Cointegration

Time Series Econometrics

Cointegration

- Let's illustrate with a special case:
- Suppose
  - $y_{1t} \sim I(1)$
  - $\Delta y_{2t} = u_{2t}$ with u_{2t} white noise (so y_{2t} is difference stationary).
  - $y_{1t} = \gamma y_{2t} + u_{1t}$
  - $\mathbf{u_t} = \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix}$, $E(\mathbf{u_t}) = \mathbf{0}$, $E(\mathbf{u_t u_t'}) = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}$

Cointegration

- Then $\Delta y_{1t} = \gamma \Delta y_{2t} + \Delta u_{1t} = \gamma u_{2t} + u_{1t} - u_{1,t-1} \sim I(0)$
  - This is a first-order moving average, with u_{2t} stationary.
- In other words, $y_{1t}, y_{2t} \sim I(1)$ but $y_{1t} - \gamma y_{2t} \sim I(0)$.
  - $[1 \quad -\gamma]'$ is the **cointegrating vector.**

Cointegration

Definition: an $(nX1)$ vector $\mathbf{y_t}$ is said to be cointegrated if each element y_{it} is $I(1)$ but there is a linear combination $\alpha'\mathbf{y_t} \sim I(0)$, where in general α is called the cointegrating vector.

- This is not spurious regression, even though we're regressing random walk on a random walk.
  - In spurious case, there is **no** choice of α that can make the vector stationary.
  - Spurious: $u_{1t} \sim I(1)$
  - Cointegrated: $u_{1t} \sim I(0)$

Cointegration

- In our example, y_{1t} will inherit the random walk that y_{2t} follows, but the cointegrating relationship keeps them close together.
  - They deviate by u_{1t}, which if $I(0)$ will always return to a fixed mean.
  - This is an appealing model of long run market (or ecological or atmospheric) relationships.

Cointegration

- Rewrite as a vector system:

$$\begin{aligned} y_{1t} - \gamma y_{2t} &= u_{1t} \\ y_{2t} &= y_{2t-1} + u_{2t} \end{aligned}$$

$$\begin{bmatrix} 1 & -\gamma \\ 0 & 1 \end{bmatrix} \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix} + \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix}$$

- premultiply by $\begin{bmatrix} 1 & -\gamma \\ 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & \gamma \\ 0 & 1 \end{bmatrix}$

$$\Longrightarrow \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} 0 & \gamma \\ 0 & 1 \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix} + \begin{bmatrix} u_{1t} + \gamma u_{2t} \\ u_{2t} \end{bmatrix}$$

- Let $\begin{bmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{bmatrix} = \begin{bmatrix} u_{1t} + \gamma u_{2t} \\ u_{2t} \end{bmatrix}$, white noise because u's are white noise.

Cointegration

- $E(\underline{\epsilon}_t) = \mathbf{0}$ and $E(\mathbf{e_t e_t'}) = \begin{bmatrix} \sigma_1^2 + \gamma^2\sigma_2^2 & \gamma\sigma_2^2 \\ \gamma\sigma_2^2 & \sigma_2^2 \end{bmatrix}$ if $t = s$, $\mathbf{0}$ otherwise.
- So we could run this vector system in levels.
- $\mathbf{y_t} = \Phi \mathbf{y_{t-1}} + \mathbf{e_t}$, where $\Phi = \begin{bmatrix} 0 & \gamma \\ 0 & 1 \end{bmatrix}$

Cointegration

- Notice a few things
  1. the vector $\mathbf{y_t}$ has a unit root (is nonstationary).

$$|\mathbf{I_2} - \Phi z| = \begin{vmatrix} 1 & -\gamma z \\ 0 & 1 - z \end{vmatrix} = 1 - z = 0 \Rightarrow z = 1$$

  2. we could write the system in rotated hybrid of changes and levels, analogous to rotated Dickey Fuller.

$$\begin{aligned} \mathbf{y_t} - \mathbf{I_2}\mathbf{y_{t-1}} &= (\Phi - \mathbf{I_2})\mathbf{y_{t-1}} + \mathbf{e_t} \\ \begin{bmatrix} \Delta y_{1t} \\ \Delta y_{2t} \end{bmatrix} &= \begin{bmatrix} -1 & \gamma \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix} + \begin{bmatrix} u_{1t} + \gamma u_{2t} \\ u_{2t} \end{bmatrix} \\ \Delta \mathbf{y_t} &= \rho \mathbf{y_{t-1}} + \mathbf{e_t} \end{aligned}$$

Cointegration

- $\Delta \mathbf{y_t} = \rho \mathbf{y_{t-1}} + \mathbf{e_t}$

1. If we should have had a system (VAR) in changes, then $\underline{\rho}$ will be $\mathbf{0}$.
2. If we should have had a system (VAR) in levels (y_{1t}, y_{2t} were stationary), then $\underline{\rho}$ is an arbitrary set of coefficients and has full rank.
3. If the system is cointegrated, $\underline{\rho}$ has rank 1 and can be written as an outer product of two vectors, one if which is the cointegrating vector.
   - $\underline{\rho} = \begin{bmatrix} -1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & -\gamma \end{bmatrix} = \mathbf{b}\alpha'$

Cointegration

- Why would we want to write the system this way: $\Delta \mathbf{y_t} = \rho \mathbf{y_{t-1}} + \mathbf{e_t}$?
- Notice we have $\Delta \mathbf{y_t} = \mathbf{b}\alpha' \mathbf{y_{t-1}} + \mathbf{e_t}$, and $\alpha' \mathbf{y_{t-1}} = u_{1,t-1}$
  - So $\Delta \mathbf{y_t} = \mathbf{b} u_{1,t-1} + \mathbf{e_t}$,
  - Today's changes $\Delta \mathbf{y_t}$ depend on how far y_1 and y_2 deviated from their cointegrating relation last period.
  - Because y_1 and y_2 are random walks, $\Delta \mathbf{y_t}$ should be white noise.
  - Indeed it is, but with a special kind of white noise: u_{1t-1}

Cointegration

- In general: An (nX1) vector $\mathbf{y_t}$ is said to exhibit $h < n$ cointegrating relations if
  1. each element of $\mathbf{y_t} \sim I(1)$
  2. there exists an (nXh) matrix $\mathbf{A}$ of full rank h such that each element of the (hX1) vector $\mathbf{A}'\mathbf{y_t} \sim I(0)$

  – The rows of $\mathbf{A}'$ are called the cointegrating relations or vectors.
- In other words, if $n > 2$, there might be more than one linear combination which is cointegrated.

Cointegration

- Implications for general case:
  - If we can write $\mathbf{y_t}$ as a vector system (VAR) then that system will have a unit root:
    1. $\mathbf{y_t} = \alpha + \Phi_1 \mathbf{y_{t-1}} + ... + \Phi_p \mathbf{y_{t-p}} + \mathbf{e_t}$
    2. $|\mathbf{I_n} - \Phi_1 z - ... - \Phi_p z^p| = 0$ when $z = 1$

Cointegration

- Implications for general case:
  - If previous slide holds, then we can write it as a rotated regression:
    1. $\Delta \mathbf{y_t} = \alpha + \rho \mathbf{y_{t-1}} + \zeta_1 \Delta \mathbf{y_{t-1}} + \ldots + \zeta_{p-1} \Delta \mathbf{y_{t-p+1}} + \mathbf{e_t}$
    2. $\rho = \mathbf{BA}'$ where $\mathbf{B}$ is an nXh matrix of rank h, $\mathbf{A}'$ is an hXn matrix whose rows are the cointegrating vectors.
    3. ρ has rank $h < n$. As before, $\rho = \mathbf{0}$ implies the VAR should be in differences $\Delta \mathbf{y_t}$ and if ρ has rank n, the VAR should be in levels $\mathbf{y_t}$
    4. the $\Delta \mathbf{y}$'s are stationary, and the $\mathbf{y}$'s are nonstationary. The rows of $\mathbf{A}'$ take exactly the linear combinations of the levels to not ruin the stationarity of the remaining differences.

Cointegration

- Implications for general case:
  - If previous slide holds, then we can think of $\mathbf{A}'\mathbf{y_t} = \mathbf{z_t}$ as the stationary residual from the cointegrating relations of $\mathbf{y_t}$ which is often called the "error correction term" and write
    1. $\Delta \mathbf{y_t} = \alpha + \mathbf{B}\mathbf{z_{t-1}} + \zeta_1 \Delta \mathbf{y_{t-1}} + ... + \zeta_{p-1} \Delta \mathbf{y_{t-p+1}} + \mathbf{e_t}$
    2. Notice that this is just a VAR in changes with an extra term (that would have been in the error if we didn't include it).
    3. $\mathbf{z_{t-1}}$ was the degree of deviation from the cointegrating relation last period. $\mathbf{B}$ measures how quickly we return to it.
- Conclusion: with a cointegrated system, we can estimate it as a VAR in levels, or as a VAR in differences with the error correction term.

Estimation and testing of a single cointegrating relation

- Let $\underline{\alpha}' = \begin{bmatrix} 1 & -\alpha_2 & \ldots & -\alpha_n \end{bmatrix}$ be the cointegrating vector.
- $\underline{\alpha}'\mathbf{y_t} \sim I(0)$ for the true α but $\sim I(1)$ for any other α.
- Consider OLS of y_{1t} on $y_{2t}, \ldots y_{n,t}$
- $y_{1t} = \alpha_2 y_{2t} + \ldots + \alpha_n y_{nt} + u_t$
- $\underset{\underline{\alpha}}{min} \sum (y_{1t} - \alpha_2 y_{2t} - \ldots - \alpha_n y_{nt})^2$

Estimation and testing of a single cointegrating relation

- Recall from spurious regression that $\frac{1}{T}\sum u_t^2 \to V$ if u_t stationary, and $\to \infty$ if not.
  - If T is large enough, we will be able to tell the difference from something going to infinity.
- Conclusion: OLS of y_{1t} on remaining y's is a good way to estimate $\underline{\alpha}$ and find out if it's really cointegrated or not.
- If it's not cointegrated, this will be a spurious regression.
- To tell the difference, take
  - Ho: no cointegration (spurious regression, residuals from OLS regression have a unit root).
  - Notice this is the same null hypothesis as the DF test for stationarity, which we can apply to the residuals (however, testing distribution is different).
  - If they are stationary, reject the null and conclude the series are cointegrated.

Estimation and testing of a single cointegrating relation

- Procedure (for example):

1. Estimate by OLS
   1.1 $y_{1t} = \alpha + \gamma_2 y_{2t} + ... + \gamma_n y_{nt} + u_t$
      1.1.1 includes a constant to account for any possible drift.
   1.2 save the residuals $\hat{u}_t$
2. Estimate by OLS
   2.1 $\hat{u}_t = \rho \hat{u}_{t-1} + \zeta_1 \Delta \hat{u}_{t-1} + ... + \zeta_{p-1} \Delta \hat{u}_{t-p+1} + \nu_t$
   2.2 no constant (Case 1).
   2.3 If spurious, $\rho = 1$, fail to reject test null
   2.4 If cointegrated, $\rho < 1$, reject test null.
   2.5 **caveat**: because the null involves a spurious regression, the distribution of ρ will have a **different nonstandard distribution** here than it does in the DF test.
   2.6 This is called a Phillips-Ouliaris-Hansen test - same setup as Dickey-Fuller, including different cases for drift and trend, but with different distributions.

Estimation and testing of multiple cointegrating relationships

- Estimating and testing of more than one cointegrating relation can be done using an MLE-type procedure called Johansen's algorithm
  - ca.jo() function in R's urca package
- Estimate with ρ as an *unrestricted* $n \times n$ matrix

$$\Delta \mathbf{y_t} = \alpha + \rho \mathbf{y_{t-1}} + \zeta_1 \Delta \mathbf{y_{t-1}} + \ldots + \zeta_{\mathbf{p-1}} \Delta \mathbf{y_{t-p+1}} + \mathbf{e_t}$$

  where ρ is an unrestricted $n \times n$ matrix.
- Estimate the restricted regression

$$\Delta \mathbf{y_t} = \alpha + \mathbf{BA}' \mathbf{y_{t-1}} + \zeta_1 \Delta \mathbf{y_{t-1}} + \ldots + \zeta_{\mathbf{p-1}} \Delta \mathbf{y_{t-p+1}} + \mathbf{e_t}$$

  where $\mathbf{B}$ is $n \times (n-1)$ and $\mathbf{A}'$ is $(n-1) \times n$. Compare the fit using a Likelihood Ratio test or compare the eigenvalues of ρ to $\mathbf{BA}'$.
- Repeat with restrictions $n \times (n-2)$, etc.