A STUDY OF ASYMPTOTIC DISTRIBUTIONS OF CANONICAL
CORRELATIONS AND VECTORS IN HIGHER-ORDER
COINTEGRATED MODELS

by

T. W. ANDERSON

Technical Report No. 2001-3
February 2001

Department of Statistics
STANFORD UNIVERSITY
Stanford, California
A STUDY OF ASYMPTOTIC DISTRIBUTIONS OF CANONICAL CORRELATIONS AND VECTORS IN HIGHER-ORDER COINTEGRATED MODELS

by

T. W. ANDERSON
Department of Statistics
Stanford University

Technical Report No. 2001-3
FEBRUARY 2001

Department of Statistics
Sequoia Hall
STANFORD UNIVERSITY
Stanford, California 94305-4065

http://www-stat.stanford.edu
A Study of Asymptotic Distributions of Canonical Correlations and Vectors in Higher-Order Cointegrated Models

T. W. Anderson

Abstract

The study of the large-sample distribution of the canonical correlations and variates in cointegrated models is extended from the first-order autoregression model [Anderson (2000a)] to autoregression of any (finite) order. The cointegrated process considered here is nonstationary in some dimensions and stationary in some other directions, but the first difference (the "error-correction form") is stationary. The asymptotic distribution of the canonical correlations between the first differences and the predictor variables as well as the corresponding canonical variables is obtained under the assumption that the process is Gaussian. The method of analysis is similar to that used for the first-order process.

1 Introduction

Cointegrated stochastic processes are used in econometrics for modelling macroeconomic time series that have both stationary and nonstationary properties. The term "cointegrated" means that in a multivariate process that appears nonstationary some linear functions are stationary. Many economic time series may show inflationary tendencies or increasing volatility, but certain relationships are not affected by these tendencies. Statistical inference is involved in identifying these relationships and estimating their importance.

The family of stochastic processes studied in this paper consists of vector autoregressive processes of finite order. A vector of contemporary measures is considered to depend linearly on earlier
values of these measures plus random disturbances or errors. The dependence may be evaluated
by the canonical correlations between the contemporary values and the earlier values.

The nonstationarity of a process may be eliminated by treating differences or higher-order
differences (over time) of the vectors. This paper treats processes in which first-order differencing
accomplishes stationarity. The first-order difference is represented as a linear combination of the
first lagged variable and lags of the difference variable. The stationary linear combinations are
the canonical variables corresponding to the nonzero process canonical correlations between the
difference variable and the first lagged variable not accounted for by the lagged differences. The
number of these is defined as the degree of cointegration.

Statistical inference of the model is based on a sample of observations; that is, a vector time
series over some period of time. The estimator of the parameters of the original autoregressive
model is a transformation of the estimator of the (stationary) error-correction form. In the latter
one coefficient matrix is of lower rank (the degree of cointegration). It is estimated efficiently by
the reduced rank regression estimator introduced by Anderson (1951). It depends on the larger
canonical correlations and corresponding canonical vectors. The smaller correlations are used to
determine the rank of this matrix. Inference is based on the large-sample distribution of these
correlations and variables.

The asymptotic distribution of the canonical correlations and coefficients of the variates for
the first-order autoregressive process was derived by Anderson (2000a). The distribution for the
higher-order process (that is, several lags) is obtained in this paper, using similar algebra. Hansen
and Johansen (2000) have independently obtained the asymptotic distribution of the canonical
correlations, but by a different method and expressed in a different form.

The likelihood ratio test for the degree of cointegration found by Anderson (1951) is given in
Section 5; its asymptotic distribution under the null hypothesis was found by Johansen (1988). To evaluate the power of such a test one needs to know the distribution or asymptotic distribution of the sample canonical correlations corresponding to process canonical correlations different from 0. See Anderson (2000c), for example.

For further background the reader is referred to Johansen (1995) and Reinsel and Velu (1998).

2 The model

The general cointegrated model is an autoregressive process \( \{Y_t\} \) of order \( m \) defined by

\[
Y_t = B_1 Y_{t-1} + \ldots + B_m Y_{t-m} + Z_t ,
\]

where \( Z_t \) is unobserved with \( EZ_t = 0, EZ_t Z_t' = \Sigma ZZ', \) and \( EY_{t-i} Z_t' = 0, i = 1, \ldots \). Let

\[
B(\lambda) = \lambda^m I - \lambda^{m-1} B_1 - \ldots - B_m .
\]

If the roots \( \lambda_1, \ldots, \lambda_{pm} \) of \( |B(\lambda)| = 0 \) satisfy \( |\lambda_i| < 1 \), a stationary process \( \{Y_t\} \) can be defined by (2.1). If some of the roots are 1, the process will be nonstationary. In this paper we assume that \( n \) (0 < \( n < p \)) roots of \( |B(\lambda)| = 0 \) are 1 (\( \lambda_1 = \ldots = \lambda_p = 1 \)) and the other \( pm - n \) roots satisfy \( |\lambda_i| < 1, i = n + 1, \ldots, pm \). The first difference of the process (the “error-correction” form) is

\[
Y_t - Y_{t-1} = \Delta Y_t = \Pi Y_{t-1} + \Pi_1 \Delta Y_{t-1} + \ldots + \Pi_{m-1} \Delta Y_{t-m+1} + Z_t
\]

\[
= \Pi Y_{t-1} + (\Pi_1, \ldots, \Pi_{m-1})(\Delta Y_{t-1}', \ldots, \Delta Y_{t-m+1}')' + Z_t
\]

\[
= \Pi Y_{t-1} + \Pi \Delta Y_{t-1} + Z_t .
\]
Here $\Pi = B_1 + \ldots + B_m - I = -B(1)$, $\Pi_j = -(B_{j+1} + \ldots + B_m)$, $j = 1, \ldots, m - 1$ [Engle and Granger (1987)], $\bar{\Pi} = (\Pi_1, \ldots, \Pi_{m-1})$, and $\bar{\Delta}Y_{t-1} = (\Delta Y_{t-1}', \ldots, \Delta Y_{t-m+1}')'$.

A sample consists of $T$ observations: $Y_1, \ldots, Y_T$. Since the rank of $\Pi$ is $k$, it is to be estimated by the reduced rank regression estimator introduced by Anderson (1951) as the maximum likelihood estimator when $Z_1, \ldots, Z_T$ are normally distributed and $Y_0, Y_1, \ldots, Y_{m+1}$ are nonstochastic and known. The matrices $\Pi_1, \ldots, \Pi_{m-1}$ are unrestricted except for the condition $|\lambda_i| < 1$, $i = n + 1, \ldots, pm$. The estimator depends on the canonical correlations and vectors of $\Delta Y_t$ and $Y_{t-1}$ conditioned on $\Delta Y_{t-1}, \ldots, \Delta Y_{t-m+1}$.

Define

$$\Delta\hat{Y}_t^+ = \Delta Y_t - S_{\Delta Y, \Delta Y}^{-1} \bar{\Delta}Y_{t-1},$$

$$\hat{Y}_{t-1}^+ = Y_{t-1} - S_{\bar{Y}, \Delta Y}^{-1} \bar{\Delta}Y_{t-1}.$$  

where $\bar{\Delta}Y_{t-1} = (\Delta Y_{t-1}', \ldots, \Delta Y_{t-m+1}')'$,

$$S_{\Delta Y, \Delta Y} = \frac{1}{T} \sum_{t=1}^{T} \bar{\Delta}Y_{t-1} \bar{\Delta}Y_{t-1}' ,$$

$$S_{\Delta Y, \bar{Y}} = \frac{1}{T} \sum_{t=1}^{T} \Delta Y_t \bar{\Delta}Y_{t-1}' ,$$

$$S_{\bar{Y}, \Delta Y} = \frac{1}{T} \sum_{t=1}^{T} Y_{t-1} \bar{\Delta}Y_{t-1}'.$$

The vectors $\Delta\hat{Y}_t^+$ and $\hat{Y}_{t-1}^+$ are the sample residuals of $\Delta Y_{t-1}$ and $Y_{t-1}$ regressed on $\bar{\Delta}Y_{t-1}$. 

4
Define

\[
\hat{S}_{\Delta Y,\Delta Y}^+ = \frac{1}{T} \sum_{t=1}^{T} \Delta \hat{Y}_t^+ \Delta \hat{Y}_{t'}^+ = S_{\Delta Y,\Delta Y} - S_{\Delta Y,\Delta Y} S_{\Delta Y,\Delta Y}^{-1} S_{\Delta Y,\Delta Y} , \quad (2.9)
\]

\[
\hat{S}_{\Delta Y,\hat{Y}}^+ = \frac{1}{T} \sum_{t=1}^{T} \Delta \hat{Y}_t^+ \hat{Y}_{t-1}^+ = S_{\Delta Y,\hat{Y}} - S_{\Delta Y,\Delta Y} S_{\Delta Y,\Delta Y}^{-1} S_{\Delta Y,\hat{Y}} , \quad (2.10)
\]

\[
\hat{S}_{\hat{Y},\hat{Y}}^+ = \frac{1}{T} \sum_{t=1}^{T} \hat{Y}_{t-1}^+ \hat{Y}_{t-1}^+ = S_{\hat{Y},\hat{Y}} - S_{\hat{Y},\Delta Y} S_{\Delta Y,\Delta Y}^{-1} S_{\Delta Y,\hat{Y}} , \quad (2.11)
\]

where

\[
\begin{bmatrix}
S_{\Delta Y,\Delta Y} & S_{\Delta Y,\hat{Y}} \\
S_{\hat{Y},\Delta Y} & S_{\hat{Y},\hat{Y}}
\end{bmatrix}
= \frac{1}{T} \sum_{t=1}^{T} \begin{bmatrix}
\Delta Y_t \\
Y_{t-1}
\end{bmatrix} (\Delta Y_t', Y_{t-1}') .
\quad (2.12)
\]

The sample canonical correlations between \(\Delta \hat{Y}_t^+\) and \(\hat{Y}_{t-1}^+\) are defined by

\[
|\hat{S}_{\hat{Y},\Delta Y}^+ \hat{S}_{\Delta Y,\Delta Y}^{+1} \hat{S}_{\Delta Y,\hat{Y}}^+ - r^2 \hat{S}_{\hat{Y},\hat{Y}}^+| = 0 ,
\quad (2.13)
\]

and the sample canonical variates by

\[
\hat{S}_{\hat{Y},\Delta Y}^+ \hat{S}_{\Delta Y,\Delta Y}^{+1} \hat{S}_{\Delta Y,\hat{Y}}^+ \hat{\gamma} = r^2 \hat{S}_{\hat{Y},\hat{Y}}^+ \hat{\gamma} , \quad \hat{\gamma} \hat{S}_{\hat{Y},\hat{Y}}^+ \hat{\gamma} = 1 .
\quad (2.14)
\]

More information on canonical analysis is covered in Chapter 12 of Anderson (1984).

One form of the reduced rank regression estimator is

\[
\hat{\Pi}_{(k)} = \hat{S}_{\Delta Y,\hat{Y}}^+ \hat{\Pi}_2 \hat{\gamma}_2 ,
\quad (2.15)
\]

where \(\hat{\Pi}_2 = (\hat{\gamma}_{n+1}, \ldots, \hat{\gamma}_p)\) and \(r^2_1 < \ldots < r^2_p\).
We shall assume that there are exactly \( n \) linearly independent solutions to \( \omega' B(1) = 0 \); that is, \( \omega' \Pi = 0 \). Then the rank of \( \Pi \) is \( p - n = k \) and there exists a \( p \times n \) matrix \( \Omega_1 \) of rank \( n \) such that \( \Omega_1' \Pi = 0 \). See Anderson (2000b). There is also a \( p \times k \) matrix \( \Omega_2 \) of rank \( k \) such that \( \Omega_2' \Pi = \Upsilon_2 \Omega_2' \), where \( \Upsilon_2 \) (\( k \times k \)) is nonsingular, and \( \Omega = (\Omega_1, \Omega_2) \) is nonsingular.

In order to distinguish between the stationary and nonstationary coordinates we make a transformation of coordinates. Define

\[
\Omega' Y_t = X_t = \begin{bmatrix} X_{1t} \\ X_{2t} \end{bmatrix}, \quad \Omega' Z_t = W_t = \begin{bmatrix} W_{1t} \\ W_{2t} \end{bmatrix}, \quad (2.16)
\]

\[
\Psi_j = \Omega' B_j (\Omega')^{-1}, \quad j = 1, \ldots, m. \quad (2.17)
\]

The process (2.1) is transformed to

\[
X_t = \Psi_1 X_{t-1} + \ldots + \Psi_m X_{t-m} + W_t. \quad (2.18)
\]

If we define \( Y = \Psi_1 + \ldots + \Psi_m - I = \Omega' \Pi (\Omega')^{-1} \), \( \Upsilon_j = - \sum_{i=1}^{m} \Psi_j = \Omega' \Pi_j (\Omega')^{-1} \), \( \Upsilon = (\Upsilon_1, \ldots, \Upsilon_{m-1}) \), and \( \Delta X_{t-1} = (\Delta X'_{t-1}, \ldots, \Delta X'_{t-m+1})' \), the form (2.3) is transformed to

\[
\Delta X_t = Y X_{t-1} + \Upsilon_1 \Delta X_{t-1} + \ldots + \Upsilon_{m-1} \Delta X_{t-m+1} + W_t
\]

\[
= Y X_{t-1} + (\Upsilon_1, \ldots, \Upsilon_{m-1})(\Delta X'_{t-1}, \ldots, \Delta X'_{t-m+1})' + W_t
\]

\[
= Y X_{t-1} + \Upsilon \Delta X_{t-1} + W_t. \quad (2.19)
\]
Note that

\[
\Upsilon = \begin{bmatrix}
0 & 0 \\
0 & \Upsilon^{22}
\end{bmatrix}.
\]  

(2.20)

Define \( \Delta \hat{X}_t^+, \Delta \hat{X}_{t-1}^+ \), \( S_{\Delta X, \Delta X}, S_{\Delta X, \Delta X}, \hat{S}_{\Delta X, \Delta X}^+, \hat{S}_{\Delta X, \Delta X}^+ \), and \( \hat{S}_{\Delta X, \Delta X}^+ \) in a manner analogous to (2.4) to (2.12). The reduced rank regression estimator of \( \Upsilon \) is based on the canonical correlations and canonical variates between \( \Delta \hat{X}_t^+ \) and \( \Delta \hat{X}_{t-1}^+ \) defined by

\[
|\hat{S}_{\Delta X, \Delta X}^+ \hat{S}_{\Delta X, \Delta X}^{+ -1} \hat{S}_{\Delta X, \Delta X}^+ - r^2 \hat{S}_{\Delta X, \Delta X}^+ | = 0, \tag{2.21}
\]

\[
\hat{S}_{\Delta X, \Delta X}^+ \hat{S}_{\Delta X, \Delta X}^{+ -1} \hat{S}_{\Delta X, \Delta X}^+ \mathbf{g} = r^2 \hat{S}_{\Delta X, \Delta X}^+ \mathbf{g}, \quad \mathbf{g}' \hat{S}_{\Delta X, \Delta X}^+ \mathbf{g} = 1. \tag{2.22}
\]

The estimator of \( \Upsilon \) of rank \( k \) is

\[
\hat{\Upsilon}_{(k)} = \hat{S}_{\Delta X, \Delta X}^+ \mathbf{G}_2 \mathbf{G}_2', \tag{2.23}
\]

where \( \mathbf{G}_2 = (g_{n+1}, \ldots, g_p) \) and \( g_i \) is the solution for \( g \) in (2.22) when \( r = r_i \), the solution to (2.21) and \( r_1 < \ldots < r_p \). The rest of this paper is devoted to finding the asymptotic distribution of \( \{g_i, r_i\} \). Note that \( \hat{\Upsilon}_{(k)} = \Omega' \hat{\Pi}_{(k)} (\Omega')^{-1} \).

The vectors \( \Delta \hat{X}_t^+ = \Delta X_t - S_{\Delta X, \Delta X} S_{\Delta X, \Delta X}^{-1} \Delta X_{t-1} \) and \( \Delta \hat{X}_{t-1}^+ = \Delta X_{t-1} - S_{\Delta X, \Delta X} S_{\Delta X, \Delta X}^{-1} \Delta X_{t-1} \) are the residuals of \( \Delta X_t \) and \( \Delta X_{t-1} \) regressed on \( \Delta X_{t-1} \), and \( r_1 \) is the maximum correlation between \( \Delta \hat{X}_t^+ \) and \( \Delta \hat{X}_{t-1}^+ \), which is the correlation between \( \Delta X_t \) and \( \Delta X_{t-1} \) after taking account of the dependence “explained” by \( \Delta X_{t-1} \). The canonical correlations are the canonical correlations between \( (\Delta X_t, \Delta X_{t-1}) \) and \( (\Delta X_{t-1}, \Delta X_{t-1}) \) other than ±1.
3 The Process

3.1 Markov model

The process \( \{X_t\} \) defined by (2.18) can be put in the form of the Markov model

\[
\begin{bmatrix}
X_t \\
X_{t-1} \\
\vdots \\
X_{t-m+1}
\end{bmatrix} =
\begin{bmatrix}
\Psi_1 & \Psi_2 & \ldots & \Psi_{m-1} & \Psi_m \\
I & 0 & \ldots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \ldots & I & 0
\end{bmatrix}
\begin{bmatrix}
X_{t-1} \\
X_{t-2} \\
\vdots \\
X_{t-m}
\end{bmatrix}
+ \begin{bmatrix}
W_t
\end{bmatrix}
\]  
(3.1)

[Sec 5.4, Anderson (1971)]. Multiplication of (3.1) on the left by

\[
\begin{bmatrix}
I & 0 & 0 & \ldots & 0 \\
I & -I & 0 & \ldots & 0 \\
0 & I & -I & \ldots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \ldots & -I
\end{bmatrix}
\]  
(3.2)

yields a form that includes the error-correction form (2.19)

\[
\begin{bmatrix}
X_t \\
\Delta X_t \\
\Delta X_{t-1} \\
\Delta X_{t-m+2}
\end{bmatrix} =
\begin{bmatrix}
\Upsilon + I & \Upsilon_1 & \Upsilon_2 & \ldots & \Upsilon_{m-2} & \Upsilon_{m-1} \\
\Upsilon & \Upsilon_1 & \Upsilon_2 & \ldots & \Upsilon_{m-2} & \Upsilon_{m-1} \\
0 & I & 0 & \ldots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \ldots & I & 0
\end{bmatrix}
\begin{bmatrix}
X_{t-1} \\
\Delta X_{t-1} \\
\Delta X_{t-2} \\
\Delta X_{t-m+1}
\end{bmatrix}
+ \begin{bmatrix}
W_t \\
\Delta X_{t-1} \\
\Delta X_{t-2} \\
\Delta X_{t-m+1}
\end{bmatrix}
\]  
(3.3)
The calculation of the coefficient matrix involves the inverse of (3.2), which is

$$
\begin{bmatrix}
1 & 0 & 0 & \ldots & 0 \\
I & -I & 0 & \ldots & 0 \\
I & -I & -I & \ldots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
I & -I & -I & \ldots & -I
\end{bmatrix}
$$

(3.4)

The first $n$ components of (3.3) constitute

$$
X_{1t} - X_{1,t-1} = \sum_{j=1}^{m-1} [\gamma_j^{11}(X_{1,t-j} - X_{1,t-j-1}) + \gamma_j^{12}(X_{2,t-j} - X_{2,t-j-1})] + W_{1t}.
$$

(3.5)

Here $\gamma_j$ has been partitioned into $n$ and $k$ rows and columns. Assume $X_{10} = X_{1,-1} = \ldots = 0$ and $W_{10} = W_{1,-1} = \ldots = 0$. The sum of (3.5) for $t = -\infty$ to $t = s$ is

$$
X_{1s} = \sum_{j=1}^{m-1} [\gamma_j^{11}X_{1,s-j} + \gamma_j^{12}X_{2,s-j}] + \sum_{t=1}^{s} W_{1t}
$$

$$
= \sum_{j=1}^{m-1} \{\gamma_j^{11}[X_{1s} - (X_{1s} - X_{1,s-j})] + \gamma_j^{12}[X_{2,s-1} - (X_{2,s-1} - X_{2,s-j})]\} + \sum_{t=1}^{s} W_{1t},
$$

(3.6)

or

$$
\left( I - \sum_{j=1}^{m-1} \gamma_j^{11} \right) X_{1s} = \sum_{j=1}^{m-1} \left( \gamma_j^{12}X_{2,s-1} - \gamma_j^{11} \sum_{i=1}^{j} \Delta X_{1,s-i} - \gamma_j^{12} \sum_{i=2}^{j} \Delta X_{2,s-i} \right) + \sum_{t=1}^{s} W_{1t}.
$$

(3.7)
Write (3.7) as
\[
X_{1t} = \Gamma \sum_{t=1}^{s} W_{1t} + H\tilde{X}_{s-1},
\]
(3.8)
where
\[
\Gamma = \left( I - \sum_{j=1}^{n-1} \gamma_{j}^{11} \right)^{-1},
\]
(3.9)
\(\Gamma^{-1}H\) is a linear combination of \(\gamma_{1}, \ldots, \gamma_{m-1}\), and \(\tilde{X}_{s} = (X_{2s}', \Delta X_{s}')'\). The right-hand side of (3.8) is the sum of a stationary process and a random walk (\(\sum_{t=1}^{s} W_{1t}\)). Note that \(I - \sum_{j=1}^{m-1} \gamma_{j}^{11}\) is nonsingular. If it were singular, there would be a vector \(\nu\) such that \(\nu' \left( I - \sum_{j=1}^{m-1} \gamma_{j}^{11} \right) = 0\) and \(\nu'\) times the right-hand side of (3.7) would be zero showing that \(\nu' \sum_{t=1}^{s} W_{1t} = 0\), which is impossible because \(W_{1t}\) has a positive definite covariance matrix.

The last \(pm - n = k + p(m - 1)\) components of (3.3) constitute a stationary process satisfying
\[
\begin{bmatrix}
X_{2t} \\
\Delta X_{1t} \\
\Delta X_{2t} \\
\Delta X_{1,t-1} \\
\Delta X_{2,t-1} \\
\vdots \\
\Delta X_{t-m+2}
\end{bmatrix} =
\begin{bmatrix}
\gamma_{22} + I & \gamma_{21}^{21} & \gamma_{22}^{22} & \cdots & 0 & 0 \\
0 & \gamma_{1}^{11} & \gamma_{1}^{12} & \cdots & 0 & 0 \\
\gamma_{22} & \gamma_{21}^{21} & \gamma_{22}^{22} & \cdots & 0 & 0 \\
0 & I & 0 & \cdots & 0 & 0 \\
0 & 0 & I & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & I & 0
\end{bmatrix}
\begin{bmatrix}
X_{2,t-1} \\
\Delta X_{1,t-1} \\
\Delta X_{2,t-1} \\
\Delta X_{1,t-2} \\
\Delta X_{2,t-2} \\
\vdots \\
\Delta X_{t-m+1}
\end{bmatrix} +
\begin{bmatrix}
W_{2t} \\
W_{1t} \\
W_{2t} \\
W_{1t} \\
W_{2t} \\
W_{1t} \\
W_{2t}
\end{bmatrix}
\]
(3.10)
Here \(\gamma_{j}^{1} = (\gamma_{j}^{11}, \gamma_{j}^{12})\) and \(\gamma_{j}^{2} = (\gamma_{j}^{21}, \gamma_{j}^{22})\). Write (3.10) as
\[
\tilde{X}_{t} = \tilde{Y} \tilde{X}_{t-1} + \tilde{W}_{t}.
\]
(3.11)
Since the eigenvalues of \( \tilde{\Psi} \) are less than 1 in absolute value [Anderson (2000b)],

\[
\tilde{X}_t = \sum_{s=0}^{\infty} \tilde{\Psi}^s \tilde{W}_{t-s} .
\]  

(3.12)

### 3.2 Second-order process moments

From (3.12) we find

\[
\epsilon \tilde{X}_t \tilde{X}_t' = \tilde{\Sigma} = \sum_{s=0}^{\infty} \tilde{\Psi}^s \tilde{\Sigma}_{WW} \tilde{\Psi}^s ,
\]

(3.13)

\[
\epsilon \tilde{X}_t \tilde{X}_{t-h} = \tilde{\Psi}^h \tilde{\Sigma} ,
\]

(3.14)

\[
\epsilon \tilde{W}_t \tilde{W}_t = \tilde{\Sigma}_{WW} =
\begin{bmatrix}
\Sigma_{WW}^{22} & \Sigma_{WW}^{21} & \Sigma_{WW}^{22} & 0 \\
\Sigma_{WW}^{12} & \Sigma_{WW}^{11} & \Sigma_{WW}^{12} & 0 \\
\Sigma_{WW}^{22} & \Sigma_{WW}^{21} & \Sigma_{WW}^{22} & 0 \\
0 & 0 & 0 & 0
\end{bmatrix}
\]

(3.15)

The covariance \( \tilde{\Sigma} \) satisfies

\[
\tilde{\Sigma} = \tilde{\Psi} \tilde{\Sigma} \tilde{\Psi}' + \tilde{\Sigma}_{WW} .
\]

(3.16)

Given \( \tilde{\Psi} \) and \( \tilde{\Sigma}_{WW} \), (3.16) can be solved for \( \tilde{\Sigma} \) [Anderson (1971), Sec. 5.5]. Further we write (3.8) as

\[
X_{1t} = \Gamma \sum_{s=0}^{t-1} W_{1,t-s} + H \sum_{s=0}^{\infty} \tilde{\Psi}^s \tilde{W}_{t-s} .
\]

(3.17)

Then

\[
\epsilon X_{1t} X_{1t}' = \epsilon \Gamma \Sigma_{WW}^{11} \Gamma' + \Gamma \Sigma_{WW}^{1-}(I - \tilde{\Psi}^{t+1})(I - \tilde{\Psi}')^{-1} H'
\]

11
\begin{align*}
+ H(I - \tilde{\mathbf{Y}})^{-1}(I - \tilde{\mathbf{Y}}^{t+1}) \Sigma_{WW}^{-1} \Gamma' + H \sum_{s=0}^{\infty} \tilde{\mathbf{Y}}^s \tilde{\mathbf{\Sigma}}_{WW} \tilde{\mathbf{Y}}^t H' \\
\sim t \Gamma \Sigma_{WW}^{11} \Gamma' + \Gamma \Sigma_{WW}^{11} (I - \tilde{\mathbf{Y}}^{-1}) H' + H(I - \tilde{\mathbf{Y}})^{-1} \Sigma_{WW}^{11} \Gamma' + H \tilde{\mathbf{\Sigma}} H' \\
(3.18)
\end{align*}

since \( I - \tilde{\mathbf{Y}}^{t+1} \rightarrow I \). Here \( \Sigma_{WW}^{11} \) denotes the second set of rows in (3.15). Then

\begin{equation}
\frac{1}{T} \mathcal{E} \bar{X} \bar{X}' = \frac{1}{T^2} \sum_{t=1}^{T} \mathcal{E} X_t X_t' \rightarrow \frac{1}{2} \Gamma \Sigma_{WW}^{11} \Gamma' \\
(3.19)
\end{equation}

because \( \sum_{t=1}^{T} t = T(T + 1)/2 \). Further

\begin{align*}
\mathcal{E} X_t X_t' &= \mathcal{E} \left( \Gamma \sum_{s=0}^{t} W_{1,t-s} + H \sum_{s=0}^{\infty} \tilde{\mathbf{Y}}^s \tilde{\mathbf{W}}_{t-s} \right) \sum_{r=0}^{\infty} \tilde{\mathbf{W}}_{t-r} \tilde{\mathbf{Y}}^r \\
&= \Gamma \Sigma_{WW}^{11} (I - \tilde{\mathbf{Y}}^{t+1})(I - \tilde{\mathbf{Y}}^{-1}) + H \tilde{\mathbf{\Sigma}} \\
&\rightarrow \Gamma \Sigma_{WW}^{11} (I - \tilde{\mathbf{Y}}^{-1}) + H \tilde{\mathbf{\Sigma}}. \\
(3.20)
\end{align*}

Define

\begin{align*}
\Delta X_t^+ &= \Delta X_t - \Sigma_{\Delta X, \Delta X} \Sigma_{\Delta X, \Delta X}^{-1} \Delta X_{t-1}, \\
X_{t-1}^+ &= X_{t-1} - \Sigma_{\tilde{X}, \Delta X} \Sigma_{\Delta X, \Delta X}^{-1} \Delta X_{t-1} \\
&= (I - \Sigma_{\tilde{X}, \Delta X} \Sigma_{\Delta X, \Delta X}^{-1}) \begin{bmatrix} X_{t-1} \\ \Delta X_{t-1} \end{bmatrix}, \\
(3.21)
\end{align*}

where \( \Sigma_{\Delta X, \Delta X} = \mathcal{E} \Delta X_t \Delta X_t' \), \( \Sigma_{\tilde{X}, \Delta X} = \mathcal{E} \Delta X_t \Delta X_t' \) depends on \( t \), and \( \Sigma_{\Delta X, \Delta X} = \mathcal{E} \Delta X_t \Delta X_t' \) does not depend on \( t \). Note that \( \Delta X_t^+ \) and \( \tilde{X}_{t-1}^+ \) correspond to \( \Delta \tilde{X}_t^+ \) and \( \tilde{X}_{t-1}^+ \) with \( \tilde{\mathbf{S}}_{\Delta X, \Delta X}, \tilde{\mathbf{S}}_{\tilde{X}, \Delta X} \) and \( \tilde{\mathbf{S}}_{\Delta X, \Delta X} \) replaced by \( \Sigma_{\Delta X, \Delta X}, \Sigma_{\tilde{X}, \Delta X} \) and \( \Sigma_{\Delta X, \Delta X} \), respectively. Then (2.19) can be written
as a regression model

\[ \Delta X_t^+ = \Sigma X_{t-1}^+ + W_t \]  

(3.23)

with \( \varepsilon X_{t-1}^+ W_t' = 0 \). Note that this model has the form of (2.10) in Anderson (2000a). In the \( X \)-coordinate system (3.23) is

\[ \Delta X_{1t}^+ = W_{1t}, \]

(3.24)

\[ \Delta X_{2t}^+ = \Sigma^{22} X_{2,t-1}^+ + W_{2t}. \]

(3.25)

The process analogs of (2.21) and (2.22) are

\[ |\Sigma_{\hat{\chi},\Delta \hat{\chi}} \Sigma^{+\cdot -1}_{\Delta \chi,\Delta \chi} \Sigma_{\Delta \chi,\hat{\chi}}^+ - \rho^2 \Sigma_{\hat{\chi},\hat{\chi}}^+ | = 0 \]

(3.26)

\[ \Sigma_{\hat{\chi},\Delta \chi} \Sigma^{+\cdot -1}_{\Delta \chi,\Delta \chi} \Sigma_{\Delta \chi,\hat{\chi}}^+ \gamma = \rho^2 \Sigma_{\hat{\chi},\hat{\chi}}^+ \gamma, \quad \gamma^T \Sigma_{\hat{\chi},\hat{\chi}}^+ \gamma = 1. \]

(3.27)

These define the process canonical correlations and variates in the \( X \)-coordinates.

\( \varepsilon \)From (3.21) and (3.22) we calculate

\[ \varepsilon \Delta X_t^+ \Delta X_t^{+'} = \begin{bmatrix} 0 & 0 \\ 0 & \Sigma_{\hat{\chi},\hat{\chi}}^{22} \end{bmatrix} + \Sigma_{WW} = \Sigma_{\Delta \Delta}^+, \]

(3.28)

\[ \varepsilon \Delta X_t^+ X_{t-1}^{+'} = \begin{bmatrix} 0 & 0 \\ \Sigma_{\hat{\chi},\hat{\chi}}^{22} \Sigma_{\hat{\chi},\hat{\chi}}^{+\cdot 21} & \Sigma_{\hat{\chi},\hat{\chi}}^{+\cdot 22} \end{bmatrix} = \Sigma_{\Delta \chi,\hat{\chi}}^+, \]

(3.29)

\[ \varepsilon X_{2,t-1}^+ X_{2,t-1}^{+'} = \Sigma_{\hat{\chi},\hat{\chi}}^{22} X_{\Delta \chi,\Delta \chi} - \Sigma_{\hat{\chi},\hat{\chi}}^{22} \Sigma_{\Delta \chi,\Delta \chi}^{-1} \Sigma_{\Delta \chi,\hat{\chi}}^+ \Sigma_{\hat{\chi},\hat{\chi}}^{22} X_{\Delta \chi,\hat{\chi}} = \Sigma_{\hat{\chi},\hat{\chi}}^{+\cdot 22} \]

(3.30)
4 Sample Statistics

The canonical correlations and vectors depend on \( \tilde{S}^{+}_{\Delta X, \Delta X}, \tilde{S}^{+}_{\Delta X, \tilde{X}}, \) and \( \tilde{S}^{+}_{\tilde{X}, \tilde{X}} \), which in turn depend on the submatrices of

\[
\begin{bmatrix}
S_{\tilde{X} \tilde{X}} & S_{\tilde{X}, \Delta X} & S_{\tilde{X} W} \\
S_{\Delta X, \tilde{X}} & S_{\Delta X, \Delta X} & S_{\Delta X, W} \\
S_{W, \tilde{X}} & S_{W, \Delta X} & S_{WW}
\end{bmatrix} = \frac{1}{T} \sum_{t=1}^{T} \begin{bmatrix} X_{1,t-1} \\ \tilde{X}_{t-1} \\ W_t \end{bmatrix} \begin{bmatrix} X_{1,t-1}, \tilde{X}_{t-1}, W_t \end{bmatrix} .
\] (4.1)

(The partitioning on the left of (4.1) is into \( p, (m-1)p, \) and \( p \) rows and columns and on the right is into \( n, (m-1)p + k, \) and \( p \) rows and columns.) The vector \( \tilde{X}_t \) satisfies the first-order stationary autoregressive model (3.9). The sample covariance matrices \( \tilde{S}_{XX}, \tilde{S}_{WX}, \) and \( S_{WW} \) are consistent estimators of \( \tilde{\Sigma}, 0, \) and \( \Sigma_{WW}, \) and \( \tilde{S}^{+}_{XX} = \sqrt{T} (\tilde{S}_{XX} - \tilde{\Sigma}), \tilde{S}^{+}_{WX} = \sqrt{T} \tilde{S}_{WX}, \)

\( \tilde{S}^{+}_{WW} = \sqrt{T} (S_{WW} - \Sigma_{WW}) \) have a limiting normal distribution with means \( 0 \) and covariances that have been given in Anderson (1999) and Anderson (2000a).

The vector \( X_{1t} \) is given by (3.8). Define

\[
I_{11} = \int_{0}^{1} W_1(u) W'_1(u) \, du ,
\] (4.2)

where \( W_1(u) \) is the Brownian motion process defined by

\[
T^{-\frac{1}{2}} \sum_{t=1}^{[T u]} W_{1t} \overset{w}{\to} W_1(u) .
\] (4.3)
See Anderson (2000a) and Theorem B.12 of Johansen (1995). Note that

\[ \frac{1}{T^2} \sum_{t=1}^{T} \sum_{r,s=1}^{t} W_{1s} W'_{1r} \xrightarrow{d} I_{11}. \]  

(4.4)

Since \( \tilde{S}_{XX} \xrightarrow{p} \tilde{\Sigma} \) and \( T^{-1}\tilde{S}_{XX} \xrightarrow{p} 0 \),

\[ \frac{1}{T} S_{XX}^{11} \xrightarrow{d} \Gamma I_{11} \Gamma' \]  

(4.5)

and by the Cauchy-Schwarz inequality

\[ \frac{1}{T^2} \sum_{t=1}^{T} W_{1t} X'_{t} \xrightarrow{p} 0. \]  

(4.6)

We shall find the limit in distribution of \( S_{XX}^{11} \) from the limit of \( S_{XX}^{11} \) by using (B.20) of Theorem B.13 of Johansen (1995). A specialization to the model here is

\[ \frac{1}{T} \sum_{t=1}^{T} \tilde{X}_{t} \sum_{s=1}^{t-1} W'_{1s} \xrightarrow{d} \left( I - \tilde{\Upsilon} \right)^{-1} \left[ \int_{0}^{1} d\tilde{W}(u) W'_{1}(u) + \Sigma_{WW}^{11} \right] - \Sigma_{WW}^{11}, \]  

(4.7)

where \( \tilde{W'}(u) = [W_{2}'(u), W'(u), 0] \) and \( \Sigma_{WW}^{11} = \mathcal{E} \tilde{W}_{1} W_{11} \). [In theorem B.13 let \( \Theta_{i} = (I, 0), \psi_{i} = (0, \tilde{\Upsilon}^{t}), \epsilon'_{t} = (W'_{1}, \tilde{W}_{t}), \) and \( \Omega = \mathcal{E} \epsilon_{t} \epsilon'_{t}, V_{t} = \tilde{X}_{t}. \)] Then

\[
S_{XX}^{11} = \frac{1}{T} \sum_{t=1}^{T} \tilde{X}_{t} X'_{1t} \\
\xrightarrow{d} \left( I - \tilde{\Upsilon} \right)^{-1} \left[ \int_{0}^{1} d\tilde{W}(u) W'_{1}(u) + \Sigma_{WW}^{11} \right] \Sigma' + \tilde{\Sigma} H' \\
= \left( I - \tilde{\Upsilon}^{t} \right) \left[ (J'_{2}, J'_{11}, J'_{21}, 0)' + \Sigma_{WW}^{11} \right] \Gamma' + \tilde{\Sigma} H'.
\]

(4.8)
Because \( \{ \tilde{X}_t \} \) is stationary, \( T^{-1} \sum_{t=1}^{T} W_t \tilde{X}_{t-1} \overset{d}{\to} 0 \) and

\[
S_{W_\tilde{X}} = \begin{bmatrix}
S_{11}^{W_{\tilde{X}}} & S_{12}^{W_{\tilde{X}}} \\
S_{21}^{W_{\tilde{X}}} & S_{22}^{W_{\tilde{X}}}
\end{bmatrix} \overset{d}{\to} \begin{bmatrix}
J_{11} \Gamma' & 0 \\
J_{21} \Gamma' & 0
\end{bmatrix}.
\]

(4.9)

Now we wish to show that \( \Delta X_t^+ \) and \( X_{t-1}^+ \) lead to the same asymptotic results as \( \Delta \tilde{X}_t^+ \) and \( \tilde{X}_{t-1}^+ \). Consider

\[
\hat{S}_{\Delta X, \Delta X}^+ - S_{\Delta X, \Delta X}^+
= \frac{1}{T} \sum_{t=1}^{T} \Delta \tilde{X}_t^+ \Delta \tilde{X}_{t'}^+ - \frac{1}{T} \sum_{t=1}^{T} \Delta X_t^+ \Delta X_{t'}^+
= -S_{\Delta X, \Delta X} \hat{S}_{\Delta X, \Delta X}^{-1} S_{\Delta X, \Delta X}
+ S_{\Delta X, \Delta X} \Sigma_{\Delta X, \Delta X}^{-1} \Sigma_{\Delta X, \Delta X} + \Sigma_{\Delta X, \Delta X} \Sigma_{\Delta X, \Delta X}^{-1} S_{\Delta X, \Delta X}
- \Sigma_{\Delta X, \Delta X} \Sigma_{\Delta X, \Delta X}^{-1} S_{\Delta X, \Delta X} \Sigma_{\Delta X, \Delta X}
\]

(4.10)

Define \( S_{\Delta X, \Delta X}^{\pm} = \sqrt{T} \left( \hat{S}_{\Delta X, \Delta X}^+ - \Sigma_{\Delta X, \Delta X} \right) \) and \( S_{\Delta X, \Delta X}^{\pm} = \sqrt{T} \left( S_{\Delta X, \Delta X}^+ - \Sigma_{\Delta X, \Delta X} \right) \). Then

\[
S_{\Delta X, \Delta X} \Sigma_{\Delta X, \Delta X}^{-1} \Sigma_{\Delta X, \Delta X} S_{\Delta X, \Delta X} = \Sigma_{\Delta X, \Delta X} \Sigma_{\Delta X, \Delta X}^{-1} \Sigma_{\Delta X, \Delta X}
+ \frac{1}{\sqrt{T}} \left( S_{\Delta X, \Delta X}^{\pm} \Sigma_{\Delta X, \Delta X}^{-1} \Sigma_{\Delta X, \Delta X} + \Sigma_{\Delta X, \Delta X} \Sigma_{\Delta X, \Delta X}^{-1} S_{\Delta X, \Delta X}^{\pm} \Sigma_{\Delta X, \Delta X}
- \Sigma_{\Delta X, \Delta X} \Sigma_{\Delta X, \Delta X}^{-1} S_{\Delta X, \Delta X}^{\pm} \Sigma_{\Delta X, \Delta X}
\right)
+ o_p \left( \frac{1}{\sqrt{T}} \right)
\]

(4.11)

and (4.10) times \( \sqrt{T} \) is \( o_p(1) \).
Consider

\[ S_{X,X}^+ = \frac{1}{T} \sum_{t=1}^{T} \left( X_{t-1} - \Sigma_{X,\Delta X} \Sigma_{\Delta X,\Delta X}^{-1} \Delta X_{t-1} \right) \left( X_{t-1} - \Sigma_{X,\Delta X} \Sigma_{\Delta X,\Delta X}^{-1} \Delta X_{t-1} \right)' \quad (4.12) \]

Then

\[ \frac{1}{T} S_{X,X}^{11} \xrightarrow{d} \Gamma_{II} \Gamma' , \quad \frac{1}{T} S_{X,X}^{11} \xrightarrow{d} \Gamma_{II} \Gamma' \quad (4.13) \]

because \((1/T)S_{X,\Delta X} \xrightarrow{p} 0\) and \((1/T)S_{\Delta X,\Delta X} \xrightarrow{p} 0\). Since \(S_{X,X}^+\) is composed of submatrices of \(\tilde{S}_{X,X}\), which has a finite probability limit

\[ S_{X,X}^{+22} \xrightarrow{p} \Sigma_{X,X}^{+22} \quad (4.14) \]

Similarly \(S_{X,X}^{+12}\) has a finite limit in distribution, and

\[ \frac{1}{\sqrt{T}} S_{X,X}^{+12} \xrightarrow{p} 0 \quad (4.15) \]

From (3.19) we obtain

\[ S_{\Delta X,X}^+ = \begin{bmatrix} 0 & 0 & \gamma^{22}S_{X,X}^{+21} & \gamma^{22}S_{X,X}^{+22} \end{bmatrix} + S_{W,X}^+ \quad (4.16) \]

where

\[ S_{W,X}^+ = S_{W,X} - S_{W,\Delta X} \Sigma_{\Delta X,\Delta X}^{-1} \Sigma_{\Delta X,X} \]

\[ \xrightarrow{d} \begin{bmatrix} J_{11} \Gamma' & 0 \\ J_{21} \Gamma' & 0 \end{bmatrix} \quad (4.17) \]
As noted above, $S^{+21}_{X,X} \xrightarrow{d} S^{+21}_\infty$. Then

$$S^{+}_{X,X} \xrightarrow{d} \begin{bmatrix} J_{11} \Gamma' & 0 \\ \Upsilon^{22} S^{+21}_\infty + J_{21} \Gamma' & \Upsilon^{22} \Sigma^{+22}_{X,X} \end{bmatrix}. \quad (4.18)$$

5  Asymptotic Distribution of the Smaller Roots

Let $Q^+ = S^+_{X,X} S^{+,-1}_{X,X} S^+_{X,X}$. The determinantal equation (2.21) for the canonical correlations is $|Q^+ - r^2 S^+_{X,X}| = 0$; that is

$$\begin{bmatrix} \frac{1}{r} Q_{11}^+ & \frac{1}{\sqrt{T}} Q_{12}^+ \\ \frac{1}{\sqrt{T}} Q_{21}^+ & Q_{22}^+ \end{bmatrix} - r^2 \begin{bmatrix} \frac{1}{T} S^{+11}_{X,X} & \frac{1}{\sqrt{T}} S^{+12}_{X,X} \\ \frac{1}{\sqrt{T}} S^{+21}_{X,X} & S^{+22}_{X,X} \end{bmatrix} = 0. \quad (5.1)$$

From (4.12), (4.14), and (4.15) the elements of the second matrix in (5.1) converge in distribution to $\text{diag}(\Gamma I_{11} \Gamma', \Sigma^{+22}_{X,X})$. From (4.18) and $S^+_{X,X} \xrightarrow{P} \Sigma^+_{X,X}$, we obtain

$$Q^+ \xrightarrow{d} \begin{bmatrix} \Gamma J_{11}^r & S^{+21}_{X,X} \Upsilon^{22} + \Gamma J_{21}^r \\ 0 & \Sigma^{+22}_{X,X} \Upsilon^{22} \end{bmatrix} (\Sigma^{+}_{X,X})^{-1} \begin{bmatrix} J_{11} \Gamma' & 0 \\ \Upsilon^{22} S^{+21}_\infty + J_{21} \Gamma' & \Upsilon^{22} \Sigma^{+22}_{X,X} \end{bmatrix} = Q^+_{\infty}, \quad (5.2)$$

say. From the above we find that the probability limit of the roots of (5.1) is the set of roots of

$$\begin{bmatrix} 0 & 0 \\ 0 & (\Sigma^{+}_{X,X} \Sigma^{+,-1}_{X,X} \Sigma^{+}_{X,X})_{22} \end{bmatrix} - r^2 \begin{bmatrix} \Gamma I_{11} \Gamma' & 0 \\ 0 & \Sigma^{+22}_{X,X} \end{bmatrix} = 0. \quad (5.3)$$
The $n$ smaller roots of (5.1) converge in probability to 0, and $k$ larger roots converge in probability to the roots of
\[
|Q_{22}^+ - r^2 \Sigma_{XX}^{+22}| = 0. \tag{5.4}
\]

Now let $d = Tr^2$ to obtain the equation $|Q^+ - d\Gamma^{-1}\Sigma_{XX}| = 0$; the limit in distribution of this equation is
\[
0 = \begin{vmatrix}
Q_{11}^{11} & Q_{12}^{12} \\
Q_{21}^{21} & Q_{22}^{22}
\end{vmatrix} - d \begin{vmatrix}
\Gamma_{11} \Gamma' & 0 \\
0 & 0
\end{vmatrix} = |Q_{\infty}^{22}| \cdot |Q_{\infty}^{11-2} - d\Gamma_{11} \Gamma'|
\tag{5.5}
\]

where
\[
Q_{\infty}^{11-2} = Q_{\infty}^{11} - Q_{\infty}^{12}(Q_{\infty}^{22})^{-1}Q_{\infty}^{21} = [(Q_{\infty}^{-1})_{11}]^{-1}. \tag{5.6}
\]

Use of (5.2) and (3.26) shows (5.5) is
\[
|\Gamma_{11}'(\Sigma_{WW}^{11})^{-1} J_{11} \Gamma' - d\Gamma_{11} \Gamma'| = 0, \tag{5.7}
\]

which is equivalent to (5.3) of Anderson (2000a). Define $B_1(u) = (\Sigma_{WW}^{11})^{-\frac{1}{2}}W_1(u)$, standard Brownian motion ($EB_1(u)B_1'(u) = uI$). Then the roots of (5.7), $d_1 < d_2 < \ldots d_r$, have the distribution of the zeros of
\[
\left| \int_0^1 B_1(u)dB_1'(u) - d \int_0^1 B_1(u)B_1'(u) du \right|. \tag{5.8}
\]

The likelihood ratio criterion for testing that the rank of $\Psi$ is $k$ is
\[
-2 \log \lambda = -T \sum_{i=1}^{p-k} \log(1 - r_i^2) = \sum_{i=1}^{p-k} d_i + o_p(1) \tag{5.9}
\]
Anderson (1951). Its limiting distribution is the distribution of

\[
\text{tr} \int_0^1 dB_1(v)B_1'(v) \left[ \int_0^1 B_1(u)B_1'(u)du \right]^{-1} \int_0^1 B_1(v)dB_1'(v)
\]  

(5.10)


6 Asymptotic Distribution of the Larger Roots

6.1 Asymptotic distribution of sample covariances

We now turn to deriving the asymptotic distribution of the \( k \) larger roots of (2.21) and the associated vectors solving (2.22). First we show that the asymptotic distribution of \( r_{n+1}^2, \ldots, r_p^2 \) is the same as the asymptotic distribution of the zeros of \( |Q_{22}^+ - r^2 S_{XX}^{-22}| \). Then we transform from the \( X \)-coordinates to the coordinates of the process canonical correlations and vectors.

Let \( \hat{R}_2^2 = \text{diag}(r_{n+1}^2, \ldots, r_p^2) \), and let \((G'_{12}, G'_{22})\)' consist of the corresponding solutions to (2.22).

We summarize the equation as

\[
\begin{bmatrix}
Q_{11}^+ & Q_{12}^+ \\
Q_{21}^+ & Q_{22}^+
\end{bmatrix}
\begin{bmatrix}
G_{12} \\
G_{22}
\end{bmatrix}
= 
\begin{bmatrix}
S_{XX}^{11} & S_{XX}^{12} \\
S_{XX}^{21} & S_{XX}^{22}
\end{bmatrix}
\begin{bmatrix}
G_{12} \\
G_{22}
\end{bmatrix}
\hat{R}_2^2.
\]  

(6.1)

The normalization of the columns of \((G'_{12}, G'_{22}))' is

\[
I = (G'_{12}, G'_{22})
\begin{bmatrix}
S_{XX}^{11} & S_{XX}^{12} \\
S_{XX}^{21} & S_{XX}^{22}
\end{bmatrix}
\begin{bmatrix}
G_{12} \\
G_{22}
\end{bmatrix}
= (\sqrt{T}G'_{12}, G'_{22})
\begin{bmatrix}
\frac{1}{\sqrt{T}}S_{XX}^{11} & \frac{1}{\sqrt{T}}S_{XX}^{12} \\
\frac{1}{\sqrt{T}}S_{XX}^{21} & S_{XX}^{22}
\end{bmatrix}
\begin{bmatrix}
\sqrt{T}G_{12} \\
G_{22}
\end{bmatrix}
\]  

(6.2)
The probability limit of (6.2) shows that $\sqrt{T}G_{12} = O_p(1)$ and $G_{22} = O_p(1)$. The submatrix equations in (6.1) can be written as

$$
\frac{1}{T}Q_{11}^+\sqrt{T}G_{12} + \frac{1}{\sqrt{T}}Q_{12}^+G_{22} = \left(\frac{1}{T}S_{XX}^{+11}\sqrt{T}G_{12} + \frac{1}{\sqrt{T}}S_{XX}^{+12}G_{22}\right)\hat{R}_2^2, \quad (6.3)
$$

$$
\frac{1}{\sqrt{T}}Q_{21}^+\sqrt{T}G_{12} + Q_{22}^+G_{22} = \left(\frac{1}{\sqrt{T}}S_{XX}^{+21}G_{12} + S_{XX}^{+22}G_{22}\right)\hat{R}_3^2. \quad (6.4)
$$

Since $T^{-1}Q_{11}^+ \xrightarrow{p} 0$, $T^{-1}Q_{12}^+ \xrightarrow{p} 0$, $T^{-1}S_{XX}^{+12} \xrightarrow{d} \Gamma_{11}\Gamma'$ and $\hat{R}_3^2 \xrightarrow{p} R_3^2 = \text{diag}(\rho_{n+1}^2, \ldots, \rho_p^2)$, the probability limit of the left-hand side of (6.3) is 0; this shows that $\sqrt{T}G_{12} \xrightarrow{p} 0$. Then the limiting distribution of $G_{22}$ is the asymptotic distribution of $G_{22}$ defined by

$$
Q_{22}^+G_{22} = S_{XX}^{+22}G_{22}\hat{R}_2^2, \quad G_{22}^{'T}S_{XX}^{+22}G_{22} = I, \quad (6.5)
$$

where the elements of $\hat{R}_2^2$ are defined by

$$
\left|Q_{20}^+ - r^2S_{XX}^{+22}\right| = 0. \quad (6.6)
$$

Note that when $\sqrt{T}G_{12} \xrightarrow{p} 0$ is combined with (6.4) we obtain

$$
Q_{22}^+G_{22} = S_{XX}^{+22}G_{22}\hat{R}_2^2 + o_p\left(\frac{1}{\sqrt{T}}\right). \quad (6.7)
$$

We proceed to find the asymptotic distribution of $G_{22}$ and $\hat{R}_2^2$ defined by (6.5) in the manner of Anderson (2000a). Let

$$
S_{XX}^{+*} = \sqrt{T}(S_{XX}^+ - \Sigma_{XX}), \quad S_{XX}^{+22} = \sqrt{T}(S_{XX}^{22} - \Sigma_{XX}^{22}). \quad (6.8)
$$
\[ S_{X}^{+2,2-1} = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} X_{t}^{+} W_{2,1,t}, \quad S_{W,W}^{+2,1,2-1} = \frac{1}{\sqrt{T}} \left( \sum_{t=1}^{T} W_{2,1,t} W_{2,1,t}' - \Sigma_{WW}^{22,1} \right), \quad (6.9) \]

where \( W_{2,1,t} = W_{2t} - \Sigma_{WW}^{21}(\Sigma_{WW}^{-1})^{-1} W_{1t} \) and \( \varepsilon W_{2,1,t} W_{2,1,t}' = \Sigma_{WW}^{22,1} = \Sigma_{WW}^{21} - \Sigma_{WW}^{21} (\Sigma_{WW}^{-1}) \Sigma_{WW}^{12} \).

We expand \( \sqrt{T} \left\{ Q_{22}^{+} - \left[ \Sigma_{X,\Delta X}(\Sigma_{\Delta X,\Delta X}^{-1})^{-1} \Sigma_{\Delta X,\Delta X}^{+} \right]_{22} \right\} \) as in the manner of Anderson (2000a) to obtain

\[
\sqrt{T} \gamma^{22} \left\{ Q_{22}^{+} - \left[ \Sigma_{X,\Delta X}(\Sigma_{\Delta X,\Delta X}^{-1})^{-1} \Sigma_{\Delta X,\Delta X}^{+} \right]_{22} \right\} \gamma^{22'} = - \gamma^{22} \Sigma_{X,X}^{+} \gamma^{22'} \Lambda^{-1} S_{WW}^{+2,1,2-1} \Lambda^{-1} \gamma^{22} \Sigma_{X,X}^{+} \gamma^{22'} + \gamma^{22} \Sigma_{WW}^{+2,1,2-1} \Lambda^{-1} \gamma^{22} \Sigma_{X,X}^{+} \gamma^{22'} + \gamma^{22} \Sigma_{WW} \Lambda^{-1} \gamma^{22} \Sigma_{X,X}^{+} \gamma^{22'} + \gamma^{22} \Sigma_{X,X}^{+} \gamma^{22'} \Lambda^{-1} \gamma^{22} \Sigma_{X,X}^{X} \gamma^{22'} + o_{p}(1), \quad (6.10)\]

where \( \Lambda = \gamma^{22} \Sigma_{X,X}^{+} \gamma^{22'} + \Sigma_{WW}^{+2,1} \). See (6.5) of Anderson (2000a).

To express the covariances of the sample matrices we use the “vec” notation. For \( A = (a_1, \ldots, a_n) \) we define \( \text{vec} A = (a'_1, \ldots, a'_n)' \). The Kronecker product of two matrices \( A = (a_{ij}) \) and \( B \) is \( A \otimes B = (a_{ij}B) \). A basic relation is \( \text{vec} ABC = (C' \otimes A) \text{vec} B \), which implies \( \text{vec} yx' = \text{vec} xy' = (y \otimes x) \text{vec} 1 = y \otimes x \). Define the commutator matrix \( K \) as the (square) permutation matrix such that \( \text{vec} A' = K \text{vec} A \) for every square matrix of the same order as \( K \).

Define \( C = \left( I, -\Sigma_{\Delta X}^{+1} \Sigma_{\Delta X,\Delta X}^{-1} \right), \ D = \left[ I, -\Sigma_{WW}^{21} (\Sigma_{WW}^{-1})^{-1}, 0 \right] \). Then \( X_{2t} = C \bar{X}_{t}, \ W_{2,1,t} = D \bar{W}_{t}, \ C \Sigma = \Sigma_{X,X}^{+22} I_{(k)}, \) and \( D \Sigma_{WW} = \Sigma_{WW}^{22,1} J_{(k)}' \), where \( I_{(k)} = (I, 0), \ J_{(k)}' = (I, 0, I, 0), \ C \Sigma C' = \Sigma_{X,X}^{+22}, \) and \( D \Sigma_{WW} D' = \Sigma_{WW}^{22,1} \).
Theorem 1. If the $W_t$ are independently normally distributed, $S_{WW}^{2,1,2}$, $S_{W}^{+*2,1,2}$ and $S_{X}^{+*22}$ have a limiting normal distribution with means $0, 0$, and $0$ and covariances

$$
\mathcal{E} \text{vec } S_{WW}^{2,1,2} \left( \text{vec } S_{WW}^{2,1,2} \right)' = (I + K)(\Sigma_{WW}^{22} \otimes \Sigma_{WW}^{22}), \quad (6.11)
$$

$$
\mathcal{E} \text{vec } S_{X}^{+*2,1,2} \left( \text{vec } S_{W}^{+*2,1,2} \right)' = \Sigma_{XX}^{+22} \otimes \Sigma_{WW}^{22}, \quad (6.12)
$$

$$
\mathcal{E} \text{vec } S_{X}^{+*2,1,2} \left( \text{vec } S_{W}^{+*2,1,2} \right)' = 0, \quad (6.13)
$$

$$
\mathcal{E} \text{vec } S_{X}^{+*22} \left( \text{vec } S_{WW}^{2,1,2} \right)' \rightarrow (I + K)(C \otimes C)[I - (\tilde{\Theta} \otimes \tilde{\Theta})]^{-1}(\tilde{\Sigma}_{WW} \otimes \tilde{\Sigma}_{WW})(D' \otimes D')
$$

$$
= (I + K)(C \otimes C)[I - (\tilde{\Theta} \otimes \tilde{\Theta})]^{-1} (I(k) \Sigma_{WW}^{22} \otimes J(k) \Sigma_{WW}^{22}) , \quad (6.14)
$$

$$
\mathcal{E} \text{vec } \Sigma_{XX}^{+*22} \left( \text{vec } \Sigma_{W}^{+*2,1,2} \right) \rightarrow (I + K)(C \otimes C)[I - (\tilde{\Theta} \otimes \tilde{\Theta})]^{-1}(\tilde{\Sigma}_{WW} \otimes \tilde{\Sigma}_{WW})(C' \otimes D')
$$

$$
= (I + K)(C \otimes C)[I - (\tilde{\Theta} \otimes \tilde{\Theta})]^{-1} (I(k) \Sigma_{XX}^{+22} \otimes J(k) \Sigma_{WW}^{22}) , \quad (6.15)
$$

$$
\mathcal{E} \text{vec } S_{X}^{+*22} \left( \text{vec } S_{X}^{+*22} \right)' = (I + K)(C \otimes C)[I - (\tilde{\Theta} \otimes \tilde{\Theta})]^{-1}(\tilde{\Sigma} \otimes \tilde{\Sigma}) + (\tilde{\Sigma}_{WW} \otimes \tilde{\Sigma})
$$

$$
- (\tilde{\Sigma}_{WW} \otimes \tilde{\Sigma}_{WW})[I - (\tilde{\Theta}' \otimes \tilde{\Theta})]'^{-1}(C' \otimes C')
$$

$$
= (I + K)(C \otimes C) \left\{ [I - (\tilde{\Theta} \otimes \tilde{\Theta})]'^{-1}(\tilde{\Sigma} \otimes \tilde{\Sigma})
$$

$$
+ (\tilde{\Sigma} \otimes \tilde{\Sigma})[I - (\tilde{\Theta}' \otimes \tilde{\Theta})]'^{-1} - (\tilde{\Sigma} \otimes \tilde{\Sigma}) \right\}(C' \otimes C')
$$

$$
= (I + K) \left\{ (C \otimes C)[I - (\tilde{\Theta} \otimes \tilde{\Theta})]^{-1} (I(k) \Sigma_{XX}^{+22} \otimes I(k) \Sigma_{XX}^{+22})
$$

$$
+ \left( \Sigma_{XX}^{+22} (I(k) \otimes I(k)) \left[I - (\tilde{\Theta}' \otimes \tilde{\Theta}') \right] (C' \otimes C') - (\Sigma_{XX}^{+22} \otimes \Sigma_{XX}^{+22}) \right) \right\} . \quad (6.16)
$$

Lemma 1. If $X$ is normally distributed with $\mathcal{E}X = 0$ and $\mathcal{E}XX' = \Sigma$, then

$$
\mathcal{E} \text{vec } XX'(vec XX')' = (I + K)(\Sigma \otimes \Sigma) + vec \Sigma(\text{vec } \Sigma)' . \quad (6.17)
$$
If X and Y are independent,

\begin{align}
\mathcal{E} \text{ vec} XX' & (\text{ vec } YY')' = \text{ vec }\mathcal{E}XX' \otimes (\text{ vec }\mathcal{E}'YY')', \\
\mathcal{E} \text{ vec } XY' & (\text{ vec } XY')' = \mathcal{E}YY' \otimes \mathcal{E}XX', \\
\mathcal{E} \text{ vec } XY' (\text{ vec } (YY')')' = K\mathcal{E}XX' \otimes \mathcal{E}YY'.
\end{align}

(6.18) (6.19) (6.20)

Proof of Theorem 1. First (6.11) is equivalent to (6.17). Next \( \text{ vec } S_{Wt}^{+2:1,2} = T^{-1/2} \sum_{t=1}^{T} (X_{2,t-1}^+ \otimes W_{2,1,t}) \) implies (6.12) because \( X_{2,t-1}^+ \) and \( W_{2,1,s} \) are independent for \( t - 1 \leq s \). Similarly (6.13) follows.

To prove (6.14), (6.15) and (6.16) we use the following lemma.

Lemma 2.

\[ \text{ vec } \tilde{S}_{XX} = [I - (\tilde{Y} \otimes \tilde{Y})]^{-1} [(I + K)(\tilde{Y} \otimes I) \text{ vec } \tilde{S}_{Wt} + \text{ vec } \tilde{S}_{Wt}] + o_p \left( \frac{1}{\sqrt{T}} \right). \]  

(6.21)

Proof of Lemma 2. We have from \( \tilde{X}_t = \tilde{Y} \tilde{X}_{t-1} + \tilde{W}_t \)

\[ \tilde{S}_{XX} = \tilde{Y} \tilde{S}_{\tilde{X} \tilde{X}} \tilde{Y}' + \tilde{Y} \tilde{S}_{\tilde{X}W} + \tilde{S}_{W \tilde{X}} \tilde{Y}' + \tilde{S}_{WW}. \]

(6.22)

Since \( \tilde{S}_{XX} - \tilde{S}_{\tilde{X} \tilde{X}} = (1/T)(\tilde{X}_T \tilde{X}_T' - \tilde{X}_0 \tilde{X}_0') \) and \( \{\tilde{X}_t\} \) is a stationary process,

\[ \tilde{S}_{\tilde{X} \tilde{X}} - \tilde{Y} \tilde{S}_{\tilde{X} \tilde{X}} \tilde{Y}' = \tilde{Y} \tilde{S}_{\tilde{X}W} + \tilde{S}_{W \tilde{X}} \tilde{Y}' + \tilde{S}_{WW} + o_p \left( \frac{1}{\sqrt{T}} \right). \]  

(6.23)

Then Lemma 2 results from \( \text{ vec } \tilde{Y} \tilde{S}_{\tilde{X} \tilde{W}} = K \text{ vec } \tilde{S}_{W \tilde{X}} \tilde{Y}'. \)  

24
Lemma 3.

\[ \mathcal{E} \text{ vec } W_{2t} W_{2t}' \left( \text{ vec } W_{2,1,t} W_{2,1,t}' \right)' = (I + K)(\Sigma_{WW}^{22} \otimes \Sigma_{W}^{22} + \text{ vec } \Sigma_{WW}^{22})'(\text{ vec } \Sigma_{WW}^{22})'. \quad (6.24) \]

Proof of Lemma 3. Write \( W_{2t} = W_{2,1,t} + \Sigma_{WW}^{21}(\Sigma_{WW}^{11})^{-1} W_{1t} \). Then

\[ \mathcal{E} \text{ vec } W_{2t} W_{2t}' \left( \text{ vec } W_{2,1,t} W_{2,1,t}' \right)' = \mathcal{E} \left\{ \left[ W_{2,1,t} + \Sigma_{WW}^{21}(\Sigma_{WW}^{11})^{-1} W_{1t} \right] \otimes \left[ W_{2,1,t} + \Sigma_{WW}^{21}(\Sigma_{WW}^{11})^{-1} W_{1t} \right] \right\} (W_{2,1,t} \otimes W_{2,1,t}) \]

\[ = (I + K) (\Sigma_{WW}^{22} \otimes \Sigma_{WW}^{22}) + \text{ vec } \Sigma_{WW}^{22} \left( \text{ vec } \Sigma_{WW}^{22} \right)' \]

\[ + \text{ vec } \left[ \Sigma_{WW}^{21}(\Sigma_{WW}^{11})^{-1} \Sigma_{WW}^{21} \right] \left( \text{ vec } \Sigma_{WW}^{22} \right)' \quad (6.25) \]

from which the lemma follows.

Proof of Theorem 1 continued. The first part of (6.14) follows from Lemma 2, (6.11) and (6.12). The second part follows from

\[ \tilde{\Sigma}_{WW} D' = \begin{bmatrix} \Sigma_{WW}^{22} & \Sigma_{WW}^{21} & \Sigma_{WW}^{22} & 0 \\ \Sigma_{WW}^{12} & \Sigma_{WW}^{11} & \Sigma_{WW}^{12} & 0 \\ \Sigma_{WW}^{22} & \Sigma_{WW}^{21} & \Sigma_{WW}^{22} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} I \\ -(\Sigma_{WW}^{11})^{-1} \Sigma_{WW}^{12} \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \Sigma_{WW}^{22} \\ 0 \\ \Sigma_{WW}^{22} \\ 0 \end{bmatrix} = \Sigma_{WW}^{22} J_{(k)} \quad (6.26) \]

The first part of (6.15) follows from Lemma 2, (6.12) and (6.13). The second part follows from (3.10) and

\[ \tilde{\Sigma} C' = \begin{bmatrix} \tilde{\Sigma}_{22} & \tilde{\Sigma}_{2,\Delta X} \\ \tilde{\Sigma}_{\Delta X,2} & \tilde{\Sigma}_{\Delta X,\Delta X} \end{bmatrix} \begin{bmatrix} I \\ -\tilde{\Sigma}_{\Delta X,\Delta X}^{-1} \Sigma_{\Delta X,2} \end{bmatrix} = \begin{bmatrix} \Sigma_{\Delta X,\Delta X}^{+22} \\ 0 \end{bmatrix}, \quad (6.27) \]
To prove (6.16) use Lemma 2, (6.11), (6.12), and (3.3) to obtain

\[
\mathcal{E} \, \text{vec} \, S^{+22} (\text{vec} \, S^{+22})' = (C \otimes C)[I - (\tilde{Y} \otimes \tilde{Y})]^{-1} \\
\{ (I + K)(\tilde{Y} \otimes I)(\tilde{X} \otimes I)(\tilde{Y} \otimes I)(I + K) \\
+ (I + K)(\tilde{\Sigma}_{WW} \otimes \tilde{\Sigma}_{WW})[I - (\tilde{Y}' \otimes \tilde{Y}'')]^{-1}(C' \otimes C') \\
= (I + K)(C \otimes C)[I - (\tilde{Y} \otimes \tilde{Y})]^{-1}[(\tilde{Y} \tilde{\Sigma} \tilde{Y}' \otimes \tilde{\Sigma}_{WW})(I + K) \\
+ (\tilde{\Sigma}_{WW} \otimes \tilde{\Sigma}_{WW})][I - (\tilde{Y}' \otimes \tilde{Y}')^{-1}(C' \otimes C').
\]

Then substitution of \( \tilde{\Sigma}_{WW} = \tilde{\Sigma} - \tilde{Y} \tilde{\Sigma} \tilde{Y}' \) yields the first form of (6.16), which leads to the second form.

6.2 Transformation to canonical form

Let \( \Xi \) be a \( k \times k \) matrix such that

\[
\Xi' (\Upsilon_2 \Sigma^{+22}_{XX} \Upsilon_2') \Xi = \Theta, \quad \Xi' \Sigma^{22}_{WW} \Xi = I,
\]

where \( \Theta = \text{diag}(\theta_{n+1}, \ldots, \theta_p) = R_2^2(I - R_2^2)^{-1} \), \( R_2^2 = \text{diag}(\rho_{n+1}^2, \ldots, \rho_p^2) \), and \( \rho_i^2 \) is a root of (3.22) with \( 0 < \rho_{n+1}^2 < \ldots < \rho_p^2 \). Let \( U_{2t}^+ = \Xi' X_{2t}, \, V_{2t} = \Xi' W_{2t}, \, V_{1t} = W_{1t}, \, \Delta_2 = \Xi' (\Upsilon_{22} + I)(\Xi')^{-1} \), \( M_2 = \Xi' \Upsilon^{22} (\Xi')^{-1} \),

\[
\tilde{\Xi} = \begin{bmatrix}
\Xi & 0 \\
0 & I_{m-1} \otimes \begin{pmatrix} I_n & 0 \\
0 & \Xi \end{pmatrix}
\end{bmatrix}.
\]

26
\[ \bar{\Delta} = \bar{\Xi}'\bar{Y}(\bar{\Xi'})^{-1}, \bar{U}_t = \bar{\Xi}'\bar{X}_t, \bar{C}_U = \bar{\Xi}'C(\bar{\Xi}')^{-1}. \]

Then \{\bar{U}_t\} is generated by

\[
\bar{U}_t = \bar{\Delta}\bar{U}_{t-1} + \tilde{V}_t ,
\]

where \(\tilde{V}_t = \bar{\Xi}'\tilde{W}_t\) and \(U^+_{2t}\) satisfies

\[
U^+_{2t} = \Delta_2 U^+_{2t-1} + V_{2t}, \quad \Delta U^+_{2t} = M_2 U^+_{2t-1} + V_{2t} .
\]

Multiplication of (6.10) on the left by \(\Xi'\) and right by \(\Xi\) yields

\[
\sqrt{T} \left[ M_2 (S^{+}_{\theta,\Delta U} S^{+\perp}_{\Delta U,\theta} S^{+\perp}_{\Delta U,\theta})_{22} M_2' - R^2_2 \Theta \right]
\]

\[
= -R^2_2 S^{+2,1,2-1}_{VV} R^2_2 + R^2_2 S^{+2,1,2}_{VV} M_2' (I - R^2_2)
\]

\[
+ (I - R^2_2) M_2 S^{+2,1,2}_{\theta V} R^2_2 + R^2_2 M_2 S^{+2,22}_{\theta \theta} M_2'
\]

\[
+ R^2_2 M_2 S^{+2,22}_{\theta \theta} M_2' - R^2_2 M_2 S^{+2,22}_{\theta \theta} M_2' R^2_2 + o_p(1) .
\]

Theorem 2. If the \(V_t\) are independently normally distributed, \(S^{+2,1,2}_{VV}\), \(S^{+2,1,2}_{V\theta}\), and \(S^{+2,22}_{\theta \theta}\) have a limiting normal distribution with means \(0, 0, 0\), and \(0\) and covariances

\[
\mathcal{E} \text{ vec } S^{+2,1,2}_{VV} (\text{ vec } S^{+2,1,2}_{VV})' = (I + K)(I \otimes I),
\]

\[
\mathcal{E} \text{ vec } S^{+2,1,2}_{V\theta} (\text{ vec } S^{+2,1,2}_{V\theta})' = M_2^{-1} \Theta M_2^{-1} \otimes I ,
\]

\[
\mathcal{E} \text{ vec } S^{+2,1,2}_{V\theta} (\text{ vec } S^{+2,1,2}_{VV})' = 0 ,
\]

\[
\mathcal{E} \text{ vec } S^{+2,22}_{\theta \theta} (\text{ vec } S^{+2,1,2}_{VV})' = (I + K)(C_U \otimes C_U)[I - (\bar{\Delta} \otimes \bar{\Delta})]^{-1} (J(k) \otimes J(k)) ,
\]

(6.37)
\[ \mathcal{E} \text{ vec } S_{vU}^{++22}(\text{ vec } S_{vL}^{++2,1,2})^t \]

\[ = (I + K)(C_U \otimes C_U)[I - (\Delta \otimes \Delta)]^{-1} \left( \bar{\Delta} I_{(k)} M_2^{-1} \Theta M_2' \otimes J_{(k)} \right) , \]  

(6.38)

\[ \mathcal{E} \text{ vec } S_{vU}^{++22}(\text{ vec } S_{vU}^{++22})^t \]

\[ = (I + K) \left\{ (C_U \otimes C_U)[I - (\Delta \otimes \Delta)]^{-1} \left( I_{(k)} M_2^{-1} \Theta M_2^{-1} \otimes I_{(k)} M_2^{-1} \Theta M_2^{-1} \right) 
+ (M_2^{-1} \Theta M_2^{-1} I_{(k)}') \otimes M_2^{-1} \Theta M_2^{-1} I_{(k)}' [I - (\Delta' \otimes \Delta')]^{-1} (C_U \otimes C_U) 
- (M_2^{-1} \Theta M_2^{-1} \otimes M_2^{-1} \Theta M_2^{-1} \right\} . \]  

(6.39)

Let \( L_{2,t-1} = M_2 U_{2,t-1} (= \Sigma' T^{22} X_{2,t-1}) \),

\[ \bar{M} = \begin{bmatrix} M_2 & 0 \\ 0 & I_{m-1} \end{bmatrix} \begin{pmatrix} I_2 & 0 \\ 0 & M_2 \end{pmatrix} \]  

(6.40)

Then (6.33) becomes

\[ \sqrt{T} \left[ \left( S_{L,\Delta U}^{+} S_{\Delta U,\Delta U}^{+-1} S_{\Delta U,L}^{+} \right)_{22} - R^2 (I - R^2) \right] 
= -R_2^2 S_{vV}^{+2,1,2} R_2^2 + R_2^2 S_{vL}^{++2,1,2} (I - R_2^2) 
+ (I - R_2^2) S_{vV}^{++2,2,1} R_2^2 + R_2^2 S_{vL}^{++22} 
+ S_{vL}^{++22} R_2^2 - R_2^2 S_{vL}^{++22} R_2^2 + O_p(1) . \]  

(6.41)

**Theorem 3.** If the \( V_t \) are independently normally distributed, \( S_{vV}^{+2,1,2} \) and \( S_{vL}^{++22} \) have
a limiting normal distribution with means 0, 0, and 0 and covariances (6.34),

\[ \mathcal{E} \text{ vec } S_{WL}^{+2,1,2}(\text{vec } S_{WL}^{+2,1,2})' = \Theta \otimes I, \quad (6.42) \]

\[ \mathcal{E} \text{ vec } S_{VL}^{+2,1,2}(\text{vec } S_{VL}^{+2,1,2})' = 0, \quad (6.43) \]

\[ \mathcal{E} \text{ vec } S_{LL}^{+22}(\text{vec } S_{VV}^{+21,2,1})' = (I + K)(M_2 C_U \otimes M_2 C_U)[I - (\tilde{\Delta} \otimes \tilde{\Delta})^{-1}(J_{(k)} \otimes J_{(k)})], \quad (6.44) \]

\[ \mathcal{E} \text{ vec } S_{LL}^{+22}(\text{vec } S_{VL}^{+2,1,2})' = (I + K)(M_2 C_U \otimes M_2 C_U)[I - (\tilde{\Delta} \otimes \tilde{\Delta})^{-1}(\tilde{\Delta} I_{(k)} M_2^{-1} \Theta \otimes J_{(k)})], \quad (6.45) \]

\[ \mathcal{E} \text{ vec } S_{LL}^{+22}(\text{vec } S_{LL}^{+22})' = (I + K) \left\{ (M_2 C_U \otimes M_2 C_U)[I - (\tilde{\Delta} \otimes \tilde{\Delta})^{-1}(I_{(k)} M_2^{-1} \Theta \otimes I_{(k)} M_2^{-1} \Theta) \right. \]

\[ + \left. (\Theta M_2^{-1} I_{(k)} \otimes \Theta M_2^{-1} I_{(k)})[I - (\tilde{\Delta} \otimes \tilde{\Delta})^{-1}(C_U M_2 \otimes C_U M_2) - (\Theta \otimes \Theta) \right] \}

\[ = (I + K) \left[ \Phi^+ (\Theta \otimes \Theta) + (\Theta \otimes \Theta) \Phi^+ - (\Theta \otimes \Theta) \right], \quad (6.46) \]

where

\[ \Phi^+ = (M_2 C_U \otimes M_2 C_U) \left[ I - (\tilde{\Delta} \otimes \tilde{\Delta}) \right]^{-1}(I_{(k)} M_2^{-1} \otimes I_{(k)} M_2^{-1}). \quad (6.47) \]

Let \( H_{22} = (M_2')^{-1} \Xi^{-1} G_{22} = \Xi^{-1} (\gamma^{22'})^{-1} G_{22} \). Then \( Q_{22}^+ G_{22} = S_{XX}^{+22} G_{22} \tilde{R}_2^2 \) transforms to

\[ (S_{L,\Delta U}^+ S_{\Delta U, \Delta U}^{-1} S_{\Delta U, L}^+)^{22} H_{22} = S_{LL}^{+22} H_{22} \tilde{R}_2^2, \quad (6.48) \]

and \( G_{22}^T S_{XX}^{+22} G_{22} = I \) transforms to

\[ H_{22}^T S_{LL}^{+22} H_{22} = I. \quad (6.49) \]

Since \( (S_{L,\Delta U}^+ S_{\Delta U, \Delta U}^{-1} S_{\Delta U, L}^+)_{22} \overset{P}{\to} \Theta R_2^2 \) and \( S_{LL}^{+22} \overset{P}{\to} \Theta \), the probability limits of (6.46) and (6.47), and \( h_t > 0 \) imply \( H_{22} \overset{P}{\to} \Theta^{-\frac{1}{2}} \).
Define $H_{22}^* = \sqrt{T}(H_{22} - \Theta^{-\frac{1}{2}})$ and $\hat{R}_2^{2*} = \sqrt{T}(\hat{R}_2^2 - R_2^2)$. Then we can write (6.47) and (6.48) as

$$\Theta^{-1} P \Theta^{-\frac{1}{2}} = \Theta^{-\frac{1}{2}} \hat{R}_2^{2*} + H_{22}^* R_2^2 - R_2^2 H_{22}^* + o_p(1) , \quad (6.50)$$

where

$$P = -R_2^2 S_{VV}^{*2,1,2,1} R_2^2 + R_2^2 S_{VL}^{++2,1,2}(I - R_2^2),$$

$$+(I - R_2^2)S_{LV}^{++2,1,2} R_2^2 + R_2^2 S_{LL}^{++22}(I - R_2^2), \quad (6.51)$$

$$H_{22}^* \Theta^{-\frac{1}{2}} + \Theta^{-\frac{1}{2}} H_{22}^* = -\Theta^{-\frac{1}{2}} S_{LL}^{++22} \Theta^{-\frac{1}{2}} + o_p(1) . \quad (6.52)$$

Lemma 4.

$$\mathcal{E}(I + K)[(I - R_2^2) \otimes R_2^2] \text{vec } S_{VL}^{++2,1,1}(\text{vec } S_{VL}^{++2,1,2})'[I - R_2^2 \otimes R_2^2](I + K)$$

$$= (I + K)[R_2^2(I - R_2^2) \otimes R_2^2](I + K) . \quad (6.53)$$

Lemma 5.

$$\mathcal{E}[(I - R_2^2) \otimes R_2^2] \text{vec } S_{LL}^{++22}\left\{-(\text{vec } S_{VV}^{*2,1,2})'(R_2^2 \otimes R_2^2) + (\text{vec } S_{VL}^{++2,1,1})'[(I - R_2^2) \otimes R_2^2](I + K)\right\}$$

$$= -[R_2^2(I - R_2^2) \otimes R_2^2](I + K) . \quad (6.54)$$

Proof of Lemma 5. We use the facts that $M_2 = \Delta_2 - I$, $J_{(k)} M_2 = \tilde{\Delta} I_{(k)} - I_{(k)} = (\tilde{\Delta} - I) I_{(k)}$, and $(I + K) K = I + K$. Then the left-hand side of (6.54) is

$$[(I - R_2^2) \otimes R_2^2](I + K)(M_2 C_U \otimes M_2 C_U)[I - (\tilde{\Delta} \otimes \tilde{\Delta})]^{-1}$$
\[
\left\{ (\Delta I_{(k)} \otimes J_{(k)} M_2) \left[(I + K) - (J_{(k)} M_2 \otimes J_{(k)} M_2)\right] (M_2^{-1} R_2^2 \otimes M_2^{-1} R_2^2) \right\} \\
= \left[(I - R_2^2) \otimes R_2^2\right] (I + K)(M_2 C U \otimes M_2 C U)[I - (\tilde{\Delta} \otimes \tilde{\Delta})]^{-1} \left\{ (\tilde{\Delta} \otimes (\tilde{\Delta} - I)) + [(\tilde{\Delta} - I) \otimes \tilde{\Delta}] \\
- [(\tilde{\Delta} - I) \otimes (\tilde{\Delta} - I)] \right\} (I_{(k)} \otimes I_{(k)})(M_2^{-1} R_2^2 \otimes M_2^{-1} R_2^2) \\
= - [(I - R_2^2) \otimes R_2^2] (I + K)(M_2 C U \otimes M_2 C U)(I_{(k)} \otimes I_{(k)})(M_2^{-1} R_2^2 \otimes M_2^{-1} R_2^2), \tag{6.55}
\]

which is the right-hand side of (6.54).

Note that (6.54) plus the transpose of (6.54) is equal to the negative of (6.53).

**Theorem 4.** If \(Z_t\) are normally distributed and the roots of (3.26) are distinct,

\[
\mathcal{E}\text{ vec } P'(\text{vec } P') = (I + K)(R_2^4 \otimes R_2^4) \\
+ [(I - R_2^2) \otimes R_2^2](I + K) \left[ \Phi^+ (\Theta \otimes \Theta) + (\Theta \otimes \Theta) \Phi^{+\prime} - (\Theta \otimes \Theta) \right] [(I - R_2^2) \otimes R_2^2] \\
= (R_2^4 \otimes R_2^4) - [(I - R_2^2) R_2^2 \otimes R_2^4] \\
+ (I + K)[(I - R_2^2) \otimes R_2^2][\Phi^+(\Theta \otimes \Theta) + (\Theta \otimes \Theta) \Phi^+] [(I - R_2^2) \otimes R_2^2]. \tag{6.56}
\]

Note that (6.56) is (6.14) of Anderson (2000a) with \(\Phi^+\) replacing \(\Phi\).

### 6.3 Distribution of larger canonical roots

Let \(\widehat{E} = \sum_{i=1}^{k} \varepsilon_i (\varepsilon'_i \otimes \varepsilon'_i)\), where \(\varepsilon_i\) is the \(k\)-vector with 1 in the \(i\)-th position and 0's elsewhere. The matrix \(\widehat{E}\) has 1 in the \(i\)-th row and \(i, i\)-th column and 0's elsewhere. Define \(r^{2*} = (r_{n+1}^{2*}, \ldots, r_p^{2*})'\).

Then

\[
r^{2*} = \widehat{E} \text{ vec } \Theta^{-\frac{1}{2}} P \Theta^{-\frac{1}{2}}. \tag{6.57}
\]
The matrix $\tilde{E}$ has the effect of selecting the $i,i$-th element of $\Theta^{-\frac{1}{2}}P\Theta^{-\frac{1}{2}}$ and placing it in the $i$-th position of $r^{2*}$.

**Theorem 5.** If the $Z_t$ vectors are independently normally distributed and the roots of (3.26) are distinct, the limiting distribution of $r^{2*}$ is normal with mean 0 and covariance matrix

$$2[(I - R^2_2)^2 \tilde{E}\Phi^+ \tilde{E}'R^4_2 + R^4_2 \tilde{E}\Phi^+ \tilde{E}'(I - R^2_2)^2]. \tag{6.58}$$

In terms of the components of $r^{2*}$ the asymptotic covariance of $r^{2*}_i$ and $r^{2*}_j$ is

$$2[(1 - \rho^2_i)^2 \phi_{ii,ij}^+ \rho_j^4 + \rho^4_i (\phi_{jj,ii}^+ (1 - \rho^2_j)^2)]. \tag{6.59}$$

Here $\phi_{ii,ij}^+$ denotes the element in the $i$-th row of the $i$-th block of rows and the $j$-th column of the $i$-th block of columns in $\Phi^+$.

### 6.4 Distribution of canonical vectors

We now derive the limiting distribution of $H^*_2 = H^*_d + H^*_n$, where $H^*_d = \text{diag}(h^*_n, h^*_n, \ldots, h^*_p)$.

From (6.48), $\text{vec} H^*_2 R^2_2 = (R^2_2 \otimes I) \text{vec} H^*_2$, and $\text{vec} R^2_2 H^*_2 = (I \otimes R^2_2) \text{vec} H^*_2$, we obtain

$$\text{vec}(H^*_2 R^2_2 - R^2_2 H^*_2) = NH^*_2 = NH^*_n, \tag{6.60}$$

where

$$N = (R^2_2 \otimes I) - (I \otimes R^2_2)$$

$$= \text{diag}(0, \rho^2_{n+1} - \rho^2_{n+2}, \ldots, \rho^2_{n+1} - \rho^2_p, \rho^2_{n+2} - \rho^2_{n+1}, 0, \ldots, \rho^2_p - \rho^2_{p-1}, 0). \tag{6.61}$$
The Moore-Penrose generalized inverse of $N$ is

$$N^+ = \text{diag}(0, (\rho_{n+1}^2 - \rho_{n+2}^2)^{-1}, \ldots, (\rho_{n+1}^2 - \rho_{n+2}^2)^{-1}, \rho_{n+1}^2 - \rho_{n+2}^2)^{-1}, 0, \ldots, (\rho_{p}^2 - \rho_{p-1}^2)^{-1}, 0).$$  \hspace{1cm} (6.62)

Note that $NN^+ = (I \otimes I) - E$, where $E = \Sigma_{i=1}^k (e_i \otimes e_i)(e_i' \otimes e_i')$. The $k^2 \times k^2$ matrix $E$ is idempotent of rank $k$; the $k^2 \times k^2$ matrix $NN^+$ is idempotent of rank $k^2 - k$; and $E$ is orthogonal to $N$ and $N^+$. From (6.62) we obtain

$$\text{vec } H_n^* = N^+ \text{vec}(\Theta^{-1}P\Theta^{-\frac{1}{2}}) = N^+ (\Theta^{-\frac{1}{2}} \otimes H^{-1}) \text{vec } P.$$  \hspace{1cm} (6.63)

From (6.52) we find $H_d^* = -\frac{1}{2} \Theta^{-3/2} \text{diag} S_{LL}^{+,22}$ and

$$\text{vec } H_d^* = -\frac{1}{2} \Theta^{-3/2} E \text{vec } S_{LL}^{+,22}.$$  \hspace{1cm} (6.64)

**Theorem 6.** If the $Z_t$ vectors are independently normally distributed and the roots of (3.26) are distinct, $\text{vec } H_n^*$ and $\text{vec } H_d^*$ have a limiting normal distribution with means $0$ and $0$ and covariances

$$N^+[(I - R_2^2) \otimes (I - R_2^2)] + N^+ [R_2^{-1}(I - R_2^2)^{3/2} \otimes (I - R_2^2)](I + K)$$
$$\left\{ \Phi^+(\Theta \otimes \Theta) + (\Theta \otimes \Theta)\Phi^+ \right\} [R_2^{-1}(I - R_2^2)^{3/2} \otimes (I - R_2^2)] N^+$$

and

$$\frac{1}{2} \text{E} \left[ (\Theta^{-1} \otimes I) + (\theta^{-3/2} \otimes I) \Phi^+ (\Theta^{\frac{1}{2}} \otimes I) + (\Theta^{\frac{1}{2}} \otimes I) \Phi^+ (\Theta^{-3/2} \otimes I) \right] E,$$  \hspace{1cm} (6.65)

respectively.
From $G_{22} = Y^{22} \Sigma H_{22}$ we can transform Theorem 6 into the asymptotic covariances of $\text{vec } G_{22} = (I \otimes Y^{22} \Xi) \text{vec } H_{22}$.

References


