A Festschrift for Herman Rubin
Institute of Mathematical Statistics
Lecture Notes – Monograph Series
Vol. 45 (2004) 378–384

A simple proof of a condition for cointegration

T. W. Anderson[1,*]

Stanford University

Abstract: A simple proof is given for a theorem concerning the first difference and some linear functions of a cointegrated autoregressive process being stationary.

1. Introduction

Many macroeconometric models are formulated in terms of autoregressive processes or autoregressive processes with moving average innovations. The most appropriate process in a given situation may not be stationary, but some linear relations of the components may be stationary; such a process is called cointegrated. Johansen (1995) has given alternative conditions for the cointegrated components and first differences of other components to be stationary. Here we give a proof of one condition that is more straightforward and transparent than what is in the literature.

A p-dimensional m-order autoregressive process $\{\mathbf{Y}_t\}$ is defined by

$$\mathbf{Y}_t = \mathbf{B}_1\mathbf{Y}_{t-1} + \mathbf{B}_2\mathbf{Y}_{t-2} + \ldots + \mathbf{B}_m\mathbf{Y}_{t-m} + \mathbf{Z}_t, \tag{1.1}$$

where the $\mathbf{Z}_t$'s are independent unobservable innovations with $\mathcal{E}\mathbf{Z}_t = \mathbf{0}$, $\mathcal{E}\mathbf{Z}_t\mathbf{Z}_t' = \mathbf{\Sigma}$, and $\mathcal{E}\mathbf{Z}_t\mathbf{Y}_{t-s}' = \mathbf{0}$, $0 < s$. Let

$$\mathbf{B}(\lambda) = \lambda^m\mathbf{I}_p - \lambda^{m-1}\mathbf{B}_1 - \ldots - \mathbf{B}_m, \tag{1.2}$$

and let the roots of $|\mathbf{B}(\lambda)| = 0$ be λ_i, $i = 1, \ldots, mp$. If $|\lambda_i| < 1$, $i = 1, \ldots, mp$, the process $\{\mathbf{Y}_t\}$ may be stationary. If one or more of the roots are 1, the process is nonstationary, but some order of differencing may yield a stationary process. When some linear functions of a nonstationary process are stationary, the model is called *cointegrated*. We call a process defined by the Equation (1.1) *stationary* if it is possible to assign a distribution to $(\mathbf{Y}_{-m+1}, \ldots, \mathbf{Y}_{-1}, \mathbf{Y}_0)$ such that (1.1) generates a process $\mathbf{Y}_{-m+1}, \mathbf{Y}_{-m+2}, \ldots$ that is stationary. Throughout this paper it is assumed that n of the roots are 1 and the other roots satisfy $|\lambda_i| < 1$, $i = n+1, \ldots, mp$.

An "error-correction form" of the autoregressive process is

$$\Delta\mathbf{Y}_t = \mathbf{\Pi}\mathbf{Y}_{t-1} + \mathbf{\Pi}_1\Delta\mathbf{Y}_{t-1} + \ldots + \mathbf{\Pi}_{m-1}\Delta\mathbf{Y}_{t-m+1} + \mathbf{Z}_t, \tag{1.3}$$

where $\Delta\mathbf{Y}_t = \mathbf{Y}_t - \mathbf{Y}_{t-1}$,

$$\mathbf{\Pi}_j = -(\mathbf{B}_{j+1} + \ldots + \mathbf{B}_m), \qquad j = 1, \ldots, m-1, \tag{1.4}$$

$$\mathbf{\Pi} = \mathbf{B}_1 + \mathbf{B}_2 + \ldots + \mathbf{B}_m - \mathbf{I}_p. \tag{1.5}$$

Note that $\mathbf{\Pi}_j = \mathbf{\Pi}_{j+1} - \mathbf{B}_{j+1}$ and $\mathbf{\Pi} = -\mathbf{B}(1)$.

---

[*]This paper is dedicated to my friend and co-author Herman Rubin, who stimulated and educated me as well as collaborated with me.

[1]Departments of Economics and Statistics, Stanford University, Stanford, CA 94305, USA. e-mail: `twa@stat.stanford.edu`

Keywords and phrases: autoregressive process, error correction form, stationarity.

AMS 2000 subject classifications: primary 62P20; secondary 60G12.

Suppose the rank of $\mathbf{\Pi}$ is k. Then $\mathbf{\Pi}$ can be written $\mathbf{\Pi} = \mathbf{A}\mathbf{\Gamma}'$, where $\mathbf{A}$ and $\mathbf{\Gamma}$ are $p \times k$ matrices of rank k. Let $\mathbf{A}_\perp$ and $\mathbf{\Gamma}_\perp$ be $p \times (p-k)$ matrices of rank $p-k$ such that $\mathbf{A}_\perp'\mathbf{A} = \mathbf{0}$ and $\mathbf{\Gamma}_\perp'\mathbf{\Gamma} = \mathbf{0}$. Then a necessary and sufficient condition that $\Delta\mathbf{Y}_t$ and $\mathbf{\Gamma}'\mathbf{Y}_t$ are stationary is that

$$\operatorname{rank}\left[\mathbf{A}_\perp'\left(\mathbf{I} - \sum_{i=1}^{m-1}\mathbf{\Pi}_i\right)\mathbf{\Gamma}_\perp\right] = p-k \tag{1.6}$$

[Theorem 4.2, Johansen (1995)]. The proof of this statement involves an expansion of $\mathbf{B}(\lambda)$ around $\lambda = 1$.

If $\{\mathbf{Y}_t\}$ is stationary, it is said to be $I(0)$. If $\{\mathbf{Y}_t\}$ is not $I(0)$, but $\{\Delta\mathbf{Y}_t\}$ is stationary, the process $\{\mathbf{Y}_t\}$ is said to be $I(1)$.

Corollary 4.3 of Johansen asserts that if k is the rank of $\mathbf{\Pi}$ and $k < p$, then the multiplicity of $\lambda = 1$ as a zero of $|\mathbf{B}(\lambda)|$ is equal to $p-k$ if and only if $\{\mathbf{Y}_t\}$ is $I(1)$. The proof of this statement depends on his Theorem 4.2 and its proof.

In this paper the condition is formulated as

Rank Condition. *There are n linearly independent solutions to*

$$\boldsymbol{\omega}'\mathbf{\Pi} = \mathbf{0}, \tag{1.7}$$

where n is the multiplicity of $\lambda = 1$ as a root of the characteristic equation $|\mathbf{B}(\lambda)| = 0$.

Let n independent solutions of (1.7) be assembled into the matrix $\mathbf{\Omega}_1 = (\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_n)$; then $\mathbf{\Omega}_1'\mathbf{\Pi} = \mathbf{0}$ and the rank of $\mathbf{\Omega}_1$ is n.

2. First-order case

First we treat the special case of $m = 1$. Then (1.1) is

$$\mathbf{Y}_t = \mathbf{B}_1\mathbf{Y}_{t-1} + \mathbf{Z}_t; \tag{2.1}$$

the error-correction form is

$$\Delta\mathbf{Y}_t = \mathbf{\Pi}\mathbf{Y}_{t-1} + \mathbf{Z}_t, \tag{2.2}$$

where $\mathbf{\Pi} = \mathbf{B}_1 - \mathbf{I}_p$; and $\mathbf{B}(\lambda) = \lambda\mathbf{I}_p - \mathbf{B}_1$.

Theorem 1 ($m = 1$). *Suppose the Rank Condition holds. Then the rank of $\mathbf{\Pi}$ is $k = p - n$, and there exists a $p \times k$ matrix $\mathbf{\Omega}_2$ such that*

$$\mathbf{\Omega}_2'\mathbf{\Pi} = \mathbf{\Upsilon}_{22}\mathbf{\Omega}_2', \tag{2.3}$$

$\mathbf{\Upsilon}_{22}$ $(k \times k)$ is nonsingular, and $\mathbf{\Omega} = (\mathbf{\Omega}_1, \mathbf{\Omega}_2)$ is nonsingular. Define

$$\mathbf{X}_t = \begin{bmatrix}\mathbf{X}_{1t}\\ \mathbf{X}_{2t}\end{bmatrix} = \begin{bmatrix}\mathbf{\Omega}_1'\mathbf{Y}_t\\ \mathbf{\Omega}_2'\mathbf{Y}_t\end{bmatrix}, \qquad \mathbf{W}_t = \begin{bmatrix}\mathbf{W}_{1t}\\ \mathbf{W}_{2t}\end{bmatrix} = \begin{bmatrix}\mathbf{\Omega}_1'\mathbf{Z}_t\\ \mathbf{\Omega}_2'\mathbf{Z}_t\end{bmatrix}. \tag{2.4}$$

Then $\Delta\mathbf{X}_{1t}, \mathbf{X}_{2t}$ defines a stationary process.

Proof. Let $\mathbf{\Omega}_1' = (\mathbf{I}_n, \mathbf{\Omega}_{21}')$ and $\mathbf{\Pi}' = (\mathbf{\Pi}_1', \mathbf{\Pi}_2')$, where $\mathbf{\Pi}_2$ is $k \times p$. (The rows of $\mathbf{\Omega}_1$ and the columns of $\mathbf{\Pi}$ can be ordered so that $\mathbf{\Omega}_{11}$ is nonsingular and can be set as $\mathbf{I}_n$.) Then the Rank Condition is

$$\mathbf{0} = \mathbf{\Omega}_1'\mathbf{\Pi} = \left(\mathbf{I}_n, \mathbf{\Omega}_{21}'\right)\begin{bmatrix}\mathbf{\Pi}_1\\ \mathbf{\Pi}_2\end{bmatrix} = \mathbf{\Pi}_1 + \mathbf{\Omega}_{21}'\mathbf{\Pi}_2, \tag{2.5}$$

which implies $\mathbf{\Pi}_1 = -\mathbf{\Omega}_{21}'\mathbf{\Pi}_2$ and

$$\mathbf{\Pi} = \begin{bmatrix} -\mathbf{\Omega}_{21}' \\ \mathbf{I}_k \end{bmatrix} \mathbf{\Pi}_2. \tag{2.6}$$

Define $\mathbf{\Omega}_2 = \mathbf{\Pi}_2'$ $(p \times k)$ and

$$\mathbf{\Upsilon}_{22} = \mathbf{\Pi}_2 \begin{bmatrix} -\mathbf{\Omega}_{21}' \\ \mathbf{I}_k \end{bmatrix} = \mathbf{\Omega}_2' \begin{bmatrix} -\mathbf{\Omega}_{21}' \\ \mathbf{I}_k \end{bmatrix}. \tag{2.7}$$

Then (2.3) is satisfied. Note that $\mathbf{\Upsilon}_{22}$ $(k \times k)$ is nonsingular, that is, of rank k, because if $\mathbf{\Upsilon}_{22}$ were singular there would exist a k-vector $\boldsymbol{\gamma}$ such that $\boldsymbol{\gamma}'\mathbf{\Upsilon}_{22} = \mathbf{0}$ and then $\boldsymbol{\gamma}'\mathbf{\Pi}_2$ would be another left-sided eigenvector of $\mathbf{\Pi}$ associated with the root 0, but that would imply more than n linearly independent vectors satisfying $\boldsymbol{\omega}'\mathbf{\Pi} = \mathbf{0}$ and hence more than n zeros of $|\mathbf{B}(\lambda)|$ at $\lambda = 1$, which is contrary to assumption. Note that (2.6) is a factorization $\mathbf{\Pi} = \mathbf{A}\mathbf{\Gamma}'$ with $\mathbf{\Gamma}' = \mathbf{\Pi}_2$.

The matrix $\mathbf{\Omega}$ satisfies

$$\mathbf{\Omega}'\mathbf{\Pi} = \begin{bmatrix} \mathbf{\Omega}_1' \\ \mathbf{\Omega}_2' \end{bmatrix} \begin{bmatrix} -\mathbf{\Omega}_{21}' \\ \mathbf{I}_k \end{bmatrix} \mathbf{\Pi}_2 = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Upsilon}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{\Omega}_1' \\ \mathbf{\Omega}_2' \end{bmatrix} = \mathbf{\Upsilon}\mathbf{\Omega}', \tag{2.8}$$

$$\mathbf{\Omega}'\mathbf{B} = \mathbf{\Omega}'(\mathbf{\Pi} + \mathbf{I}) = \begin{bmatrix} \mathbf{I}_n & \mathbf{0} \\ \mathbf{0} & \mathbf{\Psi}_{22} \end{bmatrix} \mathbf{\Omega}' = \mathbf{\Psi}\mathbf{\Omega}'. \tag{2.9}$$

where $\mathbf{\Psi}_{22} = \mathbf{\Upsilon}_{22} + \mathbf{I}_k$. Let $\mathbf{\Pi}_2 = (\mathbf{\Pi}_{21}, \mathbf{\Pi}_{22})$. Then $\mathbf{\Omega}$ is nonsingular because

$$|\mathbf{\Omega}| = \begin{vmatrix} \mathbf{I}_n & \mathbf{0} \\ -\mathbf{\Omega}_{21} & \mathbf{I}_k \end{vmatrix} \begin{vmatrix} \mathbf{I}_n & \mathbf{\Pi}_{21}' \\ \mathbf{\Omega}_{21} & \mathbf{\Pi}_{22}' \end{vmatrix} = \begin{vmatrix} \mathbf{I}_n & \mathbf{\Pi}_{21}' \\ \mathbf{0} & \mathbf{\Upsilon}_{22}' \end{vmatrix} = |\mathbf{\Upsilon}_{22}'| \neq 0. \tag{2.10}$$

Hence (2.4) is a nonsingular linear transformation.

The transformed process $\mathbf{X}_t$ satisfies the autoregressive model

$$\mathbf{X}_t = \mathbf{\Psi}\mathbf{X}_{t-1} + \mathbf{W}_t, \tag{2.11}$$

$$\Delta\mathbf{X}_t = \mathbf{\Upsilon}\mathbf{X}_{t-1} + \mathbf{W}_t, \tag{2.12}$$

where

$$\mathbf{\Psi} = \mathbf{\Omega}'\mathbf{B}_1(\mathbf{\Omega}')^{-1} = \begin{bmatrix} \mathbf{I}_n & \mathbf{0} \\ \mathbf{0} & \mathbf{\Psi}_{22} \end{bmatrix} \tag{2.13}$$

has eigenvalues λ_i, $i = 1, \ldots, p$, and $\mathbf{\Psi}_{22}$ has eigenvalues λ_i, $i = n+1, \ldots, p$, and $\mathbf{\Upsilon} = \mathbf{\Psi} - \mathbf{I}_p$. From (2.11) to (2.13) we obtain

$$\begin{aligned} \begin{bmatrix} \Delta\mathbf{X}_{1t} \\ \mathbf{X}_{2t} \end{bmatrix} &= \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Psi}_{22} \end{bmatrix} \begin{bmatrix} \Delta\mathbf{X}_{1,t-1} \\ \mathbf{X}_{2,t-1} \end{bmatrix} + \begin{bmatrix} \mathbf{W}_{1t} \\ \mathbf{W}_{2t} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{W}_{1t} \\ \mathbf{\Psi}_{22}\mathbf{X}_{2,t-1} + \mathbf{W}_{2t} \end{bmatrix} \end{aligned} \tag{2.14}$$

as generating the process $(\Delta\mathbf{X}_{1t}', \mathbf{X}_{2t}')'$. Since the eigenvalues of the coefficient matrix in (2.14) are 0 of multiplicity n and λ_i, $i = n+1, \ldots, p$, the process $(\Delta\mathbf{X}_{1t}', \mathbf{X}_{2t}')'$ is a stationary process. □

The transformation $\mathbf{X}_t = \mathbf{\Omega}'\mathbf{Y}_t$ is a change of coordinates such that the first n coordinates of $\mathbf{X}_t$ define a random walk, which is an $I(1)$ process. The other k coordinates define a stationary process. Thus $\{\mathbf{X}_t\}$ is an $I(1)$ process; that is, $\Delta\mathbf{X}_t$ is an $I(0)$ process. The process $\mathbf{Y}_t = (\mathbf{\Omega}')^{-1}\mathbf{X}_t$ is a mixture of an $I(1)$ and an $I(0)$ process.

3. General case

Theorem 2. *When the Rank Condition holds,*

$$\begin{bmatrix} \Delta \mathbf{Y}_t \\ \mathbf{\Pi}_2 \mathbf{Y}_t \end{bmatrix} \tag{3.1}$$

defines a stationary process.

Proof. For arbitrary m the models (1.1) and (1.3) can be written in "stacked" form as

$$\widetilde{\mathbf{Y}}_t = \widetilde{\mathbf{B}}_1 \widetilde{\mathbf{Y}}_{t-1} + \widetilde{\mathbf{Z}}_t \tag{3.2}$$

and

$$\Delta \widetilde{\mathbf{Y}}_t = \widetilde{\mathbf{\Pi}} \widetilde{\mathbf{Y}}_{t-1} + \widetilde{\mathbf{Z}}_t, \tag{3.3}$$

where

$$\widetilde{\mathbf{Y}}_t = \begin{bmatrix} \mathbf{Y}_t \\ \mathbf{Y}_{t-1} \\ \mathbf{Y}_{t-2} \\ \vdots \\ \mathbf{Y}_{t-m+1} \end{bmatrix}, \widetilde{\mathbf{Z}}_t = \begin{bmatrix} \mathbf{Z}_t \\ \mathbf{0} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix}, \widetilde{\mathbf{B}}_1 = \begin{bmatrix} \mathbf{B}_1 & \mathbf{B}_2 & \dots & \mathbf{B}_{m-1} & \mathbf{B}_m \\ \mathbf{I}_p & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_p & \dots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{I}_p & \mathbf{0} \end{bmatrix}, \tag{3.4}$$

and $\widetilde{\mathbf{\Pi}} = \widetilde{\mathbf{B}}_1 - \mathbf{I}_{mp}$. [See Anderson (1971), Section 5.3, for example.] Let $\widetilde{\mathbf{B}}(\lambda) = \lambda \mathbf{I}_{mp} - \widetilde{\mathbf{B}}_1$. Then $|\mathbf{B}(\lambda)| = |\widetilde{\mathbf{B}}(\lambda)|$. We shall prove Theorem 2 by using Theorem 1 with $\mathbf{Y}_t$ replaced by $\widetilde{\mathbf{Y}}_t$.

Suppose that there are n linearly independent solutions to $\widetilde{\boldsymbol{\omega}}' \widetilde{\mathbf{\Pi}} = \mathbf{0}$. Let these solutions be assembled into the $n \times mp$ matrix $\widetilde{\mathbf{\Omega}}_1' = (\widetilde{\mathbf{\Omega}}_{11}', \dots, \widetilde{\mathbf{\Omega}}_{m1}')$. Then

$$\begin{aligned} \mathbf{0} &= \widetilde{\mathbf{\Omega}}_1' \widetilde{\mathbf{\Pi}} \\ &= \Big[\widetilde{\mathbf{\Omega}}_{11}'(\mathbf{B}_1 - \mathbf{I}_p) + \widetilde{\mathbf{\Omega}}_{21}',\ \widetilde{\mathbf{\Omega}}_{11}' \mathbf{B}_2 - \widetilde{\mathbf{\Omega}}_{21}' + \widetilde{\mathbf{\Omega}}_{31}', \dots, \\ &\qquad \widetilde{\mathbf{\Omega}}_{11}' \mathbf{B}_{m-1} - \widetilde{\mathbf{\Omega}}_{m-1,1}' + \widetilde{\mathbf{\Omega}}_{m1}',\ \widetilde{\mathbf{\Omega}}_{11}' \mathbf{B}_m - \widetilde{\mathbf{\Omega}}_{m1}' \Big]. \end{aligned} \tag{3.5}$$

This equation implies

$$\widetilde{\mathbf{\Omega}}_{m1}' = \widetilde{\mathbf{\Omega}}_{11}' \mathbf{B}_m = -\widetilde{\mathbf{\Omega}}_{11}' \mathbf{\Pi}_{m-1}, \tag{3.6}$$

$$\widetilde{\mathbf{\Omega}}_{m-j,1}' = \widetilde{\mathbf{\Omega}}_{11}' \mathbf{B}_{m-j} + \widetilde{\mathbf{\Omega}}_{m-j+1}' = -\widetilde{\mathbf{\Omega}}_{11}' \mathbf{\Pi}_{m-j-1}, \qquad j = 1, \dots, m-1, \tag{3.7}$$

$$\mathbf{0} = \widetilde{\mathbf{\Omega}}_{11}'(\mathbf{B}_1 - \mathbf{I}_p) + \widetilde{\mathbf{\Omega}}_{21}' = \widetilde{\mathbf{\Omega}}_{11}' \mathbf{\Pi}. \tag{3.8}$$

It follows that

$$\widetilde{\mathbf{\Omega}}_1' = \widetilde{\mathbf{\Omega}}_{11}' [\mathbf{I}_p, -\mathbf{\Pi}_1, \dots, -\mathbf{\Pi}_{m-1}]. \tag{3.9}$$

Lemma. *There is a $pm \times n$ matrix $\widetilde{\mathbf{\Omega}}_1$ of rank n such that $\widetilde{\mathbf{\Omega}}_1' \widetilde{\mathbf{\Pi}} = \mathbf{0}$ if and only if there is a $p \times n$ matrix $\widetilde{\mathbf{\Omega}}_{11}$ of rank n such that $\widetilde{\mathbf{\Omega}}_{11}' \mathbf{\Pi} = \mathbf{0}$.*

Thus the Rank Condition on the mp-dimensional $\widetilde{\mathbf{Y}}_t$ in terms of $\widetilde{\mathbf{\Pi}}$ is equivalent to the Rank Condition on $\mathbf{Y}_t$, where $\mathbf{\Pi}$ is defined by (1.5).

It follows from Theorem 1 applied to (3.2) that the rank of $\widetilde{\mathbf{\Pi}}$ is $\tilde{k} = mp - n$. Let

$$\widetilde{\mathbf{\Pi}} = \begin{bmatrix} \widetilde{\mathbf{\Pi}}_{\cdot n} \\ \widetilde{\mathbf{\Pi}}_{\cdot \tilde{k}} \end{bmatrix} = \begin{bmatrix} (\mathbf{B}_1 - \mathbf{I}_p)_{\cdot n} & \mathbf{B}_{2\cdot n} & \dots & \mathbf{B}_{m-1\cdot n} & \mathbf{B}_{m\cdot n} \\ (\mathbf{B}_1 - \mathbf{I}_p)_{\cdot k} & \mathbf{B}_{2\cdot k} & \dots & \mathbf{B}_{m-1\cdot k} & \mathbf{B}_{m\cdot k} \\ \mathbf{I}_p & -\mathbf{I}_p & & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & & \mathbf{I}_p & -\mathbf{I}_p \end{bmatrix}, \tag{3.10}$$

where $\widetilde{\mathbf{\Pi}}_{\cdot n}$ has n rows and $(\ )_{\cdot n}$ denotes the first n rows of $(\ )$ and $(\ )_{\cdot k}$ denotes the last k rows of $(\ )$. The $pm \times \tilde{k}$ matrix $\widetilde{\mathbf{\Omega}}_2 = \widetilde{\mathbf{\Pi}}'_{\cdot \tilde{k}}$ satisfies

$$\widetilde{\mathbf{\Omega}}_2' \widetilde{\mathbf{\Pi}} = \widetilde{\mathbf{\Upsilon}}_{22} \widetilde{\mathbf{\Omega}}_2', \tag{3.11}$$

$\widetilde{\mathbf{\Upsilon}}_{22}$ is nonsingular, and $\widetilde{\mathbf{\Omega}} = (\widetilde{\mathbf{\Omega}}_1, \widetilde{\mathbf{\Omega}}_2)$ is nonsingular. Define $\widetilde{\mathbf{X}}_t = \widetilde{\mathbf{\Omega}}' \widetilde{\mathbf{Y}}_t$ and $\widetilde{\mathbf{W}}_t = \widetilde{\mathbf{\Omega}}' \widetilde{\mathbf{Z}}_t$. Then $\widetilde{\mathbf{X}}_t = (\widetilde{\mathbf{X}}_{1t}', \widetilde{\mathbf{X}}_{2t}')'$ satisfies

$$\widetilde{\mathbf{X}}_{1t} = \widetilde{\mathbf{X}}_{1,t-1} + \widetilde{\mathbf{W}}_{1t}, \tag{3.12}$$

$$\widetilde{\mathbf{X}}_{2t} = \widetilde{\mathbf{\Psi}}_{22} \widetilde{\mathbf{X}}_{2,t-1} + \widetilde{\mathbf{W}}_{2t}, \tag{3.13}$$

where the eigenvalues of $\widetilde{\mathbf{\Psi}}_{22}$ are λ_i, $i = n+1, \dots, mp$,

$$\widetilde{\mathbf{X}}_{2t} = \widetilde{\mathbf{\Omega}}_2' \widetilde{\mathbf{Y}}_t = \begin{bmatrix} (\mathbf{B}_1 - \mathbf{I}_p)_{\cdot k} \mathbf{Y}_t + \mathbf{B}_{2\cdot k} \mathbf{Y}_{t-1} + \dots + \mathbf{B}_{m\cdot k} \mathbf{Y}_{t-m+1} \\ \mathbf{Y}_t - \mathbf{Y}_{t-1} \\ \vdots \\ \mathbf{Y}_{t-m+2} - \mathbf{Y}_{t-m+1} \end{bmatrix}, \tag{3.14}$$

and

$$\widetilde{\mathbf{W}}_{2t} = \widetilde{\mathbf{\Omega}}_2' \widetilde{\mathbf{Z}}_t = \begin{bmatrix} (\mathbf{B}_1 - \mathbf{I}_p)_{\cdot k} \mathbf{W}_t \\ \mathbf{W}_t \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix}. \tag{3.15}$$

Thus $\{\widetilde{\mathbf{X}}_{1t}\}$ is an $I(1)$ process of dimension n and $\{\widetilde{\mathbf{X}}_{2t}\}$ is an $I(0)$ process of dimension $\widetilde{k}$.

Now we want to transform $\{\widetilde{\mathbf{X}}_t\}$ so that $k = p - n$ coordinates constitute the cointegrated part of $\{\mathbf{Y}_t\}$ and the other coordinates are components of $\Delta \mathbf{Y}_t, \dots, \Delta \mathbf{Y}_{t-m+1}$. In terms of $\mathbf{Y}_t$ (3.12) can be written

$$\sum_{j=1}^{m} \widetilde{\mathbf{\Omega}}_{j1}' \Delta \mathbf{Y}_{t-j+1} = \widetilde{\mathbf{\Omega}}_{11}' \left(\Delta \mathbf{Y}_t - \sum_{j=2}^{m} \mathbf{\Pi}_{j-1} \Delta \mathbf{Y}_{t-j+1} \right) = \widetilde{\mathbf{\Omega}}_{11}' \mathbf{Z}_t = \widetilde{\mathbf{W}}_{1t}. \tag{3.16}$$

Let

$$\widetilde{\mathbf{M}} = \begin{bmatrix} \mathbf{I}_k & -\mathbf{\Pi}_{1\cdot k} & \dots & -\mathbf{\Pi}_{m-1\cdot k} \\ \mathbf{0} & \mathbf{I}_p & \dots & \mathbf{0} \\ \vdots & \vdots & & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{I}_p \end{bmatrix}, \tag{3.17}$$

$$\widetilde{\mathbf{V}}_{2t} = \widetilde{\mathbf{M}} \widetilde{\mathbf{X}}_{2t} = \begin{bmatrix} \mathbf{\Pi}_{\cdot k} \mathbf{Y}_t \\ \Delta \mathbf{Y}_t \\ \vdots \\ \Delta \mathbf{Y}_{t-m+2} \end{bmatrix}, \quad \widetilde{\mathbf{U}}_{2t} = \widetilde{\mathbf{M}} \mathbf{W}_{2t} = \begin{bmatrix} \mathbf{\Pi}_{\cdot k} \mathbf{W}_t \\ \mathbf{W}_t \\ \vdots \\ \mathbf{0} \end{bmatrix}. \tag{3.18}$$

Here $\mathbf{\Pi}_{\cdot k}$ denotes the last k rows of $\mathbf{\Pi}$ defined by (1.5); that is, $\mathbf{\Pi}_{\cdot k} = \mathbf{\Pi}_2$ in (2.6). Let $\widetilde{\mathbf{\Theta}} = \widetilde{\mathbf{M}}\widetilde{\mathbf{\Psi}}_{22}\widetilde{\mathbf{M}}^{-1}$. Then $\widetilde{\mathbf{V}}_{2t}$ satisfies

$$\widetilde{\mathbf{V}}_{2t} = \widetilde{\mathbf{\Theta}}\widetilde{\mathbf{V}}_{2,t-1} + \widetilde{\mathbf{U}}_{2t}. \tag{3.19}$$

The eigenvalues of $\widetilde{\mathbf{\Theta}}$ are λ_i, $i = n+1, \ldots, mp$. Hence $\widetilde{\mathbf{V}}_{2t}$ defines a stationary process. In fact

$$\widetilde{\mathbf{V}}_{2t} = \sum_{s=0}^{\infty} \widetilde{\mathbf{\Theta}}^s \widetilde{\mathbf{U}}_{2,t-s}. \tag{3.20}$$

Since the last $m-2$ blocks of $\widetilde{\mathbf{U}}_{2t}$ are $\mathbf{0}$'s, the last $m-2$ blocks of (3.19) are identities. The first $k+p$ rows of (3.19) define a stationary process for $\mathbf{\Pi}_{\cdot k}\mathbf{Y}_t$ and $\Delta\mathbf{Y}_t$. □ □

Discussion. *The process* $\{\mathbf{Y}_t\}$ *is* cointegrated *of rank k, and* $\mathbf{\Pi}_{\cdot k}$ *is the* cointegrating matrix.

The orthogonality conditions of $\mathbf{A}_\perp$ and $\mathbf{\Gamma}_\perp$ are equivalent to $\mathbf{A}_\perp\mathbf{\Pi} = \mathbf{0}$ and $\mathbf{\Pi}\mathbf{\Gamma}_\perp = \mathbf{0}$. Hence, $\mathbf{A}_\perp$ consists of $p-k$ left-sided characteristic vectors of $\mathbf{\Pi}$ corresponding to the characteristic root of 0 and $\mathbf{\Gamma}_\perp$ consists of $p-k$ right-sided characteristic vectors corresponding to the root of 0. The matrix $\mathbf{\Gamma}$ corresponds to $\mathbf{\Omega}_2 = \mathbf{\Pi}_2'$.

4. Inference

The model (1.3) has the form of regression

$$\mathbf{Y}_t = \mathbf{A}_1\mathbf{X}_{1t} + \mathbf{A}_2\mathbf{X}_{2t} + \mathbf{Z}_t, \tag{4.1}$$

where $\mathbf{A}_1$ is of rank k. The maximum likelihood estimator of $\mathbf{A}_1$ under normality of $\mathbf{Z}_t$ is the reduced rank regression estimator introduced by Anderson (1951). Johansen (1988), (1995) also derived the estimator for (1.3) and gives some asymptotic theory suitable for the cointegrated model. Anderson (2000), (2001), (2002) has given more details of the asymptotic theory.

Acknowledgement

The author thanks Cheng Hsiao for helpful suggestions.

References

[1] Anderson, T. W. (1951). Estimating linear restrictions on regression coefficients for multivariate normal distributions, *Annals of Mathematical Statistics*, **22**, 327–351. [Correction, *Annals of Statistics*, **8**, (1980), p. 1400.] MR42664

[2] Anderson, T. W. (1971). *The Statistical Analysis of Time Series*, John Wiley and Sons, Inc., New York. MR283939

[3] Anderson, T. W. (2000). The asymptotic distribution of canonical correlations in cointegrated models. *Proceedings of the National Academy of Sciences*, **97**, 7068–7073. MR1769813

[4] Anderson, T. W. (2001). The asymptotic distribution of canonical correlations and variates in higher-order cointegrated models, *Proceedings of the National Academy of Sciences*, **98**, 4860–4865. MR1828055

[5] Anderson, T. W. (2002). Reduced rank regression in cointegrated models, *Journal of Econometrics*, **106**, 203–216. MR1884248

[6] Johansen, Soren (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, **12**, 231–254. MR986516

[7] Johansen, Soren (1995). *Likelihood-based Inference in Cointegrated Vector Autoregressive Models*, Oxford University Press, Oxford. MR1487375