E 4101/5101
Lecture 10: The cointegrated VAR—representation

Ragnar Nymoen

Department of Economics, University of Oslo

9 April 2014

Introduction I

So far we have considered:

- Stationary VAR, with deterministic extensions ("no unit roots")
  - Standard inference for dynamic models
- Non-stationary VAR with independent variables ("all unit-roots")
  - Danger of spurious relationships

We now develop the theory of *cointegration*, the case of "some, but not only" unit-roots in the VAR.

- In such systems, there exist one or more linear combinations of $I(1)$ variables that are $I(0)$—they are called *cointegration relationships*.

Introduction II

- We see that cointegration is the "flip of the coin" of spurious regression: If we have two dependent $I(1)$ variables, they are cointegrated.
- We can also guess that a test of the null hypothesis of no cointegration is going to be of the Dickey-Fuller type.
  - This is true when the cointegration relationship is unique.
- However, we want to develop the theory of cointegration more fully, and "we take it" gradually:
  - The cointegrated VAR: The representation of VARs with some but no all unit-roots (this lecture)
  - Testing the null-hypothesis of no cointegration
    - The cointegrating regression
    - The conditional ECM
    - VAR methods, testing hypotheses about rank reduction

Introduction III

- Estimating the cointegrated VAR.

References (for this and following lectures on cointegration)

- Hamilton Ch 18, 19, 20
- Davidson and MacKinnon Ch 14
- The first part of this lecture builds on a separate lecture note about the so called *representation theorem for cointegrated variables*.

The VAR with a unit root I

The Lecture note on Engle and Grangers representation theorem for the bi-variate VAR(1)

$$\mathbf{y}_t = \mathbf{\Phi}\mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t \tag{1}$$

where $\mathbf{y}_t = (x_t, y_t)$, $\mathbf{\Phi}$ is a 2×2 matrix with coefficients and $\boldsymbol{\varepsilon}_t$ is a vector with Gaussian disturbances.
The characteristic equation of $\mathbf{\Phi}$:

$$|\mathbf{\Phi} - z\mathbf{I}| = 0,$$

We consider the intermediate case of one unit-root and one stationary root. Specifically

$$z_1 = 1, \text{and } z_2 = \lambda, |\lambda| < 1. \tag{2}$$

The VAR with a unit root II

implying that both x_t and y_t are $I(1)$.
$\mathbf{\Phi}$ has full rank, equal to 2. Diagonalize in terms of its eigenvalues and the corresponding eigenvectors:

$$\mathbf{\Phi} = \mathbf{P} \begin{bmatrix} 1 & 0 \\ 0 & \lambda \end{bmatrix} \mathbf{Q} \tag{3}$$

$\mathbf{P}$ has the eigenvectors as columns:

$$\mathbf{P} = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}. \tag{4}$$

and $\mathbf{Q} = \mathbf{P}^{-1}$.

Cointegration and the Common Trends representation I

(1) with (2) implies

$$\begin{bmatrix} w_t \\ -z_t \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & \lambda \end{bmatrix} \begin{bmatrix} w_{t-1} \\ -z_{t-1} \end{bmatrix} + \boldsymbol{\eta}_t, \quad (5)$$

where $\boldsymbol{\eta}_t$ contains linear combinations of the original VAR disturbances. (Notation: z_t denotes a time series, despite the slight conflict of notation, with the roots in (2)).

- $w_t \sim I(1)$, a stochastic trend
- $z_t \sim I(0)$,notation: When z_t denotes a time series, despite the slight conflict with notation with the roots in (2)

Cointegration and the Common Trends representation II

The expression for $-z_t$ is

$$z_t = -\gamma x_t + \alpha y_t. \tag{6}$$

- There is *cointegration* between x_t and y_t,since z_t is a stationary variable, and it is a linear combination of x_t and y_t.
- $-\gamma$ and α are the cointegrating parameters in this example model

The Common Trends representation for x_t and y_t is:

$$x_t = \alpha w_t - \beta z_t \tag{7}$$

$$y_t = \gamma w_t - \delta z_t. \tag{8}$$

- x_t and y_t have one common stochastic trend, which is w_t.

Cointegration and the Common Trends representation III

"Corollaries"

1. Forecasts for $x_{T+h|T}$ and $y_{T+h|T}$ become dominated by the common stochastic trend
2. Cointegration is maintained in the forecasts, so $z_{T+h|T} = -\gamma x_{T+h|T} + \alpha y_{T+h|T} = 0$ for large h.

The MA representation I

$$\begin{bmatrix} \Delta x_t \\ \Delta y_t \end{bmatrix} = \boldsymbol{\Psi}(L)\boldsymbol{\eta}_t \tag{9}$$

where

$$\boldsymbol{\Psi}(L) = \begin{bmatrix} \alpha & \beta(1-L)/(1-\lambda L) \\ \gamma & \delta(1-L)/(1-\lambda L) \end{bmatrix}. \tag{10}$$

- The rank of $\boldsymbol{\Psi}(1)$ is 1, since $|\boldsymbol{\Psi}(1) - z\mathbf{I}| = (\alpha - z) \cdot z = 0$ have roots $z_1 = 0$ and $z_2 = \alpha$.
- Equivalently, $|\boldsymbol{\Psi}(1)| = 0$ and $|\alpha| \neq 0)$. The MA-matrix thus has *reduced rank* for $L = 1$, as a result of cointegration.

The ECM representation I

Can express $\boldsymbol{\Phi}$ as

$$\boldsymbol{\Phi} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + (1-\lambda) \begin{bmatrix} \beta \\ \delta \end{bmatrix} \begin{bmatrix} \gamma & -\alpha \end{bmatrix}. \tag{11}$$

to give

$$\begin{bmatrix} \Delta x_t \\ \Delta y_t \end{bmatrix} = (1-\lambda) \begin{bmatrix} \beta \\ \delta \end{bmatrix} \underbrace{[\gamma x_{t-1} - \alpha y_{t-1}]}_{-z_{t-1}} + \boldsymbol{\varepsilon}_t$$

$$\begin{bmatrix} \Delta x_t \\ \Delta y_t \end{bmatrix} = \boldsymbol{\alpha}\boldsymbol{\beta}' \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \boldsymbol{\varepsilon}_t, \tag{12}$$

The ECM representation II

We adopt a very popular notation:

- $\boldsymbol{\alpha}$ is known as the matrix)of equilibrium correction coefficients (loadings), here

$$\boldsymbol{\alpha} = \begin{bmatrix} (1-\lambda)\beta \\ (1-\lambda)\delta \end{bmatrix} \tag{13}$$

- $\boldsymbol{\beta}$ is the matrix of equilibrium correction coefficients, here

$$\boldsymbol{\beta} = \begin{bmatrix} \gamma \\ -\alpha \end{bmatrix}. \tag{14}$$

We often write (1) as

$$\Delta \mathbf{y}_t = \boldsymbol{\Pi} \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t$$

The ECM representation III

where

$$\mathbf{\Pi} = (\mathbf{\Phi} - \mathbf{I})$$

In this formulation we see that

- $rank(\mathbf{\Pi}) = 0$, reduced rank and no cointegration. Both eigenvalues are zero.
- $rank(\mathbf{\Pi}) = 1$, reduced rank and cointegration. One eigenvalue is different from zero.
- $rank(\mathbf{\Pi}) = 2$, full rank, both eigenvalues are different from zero and the VAR (1) is stationary.

The ECM representation IV

Cointegration and Granger causality

Since $\lambda < 1$ is equivalent with cointegration, we see from (13) that cointegration also implies Granger-causality in at least one direction: $(1-\lambda)\beta \neq 0$ and/or $(1-\lambda)\beta \neq 0$.

Cointegration and weak exogeneity

- Assume $\delta = 0$, from (13). this implies

$$\begin{bmatrix} \Delta x_t \\ \Delta y_t \end{bmatrix} = (1-\lambda) \begin{bmatrix} \beta \\ 0 \end{bmatrix} [\gamma x_{t-1} - \alpha y_{t-1}] + \boldsymbol{\varepsilon}_t$$

$$\begin{bmatrix} \Delta x_t \\ \Delta y_t \end{bmatrix} = \begin{bmatrix} (1-\lambda)\beta[\gamma x_{t-1} - \alpha y_{t-1}] + \varepsilon_{x,t} \\ \varepsilon_{y,t} \end{bmatrix}$$

- The marginal model contains no information about the cointegration parameters $(\gamma, -\alpha)'$. y_t is WE for $(\gamma, -\alpha)'$.

VAR(p) —> ECM I

$\mathbf{y}_t$ is $n \times 1$ with $I(1)$ variables. The VAR is:

$$\mathbf{y}_t = \mathbf{\Phi}(L)\mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t$$

where $\boldsymbol{\varepsilon}_t$ is multivariate Gaussian and

$$\mathbf{\Phi}(L) = \sum_{i=0}^{p} \mathbf{\Phi}_{i+1} L^i \tag{15}$$

In analogy to the scaler case, the matrix lag-polynomial is written

$$\mathbf{\Phi}(L) = \mathbf{\Phi}(1) + \Delta\mathbf{\Phi}^*(L)$$

where the $\mathbf{\Phi}_i^*$ matrices

$$\mathbf{\Phi}^*(L) = \mathbf{\Phi}_1^* + \mathbf{\Phi}_2^* L + \ldots + \mathbf{\Phi}_{p-1}^* L^{p-1}$$

VAR(p) —> ECM II

are linear transformations of $\mathbf{\Phi}_i$ $(i = 1, \ldots, p)$. Substitution yields

$$\begin{aligned} \mathbf{y}_t &= \mathbf{\Phi}^*(L)\Delta \mathbf{y}_{t-1} + \mathbf{\Phi}(1)\mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t \\ \Delta \mathbf{y}_t &= \mathbf{\Phi}^*(L)\Delta \mathbf{y}_{t-1} + \mathbf{\Pi}(1)\mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t \end{aligned} \tag{16}$$

where $\mathbf{\Pi}(1) \equiv \mathbf{\Phi}(1) - \mathbf{I}_N = \mathbf{0}$ in the case of no cointegration but

$$\mathbf{\Pi}(1) = \boldsymbol{\alpha}\boldsymbol{\beta}' \tag{17}$$

in the case of r cointegrating-vectors.

- $\boldsymbol{\beta}_{n \times r}$ contains the CI-vectors as columns, while $\boldsymbol{\alpha}_{n \times r}$ shows the strength of equilibrium correction in each of the equations for $\Delta y_{1t}, \Delta y_{2t}, \ldots, \Delta y_{nt}$. In general rank$(\boldsymbol{\beta}) = r$ and rank$(\mathbf{\Pi}) = r < n$.

VAR(p) —> ECM III

- If $\boldsymbol{\beta}$ is known, the system

$$\Delta \mathbf{y}_t = \mathbf{\Phi}^*(L)\Delta \mathbf{y}_{t-1} + \boldsymbol{\alpha}[\boldsymbol{\beta}'\mathbf{y}]_{t-1} + \boldsymbol{\varepsilon}_t \qquad (18)$$

  contains only $I(0)$ variables and conventional asymptotic inference applies.
- Moreover: If $\boldsymbol{\beta}$ can be regarded as known, *after first estimating* $\boldsymbol{\beta}$, conventional asymptotic inference also applies.
- (18) is then a stationary VAR, called the VAR-ECM or the cointegrated VAR.
- This system can be identified and modelled with the concepts that we developed for the stationary case

Restricted and unrestricted constant term I

Consider the bivariate case again with cointegration notation and assume $r = 1$:

$$\begin{bmatrix} \Delta y_{1t} \\ \Delta y_{2t} \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} + \begin{bmatrix} \alpha_{11} \\ \alpha_{21} \end{bmatrix} [\beta_{11} y_{1t-1} + \beta_{12} y_{2t-1}] + \varepsilon_t \quad (19)$$

The solution will imply that y_{1t} and y_{2t} contain trends, but the same trend even if $\mu_1 \neq \mu_2$.
There is an exception to this rule unless (Hamilton p. 581). This case is:

$$\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = - \begin{bmatrix} \alpha_{11} \\ \alpha_{21} \end{bmatrix} \underbrace{E [\beta_{11} y_{1t-1} + \beta_{12} y_{2t-1}]}_{=\mu_0}. \quad (20)$$

Restricted and unrestricted constant term II

Notice:

$$\mu_0 = E\left[\beta_{11} y_{1t-1} + \beta_{12} y_{2t-1}\right] \quad (21)$$

$E\left[\beta_{11} y_{1t-1} + \beta_{12} y_{2t-1}\right]$ is the mean of a stationary variable.
Insert (20) in (19):

$$\begin{bmatrix} \Delta y_{1t} \\ \Delta y_{2t} \end{bmatrix} = -\begin{bmatrix} \alpha_{11} \\ \alpha_{21} \end{bmatrix} \mu_0 + \begin{bmatrix} \alpha_{11} \\ \alpha_{21} \end{bmatrix} \left[\beta_{11} y_{1t-1} + \beta_{12} y_{2t-1}\right] + \boldsymbol{\varepsilon}_t$$

$$\begin{bmatrix} \Delta y_{1t} \\ \Delta y_{2t} \end{bmatrix} = \begin{bmatrix} \alpha_{11} \\ \alpha_{21} \end{bmatrix} \left[-\mu_0 + \boldsymbol{\beta}' \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix}\right] + \boldsymbol{\varepsilon}_t \quad (22)$$

Taking expectations on both sides of (22) gives:

$$E\begin{bmatrix} \Delta y_{1t} \\ \Delta y_{2t} \end{bmatrix} = \begin{bmatrix} \alpha_{11} \\ \alpha_{21} \end{bmatrix} (-\mu_0 + \mu_0) = 0$$

Restricted and unrestricted constant term III

showing that with restricted intercepts there are no drifts and therefore no trend in the levels variables.
We can write the intercept-restricted model as:

$$\begin{bmatrix} \Delta y_{1t} \\ \Delta y_{2t} \end{bmatrix} = \begin{bmatrix} \alpha_{11} \\ \alpha_{21} \end{bmatrix} \left[\mu_0 : \boldsymbol{\beta}'\right] \begin{bmatrix} 1 \\ y_{1t-1} \\ y_{2t-1} \end{bmatrix} + \boldsymbol{\varepsilon}_t, \tag{23}$$

- We often say that the intercept is restricted to be in the cointegrating vector (or in the cointegrating space).
- The decision to restrict the intercept (or not) in this way is seen to have important implications.
  - Also because different critical values apply when testing.
- A similar argument arise if a trend is included in the VAR.

Restricted and unrestricted constant term IV

- If a squared trend in the levels of variables is to be avoided, the deterministic trend has to be restricted to lies the *cointegrating space* (be part of the cointegrating vector).

Conditional ECM I

Assume that $\alpha_{21} = 0$, i.e. y_{2t} is weakly exogenous for $\boldsymbol{\beta}$.
With Gaussian disturbances $\boldsymbol{\varepsilon}_t = N(0, \boldsymbol{\Omega})$, where $\boldsymbol{\Omega}$ has elements ω_{ij},we can derive the conditional model for Δy_t:

$$\Delta y_{1t} = \underbrace{\omega_{21}\omega_{22}^{-1}}_{b}\Delta y_{2t} + \alpha_{11}\boldsymbol{\beta}' \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix} + \underbrace{\varepsilon_{1t} - \omega_{21}\omega_{22}^{-1}\varepsilon_{2t}}_{u_t} \quad (24)$$

the single equation ECM we have discussed before.
(24) is an example of an open system, since x_{t-1} is determined outside the model.
If we write it is as

$$\Delta y_{1t} = b\Delta y_{2t} + \alpha_{11}\beta_{11}y_{1t-1} + \alpha_{11}\beta_{12}y_{2t-1} + u_t$$

we see that $\boldsymbol{\Pi}$=$\alpha_{11}\beta_{11} \neq 0$, i.e., the $\boldsymbol{\Pi}$ “matrix” has full rank.

Conditional ECM II

- Open system are often relevant, ideally after first testing $\alpha_{21} = 0$, but also when this is difficult.
- The common trend is now in the non-modelled variable y_{2t-1}.
- Care must be taken: The relevant distribution for testing rank$(\mathbf{\Pi}) = 0$ is (as we shall see) different from the distribution that applies for the closed system.
- *Generalization:* Open systems can of course contain n_1 endogenous $I(1)$ variables and n_2 non-modelled $I(1)$ variables. Cointegration is then consistent with

$$0 < rank(\mathbf{\Pi}) \leq n_1$$

Identification I

- When $n = 2$, cointegration implies $rank(\mathbf{\Pi}) = 1$
  - There is one cointegration vector

$$(\beta_{11}, \beta_{12})'$$

    which is uniquely identified after normalization. For example with $\beta_{11} = -1$ the ECM variable becomes

$$ecm_{1t} = -y_{1t} + \beta_{12} y_{2t} \sim I(0)$$

- When $n > 2$, we can have $rank(\mathbf{\Pi}) > 1$, and in these case the cointegrating vectors are not identified.

Identification II

- Assume that $\mathbf{\Pi}$ is known (in practice, consistently estimated), and $\boldsymbol{\beta}$ is a $n \times r$ cointegrating vector:

$$\mathbf{\Pi} = \boldsymbol{\alpha}\boldsymbol{\beta}'$$

However for a $r \times r$ non-singular matrix $\mathbf{\Theta}$:

$$\mathbf{\Pi} = \boldsymbol{\alpha}\mathbf{\Theta}\mathbf{\Theta}^{-1}\boldsymbol{\beta}' = \boldsymbol{\alpha}_{\Theta}\boldsymbol{\beta}'_{\Theta}$$

showing that $\boldsymbol{\beta}'_{\Theta}$ is also a cointegrating vector.

This problem is equivalent to the identification problem in simultaneous equation models.

- Assume $rank(\mathbf{\Pi}) = 2$ for a $n = 3$ VAR

$$-y_{1t} + \beta_{12}y_{2t} + \beta_{13}y_{3t} = ecm_{1t}$$
$$\beta_{21}y_{1t} - y_{2t} + \beta_{13}y_{3t} = ecm_{2t}$$

Identification III

- By simply viewing these as a pair of simultaneous equations, we see that they are not identified on the order-condition.
- Exact identification requires for example 1 linear restrictions on each of the equations.
  - For example $\beta_{13} = 0$ and $\beta_{21} + \beta_{13} = 0$ will result in exact identification
  - Identification = theory !!!
- Restrictions of the loading matrix can also help identification (then hypotheses about causation?)
- A very useful estimator of $\mathbf{\Pi}$ is the Maximum-Likelihood estimator (OLS on each equation in the VAR). A natural test-statistic for any overidentifying restrictions is the LR test.
- The identification issue applies equally for open systems. Again, in direct analogy to the simultaneous equation model.

References

Engle, R. F. and C.W.J. Granger (1987) Co-integration and Error-Correction: Representation. Estimation and Testing, *Econometrica*, 55, 251-276