Time Series Analysis

Lecture 10

Review

1. Multivariate time series
2. Vector AR models
3. Vector MA models
4. Vector ARMA models

Today's Topics

1. Spurious Regression
2. Cointegration
3. Testing for Cointegration

Today's Topics

1. Spurious Regression
2. Cointegration
3. Testing for Cointegration

A toy example (Granger and Newbold, 1974JoE)

DGP: y_t and x_t are generated by the independent random walks

$$y_t = y_{t-1} + v_t, x_t = x_{t-1} + w_t, t = 1, 2, \cdots ,$$

in which v_t is $iidN(0, 1)$, w_t is $iidN(0, 1)$ and they are independent. The initial conditions are set as $x_0 = y_0 = 100$, and the sample length is 50.

Now y_t is regressed on a constant and another variate x_t giving the least squares regression

$$y_t = \hat{\alpha} + \hat{\beta} x_t + \hat{u}_t, t = 1, \cdots , T.$$

Table 1 shows values of

$$S = \frac{|\hat{\beta}|}{\hat{\sigma}_{\hat{\beta}}},$$

the customary statistic for testing the significance of β, for 100 simulations.

Table 1
Regressing two independent random walks.

S:	0–1	1–2	2–3	3–4	4–5	5–6	6–7	7–8
Frequency:	13	10	11	13	18	8	8	5
S:	8–9	9–10	10–11	11–12	12–13	13–14	14–15	15–16
Frequency:	3	3	1	5	0	1	0	1

Using the traditional t test at the 5% level, the null hypothesis of no relationship between the two series would be rejected (wrongly) on approximately three quarters of all occasions.

Today's Topics

1. Spurious Regression
   - Two independent processes
   - Extension to multiple regression
2. Cointegration
3. Testing for Cointegration

Two independent unit root processes (Phillips, 1986)

Suppose y_t and x_t are generated by the independent random walks

$$y_t = y_{t-1} + v_t, x_t = x_{t-1} + w_t, t = 1, 2, \cdots , \tag{1}$$

in which v_t is $(0, \sigma_v^2)$, w_t is $(0, \sigma_w^2)$ and they are independent.

Denote $\xi_t = (v_t, w_t)'$ and $S_t = \sum_{j=1}^{t} \xi_j$ and set $S_0 = 0$.

Assumption 1.

(a) $E(\xi_t) = 0$ for all t;

(b) $\sup_{i,t} E|\xi_{it}|^{\beta+\epsilon} < \infty$ for some $\beta > 2$ and $\epsilon > 0$;

(c) $\Sigma = \lim_{T\to\infty} T^{-1}E(S_T S_T')$ exists and is positive definite;

(d) $\{\xi_t\}_1^\infty$ is strong mixing with mixing numbers α_m satisfying $\sum_1^\infty \alpha_m^{1-2/\beta} < \infty$.

Based on what we have learned about unit root processes, we know

$$T^{-3/2} \sum_{1}^{T} x_t \Rightarrow \sigma_w \int_0^1 W(t)dt, \tag{2}$$

$$T^{-3/2} \sum_{1}^{T} y_t \Rightarrow \sigma_v \int_0^1 V(t)dt, \tag{3}$$

where $W(t)$ and $V(t)$ are independent Wiener processes on $C[0,1]$.

Now y_t is regressed on a constant and another variate x_t giving the least squares regression

$$y_t = \hat{\alpha} + \hat{\beta} x_t + \hat{u}_t, t = 1, \cdots, T. \tag{4}$$

Theorem 1. Suppose $\{y_t\}_1^\infty$ and $\{x_t\}_1^\infty$ are generated by (1), and the innovation sequence $\{v_t\}_1^\infty$ and $\{w_t\}_1^\infty$ are independent and $\{(v_t, w_t)\}_1^\infty$ satisfies Assumption 1. If (4) is estimated by least squares, then as $T \to \infty$,

(a) $\hat{\beta} \Rightarrow \frac{\sigma_v}{\sigma_w} \cdot \frac{\left\{ \int_0^1 V(t)W(t)\mathrm{d}t - \int_0^1 V(t)\mathrm{d}t \int_0^1 W(t)\mathrm{d}t \right\}}{\left\{ \int_0^1 W(t)^2\mathrm{d}t - (\int_0^1 W(t)\mathrm{d}t)^2 \right\}} \equiv \frac{\sigma_v}{\sigma_w} \cdot \zeta;$

(b) $T^{-1/2}\hat{\alpha} \Rightarrow \sigma_v \left\{ \int_0^1 V(t)\mathrm{d}t - \zeta \int_0^1 W(t)\mathrm{d}t \right\};$

(c) $T^{-1/2} t_\beta \Rightarrow \mu/\nu^{1/2}$, where

$\mu = \int_0^1 V(t)W(t)\mathrm{d}t - \int_0^1 V(t)\mathrm{d}t \int_0^1 W(t)\mathrm{d}t,$

$\nu = \left\{ \int_0^1 V(t)^2\mathrm{d}t - \left(\int_0^1 V(t)\mathrm{d}t \right)^2 \right\} \left\{ \int_0^1 W(t)^2\mathrm{d}t - \left(\int_0^1 W(t)\mathrm{d}t \right)^2 \right\} - \left\{ \int_0^1 V(t)W(t)\mathrm{d}t - \int_0^1 V(t)\mathrm{d}t \int_0^1 W(t)\mathrm{d}t \right\};$

(d) $T^{-1/2} t_\alpha \Rightarrow \left\{ \int_0^1 V(t)\mathrm{d}t - \zeta \int_0^1 W(t)\mathrm{d}t \right\} \times \left\{ \int_0^1 W(t)^2 \mathrm{d}t - \left(\int_0^1 W(t)\mathrm{d}t \right)^2 \right\} / \left[\nu \int_0^1 W(t)^2 \mathrm{d}t \right]^{1/2};$

(e) $R^2 \Rightarrow \frac{\zeta^2 \left\{ \int_0^1 W(t)^2 \mathrm{d}t - (\int_0^1 W(t)\mathrm{d}t)^2 \right\}}{\int_0^1 V(t)^2 \mathrm{d}t - (\int_0^1 V(t)\mathrm{d}t)^2};$

(f) The Durbin-Watson statistic $DW = \frac{\sum_2^T (\hat{u}_t - \hat{u}_{t-1})^2}{\sum_1^T \hat{u}_t^2} \xrightarrow{p} 0;$

Today's Topics

1. Spurious Regression
   - Two independent processes
   - Extension to multiple regression
2. Cointegration
3. Testing for Cointegration

Multiple regression

The results of the previous section are readily extended to the multiple regression of the form

$$y_t = \hat{\alpha} + \hat{\beta}' x_t + \hat{u}_t, t = 1, \cdots, T, \tag{5}$$

where y_t (a scalar) and x_t (an m-vector) are quite general integrated processes of order one.

It is not necessary to require that y_t and x_t be independent. The main requirement is that the vector time series (y_t, x_t') is not cointegrated (i.e., any linear combination of y_t and x_t remains to be integrated of order one) in the sense of Engle and Grange (1987).

Today's Topics

1. Spurious Regression
2. Cointegration
3. Testing for Cointegration

Today's Topics

Cointegration

An (n x 1) vector time series $\mathbf{y}_t$ is said to be *cointegrated* if each of the series taken individually is $I(1)$. that is, nonstationary with a unit root, while some linear combination of the series $\mathbf{a}'\mathbf{y_t}$ is stationary, or $\mathbf{I}(0)$, for some nonzero (n x 1) vector $\mathbf{a}$. When this is the case, $\mathbf{a}$ is called a *cointegrating vector*.

Example 1:

$$y_{1t} = \gamma y_{2t} + u_{1t} \tag{6}$$

$$y_{2t} = y_{2,t-1} + u_{2t}, \tag{7}$$

where u_{1t} and u_{2t} are uncorrelated white noise processes.

It is easy to show that both y_{1t} and y_{2t} are $I(1)$ processes, though the linear combination $(y_{1t} - \gamma y_{2t})$ is stationary. Hence, we would say that $\mathbf{y_t} = (y_{1t}, y_{2t})'$ is cointegrated with $\mathbf{a}' = (1, -\gamma)$

Figure 1 plots a sample realization of (6) and (7) for $\gamma = 1$ and u_{1t} and u_{2t} independent $N(0, 1)$ variables.

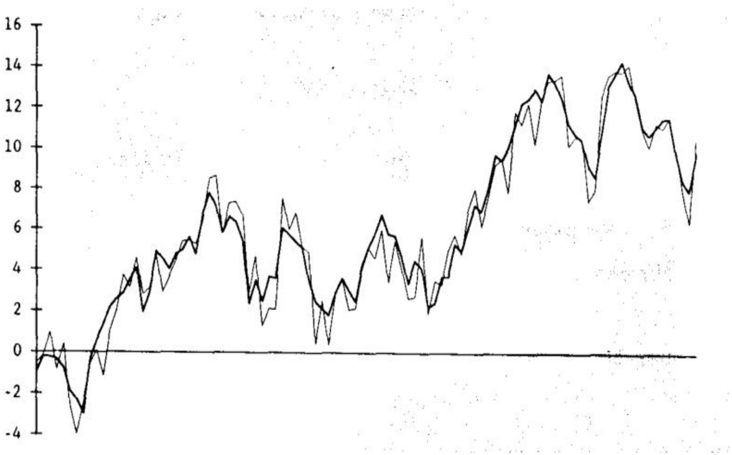

Figure 1 : Sample realization of cointegrated series

Cointegrating Vector

If there are more than two variables contained in $\mathbf{y_t}$, then there may be two nonzero (n x 1) vectors $\mathbf{a_1}$ and $\mathbf{a_2}$ such that $\mathbf{a_1'y_t}$ and $\mathbf{a_2'y_t}$ are both stationary, where $\mathbf{a_1}$ and $\mathbf{a_2}$ are linearly independent (that is, there does not exist a scalar b such that $\mathbf{a_2} = b\mathbf{a_1}$).

Indeed, there may be $h < n$ linearly independent (n x 1) vectors $(\mathbf{a_1}, \mathbf{a_2}, \ldots, \mathbf{a_h})$ such that $\mathbf{A'y_t}$ is a stationary (h x 1) vector, where $\mathbf{A'}$ is the following (h x n) matrix:

$$\mathbf{A'} \equiv \begin{bmatrix} \mathbf{a_1'} \\ \mathbf{a_2'} \\ \vdots \\ \mathbf{a_h'} \end{bmatrix} \tag{8}$$

Again, the vectors $(\mathbf{a_1}, \mathbf{a_2}, \ldots, \mathbf{a_h})$ are not unique; if $\mathbf{A'y_t}$ is stationary, then for any nonzero (1 x h) vector $\mathbf{b'}$, the scalar $\mathbf{b'A'y_t}$ is also stationary. Then the (n x 1) vector $\boldsymbol{\pi}$ given by $\boldsymbol{\pi}' = \mathbf{b'A'}$ could also be described as a cointegrating vector.

Suppose that there exists an (h x n) matrix $\mathbf{A'}$ whose rows are linearly independent such that $\mathbf{A'y_t}$ is a stationary (h x 1) vector. Suppose further that is $\mathbf{c'}$ is any (1 x n) vector that is linearly independent of the rows of $\mathbf{A'}$, then $\mathbf{c'y_t}$ is a nonstationary scalar. Then we say that there are exactly *h* cointegrating relations among the elements of $\mathbf{y_t}$ and that $(\mathbf{a_1}, \mathbf{a_2}, \ldots, \mathbf{a_h})$ form a *basis* for the space of cointegrating vectors.

Today's Topics

1. Spurious Regression
2. Cointegration
   - Definition
   - Error-Correction Representation
3. Testing for Cointegration

Revisit Example 1

The vector moving average representation for $(\Delta y_{1t}, \Delta y_{2t})'$ is

$$\begin{bmatrix} \Delta y_{1t} \\ \Delta y_{2t} \end{bmatrix} = \boldsymbol{\Psi}(L) \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}, \tag{9}$$

where $\varepsilon_{2t} \equiv u_{2t}$, $\varepsilon_{1t} \equiv \gamma u_{2t} + u_{1t}$,

$$\boldsymbol{\Psi}(L) = \begin{bmatrix} 1 - L & \gamma L \\ 0 & 1 \end{bmatrix} \tag{10}$$

A *VAR* for the differenced data, if it existed, would take the form

$$\boldsymbol{\Phi}(\mathbf{L})\boldsymbol{\Delta}\mathbf{y_t} = \boldsymbol{\varepsilon}_t$$

where $\boldsymbol{\Phi}(\mathbf{L}) = [\boldsymbol{\Psi}(\mathbf{L})]^{-\mathbf{1}}$.

But $\mathbf{\Psi}(z)$ has a root at unity,

$$|\mathbf{\Psi}(1)| = \begin{vmatrix} (1-1) & \gamma \\ 0 & 1 \end{vmatrix} = 0.$$

Hence the matrix moving average operator is noninvertible, and no finite-order vector autoregression could describe $\mathbf{\Delta y_t}$.

The reason is that the level of y_2 contains information that is useful for forecasting y_1 beyond that contained in a finite number of lagged *changes* in y_2 alone.

Although a VAR in difference is not consistent with a cointegrated system, a VAR in levels could be.

VAR in levels

Suppose that the level of $\mathbf{y_t}$ can be represented as a nonstationary pth-order vector autoregression:

$$\mathbf{y_t} = \boldsymbol{\alpha} + \boldsymbol{\Phi}_1 \mathbf{y}_{t-1} + \boldsymbol{\Phi}_2 \mathbf{y}_{t-2} + \cdots + \boldsymbol{\Phi}_p \mathbf{y}_{t-p} + \boldsymbol{\varepsilon}_t, \qquad (11)$$

or

$$\boldsymbol{\Phi}(L)\mathbf{y}_t = \boldsymbol{\alpha} + \boldsymbol{\varepsilon}_t. \qquad (12)$$

Suppose that $\Delta \mathbf{y}_t$ has the Wold decomposition

$$(1 - L)\mathbf{y}_t = \boldsymbol{\delta} + \boldsymbol{\Psi}(L)\boldsymbol{\varepsilon}_t.$$

If there are h cointegrations in $\mathbf{y}_t$, then it can be shown that (i)

$$\boldsymbol{\Phi}(1)\boldsymbol{\delta} = 0, \quad \boldsymbol{\Phi}(1)\boldsymbol{\Psi}(1) = 0;$$

and (ii) there exists an $(n \times h)$ matrix $\mathbf{B}$ such that

$$\boldsymbol{\Phi}(1) = \mathbf{B}\mathbf{A}', \qquad (13)$$

where rows of $\mathbf{A}'$ $(h \times n)$ form a basis for the space of cointegrating vectors.

Any *VAR* in the form of (11) can be equivalently be written as

$$\mathbf{y_t} = \zeta_1 \mathbf{\Delta y_{t-1}} + \zeta_2 \mathbf{\Delta y_{t-2}} + \cdots + \zeta_{\mathbf{p-1}} \mathbf{\Delta y_{t-p+1}} + \boldsymbol{\alpha} + \boldsymbol{\rho} \mathbf{y_{t-1}} + \boldsymbol{\varepsilon}_\mathbf{t}, \tag{14}$$

where

$$\boldsymbol{\rho} \equiv \mathbf{\Phi_1} + \mathbf{\Phi_2} + \cdots + \mathbf{\Phi_p},$$

$$\zeta_\mathbf{s} \equiv -[\mathbf{\Phi_{s+1}} + \mathbf{\Phi_{s+2}} + \cdots + \mathbf{\Phi_p}] \quad \textit{for } s = 1, 2, \ldots, p-1.$$

Substracting $\mathbf{y_{t-1}}$ from both sides of (14) produces

$$\mathbf{\Delta y_t} = \zeta_1 \mathbf{\Delta y_{t-1}} + \zeta_2 \mathbf{\Delta y_{t-2}} + \cdots + \zeta_{\mathbf{p-1}} \mathbf{\Delta y_{t-p+1}} + \boldsymbol{\alpha} + \zeta_\mathbf{0} \mathbf{y_{t-1}} + \boldsymbol{\varepsilon}_\mathbf{t}, \tag{15}$$

where

$$\zeta_\mathbf{0} \equiv \boldsymbol{\rho} - \mathbf{I_n} = -(\mathbf{I_n} - \mathbf{\Phi_1} - \mathbf{\Phi_2} - \cdots - \mathbf{\Phi_p}) = -\mathbf{\Phi}(1). \tag{16}$$

Note that if $\mathbf{y_t}$ has h cointegrating relations, then substitution of $\Phi(1) = \mathbf{BA}'$ and (16) into (15) results in

$$\mathbf{\Delta y_t} = \zeta_1 \mathbf{\Delta y_{t-1}} + \zeta_2 \mathbf{\Delta y_{t-2}} + \cdots + \zeta_{\mathbf{p-1}} \mathbf{\Delta y_{t-p+1}} + \alpha - \mathbf{Bz_{t-1}} + \varepsilon_\mathbf{t}. \tag{17}$$

where $\mathbf{z}_t \equiv \mathbf{A}'\mathbf{y}_t$ is a stationary $(h \times 1)$ vector.

Expression (17) is known as the error-correction representation of the cointegrated system.

Example 1

The system of (6) and (7) can be written in the form

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

The error-correction form is

$$\begin{bmatrix} \Delta y_{1t} \\ \Delta y_{2t} \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix} z_t + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix},$$

where $z_t \equiv y_{1t} - \gamma y_{2t}$, $\varepsilon_{1t} = \gamma u_{2t} + u_{1t}$, $\varepsilon_{2t} = u_{2t}$.

Summary

Granger Representation Theorem *Consider an $(n \times 1)$ vector $\mathbf{y_t}$ where $\mathbf{\Delta y_t}$ satisfies $(1-L)\mathbf{y_t} = \boldsymbol{\delta} + \boldsymbol{\Psi}(L)\boldsymbol{\varepsilon}_\mathbf{t}$ for $\boldsymbol{\varepsilon}_\mathbf{t}$ white noise with positive definite variance-covariance matrix and $\{s \cdot \boldsymbol{\Psi}_s\}_{s=0}^{\infty}$ absolutely summable. Suppose that there are exactly h cointegrating relations among the elements of $\mathbf{y_t}$. Then there exists an $(h \times n)$ matrix $\mathbf{A'}$ whose rows are linearly independent such that the $(h \times 1)$ vector $\mathbf{z_t}$ defined by*

$$\mathbf{z_t} \equiv \mathbf{A'y_t},$$

is stationary. The matrix $\mathbf{A'}$ has the property that

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

If, moreover, the process can be represented as the pth-order VAR in levels as $\Phi(L)\mathbf{y}_t = \boldsymbol{\alpha} + \boldsymbol{\varepsilon}_t$*, then there exists an* $(n \times h)$ *matrix* $\mathbf{B}$ *such that*

$$\Phi(1) = \mathbf{B}\mathbf{A}',$$

and there further exist $(n \times n)$ *matrices* $\boldsymbol{\zeta}_1, \boldsymbol{\zeta}_2, \ldots, \boldsymbol{\zeta}_{p-1}$ *such that*

$$\boldsymbol{\Delta}\mathbf{y}_\mathbf{t} = \boldsymbol{\zeta}_1\boldsymbol{\Delta}\mathbf{y}_{\mathbf{t-1}} + \boldsymbol{\zeta}_2\boldsymbol{\Delta}\mathbf{y}_{\mathbf{t-2}} + \cdots + \boldsymbol{\zeta}_{\mathbf{p-1}}\boldsymbol{\Delta}\mathbf{y}_{\mathbf{t-p+1}} + \boldsymbol{\alpha} - \mathbf{B}\mathbf{z}_{\mathbf{t-1}} + \boldsymbol{\varepsilon}_\mathbf{t}.$$

Today's Topics

1. Spurious Regression
2. Cointegration
3. Testing for Cointegration

Today's Topics

1. Spurious Regression
2. Cointegration
3. Testing for Cointegration
   - When the Cointegrating Vector is Known
   - Estimating the Cointegrating Vector
   - Phillips-Ouliaris-Hansen Tests for Cointegration

Testing Cointegration with Known Cointegrating Vector

If the interest in cointegration is motivated by the possibility of a particular known cointegrating vector $\mathbf{a}$, then by far the best method is to use this value directly to construct a test for cointegration.

To implement this approach, we

(i) first test whether each of the elements of $\mathbf{y}_t$ is individually $I(1)$. If yes, go to step (ii).

(ii) next construct the scalar $\mathbf{z}_t = \mathbf{a}'\mathbf{y}_t$ and test if it is $I(1)$ using, e.g., the ADF test.

Thus, a test of the null hypothesis that $\mathbf{z}_t$ is $I(1)$ is equivalent to a test of the null hypothesis that $\mathbf{y}_t$ is not cointegrated.

Example 2: Purchasing Power Parity

This theory holds that, apart from transportation costs, goods should sell for the same effective price in two countries. Let P_t denote an index of the price level in the United States (in dollars per good), P_t^* a price index for Italy (in lire per good), and S_t the rate of exchange between the currencies (in dollars per lira). Then purchasing power parity holds that

$$P_t = S_t P_t^*,$$

or, taking logarithms,

$$p_t = s_t + p_t^*,$$

where $p_t \equiv log P_t$, $s_t \equiv log S_t$. and $p_t^* \equiv log P_t^*$.

A weaker version of the hypothesis is that the variable z_t defined by

$$z_t \equiv p_t - s_t - p_t^*, \tag{18}$$

, i.e., the (log) real exchange rate, is stationary, even though the individual elements $(p_t,\ s_t,\ or\ p_t^*)$ are all $I(1)$.

For example, Figure 19.2 plots monthly data from 1973:1 to 1989:10 for the CPI for the United States (p_t) and Italy (p_t^*), along with the exchange rate (s_t). Natural logs of the raw data were taken and multiplied by 100, and initial value for 1973:1 was then subtracted.

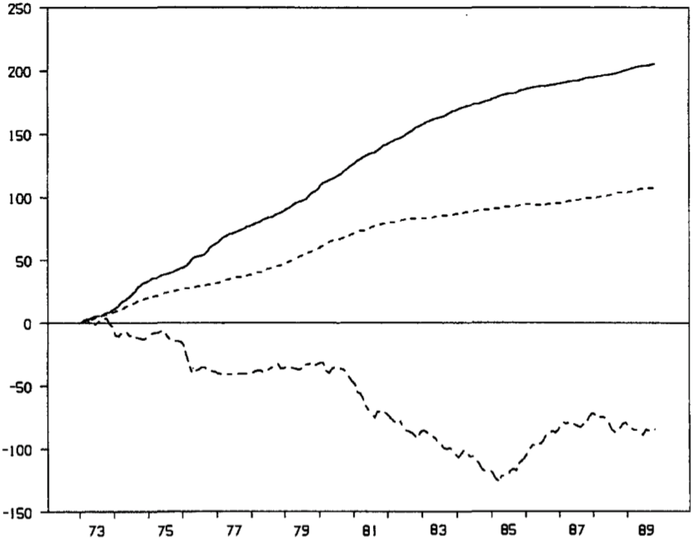

FIGURE 19.2 One hundred times the log of the price level in the United States (p_t), the dollar-lira exchange rate (s_t), and the price level in Italy (p_t^*), monthly, 1973–89. Key: ---- p_t; —-— s_t; ——— p_t^*.

Figure 19.3 plots the (log) real exchange rate,

$$z_t \equiv p_t - s_t - p_t^*$$

FIGURE 19.3 The real dollar-lira exchange rate, monthly, 1973–89.

To test for cointegration, we first verify that p_t, p_t^*, and s_t are each individually $I(1)$. These results are omitted here.

The next step is to test whether $z_t = p_t - s_t - p_t^*$ is stationary. According to the theory, there should not be any trend in z_t, and none appears evident in Figure 19.3. Thus ADF test without trend might be used. The following estimates were obtained by *OLS*:

$$\begin{aligned} z_t = {} & \underset{(0.07)}{0.32}\,\Delta z_{t-1} - \underset{(0.08)}{0.01}\,\Delta z_{t-2} + \underset{(0.08)}{0.01}\,\Delta z_{t-3} + \underset{(0.08)}{0.02}\,\Delta z_{t-4} \\ & + \underset{(0.08)}{0.08}\,\Delta z_{t-5} - \underset{(0.08)}{0.00}\,\Delta z_{t-6} + \underset{(0.08)}{0.03}\,\Delta z_{t-7} + \underset{(0.08)}{0.08}\,\Delta z_{t-8} \\ & - \underset{(0.08)}{0.05}\,\Delta z_{t-9} + \underset{(0.08)}{0.08}\,\Delta z_{t-10} + \underset{(0.08)}{0.05}\,\Delta z_{t-11} - \underset{(0.08)}{0.01}\,\Delta z_{t-12} \\ & + \underset{(0.18)}{0.00} + \underset{(0.01410)}{0.97124}\, z_{t-1} \end{aligned}$$

Here the ADF t test is

$$t = (0.97124 - 1.0)/(0.01410) = -2.04.$$

Comparing this with the 5% critical value for case 2 of Table B.6 (Hamilton, 1994), we see that -2.04>-2.88, and so the null hypothesis of a unit root is accepted. The F test of the joint null hypothesis that $\rho = 1$ and that the constant term is zero is 2.19<4.66, which is again accepted. Thus, we could accept the null hypothesis that the series are not cointegrated.

Alternatively, the null hypothesis that z_t is nonstationary could be tested using the Phillips-Perron tests. The Phillips-Perron Z_p test and the Phillips-Perron Z_t test give the same conclusion. Details are omitted here.

Today's Topics

1. Spurious Regression
2. Cointegration
3. Testing for Cointegration
   - When the Cointegrating Vector is Known
   - Estimating the Cointegrating Vector
   - Phillips-Ouliaris-Hansen Tests for Cointegration

OLS when $h = 1$

If the theoretical model of the system dynamics does not suggest a particular value for the cointegrating vector $\mathbf{a}$, then one approach to testing for cointegration is first to estimate $\mathbf{a}$ by *OLS*.

Let $z_t = \mathbf{a'y_t}$. We can obtain a consistent estimate of a cointegrating vector by choosing $\mathbf{a}$ so as to minimize

$$T^{-1} \sum_{t=1}^{T} z_t^2 = T^{-1} \sum_{t=1}^{T} (\mathbf{a'y_t})^2$$

subject to some normalization condition on $\mathbf{a}$.

The consistency of *OLS* can be seen by noting that $T^{-1} \sum_{t=1}^{T} (\mathbf{a'y_t})^2$ converges in probability when $\mathbf{a}$ is a cointegrating vector and diverges to infinity while it is not. Such an estimator of $\mathbf{a}$ turns out to be converging at rate T rather than $T^{1/2}$.

If it is known for certain that $a_1 \neq 0$, then a particularly convenient normalization is to set $a_1 = 1$ and represent subsequent entries of $\mathbf{a}(a_2, a_3, \cdots, a_n)$ as the negatives of a set of unknown parameters $(\gamma_2, \gamma_3, \cdots, \gamma_n)$:

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} 1 \\ -\gamma_2 \\ -\gamma_3 \\ \vdots \\ -\gamma_n \end{bmatrix}. \tag{19}$$

In this case, the objective is to choose $(\gamma_2, \gamma_3, \cdots, \gamma_n)$ so as to minimize

$$T^{-1} \sum_{t=1}^{T} (\mathbf{a}'\mathbf{y_t})^2 = T^{-1} \sum_{t=1}^{T} (y_{1t} - \gamma_2 y_{2t} - \gamma_3 y_{3t} - \cdots - \gamma_n y_{nt})^2. \tag{20}$$

This minimization is achieved by an *OLS* regression of the first element of $\mathbf{y_t}$ on all of the others:

$$y_{1t} = \gamma_2 y_{2t} + \gamma_3 y_{3t} + \cdots + \gamma_n y_{nt} + u_t \tag{21}$$

Consistent estimates of $\gamma_2, \gamma_3, \cdots, \gamma_n$ are also obtained when a constant term is included in (21), as in

$$y_{1t} = \alpha + \gamma_2 y_{2t} + \gamma_3 y_{3t} + \cdots + \gamma_n y_{nt} + u_t \tag{22}$$

or

$$y_{1t} = \alpha + \boldsymbol{\gamma'} \mathbf{y}_{2t} + u_t \tag{23}$$

where $\boldsymbol{\gamma}' \equiv (\gamma_2, \gamma_3, \cdots, \gamma_n)$ and $\mathbf{y}_{2t} \equiv (y_{2t}, y_{3t}, \cdots, y_{nt})'$.

The OLS estimates, $\hat{\alpha}_T$ is $\sqrt{T}$-consistent, and $\hat{\gamma}_T$ is T-consistent (Phillips and Durlauf, 1986; Stock, 1987).

The role of normalization

For example, with $n = 2$, if we normalize $a_1 = 1$, we would regress y_{1t} on y_{2t}:

$$y_{1t} = \alpha + \gamma y_{2t} + u_t$$

Obviously, we might equally well have normalized $a_2 = 1$ and used the same argument to suggest a regression of y_{2t} on y_{1t}:

$$y_{2t} = \theta + \aleph y_{1t} + v_t$$

The *OLS* estimate $\hat{\aleph}$ is not simply the inverse of $\hat{\gamma}$ meaning that these two regressions will give different estimates of the cointegrating vector:

$$\begin{bmatrix} 1 \\ -\hat{\gamma} \end{bmatrix} \neq -\hat{\gamma} \begin{bmatrix} -\hat{\aleph} \\ 1 \end{bmatrix} .$$

Only in the limiting case when $R^2 = 1$ would the two estimates coincide.

Thus, choosing which variable to call y_1 and which to call y_2 might end up making a material difference for the estimate of $\mathbf{a}$ as well as for the evidence one finds for cointegration among the series.

One approach that avoids this normalization problem is the full-information maximum likelihood estimate (FIMLE) proposed by Johansen (1988, 1991).

$h > 1$

In the more general case with $h > 1$, *OLS* estimation of (22) should still provide a consistent estimate of a cointegrating vector. But which cointegrating vector is it?

Among the set of possible cointegrating relations, *OLS* estimation of (22) selects the relation whose residuals are uncorrelated with any other $I(1)$ linear combinations of $(y_{2t}, y_{3t}, \cdots, y_{nt})$ (Wooldridge, 1991).

$h = 0$

We have seen that if there is at least one cointegrating relation involving y_{1t}, then *OLS* estimation of (22) gives a consistent estimate of a cointegrating vector.

Let us now consider the properties of *OLS* estimation when there is no cointegrating relation. Then (22) is a regression of an $I(1)$ variable on a set of $(n-1)I(1)$ variables for which no coefficients produce an $I(0)$ error term. The regression is therefore subject to the spurious regression problem.

In particular, the *OLS* sample residuals $\hat{u}_t$ will be nonstationary. This property can be exploited to test for cointegration. If there is no cointegration, then a regression of $\hat{u}_t$ on $\hat{u}_{t-1}$ should yield a unit coefficient. If there is cointegration, then a regression of $\hat{u}_t$ on $\hat{u}_{t-1}$ should yield a coefficient that is less than 1.

Today's Topics

Phillips-Ouliaris-Hansen Tests for Cointegration

The proposal is to estimate

$$y_{1t} = \alpha + \gamma_2 y_{2t} + \gamma_3 y_{3t} + \cdots + \gamma_n y_{nt} + u_t \tag{24}$$

by *OLS* and then construct one of the standard unit root tests on the estimated residuals, such as the augmented Dickey-Fuller t test or the Phillips (1987) Z_ρ or Z_t test.

Although these test statistics are constructed in the same way as when they are applied to an individual series y_t, when the tests are applied to the residuals $\hat{u}_t$ from a spurious regression, the critical values that are used to interpret the test statistics are different.

Residual-based Tests for Cointegration

Specifically, let $\mathbf{y_t}$ be an $(n \times 1)$ vector partitioned as

$$\underset{(n\times 1)}{\mathbf{y_t}} = \begin{bmatrix} \underset{(1\times 1)}{y_{1t}} \\ \underset{(g\times 1)}{\mathbf{y}_{2t}} \end{bmatrix} \tag{25}$$

for $g \equiv (n-1)$. Consider the regression

$$y_{1t} = \alpha + \gamma' \mathbf{y}_{2t} + u_t. \tag{26}$$

Let $\hat{u}_t$ be the sample residual associated with *OLS* estimation of (26) in a sample of size T:

$$\hat{u}_t = y_{1t} - \hat{\alpha}_T - \hat{\gamma}_{\mathbf{T}}' \mathbf{y}_{2t} \quad \text{for } t = 1, 2, \ldots, T. \tag{27}$$

where

$$\begin{bmatrix} \hat{\alpha}_T \\ \hat{\gamma}_{\mathbf{T}} \end{bmatrix} = \begin{bmatrix} T & \sum \mathbf{y}_{2t}' \\ \sum \mathbf{y}_{2t} & \sum \mathbf{y}_{2t}\mathbf{y}_{2t}' \end{bmatrix}^{-1} \begin{bmatrix} \sum y_{1t} \\ \sum \mathbf{y}_{2t} y_{1t} \end{bmatrix},$$

and where $\sum$ indicates summation over t from 1 to T.

Residual regression

The residual $\hat{u}_t$ can then be regressed on its own lagged value $\hat{u}_{t-1}$ without a constant term:

$$\hat{u}_t = \rho \hat{u}_{t-1} + e_t \quad \text{for} t = 2, 3, \ldots, T. \tag{28}$$

yielding the estimate

$$\hat{\rho}_T = \frac{\sum_{t=2}^{T} \hat{u}_{t-1} \hat{u}_t}{\sum_{t=2}^{T} \hat{u}_{t-1}^2}.$$

Phillips's Z_ρ statistic

Phillips's Z_ρ statistic (1987) can be calculated as:

$$Z_{\rho,T} = (T-1)(\hat{\rho}_T - 1) - (1/2) \cdot \{(T-1)^2 \cdot \hat{\sigma}^2_{\hat{\rho}T} \div s^2_T\} \cdot \{\hat{\lambda}^2_T - \hat{c}_{0,T}\},$$

where s^2_T is the *OLS* estimate of the variance of e_t for the regression of (27):

$$s^2_T = (T-2)^{-1} \sum_{t=2}^{T} (\hat{u} - \hat{\rho}_T \hat{u}_{t-1})^2,$$

$\hat{\sigma}_{\hat{\rho}T}$ is the standard error of $\hat{\rho}_T$:

$$\hat{\sigma}^2_{\hat{\rho}T} = s^2_T \div \{ \sum_{t=2}^{T} \hat{u}^2_{t-1} \}.$$

and $\hat{c}_{j,T}$ is the jth sample autocovariance of the estimated residuals associated with (28):

$$\hat{c}_{j,T} = (T-1)^{-1} \sum_{t=j+2}^{T} \hat{e}_t \hat{e}_{t-j} \quad \text{for } j = 0, 1, 2, \ldots, T-2$$

for $\hat{e} \equiv \hat{u}_t - \hat{\rho}_T \hat{u}_{t-1}$; and the square of $\hat{\lambda}_T$ is given by

$$\hat{\lambda}^2_T = \hat{c}_{0,T} + 2 \cdot \sum_{j=1}^{q} [1 \ - \ j/(q+1)] \hat{c}_{j,T},$$

where q is the number of autocovariances to be used.

Phillips's Z_t statistic

Phillips's Z_t statistic associated with (28) would be

$$Z_{t,T} = (\hat{c}_{0,T}/\hat{\lambda}_T^2)^{1/2} \cdot t_T - (1/2) \cdot \{(T-1) \cdot \hat{\sigma}_{\hat{\rho}T} \div s_T\} \cdot \{\hat{\lambda}_T^2 - \hat{c}_{0,T}\}/\hat{\lambda}_T \tag{29}$$

for t_T the usual OLS t statistic for testing the hypothesis $\rho = 1$:

$$t_T = (\hat{\rho}_T - 1)/\hat{\sigma}_{\hat{\rho}T}.$$

Augmented Dickey-Fuller Tests

Alternatively, lagged changes in the residuals could be added to the regression of (28) as in the augmented Dickey-Fuller test with no constant term:

$$\hat{u}_t = \zeta_1 \Delta \hat{u}_{t-1} + \zeta_2 \Delta \hat{u}_{t-2} + \cdots + \zeta_{p-1} \Delta \hat{u}_{t-p+1} + \rho \hat{u}_{t-1} + e_t.$$

Again, this is estimated by *OLS* for $t = p+1, p+2, \ldots, T$, and the *OLS* t test of $\rho = 1$ is calculated using the standard *OLS* formula. If this t statistic or the Z_t statistic in (29) is negative and sufficiently large in absolute value, this again casts doubt on the null hypothesis of no cointegration.

Summary

Case 1:

- Estimated cointegrating regression:

$$y_{1t} = \gamma_2 y_{2t} + \gamma_3 y_{3t} + \cdots + \gamma_n y_{nt} + u_t$$

- True process for $\mathbf{y_t} = (y_{1t}, y_{2t}, \ldots, y_{nt})'$:

$$\mathbf{\Delta y_t} = \sum_{s=0}^{\infty} \mathbf{\Psi}_s \varepsilon_{t-s}$$

- $Z_\rho \sim$ Case 1 in Table B.8.
- $Z_t \sim$ and the augmented Dickey-Fuller t test $\sim$ Case 1 in Table B.9.

Case 2:

- Estimated cointegrating regression:

$$y_{1t} = \alpha + \gamma_2 y_{2t} + \gamma_3 y_{3t} + \cdots + \gamma_n y_{nt} + u_t$$

- True process for $\mathbf{y_t} = (y_{1t}, y_{2t}, \ldots, y_{nt})'$:

$$\mathbf{\Delta y_t} = \sum_{s=0}^{\infty} \mathbf{\Psi}_s \varepsilon_{t-s}$$

- $Z_\rho \sim$ Case 2 in Table B.8.
- $Z_t$and the augmented Dickey-Fuller t test $\sim$ Case 2 in Table B.9.

Case 3:

- Estimated cointegrating regression:

$$y_{1t} = \alpha + \gamma_2 y_{2t} + \gamma_3 y_{3t} + \cdots + \gamma_n y_{nt} + \mathbf{u_t}$$

- True process for $\mathbf{y_t} = (y_{1t}, y_{2t}, \ldots, y_{nt})'$:

$$\boldsymbol{\Delta}\mathbf{y_t} = \boldsymbol{\delta} + \sum_{s=0}^{\infty} \boldsymbol{\Psi}_s \varepsilon_{t-s}$$

  with at least one element of $\delta_2, \delta_3, \ldots, \delta_n$ nonzero.

- $Z_\rho \sim$ Case 3 in Table B.8.
- Z_t and the augmented Dickey-Fuller t test $\sim$ Case 3 in Table B.9.

TABLE B.8
Critical Values for the Phillips Z_ρ Statistic When Applied to Residuals from Spurious Cointegrating Regression

Number of right-hand variables in regression, excluding trend or constant $(n-1)$	Sample size (T)	Probability that $(T-1)(\hat{\rho}-1)$ is less than entry						
		0.010	0.025	0.050	0.075	0.100	0.125	0.150
		Case 1						
1	500	−22.8	−18.9	−15.6	−13.8	−12.5	−11.6	−10.7
2	500	−29.3	−25.2	−21.5	−19.6	−18.2	−17.0	−16.0
3	500	−36.2	−31.5	−27.9	−25.5	−23.9	−22.6	−21.5
4	500	−42.9	−37.5	−33.5	−30.9	−28.9	−27.4	−26.2
5	500	−48.5	−42.5	−38.1	−35.5	−33.8	−32.3	−30.9
		Case 2						
1	500	−28.3	−23.8	−20.5	−18.5	−17.0	−15.9	−14.9
2	500	−34.2	−29.7	−26.1	−23.9	−22.2	−21.0	−19.9
3	500	−41.1	−35.7	−32.1	−29.5	−27.6	−26.2	−25.1
4	500	−47.5	−41.6	−37.2	−34.7	−32.7	−31.2	−29.9
5	500	−52.2	−46.5	−41.9	−39.1	−37.0	−35.5	−34.2
		Case 3						
1	500	−28.9	−24.8	−21.5	—	−18.1	—	—
2	500	−35.4	−30.8	−27.1	−24.8	−23.2	−21.8	−20.8
3	500	−40.3	−36.1	−32.2	−29.7	−27.8	−26.5	−25.3
4	500	−47.4	−42.6	−37.7	−35.0	−33.2	−31.7	−30.3
5	500	−53.6	−47.1	−42.5	−39.7	−37.7	−36.0	−34.6

The probability shown at the head of the column is the area in the left-hand tail.

TABLE B.9
Critical Values for the Phillips Z_t Statistic or the Dickey-Fuller t Statistic When Applied to Residuals from Spurious Cointegrating Regression

Number of right-hand variables in regression, excluding trend or constant $(n-1)$	Sample size (T)	Probability that $(\hat{\rho}-1)/\hat{\sigma}_{\hat{\rho}}$ is less than entry						
		0.010	0.025	0.050	0.075	0.100	0.125	0.150
		Case 1						
1	500	−3.39	−3.05	−2.76	−2.58	−2.45	−2.35	−2.26
2	500	−3.84	−3.55	−3.27	−3.11	−2.99	−2.88	−2.79
3	500	−4.30	−3.99	−3.74	−3.57	−3.44	−3.35	−3.26
4	500	−4.67	−4.38	−4.13	−3.95	−3.81	−3.71	−3.61
5	500	−4.99	−4.67	−4.40	−4.25	−4.14	−4.04	−3.94
		Case 2						
1	500	−3.96	−3.64	−3.37	−3.20	−3.07	−2.96	−2.86
2	500	−4.31	−4.02	−3.77	−3.58	−3.45	−3.35	−3.26
3	500	−4.73	−4.37	−4.11	−3.96	−3.83	−3.73	−3.65
4	500	−5.07	−4.71	−4.45	−4.29	−4.16	−4.05	−3.96
5	500	−5.28	−4.98	−4.71	−4.56	−4.43	−4.33	−4.24
		Case 3						
1	500	−3.98	−3.68	−3.42	—	−3.13	—	—
2	500	−4.36	−4.07	−3.80	−3.65	−3.52	−3.42	−3.33
3	500	−4.65	−4.39	−4.16	−3.98	−3.84	−3.74	−3.66
4	500	−5.04	−4.77	−4.49	−4.32	−4.20	−4.08	−4.00
5	500	−5.36	−5.02	−4.74	−4.58	−4.46	−4.36	−4.28

The probability shown at the head of the column is the area in the left-hand tail.

Other tests

The tests that have seen described in this section can give different answers depending on which variable is labeled y_1.

Important tests for cointegration that are invariant to the ordering of variables are the full-information maximum likelihood test of Johansen (1988, 1991) and the related tests of Stock and Watson (1988) and Ahn and Reinsel (1990).

Other useful tests for cointegration have been proposed by Phillips and Ouliaris (1990), Park, Ouliaris, and Choi (1988), Stock (1990), and Hansen (1990).