Lecture: Cointegration

222061-1617: Time Series Econometrics

Spring 2023

Jacek Suda

Outline

Outline:

1. Cointegration
2. VECM

Cointegration

Tale of Two Econometricians

- Stock, James H. and Mark W. Watson (1988), “Variable Trends in Economic Time Series”, Journal of Economic Perspective 2 (3), 147-174.
- Stock and Watson simulate 1000 series of 150 observations.
- Give the data to two econometricians.

Data Generating Process

True Data Generating Process

- Income (à la Friedman (1957))

$$Y_t = Y_t^P + Y_t^T$$

  - Y_t^T is a transitory income:

$$Y_t^T \sim iidN(0, 1),$$

    $\rightarrow$ transitory income is covariance stationary,
  - Y_t^P is a permanent income, non-observable by the econometrician but possibly observable by the agent,

$$Y_t^P = Y_{t-1}^P + u_t, \qquad u_t \sim iidN(0, 1)$$

    $\rightarrow$ permanent income follows a random walk.
    (consistent with rational expectation and Hall (1978)).

Data Generating Process

True Data Generating Process

- Consumption:

$$C_t = Y_t^P$$

  - Consumption and permanent income have the same permanent (random walk) component (consistently with PIH)
  - We will say they are cointegrated.

- Prices:

$$P_t = P_{t-1} + v_t, \qquad v_t \sim iidN(0,1)$$

  - Prices also have random walk (but different than consumption/permanent income).

Econometrician #1

1. Does the consumption depend on price level ?

$$C_t = 9.16 + 0.40 P_t, \qquad (t = 5.12)$$

$\Longrightarrow$ money illusion
But $C_t \sim I(1)$, $P_t \sim I(1)$ and no cointegration $(u_t, v_t \sim iid) \Longrightarrow$ spurious regression

2. Is there a time trend in consumption ?

$$C_t = 2.48 + 0.069 \cdot t, \qquad (t_t = 16.9,\ DW = 0.16)$$

$\Longrightarrow$ time trend and serial correlation (recall DW captures serial correlation)
But $C_t \sim I(1)$ without drift $\Longrightarrow$ spurious cycle.

- *If you regress random walk on time trend, you get something looking like trend-stationary process.*
- *Low DW may indicate that a series is in fact I(1).*

Econometrician #1

3. What is marginal propensity to consume ?

$$\Delta C_t = 0.048 + 0.28\Delta Y_t, \qquad (t = 8.06, DW = 2.27)$$

$\Longrightarrow$ no serial correlation and marginal propensity to consume, $MPC = 28\%$

- But measurement error $< \Delta Y_t = \Delta Y_t^P + \Delta Y_t^T >$
- *MPC =1 => transitory component only; MPC =0 => PIH*

4. Does consumption follow random walk (Hall's (1978) argument)?

$$\Delta C_t = 0.41 - 0.041 C_{t-1}, \qquad (t = -2.15, DW = 1.98)$$

$\Longrightarrow$ at 5% level PIH fails, lagged consumption is a useful predictor of future changes in consumption.

- But, $t_{\beta=0} \sim DF \neq N(0,1)$ and with $t = -2.15$ cannot reject the unit root.
- DW critical values are -2.88 at 5% and -2.57 at 10%.

Econometrician #2

1. $$C_t = 0.31 + 0.94Y_t, \qquad (t_Y = -2.274)$$

$\Longrightarrow$ MPC large but statistically significant less than 1

- $C_t \sim I(1)$ and $Y_t \sim I(1)$ are cointegrated ($Y_t - C_t \sim I(0)$).
- Despite $\hat{\beta} \xrightarrow{p} 1$ at rate T (superconsistency), there is asymptotic bias.

2. $$C_t = 0.45 + 0.97C_{t-1} - 0.01C_{t-2}, \qquad (t_{C_1} = 11.7,\ t_{C_2} = -0.13)$$

$\Longrightarrow$ PIH,
(once we control for lagged consumption, there is no predictability left)

$$C_t = (\alpha + \beta)C_{t-1} - \beta(C_{t-1} - C_{t-2}) + \varepsilon_t,\ t_\beta \sim N(0, 1)$$

since $\Delta C_{t-1} = C_{t-1} - C_{t-2} \sim I(0)$

Econometrician #2

3.

$$C_t = 0.41 + 1.03C_{t-1} - 0.07Y_{t-1}, \qquad (t_Y = -0.97)$$

$\Longrightarrow$ PIH

$$C_t = (\beta + \beta_Y)C_{t-1} + \beta_Y(Y_{t-1} - C_{t-1}) + \xi_t, \; t_{\beta_Y} \sim N(0,1)$$

since $Y_{t-1} - C_{t-1} \sim I(0)$

4. Do prices help?

$$C_t = 0.47 + 0.95C_{t-1} + 0.004P_{t-1} + 0.06\Delta P_{t-1}, \quad (t_P = 0.17 \;\; t_{\Delta P} = 0.87)$$

$\Longrightarrow$ PIH and no money illusion

- Note: consistent estimates, no spurious regression despite I(1)

Cointegration

Definition

A $(n \times 1)$ vector y_t is cointegrated if $y_t \sim I(1)$ and there exists $(n \times 1)$ vector $a = [a_{11} \quad a_{12} \quad \ldots \quad a_{1n}] \neq 0$ such that

$$a' y_t = a_{11} y_{1t} + a_{12} y_{2t} + \ldots + a_{1n} y_{nt} \sim I(0).$$

Cointegration

- Problem: A cointegrating vector, a, is not unique since $c \cdot a' y_t \sim I(0)$ for any scalar c.

- For $(n \times 1)$ vector y_t there may be up to $n-1$ linearly independent vectors, $a_1, a_2, \ldots, a_{n-1}$ such that $a_i' y_t \sim I(0)$ for $i = 1, \ldots, n-1$, i.e. $A' y_t \sim I(0)$ with A forming "basis for cointegrating space".

Representation

Beveridge-Nelson Decomposition

Beveridge-Nelson decomposition:

If $(1-L)y_t = \mu + a(L)\varepsilon_t$ then we can write

$$y_t = c_t + \tau_t$$

where

$$\begin{aligned} \tau_t &= \mu + \tau_{t-1} + a(1)\varepsilon_t \\ c_t &= a^*(L)\varepsilon_t, \qquad a_j^* = -\sum_{k=j+1}^{\infty} a_k \end{aligned}$$

Beveridge-Nelson Decomposition: Proof

Proof:

Any lag polynomial $a(L)$ can be written as

$$a(L) = a(1) + (1-L)a^*(L), \qquad a_j^* = -\sum_{k=j+1}^{\infty} a_k = -a_{j+1} - a_{j+2} - a_{j+3} - \ldots .$$

since

$$\begin{array}{rccccccc}
a(1) = & a_0 & +a_1 & +a_2 & +a_3 & +a_4 & \ldots \\
(1-L)a^*(L) = & & -a_1 & -a_2 & -a_3 & -a_4 & \ldots \\
 & & +a_1 L & +a_2 L & +a_3 L & +a_4 L & \ldots \\
 & & & -a_2 L & -a_3 L & -a_4 L & \ldots \\
 & & & +a_2 L^2 & +a_3 L^2 & +a_4 L^3 & \ldots \\
 & & & & -a_3 L^2 & -a_4 L^3 & \ldots
\end{array}$$

Then,

$$(1-L)y_t = \mu + a(L)\varepsilon_t = \mu + a(1)\varepsilon_t + (1-L)a^*(L)\varepsilon_t = (1-L)\tau_t + (1-l)c_t \qquad \square$$

Representation 1: Vector MA

- If $y_t \sim I(1)$ then $\Delta y_t \sim I(0)$ with the Wold form

$$\Delta y_t = \delta + \Psi(L) e_t, \qquad e_t \sim WN(0, \Omega).$$

- Beveridge-Nelson decomposition allows us to write it down as

$$y_t = y_0 + \delta \cdot t + \Psi(1) \sum_{j=1}^{t} e_j + (\hat{e}_t - \hat{e}_0)$$

where

$$\hat{e}_t = \sum_{k=0}^{\infty} \Psi_k^* e_{t-k} \sim I(0)$$

Representation 1: Vector MA

- If $A'y_t \sim I(0)$ then

$$A'y_t = A'y_0 + A'\delta \cdot t + A'\Psi(1)\sum_{j=1}^{t} e_j + A'\hat{e}_t - A'\hat{e}_0 \sim I(0)$$

- For $A'y_t$ to be stationary we need
  - (i) $A'\delta = 0$, and
  - (ii) $A'\Psi(1) = 0$.
- Since A is a non-zero matrix, $\Psi(1)$ must be singular and $\Lambda = \Psi(1)\Omega\Psi(1)'$ is singular $\Longrightarrow$ $rank(\Lambda) < n$. $\Longrightarrow$ not all shocks have long run impact on y_t.
- Also, if $\Psi(1)$ is singular, $\Psi(L)^{-1}$ does not exist and, therefore, there is no finite-order VAR representation for Δy_t.

Phillips' Triangular Representation

Structural Equations
(Static long-run equilibrium)

$$
\begin{aligned}
y_{1t} &= \beta_{12}y_{2t} + \beta_{13}y_{3t} + \ldots + \beta_{1n}y_{nt} + u_{1t}, \quad u_{1t} \sim I(0) \\
y_{2t} &= \beta_{23}y_{3t} + \beta_{24}y_{4t} + \ldots + \beta_{2n}y_{nt} + u_{2t}, \quad u_{2t} \sim I(0)
\end{aligned}
$$

- there can be up to $n-1$ such equations.

Phillips' Triangular Representation

Reduced-form Equations

$$
\begin{aligned}
y_{3t} &= y_{3t-1} + u_{3t}, \qquad & u_{3t} \sim I(0) \\
y_{4t} &= y_{4t-1} + u_{4t}, \qquad & u_{4t} \sim I(0) \\
u_t &= \Psi(L)e_t, \qquad & e_t \sim iid(0, \Omega)
\end{aligned}
$$

- *If y_{3t} relevant for y_{4t} it will be reflected in correlation between u_{3t} and u_{4t}.*

Example: PPP

Example: Purchasing Power Parity

$$e_t = p_t - p_t^*$$

with $p_t, p_t^* \sim I(1)$, has Phillips' triangular representation with

$$\begin{aligned}
&\text{Structural equation} \\
e_t &= \beta_1 p_t + \beta_2 p_t^* + u_{1t} \\
&\text{Reduced-form equations} \\
p_t^* &= p_{t-1}^* + u_{2t} \\
p_t &= p_{t-1} + u_{3t}, \qquad u_{it} \sim I(0)
\end{aligned}$$

with $\beta_1 = 1$, $\beta_2 = -1$ and the cointegrating vector $A' = [1 \quad -1 \quad 1]$ for $y_t = [e_t \quad p_t \quad p_t^*]'$.

- Note that the # of reduced-form equations is at most $n-1$, and # of structural equation equals # of cointegrating relations.

Example: PPP 2

- Assume that price level in foreign country is linked to price level in domestic country. Then

$$
\begin{array}{rcl}
\text{Structural equations} & & \\
e_t & = & p_t - p_t^* + u_{1t} \\
p_t^* & = & p_{t-1} + u_{2t} \\
\text{Reduced-form equation} & & \\
p_t & = & p_{t-1} + u_{3t}, \qquad u_{it} \sim I(0).
\end{array}
$$

- There are two cointegrating vectors and matrix A has triangular representation with

$$
A'z = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} e_t \\ p_t^* \\ p_t \end{bmatrix} \sim I(0).
$$

- This model has one stochastic trend and two cointegrating relationships.

Stock and Watson's Common Trend Representation

- Cointegrated series share the same unobserved random walk component.
- Example
  - Bivariate system

$$
\begin{aligned}
y_{1t} &= \beta\tau_t + u_{1t} \\
y_{2t} &= \tau_t + u_{2t} \\
\tau_t &= \tau_{t-1} + v_t, \quad v_t \sim iid.
\end{aligned}
$$

  - y_{1t} and y_{2t} are cointegrated because

$$y_{1t} - \beta y_{2t} = u_{1t} - \beta u_{2t} \sim I(0).$$

TESTING

Pre-specified cointegrating vector

- Consider that cointegrating vector a is given by theory.
- Then $z_t = a'y_t$ is a cointegrating residual.

Example: real exchange rate

$$z_t = e_t - p_t + p_t^*$$

Hypothesis

$$\begin{aligned} H_0 &: \quad \text{no cointegration: } z_t \sim I(1) \\ H_1 &: \quad \text{cointegration: } z_t \sim I(0) \end{aligned}$$

$\Longrightarrow$ unit root test on z_t.

- *We could take alternatively* $H_0: \; z_t \sim I(0)$ *and do the stationarity test.*

Unknown cointegrating vector

Engle-Granger Test

- Step 1: Estimate a by static OLS regression

$$y_{1t} = \textit{deteministic term } + \hat{\beta}_2 y_{2t} + \ldots + \hat{\beta}_n y_{nt} + \hat{z}_t$$

so $a' = [1 \quad -\beta_1 \quad \ldots \quad -\beta_n]$

- Step 2: ADF unit root test on $\hat{z}_t$.

Unknown cointegrating vector

ADF regression (no constant, no drift)

$$\hat{z}_t = \rho \hat{z}_{t-1} + \sum_{j=1}^{k} \gamma_j \Delta \hat{z}_{t-j} + \varepsilon_t$$

and

$$t_{\rho=1} = \frac{\hat{\rho} - 1}{\widehat{SE}(\hat{\rho})}.$$

- Superconsistency makes it feasible to estimate
- The distribution of the test statistics is shifted further left than DF case (see Hamilton, Table B.9).

  Case I: no deterministic trend
- Case II: constant + no drift in $y_{1t}, \ldots, y_{nt}$
- Case III: constant + drift in $y_{1t}, \ldots, y_{nt}$

Recap

- Any stationary process has Wold form representation.
- Cointegration has a special version

$$\Delta y_t = \begin{bmatrix} \Delta y_{1t} \\ \Delta y_{2t} \end{bmatrix} = \mu + \bar{\Psi}(L)e_t, \quad e_t \sim WN(0, \Omega),$$

  where $\bar{\Psi}(L)^{-1}$ does not exist.
- It does not have VAR representation but VEC (vector error correction).
- Cointegration implies that there is a long run stochastic equilibrium relationship between variables even though they are $I(1)$.
- For example, $y_t = c_t + v_t$, $c_t = c_{t-1} + u_{2t}$.
  - General Stock-Watson representation: y_t and c_t share the same underlying random walk component:

$$\begin{aligned} y_t &= \tau_t + u_{1t} \\ \tau_t &= \tau_{t-1} + \mu + v_t \\ c_t &= \tau_t + u_{2t} \end{aligned}$$

    with $u_{1t}, u_{2t} \sim I(0)$.

Recap

Testing for cointegration:

1. Known cointegrating vector
   - Construct $z_t = y_t - c_t$.
   - Do DF for $H_0 : \ z_t \sim I(1)$: no cointegration
2. If vector unknown
   - Estimate $y_t = \alpha + \beta c_t + z_t$.
   - Construct $\hat{z}_t = y_t - \hat{\alpha} - \hat{\beta} c_t$
   - Test for unit root in $\hat{z}_t$ but use corrected statictics (not DF)

Estimation

Levels Estimator

- Consider

$$y_{1t} = \beta y_{2t} + u_t$$

- $\hat{\beta}^{OLS}$ is superconsistent (collapse to β at rate T) but is asymptotically bias

$$\begin{aligned} T(\hat{\beta}^{OLS} - \beta) &\xrightarrow{F} \text{mixture of N and DF} \\ &\nsim N(0, \Upsilon) \end{aligned}$$

- To solve the bias problem:
  - Use alternative estimator.
  - For example $\hat{\beta}^{median}$ - no bias but difficult to compute.

Error Correction Estimator

- Consider an equation

$$\Delta y_{1t} = \beta_1 \Delta y_{2t} + \gamma(y_{1,t-1} - \beta y_{2,t-1}) + u_t,$$

  where γ is the error correction coefficient, β is the long-run coefficient and β_1 is short-run multiplier.
- If y_{1t}, y_{2t} cointegrated, all terms in the equation are $I(0)$.

Error Correction Estimator

- Identification:

$$\Delta y_{1t} = \beta_1 \Delta y_{2t} + \delta_1 y_{1,t-1} + \delta_2 y_{2,t-1} + u_t$$

  where $\delta_1 = \gamma$, $\delta_2 = \gamma\beta$ and $\beta = -\frac{\delta_2}{\delta_1}$.
- Then $\hat{\beta} = -\frac{\hat{\delta}_2^{OLS}}{\hat{\delta}_1^{OLS}}$ is superconsistent and less biased than $\hat{\beta}^{OLS}$.
- *To use error correction model, we need a correct structure of the long-run relationship.*
- *If the error correction works the other way, i.e.* $\Delta y_{2t} = \ldots$, *we can get* $\gamma = 0$ *and misspecification.*

Dynamic OLS/GLS

- Stock and Watson (1993, *Econometrica*), Saikkonen (1991, *Econometric Theory*)
- Consider

$$y_{1,t} = \beta y_{2,t} + \sum_{j=-p}^{p} \gamma_j \Delta y_{2,t+j} + u_t$$

  - first term capture LR relationship, and
  - second term captures serial correlation in lags (y_{2t} predicts y_{1t}) and leads (y_{1t} predicts y_{2t}).
- Then, $\hat{\beta}^{OLS}$ is superconsistent and asymptotically unbiased.
- u_t can be serially correlated but then use Newton-West or Cochrane-Orcurr standard errors when testing

$$t = \frac{\hat{\beta}^{OLS} - \beta_0}{\widehat{SE}(\hat{\beta})}.$$

VECTOR ERROR CORRECTION MODEL

VAR

- Assume that $y_t \sim I(1), A'y_t \sim I(0)$ and that y_t has finite-order VAR representation.

$$y_t = \Phi_1 y_{t-1} + \ldots \Phi_p y_{t-p} + e_t, e_t \sim iid(0, \Omega)$$

- Recall that $A'y_t \sim I(0)$ implies that Δy_t has no VAR representation.
- But it has VECM (vector error correction model) representation.

- DF transformation:

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

- Subtract y_{t-1} to get a vector generalization of error correction model

$$\begin{aligned} \Delta y_{t_{n\times 1}} &= \Pi_{n\times n} y_{t-1} + \zeta_1 \Delta y_{t-1} + \ldots + \zeta_{p-1} \Delta y_{t-p+1} + e_t \\ \Pi &= \rho - \mathbb{I} = \Phi_1 + \Phi_2 + \ldots + \Phi_p - \mathbb{I} = -\Phi(1) \end{aligned}$$

- If $rank(\Pi) = 0 \Longrightarrow \Pi = 0, y_t \sim I(1)$ with n unit roots (no cointegration), $\Longrightarrow \Delta y_t \sim VAR(p-1)$.
- If $rank(\Pi) = n$ (full rank) $\Longrightarrow y_t \sim I(0)$, 0 unit roots (no cointegration), $\Longrightarrow y_t \sim VAR(p)$.
- If $rank(\Pi) = r < n \Longrightarrow h$ linearly independent cointegrating vectors $(h = n - r)$ with basis $A = [a'_1 \ \ldots \ a'_h]$.

$$\Pi = \underset{n \times h}{B} \underset{h \times n}{A'}$$

$$\Longrightarrow \Delta y_t = B(A' y_{t-1}) + \zeta_1 \Delta y_{t-1} + \ldots + \zeta_{p-1} \Delta y_{t-p+1} + e_t,$$

with $A' y_{t-1} \sim I(0)$ providing levels information.

- Johansen: Given normality, estimate via MLE.

VECM: Example

- Let

$$\begin{bmatrix} c_t \\ y_t \end{bmatrix} \sim VAR(1) \sim I(1)$$

- but

$$c_t - y_t \sim I(0)$$

- As VAR(1) in level, we have no lags in VECM.

VECM: Example

- Then

$$
\begin{aligned}
\Delta c_t &= b_{11}(c_{t-1} - y_{t-1}) + e_{1t} \\
\Delta y_t &= b_{12}(c_{t-1} - y_{t-1}) + e_{2t}
\end{aligned}
$$

- b_{11}, b_{12} determine adjustment of y_t, c_t to cointegrating (long-run) relationship.
- *e.g. if c_t above its long run value, it should fall down, i.e. $b_{11} < 0$.*

- Companion form: VAR(1)$\sim I(0)$

$$\begin{bmatrix} \Delta c_t \\ \Delta y_t \\ c_t - y_t \end{bmatrix} = \begin{bmatrix} 0 & 0 & b_{11} \\ 0 & 0 & b_{12} \\ 0 & 0 & 1 + b_{11} - b_{12} \end{bmatrix} \begin{bmatrix} \Delta c_{t-1} \\ \Delta y_{t-1} \\ c_{t-1} - y_{t-1} \end{bmatrix} + \begin{bmatrix} e_{1t} \\ e_{2t} \\ e_{1t} - e_{2t} \end{bmatrix}$$

- Or

$$\beta_t = F \cdot \beta_{t-1} + v_t,$$

with the Wold form

$$\beta_t = (\mathbb{I} - F)^{-1} v_t.$$

- *Here, v_t is a forecast error.*
- Companion to general form:

$$A' = [1 \quad -1], \quad \text{so that } A' y_{t-1} = c_{t-1} - y_{t-1} \sim I(0)$$

and

$$B' = [b_{11} \quad b_{12}]$$

TESTING

Cointegrating vector known

- Estimate VECM rather than VAR in levels.

$$H_0: \text{ no cointegration } \iff B = 0$$

  e.g. $(b_{11} = b_{12} = 0)$.

- *If we estimate VECM equation by equation, b_{11} and b_{12} have standard distribution but the joint distribution of b_{11} and b_{12} (i.e. to test $b_{11} = b_{12} = 0$) is not standard.*
- Horvath and Watson's Wold test

$$W = \hat{B}' \widehat{var}(\hat{B})^{-1} \hat{B}, \qquad \hat{B} = \begin{pmatrix} \hat{b}_{11} \\ \hat{b}_{12} \end{pmatrix}$$

- Under $H_0: A'y_t \sim I(1)$, $W \sim$ Multivariate DF.
- Wald test has higher power than ADF residual test if data from VAR.

Cointegrating vector unknown

- The model

$$\Delta y_t = \gamma X_t + \Pi y_{t-1} + \zeta_1 \Delta y_{t-1} + \ldots + \zeta_{p-1} \Delta y_{t-p+1} + e_t$$

  - X_t is deterministic and $e_t \sim N(0, \Omega)$.
- If y_t is cointegrated then $A' y_t = 0$ and $\Pi = -BA'$.
- Johansen's maximum likelihood approach is to maximize $L(\gamma, \Pi, \zeta_1, \ldots, \zeta_{p-1}, \Omega)$ subject to $\Pi = -BA'$.

Johansen's MLE

Step 1

- Regress Δy_t and y_{t-1} on $X_t, \Delta y_{t-1}, \ldots, \Delta y_{t-p+1}$ via OLS.

$$
\begin{aligned}
\Delta y_t &= \xi_0 + \Gamma_1 X_t + \Xi_1 \Delta y_{t-1} + \Xi_2 \Delta y_{t-2} + \ldots + \Xi_{p-1} \Delta y_{t-p+1} + v_{1t} \\
y_{t-1} &= \upsilon_0 + \Gamma_2 X_t + \Upsilon_1 \Delta y_{t-1} + \Upsilon_2 \Delta y_{t-2} + \ldots + \Upsilon_{p-1} \Delta y_{t-p+1} + v_{2t}
\end{aligned}
$$

Johansen's MLE

Step 2

- Construct residuals

$$\begin{aligned} \hat{v}_{1t} &= \Delta y_t - \Delta \hat{y}_t \\ \hat{v}_{2t} &= y_{t-1} - \hat{y}_{t-1} \end{aligned}$$

  and calculate the sample variance-covariance matrixes of the OLS residuals

$$\hat{\Sigma}_{ij} = \frac{1}{T} \sum_{t=1}^{T} \hat{v}_{it} \hat{v}'_{jt}, \qquad i, j = 1, 2.$$

- Find eigenvalues $\hat{\lambda}_n$ and eigenvectors $\hat{a}_n$ for

$$A = \hat{\Psi}'_{22} \hat{\Sigma}_{21} \hat{\Sigma}_{11}^{-1} \hat{\Sigma}_{12} \hat{\Psi}_{22}$$

  where $\hat{\Psi}_{22} \hat{\Psi}'_{22} = \hat{\Sigma}_{22}^{-1}$. (Cholesky factorization)
- Order eigenvalues from largest to smallest

$$\lambda_i > \lambda_j \qquad \text{for } i < j$$

Johansen's MLE

Step 3

- Calculate maximum likelihood for restricted (under cointegration) model

$$L^* = -\frac{nT}{2}(\ln(2\Pi) + 1) - \frac{T}{2}\sum_{i=1}^{r}\ln(1-\lambda_i),$$

where r is the number of non-zero eigenvalues of Π for restricted model.

Johansen's MLE

Step 4

- Trace statistics:

$$\begin{aligned} H_0 &: \quad h = h_1 \\ H_1 &: \quad h = h_2, \qquad \text{for } h_1 < h_2 < n \end{aligned}$$

h: number of cointegrating vectors. The test statistic

$$LR = -T \sum_{i=h_1+1}^{h2} \ln(1 - \lambda_i)$$

has a non-standard distribution (depends on X_t).

Johansen's MLE

Step 4

- $\lambda_{max}-$statistics

$$\begin{aligned} H_0 &: h \quad \text{cointegrating vectors} \\ H_0 &: h+1 \quad \text{cointegrating vectors} \end{aligned}$$

The test statistic

$$LR = -T\ln(1-\lambda_{max})$$

where λ_{max} is the largest eigenvalue consider to be zero under the null (i.e. $\lambda_{max} = \lambda_{h+1}$).

ESTIMATION

Johansen's MLE

Step 5

- Let $\hat{a}_1, \ldots, \hat{a}_h$ denote $(n \times 1)$ eigenvectors of A associated with the h largest eigenvalues, forming a base for the space of cointegrating relations, $a = b_1\hat{a}_1 + \ldots + b_h\hat{a}_h$ for some $b = (b_1, \ldots, b_h)$.
- Normalize those vectors so that $\hat{a}_i'\widehat{\Sigma}_{22}\hat{a}_i = 1$ and collect first h normalized vectors into

$$\hat{A} \equiv [\hat{a}_1 \quad \hat{a}_2 \quad \cdots \quad \hat{a}_h]$$

- MLE of parameters can be computed from

$$\begin{aligned}\hat{\Pi} &= \hat{\Sigma}_{12}\hat{A}\hat{A}' \\ \hat{\Omega} &= \frac{1}{T}\sum_{t=1}^{T}[(\hat{u}_{1t} - \hat{\Pi}\hat{u}_{2t})(\hat{u}_{1t} - \hat{\Pi}\hat{u}_{2t})']\end{aligned}$$

GARCH

GARCH in a nutshell

- Consider a stochastic process $\{y_t\}$ such that

$$y_t = \mu + \varepsilon_t$$

- and the conditional mean of the process to be

$$E_{t-1} y_t = \mu \Longrightarrow E_{t-1} \varepsilon_t = 0.$$

GARCH in a nutshell

- Assume that ε_t is conditionally heteroskedastic

$$\varepsilon_t = \sigma_t z_t, \qquad z_t \sim iid(0,1)$$

- Then, the conditional variance of the process is

$$Var_{t-1} y_t = E_{t-1}\varepsilon_t^2 = \sigma_t^2 E_{t-1} z_t^2 = \sigma_t^2$$

- We will call $\{y_t\}$ a ARMA(P,Q)-GARCH(p,q) process if

$$\begin{aligned}
y_t &= \mu + \sum_{i=1}^{P} \phi_i y_{t-i} + \sum_{j=1}^{Q} \theta_j \varepsilon_{t-j} + \varepsilon_t \\
\varepsilon_t &= \sigma_t z_t, \qquad z_t \sim iid(0,1) \\
\sigma_t^2 &= \omega_0 + \sum_{l=1}^{p} \alpha_l \varepsilon_{t-l}^2 + \sum_{k=1}^{q} \alpha_k \sigma_{t-k}^2
\end{aligned}$$