# The Engle-Granger representation theorem

**Reference note to lecture 10 in ECON 5101, Time Series Econometrics***

Ragnar Nymoen

9 April 2014

## 1 Introduction

The Granger-Engle representation theorem is in their *Econometrica* paper from 1987.

The main thesis is that systems with cointegrated $I(1)$ variables have three equivalent representations

- Common Trends
- Moving average, MA
- Equilibrium correction (*aka* error correction), ECM

In this note we give a simplified version of this famous theorem. We start from the VAR-representation of a first order bi-variate system and assume that one of the roots of the autoregressive matrix is unity while the other root is less than one. The Common Trends, MA and ECM representations are then shown. Note that in the literature, the theorem is also shown "the other way", by starting from the MA representation.

As noted we use a set-up with first order dynamics and only two time series variables. We abstract from deterministic variables. This enough to show the different representastions of cointegrated systems that have become standard in the literature. The consequences of higher order dynamics, higher dimensionality and deterministic terms are discussed in the lecture.

## 2 The VAR

Consider the bivariate *VAR(1)*

$$\boldsymbol{y}_t = \boldsymbol{\Phi}\boldsymbol{y}_{t-1} + \boldsymbol{\varepsilon}_t \tag{2.1}$$

where $\boldsymbol{y}_t = (x_t, y_t)$, $\boldsymbol{\Phi}$ is a $2 \times 2$ matrix with coefficients and $\boldsymbol{\varepsilon}_t$ is a vector with two Gaussian disturbances.

The characteristic equation of $\boldsymbol{\Phi}$:

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

*This lecture note is a translation of Chapter 11.2 in Bårdsen G. and R. Nymoen *Videregående emner i økonometri*, 2014, Fagbokforlaget

In the spurious regression case we imposed two unit-roots. In the stationary case neither root is equal to one. We now consider the intermediate case of one unit-root and one stationary root. Specifically, write the two characteristic roots as:

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

$z_1 = 1$ implies that both $x_t$ and $y_t$ are $I(1)$. $\boldsymbol{\Phi}$ has full rank, equal to 2, since both eigenvalues are different from zero. $\boldsymbol{\Phi}$ can therefore be diagonalized in terms of its eigenvalues and the corresponding eigenvectors. We write it as:

$$\boldsymbol{\Phi} = \boldsymbol{P} \begin{bmatrix} 1 & 0 \\ 0 & \lambda \end{bmatrix} \boldsymbol{Q} \tag{2.2}$$

where $\boldsymbol{P}$ is the matrix with the eigenvectors as columns:

$$\boldsymbol{P} = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}. \tag{2.3}$$

and $\boldsymbol{Q} = \boldsymbol{P}^{-1}$. Assume, without loss of generality that $|\boldsymbol{P}| = 1$, i.e.,

$$\alpha\delta - \gamma\beta = 1. \tag{2.4}$$

and thus

$$\boldsymbol{P}^{-1} = \begin{bmatrix} \delta & -\beta \\ -\gamma & \alpha \end{bmatrix} = \boldsymbol{\Phi}$$

By using this in (2.2), $\boldsymbol{\Phi}$ can be expressed as

$$\boldsymbol{\Phi} = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} \begin{bmatrix} \delta & -\beta \\ -\gamma\lambda & \alpha\lambda \end{bmatrix}. \tag{2.5}$$

By carrying out the multiplication on the left hand side of the equality sign, we also get:

$$\boldsymbol{\Phi} = \begin{bmatrix} (\alpha\delta - \lambda\beta\gamma) & -\alpha\beta(1-\lambda) \\ \gamma\delta(1-\lambda) & (-\gamma\beta + \lambda\alpha\delta) \end{bmatrix}. \tag{2.6}$$

We refer to these expressions later.

## 2.1 The Common-trends representation

Multiplication of (2.1) with $\boldsymbol{Q}$ yields

$$\boldsymbol{Q}\boldsymbol{x}_t = \boldsymbol{Q}\boldsymbol{\Phi}\boldsymbol{x}_{t-1} + \boldsymbol{Q}\boldsymbol{\varepsilon}_t. \tag{2.7}$$

Observe that

$$\boldsymbol{Q}\boldsymbol{\Phi} = \boldsymbol{Q}\boldsymbol{P} \begin{bmatrix} 1 & 0 \\ 0 & \lambda \end{bmatrix} \boldsymbol{Q} = \begin{bmatrix} 1 & 0 \\ 0 & \lambda \end{bmatrix} \boldsymbol{Q},$$

from the definitions above, implying that (2.7) can be written in terms of two new variables $w_t$ and $z_t$:

$$\begin{bmatrix} w_t \\ -z_t \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & \lambda \end{bmatrix} \begin{bmatrix} w_{t-1} \\ -z_{t-1} \end{bmatrix} + \boldsymbol{\eta}_t, \tag{2.8}$$

where

$$\begin{bmatrix} w_t \\ -z_t \end{bmatrix} = \boldsymbol{Q} \begin{bmatrix} x_t \\ y_t \end{bmatrix} \tag{2.9}$$

and

$$\boldsymbol{Q}\boldsymbol{\varepsilon}_t = \boldsymbol{\eta}_t = \begin{bmatrix} \eta_{1t} \\ \eta_{2t} \end{bmatrix}. \tag{2.10}$$

To solve for $w_t$ and $z_t$, we multiply out the right hand side of (2.9). Notice that since

$$\boldsymbol{Q} = \begin{bmatrix} \delta & -\beta \\ -\gamma & \alpha \end{bmatrix} \tag{2.11}$$

we obtain $w_t$ as

$$w_t = \delta x_t - \beta y_t \tag{2.12}$$

and $z_t$ as:

$$-z_t = \gamma x_t - \alpha y_t. \tag{2.13}$$

Conversely, from (2.9):

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \boldsymbol{P} \begin{bmatrix} w_t \\ -z_t \end{bmatrix}$$

yielding

$$x_t = \alpha w_t - \beta z_t \tag{2.14}$$

and

$$y_t = \gamma w_t - \delta z_t. \tag{2.15}$$

Equation (2.8) gives $w_t$ as a random-walk variable and, from the same equation, $z_t$ is seen to be stationary since $|\lambda| < 1$.

$z_t$, defined by (2.13) is therefore a stationary linear combination of the two $I(1)$ variables $x_t$ and $y_t$. We say that $x_t$ and $y_t$ are two cointegrated variables, with cointegrating vector $(\gamma, -\alpha)'$.

On the other hand: equation (2.14) and (2.15) show that $x_t$ and $y_t$ contain an $I(0)$ component ($z_t$), *and* an $I(1)$ component ($w_t$)—hence the name Common trend representation.

## 2.2 Integration and cointegration is maintained in forecasts

Let $\boldsymbol{y}_{T,h}$ denote the $h$-period ahead forecast, conditional on $\boldsymbol{x}_T$. From (2.2) we obtain:

$$\boldsymbol{y}_{T,h} = \boldsymbol{P} \begin{bmatrix} 1 & 0 \\ 0 & \lambda^h \end{bmatrix} \boldsymbol{Q}\boldsymbol{y}_T = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \lambda^h \end{bmatrix} \begin{bmatrix} w_T \\ -z_T \end{bmatrix}. \tag{2.16}$$

and, if $h$ is large enough

$$\boldsymbol{y}_{T,h} = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} w_T \\ -z_T \end{bmatrix} = \begin{bmatrix} \alpha \\ \gamma \end{bmatrix} w_T. \tag{2.17}$$

For large $h$, the forecasts for each individual variables are dominated by the common trend $w_T$.

But for $x_{T,h} = \alpha w_T$ and $y_{T,h} = \gamma w_T$ we get

$$\gamma x_{T,h} - \alpha y_{T,h} = (\gamma\alpha - \gamma\alpha) w_T = 0$$

showing that the forecasts obey the cointegrating relationship and the forecast of the $I(0)$ variable $z$ is $z_{T,h} = 0$, with finite variance.

## 2.3 The mean of the cointegrating relationship

Note that in this simple example, the mean of the cointegration relationship is 0. In practice, this *long-run mean* can be any number, and appears as an important parameter in the econometric cointegrating model. It is in fact a critical parameter for the forecast performance of econometric models.

In many cases, the long-run cointegrating mean also has an economic interpretation. If $x_t$ and $y_t$ are logs of income and consumption, the mean of the cointegration relationship may for example be interpreted as the long-run (steady-state) savings rate.

# 3 MA representation

Using lag-polynomials, write $w_t$ and $z_t$ as:

$$w_t = \frac{\eta_{1t}}{1 - L}$$

and

$$z_t = \frac{-\eta_{2t}}{1 - \lambda L}.$$

Inserting back in (2.14) and (2.15) gives

$$\begin{bmatrix} \Delta x_t \\ \Delta y_t \end{bmatrix} = \boldsymbol{\Psi}(L)\boldsymbol{\eta}_t \tag{3.18}$$

where

$$\boldsymbol{\Psi}(L) = \begin{bmatrix} \alpha & \beta(1-L)/(1-\lambda L) \\ \gamma & \delta(1-L)/(1-\lambda L) \end{bmatrix}. \tag{3.19}$$

The rank of $\boldsymbol{\Psi}(1)$ is 1, since $|\boldsymbol{\Psi}(1) - z\boldsymbol{I}| = (\alpha - z) \cdot z = 0$ 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.

## 4 Equilibrium correction representation (ECM)

Rewrite (2.4)

$$\alpha\delta - \gamma\beta = 1,$$

first as

$$\alpha\delta - \lambda\beta\gamma = \alpha\delta - \gamma\beta + \gamma\beta - \lambda\beta\gamma = 1 + \gamma\beta(1-\lambda),$$

and next,

$$-\gamma\beta + \lambda\alpha\delta = 1 - \alpha\delta + \lambda\alpha\delta = 1 - \alpha\delta(1-\lambda),$$

and use this to write $\boldsymbol{\Phi}$ in (2.6) as

$$\boldsymbol{\Phi} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + (1-\lambda)\begin{bmatrix} \beta \\ \delta \end{bmatrix}\begin{bmatrix} \gamma & -\alpha \end{bmatrix}. \tag{4.20}$$

The ECM representation follows directly from the VAR, by using (4.20) in

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

We obtain:

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

which can be written as

$$\begin{bmatrix} \Delta x_t \\ \Delta y_t \end{bmatrix} = \boldsymbol{\alpha}\boldsymbol{\beta}' \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \boldsymbol{\varepsilon}_t, \tag{4.21}$$

or

$$\Delta\boldsymbol{y}_t = \boldsymbol{\alpha}\boldsymbol{\beta}'\boldsymbol{y}_{t-1} + \boldsymbol{\varepsilon}_t, \tag{4.22}$$

when the vector of equilibrium correction coefficients ($\boldsymbol{\alpha}$) are

$$\boldsymbol{\alpha} = \begin{bmatrix} (1-\lambda)\beta \\ (1-\lambda)\delta \end{bmatrix} = \begin{bmatrix} \alpha_{11} \\ \alpha_{21} \end{bmatrix} \tag{4.23}$$

and the cointegrating vector is:

$$\boldsymbol{\beta} = \begin{bmatrix} \gamma \\ -\alpha \end{bmatrix}. \tag{4.24}$$

When we work with a VAR, we often re-write (2.1) as

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

where $\boldsymbol{\Pi} = (\boldsymbol{\Phi} - \boldsymbol{I})$ is the matrix with the coefficients of the lagged levels variables.

Note that we have

$$\boldsymbol{\alpha}\boldsymbol{\beta}' = \boldsymbol{\Pi} = \boldsymbol{\Phi} - \boldsymbol{I} \tag{4.25}$$

which implies that the $\boldsymbol{\Pi}$ matrix with coefficients of the lagged levels variables has reduced rank ($rank = 1$) in the case of cointegration. (4.25) implies that the roots of $\boldsymbol{\Pi}$ are 0 and $(1-\lambda)$. $\boldsymbol{\beta}$ is the matrix of cointegration coefficients.

Note: If $\lambda = 1$, the rank of $\mathbf{\Pi}$ is zero (both roots are 0). Moreover:

$$\mathbf{\Phi} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

from (4.20) and the ECM representations collapses to a VAR in differences, a $dVAR$. Cointegration does not occur.

We see that there is a close relationship between the eigenvalues of $\mathbf{\Pi}$, its rank and cointegration:

- $rank(\mathbf{\Pi}) = 0$, reduced rank and no cointegration. Both eigenvalues of $\mathbf{\Pi}$ 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 (2.1) is stationary.

Hence, if there is a way of testing formally whether the eigenvalues of $\mathbf{\Pi}$ are significantly different from zero, we would have a logically very satisfying method of testing the hypotheses about the absence of cointegration.

This has method has actually been developed under the name of the Johansen (1995) procedure and the cointegrated VAR

## 4.1 Cointegration and Granger causality

Since $\lambda < 1$ is equivalent with cointegration, we see from (4.23) that cointegration also implies Granger-causality in at least one direction: $\alpha_{11} \neq 0$ and/or $\alpha_{21} \neq 0$. Conversely If $\lambda = 1, \boldsymbol{\alpha} = \mathbf{0}$.

## 4.2 Cointegration and weak-exogeneity

Assume $\delta = 0$, from (4.23). this implies $\alpha_{21} = 0$.

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

The marginal model of $y_t$ contains no information about the cointegration parameters $(\gamma, -\alpha)'$ in this case. $y_t$ is therefore weakly exogenous for the cointegration parameters $(\gamma, -\alpha)'$.

# 5 Some exercises

**Exercise 1** *Consider the model*

$$\begin{aligned} \Delta y_t &= (\rho - 1) y_{t-1} + b_0 \Delta x_t + b_1 x_{t-1} + \varepsilon_t, \qquad 0 < \rho < 1 \\ \Delta x_t &= u_t, \end{aligned}$$

*where $\varepsilon_t$ and $u_t$ are independent white-noise processes.*

1. *Show that* $y_t \sim I(1)$.

2. *Write the system in AR-form:*

$$\begin{pmatrix} y_t \\ x_t \end{pmatrix} = \mathbf{A} \begin{pmatrix} y_{t-1} \\ x_{t-1} \end{pmatrix} + \begin{pmatrix} b_0 u_t + \varepsilon_t \\ u_t \end{pmatrix}$$

   *where*

$$\mathbf{A} = \begin{pmatrix} \rho & b_1 \\ 0 & 1 \end{pmatrix}.$$

3. *Show that the characteristic roots of* $\mathbf{A}$ *is* $\lambda_1 = 1$ *and* $\lambda_2 = \rho$.

4. *Show that* $\mathbf{A}^*$ *in the ECM form*

$$\begin{pmatrix} \Delta y_t \\ \Delta x_t \end{pmatrix} = A^* \begin{pmatrix} y_{t-1} \\ x_{t-1} \end{pmatrix} + \begin{pmatrix} b_0 u_t + \varepsilon_t \\ u_t \end{pmatrix}$$

   *has roots* $z_1 = 0$ *and* $z_2 = \rho - 1$.

5. *Show that* $\mathbf{A}^*$ *can be written*

$$\mathbf{A}^* = \begin{pmatrix} \rho - 1 \\ 0 \end{pmatrix} (1, \gamma)$$

   *where* $\gamma = b_1/(\rho - 1)$.

6. *Show that* $\beta' = (1, \gamma)$ *is the cointegrating vector. Hint: Show that*

$$w_t = \beta' \begin{pmatrix} y_t \\ x_t \end{pmatrix} = y_t + \gamma x_t$$

   *is a stationary process by multiplying the AR-equation with* $\beta'$ *from the left.*

## References


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

Johansen, S. (1995). *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford University Press, Oxford.