ASYMPTOTIC THEORY FOR CANONICAL CORRELATION ANALYSIS

by

T. W. Anderson

TECHNICAL REPORT NO. 520
November 1997

Approved for public release; distribution unlimited

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
Asymptotic Theory for Canonical Correlation Analysis

by

T. W. Anderson
Department of Statistics
Stanford University

TECHNICAL REPORT NO. 520
November 1997

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
Asymptotic Theory for Canonical Correlation Analysis

T. W. Anderson

Department of Statistics, Stanford University

November 12, 1997

Abstract

The asymptotic distribution of the sample canonical correlations and coefficients of the canonical variates are obtained when the nonzero population canonical correlations are distinct and sampling is from the normal distribution. The asymptotic distributions are also obtained for reduced rank regression when one set of variables is treated as independent (stochastic or nonstochastic) and the other set as dependent. Earlier work is corrected.

Key Words: Canonical variates, reduced rank regression, maximum likelihood estimators, test of rank.

1 Introduction

Hotelling (1936) proposed canonical correlations as invariant measures of relationships between two sets of variates. Suppose that the two random vectors \( \mathbf{Y} \) and \( \mathbf{X} \) of \( p \) and \( q \) components \( (p \leq q) \), respectively, have the covariance matrix

\[
\Sigma = \begin{pmatrix}
\Sigma_{YY} & \Sigma_{YX} \\
\Sigma_{XY} & \Sigma_{XX}
\end{pmatrix}
\]

(1.1)
with $\Sigma_{YY}$ and $\Sigma_{XX}$ nonsingular. The first canonical correlation, say $\rho_1$, is the maximum correlation between linear combinations $U = \alpha' Y$ and $V = \gamma' X$. The maximizing linear combinations normalized so $\text{var} U = \text{var} V = 1$ and denoted $U_1 = \alpha_1' Y$ and $V_1 = \gamma_1' X$ are the first canonical variates. The second canonical correlation $\rho_2$ is the maximum correlation between linear combinations $U = \alpha' Y$ and $V = \gamma' X$ uncorrelated with $U_1$ and $V_1$. The maximizing linear combinations, say $U_2 = \alpha_2' Y$ and $V_2 = \gamma_2' X$ are the second canonical variates. Other pairs, $(U_j, V_j)$, are defined similarly, $j = 3, \ldots, \min(p, q)$. If $p < q$, $V_{p+1} = \gamma_{p+1}' X$, ..., $V_q = \gamma_q' X$ are defined so $\text{var}(V_j) = 1$ and $\text{cov}(V_j, V_k) = 0$, $k = 1, \ldots, j - 1$, $j = p + 1, \ldots, q$. See Anderson (1984, Sec. 12.2) for more details.

From a sample $(y_1, x_1), \ldots, (y_N, x_N)$ the sample covariance matrices $S_{YY}$, $S_{YX}$, and $S_{XX}$ are calculated. Then the sample canonical correlations are the maximal sample correlations of linear combinations of the vectors $(\alpha'y_j, c'x_j)$, and the sample canonical variates are those maximizing linear combinations. The sample canonical correlations and the coefficients of the sample canonical variates are estimators of the population canonical correlations and coefficients. Except for a factor $(N - 1)/N$ these estimators are maximum likelihood estimators under the assumption that $(Y', X')'$ has a normal distribution.

The distribution of the canonical correlations under normality when $\Sigma_{YX} = 0$ (that is, $Y$ and $X$ uncorrelated) has a simple form (Hsu, 1939), but when $\Sigma_{YX} \neq 0$ the distribution involves zonal polynomials (Constantine, 1963) and is of limited use. Hsu (1941) found the asymptotic distribution of the canonical correlations including models in which the population canonical correlations were of arbitrary multiplicities. Lawley (1956) treated the asymptotic expansion of the distribution of the correlations.

The coefficients of the canonical vectors have received less attention. Asymptotic distributions
of the coefficients have been obtained, but they have pertained to modification of the canonical vectors, the derivations are incomplete, or they are incorrect. One purpose of this paper is to obtain completely the asymptotic distribution of the correlations and coefficients as usually defined. The relation to other contributions is discussed in Section 6.

If the vector \((Y', X')'\) has a normal distribution with \(EY = \mu_Y\), \(EX = \mu_X\), the conditional distribution of \(Y\) given \(X = x\) is normal and can be represented as

\[
Y = \mu_Y + B(x - \mu_X) + Z, \tag{1.2}
\]

where

\[
B = \Sigma_{YX} \Sigma_{XX}^{-1}, \tag{1.3}
\]

and the unobservable vector

\[
Z = Y - \mu_Y - B(X - \mu_X) \tag{1.4}
\]

has mean 0 and covariance matrix

\[
\Sigma_{ZZ} = \Sigma_{YY} - B \Sigma_{XX} B' = \Sigma_{YY} - \Sigma_{YX} \Sigma_{XX}^{-1} \Sigma_{XY}, \tag{1.5}
\]

and is normally distributed. The variance of a linear combination \(\phi'Y\) of the dependent variable \(Y\) can be decomposed into the “effect” variance due to the regression on \(X\) and the “error” variance due to \(\phi'Z\). (Note \(E XZ' = 0\)). The linear combination, say \(\phi_1' Y\), maximizing the ratio of the effect variance to the error variance is a multiple of \(U_1 = \alpha_1' Y\). Similarly, the linear combination, say \(\phi_2' Y\), maximizing this ratio among linear combination uncorrelated with \(U_1\) is a multiple of \(U_2 = \alpha_2' Y\), etc. The vectors \(\phi_1, \ldots, \phi_p\) are the characteristic vectors of \(B \Sigma_{XX} B'\) in the metric of \(\Sigma_{ZZ}\), and the characteristic roots are \(\theta_1 = \rho_1^2/(1 - \rho_1^2), \ldots, \theta_p = \rho_p^2/(1 - \rho_p^2)\). The asymptotic theory for these linear combinations and ratios is also obtained and compared with earlier work.
In “reduced rank regression” the matrix $B$ may be restricted to a rank $r$, say, less than its numbers of rows and columns. The maximum likelihood estimator of $B$ then involves the estimators of the vectors $\phi_1, ..., \phi_r$ and roots $\theta_1, ..., \theta_r$ or equivalently $\alpha_1, ..., \alpha_r, \gamma_1, ..., \gamma_r, \rho_1, ..., \rho_r$ (Anderson 1951a). Reduced rank regression has been used to estimate cointegrating relations in some nonstationary models (Johansen, 1988).

In econometric models the interest is in the linear restrictions that hold on $B$; these can be represented by the vectors $\phi_{r+1}, ..., \phi_p$ associated with the $p - r$ zero characteristic roots and are estimated by the sample vectors associated with the $p - r$ smallest roots. In stationary econometric models, the estimation of a single restriction is known as the Limited Information Maximum Likelihood method (Anderson and Rubin, 1950). Anderson (1951b) found the asymptotic distribution of the estimators of $\phi_1, ..., \phi_p$, $\theta_1, ..., \theta_p$ for the model (1.2) when the observations $x_1, ..., x_N$ are nonstochastic and $Z$ has a normal distribution. In this paper $X$ is a normally distributed random vector, leading to different results.

Robinson (1973) claimed to have found the asymptotic distribution of these estimators when $X$ and $Z$ are random with arbitrary distributions, but his results are incorrect. Izenman (1975) treated the asymptotic distribution of the estimators of $\alpha_1, ..., \alpha_p, \gamma_1, ..., \gamma_q$ and $\rho_1, ..., \rho_p$. However, he did not succeed in obtaining an explicit expression for the asymptotic covariances of the estimators. Velu, Reinsel, and Wichern (1986) have further studied reduced rank regression. Brillinger (1975) found the asymptotic distribution of vectors related to the canonical variates, but is not correct.

In Sections 2 and 3 $Y$ and $X$ are treated as a pair of random vectors having a joint distribution; the asymptotic distribution of the canonical correlations and canonical vectors is found. In Section 4 the vector $Y$ is treated as a dependent vector generated by (1.2), where $X$ is the independent
(random) vector; the asymptotic distribution of the characteristic roots and vectors of $\mathbf{B}\Sigma_{XX}\mathbf{B}'$
in the metric of $\Sigma_{ZZ}$ is obtained. In Section 5 each of these problems is studied in the context
of $\mathbf{X}$ being a nonstochastic vector. Section 6 comments on the relations between the results in
this paper with results in other papers. Section 7 treats the more general models admitting zero
canonical correlations.

2 Maximum likelihood estimation

The equations defining the canonical correlations and variates (in the population) are

$$
\begin{pmatrix}
-\rho \Sigma_{YY} & \Sigma_{YX} \\
\Sigma_{XY} & -\rho \Sigma_{XX}
\end{pmatrix}
\begin{pmatrix}
\alpha \\
\gamma
\end{pmatrix}
= 0,
$$

(2.1)

where $\rho$ satisfies

$$
\begin{vmatrix}
-\rho \Sigma_{YY} & \Sigma_{YX} \\
\Sigma_{XY} & -\rho \Sigma_{XX}
\end{vmatrix}
= 0,
$$

(2.2)

and

$$
\alpha' \Sigma_{YY} \alpha = 1, \quad \gamma' \Sigma_{XX} \gamma = 1.
$$

(2.3)

The number of nonzero canonical correlations is the rank of $\Sigma_{XX}$. The canonical correlations are
ordered $\rho_1 \geq \cdots \geq \rho_p \geq -\rho_p \geq \cdots \geq -\rho_1$ with $q - p$ additional roots of 0. To eliminate the
indeterminacy of a solution $\alpha, \gamma$ in (2.1) we require $\alpha_{ii} > 0$. (Since the matrix $\mathbf{A} = (\alpha_1, \ldots, \alpha_p)$ is
nonsingular, the components of $\mathbf{Y}$ can be numbered in such a way that the $i$th component of $\alpha_i$
is nonzero.)

From (2.1) we obtain $\gamma = (1/\rho) \Sigma_{XX}^{-1} \Sigma_{XY} \alpha$, $\alpha = (1/\rho) \Sigma_{YY}^{-1} \Sigma_{YX} \gamma$,

$$
\rho^2 \Sigma_{YY} \alpha = \Sigma_{YX} \Sigma_{XX}^{-1} \Sigma_{XY} \alpha = \mathbf{B} \Sigma_{XX} \mathbf{B}' \alpha,
$$

(2.4)

$$
\rho^2 \Sigma_{XX} \gamma = \Sigma_{YX} \Sigma_{YY}^{-1} \Sigma_{XY} \gamma.
$$

(2.5)
The solutions of (2.1) corresponding to $\rho_1, \ldots, \rho_p$ can be assembled as $A = (\alpha_1, \ldots, \alpha_p)$ and $(\gamma_1, \ldots, \gamma_p)$. If $p > q$, there are $p - q$ additional solutions $(\gamma_{p+1}, \ldots, \gamma_q)$ with $\alpha = 0$. Let $\Gamma = (\gamma_1, \ldots, \gamma_q)$ and let $R = \text{diag}(\rho_1^2, \ldots, \rho_p^2), \bar{R} = (R, 0)$. Then the solutions can be chosen to satisfy

\begin{equation}
\begin{pmatrix}
A' & 0 \\
0 & \Gamma'
\end{pmatrix}
\begin{pmatrix}
\Sigma_{YY} & \Sigma_{YX} \\
\Sigma_{XY} & \Sigma_{XX}
\end{pmatrix}
\begin{pmatrix}
A \\
0
\end{pmatrix}
\begin{pmatrix}
\Gamma
\end{pmatrix}
\end{equation}

\begin{equation}
= \begin{pmatrix}
A'\Sigma_{YY}A & A'\Sigma_{YX}\Gamma' \\
\Gamma'\Sigma_{XY}A & \Gamma'\Sigma_{XX}\Gamma
\end{pmatrix} = \begin{pmatrix}
I & \bar{R} \\
\bar{R}' & I
\end{pmatrix}.
\end{equation}

This is the covariance matrix of the canonical variates $U = A'Y$ and $V = \Gamma'Y$.

If the sample is from a normal distribution $N(\mu, \Sigma)$ with $\mu = (\mu'_y, \mu'_x)'$, the maximum likelihood estimators of the means and covariances are

\begin{equation}
\hat{\mu} = \left(\begin{array}{c}
\hat{\mu}_y \\
\hat{\mu}_x
\end{array}\right) = \left(\begin{array}{c}
\bar{y} \\
\bar{x}
\end{array}\right) = \frac{1}{N} \sum_{\alpha=1}^N \left(\begin{array}{c}
y_\alpha \\
x_\alpha
\end{array}\right),
\end{equation}

\begin{equation}
\hat{\Sigma} = \frac{1}{N} \sum_{\alpha=1}^N \left(\begin{array}{c}
y_\alpha - \bar{y} \\
x_\alpha - \bar{x}
\end{array}\right) \left(\begin{array}{c}
y'_\alpha - \bar{y}' \\
x'_\alpha - \bar{x}'
\end{array}\right) = \begin{pmatrix}
\hat{\Sigma}_{YY} & \hat{\Sigma}_{YX} \\
\hat{\Sigma}_{XY} & \hat{\Sigma}_{XX}
\end{pmatrix} = \frac{1}{n} S,
\end{equation}

where $n = N - 1$. The statistics $\hat{\mu}$ and $\hat{\Sigma}$ are sufficient; $S$ is an unbiased estimator of $\Sigma$.

The sample equations corresponding to (2.1) and (2.3) defining the population canonical correlations and variates are

\begin{equation}
\begin{pmatrix}
-S_{YY} & S_{YX} \\
S_{XY} & -S_{XX}
\end{pmatrix}
\begin{pmatrix}
a \\
c
\end{pmatrix} = 0,
\end{equation}

\begin{equation}
a'S_{YY}a = 1, \quad c'S_{XX}c = 1.
\end{equation}
The solutions with $a_{ii} > 0$, $i = 1, ..., p$, and $r_1 > r_2 > ... > r_p > 0$ define the estimators 
\[ \hat{A} = (a_1, ..., a_p), \quad \hat{c} = (c_1, ..., c_q), \quad \hat{R} = \text{diag}(r_1, ..., r_p). \]
These are uniquely defined except that if $q > p$, $c_{p+1}, ..., c_q$ satisfy $c'S_{XX}c = 1$, $c'S_{XX}c_j = 0$, $j = 1, ..., p$, and some other $(q-p)(q-p-1)$ arbitrary conditions. From (2.9) and (2.10) we obtain 
\[ c = (1/r)S_{XX}^{-1}S_{XY}a, \quad a = (1/r)S_{YY}^{-1}S_{YY}c \]
\[ S_{YY}S_{XX}^{-1}S_{XY}a = r^2S_{YY}a, \quad \tag{2.11} \]
\[ S_{XY}S_{YY}^{-1}S_{YY}c = r^2S_{XX}c. \quad \tag{2.12} \]

The equation (2.4) can be written

\[ \Sigma_{YY}\Sigma_{XX}^{-1}\Sigma_{XY} \alpha = \frac{\rho^2}{1 - \rho^2} \left( \Sigma_{YY} - \Sigma_{YY}\Sigma_{XX}^{-1}\Sigma_{XY} \right) \alpha = \theta \Sigma_{ZZ} \alpha, \quad \tag{2.13} \]
where $\theta = \rho^2/(1 - \rho^2)$. Let $\phi = (1/\sqrt{1 - \rho^2})\alpha$. Then

\[ B \Sigma_{XX}B' \phi = \theta \Sigma_{ZZ} \phi, \quad \phi' \Sigma_{ZZ} \phi = 1, \quad \tag{2.14} \]
and $|B \Sigma_{XX}B' - \theta \Sigma_{ZZ}| = 0$. It may be convenient to treat $\phi_1, \theta_1, ..., \phi_p, \theta_p$ based on the model (1.2).

The density of $(Y', X')'$, can be written

\[ n \left[ \begin{bmatrix} (y) \\ (x) \end{bmatrix} \right] \left( \begin{bmatrix} \mu_Y \\ \mu_X \end{bmatrix} \begin{bmatrix} \Sigma_{YY} & \Sigma_{XY} \\ \Sigma_{XY} & \Sigma_{XX} \end{bmatrix} \right)^{-1} \]
\[ = n(y|\mu_Y + B(x - \mu_X), \Sigma_{ZZ}) \cdot n(x|\mu_X, \Sigma_{XX}). \quad \tag{2.15} \]

The maximum likelihood estimators of $B$ and $\Sigma_{ZZ}$ are determined from the first factor on the right-hand side of (2.15); they depend on normality of $z_\alpha$ independent of $x_\alpha, x_{\alpha-1}, z_{\alpha-1}, ..., x_1, z_1$. In econometric terms $x_\alpha$ is "predetermined"; components of $x_\alpha$ can be functions of $y_{a-1}, x_{a-1}, ..., y_1, x_1$, for example, lagged values of the dependent vector as in time series analysis. This was pointed out
by Anderson (1951a). The estimators are

\[ \hat{B} = \hat{\Sigma}_{YY} \hat{\Sigma}_{XX}^{-1} = S_{YY} S_{XX}^{-1}, \]

\[ \frac{n}{N} S_{ZZ} = \hat{\Sigma}_{ZZ} = \hat{\Sigma}_{YY} - \hat{B} \hat{\Sigma}_{XX} B' = \frac{n}{N} \left( S_{YY} - S_{XX} S_{XX}^{-1} S_{XY} \right). \]

The sample analogs of (2.14) are

\[ \hat{B} S_{XX} \hat{B}' f = t S_{ZZ} f, \quad f' S_{ZZ} f = 1. \]

The solutions are \( t_1 = r_1^2/(1 - r_1^2) > ... > t_p = r_p^2/(1 - r_p^2) \) and \( f_1 = a_1(1 - r_1^2)^{-\frac{1}{2}}, \ldots, f_p = a_p(1 - r_p^2)^{-\frac{1}{2}} \), with \( f_i > 0, \ i = 1, \ldots, p \).

In some problems it may be required that \( \Sigma_{XY} \) and hence \( B \) have rank not greater than a specified number, say \( r \). Anderson (1951a) showed that the maximum likelihood estimator of \( B \) of rank \( r \) is \( \hat{B}_r = S_{ZZ} \hat{\Phi}_1 \hat{\Phi}_1' \hat{B} \); alternative forms are

\[ \hat{B}_r = S_{ZZ} \hat{\Phi}_1 \hat{\Phi}_1' \hat{B} = S_{ZZ} \hat{A}_1 R_1 \left( 1 - R_1^2 \right)^{-1} \hat{A}_1 \hat{\Gamma}_1 = S_{XX} \bar{A}_1 \bar{\Gamma}_1, \]

where \( \hat{\Phi}_1 = (\hat{\phi}_1, \ldots, \hat{\phi}_r) \), \( \hat{A}_1 = (a_1, \ldots, a_r) \), \( \hat{\Gamma}_1 = (\hat{\gamma}_1, \ldots, \hat{\gamma}_r) \) and \( R_1 = \text{diag}(r_1, \ldots, r_r) \). This result has been repeated by Hannan (1967), Chow (1967), and Johansen (1988). Robinson (1973), Izenman (1975) and Velu, Reinsel, and Wichern (1986) have found the same estimator by minimizing \( \text{tr} (S_{YY} - EFS_{XY} - S_{XX} F'E' + EFS_{XX} F'E') S_{ZZ}^{-1} \) with respect to \( E \) and \( F \) (of orders \( p \times r \) and \( r \times p \) respectively).

Anderson (1951a) actually treated a more general case. Partition \( X, B \) and \( \Sigma_{YY} \) into \( q_1 \) and \( q_2 \) columns: \( X = (X^{(1)'}, X^{(2)'})' \), \( B = (B_1, B_2) \), \( \Sigma_{YY} = (\Sigma_{Y1}, \Sigma_{Y2}) \), and

\[ S_{XX} = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}. \]
Suppose $B_2(p \times q_2)$ is specified to have rank $r(< q_2)$. Then the maximum likelihood estimator of $B_2$ has the form (2.17) with $\hat{B}$ replaced by $\hat{B}_2$ defined by $\hat{B} = (\hat{B}_1, \hat{B}_2)$ and $\hat{\phi}_j$ replaced by the solution of

$$
\hat{B}_2(S_{22} - S_1S_{11}^{-1}S_{12})\hat{B}_2'\phi' = t S_{ZZ}\phi, \quad \phi'S_{ZZZZ}\phi = 1
$$

(2.21)

for the $r$ largest values of $\theta$.

Another question of interest is to test the null hypothesis that the rank of $B$ (or $B_2$) is $r$. Anderson (1951a) found the likelihood ratio criterion to be

$$
-2 \log \lambda = -N \sum_{i=r+1}^{p} \log \left(1 - \frac{r_i^2}{N}\right).
$$

(2.22)

The covariance matrices $S_{YY}$, $S_{YX}$, $S_{XX}$ under normality have the same distribution as

$$
(1/n) \sum_{\alpha=1}^{n} y_{\alpha}y'_{\alpha}, \quad (1/n) \sum_{\alpha=1}^{n} y_{\alpha}x'_{\alpha}, \quad (1/n) \sum_{\alpha=1}^{n} x_{\alpha}x'_{\alpha}
$$

which are the maximum likelihood estimators when $\varepsilon y_{\alpha} = 0$, $\varepsilon x_{\alpha} = 0$. Since the focus of the paper is on asymptotic theory, we shall henceforth assume $\varepsilon y_{\alpha} = 0$, $\varepsilon x_{\alpha} = 0$ and define $S_{YY} = (1/n) \sum_{\alpha=1}^{n} y_{\alpha}y'_{\alpha}$, etc.

3 Sample canonical variables

The sample covariances of $U = A'Y$ and $V = \Gamma'X$ are

$$
S_{UU} = A'S_{YY}A, \quad S_{UV} = A'S_{YX}\Gamma, \quad S_{VV} = \Gamma'S_{XX}\Gamma.
$$

(3.1)

The sample canonical variates in terms of the population canonical variates satisfy

$$
\begin{pmatrix}
-rS_{UU} & S_{UV} \\
S_{UV} & -rS_{VV}
\end{pmatrix}
\begin{pmatrix}
g \\
h
\end{pmatrix}
= 0,
$$

(3.2)

$$
g'S_{UU}g = 1, \quad h'S_{VV}h = 1.
$$

(3.3)

Let $G = (g_1, \ldots, g_p)$, $H = (h_1, \ldots, h_q)$. Then

$$
S_{UV}H = S_{UU}G(\bar{R}, 0),
$$

(3.4)
\[ S_{VV} G = S_{VV} H \left( \begin{array}{c} \hat{R} \\ 0 \end{array} \right), \]  

(3.5)

where \( \hat{R} = \text{diag}(r_1, ..., r_p) \). Normalization leads to

\[ G' S_{UU} G = I, \quad H' S_{VV} H = I. \]  

(3.6)

Since \( S_{UU} \xrightarrow{p} \Sigma_{UU} = I \), \( S_{VV} \xrightarrow{p} \Sigma_{VV} = I \), and \( S_{UV} \xrightarrow{p} \Sigma_{UV} = \hat{R} \), then \( G \xrightarrow{p} I \) if \( \rho_1 > ... > \rho_p \)

and \( g_{ii} > 0 \), and

\[ H \xrightarrow{p} \left( \begin{array}{cc} I & 0 \\ 0 & H_{22} \end{array} \right), \]  

(3.7)

and \( H_{22} \) is orthogonal. To avoid further complications we shall assume \( p = q \).

If the fourth-order moments of \( Y \) and \( X \) are bounded, then the matrices \( \sqrt{n}(S_{YY} - \Sigma_{YY}) \), \( \sqrt{n}(S_{XY} - \Sigma_{XY}) \), and \( \sqrt{n}(S_{XX} - \Sigma_{XX}) \) have a limiting normal distribution with means zero, and so do \( \sqrt{n}(S_{UU} - I) \), \( \sqrt{n}(S_{UV} - R) \) and \( \sqrt{n}(S_{VV} - I) \). The covariances of these normal distributions are linear functions of the fourth-order moments of \( Y \) and \( X \) or \( U \) and \( V \).

Since the set of components of \( G, H, \) and \( R \) is a single-valued, invertible, differentiable function of the components of \( S_{UU}, S_{VV}, \) and \( S_{UV} \) in the neighborhood of \( I, I, \) and \( R \), the set of components of

\[ G^* = \sqrt{n}(G - I), \quad H^* = \sqrt{n}(H - I), \quad R^* = \sqrt{n}(\hat{R} - R) \]  

(3.8)

has a limiting normal distribution. The variances and covariances of the limiting normal distribution can be found by a Taylor's expansion of \( (G, H, \hat{R}) \) around the point \( (S_{UU}, S_{UV}, S_{VV}) = (I, R, I) \). We shall calculate this linear transformation as the inverse of the expansion of \( (S_{UU}, S_{UV}, S_{VV}) \) around the point \( (G, H, \hat{R}) = (I, I, R) \). We substitute in (3.4) and (3.5) for \( G, H \) and \( R \) from (3.8) and retain the terms linear in \( G^*, H^*, \) and \( R^* \). Note that this procedure amounts to finding the derivatives of \( G^*, H^*, \) and \( R^* \) with respect to the components of \( (S_{UU}, S_{VV}, S_{UV}) \).
After rearrangement of terms, we obtain

\[
\sqrt{n} \left[ (S_{UU} - R) - (S_{UU} - I)R \right] = G^*R + R^* - RH^* + o_p(1), \tag{3.9}
\]

\[
\sqrt{n} \left[ (S_{VV} - R) - (S_{VV} - I)R \right] = -RG^* + H^*R + R^* + o_p(1). \tag{3.10}
\]

We solve (3.9) and (3.10) for the nondiagonal elements of $G^*$ and $H^*$ and the diagonal elements of $R^*$. The sum of (3.9) multiplied by $R$ on the right and (3.10) multiplied by $R$ on the left and the sum of (3.9) multiplied by $R$ on the left and (3.10) multiplied by $R$ on the right are

\[
\sqrt{n} \left[ (S_{UU} - R)R + R(S_{VV} - R) - R(S_{VV} - I)R - (S_{UU} - I)R^2 \right] \tag{3.11}
\]

\[= 2RR^* + G^*R^2 - R^2G^* + o_p(1), \]

\[
\sqrt{n} \left[ R(S_{UU} - R) + (S_{UU} - R)R - R(S_{UU} - I)R - (S_{VV} - R)R^2 \right] \tag{3.12}
\]

\[= 2RR^* + H^*R^2 - R^2H^* + o_p(1). \]

[These equations can also be obtained from (2.4) and (2.5)]. When $G = I + (1/\sqrt{n})G^*$ and $H = I + (1/\sqrt{n})H^*$ are substituted into (3.6), we obtain

\[
\sqrt{n} (S_{UU} - I) = -(G^{*'} + G^*) + o_p(1), \tag{3.13}
\]

\[
\sqrt{n} (S_{VV} - I) = -(H^{*'} + H^*) + o_p(1). \tag{3.14}
\]

The $i$th diagonal term in each of (3.11) and (3.12) is $2\rho_i r_i^* + o_p(1), \ i = 1, ..., p$. The $i, j$th term of (3.11) is $g_{ij}^* (\rho_j^2 - \rho_i^2) + o_p(1)$ and the $i, j$th term of (3.12) is $h_{ij}^* (\rho_j^2 - \rho_i^2) + o_p(1), \ i \neq j, i, j = 1, ..., p$. The $i$th diagonal term of $\sqrt{n}(S_{UU} - I)$ is $-2g_{ii}^* + o_p(1)$ and the $i$th diagonal term of $\sqrt{n}(S_{VV} - I)$ is $-2h_{ii}^* + o_p(1)$.
is $-2h_{ii}^* + o_p(1)$. From now on we shall suppose that $\rho_1, \ldots, \rho_p$ are distinct and positive. Then the equations (3.11), (3.12), (3.13), and (3.14) can be solved for the elements of $\mathbf{R}^*, \mathbf{G}^*$, and $\mathbf{H}^*$.

The covariances of these components are linear functions of the fourth-order moments of the components of $\mathbf{Y}$ and $\mathbf{X}$. We shall now assume that $\mathbf{Y}$ and $\mathbf{X}$ are normal in order to compute the covariances of the sample covariances, which are quadratic in the elements of $\Sigma$. We use the fact that the 2-vectors $(U_1, V_1), \ldots, (U_p, V_p)$ are uncorrelated and hence independent.

From (3.11) and (3.12) we obtain

$$
\frac{1}{\sqrt{n}} \sum_{\alpha=1}^{n} \left( -\rho_i \rho_j v_{\alpha a} v_{ja} + \rho_j u_{\alpha a} v_{ja} + \rho_i v_{\alpha a} u_{ja} - \rho_j^2 u_{\alpha a} u_{ja} \right) \tag{3.15}
$$

$$
= \frac{1}{\sqrt{n}} \sum_{\alpha=1}^{n} \left[ - (u_{\alpha a} - \rho_i u_{\alpha a})(u_{ja} - \rho_j v_{ja}) + (1 - \rho_j^2) u_{\alpha a} u_{ja} \right] \nonumber
$$

$$
= 2 \rho_i r_i^* \delta_{ij} + g_{ij}^*(\rho_j^2 - \rho_i^2) + o_p(1), \nonumber
$$

$$
\frac{1}{\sqrt{n}} \sum_{\alpha=1}^{n} \left( -\rho_i \rho_j u_{\alpha a} u_{ja} + \rho_i u_{\alpha a} v_{ja} + \rho_j v_{\alpha a} u_{ja} - \rho_j^2 v_{\alpha a} v_{ja} \right) \tag{3.16}
$$

$$
= \frac{1}{\sqrt{n}} \sum_{\alpha=1}^{n} \left[ - (u_{\alpha a} - \rho_i u_{\alpha a})(v_{ja} - \rho_j u_{ja}) + (1 - \rho_j^2) v_{\alpha a} v_{ja} \right] \nonumber
$$

$$
= 2 \rho_i r_i^* \delta_{ij} + h_{ij}^*(\rho_j^2 - \rho_i^2) + o_p(1). \nonumber
$$

The $np$ pairs $(u_{\alpha a}, v_{\alpha a})$ are uncorrelated and independent. The covariances of the left-hand sides can be calculated from the following covariance matrix of $u_{\alpha a}, v_{\alpha a}, u_{\alpha a} - \rho_i u_{\alpha a}, v_{\alpha a} - \rho_i v_{\alpha a}$

$$
\begin{bmatrix}
1 & \rho_i & 1 - \rho_i^2 & 0 \\
\rho_i & 1 & 0 & 1 - \rho_i^2 \\
1 - \rho_i^2 & 0 & 1 - \rho_i^2 & -\rho_i(1 - \rho_i^2) \\
0 & 1 - \rho_i^2 & -\rho_i(1 - \rho_i^2) & 1 - \rho_i^2
\end{bmatrix}. \tag{3.17}
$$

The quadruples for $i \neq j$ ($g_{ij}^*, h_{ij}^*$, $h_{ji}^*$, $h_{jj}^*$) are asymptotically uncorrelated. The asymptotic covari-
The asymptotic covariance matrix of \((\rho_j^2 - \rho_i^2)g_{ij}^*\) and \((\rho_i^2 - \rho_j^2)g_{ji}^*\) is
\[
\begin{bmatrix}
(1 - \rho_j^2) \left( \rho_i^2 + \rho_j^2 - 2\rho_i^2 \rho_j^2 \right) & (1 - \rho_i^2) \left( \rho_i^2 + \rho_j^2 \right) \\
(1 - \rho_i^2) \left( 1 - \rho_j^2 \right) \left( \rho_i^2 + \rho_j^2 \right) & (1 - \rho_i^2) \left( \rho_i^2 + \rho_j^2 - 2\rho_i^2 \rho_j^2 \right)
\end{bmatrix}
\]
(3.18)
which is also the asymptotic covariance matrix for \((\rho_j^2 - \rho_i^2)h_{ij}^*\) and \((\rho_i^2 - \rho_j^2)h_{ji}^*\). The asymptotic covariance matrix of the row vector \([\(\rho_j^2 - \rho_i^2\)g_{ij}^*, (\rho_i^2 - \rho_j^2)g_{ji}^*]\) and the column vector \([\(\rho_j^2 - \\
\rho_i^2\)h_{ij}^*, (\rho_i^2 - \rho_j^2)h_{ji}^*]\) is
\[
\begin{bmatrix}
\rho_i \rho_j \left( 2 - \rho_i^2 - 3\rho_j^2 + 2\rho_i^2 \rho_j^2 \right) & (1 - \rho_i^2) \left( 1 - \rho_j^2 \right) \left( \rho_i^2 + \rho_j^2 \right) \\
(1 - \rho_i^2) \left( 1 - \rho_j^2 \right) \left( \rho_i^2 + \rho_j^2 \right) & \rho_i \rho_j \left( 2 - \rho_i^2 - 3\rho_j^2 + 2\rho_i^2 \rho_j^2 \right)
\end{bmatrix}
\]
(3.19)

From (3.11) and/or (3.12) we obtain
\[
-2\rho_i r_i^* = \frac{1}{\sqrt{n}} \sum_{a=1}^{n} \left( \rho_i^2 v_{ia}^2 - 2\rho_i v_{ia} u_{ia} + \rho_i^2 u_{ia}^2 \right) + o_p(1)
\]
(3.20)
\[
= \frac{1}{\sqrt{n}} \sum_{a=1}^{n} \left[ (u_{ia} - \rho_i v_{ia})^2 - (1 - \rho_i^2) u_{ia}^2 \right] + o_p(1)
\]

From (3.13) and (3.14) we obtain
\[
2g_{ii}^* = - \frac{1}{\sqrt{n}} \sum_{a=1}^{n} \left( u_{ia}^2 - 1 \right) + o_p(1)
\]
(3.21)
\[
2h_{ii}^* = - \frac{1}{\sqrt{n}} \sum_{a=1}^{n} \left( v_{ia}^2 - 1 \right) + o_p(1)
\]

The asymptotic covariance matrix of \(r_i^*, g_{ii}^*, h_{ii}^*\) is
\[
\begin{bmatrix}
(1 - \rho_i^2)^2 & -\frac{1}{2}\rho_i \left( 1 - \rho_i^2 \right) & -\frac{1}{2}\rho_i \left( 1 - \rho_i^2 \right) \\
\frac{1}{2}\rho_i \left( 1 - \rho_i^2 \right) & \frac{1}{2} & \frac{1}{2} \rho_i^2 \\
\frac{1}{2}\rho_i \left( 1 - \rho_i^2 \right) & \frac{1}{2} \rho_i^2 & \frac{1}{2}
\end{bmatrix}
\]
(3.22)
To summarize: the variables \((r_i^*, g_{ii}^*, h_{ii}^*), i = 1, ..., p)\) and \((g_{ij}^*, g_{ji}^*, h_{ij}^*, h_{ji}^*), i < j, i, j = 1, ..., p)\) have a limiting normal distribution, the covariance matrix of which has blocks (3.18), (3.19), and (3.22) with the rest of the entries being zeros.
A solution to (3.2) and (3.3) is related to the solution to (2.18) and (2.19) by

\[ a = A g, \quad c = \Gamma h. \]  

(3.23)

Thus

\[
\sqrt{n} \left( a_j - \alpha_j \right) = \sqrt{n} \ A (g_j - \varepsilon_j) = A \left( g_{1j}^*, ..., g_{pj}^* \right)' + o_p(1),
\]

\[
= \sum_{k=1}^{p} \alpha_k g_{kj}^* + o_p(1) = \alpha_j g_{jj}^* + \sum_{k=1 \atop k \neq j}^{p} \alpha_k g_{kj}^* + o_p(1),
\]

(3.24)

\[
\sqrt{n} (c_j - \gamma_j) = \sqrt{n} \Gamma (h_j - \varepsilon_j) = \Gamma \left( h_{1j}^*, ..., h_{pj}^* \right)' + o_p(1)
\]

\[
= \sum_{k=1}^{p} \gamma_k h_{kj}^* + o_p(1) = \gamma_j h_{jj}^* + \sum_{k=1 \atop k \neq j}^{p} \gamma_k h_{kj}^* + o_p(1),
\]

where \( \varepsilon_j \) is the \( j \)th column of \( I \). The asymptotic variances and covariances are

\[
n \mathcal{E} \left( a_j - \alpha_j \right) (a_j - \alpha_j)' = \frac{1}{2} \alpha_j \alpha_j' + \left( 1 - \rho_j^2 \right) \sum_{k=1 \atop k \neq j}^{p} \frac{\left( \rho_j^2 + \rho_k^2 - 2 \rho_j \rho_k \rho_{jk}^* \right)}{\left( \rho_j^2 - \rho_k^2 \right)^2} \alpha_k \alpha_k',
\]

(3.25)

\[
n \mathcal{E} \left( c_j - \gamma_j \right) (c_j - \gamma_j)' = \frac{1}{2} \gamma_j \gamma_j' + \left( 1 - \rho_j^2 \right) \sum_{k=1 \atop k \neq j}^{p} \frac{\left( \rho_j^2 + \rho_k^2 - 2 \rho_j \rho_k \rho_{jk}^* \right)}{\left( \rho_j^2 - \rho_k^2 \right)^2} \gamma_k \gamma_k',
\]

(3.26)

\[
n \mathcal{E} \left( a_j - \alpha_j \right) (a_\ell - \alpha_\ell)' = - \frac{\left( 1 - \rho_j^2 \right) \left( 1 - \rho_\ell^2 \right) \left( \rho_j^2 + \rho_\ell^2 \right)}{\left( \rho_j^2 - \rho_\ell^2 \right)^2} \alpha_j \alpha_\ell', \quad j \neq \ell,
\]

(3.27)

\[
n \mathcal{E} \left( c_j - \gamma_j \right) (c_\ell - \gamma_\ell)' = - \frac{\left( 1 - \rho_j^2 \right) \left( 1 - \rho_\ell^2 \right) \left( \rho_j^2 + \rho_\ell^2 \right)}{\left( \rho_j^2 - \rho_\ell^2 \right)^2} \gamma_j \gamma_\ell', \quad j \neq \ell,
\]

(3.28)

\[
n \mathcal{E} \left( a_j - \alpha_j \right) (c_\ell - \gamma_\ell)' = - \frac{\left( 1 - \rho_j^2 \right) \left( 1 - \rho_\ell^2 \right) \left( \rho_j^2 + \rho_\ell^2 \right)}{\left( \rho_j^2 - \rho_\ell^2 \right)^2} \gamma_j \alpha_\ell',
\]

(3.29)
Note that the contribution of \( \alpha_k \alpha'_k \) to the asymptotic covariance matrix of \( a_j \) depends on \( \rho_j^2 \) and \( \rho_k^2 \). The contribution is small if \( \rho_j^2 - \rho_k^2 \) is numerically large or if \( \rho_j^2 \) is large.

To obtain the asymptotic covariances of \( g_{ij}^* \) and \( h_{ii}^* \), \( i \neq j \), we have used

\[
\mathcal{E} \xi_i \xi_j \eta_i \eta_j = \mathcal{E} \xi_i \eta_i \mathcal{E} \xi_j \eta_j
\]

(3.30)

when \( \mathcal{E} \xi_i \xi_j = \mathcal{E} \xi_i \eta_j = \mathcal{E} \eta_i \eta_j = 0 \) and \( \mathcal{E} \xi_i = \mathcal{E} \xi_j = \mathcal{E} \eta_i = \mathcal{E} \eta_j = 0 \). That is, that zero correlation of \((\xi_i, \eta_i)\) and \((\xi_j, \eta_j)\) implies factoring of fourth-order moments. This is the case when \( \xi_i, \eta_i, \xi_j, \eta_j \) have a joint normal distribution. In obtaining the asymptotic covariances of \( r_i^*, g_{ii}^*, \) and \( h_{ii}^* \) we have used the more general

\[
\mathcal{E} \eta_1 \eta_2 \eta_3 \eta_4 = \mathcal{E} \eta_1 \eta_2 \mathcal{E} \eta_3 \eta_4 + \mathcal{E} \eta_1 \eta_3 \mathcal{E} \eta_2 \eta_4 + \mathcal{E} \eta_1 \eta_4 \mathcal{E} \eta_2 \eta_3.
\]

(3.31)

These relations between fourth-order and second-order moments hold for normally distributed variables. If normality is not assumed, the asymptotic covariances would have to include more general fourth-order moments which would not necessarily depend only on \( \Sigma \).

4 Asymptotic distribution in a regression model

Consider the equations (2.16). In terms of the canonical variates \( U = A'Y \) and \( V = \Gamma'X \) (2.16) becomes \((Ak = f)\)

\[
S_{UV} S_{VV}^{-1} S_{VU} k = t S_{WW} k, \quad k' S_{WW} k = 1,
\]

(4.1)

where

\[
S_{WW} = S_{UU} - S_{UV} S_{VV}^{-1} S_{VU}.
\]

(4.2)

For a solution to (4.1) \( t \) has to satisfy

\[
0 = |S_{UV} S_{VV}^{-1} S_{VU} - t S_{WW}|
\]

(4.3)
\[(1 + t)S_{UV} S_{VV}^{-1} S_{VU} - t S_{UV}\].

Thus \(t/(1 + t) = r^2\) or \(t = r^2/(1 - r^2)\). Then (4.1) becomes

\[S_{UV} S_{VV}^{-1} S_{VU} k = r^2 S_{UU} k, \tag{4.4}\]

which is the equation for \(g\) solving (3.2). Hence a solution \(k_i\) to (4.4) is proportional to a solution \(g_i\) to (3.2) after eliminating \(h\). The normalization of \(g_i' S_{UU} g_i = 1\) implies \(g_i S_{WW} g_i = 1 - r_i^2\). Hence \(k_i = (1 - r_i^2)^{-1/2} g_i\), and \(k_i)^P = (1 - r_i^2)^{-1/2} \epsilon_i\).

Define \(K = (k_1, \ldots, k_p)\) and \(T = \text{diag}(t_1, \ldots, t_p)\). Then \(K = G (I - R^2)^{-1/2}\) and \(T = R^2 (I - R^2)^{-1}\). Define

\[
K^* = \sqrt{n} \left[ K - (I - R^2)^{-1/2} \right], \quad T^* = \sqrt{n} \left[ T - \Theta \right]. \tag{4.5}\]

Then

\[
K^* = G^* (I - R^2)^{-1/2} + R (I - R^2)^{-3/2} R^* + o_p(1) \tag{4.6}\]

\[
T^* = 2RR^* (I - R^2)^{-2} + o_p(1). \tag{4.7}\]

The asymptotic covariance matrix of \(t_i^* = 2\rho_i r_i^* (1 - \rho_i^2)^{-2} + o_p(1)\) and \(k_i^* = g_i^* (1 - \rho_i^2)^{-1/2} + \rho_i r_i^* (1 - \rho_i^2)^{-3/2} + o_p(1)\) is

\[
\begin{bmatrix}
\frac{4\rho_i^2}{(1 - \rho_i^2)^2} & \frac{\rho_i^2}{(1 - \rho_i^2)^{3/2}} \\
\frac{\rho_i^2}{(1 - \rho_i^2)^{3/2}} & \frac{1}{2(1 - \rho_i^2)}
\end{bmatrix}. \tag{4.8}\]

The asymptotic covariance matrix of \((\rho_j^2 - \rho_i^2)k_{ij}^* = (\rho_j^2 - \rho_i^2)g_{ij}^* (1 - \rho_j^2)^{-1/2} + o_p(1)\) and \((\rho_i^2 - \rho_j^2)k_{ij}^* = (\rho_i^2 - \rho_j^2)g_{ij}^* (1 - \rho_i^2)^{-1/2} + o_p(1)\), \(i \neq j\), is

\[
\begin{bmatrix}
\rho_i^2 + \rho_j^2 - 2\rho_i^2 \rho_j^2 & (\rho_i^2 + \rho_j^2) \sqrt{1 - \rho_i^2} \sqrt{1 - \rho_j^2} \\
(\rho_i^2 + \rho_j^2) \sqrt{1 - \rho_i^2} \sqrt{1 - \rho_j^2} & \rho_i^2 + \rho_j^2 - 2\rho_i^2 \rho_j^2
\end{bmatrix}. \tag{4.9}\]
A solution $f$ to (4.1) is related to a solution $k$ of (4.4) by $f = A k$. Thus

$$\sqrt{n} \left( f_j - \phi_j \right) = \sqrt{n} A \left[ k_j - \left( 1 - \rho_j^2 \right)^{-\frac{1}{2}} \varepsilon_j \right]$$

$$= A \left( k_{1j}^*, ..., k_{pj}^* \right)' + o_p(1),$$

Thus

$$n \mathcal{E} \left( f_j - \phi_j \right) (\bar{f}_j - \phi_j)' \rightarrow \frac{1}{2 \left( 1 - \rho_j^2 \right)} \alpha_j \alpha_j' + \sum_{k=1 \atop k \neq j}^p \frac{\rho_k^2 + \rho_j^2 - 2 \rho_k^2 \rho_j^2}{\left( \rho_j^2 - \rho_k^2 \right)^2} \alpha_k \alpha_k'$$

$$= \frac{1}{2} \phi_j \phi_j' + \sum_{k=1 \atop k \neq j}^p \frac{\left( \rho_k^2 + \rho_j^2 - 2 \rho_k^2 \rho_j^2 \right) (1 - \rho_k^2)}{\left( \rho_j^2 - \rho_k^2 \right)^2} \phi_k \phi_k',$$

$$n \mathcal{E} \left( f_j - \phi_j \right) (\bar{f}_\ell - \phi_\ell)' \rightarrow \frac{\sqrt{1 - \rho_j^2} \sqrt{1 - \rho_\ell^2} \left( \rho_j^2 + \rho_\ell^2 \right)}{\left( \rho_j^2 - \rho_\ell^2 \right)^2} \alpha_j \alpha_j'$$

$$= \frac{1}{2} \phi_j \phi_j' + \frac{(1 - \rho_j^2)(1 - \rho_\ell^2)}{\left( \rho_j^2 - \rho_\ell^2 \right)^2} \phi_\ell \phi_\ell'.$$

5 Case of one set of variables nonstochastic

5.1 Canonical variables

If the observations on $X$ are nonstochastic, the asymptotic theory of the estimators is different.

We take as the model

$$Y = Bx + Z, \quad (5.1)$$

with $Z$ distributed according to $N(0, \Sigma_{ZZ})$. We denote the observations as $(y_1, x_1), ..., (y_n, x_n)$. 

17
A linear combination $\phi'Y_\alpha$ has mean $\phi'Bx_\alpha$ and variance $\phi'\Sigma_{ZZ}\phi$. The mean sum of squares due to $x$ is

$$\frac{1}{n} \sum_{\alpha=1}^{n} (E\phi'Y_\alpha)^2 = \phi'BS_{XX}B'\phi,$$  \hspace{1cm} (5.2)$$

where $S_{XX} = (1/n) \sum_{\alpha=1}^{n} x_\alpha x_\alpha$. Maximization of (5.2) relative to the error variance of $\phi'\Sigma_{ZZ}\phi$ leads to

$$BS_{XX}B'\phi = \theta \Sigma_{ZZ}\phi, \quad \phi'\Sigma_{ZZ}\phi = 1,$$  \hspace{1cm} (5.3)$$

$$|BS_{XX}B' - \theta \Sigma_{ZZ}| = 0. \quad \hspace{1cm} (5.4)$$

These equations are similar to (2.12), but $\Sigma_{XX}$ in (2.12) has been replaced by (nonstochastic) $S_{XX}$ in (5.3).

We shall find a suitable canonical form by replacing (2.1) and (2.3) by

$$\begin{bmatrix}
-\rho(\Sigma_{ZZ} + BS_{XX}B') & BS_{XX} \\
S_{XX}B' & -\rho S_{XX}
\end{bmatrix} \begin{bmatrix}
\alpha \\
\gamma
\end{bmatrix} = 0, \quad \hspace{1cm} (5.5)$$

$$\alpha'(\Sigma_{ZZ} + BS_{XX}B')\alpha = 1, \quad \gamma' S_{XX} \gamma = 1. \quad \hspace{1cm} (5.6)$$

Solving the second vector equation in (5.5) for $\rho \gamma = B'\alpha$ and substituting in the first gives

$$BS_{XX}B'\alpha = \rho^2 (\Sigma_{ZZ} + BS_{XX}B') \alpha. \quad \hspace{1cm} (5.7)$$

This equation and (5.6) imply $\alpha'\Sigma_{ZZ}\alpha = 1 - \rho^2$, $\theta = \rho^2/(1 - \rho^2)$, and $\phi = \alpha(1 - \rho^2)^{-\frac{1}{2}}$.

The solutions to (5.5) and (5.6) and $\alpha_{ii} > 0, \rho_1 > ... > \rho_p$ define the matrices $A_n = (\alpha_1, ..., \alpha_p)$, $\Gamma_n = (\gamma_1, ..., \gamma_p)$, $R_n = \text{diag}(\rho_1, ..., \rho_p)$. Where convenient, the subscript $n$ is used to emphasize that the matrices of transformed parameters depend on $n$ through $S_{XX}$. Now
define $U = A'_n Y$, $v_\alpha = \Gamma'_n x_\alpha$, $W = A'_n Z$. Then

$$\Sigma_{WW} = A'_n \Sigma_{ZZ} A_n = I - R_n^2, \quad \text{(5.8)}$$

$$S_{VV} = \Gamma'_n S_{XX} \Gamma_n = I, \quad \text{(5.9)}$$

$$A'_n B S_{XX} \Gamma_n = R_n = A'_n B (\Gamma'_n)^{-1}. \quad \text{(5.10)}$$

We write the model for $U$ in terms of $v$ and $W$ as

$$U = \Psi v + W, \quad \text{(5.11)}$$

where $R_n = A'_n B (\Gamma'_n)^{-1}$ has been replaced by $\Psi$.

The maximum likelihood estimators of $B$ and $\Sigma_{ZZ}$ are given by (2.14) and

$$S_{ZZ} = \frac{1}{n} \sum_{\alpha=1}^{n} \left( y_\alpha - \hat{B} x_\alpha \right) \left( y_\alpha - \hat{B} x_\alpha \right)' = S_{YY} - \hat{B} S_{XX} \hat{B}'. \quad \text{(5.12)}$$

The estimators of $A_n$ and $R_n^2$ are formed from the solution of

$$\hat{B} S_{XX} \hat{B} a = r^2 S_{YY} a, \quad \text{(5.13)}$$

$$a' S_{YY} a = 1. \quad \text{(5.14)}$$

The estimators are $\hat{A} = (a_1, \ldots, a_p)$, $\hat{R}^2 = \text{diag}(r_1^2, \ldots, r_p^2)$, $r_1^2 > \ldots > r_p^2 > 0$.

When we transform from $Y$, $X$, and $Z$ to $U$, $V$, and $W$, the estimators of $\Psi$ and $\Sigma_{WW}$ are

$$\hat{\Psi} = S_{UV} S_{VV}^{-1} = S_{UV}, \quad \text{(5.15)}$$

$$S_{WW} = \frac{1}{n} \sum_{\alpha=1}^{n} \left( u_\alpha - \hat{\Psi} v_\alpha \right) \left( u_\alpha - \hat{\Psi} v_\alpha \right)' \quad \text{(5.16)}$$

$$= S_{UU} - \hat{\Psi} S_{VV} \hat{\Psi}' = S_{UU} - \hat{\Psi} \hat{\Psi}'$$

$$= S_{UU} - S_{UV} S_{VV}^{-1} S_{VV} = S_{UU} - S_{UV} S_{VV},$$

19
where

\[ S_{UU} = A'_n S_{YY} A_n, \quad S_{UV} = A'_n S_{YX} \Gamma_n. \] (5.17)

Let \( g = A_n^{-1} a \) and \( G = A_n^{-1} \hat{A} \). Then (5.13) and (5.14) yield \( \hat{\Psi} \hat{\Psi}' g = r^2 S_{UU} g \) and \( g' S_{UU} g = 1 \),

\[ \hat{\Psi} \hat{\Psi}' G = S_{UU} G \hat{R}^2, \] (5.18)

\[ G' S_{UU} G = I. \] (5.19)

For the asymptotic theory we assume \( S_{XX} \to \Sigma_{XX} \) and that the roots of

\[ |B \Sigma_{XX} B' - \theta \Sigma_{ZZ}| = 0 \] (5.20)

are distinct. Then \( A_n \to A \) and \( \Theta_n \to \Theta \) satisfying

\[ B \Sigma_{XX} B'A = \Sigma_{ZZ} A \Theta, \] (5.21)

\[ A' \Sigma_{ZZ} A = (I + \Theta)^{-1}. \] (5.22)

Let \( G = I + (1/\sqrt{n}) G^* \) and \( \hat{R} = R_n + (1/\sqrt{n}) R^* \). Then (5.18) and (5.19) yield

\[ \sqrt{n} \left\{ (S_{WW} R_n + R_n S_{VV}) (I - R_n^2) - \left[ S_{WW} - (I - R_n^2) \right] R_n^2 \right\} \]
\[ = \ G^* R_n^2 - R_n^2 G^* + 2R_n R^* + o_p(1), \] (5.23)

\[ \sqrt{n} \sum_{\alpha=1}^{n} \left[ S_{WW} R_n + R_n S_{VV} + S_{WW} - (I - R_n^2) \right] = - (G^* + G^*) + o_p(1). \] (5.24)

We have used \( S_{UV} = R_n + S_{VV} \) and (5.17).

In components these equations include

\[ \frac{1}{\sqrt{n}} \sum_{\alpha=1}^{n} \left\{ (w_\alpha v_\alpha \rho_j + \rho_i v_\alpha w_\alpha) \left( 1 - \rho_j^2 \right) - \left[ w_\alpha w_\alpha \delta_{ij} - \left( 1 - \rho_j^2 \right) \delta_{ij} \right] \rho_j^2 \right\} \]
\[ = g_{ij} \left( \rho_j^2 - \rho_i^2 \right) + 2r_j \rho_j \delta_{ij} + o_p(1), \] (5.25)
\[
\frac{1}{\sqrt{n}} \sum_{a=1}^{n} \left[ 2w_{ia}v_{ia} \rho_i + w_{ia}^2 \right] = -2g_{ii} + o_p(1). \tag{5.26}
\]

From these we find the asymptotic covariances of \((\rho_j^2 - \rho_i^2)g_{ij}^*\) and \((\rho_i^2 - \rho_j^2)g_{ji}^*\), \(i \neq j\), as

\[
\begin{bmatrix}
(1 - \rho_j^2) \left( \rho_i^2 + \rho_j^2 - 3 \rho_i^2 \rho_j^2 + \rho_i^4 \rho_j^4 \right) & (1 - \rho_i^2) \left( \rho_i^2 + \rho_j^2 - 3 \rho_i^2 \rho_j^2 + \rho_i^4 \rho_j^4 \right) \\
(1 - \rho_i^2) \left( \rho_i^2 + \rho_j^2 - 3 \rho_i^2 \rho_j^2 + \rho_i^4 \rho_j^4 \right) & (1 - \rho_j^2) \left( \rho_i^2 + \rho_j^2 - 3 \rho_i^2 \rho_j^2 + \rho_i^4 \rho_j^4 \right)
\end{bmatrix}. \tag{5.27}
\]

The asymptotic covariances of \(r_i^*\) and \(g_{ii}^*\) are

\[
\begin{bmatrix}
(1 - \frac{1}{2} \rho_i^2) \left( 1 - \rho_i^2 \right)^2 & -\frac{1}{2} \rho_i \left( 1 - \rho_i^2 \right) \\
-\frac{1}{2} \rho_i \left( 1 - \rho_i^2 \right) & \frac{1}{2} \left( 1 - \rho_i^4 \right)
\end{bmatrix}. \tag{5.28}
\]

The covariances in (5.27) and (5.28) compared with those in (3.18) and (3.22) show the effect of treating \(x_{ia}\) as nonstochastic. The variance of \(g_{ij}^*\) in (5.27) is smaller than that in (3.18) by \(\rho_i^2 \rho_j^2 (1 - \rho_j^2)\); the variance of \(g_{ii}^*\) in (5.28) is smaller than that in (3.22) by \(\frac{1}{2} \rho_i^4\); and the variance of \(r_i^*\) in (5.28) is smaller than that in (3.22) by \(\frac{1}{2} \rho_i^2 (1 - \rho_i^2)^2\).

Since \(a_j = A_n g_j\) and \(\alpha_j = A_n \varepsilon\) as in (3.24), the asymptotic variances and covariances are

\[
nEC (a_j - \alpha_j) (a_j - \alpha_j)' = \frac{1}{2} \left( 1 - \rho_i^4 \right) \alpha_j \alpha_j' + \left( 1 - \rho_i^2 \right) \sum_{i=1}^{p} \frac{\rho_i^2 + \rho_i^2 - 3 \rho_i^2 \rho_j^2 + \rho_i^4 \rho_j^4}{\left( \rho_j^2 - \rho_i^2 \right)^2} \alpha_i \alpha_i', \tag{5.29}
\]

\[
nEC (a_j - \alpha_j) (a_\ell - \alpha_\ell)' = -\frac{\left( 1 - \rho_j^2 \right) \left( 1 - \rho_i^2 \right) \left( \rho_i^2 + \rho_j^2 - 3 \rho_i^2 \rho_j^2 \right)}{\left( \rho_j^2 - \rho_i^2 \right)^2} \alpha_j \alpha_\ell, \quad j \neq \ell. \tag{5.30}
\]

### 5.2 Regression model

Now consider the estimation of \(\Phi = (\phi_1, \ldots, \phi_p)\) and \(\Theta = \text{diag}(\theta_1, \ldots, \theta_p)\) composed of solutions to (5.3). As before, \(\phi = (1 - \rho^2)^{-\frac{1}{2}} \alpha\) and \(\theta = \rho^2 / (1 - \rho^2)\). The estimators of \(\phi\) and \(\theta\) are
the solutions \( f \) of (2.16). The transformation \( A_n k = f \) leads to (4.1), where \( S_{V V} = I \). The matrices \( K = (k_1, \ldots, k_p) \) and \( T = \text{diag}(t_1, \ldots, t_p) \) satisfy

\[
S_{U V} S_{V V} K = S_{W W} K T,
\]

(5.31)

\[
K' S_{W W} K = I.
\]

(5.32)

As in Section 4, \( K = G(I - \bar{R}^2)^{-\frac{1}{2}} \) and \( T = \bar{R}^2 (I - \bar{R}^2)^{-1} \). Define \( K^* \) and \( T^* \) by (4.5). Then (4.6) and (4.7) hold. The asymptotic covariance matrix of \( (\rho^2_j - \rho^2_i) k_{ji}^* = (\rho^2_j - \rho^2_i) g_{ji}^* (1 - \rho^2_i)^{-\frac{1}{2}} + o_p(1) \) and \( (\rho^2_i - \rho^2_j) k_{ji}^* = (\rho^2_j - \rho^2_i) g_{ji}^* (1 - \rho^2_i)^{-\frac{1}{2}} + o_p(1), i \neq j \), is

\[
\begin{bmatrix}
\rho^2_i + \rho^2_j - 3 \rho^2_i \rho^2_j + \rho^2_i \rho^2_j & \sqrt{1 - \rho^2_i} \sqrt{1 - \rho^2_j} (\rho^2_i + \rho^2_j - \rho^2_i \rho^2_j) \\
\sqrt{1 - \rho^2_i} \sqrt{1 - \rho^2_j} (\rho^2_i + \rho^2_j - \rho^2_i \rho^2_j) & \rho^2_i + \rho^2_j - 3 \rho^2_i \rho^2_j + \rho^2_i \rho^2_j
\end{bmatrix}
\]

(5.33)

The asymptotic covariance matrix of \( t_i^* = 2 \rho_i r_i^* (1 - \rho^2_i)^{-\frac{1}{2}} + o_p(1) \) and \( k_{ii}^* = g_{ii}^* (1 - \rho^2_i)^{-\frac{1}{2}} + \rho_i r_i^* (1 - \rho^2_i)^{-\frac{3}{2}} + o_p(1) \) is

\[
\begin{bmatrix}
\frac{4 \rho_i^2 (1 - \rho_i^2)}{(1 - \rho_i^2)^2} & \frac{\rho_i^2}{(1 - \rho_i^2)^{\frac{3}{2}}} \\
\frac{\rho_i^2}{(1 - \rho_i^2)^{\frac{3}{2}}} & \frac{1}{2 (1 - \rho_i^2)}
\end{bmatrix}
\]

(5.34)

Note that the asymptotic variance of \( k_{ij}^* (i \neq j) \) in (5.33) is that in (4.9) minus \( (1 - \rho^2_j) \rho^2_i \rho^2_j \); the asymptotic variance of \( k_{ii}^* \) in (5.34) is the same as that in (4.8); and the variance of \( t_i^* \) in (5.34) is that in (4.8) minus \( 2 \rho_i^2/(1 - r_i^2)^2 \). We have used the fact that the limiting distribution of \( \sqrt{n} \left( S_{U U} - S_{U V} S_{V V}^{-1} S_{U V} \right) \) is the same as the limiting distribution of

\[
(1/\sqrt{n}) \sum_{\alpha=1}^{n} (u_\alpha - \Psi v_\alpha) (u_\alpha - \Psi v_\alpha)'.
\]

(5.35)

Since \( f_j = A_n k_j \) and \( \phi_j = A_n \varepsilon_j (1 - \rho^2_j)^{-\frac{1}{2}} \),

\[
\sqrt{n} \left( f_j - \phi_j \right) = \sqrt{n} A_n \left[ k_j - (1 - \rho^2_j)^{-\frac{1}{2}} \varepsilon_j \right]
\]

22
\[ A_n \left[ k_j - (1 - \rho_j^2)^{-\frac{1}{2}} \varepsilon_j \right] \]

\[ = \sum_{i=1}^{p} \alpha_i k_{ij}^* + o_p(1) = \alpha_j k_{jj}^* + \sum_{j \neq i}^{p} \alpha_i k_{ij}^* + o_p(1) \]  

\[ = \phi_j \left( 1 - \rho_j^2 \right)^{-\frac{1}{2}} k_{jj}^* + \sum_{i=1}^{p} \phi_i \left( 1 - \rho_i^2 \right)^{-\frac{1}{2}} k_{ij}^* + o_p(1) \]  

\[ \phi_j \left( 1 - \rho_j^2 \right)^{-\frac{1}{2}} k_{jj}^* + \sum_{i \neq j}^{p} \phi_i \left( 1 - \rho_i^2 \right)^{-\frac{1}{2}} g_{ij}^* \left( 1 - \rho_j^2 \right)^{-\frac{1}{2}} + o_p(1). \]  

The asymptotic variances and covariances are

\[ n \mathcal{E} \left( f_j - \phi_j \right) \left( f_j - \phi_j \right)' \]

\[ = \frac{1}{2} \phi_j^2 \phi_j' + \sum_{j \neq i}^{p} \frac{\left( \rho_j^2 + \rho_i^2 - 3 \rho_i^2 \rho_j^2 + \rho_i^2 \rho_j^2 \right) \left( 1 - \rho_i^2 \right)}{\left( \rho_j^2 - \rho_i^2 \right)^2} \phi_i \phi_i', \]  

\[ = \frac{1}{2} \phi_j^2 \phi_j' + \sum_{j \neq i}^{p} \frac{\left( \rho_j^2 + \rho_i^2 - 3 \rho_i^2 \rho_j^2 + \rho_i^2 \rho_j^2 \right) \left( 1 - \rho_i^2 \right)}{\left( \rho_j^2 - \rho_i^2 \right)^2} \phi_i \phi_i', \]  

\[ n \mathcal{E} \left( f_j - \phi_j \right) \left( f_i - \phi_i \right)' = \frac{\left( 1 - \rho_j^2 \right) \left( 1 - \rho_i^2 \right) \left( \rho_j^2 + \rho_i^2 - 3 \rho_i^2 \rho_j^2 \right)}{\left( \rho_j^2 - \rho_i^2 \right)^2} \phi_i \phi_i', \]  

\[ n \mathcal{E} \left( f_j - \phi_j \right) \left( f_i - \phi_i \right)' = \frac{\left( 1 - \rho_j^2 \right) \left( 1 - \rho_i^2 \right) \left( \rho_j^2 + \rho_i^2 - 3 \rho_i^2 \rho_j^2 \right)}{\left( \rho_j^2 - \rho_i^2 \right)^2} \phi_i \phi_i', \]  

\[ i \neq j. \]  

Anderson (1951b) made a slightly different transformation to a canonical form. He defined

\[ \bar{U} = \Phi_n' Y = (I - R_n^2)^{-\frac{1}{2}} A_n' Y = (I - R_n^2)^{-\frac{1}{2}} U \]  

and \[ \bar{W} = \Phi_n' Z = (I - R_n^2)^{-\frac{1}{2}} Z. \]  

Then instead of (5.31) and (5.32) we obtain

\[ \bar{S}_{UV} \bar{S}_{VU} \bar{K} = \bar{S}_{WW} \bar{K} \bar{T}, \]  

\[ \bar{K}' \bar{S}_{WW} \bar{K} = I, \]  

and \[ \bar{K} = (I - R^2)^{\frac{1}{2}} \bar{K}. \]  

The asymptotic covariances of \[ \left( \rho_j^2 - \rho_i^2 \right) \bar{k}_{ij} = \left( \rho_j^2 - \rho_i^2 \right) \sqrt{n} \bar{k}_{ij} + o_p(1) \]  

and \[ \left( \rho_i^2 \rho_j^2 \right) \bar{k}_{ji}, i \neq j, \]  

are

\[ \begin{bmatrix}
(1 - \rho_j^2) \left( \rho_i^2 + \rho_j^2 - 3 \rho_i^2 \rho_j^2 + \rho_i^2 \rho_j^2 \right) & (1 - \rho_j^2) \left( \rho_i^2 + \rho_j^2 - 3 \rho_i^2 \rho_j^2 + \rho_i^2 \rho_j^2 \right) \\
(1 - \rho_j^2) \left( \rho_i^2 + \rho_j^2 - 3 \rho_i^2 \rho_j^2 + \rho_i^2 \rho_j^2 \right) & (1 - \rho_j^2) \left( \rho_i^2 + \rho_j^2 - 3 \rho_i^2 \rho_j^2 + \rho_i^2 \rho_j^2 \right)
\end{bmatrix}. \]  

\[ \begin{bmatrix}
(1 - \rho_j^2) \left( \rho_i^2 + \rho_j^2 - 3 \rho_i^2 \rho_j^2 + \rho_i^2 \rho_j^2 \right) & (1 - \rho_j^2) \left( \rho_i^2 + \rho_j^2 - 3 \rho_i^2 \rho_j^2 + \rho_i^2 \rho_j^2 \right) \\
(1 - \rho_j^2) \left( \rho_i^2 + \rho_j^2 - 3 \rho_i^2 \rho_j^2 + \rho_i^2 \rho_j^2 \right) & (1 - \rho_j^2) \left( \rho_i^2 + \rho_j^2 - 3 \rho_i^2 \rho_j^2 + \rho_i^2 \rho_j^2 \right)
\end{bmatrix}. \]  

\[ \begin{bmatrix}
(1 - \rho_j^2) \left( \rho_i^2 + \rho_j^2 - 3 \rho_i^2 \rho_j^2 + \rho_i^2 \rho_j^2 \right) & (1 - \rho_j^2) \left( \rho_i^2 + \rho_j^2 - 3 \rho_i^2 \rho_j^2 + \rho_i^2 \rho_j^2 \right) \\
(1 - \rho_j^2) \left( \rho_i^2 + \rho_j^2 - 3 \rho_i^2 \rho_j^2 + \rho_i^2 \rho_j^2 \right) & (1 - \rho_j^2) \left( \rho_i^2 + \rho_j^2 - 3 \rho_i^2 \rho_j^2 + \rho_i^2 \rho_j^2 \right)
\end{bmatrix}. \]
The asymptotic covariances of $t_i^*$ and $k_{ii}^*$ are

$$
\begin{bmatrix}
4\theta_i + 2\theta_i^2 & \theta_i \\
\theta_i & \frac{1}{2}
\end{bmatrix} = \begin{bmatrix}
\frac{4\rho_i^2}{(1-\rho_i^2)^2} & \frac{\rho_i^2}{1-\rho_i^2} \\
\frac{\rho_i^2}{1-\rho_i^2} & \frac{1}{2}
\end{bmatrix}
$$

(5.42)

Anderson gave the asymptotic covariances of $(\theta_j - \theta_i)k_{ij}^*$ and $(\theta_i - \theta_j)k_{ji}^*$, $i \neq j$, as

$$
\begin{bmatrix}
\theta_i + \theta_j + \theta_i^2 & \theta_i^2 + \theta_i \theta_j \\
\theta_i^2 + \theta_i \theta_j & \theta_i + \theta_j + \theta_i^2
\end{bmatrix}
$$

$$
= \begin{bmatrix}
\rho_i^2 + \rho_i^2 - 3\rho_i^2 \rho_j^2 + \rho_i^4 & \rho_i^2 - \rho_i^2 \rho_j^2 \\
\rho_i^2 + \rho_i^2 - 3\rho_i^2 \rho_j^2 + \rho_i^4 & \rho_i^2 - \rho_i^2 \rho_j^2
\end{bmatrix}
\begin{bmatrix}
(1-\rho_i^2)(1-\rho_j^2) \\
(1-\rho_i^2)(1-\rho_j^2)
\end{bmatrix}
$$

(5.43)

These asymptotic covariances in this (current) paper agree with those in the earlier paper.

However, in the earlier paper a more general asymptotic theory was developed to include cases of roots of $|BS_{XX}B' - \theta\Sigma_{ZZ}| = 0$ having multiplicities greater than 1.

6 Other work on asymptotic theory

6.1 Robinson

Robinson (1973) also treats $(Y', X')'$ with covariance matrix (1.1) and transforms to the model (1.2) in the notation of this paper. He then defines $\bar{\phi}_j$, $j = 1, ..., p$, as solutions to

$$
\Sigma_{ZZ}^{-1} \Sigma_{YX} \Sigma_{XX}^{-1} \Sigma_{XY} \Sigma_{ZZ}^{-1} \bar{\phi} = \theta \Sigma_{ZZ}^{-1} \bar{\phi},
$$

(6.1)

$$
\bar{\phi}^t \Sigma_{ZZ}^{-1} \bar{\phi} = 1,
$$

(6.2)

where $\theta = \rho^2/(1-\rho^2)$. The solutions of (6.1) are $\bar{\phi}_j = \Sigma_{ZZ} \phi_j = \Sigma_{ZZ} \alpha_j / \sqrt{1-\rho_j^2}$, $j = 1, ..., p$, as in (2.6) and (2.9). He further defines

$$
\bar{\gamma}_j = \theta_j^{-1} B' \Sigma_{ZZ}^{-1} \bar{\phi}_j = \theta_j^{-1} B' \phi_j = \left(\sqrt{1-\rho_j^2} / \rho_j\right) \gamma_j.
$$

(6.3)
Then
\[ B = \sum_{j=1}^{p} \theta_j^2 \phi_j \tilde{v}_j'. \] (6.4)

The sample vectors corresponding to \( \tilde{\phi}_j, \ j = 1, ..., p, \) are the solutions to
\[ S_{YY} S_{XX}^{-1} S_{XY} S_{ZZ}^{-1} \tilde{f} = t \tilde{f}, \] (6.5)
\[ \tilde{f}' S_{ZZ}^{-1} \tilde{f} = 1. \] (6.6)

Let \( \tilde{F} = (\tilde{f}_1, ..., \tilde{f}_p). \) Then (6.5) and (6.6) are
\[ S_{YY} S_{XX}^{-1} S_{XY} S_{ZZ}^{-1} \tilde{F} = \tilde{F} T, \] (6.7)
\[ \tilde{F}' S_{ZZ}^{-1} \tilde{F} = I. \] (6.8)

Let \( A' \tilde{F} = \tilde{K}. \) Then (6.7) and (6.8) become
\[ S_{UV} S_{VY}^{-1} S_{VU} S_{WW}^{-1} \tilde{K} = \tilde{K} T, \] (6.9)
\[ \tilde{K}' S_{WW}^{-1} \tilde{K} = I. \] (6.10)

Then \( \tilde{K} \overset{p}{\to} (I - R^2)^{1/2} \) and \( T \overset{p}{\to} \Theta. \) Define \( \tilde{K}^* \) by \( \tilde{K} = (I - R^2)^{1/2} + (1/\sqrt{n})K^*. \) Then (6.9) yields
\[ \sqrt{n}\left\{S_{VV}R + RS_{VV} + R(S_{VV} - I)R - \Theta \left[S_{WW} - (I - R^2)\right]\right\} \] (6.11)
\[ = [\tilde{K}^* \Theta - \Theta \tilde{K}^*] (I - R^2)^{1/2} + T^* (I - R^2) + o_p(1). \]

In components (6.11) is
\[ \frac{1}{\sqrt{n}} \sum_{\alpha=1}^{n} (\omega_{\alpha} v_{\alpha j} \rho_j + \rho_{\alpha} \omega_{\alpha} v_{\alpha j} + \rho_{\alpha} \omega_{\alpha} v_{\alpha j} \rho_j - \theta_{\alpha} \omega_{\alpha} w_{\alpha j}) \] (6.12)
\[ = (\theta_j - \theta_i) \tilde{k}_{ij} \sqrt{1 - \rho_j^2} + o_p(1), \ i \neq j. \]
From this we calculate the asymptotic covariance matrix of \((\theta_j - \theta_i)\hat{k}^*_{ij}\) and \((\theta_i - \theta_j)\hat{k}^*_{ji}\) as

\[
\begin{bmatrix}
\theta_i + \theta_j & \frac{\theta_i + \theta_j + 2\theta_i \theta_j}{\sqrt{\theta_i} \sqrt{\theta_j}} \\
\frac{\theta_i + \theta_j + 2\theta_i \theta_j}{\sqrt{\theta_i} \sqrt{\theta_j}} & \theta_i + \theta_j
\end{bmatrix}
= \begin{bmatrix}
\frac{\rho_i^2 + \rho_j^2 - 2\rho_i \rho_j}{(1 - \rho_i)(1 - \rho_j)} & \frac{\rho_i^2 + \rho_j^2}{\sqrt{1 - \rho_i} \sqrt{1 - \rho_j}} \\
\frac{\rho_i^2 + \rho_j^2}{\sqrt{1 - \rho_i} \sqrt{1 - \rho_j}} & \frac{\rho_i^2 + \rho_j^2 - 2\rho_i \rho_j}{(1 - \rho_i)(1 - \rho_j)}
\end{bmatrix}.
\quad (6.13)
\]

The asymptotic covariance matrix of \((\rho_j^2 - \rho_i^2)\hat{k}^*_{ij}\) and \((\rho_i^2 - \rho_j^2)\hat{k}^*_{ji}\) is

\[
\begin{bmatrix}
(1 - \rho_i^2) \left(1 - \rho_j^2\right) \left(\rho_i^2 + \rho_j^2 - 2\rho_i \rho_j\right) & (1 - \rho_i^2) \frac{3}{2} \left(1 - \rho_j^2\right) \left(\rho_i^2 + \rho_j^2\right) \\
(1 - \rho_i^2) \frac{3}{2} \left(1 - \rho_j^2\right) \left(\rho_i^2 + \rho_j^2\right) & (1 - \rho_i^2) \left(1 - \rho_j^2\right) \left(\rho_i^2 + \rho_j^2 - 2\rho_i \rho_j\right)
\end{bmatrix}.
\quad (6.14)
\]

From (6.10) we obtain

\[
2\sqrt{1 - \rho_i^2} \hat{k}^*_{ii} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} w_i^2 + o_p(1).
\quad (6.15)
\]

and \(\text{Var} \hat{k}^*_{ii} \to \frac{1}{2}(1 - \rho_i^2)\). The asymptotic covariance matrix of \(t_i^*\) and \(\hat{k}^*_{ii}\) is

\[
\begin{bmatrix}
\frac{4\rho_i^2}{(1 - \rho_i)^2} & -\frac{\rho_i^2}{2\sqrt{1 - \rho_i}} \\
-\frac{\rho_i^2}{2\sqrt{1 - \rho_i}} & \frac{1}{2} (1 - \rho_i^2)
\end{bmatrix}.
\quad (6.16)
\]

From \(\Phi^* = \Sigma_{ZZ}A(I - R^2)^{-\frac{1}{2}}\) and \(A'\Sigma_{ZZ}A = I - R^2\) we find \(\Phi^* = (A')^{-1}(I - R^2)^{-\frac{1}{2}}\) and

\[
\sqrt{n} (F^* - \Phi^*) = \Phi^* (I - R^2)^{-\frac{1}{2}} \tilde{K}^*.
\quad (6.17)
\]

Thus

\[
n \mathcal{E} \left(\ell_j^* - \phi_j^*\right) (\ell_j^* - \phi_j^*)' \rightarrow \frac{1}{2} \phi_j^* \phi_j^* + \sum_{k=1}^{p} \frac{(1 - \rho_j^2) \left(\rho_k^2 + \rho_j^2 - 2\rho_k \rho_j\right)}{(\rho_j^2 - \rho_k^2)^2} \phi_k^* \phi_k^* \quad (6.18)
\]

\[
= \frac{1}{2} \phi_j^* \phi_j^* + \sum_{k=1}^{p} \frac{(1 + \theta_k)(\theta_k + \theta_j)}{(\theta_j - \theta_k)^2} \phi_k^* \phi_k^*,
\]

\[
n \mathcal{E} (\ell_j^* - \phi_j^*) (\ell_j - \phi_j')' = -\frac{(1 - \rho_j^2) \left(1 - \rho_j^2\right) \left(\rho_j^2 + \rho_j^2\right)}{(\rho_j^2 - \rho_j^2)^2} \quad (6.19)
\]
\[ = - \frac{\theta_\ell + \theta_\ell + \theta_j \theta_\ell}{(\theta_j - \theta_\ell)^2}. \]

Robinson gives as the asymptotic covariance matrix of \( \hat{f}^*_j \)

\[
\sum_{k \neq j}^{p} \frac{\theta_j + \theta_k}{(\theta_j - \theta_k)^2} \phi_k^* \phi_k'^*.
\]

This expression differs from (6.18) in two respects. The coefficient in the \( k \)th term of (6.18) is \( 1 + \theta_k \) times the coefficient in (6.20); this reflects the fact that Robinson calculates the asymptotic variance of \( \hat{k}^*_j \) using only the first two terms in (6.11). Secondly, the omission of the term \( \frac{1}{2} \phi_j^* \phi_j'^* \) is correct only if the normalization is \( \hat{k}' \hat{k} = 1 \), but Robinson’s normalization is \( \hat{f}' S^i_{zz} \hat{f} = 1 \).

(Robinson refers to Wilkinson (1965), Chapter 2, for a “perturbation expansion,” but Wilkinson writes “... we are not interested in a multiplying factor,” meaning no normalization is imposed.)

6.2 Brillinger

Brillinger (1975, Ch. 10) scaled the vectors, say \( \hat{\alpha} = \omega \alpha \) and \( \hat{\gamma} = \psi \gamma \) so \( \hat{\alpha}' \hat{\alpha} = 1 \) and \( \hat{\gamma}' \hat{\gamma} = 1 \); that is \( \omega^2 = 1/\alpha' \alpha = \hat{\alpha}' \Sigma_{YY} \hat{\alpha} \) and \( \psi^2 = 1/\gamma' \gamma = \hat{\gamma}' \Sigma_{XX} \hat{\gamma} \). His estimators \( \bar{\alpha} \) and \( \bar{\epsilon} \) satisfy (2.11), (2.12),

\[
\bar{\alpha}' \bar{\alpha} = 1, \quad \bar{\epsilon}' \bar{\epsilon} = 1.
\]

Then \( \bar{g} = A^{-1} \bar{\alpha} \) and \( \bar{h} = \Gamma^{-1} \bar{\epsilon} \) satisfy (3.2) [or (3.4) and (3.5)] and

\[
\bar{g}' A' A \bar{g} = 1, \quad \bar{h}' \Gamma' \Gamma \bar{h} = 1.
\]

As above, \( \bar{\alpha}_j \xrightarrow{P} \alpha_j^* \) \( \omega_j \alpha_j \), \( \bar{\gamma}_j \xrightarrow{P} \gamma_j = \psi_j \gamma_j \), \( j = 1, ..., p \). The probability limits of (2.11), (2.12), and (6.22) imply that \( \bar{g}_j \xrightarrow{P} \omega_j \varepsilon_j \) and \( \bar{h} \xrightarrow{P} \psi_j \varepsilon_j \). Define \( \bar{G} = (\bar{g}_1, ..., \bar{g}_p) \), \( \Omega = \text{diag}(\omega_1, ..., \omega_p) \), \( \bar{G}^* = \)
\[
\sqrt{T} (\mathbf{G} - \Omega). \quad \text{Then}
\]
\[
\sqrt{T} \left[ (S_{UV} - R)R + R(S_{VU} - R)R - R(S_{VV} - I)R - (S_{UU} - I)R^2 \right] \Omega
\]
\[
= \mathbf{G}^* R^2 - R^2 \mathbf{G}^* + \Omega Q + o_p(1). \tag{6.23}
\]

Thus \( \mathbf{g}_{ij}^* = \omega_j \mathbf{g}_{ij}^* + o_p(1), \ i \neq j. \) From (6.22) we obtain
\[
1 = \left( \omega_j \mathbf{e}_j + \frac{1}{\sqrt{T}} \mathbf{g}_{ij}^* \right)' \mathbf{A}' \mathbf{A} \left( \omega_j \mathbf{e}_j + \frac{1}{\sqrt{T}} \mathbf{g}_{ij}^* \right)
\]
\[
= \omega_j^2 \alpha_j \mathbf{A} \mathbf{g}_{ij}^* + \frac{2}{\sqrt{T}} \alpha_j' \mathbf{A} \mathbf{g}_{ij}^* + o_p \left( \frac{1}{\sqrt{T}} \right) \tag{6.24}
\]
or
\[
0 = \alpha_j' \mathbf{A} \mathbf{g}_{ij}^* + o_p(1)
\]
\[
= \alpha_j' \sum_{k=1}^{p} \alpha_k \mathbf{g}_{kj}^* + o_p(1). \tag{6.25}
\]
Hence
\[
\alpha_j' \alpha_j \mathbf{g}_{jj}^* = - \sum_{k=1 \atop k \neq j}^{p} \alpha_j \alpha_k \mathbf{g}_{kj}^* + o_p(1). \tag{6.26}
\]

From \( \bar{\alpha}_j = \mathbf{A} \bar{\mathbf{g}}_j \) and \( \bar{\alpha}_j = A \omega_j \mathbf{e}, \) we obtain
\[
\sqrt{T} (\bar{\mathbf{a}}_j - \bar{\alpha}_j) = \sum_{k=1}^{p} \alpha_k \mathbf{g}_{kj}^* + o_p(1)
\]
\[
= \sum_{k=1 \atop k \neq j}^{p} \left( \alpha_k - \omega_j^2 \alpha_j' \alpha_k \alpha_j \right) \mathbf{g}_{kj}^* + o_p(1). \tag{6.27}
\]

Brillinger calculated the asymptotic covariance of \( \bar{\mathbf{a}}_j \) by representing (6.27) as \( \sum_{k \neq j}^{p} \alpha_k \mathbf{g}_{kj}^* \), which is incorrect. (Note that this would be correct if \( \alpha_j' \alpha_k = 0, \ j \neq k. \)) The corresponding error is made in the asymptotic covariance matrix of \( \bar{\mathbf{e}}_j. \) The asymptotic covariance matrix of \( \bar{\mathbf{a}}_j \) and \( \bar{\mathbf{a}}_k \) contains the term \( 4 \rho_j^2 \rho_k^2 \) instead of the correct \( 2 \rho_j^2 \rho_k^2. \)
6.3 Velu, Reinsel, and Wichern

Velu, Reinsel, and Wichern (1986) have treated this matter. They define vectors \( \tilde{\phi}_j, j = 1, \ldots, p \), as solutions to

\[
\Sigma_{ZZ}^{-\frac{1}{2}} \Sigma_{YX} \Sigma_{XX}^{-1} \Sigma_{YX} \Sigma_{ZZ}^{-\frac{1}{2}} \tilde{\phi} = \theta \phi, \tag{6.28}
\]

\[
\tilde{\phi}' \phi = 1. \tag{6.29}
\]

As before, \( \theta \) is a solution to \( |B \Sigma_{XX} B - \theta \Sigma_{ZZ}| = 0 \). The estimators satisfy

\[
S_{ZZ}^{-\frac{1}{2}} S_{YX} S_{XX}^{-1} S_{YX} S_{ZZ}^{-\frac{1}{2}} \bar{f} = t \bar{f}, \tag{6.30}
\]

\[
\bar{f}' \bar{f} = 1. \tag{6.31}
\]

Because the Taylor's expansion of \( \Sigma_{ZZ}^{-\frac{1}{2}} \) is complicated, for the sake of showing the error we shall study the special case of \( \Sigma_{ZZ} = I \), \( \Sigma_{XX} = I \), \( B = \Delta \) diagonal. Then \( \Sigma_{YX} = \Delta \), \( \Sigma_{YY} = \Delta^2 + I \), \( \Theta = \Delta^2 \), \( \Phi = I \). We use

\[
S_{ZZ}^{-\frac{1}{2}} = I - \frac{1}{2} (S_{ZZ} - I) + o_p(1). \tag{6.32}
\]

Let \( \bar{F} = (\bar{f}_1, \ldots, \bar{f}_p) = I + (1/\sqrt{n}) \bar{F}^{**} \). Then (6.30) leads to

\[
\sqrt{n} \left[ S_{XX} \Delta + \Delta S_{XX} + \Delta (S_{XX} - I) \Delta - \frac{1}{2} (S_{ZZ} - I) \Delta^2 - \frac{1}{2} \Delta^2 (S_{ZZ} - I) \right] \tag{6.33}
\]

\[
= \bar{F}^{**} \Theta - \Theta \bar{F} + T^* + o_p(1),
\]

and (6.29) leads to \( \bar{F}^{**} + \bar{F}^{*} = o_p(1) \). From (6.33) we calculate the asymptotic variance of \( (\theta_j - \theta_i) \bar{f}_{ij} \) as

\[
\text{var} \left[ (\theta_j - \theta_i) \bar{f}_{ij} \right] = \theta_i + \theta_j + \theta_i \theta_j + (\theta_i + \theta_j)^2 / 4. \tag{6.34}
\]

Velu, Reinsel, and Wichern have \( \theta_i + \theta_j \) for (6.34); as did Robinson, they neglect the terms due to \( S_{XX} \) and \( S_{ZZ} \). Velu et al. state incorrectly that the asymptotic distribution is the same for
\( \bar{f} \) defined by (6.30) as for \( \bar{f} \) defined by (6.30) with \( S_{ZZ}^{-1/2} \) replaced by \( S_{ZZ}^{-1/2} \). (However, Velu et al. purport to treat an autoregressive model with \( \mathbf{X}_t = \mathbf{Y}_{t-1} \). In that case the covariances of \( \text{vec} \sum_{t=1}^{T} \mathbf{Y}_t \mathbf{Y}_t' \) is more complicated than for \( \mathbf{Y}_t \) and \( \mathbf{X}_t \) uncorrelated.)

### 6.4 Cointegration

Canonical correlations and vectors have been used in time series analysis. In a stationary first-order vector autoregressive process

\[
\mathbf{Y}_t = \mathbf{B} \mathbf{Y}_{t-1} + \mathbf{Z}_t, \quad t = \ldots, -1, 0, 1, \ldots, \tag{6.35}
\]

the covariance matrix of \( (\mathbf{Y}_t', \mathbf{Y}_{t-1}') \) is

\[
\mathbb{E} \left[ \begin{bmatrix} \mathbf{Y}_t \\ \mathbf{Y}_{t-1} \end{bmatrix} \right] \left[ \begin{bmatrix} \mathbf{Y}_t' \\ \mathbf{Y}_{t-1}' \end{bmatrix} \right] = \begin{bmatrix} \Sigma_{YY} & \mathbf{B} \Sigma_{YY} \\ \Sigma_{XY} \mathbf{B}' & \Sigma_{YY} \end{bmatrix}, \tag{6.36}
\]

where \( \Sigma_{YY} = \Sigma_{ZZ} + \mathbf{B} \Sigma_{YY} \mathbf{B}' = \sum_{i=0}^{\infty} \mathbf{B}^i \Sigma_{ZZ} \mathbf{B}^i \). A necessary condition for stationarity is that the characteristic roots of \( \mathbf{B} \) be less than 1 in absolute value. A nonstationary process, \( t = 1, 2, \ldots \), may be defined by (6.36) and \( \mathbf{Y}_0 = \mathbf{0} \) when some of the characteristic roots are 1. The left-sided characteristic vectors of \( \mathbf{B} \) define linear combinations of \( \mathbf{Y}_t \) that are stationary; these are the cointegrating relations (Granger, 1981). Suppose each root of \( \mathbf{B} \) is 1 or is less than 1 in absolute value and that \( \mathbf{B} - \mathbf{I} = \Pi \) is of rank \( r \). Let \( \Delta \mathbf{Y}_t = \mathbf{Y}_t - \mathbf{Y}_{t-1} \). Then (6.35) can be written

\[
\Delta \mathbf{Y}_t = \Pi \mathbf{Y}_{t-1} + \mathbf{Z}_t, \tag{6.37}
\]

and \( \Pi \mathbf{Y}_t \) are the cointegrating relations. The maximum likelihood estimator of \( \Pi \) is the reduced rank regression estimator described in Section 2 of Johansen (1988).
6.5 Anderson

In the model \( \mathbf{Y} = \mathbf{B} \mathbf{X} + \mathbf{Z} \) Anderson (1951b) treated the general case in which the \( \mathbf{X} \)'s are nonstochastic and the distinct roots of \(|\mathbf{B} \mathbf{S}_{XX} \mathbf{B}' - \theta \mathbf{S}_{ZZ}| = 0\) are of arbitrary multiplicities; the limiting distributions are not normal distributions. In the next section of this paper we treat canonical correlations and vectors when \( \theta = 0 \) has arbitrary multiplicity.

6.6 Izenman

Izenman (1975) defines vectors by

\[
\Sigma_{YY}^{-1/2} \Sigma_{YX} \Sigma_{XX}^{-1} \Sigma_{XY} \Sigma_{YY}^{-1/2} \tilde{\phi} = \theta \tilde{\phi}
\]

(6.38)

and \( \phi^* \tilde{\phi} = 1 \) and treats the corresponding estimators but does not get explicit expressions for the asymptotic variances.

7 Case of zero canonical correlations

Suppose \( \rho_{r+1} = \ldots = \rho_p = 0 \). We write

\[
\mathbf{R} = \begin{pmatrix} \mathbf{R}_1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \mathbf{A} = (\mathbf{A}_1, \mathbf{A}_2), \quad \Gamma = (\Gamma_1, \Gamma_2).
\]

(7.1)

Then (2.4) leads to

\[
\begin{pmatrix} \Sigma_{YY} \Sigma_{XX}^{-1} \Sigma_{YX} \mathbf{A}_1, & \Sigma_{YY} \Sigma_{XX}^{-1} \Sigma_{YX} \mathbf{A}_2 \end{pmatrix} = \begin{pmatrix} \Sigma_{YY} \mathbf{A}_1 \mathbf{R}_1^2, & 0 \end{pmatrix},
\]

(7.2)

\[
\mathbf{A}' \Sigma_{YY} \mathbf{A} = \mathbf{I}.
\]

(7.3)

The matrix \( \mathbf{A}_2 \) has the indeterminacy of multiplication on the right by an arbitrary orthogonal matrix.
Let \( U = A'Y, \ V = \Gamma'X, \ W = A'Z, \) where \( A_2 \) has been made unique by imposing suitable conditions. Then
\[
\Sigma_{UU} = I, \quad \Sigma_{UV} = R, \quad \Sigma_{VV} = I, \quad \Sigma_{WW} = I - R^2.
\] (7.4)

We write
\[
S_{UV} S_{VV}^{-1} S_{VV} G = \left[ R^2 + R S_{VW} + S_{VW} R + R (S_{VV} - I) R + S_{WV} S_{VV}^{-1} S_{VV} \right] G \quad (7.5)
\]
\[
= S_{UU} G \tilde{R}^2 = \left[ R^2 + R S_{VW} + S_{VW} R + R (S_{VV} - I) R + S_{WV} \right] G \tilde{R}^2.
\]

Let \( G = M + (1/\sqrt{n}) G^* \), where
\[
M = \begin{pmatrix}
I & 0 \\
0 & M_2
\end{pmatrix}
\] (7.6)
and \( M_2 \) is an orthogonal matrix defined as follows. If the singular value decomposition of the \((p-r) \times (p-r)\) lower right-hand submatrix of \( G \) is \( G_{22} = E_1 D E_2 \), where \( E_1 \) and \( E_2 \) are orthogonal and \( D \) is diagonal, then \( M_2 = E_1 E_2 \). Note that \( M_2' G_{22} = G_{22}' M_2 \) and hence \( M_2' G_{22}^* = G_{22}' M_2 \).

Let \( \tilde{R}^2 = R^2 + (1/\sqrt{n}) Q \), where
\[
Q = \begin{pmatrix}
Q_1 & 0 \\
0 & \frac{1}{\sqrt{n}} Q_2
\end{pmatrix}
\] (7.7)

We write (7.5) as
\[
\left[ R^2 + R S_{VW} + S_{VW} R + R (S_{VV} - I) R \right] G \left( I - \tilde{R}^2 \right) \quad (7.8)
\]
\[
+ S_{WV} S_{VV}^{-1} S_{VV} G = S_{WW} G \tilde{R}^2.
\]

As before, we substitute \( G = M + (1/\sqrt{n}) G^* \) and \( \tilde{R}^2 = R^2 + (1/\sqrt{n}) Q \) into (7.8) to obtain the equation
\[
\sqrt{n} \left[ R (S_{VV} - I) R + R S_{VW} + S_{VW} R \right] \left[ M (I - R^2) + \frac{1}{\sqrt{n}} G^* (I - R^2) - \frac{1}{\sqrt{n}} M Q \right]
\]

32
\[ + R^2 \left[ G^* \left( I - R^2 \right) - MQ + \frac{1}{\sqrt{n}} G^* Q \right] + \sqrt{n} S_{WV} S_{VV}^{-1} S_{VW} M \]
\[ = \sqrt{n} \left[ S_{WW} - \left( I - R^2 \right) \right] \left[ MR^2 + \frac{1}{\sqrt{n}} MQ + \frac{1}{\sqrt{n}} G^* R^2 \right] \]
\[ + \left( I - R^2 \right) \left[ MQ + G^* R^2 + \frac{1}{\sqrt{n}} G^* Q \right] + o_p \left( \frac{1}{\sqrt{n}} \right). \] (7.9)

The first \( r \) columns of (7.9) are the same as the first \( r \) columns of (3.11) to order \( o_p(1) \) and can be solved for \( g_{ij}^* \), \( i = 1, \ldots, p, \ j = 1, \ldots, r, \ i \neq j \). The upper right-hand corner of (7.9) is

\[ \sqrt{n} R_{11} S_{VW}^{12} M_2 + R_{11}^2 G_{12}^* = 0 + o_p(1). \] (7.10)

This gives \( \sqrt{n} S_{VW}^{12} = -R_{11} G_{12}^* M_2' + o_p(1) \). The lower right-hand corner of (7.9) is

\[ n \left( S_{WV} S_{VV}^{-1} S_{VW} \right)_{22} = -\sqrt{n} S_{WV}^{21} R_{11} G_{12}^* M_2 + M_2 Q_2 M_2' + o_p(1) \]
\[ = n S_{WV}^{21} S_{VV}^{12} + M_2 Q_2 M_2' + o_p(1). \] (7.11)

Since \( S_{VV} \xrightarrow{P} I \).

\[ n \left( S_{WV} S_{VV}^{-1} S_{VW} \right)_{22} - n S_{WV}^{21} S_{VV}^{12} = n S_{WV}^{22} S_{VW}^{22} + o_p(1) \]
\[ = M_2 Q_2 M_2' + o_p(1). \] (7.12)

By the central limit theorem

\[ \text{vec} \sqrt{n} S_{WV}^{22} = \sqrt{n} \sum_{\alpha=1}^{n} v_{\alpha}^{(2)} \otimes w_{\alpha}^{(2)} \xrightarrow{P} N \left( 0, I \otimes I \right), \] (7.13)

and \( n S_{WV}^{22} S_{VW}^{22} \xrightarrow{P} W_{p_2}(p_2, I) \). The limiting distribution of the \( q_{i} = n r_{i}^2 \), \( i = r + 1, \ldots, p, \) is

\[ \frac{\sqrt{\frac{1}{2} p_2}}{2^{\frac{1}{2} p_2} p_2 \left( \frac{1}{2} p_2 \right)} \prod_{i=r+1}^{p} q_i^{\frac{1}{2}} \exp \left( -\frac{1}{2} \sum_{i=r+1}^{p} q_i \right) \prod_{i,j=r+1}^{p} (q_i - q_j), \] (7.14)

where \( p_2 = p - r \), and the distribution of \( M_2 \) is the Haar measure in the space of \( p_2 \)-order orthogonal matrices. See Anderson (1951b), (1984), and (1989).
From \( G' S_{UU} G = I \) we derive

\[
\sqrt{n} (S_{UU} - I)_{22} = - (G^*_2 M_2' + M_2 G^*_2) + o_p(1) = - 2G^*_2 M_2' + o_p(1). \tag{7.15}
\]

The density of the limiting distribution of \( X = \sqrt{n} (S_{UU} - I)_{22} \) is const \( \exp\left(-\frac{1}{4} \text{tr} X'X\right) \). The density of the limiting distribution of \( G^*_2 \) is const \( \exp\left(-\frac{1}{2} \text{tr} G^*_2 G^*_2\right) \).

The density of the limiting distribution of \( Y = \sqrt{n} S_{VS}^{12} \) is const \( \exp\left(-\frac{1}{2} \text{tr} Y'Y\right) \); the density of \( G^*_1 \) is const \( \exp\left(-\frac{1}{2} G^*_1 R^2 G^*_1\right) \).

The matrices \( H \) and \( \hat{R}^2 \) are treated similarly. In the development \( G, G^*, V, W, M_2 \) are replaced by \( H, H^*, U, T = V - RU, L_2 \), respectively. Note that \( W_2 = U_2 \) and \( T_2 = V_2 \) so that \( S_{UV}^{22} = S_{UV}^{22} \) and \( S_{TV}^{22} = S_{TV}^{22} \) in (7.12) and

\[
L_2 T_2^* L_2' = n S_{TV}^{22} S_{UT}^{22} + o_p(1) = n S_{UV}^{22} S_{VU}^{22} + o_p(1), \tag{7.16}
\]

as compared to \( M_2 T_2^* M_2' = n S_{UV}^{22} S_{UV}^{22} + o_p(1) \).

From (7.14) it follows that

\[
-2 \log \lambda = n \sum_{i=r+1}^{p} \log \left(1 + \frac{q_i}{n} \frac{q_i/n}{1-(q_i/n)}\right) \tag{7.17}
= -n \sum_{i=r+1}^{p} \log \left(1 - \frac{q_i}{n}\right)
\]

has a limiting \( \chi^2 \)-distribution with \( (p-r)^2 \) degrees of freedom.

The results of this section have been derived rather rather heuristically. A rigorous proof follows from Theorem 4.1 and Corollary 4.2 of Anderson (1987). In fact, from the theorem and corollary, one can obtain the asymptotic theory for roots of arbitrary multiplicities, specifically in the case that \( R \) is the block diagonal matrix with \( i \)-th block \( \lambda_i I_{pi}, i = 1, ..., q \), and \( \lambda_1 > \lambda_2 > ... > \lambda_q \geq 0 \).
Acknowledgment

The author has benefited from helpful comments of Yasuo Amemiya.

References


