (10 points) Suppose that y_t is a time series process and we are interested in testing for a *unit root* in y_t. That is to test if $\lambda = 1$ in

$$y_t = \lambda y_{t-1} + u_t.$$

Under the null hypothesis the estimator of λ is *super consistent.* This means that serial correlation in u_t will not influence the limiting distribution of the test statistic.

(c) (10 points) Suppose that for the panel data variable $y_{i,t}$, we have the following model

$$y_{i,t} = \beta x_{i,t} + u_{i,t},$$

for $i = 1, \ldots, N$ and $t = 1, \ldots, T$. Let N represent the number of cross-section units and T represent the number of time series periods. Suppose that $u_{i,t}$ follows

$$u_{i,t} = \lambda_i f_t + \epsilon_{i,t},$$

where $\epsilon_{i,t} \sim i.i.d(0, \sigma_\epsilon^2)$. This assumption on $u_{i,t}$ is used to model *spatial correlation* in panel data models. Unfortunately, *pooled OLS* estimator of β is inconsistent if we ignore f_t.

Question 2: Modeling and stationarity (25 points out of 100 points)

(a) (15 points) Consider two processes y_t and x_t with the DGPs

$$\Delta y_t = A_{11}\Delta y_{t-1} + A_{12}\Delta x_{t-1} + u_{y,t},$$
$$\Delta x_t = \alpha(\Delta y_{t-1} - \beta x_{t-1}) + A_{21}\Delta^2 y_{t-1} + A_{22}\Delta x_{t-1} + u_{x,t}.$$

where for $\mathbf{u}_t = (u_{y,t}, u_{x,t})'$,

$$\mathbf{u}_t = \begin{pmatrix} u_{y,t} \\ u_{x,t} \end{pmatrix} \sim IN\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix}\right].$$

(i) Discuss the nature of the series in terms of their (non-)stationarity, integration order and cointegration properties if $A_{11} = 1$, $A_{12} = 0$, $A_{21} \neq 0$, $A_{22} = 0$ and $-2 < \alpha < 0$.

(ii) Now ignoring the restrictions of part **(i)**, obtain a VAR model for the 2×1 vector $(\Delta y_t, x_t)$ in terms of the model parameters A_{11}, A_{12}, A_{21}, A_{22}, α and β.

(b) (10 points) Consider the bivariate system for $\mathbf{y}_t = (y_{1,t}, y_{2,t})'$. The DGP for $\mathbf{y}_t$ can be written as

$$\mathbf{y}_t = \begin{bmatrix} 0.2 \\ 0.1 \end{bmatrix} + \begin{bmatrix} 1 & -0.6 \\ 0.5 & -0.7 \end{bmatrix} \mathbf{y}_{t-1} + \mathbf{u}_t$$

where $\{\mathbf{u}_t\}$ is a vector white noise. Show that if this process is a weakly stationary process.

Question 3: Asymptotic Derivations (25 points out of 100 points)

(a) (15 points) **A test for unit roots:**

Suppose that we have the following data generating processes (DGP) for $\{y_t\}$

$$y_t = \rho y_{t-1} + u_t,$$

for $t = 1, \ldots, T$. We assume:

- $u_t = \psi(L)\epsilon_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}$ where $\sum_{j=0}^{\infty} j \cdot |\psi_j| < \infty$ and $\{\epsilon_t\}$ is an $i.i.d$ sequence with mean zero and variance σ_ϵ^2, and finite fourth moment. σ^2 denotes the long run variance of $\{u_t\}$ and σ_u^2 denotes the contemporaneous variance of $\{u_t\}$;
- $y_0 = 0$.

We would like to test for unit roots. So we are interested in testing

$$\mathcal{H}_0 : \ \rho = 1,$$

against the alternative

$$\mathcal{H}_1 : \ |\rho| < 1.$$

For this reason, we consider the test statistic

$$DW = \frac{T \sum_{t=1}^{T} (y_t - y_{t-1})^2}{\sum_{t=1}^{T} y_t^2}$$

(i) Derive carefully the order of probability and the limiting distribution of DW under the null hypothesis.

(ii) Obtain the order of probability of DW under the alternative hypothesis.

(iii) What do you conclude about the consistency of this test?

(b) (10 points) Suppose that we have the following data generating process (DGP) for $\{y_t\}$

$$y_t = y_{t-1} + u_t$$

for $t = 1, \ldots, T$. We assume:

- $u_t = \psi(L)\epsilon_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}$ where $\sum_{j=0}^{\infty} j \cdot |\psi_j| < \infty$ and $\{\epsilon_t\}$ is an $i.i.d$ sequence with mean zero and variance σ_ϵ^2, and finite fourth moment. σ^2 denotes the long run variance of $\{u_t\}$ and σ_u^2 denotes the contemporaneous variance of $\{u_t\}$;
- $y_0 = 0$.

We consider the estimation of the regression model

$$y_t = \delta t + error,$$

using a sample of T observations.

Consider the least squares estimator

$$\widehat{\delta} = \frac{\sum_{t=1}^{T} t y_t}{\sum_{t=1}^{T} t^2}.$$

Derive and discuss the limiting distribution $\widehat{\delta}$. Write down any necessary additional assumptions. Explain shortly what is being done in this exercise.

Question 4: Empirical Application (25 points out of 100 points)

(a) (10 points) Marty McFly is an econometrics student from Hill Valley University. He wants to analyze the relationship between the insurance claims and weather conditions. He collects annual data on related variables for the period 1955 - 2015. The retained variables are

- log of total annual insurance claims made by the residents of Hill Valley (i_t),
- log of annual average wind speed in Hill Valley (w_t)
- log of annual average rainfall in Hill Valley (r_t)

He is suspecting that there might be unit roots in the time series variables that he is considering. He is asking for an advice on how to proceed with his analysis.

(i) Explain **in detail** to Marty McFly what he should do to confirm or refute his suspicions of unit roots in the processes that he is considering.

(ii) If there are indeed unit roots in all the variables he is considering, explain **in detail** to Marty McFly how he should proceed with the choice of the model and estimation method. Do not forget to tell him about the advantages and limitations of the models/methods you are proposing.

(b) (15 points) A professor from Hill Valley University named Dr Emmett Brown is analyzing a panel data set of annual gasoline consumption of 1000 households over 4 years. Let $g_{i,t}$ denote the logarithm of annual gasoline consumption of household i in year t. Dr Emmett Brown considers the model

$$g_{i,t} = \rho g_{i,t-1} + u_{i,t}.$$

He suspects that the gasoline consumption of each household might be affected by household specific factors, so he assumes

$$u_{i,t} = \alpha_i + \epsilon_{i,t},$$

where $\epsilon_{i,t} \sim i.i.d(0, \sigma^2)$. He starts looking for an advice on how to analyze this panel data set. As he is an inventor but not an econometrician, he can really use your help.

(i) One way to estimate this model is to first apply a within transformation and estimate the transformed model by pooled OLS. This is called the fixed effects estimation. Explain this method in detail to Dr Emmett Brown.

(ii) What is the problem with the method proposed in **(i)**, explain in detail.

(iii) Propose an alternative approach to the fixed effects method that does not suffer from the problems that the fixed effects estimator suffers. Explain in detail.

End of the exam.