CH5 Cointegration and Vector Error Correction

Long-run equilibrium
Cointegration, Cointegrating vector
Vector Error Correction and VAR
Testing for cointegration
Residual-based Engle-Granger method
Johanson-Stock-Waston method
Characteristic Roots and Rank

An Example: VAR or VEC ?

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

u_{1t} and u_{2t} uncorrelated white noice processes.

Both y_{1t} and y_{2t} are I(1) processes:

$$\begin{aligned} \Delta y_t &\equiv \begin{pmatrix} \Delta y_{1t} \\ \Delta y_{2t} \end{pmatrix} = \begin{pmatrix} u_{1t} - u_{1,t-1} + \gamma u_{2t} \\ u_{2t} \end{pmatrix} \\ &= \begin{pmatrix} 1-L & \gamma \\ 0 & 1 \end{pmatrix} \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix} \equiv \Psi(L) u_t. \end{aligned}$$

noninvertible

$$\begin{aligned} \Delta y_t &= \begin{pmatrix} -1 & \gamma \\ 0 & 0 \end{pmatrix} y_{t-1} + u_t \\ &= \begin{pmatrix} -1 \\ 0 \end{pmatrix} (y_{1,t-1} - \gamma y_{2,t-1}) + u_t \\ &= \begin{pmatrix} -1 \\ 0 \end{pmatrix} (1, -\gamma)\, y_{t-1} + u_t \equiv \alpha\beta' y_{t-1} + u_t. \end{aligned}$$

Long-run equilibrium and Cointegration

Long-run equilibrium: $\beta_1 x_{1t} + \beta_2 x_{2t} + \cdots + \beta_n x_{nt} = 0$ or $\beta' x_t = 0$,

where $\beta = (\beta_1, \beta_2, \cdots, \beta_n)'$, $x_t = (x_{1t}, x_{2t}, \cdots, x_{nt})'$.

Equilibrium error——the deviation from the long-run equilibrium: $\beta' x_t = e_t$,

where $\{e_t\}$ is stationary.

Cointegration: If $z_t \sim I(d)$ and $y_t \sim I(d)$, it is generally true that $z_t - a y_t \sim I(d)$. Further, when $z_t - a y_t \sim I(d-1)$, we say that z_t and y_t are cointegrated. More formally, the components of vector x_t are said to be **Cointegrated of order** d, b, denoted $x_t \sim CI(d, b)$, if

1) all components of x_t are $I(d)$;

2) $\exists$ a vector $\beta = (\beta_1, \beta_2, \cdots, \beta_n)' \neq 0$ such that the **linear** combination $\beta' x_t = \beta_1 x_{1t} + \beta_2 x_{2t} + \cdots + \beta_n x_{nt} \sim I(d-b)$ where $b > 0$.

The vector β is called **the cointegrating vector,** which represents the long-run equilibrium relationship among variables.

Long-run equilibrium and Cointegration

Remarks: (1) Cointegration refers to a **linear combination** of nonstationary variables. (2) The cointegrating vector is not unique: the set of cointegrating vectors constitutes a vector subspace satisfying $\beta_1 x_{1t} + \beta_2 x_{2t} + \cdots + \beta_n x_{nt} \sim I(d-b)$; that is,

$$\{\beta = (\beta_1, \beta_2, \cdots, \beta_n)' \in R^n \backslash \{0\} : \beta' x_t \sim I(d-b)\}$$

When β is a cointegrating vector, $\lambda\beta$ is also a cointegrating vector for all $\lambda \neq 0$. **A normalized integrating vector** is $\beta/\beta_1 = (1, \beta_2/\beta_1, \cdots, \beta_n/\beta_1)'$ if $\beta_1 \neq 0$. There may be at most $n-1$ linearly independent cointegrating vectors. The number of linearly independent cointegrating vectors is called **the cointegrating rank of** x_t. (3) Convention: Here assume that $x_t \sim CI(1,1)$ s.t. $\beta' x_t \sim I(0)$.

Cointegration: Example 1

Money demand is proportional to the price level; as income increases, indviduals will want to hold increased money balances; money demand is negatively related to the interest rate.

$$m_t = \beta_0 + \beta_1 p_t + \beta_2 y_t + \beta_3 r_t + e_t$$

where m_t is the money demand (= supply, in equilibrium), p_t is the price level, y_t is the real income, r_t is the interest rate, and e_t is stationary disturbance term. Here $\beta_1 = 1$, $\beta_2 > 0$ and $\beta_3 < 0$ by the behavioral assumptions. When all the variables are I(1), $(m_t, p_t, y_t, r_t)'$ is cointegrated with a cointegration vector $(1, -\beta_1, -\beta_2, -\beta_3)'$.

Also, suppose that the monetary authorities followed a feedback rule such that they decreased the money supply when nominal GDP was high and increased the money supply when nominal GDP was low. Then

$$\begin{aligned} m_t &= \gamma_0 - \gamma_1 (p_t + y_t) + e_{1t} \\ &= \gamma_0 - \gamma_1 p_t - \gamma_1 y_t + 0 \cdot r_t + e_{1t} \end{aligned}$$

where e_{1t} is stationary. Then $(m_t, p_t, y_t, r_t)'$ is also cointegrated with another cointegration vector $(1, \gamma_1, \gamma_1, 0)'$, which is linearly independent of $(1, -\beta_1, -\beta_2, -\beta_3)'$.

Cointegration: Example 2

PPP: $P_t = S_t P_t^*$ or taking log, $p_t = s_t + p_t^*$.

Each of the three variables p_t, s_t, and p_t^* is $I(1)$.

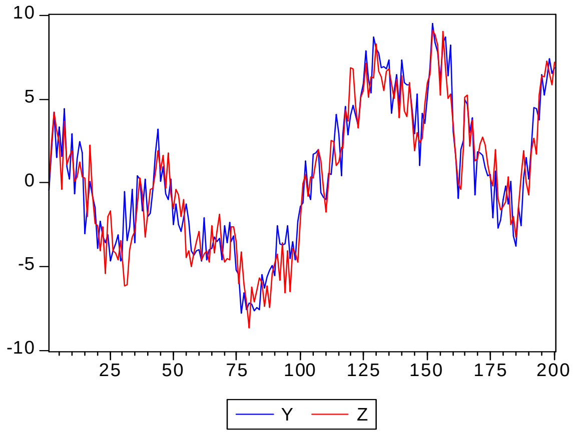
A weak version of the hypothesis is that the variable $z_t \equiv (p_t - s_t - p_t^*)$ is stationary, i.e. $(p_t, s_t, p_t^*)'$ is cointegrated with $(1, -1, -1)'$.

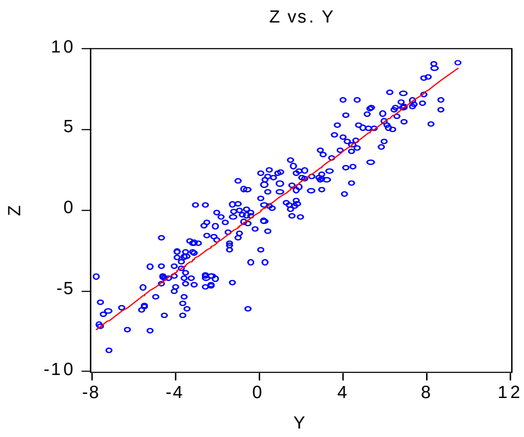
Cointegration and Trend: CH6-ex1:

Graph for two random walk plus noise processes:

$$y_t = \mu_t + \varepsilon_{yt}, \quad z_t = \mu_t + \varepsilon_{zt}$$

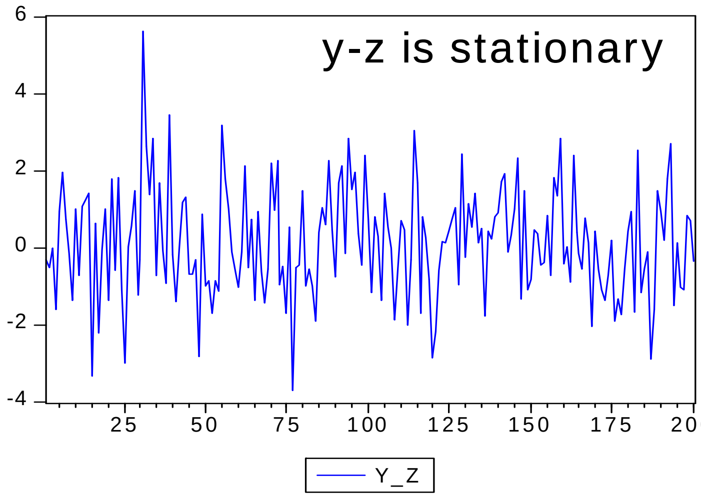
where $\mu_t = \mu_{t-1} + u_t.$

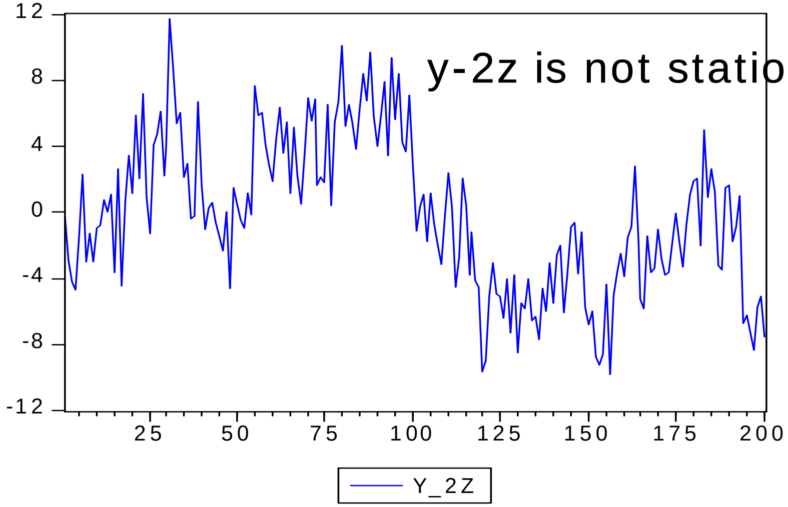




$\boxed{When\ is\ \beta_1 y_t + \beta_2 z_t\ stationary?}$

The parameters of the cointegration vector purge the trend from the linear combination of the cointegrated variables



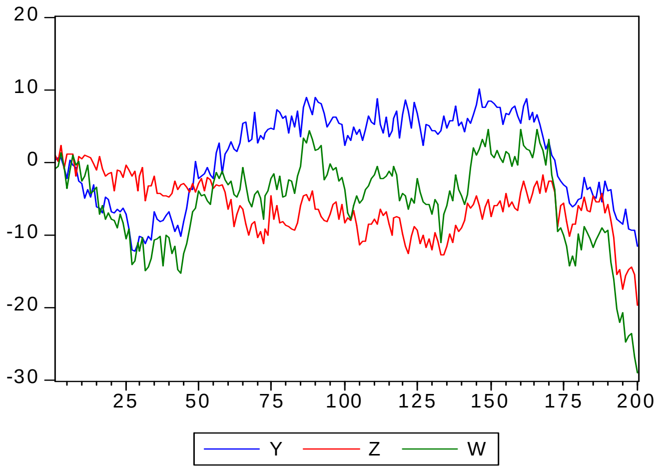


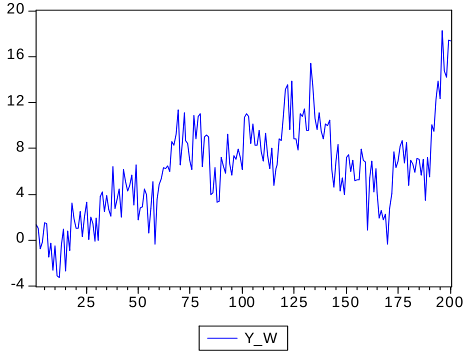
Cointegration and Trend: CH6-ex2

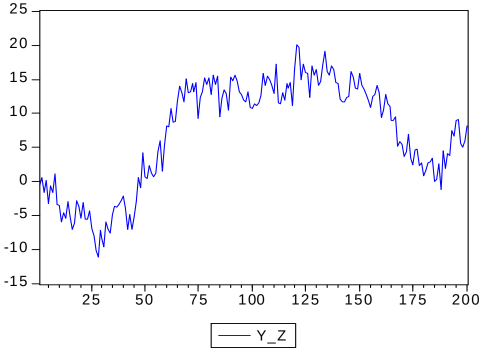
Graph for three random walk plus noise processes

$$y_t = \mu_{yt} + \varepsilon_{yt}, z_t = \mu_{zt} + \varepsilon_{zt}, w_t = \mu_{wt} + \varepsilon_{wt}$$

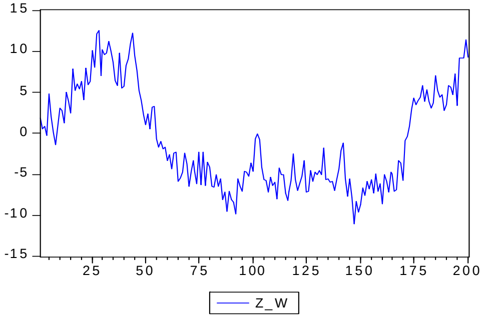
where $\mu_{wt} = \mu_{yt} + \mu_{zt}$.

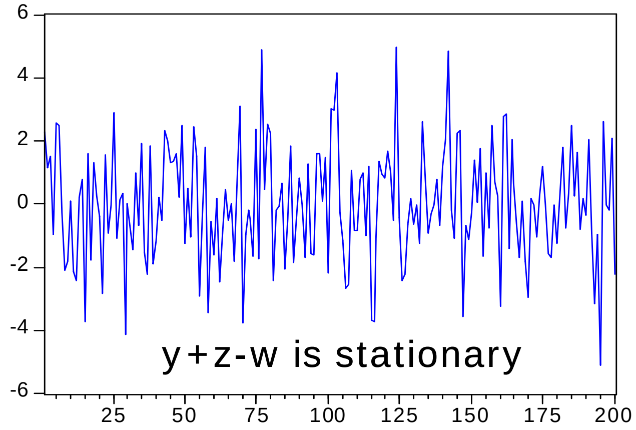






not stationary





$$y_t + z_t - w_t = \varepsilon_{yt} + \varepsilon_{zt} - \varepsilon_{wt} \sim I(0)$$

Cointegration: Purge the trend

Consider the vector representation: $x_t = \mu_t + e_t$

where $x_t = (x_{1t}, \cdots, x_{nt})'$, $\mu_t = (\mu_{1t}, \cdots, \mu_{nt})'$ is the vector of stochastic trends, and e_t is an $n \times 1$ vector of stationary components. If one trend can be expressed as a linear combination of the other trends in the system, i.e. there exists a vector $\beta = (\beta_1, \cdots, \beta_n)'$ such that $\beta'\mu_t = 0$, then $\beta' x_t = \beta' e_t \sim I(0)$. That is, x_t is integrated with $(\beta_1, \cdots, \beta_n)'$.

Suppose $\varepsilon_{yt}, \varepsilon_{zt}, \varepsilon_{wt}, \varepsilon_t$ are i.i.d. white noise processes, and

$$\begin{aligned} y_t &= \mu_{yt} + \varepsilon_{yt}, \quad z_t = \mu_{zt} + \varepsilon_{zt}, \quad w_t = \mu_{wt} + \varepsilon_{wt}, \\ \mu_{yt} &= \mu_{y,t-1} + \varepsilon_t, \ \ \mu_{zt} = \mu_{z,t-1} + \varepsilon_t, \ \ \mu_{wt} = \mu_{yt} + \mu_{zt}. \end{aligned}$$

Then $y_t + z_t - w_t = \varepsilon_{yt} + \varepsilon_{zt} - \varepsilon_{wt} \sim I(0)$, i.e. $(1, 1, -1)\,(y_t, z_t, w_t)' \sim I(0)$. $(y_t, z_t, w_t)'$ is integrated with $(1, 1, -1)'$. Here the stochastic trend in the cointegration is also purged.

Cointegration and Error Correction

An example: Term structure of the long- and short-term interest rates.

$$\boxed{\begin{aligned} \Delta r_{St} &= \alpha_S(r_{Lt-1} - \beta r_{St-1}) + \varepsilon_{St}, \qquad \alpha_S > 0 \\ \Delta r_{Lt} &= -\alpha_L(r_{Lt-1} - \beta r_{St-1}) + \varepsilon_{Lt}, \qquad \alpha_L > 0 \end{aligned}}$$

where r_{St} and r_{Lt} are the long- and short-term interest rates, and ε_{St} and ε_{Lt} are white-noise disturbance terms which may be correlated. The long- and short-term interest rates change in response to stochastic shocks (ε_{St} and ε_{Lt}) and in response to the previous period's deviation from long-run equilibrium $(r_{Lt-1} - \beta r_{St-1})$.

$$\boxed{\begin{aligned} \Delta r_{St} &= \alpha_S(r_{Lt-1} - \beta r_{St-1}) + \sum a_{11}(i)\Delta r_{St-i} + \sum a_{12}(i)\Delta r_{Lt-i} + \varepsilon_{St}, \\ \Delta r_{Lt} &= -\alpha_L(r_{Lt-1} - \beta r_{St-1}) + \sum a_{21}(i)\Delta r_{St-i} + \sum a_{22}(i)\Delta r_{Lt-i} + \varepsilon_{Lt}. \end{aligned}}$$

This is a bivariate VAR in first difference augmented by the error-correction trems $\alpha_S(r_{Lt-1} - \beta r_{St-1})$ and $-\alpha_L(r_{Lt-1} - \beta r_{St-1})$. α_S and α_L have the interpretation of speed of adjustment parameters. The larger α_S is, the greater the response of the previous period's deviation from long-run equilibrium $(r_{Lt-1} - \beta r_{St-1})$.

Cointegration and Error Correction

The vector x_t has an **error-correction representation** if it can be expressed in the form

$$\boxed{\Delta x_t = \pi_0 + \pi x_{t-1} + \pi_1 \Delta x_{t-1} + \pi_2 \Delta x_{t-2} + \cdots + \pi_p \Delta x_{t-p} + \varepsilon_t}$$

where $\pi_0 = (\pi_{i0})_{n\times 1}$, $\pi = (\pi_{jk})_{n\times n} \neq 0$, $\pi_i = (\pi_{jk}(i))_{n\times n}, i = 1, 2, \cdots, p$. The components in the error term vector $\varepsilon_t = (\varepsilon_{it})_{n\times 1}$ may be correlated with each other, but are stationary. Suppose that $x_t \sim I(1)$. Then

πx_{t-1} is stationary. Each row of π is a cointegrating vector of x_t.

If $\pi = 0$, the model is a VAR in first difference of x_t. There is no error correction term, implying that Δx_t does not respond to the previous period's deviation from long-run equilibrium (or disequilibrium error).

If $\pi \neq 0$, Δx_t responds to the previous period's deviation from long-run equilibrium and estimating x_t as a VAR in the first difference by omitting the error correction term πx_{t-1} is inappropriate.

VAR(1) and CI(1,1)

$$\begin{cases} y_t = a_{11}y_{t-1} + a_{12}z_{t-1} + \varepsilon_{yt} \\ z_t = a_{21}y_{t-1} + a_{22}z_{t-1} + \varepsilon_{zt} \end{cases} \qquad \begin{cases} \Delta y_t = (a_{11}-1)y_{t-1} + a_{12}z_{t-1} + \varepsilon_{yt} \\ \Delta z_t = a_{21}y_{t-1} + (a_{22}-1)z_{t-1} + \varepsilon_{zt}. \end{cases}$$

$$y_t = \frac{(1-a_{22}L)\varepsilon_{yt} + a_{12}L\varepsilon_{zt}}{(1-a_{11}L)(1-a_{22}L) - a_{12}a_{21}L^2} \qquad z_t = \frac{a_{21}L\varepsilon_{yt} + (1-a_{11}L)\varepsilon_{zt}}{(1-a_{11}L)(1-a_{22}L) - a_{12}a_{21}L^2}.$$

$$\lambda^2 - (a_{11}+a_{22})\lambda + (a_{11}a_{22} - a_{12}a_{21}) = 0. \quad \lambda_1 \text{ and } \lambda_2.$$

When are yt and zt CI(1,1)?

$$\pi = \begin{pmatrix} a_{11}-1 & a_{12} \\ a_{21} & a_{22}-1 \end{pmatrix}$$

Lemma:

1) If $a_{12} = a_{21} = 0$, then y_t and z_t cannot be $CI(1,1)$.

2) If $a_{11} = a_{22} = 1$, then y_t and z_t cannot be $CI(1,1)$.

3) If y_t and z_t are $CI(1,1)$, then $det(\pi) = 0$ and $rank(\pi) = 1$.

VAR(1): Conditions for CI(1,1)

$$\begin{cases} \Delta y_t = (a_{11} - 1)y_{t-1} + a_{12}z_{t-1} + \varepsilon_{yt} \\ \Delta z_t = a_{21}y_{t-1} + (a_{22} - 1)z_{t-1} + \varepsilon_{zt}. \end{cases}$$

Discussion for λ_1, λ_2 for $\{y_t\}$ and $\{z_t\}$ to be $CI(1,1)$:

1) If λ_1 and λ_2 lie inside the unit circle, $\{y_t\}$ and $\{z_t\}$ are stationary and cannot be integrated of order $(1,1)$.

2) If either of λ_1 and λ_2 lies outside the unit circle, the solution is explosive. Neither variable is difference stationary and so they cannot be $CI(1,1)$.

3) If both λ_1 and λ_2 are unity, $\{y_t\}$ and $\{z_t\}$ are I(2) and cannot be integrated of order $(1,1)$.

4) For $\{y_t\}$ and $\{z_t\}$ to be $CI(1,1)$, **it is necessary that** $\lambda_1 = 1$ and $|\lambda_2| < 1$.

VAR(1): Granger Representation for CI(1,1) 1/4

$$\begin{cases} \Delta y_t = (a_{11}-1)y_{t-1} + a_{12}z_{t-1} + \varepsilon_{yt} \\ \Delta z_t = a_{21}y_{t-1} + (a_{22}-1)z_{t-1} + \varepsilon_{zt}. \end{cases} \qquad a_{11} = 1 - \frac{a_{12}a_{21}}{1-a_{22}}.$$

$$\begin{cases} \Delta y_t = -\frac{a_{12}a_{21}}{1-a_{22}} y_{t-1} + a_{12}z_{t-1} + \varepsilon_{yt} = -\frac{a_{12}a_{21}}{1-a_{22}}\left(y_{t-1} - \frac{1-a_{22}}{a_{21}} z_{t-1}\right) + \varepsilon_{yt} \\ \Delta z_t = a_{21}y_{t-1} + (a_{22}-1)z_{t-1} + \varepsilon_{zt} = a_{21}\left(y_{t-1} - \frac{1-a_{22}}{a_{21}} z_{t-1}\right) + \varepsilon_{zt} \end{cases}$$

$$\begin{pmatrix} \Delta y_t \\ \Delta z_t \end{pmatrix} = \begin{pmatrix} \alpha_y \\ \alpha_z \end{pmatrix} (y_{t-1} - \beta_1 z_{t-1}) + \begin{pmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{pmatrix}$$

$$= \begin{pmatrix} \alpha_y \\ \alpha_z \end{pmatrix} (1, -\beta_1) \begin{pmatrix} y_{t-1} \\ z_{t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{pmatrix} = \alpha\beta' x_{t-1} + \varepsilon_t$$

where $\alpha - \left(-\frac{a_{12}a_{21}}{1-a_{22}}, a_{21}\right)'$, $\beta - (1, -\frac{1-a_{22}}{a_{21}})'$ with $rank(\beta) - 1$.

VAR(1): Granger Representation for CI(1,1) 2/4

1) Here $\beta_1 = \frac{1-a_{22}}{a_{21}} \neq 0$. $\{y_{t-1} - \beta_1 z_{t-1}\}$ is stationary and $y_t - \beta_1 z_t = 0$ is the long-run equilibrium. The normalized cointegrating vector is $\beta = (1, -\beta_1)'$. We can see that y_t and z_t change in response to the previous period's deviation $(y_{t-1} - \beta_1 z_{t-1})$ from the long-run equilibrium $y_t = \beta_1 z_t$. Also, at least one of adjustment speed parameters $\alpha_y = -\frac{a_{12}a_{21}}{1-a_{22}}$ and $\alpha_z = a_{21}$ is not equal to zero, meaning that the adjustment in the system plays a role in the response to the deviation. In the long-run equilibrium (there is no deviation), y_t and z_t change only in response to ε_{yt} and ε_{zt} shocks.

2) Granger representation theorem says that error correction and cointegration are two equivalent ways in representing $CI(1,1)$. $CI(1,1)$ guarantees the existence of an error-correction model and an error-correction model for $I(1)$ variables implies cointegration.

3) Here, a cointegrated system can be viewed as a restricted form of a general VAR model: $x_t = \alpha\beta' x_{t-1} + \varepsilon_t$ with $rank(\beta) = 1$. It is inappropriate to estimate a VAR of cointegrated variables using only first differences but ignoring the error-correction portion of the model.

VAR(1): Granger Representation for CI(1,1) 3/4

4) If $rank(\pi) = 0$, we have $a_{12} = a_{21} = 0, a_{11} = a_{22} = 1$. Then $\Delta x_t = \varepsilon_t$ or x_t is not $CI(1,1)$. If the variables are cointegrated, the rows of π must be linearly dependent, and hence $\det(\pi) = 0$ or the rank of π is 1.

5) In general both variables in a cointegration system will response to a deviation from the long run equilibrium. Possibly, one of the two djustment speed parameters α_y and α_z (not both) may be equal to zero. In our stated case $a_{12} = 0$ such that $\alpha_y = -\frac{a_{12}a_{21}}{1-a_{22}} = 0$. That is, the change of $\{y_t\}$ does not respond to the discrepancy from long-run equilibrium and $\{z_t\}$ does all of the adjustment:

$$\begin{cases} \Delta y_t = 0 \cdot (y_{t-1} - \beta_1 z_{t-1}) + \sum c_{11}(i)\Delta y_{t-i} + \sum c_{12}(i)\Delta z_{t-i} + \varepsilon_{yt} \\ \Delta z_t = \alpha_z \cdot (y_{t-1} - \beta_1 z_{t-1}) + \sum c_{21}(i)\Delta y_{t-i} + \sum c_{22}(i)\Delta z_{t-i} + \varepsilon_{zt} \end{cases}$$

This means that $\{y_t\}$ is weakly exogenous. Estimation and testing can be conducted only for $\{z_t\}$ without reference to the model for $\{y_t\}$.

VAR(1): Granger Representation for CI(1,1) 4/4

6) Granger causality in a cointegrated system is reinterpreted: $\{y_t\}$ **does not Granger cause** $\{z_t\}$ if lagged values Δy_{t-i} do not enter the Δz_t equation and if z_t does not respond to the deviation from the long-run equilibrium, that is, if $\alpha_z = 0$ and all $c_{21}(i) = 0$ in

$$\begin{cases} \Delta y_t = \alpha_y \cdot (y_{t-1} - \beta_1 z_{t-1}) + \sum c_{11}(i)\Delta y_{t-i} + \sum c_{12}(i)\Delta z_{t-i} + \varepsilon_{yt} \\ \Delta z_t = \alpha_z \cdot (y_{t-1} - \beta_1 z_{t-1}) + \sum c_{21}(i)\Delta y_{t-i} + \sum c_{22}(i)\Delta z_{t-i} + \varepsilon_{zt}. \end{cases}$$

VAR(n): Granger Representation for CI(1,1) 1/2

$$x_t = A_1 x_{t-1} + \varepsilon_t$$

where $x_t = (x_{1t}, x_{2t}, \cdots, x_{nt})'$, $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t}, \cdots, \varepsilon_{nt})', i.i.d.(0, \Omega)$,

$$\begin{aligned}\Delta x_t &= (A_1 - I)x_{t-1} + \varepsilon_t \\ &= \pi x_{t-1} + \varepsilon_t.\end{aligned}$$

The rank of π determine the number of cointegration vectors. For example,

1) $rank(\pi) = 0 : \pi = 0$ and $\Delta x_t = \varepsilon_t$. All the sequences $\{x_{it}\}$ are unit root processes and there is no linear combination of the variables that is stationary. Hence $x_t \nsim CI(1,1)$.

2) $rank(\pi) = n : det(\pi) \neq 0$ (there is no unit root) and each row of $\pi x_{t-1} = 0$ is an independent restriction on the long-run solution of the variables. Each of the n variables in x_t must be stationary with the corresponding long-run value constrain in $\pi x_{t-1} = 0$. Hence $x_t \nsim CI(1,1)$.

VAR(n): Granger Representation for CI(1,1) 2/2

3) $rank(\pi) = 1$: there is a single cointegrating vector given by any row of π, e.g. for

$$\Delta x_{1t} = \pi_{11}(x_{1t-1} + \beta_{12}x_{2t-1} + \cdots + \beta_{1n}x_{nt-1}) + \varepsilon_{1t},$$

the normalized cointegrating vector is $(1, \beta_{12}, \cdots, \beta_{1n})'$, where $\beta_{ij} = \pi_{ij}/\pi_{11}$, and the speed of adjustment parameter is π_{11}.

4) $rank(\pi) = 2$: Assume, for example, the first two lines of π are linearly independent. The two cointegration vectors are $(\pi_{11}, \pi_{12}, \cdots, \pi_{1n})'$ and $(\pi_{21}, \pi_{22}, \cdots, \pi_{2n})'$ without normaliztion. In the long-run equilibrium, the cointegrated variables $\{x_{1t}\}$, $\{x_{2t}\}, \cdots, \{x_{nt}\}$ satisfy the two relationships:

$$\begin{aligned} \pi_{11}x_{1t} + \pi_{12}x_{2t} + \cdots + \pi_{1n}x_{nt} &= 0 \\ \pi_{21}x_{1t} + \pi_{22}x_{2t} + \cdots + \pi_{2n}x_{nt} &= 0. \end{aligned}$$

5) $0 < rank(\pi) = r < n$: there are r cointegrating vectors and $n - r$ stochastic trends in the system.

Test for Cointegration: Engle-Granger 1/5

Consider the linear regression model

$$y_t = z_t'\beta + e_t$$

where the $k \times 1$ vector $z_t \sim I(1)$ and elements of z_t are not cointegrated. Further, assume

$$e_t = a_1 e_{t-1} + v_t$$

where $v_t \sim I(0)$. Testing the null

$$H_0 : a_1 = 1$$

amounts to testing the null of non-cointegration of y_t and z_t.

The Engle-Granger method has the following shortcomings:

in practice, it is possible to find that one regression indicates that the variables are cointegrated, whereas reversing the order indicates no cointegration.

e_t is carried into the estimation of a_1 and the no-cointegration test.

Test for Cointegration: Engle-Granger 2/5

Four steps to test for the cointegration of y_t and z_t :

1) pretest each variable to determine its integration order (Augmented Dickey-Fuller test infers the number of unit roots). If the two are stationary, it is not necessary to proceed. If they are integrated of different orders, they are not cointegrated.

2) Estimate the long-run equilibrium relationship. If the above test indicates that both y_t and z_t are $I(1)$, estimate the long-run equilibrium relationship:

$$y_t = \beta_0 + \beta_1 z_t + e_t.$$

Conduct unit root test for the AR(1) model of the above residuals $\hat{e}_t$:

$$\Delta \hat{e}_t = a_1 \hat{e}_{t-1} + \varepsilon_t, \quad H_0 : a_1 = 0$$

$$\Delta \hat{e}_t = a_1 \hat{e}_{t-1} + \sum_{i=1}^{p} a_{i+1} \Delta \hat{e}_{t-i} + \varepsilon_t, \quad H_0 : a_1 = 0 \text{ (no cointegration)}$$

Here the critical values for the test are provided by Engle-Granger in **Table C** at the end of the text

Test for Cointegration: Engle-Granger 3/5

3) Estimate the error correction model. Use the saved residuals from the equilibrium regression (4) as the deviation from the long-run equilibrium and estimate the following error correction model

$$\begin{aligned}
\Delta y_t &= \alpha_1 + \alpha_y \hat{e}_{t-1} + \sum_{i=1} \alpha_{11}(i) \Delta y_{t-i} + \sum_{i=1} \alpha_{12}(i) \Delta z_{t-i} + \varepsilon_{yt} \\
\Delta z_t &= \alpha_2 + \alpha_z \hat{e}_{t-1} + \sum_{i=1} \alpha_{21}(i) \Delta y_{t-i} + \sum_{i=1} \alpha_{22}(i) \Delta z_{t-i} + \varepsilon_{zt}
\end{aligned}$$

which constitutes VAR in first difference other than the error-correction term $\hat{e}_{t-1}$. This can be efficiently estimated since each equation contains the same set of regressors. Since all terms are stationary, t-test and F-test are appropriate.

Test for Cointegration: Engle-Granger 4/5

4) Assess model adequacy.

(i) Estimate the error correction model by adjusting the lag lengths such that the residuals of the error-correction equations approximate white noise. If you need to allow longer lags of some variables than of others, efficiency can be gained by estimating the near-VAR using the SUR method.

(ii) The speed of adjustment coefficients α_y and α_z are of particular interest in that they have important implications for the dynamics of the system. Some cases are interesting: if $\alpha_z = 0$ the change in z_t does not at all respond to the deviation from the long-run equilibrium in $(t-1)$; if $\alpha_z = 0$ and all $\alpha_{21}(i)$ are equal to zero, $\{\Delta y_t\}$ does not Granger cause $\{\Delta z_t\}$.The estimates of α_y and/or α_z should be significantly different from zero if the variables are cointegrated. Moreover, the absolute values of α_y and α_z must not be too large.

Test for Cointegration: Engle-Granger 5/5

(iii) Impulse responses and variance decomposition analysis can be used to obtain information concerning the interactions among the variables (some method, such as Choleski Decomposition, must be used to orthogonalize the innovations), which can indicate whether the dynamic responses of the variables conform to theory. The impulse responses of $\{\Delta y_t\}$ and $\{\Delta z_t\}$ should convergen to zero. You should reexamine your results from each step if you obtain a nondecaying or explosive impulse response function.

(iv) Generally, it is inappropriate to use t-statistic to perform significance tests on the cointegrating vector: $y_t = z_t'\beta + e_t$ since the coefficients are super-consistent but the standard errors are not. See Textbook: Appendix 6.1 on P378-380. The exception is the case that the residuals from both equations are serially uncorrelated and the cross-correlations are zero, i.e. the cointegration relationship between $\{y_t\}$ and $\{z_t\}$ is such that

$$\begin{aligned} y_t &= \beta_0 + \beta_1 z_t + \varepsilon_{1t}, \qquad && cov(\varepsilon_{1t}, \varepsilon_{1s}) = 0, cov(\varepsilon_{2t}, \varepsilon_{2s}) = 0, t \neq s \\ \Delta z_t &= \varepsilon_{2t} && cov(\varepsilon_{1t}, \varepsilon_{2s}) = 0, \forall t, s \end{aligned}$$

Engle-Granger Test: Example

Example: see Ch6-ex3. Test the cointegration of y_t, z_t and w_t, where the data are generated from

$$\begin{aligned} y_t &= \mu_{yt} + \delta_{yt}, \quad \mu_{yt} = \mu_{y,t-1} + \varepsilon_{yt}, \;\; \delta_{yt} = 0.5\delta_{y,t-1} + \eta_{yt} \\ z_t &= \mu_{zt} + \delta_{zt} + 0.5\delta_{yt}, \quad \mu_{zt} = \mu_{t,t-1} + \varepsilon_{zt}, \;\; \delta_{zt} = 0.5\delta_{z,t-1} + \eta_{zt} \\ w_t &= \mu_{wt} + \delta_{wt} + 0.5\delta_{yt} + +0.5\delta_{zt}, \quad \mu_{wt} = \mu_{yt} + \mu_{zt}, \;\; \delta_{wt} = 0.5\delta_{w,t-1} + \eta_{wt} \end{aligned}$$

where $\varepsilon_{yt}, \eta_{yt}, \varepsilon_{zt}, \eta_{zt}$ and η_{wt} are all white noise processes. The true relationship is that $(y_t, z_t, w_t)'$ is cointegrated with a coinegrating vector $(1, 1, -1)'$.

Test with I(2): Multicointegration 1/2

Multicointegration:

A linear combination of I(2) and I(1) variables is integrated of order zero.

Suppose that x_{1t} and x_{2t} are $I(2)$ and that z_t is I(1). Three cases:

(i) x_{1t} and x_{2t} are CI(2, 1). $x_{1t} - \beta_2 x_{2t} - \gamma_1 z_t$ is $I(0)$.

(ii) x_{1t} and x_{2t} are CI(2, 1). $x_{1t} - \beta_2 x_{2t} - \alpha_1 \Delta x_{2t} - \gamma_1 z_t$ is $I(0)$.

(iii) x_{1t} and x_{2t} are CI(2, 2). $x_{1t} - \beta_2 x_{2t} - 0 \cdot z_t$ is $I(0)$.

In principle, check for multicointegration using a two-step procedure: First find a cointegrating relationship among I(2) variables and then use this relationship to check for a possible cointegrating relationship with the remaining I(1) variables.

Test with I(2): Multicointegration 2/2

Engsted, Gonzalo znd Haldrup (1997) show that this procedure is effective only if the cointegrating vector for the first step is known. Otherwise, the second step is contaminated with the errors generated in the first step.

$$\boxed{x_{1t} = a_0 + a_1 t + a_2 t^2 + \beta_2 x_{2t} + \beta_3 x_{3t} + \gamma_1 \Delta x_{2t} + \gamma_2 \Delta x_{3t} + \alpha_1 z_t + e_t}$$

where x_{1t}, x_{2t} and x_{3t} are $I(2)$,and z_t is I(1).

$$\boxed{\Delta \hat{e}_t - \rho \Delta \hat{e}_{t-1} + \sum_{i=1}^{p} \rho_i \Delta \hat{e}_{t-i} + v_i}$$

test the null $\rho = 0$

Rejecting → multicointegration

critical values of this t-statistic for the null $\rho = 0$ depend on the sample size, the number of I(2) regressors ($m_2 = 1$ or 2), the number of I(1) regressors ($m_1 = 0$ to 4), and the form of the deterministic regressors. See **Table D** on P442 of the Textbook.

Forms of Models for Cointegration 1/4

Johansen test: Generalize DF test to the vector.

Some forms of the models for the cointegration test:

1) The model without a drift is

$$Model\ 1: x_t = A_1 x_{t-1} + \varepsilon_t \ \text{ or } \ \Delta x_t = \pi x_{t-1} + \varepsilon_t,$$

where $\pi = A_1 - I$. The rank of π equals the number of cointegrating vectors

2) Add a drift term, if the variables exhibit a decided tendency to increase or decrease:

$$Model\ 2: x_t = A_0 + A_1 x_{t-1} + \varepsilon_t \ \text{ or } \Delta x_t = A_0 + \pi x_{t-1} + \varepsilon_t.$$

Forms of Models for Cointegration 2/4

3) Include a constant in the cointegrating relationship

Model 3 : e.g. assume that $rank(\pi) = 1$ and $a_{i0} = s_i a_{10}$

$$\Delta x_{1t} = (\pi_{11} x_{1t-1} + \pi_{12} x_{2t-1} + \cdots + \pi_{1n} x_{nt-1} + a_{10}) + \varepsilon_{1t}$$

$$\Delta x_{2t} = s_2 (\pi_{11} x_{1t-1} + \pi_{12} x_{2t-1} + \cdots + \pi_{1n} x_{nt-1} + a_{10}) + \varepsilon_{2t}$$

$$\cdots$$

$$\Delta x_{1t} = s_n (\pi_{11} x_{1t-1} + \pi_{12} x_{2t-1} + \cdots + \pi_{1n} x_{nt-1} + a_{10}) + \varepsilon_{nt}$$

$\boxed{\Delta x_t = \pi^* x^*_{t-1} + \varepsilon_t,}$ where $x^*_{t-1} = (x_{1t-1}, x_{2t-1}, \cdots, x_{nt-1}, 1)'$ and $\pi^* = (\pi, A_0)$.

Generally, for $1 \le r < n$, $\pi = \alpha\beta'$,where α and β are two $n \times r$ matrix with $rank(\beta) = r$.

$$\begin{aligned} \Delta x_t &= A_0 + \pi x_{t-1} + \varepsilon_t = \alpha\left(\beta'_0 + \beta' x_{t-1}\right) + \varepsilon_t \\ &= \alpha(\beta', \beta'_0) \begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix} + \varepsilon_t = \pi^* x^*_{t-1} + \varepsilon_t \end{aligned}$$

The linear trend is purged from the system.

Forms of Models for Cointegration 3/4

4) Include an intercept term in the cointegrating vector along with a drift term: e.g. assume that $rank(\pi) = 1$ and $s_i b_{10} + b_{i1} = a_{i0}$,then

$$\begin{aligned} Model\ 4 \quad : \quad & \Delta x_{1t} = (\pi_{11} x_{1t-1} + \pi_{12} x_{2t-1} + \cdots + \pi_{1n} x_{nt-1} + b_{10}) + b_{11} + \varepsilon_{1t} \\ & \Delta x_{2t} = s_2 (\pi_{11} x_{1t-1} + \pi_{12} x_{2t-1} + \cdots + \pi_{1n} x_{nt-1} + b_{10}) + b_{21} + \varepsilon_{2t} \\ & \cdots \\ & \Delta x_{1t} = s_n (\pi_{11} x_{1t-1} + \pi_{12} x_{2t-1} + \cdots + \pi_{1n} x_{nt-1} + b_{10}) + b_{n1} + \varepsilon_{nt} \end{aligned}$$

Generally, for $1 \leq r < n$, $\pi = \alpha\beta'$,where α and β are two $n \times r$ matrix with $rank(\beta) = r$. **If A_0 has a decomposition:** $A_0 = (A_0 - \alpha\beta_0') + \alpha\beta_0'$, where β_0 is a $1 \times r$ matrix, then

$$\begin{aligned} \Delta x_t &= A_0 + \pi x_{t-1} + \varepsilon_t = (A_0 - \alpha\beta_0') + \alpha (\beta_0' + \beta' x_{t-1}) + \varepsilon_t \\ &= A_0^* + \pi^* x_{t-1}^* + \varepsilon_t \end{aligned}$$

where $A_0^* = A_0 - \alpha\beta_0'$, $\pi^* = \alpha(\beta', \beta_0')$,and $x_{t-1}^* = \begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix}$

Forms of Models for Cointegration 4/4

5) Higher-order AR process:

$$Model\ 5: x_t = A_1 x_{t-1} + A_2 x_{t-2} + \cdots + A_p x_{t-p} + e_t$$

or
$$\Delta x_t = \pi x_{t-1} + \sum_{i=1}^{p-1} \pi_i \Delta x_{t-i} + e_t$$

where $\pi = \sum_{i=1}^{p} A_i - I$, $\pi_i = -\sum_{j=i+1}^{p} A_j$

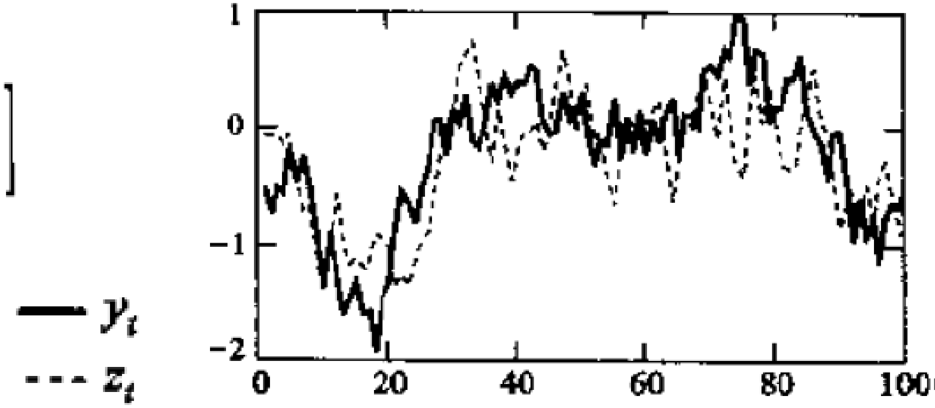
if $1 < r \equiv rank(\pi) < n$, π can be decomposed as $\boxed{\pi = \alpha\beta'}$

where α and β are two $n \times r$ matrix with $rank(\beta) = r$.

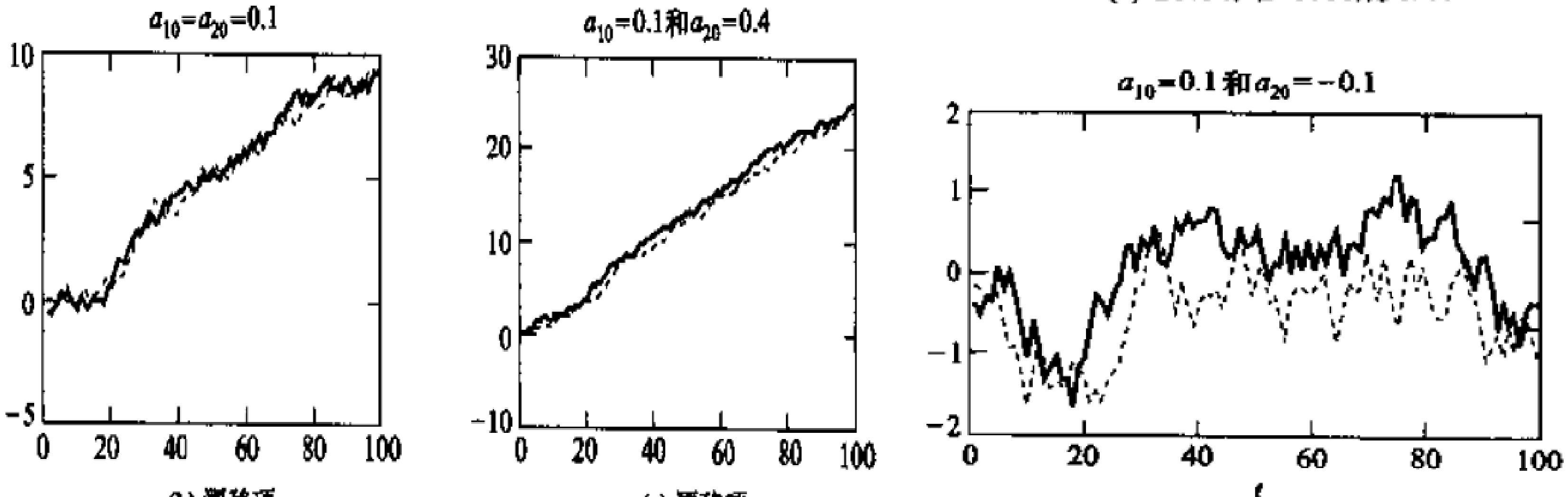
$\beta' x_t = 0$ is the long run equilibrium.

An Example on Drift and Intercept

$$\begin{bmatrix}\Delta y_t\\ \Delta z_t\end{bmatrix} = \begin{bmatrix}-0.2 & 0.2\\ 0.2 & -0.2\end{bmatrix}\begin{bmatrix}y_{t-1}\\ z_{t-1}\end{bmatrix} + \begin{bmatrix}\varepsilon_{yt}\\ \varepsilon_{zt}\end{bmatrix}$$



(a) 没有漂移项或截距项



(b) 漂移项　(c) 漂移项　(d) 在协整中的常量

Concentration of Likelihood (No drift) 1/4

Likelihood ratio test of cointegration rank for Model 5 (No drift):

$$\Delta x_t = \pi x_{t-1} + \sum_{i=1}^{p-1} \pi_i \Delta x_{t-i} + e_t \qquad \text{where } e_t \sim iidN(0, \Sigma)$$

$$t = 1, 2, \cdots, T, \qquad X_{1-p}, \cdots, X_0 \text{ are given constant vectors.}$$

Likelihood function:

$$\ln L(x_1, x_2, \cdots, x_T; \pi_1, \pi_2, \cdots, \pi_{p-1}, \pi, \Sigma) = \frac{-nT}{2} \ln(2\pi) - \frac{T}{2} \ln|\Sigma| - \frac{1}{2} \sum_{t=1}^{T} e_t' \Sigma^{-1} e_t.$$

(i) Concentrate $\ln L$ with respect to Σ : by $\hat{\Sigma} = \frac{1}{T} \sum_{t=1}^{T} e_t e_t'$,

$$\ln L^*(x_1, x_2, \cdots, x_T; \pi_1, \pi_2, \cdots, \pi_{p-1}, \pi) = C - \frac{T}{2} \ln \left| \sum_{t=1}^{T} e_t e_t' \right|.$$

Concentration of Likelihood (No drift) 2/4

(ii) Concentrate $\ln L^*$ with respect to $\pi_1, \pi_2, \cdots, \pi_{p-1}$.

Let $q_t = (\Delta x_{t-1}, \Delta x_{t-2}, \cdots, \Delta x_{t-p+1})'$.

$$R_{0t} = \Delta x_t - \sum_{i=1}^{p-1} \hat{\pi}_i \Delta x_{t-i} \qquad R_{1t} = x_{t-1} - \sum_{i=1}^{p-1} \tilde{\pi}_i \Delta x_{t-i}$$

$$(\hat{\pi}_1, \hat{\pi}_2, \cdots, \hat{\pi}_{p-1}) = \left(\sum_{t=1}^{T} q_t q_t'\right)^{-1} \left(\sum_{t=1}^{T} \Delta x_t q_t'\right)$$

$$(\tilde{\pi}_1, \tilde{\pi}_2, \cdots, \tilde{\pi}_{p-1}) = \left(\sum_{t=1}^{T} q_t q_t'\right)^{-1} \left(\sum_{t=1}^{T} x_{t-1} q_t'\right).$$

$$\begin{aligned} \ln L^{**}(x_1, x_2, \cdots, x_T; \pi) &= C_0 - \frac{T}{2} \ln \left| \sum_{t=1}^{T} (R_{0t} - \pi R_{1t})(R_{0t} - \pi R_{1t})' \right| \\ &= C_{00} - \frac{T}{2} \ln |s_{00} - \pi s_{10} - s_{01}\pi' + \pi s_{11}\pi'| \end{aligned}$$

Concentration of Likelihood (No drift) 2/4

Alternatively,

By Anderson(1958)'s canonical correlation analysis, if let $\hat{r}_i$ be the **canonical correlation** of the residual series R_{0t} and R_{1t}, $-1 < \hat{r}_i < 1$, we have

$$\min_{\pi} \left| \frac{1}{T} \sum_{t=1}^{T} (R_{0t} - \pi R_{1t}) (R_{0t} - \pi R_{1t})' \right| = \prod_{i=1}^{n} (1 - \lambda_i)$$

with $\lambda_i \equiv \hat{r}_i^2, 0 \leq \lambda_i \leq 1$, So when π is not restricted, the minimum value of $\ln L^{**}(x_1, x_2, \cdots, x_T; \pi)$ is

$$\ln L^{**}(x_1, x_2, \cdots, x_T; \hat{\pi}) = C_0 - \frac{T}{2} \sum_{i=1}^{n} \ln(1 - \lambda_i).$$

Concentration of Likelihood (No drift) 3/4

(iii) Imposing the restriction $\pi = \alpha\beta'$,

$$\ln L^{**}(x_1, x_2, \cdots, x_T; \alpha, \beta) = C_{00} - \frac{T}{2} \ln |s_{00} - \alpha\beta' s_{10} - s_{01}\beta\alpha' + \alpha\beta' s_{11}\beta\alpha'| .$$

FOC:

$$\frac{\partial \ln L^{**}(x_1, x_2, \cdots, x_T; \alpha, \beta)}{\partial \alpha} = 0 \Longrightarrow \hat{\alpha} = s_{01}\beta\left(\beta' s_{11}\beta\right)^{-1},$$

$$\ln L^{***}(x_1, x_2, \cdots, x_T; \beta) = C_1 - \frac{T}{2} \ln |s_{00}| - \frac{T}{2} \ln \left|\beta' \left(s_{11} - s_{10} s_{00}^{-1} s_{01}\right) \beta\right| .$$

$$\boxed{\begin{array}{l} \text{minimize } \left|\beta' \left(s_{11} - s_{10} s_{00}^{-1} s_{01}\right) \beta\right| \\ \text{subject to } \beta' s_{11} \beta = I. \end{array}} \qquad (8)$$

Concentration of Likelihood (No drift) 3/4

$$\boxed{\begin{array}{l} \text{minimize } \left|\beta' \left(s_{11} - s_{10}s_{00}^{-1}s_{01}\right)\beta\right| \\ \text{subject to } \beta' s_{11}\beta = I. \end{array}} \tag{8}$$

Anderson (1958) shows that the above problem is equivalent to solving β in the equation

$$(\lambda s_{11} - s_{10}s_{00}^{-1}s_{01})\beta = 0, \tag{9}$$

where λ is the root of the characteristic equation $\left|\lambda s_{11} - s_{10}s_{00}^{-1}s_{01}\right| = 0.$ (10)

It has n roots equal to $\lambda_i = \hat{r}_i^2$. Order them in $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$. Correspondingly, the characteristic vectors are $V_1, V_2, \cdots, V_n$. Choose $\hat{\beta} = (V_1, V_2, \cdots, V_r)$, where $r = rank(\beta)$. Then $\hat{\beta}$ is the solution to (8). And the minimum value of $\ln L^{***}(x_1, x_2, \cdots, x_T; \beta)$ is

$$\ln L^{***}(x_1, x_2, \cdots, x_T; \hat{\beta}) = C_0 - \frac{T}{2}\sum_{i=1}^{r}\ln(1 - \lambda_i). \tag{11}$$

Concentration of Likelihood (No drift) 4/4

$$|\lambda s_{11} - s_{10}s_{00}^{-1}s_{01}| = 0. \qquad (10)$$

(iv) Suppose $rank(\pi) = r$. Order the characteristic roots of (10) such that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0$ and $\lambda_{r+1} = \cdots = \lambda_n = 0$. If the variables in x_t are not cointegrated, $rank(\pi) = 0$, all $\lambda_i = 0$ and $\ln(1-\lambda_i) = 0$. If $rank(\pi) = 1, 0 < \lambda_1 < 1$, then $\ln(1-\lambda_1) < 0$ and $\ln(1-\lambda_i) = 0$ for $i = 2, 3, \cdots, n$. It is deduced from (11) that the maximum value of the likelihood function is given by

$$\begin{aligned}\ln L^{****}(x_1, x_2, \cdots, x_T; r) &= C_2 - \frac{T}{2}\ln\left|\hat{\beta}'\left(s_{11} - s_{10}s_{00}^{-1}s_{01}\right)\hat{\beta}\right| \\ &= K - \frac{T}{2}\sum_{i=1}^{r}\ln(1-\lambda_i),\end{aligned}$$

where K is a constant related with n and T.

Johansen's Method: Calculate the characteristic roots and vectors

Take $x_t = A_1 x_{t-1} + A_2 x_{t-2} + \cdots + A_p x_{t-p} + e_t$ as an example. Write

$$\Delta x_t = \pi x_{t-1} + \sum_{i=1}^{p-1} \pi_i \Delta x_{t-i} + e_t$$

where $\pi = \sum_{i=1}^{p} A_i - I$, $\pi_i = -\sum_{j=i+1}^{p} A_j$, and $rank(\pi)$ = the number of independent cointegrating vector.

Step 1: Estimate the VAR in first difference: $\Delta x_t = \sum_{i=1}^{p-1} B_i \Delta x_{t-i} + e_{0t}$

Step 2: Estimate $x_{t-1} = \sum_{i=1}^{p-1} C_i \Delta x_{t-i} + e_{1t}$

Step 3: Compute the squared canonical correlations between e_{0t} and e_{1t}, λ_i. That is, calculate λ_i as the solution to $\left|\lambda s_{11} - s_{10} s_{00}^{-1} s_{01}\right| = 0$, where $s_{ij} = \frac{1}{T}\sum_{t=1}^{T} e_{it} e'_{jt}$, $i, j = 0, 1$.

Step 4: The maximum likelihood estimates of the cointegrating vectors are the non-trivial solutions to $\lambda_i s_{11} \beta_i = s_{10} s_{00}^{-1} s_{01} \beta_i$.

Test for Cointegration: Johansen (**No drift**) 1/2

1) The first statistic tests the null hypothesis that the number of distinct cointegrating vectors is less than or equal to r against a general alternative. That is, to test

$$\boxed{H_0 : rank(\pi) \leq r \quad (that\ is,\ \lambda_{r+1} = \cdots = \lambda_n = 0)}$$

against $H_1 : rank(\pi) > r$ (i.e. at least one of $\lambda_{r+1}, \cdots, \lambda_n$ is not equal to zero).

$$L_r = K - \frac{T}{2}\sum_{i=1}^{r} \ln(1-\lambda_i). \qquad L_u = K - \frac{T}{2}\sum_{i=1}^{n} \ln(1-\lambda_i).$$

the trace-statistic $\lambda_{trace}(r)$: $\boxed{\lambda_{trace}(r) = -T \sum_{i=r+1}^{n} \ln(1-\hat{\lambda}_i).}$

When all $\lambda_i = 0$, $\lambda_{trace}(r) = 0$; The further the estimated characteristic roots are from zero, the more neagtive is $\ln(1-\hat{\lambda}_i)$, and the larger is the trace-statistic $\lambda_{trace}(r)$. Critical values of $\lambda_{trace}(r)$ are obtained by simulation, see Table E in Enders's textbook.

Test for Cointegration: Johansen **(No drift)** 2/2

2) The second statistic tests the null hypothesis that the number of distinct cointegrating vectors is r against the alternative of $r+1$ cointegrating vectors. That is, to test $\boxed{H_0 : rank(\pi) = r}$ $\boxed{\lambda_{r+1} = \cdots = \lambda_n = 0}$

against the alternative hypothesis $H_1 : rank(\pi) = r+1.$ $\lambda_{r+2} = \cdots = \lambda_n = 0.$

$$L_r = K - \frac{T}{2}\sum_{i=1}^{r} \ln(1-\lambda_i) \qquad L_u = K - \frac{T}{2}\sum_{i=1}^{r+1} \ln(1-\lambda_i).$$

the maximum-eigenvalue statistic $\boxed{\lambda_{\max}(r, r+1) = -T\ln(1-\hat{\lambda}_{r+1})}$

When the estimated $\hat{\lambda}_{r+1}$ is close to zero, $\lambda_{\max}(r, r+1)$ will be small. The further the estimated characteristic root are from zero, the more neagtive is $\ln(1-\hat{\lambda}_{r+1})$, and the larger is $\lambda_{\max}(r, r+1)$. Critical values of the test are in Table E (the case without any deterministic regressors: the first cell) at the end of the text.

表 E λ_{max} 和 λ_{trace} 统计量的经验分布

显著性水平	10%	5%	2.5%	1%	10%	5%	2.5%	1%
无任何确定性回归变量的 λ_{max} 和 λ_{trace} 统计量								
	λ_{max}				λ_{trace}			
1	2.86	3.84	4.93	6.51	2.86	3.84	4.93	6.51
2	9.52	11.44	13.27	15.69	10.47	12.53	14.43	16.31
3	15.59	17.89	20.02	22.99	21.63	24.31	26.64	29.75
4	21.56	23.80	26.14	28.82	36.58	39.89	42.30	45.58
5	27.62	30.04	32.51	35.17	54.44	59.46	62.91	66.52
带漂移的 λ_{max} 和 λ_{trace} 统计量								
	λ_{max}				λ_{trace}			
1	2.69	3.76	4.95	6.65	2.69	3.76	4.95	6.65
2	12.07	14.07	16.05	18.63	13.33	15.41	17.52	20.04
3	18.60	20.97	23.09	25.52	26.79	29.68	32.56	35.65
4	24.73	27.07	28.98	32.24	43.95	47.21	50.35	54.46
5	30.90	33.46	35.71	38.77	64.84	68.52	71.80	76.07
协整向量中带常数项的 λ_{max} 和 λ_{trace} 统计量								
	λ_{max}				λ_{trace}			
1	7.52	9.24	10.08	12.97	7.52	9.24	10.80	12.95
2	13.75	15.67	17.63	20.20	17.85	19.96	22.05	24.60
3	19.77	22.00	24.07	26.81	32.00	34.91	37.61	41.07
4	25.56	28.14	30.32	33.24	49.65	53.12	56.06	60.16
5	31.66	34.40	36.90	39.79	71.86	76.07	80.06	84.45

Test for Cointegration: Johansen **(with drift)**

$$\Delta x_t = \mu + \pi x_{t-1} + \sum_{i=1}^{p-1} \pi_i \Delta x_{t-i} + e_t \quad \text{where } \mu \text{ is unrestricted.}$$

Replacing R_{0t} and R_{1t}, respectively, by

$$\begin{aligned} R_{0t} &= \Delta x_t - \sum_{i=1}^{p-1} \hat{\pi}_i \Delta x_{t-i} - \hat{\mu} \\ R_{1t} &= x_{t-1} - \sum_{i=1}^{p-1} \tilde{\pi}_i \Delta x_{t-i} - \tilde{\mu}, \end{aligned}$$

we proceed in the same manner. The limiting distributions of the trace and maximum-eigenvalue tests change. Their critical values are listed in Table E (the case with drift: the second cell) at the end of the text according to the form of the vector A_0.

Test for Cointegration: Johansen

(with a constant in Cointegrating Vector)

$$\Delta x_t = \mu + \pi x_{t-1} + \sum_{i=1}^{p-1} \pi_i \Delta x_{t-i} + e_t$$

with $\pi = \alpha\beta'$,where α and β are two $n \times r$ matrix with $rank(\beta) = r$. The intercept is restricted: $\mu = \alpha\beta_0'$ for any arbitrary $1 \times r$ vector β_0. Then

$$\begin{aligned} \Delta x_t &= \alpha\left(\beta_0' + \beta' x_{t-1}\right) + \sum_{i=1}^{p-1} \pi_i \Delta x_{t-i} + e_t \\ &= \alpha\left(\beta_0', \beta'\right)\begin{pmatrix} 1 \\ x_{t-1} \end{pmatrix} + \sum_{i=1}^{p-1} \pi_i \Delta x_{t-i} + e_t \equiv \pi^* x_{t-1}^* + \sum_{i=1}^{p-1} \pi_i \Delta x_{t-i} + e_t. \end{aligned}$$

replace R_{1t} by $R_{1t}^* = x_{t-1}^* - \sum_{i=1}^{p-1} \tilde{\pi}_i \Delta x_{t-i}$.

Critical values for these tests should change, and are listed in Table E (the case with a constant in the cointegrating vector: the third cell) at the end of the text.

Hypothesis Testing in Cointegrating Vector 1/2

Hypothesis testing about some restricted forms of cointegrating vectors:

1) Test H_0 : including an intercept in the cointegrating vector as opposed to the unrestricted drift A_0 (i.e. a linear time trend). Estimate the two forms of the model and obtain the ordered characteristic roots:

$$\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \cdots \geq \hat{\lambda}_n \qquad\qquad \hat{\lambda}_1^* \geq \hat{\lambda}_2^* \geq \cdots \geq \hat{\lambda}_n^*$$

with unrestricted A_0 with restricted $A_0 = \alpha\beta_0'$

$$2\left[\left(-\frac{T}{2}\sum_{i=1}^{r}\ln(1-\hat{\lambda}_i)\right)-\left(-\frac{T}{2}\sum_{i=1}^{r}\ln(1-\hat{\lambda}_i^*)\right)\right]= T\sum_{i=1}^{r}\ln\left(\frac{1-\hat{\lambda}_i^*}{1-\hat{\lambda}_i}\right) \sim \chi^2(n-r)$$

Large values of the test statistic imply that it is possible to reject the null of including an intercept in the cointegrating vectors, and that there is a linear trend in the variables.

Hypothesis Testing in Cointegrating Vector 2/2

2) Test for restrictions on β or α, where $\pi = \alpha\beta'$, α $(n \times r)$ is the matrix of the speed of adjustment parameters, and β' $(r \times n)$ is the matrix of cointegrating parameters. Use MLE to estimate VAR(p) model (6), where $e_t \sim iidN(0, \Sigma)$, $t - 1, 2, \cdots, T$, and $X_{1-p}, \cdots, X_0$ are given constant vectors, determine the rank of π, use the r most significant cointegrating vectors to form α, then select β such that $\pi = \alpha\beta'$.

The test statistic is

$$\left(-T\sum_{i=1}^{n}\ln(1-\hat{\lambda}_i)\right) - \left(-T\sum_{i=1}^{n}\ln(1-\hat{\lambda}_i^*)\right)$$

$$= \quad T\sum_{i=1}^{n}\left[\ln(1-\hat{\lambda}_i^*) - \ln(1-\hat{\lambda}_i)\right]$$

$$\sim \quad \chi^2(\text{the number of restrictions}), \ \text{asymptotically}.$$

we can test $\alpha_i = 0$, the variable x_{it} is weakly exogenous.

Notations for α, β. The VEC model is $x_t = A_0 + \alpha\beta' x_{t-1} + \sum_{i=1}^{p-1} \pi_i \Delta x_{t-i} + e_t$.Set $n = 3, r = 2$. Then

$$\begin{pmatrix} \Delta y_t \\ \Delta z_t \\ \Delta w_t \end{pmatrix} = \alpha\beta' x_{t-1} + \cdots = \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \\ \alpha_{31} & \alpha_{32} \end{pmatrix} \begin{pmatrix} \beta_{11} & \beta_{12} & \beta_{13} \\ \beta_{21} & \beta_{22} & \beta_{23} \end{pmatrix} \begin{pmatrix} y_{t-1} \\ z_{t-1} \\ w_{t-1} \end{pmatrix} + \cdots$$

$$= \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \\ \alpha_{31} & \alpha_{32} \end{pmatrix} \begin{pmatrix} \beta_{11} y_{t-1} + \beta_{12} z_{t-1} + \beta_{13} w_{t-1} \\ \beta_{21} y_{t-1} + \beta_{22} z_{t-1} + \beta_{23} w_{t-1} \end{pmatrix} + \cdots$$

$$= \begin{pmatrix} \alpha_{11}[\beta_{11} y_{t-1} + \beta_{12} z_{t-1} + \beta_{13} w_{t-1}] + \alpha_{12}[\beta_{21} y_{t-1} + \beta_{22} z_{t-1} + \beta_{23} w_{t-1}] \\ \alpha_{21}[\beta_{11} y_{t-1} + \beta_{12} z_{t-1} + \beta_{13} w_{t-1}] + \alpha_{22}[\beta_{21} y_{t-1} + \beta_{22} z_{t-1} + \beta_{23} w_{t-1}] \\ \alpha_{31}[\beta_{11} y_{t-1} + \beta_{12} z_{t-1} + \beta_{13} w_{t-1}] + \alpha_{32}[\beta_{21} y_{t-1} + \beta_{22} z_{t-1} + \beta_{23} w_{t-1}] \end{pmatrix} + \cdots$$

In EViews the rule in notations is the same on $A(i, j)$ and $B(i, j)$.For example, $A(2, 1)$ indicates the adjustment coefficient of the first cointegraion equation in the second equation of the VEC model. $B(2, 1)$ indicates the first coefficient in the second cointegraion equation.

Four Steps in Johansen-Stock-Watson Test

1. Pretest all variables to assess their order of integration.
2. Estimate the model and determine the rank of π.
3. Analyze the normalized cointegrating vectors and speed of adjustment coefficients and test the restrictions about each of both.
4. Conduct innovation accounting and causality tests on the error-correction model to identify a structural model and determine whether the estimated model appears to be reasonable.

Difference or Not Difference? VAR or ECM?

If the I(1) variables x_t are cointegrated, differencing them and estimating a VAR:

$$\Delta x_t = \sum_{i=1}^{p-1} \pi_i \Delta x_{t-i} + e_t$$

will lead to a misspecification error since it excludes the long-run equilibrium relationship among the variables that are included in πx_{t-1}.If the I(1) variables x_t are not cointegrated, it is preferable to estimate the VAR in first differences.

If I(1) x_t are not cointegrated $\Longrightarrow$ estimate VAR in first differences

If I(1) x_t are cointegrated $\Longrightarrow$ estimate the error-correction model