The asymptotic distribution of canonical correlations and variates in cointegrated models

T. W. Anderson*

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

Contributed by T. W. Anderson, February 24, 2000

The cointegrated model considered here is a nonstationary vector autoregressive process in which some linear functions are stationary and others are random walks. The first difference of the process (the "error-correction form") is stationary. Statistical inference, such as reduced rank regression estimation of the coefficients of the process and tests of hypotheses of dimensionality of the stationary part, involves the canonical correlations between the difference vector and the relevant vector of the past of the process. The asymptotic distributions of the canonical correlations and the canonical vectors under the assumption that the process is Gaussian are found.

1. Introduction

A relatively new multivariate time series model, which has become very important in econometrics, is the cointegrated autoregressive process, which includes stationary and nonstationary aspects. Some linear combinations of the variables are stationary, and some linear combinations are nonstationary. A crucial statistical problem is to distinguish these two sets of linear combinations.

A general mathematical model that is used for these data is an autoregressive model. One set of statistical variables consisting of the present elements of the process is predicted or "caused" by another set consisting of earlier observations. The relations between these two sets can be clarified by canonical correlation analysis. See ref. 1, for example.

A method of handling nonstationary elements in a process, such as trends, is to difference the process (2). The discovery of cointegrating relations involves the differencing of the vector series and using the canonical correlations between the differences and lags of the original series. The nonzero canonical correlations correspond to the cointegrating relations and the zero correlations, to the nonstationary relations.

When some canonical correlations are zero, the reduced rank regression estimator introduced by Anderson (3), which is based on the sample canonical correlations and variates, is a more efficient estimator of the regression or autoregression than the least squares estimator. Here the estimation is of the regression of the differenced series on the lagged variables, which implies estimation of the regression of the original series on the lagged series. The sample canonical correlations may be used to determine the number of process canonical correlations different from 0, which is the rank of the regression of the difference series on the lagged series. Inference is based on the large-sample distribution of the sample canonical correlations and variates.

The model we study is similar to that of Johansen (4). However, we focus on the first-order case, indicating later how it generalizes to higher order.

2. The Model

A general cointegrated model is an autoregressive process

$$\mathbf{Y}_t = \mathbf{B}\mathbf{Y}_{t-1} + \mathbf{Z}_t\,, \qquad [2.1]$$

where $\mathbf{Z}_t$ is unobserved with $\mathcal{E}\mathbf{Y}_{t-1}\mathbf{Z}_t' = \mathbf{0}$, $\mathcal{E}\mathbf{Z}_t\mathbf{Z}_t' = \mathbf{\Sigma}_{ZZ}$. If the eigenvalues of $\mathbf{B}$, that is, the roots $\lambda_1, \ldots, \lambda_p$ of $|\mathbf{B} - \lambda\mathbf{I}| = 0$, satisfy $|\lambda_j| < 1$, $j = 1, \ldots, p$, a stationary process can be defined with $\mathbf{Y}_t = \sum_{s=0}^{\infty} \mathbf{B}^s\mathbf{Z}_{t-s}$, $t = \ldots, -1, 0, 1, \ldots$. If some or all of the eigenvalues are 1, the process will be nonstationary. In this paper, we treat processes $\{\mathbf{Y}_t\}$ with $\mathbf{Y}_0 = \mathbf{0}$ such that some roots, say n roots, are 1, $\lambda_i = 1$, $i = 1, \ldots, n$ $(< p)$, and the other $k = p - n$ roots satisfy $|\lambda_i| < 1$. The first difference of the process is

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

where $\mathbf{\Pi} = \mathbf{B} - \mathbf{I}$ has characteristic roots $\lambda_i - 1$, $i = 1, \ldots, p$, since $|\mathbf{B} - \lambda\mathbf{I}| = |\mathbf{\Pi} - (\lambda - 1)\mathbf{I}|$. Of these roots n are 0 and k are not 0. The form **2.2** is known as the error-correction form (5). As we shall see some linear combinations of $\mathbf{Y}_t$ are stationary. Granger (6) called such models "cointegrated." If a process $\{\mathbf{Y}_t\}$ is stationary, we say it is integrated of order 0 ($\{\mathbf{Y}_t\} \in I(0)$). If $\{\mathbf{Y}_t\}$ is not stationary, but $\Delta\mathbf{Y}_t = \mathbf{Y}_t - \mathbf{Y}_{t-1}$ is stationary, we say it is integrated of order 1 ($\{\mathbf{Y}_t\} \in I(1)$).

A sample consists of T observations: $\mathbf{y}_1, \ldots, \mathbf{y}_T$. An estimator of $\mathbf{B}$ can be obtained from an estimator of $\mathbf{\Pi}$ by adding $\mathbf{I}$ to the estimator of $\mathbf{\Pi}$. Under the assumptions that n of the eigenvalues of $\mathbf{B}$ are 1 and k satisfy $|\lambda_i| < 1$ and that $\boldsymbol{\omega}'\mathbf{B} = \mathbf{0}$ has n linearly independent solutions, $\mathbf{\Pi}$ has rank k and is to be estimated by the reduced rank regression estimator introduced by Anderson in ref. 3. It involves the sample canonical correlations and vectors between $\Delta\mathbf{Y}_t$ and $\mathbf{Y}_{t-1}$. One form of the estimator is

$$\widehat{\mathbf{\Pi}}_k = \mathbf{S}_{\Delta Y,\bar{Y}}\widehat{\mathbf{\Gamma}}\widehat{\mathbf{\Gamma}}'\,, \qquad [2.3]$$

where $\mathbf{S}_{\Delta Y,\bar{Y}} = T^{-1}\sum_{t=1}^{T} \Delta\mathbf{Y}_t\mathbf{Y}_{t-1}'$, etc., and the k columns of $\widehat{\mathbf{\Gamma}}$ satisfy

$$\mathbf{S}_{\bar{Y},\Delta Y}\mathbf{S}_{\Delta Y,\Delta Y}^{-1}\mathbf{S}_{\Delta Y,\bar{Y}}\hat{\boldsymbol{\gamma}} = r^2\mathbf{S}_{\bar{Y}\bar{Y}}\hat{\boldsymbol{\gamma}} \qquad [2.4]$$

and $\hat{\boldsymbol{\gamma}}'\mathbf{S}_{\bar{Y},\bar{Y}}\hat{\boldsymbol{\gamma}} = 1$ for the k largest values of r^2 satisfying

$$|\mathbf{S}_{\bar{Y},\Delta Y}\mathbf{S}_{\Delta Y,\Delta Y}^{-1}\mathbf{S}_{\Delta Y,\bar{Y}} - r^2\mathbf{S}_{\bar{Y}\bar{Y}}| = 0\,. \qquad [2.5]$$

The estimator $\widehat{\mathbf{\Pi}}_k$ is the product of a $p \times k$ matrix and a $k \times p$ matrix. This paper is devoted to finding the asymptotic distribution of the k larger roots of **2.5** and the corresponding vectors $\hat{\boldsymbol{\gamma}}$.

2.1. Transformation of Coordinates. To study the properties of the estimators, tests, and canonical correlations and variables, we transform $\mathbf{Y}_t$ so that the first n coordinates are $I(1)$ and the other k coordinates are $I(0)$.

LEMMA 1. *Given that n roots of $\mathbf{\Pi}$ are 0 and $k = p - n$ roots are different from 0, the rank of $\mathbf{\Pi}$ is k if and only if there exists a $p \times n$ matrix $\mathbf{\Omega}_1$ of rank n such that*

$$\mathbf{\Omega}_1'\mathbf{\Pi} = \mathbf{0}\,. \qquad [2.6]$$

Proof: Since k roots of $\mathbf{\Pi}$ are different from 0 and **2.6** holds, the rank of $\mathbf{\Pi}$ is k. Conversely, if the rank of $\mathbf{\Pi}$ is k, there exist $p \times k$ matrices $\mathbf{A}$ and $\mathbf{C}$ of rank k such that $\mathbf{\Pi} = \mathbf{A}\mathbf{C}'$. ($\mathbf{C}'$ is composed of any k linearly independent rows of $\mathbf{\Pi}$.) Then any $p \times n$ matrix $\mathbf{\Omega}_1$ such that $\mathbf{\Omega}_1'\mathbf{A} = \mathbf{0}$ satisfies **2.6**. ■

*E-mail: twa@stat.stanford.edu.

In the rest of this paper, it is assumed that a matrix $\mathbf{\Omega}_1$ of rank n satisfying **2.6** exists. Define $\mathbf{\Omega}_2 = \mathbf{C}$ and $\mathbf{\Omega} = (\mathbf{\Omega}_1, \mathbf{\Omega}_2)$. Further define $\mathbf{Y}_2 = \mathbf{C}'\mathbf{A}$ (of rank k). Then

$$\mathbf{\Omega}_2'\mathbf{\Pi} = \mathbf{Y}_2\mathbf{\Omega}_2'. \qquad [2.7]$$

The columns of $\mathbf{\Omega}_1$ are left-sided characteristic vectors of $\mathbf{\Pi}$ and $\mathbf{B}$ associated with the roots of 0 and 1, respectively.

Define

$$\mathbf{\Omega}'\mathbf{Y}_t = \mathbf{X}_t = \begin{bmatrix} \mathbf{X}_{1t} \\ \mathbf{X}_{2t} \end{bmatrix}, \quad \mathbf{\Omega}'\mathbf{Z}_t = \mathbf{W}_t = \begin{bmatrix} \mathbf{W}_{1t} \\ \mathbf{W}_{2t} \end{bmatrix}, \qquad [2.8]$$

$$\mathbf{\Psi} = \mathbf{\Omega}'\mathbf{B}(\mathbf{\Omega}')^{-1} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Psi}_2 \end{bmatrix}. \qquad [2.9]$$

Then $\mathbf{X}_{1t} = \sum_{s=0}^{t-1}\mathbf{W}_{1,t-s}$ is a random walk, and $\mathbf{X}_{2t} = \sum_{s=0}^{t-1}\mathbf{\Psi}_2^s \cdot \mathbf{W}_{2,t-s}$ converges to the stationary process $\sum_{s=0}^{\infty}\mathbf{\Psi}_2^s\mathbf{W}_{2,t-s}$. We say that $\mathbf{X}_{2t} = \mathbf{\Omega}_2'\mathbf{Y}_t$ are the cointegrating relations (6). The characteristic roots of $\mathbf{\Psi}_2$ are λ_i for $|\lambda_i| < 1$ because the transformation **2.9** leaves the roots invariant.

Then, let $\mathbf{Y}_2 = \mathbf{\Psi}_2 - \mathbf{I}$ and $\mathbf{Y} = \operatorname{diag}(\mathbf{0}, \mathbf{Y}_2)$. The model in terms of $\Delta\mathbf{X}_t$ and $\mathbf{X}_{t-1}$ is

$$\Delta\mathbf{X}_t = \mathbf{Y}\mathbf{X}_{t-1} + \mathbf{W}_t. \qquad [2.10]$$

3. Process Parameters

From the definitions of $\mathbf{X}_{1t}$ and $\mathbf{X}_{2t}$ we find

$$\begin{aligned} \mathcal{E}\mathbf{X}_{1,t-1}\mathbf{X}_{1,t-1}' &= (t-1)\mathbf{\Sigma}_{WW}^{11}, \\ \mathcal{E}\mathbf{X}_{2,t-1}\mathbf{X}_{1,t-1}' &= \sum_{s=0}^{t-1}\mathbf{\Psi}_2^s\mathbf{\Sigma}_{WW}^{21} \\ &\to (\mathbf{I}-\mathbf{\Psi}_2)^{-1}\mathbf{\Sigma}_{WW}^{21} = -\mathbf{Y}_2^{-1}\mathbf{\Sigma}_{WW}^{21} \end{aligned}$$

as $t \to \infty$, which we define as $\mathbf{\Sigma}_{--}^{21}$,

$$\mathcal{E}\mathbf{X}_{2,t-1}\mathbf{X}_{2,t-1}' = \sum_{s=0}^{t-1}\mathbf{\Psi}_2^s\mathbf{\Sigma}_{WW}^{22}\mathbf{\Psi}_2'^s \to \sum_{s=0}^{\infty}\mathbf{\Psi}_2^s\mathbf{\Sigma}_{WW}^{22}\mathbf{\Psi}_2'^s,$$

which we define as $\mathbf{\Sigma}_{--}^{22}$. Then

$$\mathcal{E}\Delta\mathbf{X}_t\mathbf{X}_{t-1}' \to \mathbf{Y}\mathbf{\Sigma}_{--} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{Y}_2\mathbf{\Sigma}_{--}^{21} & \mathbf{Y}_2\mathbf{\Sigma}_{--}^{22} \end{bmatrix} = \mathbf{\Sigma}_{\Delta-}$$

$$\mathcal{E}\Delta\mathbf{X}_t\Delta\mathbf{X}_t' \to \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{Y}_2\mathbf{\Sigma}_{--}^{22}\mathbf{Y}_2' \end{bmatrix} + \mathbf{\Sigma}_{WW} = \mathbf{\Sigma}_{\Delta\Delta}.$$

Define $\mathbf{\Sigma}_{--}^{11}(T) = T^{-1}\sum_{t=1}^{T}\mathcal{E}\mathbf{X}_{1,t-1}\mathbf{X}_{1,t-1} = \frac{1}{2}(T-1)\mathbf{\Sigma}_{WW}^{11}$, and

$$\mathbf{\Sigma}_{--}(T) = \begin{bmatrix} \mathbf{\Sigma}_{--}^{11}(T) & \mathbf{\Sigma}_{--}^{12} \\ \mathbf{\Sigma}_{--}^{21} & \mathbf{\Sigma}_{--}^{22} \end{bmatrix}.$$

We use subscripts Δ and $-$ for $\Delta\mathbf{X}_t$ and $\mathbf{X}_{t-1}$. Matrices are partitioned into n and k rows and columns.

The canonical correlations and variables for $\Delta\mathbf{X}_t$ and $\mathbf{X}_{t-1}$ are defined by

$$\begin{bmatrix} -\rho\mathbf{\Sigma}_{\Delta\Delta} & \mathbf{\Sigma}_{\Delta-} \\ \mathbf{\Sigma}_{-\Delta} & -\rho\mathbf{\Sigma}_{--}(T) \end{bmatrix}\begin{bmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\gamma} \end{bmatrix} = 0.$$

Elimination of $\boldsymbol{\alpha}$ yields

$$\mathbf{\Sigma}_{\Delta-}\mathbf{\Sigma}_{\Delta\Delta}^{-1}\mathbf{\Sigma}_{\Delta-}\boldsymbol{\gamma} = \rho^2\mathbf{\Sigma}_{--}(T)\boldsymbol{\gamma}, \qquad [3.1]$$

where ρ^2 (which may depend on T) satisfies

$$|\mathbf{\Sigma}_{-\Delta}\mathbf{\Sigma}_{\Delta\Delta}^{-1}\mathbf{\Sigma}_{\Delta-} - \rho^2\mathbf{\Sigma}_{--}(T)| = 0. \qquad [3.2]$$

Algebraic calculation yields

$$\begin{aligned} \mathbf{\Sigma}_{-\Delta}\mathbf{\Sigma}_{\Delta\Delta}^{-1}\mathbf{\Sigma}_{\Delta-} &= \begin{bmatrix} \mathbf{\Sigma}_{--}^{12}\mathbf{Y}_2' \\ \mathbf{\Sigma}_{--}^{22}\mathbf{Y}_2' \end{bmatrix}(\mathbf{Y}_2\mathbf{\Sigma}_{--}^{22}\mathbf{Y}_2' + \mathbf{\Sigma}_{WW}^{22\cdot1})^{-1} \\ &\quad \cdot (\mathbf{Y}_2\mathbf{\Sigma}_{--}^{21}, \mathbf{Y}_2\mathbf{\Sigma}_{--}^{22}), \end{aligned} \qquad [3.3]$$

where $\mathbf{\Sigma}_{WW}^{22\cdot1} = \mathbf{\Sigma}_{WW}^{22} - \mathbf{\Sigma}_{WW}^{21}(\mathbf{\Sigma}_{WW}^{11})^{-1}\mathbf{\Sigma}_{WW}^{12}$. The determinantal equation **3.2** has n roots of 0 and the corresponding vectors $\boldsymbol{\gamma} = (\boldsymbol{\gamma}_1', \boldsymbol{\gamma}_2')'$ satisfy $\mathbf{\Sigma}_{--}^{21}\boldsymbol{\gamma}_1 + \mathbf{\Sigma}_{--}^{22}\boldsymbol{\gamma}_2 = \mathbf{0}$. There are n linearly independent solutions to this equation.

Multiplication of **3.1** on the left by $[\mathbf{I}, -\mathbf{\Sigma}_{--}^{12}(\mathbf{\Sigma}_{--}^{22})^{-1}]$ gives

$$0 = \rho^2\left[\frac{1}{2}T\mathbf{\Sigma}_{WW}^{11} - \mathbf{\Sigma}_{--}^{12}(\mathbf{\Sigma}_{--}^{22})^{-1}\mathbf{\Sigma}_{--}^{21}\right]\boldsymbol{\gamma}_1. \qquad [3.4]$$

For $T \geq 3$, the matrix in **3.3** is nonsingular; hence, $\boldsymbol{\gamma}_1 = \mathbf{0}$ for $\rho^2 > 0$. The second submatrix equation in **3.1** is

$$\mathbf{\Sigma}_{--}^{22}\mathbf{Y}_2'(\mathbf{Y}_2\mathbf{\Sigma}_{--}^{22}\mathbf{Y}_2' + \mathbf{\Sigma}_{WW}^{22\cdot1})^{-1}\mathbf{Y}_2\mathbf{\Sigma}_{--}^{22}\boldsymbol{\gamma}_2 = \rho^2\mathbf{\Sigma}_{--}^{22}\boldsymbol{\gamma}_2. \qquad [3.5]$$

For a nontrivial solution to **3.4** ρ^2 must satisfy

$$\left|\mathbf{\Sigma}_{--}^{22}\mathbf{Y}_2'(\mathbf{Y}_2\mathbf{\Sigma}_{--}^{22}\mathbf{Y}_2' + \mathbf{\Sigma}_{WW}^{22\cdot1})^{-1}\mathbf{Y}_2\mathbf{\Sigma}_{--}^{22} - \rho^2\mathbf{\Sigma}_{--}^{22}\right| = 0. \qquad [3.6]$$

We normalize $\boldsymbol{\gamma}_2$ by $\boldsymbol{\gamma}_2'\mathbf{\Sigma}_{--}^{22}\boldsymbol{\gamma}_2 = 1$. (Since $\boldsymbol{\gamma}_1 = \mathbf{0}$, an equivalent normalization is $\boldsymbol{\gamma}'\mathbf{\Sigma}_{--}(T)\boldsymbol{\gamma} = 1$.)

4. Sample Statistics

The observable sample covariance matrices $\mathbf{S}_{\Delta\Delta}$, $\mathbf{S}_{\Delta-}$, $\mathbf{S}_{--}$ are linear functions of $\mathbf{S}_{--}$, $\mathbf{S}_{-W}$ and $\mathbf{S}_{WW}$. The matrix $\mathbf{S}_{WW}$ is composed of independently identically distributed vectors; $\mathbf{S}_{WW} \xrightarrow{p} \mathbf{\Sigma}_{WW}$. We assume the fourth-order moments are finite; hence, $T^{\frac{1}{2}}(\mathbf{S}_{WW} - \mathbf{\Sigma}_{WW})$ has a limiting normal distribution. Since $\Delta\mathbf{X}_t$ depends only on $\mathbf{X}_{2,t-1}$ and $\mathbf{W}_t$ and $\mathbf{X}_{2t}$ is stationary, $\mathbf{S}_{\Delta\Delta} \xrightarrow{p} \mathbf{\Sigma}_{\Delta\Delta}$ and $T^{\frac{1}{2}}(\mathbf{S}_{\Delta\Delta} - \mathbf{\Sigma}_{\Delta\Delta})$ has a limiting normal distribution.

The lower left-hand corner of $\mathbf{S}_{\Delta-}$ is

$$\begin{aligned} \mathbf{S}_{\Delta-}^{21} &= \frac{1}{T}\sum_{t=1}^{T}(\mathbf{X}_{2t} - \mathbf{X}_{2,t-1})\mathbf{X}_{1,t-1}' \\ &= \frac{1}{T}\left[\sum_{t=2}^{T}\mathbf{X}_{2t}(\mathbf{X}_{1,t-1}' - \mathbf{X}_{1t}') + \mathbf{X}_{2T}\mathbf{X}_{1T}' - \mathbf{X}_{21}\mathbf{X}_{11}'\right] \\ &= \frac{1}{T}\left[-\sum_{t=2}^{T}(\mathbf{Y}_2\mathbf{X}_{2,t-1} + \mathbf{W}_{2t})\mathbf{W}_{1t}' + \mathbf{X}_{2T}\mathbf{X}_{1T}' - \mathbf{X}_{21}\mathbf{X}_{11}'\right] \\ &\xrightarrow{p} -\mathbf{\Sigma}_{WW}^{21}, \end{aligned} \qquad [4.1]$$

because the first term on the right-hand side of **4.1** converges stochastically to $-\mathcal{E}\mathbf{W}_{2t}\mathbf{W}_{1t}' = -\mathbf{\Sigma}_{ZZ}^{21}$ and the second term converges stochastically to $\mathbf{0}$. Note that the sum of squares of the elements of $T^{-1}\mathbf{X}_{2T}\mathbf{X}_{1T}'$ is $T^{-2}\operatorname{tr}\mathbf{X}_{2T}\mathbf{X}_{1T}'(\mathbf{X}_{2T}\mathbf{X}_{1T}')' = T^{-3/2}\mathbf{X}_{1T}'\mathbf{X}_{1T} \cdot T^{-1/2}\mathbf{X}_{2T}'\mathbf{X}_{2T}$; each of the two factors converges in probability to 0.

To obtain the limit of $\mathbf{S}_{\Delta-}$ in distribution, we need to use the vector Brownian motion process. Let $\mathbf{V}_1, \mathbf{V}_2, \ldots$ be independently identically distributed with $\mathcal{E}\mathbf{V}_t = \mathbf{0}$, $\mathcal{E}\mathbf{V}_t\mathbf{V}_t' = \mathbf{\Sigma}$. Then $T^{-\frac{1}{2}}\sum_{t=1}^{[Tu]}\mathbf{V}_t \xrightarrow{w} \mathbf{V}(u)$, $0 \leq u \leq 1$, the Brownian motion process. If $\mathbf{\Sigma} = \mathbf{I}$, we term $\mathbf{V}(u)$ the standard Brownian motion. For more detail see ref. 4, appendix B.7, or ref. 7.

Define $\mathbf{J}_{11}$ and $\mathbf{J}_{21}$ by

$$\begin{aligned} \mathbf{S}_{W-}^{j1} &= \frac{1}{T}\sum_{t=1}^{T}\mathbf{W}_{jt}\mathbf{X}_{1,t-1}' \\ &\xrightarrow{d} \int_0^1 d\mathbf{W}_j(u)\mathbf{W}_1'(u) = \mathbf{J}_{j1}, \qquad j = 1, 2. \end{aligned} \qquad [4.2]$$

ECONOMIC SCIENCES

STATISTICS

Then

$$\mathbf{S}_{\Delta-} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{Y}_2\mathbf{S}_{--}^{21} & \mathbf{Y}_2\mathbf{S}_{--}^{22} \end{bmatrix} + \begin{bmatrix} \mathbf{S}_{W-}^{11} & \mathbf{S}_{W-}^{12} \\ \mathbf{S}_{W-}^{21} & \mathbf{S}_{W-}^{22} \end{bmatrix} \xrightarrow{d} \begin{bmatrix} \mathbf{J}_{11} & \mathbf{0} \\ -\mathbf{\Sigma}_{WW}^{21} & \mathbf{Y}_2\mathbf{\Sigma}_{--}^{22} \end{bmatrix}.$$

Let $\mathbf{Q} = \mathbf{S}_{-\Delta}\mathbf{S}_{\Delta\Delta}^{-1}\mathbf{S}_{\Delta-}$. Then

$$\mathbf{Q} \xrightarrow{d} \begin{bmatrix} \mathbf{J}_{11}' & -\mathbf{\Sigma}_{WW}^{12} \\ \mathbf{0} & \mathbf{\Sigma}_{--}^{22}\mathbf{Y}_2' \end{bmatrix} \mathbf{\Sigma}_{\Delta\Delta}^{-1} \begin{bmatrix} \mathbf{J}_{11} & \mathbf{0} \\ -\mathbf{\Sigma}_{WW}^{21} & \mathbf{Y}_2\mathbf{\Sigma}_{--}^{22} \end{bmatrix}.$$

From $\mathbf{S}_{\Delta-}^{21} = \mathbf{Y}_2\mathbf{S}_{--}^{21} + \mathbf{S}_{W-}^{21}$, we have

$$\mathbf{S}_{--}^{21} = \mathbf{Y}_2^{-1}(\mathbf{S}_{\Delta-}^{21} - \mathbf{S}_{W-}^{21}) \xrightarrow{d} -\mathbf{Y}_2^{-1}(\mathbf{\Sigma}_{WW}^{21} + \mathbf{J}_{21}).$$

Then

$$\begin{bmatrix} \frac{1}{T}\mathbf{S}_{--}^{11} & \mathbf{S}_{--}^{12} \\ \mathbf{S}_{--}^{21} & \mathbf{S}_{--}^{22} \end{bmatrix} \xrightarrow{d} \begin{bmatrix} \mathbf{I}_{11} & -(\mathbf{\Sigma}_{WW}^{12} + \mathbf{J}_{21}')\mathbf{Y}_2'^{-1} \\ -\mathbf{Y}_2^{-1}(\mathbf{\Sigma}_{WW}^{21} + \mathbf{J}_{21}) & \mathbf{\Sigma}_{--}^{22} \end{bmatrix},$$

where $\mathbf{I}_{11}$ is defined as follows:

$$\frac{1}{T}\mathbf{S}_{--}^{11} = \frac{1}{T^2}\sum_{t=1}^{T}\mathbf{X}_{1,t-1}\mathbf{X}_{1,t-1}' \xrightarrow{d} \int_0^1 \mathbf{W}_1(u)\mathbf{W}_1'(u)\,du = \mathbf{I}_{11}. \quad \textbf{[4.3]}$$

5. Asymptotic Distribution of the Smaller Roots

We are interested in the asymptotic distribution of the eigenvalues and eigenvectors satisfying $\mathbf{Q}\mathbf{g} = r^2\mathbf{S}_{--}\mathbf{g}$. First consider the roots of $|\mathbf{Q} - r^2\mathbf{S}_{--}| = 0$. Multiplication of this determinant on the left and right by the determinant of $\mathrm{diag}(T^{-\frac{1}{2}}\mathbf{I}, \mathbf{I})$ yields

$$\left| \begin{bmatrix} \frac{1}{T}\mathbf{Q}_{11} & \frac{1}{\sqrt{T}}\mathbf{Q}_{21} \\ \frac{1}{\sqrt{T}}\mathbf{Q}_{21} & \mathbf{Q}_{22} \end{bmatrix} - r^2 \begin{bmatrix} \frac{1}{T}\mathbf{S}_{--}^{11} & \frac{1}{\sqrt{T}}\mathbf{S}_{--}^{12} \\ \frac{1}{\sqrt{T}}\mathbf{S}_{--}^{21} & \mathbf{S}_{--}^{22} \end{bmatrix} \right| = 0. \quad \textbf{[5.1]}$$

Since $T^{-1}\mathbf{Q}_{11} \xrightarrow{p} \mathbf{0}$, $T^{-\frac{1}{2}}\mathbf{Q}_{21} \xrightarrow{p} \mathbf{0}$, $\mathbf{Q}_{22} \xrightarrow{p} (\mathbf{\Sigma}_{-\Delta}\mathbf{\Sigma}_{\Delta\Delta}^{-1}\mathbf{\Sigma}_{\Delta-})_{22}$, $T^{-1}\mathbf{S}_{--}^{11} \xrightarrow{d} \mathbf{I}_{11}$, $T^{-\frac{1}{2}}\mathbf{S}_{--}^{12} \xrightarrow{p} \mathbf{0}$, and $\mathbf{S}_{--}^{22} \xrightarrow{p} \mathbf{\Sigma}_{--}^{22}$, the n smaller roots of **5.1** converge in probability to 0 and the k larger roots converge to the roots of **3.5**.

To study the behavior of the smaller roots let $Tr^2 = d$ to obtain the equation $|\mathbf{Q} - dT^{-1}\mathbf{S}_{--}| = 0$. The n smaller roots of this equation converge in distribution to the roots of the limit in distribution of

$$0 = \left| \begin{bmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} \\ \mathbf{Q}_{21} & \mathbf{Q}_{22} \end{bmatrix} - d \begin{bmatrix} \mathbf{I}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \right| = |\mathbf{Q}_{22}| \cdot |\mathbf{Q}_{11\cdot2} - d\mathbf{I}_{11}|, \quad \textbf{[5.2]}$$

where $\mathbf{Q}_{11\cdot2} = \mathbf{Q}_{11} - \mathbf{Q}_{12}\mathbf{Q}_{22}^{-1}\mathbf{Q}_{21}$. Some algebra shows that **5.2** is asymptotically

$$|\mathbf{J}_{11}'(\mathbf{\Sigma}_{WW}^{11})^{-1}\mathbf{J}_{11} - d\mathbf{I}_{11}| = 0. \quad \textbf{[5.3]}$$

We can write $\mathbf{W}_1(u) = (\mathbf{\Sigma}_{WW}^{11})^{\frac{1}{2}}\mathbf{B}_1(u)$, where $\mathbf{B}_1(u)$ is standard Brownian motion with $\mathcal{E}\mathbf{B}_1(u)\mathbf{B}_1'(u) = u\mathbf{I}$. Then the zeros of **5.3** are the zeros of

$$\left| \int_0^1 \mathbf{B}_1(u)\,d\mathbf{B}_1'(u) \int_0^1 d\mathbf{B}_1(v)\mathbf{B}_1'(v) - d\int_0^1 \mathbf{B}_1(u)\mathbf{B}_1'(u)\,du \right|.$$

The distribution does not depend on any parameters nor does it require normality of $\mathbf{Z}_t$.

The likelihood ratio criterion for testing rank $\mathbf{Y} = k$ found by Anderson (3) is

$$\begin{aligned} -2\log\lambda &= -T\sum_{i=1}^{p-k}\log(1 - r_i^2) \\ &= T\sum_{i=1}^{p-k} r_i^2 + o_p(1) = \sum_{i=1}^{p-k} d_i + o_p(1) \\ &= \mathrm{tr}\int_0^1 d\mathbf{B}_1(v)\mathbf{B}_1'(v)\left[\int_0^1 \mathbf{B}_1(u)\mathbf{B}_1'(u)\,du\right]^{-1} \\ &\quad \cdot \int_0^1 \mathbf{B}_1(v)\,d\mathbf{B}_1'(v) + o_p(1); \end{aligned} \quad \textbf{[5.4]}$$

see Johansen (4, 8). The roots are ordered $r_1^2 < \cdots < r_p^2$.

Let the p solutions to $\mathbf{Q}\mathbf{g} = r^2\mathbf{S}_{--}\mathbf{g}$ be $\mathbf{G} = (\mathbf{g}_1, \ldots, \mathbf{g}_p)$ and $\widehat{\mathbf{R}}^2 = \mathrm{diag}(r_1^2, \ldots, r_p^2)$ with the normalization of the columns of $\mathbf{G}$ to be determined. Then the partitioned form of $\mathbf{Q}\mathbf{G} = \mathbf{S}_{--}\mathbf{G}\widehat{\mathbf{R}}^2$ is

$$\begin{bmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} \\ \mathbf{Q}_{21} & \mathbf{Q}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{G}_{11} & \mathbf{G}_{12} \\ \mathbf{G}_{21} & \mathbf{G}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{S}_{--}^{11} & \mathbf{S}_{--}^{12} \\ \mathbf{S}_{--}^{21} & \mathbf{S}_{--}^{22} \end{bmatrix} \begin{bmatrix} \mathbf{G}_{11} & \mathbf{G}_{12} \\ \mathbf{G}_{21} & \mathbf{G}_{22} \end{bmatrix} \begin{bmatrix} \widehat{\mathbf{R}}_1^2 & \mathbf{0} \\ \mathbf{0} & \widehat{\mathbf{R}}_2^2 \end{bmatrix}. \quad \textbf{[5.5]}$$

Since $T\widehat{\mathbf{R}}_1^2 = \mathbf{D}_1$ has a limiting distribution, the weak limits of the first n columns of **5.5** give

$$\begin{bmatrix} \mathbf{Q}_{11}\mathbf{G}_{11} + \mathbf{Q}_{12}\mathbf{G}_{21} \\ \mathbf{Q}_{21}\mathbf{G}_{11} + \mathbf{Q}_{22}\mathbf{G}_{21} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_{11}\mathbf{G}_{11}\mathbf{D}_1 \\ \mathbf{0} \end{bmatrix} + o_p(1). \quad \textbf{[5.6]}$$

The last k rows of **5.6** give $\mathbf{G}_{21} = -\mathbf{Q}_{22}^{-1}\mathbf{Q}_{21}\mathbf{G}_{11} + o_p(1)$. Insertion of this in the first n rows of **5.6** yields

$$\mathbf{Q}_{11\cdot2}\mathbf{G}_{11} = \mathbf{I}_{11}\mathbf{G}_{11}\mathbf{D}_1 + o_p(1). \quad \textbf{[5.7]}$$

The columns of $\mathbf{G}_{11}$ are the characteristic vectors of $\mathbf{Q}_{11\cdot2}$ in the metric of $\mathbf{I}_{11}$; the diagonal elements of $\mathbf{D}_1$ are the ordered roots of **5.2**.

Now multiply **5.7** on the left by $(\mathbf{\Sigma}_{WW}^{11})^{\frac{1}{2}}$ and define $(\mathbf{\Sigma}_{WW}^{11})^{-\frac{1}{2}}\mathbf{G}_{11} = \mathbf{L}$ to obtain

$$\int_0^1 \mathbf{B}_1(u)\,d\mathbf{B}_1'(u)\int_0^1 d\mathbf{B}_1(v)\mathbf{B}_1'(v)\mathbf{L} = \int_0^1 \mathbf{B}_1(u)\mathbf{B}_1'(u)\,du\,\mathbf{L}\mathbf{D}_1. \quad \textbf{[5.8]}$$

At this point we normalize $\mathbf{L}$ by $\mathbf{L}'\mathbf{I}_{11}\mathbf{L} = \mathbf{I}$. The distribution of the integrals is invariant with respect to multiplying on the left by an arbitrary orthogonal matrix and on the right by the transpose of that matrix. This fact implies that the distribution of $\mathbf{L}$ is invariant with respect to multiplication on the left by an arbitrary orthogonal matrix. It has the Haar uniform measure in the space of orthogonal matrices.

6. Asymptotic Distribution of the Larger Roots

Let $r_i^{2*} = \sqrt{T}\,(r_i^2 - \rho_i^2)$, $i = n+1, \ldots, p$. We shall find the limiting normal distribution of $r_{n+1}^{2*}, \ldots, r_p^{2*}$ from the limiting normal distribution of $\left\{\sqrt{T}[\mathbf{Q}_{22} - (\mathbf{\Sigma}_{-\Delta}\mathbf{\Sigma}_{\Delta\Delta}^{-1}\mathbf{\Sigma}_{\Delta-})_{22}],\ \sqrt{T}(\mathbf{S}_{--}^{22} - \mathbf{\Sigma}_{--}^{22})\right\}$. To carry out this program, we make a series of transformations to bring $(\mathbf{\Sigma}_{-\Delta}\mathbf{\Sigma}_{\Delta\Delta}^{-1}\mathbf{\Sigma}_{\Delta-})_{22}$ and $\mathbf{\Sigma}_{--}^{22}$ to diagonal forms.

The last k columns of **5.5** can be written

$$\frac{1}{T}\mathbf{Q}_{11}\sqrt{T}\,\mathbf{G}_{12} + \frac{1}{\sqrt{T}}\mathbf{Q}_{12}\mathbf{G}_{22} = \left(\frac{1}{T}\mathbf{S}^{11}_{--}\sqrt{T}\,\mathbf{G}_{12} + \frac{1}{\sqrt{T}}\mathbf{S}^{12}_{--}\mathbf{G}_{22}\right)\widehat{\mathbf{R}}^2_2, \qquad [6.1]$$

$$\frac{1}{\sqrt{T}}\mathbf{Q}_{21}\sqrt{T}\,\mathbf{G}_{12} + \mathbf{Q}_{22}\mathbf{G}_{22} = \left(\frac{1}{\sqrt{T}}\mathbf{S}^{21}_{--}\sqrt{T}\,\mathbf{G}_{12} + \mathbf{S}^{22}_{--}\mathbf{G}_{22}\right)\widehat{\mathbf{R}}^2_2. \qquad [6.2]$$

We normalize $\mathbf{G}_{12}$ and $\mathbf{G}_{22}$ by

$$\mathbf{I} = (\mathbf{G}'_{12}, \mathbf{G}'_{22})\begin{bmatrix}\mathbf{S}^{11}_{--} & \mathbf{S}^{12}_{--}\\ \mathbf{S}^{21}_{--} & \mathbf{S}^{22}_{--}\end{bmatrix}\begin{bmatrix}\mathbf{G}_{12}\\ \mathbf{G}_{22}\end{bmatrix} = (\sqrt{T}\,\mathbf{G}'_{12}, \mathbf{G}'_{22})\begin{bmatrix}\frac{1}{T}\mathbf{S}^{11}_{--} & \frac{1}{\sqrt{T}}\mathbf{S}^{12}_{--}\\ \frac{1}{\sqrt{T}}\mathbf{S}^{21}_{--} & \mathbf{S}^{22}_{--}\end{bmatrix}\begin{bmatrix}\sqrt{T}\,\mathbf{G}_{12}\\ \mathbf{G}_{22}\end{bmatrix}. \qquad [6.3]$$

Since the second matrix in **6.3** is positive definite with probability 1 and converges in distribution to $\mathrm{diag}(\mathbf{I}_{11}, \boldsymbol{\Sigma}^{22}_{--})$, $\sqrt{T}\,\mathbf{G}_{12} = O_p(1)$ and $\mathbf{G}_{22} = O_p(1)$. The probability limit of **6.1** shows that $\sqrt{T}\,\mathbf{G}_{12} \xrightarrow{p} \mathbf{0}$ and

$$\mathbf{Q}_{22}\mathbf{G}_{22} = \mathbf{S}^{22}_{--}\mathbf{G}_{22}\widehat{\mathbf{R}}^2_2 + o_p(1). \qquad [6.4]$$

Then the probability limit of **6.3** and $\mathbf{S}^{22}_{--} \xrightarrow{p} \boldsymbol{\Sigma}^{22}_{--}$ shows $\mathbf{G}'_{22}\boldsymbol{\Sigma}^{22}_{--}\mathbf{G}_{22} = \mathbf{I} + o_p(1)$. This and **6.4** imply that $\mathbf{G}_{22} \xrightarrow{p} \boldsymbol{\Gamma}_{22}$.

When $\mathrm{plim}_{T\to\infty}\sqrt{T}\,\mathbf{G}_{12} = \mathbf{0}$ is put into **6.2**, we obtain $\sqrt{T}\,\mathbf{Q}_{22}\mathbf{G}_{22} = \sqrt{T}\,\mathbf{S}^{22}_{--}\mathbf{G}_{22}\widehat{\mathbf{R}}^2_2 + o_p(1)$. This justifies expanding **6.4** to terms $O_p(1/\sqrt{T})$.

Now we want to expand $\sqrt{T}[\mathbf{Q}_{22} - (\boldsymbol{\Sigma}_{-\Delta}\boldsymbol{\Sigma}^{-1}_{\Delta\Delta}\boldsymbol{\Sigma}_{\Delta-})_{22}]$. Define $\mathbf{S}^{*}_{\Delta-} = \sqrt{T}(\mathbf{S}_{\Delta-} - \boldsymbol{\Sigma}_{\Delta-})$, $\mathbf{S}^{*22}_{--} = \sqrt{T}\,(\mathbf{S}^{22}_{--} - \boldsymbol{\Sigma}^{22}_{--})$, $\mathbf{S}^{*2,2\cdot1}_{-W} = T^{-\frac{1}{2}}\sum_{t=1}^{T}\mathbf{X}_{2,t-1}\mathbf{W}'_{2\cdot1,t}$, and $\mathbf{S}^{*2\cdot1,2\cdot1}_{WW} = T^{-\frac{1}{2}}\sum_{t=1}^{T}(\mathbf{W}_{2\cdot1,t}\mathbf{W}'_{2\cdot1,t} - \boldsymbol{\Sigma}^{22\cdot1}_{WW})$, where $\mathbf{W}_{2\cdot1,t} = \mathbf{W}_{2t} - \boldsymbol{\Sigma}^{21}_{WW}(\boldsymbol{\Sigma}^{11}_{WW})^{-1}\mathbf{W}_{1t}$. Then

$$\boldsymbol{\Upsilon}_2(\mathbf{B}_{-\Delta}\mathbf{S}^{*}_{\Delta-})_{22}\boldsymbol{\Upsilon}'_2 = \boldsymbol{\Upsilon}_2\boldsymbol{\Sigma}^{22}_{--}\boldsymbol{\Upsilon}'_2\boldsymbol{\Lambda}^{-1}(\boldsymbol{\Upsilon}_2\mathbf{S}^{*22}_{--}\boldsymbol{\Upsilon}'_2 + \mathbf{S}^{*2\cdot1,2}_{W-}\boldsymbol{\Upsilon}'_2),$$

$$\boldsymbol{\Upsilon}_2(\mathbf{B}_{-\Delta}\mathbf{S}^{*}_{\Delta\Delta}\mathbf{B}'_{-\Delta})_{22}\boldsymbol{\Upsilon}'_2 = \boldsymbol{\Upsilon}_2\boldsymbol{\Sigma}^{22}_{--}\boldsymbol{\Upsilon}'_2\boldsymbol{\Lambda}^{-1}\cdot(\boldsymbol{\Upsilon}_2\mathbf{S}^{*22}_{--}\boldsymbol{\Upsilon}'_2 + \boldsymbol{\Upsilon}_2\mathbf{S}^{*2,2\cdot1}_{-W} + \mathbf{S}^{*2\cdot1,2}_{W-}\boldsymbol{\Upsilon}'_2 + \mathbf{S}^{*2\cdot1,2\cdot1}_{WW})\cdot\boldsymbol{\Lambda}^{-1}\boldsymbol{\Upsilon}_2\boldsymbol{\Sigma}^{22}_{--}\boldsymbol{\Upsilon}'_2,$$

where $\boldsymbol{\Lambda} = \left(\boldsymbol{\Upsilon}_2\boldsymbol{\Sigma}^{22}_{--}\boldsymbol{\Upsilon}'_2 + \boldsymbol{\Sigma}^{22\cdot1}_{WW}\right)$ and $\mathbf{B}_{-\Delta} = \boldsymbol{\Sigma}_{-\Delta}\boldsymbol{\Sigma}^{-1}_{\Delta\Delta}$. Then

$$\begin{aligned}&\sqrt{T}\boldsymbol{\Upsilon}_2[\mathbf{Q}_{22} - (\boldsymbol{\Sigma}_{-\Delta}\boldsymbol{\Sigma}^{-1}_{\Delta\Delta}\boldsymbol{\Sigma}_{\Delta-})_{22}]\boldsymbol{\Upsilon}'_2 + o_p(1)\\ &= \boldsymbol{\Upsilon}_2\left(\mathbf{S}^{*}_{-\Delta}\mathbf{B}'_{-\Delta} - \mathbf{B}_{-\Delta}\mathbf{S}^{*}_{\Delta\Delta}\mathbf{B}'_{-\Delta} + \mathbf{B}_{-\Delta}\mathbf{S}^{*}_{\Delta-}\right)_{22}\boldsymbol{\Upsilon}'_2\\ &= -\boldsymbol{\Upsilon}_2\boldsymbol{\Sigma}^{22}_{--}\boldsymbol{\Upsilon}'_2\boldsymbol{\Lambda}^{-1}\mathbf{S}^{*2\cdot1,2\cdot1}_{WW}\boldsymbol{\Lambda}^{-1}\boldsymbol{\Upsilon}_2\boldsymbol{\Sigma}^{22}_{--}\boldsymbol{\Upsilon}'_2\\ &\quad + \boldsymbol{\Upsilon}_2\boldsymbol{\Sigma}^{22}_{--}\boldsymbol{\Upsilon}'_2\boldsymbol{\Lambda}^{-1}\mathbf{S}^{*2\cdot1,2}_{W-}\boldsymbol{\Upsilon}'_2\boldsymbol{\Lambda}^{-1}\boldsymbol{\Sigma}^{22\cdot1}_{WW}\\ &\quad + \boldsymbol{\Sigma}^{22\cdot1}_{WW}\boldsymbol{\Lambda}^{-1}\boldsymbol{\Upsilon}_2\mathbf{S}^{*2,2\cdot1}_{-W}\boldsymbol{\Lambda}^{-1}\boldsymbol{\Upsilon}_2\boldsymbol{\Sigma}^{22}_{--}\boldsymbol{\Upsilon}'_2\\ &\quad + \boldsymbol{\Upsilon}_2\boldsymbol{\Sigma}^{22}_{--}\boldsymbol{\Upsilon}'_2\boldsymbol{\Lambda}^{-1}\boldsymbol{\Upsilon}_2\mathbf{S}^{*22}_{--}\boldsymbol{\Upsilon}'_2\\ &\quad + \boldsymbol{\Upsilon}_2\mathbf{S}^{*22}_{--}\boldsymbol{\Upsilon}'_2\boldsymbol{\Lambda}^{-1}\boldsymbol{\Upsilon}_2\boldsymbol{\Sigma}^{22}_{--}\boldsymbol{\Upsilon}'_2\\ &\quad - \boldsymbol{\Upsilon}_2\boldsymbol{\Sigma}^{22}_{--}\boldsymbol{\Upsilon}'_2\boldsymbol{\Lambda}^{-1}\boldsymbol{\Upsilon}_2\mathbf{S}^{*22}_{--}\boldsymbol{\Upsilon}'_2\boldsymbol{\Lambda}^{-1}\boldsymbol{\Upsilon}_2\boldsymbol{\Sigma}^{22}_{--}\boldsymbol{\Upsilon}'_2. \qquad [6.5]\end{aligned}$$

Let $\boldsymbol{\Xi}$ be a $k \times k$ matrix such that $\boldsymbol{\Xi}'(\boldsymbol{\Upsilon}_2\boldsymbol{\Sigma}^{22}_{--}\boldsymbol{\Upsilon}'_2)\boldsymbol{\Xi} = \boldsymbol{\Theta}$, $\boldsymbol{\Xi}'\boldsymbol{\Sigma}^{22\cdot1}_{WW}\boldsymbol{\Xi} = \mathbf{I}$, where $\boldsymbol{\Theta} = \mathrm{diag}(\theta_{n+1}, \ldots, \theta_p) = \mathbf{R}^2_2(\mathbf{I} - \mathbf{R}^2_2)^{-1}$, $\mathbf{R}^2_2 = \mathrm{diag}(\rho^2_{n+1}, \ldots, \rho^2_p)$ and ρ^2_i is a root of **3.6**. Let $\mathbf{U}_{2t} = \boldsymbol{\Xi}'\mathbf{X}_{2t}$, $\mathbf{V}_{2t} = \boldsymbol{\Xi}'\mathbf{W}_{2t}$, $\mathbf{V}_{1t} = \mathbf{W}_{1t}$. Then $\mathbf{U}_{2t}$ satisfies

$$\mathbf{U}_{2t} = \boldsymbol{\Delta}_2\mathbf{U}_{2,t-1} + \mathbf{V}_{2t}, \quad \Delta\mathbf{U}_{2t} = \mathbf{M}_2\mathbf{U}_{2,t-1} + \mathbf{V}_{2t}, \qquad [6.6]$$

where $\boldsymbol{\Delta}_2 = \boldsymbol{\Xi}'\boldsymbol{\Psi}_2(\boldsymbol{\Xi}')^{-1}$, and $\mathbf{M}_2 = \boldsymbol{\Delta}_2 - \mathbf{I} = \boldsymbol{\Xi}'\boldsymbol{\Upsilon}_2(\boldsymbol{\Xi}')^{-1}$. Note that $\boldsymbol{\Xi}'\boldsymbol{\Lambda}\boldsymbol{\Xi} = \boldsymbol{\Theta} + \mathbf{I}$. From **3.3** we obtain $\boldsymbol{\Xi}'\boldsymbol{\Upsilon}_2\left(\boldsymbol{\Sigma}_{-\Delta}\boldsymbol{\Sigma}^{-1}_{\Delta\Delta}\boldsymbol{\Sigma}_{\Delta-}\right)_{22}\cdot\boldsymbol{\Upsilon}'_2\boldsymbol{\Xi} = \boldsymbol{\Theta}(\boldsymbol{\Theta}+\mathbf{I})^{-1}\boldsymbol{\Theta}$. Multiplication of **6.5** on the left by $\boldsymbol{\Xi}'$ and the right by $\boldsymbol{\Xi}$ gives

$$\begin{aligned}&\sqrt{T}\left[\mathbf{M}_2(\mathbf{S}_{\bar{U},\Delta U}\mathbf{S}^{-1}_{\Delta U,\Delta U}\mathbf{S}_{\Delta U,\bar{U}})_{22}\mathbf{M}'_2 - \boldsymbol{\Theta}(\boldsymbol{\Theta}+\mathbf{I})^{-1}\boldsymbol{\Theta}\right]\\ &= -\boldsymbol{\Theta}(\mathbf{I}+\boldsymbol{\Theta})^{-1}\mathbf{S}^{*2\cdot1,2\cdot1}_{VV}(\boldsymbol{\Theta}+\mathbf{I})^{-1}\boldsymbol{\Theta}\\ &\quad + \boldsymbol{\Theta}(\mathbf{I}+\boldsymbol{\Theta})^{-1}\mathbf{S}^{*2\cdot1,2}_{V\bar{U}}\mathbf{M}'_2(\boldsymbol{\Theta}+\mathbf{I})^{-1}\\ &\quad + (\mathbf{I}+\boldsymbol{\Theta})^{-1}\mathbf{M}_2\mathbf{S}^{*2,2\cdot1}_{\bar{U}V}(\boldsymbol{\Theta}+\mathbf{I})^{-1}\boldsymbol{\Theta}\\ &\quad + \boldsymbol{\Theta}(\mathbf{I}+\boldsymbol{\Theta})^{-1}\mathbf{M}_2\mathbf{S}^{*22}_{\bar{U}\bar{U}}\mathbf{M}'_2 + \mathbf{M}_2\mathbf{S}^{*22}_{\bar{U}\bar{U}}\mathbf{M}'_2\boldsymbol{\Theta}(\mathbf{I}+\boldsymbol{\Theta})^{-1}\\ &\quad - \boldsymbol{\Theta}(\mathbf{I}+\boldsymbol{\Theta})^{-1}\mathbf{M}_2\mathbf{S}^{*22}_{\bar{U}\bar{U}}\mathbf{M}'_2(\mathbf{I}+\boldsymbol{\Theta})^{-1}\boldsymbol{\Theta} + o_p(1). \qquad [6.7]\end{aligned}$$

Let $\mathbf{M}_2\mathbf{U}_{2,t-1} = \boldsymbol{\Xi}'\boldsymbol{\Upsilon}_2\mathbf{X}_{2,t-1} = \mathbf{L}_{2,t-1}$. Then **6.7** transforms to

$$\begin{aligned}&\sqrt{T}\left[(\mathbf{S}_{\bar{L},\Delta U}\mathbf{S}^{-1}_{\Delta U,\Delta U}\mathbf{S}_{\Delta U,\bar{L}})_{22} - \boldsymbol{\Theta}^2(\mathbf{I}+\boldsymbol{\Theta})^{-1}\right]\\ &= -\mathbf{R}^2_2\mathbf{S}^{*2\cdot1,2\cdot1}_{VV}\mathbf{R}^2_2 + \mathbf{R}^2_2\mathbf{S}^{*2\cdot1,2}_{V\bar{L}}(\mathbf{I}-\mathbf{R}^2_2)\\ &\quad + (\mathbf{I}-\mathbf{R}^2_2)\mathbf{S}^{*2,2\cdot1}_{\bar{L}V}\mathbf{R}^2_2 + \mathbf{R}^2_2\mathbf{S}^{*22}_{\bar{L}\bar{L}}\\ &\quad + \mathbf{S}^{*22}_{\bar{L}\bar{L}}\mathbf{R}^2_2 - \mathbf{R}^2_2\mathbf{S}^{*22}_{\bar{L}\bar{L}}\mathbf{R}^2_2 + o_p(1).\end{aligned}$$

Let $\mathbf{H}_{22} = (\mathbf{M}'_2)^{-1}\boldsymbol{\Xi}^{-1}\mathbf{G}_{22} = \boldsymbol{\Xi}^{-1}(\boldsymbol{\Upsilon}'_2)^{-1}\mathbf{G}_{22}$. Then $\mathbf{Q}_{22}\mathbf{G}_{22} = \mathbf{S}^{22}_{--}\mathbf{G}_{22}\widehat{\mathbf{R}}^2_2$ transforms to

$$\left(\mathbf{S}_{\bar{L},\Delta U}\mathbf{S}^{-1}_{\Delta U,\Delta U}\mathbf{S}_{\Delta U,\bar{L}}\right)_{22}\mathbf{H}_{22} = \mathbf{S}^{22}_{\bar{L}\bar{L}}\mathbf{H}_{22}\widehat{\mathbf{R}}^2_2, \qquad [6.8]$$

and $\mathbf{G}'_{22}\mathbf{S}^{22}_{--}\mathbf{G}_{22} = \mathbf{I}$ transforms to

$$\mathbf{H}'_{22}\mathbf{S}^{22}_{\bar{L}\bar{L}}\mathbf{H}_{22} = \mathbf{I}. \qquad [6.9]$$

Then **6.8** converges in probability to $\boldsymbol{\Theta}^2(\mathbf{I}+\boldsymbol{\Theta})^{-1}\mathbf{H}_{22} = \boldsymbol{\Theta}\mathbf{H}_{22}\mathbf{R}^2_2$, which implies $\mathrm{plim}\,\mathbf{H}_{22}$ is diagonal (since the diagonal elements of $\boldsymbol{\Theta}$ are distinct). Then the probability limit of **6.9** and $h_{ii} > 0$ implies $\mathbf{H}_{22} \xrightarrow{p} \boldsymbol{\Theta}^{-\frac{1}{2}}$.

Define $\mathbf{H}^{*}_{22} = \sqrt{T}(\mathbf{H}_{22} - \boldsymbol{\Theta}^{-\frac{1}{2}})$ and $\widehat{\mathbf{R}}^{2*}_2 = \sqrt{T}(\widehat{\mathbf{R}}^2_2 - \mathbf{R}^2_2)$. Then **6.8** is

$$\begin{aligned}&\boldsymbol{\Theta}^{-1}\Big[-\mathbf{R}^2_2\mathbf{S}^{*2\cdot1,2\cdot1}_{VV}\mathbf{R}^2_2 + \mathbf{R}^2_2\mathbf{S}^{*2\cdot1,2}_{V\bar{L}}(\mathbf{I}-\mathbf{R}^2_2) + (\mathbf{I}-\mathbf{R}^2_2)\mathbf{S}^{*2,2\cdot1}_{\bar{L}V}\mathbf{R}^2_2\\ &\quad + \mathbf{R}^2_2\mathbf{S}^{*22}_{\bar{L}\bar{L}}(\mathbf{I}-\mathbf{R}^2_2)\Big]\boldsymbol{\Theta}^{-\frac{1}{2}}\\ &= \boldsymbol{\Theta}^{-\frac{1}{2}}\widehat{\mathbf{R}}^{2*}_2 + \mathbf{H}^{*}_{22}\mathbf{R}^2_2 - \mathbf{R}^2_2\mathbf{H}^{*}_{22} + o_p(1), \qquad [6.10]\end{aligned}$$

and **6.9** is

$$\mathbf{H}^{*\prime}_{22}\boldsymbol{\Theta}^{\frac{1}{2}} + \boldsymbol{\Theta}^{\frac{1}{2}}\mathbf{H}^{*\prime}_{22} = -\boldsymbol{\Theta}^{-\frac{1}{2}}\mathbf{S}^{*22}_{\bar{L}\bar{L}}\boldsymbol{\Theta}^{-\frac{1}{2}} + o_p(1). \qquad [6.11]$$

Define the left-hand side of **6.10** as $\boldsymbol{\Theta}^{-1}\mathbf{P}\boldsymbol{\Theta}^{-\frac{1}{2}}$. To find the limiting distribution of $\widehat{\mathbf{R}}^2_2$ and $\mathbf{H}^{*}_{22}$, we use the limiting distribution of $\mathbf{P}$. To describe the limiting distributions of matrices, we use the vec notation: $\mathrm{vec}(\mathbf{a}_1, \ldots, \mathbf{a}_m) = (\mathbf{a}'_1, \ldots, \mathbf{a}'_m)'$ and the relation $\mathrm{vec}\,\mathbf{ABC} = (\mathbf{C}' \otimes \mathbf{A})\,\mathrm{vec}\,\mathbf{B}$, where $\otimes$ denotes the Kronecker product. $\mathbf{K}$ is the permutation matrix such that $\mathbf{K}\,\mathrm{vec}\,\mathbf{A} = \mathrm{vec}\,\mathbf{A}'$.

THEOREM 2. *If the $\mathbf{V}_t$ values are independently normally distributed, the limiting distribution of $\mathbf{S}^{*2\cdot1,2\cdot1}_{VV} = \sqrt{T}(\mathbf{S}^{2\cdot1,2\cdot1}_{VV} - \mathbf{I})$, $\mathbf{S}^{*2\cdot1,2}_{V\bar{U}} = \sqrt{T}\,\mathbf{S}^{2\cdot1,2}_{V\bar{U}}$, and $\mathbf{S}^{*22}_{\bar{U}\bar{U}} = \sqrt{T}(\mathbf{S}^{22}_{\bar{U}\bar{U}} - \boldsymbol{\Sigma}^{22}_{\bar{U}\bar{U}})$ is normal with means* $\mathbf{0}$, $\mathbf{0}$, *and* $\mathbf{0}$ *and covariances*

$$\begin{aligned}
&\mathcal{E}\operatorname{vec}\mathbf{S}^{*22}_{\bar{U}\bar{U}}\big(\operatorname{vec}\mathbf{S}^{*22}_{\bar{U}\bar{U}}\big)' \\
&\quad\to [\mathbf{I}-(\boldsymbol{\Delta}_2\otimes\boldsymbol{\Delta}_2)]^{-1}(\mathbf{I}+\mathbf{K}) \\
&\quad\cdot\big[(\boldsymbol{\Sigma}^{22}_{\bar{U}\bar{U}}\otimes\boldsymbol{\Sigma}^{22}_{VV})+(\boldsymbol{\Sigma}^{22}_{VV}\otimes\boldsymbol{\Sigma}^{22}_{\bar{U}\bar{U}})-(\boldsymbol{\Sigma}^{22}_{VV}\otimes\boldsymbol{\Sigma}^{22}_{VV})\big] \\
&\quad\cdot\big[\mathbf{I}-(\boldsymbol{\Delta}'_2\otimes\boldsymbol{\Delta}'_2)\big]^{-1}, \qquad \textbf{[6.12]} \\
&\mathcal{E}\operatorname{vec}\mathbf{S}^{*2\cdot1,2}_{V\bar{U}}\big(\operatorname{vec}\mathbf{S}^{*2\cdot1,2}_{V\bar{U}}\big)' = \boldsymbol{\Sigma}^{22}_{\bar{U}\bar{U}}\otimes\mathbf{I}, \\
&\mathcal{E}\operatorname{vec}\mathbf{S}^{*2\cdot1,2\cdot1}_{VV}\big(\operatorname{vec}\mathbf{S}^{*2\cdot1,2\cdot1}_{VV}\big)' = (\mathbf{I}+\mathbf{K})(\mathbf{I}\otimes\mathbf{I}), \\
&\mathcal{E}\operatorname{vec}\mathbf{S}^{*22}_{\bar{U}\bar{U}}\big(\operatorname{vec}\mathbf{S}^{*2\cdot1,2}_{V\bar{U}}\big)' \\
&\quad\to [\mathbf{I}-(\boldsymbol{\Delta}_2\otimes\boldsymbol{\Delta}_2)]^{-1}(\mathbf{I}+\mathbf{K})(\boldsymbol{\Delta}_2\boldsymbol{\Sigma}^{22}_{\bar{U}\bar{U}}\otimes\mathbf{I}), \\
&\mathcal{E}\operatorname{vec}\mathbf{S}^{*22}_{\bar{U}\bar{U}}\big(\operatorname{vec}\mathbf{S}^{*2\cdot1,2\cdot1}_{VV}\big)' \to [\mathbf{I}-(\boldsymbol{\Delta}_2\otimes\boldsymbol{\Delta}_2)]^{-1}(\mathbf{I}+\mathbf{K})(\mathbf{I}\otimes\mathbf{I}), \\
&\mathcal{E}\operatorname{vec}\mathbf{S}^{*2\cdot1,2}_{V\bar{U}}\big(\operatorname{vec}\mathbf{S}^{*2\cdot1,2\cdot1}_{VV}\big)' = \mathbf{0}, \\
&\mathcal{E}\operatorname{vec}\mathbf{S}^{*2\cdot1,2}_{V\bar{U}}\big(\operatorname{vec}\mathbf{S}^{*2,2\cdot1}_{\bar{U}V}\big)' = \mathbf{K}(\mathbf{I}\otimes\boldsymbol{\Sigma}^{22}_{\bar{U}\bar{U}}).
\end{aligned}$$

Proof: These asymptotic covariances are adaptations of theorem 1 from ref. 9 with $\mathcal{E}\mathbf{V}_{2\cdot1,t}\mathbf{V}'_{2t} = \mathcal{E}\big[\mathbf{V}_{2t}-\boldsymbol{\Sigma}^{21}_{VV}(\boldsymbol{\Sigma}^{11}_{VV})^{-1}\mathbf{V}_{1t}\big]\mathbf{V}'_{2t} = \boldsymbol{\Sigma}^{22\cdot1}_{VV} = \mathbf{I}$. The details of the proof are left to the reader. ■

Since $\mathbf{L}_{2,t-1} = \mathbf{M}_2\mathbf{U}_{2,t-1}$, $\mathbf{S}^{*2\cdot1,2}_{V\bar{L}} = \mathbf{S}^{*2\cdot1,2}_{V\bar{U}}\mathbf{M}'_2$, $\mathbf{S}^{*2,2\cdot1}_{\bar{L}V} = \mathbf{M}_2\mathbf{S}^{*2,2\cdot1}_{\bar{U}V}$, and $\mathbf{S}^{*22}_{\bar{L}\bar{L}} = \mathbf{M}_2\mathbf{S}^{*22}_{\bar{U}\bar{U}}\mathbf{M}'_2$, the vec of the left-hand side of **6.10** multiplied on the left by $\boldsymbol{\Theta}$ and the right by $\boldsymbol{\Theta}^{\frac{1}{2}}$ is

$$\begin{aligned}
\operatorname{vec}\mathbf{P} = {}&-(\mathbf{R}^2_2\otimes\mathbf{R}^2_2)\operatorname{vec}\mathbf{S}^{*2\cdot1,2\cdot1}_{VV} + [(\mathbf{I}-\mathbf{R}^2_2)\otimes\mathbf{R}^2_2]\operatorname{vec}\mathbf{S}^{*2\cdot1,2}_{V\bar{L}} \\
&+[\mathbf{R}^2_2\otimes(\mathbf{I}-\mathbf{R}^2_2)]\operatorname{vec}\mathbf{S}^{*2,2\cdot1}_{\bar{L}V} + [(\mathbf{I}-\mathbf{R}^2_2)\otimes\mathbf{R}^2_2]\operatorname{vec}\mathbf{S}^{*22}_{\bar{L}\bar{L}}.
\end{aligned}$$

Algebraic calculation shows that the covariance matrix of the limiting distribution of $\operatorname{vec}\mathbf{P}$ is

$$\begin{aligned}
&(\mathbf{I}+\mathbf{K})(\mathbf{R}^4_2\otimes\mathbf{R}^4_2) + [(\mathbf{I}-\mathbf{R}^2_2)\otimes\mathbf{R}^2_2](\mathbf{I}+\mathbf{K})\big[\mathbf{I}-(\boldsymbol{\Delta}_2\otimes\boldsymbol{\Delta}_2)\big]^{-1} \\
&\quad\cdot\big[(\boldsymbol{\Theta}\otimes\boldsymbol{\Theta})-(\boldsymbol{\Delta}_2\otimes\boldsymbol{\Delta}_2)(\boldsymbol{\Theta}\otimes\boldsymbol{\Theta})(\boldsymbol{\Delta}'_2\otimes\boldsymbol{\Delta}'_2)\big] \\
&\quad\cdot\big[\mathbf{I}-(\boldsymbol{\Delta}'_2\otimes\boldsymbol{\Delta}'_2)\big]^{-1}[(\mathbf{I}-\mathbf{R}^2_2)\otimes\mathbf{R}^2_2] \\
&= (\mathbf{I}+\mathbf{K})(\mathbf{R}^4_2\otimes\mathbf{R}^4_2) + [(\mathbf{I}-\mathbf{R}^2_2)\otimes\mathbf{R}^2_2](\mathbf{I}+\mathbf{K}) \\
&\quad\cdot\Big\{\big[\mathbf{I}-(\boldsymbol{\Delta}_2\otimes\boldsymbol{\Delta}_2)\big]^{-1}(\boldsymbol{\Theta}\otimes\boldsymbol{\Theta}) \\
&\qquad+(\boldsymbol{\Theta}\otimes\boldsymbol{\Theta})\big[\mathbf{I}-(\boldsymbol{\Delta}'_2\otimes\boldsymbol{\Delta}'_2)\big]^{-1}-(\boldsymbol{\Theta}\otimes\boldsymbol{\Theta})\Big\} \\
&\quad\cdot[(\mathbf{I}-\mathbf{R}^2_2)\otimes\mathbf{R}^2_2]. \qquad \textbf{[6.13]}
\end{aligned}$$

Each matrix in **6.13** is diagonal except $\mathbf{K}$ and $[\mathbf{I}-(\boldsymbol{\Delta}_2\otimes\boldsymbol{\Delta}_2)]^{-1}$, which we define as $\boldsymbol{\Phi} = [\mathbf{I}-(\boldsymbol{\Delta}_2\otimes\boldsymbol{\Delta}_2)]^{-1}$. Then the asymptotic covariance matrix of $\operatorname{vec}\mathbf{P}$ is

$$\begin{aligned}
&(\mathbf{I}+\mathbf{K})(\mathbf{R}^4_2\otimes\mathbf{R}^4_2) + [(\mathbf{I}-\mathbf{R}^2_2)\otimes\mathbf{R}^2_2](\mathbf{I}+\mathbf{K}) \\
&\quad\cdot\Big[\boldsymbol{\Phi}(\boldsymbol{\Theta}\otimes\boldsymbol{\Theta})+(\boldsymbol{\Theta}\otimes\boldsymbol{\Theta})\boldsymbol{\Phi}'-(\boldsymbol{\Theta}\otimes\boldsymbol{\Theta})\Big] \\
&\quad\cdot[(\mathbf{I}-\mathbf{R}^2_2)\otimes\mathbf{R}^2_2]. \qquad \textbf{[6.14]}
\end{aligned}$$

From **6.10**, we see that r^{2*}_i is the ith diagonal element of $\boldsymbol{\Theta}^{-\frac{1}{2}}\mathbf{P}\boldsymbol{\Theta}^{-\frac{1}{2}}$ and the i, ith element of $\operatorname{vec}(\boldsymbol{\Theta}^{-\frac{1}{2}}\mathbf{P}\boldsymbol{\Theta}^{-\frac{1}{2}}) = (\boldsymbol{\Theta}^{-\frac{1}{2}}\otimes\boldsymbol{\Theta}^{-\frac{1}{2}})\operatorname{vec}\mathbf{P}$. Hence the asymptotic covariance of r^{2*}_i and r^{2*}_j is $\theta^{-1}_i\theta^{-1}_j$ times the ii, jjth element of **6.14**, that is,

$$2\Big[(1-\rho^2_i)^2\Phi_{ii,jj}\rho^4_j + \rho^4_i\Phi_{jj,ii}(1-\rho^2_j)^2\Big]. \qquad \textbf{[6.15]}$$

We can put this into matrix form by use of $\tilde{\mathbf{E}} = \sum_{i=1}^k \boldsymbol{\epsilon}_i(\boldsymbol{\epsilon}'_i\otimes\boldsymbol{\epsilon}'_i)$, where $\boldsymbol{\epsilon}_i$ is the k-vector with 1 in the ith position and 0s elsewhere. The matrix $\tilde{\mathbf{E}}$ has 1 in the ith row and i, ith column, $i = 1, \ldots, k$, and 0s elsewhere. Define $\mathbf{r}^{2*} = (r^{2*}_{n+1}, \ldots, r^{2*}_p)'$.

THEOREM 3. *If the $\mathbf{V}_t$ vectors are independently normally distributed and the diagonal elements of $\boldsymbol{\Theta}$ are different, the limiting distribution of $\mathbf{r}^{2*}$ is normal with mean* $\mathbf{0}$ *and covariance matrix*

$$2(\mathbf{I}-\mathbf{R}^2_2)^2\tilde{\mathbf{E}}\boldsymbol{\Phi}\tilde{\mathbf{E}}'\mathbf{R}^4_2 + 2\mathbf{R}^4_2\tilde{\mathbf{E}}\boldsymbol{\Phi}'\tilde{\mathbf{E}}'(\mathbf{I}-\mathbf{R}^2_2)^2. \qquad \textbf{[6.16]}$$

Hansen and Johansen (10) have independently developed the asymptotic distribution of the larger roots; their method is different and the result is expressed differently, including a more general model. The development here has benefited by comparison with their work.

In the limiting distribution of $\mathbf{r}^{2*}$, the components of $\mathbf{r}^{2*}$ are correlated in contrast to the asymptotic distribution of the canonical correlations between a dependent vector $\mathbf{Y}_t$ and an independent vector $\mathbf{X}_t$. Although the distribution of the stationary part of $\mathbf{Y}_t$ (namely, $\mathbf{X}_{2t}$) depends directly on the lagged stationary part ($\mathbf{X}_{2,t-1}$), the limiting distribution in *Theorem 3* depends indirectly on the nonstationary part (through $\mathbf{W}_{1t}$).

From **6.10** and **6.11** we can derive the limiting distribution of $\mathbf{H}^*_{22}$. Let $\mathbf{H}^*_{22} = \mathbf{H}^*_d + \mathbf{H}^*_n$, where $\mathbf{H}^*_d = \operatorname{diag}(h^*_{n+1,n+1}, \ldots, h^*_{pp})$. Note that $\operatorname{vec}\mathbf{H}^*_{22}\mathbf{R}^2_2 = (\mathbf{R}^2_2\otimes\mathbf{I})\operatorname{vec}\mathbf{H}^*_{22}$ and $\mathbf{R}^2_2\mathbf{H}^*_{22} = (\mathbf{I}\otimes\mathbf{R}^2_2)\cdot\operatorname{vec}\mathbf{H}^*_{22}$. Then $\operatorname{vec}(\mathbf{H}^*_{22}\mathbf{R}^2_2 - \mathbf{R}^2_2\mathbf{H}^*_{22}) = \mathbf{N}\operatorname{vec}\mathbf{H}^*_{22} = \mathbf{N}\mathbf{H}^*_n$, where

$$\begin{aligned}
\mathbf{N} &= (\mathbf{R}^2_2\otimes\mathbf{I})-(\mathbf{I}\otimes\mathbf{R}^2_2) \\
&= \operatorname{diag}(0, \rho^2_{n+1}-\rho^2_{n+2}, \ldots, \rho^2_{n+1}-\rho^2_p, 0, \ldots, \rho^2_p-\rho^2_{p-1}, 0).
\end{aligned} \qquad \textbf{[6.17]}$$

The Moore–Penrose generalized inverse of $\mathbf{N}$ is $\mathbf{N}^+$, which has 0 in the i, ith position, $i = 1, \ldots, k$, and $1/(\rho^2_j - \rho^2_i)$ as a typical diagonal element. Note $\tilde{\mathbf{E}}\mathbf{N} = \mathbf{0}$ and $\mathbf{N}\mathbf{N}^+ = (\mathbf{I}\otimes\mathbf{I}) - \tilde{\mathbf{E}}'\tilde{\mathbf{E}}$. From **6.8**, we obtain $\operatorname{vec}\mathbf{H}^*_n = \mathbf{N}^+(\boldsymbol{\Theta}^{-\frac{1}{2}}\otimes\boldsymbol{\Theta}^{-1})\operatorname{vec}\mathbf{P}$. Then

$$\begin{aligned}
\mathcal{E}\operatorname{vec}\mathbf{H}^*_n(\operatorname{vec}\mathbf{H}^*_n)' = {}&\mathbf{N}^+(\mathbf{I}+\mathbf{K})\big[(\mathbf{I}-\mathbf{R}^2_2)\mathbf{R}^2_2\otimes(\mathbf{I}-\mathbf{R}^2_2)^2\big]\mathbf{N}^+ \\
&+\mathbf{N}^+\big[\mathbf{R}^{\frac{1}{2}}_2(\mathbf{I}-\mathbf{R}^2_2)^{\frac{3}{2}}\otimes(\mathbf{I}-\mathbf{R}^2_2)\big](\mathbf{I}+\mathbf{K}) \\
&\big[\boldsymbol{\Phi}(\boldsymbol{\Theta}\otimes\boldsymbol{\Theta})+(\boldsymbol{\Theta}\otimes\boldsymbol{\Theta})\boldsymbol{\Phi}'-(\boldsymbol{\Theta}\otimes\boldsymbol{\Theta})\big] \\
&\big[\mathbf{R}^{\frac{1}{2}}_2(\mathbf{I}-\mathbf{R}^2_2)^{\frac{3}{2}}\otimes(\mathbf{I}-\mathbf{R}^2_2)\big]\mathbf{N}^+. \qquad \textbf{[6.18]}
\end{aligned}$$

From **6.11**, we obtain $\mathbf{H}^*_d = -\frac{1}{2}\boldsymbol{\Theta}^{-\frac{3}{2}}\operatorname{diag}(\mathbf{S}^{*22}_{\bar{L}\bar{L}}) + o_p(1)$, where $\operatorname{diag}(\mathbf{S}^{*22}_{\bar{L}\bar{L}})$ is the diagonal part of $\mathbf{S}^{*22}_{\bar{L}\bar{L}} = \mathbf{M}'_2\mathbf{S}^{*22}_{\bar{U}\bar{U}}\mathbf{M}_2$. The asymptotic covariance of $\operatorname{vec}\mathbf{S}^{*22}_{\bar{U}\bar{U}}$ is given in **6.12**.

Since $\mathbf{G}_{22} = \mathbf{Y}'_2\boldsymbol{\Xi}\mathbf{H}_{22}$, the asymptotic covariance of $\operatorname{vec}\mathbf{G}_{22}$ can be found from the asymptotic covariance of $\mathbf{H}_{22}$.

7. Higher-Order Processes

Now we consider an autoregressive process of order m

$$\mathbf{Y}_t = \sum_{j=1}^m \mathbf{B}_j\mathbf{Y}_{t-j} + \mathbf{Z}_t. \qquad \textbf{[7.1]}$$

If the roots $\lambda_1, \ldots, \lambda_{pm}$ of

$$|\lambda^m\mathbf{I} - \lambda^{m-1}\mathbf{B}_1 - \cdots - \mathbf{B}_m| = 0 \qquad \textbf{[7.2]}$$

satisfy $|\lambda_j| < 1$, $j = 1, \ldots, pm$, **7.1** defines a stationary process. In a cointegrated process $\lambda_j = 1$ for one or more values of j.

Suppose $\{\mathbf{Y}_t\} \in I(1)$ and $\{\Delta\mathbf{Y}_t\} \in I(0)$. The model can be put in a error-correction form

$$\Delta\mathbf{Y}_t = \mathbf{\Pi}\mathbf{Y}_{t-1} + \sum_{j=1}^{m-1} \mathbf{\Pi}_j \Delta\mathbf{Y}_{t-j} + \mathbf{Z}_t\,, \qquad [7.3]$$

where $\mathbf{\Pi} = \sum_{j=1}^{m} \mathbf{B}_j - \mathbf{I}$ and $\mathbf{\Pi}_i = -\sum_{j=i+1}^{m} \mathbf{B}_j,\; i = 1, \ldots, m-1$. Note that $\mathbf{\Pi}$ is the matrix in **7.2** for $\lambda = 1$. Suppose that the multiplicity of the root $\lambda = 1$ is n, that there are n linearly independent solutions of $\boldsymbol{\omega}'\mathbf{\Pi} = \mathbf{0}$, say $\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_n$, and that the other $pm - n$ roots of **7.2** satisfy $|\lambda_i| < 1$. Define $\mathbf{\Omega}_1 = (\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_n)$ and let $\mathbf{\Omega}_2$ be a $p \times k$ matrix, where $k = p - n$, such that $\mathbf{\Omega}_2'\mathbf{\Pi} = \mathbf{\Upsilon}_2\mathbf{\Omega}_2'$. Then $\Delta\mathbf{Y}_t$ and $\mathbf{\Omega}_2'\mathbf{Y}_t$ can be given initial distributions such that they are stationary.

Define $\mathbf{X}_t$ and $\mathbf{W}_t$ by **2.8**, $\mathbf{\Upsilon}$ by $\mathbf{\Omega}'\mathbf{\Pi}(\mathbf{\Omega}')^{-1}$ and $\mathbf{\Upsilon}_i$ by $\mathbf{\Omega}'\mathbf{\Pi}_i(\mathbf{\Omega}')^{-1}$. Then the model **7.3** is transformed to

$$\Delta\mathbf{X}_t = \mathbf{\Upsilon}\mathbf{X}_{t-1} + \bar{\mathbf{\Upsilon}}\Delta\bar{\mathbf{X}}_{t-1} + \mathbf{W}_t\,, \qquad [7.4]$$

where $\bar{\mathbf{\Upsilon}} = (\mathbf{\Upsilon}_1, \ldots, \mathbf{\Upsilon}_{m-1})$ and $\Delta\bar{\mathbf{X}}_{t-1} = (\Delta\mathbf{X}_{t-1}', \ldots, \Delta\mathbf{X}_{t-m+1}')'$. The reduced rank regression estimator for the form **7.4** was found by Anderson (3). By writing **7.4** in terms of $\Delta\bar{\mathbf{X}}_{t-1}$ and the residual of $\mathbf{X}_{t-1}$ regressed on $\Delta\bar{\mathbf{X}}_{t-1}$, the estimator, eigenvalues, eigenvectors, and their asymptotic distributions follow from the preceeding developments.

I am indebted to Naoto Kunitomo and Soren Johansen for assistance in preparing this paper.

1. Tiao, G. C., & Tsay, R. S. (1989) *J. R. Stat. Soc.* **51,** 157–213.
2. Box, G. E. P. & Jenkins, G. M. (1971) *Time Series: Forecasting and Control* (Holden-Day, San Francisco).
3. Anderson, T. W. (1951) *Ann. Math. Stat.* **22,** 327–351.
4. Johansen, S. (1995) *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models* (Oxford Univ. Press, Oxford).
5. Engle, R. F. & Granger, C. W. J. (1987) *Econometrica* **55,** 251–276.
6. Granger, C. W. J. (1981) *J. Econ.* **16**, 121–130.
7. Billingsley, P. (1968) *Convergence of Probability Measures* (Wiley, New York).
8. Johansen, S. (1988) *J. Econ. Dyn. Control* **12**, 231–254.
9. Anderson, T. W. (2000) *Ann. Statist.*, in press.
10. Hansen, H. & Johansen, S. (2000) *Econ. J.*, in press.