Second midterm test in Advanced Econometrics: Tentative answers

Robert M. Kunst

May 23, 2013

1. A vector autoregression is given in its standard form ($(\varepsilon_{1t}, \varepsilon_{2t})'$ is white noise)
$$\begin{bmatrix} X_t \\ Y_t \end{bmatrix} = \begin{bmatrix} 0.5 & 0.5 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_{t-1} \\ Y_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}.$$

   (a) Defining $x = (X, Y)'$, transform the system to its error-correction form $\Delta x_t = \mathbf{\Pi} x_{t-1} + \varepsilon_t$ and determine $\mathbf{\Pi} = \gamma\beta'$, γ, and β. What are the cointegrating vectors in this example? What is the cointegrating rank here?

   Answer: When transformed to the EC form, the system becomes
$$\begin{bmatrix} \Delta X_t \\ \Delta Y_t \end{bmatrix} = \begin{bmatrix} -0.5 & 0.5 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} X_{t-1} \\ Y_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}.$$
   The matrix
$$\begin{bmatrix} -0.5 & 0.5 \\ 0 & 0 \end{bmatrix}$$
   is the impact matrix Π, which clearly has rank one, which is then also the *cointegrating rank*. One possible representation $\Pi = \gamma\alpha'$ is $\gamma = (0.5, 0)'$ and $\alpha = (-1, 1)'$. Cointegrating vectors are multiples of α: $X_t - Y_t$ is stationary.

   (b) The characteristic polynomial for a cointegrated VAR has some unit roots and some stable roots. What is this characteristic polynomial for the system in (a)?

   Answer: Usually, the scalar polynomial $\det \Theta(z)$ is called the characteristic polynomial, although the word may also apply to the matrix polynomial $\Theta(z)$. The former one implies
$$\det \Theta(z) = (1 - 0.5z)(1 - z),$$

which has the two roots $\zeta = 1$ and $\zeta = 2$, one unstable and one stable root, thus confirming cointegration in the system.

2. In a vector autoregressive system of dimension 3, you find evidence that the cointegrating rank is 1. Inspired by this finding, you regress the first variable X on the other two variables Y and Z. You think that this is a cointegrating regression.

   (a) Why can this procedure fail entirely? Presuming that the DF test on the residuals fails to reject, can it make sense to try again and to regress Y on X and Z?

   Answer: Yes. Cointegration among the three variables X, Y, Z does not necessarily imply that X is included in the relation. If the coefficient on X is zero, the outlined regression fails to yield a consistent estimate of the cointegrating vector. However, at least one variable must be in that vector. Regressing Y on X, Z may do the trick.

   (b) If you find autocorrelation in the residuals of this cointegrating regression, do you think it makes sense to add lags of Y and Z to the regressors?

   Answer: No. There is some interest in dynamic cointegration in special applications such as seasonal integration or I(2), but within the outlined field cointegration is static in principle. Autocorrelation in the residuals does not invalidate cointegration, and the regression does not improve by GLS steps or inclusion of lags. Just to the contrary, after including lagged level terms, the regression becomes unusable. Without those lags, the method is inefficient but correct.

   (c) Imagine you join another variable to the system, such that it has a dimension of 4. If that variable is stationary, what is the cointegrating rank of the 4–dimensional system?

   Answer: Two. A stationary variable in an otherwise I(1) system defines a formal cointegrating unit vector. This is called 'self-cointegration'. In the four-variables system, there are two cointegrating vectors: the original in the subsystem of the first three variables and the unit vector for self-cointegration of the fourth variable.

3. A panel consists of 100 yearly measurements at five woodland locations. Available variables are an index of tree species diversity and

three covariates: concentration of carbon dioxide in the air, concentration of sulphur in the air, average temperature over the year. You wish to determine the effects of the pollutants and of climate on the tree diversity.

(a) Do you feel that a fixed-effects or a random-effects model is more adequate here? Why?

Answer: Random-effect estimation can be supported by some arguments in this case, but the more natural suggestion is to use fixed-effect estimation. The cross-section dimension is small, and the time dimension is so large that RE estimates will be close to the simpler FE estimate.

(b) You run a Hausman test, and the test rejects. Would you revise your answer to (a) in the light of this result?

Answer: If (a) has been answered as suggested above, there will be no change, as rejection by the Hausman test implies a preference for the FE estimator. Note that the null of the Hausman test is that a RE model is valid, with effects and covariates uncorrelated, which would cause the RE estimator to be consistent and even efficient. Under the alternative, RE becomes inconsistent, while FE remains consistent.

(c) Try and provide an interpretation of the individual 'effect' in this example.

Answer: Several suggestions are possible here. The effects should be characteristics of the location, such as omitted indicators (precipitation, amplitude of the temperature cycle, soil) that may be influential additionally to the given covariates, or the policy pursued by forestry. The effects should not reflect the covariates, such as air-polluting factories.