Journal of Korean
Data & Information Science Society
2007, Vol. 18, No. 1, pp. 73 ~ 80

Analysis of Multivariate Financial Time Series Using Cointegration : Case Study[1)]

M. S. Choi[2)] · S. Y. Hwang[3)] · J. A. Park[4)]

Abstract

Cointegration(together with VARMA(vector ARMA)) has been proven to be useful for analyzing multivariate non-stationary data in the field of financial time series. It provides a linear combination (which turns out to be stationary series) of non-stationary component series. This linear combination equation is referred to as long term equilibrium between the component series. We consider two sets of Korean bivariate financial time series and then illustrate cointegration analysis. Specifically estimated VAR(vector AR) and VECM(vector error correction model) are obtained and CV(cointegrating vector) is found for each data sets.

Keywords : Cointegration, Error Correction Model, Vector AR

1. 서 론

시계열자료 분석의 목적은 자료의 생성구조에 대한 이해와 예측에 있다. 분석대상이 되는 시계열이 단변량 시계열 자료일 경우, 한 변수에 대해 시간에 따라 얻어진 값들이므로 예측모형도 자기의 과거와 현재의 값들만을 이용하게 된다. 그러나 실제의 경우 동시에 두 개 또는 그 이상의 변수들의 시계열에 관심을 갖고 각 시계열의 특성과 이 시계열들의 상호 관련성을 파악한 후 예측모형에 이용하는 것이 효과적일 때가 많다. 이와 같은 다변량 시계열 자료를 분석하는 목적은 좀 더 신뢰할 수 있는 예측시스템(forecasting system)을 얻기 위한 것과 여러 변수들 사이의 동적 관계(dynamic interrelationships)를 밝혀내는 것이다.

다변량 시계열 자료를 분석하기 위해서 여러 가지 다양한 방법들이 연구, 개발되고

---

1) This work was supported by a grant from 2006 Sookmyung Women's Univ. Support from BK 21 Project is also acknowledged.

2) Graduate student, Department of Statistics, Sookmyung Women's Univ.

3) Corresponding author : Professor, Department of Statistics, Sookmyung Women's Univ.
Email : shwang@sookmyung.ac.kr

4) Doctoral student, Department of Statistics, Sookmyung Women's Univ.

있다. 기초적인 모형은 Box와 Jenkins(1976)가 제안한 전이함수모형(transfer function model ; TFM)으로 입력계열(input series)과 출력계열(output series)이 인과관계(causal relationship)에 있는 경우 적용할 수 있다. 전이함수모형의 특별한 경우인 시계열회귀모형(regression time series model)은 가장 일반적으로 고려되는 모형이다. 그러나 실제 자료들에서는 시계열 변수들 사이에 복잡한 feedforward 와 feedback 관계가 존재하기 때문에 전이함수모형이나 시계열회귀모형만으로는 분석에 어려움이 있다. 변수들 사이의 복잡한 관계들을 모형화하는 방법으로는 벡터시계열(multivariate vector time series)에서 VARMA(vector ARMA), VAR(vector AR)모형 등이 주로 이용되며, 벡터시계열의 각 성분시계열들은 비정상(non-stationary)이지만 성분시계열들 간의 선형결합이 정상이 되는 경우가 존재할 경우, Engle and Granger(1987)에 의해 제안된 공적분(cointegration)모형이 이용된다. 본 논문에서는 공적분관계가 존재하는 국내 다변량 시계열들의 분석방법(공적분 분석)에 대하여 알아보고 실제 금융시계열 자료에 적용하고자 한다.

2. 벡터자기회귀(VAR ; Vector AR)모형

본 절에서 VAR 모형 및 공적분, 오차수정모형에 대한 기본 개념 및 용어는 김해경과 이명숙(2005), Wei(2006)를 참고하였으며 수식은 Wei(2006)를 중심으로 정리해 보았다. 벡터시계열에 대한 모형으로는 VAR(vector AR)모형, VMA(vector MA)모형 그리고 VARMA(vector ARMA)모형이 있다. 이 중에서 VAR모형은 다변량시계열을 모형화하는데 이용되는 가장 간편하고 적용성이 높은 모형이다. 이 모형은 벡터시계열을 이루는 한 성분 시계열의 현재의 값이 자기의 과거의 값과 함께 다른 성분 시계열들의 과거의 값들에 의해 영향을 받고 있음(feedforward)을 설명하는 모형으로서 각각의 성분 시계열은 다른 성분 시계열의 과거값에 상호 영향을 받는 특징이 있으며 이를 feedback 효과라 부른다. VAR 모형의 정의 식은 다음과 같다.

$$\left(\boldsymbol{I} - \boldsymbol{\Phi}_1\boldsymbol{B} - \cdots - \boldsymbol{\Phi}_p\boldsymbol{B}^p\right)\boldsymbol{Z}_t = \boldsymbol{\mu} + \boldsymbol{a}_t$$

$\boldsymbol{\mu}$: m-차원 벡터, $\boldsymbol{\Phi}_i$: $m \times m$ 행렬, $\boldsymbol{a}_t \sim i.i.d.(\boldsymbol{0}, \boldsymbol{\Sigma})$

정상성(stationarity)을 만족하기 위해서는 $\left|\boldsymbol{I} - \boldsymbol{\Phi}_1\boldsymbol{B} - \cdots - \boldsymbol{\Phi}_p\boldsymbol{B}^p\right| = 0$ 의 모든 근들이 1보다 커야한다. 정상조건을 만족한다면 위의 모형은 다음과 같이 표현할 수 있다. 여기서 $\boldsymbol{Z}_t^* = \boldsymbol{Z}_t - \boldsymbol{\mu}$ 을 의미한다.

$$\boldsymbol{Z}_t^* = \sum_{s=0}^{\infty} \boldsymbol{\Psi}_s \boldsymbol{a}_{t-s}$$

3. 공적분(Cointegration)

공적분은 Granger(1986)와 Engle과 Granger(1987)에 의해 도입된 개념으로 경제나

금융현상을 분석하는데 효과적으로 이용된다. 특히 벡터시계열에서 비정상인 개별 시계열들 사이의 장기균형(long-term equilibrium)관계 등을 설명하기 위해 적용되고 있다. 차수 1의 적분계열(integrated process of order 1)인 시계열 즉, $\{\boldsymbol{Z}_t\}$가 1차 차분에 의하여 정상시계열이 되는 경우를 $I(1)$으로, 그리고 정상시계열 즉, 시계열 $Z_t = \psi(B)a_t = \sum_{j=0}^{\infty} \psi_j a_{t-j}$ (이 때 $\psi(1) \neq 0$) 을 만족하는 경우인 차수 0의 적분계열(integrated process of order 0)을 $I(0)$로 표현하기로 하자. 이 때, m차원 벡터시계열 $\{\boldsymbol{Z}_t\}$, $Z_t = (Z_{1t}, Z_{2t}, \cdots, Z_{mt})'$ 의 각 성분계열들이 $I(1)$ 일 때, 상수벡터 $\boldsymbol{\beta} = (\beta_1, \beta_2, \cdots, \beta_m)' \ (\neq \boldsymbol{0})$가 존재하여 선형결합 $\{\boldsymbol{\beta}'\boldsymbol{Z}_t\}$가 정상, 즉

$$\boldsymbol{\beta}'\boldsymbol{Z}_t = \beta_1 Z_{1t} + \beta_2 Z_{2t} + \cdots + \beta_m Z_{mt} \sim I(0)$$

이면 벡터시계열 $\{\boldsymbol{Z}_t\}$는 벡터 $\boldsymbol{\beta}'\boldsymbol{Z}$ 로 공적분되었다고 한다. 여기서, 벡터 $\boldsymbol{\beta}$를 공적분벡터(CV : cointegrating vector)라 한다. 주로 표준화된 공적분벡터

$$\boldsymbol{\beta}^* = \left(1, -\frac{\beta_2}{\beta_1}, -\frac{\beta_3}{\beta_1}, \cdots, -\frac{\beta_m}{\beta_1}\right)' = \left(1, -\beta_2^*, -\beta_3^*, \cdots, -\beta_m^*\right)'$$

이 이용되며 이때의 공적분 관계는 다음과 같은 선형회귀식이 된다.

$$Z_{1t} = \beta_2^* Z_{2t} + \beta_3^* Z_{3t} + \cdots + \beta_m^* Z_{mt} + a_t \quad , \quad a_t \sim I(0)$$

$\{a_t\}$: 공적분잔차(cointegrating residual)

각 시계열들 사이에 공적분관계가 존재한다는 것은 단위근을 갖는 비정상시계열들은 각기 자신만의 고유한 확률추세를 갖지만 이 시계열들의 선형결합인 공적분잔차는 정상이 되어 시간이 흐름에 따라 그 값이 평균으로 회귀하는 것을 의미한다. 즉, 비정상시계열들이 어떠한 공통의 확률추세를 갖게 되는 것이다. 실제 경제나 금융 분야의 연구에 따르면 소득(income)과 소비(consumption), 선물이자율(forward rate)과 현물이자율(spot rate), 주가지수와 주가지수선물, 두 나라의 각각의 GDP, 같은 업종의 주가지수 등은 확률적 추세를 공유하는 공적분관계를 갖는다고 알려져 있다. 경제학적 관점에서 공적분관계가 있는 시계열들 사이에는 장기균형(long-term equilibrium)이 존재한다고 하며 공적분잔차(오차항) $\{a_t\}$를 경제체제 내의 균형력(equilibrium force)으로 해석한다.

4. 오차수정모형(Error Correction Model ; ECM)

공적분은 벡터시계열의 개별 계열 사이의 균형관계를 설명하는 개념으로서 한 특정 시점이 아닌 일정기간 동안 문제의 상황이 변하지 않는다는 가정하에 여러 기간에 걸쳐 성립하는 성질이다. 따라서 공적분 개념은 정적(static)개념으로 생각할 수 있으며,

이러한 공적분을 여러 기간에 걸쳐 서로 다르게 발생하는 성질들을 상호 결합하여 설명하는 동적(dynamic)인 개념을 표현하는 것으로 오차수정모형(ECM)(Davidson et al.(1978))을 생각할 수 있다. Engle과 Granger(1987)에 의해 제안된 공적분관계에 있는 벡터시계열들의 오차수정모형(VECM)은 다음과 같다.

$$\Delta \boldsymbol{Z}_t = \boldsymbol{\theta}_0 - \boldsymbol{M}\boldsymbol{Y}_{t-1} + \boldsymbol{\Phi}_1^{*}\Delta \boldsymbol{Z}_{t-1} + \cdots + \boldsymbol{\Phi}_{p-1}^{*}\Delta \boldsymbol{Z}_{t-p+1} + \boldsymbol{a}_t$$

여기서

$\boldsymbol{M}_{m\times k}$: 오차수정계수(error correction coefficient)
$\boldsymbol{Y}_{t-1} = \boldsymbol{\beta}'\boldsymbol{Z}_{t-1}$: $(k\times 1)$인 정상계열.

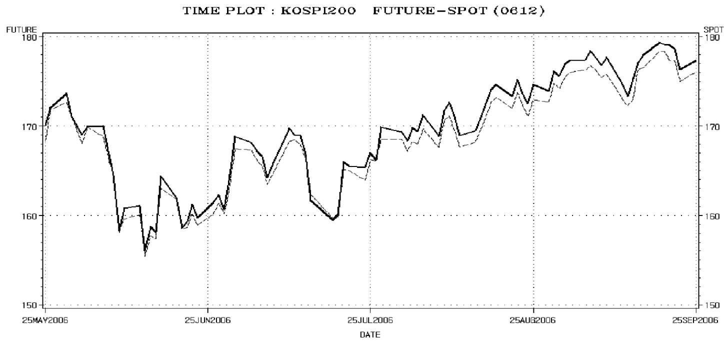
이 모형은 공적분관계가 존재하는 벡터시계열 $\{\boldsymbol{Z}_t\}$의 차분계열 $\{\Delta \boldsymbol{Z}_t\}$은 과거시차차분(past lagged differences) $\Delta \boldsymbol{Z}_j$, $j < t$ 의 값들만을 가지고는 설명될 수 없으며 반드시 오차수정(error-correction)항인 $\boldsymbol{M}\boldsymbol{Y}_{t-1} = \boldsymbol{M}\boldsymbol{\beta}'\boldsymbol{Z}_{t-1}$ 을 포함하고 있어야함을 뜻한다. $\Delta \boldsymbol{Z}_t$ 와 $\Delta \boldsymbol{Z}_j$, $j < t$ 값들 간의 장기균형관계를 고려할 때, $\boldsymbol{Y}_{t-1} = \boldsymbol{\beta}'\boldsymbol{Z}_{t-1}$ 은 균형관계로부터의 오차라고 볼 수 있으며, 이 때 계수행렬 $\boldsymbol{M}$ 은 이러한 오차를 수정하는 역할을 한다. 결국, 오차수정모형은 시계열 $\{\boldsymbol{Z}_t\}$가 장기균형상태에서 벗어났을 때 그 변수들이 공적분의 단기동적조절기능을 통하여 어떻게 반응하여 장기균형관계를 유지하는지를 설명한다. 이런 관점에서 이 모형은 동적모형(dynamic model)이다(cf. 박준용 et al.(2002)).

5. 국내 금융시계열 사례분석

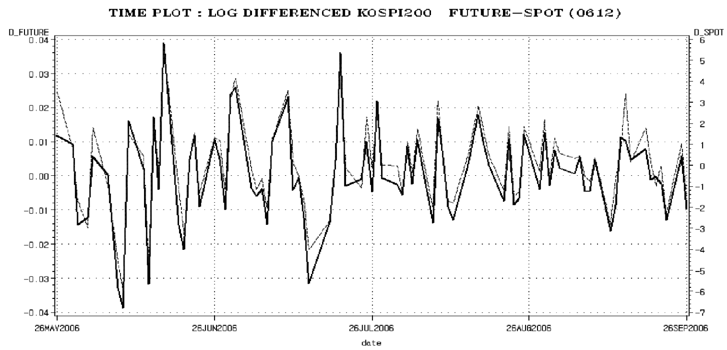
이 장에서는 기존의 계량경제 분야에서 공적분관계가 존재한다고 알려진 시계열 변수들 중 KOSPI200지수와 KOSPI200선물지수의 이변량시계열자료와 KOSPI와 일본의 주가지수 중의 하나인 Nikkei225의 이변량시계열자료를 이용하여 먼저 VAR모형에 적합 시켜 본 후 다시 단위근검정, 공적분검정을 수행한 다음, 벡터오차수정모형(VECM)에 적용해 보았다. 자료 분석에는 SAS/ETS의 PROC VARMAX 프로그램을 사용하였다.

사례분석 I - KOSPI200지수와 KOSPI200선물지수

분석에 사용한 자료는 KOSPI200의 2006년 12월물에 대한 현물지수와 선물지수에 대한 2006년 5월 25일부터 10월 31일까지의 일별자료이다. 두 계열에 대한 시계열도는 <그림 1>과 같다. 그림에서 실선은 선물지수를 나타내며 점선은 현물지수를 나타내고 있다. 그림에서 확인할 수 있는 것처럼 시간이 지남에도 안정적 성향이 나타나지 않고 불규칙하게 변동하므로 정상성을 만족하지 않기 때문에 VAR모형에 적합시키기 위해 두 계열을 로그차분시킴으로서 정상적이 되도록 하였다.



<그림 1> KOSPI200 선물지수(FUTURE)와 현물지수(SPOT)에 대한 시도표



<그림 2> 로그차분 후의 KOSPI200 선물지수와 현물지수에 대한 시도표

VAR(2)에 자료를 적합시켜서 얻은 모형식은 다음과 같다.(선물지수 : $\boldsymbol{F_t}$, 현물지수 : $\boldsymbol{S_t}$)

$$\begin{bmatrix} F_t \\ S_t \end{bmatrix} = \begin{bmatrix} -0.47 & 0.53 \\ 0.19 & 0.11 \end{bmatrix} \begin{bmatrix} F_{t-1} \\ S_{t-1} \end{bmatrix} + \begin{bmatrix} -0.89 & 0.74 \\ 0.62 & 0.52 \end{bmatrix} \begin{bmatrix} F_{t-2} \\ S_{t-2} \end{bmatrix}$$

모형 타당성 확인을 위한 포트맨토 검정 결과 유의수준 0.05에서 적합된 모형이 알맞다는 것을 확인할 수 있었으나 잔차에 대한 정규성 검정 결과, 잔차가 정규분포를 따른다는 가정에 문제가 있으며, 각 계수들 역시 모두 유의하지 않다는 결과를 얻었다. 따라서 오차항이 정규분포를 따른다는 가정에 문제가 있거나 혹은 자료의 시계열들을 정상적으로 만들었다고 할지라도 VAR모형에 적합시키는 것에 문제가 있다. 앞의 <그림 1>에서 두 계열이 동행하는 관계를 보이며, 두 계열 모두 비정상시계열이므로 KOSPI200 선물지수와 현물지수를 모형화하기 위해서 공적분관계를 이용한 모형

을 고려해 볼 수 있다.

공적분검정에 앞서 두 시계열이 각각 확률적 추세를 갖는 비정상시계열임을 확인하기 위해 각각에 대한 단위근검정 결과 선물지수와 현물지수 두 시계열모두 단위근을 갖는 비정상시계열임을 확인할 수 있다. 또한, 두 시계열의 선형결합이 정상시계열이 되는 공적분관계를 갖고 있는지를 검정하기 위해 Johansen의 대각합 통계량(Johansen's trace statistic)을 이용한 검정 결과, $\alpha = 0.05$에서 두 시계열 사이에 rank=1의 공적분관계가 존재함을 확인할 수 있다.

공적분관계가 존재하므로 선물지수와 현물지수 두 시계열의 벡터오차수정모형(VECM)의 적합 결과 얻어지는 모형식은 다음과 같다.(△기호는 1차 차분을 의미)

VECM(4) :

$$\begin{bmatrix} \triangle F_t \\ \triangle S_t \end{bmatrix} = \begin{bmatrix} 1.41 & -1.47 & 9.36 \\ 1.80 & -1.88 & 11.94 \end{bmatrix} \begin{bmatrix} F_{t-1} \\ S_{t-1} \\ 1 \end{bmatrix} + \begin{bmatrix} -1.69 & 1.76 \\ -1.25 & 1.32 \end{bmatrix} \begin{bmatrix} \triangle F_{t-1} \\ \triangle S_{t-1} \end{bmatrix}$$
$$+ \begin{bmatrix} -1.88 & 1.78 \\ -1.64 & 1.56 \end{bmatrix} \begin{bmatrix} \triangle F_{t-2} \\ \triangle S_{t-2} \end{bmatrix} + \begin{bmatrix} -0.21 & 0.02 \\ -0.22 & 0.06 \end{bmatrix} \begin{bmatrix} \triangle F_{t-3} \\ \triangle S_{t-3} \end{bmatrix}$$

모형에서 두 시계열의 공통 확률추세를 살펴보기 위해 공적분벡터만을 따로 분리하여 살펴보면 다음과 같다.

$$\beta' = (1, -1.05, 6.64)$$

선물지수를 반응변수처럼 보기 위해 선물지수를 1로 표준화시켰으며 위의 공적분벡터에 의해 선물지수와 현물지수간의 확률적 추세 관계는 다음과 같이 설명될 수 있다

$$F_t = 1.05 S_t - 6.64$$

현대 투자이론의 연구 결과에 의하면 현시점에서의 선물가격(F_0)과 현물가격(S_0) 사이에는 다음 관계가 성립함이 알려져 있다. R_f 는 이자율이며, D 는 기간 중 배당수입을 나타낸다(장영광, 2006).

$$F_0 = S_0(1 + R_f) - D$$

위의 관계식을 고려해보았을 때, 적합된 벡터오차수정모형이 선물지수와 현물지수 자료를 분석하는데 있어서 기존의 연구 결과에 어느 정도 부합한다고 볼 수 있다.

사례분석Ⅱ - KOSPI지수와 Nikkei225지수

2000년 1월부터 2006년 1월까지의 일별 KOSPI지수와 Nikkei225지수를 VAR(2) 모형에 적합 시킨 결과는 다음과 같다.

$$\begin{bmatrix} KOSPI_t \\ NK225_t \end{bmatrix} = \begin{bmatrix} 5.45 \\ -89.64 \end{bmatrix} + \begin{bmatrix} -0.09 & -0.01 \\ 1.59 & -0.03 \end{bmatrix} \begin{bmatrix} KOSPI_{t-1} \\ NK225_{t-1} \end{bmatrix} + \begin{bmatrix} 0.16 & 0 \\ 3.91 & -0.04 \end{bmatrix} \begin{bmatrix} KOSPI_{t-2} \\ NK225_{t-2} \end{bmatrix}$$

포트맨토 검정 결과, 추정된 모형은 알맞다고 볼 수 있으나 앞의 사례연구에서와 마찬가지로 각 계수들에 대한 유의성 검정에서는 추정된 계수들이 유의하지 않다. 잔차에 대한 정규성 검정 결과 오차항에 대한 정규분포 가정에는 문제가 없다고 나타났으나 Nikkei225 계열에 대한 ARCH검정 결과 ARCH효과를 가지고 있다고 판단된다. 시도표(생략)를 보면 두 시계열간에 공적분관계가 존재한다고 생각되므로 단위근검정, 공적분검정을 수행한 다음, 벡터오차수정모형(VECM)에 적합시켜 보았다. 적합된 모형식은 다음과 같다.

VECM(1) :

$$\begin{bmatrix} \triangle KOSPI_t \\ \triangle NK225_t \end{bmatrix} = \begin{bmatrix} 0.0610 & -0.0033 & -10.0233 \\ 1.5247 & -0.0820 & -250.5322 \end{bmatrix} \begin{bmatrix} KOSPI_{t-1} \\ NK225_{t-1} \\ 1 \end{bmatrix}$$

KOSPI와 Nikkei225사이의 공통 확률추세를 보기 위해 공적분벡터 부분을 분리하여 보면 다음과 같으며 KOSPI변수를 종속변수로 고려하기 위해 1로 표준화시켰다.

$$\boldsymbol{\beta}' = (1, -0.0538, -164.3190)$$

즉, KOSPI와 Nikkei225는 다음과 같은 확률적 추세 관계를 유지하고 있다.

$$KOSPI_t = 0.0538\, NK225_t + 164.3190$$

6. 결 론

분석하고자 하는 변수가 두 개 이상인 다변량 시계열 자료의 경우 일반적으로 전이함수모형이나 시계열 회귀분석을 사용할 수 있다. 그러나 변수들 간에 인과관계만을 고려하는 전이함수모형보다는 변수들 간의 다양한 관계를 살펴볼 수 있는 VAR모형이나 성분 시계열들의 장기적, 단기적 관계를 함께 설명해 줄 수 있는 오차수정모형 등이 분석에 더 유용하다고 보여진다.

본 논문에서는 실제 우리나라의 다변량 시계열 자료를 이용하여 VAR모형과 오차수정모형을 적합시켜 보았다. VAR모형을 적합시에는 포트맨토 검정 결과와 계수의 유의성 검정 결과가 서로 다른 적합성을 보여주었다. KOSPI200선물지수와 현물지수에 대해 VAR모형을 적합시킬 때는, 오차항의 분포를 t 분포와 같이 정규분포보다 꼬리가 두꺼운 분포를 가정한 후 분석하는 것이 필요하며 KOSPI와 Nikkei225의 VAR

모형에 적합시키는 경우에는 Nikkei225 계열의 오차항에 ARCH효과를 고려한 항을 포함하여 분석하는 것이 더 필요할 것으로 생각된다. 확률적 추세를 갖는 비정상시계열의 경우에는 VAR모형에 적합시키기 위해 차분 등의 정상화 과정을 거치면서 본래 자료에 존재하는 확률적 추세를 작위적으로 제거하여 정보의 손실을 가져올 수 있다. 따라서 분석에 이용한 자료들의 경우 확률적 추세를 가지고 있다고 판단되므로 VAR 모형보다는 오차수정모형에 적합 시키는 것이 더 바람직하다고 판단된다.

감사의 글

본 논문을 심사해주신 세 분의 심사위원께 감사드립니다. 본 연구는 2006년도 숙명여자대학교 교내 연구비 및 BK21 사업의 지원에 의한 것입니다.

참 고 문 헌

1. 김해경, 이명숙 (2005). 경제 및 금융자료를 위한 시계열분석, 경문사.
2. 박준용, 장유순, 한상범(2002). 경제시계열분석. 세경사.
3. 장영광 (2006). 증권투자론, 신영사.
4. Box, G. E. P. and Jenkins, G. M. (1976). *Time Series Analysis Forecasting and Control*, 2nd ed., Holden-Day, San Francisco.
5. Davidson, J. E. H., Henry, D. F., Srba, F. and Yeo, S. (1978). Econometric modeling of the aggregate time series relationship between consumer's expenditure and income in the United Kingdom, *Economic Journal*, **88**, 661-692.
6. Engle, R. F. and Granger, C. W. J. (1987). Co-integration and error correction representation, estimation and testing, *Econometrica,* **55**, 251-276.
7. Granger, C. W. J. (1986) Developments in the study of co-integrated economic variables, *Oxford Bulletin of Economics and Statistics*, **48**, 213-228.
8. Nelson, C. R. and Plosser, C. I. (1982). Trends and random walk in macroeconomic time series : some evidence and implications, *Journal of Monetary Economics*, **10**, 139-162.
9. Wei, W. W. S. (2006). *Time Series Analysis*, 2nd ed., Pearson International Edition.

[2007년 1월 접수, 2007년 2월 채택]