ON SOME TRANSFORMATIONS FOR ACHIEVING STATIONARITY AND INVERTIBILITY IN COINTEGRATED SYSTEMS

Mitsuhiro Odaki

1. Introduction

In spite of some recent development of the asymptotic theory for nonstationary time series, stationarity property is still one of useful requirements for the statistical analysis of time series. On the other hand, invertibility property, which guarantees the derivation of an autoregressive representation with infinite or finite order, is also convenient for modeling and forecasting in time series. It is widely accepted that many time series are not stationary but can be transformed into stationary series by differencing. As for a univariate time series, the differenced series is also invertible if an appropriate degree of differencing is applied to the original series. However, when the data transformation by differencing is used for a multiple time series, invertibility of the transformed series is not necessarily assured even if the original series is not overdifferenced.

The concept of cointegration, which has been studied by Engle and Granger (1987), Engle and Yoo (1987), Stock and Watson (1988) and others, defines multiple time series such that: (i) each component is nonstationary; but (ii) there exists at least one linear combination of the components that is stationary. At the same time, it formulates that the differenced series are not invertible when they are considered as multiple systems. In conintegrated multiple time series, it is impossible to derive a vector autoregressive (VAR) representation validly by differencing. As shown by Engle and Granger (1987), one of useful representations we should derive is not a VAR but an error correction model (ECM), in which all variables are stationary. However, the ECM is considerably inconvenient for spectral analysis and multi-period forecasting. We may face some situations which need to derive a stationary VAR representation by applying other transformations than differencing to the original series.

The purpose of this paper is to examine what transformation should be applied to the original data of cointegrated multiple time series in order to achieve stationarity and invertibility. For this, we first extend the concept of differencing. Based on the extended concept, we provide a class of transformations and show that any transformation of this class yields a stationary VAR representations. Next, we consider how such transformations are estimated in practice. It is explained that some transformations can be easily constructed from the estimates of the cointegrating vectors. As another useful method, the estimation of the first order serial correlation matrix of the original series is considered. It is shown that its estimator converges in probability to a matrix which forms one of such transformations. It is also emphasized that for large samples the transformation based on this estimator leads to a stationary VAR representation whether the multiple time series is cointegrated or not.

In Section 2, using Engle and Granger's (1987) difinitions and notations, the concept of cointegration and its related time seires representations are presented. The class of transformations to achieve stationarity and invertibility is formaulated in Section 3. Some of such transformations are concretely presented. Section 4 deals with the methods to estimate these transformations. In Section 5, the results established in this paper are summarized. Proofs of the theorems and lemma presented in Section 3 and 4 are given in the Appendix.

2. Characterization of Cointegrated Multiple Time Series

Let y_t denote an $N\times 1$ multiple time series that is cointegrated with cointegrating rank r (provided $0<r<N$). For the sake of simplicity, suppose that each component of y_t is stationary after differencing once. Then, following Engle and Granger's (1987) definitions and notations, y_t is expressed as

$$(2.1) \qquad (1-B)y_t=C(B)\varepsilon_t,$$

where B is the backward shift operator, $C(\lambda)=\Sigma_{j=0}^{\infty}C_j\lambda^j$ with $C(0)=I_N$ (the $N\times N$ identity matrix) and the unobservable time series ε_t are i. i. d. with mean zero and positive definite covariance matrix Ω. Further, in (2.1), all zeroes of $\det[C(\lambda)]$ lie on or outside the unit circle, $\Sigma_{j=1}^{\infty}|C_j|<\infty$ and rank $C(1)=k$, where $k=N-r$. That the rank of $C(1)$ is k implies that the differenced series y_t-y_{t-1} is not invertible.

In order to facilitate the derivation of another representation of y_t, it is assumed throughout the paper that $y_s=\varepsilon_s=0$ for $s\leq 0$, which can be interpreted as the conventional initial condition. Also, assume that the components of ε_t have finite fourth moments:

$$E|\varepsilon_{it}\varepsilon_{jt}\varepsilon_{ht}\varepsilon_{st}|\leq\mu<\infty, \quad i,\ j,\ h,\ s=1,\ \ldots,\ N.$$

where ε_{it} is the i-th component of ε_t. This assumption is required to establish the consistency properties of estimates presented in Section 4. Then, (2.1) is also written as

$$y_t = C(1) \sum_{s=1}^{t} \varepsilon_s + C^*(B)\varepsilon_t, \tag{2.2}$$

where $C^*(\lambda) = (-1)\Sigma_{j=0}^{\infty}\Sigma_{h=j+1}^{\infty} C_h \lambda^j$.

Engle and Granger (1987) proved that the following representation is derived from (2.1)

$$A(B)y_t = d(B)\varepsilon_t, \tag{2.3}$$

where $A(\lambda)$ and $d(\lambda)$ are defined as

$$\mathrm{Adj}[C(\lambda)] = (1-\lambda)^{r-1}A(\lambda)$$

$$\det[C(\lambda)] = (1-\lambda)^r d(\lambda).$$

Then, it is shown that there exist $N \times r$ matrices, A, Γ of rank r such that $A'C(1)=0$, $C(1)\Gamma=0$ and $A(1)=\Gamma A'$. The row vectors of A' are exactly r linearly independent cointegrating vectors. In this paper, we assume that $d(\lambda) \neq 0$ for $|\lambda| \leq 1$, which implies that zeroes of $\det[C(\lambda)]$ lying on the unit circle are restricted to one. This seems to be natural assumption in the framework of cointegrated multiple time series.

Defining $A^*(\lambda)$ as an $N \times N$ matrix satisfying $A(\lambda) = A^*(\lambda)(1-\lambda) + A(1)\lambda$, (2.3) is straightforwardly rewritten as an ECM:

(2.4) $$A^*(B)(1-B)y_t+\Gamma A' y_{t-1}=d(B)\varepsilon_t$$

It should be noted that y_t-y_{t-1}, $A'y_{t-1}$ and ε_t are stationary series. (2.4) is different from VAR representations because of the existence of the term $\Gamma A' y_{t-1}$.

3. Transformations for Achieving Stationarity and Invertibility

Let R denote an $N \times N$ matrix such that $R \neq 0$. For y_t generated by (2.1), the transformation expressed as $y_t - Ry_{t-1}$ may be interpreted as an extension of the concept of differencing. In this paper, we restrict our analysis to such transformations.[1] The following theorem formulates a class of transformations which leads to a stationary and invertible series from y_t.

THEOREM 1: *Suppose that for y_t generated by* (2.1), *the assumptions given in Section* 2 *are satisfied. Then, a necessary and sufficient condition for $y_t - Ry_{t-1}$ to be stationary and invertible is that R satisfy the following conditions* (i) *and* (ii):

(i) *There exists an $N \times r$ matrix F such that $R = I_N - FA'$.*

(ii) *R is expressed as the sum of $N \times N$ matrices R_1 and R_2 such that $R_1 = P(A)D$,* $\det(I_N - R_2\lambda) \neq 0$ for $|\lambda| \leq 1$ *and $R_2R_1 = 0$, where $P(A) = I_N - A(A'A)^{-1}A'$* and *$D$ is an $N \times N$ matrix.*

Proofs of the theorems and lemma in this paper are given in the Appendix.

We shall call the class of R which satisfy the conditions (i) and (ii) of the above theorem T hereafter. Also, note that the condition (ii) implies that $(I_N - R\lambda) = (I_N - R_2\lambda)(I_N - R_1\lambda)$.

This theorem is not so practical. Now, we shall formulate R of T more concretely and derive some time series representations for the transformed series $y_t - Ry_{t-1}$. First, as some candidates of R, consider a class of matrices which are expressed as

$$(3.1)\qquad G_1 = P(A) + P(A)DQ(A),$$

where $P(A)$ is given in the above theorem, $Q(A) = A(A'A)^{-1}A'$ and D is an $N \times N$ matrix. Letting $R_1 = G_1$ and $R_2 = 0$, it can be easily checked that Theorem 1's (i) and (ii) are satisfied. Let $x_t = y_t - G_1 y_{t-1}$ and $H_1 = I_N - G_1$. Noting that $(1-\lambda)I_N = (I_N - H_1\lambda)(I_N - G_1\lambda)$ and $A(1)y_t = A(1)x_t$, it follows from (2.4) that the stationary series x_t possesses

$$(3.2)\qquad [A^*(B)(I_N - H_1 B) + A(1)B]x_t = d(B)\varepsilon_t.$$

since all zeroes of $d(\lambda)$ are assumed to lie outside the unit circle, (3.2) leads to a VAR representation for x_t. Similarly, noting that $Q(A)C(1) = 0$ and $I_N - G_1\lambda = (1-\lambda)I_N - [I_N + P(A)D]Q(A)$, from (2.2), we derive the multivariate Wold representation for x_t:

$$(3.3)\qquad x_t = [C(1) + (I_N - G_1 B)C^*(B)]\varepsilon_t$$

Next, we construct R such that $R_2 \neq 0$. Suppose that there exists an $N \times N$ matrix G_2 satisfying

$$(3.4)\qquad G_2 = LA',\ \det(I_N - G_2\lambda) \neq 0 \text{ for } |\lambda| \leq 1,$$

where L is an $N \times r$ matrix.[2)] For G_1 and G_2 given in (3.1) and (3.4) respectively, let $G=G_1+G_2$ and $z_t=y_t-Gy_{t-1}$. It is obvious that G is included in T. Then, noting that $z_t=x_t-G_2x_{t-1}$, (3.2) leads to a VAR representation for z_t:

$$(3.5) \qquad [A^*(B)(I_N-H_1B)+A(1)B](I_N-G_2B)^{-1}z_t=d(B)\varepsilon_t$$

Similarly, from (3.3)

$$(3.6) \qquad z_t=(I_N-G_2B)[C(1)+(I_N-G_1B)C^*(B)]\varepsilon_t$$

(3.6) is the multivariate Wold representation for z_t. We note that $A^*(0)=A(0)=I_N$ in (3.2) and (3.5) and that $C(1)+C^*(0)=I_N$ in (3.3) and (3.6).

4. Some Estimation Methods

Let us consider how R of T formulated in the former section can be estimated in practice. (3.1) and (3.4) suggest that G_1 and G may be estimated based on the estimates of A. For example, letting $D=0$ in (3.1), $G_1(=P(A))$ is calculated from A only. However, the rows of A' are the conintegrating vectors, and it is impossible to determine them uniquely unless some normalizations are imposed.

Engle and Granger (1987) and Stock (1987) discussed some normalizations and estimation methods for A. The approach proposed in this paper is different from them. Before starting the estimation, we shall choose one of A. For this, put $C(1)'=(c_{\cdot 1}, \ldots, c_{\cdot N})$, where $c'_{\cdot j}$ is the j-th row vector of $C(1)$, which is given in (2.1) or (2.2). Since rank $C(1)=k$, there must exist c_{i_s}, $s=1, \ldots, k$ and an $r\times k$ matrix M such that rank $[c_{\cdot i_1}, \ldots, c_{\cdot i_k}]=k$ and $[c_{\cdot j_1}, \ldots, c_{\cdot j_r}]'=M[c_{\cdot i_1}, \ldots, c_{\cdot i_k}]'$, where $\{i_1, \ldots, i_k, j_1, \ldots j_r\}$ is a permutation of $\{1, 2, \ldots, N\}$. Put

$$S_1=(e_{j_1}, \ldots, e_{j_r})', \ S_2=(e_{i_1}, \ldots, e_{i_k})' \text{ and } S=[S_1' \vdots S_2']',$$

where e_j' is the j-th row vector of I_N. Also, for y_t generated by (2.1), define

$$y_{it}=S_i y_t \quad (i=1,2), \ \psi_t=y_{1t}-My_{2t} \text{ and } \varphi_t=y_{2t}-y_{2t-}-y_{2t-1}.$$ [3)]

Obviously, ψ_t is stationary, and the row vectors of $[I_r \vdots -M]$ are the cointegrating vectors. On the other hand, y_{2t} is not cointegrated and φ_t is a k-dimensional stationary and invertible series. Now, one of A' is given as $[I_r \vdots -M]S$.

Consider

$$(4.1) \qquad P(M)=S'\begin{bmatrix} M \\ I_k \end{bmatrix}[C \vdots I_k-CM]S, \text{ with an } k\times r \text{ matrix } C.$$

Noting that $SS'=I_N$, it is easily checked that $P(M)P(M)=P(M)$ and $[I_r \vdots -M]SP(M)=0$. Therefore, $P(M)$ can be considered as $P(A)$. Since $P(A)$ is G_1 as $D=I_N$ or 0, it is obvious that $P(M)$ satisfies the conditions of Theorem 1.

If we let $C=0$ in (4.1), we derive

$$(4.2) \qquad y_t-P(M)y_{t-1}=S'\begin{pmatrix} \psi_t-M\varphi_t \\ \varphi_t \end{pmatrix},$$

where gives an implication of the series transformed by R of T.[4)]

M can be consistently estimated by regressing y_{1t} on y_{2t}, provided that S is known. It is already shown by the results established in Engle and Granger (1987) or Stock (1987).[5)] Given C arbitrarily, $P(M)$ can be estimated using such estimates of M.

The testing method for cointegrated systems Stock and Watson (1988) proposed is constructed based on the estimates of the first order serial correlation matrix of y_t. As another useful method to estimate R

of T, we also focus our attention on it. On the basis of a sample $y_1, \ldots, y_T$ for y_t generated in (2.1), it is usually given as

$$\text{(4.3)} \qquad \hat{R} = \left(\sum_{t=2}^{T} y_t y'_{t-1}\right)\left(\sum_{t=2}^{T} y_{t-1} y'_{t-1}\right)^{-1},$$

which may be considered as the ordinary least squares (OLS) estimator derived by regressing y_t on y_{t-1}. Now, we are interested in whether the probability limit of $\hat{R}$ satisfies the conditions of Theorem 1 or not.

Before investigating the asymptotic perfrmance of $\hat{R}$, consider the following matrix $\bar{R}$:

$$\text{(4.4)} \qquad \bar{R} = \bar{R}_1 + \bar{R}_2, \text{ with } \bar{R}_1 \text{ and } \bar{R}_2 \text{ such that}$$

$$\bar{R}_1 = S' \begin{bmatrix} M \\ I_k \end{bmatrix} [R_2(1) R_1(0)^{-1} \mid I_k - R_2(1) R_1(0)^{-1} M] S,$$

$$\bar{R}_2 = S' \begin{bmatrix} R_1(1) R_1(0)^{-1} \\ 0 \end{bmatrix} [I_r \mid -M] S,$$

where $R_1(0) = E[\psi_t \psi'_t]$, $R_1(1) = E[\psi_{t+1} \psi'_t]$ and $R_2(1) = E[\varphi_{t+1} \psi'_t]$. Since $\bar{R}_1$ is $P(M)$ as $C = R_2(1) R_1(0)^{-1}$, $\bar{R}_1$ must satisfy the condition (ii) of Theorem 1. Also, it follows immediately that $\bar{R}_2 \bar{R}_1 = 0$. In the following lemma, it is shown that $\bar{R}_2$ can be considered as G_2 given in (3.4).

LEMMA: *For $\bar{R}_2$ given in* (4.4), $\det(I_N - \bar{R}_2 \lambda) \neq 0$ *for* $|\lambda| \leq 1$.

From the above results, we can see that $\bar{R}$ is included in T. As for such $\hat{R}$ and $\bar{R}$, we establish the following theorem.

THEOREM 2: *Suppose that for y_t generated by* (2.1), *the assumptions given in Section* 2 *are satisfied. Also, suppose that $\hat{R}$ and $\bar{R}$ are given in* (4.3) *and* (4.4) *respectively. Then,* $\hat{R} - \bar{R} = O_p(T^{-1/2})$.

In the above theorem, that $\hat{R}$ converges in probability to $\bar{R}$ at the rate $T^{1/2}$ is established. It implies that for the cointegrated multiple time series y_t, the transformation based on $\hat{R}$ yields a stationary VAR representation in large samples. When y_t is not cointegrated (rank $C(1)=k$), y_t may be considered as y_{2t}. Then, from the results established in Phillips and Durlauf (1986) and others, it is shown that $\hat{R}$ converges in probability to I_N. This implies that $y_t-\hat{R}y_{t-1}$ tends to a stationary and invertible series as $T\rightarrow\infty$ whether y_t is cointegrated or not.

5. Summary

The transformation of nonstationary time series into a stationary and invertible series is still useful in many aspects. As such a transformation, differencing is usually used. However, in cointegrated multiple time seris, any stationary and invertible series never be brought by differencing. In this paper, a class of transformations which lead to a stationary VAR representation is formulated. These transformations are motivated by an extension of differencing.

Some of the transformations can be constructed based on the cointegrating vectors. It suggests a method to estimate the matrix which forms such a transformation. It is pointed out that this method is dependent on the identifiability and estimation methods of the cointegrating vectors. One of the methods to estimate the cointegrating vectors is outlined in this paper.

As another method recommeded, the estimation of the first order serial correlation matrix of the original series is proposed. The estimator $\hat{R}$ is usually derived by fitting a first-order VAR and applying the OLS for the original data. For cointegrated series, it is shown that $\hat{R}$ converges in probability to a matrix to transform the original series

into a stationary and invertible series at the rate $T^{1/2}$. It is already established in other papers that $\hat{R}$ converges in probability to I_N at the rate T when this time series is not cointegrated. These results implies that the transformation based on $\hat{R}$ leads to a stationary VAR representation in large samples regardless of the existence of cointegration. In this respect, the transformation basedon $\hat{R}$ is strongly jutified.

Appendix

PROOF OF THEOREM 1: To prove that (i) is necessary and sufficient for the stationarity of y_t-Ry_{t-1}, consider

$$(I_N-RB)y_t=(1-B)y_t+B(I_N-R)y_t \tag{A.1}$$

Noting that the first term in the right side of (A.1) is stationary, it is obvious that y_t-Ry_{t-1} is stationary if and only if (i) holds.

On the other hand, in view of (2.3), y_t-Ry_{t-1} is invertible if and only if the power series expansion of $A(\lambda)(I_N-R\lambda)^{-1}$ is absolutely summable for $|\lambda|\leq 1$. Noting that $A(\lambda)=A^*(\lambda)(1-\lambda)+\lambda\Gamma A'=A^*(\lambda)(I_N-R\lambda)+\lambda\{\Gamma A'-A^*(\lambda)(I_N-R)\}$, from (i)

$$A(\lambda)(I_N-R\lambda)^{-1}=A^*(\lambda)+\lambda(\Gamma A'-A^*(\lambda)FA')(I_N-R\lambda)^{-1} \tag{A.2}$$

Since $A^*(\lambda)$ is absolutely summable for $|\lambda|\leq 1$, the invertibility of y_t-Ry_{t-1} is equivalent to the absolute summability of the power series expansions of $\Gamma A'(I_N-R\lambda)^{-1}$ and $FA'(I_N-R\lambda)^{-1}$ on $|\lambda|\leq 1$. Noting that rank $\Gamma=r$, their absolute summability is achieved if and only if

$$\sum_{j=1}^{\infty}|A'R^j|<\infty \tag{A.3}$$

Now, we shall prove that (ii) is necessary and sufficient for (A.3).

Since the sufficiency can be directly checked, it is sufficient to show the necessity. For this, suppose that (ii) does not hold. Then, defining $R_2=R-R_1$ for any $N\times N$ matrix R_1 such that $R_1=P(A)D$, we have

$$AR_2\neq 0,\quad R_2\neq 0 \tag{A.4}$$

Also, for at least one λ such that $|\lambda|\leq 1$,

$$\begin{array}{lll} \text{either} & \text{(a)} & \det(I_N-R_2\lambda)=0 \\ \text{or} & \text{(b)} & R_2R_1\neq 0 \end{array} \tag{A.5}$$

must hold. Since $A'(I_N-R_1\lambda)^{-1}=A'(I_N-R_1\lambda)=A'$, we derive

$$\begin{aligned} A'(I_N-R\lambda)^{-1} &= A'\{(I_N-R\lambda)(I_N-R_1\lambda)^{-1}\}^{-1} \\ &= A'\{I_N-R_2(I_N-R_1\lambda)^{-1}\lambda\}^{-1} \\ &= A'\left\{I_N+\sum_{j=1}^{\infty}R_2^j\left(I_N+\sum_{i=1}^{\infty}R_1^i\lambda^i\right)^j\lambda^j\right\}. \end{aligned} \tag{A.6}$$

It is obvious that $A'\Sigma\sum_{j=1}^{\infty}R_2^j$ does not converge when (a) holds. On the other hand, under (b), it can be also shown that $R_2^n(I_N+\Sigma_{i=1}^{\infty}R_1^i)^n$ does not converge to zero matrix as $n\to\infty$. In view of (A.6), these results imply that (A.3) does not hold. *Q. E. D.*

PROOF OF LEMMA: First, note that $\det(I_N-\bar{R}_2\lambda)=\det(I_N-S\bar{R}_2S'\lambda)=\det\{I_r-R_1(1)R_1(0)^{-1}\lambda\}$. If $\det\{I_r-R_1(1)R_1(0)^{-1}\lambda\}=0$ for one λ such that $|\lambda|\leq 1$, there must exist $\beta\neq 0\in R^r$ satisfying

$$B'\{R_1(0)-R_1(1)\lambda\}=0', \tag{A.7}$$

which implies that $\rho(1)\lambda=1$, where $\rho(1)=\beta' R_1(1)\beta/\beta' R_1(0)\beta$. This contradicts the stationarity of $\beta'\psi_t$. *Q. E. D.*

PROOF OF THEOREM 2: Put

$$Y_i=(y_{i1},\ \dots,\ y_{i,T-1})',\quad Y_i^{(+)}=(y_{i2},\ \dots,\ y_{iT})',\quad i=1,\ 2,$$

$$\Psi=(\psi_1,\ \dots,\ \psi_{T1})',\ \Psi^{(+)}=(\psi_2,\ \dots,\ \psi_T)' \text{ and } \Phi=(\varphi_2,\ \dots,\ \varphi_T)'.$$

Using these notations, $\hat{R}$ can be expressed as

$$\text{(A.8)}\qquad \hat{R}=S'\left(\sum_{t=2}^{T} Sy_t y_{t-1}' S'\right)\left(\sum_{t=2}^{T} Sy_{t-1}y_{t-1}'S'\right)^{-1}S$$

$$=S'\left[\begin{array}{c|c} Y_1^{(+)\prime}P_2Y_1Q^{-1} & \hat{M}-Y_1^{(+)\prime}P_2Y_1Q^{-1}\hat{M} \\ \hline Y_2^{(+)\prime}P_2Y_1Q^{-1} & \hat{W}-Y_2^{(+)\prime}P_2Y_1Q^{-1}\hat{M} \end{array}\right]S,$$

where $P_2=I_{T-1}-Y_2(Y_2'Y_2)^{-1}Y_2'$, $Q=Y_1'P_2Y_1$, $\hat{M}=Y_1^{(+)\prime}Y_2(Y_2'Y_2)^{-1}$, and $\hat{W}=Y_2^{(+)\prime}Y_2(Y_2'Y_2)^{-1}$.

By the standard asymptotic theory for stationary series satisfying the appropriate assumptions (see, for example, Hannan (1970, p. 228), it is shown that

$$\text{(A.9)}\qquad \Psi'\Psi_{/T}-R_1(0)=O_p(T^{-1/2}),\ \Psi^{(+)\prime}\Psi_{/T}-R_1(1)=O_p(T^{-1/2}),$$
$$\text{and } \Phi'\Psi_{/T}-R_2(1)=O_p(T^{-1/2}).$$

On the other hand, the asymptotic results for the nonstationary multiple time series y_{2t}, which is not cointegrated, are established in Phillips and Durlauf (1986), Stock (1987), and Stock and Watson (1988):

$$\text{(A.10)}\qquad Y_2'Y_2=O_p(T^2),\ \Psi^{(+)\prime}Y_2=O_p(T),\ \Psi'Y_2=O_p(T),$$

$$\Phi' Y_2 = O_P(T).$$

Therefore, noting that $Y_1^{(+)} = Y_2^{(+)}M' + \Psi^{(+)}$, $Y_1 = Y_2M' + \Psi$, and $Y_2^{(+)} = Y_2 + \Phi$, which follow directly from the definitions of ψ_t, φ_t, and M, from (A.9) and (A.10) , we have

$$\text{(A.11)} \qquad Y_1^{(+)\prime} P_2 Y_{1\ T} - \{R_1(1) + MR_2(1)\} = O_p(T^{-1/2}),$$

$$Y_2^{(+)\prime} P_2 Y_{1/T} - R_2(1) = O_p(T^{-1/2}),\ Q_{/T} - R_1(0) = O_p(T^{-1/2}),$$

$$\hat{M} - M = O_p(T^{-1}), \text{ and } \hat{W} - I_k = O_p(T^{-1}).$$

Thus the desired result is established. *Q. E. D.*

References

Campbell, J. Y., and R. J. Shiller (1988): "Interpreting Cointegrated Models," Journal of Economic Dynamics and Control, 12, 505-522.

Engle, R. F., and C. W. J. Granger (1987): "Co-Integration and Error-Correction: Representation, Estimation and Testing," Econometrica, 55, 251-276.

Engle, R. F., and B. S. Yoo (1987): "Forecasting and Testing in Co-integrated Systems," Journal of Econometrics, 35, 143-159.

Hannan, E. J. (1970): *Multipul Time Series.* New York: Wiley.

Phillips, P. C. B., and S. N. Durlauf (1986): "Multiple Time Series Regression with Integrated Processes," Review of Economic Studies, 53, 473-496.

Stock, J. H. (1987): "Asymptotic Properties of Least Squares Estimators of Cointegrating Vectors," Econometrica, 55, 1035-1056.

Stock, J. H., and M. W. Watson (1988): "Testing for Common Trends," Journal of the American Statistical Association, 83, 1097-1107.

Notes

1) The covariance matrix of the error process is invariant under such a transformation. In other words, the series derived by such a transformation has the same covariance matrix of the error process as that of the original series, i. e., Ω.

2) We can give some examples of G_2. For example, let $L=\bar{\lambda}A(A'A)^{-1}$ for a real number $\bar{\lambda}$ such that $0<\bar{\lambda}<1$. Then, that $\det(I_N-LA'\lambda)=(1-\bar{\lambda}\lambda)^r$ can be easily checked.

3) Note that $S=I_N$ and $y_t=(y_{1t}', y_{2t}')'$ when $\{i_1, ..., i_k\}=\{1, ..., k\}$. From a practical viewpoint, our analysis may be restricted to such a case.

4) In connection with this transformation, Campbell and Shiller (1988) showed that $(\psi_t', \varphi_t')'$ possesses a stationary VAR representation. It can be related with this transformation as

$$\begin{pmatrix}\psi_t \\ \varphi_t\end{pmatrix}=H\begin{pmatrix}\psi_t-M\varphi_t \\ \varphi_t\end{pmatrix}, \text{ with } H=\left[\begin{array}{c|c} I_r & -\mathrm{M} \\ \hline 0 & \mathrm{I_k}\end{array}\right]$$

However, the transformed series $(\psi'_t, \varphi'_t)'$ is not derived by any R of T, and the covarinace matrix of the error process is not Ω but $H\Omega H'$.

5) This result is also used in Theorem 2 below.