Nonstationary Time Series Models

Yanping YI

October 19, 2012

Table of contents

Why do we need to test for Non-Stationarity?

- The stationarity or otherwise of a series can strongly influence its behavior and properties - e.g. persistence of shocks will be infinite for non-stationary series
- Spurious regressions : if two variables are trending over time, a regression of one on the other could have a high R^2 even if the two are totally unrelated
- If the variables in the regression model are not stationary, then it can be proved that the standard assumptions for asymptotic analysis will not be valid. In other words, the usual "t-ratios" will not follow a t-distribution, so we cannot validly undertake hypothesis tests about the regression parameters.

Introduction to Non-Stationary Time Series

Stationary Time Series :

- Exhibits mean reversion in that it fluctuates around a constant long run mean
- Has a finite variance that is time invariant
- Has a theoretical covariance between values of y_t that depends only on the difference apart in time

Introduction to Non-Stationary Time Series

However, in many economic contexts the stationarity assumptions are rather implausible.

- There is no long-run mean to which the series returns
- The variance is time dependent and goes to infinity as time approaches to infinity
- Theoretical autocorrelations do not decay but, in finite samples, the sample correlogram dies out slowly
- For example, U.S. quarterly GDP 1947 to 2008 ; foreign exchange rates; price series of an asset

Introduction to Non-Stationary Time Series

There are two common models for nonstationarity in economic time series. The two will require different treatments to induce stationarity.

1. Deterministic time trends (also called trend-stationary) : what is required is detrending.
2. Unit root processes (also known as stochastic non-stationarity): we induce stationarity by differencing once .

Introduction to Non-Stationary Time Series

The general formulation

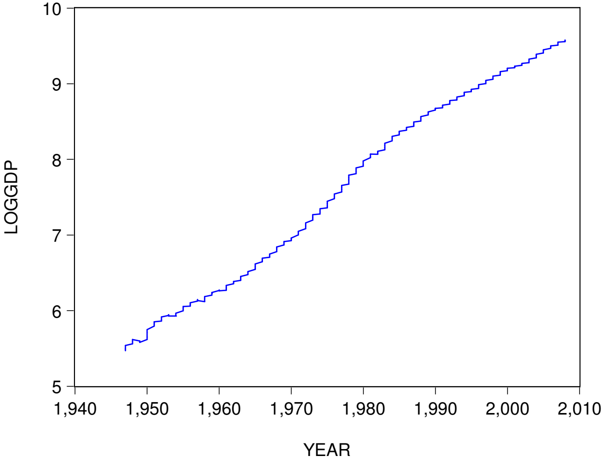
$$y_t = \alpha + \delta t + u_t$$

where u_t follows a zero mean ARMA process:

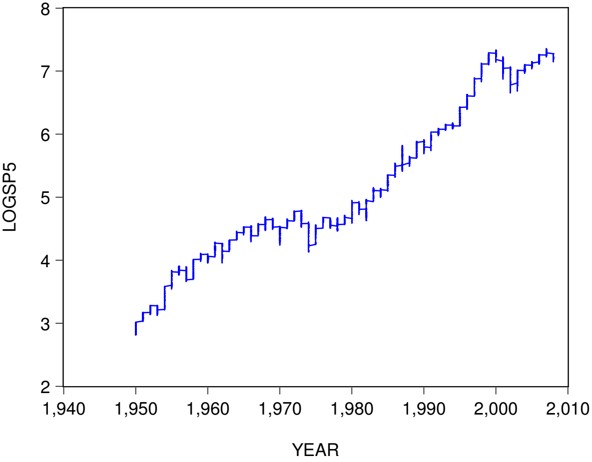
$$(1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p) u_t = (1 + \theta_1 L + \theta_2 L^2 + \cdots + \theta_q L^q)\epsilon_t$$

$$\phi(L) u_t = \theta(L)\epsilon_t$$

Time plot of logged U.S. quarterly GDP from 1947 to 2008



Time plot of logarithm of daily S&P 500 index from 01/03/1950 to 04/16/2008



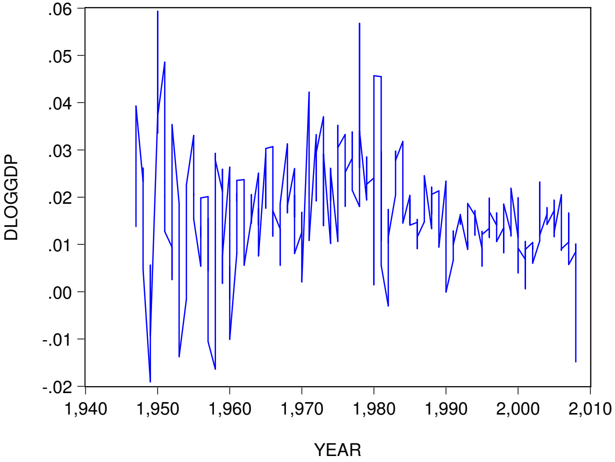
Types of Model from the General Formulation $y_t = \alpha + \delta t + u_t$

Parameter Set	Description
$\delta = 0$, The roots of $\phi(z) = 0$ lie outside the unit circle	ARMA (p, q) process
$\delta \neq 0$, The roots of $\phi(z) = 0$ lie outside the unit circle	trend stationary process
$\delta = 0$, $\phi(L) = (1 - L)\psi(L)$ The roots of $\psi(z) = 0$ lie outside the unit circle	ARIMA(p-1, 1, q) process or unit root process or $I(1)$
$\delta \neq 0$, $\phi(L) = (1 - L)\psi(L)$ The roots of $\psi(z) = 0$ lie outside the unit circle	ARIMA(p-1, 1, q) process or $I(1)$ or unit root process with deterministic trend

Types of Model from the General Formulation $y_t = \alpha + \delta t + u_t$

- We can generalize this concept to consider the case where the series contains more than one "unit root". That is, we would need to apply the first difference $\triangle$ more than once to induce stationarity.
- If a non-stationary series, y_t must be differenced d times before it becomes stationary, then it is said to be integrated of order d. We write $y_t \sim I(d)$. So if $y_t \sim I(d)$, then $\triangle^d y_t \sim I(0)$.
- Taking *dth* differences of an ARIMA(p, d, q) process produces a stationary ARMA (p, q) process.
- Sometimes such models are formulated in logs so in these cases such models may be thought of exhibiting exponential growth.

Time plot of first log difference series of U.S. quarterly GDP from 1947 to 2008



Detrending a Series: Using the Right Method

- Although trend-stationary and difference-stationary series are both "trending" over time, the correct approach needs to be used in each case.
- If we first difference the trend-stationary series, it would "remove" the non-stationarity, but at the expense on introducing an MA(1) structure into the errors.
- Conversely if we try to detrend a series which has stochastic trend, then we will not remove the non-stationarity.
- We will now concentrate on the stochastic non-stationarity model since deterministic non-stationarity does not adequately describe most series in economics or finance.

Comparison of Stationary Processes and Unit Root Processes

- Unit Root Processes can wander a long way from their mean value and cross this mean value rarely.
- Stationary series should cross the mean frequently.
- The majority of economic and financial series contain a single unit root, although some are stationary and consumer prices have been argued to have 2 unit roots.
- Forecasts
- Forecast errors
- Sample ACFs approach 1 for any finite lag (strong memory)

How do we test for a unit root?

The early and pioneering work on testing for a unit root in time series was done by Dickey and Fuller (Dickey and Fuller 1979, Fuller 1976).

No constant term in the test regression

- True Process (H_0) : $y_t = y_{t-1} + u_t$, where $u_t \sim i.i.d.\ N(0, \sigma^2)$ and $y_0 = 0$.
- No constant term in the test regression: $y_t = \rho y_{t-1} + u_t$. Then the OLS estimator $\hat{\rho_T} = \frac{\sum_{t=1}^{T} y_{t-1} y_t}{\sum_{t=1}^{T} y_{t-1}^2}$.
- To get better sense of the asymptotic of $\hat{\rho_T}$, we derive

$$\hat{\rho_T} - 1 = \frac{\sum_{t=1}^{T} y_{t-1} u_t}{\sum_{t=1}^{T} y_{t-1}^2}$$

The denominator is $O_p(T^2)$, while the numerator is $O_p(T)$ (FCLT).

No constant term in the test regression

- So it makes sense to multiply this equation by T

$$T(\hat{\rho_T} - 1) = \frac{\frac{1}{T}\sum_{t=1}^{T} y_{t-1}u_t}{\frac{1}{T^2}\sum_{t=1}^{T} y_{t-1}^2}$$

- By FCLT and continuous mapping theorem,

$$T(\hat{\rho_T} - 1) \xrightarrow{\mathcal{D}} \frac{\frac{1}{2}(W^2(1) - 1)}{\int_0^1 W^2(r)dr} \quad \textbf{superconsistent} \qquad (1)$$

- The probability that $T(\hat{\rho_T} - 1)$ is negative approaches 0.68 as T becomes large. The limiting distribution of $T(\hat{\rho_T} - 1)$ is skewed to the left.

No constant term in the test regression

- OLS t test of $H_0 : \rho = 1$:

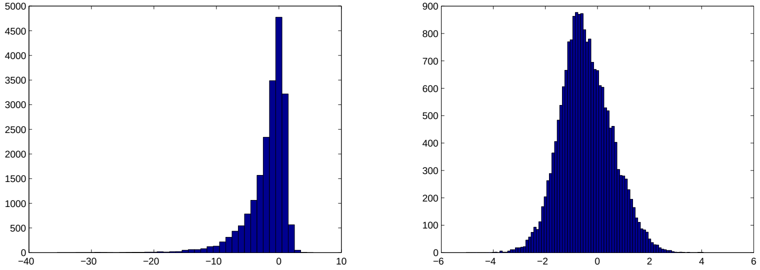
$$t_T = \frac{\hat{\rho_T} - 1}{s_T / \sqrt{\sum_{t=1}^{T} y_{t-1}^2}} \quad \textbf{Dickey Fuller test}$$

where $s_T / \sqrt{\sum_{t=1}^{T} y_{t-1}^2}$ is the usual OLS standard error for the estimated coefficient, and $s_T^2 = \sum_{t=1}^{T} (y_t - \hat{\rho_T} y_{t-1})^2 / (T-1)$ is the OLS estimate of the residual variance.

- Again, we can use FCLT and continuous mapping theorem to show

$$t_T \xrightarrow{\mathcal{D}} \frac{\frac{1}{2}(W^2(1) - 1)}{\left[\int_0^1 W^2(r) dr\right]^{1/2}} \tag{2}$$

No constant term in the test regression



(a)　　(b)

Figure: (a) Simulated Histogram for the limiting distribution in equation (1) ; (b) Simulated Histogram for the limiting distribution in equation (2).

Constant term is included in the test regression

- True Process (H_0) : $y_t = y_{t-1} + u_t$, where $u_t \sim i.i.d.\ N(0, \sigma^2)$ and $y_0 = 0$.
- Constant term is included in the test regression: $y_t = \alpha + \rho y_{t-1} + u_t$.
- Also by FCLT and continuous mapping theorem,

$$T(\hat{\rho_T} - 1) \xrightarrow{\mathcal{D}} \frac{\frac{1}{2}(W^2(1) - 1) - W(1)\int_0^1 W(r)dr}{\int_0^1 W^2(r)dr - \left[\int_0^1 W(r)dr\right]^2}$$

Constant term is included in the test regression

The asymptotic distribution of the estimate of ρ is not the same as the asymptotic distribution in equation (1). Note that this distribution is even more strongly skewed than that for the previous case, so that when a constant term is included in the regression, the estimated coefficient on y_{t-1} must be farther from unity in order to reject the null hypothesis of a unit root.

DF test of $H_0 : \rho = 1$

$$t_T = \frac{\hat{\rho_T} - 1}{\hat{\sigma}_{\hat{\rho_T}}} \xrightarrow{\mathcal{D}} \frac{\frac{1}{2}(W^2(1) - 1) - W(1) \int_0^1 W(r)dr}{\sqrt{\int_0^1 W^2(r)dr - \left[\int_0^1 W(r)dr\right]^2}}$$

- So far we have only considered the case $y_t = y_{t-1} + u_t$, where $u_t \sim i.i.d.\ N(0, \sigma^2)$ and $y_0 = 0$. What if $\{u_t\}$ is not $i.i.d.$?
- Unit root processes with general serial correlation

$$
\begin{aligned}
y_t &= y_{t-1} + u_t \\
u_t &= \psi(L)\epsilon_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j} \quad \sum_{j=0}^{\infty} j|\psi_j| < \infty \quad 1 \ \textit{summable}
\end{aligned}
$$

where $\{\epsilon_t\}$ is an $i.i.d.$ sequence with mean zero, variance σ^2 and finite fourth moment.
In other words, u_t will be **autocorrelated** if there was autocorrelation in the dependent variable of the regression $(\triangle y_t)$ which we have not modeled.

Augmented Dickey-Fuller Tests for Unit Roots

- The basic idea of ADF tests is to control the serial correlation by including higher-order autoregressive terms in the regression.
- Suppose the true process is $(1-\phi_1 L-\phi_2 L^2-\cdots-\phi_p L^p)y_t=\epsilon_t$ and $\{\epsilon_t\}$ is an $i.i.d.$ sequence with mean zero, variance σ^2 and finite fourth moment. (What is the relation between this process and $y_t-y_{t-1}=\psi(L)\epsilon_t$ we studied before?)

Augmented Dickey-Fuller Tests for Unit Roots

- We can rewrite the model as

$$\begin{aligned} y_t &= \rho y_{t-1} + \eta_1 \Delta y_{t-1} + \cdots + \eta_{p-1} \Delta y_{t-p+1} + \epsilon_t \\ or\ \Delta y_t &= \tilde{\rho} y_{t-1} + \eta_1 \Delta y_{t-1} + \cdots + \eta_{p-1} \Delta y_{t-p+1} + \epsilon_t \end{aligned}$$

  where $\rho \equiv \phi_1 + \phi_2 + \cdots + \phi_p$, $\eta_j \equiv -[\phi_{j+1} + \phi_{j+2} + \cdots + \phi_p]$ for $j = 1, 2, \cdots, p-1$, and $\tilde{\rho} \equiv \rho - 1$.
- It can be shown that testing unit roots is equivalent to testing $\rho = 1$ (or $\tilde{\rho} = 0$).
- We consider the regression including a constant term. $y_t = \alpha + \rho y_{t-1} + \eta_1 \Delta y_{t-1} + \cdots + \eta_{p-1} \Delta y_{t-p+1} + \epsilon_t$.

Augmented Dickey-Fuller Tests for Unit Roots

- Augmented Dickey-Fuller t Test for unit roots:

$$t_T = \frac{\hat{\rho_T} - 1}{\hat{\sigma}_{\hat{\rho_T}}} \xrightarrow{\mathcal{D}} \frac{\frac{1}{2}(W^2(1) - 1) - W(1)\int_0^1 W(r)dr}{\sqrt{\int_0^1 W^2(r)dr - \left[\int_0^1 W(r)dr\right]^2}}$$

- The usual t statistic of the hypothesis $\rho = 1$ does not need to be adjusted for sample size or serial correlation and has the same asymptotic distribution as the case without serial correlation.

Augmented Dickey-Fuller Tests for Unit Roots

A problem now arises in determining the optimal number of lags of the dependent variable. There are 2 ways

- use the frequency of the data to decide
- use information criteria

Example 2.2.

- Transformation of the data : $loggdp = log(gdp)$
- Quick $\rightarrow$ Series Statistics $\rightarrow$ Unit Root Test
  1. Test Type: Augmented Dickey–Fuller
  2. Test for Unit Root in : Levels
  3. Include in test equation : Intercept
  4. Maximum lags :

The Explosive Case

- The explosive case : $\rho > 1$
- Typically, the explosive case is ignored and we use $\rho = 1$ to characterize the non-stationarity because
  1. $\rho > 1$ does not describe many data series in economics and finance.
  2. $\rho > 1$ has an intuitively unappealing property: shocks to the system are not only persistent through time, they are propagated so that a given shock will have an increasingly large influence.

Testing for Higher Orders of Integration

- Consider the simple regression:

$$\triangle y_t = \tilde{\rho} y_{t-1} + \epsilon_t$$

  We test $H_0 : \tilde{\rho} = 0$ v.s. $H_1 : \tilde{\rho} < 0$.
- If H_0 is rejected we simply conclude that y_t does not contain a unit root.

Testing for Higher Orders of Integration

- But what do we conclude if H_0 is not rejected? The series contains a unit root, but is that it? No! What if $y_t \sim I(2)$? We would still not have rejected. So we now need to test $H_0 : y_t \sim I(2)$ v.s. $H_1 : y_t \sim I(1)$. We would continue to test for a further unit root until we rejected H_0.
- We now regress $\triangle^2 y_t$ on $\triangle y_{t-1}$ (plus lags of $\triangle^2 y_t$ if necessary)
- Now we test $H_0 : \triangle y_t \sim I(1)$ which is equivalent to $H_0 : y_t \sim I(2)$.
- So in this case, if we do not reject (unlikely), we conclude that y_t is at least I(2).

Co-integration

- The VAR models discussed so far are appropriate for modeling I(0) data, like asset returns or growth rates of macroeconomic time series.
- Economic theory, however, often implies equilibrium relationships between the levels of time series variables that are best described as being $I(1)$.
- Similarly, arbitrage arguments imply that the $I(1)$ prices of certain financial time series are linked.
- The statistical concept of cointegration is required to make sense of regression models and VAR models with $I(1)$ data.

Co-integration

Basic ideas :

- x_{1t} and x_{2t} are unit-root nonstationary
- a linear combination of x_{1t} and x_{2t} is unit-root stationary

That is, x_{1t} and x_{2t} share a single unit root!

Co-integration

- Let y_t denote a $(k \times 1)$ vector of $I(1)$ time series. y_t is co-integrated if there exists a $(k \times 1)$ vector β_0 such that $\beta_0' y_t \sim I(0)$.
- In words, the nonstationary time series in y_t are cointegrated if there is a linear combination of them that is stationary or $I(0)$.
- The linear combination $\beta_0' y_t$ is often motivated by economic theory and referred to as a long-run equilibrium relationship.
- Intuition: $I(1)$ time series with a long-run equilibrium relationship cannot drift too far apart from the equilibrium because economic forces will act to restore the equilibrium relationship.

Normalization

- The cointegration vector β_0 is not unique since for any scalar c, $c\beta_0' y_t = \beta^{*'} y_t \sim I(0)$
- Some normalization assumption is required to uniquely identify β_0. A typical normalization is $\beta_0 = (1, -\beta_{0,2}, \cdots, -\beta_{0,k})'$, so that $\beta_0' y_t = y_{t,1} - \beta_{0,2} y_{t,2} - \cdots - \beta_{0,k} y_{t,k} \sim I(0)$.
- Or $y_{t,1} = \beta_{0,2} y_{t,2} + \cdots + \beta_{0,k} y_{t,k} + u_t$
  $u_t \sim I(0) =$ cointegrating residual
- In long-run equilibrium, $u_t = 0$ and the long-run equilibrium relationship is $y_{t,1} = \beta_{0,2} y_{t,2} + \cdots + \beta_{0,k} y_{t,k}$

Multiple Cointegrating Relationships

- If the $(k \times 1)$ vector y_t is cointegrated there may be $0 < m < k$ linearly independent cointegrating vectors $\{\beta_1, \cdots, \beta_m\}$.
- $\beta_i' y_t \sim I(0)$ for $i = 1, \cdots, m$.
- The $(k \times m)$ matrix $\beta = (\beta_1, \cdots, \beta_m)$ forms a basis for the space of cointegrating vectors.
- Any linear combination of $\{\beta_1, \cdots, \beta_m\}$ is also a cointegrating vector.

Examples of Cointegration and Common Trends in Economics and Finance

Cointegration naturally arises in economics and finance. In economics, cointegration is most often associated with economic theories that imply equilibrium relationships between time series variables:

- The permanent income model implies cointegration between consumption and income, with consumption being the common trend.
- Growth theory models imply cointegration between income, consumption and investment, with productivity being the common trend.
- Purchasing power parity implies cointegration between the nominal exchange rate and foreign and domestic prices.
- The Fisher equation implies cointegration between nominal interest rates and inflation ...

Examples of Cointegration and Common Trends in Economics and Finance

In finance, cointegration may be a high frequency relationship or a low frequency relationship. Cointegration at a high frequency is motivated by arbitrage arguments.

- The Law of One Price implies that identical assets must sell for the same price to avoid arbitrage opportunities. This implies cointegration between the prices of the same asset trading on different markets, for example.
- Similar arbitrage arguments imply cointegration between spot and futures prices, and spot and forward prices, and bid and ask prices.

Co-integration

Why is it of interest?

- Stationary series is mean reverting.
- Long term forecasts of the "linear" combination converge to a mean value, implying that the long-term forecasts of x_{1t} and x_{2t} must be linearly related.

VAR Containing Unit Roots

- Assume each component series contains only one unit root.
- Let y_t be an $(k \times 1)$ vector satisfying

$$(I_k - \Phi_1 L - \Phi_2 L^2 - \cdots - \Phi_p L^p) y_t = \mu_t + \epsilon_t$$

where Φ_s denotes an $(k \times k)$ matrix for $s = 1, 2, \cdots, p$ and μ_t and ϵ_t are $(k \times 1)$ vectors.

- As we did in the univariate case, we can rewrite the model as

$$\begin{aligned} \Delta y_t &= \eta_1 \Delta y_{t-1} + \cdots + \eta_{p-1} \Delta y_{t-p+1} + \mu_t + \epsilon_t \\ &+ \underbrace{[\Phi_1 + \Phi_2 + \cdots + \Phi_p - I_k]\, y_{t-1}}_{error\ \ correction\ \ term} \end{aligned}$$

where $\eta_j \equiv -[\Phi_{j+1} + \Phi_{j+2} + \cdots + \Phi_p]$ for $j = 1, 2, \cdots, p-1$.

Case 1 : No cointegration

- The first difference of y follows a stationary VAR (p-1) process (possible with deterministic trend μ_t)

$$\Delta y_t = \eta_1 \Delta y_{t-1} + \cdots + \eta_{p-1} \Delta y_{t-p+1} + \mu_t + \epsilon_t \qquad (3)$$

or $\Phi_1 + \Phi_2 + \cdots + \Phi_p - I_k = zero\ matrix$
or No cointegration.

- The characteristic equation of the system

$$\begin{aligned}
&|I_k - \Phi_1 z - \Phi_2 z^2 - \cdots - \Phi_p z^p| = 0 \\
&|\Phi_1 + \Phi_2 + \cdots + \Phi_p - \Phi_1 z - \Phi_2 z^2 - \cdots - \Phi_p z^p| = 0 \\
&|\Phi_1(1-z) + \Phi_2(1-z^2) + \cdots + \Phi_p(1-z^p)| = 0 \\
&(1-z)^k |\Phi_1 + \Phi_2(1+z) + \cdots + \Phi_p(1+z+\cdots+z^{p-1})| = 0
\end{aligned}$$

This system contains exactly k unit roots.

Case 2 : Cointegration

- $\Pi \equiv \Phi_1 + \Phi_2 + \cdots + \Phi_p - I_k$ is singular but nonzero, say $rank(\Pi) = m < k$.
- $|I_k - \Phi_1 z - \Phi_2 z^2 - \cdots - \Phi_p z^p| = 0$ contains $r = k - m$ unit roots.

Case 2 : a bit linear algebra (optional)

- It is equivalent to consider the eigenvalues of Φ^* in equation (8.15) (Tsay, pp.404)
- $\Pi \equiv \Phi_1 + \Phi_2 + \cdots + \Phi_p - I_k$ being singular implies **unity** is the eigenvalue of Φ^*.
- We need to determine the multiplicity of the eigenvalue 1, i.e. the number of linearly independent eigenvectors with the eigenvalue 1.

$$I_{kp} - \Phi^* = \begin{bmatrix} I_k & -I_k & 0 & 0 & \cdots & 0 \\ 0 & I_k & -I_k & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 & I_k & -I_k \\ -\Phi_p & -\Phi_{p-1} & -\Phi_{p-2} & \cdots & -\Phi_2 & I_k - \Phi_1 \end{bmatrix}$$

Case 2 : a bit linear algebra (optional)

- We calculate the rank of $I_{kp} - \Phi^*$ by using elementary column operations. The following matrix has the same rank as $I_{kp} - \Phi^*$

$$\begin{bmatrix} I_k & -I_k & 0 & 0 & \cdots & 0 \\ 0 & I_k & -I_k & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & I_k & -I_k & 0 \\ 0 & 0 & \cdots & 0 & I_k & 0 \\ -\Phi_p & -\Phi_{p-1} & -\Phi_{p-2} & \cdots & -\Phi_2 & -\Pi \end{bmatrix}$$

- $rank(I_{kp} - \Phi^*) = k \times (p-1) + m.$
- the multiplicity of the eigenvalue unity $= kp - rank(I_{kp} - \Phi^*) = k - m.$

Case 2 : Cointegration

- y_t contains $k-m$ unit roots that give $k-m$ common stochastic trends.
- There are **m** linear combinations of y_t that are unit-root stationary.
- $\Pi = \alpha\beta'$, where α and β are $k \times m$ matrices with full column rank.
- $Z_t = \beta' y_t$ is unit-root stationary. $\beta = [\beta_1, \beta_2, \cdots, \beta_m]$ are **m** linearly independent cointegrating vectors.
- $\Delta y_t = \eta_1 \Delta y_{t-1} + \cdots + \eta_{p-1} \Delta y_{t-p+1} + \mu_t + \epsilon_t + \alpha Z_{t-1}$, where α are adjustment parameters in VEC model.

Case 2 : Cointegration

- For a cointegrated system, you can not apply VAR model directly to the first difference process.
- Overdifferencing
- $Z_{t-1} = \beta' y_{t-1}$ is a "compensation" term for the overdifferenced system Δy_t.

Case 3: Stationary System

- $rank(\Pi) = k$
- 1 does not solve $|I_k - \Phi_1 z - \Phi_2 z^2 - \cdots - \Phi_p z^p| = 0$
- y_t contains no unit roots
- The ECM model is not informative and one studies y_t directly

Cointegration Test Specification

The limiting distributions of cointegration tests depend on the deterministic function μ_t

- Case 1 $\mu_t = 0$: In this case, all the component series of y_t are $I(1)$ without drift and the stationary series $Z_t = \beta' y_t$ has mean zero.
- Case 2 $\mu_t = \mu_0 = \alpha c_0$, where c_0 is an m-dimensional nonzero constant vector. The ECM becomes

$$\Delta y_t = \alpha(\beta' y_{t-1} + c_0) + \eta_1 \Delta y_{t-1} + \cdots + \eta_{p-1} \Delta y_{t-p+1} + \epsilon_t$$

so that the components of y_t are $I(1)$ without drift, but Z_t have a nonzero mean $-c_0$. This is referred to as the case of restricted constant.

Cointegration Test Specification

- Case 3 $\mu_t = \mu_0$, which is nonzero. Here the component series of y_t are $I(1)$ with drift μ_0 and Z_t may have a nonzero mean.
- Case 4 $\mu_t = \mu_0 + \alpha c_1 t$, where c_1 is an m-dimensional nonzero constant vector. The ECM becomes

$$\Delta y_t = \mu_0 + \alpha(\beta' y_{t-1} + c_1 t) + \eta_1 \Delta y_{t-1} + \cdots + \eta_{p-1} \Delta y_{t-p+1} + \epsilon_t$$

  so that the components of y_t are $I(1)$ with drift μ_0 and Z_t has a linear time trend related to $c_1 t$. This is the case of restricted trend.
- Case 5 (not common in empirical work) $\mu_t = \mu_0 + \mu_1 t$, where μ_i are nonzero. Here both the constant and trend are unrestricted. The components of y_t are $I(1)$ and have a quadratic time trend and Z_t have a linear trend.

Johansen's test

- Originally considered by Johansen (1988), "Statistical Analysis of Cointegration Vectors," Journal of Economics Dynamics and Control.
- He developed a sophisticated sequential procedure for determining the existence of cointegration and for determining the number of cointegrating relationships based on maximum likelihood techniques.

Johansen's test

To test for co-integration:

- Fit the model in

$$\Delta y_t = \Pi y_{t-1} + \eta_1 \Delta y_{t-1} + \cdots + \eta_{p-1} \Delta y_{t-p+1} + \mu_t + \epsilon_t$$

- Test for the rank of Π : $rank(\Pi) =$ the number of cointegrating relations
- The rank of a matrix is equal to the number of its characteristic roots (eigenvalues) that are different from zero
- The test for cointegration between y_t is calculated by looking at the rank of the Π matrix via its eigenvalues

Johansen's test

- The eigenvalues denoted λ_i are put in order:

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{k-1} \geq \lambda_k$$

- If y_t are not cointegrated, the rank of Π will not be significantly different from zero, so $\lambda_i = 0 \ \forall i$
  Then if $\lambda_i = 0$, $\log(1 - \lambda_i) = 0$
- Say $rank(\Pi) = 1$, then $\log(1 - \lambda_1)$ will be negative and $\log(1 - \lambda_i) = 0$ for $i \geq 2$
- If Π is of full rank, then $\log(1 - \lambda_i) < 0 \ \forall i$

Johansen's test : trace cointegration tests

A sequential procedure to determine the number of cointegrating vectors.

- $LR_{trace}(m) = -(T-p)\sum_{i=m+1}^{k}\log(1-\hat{\lambda}_i)$, where $\hat{\lambda}_i$ is the estimated value for the *ith* ordered eigenvalue from the Π matrix.
- $LR_{trace}(m)$ is used to test $H_0 : rank(\Pi) = m$ against $H_a : rank(\Pi) > m$ for $m = 0,\ 1,\ 2,\ \cdots, k-1$
- Due to the presence of unit roots, the asymptotic distribution of the test statistics is a function of standard Brownian motions.
- The critical values depend on :
    1. the value of $k-m$, the number of non-stationary components
    2. the specification of μ_t

Johansen's test : trace cointegration tests

The testing sequence under the null is $m = 0,\ 1,\ 2,\ \cdots, k-1$ so that the hypotheses for trace cointegration tests are

$$\begin{array}{rcl}
H_0 : rank(\Pi) = 0 & v.s. & H_a : 0 < rank(\Pi) \leq k \\
H_0 : rank(\Pi) = 1 & v.s. & H_a : 1 < rank(\Pi) \leq k \\
H_0 : rank(\Pi) = 2 & v.s. & H_a : 2 < rank(\Pi) \leq k \\
\dots & \dots & \dots \\
H_0 : rank(\Pi) = k-1 & v.s. & H_a : rank(\Pi) = k
\end{array}$$

We keep increasing the value of m until we no longer reject the null.

Johansen's test : maximum eigenvalue tests

A sequential procedure to determine the number of cointegrating vectors.

- $LR_{max}(m) = -(T - p)\log(1 - \hat{\lambda}_{m+1})$
- $LR_{max}(m)$ is used to test $H_0 : rank(\Pi) = m$ against $H_a : rank(\Pi) = m + 1$ for $m = 0,\ 1,\ 2,\ \cdots, k - 1$
- Critical values of the test statistics are nonstandard and depend on the specification of μ_t. And it must be evaluated via simulation.

Johansen's test : maximum eigenvalue tests

The testing sequence under the null is $m = 0,\ 1,\ 2,\ \cdots, k-1$ so that the hypotheses for maximum eigenvalue tests are

$$
\begin{array}{rcl}
H_0 : rank(\Pi) = 0 & v.s. & H_a : rank(\Pi) = 1 \\
H_0 : rank(\Pi) = 1 & v.s. & H_a : rank(\Pi) = 2 \\
H_0 : rank(\Pi) = 2 & v.s. & H_a : rank(\Pi) = 3 \\
\cdots & \cdots & \cdots \\
H_0 : rank(\Pi) = k-1 & v.s. & H_a : rank(\Pi) = k
\end{array}
$$

We keep increasing the value of m until we no longer reject the null.

Johansen's test : remarks

- The critical values depend on the trend assumptions and may not be appropriate for models that contain other deterministic regressors.
- The trace statistic and the maximum eigenvalue statistic may yield conflicting results. For such cases, we recommend that you examine the estimated cointegrating vector and base your choice on the interpretability of the cointegrating relations
- In some cases, the individual unit root tests will show that some of the series are integrated, but the cointegration test will indicate that the matrix has full rank ($m = k$). This apparent contradiction may be the result of low power of the cointegration tests, stemming perhaps from a small sample size or serving as an indication of specification error.

Error Correction Model

The general formulation

$$\Delta y_t = \mu_0 + \Pi y_{t-1} + \eta_1 \Delta y_{t-1} + \cdots + \eta_{p-1} \Delta y_{t-p+1} + \epsilon_t$$

Consider a trivariate $I(1)$ vector $y_t = (y_{1t}, y_{2t}, y_{3t})'$ and assume that y_t is cointegrated with cointegrating vectors $(1, 0, -\beta_1)'$ and $(0, 1, -\beta_2)'$

$$\begin{aligned} \Delta y_{1t} &= \mu_{01} + \alpha_{11}(y_{1t-1} - \beta_1 y_{3t-1}) + \alpha_{12}(y_{2t-1} - \beta_2 y_{3t-1}) \\ &+ \sum_{j=1}^{p-1} \eta_{11}^{j} \Delta y_{1t-j} + \sum_{j=1}^{p-1} \eta_{12}^{j} \Delta y_{2t-j} + \sum_{j=1}^{p-1} \eta_{13}^{j} \Delta y_{3t-j} + \epsilon_{1t} \end{aligned}$$

Error Correction Model

$$\begin{aligned}\Delta y_{2t} &= \mu_{02} + \alpha_{21}(y_{1t-1} - \beta_1 y_{3t-1}) + \alpha_{22}(y_{2t-1} - \beta_2 y_{3t-1}) \\ &+ \sum_{j=1}^{p-1} \eta_{21}^j \Delta y_{1t-j} + \sum_{j=1}^{p-1} \eta_{22}^j \Delta y_{2t-j} + \sum_{j=1}^{p-1} \eta_{23}^j \Delta y_{3t-j} + \epsilon_{2t}\end{aligned}$$

$$\begin{aligned}\Delta y_{3t} &= \mu_{03} + \alpha_{31}(y_{1t-1} - \beta_1 y_{3t-1}) + \alpha_{32}(y_{2t-1} - \beta_2 y_{3t-1}) \\ &+ \sum_{j=1}^{p-1} \eta_{31}^j \Delta y_{1t-j} + \sum_{j=1}^{p-1} \eta_{32}^j \Delta y_{2t-j} + \sum_{j=1}^{p-1} \eta_{33}^j \Delta y_{3t-j} + \epsilon_{3t}\end{aligned}$$

Error Correction Model

- The ECM links the long-run equilibrium relationship implied by cointegration with the short-run dynamic adjustment mechanism that describes how the variables react when they move out of long-run equilibrium.
- Estimation : MLE
- Forecasting

Example : Section 8.6.5

Two weekly U.S. short-term interest rates : the 3-month Treasury bill (TB) rate and 6-month Treasury bill rate

- The ADF tests fail to reject the hypothesis of a unit root in the individual series [AR(3) used]
- VAR modeling : BIC $\Rightarrow$ VAR(3)
- Cointegration Test :
  Quick $\rightarrow$ Group Statistics $\rightarrow$ Cointegration Test
  Cointegration Test Specification :
  1. a restricted constant for μ_t
  2. Lag intervals 1 2. why? (Lag specification for first differenced process Δy_t)

$\Rightarrow$ One cointegrating relation

Example : Section 8.6.5

- Maximum Likelihood Estimation : Cointegrated VAR(3) model using an ECM presentation
  Quick → Estimate VAR
  VAR Specification:
  1. VAR Type : Vector Error Correction
  2. Endogenous Variables : 3m 6m
  3. Lag Intervals for D : 1 2.
  4. Number of cointegrating :1
  5. a restricted constant for μ_t
- Forecasting:
  1. Workfile: Proc → Resize 2393
  2. Var : Proc → Make Model → Solve
     Solution sample : 2384 2393

Discussion

- ECM formulation is useful
- Co-integration tests have some weaknesses, e.g. robustness
- Co-integration overlooks the effect of scale of the series

Assignment 3

- Tsay :
  pp 107 : 2.15.
  pp 463 : 8.7. (a) (c) (d) (e) (f) (g)
- Brooks :
  pp 376 : Q1, Q2, Q3
- Due Oct. 26 in class