An introduction to

Cointegration

Jacob Østergaard

DSIN forum

October 30, 2014

Agenda

- A man and his dog
- Nonstationary processes: unit roots and integration order
- Defining cointegration
- Model parameters
- Interpreting cointegration
- Simulation
- Cointegration and phase synchronization
- Future work

A man and his dog

Nonstationarity

Let $p \in \mathbf{N}$ and

$$x_t = a_1 x_{t-1} + \cdots + a_p x_{t-p} + \varepsilon_t,$$

The associated characteristic polynomial for x_t is

$$C(z) = 1 - a_1 z - \cdots - a_p z^p$$

If $C(z) \neq 0$ for $|z| \leq 1$ the process is *stationary*.

If $C(z) = 0$ for $|z| = 1$, the process contains a *unit root*, and is *nonstationary*.

Nonstationarity

A random walk $x_t = x_{t-1} + \varepsilon_t$ contains a unit root.

Applying the difference operator Δ to x_t, we obtain stationarity:

$$\Delta x_t = x_t - x_{t-1} = \varepsilon_t,$$

and call x_t *integrated of order 1*, denoted $I(1)$.

Generally: A process is $I(d)$, if it has d unit roots, and hence a stationary process as $I(0)$.

Nonstationarity

A random walk $x_t = x_{t-1} + \varepsilon_t$ contains a unit root.

Applying the difference operator Δ to x_t, we obtain stationarity:

$$\Delta x_t = x_t - x_{t-1} = \varepsilon_t,$$

and call x_t *integrated of order 1*, denoted $I(1)$.

Generally: A process is $I(d)$, if it has d unit roots, and hence a stationary process as $I(0)$.

Disclaimer: Strictly speaking, a stationary process is *stable*, i.e. $I(0)$, but the opposite is not necessarily true.

A stable process is however *asymptotically stationary* and thus the terms stable and stationary are often used interchangeably in literature.

Defining cointegration

Let $\boldsymbol{y_t} = (y_{1,t}, \ldots, y_{k,t})'$ be a k-dimensional process

$$\boldsymbol{y}_t = A_1 \boldsymbol{y}_{t-1} + \cdots + A_p \boldsymbol{y}_{t-p} + \varepsilon_t,$$

where A_i are $k \times k$ matrices and ε_t is a white noise process.

The characteristic polynomial in the multivariate case is

$$C(z) = \det(I_k - A_1 z - \cdots - A_p z^p)$$

For a multivariate process, the number of units roots may be larger than the integration order.

If $y_{1,t}, \ldots, y_{k,t}$ are integrated of order max. d, then $\boldsymbol{y}_t$ is $I(d)$.

Defining cointegration

We can rewrite

$$y_t = A_1 y_{t-1} + \cdots + A_p y_{t-p} + \varepsilon_t,$$

as

$$\Delta y_t = \Pi y_{t-1} + \Gamma_1 \Delta y_{t-1} + \cdots + \Gamma_{p-1} \Delta y_{t-p+1} + \varepsilon_t,$$

where

$$\Pi = -(I_k - A_1 - \cdots - A_p)$$
$$\Gamma_j = -(A_{j+1} + \cdots + A_p), \quad j = 1, \dots p-1.$$

If y_t is $I(1)$ then Δy_t is $I(0)$ and thus Πy_{t-1} must be $I(0)$!

Defining cointegration

3 possibilities:

- Π has full rank k.
- Π has reduced rank $0 < r < k$.
- Π has rank 0.

Rank k implies that y_t *must* be $I(0)$ and rank 0 implies no stationary combinations of y_t.

Rank $0 < r < k$ implies r linearly independent stationary combinations of y_t variables.

We then define y_t as a *cointegrated* process.

Model parameters

If Π has rank $0 < r < k$, then

$$\Pi = \alpha\beta',$$

where α and β are $k \times r$, with full column rank $r < k$.

We find

$$(\alpha'\alpha)^{-1}\alpha'\Pi\boldsymbol{y}_{t-1} = (\alpha'\alpha)^{-1}\alpha'\alpha\beta'\boldsymbol{y}_{t-1} = \beta'\boldsymbol{y}_{t-1}$$

is $I(0)$.

Hence the r linearly independent columns of β correspond to r stationary linear combinations of $\boldsymbol{y}_t$.

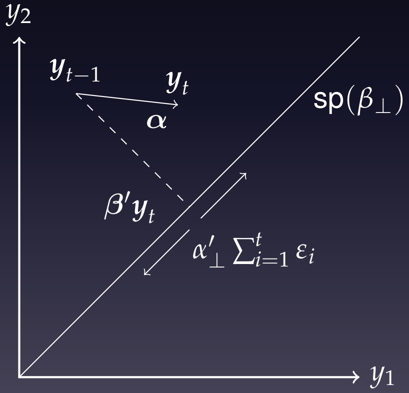
Interpreting cointegration

Assume the simple case

$$\boldsymbol{y}_t = (y_{1,t}, y_{2,t})'$$
$$\Delta \boldsymbol{y}_t = \alpha\beta' \boldsymbol{y}_{t-1} + \varepsilon_t$$
$$\beta = (1, -1)',$$

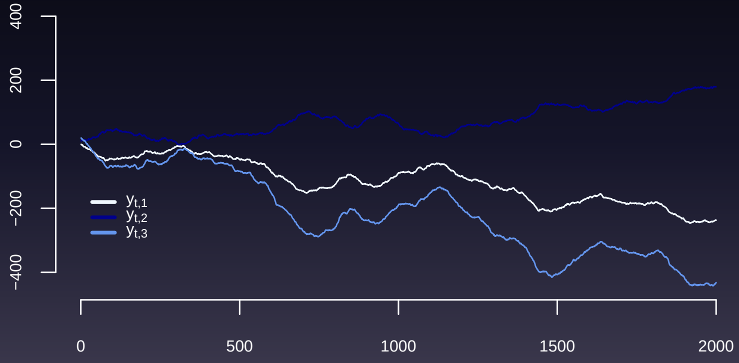
Here $\beta_\perp = (1, 1)'$ is the orthogonal complement to β, such that $\beta'\beta_\perp = 0$.

Same goes for $\alpha_\perp$.



Simulation

2000 observations from a 3-dim process y_t



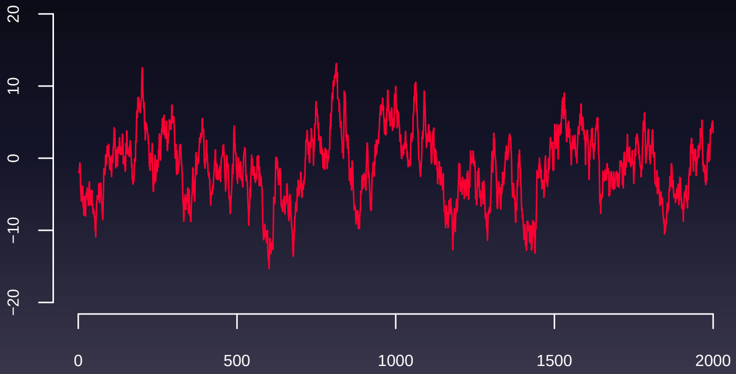
A simulated visualization

The 3-dim process y_t has been simulated with the following parameters

$$\alpha = (0.05,\ \ -0.1,\ \ 0.15)'$$
$$\beta = (1,\ \ 0.5,\ \ -0.35)'$$
$$\Gamma_1 = \begin{bmatrix} 0.1 & 0.1 & 0.1 \\ 0.1 & 0.1 & 0.1 \\ 0.1 & 0.1 & 0.1 \end{bmatrix}$$
$$y_0 = (0,\ \ 10,\ \ 20)'$$
$$y_1 = (-2,\ \ 12,\ \ 17)'$$

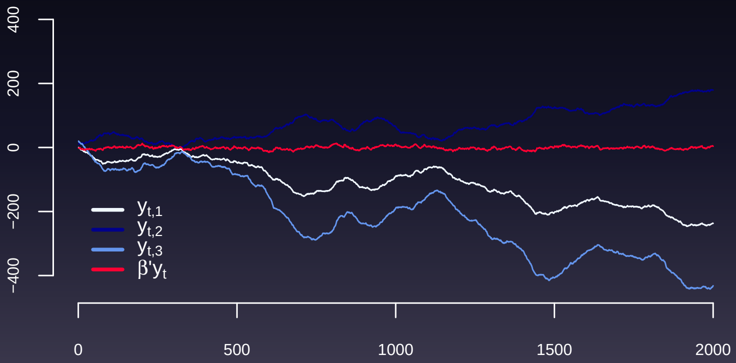
Simulation

We obtain the following $\beta' y_t$-series



Simulation

Plotting $\beta' \boldsymbol{y}_t$ along with the process...



Simulation

Analyzing $\boldsymbol{y}_t$ would consist of the following steps:

1. Determine the order of integration.
2. Choose the lag-order p, for the process.
3. Determine the rank of Π and estimate α, β, $\Gamma_1, \ldots, \Gamma_{p-1}$
   - Auxiliary regression of $\Delta\boldsymbol{y}_t$ and $\boldsymbol{y}_{t-1}$ on $\Delta\boldsymbol{y}_{t-1}$.
   - Estimate β.
   - Use $\hat{\beta}$ to estimate $\hat{\alpha}$, $\hat{\Gamma}_j$, $j = 1, \ldots, p-1$.
   - LR test for the rank of Π:
     H_0 : rank$(\Pi) = r$ vs. H_a : rank$(\Pi) > r$ (or r vs. $r+1$)
4. Model diagnostics.
5. Parameter inference.

Cointegration and phase synchronization

In their 2012 paper

On the relationship between the theory of cointegration and the theory of phase synchronization

Rainer Dahlhaus and Jan C. Neddermeyer introduced cointegration to the realm of phase synchronization.

Cointegration and phase synchronization

They discuss the similarity between the (extended) *Kuramoto model*

$$\dot{\phi}^{(i)} = \omega + \frac{K}{d}\sum_{j=1}^{d}\sin(\phi^{(j)} - \phi^{(i)}) + \eta_i, \quad i = 1, \ldots, d$$

where η_i is a white noise, and a cointegrated system

$$\Delta\phi_t = \omega + \Pi\phi_{t-1} + \eta_t$$
$$\phi_t = (\phi_t^{(1)}, \ldots, \phi_t^{(d)})$$
$$\omega = (\omega_1, \ldots, \omega_d)'$$

using $\sin(\phi^{(j)} - \phi^{(i)}) \approx \phi^{(j)} - \phi^{(i)}$ for small differences.

Cointegration and phase synchronization

They also determine the coupling properties of a simulated coupled Rössler system

$$\begin{aligned}
\dot{x}_1 &= -12(y_1 + z_1) + w_1, \\
\dot{y}_1 &= 12(x_1 + 0.2y_1), \\
\dot{z}_1 &= 12(0.2 + z_1(x_1 - 5.7)),
\end{aligned}$$

and Lorenz system,

$$\begin{aligned}
\dot{x}_2 &= 16(y_2 - x_2) - \varepsilon(x_2 - x_1) + w_2, \\
\dot{y}_2 &= 45.92x_2 - y_2 - x_2 z_2, \\
\dot{z}_2 &= x_2 y_2 - 4z_2,
\end{aligned}$$

using cointegration analysis.

Future work

Some possible extensions to the Dahlhaus paper are

- Formalizing the heuristic arguments.
- Higher dimensionality analysis.
- LASSO algorithm for parameter estimation.
- Bootstrapping procedures to estimate asymptotic distributions for LR tests.

?

Thank you!

Jacob Østergaard

jacobostergaard@gmail.com