Università di Pavia

Cointegration

Eduardo Rossi

Integrated variables exhibit a systematic variation. But the variation is hardly predictable, though the variation is systematic. This type of variation is called *stochastic trend*. On the other hand trends which are completely predictable are known as *deterministic trends*.

Beveridge and Nelson (1981) show that any ARIMA(p,1,q) model can be represented as a stochastic trend plus a stationary component.

$$\Delta \mathbf{y}_t = \mathbf{\Psi}(L)\,\varepsilon_t$$

$\mathbf{y}_t$ is $(n \times 1)$. $\varepsilon_t \sim VWN$.

$$\mathbf{\Psi}(1) \neq 0$$

$$\Delta \mathbf{y}_t = [\mathbf{\Psi}(L) - \mathbf{\Psi}(1) + \mathbf{\Psi}(1)]\,\varepsilon_t$$

$$\mathbf{\Psi}(L) - \mathbf{\Psi}(1) = \mathbf{\Psi}^*(L)\,\Delta$$

$\mathbf{\Psi}^*(L)$ with roots outside the unit circle,

$$\mathbf{\Psi}^*(L) = \frac{\mathbf{\Psi}(L) - \mathbf{\Psi}(1)}{\Delta}$$

$$\Delta \mathbf{y}_t = \mathbf{\Psi}^*(L)\,\Delta\varepsilon_t + \mathbf{\Psi}(1)\,\varepsilon_t$$

$$\mathbf{y}_t = \mathbf{y}_{t-1} + \mathbf{\Psi}^* (L) \Delta \varepsilon_t + \mathbf{\Psi} (1) \varepsilon_t$$

Recursive substitution with $\mathbf{y}_0 = 0$, $\varepsilon_0 = 0$:

$$\begin{aligned}
\mathbf{y}_t &= \mathbf{y}_{t-2} + \mathbf{\Psi}^* (L) \Delta \varepsilon_{t-1} + \mathbf{\Psi} (1) \varepsilon_{t-1} + \mathbf{\Psi}^* (L) \Delta \varepsilon_t + \mathbf{\Psi} (1) \varepsilon_t \\
\mathbf{y}_t &= \cdots \\
\mathbf{y}_t &= \mathbf{y}_0 + \mathbf{\Psi}^* (L) \Delta \sum_{s=1}^{t} \varepsilon_s + \mathbf{\Psi} (1) \sum_{s=1}^{t} \varepsilon_s \\
\mathbf{y}_t &= \mathbf{\Psi}^* (L) \Delta \sum_{s=1}^{t} \varepsilon_s + \mathbf{\Psi} (1) \sum_{s=1}^{t} \varepsilon_s
\end{aligned}$$

$$\Delta \sum_{s=1}^{t} \varepsilon_s = \mathbf{\Delta}\varepsilon_t + \mathbf{\Delta}\varepsilon_{t-1} + \ldots + \mathbf{\Delta}\varepsilon_1 = \varepsilon_t$$

$$\mathbf{y}_t = \mathbf{\Psi}(1) \sum_{s=1}^{t} \varepsilon_s + \mathbf{\Psi}^*(L)\,\varepsilon_t$$

$$\mathbf{\Psi}(1) \neq 0 \quad rank\left(\mathbf{\Psi}(1)\right) = n - r$$

$$\mathbf{\Psi}(1) \sum_{s=1}^{t} \varepsilon_s \quad : \quad \text{permanent component (stochastic trend)}$$

$$\mathbf{\Psi}^*(L)\,\varepsilon_t \quad : \quad \text{transitory component}$$

Consider two uncorrelated random walk processes

$$\begin{aligned} y_t &= y_{t-1} + u_t \quad u_t \sim i.i.d.\left(0, \sigma_u^2\right) \\ x_t &= x_{t-1} + v_t \quad v_t \sim i.i.d.\left(0, \sigma_v^2\right) \end{aligned}$$

where u_t and v_t are assumed to be serially uncorrelated as well as mutually uncorrelated. Consider the regression

$$y_t = \beta_0 + \beta_1 x_t + \varepsilon_t$$

since y_t, x_t are uncorrelated rw processes, we would expect the R^2 from this regression would tend to zero. However, this is not the case. The parameter β_1 detects correlation and Yule (1927) showed that spurious correlation can persist even in large samples in nonstationary time series. If the two series are growing over time, they can be correlated even if the increments in each series are uncorrelated.

Given a linear process $\mathbf{y}_t$

$$\mathbf{y}_t = \sum_{i=1}^{\infty} \boldsymbol{\Psi}_i \boldsymbol{\epsilon}_{t-i} = \boldsymbol{\Psi}(L)\boldsymbol{\epsilon}_t \;\; t = 0, 1, \ldots$$

$$\boldsymbol{\epsilon}_t \sim i.i.d.(\mathbf{0}, \boldsymbol{\Omega})$$

$\boldsymbol{\Psi}(z) = \sum_i^{\infty} \boldsymbol{\Psi}_i z^i$ is convergent for $|z| \leq 1 + \delta$ for some $\delta > 0$.

A stochastic process $\mathbf{y}_t$ which satisfies that

$$\mathbf{y}_t - E(\mathbf{y}_t) = \sum_{i=0}^{\infty} \boldsymbol{\Psi}_i \boldsymbol{\epsilon}_{t-i}$$

is called I(0) if

$$\boldsymbol{\Psi}(1) = \sum_{i=0}^{\infty} \boldsymbol{\Psi}_i \neq 0$$

Univariate linear process consider the AR(1) process:

$$y_{1t} = \rho y_{1t-1} + \epsilon_t$$

for $|\rho| < 1$

$$y_{1t} = \sum_{i=0}^{\infty} \rho^i \epsilon_{t-i}$$

the sum of coefficients is

$$\sum_{i=0}^{\infty} \rho^i = \frac{1}{1-\rho}$$

the process is I(0). The process

$$y_{2t} = \epsilon_{2t} - \theta \epsilon_{2t-1}$$

is stationary $\forall \theta$, but for $\theta \neq 1$ it is an I(0) process.

The process

$$\mathbf{y}_t = \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix}$$

is called an I(0) process even if $\theta = 1$ since

$$\mathbf{\Psi}(1) = \begin{bmatrix} \frac{1}{1-\rho} & 0 \\ 0 & 1-\theta \end{bmatrix} \neq 0 \;\; \forall\theta$$

A stochastic process $\mathbf{y}_t$ is called integrated of order d, I(d), $d = 0, 1, 2, \ldots$ if

$$\Delta^d(\mathbf{y}_t - E(\mathbf{y}_t))$$

is I(0).

If $\mathbf{y}_t$ is I(1) and $\mathbf{A}$ is full rank then $\mathbf{A}\mathbf{y}_t \sim I(1)$.

A consequence of the definition is that the stochastic part of an I(1) process $\mathbf{y}_t$ is nonstationary, since if

$$\mathbf{y}_t = \mathbf{y}_0 + \sum_{i=1}^{t} \mathbf{x}_i$$

for some I(0) process

$$\mathbf{x}_t = \sum_{i=0}^{\infty} \mathbf{\Psi}_i \boldsymbol{\epsilon}_{t-i}$$

the condition

$$\sum_{i=0}^{\infty} \mathbf{\Psi}_i \neq \mathbf{0}$$

implies that $\mathbf{y}_t$ is nonstationary.

Let

$$\mathbf{\Psi}(z)=\sum_{i}^{\infty}\mathbf{\Psi}_i z^i$$

We define the function

$$\mathbf{\Psi}^*(z)=\frac{\mathbf{\Psi}(z)-\mathbf{\Psi}(1)}{1-z}=\sum_{i=0}^{\infty}\mathbf{\Psi}_i^* z^i$$

$$\mathbf{\Psi}_i^*=-\sum_{j=i+1}^{\infty}\mathbf{\Psi}_j$$

the power series for $\mathbf{\Psi}^*$ is convergent for $|z|<1+\delta$

$$\mathbf{\Psi}(z)=\mathbf{\Psi}(1)+\mathbf{\Psi}^*(z)(1-z) \tag{1}$$

$$=\sum_{i}^{\infty}\mathbf{\Psi}_i+\mathbf{\Psi}^*(z)(1-z) \tag{2}$$

Define the stationary process

$$\mathbf{x}_t^* = \sum_{i=1}^{\infty} \mathbf{\Psi}_i^* \boldsymbol{\epsilon}_{t-i} = \mathbf{\Psi}^*(L)\boldsymbol{\epsilon}_t$$

$$\mathbf{x}_t = \mathbf{\Psi}(L)\boldsymbol{\epsilon}_t = \mathbf{\Psi}(1)\boldsymbol{\epsilon}_t + (1-L)\mathbf{x}_t^*$$

hence

$$\mathbf{y}_t = \mathbf{y}_0 + \sum_{i=1}^{t} \mathbf{x}_i = \mathbf{y}_0 + \mathbf{\Psi}(1)\sum_{i=1}^{t} \boldsymbol{\epsilon}_i + \sum_{i=1}^{t}(1-L)\mathbf{x}_i^*$$

$$\mathbf{y}_t = \mathbf{y}_0 + \mathbf{\Psi}(1)\sum_{i=1}^{t} \boldsymbol{\epsilon}_i + \mathbf{x}_t^* - \mathbf{x}_0^*$$

This process is nonstationary since

$$\begin{aligned}\Delta \mathbf{y}_t &= \Delta \mathbf{y}_0 + \sum_{i=1}^{t} \Delta \mathbf{x}_i \\ &= \mathbf{y}_0 - \mathbf{y}_{-1} - \mathbf{x}_0 + \mathbf{x}_t \\ &= \mathbf{y}_0 - (\mathbf{y}_{-1} + \mathbf{x}_0) + \mathbf{x}_t \\ &= \mathbf{x}_t \sim I(0)\end{aligned}$$

Linear combinations of $\mathbf{y}_t$. Let $\boldsymbol{\beta} \in \mathbb{R}^N$, then

$$\boldsymbol{\beta}'\mathbf{y}_t = \boldsymbol{\beta}'\mathbf{y}_0 + \boldsymbol{\beta}'\boldsymbol{\Psi}(1)\sum_{i=1}^{t}\boldsymbol{\epsilon}_t + \boldsymbol{\beta}'\mathbf{x}_t^* - \boldsymbol{\beta}'\mathbf{x}_0^*$$

if we want $\boldsymbol{\beta}'\mathbf{y}_t$ to be stationary we must have

$$\boldsymbol{\beta}'\boldsymbol{\Psi}(1) = \mathbf{0}'$$

then

$$\boldsymbol{\beta}'\mathbf{y}_t = \boldsymbol{\beta}'\mathbf{y}_0 + \boldsymbol{\beta}'\mathbf{x}_t^* - \boldsymbol{\beta}'\mathbf{x}_0^*$$

This is stationary only if we choose the initial values of the process $\boldsymbol{\beta}'\mathbf{y}_t$ such that

$$\boldsymbol{\beta}'\mathbf{y}_0 = \boldsymbol{\beta}'\mathbf{x}_0^*$$

in this case

$$\boldsymbol{\beta}'\mathbf{y}_t = \boldsymbol{\beta}'\mathbf{x}_t^*$$

The definition of I(1) gives no condition on the initial values or the level of the process. Thus we need an extra condition that the initial values can be chosen as indicated in order to have the stationarity conditions of linear combinations of the levels.

Cointegration: Let $\mathbf{y}_t$ be I(1). We call $\mathbf{y}_t$ *cointegrated* with cointegrating vector $\boldsymbol{\beta} \neq 0$ if $\boldsymbol{\beta}'\mathbf{y}_t$ can be made stationary by a suitable choice of its initial distribution.
The *cointegrating rank* is the number of linearly independent cointegrating relations, and the space spanned by the cointegrating relations is the cointegrating space.

Note that $\boldsymbol{\beta}'\mathbf{y}_t$ need not be I(0), but for AR processes the cointegrating relations we find are in fact I(0).

Two-dimensional process $\mathbf{y}_t = (y_{1t}, y_{2t})'$

$$y_{1t} = \sum_{i=1}^{t} \epsilon_{1i} + \epsilon_{2t}$$

$$y_{2t} = a\sum_{i=1}^{t} \epsilon_{1i} + \epsilon_{3t}$$

y_{1t}, y_{2t} and $\mathbf{y}_t$ are also I(1) processes. They cointegrate with cointegrating vector $\boldsymbol{\beta}' = (a, -1)$

$$\begin{aligned}\boldsymbol{\beta}'\mathbf{y}_t = ay_{1t} - y_{2t} &= a\sum_{i=1}^{t} \epsilon_{1i} - a\sum_{i=1}^{t} \epsilon_{1i} + a\epsilon_{2t} - \epsilon_{3t} \\ &= a\epsilon_{2t} - \epsilon_{3t}\end{aligned}$$

If further

$$y_{3t} = \epsilon_{4t}$$

then $y_{3t} \sim I(0)$ but the vector process $\mathbf{y}_t = (y_{1t}, y_{2t}, y_{3t})'$ is I(1), with cointegrating vectors

$$(a, -1, 0)$$

$$(0, 0, 1)$$

We allow for unit vectors (abuse of language). The inclusion a stationary variable in the process increases the dimension of the cointegrating space by one.

Three-dimensional process

$$y_{1t} = \sum_{i=1}^{t}\sum_{j=1}^{i}\epsilon_{1j} + \sum_{i=1}^{t}\epsilon_{2i}$$

$$y_{2t} = a\sum_{i=1}^{t}\sum_{j=1}^{i}\epsilon_{1j} + b\sum_{i=1}^{t}\epsilon_{2i} + \epsilon_{3t}$$

$$y_{3t} = c\sum_{i=1}^{t}\epsilon_{2i} + \epsilon_{4t}$$

y_{1t} and y_{2t} are I(2), y_{3t} is I(1). The process $\mathbf{y}_t = (y_{1t}, y_{2t}, y_{3t})'$ is I(2) and cointegrates since

$$ay_{1t} - y_{2t} = (a-b)\sum_{i=1}^{t}\epsilon_{2i} - \epsilon_{3t} \sim I(1)$$

and

$$acy_{1t} - cy_{2t} - (a-b)y_{3t} = -c\epsilon_{3t} - (a-b)\epsilon_{4t} \sim I(0)$$

thus $(a, -1, 0)$ is a cointegrating vector that changes the order of the process from two to one, $[ac, -c, -(a-b)]$ changes the integration order from 2 to 0.

General process that illustrates the various possibilities in N dimensions for $d = 2$:

$$\mathbf{y}_t = \mathbf{\Psi}_2 \sum_{s=1}^{t} \sum_{i=1}^{s} \boldsymbol{\epsilon}_i + \mathbf{\Psi}_1 \sum_{i=1}^{t} \boldsymbol{\epsilon}_i + \tau_0 + \tau_1 t + \frac{1}{2}\tau_2 t^2 + \mathbf{x}_t \qquad t = 1, 2, \ldots$$

$\mathbf{y}_t \sim I(2)$ if $\mathbf{\Psi}_2 \neq \mathbf{0}$, it is nonstationary and differencing it twice makes it $I(0)$.

The MA representation models the variation of the economic data through the matrices $\mathbf{\Psi}_1$ and $\mathbf{\Psi}_1$ as the results of the influence of its *unobserved common trends* given by the cumulated sums of $\boldsymbol{\epsilon}$.

If $\boldsymbol{\beta}'\boldsymbol{\Psi}_2 = \mathbf{0}'$ then the order of integration is reduced from 2 to 1, Granger (1981) called this case *cointegration*, denoted by $CI(2,1)$.

The idea is to describe the stable relations in the economy by linear relations that are most stationary than the original variables.

Another way of modeling cointegrating variables is through the error correction model. An example of a reduced form ECM is:

$$\Delta \mathbf{y}_t = \boldsymbol{\alpha}\boldsymbol{\beta}' \mathbf{y}_{t-1} + \boldsymbol{\mu} + \boldsymbol{\epsilon}_t \qquad t = 1, 2, \ldots$$

with initial value $\mathbf{y}_0$,

$$\begin{aligned} \boldsymbol{\alpha} : (N \times r) &\qquad \text{rank}(\boldsymbol{\alpha}) = r \\ \boldsymbol{\beta} : (N \times r) &\qquad \text{rank}(\boldsymbol{\beta}) = r \\ \boldsymbol{\alpha}_\perp : (N \times (N - r)) &\qquad \text{rank}(\boldsymbol{\alpha}_\perp) = N - r \end{aligned}$$

such that

$$\boldsymbol{\alpha}' \boldsymbol{\alpha}_\perp = \mathbf{0}_{r \times N - r}$$

then

$$\text{rank}(\boldsymbol{\alpha}, \boldsymbol{\alpha}_\perp) = N$$

The matrix $\boldsymbol{\alpha}_{\perp}$ is not uniquely defined. The relation

$$\boldsymbol{\beta}_{\perp}(\boldsymbol{\alpha}'_{\perp}\boldsymbol{\beta}_{\perp})^{-1}\boldsymbol{\alpha}'_{\perp} + \boldsymbol{\alpha}(\boldsymbol{\beta}'\boldsymbol{\alpha})^{-1}\boldsymbol{\beta}' = \mathbf{I}_N$$

which expresses that if $\boldsymbol{\beta}'\boldsymbol{\alpha}$ has full rank then any vector $\mathbf{v} \in \mathbb{R}^N$ can be decomposed in $\mathbf{v}_1 \in \mathrm{Col}(\boldsymbol{\beta}_{\perp})$ and $\mathbf{v}_2 \in \mathrm{Col}(\boldsymbol{\alpha})$:

$$\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2 = \boldsymbol{\beta}_{\perp}(\boldsymbol{\alpha}'_{\perp}\boldsymbol{\beta}_{\perp})^{-1}\boldsymbol{\alpha}'_{\perp}\mathbf{v} + \boldsymbol{\alpha}(\boldsymbol{\beta}'\boldsymbol{\alpha})^{-1}\boldsymbol{\beta}'\mathbf{v}$$

If $\mathrm{rank}(\boldsymbol{\alpha}'_{\perp}\boldsymbol{\beta}_{\perp}) = N - r$ we can show that the solution of ECM is the MA (common trends) representation:

$$\mathbf{y}_t = \boldsymbol{\Psi}\sum_{i=1}^{t}\boldsymbol{\epsilon}_i + \tau_0 + \tau_1 t + \mathbf{x}_t$$

where

$$\boldsymbol{\Psi} = \boldsymbol{\beta}_{\perp}(\boldsymbol{\alpha}'_{\perp}\boldsymbol{\beta}_{\perp})^{-1}\boldsymbol{\alpha}'_{\perp} \qquad \text{and} \qquad \tau_1 = \boldsymbol{\Psi}\boldsymbol{\mu}.$$

To see this multiply by $\boldsymbol{\beta}'$

$$\begin{aligned}\boldsymbol{\beta}'\Delta\mathbf{y}_t &= \boldsymbol{\beta}'\boldsymbol{\alpha}\boldsymbol{\beta}'\mathbf{y}_{t-1} + \boldsymbol{\beta}'\boldsymbol{\mu} + \boldsymbol{\beta}'\boldsymbol{\epsilon}_t \\ \boldsymbol{\beta}'\mathbf{y}_t &= \boldsymbol{\beta}'\mathbf{y}_{t-1} + \boldsymbol{\beta}'\boldsymbol{\alpha}\boldsymbol{\beta}'\mathbf{y}_{t-1} + \boldsymbol{\beta}'\boldsymbol{\mu} + \boldsymbol{\beta}'\boldsymbol{\epsilon}_t \\ &= (\mathbf{I} + \boldsymbol{\beta}'\boldsymbol{\alpha})\boldsymbol{\beta}'\mathbf{y}_{t-1} + \boldsymbol{\beta}'\boldsymbol{\mu} + \boldsymbol{\beta}'\boldsymbol{\epsilon}_t\end{aligned}$$

the process is stationary if the matrix $(\mathbf{I} + \boldsymbol{\beta}'\boldsymbol{\alpha})$ has its eigenvalues inside the unit circle. The stationary representation

$$\boldsymbol{\beta}'\mathbf{y}_t = \sum_{i=0}^{\infty}(\mathbf{I} + \boldsymbol{\beta}'\boldsymbol{\alpha})^i\boldsymbol{\beta}'(\boldsymbol{\mu} + \boldsymbol{\epsilon}_{t-i})$$

Multiplying the ECM by $\boldsymbol{\alpha}'_{\perp}$

$$\boldsymbol{\alpha}'_{\perp}\Delta\mathbf{y}_t = \boldsymbol{\alpha}'_{\perp}\boldsymbol{\alpha}\boldsymbol{\beta}'\mathbf{y}_{t-1} + \boldsymbol{\alpha}'_{\perp}\boldsymbol{\mu} + \boldsymbol{\alpha}'_{\perp}\boldsymbol{\epsilon}_t$$

$$\boldsymbol{\alpha}'_{\perp}\Delta\mathbf{y}_t = \boldsymbol{\alpha}'_{\perp}\boldsymbol{\mu} + \boldsymbol{\alpha}'_{\perp}\boldsymbol{\epsilon}_t$$

which has solution

$$\boldsymbol{\alpha}'_{\perp}\mathbf{y}_t = \boldsymbol{\alpha}'_{\perp}\mathbf{y}_0 + \sum_{i=1}^{t}\boldsymbol{\alpha}'_{\perp}(\boldsymbol{\mu} + \boldsymbol{\epsilon}_i)$$

Combining these results

$$\begin{aligned}\mathbf{y}_t &= (\boldsymbol{\beta}_{\perp}(\boldsymbol{\alpha}'_{\perp}\boldsymbol{\beta}_{\perp})^{-1}\boldsymbol{\alpha}'_{\perp} + \boldsymbol{\alpha}(\boldsymbol{\beta}'\boldsymbol{\alpha})^{-1}\boldsymbol{\beta}')\mathbf{y}_t \\ &= \boldsymbol{\Psi}\mathbf{y}_0 + \boldsymbol{\Psi}\sum_{i=1}^{t}(\boldsymbol{\mu} + \epsilon_i) + \boldsymbol{\alpha}(\boldsymbol{\beta}'\boldsymbol{\alpha})^{-1}\sum_{i=0}^{\infty}(\mathbf{I} + \boldsymbol{\beta}'\boldsymbol{\alpha})^{i}\boldsymbol{\beta}'(\boldsymbol{\mu} + \boldsymbol{\epsilon}_{t-i})\end{aligned}$$

this is an instance of the Granger Representation Theorem.

This is a way of finding the MA representation from the AR representation and vice versa when there are $I(1)$ variables in the system.
The non stationarity in the process $\mathbf{y}_t$ is created by the common trends $\boldsymbol{\alpha}'_\perp \sum_{i=1}^t \epsilon_i$.