kolla Task 3

Solutions to TSE Exam, 141030

Sebastian Andersson
Department of Statistics, Uppsala University

November 20, 2014

Task 1

(20p in total)

1. (12p) Let $y_1, y_2, \ldots$ be a covariance stationary process with $E(y_t) = \mu$, variance $V(y_t) = \gamma_0$, $Cov(y_{t+1}, y_t) = \gamma_1$ and $\rho_1 = \frac{\gamma_1}{\gamma_0}$. Let the best one step ahead linear predictor be written on the form

$$\widehat{y_{t+1}} = \phi_0 + \phi_1 y_t$$

such that

$$E[y_{t+1} - \widehat{y}_{t+1}] = 0$$

and

$$E[(y_{t+1} - \widehat{y}_{t+1})\, y_t] = 0.$$

Show that the best predictor of Y_{t+1} is obtained by choosing

$$\phi_0 = \mu(1 - \rho_1)$$

and

$$\phi_1 = \rho_1.$$

2. (8p) Consider the AR(2) process

$$(\lambda^2 - \lambda\phi_1 - \phi_2) \leq 0$$

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + w_t \quad \Rightarrow$$

First, write the AR(2) process on an AR(1) vector form. Now, show that the stationarity condition that the roots of the characteristic polynomial is *outside* the unit circle, is equivalent to that the eigenvalues of the parameter matrix $\mathbf{F}$ (of the AR(1) vector representation) being *inside* the unit circle.

$$\begin{pmatrix} y_t \\ y_{t-1} \end{pmatrix} = \begin{pmatrix} \phi_1 & \phi_2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} y_{t-1} \\ y_{t-2} \end{pmatrix} + \begin{pmatrix} w_t \\ 0 \end{pmatrix}$$

$$= Y_t = F\, Y_{t-1} + \varepsilon_t$$

Task 2

(40p in total) Consider the following bivariate VAR:

$$
\begin{aligned}
y_{1,t} &= -0.25y_{1,t-1} + 0.125y_{1,t-2} + u_{1,t} \\
y_{2,t} &= 0.6y_{1,t-1} + 0.375y_{2,t-1} + 0.75y_{1,t-2} + 0.0625y_{2,t-2} + u_{2,t}
\end{aligned}
$$

where

$$
E(\mathbf{u}_t\mathbf{u}_s') = \begin{cases} \mathbf{\Omega}, & \text{for } t = s, \\ \mathbf{0}, & \text{for } t \neq s. \end{cases}
$$

for $\mathbf{u}_t' = (u_{1,t}, u_{2,t})$.

1. (8p) Is the system covariance-stationary?
2. (8p)The moving average weights $\mathbf{\Psi}_s$ can be calculated recursively. Write the equation for this as a first order difference equation. Using this first order difference equation, consider $\lim_{s\to\infty} \mathbf{\Psi}_s$, and explain why your result is reasonable in light of your answer to the previous question.
3. (8p) Suppose we do not know the true system, but are choosing between a VAR(2) and a VAR(4) model. For a sample of $T = 81$, both models are estimated. Test the null hypothesis that the true model is a VAR(2) against the alternative that the true model is a VAR(4), given
$$
\hat{\mathbf{\Omega}}_0 = \begin{pmatrix} 3 & 1 \\ 1 & 1.5 \end{pmatrix}, \quad \hat{\mathbf{\Omega}}_1 = \begin{pmatrix} 2 & 0.5 \\ 0.5 & 1.5 \end{pmatrix},
$$
where $\hat{\mathbf{\Omega}}_0$ and $\hat{\mathbf{\Omega}}_1$ are the estimated error covariance matrices under the null and under the alternative, respectively.
4. (16p) Given your conclusion to the previous test, pick an appropriate lag length and test that the covariance between $u_{1,t}$ and $u_{2,t}$ is equal to 0. Briefly discuss how a non-zero covariance between the error terms affects the meaning and interpretation of impulse responses. Suggest a remedy and explain the idea behind it (note: you are not required to show it mathematically).

Task 3

(40p in total) Assume the error correction representation

$$
\Delta y_t = \zeta_0 y_{t-1} + u_t, \tag{1}
$$

where $y_t = [y_{1t}\ y_{2t}\ y_{3t}]'$, ζ_0 is a 3×3 matrix of coefficients, and $u_t = [u_{1t}\ u_{2t}\ u_{3t}]'$ is a stationary vector error term.

1. (8p) If $y_t \sim I(0)$, can (1) be a cointegrated system?
2. (8p) If $y_t \sim I(1)$ but not a cointegrated system, does the representation in (1) make sense?
3. Assume

$$\zeta_0 = \begin{bmatrix} b_1 & -b_1 & -b_1 \\ b_2 & -b_2 & -b_2 \\ 0 & 0 & 0 \end{bmatrix}.$$

   (a) (8p) Show that (1) is a cointegrated system.

   (b) (8p) Show that there is only a single (independent) cointegrating vector.

   (c) (8p) Suppose that the system (1) is at the *off* equilibrium point ($y_1 = 4, y_2 = 1, y_3 = 1$). Describe the adjustment mechanisms, for the first time period, toward the equilibrium (ignore the impact of u_t).

Solution to Task 1

Part 1.1

Given the conditions, we can justify the choices for ϕ_0 and ϕ_1 by substitution.[1] From the first condition, $E(y_t - \hat{y}_t) = 0$, we get that

$$E(y_t - \hat{y}_t) = E(y_t - \phi_0 - \phi_1 y_t) = \mu - \phi_0 - \phi_1 \mu = 0$$

which means that

$$\mu(1 - \phi_1) = \phi_0.$$

Thus, given ϕ_1 we know ϕ_0. With the second condition we get another restriction:

$$\begin{aligned} E[(y_{t+1} - \hat{y}_{t+1})y_t] &= E[(y_{t+1} - \phi_0 - \phi_1 y_t)y_t] = E(y_{t+1}y_t) - \phi_0 E(y_t) - \phi_1 E(y_t^2) \\ &= \gamma_1 + \mu^2 - \phi_0\mu - \phi_1(\gamma_0 + \mu^2) \end{aligned}$$

where we can use that $\phi_0 = \mu(1 - \phi_1)$:

$$\begin{aligned} 0 = E[(y_{t+1} - \hat{y}_{t+1})y_t] &= \gamma_1 + \mu^2 - \mu(1 - \phi_1)\mu - \phi_1(\gamma_0 + \mu^2) \\ &= \gamma_1 + \mu^2 - \mu^2 + \phi_1\mu^2 - \phi_1\gamma_0 - \phi_1\mu^2 \\ &= \gamma_1 - \phi_1\gamma_0 \\ &= 0. \end{aligned}$$

[1] There were two typos in the original exam: the linear predictor should have a hat, i.e. $\hat{y}_{t+1} = \phi_0 + \phi_1 y_t$, and the second orthogonality condition should be $E[(y_{t+1} - \hat{y}_{t+1})y_t] = 0$.

Thus, we have

$$\gamma_1 = \phi_1\gamma_0,$$
$$\phi_1 = \frac{\gamma_1}{\gamma_0} = \rho_1.$$

Hence, the best linear predictor (which is the one satisfying our two conditions), is obtained by choosing

$$\begin{aligned}\phi_1 &= \rho_1 \\ \phi_0 &= \mu(1-\phi_1) = \mu(1-\rho_1).\end{aligned}$$

Part 1.2

In AR(1) vector form the AR(2) is:

$$\begin{pmatrix} y_t \\ y_{t-1}\end{pmatrix} = \begin{pmatrix}\phi_1 & \phi_2 \\ 1 & 0\end{pmatrix}\begin{pmatrix} y_{t-1} \\ y_{t-2}\end{pmatrix} + \begin{pmatrix} w_t \\ w_{t-1}\end{pmatrix}.$$

The eigenvalues are given by the solutions to

$$\left|\begin{pmatrix}\phi_1 & \phi_2 \\ 1 & 0\end{pmatrix} - \begin{pmatrix}\lambda & 0 \\ 0 & \lambda\end{pmatrix}\right| = 0,$$

that is

$$\begin{aligned}(\lambda-\phi_1)\lambda - \phi_2 &= 0 \\ \lambda^2 - \phi_1\lambda - \phi_2 &= 0 \qquad (2)\end{aligned}$$

For the characteristic polynomial, we have

$$\begin{aligned}\phi(L)y_t &= w_t \\ \phi(L) &= (1-\phi_1 L - \phi_2 L^2)\end{aligned}$$

so the roots are given by the solutions to

$$\phi(z) = (1-\phi_1 z - \phi_2 z^2) = 0. \qquad (3)$$

Go back to (2):

$$\begin{aligned}\lambda^2 - \phi_1\lambda - \phi_2 &= 0 \\ \lambda^2(1-\phi_1\lambda^{-1} - \phi_2\lambda^{-2}) &= 0\end{aligned}$$

and since $\lambda^2 \neq 0$,

$$\begin{aligned}\lambda^2(1-\phi_1\lambda^{-1} - \phi_2\lambda^{-2}) &= 0 \\ (1-\phi_1\lambda^{-1} - \phi_2\lambda^{-2}) &= 0. \qquad (4)\end{aligned}$$

which is (3) again. Equation (4) is strikingly similar to (3), and if both are to be equal to zero, it must be that $z = \lambda^{-1}$. You can see this more clearly by for example using the fact that both are equal to zero, so we can equate them:

$$\begin{aligned}(1 - \phi_1 z - \phi_2 z^2) &= (1 - \phi_1 \lambda^{-1} - \phi_2 \lambda^{-2}) \\ (1-1) - \phi_1(z - \lambda^{-1}) - \phi_2(z^2 - \lambda^{-2}) &= 0 \\ -\phi_1(z - \lambda^{-1}) - \phi_2(z^2 - \lambda^{-2}) &= 0.\end{aligned}$$

Therefore, it should be obvious that $z = \lambda^{-1}$.

Task 2

Part 1.1

In matrix form:

$$\begin{pmatrix} y_{1,t} \\ y_{2,t} \end{pmatrix} = \begin{pmatrix} -0.25 & 0 \\ 0.6 & 0.375 \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \end{pmatrix} + \begin{pmatrix} 0.125 & 0 \\ 0.75 & 0.0625 \end{pmatrix} \begin{pmatrix} y_{1,t-2} \\ y_{2,t-2} \end{pmatrix} + \begin{pmatrix} u_{1,t} \\ u_{2,t} \end{pmatrix}$$

$$\mathbf{y}_t = \mathbf{\Phi}_1 \mathbf{y}_{t-1} + \mathbf{\Phi}_2 \mathbf{y}_{t-2} + \mathbf{u}_t$$

for

$$\mathbf{y}_t = \begin{pmatrix} y_{1,t} \\ y_{2,t} \end{pmatrix}, \quad \mathbf{\Phi}_1 = \begin{pmatrix} -0.25 & 0 \\ 0.6 & 0.375 \end{pmatrix}, \quad \mathbf{\Phi}_2 = \begin{pmatrix} 0.125 & 0 \\ 0.75 & 0.0625 \end{pmatrix}.$$

From Proposition 10.1 (page 259) it follows that the system is stationary if the solutions to

$$|\mathbf{I}_2 \lambda^2 - \mathbf{\Phi}_1 \lambda - \mathbf{\Phi}_2| = 0$$

are inside the unit circle, i.e. if $|\lambda| < 1$. We have that

$$\begin{aligned}|\mathbf{I}_2 \lambda^2 - \mathbf{\Phi}_1 \lambda - \mathbf{\Phi}_2| &= \left| \begin{pmatrix} \lambda^2 & 0 \\ 0 & \lambda^2 \end{pmatrix} - \begin{pmatrix} -0.25\lambda & 0 \\ 0.6\lambda & 0.375\lambda \end{pmatrix} - \begin{pmatrix} 0.125 & 0 \\ 0.75 & 0.0625 \end{pmatrix} \right| \\ &= \begin{vmatrix} \lambda^2 + 0.25\lambda - 0.125 & 0 \\ -0.6\lambda - 0.75 & \lambda^2 - 0.375\lambda - 0.0625 \end{vmatrix} \\ &= (\lambda^2 + 0.25\lambda - 0.125)(\lambda^2 - 0.375\lambda - 0.0625) \\ &= 0.\end{aligned}$$

The first polynomial has roots

$$\begin{aligned}\lambda &= \frac{-0.25 \pm \sqrt{0.25^2 + 4 \times 0.125}}{2} = \frac{-0.25 \pm \sqrt{0.5625}}{2} \\ &= \frac{-0.25 \pm 0.75}{2}.\end{aligned}$$

so the roots are $-1/2$ and $1/4$. For the second polynomial, we get

$$\begin{aligned}\lambda &= \frac{\frac{3}{8} \pm \sqrt{\left(\frac{3}{8}\right)^2 + 4\frac{1}{16}}}{2} = \frac{\frac{3}{8} \pm \sqrt{\frac{3^2+4^2}{8^2}}}{2} \\ &= \frac{\frac{3}{8} \pm \frac{5}{8}}{2} = \frac{3}{16} \pm \frac{5}{16}\end{aligned}$$

such that the roots are $1/2$ and $-1/8$. In conclusion, we have the distinct set of roots

$$\lambda_1 = -1/2, \quad \lambda_2 = -1/8, \quad \lambda_3 = 1/4, \quad \lambda_4 = 1/2,$$

where $|\lambda_i| < 1$ for all $i = 1, 2, 3, 4$. Hence, the system is covariance-stationary.

Part 1.2

For a VAR(2), the MA weights are given by

$$\begin{aligned}\mathbf{\Psi}_1 &= \mathbf{\Phi}_1 \\ \mathbf{\Psi}_2 &= \mathbf{\Phi}_1\mathbf{\Psi}_1 + \mathbf{\Phi}_2 \\ \mathbf{\Psi}_s &= \mathbf{\Phi}_1\mathbf{\Psi}_{s-1} + \mathbf{\Phi}_2\mathbf{\Psi}_{s-2}, \quad s = 3, 4, \ldots\end{aligned}$$

If we define $\mathbf{\Psi}_0 = \mathbf{I}$ and $\mathbf{\Psi}_{-1} = \mathbf{0}$ we can use the last equation for $s = 1, 2, \ldots$. With these definitions our second-order difference equation is

$$\mathbf{\Psi}_s = \mathbf{\Phi}_1\mathbf{\Psi}_{s-1} + \mathbf{\Phi}_2\mathbf{\Psi}_{s-2}, \quad s = 1, 2, \ldots$$

To express this as a first-order difference equation, we use the companion form:

$$\begin{aligned}\begin{pmatrix}\mathbf{\Psi}_s \\ \mathbf{\Psi}_{s-1}\end{pmatrix} &= \begin{pmatrix}\mathbf{\Phi}_1 & \mathbf{\Phi}_2 \\ \mathbf{I} & \mathbf{0}\end{pmatrix}\begin{pmatrix}\mathbf{\Psi}_{s-1} \\ \mathbf{\Psi}_{s-2}\end{pmatrix} \\ \boldsymbol{\xi}_s &= \mathbf{F}\boldsymbol{\xi}_{s-1}\end{aligned}$$

where

$$\boldsymbol{\xi}_s = \begin{pmatrix}\mathbf{\Psi}_s \\ \mathbf{\Psi}_{s-1}\end{pmatrix}, \quad \mathbf{F} = \begin{pmatrix}\mathbf{\Phi}_1 & \mathbf{\Phi}_2 \\ \mathbf{I} & \mathbf{0}\end{pmatrix}.$$

This is the same $\mathbf{F}$ matrix that we get from the companion form of the VAR model (see equation [10.1.10]). Hence, by the already used Proposition 10.1 the roots from part 1 are the eigenvalues of $\mathbf{F}$. Decomposing $\mathbf{F}$ into $\mathbf{F} = \mathbf{T}\mathbf{\Lambda}\mathbf{T}^{-1}$, where $\mathbf{\Lambda}$ is a diagonal matrix containing the eigenvalues. By recursive substitution we find

$$\begin{aligned}\boldsymbol{\xi}_s &= \mathbf{F}\boldsymbol{\xi}_{s-1} = \mathbf{F}^2\boldsymbol{\xi}_{s-2} = \cdots = \mathbf{F}^s\boldsymbol{\xi}_0 \\ &= \mathbf{T}\mathbf{\Lambda}^s\mathbf{T}^{-1}.\end{aligned}$$

Since

$$\lim_{s\to\infty} \mathbf{\Lambda}^s = \begin{pmatrix} -1/2^s & 0 & 0 & 0 \\ 0 & -1/8^s & 0 & 0 \\ 0 & 0 & 1/4^s & 0 \\ 0 & 0 & 0 & 1/2^s \end{pmatrix} \longrightarrow \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Thus,

$$\lim_{s\to\infty} \boldsymbol{\xi}_s = \begin{pmatrix} \lim_{s\to\infty} \mathbf{\Psi}_s \\ \lim_{s\to\infty} \mathbf{\Psi}_{s-1} \end{pmatrix} = \mathbf{0}.$$

This means that the moving average weights for increasing s will be smaller and smaller. If we interpret these as dynamic multipliers, then this is a very reasonable result that is consistent with the result of stationarity in part 1. If they would not tend to zero, then shocks in the very distant future would have persistent (if we have at least one $|\lambda_i| = 1$) or explosive (if at least one $|\lambda_i| > 1$) effects on the current value of the process. So that the moving average weights decrease is therefore expected.

Part 1.3

From p. 296-297, we have the likelihood ratio test of $H_0 : p = p_0$ vs. $H_1 : p = p_1$, where $p_1 > p_0$. The test statistic is

$$2(\log L_1^* - \log L_0^*) = T(\log |\hat{\mathbf{\Omega}}_0| - \log |\hat{\mathbf{\Omega}}_1|)$$

which under the null is asymptotically χ^2 with degrees of freedom equal to the number of restrictions. There are n equations, and in each equation $n(p_1 - p_0)$ zero restrictions are imposed under the null. Thus, there are in total $n^2(p_1 - p_0)$.

In our case, $n = 2$, $p_1 = 4$ and $p_0 = 2$. Hence, the test statistic should be compared to quantiles of a $\chi^2(8)$ distribution. From the information given, we get

$$\log |\hat{\mathbf{\Omega}}_0| = \log 3.5 \approx 1.25$$
$$\log |\hat{\mathbf{\Omega}}_1| = \log 2.75 \approx 1.01.$$

The test is

$$2(\log L_1^* - \log L_0^*) = T(\log |\hat{\mathbf{\Omega}}_0| - \log |\hat{\mathbf{\Omega}}_1|) \approx 81 \times (1.25 - 1.01) = 19.44.$$

In a $\chi^2(8)$ distribution and a significance level of 5 %, the critical value is 15.5, and thus we reject the null hypothesis in this case. A model with four lags appears to fit the data better.

Part 1.4

We choose the model with four lags, in which the estimated error covariance is 0.5. To test this, we can use the results on p. 301. The hypotheses are given by $H_0 : \sigma_{12} = 0$ vs $H_1 : \sigma_{12} \neq 0$, where $\sigma_{12} = E(u_{1,t}u_{2,t})$. The test is given by

$$\frac{\sqrt{T}\hat{\sigma}_{12}}{(\hat{\sigma}_{11}\hat{\sigma}_{22} + \hat{\sigma}_{12}^2)^{1/2}} \approx N(0, 1).$$

The result is

$$\frac{9 \times 0.5}{(3 + 0.25)^{1/2}} \approx 2.48.$$

This exceeds the right-tail critical value of 1.96, so we reject the null hypothesis. The test indicates that there is a non-zero covariance between the error terms.

The impulse responses are affected in the sense that the elements of the moving average weights $\mathbf{\Psi}_s$ describe the responses of $y_{i,t+s}$ to a one-time shock in $u_{j,t}$. If the error covariance is actually zero, then these are easily interpreted. But, if the covariance is non-zero, as in this case, a shock to $u_{1,t}$ tells us something about $u_{2,t}$ as well. Thus, since the shocks are not isolated it does not make sense to interpret them as such. One possible remedy is orthogonalization of the errors. This can be done by using a Cholesky decomposition of the error covariance matrix. The idea is that we by this approach can obtain shocks which are in fact uncorrelated, and by such an adjustment we can interpret the impulse responses as responses to these isolated shocks.

Task 3

Part 3.1

If $y_t \sim I(0)$, then the system cannot be cointegrated. A vector y_t is said to be cointegrated (p. 574) if its elements are individually $I(1)$ and if there exists a non-zero vector a such that $a'y_t$ is stationary. The first condition, that the elements are $I(1)$ is not fulfilled, so the system cannot be cointegrated.

Part 3.2

If $y_t \sim I(1)$ but not cointegrated, the representation does not make sense. This is because if $y_t \sim I(1)$ but not cointegrated, then the elements are individual random walks:

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

The matrix ζ_0 is defined as

$$\zeta_0 = \Phi_1 + \Phi_2 + \cdots + \Phi_p - I_n,$$

where we have from the representation above that $p = 1$ and $\Phi_1 = I_n$. It thus follows that

$$\zeta_0 = I_n - I_n = 0$$

and the representation including ζ_0 does not make sense as it should be excluded in this case.

Part 3.3

Part 3.3.a

To show that it is a cointegrated system, we decompose the matrix into two reduced rank matrices:

$$\zeta_0 = \begin{bmatrix} b_1 & -b_1 & -b_1 \\ b_2 & -b_2 & -b_2 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & -1 & -1 \end{bmatrix}$$

The reduced rank implies that there is cointegration. Alternatively, one can see that $\det(\zeta_0) = 0$, which is the same as reduced rank.

Part 3.3.b

From the decomposition above, it is evident that there is only one linearly independent cointegrating vector as the dimension of the reduced rank matrices is 3×1 (and 1×3). This means that h, the number of cointegrating relations, is equal to 1. This can also be seen from ζ_0 directly, as multiplying the first row by b_2/b_1 gives the second row, meaning that there is a linear dependence.

Part 3.3.c

The adjustment that occurs is given by

$$\begin{bmatrix} \Delta y_1 \\ \Delta y_2 \\ \Delta y_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} 4 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2b_1 \\ 2b_2 \\ 0 \end{bmatrix}$$

meaning that y_1 is adjusted by $2b_1$, y_2 by $2b_2$ and y_3 is unadjusted.