A Condition for Cointegration

T. W. Anderson

Department of Statistics, Stanford University, Stanford, CA 94305-4065

March 19, 2001

Abstract

A simple condition is given for the first difference and some linear functions of a cointegrated autoregressive process to be stationary.

Key Words: *autoregressive process,* *error correction form,* *stationarity.*

1 Introduction

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

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

where $\mathbf{Z}_t$ is an unobservable innovation with $\mathcal{E}\mathbf{Z}_t = \mathbf{0}$ and $\mathcal{E}\mathbf{Z}_t\mathbf{Z}_t' = \mathbf{\Sigma}$, may be stationary if the roots $\lambda_i,\ i = 1, \ldots, 2p$, of $|\mathbf{B}(\lambda)| = 0$, where $\mathbf{B}(\lambda) = \lambda^2\mathbf{I} - \lambda\mathbf{B}_1 - \mathbf{B}_2$, satisfy $|\lambda_i| < 1,\ i = 1, \ldots, 2p$. If one or more of the roots are 1, the process is nonstationary, but some order of differencing will yield a stationary process. If some linear functions of a nonstationary process are stationary, the model is called *cointegrated.* The purpose of this paper is to give a simple condition on $\mathbf{B}_1$ and $\mathbf{B}_2$ to ensure stationarity of the differenced process $\Delta\mathbf{Y}_t = \mathbf{Y}_t - \mathbf{Y}_{t-1}$ and these linear combinations of $\mathbf{Y}_t$ and provide a direct proof. We call a process defined by the equation (1.1) *stationary* if it is possible to assign a distribution to $(\mathbf{Y}_{-1}, \mathbf{Y}_0)$ such that (1.1) generates a process $\mathbf{Y}_1, \mathbf{Y}_2, \ldots$ that is stationary. Throughout this paper it is assumed that n of the roots are 1 and the other $2p - n$ roots satisfy $|\lambda_i| < 1$.

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

$$\Delta\mathbf{Y}_t = \mathbf{\Pi}\mathbf{Y}_{t-1} + \mathbf{\Pi}_1\Delta\mathbf{Y}_{t-1} + \mathbf{Z}_t, \tag{1.2}$$

where

$$\mathbf{\Pi} = \mathbf{B}_1 + \mathbf{B}_2 - \mathbf{I} = -\mathbf{B}(1), \quad \mathbf{\Pi}_1 = -\mathbf{B}_2. \tag{1.3}$$

Johansen (1995, Theorem 4.2) has shown that $\Delta\mathbf{Y}_t$ and the cointegrated linear combinations of $\mathbf{Y}_t$ define a stationary process if a certain condition (to be given later after more quantities have been defined) is satisfied. In this paper we give a simpler condition and prove it more directly. The generalization of the theorem to an autoregressive process of arbitrary order is stated in Section 3.3.

2 The theorem.

Condition A. There are n linearly independent solutions to

$$\boldsymbol{\omega}'\mathbf{\Pi} = \mathbf{0}, \tag{2.1}$$

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

Let the solutions of (2.1) be assembled into the matrix $\mathbf{\Omega}_1 = (\omega_1, \ldots, \omega_n)$; then $\mathbf{\Omega}_1'\mathbf{\Pi} = \mathbf{0}$ and the rank of $\mathbf{\Omega}_1$ is n. Note that (2.1) is equivalent to $\boldsymbol{\omega}'\mathbf{B}(1) = \boldsymbol{\omega}$. This assumption implies that $\{\mathbf{Y}_t\}$ is $I(1)$ (as will be seen later).

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

$$\mathbf{\Omega}_2'\mathbf{\Pi} = \mathbf{\Upsilon}_{22}\mathbf{\Omega}_2', \tag{2.2}$$

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

$$\mathbf{X}_t = \begin{bmatrix} \mathbf{X}_{1t} \\ \mathbf{X}_{2t} \end{bmatrix} = \begin{bmatrix} \mathbf{\Omega}_1'\mathbf{Y}_t \\ \mathbf{\Omega}_2'\mathbf{Y}_t \end{bmatrix}, \quad \mathbf{W}_t = \begin{bmatrix} \mathbf{W}_{1t} \\ \mathbf{W}_{2t} \end{bmatrix} = \begin{bmatrix} \mathbf{\Omega}_1'\mathbf{Z}_t \\ \mathbf{\Omega}_2'\mathbf{Z}_t \end{bmatrix}. \tag{2.3}$$

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

Proof. Let $\mathbf{\Omega}_1' = (\mathbf{\Omega}_{11}', \mathbf{\Omega}_{21}')$, where $\mathbf{\Omega}_{11}$ is nonsingular, 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.) Then Condition

A is

$$\mathbf{0} = \mathbf{\Omega}_1'\mathbf{\Pi} = (\mathbf{\Omega}_{11}', \mathbf{\Omega}_{21}') \begin{bmatrix} \mathbf{\Pi}_1 \\ \mathbf{\Pi}_2 \end{bmatrix} = \mathbf{\Omega}_{11}'\mathbf{\Pi}_1 + \mathbf{\Omega}_{21}'\mathbf{\Pi}_2, \tag{2.4}$$

$\mathbf{\Pi}_1 = -(\mathbf{\Omega}_{11}')^{-1}\mathbf{\Omega}_{21}'\mathbf{\Pi}_2$, and

$$\mathbf{\Pi} = \begin{bmatrix} -(\mathbf{\Omega}_{11}')^{-1}\mathbf{\Omega}_{21}' \\ \mathbf{I} \end{bmatrix} \mathbf{\Pi}_2. \tag{2.5}$$

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

$$\mathbf{\Upsilon}_{22} = \mathbf{\Pi}_2 \begin{bmatrix} -(\mathbf{\Omega}_{11}')^{-1}\mathbf{\Omega}_{22}' \\ \mathbf{I} \end{bmatrix}. \tag{2.6}$$

Then (2.2) 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 γ such that $\gamma'\mathbf{\Upsilon}_{22} = \mathbf{0}$ and then $\gamma'\mathbf{\Pi}_2$ would be another left-sided characteristic vector of $\mathbf{\Pi}$ associated with the root 0, but that would imply more than n independent vectors satisfying $\omega'\mathbf{B}(1) = \omega'$ and hence more than n zeros of $|\mathbf{B}(\lambda)| = 0$ at $\lambda = 1$, which is contrary to assumption.

The matrix $\mathbf{\Omega} = (\mathbf{\Omega}_1, \mathbf{\Omega}_2)$ is nonsingular because

$$|\mathbf{\Omega}| = \begin{vmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{\Omega}_{21}\mathbf{\Omega}_{11}^{-1} & \mathbf{I} \end{vmatrix} \begin{vmatrix} \mathbf{\Omega}_{11} & \mathbf{\Pi}_{21}' \\ \mathbf{\Omega}_{21} & \mathbf{\Pi}_{22}' \end{vmatrix} = \begin{vmatrix} \mathbf{\Omega}_{11} & \mathbf{\Pi}_{21}' \\ \mathbf{0} & \mathbf{\Upsilon}_{22}' \end{vmatrix} = |\mathbf{\Omega}_{11}| \cdot |\mathbf{\Upsilon}_{22}'| \neq 0. \tag{2.7}$$

Hence (2.3) is a nonsingular linear transformation.

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

$$\mathbf{X}_t = \mathbf{\Psi}_1\mathbf{X}_{t-1} + \mathbf{\Psi}_2\mathbf{X}_{t-2} + \mathbf{W}_t, \tag{2.8}$$

$$\Delta\mathbf{X}_t = \mathbf{\Upsilon}\mathbf{X}_{t-1} + \mathbf{\Upsilon}_1\Delta\mathbf{X}_{t-1} + \mathbf{W}_t, \tag{2.9}$$

where

$$\mathbf{\Psi}_1 = \mathbf{\Omega}'\mathbf{B}_1(\mathbf{\Omega}')^{-1}, \qquad \mathbf{\Psi}_2 = \mathbf{\Omega}'\mathbf{B}_2(\mathbf{\Omega}')^{-1}, \tag{2.10}$$

$$\mathbf{\Upsilon}_1 = \mathbf{\Omega}'\mathbf{\Pi}_1(\mathbf{\Omega}')^{-1} = \mathbf{\Psi}_2, \tag{2.11}$$

$$\boldsymbol{\Upsilon} = \boldsymbol{\Omega}'\boldsymbol{\Pi}(\boldsymbol{\Omega}')^{-1} = \boldsymbol{\Psi}_1 + \boldsymbol{\Psi}_2 - \mathbf{I} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Upsilon}_{22} \end{bmatrix}. \tag{2.12}$$

Partition each of the coefficient matrices $\boldsymbol{\Psi}_1, \boldsymbol{\Psi}_2, \boldsymbol{\Upsilon}, \boldsymbol{\Upsilon}_1$ into n and k rows and columns. Then from (2.8) to (2.12) we obtain

$$\begin{bmatrix} \Delta\mathbf{X_{1t}} \\ \mathbf{X_{2t}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Upsilon}_{11}^{(1)} & \boldsymbol{\Upsilon}_{12}^{(1)} \\ \boldsymbol{\Upsilon}_{21}^{(1)} & \boldsymbol{\Upsilon}_{22}^{(1)} \end{bmatrix} \begin{bmatrix} \Delta\mathbf{X}_{1,t-1} \\ \mathbf{X}_{2,t-1} \end{bmatrix} + \begin{bmatrix} \mathbf{0} & \boldsymbol{\Psi}_{12}^{(2)} \\ \mathbf{0} & \boldsymbol{\Psi}_{22}^{(2)} \end{bmatrix} \begin{bmatrix} \Delta\mathbf{X}_{1,t-2} \\ \mathbf{X}_{2,t-2} \end{bmatrix} + \begin{bmatrix} \mathbf{W}_{1t} \\ \mathbf{W}_{2t} \end{bmatrix} \tag{2.13}$$

as generating the process $\{\Delta\mathbf{X}_{it}', \mathbf{X}_{2,t=1}'\}$. Let $\mathbf{X}_{10} = \mathbf{X}_{1,-1} = \ldots = \mathbf{0}$ and $\mathbf{W}_{10} = \mathbf{W}_{1,-1} = \ldots = \mathbf{0}$. The characteristic polynomial of $\{\mathbf{X}_t\}$ is

$$\begin{aligned}
|\lambda^2\mathbf{I} - \lambda\boldsymbol{\Psi}_1 - \boldsymbol{\Psi}_2| &= \begin{vmatrix} \lambda^2\mathbf{I} - \lambda(\boldsymbol{\Upsilon}_{11} + \boldsymbol{\Upsilon}_{11}^{(1)} + \mathbf{I}) + \boldsymbol{\Upsilon}_{11}^{(1)} & -\lambda\boldsymbol{\Psi}_{12}^{(1)} - \boldsymbol{\Psi}_{12}^{(2)} \\ -\lambda(\boldsymbol{\Upsilon}_{21} + \boldsymbol{\Upsilon}_{21}^{(1)}) + \boldsymbol{\Upsilon}_{21}^{(1)} & \lambda^2\mathbf{I} - \lambda\boldsymbol{\Psi}_{22}^{(1)} - \boldsymbol{\Psi}_{22}^{(2)} \end{vmatrix} \\
&= \begin{vmatrix} (\lambda-1)(\lambda\mathbf{I} - \boldsymbol{\Upsilon}_{11}^{(1)}) & -\lambda\boldsymbol{\Psi}_{12}^{(1)} - \boldsymbol{\Psi}_{12}^{(2)} \\ -(\lambda-1)\boldsymbol{\Upsilon}_{21}^{(1)} & \lambda^2\mathbf{I} - \lambda\boldsymbol{\Psi}_{22}^{(1)} - \boldsymbol{\Psi}_{22}^{(2)} \end{vmatrix} \\
&= (\lambda-1)^n \begin{vmatrix} \lambda\mathbf{I} - \boldsymbol{\Upsilon}_{11}^{(1)} & -\lambda\boldsymbol{\Psi}_{12}^{(1)} - \boldsymbol{\Psi}_{12}^{(2)} \\ -\boldsymbol{\Upsilon}_{21}^{(1)} & \lambda^2\mathbf{I} - \lambda\boldsymbol{\Psi}_{22}^{(1)} - \boldsymbol{\Psi}_{22}^{(2)} \end{vmatrix}.
\end{aligned} \tag{2.14}$$

The characteristic polynomial of $(\Delta\mathbf{X}_{1t}', \mathbf{X}_{2t}')'$ is

$$\begin{vmatrix} \lambda^2\mathbf{I} - \lambda\boldsymbol{\Upsilon}_{11}^{(1)} & -\lambda\boldsymbol{\Psi}_{12}^{(1)} - \boldsymbol{\Psi}_{12}^{(2)} \\ -\lambda\boldsymbol{\Upsilon}_{21}^{(1)} & \lambda^2\mathbf{I} - \lambda\boldsymbol{\Psi}_{22}^{(1)} - \boldsymbol{\Psi}_{22}^{(2)} \end{vmatrix} = \lambda^n \begin{vmatrix} \lambda\mathbf{I} - \boldsymbol{\Upsilon}_{11}^{(1)} & -\lambda\boldsymbol{\Psi}_{12}^{(1)} - \boldsymbol{\Psi}_{12}^{(2)} \\ -\boldsymbol{\Upsilon}_{21}^{(1)} & \lambda^2\mathbf{I} - \lambda\boldsymbol{\Psi}_{22}^{(1)} - \boldsymbol{\Psi}_{22}^{(2)} \end{vmatrix}. \tag{2.15}$$

The factor $(\lambda - 1)^n$ has been replaced by λ^n; the zero of 1 of multiplicity n has been replaced by the zero of 0 of multiplicity n. Thus all of the zeros of (2.15) are less than 1 in absolute value. ■

3 Discussion

3.1 Johansen's condition.

Johansen (1995) uses the error-correction form (1.2). Denote the rank of $\mathbf{\Pi}$ by r. Then $\mathbf{\Pi}$ can be written as $\mathbf{\Pi} = \alpha\beta'$, where α and β are $p \times r$ matrices of rank r. Let $\alpha_\perp$ and $\beta_\perp$ be $p \times (p-r)$ matrices of rank $p-r$ such that $\alpha_\perp'\alpha = \mathbf{0}$ and $\beta_\perp'\beta = \mathbf{0}$. Then Johansen's condition that $\Delta\mathbf{Y}_t$ and $\alpha'\mathbf{Y}_t$ define a stationary process is

Condition J.

$$\text{rank}\,[\alpha_\perp'(\mathbf{I} + \mathbf{B}_2)\beta_\perp] = p - r. \tag{3.1}$$

The orthogonality conditions of $\alpha_\perp$ and $\beta_\perp$ are equivalent to $\alpha_\perp'\mathbf{\Pi} = \mathbf{0}$ and $\mathbf{\Pi}\beta_\perp = \mathbf{0}$. Hence, $\alpha_\perp$ consists of $p-r$ left-sided characteristic vectors of $\mathbf{\Pi}$ corresponding to the characteristic root of 0 and $\beta_\perp$ consists of $p-r$ right-sided characteristic vectors corresponding to the root of 0. Thus $\beta_\perp$ is $\mathbf{\Omega}_1$ multiplied on the right by a nonsingular $n \times n$ matrix with $n = p - r$.

We shall show that Condition A implies Condition J with $n = p - r$. First, Condition A implies that the rank of $\mathbf{\Pi}$ is $p - n = r$ $(= k)$, which implies that $\mathbf{\Pi} = \alpha\beta'$ for α $(p \times r)$ and β $(p \times r)$ of rank r. Hence $\beta_\perp$ is of rank $p - r = n$. We can take $\alpha_\perp = \mathbf{\Omega}_1$ and hence $\mathbf{X}_{1t} = \alpha_\perp'\mathbf{Y}_t$ and α is the first factor in (2.5). The second factor is $\beta' = \mathbf{\Pi}_2 = \mathbf{\Omega}_2'$, and $\mathbf{X}_{2t} = \beta_\perp'\mathbf{Y}_t$. Then $\mathbf{\Omega} = (\alpha_\perp, \beta)$ and $\mathbf{\Omega}^{-1} = (\alpha, \beta_1)$ and

$$\alpha_\perp'(\mathbf{I} + \mathbf{B}_2)\beta_\perp = \mathbf{I} - \mathbf{\Pi}_{11}^{(1)}. \tag{3.2}$$

Anderson (2001a) has shown that $\mathbf{I} - \mathbf{\Pi}_{11}^{(1)}$ is nonsingular under Condition A. (See Anderson (2001b) for more details.) The conclusion is that Condition A implies Condition J. Either condition implies that $(\Delta\mathbf{X}_{1t}', \mathbf{X}_{2t}')'$ is stationary. Note that Condition A is stated in terms of one matrix while Condition J is the produce of three matrices.

3.2 Higher order autoregressive processes

Let $\mathbf{Y}_t$ satisfy

$$\mathbf{Y}_t = \mathbf{B}_1\mathbf{Y}_{t-1} + \ldots + \mathbf{B}_m\mathbf{Y}_{t-m} + \mathbf{Z}_t, \quad t = 1, 2, \ldots, \tag{3.3}$$

with $\mathbf{Y}_0 = \ldots = \mathbf{Y}_{-m+1} = 0$. Suppose the root λ_i of $|\mathbf{B}(\lambda)| = 0$, where

$$B(\lambda) = \lambda^m \mathbf{I} - \lambda^{m-1}\mathbf{B}_1 - \ldots - \mathbf{B}_m \tag{3.4}$$

satisfies $\lambda_i = 1$ or $|\lambda_i| < 1, i = 1, \ldots, mp$. An error-correction form is

$$\Delta\mathbf{Y}_t = \mathbf{\Pi}\mathbf{Y}_{t-1} + \sum_{j=1}^{m-1} \mathbf{\Pi}_j \Delta\mathbf{Y}_{t-j} + \mathbf{Z}_t, \tag{3.5}$$

where

$$\mathbf{\Pi}_j = \sum_{i=1}^{j} \mathbf{B}_i - \mathbf{I}, \qquad j = 1, \ldots, m-1, \tag{3.6}$$

$$\mathbf{\Pi} = \sum_{i=1}^{m} \mathbf{B}_i - \mathbf{I}. \tag{3.7}$$

Then $\mathbf{\Pi} = -\mathbf{B}(1)$. It is assumed that $\lambda = 1$ is a root of $|\mathbf{B}(\lambda)| = 0$ of multiplicity $n < p$ and the other $pm - n$ roots satisfy $|\lambda_i| < 1$.

For the generalized theorem assume the existence of $\mathbf{\Omega}_1$ $(p \times n)$ of rank n satisfying (2.1). Then the conclusions of the theorem in Section 2 follow. The proof follows the pattern of Section 2.

3.3 Inference

The models (1.3) and (3.8) have the form

$$\mathbf{Y}_t = \mathbf{B}_1\mathbf{X}_{1t} + \mathbf{B}_2\mathbf{X}_{2t} + \mathbf{Z}_t, \tag{3.8}$$

where $\mathbf{B}_2$ is of rank k. The maximum likelihood estimator of $\mathbf{B}_2$ is the reduced rank regression estimator introduced by Anderson (1951) and applied by Johansen (1988), (1995), who gave some asymptotic theory suitable for the cointegrated model. Anderson (2001b) has given more details.

References

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.]

Anderson, T.W. (2001a). The asymptotic distribution of canonical correlations and variates in higher-order cointegrated models, *Proceedings of the National Academy of Sciences*, **98**, in press.

Anderson, T. W. (2001b). A study of asymptotic distributions of canonical correlations and variates in higher-order cointegrated models. Technical Report 2001-3, Department of Statistics, Stanford University, February 2001.

Johansen, Soren (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, **12**, 231-254.

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