Lecture 9
Co-integration
and Error Correction Models

CHEN Ying
FE5209 Financial Econometrics

Outline

- Co-integration.
- Error correction models.
- Cointegration tests.
- Trading strategy.

Readings

FE chapter 11
SDA chapter 15
MTS chapter 5

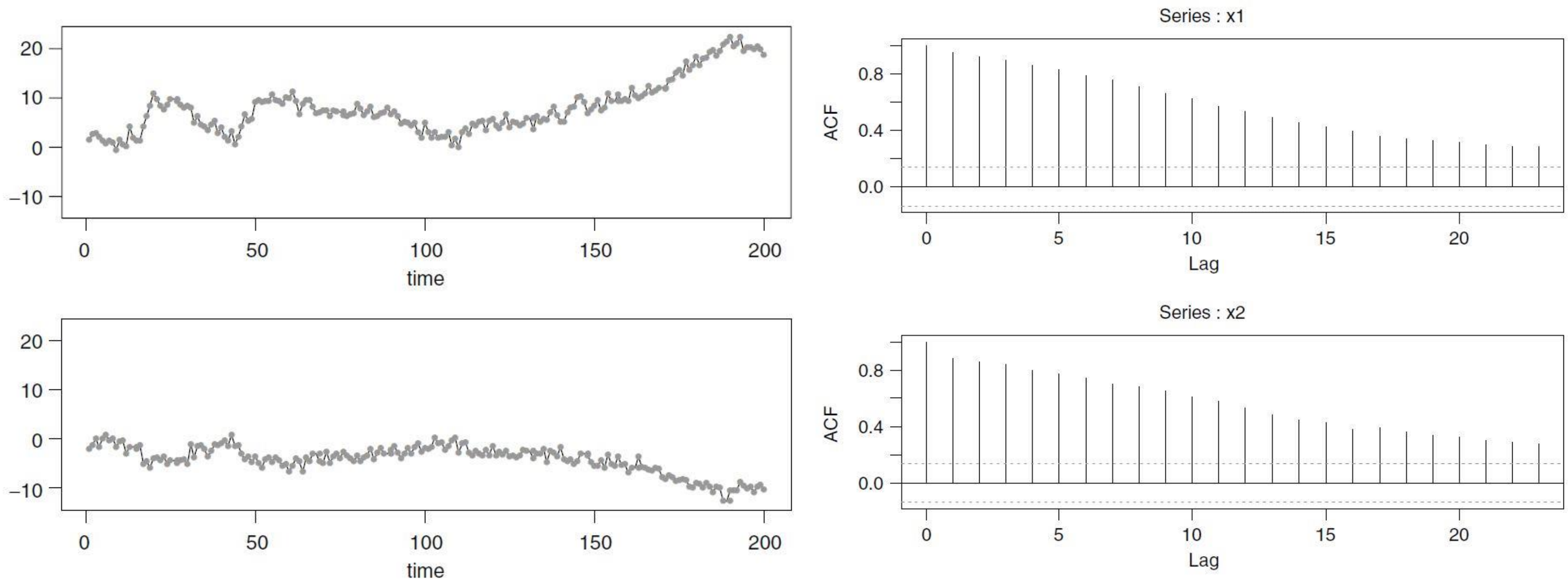
MTS: Multivariate Time Series Analysis with R and Financial Applications by Ruey Tsay

Unit root nonstationarity and cointegration

Consider the bivariate VARMA(1,1) model:

$$\begin{bmatrix} Y_{1t} \\ Y_{2t} \end{bmatrix} - \begin{bmatrix} 0.5 & -1.0 \\ -0.25 & 0.5 \end{bmatrix} \begin{bmatrix} Y_{1,t-1} \\ Y_{2,t-1} \end{bmatrix} = \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} - \begin{bmatrix} 0.2 & -0.4 \\ -0.1 & 0.2 \end{bmatrix} \begin{bmatrix} \varepsilon_{1,t-1} \\ \varepsilon_{2,t-1} \end{bmatrix}$$

where the covariance matrix $\mathbf{\Sigma}$ of the shock $\boldsymbol{\varepsilon}_t$ is positive definite.



It is easy to see from the simulation that the two series have high autocorrelations and exhibit features of unit-root nonstationarity. The two marginal models of $\boldsymbol{Y}_t$ are indeed unit-root nonstationary.

Unit root nonstationarity and cointegration

Rewrite the model as:

$$\begin{bmatrix} 1-0.5B & B \\ 0.25B & 1-0.5B \end{bmatrix}\begin{bmatrix} Y_{1t} \\ Y_{2t} \end{bmatrix} = \begin{bmatrix} 1-0.2B & 0.4B \\ 0.1B & 1-0.2B \end{bmatrix}\begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}.$$

Pre-multiplying the above equation by

$$\begin{bmatrix} 1-0.5B & -B \\ -0.25B & 1-0.5B \end{bmatrix}$$

we obtain the result

$$\begin{bmatrix} 1-B & 0 \\ 0 & 1-B \end{bmatrix}\begin{bmatrix} Y_{1t} \\ Y_{2t} \end{bmatrix} = \begin{bmatrix} 1-0.7B & -0.6B \\ -0.15B & 1-0.7B \end{bmatrix}\begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}.$$

This is not a weakly stationary model because the roots of the AR polynomial are 0 and 1. Y_{it} is unit root series that follows ARIMA(0,1,1) process.

Co-integration

Now consider a linear transformation by defining:

$$\boldsymbol{X}_t = \begin{bmatrix} X_{1t} \\ X_{2t} \end{bmatrix} = \begin{bmatrix} 1.0 & -2.0 \\ 0.5 & 1.0 \end{bmatrix} \begin{bmatrix} Y_{1,t} \\ Y_{2,t} \end{bmatrix} = \boldsymbol{L}\boldsymbol{Y}_t$$

$$\begin{bmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{bmatrix} = \begin{bmatrix} 1.0 & -2.0 \\ 0.5 & 1.0 \end{bmatrix} \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix} = \boldsymbol{L}\boldsymbol{\varepsilon}_t$$

The model for $\boldsymbol{Y}_t$: $\boldsymbol{L} \begin{bmatrix} Y_{1t} \\ Y_{2t} \end{bmatrix} - \boldsymbol{L} \begin{bmatrix} 0.5 & -1.0 \\ -0.25 & 0.5 \end{bmatrix} \boldsymbol{L}^{-1}\boldsymbol{L} \begin{bmatrix} Y_{1,t-1} \\ Y_{2,t-1} \end{bmatrix} = \boldsymbol{L} \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} -$
$\boldsymbol{L} \begin{bmatrix} 0.2 & -0.4 \\ -0.1 & 0.2 \end{bmatrix} \boldsymbol{L}^{-1}\boldsymbol{L} \begin{bmatrix} \varepsilon_{1,t-1} \\ \varepsilon_{2,t-1} \end{bmatrix}$

can be represented as:

$$\begin{bmatrix} X_{1t} \\ X_{2t} \end{bmatrix} - \begin{bmatrix} 1.0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} X_{1,t-1} \\ X_{2,t-1} \end{bmatrix} = \begin{bmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{bmatrix} - \begin{bmatrix} 0.4 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \epsilon_{1,t-1} \\ \epsilon_{2,t-1} \end{bmatrix}$$

(a) X_{1t} and X_{2t} are uncoupled series with concurrent correlation equal to that between the shocks ϵ_{1t} and ϵ_{2t},

(b) X_{1t} follows a univariate ARIMA(0,1,1) model, and there is only one single unit root in the system of Y. Moreover, the unit roots of Y_{1t} and Y_{2t} are introduced by the unit root of X_{1t}.

(c) X_{2t} is a white noise series. $X_{2t} = 0.5\ Y_{1t} + Y_{2t}$ does not have a unit root.

Co-integration

Co-integration:
Both Y_{1t} and Y_{2t} are unit-root nonstationary. But there is only one single unit root in the vector series.

Remark:
A linear combination of Y_{1t} and Y_{2t} is unit-root stationary.
Y_{1t} and Y_{2t} share a common unit root!

The vector $(0.5,1.0)'$ is a **cointegrating vector** for the system. It gives stationary series.

Why is co-integration of interest?
Because stationary series is mean reverting. Thus, long term forecasts of the ``linear'' combination converge to a mean value, implying that the long-term forecasts of Y_{1t} and Y_{2t} must be linearly related.

Cointegration: Three Key Features

First, consider ***reduction of order of integration***, two or more stochastic processes that are integrated of order one or higher are said to be cointegrated if there are linear combinations of the processes with a ***lower*** order of integration.

Second, the concept of cointegration can be also stated in terms of ***linear regression***. Two or more processes integrated of order one are said to be cointegrated if it is possible to make a meaningful linear regression of one process on the other(s).

Third, a property of cointegrated processes is the presence of integrated ***common trends***. Given n processes with r cointegrating relationships, it is possible to determine $\mathrm{d} - r$ common trends. Common trends are integrated processes such that any of the d original processes can be expressed as a linear regression on the common trends.

Common trends model

Two time series, Y_{1t} and Y_{2t}, are cointegrated, where each is *I*(1) but shares the common trends:

$$Y_{1t} = \beta_1 W_t + \epsilon_{1t} \text{ and } Y_{2t} = \beta_2 W_t + \epsilon_{2t}$$

where β_1 and β_2 are nonzero, the trend W_t common to both series is *I*(1), and the noise processes ϵ_{1t} and ϵ_{2t} are *I*(0).

Because of the common trend, Y_{1t} and Y_{2t} are nonstationary but there is a linear combination of these two series that is free of the trend so they are cointegrated.

If $\lambda = \beta_1/\beta_2$, then

$$\beta_2(Y_{1t} - \lambda Y_{2t}) = \beta_2 Y_{1t} - \beta_1 Y_{2t} = \beta_2 \epsilon_{1t} - \beta_1 \epsilon_{2t}$$

is free of the trend W_t and therefore is *I*(0).

The **Phillips-Ouliaris cointegration test** regresses one integrated series on others and applies the Phillips-Perron unit root test to the residuals.

H_0: The residuals are unit root nonstationary,

which implies that the series are *not* cointegrated.

Example: Phillips-Ouliaris test on bond yields

We consider three-month, six-month, one-year, two-year, and three-year bond yields recorded daily from January 2, 1990 to October 31, 2008, for a total of 4714 observations.

The five yields series track each other somewhat closely. This suggests that the five series may be cointegrated.

The one-year yields were regressed on the four others and the residuals and their ACF are also plotted in Figure 15.1. The two residual plots are ambiguous about whether the residuals are stationary, so a test of cointegration would be helpful.

R: 8_Phillips-OuliarisTest.R

Example: Phillips-Ouliaris test on bond yields

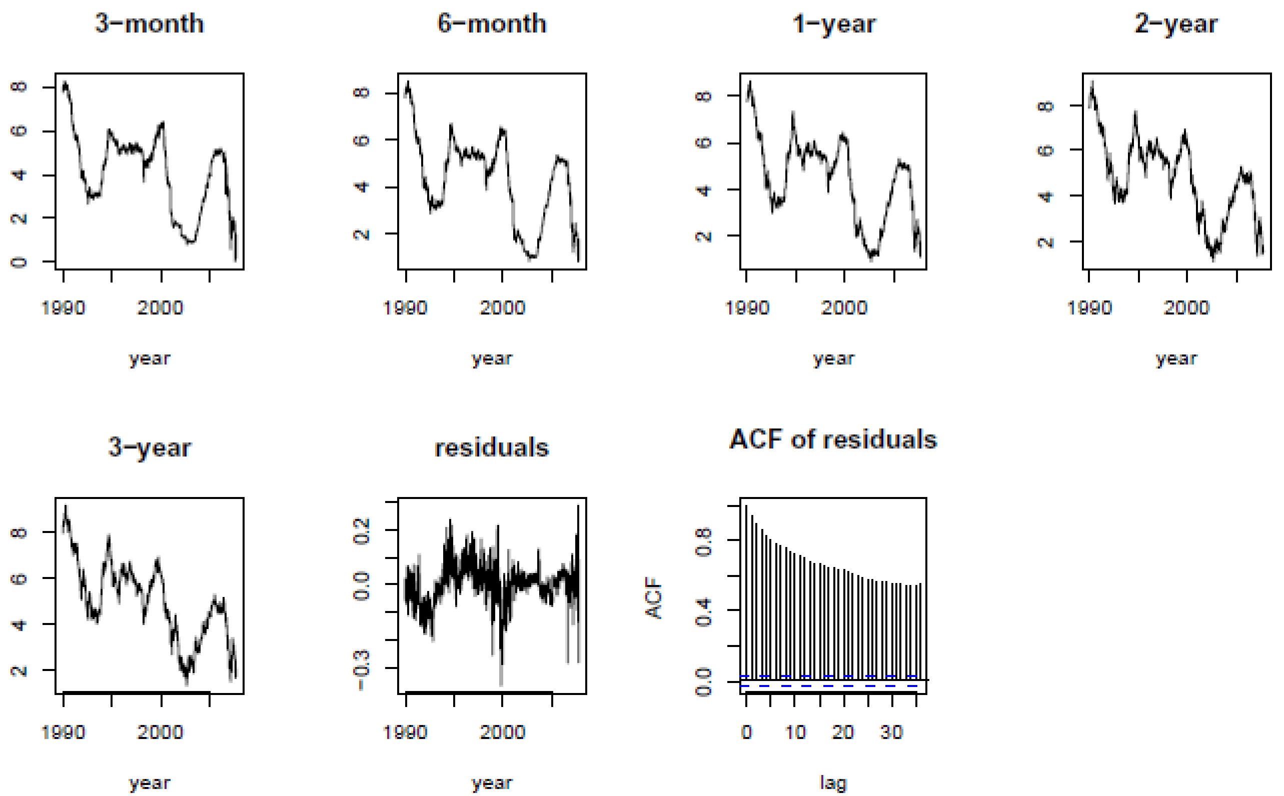


Fig. 15.1. *Time series plots of the five yields and the residuals from a regression of the 1-year yields on the other four yields. Also, a ACF plot of the residuals.*

R: 8_Phillips-OuliarisTest.R

Example: Phillips-Ouliaris test on bond yields

Next, the Phillips-Ouliaris test was run using the R function po.test in the tseries package.

Phillips-Ouliaris Cointegration Test
data: dat[, c(3, 1, 2, 4, 5)]
Phillips-Ouliaris demeaned = -323.546, Truncation lag
parameter = 47, p-value = 0.01
Warning message:
In po.test(dat[, c(3, 1, 2, 4, 5)]) : p-value smaller
than printed p-value

The small *p*-value leads to the conclusion that the residuals are stationary and so the five series are cointegrated. Though stationary, the residuals have a large amount of autocorrelation and may have long-term memory. They take a long time to revert to their mean of zero. Devising a profitable trading strategy from these yields seems problematic.

R: 8_Phillips-OuliarisTest.R

Vector Error Correction Models

The regression approach to co-integration is unsatisfactory, since one series must be chosen as the dependent variable, and this choice must be somewhat arbitrary. Moreover, regression will find only one co-integrating vector, but there could be more than one.

An alternative approach to cointegration that treats the series symmetrically uses a ***vector error correction model*** **(VECM).**

- "Error": the deviation from the mean
- "Correction": whenever the stationary linear combination deviates from its mean, then it is pushed back toward its mean (the error is corrected).

Vector Error Correction Models

Let $\boldsymbol{Y_t} = \left(Y_{1,t}, Y_{2,t}\right)^T$ and $\boldsymbol{\epsilon_t} = \left(\epsilon_{1,t}, \epsilon_{2,t}\right)^T$.

$$\Delta Y_{1t} = \phi_1(Y_{1,t-1} - \lambda Y_{2,t-1}) + \epsilon_{1t}$$
$$\Delta Y_{2t} = \phi_2(Y_{1,t-1} - \lambda Y_{2,t-1}) + \epsilon_{2t}$$

where ϵ_{1t} and ϵ_{2t} are white noise.

Rearranging gives

$$\Delta(Y_{1t} - \lambda Y_{2t}) = (\phi_1 - \lambda\phi_2)\left(Y_{1,t-1} - \lambda Y_{2,t-1}\right) + (\epsilon_{1t} - \lambda\epsilon_{2t}).$$

If $\phi_1 - \lambda\phi_2 < 0$, then the expected change has an opposite sign to $Y_{1,t-1} - \lambda Y_{2,t-1}$, this causes error correction.

$Y_{1,t-1} - \lambda Y_{2,t-1}$ is an AR(1) process with coefficient $1 + \phi_1 - \lambda\phi_2$

- If $\phi_1 - \lambda\phi_2 > 0$, then $1 + \phi_1 - \lambda\phi_2 > 1$ and $Y_{1,t-1} - \lambda Y_{2,t-1}$ is explosive.
- If $\phi_1 - \lambda\phi_2 = 0$, then $1 + \phi_1 - \lambda\phi_2 = 1$ and $Y_{1,t-1} - \lambda Y_{2,t-1}$ is random walk.
- If $-2 < \phi_1 - \lambda\phi_2 < 0$, then $-1 < 1 + \phi_1 - \lambda\phi_2 < 1$ and $Y_{1,t-1} - \lambda Y_{2,t-1}$ is stationary.

Vector Error Correction Models

The VECM for $\boldsymbol{Y_t} = \left(Y_{1,t}, Y_{2,t}\right)^T$:

$$\Delta Y_{1t} = \phi_1(Y_{1,t-1} - \lambda Y_{2,t-1}) + \epsilon_{1t}$$
$$\Delta Y_{2t} = \phi_2(Y_{1,t-1} - \lambda Y_{2,t-1}) + \epsilon_{2t}$$

$$\boldsymbol{\Delta Y_t = \alpha \beta^T Y_{t-1} + \epsilon_t}$$

where

$$\boldsymbol{\alpha} = \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} \; and \, \boldsymbol{\beta} = \begin{pmatrix} 1 \\ -\lambda \end{pmatrix}$$

- $\boldsymbol{\beta}$ is the co-integrating vector and
- $\boldsymbol{\alpha}$ specifies the speed of mean-reversion and is called the ***loading matrix*** or ***adjustment matrix***.

This VECM also applies for d series $\boldsymbol{Y_t} = (Y_{1t}, \cdots, Y_{dt})^T$ and higher order VAR

$$\boldsymbol{\Delta Y_t = \Gamma_1 \Delta Y_{t-1} + \cdots + \Gamma_{p-1} \Delta Y_{t-p+1} + \Pi Y_{t-1} + \mu + \Phi D_t + \epsilon_t}$$

where $\boldsymbol{\mu}$ is a mean vector, $\boldsymbol{D_t}$ is a vector of nonstochastic regressors, and

$$\boldsymbol{\Pi = \alpha \beta^T}$$

As before, $\boldsymbol{\beta}$ is the co-integrating vectors and $\boldsymbol{\alpha}$ is called the loading matrix. They are both full-rank $d \times m$ matrices.

Example: Simulation of ECM

The parameters are $\phi_1 = 0.5$, $\phi_2 = 0.55$, and $\lambda = 1$. A total of 5000 observations was simulated, but, for visual clarity, only every 10th observation (500 observations) is plotted in Figure 15.2. Neither $Y_{1,t}$ nor $Y_{2,t}$ is stationary, but $Y_{1,t} - \lambda Y_{2,t}$ is stationary. Notice how closely $Y_{1,t}$ and $Y_{2,t}$ track each other.

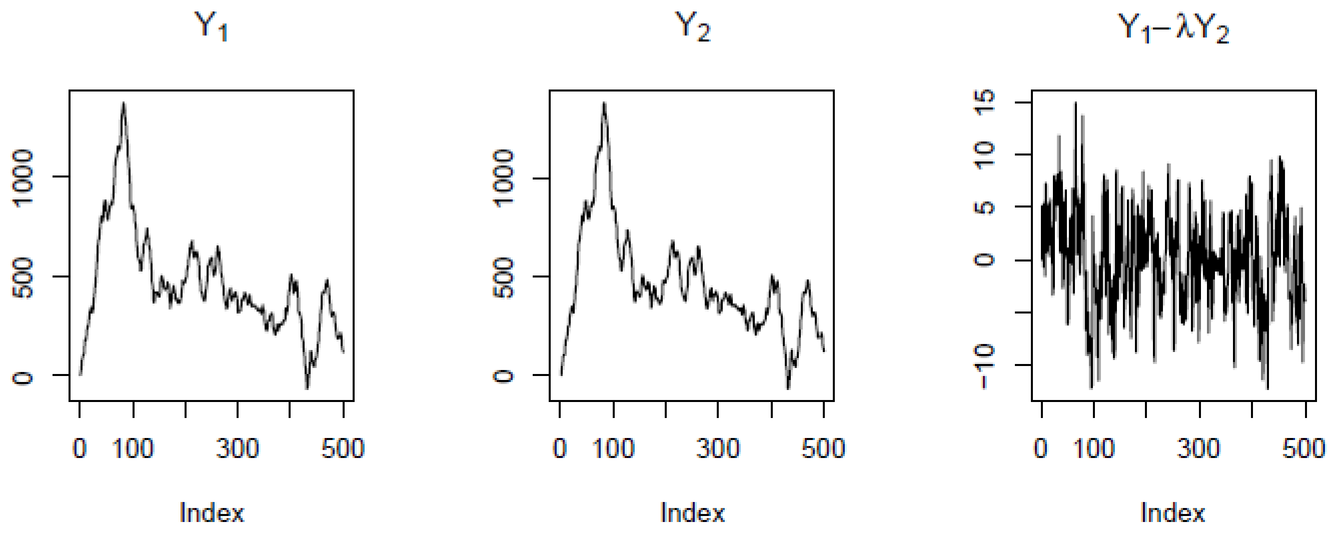


Fig. 15.2. *Simulation of an error correction model. 5000 observations were simulated but only every 10th is plotted.*

R: 8_SimulationErrorCorrectionModel.R $1 + \phi_1 - \lambda\phi_2$=1+0.5-0.55=0.95<1.

Review: Unit root test

Consider a **univariate** AR(1) model

$$Y_t = \phi_1 Y_{t-1} + \varepsilon_t.$$

Let $\Delta Y_t = Y_t - Y_{t-1}$. Substract Y_t from both sides and rearrange terms

$$\Delta Y_t = \gamma Y_{t-1} + \epsilon_t,$$

where $\gamma = \phi_1 - 1$ is associated with the AR polynomial with opposite sign $-\phi(L)$.
If there is unit root, $\gamma = -\phi(1) = 0$.

Consider a **univariate** AR(2) model

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t.$$

Let $\Delta Y_t = Y_t - Y_{t-1}$. Substract Y_t from both sides and rearrange terms

$$\Delta Y_t = \gamma Y_{t-1} + \phi_1^* \Delta Y_{t-1} + \epsilon_t,$$

where $\phi_1^* = -\phi_2$ and $\gamma = \phi_2 + \phi_1 - 1 = -\phi(1).$

Y_t is unit-root nonstationary if and only if $\gamma = 0$.

Testing for Y_t has a unit root is equivalent to testing for $\gamma = 0$ in the above model.

VECM

Vector Error Correction Model (VECM) for an VARMA(p,q) process:

$$\boldsymbol{Y}_t = \sum_{j=1}^{p} \boldsymbol{\phi}_j \boldsymbol{Y}_{t-j} + \boldsymbol{\varepsilon}_t - \sum_{j=1}^{q} \boldsymbol{\Theta}_j \boldsymbol{\varepsilon}_{t-j}$$

Let $\Delta \boldsymbol{Y}_t = \boldsymbol{Y}_t - \boldsymbol{Y}_{t-1}$ and rewrite the model (Engle and Granger 1987) as

$$\Delta \boldsymbol{Y}_t = \boldsymbol{\alpha}\boldsymbol{\beta}' \, \boldsymbol{Y}_{t-1} + \sum_{i=1}^{p-1} \boldsymbol{\phi}_i^* \Delta \boldsymbol{Y}_{t-i} + \boldsymbol{\varepsilon}_t - \sum_{j=1}^{q} \boldsymbol{\Theta}_j \boldsymbol{\varepsilon}_{t-j} \qquad (1)$$

- ❑ $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are $d \times m$ full-rank matrices for some $m \leq d$, with m denoting the number of linearly independent cointegration vectors in the system.
- ❑ $\boldsymbol{\beta}' \boldsymbol{Y}_t$ is unit-root stationary, with $\boldsymbol{\beta}$ as the co-integrating vector.
- ❑ In addition, we have

$$\boldsymbol{\phi}_j^* = -\sum_{i=j+1}^{p} \boldsymbol{\phi}_i, \qquad j = 1, \cdots, p-1.$$

$$\Pi = \boldsymbol{\alpha}\boldsymbol{\beta}' = \boldsymbol{\phi}_p + \cdots + \boldsymbol{\phi}_1 - \boldsymbol{I} = -\boldsymbol{\phi}(1).$$

VECM derivation

Consider the VARMA(p,q) model with drift

$$\boldsymbol{Y}_t = \sum_{j=1}^{p} \boldsymbol{\phi}_j \boldsymbol{Y}_{t-j} + c_0 + c_1 t + \boldsymbol{\varepsilon}_t - \sum_{j=1}^{q} \boldsymbol{\Theta}_j \boldsymbol{\varepsilon}_{t-j}$$

Write it as $\Phi(L)Y_t = c(t) + \Theta(L)\epsilon_t$, where $\Phi(L) = I_d - \Phi_1 L - \Phi_2 L^2 - \cdots - \Phi_p L^p$, $\mathrm{c_t} = c_0 + c_1 t$, $\Theta(L) = I_d - \Theta_{\,1} L - \Theta_2 L^2 - \cdots - \Theta_q L^q$ and $\boldsymbol{\varepsilon}_t$ are IID random vectors with mean zero and positive-definite covariance matrix.

$$\Delta \boldsymbol{Y}_t = \boldsymbol{\alpha\beta}' \ \boldsymbol{Y}_{t-1} + \sum_{i=1}^{p-1} \boldsymbol{\phi}_i^* \Delta \boldsymbol{Y}_{t-i} + \boldsymbol{\varepsilon}_t - \sum_{j=1}^{q} \boldsymbol{\Theta}_j \boldsymbol{\varepsilon}_{t-j}$$

To begin, we consider a special case of $p = 3$:

$$\Phi(L) = I_d - \Phi_1 \boldsymbol{L} - \Phi_2 L^2 - \Phi_3 L^3$$
$$= I_d - \Phi_1 \boldsymbol{L} - \Phi_2 L^2 - \Phi_3 L^2 + \Phi_3 L^2 - \Phi_3 L^3 = I_d - \Phi_1 L - (\Phi_2 + \Phi_3) L^2 + \Phi_3 L^2 (1 - L)$$
$$= I_d - \Phi_1 \boldsymbol{L} - (\Phi_2 + \Phi_3) L + (\Phi_2 + \Phi_3) L - (\Phi_2 + \Phi_3) L^2 + \Phi_3 L^2 (1 - L)$$
$$= I_d - (\Phi_1 + \Phi_2 + \Phi_3) \boldsymbol{L} + (\Phi_2 + \Phi_3) L (1 - L) + \Phi_3 L^2 (1 - L)$$

VECM derivation

Let $\Delta \boldsymbol{Y}_t = \boldsymbol{Y}_t - \boldsymbol{Y}_{t-1} = \boldsymbol{Y}_{\boldsymbol{t}}(\mathbf{1} - \boldsymbol{L})$ and define

$$\boldsymbol{\Pi} = I_d - (\Phi_1 + \Phi_2 + \Phi_3)L$$

$$\boldsymbol{\Phi}_2^* = -\Phi_3 = -\sum_{i=j+1}^{p} \boldsymbol{\phi}_i, \qquad j = 2,$$

$$\boldsymbol{\Phi}_1^* = -(\Phi_2 + \Phi_3) = -\sum_{i=j+1}^{p} \boldsymbol{\phi}_i, \qquad j = 1.$$

It gives VECM:

$$\Delta \boldsymbol{Y}_t = \Pi\, \boldsymbol{Y}_{t-1} + \sum_{i=1}^{p-1} \boldsymbol{\phi}_i^* \Delta \boldsymbol{Y}_{t-i} + c(t) + \boldsymbol{\varepsilon}_t - \sum_{j=1}^{q} \boldsymbol{\Theta}_j \boldsymbol{\varepsilon}_{t-j}$$

Clearly these transformations can be immediately generalized to any number of lags, they are simple rearrangements of terms which are always possible.

Co-integration is expressed as restrictions on the matrix $\boldsymbol{\Pi}$.

Co-integration test

If $\boldsymbol{Y}_t$ contains unit roots, then $|\boldsymbol{\Phi}(1)| = \mathbf{0}$ and $|\Pi| = |-\boldsymbol{\Phi}(1)| = \mathbf{0}$. However $|\Pi| = 0$ is not equal $|\boldsymbol{\Phi}(1)| = \mathbf{0}$. Nevertheless $\Pi = 0$ implies that $\boldsymbol{\Phi}_1 + \cdots + \boldsymbol{\Phi}_p = \boldsymbol{I_d}$ and $|\boldsymbol{\Phi}(1)| = 0$.

Let m be the rank of Π:

- ❑ $Rank(\Pi) = 0$: This implies that $\Pi = 0$, there is no co-integrating vector. In this case, $\boldsymbol{Y}_t$ has d unit roots and we can work directly on the differenced series:

$$\Delta\boldsymbol{Y}_t = \boldsymbol{\Pi}\boldsymbol{Y}_{t-1} + \sum_{i=1}^{p-1} \boldsymbol{\phi}_i^* \Delta\boldsymbol{Y}_{t-i} + \boldsymbol{\varepsilon}_t - \sum_{j=1}^{q} \boldsymbol{\Theta}_j \boldsymbol{\varepsilon}_{t-j},$$

  $\Delta\boldsymbol{Y}_t$ follows a VAR(p-1) model.

- ❑ $Rank(\Pi) = d$: $|\boldsymbol{\Phi}(1)| \neq \mathbf{0}$ and $\boldsymbol{Y}_t$ contains no unit root. $\boldsymbol{Y}_t$ is $I(0)$ and one studies $\boldsymbol{Y}_t$ directly.

- ❑ $0 < Rank(\Pi) = m < d$: $\boldsymbol{Y}_t$ is co-integrated with m linearly independent cointegrating vectors $\boldsymbol{\beta}'\boldsymbol{Y}_t$. There are $d - m$ unit roots.

VECM formulation is useful ⇒ Johansen (1988)

- ❑ Co-integration tests have some weaknesses, e.g. robustness.
- ❑ Co-integration overlooks the effect of scale of the series.

Cointegration: Johansen methodology

Let's run the cointegration analysis using the Johansen methodology.

First, determining the number of cointegrating relationships, then specify the number of lags.

The Johansen methodology offers two tests for testing the number of cointegrating relationships: the trace test and the eigenvalue test.

- The ***trace test*** tests the null hypothesis that there are at most r cointegrating relationships. That is, rejecting the null means that there are more than r cointegrating relationships.
- The ***eigenvalue test*** tests the null hypothesis of r versus $r + 1$ cointegrating relationships. The test rejects the null hypothesis if the eigenvalue test statistic exceeds the respective critical value.

Example: VECM test on bond yields

A VECM was fit to the bond yields using R's ca.jo function. The output is below. The eigenvalues are used to test null hypotheses of the form $H_0: r \leq r_0$. The values of the test statistics and critical values (for 1%, 5%, and 10% level tests) are listed below the eigenvalues. The null hypothesis is rejected when the test statistic exceeds the critical level.

In this case, regardless of whether one uses a 1%, 5%, or 10% level test, one accepts that r is less than or equal to 3 but rejects that r is less than or equal to 2, so one concludes that r = 3

```
######################
# Johansen-Procedure #
######################

Test type: maximal eigenvalue statistic (lambda
with linear trend

Eigenvalues (lambda):
[1] 0.03436 0.02377 0.01470 0.00140 0.00055

Values of test statistic and critical values of

           test 10pct  5pct 1pct
r <= 4 |   2.59   6.5  8.18 11.6
r <= 3 |   6.62  12.9 14.90 19.2
r <= 2 |  69.77  18.9 21.07 25.8
r <= 1 | 113.36  24.8 27.14 32.1
r = 0  | 164.75  30.8 33.32 38.8
```

Parameterization of co-integrating vectors

Since $\Phi(1)$ is a singular matrix for a co-integrated system, $\Pi = -\Phi(1)$ is not full rank. Assume that $Rank(\Pi) = m > 0$. Then, there exist $d \times m$ matrices α and β of rank m such that $\Pi = \alpha\beta'$.

This decomposition, however, is not unique. In fact, for any $m \times m$ orthogonal matrix P such that $PP' = I$, it is easy to see that

$$\alpha\beta' = \alpha PP'\beta' = (\alpha P)(\beta P)'$$

Thus, αP and βP are also of rank m and may serve as another decomposition of Π. Consequently, some additional normalization is needed to uniquely identify α and β.

An approach considered by Phillips (1991) is to rewrite the cointegrating matrix as

$$\beta = \begin{bmatrix} I_m \\ \beta_1 \end{bmatrix}$$

where β_1 is an arbitrary $(d-m) \times m$ matrix. This parameterization can be achieved by reordering the components of Y_t if necessary and is similar to those commonly used in statistical factor models.

Parameterization of co-integrating vectors

For example, given $\begin{bmatrix} Y_{1t} \\ Y_{2t} \end{bmatrix} - \begin{bmatrix} 0.5 & -1.0 \\ -0.25 & 0.5 \end{bmatrix} \begin{bmatrix} Y_{1,t-1} \\ Y_{2,t-1} \end{bmatrix} = \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} - \begin{bmatrix} 0.2 & -0.4 \\ -0.1 & 0.2 \end{bmatrix} \begin{bmatrix} \varepsilon_{1,t-1} \\ \varepsilon_{2,t-1} \end{bmatrix}$

We obtain $\Pi = \begin{bmatrix} 0.5 & -1.0 \\ -0.25 & 0.5 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -0.5 & -1.0 \\ -0.25 & -0.5 \end{bmatrix} = \alpha\beta'$.

Let $\beta = \begin{bmatrix} 1 \\ \beta_1 \end{bmatrix}$ and $\alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}$

$$\alpha_1 \times 1 = -0.5$$
$$\alpha_1 \times \beta_1 = -1.0$$
$$\alpha_2 \times 1 = -0.25$$
$$\alpha_2 \times \beta_1 = -0.5$$

Solve the equations, we have

$\alpha = \begin{bmatrix} -0.5 \\ -0.25 \end{bmatrix}$, $\beta = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\beta' \begin{bmatrix} Y_{1t} \\ Y_{2t} \end{bmatrix}$ is stationary

Example: VECM test on bond yields

Although five cointegration vectors are printed, only the first three would be meaningful. The cointegration vectors are the columns of the matrix labeled “Eigenvectors, normalised to first column." The cointegration vectors are determined only up to multiplication by a nonzero scalar and so can be normalized so that their first element is 1.

R: 8_Phillips-OuliarisTest.R

```
Eigenvectors, normalised to first column:
(These are the cointegration relations)

        X3mo.l2 X6mo.l2 X1yr.l2 X2yr.l2 X3yr.l2
X3mo.l2   1.000    1.00    1.00  1.0000   1.000
X6mo.l2  -1.951    2.46    1.07  0.0592   0.897
X1yr.l2   1.056   14.25   -3.95 -2.5433  -1.585
X2yr.l2   0.304  -46.53    3.51 -3.4774  -0.118
X3yr.l2  -0.412   30.12   -1.71  5.2322   1.938

Weights W:
(This is the loading matrix)

        X3mo.l2   X6mo.l2   X1yr.l2   X2yr.l2   X3yr.l2
X3mo.d -0.03441 -0.002440 -0.011528 -0.000178 -0.000104
X6mo.d  0.01596 -0.002090 -0.007066  0.000267 -0.000170
X1yr.d -0.00585 -0.001661 -0.001255  0.000358 -0.000289
X2yr.d  0.00585 -0.000579 -0.003673 -0.000072 -0.000412
X3yr.d  0.01208 -0.000985 -0.000217 -0.000431 -0.000407
```

VECM estimation

The ML estimation methodology is often referred to as the Johansen method.

Assume that innovations are independent identically distributed (IID) multivariate, correlated, Gaussian variables. The methodology can be extended to nonnormal distributions for innovations but computations become more complex and depend on the distribution.

We will use the ECM formulation of the VAR model.

Cointegrated VAR as follows:

$$\Delta Y_t = \boldsymbol{\Pi} Y_{t-1} + \boldsymbol{F_1} \Delta Y_{t-1} + \boldsymbol{F_2} \Delta Y_{t-2} + \cdots + \boldsymbol{F_{p-1}} \Delta Y_{t-p+1} + c(t) + \epsilon_t$$

Concentrated likelihood is a mathematical technique through which the original likelihood function (LF) is transformed into a function of a smaller number of variables, called the ***concentrated likelihood function (CLF)***.

The CLF is also known in statistics as the **profile likelihood**.

To see how CLF works, suppose that the LF is a function of two separate sets of parameters:

$$L = L(\vartheta_1, \vartheta_2)$$

ML Estimators

The MLE principle can be established as follows:

$$\max_{\vartheta_1,\vartheta_2} L(\vartheta_1,\vartheta_2) = \max_{\vartheta_1}\left(\max_{\vartheta_2} L(\vartheta_1,\vartheta_2)\right) = \max_{\vartheta_1}\left(L^C(\vartheta_1)\right)$$

where $L^C(\vartheta_1)$ is the CLF which is a function of the parameters ϑ_1 only.

Assuming usual regularity conditions, the maximum of the LF is attained where the partial derivatives of the log-likelihood function l are zero. In particular:

$$\frac{\partial l(\vartheta_1,\vartheta_2)}{\partial\vartheta_2} = 0$$

If we can solve this system of functional equations, we obtain: $\vartheta_2 = \vartheta_2(\vartheta_1)$. Now the following relationship must hold between the two sets of estimated parameters: $\hat{\vartheta}_2 = \vartheta_2(\hat{\vartheta}_1)$

We see that the original likelihood function has been *concentrated* in a function of a smaller set of parameters. We now apply this idea to the ML estimation of cointegrated systems.

ML Estimators

The following notation is adapted to the special form of the co-integrated VAR model:

$$\Delta Y_t = \boldsymbol{\Pi} Y_{t-1} + \boldsymbol{F_1} \Delta Y_{t-1} + \boldsymbol{F_2} \Delta Y_{t-2} + \cdots + \boldsymbol{F_{p-1}} \Delta Y_{t-p+1} + c(t) + \epsilon_t$$

We define:

$$Y = (y_0, \dots, y_{t-1})$$

$$\boldsymbol{\Delta y_t} = \begin{pmatrix} \Delta y_{1,t} \\ \vdots \\ \Delta y_{d,t} \end{pmatrix}$$

$$\boldsymbol{\Delta Y_t} = (\boldsymbol{\Delta y_1}, \dots, \boldsymbol{\Delta y_T}) = \begin{pmatrix} \Delta y_{1,1} & \cdots & \Delta y_{1,T} \\ \vdots & \ddots & \vdots \\ \Delta y_{d,1} & \cdots & \Delta y_{d,T} \end{pmatrix}$$

$$\boldsymbol{\Delta Z_t} = \begin{pmatrix} \boldsymbol{\Delta y_t} \\ \vdots \\ \boldsymbol{\Delta y_{t-p+2}} \end{pmatrix}$$

ML Estimators

$$\boldsymbol{\Delta Z} = \begin{pmatrix} \boldsymbol{\Delta y_0} & \cdots & \boldsymbol{\Delta y_{T-1}} \\ \vdots & \ddots & \vdots \\ \boldsymbol{\Delta y_{-p+2}} & \cdots & \boldsymbol{\Delta y_{T-p+1}} \end{pmatrix} = \begin{pmatrix} \Delta y_{1,0} & \cdots & \Delta y_{1,T-1} \\ \vdots & \ddots & \vdots \\ \Delta y_{d,0} & \cdots & \Delta y_{d,T-1} \\ \vdots & \ddots & \vdots \\ \Delta y_{1,-p+2} & \cdots & \Delta y_{1,T} \\ \vdots & \ddots & \vdots \\ \Delta y_{d,-p+2} & \cdots & \Delta y_{d,T} \end{pmatrix}$$

$$\boldsymbol{F} = (\boldsymbol{F_1}, \boldsymbol{F_2}, \dots, \boldsymbol{F_{p-1}})$$

Using the matrix notation, as we assume $\boldsymbol{\Pi} = \boldsymbol{\alpha\beta'}$ we can compactly write our model in the following form:

$$\Delta \boldsymbol{Y} = \boldsymbol{F\Delta Z} - \boldsymbol{\alpha\beta' Y} + \boldsymbol{U}$$

ML Estimators

The log-likelihood function can be written:

$$\log(l) = -\frac{nT}{2}\log(2\pi) - \frac{T}{2}\log|\boldsymbol{\Sigma_u}| - \frac{1}{2}\sum_{t=1}^{T}\boldsymbol{\epsilon_t'\Sigma^{-1}\epsilon_t}$$
$$= -\frac{nT}{2}\log(2\pi) - \frac{T}{2}\log|\boldsymbol{\Sigma_u}| - \frac{1}{2}trace\left(\boldsymbol{U'\Sigma_u^{-1}U}\right)$$
$$= -\frac{nT}{2}\log(2\pi) - \frac{T}{2}\log|\boldsymbol{\Sigma_u}| - \frac{1}{2}trace\left(\boldsymbol{\Sigma_u^{-1}UU'}\right)$$
$$= -\frac{nT}{2}\log(2\pi) - \frac{T}{2}\log|\boldsymbol{\Sigma_u}|$$
$$-\frac{1}{2}trace\left((\boldsymbol{\Delta Y} - \boldsymbol{F\Delta Z} + \boldsymbol{\alpha\beta' Y})'\boldsymbol{\Sigma_u^{-1}}(\boldsymbol{\Delta Y} - \boldsymbol{F\Delta Z} + \boldsymbol{\alpha\beta' Y})\right)$$

Eliminating $\boldsymbol{\Sigma}$ and $\boldsymbol{F}$, entails taking partial derivatives, equating them to zero, and expressing $\boldsymbol{\Sigma}$ and $\boldsymbol{F}$ in terms of the other parameters.

By equating the derivatives with respect to $\boldsymbol{\Sigma}$ to zero, it can be demonstrated that $\boldsymbol{\Sigma}_{\mathrm{C}} = \boldsymbol{T^{-1}UU'}$

ML Estimators

We obtain the concentrated likelihood after removing $\mathbf{\Sigma}$:

$$I^{CI} = K - \frac{T}{2}\log|\boldsymbol{U}\boldsymbol{U}'|$$
$$= K - \frac{T}{2}\log|(\boldsymbol{\Delta Y} - \boldsymbol{F}\boldsymbol{\Delta Z} + \boldsymbol{\alpha}\boldsymbol{\beta}'\boldsymbol{Y})(\boldsymbol{\Delta Y} - \boldsymbol{F}\boldsymbol{\Delta Z} + \boldsymbol{\alpha}\boldsymbol{\beta}'\boldsymbol{Y})'|$$

where K is a constant that includes all the constant terms left after concentrating.

Next eliminating the $\boldsymbol{F}$ terms, take derivatives of l with respect to $\boldsymbol{F}$, equating them to zero, and evaluating them at $\mathbf{\Sigma}_{\mathrm{C}}$.

Performing all the calculations, the evaluation at I^{CI} is irrelevant and the following formula holds:

$$\boldsymbol{F_C} = (\boldsymbol{\Delta Y} + \boldsymbol{\alpha}\boldsymbol{\beta}'\boldsymbol{Y})\boldsymbol{\Delta Z}'(\boldsymbol{\Delta Z}\boldsymbol{\Delta Z}')^{-\mathbf{1}}$$

Substituting this expression in the formula for I^{CI} (the log-likelihood after eliminating I^{CI}) we obtain the expression on the next slide.

ML Estimators

$$l^{C\Pi} = K - \frac{T}{2} log\,|\left(\left(\Delta Y - \left((\Delta Y + \alpha\beta' Y)\Delta Z'(\Delta Z\Delta Z')^{-1}\right)\Delta Z + \alpha\beta' Y\right)\right)$$
$$\left(\Delta Y - (\Delta Y + \alpha\beta' Y)\Delta Z'(\Delta Z\Delta Z')^{-1}\Delta Z\; + \alpha\beta' Y'\right)|$$
$$= K - \frac{T}{2} log\,|\left(\Delta Y + \alpha\beta' Y - \left((\Delta Y + \alpha\beta' Y)\Delta Z'(\Delta Z\Delta Z')^{-1}\right)\Delta Z\right)$$
$$\left(\Delta Y + \alpha\beta' Y - (\Delta Y + \alpha\beta' Y)\Delta Z'(\Delta Z\Delta Z')^{-1}\Delta Z\right)'|$$
$$= K - \frac{T}{2} log\,|\left((\Delta Y + \alpha\beta' Y)\left(I_T - \Delta Z'(\Delta Z\Delta Z')^{-1}\Delta Z\right)\right)$$
$$\left((\Delta Y + \alpha\beta' Y)\left(I_T - \Delta Z'(\Delta Z\Delta Z')^{-1}\Delta Z\right)\right)'\,|$$
$$= K - \frac{T}{2} log|(\Delta Y + \alpha\beta' Y)M(\Delta Y + \alpha\beta' Y)'|$$
$$= K - \frac{T}{2} log|\Delta Y M \Delta Y' + \alpha\beta' Y M \Delta Y' + \Delta Y M(\alpha\beta' Y)' + \alpha\beta' Y M(\alpha\beta' Y)'|$$

Where $M = I_T - \Delta Z'(\Delta Z\Delta Z')^{-1}\Delta Z$.

ML Estimators

Matrices of the form $\boldsymbol{A} = \boldsymbol{I} - \boldsymbol{B}'(\boldsymbol{B}\boldsymbol{B}')^{-1}\boldsymbol{B}$ are called ***projection matrices***. They are idempotent and symmetric, that is $\boldsymbol{A}\boldsymbol{A} = \boldsymbol{A}^2 = \boldsymbol{A}$ and $\boldsymbol{A} = \boldsymbol{A}'$.

Defining $\boldsymbol{R_0} = \boldsymbol{\Delta Y M}$,$\boldsymbol{R_1} = \boldsymbol{Y M}$ and

$$S_{ij} = \frac{\boldsymbol{R_i R_j}}{T}, i, j = 1,2$$

we can rewrite the CLF as follows:

$$I^{C\Pi}(\boldsymbol{\alpha\beta'}) = K - \frac{T}{2}\log|\boldsymbol{S_{00}} - \boldsymbol{S_{10}\alpha\beta'} - \boldsymbol{S_{01}}(\boldsymbol{\alpha\beta'})' + \boldsymbol{\alpha\beta' S_{11}}(\boldsymbol{\alpha\beta'})'|$$

The original analysis of Johansen obtained the same result applying the method of ***reduced rank regression***.

Reduced rank regressions are multiple regressions where the coefficient matrix is subject to constraints.

ML Estimators

The Johansen method eliminates the terms $\boldsymbol{F}$ by regressing Δy_t and y_{t-1} on $(\Delta y_{t-1}, \Delta y_{t-2}, \ldots, \Delta y_{t-p+1})$ to obtain the following residuals:

$$R_{0t} = \Delta y_t + D_1 \Delta y_{t-1} + D_2 \Delta y_{t-2} + \cdots + D_{p-1} \Delta y_{t-p+1}$$
$$R_{1t} = \Delta y_{t-1} + E_1 \Delta y_{t-1} + E_2 \Delta y_{t-2} + \cdots + E_{p-1} \Delta y_{t-p+1}$$

where

$$\boldsymbol{D} = \left(\boldsymbol{D_1}, \boldsymbol{D_2}, \ldots, \boldsymbol{D_{p-1}}\right) = \boldsymbol{\Delta Y \Delta Z'} (\boldsymbol{\Delta Z \Delta Z'})^{-1}$$
$$\boldsymbol{E} = \left(\boldsymbol{E_1}, \boldsymbol{E_2}, \ldots, \boldsymbol{E_{p-1}}\right) = \boldsymbol{Y \Delta Z'} (\boldsymbol{\Delta Z \Delta Z'})^{-1}$$

The original model is therefore reduced to the following ”simpler model”:

$$R_{0t} = \boldsymbol{\alpha\beta'} R_{1t} + u_t$$

The likelihood function of this model depends only on R_{0t} , R_{1t}. It can be written as follows:

$$l(\boldsymbol{\alpha\beta'}) = K_1 - \frac{T}{2} \log \left| \left(R_0 + R_1(\alpha\beta')\right)' \left(R_0 + R_1(\alpha\beta')\right) \right|$$

where we define R_0, R_1 as above.

Defining S_{ij} as above, we obtain exactly the same form for the CLF:

$$I^{C\Pi}(\boldsymbol{\alpha\beta'}) = K - \frac{T}{2} \log |\boldsymbol{S_{00}} - \boldsymbol{S_{10}\alpha\beta'} - \boldsymbol{S_{01}}(\boldsymbol{\alpha\beta'})' + \boldsymbol{\alpha\beta' S_{11}}(\boldsymbol{\alpha\beta'})'|$$

ML Estimators

If the matrix $\mathbf{\Pi} = \boldsymbol{\alpha\beta}'$ were unrestricted, then maximization would yield

$$\mathbf{\Pi} = \boldsymbol{S}_{\mathbf{01}}\boldsymbol{S}_{\mathbf{11}}^{-\mathbf{1}}.$$

By performing the rather lengthy computations, it can be demonstrated that we obtain a solution by solving the following eigenvalue problem:

$$|\boldsymbol{S}_{\mathbf{10}}\boldsymbol{S}_{\mathbf{00}}^{-\mathbf{1}}\boldsymbol{S}_{\mathbf{01}} - \boldsymbol{\lambda S}_{\mathbf{11}}| = 0$$

This eigenvalue problem, together with normalizing conditions, will yield d eigenvalues λ_i and d eigenvectors Λ_i .

In order to make this problem well determined, Johansen imposed the normalizing conditions: $\boldsymbol{\Lambda}'\boldsymbol{S}_{\mathbf{11}}\boldsymbol{\Lambda} = \boldsymbol{I}$

Order the eigenvalues and choose the r eigenvectors Λ_i corresponding to the largest r eigenvalues.

It can be demonstrated that a ML estimator of the matrix $\boldsymbol{C}$ is given by $\widehat{\boldsymbol{\beta}}' = (\widehat{\boldsymbol{\Lambda}}_{\mathbf{1}}, \dots, \widehat{\boldsymbol{\Lambda}}_{\boldsymbol{r}})$ and an estimator of the matrix α by $\widehat{\boldsymbol{\alpha}} = \boldsymbol{S}_{\mathbf{00}}\widehat{\boldsymbol{C}}$.

The maximum of the log-likelihood is

$$l_{max} = K - \frac{T}{2}\log|\boldsymbol{S}_{\mathbf{00}}| - \frac{T}{2}\sum_{i=1}^{r}\log(1 - \lambda_i)$$

Pairs trading

Pairs trading was developed in the 1980's by a group of Quants at Morgan Stanley, who reportedly made over $50 million profit for the firm in 1987.

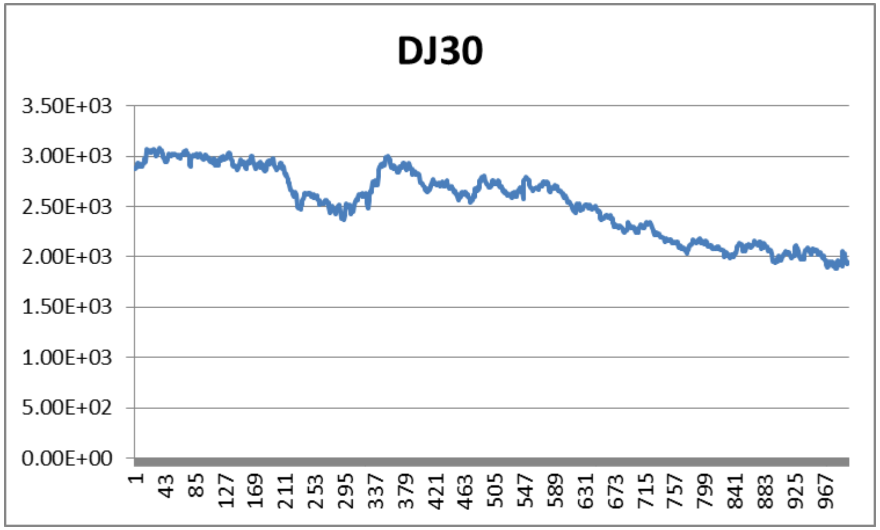
General idea of trading is to sell overvalued securities and buy undervalued ones. But the true value of the security is hard to determine in practice.

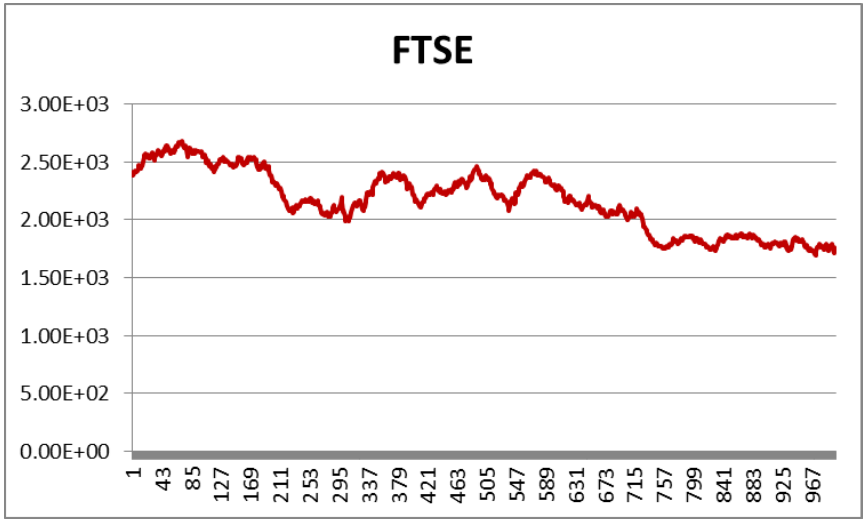
Pairs trading attempts to resolve this difficulty by using relative pricing. If two securities have similar characteristics, then the price movements of both securities must be more or less the same. A relative large deviation of one security indicates an "over-value" relative to the other. Here the true price is not important.

Specifically, pairs trading tries to profit from the principles of mean-reversion processes, a property of stationary series. The idea is to trade a pair or more of related stocks whose combination is stationary. As the mean of the series is not changing over time, any deviation is temporary and the series moves back to the average.

Example

Date: Daily close prices of DJ30 and FTSE from 2007-2011.

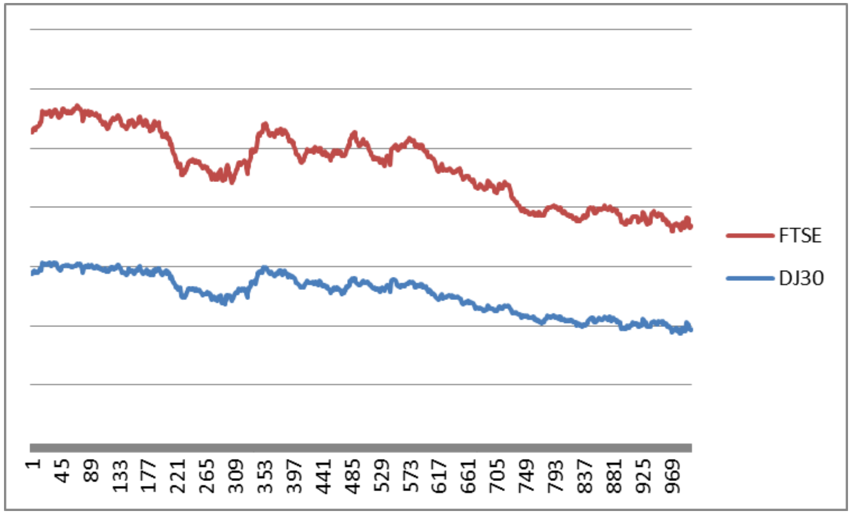




The series of both prices behave like random-walk processes (non-stationary), but with a similar movement.

Example of Pairs Trade

Daily close prices of DJ30 and FTSE from 2007-2011.



It shows DJ30 and FTSE move together, indicating that these two integrated processes share the similar errors. Possibly, a linear combination of them (regression residuals) is stationary with mean-reverting. Deviations from the mean lead to trading opportunities.

Wait until the prices diverge beyond a certain threshold, then short the “winner” and buy the “loser”.

Reverse your positions when the two prices converge --> Profit from the reversal in trend.

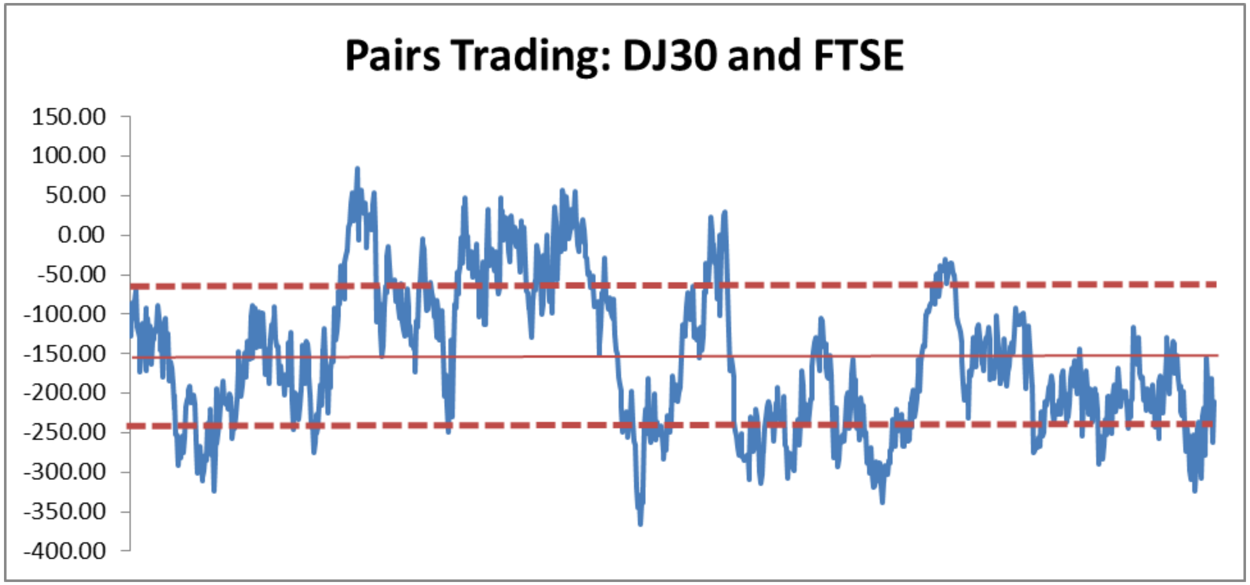
Example of Pairs Trade

Consider two indices: DJ30 and FTSE. The two log-price series are co-integrated. Therefore, there exists a linear combination that is stationary (mean-reverting).

$$Y_t = DJ_t - \beta \times FTSE_t$$

Buy one share of DJ30 and short sell β share FTSE. The return of the portfolio for a given period h is

$$r_h = (DJ_{t+h} - DJ_t) - \beta(FTSE_{t+h} - FTSE_t) = Y_{t+h} - Y_t.$$



R lab

The data set midcapD.ts in the fEcofin package has daily returns on 20 midcap stocks in columns 2-21. Columns 1 and 22 contain the date and market returns, respectively. In this section, we will use returns on the first 10 stocks. To find the stock prices from the returns, we use the relationship

$$P_t = P_0 \exp(r_1 + \cdots + r_t)$$

where P_t and r_t are the price and log return at time t. The returns will be used as approximations to the log returns. The prices at time 0 are unknown, so we will use $P_0 = 1$ for each stock. This means that the price series we use will be off by multiplicative factors. This does not affect the number of cointegration vectors. If we find that there are cointegration relationships, then it would be necessary to get the price data to investigate trading strategies. Johansen's cointegration analysis will be applied to the prices with the ca.jo function in the urca package. Run

```
library(fEcofin)
library(urca)
x = midcapD.ts[,2:11]
prices= exp(apply(x,2,cumsum))
options(digits=3)
summary(ca.jo(prices))
```

Problem 1 How many cointegration vectors were found?

R lab

This example is similar to Example 15.3 but uses different yield data. The data are in the mk.zero2 data set in the fEcofin package. There are 55 maturities and they are in the vector mk.maturity. We will use only the first 10 yields. Run

```
library("fEcofin")
library(urca)
mk.maturity[2:11,]
summary(ca.jo(mk.zero2[,2:11]))
```

Problem 2 What maturities are being used? Are they short-, medium-, or long-term, or a mixture of short- and long-term maturities?

Problem 3 How many cointegration vectors were found? Use 1% level tests.
Problem 4 Which trading strategy do you recommend?

- Data: 9_midcapD.ts.csv, 9_mk.maturity.csv, 9_mk.zero2.csv
- R: Rlab9.R

R lab

Consider the monthly yields of Moody's seasoned corporate Aaa and Baa bonds from July 1954 to February 2005 for 609 observations. The data m-bnd.txt were obtained from Federal Reserve Bank of St. Louis.

```
>library("fUnitRoots")
>library("urca")
>library("MTS")
>da=read.table("m-bnd.txt")
>bnd=da[,4:5]
>colnames(bnd) <- c("Aaa","Baa")
>adfTest(bnd[,1],lags=3,type="c")
>adfTest(bnd[,2],lags=2,type="c")
>m1=VARorder(bnd)
>m2=ca.jo(bnd,K=2,ecdet=c("none"))
>summary(m2)
```

Problem 5 Are the bond yields stationary? Justify using the augmented Dicky-Fuller unit root test.
Problem 6 Choose an VAR model to fit the data. Specify the order using the information criteria.
Problem 7 Conduct cointegration test.
Problem 8 Propose a trading strategy.

- Rlab9_2.R

R lab – Results & Discussions

Problem 1 How many cointegration vectors were found?

Selected output is below. In no case is a test statistic greater than even the 10% level critical value. This suggests that there are no cointegration vectors.

```
######################
# Johansen-Procedure #
######################

Test type: maximal eigenvalue statistic (lambda max) , with linear trend

Eigenvalues (lambda):
[1] 0.11064 0.06903 0.04509 0.04255 0.03916 0.03029 0.02339 0.01831 0.01259 0.00120

Values of test statistic and critical values of test:

test 10pct 5pct 1pct
r <= 9  0.60  6.5 8.18 11.6
r <= 8  6.31 12.9 14.90 19.2
r <= 7  9.20 18.9 21.07 25.8
r <= 6 11.79 24.8 27.14 32.1
r <= 5  15.32 30.8 33.32 38.8
r <= 4 19.89 36.2 39.43 44.6
r <= 3  21.65 42.1 44.91 51.3
r <= 2  22.98 48.4 51.07 57.1
r <= 1  35.62 54.0 57.00 63.4
r = 0  58.39 59.0 62.42 68.6
```

R lab – Results & Discussions

Problem 2 What maturities are being used? Are they short-, medium-, or long-term, or a mixture of short- and long-term maturities?

The maturities (in years) are given below. We see that they are all short term, less than one year.

```
>  mk.maturity[2:11,]
[1)0.167 0.250 0.333 0.417 0.500 0.583 0.667 0.750 0.833 0.917
```

R lab – Results & Discussions

Problem 3 How many cointegration vectors were found? Use 1% level tests.
From the (selected) output below, we accept that $r \leq 4$ but reject that $r \leq 3$, so 4 cointegration vectors were found.

```
> summary(ca.jo(mk.zero2[,2:11]))

######################
# Johansen-Procedure #
######################

Test type: maximal eigenvalue statistic (lambda max) , with linear trend

Eigenvalues (lambda):
[1] 0.7518 0.6631 0.5981 0.5498 0.4675 0.4263 0.3486 0.2474 0.1162 0.0218

Values of teststatistic and critical values of test:

test 10pct 5pct 1pct
r <= 6 27.87 24.8 27.14 32.1
r <= 5  36.12 30.8 33.32 38.8
r <= 4 40.96 36.2 39.43 44.6
r <= 3  51.87 42.1 44.91 51.3
r <= 2  59.25 48.4 51.07 57.1
r <= 1  70.71 54.0 57.00 63.4
```

R lab – Results & Discussions

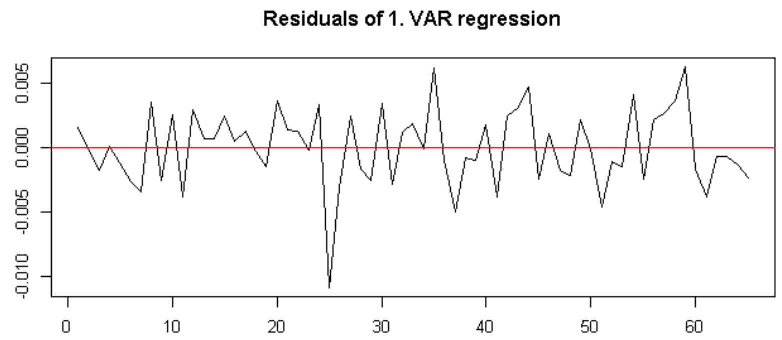
Problem 4 Which trading strategy do you recommend?

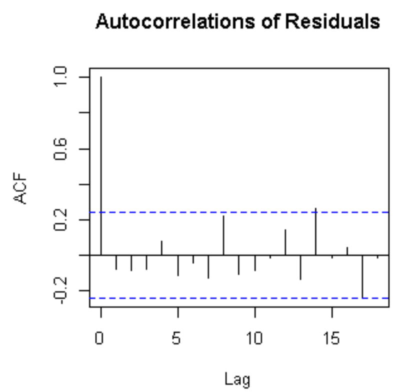
From the (selected) output below, we choose the first 3 columns as co-integrating vectors. One trading strategy is e.g.

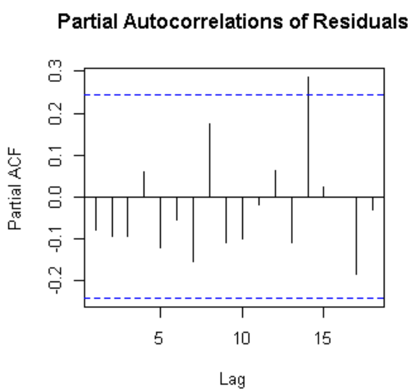
$$M_1 - 4.73M_2 + 10.45M_3 - 11.40M_4 + 0.08\,M_5 + \cdots - 1.40M_{10}$$

```
Eigenvectors, normalised to first column:
(These are the cointegration relations)

           M.1.12   M.2.12  M.3.12  M.4.12  M.5.12
M.1.12     1.0000   1.0000   1.000   1.000    1.00
M.2.12    -4.7341   0.0734  -5.227  -2.686   -1.11
M.3.12    10.4466 -16.1591  14.437  -1.213  -14.93
M.4.12   -11.4036  55.2255 -27.043  11.818   59.12
M.5.12     0.0821 -95.7229  37.195 -14.865  -92.69
M.6.12    16.0789  92.4944 -38.014  -1.409   18.43  -
M.7.12   -18.4762 -46.0821  23.980  15.686  130.51  :
M.8.12     6.4916  11.3169  -4.378  -0.998 -160.41 -:
M.9.12     1.9201  -3.1889  -2.661 -14.817   70.33
M.10.12   -1.4041   1.0603   0.708   7.486  -10.25
         M.9.12 M.10.12
M.1.12      1.0   1.000
M.2.12    -13.8  -0.215
M.3.12     72.0 -16.884
M.4.12    -99.1  39.882
M.5.12    -29.7 -20.449
M.6.12     46.0 -32.564
M.7.12    254.0  42.682
M.8.12   -427.4  -1.593
M.9.12    258.6 -23.037
M.10.12   -61.9  11.081
```



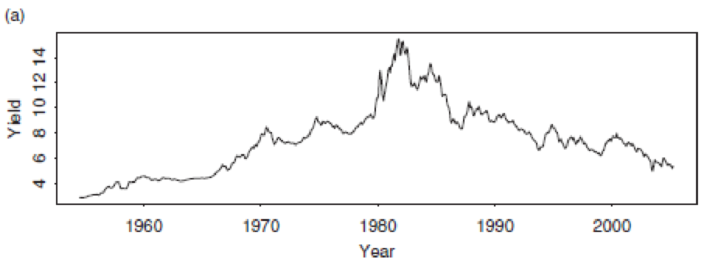


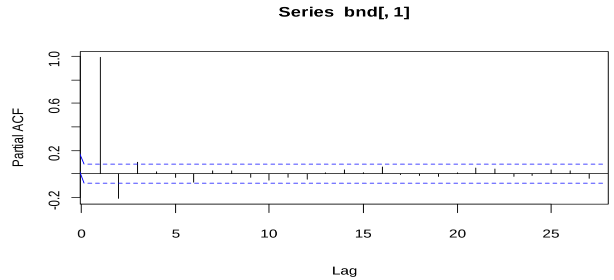


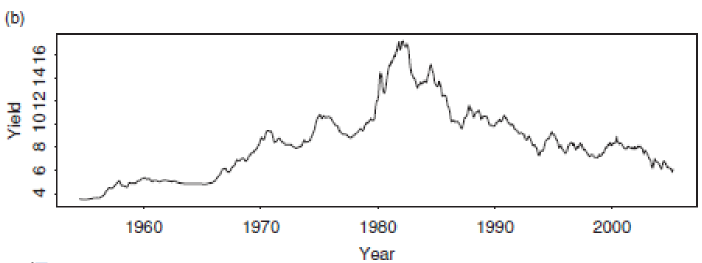
R lab – Results & Discussions

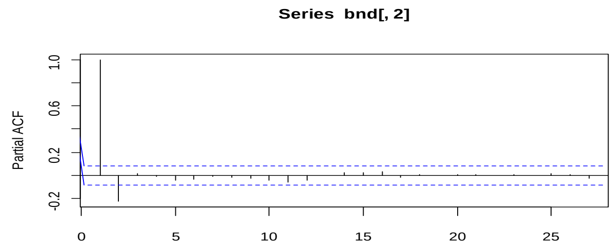
Problem 5 Are the bond yields stationary? Justify using the augmented Dicky-Fuller unit root test.

The time series of the bond yields move in a parallel manner. The augmented Dickey–Fuller unit-root test confirms that the two bond yields are unit-root nonstationary.










```
> adfTest(bnd[,1],lags=3,type="c")

Title:
 Augmented Dickey-Fuller Test

Test Results:
  PARAMETER:
    Lag Order: 3
  STATISTIC:
    Dickey-Fuller: -1.7007
  P VALUE:
    0.425
```

```
> adfTest(bnd[,2],lags=2,type="c")

Title:
 Augmented Dickey-Fuller Test

Test Results:
  PARAMETER:
    Lag Order: 2
  STATISTIC:
    Dickey-Fuller: -1.6221
  P VALUE:
    0.4544
```

R lab – Results & Discussions

Problem 6 Choose an VAR model to fit the data. Specify the order using the information criteria.
The order *p* = 3 is selected by both BIC and HQ. We will also employ a VAR(3) model in the cointegration test.

```
> m1=VARorder(bnd)
selected order: aic =  11
selected order: bic =  3
selected order: hq =  3
Summary table:
       p   AIC   BIC    HQ    M(p) p-value
 [1,]  0 -0.57 -0.57 -0.57    0.00  0.0000
 [2,]  1 -7.87 -7.84 -7.86 4331.08  0.0000
 [3,]  2 -8.18 -8.13 -8.16  195.53  0.0000
 [4,]  3 -8.26 -8.17 -8.23   51.61  0.0000
 [5,]  4 -8.26 -8.14 -8.21    5.51  0.2387
 [6,]  5 -8.25 -8.10 -8.19    3.51  0.4768
 [7,]  6 -8.28 -8.10 -8.21   23.42  0.0001
 [8,]  7 -8.28 -8.08 -8.20   10.63  0.0311
 [9,]  8 -8.28 -8.05 -8.19    8.91  0.0634
[10,]  9 -8.28 -8.02 -8.18    5.16  0.2711
[11,] 10 -8.28 -7.99 -8.16    7.05  0.1335
[12,] 11 -8.28 -7.97 -8.16   11.22  0.0242
[13,] 12 -8.28 -7.93 -8.15    5.90  0.2071
[14,] 13 -8.27 -7.89 -8.13    2.04  0.7292
```

R lab – Results & Discussions

Problem 7 Conduct cointegration test.

Since the mean vector of the differenced series Δy_t is not significantly different from 0, we do not consider the constant in the test. Compared with critical values, we reject r=0, but cannot reject r=1. Therefore there is one cointegrating vector.

```
######################
# Johansen-Procedure #
######################

Test type: maximal eigenvalue statistic (lambda max) , with linear trend

Eigenvalues (lambda):
[1] 0.05477 0.00467

Values of teststatistic and critical values of test:

          test 10pct  5pct 1pct
r <= 1 |  2.84   6.5  8.18 11.6
r = 0  | 34.19  12.9 14.90 19.2

Eigenvectors, normalised to first column:
(These are the cointegration relations)

       Aaa.l2 Baa.l2
Aaa.l2  1.000   1.00
Baa.l2 -0.886  -2.72

Weights W:
(This is the loading matrix)

       Aaa.l2  Baa.l2
Aaa.d -0.0470 0.00248
Baa.d  0.0405 0.00214
```

R lab – Results & Discussions

Problem 8 Propose a trading strategy.

From the R output, we choose the first column as the co-integrating vector. The trading strategy is Aaa – 0.886Baa. The ADF test of the combination confirms that the series does not have any unit root. The test statistic is −4.61 with p-value 0.01.

As expected, the time plot of the cointegrated series shows the characteristics of a stationary time series.