A New Approach of Determining Cointegrating Vectors

Yan-Xia Lin
School of mathematics and Applied Statistics, The University of Wollongong
Michael McCrae
Department of Accounting & Finance, The University of Wollongong

Abstract: A new approach of determining cointegrating vectors based on the rank of covariance matrix of error processes of ARIMA model is presented. By noting that, in practice, exactly cointegration behaviours rarely exist, a new concept of asymptotic cointegration is introduced. Therefore a new procedure of determining asymptotic conitgration vectors is described in this paper. Under this procedure, the significance of the cointegration behaviours can be easily identified.

1 Introduction

The analysis of long run equilibrium processes among components of a given time series vector under Univariate ARIMA type modelling has been limited by two aspects. One is the lack of a rigorous statistical definition for the concept of long run equilibrium and the other is the requirement for differencing out long run dynamics to ensure stationarity. The development of multivariate cointegration overcame these limitations by defining long run equilibrium in terms of the existence of stationary behaviour in linear combinations of series of interest and by introducing techniques which obviate the need for differencing.

Two fundamental issues in cointegration are the problems of how to identify cointegration relationships among time series of interest and how to estimate the cointegration vector. Currently, several procedures have been developed to solve these issues. Among them two basic procedures are the Engle-Granger two-step procedure and the Johansen's procedure. The limitation of earlier cointegration procedures based on the the Engle-Granger (1987) approach to the identification of one independent coingtegrating vector was overcome by the Johansen (1988) maximum-liklihood procedure for identifying n independent cointegrating vectors.

By understanding Johansen's procedure (see Harris, 1996), we realize that any application of Johansen's method for estimating cointegrating vectors to real data must always be conditional upon the independence condition or, at most, weak correlation, between the summand I(0) series. The difficulty of checking these conditions in practice necessitates further unit root testing of all Johansen type estimates of cointegrating vectors to ensure the necessary conditions before they can be accepted as eligible cointegrating vectors.

To overcome these difficulties, we now develop an alternative approach for identifying and estimating cointegrating vectors based on a connection between univariate modelling processes and multivariate cointegration. The application of such connection to cointegration has not been discussed in literature.

In Section 2 it shows that there is an essential relationship between cointegration and univeriate ARIMA modelling and to suggest a new way to identify cointegrating vectors by using the information of covariance matrix of error proceses. A new concept of 'asymptotic' cointegration is given in Section 3. Several simulated examples are presented in Section 3 to show the application of the new approach and the new concept of 'asymptotic' cointegration.

2 Cointegrating Vectors and Error Processes

An alternative approach for identifying and estimating cointegrating vectors may be found in the inherent theoretical relationship between cointegrating vectors and the covariance matrix of error processes. The idea used to establish such approach is similar to that used by Bierens (1997) and Engle and Yoo (1987), but the approach is different.

We first state the definitions of $I(d)$ time series and cointegration.

Definition 1 *A time series X_t is called a $I(d)$ se-*

ries, if after d differencing, X_t is stationary.

Definition 2 *Let $\mathbf{X_t}$ be a $I(d)$ time series vector, $d > 0$. The time series vector $\mathbf{X_t}$ is said to be cointegrated in order $d-b > 0$ for integer b, if there is a vector β such that $\beta^T\mathbf{X_t}$ is a $I(d-b)$ time series vector.*

Definitions 1 and 2 are slightly different from those given by Engle and Granger (1987) and Granger and Weiss (1983). Our definition that X_t is $I(d)$ does not requir that $(1-B)^d X_t$ is invertible. This limitation is not significant since our only concern is with the stationarity property of the relevant time series.

For simplicity, we limit our discussion to the case $d = b = 1$.

First, two simple but basic lemmas are given below, which will be used during the proof of the following theorems.

Lemma 1 *If $\phi_i(B)$ and $\theta_i(B)$, $i = 1,2$ are finite-order polynomial, and the roots of $\phi_i(B) = 0$, $i = 1,2$, are outside the complex unit circle, then $\phi_1^{-1}(B)\theta_1(B)v_t + \phi_2^{-1}(B)\theta_2(B)v_t$ is stationary, where v_t is an i.i.d series and B the back-shift operator.*

The proof of Lemma 1 is straightforward.

From Lemma 1, we can obtain the following lemma

Lemma 2 *Assume that $\mathbf{W_t} = C_1^{-1}(B)C_2(B)\mathbf{v_t}$, where $C_1(B)$ and $C_2(B)$ are two matrices and $C_1(B)$ is a diagonal matrix; $\{\mathbf{v_t}\}$ are independent and $var(\mathbf{v_t})$ is a diagonal matrix. If the roots of $det(C_1(B)) = 0$ lie outside the complex unit circle and, $C_1(B)$ and $C_2(B)$ are finite-order lag polynomials, then, for any vector ξ, $\xi^{\mathbf{T}}\mathbf{W_t}$ is a stationary process.*

The relationship between the number of cointegrating vectors and the rank of covariance matrix of error process is described by Theorems 1 and 2.

Theorem 1 *Assume that $\mathbf{X_t}$ is a $p \times 1$ $I(1)$ time series vector and each component of $\mathbf{X_t}$ has the following expression:*

$$\Phi_i(B)(1-B)X_{it} = \Theta_i(B)\epsilon_{\mathbf{it}} \quad (1)$$

where ε_{it} are white noise, $i = 1, 2, \cdots, p$, both $\Phi_i(B)$ and $\Theta_i(B)$ are finite-order polynomial and have roots outside the unit circle. Assume, for any $s \neq t$, $Cov(\epsilon_{\mathbf{t}}, \epsilon_{\mathbf{s}}) = \mathbf{0}$ and $Var(\epsilon_{\mathbf{t}}) = \Sigma$ is free from t, where $\epsilon_{\mathbf{t}} = (\epsilon_{1,t}, \cdots, \epsilon_{p,t})^T$. If $rank(\Sigma) = p - r < p$, then $\mathbf{X_t}$ has $r > 0$ cointegrating vectors.

Inversely, if $\mathbf{X_t}$ are cointegrated, then $\epsilon_{\mathbf{t}}$ will be correlated, i.e. $rank(\Sigma) < p$.

Proof: First we prove that, if the $rank(\Sigma) = p - r < p$, then $\mathbf{X_t}$ has $r > 0$ cointegrating vectors.

Since the $rank(\Sigma) = p - r < p$, there is a nonsingular matrix A such that

$$Var(A\epsilon_{\mathbf{t}}) = \begin{pmatrix} \Lambda & \mathbf{0} \\ \mathbf{0} & \mathbf{D} \end{pmatrix}$$

where $\mathbf{D}$ is a $r \times r$ zero matrix and Λ is a diagonal matrix containing all non zero eigenvalues of Σ.

Let $A\epsilon_{\mathbf{t}} = \mathbf{v}_t^* = (v_{1,t}^*, v_{2,t}^*, \cdots, v_{p,t}^*)^T$. Then $v_{i,t}^* = 0$, $i = p-r+1, \cdots, p$, because $Var(v_{i,t}^*) = 0$ for all $i \geq p - r + 1$. Therefore

$$\epsilon_{\mathbf{t}} = A^{-1}\begin{pmatrix} v_{1,t}^* \\ \vdots \\ v_{p-r,t}^* \\ 0 \\ \vdots \\ 0 \end{pmatrix} = (A_{p-r}^*)\begin{pmatrix} v_{1,t}^* \\ \vdots \\ v_{p-r,t}^* \end{pmatrix},$$

where A_{p-r}^* is a $p \times (p-r)$ matrix. Denote $\mathbf{v_t} = (v_{1,t}^*, \cdots, v_{p-r,t}^*)^T$. Then $\mathbf{v_t}$ has mutually independent components and

$$(1-B)\mathbf{X_t} = C(B)\mathbf{v_t},$$

where

$$C(B) = \begin{pmatrix} \Phi_1^{-1}(B) & & 0 \\ & \ddots & \\ 0 & & \Phi_p^{-1}(B) \end{pmatrix} \times \begin{pmatrix} \Theta_1(B) & & 0 \\ & \ddots & \\ 0 & & \Theta_p(B) \end{pmatrix} A_{p-r}^*$$

Therefore,

$$\mathbf{X_t} = \mathbf{X_0} + (1-B)\frac{C(B) - C(1)}{1-B}\sum_{i=1}^{t}\mathbf{v_i} + C(1)\sum_{i=1}^{t}\mathbf{v_i}$$

$$= \mathbf{X_0} - \mathbf{W_0} + \mathbf{W_t} + C(1)\sum_{i=1}^{t}\mathbf{v_i},$$

in which

$$\mathbf{W_t} = \frac{C(B) - C(1)}{1-B}\mathbf{v_t}. \quad (2)$$

By noting that the rank of $C(1)$ is equal to the rank of A_{p-r}^*, i.e. $p-r$, there exist r vectors, say $\xi_1, \cdots, \xi_r$, such that $\xi_i^T C(1) = \mathbf{0}$, $i = 1, 2, \cdots, r$. Thus

$$\xi_i^T \mathbf{X}_t = \xi_i^T(\mathbf{X}_0 - \mathbf{W}_0) + \xi_i^T \mathbf{W}_t,$$

$i = 1, 2, \cdots, r$. From Lemma 2 and (2), $\xi_i^T \mathbf{W_t}$ is stationary, for all $i = 1, \cdots, r$. This implies that $\mathbf{X_t}$ are cointegrated with cointegrating vectors ξ_i, $i = 1, \cdots, r$.

Now we prove the inverse part indirectly. Assume that $\mathbf{X_t}$ are cointegrated, but $\{\varepsilon_{i,t}\}$ are independent, for $i = 1, 2, \cdots, p$.

Since $\mathbf{X_t}$ are cointegrated, there is a vector $\xi = (\xi_1^*, \xi_2^*, \cdots, \xi_p^*)^T$ such that $\xi^T \mathbf{X_t}$ is $I(0)$. By noting that

$$(1-B)X_{i,t} = \Phi_i^{-1}(B)\Theta_i(B)\varepsilon_{i,t},$$

we obtain that

$$(1-B)\xi_i^* X_{i,t} = \Phi_i^{-1}(B)\Theta_i(B)\xi_i^*\varepsilon_{i,t},$$

$i = 1, 2, \cdots, p$, then $\{\xi_i^* X_{i,t}\}$ are independent $I(1)$ series. It leads to $\xi^T \mathbf{X_t}$ an $I(1)$ time series. We obtain contradiction. Thus $\{\varepsilon_{i,t}\}$ have to be correlated. □

The proof of Theorem 1 gives a procedure for determining cointegrating vectors via the covariance matrix of ϵ_t. The procedure consists the following steps.

(i) Fit the individual time series of interest with a ARIMA model and obtaining error processes;

(ii) identify the covariance matrix of the error processes, Σ;

(iii) use eigenvectors to determine matrix A_{p-r}^*(in the proof of Theorem 1), if $rank(\Sigma) = p - r < p$;

(iv) determine all independent cointegrating vectors by solving the equation $\xi^T C(1) = 0$, where $C(1)$ is defined in Theorem 1.

The following example demonstrates this procedure.

Example 1 Consider two time series x_t and y_t, which satisfy the following models:

$$(1-B)\begin{pmatrix} x_t \\ y_t \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1-0.2B \end{pmatrix}\begin{pmatrix} \varepsilon_t \\ \varepsilon_t \end{pmatrix},$$

where ε_t is white noise with variance 1. From the model we have

$$\Sigma = Cov(\begin{pmatrix} \varepsilon_t \\ \varepsilon_t \end{pmatrix}, \begin{pmatrix} \varepsilon_t \\ \varepsilon_t \end{pmatrix}) = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix},$$

which has eigenvalues 1 and 0. Thus

$$\begin{pmatrix} \varepsilon_t \\ \varepsilon_t \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}^{-1}\begin{pmatrix} v_{1t}^* \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} v_{1,t}^*$$

and

$$C(B) = \begin{pmatrix} 1 \\ 1-0.2B \end{pmatrix}, C(1) = \begin{pmatrix} 1 \\ 0.8 \end{pmatrix}.$$

Let $\xi^T = (-0.8, 1)$, then $\xi^T C(1) = 0$ and $\mathbf{W}_t = (0, 0.2v_{1,t}^*)^T$.

From Theorem 1, x_t and y_t are cointegrated with cointegrating vector $\xi = (0, 0.2)^T$.

Theorem 1 can be generalised as follows:.

Theorem 2 *Assume that $\mathbf{X_t}$ is an $p \times 1$ $I(1)$ time series vector and each component of $\mathbf{X_t}$ has the following expression:*

$$\Phi_i(B)(1-B)X_{it} = a_i + \Theta_i(B)\varepsilon_{it}$$

where ϵ_{it} are white noise, $i = 1, 2, \cdots$, both $\Phi_i(B)$ and $\Theta_i(B)$ are finite-order polynomial and have roots outside the unit circle. Assume, for any $s \neq t$, $Cov(\epsilon_t, \epsilon_s) = 0$ and $Var(\epsilon_t) = \Sigma$ is free from t, where $\epsilon_t = (\epsilon_{1,t}, \cdots, \epsilon_{p,t})^T$. If the $rank(\Sigma) = p - r < p$, then there are r vectors, say ξ_j, $j = 1, 2, \cdots, r$, such that $\xi^T(\mathbf{X_t} - \mathbf{A_0}t)$ is a stationary time series vector, where

$$\mathbf{A_0} = \begin{pmatrix} \Phi_1^{-1}(1) & & 0 \\ & \ddots & \\ 0 & & \Phi_p^{-1}(1) \end{pmatrix}\begin{pmatrix} a_1 \\ \vdots \\ a_p \end{pmatrix}.$$

In the proof of Theorem 1, the structure of the matrix $C(B)$ is special, where the rank of $C(1)$ is determined by the rank of A_{p-r}^*. A further interesting question is whether there are any relationships between cointegrating of $\mathbf{X_t}$ and the rank of $C(1)$ if the matrix $C(B)$ is an any $p \times p$ matrix. In general, under some minor conditions, it can prove that $\mathbf{X_t}$ are cointegrated iff the rank of $C(1)$ is less than p. This result will be discussed in somewhere else.

3 The Asymptotic Cointegration Behaviours

In this section, we discuss how to determine cointegration behaviours among time series of interest in practice via the procedure mentioned in Section 2.

Assume that time series $X_{i,t}$, $i = 1, 2, \cdots, p$ and $t = 1, 2, \cdots$, have been correctly fitted by ARIMA models. Let $\varepsilon_{i,t}$ be error processes corresponding to $X_{i,t}$, and, for each i fixed, $\varepsilon_{i,t}$ are independent, $i = 1, 2, \cdots, p$ and $t = 1, 2, \cdots$. Then covariance matrix $\Sigma = Var(\varepsilon_t)$ can be estimated by sample covariance matrix $\frac{1}{n}\sum_{t=1}^{n} \varepsilon_t \varepsilon_t^T$. It can prove that the sample covariance matrix converges to the covariance matrix Σ in probability.

According to the following theorem,

Theorem 3 *(Anderson, Brons and Jensen, 1983) If for a pair of square random matrices P_n, Q_n, (P_n, Q_n) converges in distribution to (P, Q), where Q is a.s. nonsingular, then the ordered solutions of the generalized eigenvalue problem $det(P_n - \lambda Q_n) = 0$ converge in distribution to the ordered solutions of the generalized eigenvalue problem $det(P - \lambda Q) = 0$.*

the eigenvalues of Σ can be estimated by the eigenvalues of the sample covariance matrix

$$\Sigma_n = \frac{1}{n}\sum_{t=1}^{n} \varepsilon_t \varepsilon_t{}^T.$$

When n is large enough, if the rank of Σ_n is less than p, then based on the result showed in Section 2, we can accept that $X_{1,t}$, $X_{2,t}$, $\cdots$, $X_{p,t}$ as cointergated; if all the eigenvalues of Σ_n are significantly different from 0, then we can reject $X_{1,t}$, $X_{2,t}$, $\cdots$, $X_{p,t}$ cointegrated; if some eigenvalues of Σ_n are not significantly different from 0, then there are two possibilities. One possibility is that $X_{1,t}$, $X_{2,t}$, $\cdots$, $X_{p,t}$ are not essentially cointegrated. The other possibility is that $X_{1,t}$, $X_{2,t}$, $\cdots$, $X_{p,t}$ are cointegrated but, due to the effect of sample size or the bias of ARIMA model fitting, the rank of Σ_n is still p .

Here we are interested in the situation that some eigenvalues of Σ_n are not significantly different from 0, because we need to understand how to make a decision if we can accept $X_{1,t}$, $X_{2,t}$, $\cdots$, $X_{p,t}$ cointegrated. In the following we replace Σ_n by Σ because given a data set the sample size n is fixed. It is not an important issue in the following discussion.

Consider Model (1) in Theorem 1 with a full rank matrix $var(\varepsilon_t) = \Sigma$. Assume that the last r eigenvalues of Σ are small enough to be negligible. Then

$$\varepsilon_t = A^* \begin{pmatrix} v^*_{1,t} \\ \vdots \\ v^*_{p-r,t} \\ v^*_{p-r+1} \\ \vdots \\ v^*_p \end{pmatrix} = (A^*_1, A^*_2) \begin{pmatrix} v^*_{1,t} \\ \vdots \\ v^*_{p-r,t} \\ v^*_{p-r+1} \\ \vdots \\ v^*_p \end{pmatrix}$$

$$= A^*_1 \mathbf{v}^*_{(1),t} + A^*_2 \mathbf{v}^*_{(2),t}.$$

Thus

$$(1-B)\mathbf{X}_t = C(B) v^*_{(1),1} + C_1(B) v^*_{(2),t}.$$

i.e.

$$\mathbf{X}_t = \mathbf{X}_0 + (1-B)\frac{C(B) - C(1)}{1-B}\sum_{i=1}^{t} \mathbf{v}^*_{(1),i}$$

$$+C(1)\sum_{i=1}^{t} \mathbf{v}^*_{(1),i} + C_1(B)\sum_{i=1}^{t} \mathbf{v}^*_{(2),i}. \quad (3)$$

Since the rank of $C(1)$ is $p - r$, there is a vector ξ such that $\xi^T C(1) = 0$ and

$$\xi^T \mathbf{X}_t = \xi^T(\mathbf{X}_0 - \mathbf{W}_0) + \xi^T \mathbf{W}_t$$

$$+\xi^T C_1(B)\sum_{i=1}^{t} \mathbf{v}^*_{(2),i}.$$

For given $\mathbf{X}_0$ and $\mathbf{W}_0$, $\xi^T(\mathbf{X}_0 - \mathbf{W}_0) + \xi^T \mathbf{W}_t$ is stationary and the difference between $\xi^T \mathbf{X}_t$ and $\xi^T(\mathbf{X}_0 - \mathbf{W}_0) + \xi^T \mathbf{W}_t$ is evaluated by $\xi^T C_1(B)\sum_{i=1}^{t} \mathbf{v}^*_{(2),i}$. If the variance of $\xi^T C_1(B)\sum_{i=1}^{t} \mathbf{v}^*_{(2),i}$ is not significant large, it may be reasonable to accept $\xi^T \mathbf{X}_t$ as a stationary time series, that is, $\mathbf{X}_t$ is accepted as cointegrated.

To distinguish the above "cointegration" from the exact definition of cointegration, we call it as "asymptotic cointegration".

Since "asymptotic cointegration" relations are not exact "cointegration" relation, the "cointergation" behaviours will be controlled by the factor of $\xi^T C_1(B)\sum_{i=1}^{t} \mathbf{v}^*_{(2),i}$, which, in fact, is significantly affected by sample size. The important question is how to measure the effect of $\xi^T C_1(B)\sum_{i=1}^{t} \mathbf{v}^*_{(2),i}$. In this paper, we will indicate a partial answer to the question and demonstrate the effect of $\xi^T C_1(B)\sum_{i=1}^{t} \mathbf{v}^*_{(2),i}$ via several examples. A detailed study will appear in our another papers.

In this paper we use the value of the variance of $\xi^T C_1(B)\sum_{i=1}^{t} \mathbf{v}^*_{(2),i}$ to measure the effect caused by $\xi^T C_1(B)\sum_{i=1}^{t} \mathbf{v}^*_{(2),i}$. First we deduce a formula for evaluating the variance of $C_1(B)\sum_{i=1}^{t} \mathbf{v}^*_{(2),i}$. To evaluate the variance of $C_1(B)\sum_{i=1}^{t} \mathbf{v}^*_{(2),i}$, the following notations are needed.

Let

$$f_i(B) = \sum_{k=0}^{\infty} \phi_{i,k} B^k,$$

$i = 1, 2$, are polynomial functions of B, where B is a back-shift operator. The inner product of f_1 and f_2 is defined as

$$(f_1(B), f_2(B)) = \sum_{k=0}^{\infty} \phi_{1,k}\phi_{2,k}.$$

Sometimes $(f_1(B), f_2(B))$ is briefly denoted as (f_1, f_2).

Now we consider $C_1(B)\sum_{i=1}^{t}\mathbf{v}^*_{(2),i}$ in (3). Let

$$H(B) = C_1(B)(1+B+B^2+\cdots+B^{t-1}),$$

where the element $h_{i,j}$ of $H(B)$ can be expressed as

$$h_{i,j} = \sum_{k=0}^{\infty}\phi_{i,j,k}B^k(\sum_{s=0}^{t-1}B^s)$$

$$= \sum_{s=0}^{t-1}(\sum_{k=0}^{s}\phi_{i,j,k})B^s + \sum_{s=t}^{\infty}(\sum_{k=s-t+1}^{s}\phi_{i,j,k})B^s.$$

Thus $cov(C_1(B)\sum_{i=1}^{t}\mathbf{v}^*_{(2),i}, C_1(B)\sum_{i=1}^{t}\mathbf{v}^*_{(2),i})$ can be expressed as $(\sum_{j=p-r+1}^{p}(h_{i,j},h_{k,j})\sigma_j^2)$ where

$$(h_{i,j},h_{k,j}) = \sum_{s=0}^{t-1}[(\sum_{m=0}^{s}\phi_{i,j,m})(\sum_{n=0}^{s}\phi_{i,k,n})]$$

$$+\sum_{s=t}^{\infty}[(\sum_{m=s-t+1}^{s}\phi_{i,j,m})(\sum_{n=s-t+1}^{s}\phi_{i,k,n})],$$

and $Var(\xi^{\mathbf{T}}C_1(B)\sum_{i=1}^{t}\mathbf{v}^*_{(2),i})$ can be evaluated by

$$\xi^{\mathbf{T}}cov(C_1(B)\sum_{i=1}^{t}\mathbf{v}^*_{(2),i}, C_1(B)\sum_{i=1}^{t}\mathbf{v}^*_{(2),i})\xi. \quad (4)$$

The formula (4) can be used to estimate the effect of $\xi^{\mathbf{T}}C_1(B)\sum_{i=1}^{t}\mathbf{v}^*_{(2),i}$. Example 2 below shows the connection between asymptotic cointegration and the effect of $\xi^{\mathbf{T}}C_1(B)\sum_{i=1}^{t}\mathbf{v}^*_{(2),i}$.

Example 2 Assume time series X_t and Y_t satisfy the following model

$$X_t = X_{t-1} + v_t - 0.2v_{t-1} \quad (5)$$
$$Y_t = Y_{t-1} + \sqrt{2}v_t + u_t \quad (6)$$

$t = 1,2,\cdots,T$, where $v_t \sim N(0,1)$ and $u_t \sim N(0,\sigma^2)$ with a small variance, and v_t and u_t are independent.

We express X_t and Y_t as follow:

$$(1-B)\begin{pmatrix} X_t \\ Y_t \end{pmatrix} = \begin{pmatrix} 1-0.2B & 0 \\ \sqrt{2} & 1 \end{pmatrix}\begin{pmatrix} v_t \\ u_t \end{pmatrix}$$

$$= C(B)\begin{pmatrix} v_t \\ u_t \end{pmatrix},$$

where

$$\Sigma = Var(\begin{pmatrix} v_t \\ u_t \end{pmatrix}) = \begin{pmatrix} 1 & 0 \\ 0 & \sigma^2 \end{pmatrix}.$$

In this particular example, the eigenvalues of Σ are 1 and σ^2 and the matrix A (in the proof of Theorem 1) is an identity matrix. By assuming that the variance for u_t is fairly small, we partition $C(B)$ as two parts, $C_1(B)$ and $C_2(B)$, that is

$$(1-B)\begin{pmatrix} X_t \\ Y_t \end{pmatrix} = \begin{pmatrix} 1-0.2B \\ \sqrt{2} \end{pmatrix}v_t + \begin{pmatrix} 0 \\ 1 \end{pmatrix}u_t$$

$$= C_1(B)v_t + C_2(B)u_t$$

For $C_1(B)$, we can determine a vector $\xi = (1,-0.5657)^T$ such that $\xi^{\mathbf{T}}\mathbf{C_1(1)} = \mathbf{0}$. For this vector ξ, the variance of the remaining term is evaluated as follows,

$$Var(\xi^T C_2(B)\sum_{i=1}^{t}u_i) \le Var(\xi^T C_2(B)\sum_{i=1}^{T}u_i)$$

$$= 0.32001649T\sigma^2.$$

It expects that, if $0.32001649T\sigma^2$ is smaller enough, we shall be able to accept that X_t and Y_t are cointegrated; if $0.32001649T\sigma^2$ is not small, the acceptance of cointegrated relationship between X_t and Y_t becomes critical. This conclusion is confirmed by the following simulating examples.

In the following several data sets are simulated from models (5) and (6) with different sample size and different value of σ. We try to use the following examples to indicate that the conclusion whether X_t and Y_t can be accepted as cointegrated within a certain time period can be judged by the variance of $\xi^{\mathbf{T}}C_1(B)\sum_{i=1}^{t}\mathbf{v}^*_{(2),i}$.

In the examples we also apply the Johansen's procedure (by using PcGive package) to each data set. The purpose of carrying out such practices is to use the outputs as benchmarks.

In the following, firstly, we simulate data from the model; secondly make conclusion based on the value of variance $\xi^{\mathbf{T}}C_1(B)\sum_{i=1}^{t}\mathbf{v}^*_{(2),i}$; then apply Johanson's procedure to data; finally summarize the examples.

Example 3 We simulated several samples with size 150, 950, 1000 and 2000 for $\sigma = 0.001$ and a sample with size 1000 for $\sigma = 0.01$ from models (5) and (6). The variance of

$$Var(\xi^T C_2(B)\sum_{i=1}^{t}u_i) \le Var(\xi^T C_2(B)\sum_{i=1}^{T}u_i),$$

for each case, is bounded by 0.04800, 0.304, 0.3200, 0.6400 and 3.2002 respectively. Having a quick look at these values, it expects that, for the first four samples, X_t and Y_t are coinegrated. However, for the last one, there is a risk to accept that X_t and Y_t are cointegrated. This conclusion is slightly different from the output of PcGive which indicates, at the level of 0.05 % for all the cases, X_t and Y_t

are accepted as cointegrated under Johansen's procedure although they are not in fact. The cointegrating vectors are $(1, -0.56369)$, $(1, -0.56887)$, $(1, -0.55250)$, $(1, -0.57223)$ and $(1, -0.42799)$ respectively. However after applying unit-root test to these cointegrating vector, it turns out X_t and Y_t should not be cointegrated for the last case. It coincides with the conclusion under the criterion given by this paper.

The following is an other example where x_t and y_t are essentially cointegrated. However due model fitting, sample and sample size, the sample covariance matrix turns out to be a full rank matrix. In this example we show how to use the procedure to identify the cointegrating vector for x_t and y_t.

Example 4 In this example we simulate a sample (x_t, y_t) with size 1000 from the model given in Example 1, where ϵ_t are independent normally distributed with mean 0 and variance 1.

Based on the sample, we fit the data by using the following model

$$(1-B)x_t = -0.021331 + \frac{\epsilon_{1,t}}{1+0.0074721B} \quad (7)$$

$$(1-B)y_t = -0.01701 + (1-0.20119B)\epsilon_{2,t} \quad (8)$$

and save residuals for each model. The estimate of the covariance matrix $cov(\epsilon_{1,t}, \epsilon_{2,t})$ is given by sample covariance

$$\hat{\Sigma} = \begin{pmatrix} 0.9745578557 & 0.9746402982 \\ 0.974602982 & 0.974770822 \end{pmatrix}$$

with eigenvalues 1.949930 and 0.00002. Use eigenvectors to determine the matrix A^* (mentioned the proof of Theorem 1) and obtain

$$\begin{pmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{pmatrix} = \begin{pmatrix} 0.707068 & 0.707145 \\ 0.707145 & -0.707068 \end{pmatrix} \begin{pmatrix} v_{1,t} \\ v_{2,t} \end{pmatrix}$$

where $var(v_{1,t}) = 1.949930$ and $var(v_{2,t}) = 0.00002$. Since the variance of $v_{2,t}$ is relatively small, we may consider it as unimportant factor. Then we express the model as below

$$(1-B)\begin{pmatrix} x_t \\ y_t \end{pmatrix} = \begin{pmatrix} -0.02131 \\ -0.01701 \end{pmatrix} + C(B)v_{1,t}$$

$$+C_1(B)v_{2,t}.$$

From $C(1) = (0.706540, 0.564877)^T$ and a vector $\xi = (-0.799498, 1)$ can be determined such that $\xi^T C(1) = 0$.

By applying the formula (4), we find that the variance of $\xi^T C_1(B) \sum_{i=1}^{t} v_{2,i}$ are reasonable small, which is less than 0.0507 for $0 \le t \le 1000$. Thus, from Theorem 2, for $0 \le t \le 1000$, it reasonable to accept that $\xi^T(\mathbf{X_t} - \mathbf{A_0}t) = (-0.07994898x_t + y_t) - 0.000027128t$ is stationary. As we can see the conintegrating vector $(-0.799498, 1)$ obtained in here is very close to the theoretical value $(-0.8, 1)$ given by Example 1 and the non random term only makes a small amount of contribution .

From this paper and examples, we can see the remaining term in (3) plays important role in the cointegration inference. Once the effect of the remaining term is precisely evaluated, it will give a clear picture that how long period a cointegration phenomenon, or, precisely say, asymptotic cointegration phenomenon, can be lasted. This issue is important, especially when the cointegration relation is applied to forecasting. In this paper, we only show the relation between the covariance matrix of error process and conitegrating vector(s), and the role of the remaining term in (3). Several other interesting questions still need to be answered. They will be carried out in our subsequence papers.

References

Anderson, S.A., H. K. Brons and S.T. Jensen(1983). Distribution of eigenvalues in Multivariate statistical analysis. Annals of Statistics, 11, 392-415.

Engle, R. F. and C. W. Granger(1987). Co-integrated and error correction: representation, estimation and testing. Econometrica, 55, 251-76.

Johansen, S. (1988). Statistical Analysis of cointegration vectors. Journal of Economic Dynamics and Control, 12. 231-54.

Harris, R.I.D. (1996). Using Cointegration Analysis in Econometric Modelling, Prentice Hall, New York.

Granger, C.W.J. and A. Weiss(1983). Time Series Analysis of Error Correction Models. In Karlin, S., T. Ameniya and L. Goodman, (eds) Studies in Econometrics, Time Series and Multivariate Statistics (New York, Academic Press), 255-278.

Engle, Robert F. and Byung Sam Yoo (1987). Forecasting and testing in cointegrated systems. Journal of Econometrics, 35, 143-159.

Bierens, Herman J. (1997). Nonparametric cointegration analysis. Journal of Econometrics, 77, 379-404.