PROCEEDINGS of the SECOND
BERKELEY SYMPOSIUM ON
MATHEMATICAL STATISTICS
AND PROBABILITY

Held at the Statistical Laboratory
Department of Mathematics
University of California
July 31–August 12, 1950

EDITED BY JERZY NEYMAN

UNIVERSITY OF CALIFORNIA PRESS
BERKELEY AND LOS ANGELES
1951
THE ASYMPTOTIC DISTRIBUTION OF CERTAIN CHARACTERISTIC ROOTS AND VECTORS

T. W. ANDERSON
COLUMBIA UNIVERSITY

1. Introduction

In a number of problems in multivariate statistical analysis use is made of characteristic roots and vectors of one sample covariance matrix in the metric of another. If \( A^* \) and \( D^* \) are the sample matrices, we are interested in the roots \( \phi^* \) of \( |D^* - \phi^*A^*| = 0 \) and the associated vectors satisfying \( D^*c^* = \phi^*A^*c^* \). In the cases we consider \( A^* \) and \( D^* \) have independent distributions. Each is distributed like a sum \( \sum_{a=1}^{q} y_ay'_a \) where \( y_1, \ldots, y_q \) are independently normally distributed with common covariance matrix. In the case of \( A^* \) the means of the vectors are zero; in the case of \( D^* \) the means may not be zero. We are interested in the asymptotic distribution of the characteristic roots and vectors when the number of vectors defining \( A^* \) increases indefinitely and when the means of the vectors defining \( D^* \) change in a certain way. The form of the limiting distribution depends on the multiplicity of the roots of a certain determinantal equation involving the parameters. If these roots are simple and different from zero, the asymptotic distribution is joint normal. If the roots are not simple, the asymptotic distribution is expressed in terms of "uniform distributions" on orthogonal matrices and a normal distribution.

We shall first state our problem in a general form and show in what kinds of statistical problems there is interest in these characteristic roots and vectors. Suppose \( x_a(a = 1, \ldots, N) \) of \( \rho \) components is normally distributed independently of \( x_\beta(a \neq \beta) \) with mean

\[
\mathbf{E}x_a = B_1z_{1a} + B_2z_{2a}
\]

and covariance

\[
\mathbf{E}(x_a - \mathbf{E}x_a)(x_a - \mathbf{E}x_a)' = \Sigma
\]

where \( z_{1a} \) and \( z_{2a} \) are vectors of fixed variates of \( q_1 \) and \( q_2 \) components, respectively, and \( B_1 \) and \( B_2 \) are \( \rho \times q_1 \) and \( \rho \times q_2 \) matrices, respectively. We shall use the notation \( N(B_1z_{1a} + B_2z_{2a}, \Sigma) \) for the distribution of \( x_a \).

Most of the research contained in this paper was done while the author was Fellow of the John Simon Guggenheim Memorial Foundation (at the Institute of Mathematical Statistics, University of Stockholm, and the Department of Applied Economics, University of Cambridge). The work was sponsored in part by the Office of Naval Research.

1 Unless specifically indicated otherwise, a vector is a column vector; a prime indicates the transpose of a vector or matrix. Vectors and matrices are indicated by bold face type.
On the basis of a sample \((x_1, z_{11}, z_{21}), \ldots, (x_N, z_{1N}, z_{2N})\) the usual estimate of \(B = (B_1B_2)\) is

\[
B = \sum_{a=1}^{N} x_a z_a' \left( \sum_{a=1}^{N} z_a z_a' \right)^{-1},
\]

where \(z_a' = (z_{1a}, z_{2a})\) and \(\sum_{a=1}^{N} z_a z_a'\) is assumed to be nonsingular. The columns of \(B\), say \(b_u\), are normally distributed with means \(\beta_u\), the corresponding columns of \(B\), and covariance

\[
E (b_u - \beta_u) (b_v - \beta_v)' = \left( \sum_{a=1}^{N} z_a z_a' \right)^{-1} \Sigma,
\]

where \(\left( \sum_{a=1}^{N} z_a z_a' \right)^{-1}\) indicates the element of the inverse matrix in the \(u\)-th row and \(v\)-th column.

Let \(Q^{-1}\) be the submatrix of \(\left( \sum_{a=1}^{N} z_a z_a' \right)^{-1}\) consisting of the last \(q_2\) rows and \(q_2\) columns; this is also given by

\[
Q = \sum_{a=1}^{N} z_{2a} z_{2a}' - \sum_{a=1}^{N} z_{2a} z_{1a}' \left( \sum_{a=1}^{N} z_{1a} z_{1a}' \right)^{-1} \sum_{a=1}^{N} z_{1a} z_{2a}'.
\]

Then \((B_2 - B_2') Q (B_2 - B_2)'\) has a Wishart distribution with covariance matrix \(\Sigma\) and \(q_2\) degrees of freedom, denoted by \(W(\Sigma, q_2)\). If \(q_2 < p\), this distribution, called *singular*, is the distribution of \(\sum_{u=1}^{q_1} y_u y_u'\) where \(y_u\) is distributed according to \(N(0, \Sigma)\) independently of \(y_v\) \((u \neq v)\). The usual estimate of \(\Sigma\) is

\[
A^* = \sum_{a=1}^{N} (x_a - B z_a) (x_a - B z_a)' = \sum_{a=1}^{N} x_a x_a' - B \sum_{a=1}^{N} z_a z_a' B',
\]

divided by \(N - (q_1 + q_2)\). This matrix is distributed according to \(W(\Sigma, N - q_1 - q_2)\) independently of \(B\).

Many statistical problems, for example [3], [6], involve the roots of

\[
|B_2 Q B_2' - \phi^* A^*| = 0,
\]

or the vectors, for example [1], [2], satisfying

\[
(B_2 Q B_2' - \phi^* A^*) c^* = 0.
\]

The \(p\) algebraically independent vectors \(c_g\) satisfying (1.8) may be normalized by

\[
c_g^* A^* c_h^* = n \delta_{gh},
\]

where \(n = N - q_1 - q_2\) and \(\delta_{gh} = 1\) and \(\delta_{gh} = 0\) for \(g \neq h\). We say that the solutions of (1.7) and (1.8) are the "characteristic roots and vectors of \(B_2 Q B_2'\) in the metric of \(A^*". If we wish to test the hypothesis that the rank of \(B_2\) is \(r\) against the alternatives that it is greater than \(r\) we use the \(p - r\) smallest roots of (1.7).

If we assume that the rank of \(B_2\) is \(r\) and we wish to estimate \(B_2\) (or, equivalently,
estimate the linear restrictions on \( B_2 \) we make use of the vectors \( e^* \) satisfying (1.8) for the \( p - r \) smallest roots of (1.7).

In this paper we shall study the joint asymptotic distribution of the roots and vectors defined by (1.7), (1.8), and (1.9) when \( n = N - q_1 - q_2 \to \infty \) and 
\[
\frac{1}{n} \sum_{i=1}^{N} z_i z'_i \text{ approaches a nonsingular limit. The asymptotic distribution of the roots alone has been given by Hsu [8]. We find it convenient to make use of some of the results in [8] to obtain the joint asymptotic distribution of roots and vectors; however, the method used in the present paper could be used independently of [8]. We shall assume throughout the paper that } q_2 \geq p.
\]

2. Reduction of the problem to canonical form

To simplify the following derivations we shall transform the matrices \( B_2 \beta B_2^t \) and \( A^* \) so that they have distributions with fewer parameters. Corresponding to (1.7) and (1.8) in the sample, we have the population equations
\[
(2.1) \quad |B_2 \bar{Q}_n B_2 - \tau \Sigma| = 0
\]
and
\[
(2.2) \quad (B_2 \bar{Q}_n B_2 - \tau \Sigma) \gamma = 0,
\]
where \( \bar{Q}_n = \frac{1}{n} Q \). Let the roots of (2.1) be \( \tau_1(n) \geq \tau_2(n) \geq \ldots \geq \tau_p(n) \geq 0 \). The number of zero roots is the difference of \( p \) and the rank of \( B_2 \) for each \( n \) for which \( \bar{Q}_n \) is nonsingular (in particular for \( n \) sufficiently large). Let \( \gamma_1(n), \ldots, \gamma_p(n) \) be a set of corresponding solutions of (2.2) satisfying
\[
(2.3) \quad \gamma'_j(n) \Sigma \gamma_h(n) = \delta_{jh}.
\]
Let \( \Gamma_n = [\gamma_1(n), \ldots, \gamma_p(n)] \). Then we can make a transformation, for example [7], so that \( A^* \) is replaced by
\[
(2.4) \quad A_n = \sum_{\beta=1}^{n} y_\beta^* y_\beta^*',
\]
where \( y_\beta^* \) is distributed according to \( N(0, I) \) independently of \( y_\alpha^* (\beta \neq \alpha) \), and \( B_2 \bar{Q}_n B_2^t \) is replaced by
\[
(2.5) \quad D_n = \sum_{\gamma=1}^{n} y_\gamma^* y_\gamma^* (n)
\]
where \( y_\gamma^* (n) \) is distributed independently of \( y_h^* (n) (g \neq h) \) according to \( N(\sqrt{n} \tau_\gamma(n) \varepsilon_\gamma, I) \) where \( \tau_\gamma(n) \) is the nonnegative square root of \( \tau_\gamma^2(n) \) and \( \varepsilon_\gamma \) is a vector with all components 0 except the \( g \)-th (for \( g \leq p \) which is 1. The roots of (1.7) are the roots of
\[
(2.6) \quad |D_n - \phi A_n| = 0,
\]
and the vectors satisfying (1.8) and (1.7) are related to the vectors \( c_1(n), \ldots, c_p(n) \) satisfying
\[
(2.7) \quad (D_n - \phi A_n^*) c = 0
\]
and
\begin{equation}
(2.8) \quad c'_n A_n c_n = n \delta_{\phi_n}
\end{equation}

by
\begin{equation}
(2.9) \quad c_n^* (n) = T_n c_n.
\end{equation}

It should be observed that $T_n$ and $\tau_2(n)$ depend on $n$ because $Q_n$ depends on $n$.

We shall first find the limiting distribution of $c \phi(n)$ and $\phi_0(n)$ as $n \to \infty$ (that is, as $N \to \infty$). Let $y_0 = y_0^*(n) - \sqrt{n} \tau_\phi(n) \varepsilon_\phi$. Then
\begin{equation}
(2.10) \quad D_n = \sum_{i=1}^{q_2} [y_0 + \sqrt{n} \tau_\phi(n) \varepsilon_\phi] [y_0 + \sqrt{n} \tau_\phi(n) \varepsilon_\phi]',
\end{equation}

and $y_0$ is distributed according to $N(0, I)$.

Let $C_n = [c_1(n), \ldots, c_p(n)]$. Then (2.7) can be written
\begin{equation}
(2.11) \quad D_n C_n^* = A_n^* \Phi_n,
\end{equation}

where $\Phi_n = [\phi_i(n) \delta_{ij}]$ and $\phi_1(n), \ldots, \phi_p(n)$ are the roots of (2.6) and (2.8) can be written
\begin{equation}
(2.12) \quad C' A_n C_n = n I.
\end{equation}

If
\begin{equation}
(2.13) \quad X_n = C_n^{-1},
\end{equation}

we have
\begin{equation}
(2.14) \quad \frac{1}{n} A_n = X_n' X_n,
\end{equation}
\begin{equation}
(2.15) \quad \frac{1}{n} D_n = X_n' \Phi_n X_n.
\end{equation}

We shall set out to find the limiting distribution of $\Phi_n$ and $X_n$ for $\tau_\phi(n)$ approaching limits as $n \to \infty$. To make $\Phi_n$ and $X_n$ unique we require $\phi_1(n) > \phi_2(n) > \ldots > \phi_p(n)$ and $x_{ii}(n) > 0$. The probability is 0 of a $D_n$ and $A_n$ for which $X_n$ and $\Phi_n$ are not uniquely defined.

Throughout this paper we shall make use of the following special case of a theorem of Rubin [9]:

**Rubin's Theorem:** Let $F_n(u)$ be the cumulative distribution function of a random vector $u_n$. Let $v_n$ be a (vector valued) function of $u_n$, $v_n = f_n(u_n)$, and let $G_n(v)$ be the (induced) distribution of $v_n$. Suppose $\lim F_n(u) = F(u)$ [in every continuity point of $F(u)$] and suppose for every continuity point $u$ of $f(u)$, $\lim f_n(u_n) = f(u)$, when $\lim u_n = u$. Let $G(v)$ be the distribution of the random vector $v = f(u)$, where $u$ has the distribution $F(u)$. If the probability of the set of discontinuities of $f(u)$ in terms of $F(u)$ is 0, then
\begin{equation}
\lim_{n \to \infty} G_n(v) = G(v).
\end{equation}

We could justify the limiting procedures by another method that consists of extending a theorem of L. C. Young ("Limits of Stieljes integrals," *Jour. London Math. Soc.*, Vol. 9 [1934], pp. 119-126), concerning the limit of $\int f_n(u) dF_n(u)$, applying this to the characteristic function of $f_n(u_n)$, and thus obtaining a restricted form of Rubin's theorem.
CHARACTERISTIC ROOTS AND VECTORS

In our case the components of \( u_n \) are linear combinations of the components of the matrices \( A_n \) and \( D_n \); the components of \( v_n \) are linear combinations of the characteristic roots and the components of the characteristic vectors. The distribution of \( u_n \) approaches a limit and the function \( f_n(u) \) approaches a limit (in the above sense). We shall verify that the discontinuities of the limiting function are of limiting probability zero. Thus we can deduce the asymptotic distribution of the characteristic roots and (normalized) vectors by using the asymptotic distribution of \( A_n \) and \( D_n \) and the limiting function.

3. Derivation of two special distributions

In order to derive the desired asymptotic distributions we need to obtain the distributions of the characteristic roots and vectors (in the metric of \( I \)) of a symmetric matrix \( B \) in two special cases. Let the roots of

\[
|B - \psi I| = 0
\]

be \( \psi_1 \geq \psi_2 \geq \ldots \geq \psi_p \). Let the characteristic vector satisfying

\[
Bh = \psi_i h,
\]

and \( h'h = 1 \) be \( h_i \) (\( i = 1, \ldots, p \)). If \( \psi_1, \ldots, \psi_p \) are different \( h_1, \ldots, h_p \) are uniquely defined except for multiplication of a vector by \(-1\), and \( h_i h_j = 0 \), \( i \neq j \). Let \( H = (h_1, \ldots, h_p) \). Then \( BH = H\Psi \), where \( \Psi = (\psi_i \delta_{ij}) \). Let \( H' = G \). Then \( G \) satisfies

\[
G'\Psi G = B,
\]

\[
G'G = I.
\]

These equations define \( \Psi \) and \( G \) uniquely if we require \( g_{ii} \geq 0 \) except for a set of \( B \) of measure zero. Since it is trivial to obtain the distribution of \( H \) and \( \Psi \) from that of \( G \) and \( \Psi \), we shall now obtain the distribution of \( G \) and \( \Psi \).

First we consider the case that the distribution of \( B \) is \( W(I, m) \) (\( m \geq p \)); that is, the density is

\[
C(m, p) |B|^{(m-p-1)/2} e^{-\text{tr} B/2}
\]

where

\[
C^{-1}(m, p) = 2^{mp/2} \pi^{p(p-1)/4} \prod_{i=1}^p \Gamma \left( \frac{1}{2} (m + 1 - i) \right)
\]

and "\( \text{tr} \)" denotes trace. This is the distribution of

\[
B = \sum_{j=1}^n u_j u_j',
\]

where \( u_1, \ldots, u_m \) are independently distributed according to \( N(0, I) \).

**Theorem 1.** Let \( B \) have the distribution \( W(I, m) \). Then \( G \) and \( \Psi \), defined by (3.3), (3.4), the restriction that \( \Psi \) is diagonal with diagonal elements in descending order
and \( g_{ij} \geq 0 \), are independently distributed. The density of the diagonal elements of \( \Psi \) is

\[
(3.8) \quad \pi^{p/2} \prod_{i=1}^{p} \psi^{i(m-1)/2} e^{-\sum_{j=1}^{p} \psi^{j}/2} \]

\[
2^{p^{2}/2} \prod_{i=1}^{p} \left\{ \Gamma \left( \frac{1}{2} (m + 1 - i) \right) \Gamma \left( \frac{1}{2} (p + 1 - i) \right) \right\} \times \prod_{i} \prod_{j=i+1}^{p} \left( \psi_{i} - \psi_{j} \right)
\]

for \( \psi_{1} \geq \ldots \geq \psi_{p} > 0 \) and is 0 elsewhere. The distribution of \( G \) is "uniform."

Proof. That the marginal density of \( \psi_{1}, \ldots, \psi_{p} \) is (3.8) has been proved by Hsu [5]. It remains to show that \( G \) is distributed independently of \( \Psi \) and "uniformly." The "uniform" distribution of all orthogonal \( p \)-dimensional matrices is given by the (normalized) Haar measure on the orthogonal group; that is, the (normalized) Haar measure is the only probability measure on the group that is invariant under the group operation on the right [4]. Since we require \( g_{ii} \geq 0 \), our definition of "uniform distribution" is the conditional distribution obtained from the Haar measure by requiring \( g_{ii} \geq 0 \). For this part of the space the probability measure is \( 2^{p} \) times the normalized Haar measure.

The measure on the space of \( \mathbf{u} \) defines a measure on the space of \( G, g_{ii} \geq 0 \). Consider any measurable set \( H \) in the space of all orthogonal matrices. Let the diagonal matrices with diagonal elements \( +1 \) and \( -1 \) be \( J_{1}, \ldots, J_{p} \). Let \( H = \sum_{i=1}^{p} H_{i} \), where \( J_{i}H_{i} \) is a set in the space of \( G, g_{ii} \geq 0 \). Define the measure of \( H \) as the sum of the measures of \( J_{i}H_{i} \). Now let us show that this measure is invariant with respect to multiplication on the right. Let \( E_{i} \) be the set in the space of \( \mathbf{u} \) that maps into \( J_{i}H_{i} \). Let \( H^{*} = HP_{i} \) that is, \( H^{*} \) is the set obtained by multiplying each element of \( H \) on the right by the orthogonal matrix \( P \). Then \( H^{*} = \sum H^{*}_{i} = \sum \mathbf{H}H_{i}P \). We now show that the measure of \( H^{*}_{i} \) is the same as \( H_{i} \). Let \( H^{*}_{i} = \sum_{i} H^{*}_{ij} \) such that \( J_{i}H^{*}_{ij} \) is in the space \( g_{ii} \geq 0 \). Let \( E^{*}_{ij} \) be the set in the space \( \mathbf{u} \) that maps into \( J_{i}H^{*}_{ij} \). Then \( \sum E^{*}_{ij} = P^{*}E_{i} \); that is, \( \sum E^{*}_{ij} \) is the set obtained by multiplying each \( (\mathbf{u}_{1}, \ldots, \mathbf{u}_{m}) \) by \( P^{*} \) on the left. The measure of \( P^{*}E_{i} \) is the integral of the density of \( P^{*}\mathbf{u}_{1}, \ldots, P^{*}\mathbf{u}_{m} \) over \( E_{i} \). Since the density of \( P^{*}\mathbf{u}_{1}, \ldots, P^{*}\mathbf{u}_{m} \) is the same as that of \( \mathbf{u}_{1}, \ldots, \mathbf{u}_{m} \), the measure of \( P^{*}E_{i} \) is that of \( E_{i} \). Thus the measure of \( H^{*} \) is that of \( H \). This proves that the measure is invariant with regard to the group operation on the right. Since there is only one such measure on the group of orthogonal matrices with total measure \( 2^{p} \), this is it. The joint distribution of \( \Phi \) and \( G(g_{ii} \geq 0) \) has a density. This density does not depend on \( G \) because the density at \( \Phi \) and \( G \) is the same as at \( \Phi \) and \( G^{*} \) since \( G^{*} \) can be obtained from \( G \) by multiplication on the right by some orthogonal matrix \( P \) and this is equivalent to transforming \( B \) to \( P^{*}BP \) which has the same characteristic roots as \( B \). This proves the theorem.

Now suppose the density function of \( B = B' \) is

\[
(3.9) \quad \pi^{-p(p+1)/4} \pi^{-p/2} e^{-t$2B'/2$
\]
that is, \( b_{ij} (i = 1, \ldots, p; j = i, i + 1, \ldots, p) \) are independently and normally distributed with means zero; the variance of \( b_{ii} \) is 1 and that of \( b_{ij} (i < j) \) is \( \frac{1}{2} \).

Now define \( G \) and \( \Psi \) (diagonal) by (3.3) and (3.4) with the understanding that the elements of the first column of \( G \) are nonnegative. The ordered roots \( \psi_i \) are not restricted to being nonnegative.

**Theorem 2.** Let the symmetric matrix \( B \) have the distribution with density (3.9). Then \( G \) and \( \Psi \), defined by (3.3), (3.4), the restriction that \( \Psi \) is diagonal with diagonal elements in descending order and \( q_{ii} \geq 0 \), are independently distributed. The density of the diagonal elements of \( \Psi \) is

\[
2^{-p/2} \left\{ \prod_{i=1}^{p} \Gamma \left( \frac{1}{2} (p + 1 - i) \right) \right\}^{-1} e^{-\sum_{i=1}^{p} \psi_{ii}^{1/2} \prod_{j=i+1}^{p} (\psi_i - \psi_j)}
\]

for \( \psi_1 \geq \ldots \geq \psi_p \) and 0 elsewhere. The distribution of the orthogonal matrix \( G \) is uniform.

**Proof.** The proof that the marginal density of \( \psi_1, \ldots, \psi_p \) is (3.10) has been given by Hsu [8]. The remainder of the proof is the same as for theorem 1 since the density of \( P'BP \) for \( P \) orthogonal is the same as \( B \).

**4. An asymptotic distribution when all population roots are zero**

A simple case of our main problem is the case where \( \tau_i^2(n) = 0 \) for all \( g \) and \( n \). Then \( D_n = D \) has a Wishart distribution with \( q_2 (\geq p) \) degrees of freedom which does not depend on \( n \). In this section we shall find the asymptotic distribution of \( X_n \) and \( \Phi_n \) in this special case.

In all of the asymptotic theory we use the result [8] that as \( n \to \infty \)

\[
U_n = \frac{1}{\sqrt{n}} (A_n - nI)
\]

is asymptotically normally distributed with mean zero. The functionally independent variables are statistically asymptotically independent and the variances are given by

\[
\mathcal{E} u_{ii}^2 = 2, \quad \mathcal{E} u_{ij}^2 = 1, \quad i \neq j.
\]

The matrices \( X_n \) and \( \Phi_n \) are defined by

\[
\frac{1}{n} D = X_n' \Phi_n X_n,
\]

\[
\frac{1}{n} A_n = X_n' X_n,
\]

where \( x_{ii}(n) \geq 0, \Phi_n \) is diagonal and the diagonal elements of \( \Phi_n \) are labelled in descending order. For each \( n, X_n \) and \( \Phi_n \) are defined uniquely except on a set of probability zero.

As \( n \to \infty \), \( \frac{1}{n} A_n \) approaches the stochastic limit \( I \) and \( \frac{1}{n} D \) approaches the stochastic limit \( 0. \) In the limit \( X_n \) must satisfy

\[
I = X'X,
\]

and \( \Phi_n \) must approach \( 0 \) stochastically.
To obtain the full asymptotic theory we define new matrices $W_n$, $Z_n$ and $\Theta_n$. For any matrix $X_n$ we have an orthogonal matrix $O_n$ and a diagonal matrix $\Delta_n$ defined by

\[(4.6) \quad X_n^TX_n = O_n\Delta_nO_n,\]

where the diagonal elements of $\Delta$ are ordered in descending size and $o_{1}(n) \geq 0$. Let

\[(4.7) \quad G_n = O_n\Delta_n^{1/2}O_n,\]

where the elements of $\Delta_n^{1/2}$ are the positive square roots of the corresponding elements of $\Delta_n$ (the roots are different from 0 when $X_n^TX_n$ is nonsingular). Let

\[(4.8) \quad W_n = X_nG_n^{-1}.\]

This is an orthogonal matrix; that is,

\[(4.9) \quad W_n^TW_n = I.\]

Let

\[(4.10) \quad Z_n = \sqrt{n}W_n(G_n - I).\]

Then

\[(4.11) \quad X_n = W_nG_n = W_n\left(I + W_n\frac{1}{\sqrt{n}}Z_n\right) = W_n + \frac{1}{\sqrt{n}}Z_n.\]

We notice that

\[(4.12) \quad W_n^TZ_n = Z_n^TW_n,\]

because

\[(4.13) \quad W_n^TZ_n = \sqrt{n}W_n^TW_n(G_n - I) = G_n - I = G_n' - I = \sqrt{n}(G_n' - I)W_n^TW_n = Z_nW_n.\]

Now let us show that (4.9), (4.11), and (4.12) define $W_n$ and $Z_n$ in terms of $X_n$ (except for a set of measure 0). We have

\[(4.14) \quad X_n = W_nO_n\Delta_n^{1/2}O_n.\]

Let $W^*$ be another matrix satisfying (4.9), (4.11) and (4.12), with possibly a different $Z_n$. Then

\[(4.15) \quad W^*X_n = X_n^TW^*,\]

\[(4.16) \quad X_nW^* = W^*X_n'.\]

Equation (4.15) is

\[(4.17) \quad W^*W_nO_n\Delta_n^{1/2}O_n = O_n\Delta_n^{1/2}O_nW_n^TW^*.\]

From this we derive

\[(4.18) \quad O_nW^*W_nO_n\Delta_n^{1/2} = \Delta_n^{1/2}O_nW_n'W^*O_n'.\]

Let

\[(4.19) \quad O_nW^*W_nO_n' = O*.\]

Then

\[(4.20) \quad O*\Delta_n^{1/2} = \Delta_n^{1/2}O*.\]
The component equations are
\( (4.21) \quad \delta^*_{ij} \theta^{1/2}_{ij} = \theta^{1/2}_{ij} \delta^*_{ij} \). 
This gives us
\( (4.22) \quad \delta^*_{ij} = \frac{\delta^{1/2}_{ij}}{\theta^{1/2}_{ij}} \delta^*_{ij} \).
From (4.16) we derive
\( (4.23) \quad \delta^*_{ij} = \frac{\delta^{1/2}_{ij}}{\theta^{1/2}_{ij}} \delta^*_{ij} \).
If \( \delta_i \neq \delta_j, \delta^*_{ij} = 0 \). Therefore, if the \( \delta_i \) are all different
\( (4.24) \quad \delta^* = I \),
and
\( (4.25) \quad W^* = W \).
Therefore, except for a set of measure zero of \( X_n \), (4.9), (4.11), and (4.12) define \( W_n \) and \( Z_n \) uniquely. Let
\( (4.26) \quad \Theta_n = n\Phi_n \).
Now let us substitute into (4.3) and (4.4). We obtain
\( (4.27) \quad D = W'_n \Theta_n W_n + \frac{1}{\sqrt{n}} (Z'_n \Theta_n W_n + W_n \Theta_n Z_n) + \frac{1}{n} Z'_n \Theta_n Z_n \),
\( (4.28) \quad U_n = W'_n Z_n + Z'_n W_n + \frac{1}{\sqrt{n}} Z'_n Z_n \).
Together with (4.9), (4.12) and
\( (4.29) \quad w_{i1} (n) + \frac{1}{\sqrt{n}} z_{i1} (n) \geq 0, \quad i = 1, \ldots, p \),
(4.27) and (4.28) define \( \Theta_n, W_n \) and \( Z_n \) uniquely for each \( n \).
For given \( W_n = W, Z_n = Z \) and \( \Theta_n = \Theta \) the limits of (4.27) and (4.28) expressing \( D \) and \( U \) in terms of \( W_n, Z_n \) and \( \Theta_n \) are
\( (4.30) \quad D = W'\Theta W \),
\( (4.31) \quad U = W'Z + Z'W = 2W'Z \).
If
\( (4.32) \quad w_{i1} \geq 0 \),
and \( \theta_i > \theta_j \) for \( i > j \), then (4.9), (4.12), (4.30) and (4.31) define \( W, Z \) and \( \Theta \) uniquely in terms of \( D \) and \( U \) (except for a set of \( D \) and \( U \) of measure 0). Now we wish to argue that if we take (4.9), (4.12), (4.27), (4.28), and (4.29) as defining \( W_n \) (diagonal) \( \Theta_n \), \( Z_n \) in terms of (nonrandom) \( D = D_n \) and \( U_n \), the limit of \( W_n, \Theta_n \) and \( Z_n \) is the solution of (4.9), (4.12), (4.30), (4.31), and (4.32) as \( n \to \infty \) for \( D_n \to D \) and \( U_n \to U \) where \( D \) and \( U \) are such that the solution is
unique (the exceptional $D$ and $U$ are of measure 0). A diagonal element of $\Theta_n$ is a root of

$$D_n - \theta \left( I + \frac{1}{\sqrt{n}} U_n \right) = 0. \tag{4.33}$$

As $n \to \infty$, this root approaches the root of

$$|D - \theta I| = 0, \tag{4.34}$$

and this is an element of $\Theta$ defined by (4.9) and (4.30). $Z_n$ is defined (equivalently) by

$$Z_n = \sqrt{n} (X_n - W_n) = X_n' \sqrt{n} (I - \Delta_n^{-1/2}) O_n, \tag{4.35}$$

where the diagonal elements of $\Delta_n$ are roots of

$$\left| \frac{1}{\sqrt{n}} U_n + I - \delta I \right| = 0. \tag{4.36}$$

Let $\psi_i(n)$ be the $i$-th root of

$$|U_n - \psi I| = 0. \tag{4.37}$$

Then

$$\delta_i(n) = 1 + \frac{1}{\sqrt{n}} \psi_i(n). \tag{4.38}$$

Clearly

$$\lim_{n \to \infty} \sqrt{n} \left[ 1 - \delta_i^{-1/2}(n) \right] = \frac{1}{2} \lim_{n \to \infty} \psi_i(n) = \frac{1}{2} \psi_i. \tag{4.39}$$

Since $X_n X_n \to I$ and $O_n$ is orthogonal, each element of $Z_n$ is bounded in the limit. Thus the norm (any standard norm) of $\frac{1}{\sqrt{n}} Z_n Z_n$ and the norm of $\frac{1}{\sqrt{n}} (Z_n' \Theta_n W_n + W_n' \Theta_n Z_n) + \frac{1}{n} Z_n' \Theta_n Z_n$ go to zero as $n \to \infty$. Thus each element of $D_n - W_n' \Theta_n W_n$ and each element of $U_n - 2W_n' Z_n$ go to 0 as $n \to \infty$. Consider the matrix function $(P, Q) = (D - W^* \Theta^* W^*, U - 2W^* Z^*)$, where $W^*$ and $\Theta^*$ satisfy our usual conditions including (4.32). The inverse functions $W^*, \Theta^*, Z^*$ (as functions of $P$ and $Q$) are continuous in the proper domain (except on the exceptional set). Hence, if the norm of $(P, Q)$ is sufficiently small the norm of $(W^* - W, \Theta^* - \Theta, Z^* - Z)$ must be arbitrarily small. If $w_{ii} > 0$, then $w_{ii} > 0$ for norm of $(P, Q)$ sufficiently small. Then $w_{ii}(n)$ for $n$ sufficiently large is bounded away from 0, and for $n$ sufficiently large $w_{ii}(n)$ satisfying (4.29) must satisfy (4.32). Thus $W_n, \Theta_n, Z_n$ converge to $W, \Theta, Z$ defined by (4.30) and (4.31).

The limiting equations (4.30) and (4.31) define $W, Z,$ and $\Theta$ uniquely except on a set of Lebesgue measure zero. The discontinuities can only occur on this set. Now considering $D$ and $U_n$ as random matrices we observe that the limiting distribution of $D$ and $U_n$ is absolutely continuous. Thus the conditions of Rubin's theorem are fulfilled. To obtain the limiting distribution of the random matrices $W_n, Z_n$ and $\Theta_n$ defined in terms of the random matrices $D$ and $U_n$ we need only find the distribution of $W, Z$ and $\Theta$ defined by (4.9), (4.12), (4.30), (4.31) and (4.32), where $U$ has the limiting distribution of $U_n$.

The distribution of $W$ and $\Theta$ is that of theorem 1. The conditional distribution
of \( Z \) given \( W \) and \( \Theta \) is obtained from
\[
(4.40) \quad Z = \frac{1}{2} W U.
\]
Thus
\[
(4.41) \quad E[Z \mid W] = \frac{1}{2} W E U = 0.
\]
Let
\[
(4.42) \quad U = (u_1, \ldots, u_p),
Z = (z_1, \ldots, z_p),
W = (w_1, \ldots, w_p).
\]
Then
\[
(4.43) \quad E[z_i z'_j \mid W] = \frac{1}{2} W E(u_i u'_j) W.
\]
Since \( E u_i^2 = 2 \) and \( E u_i u_j = 1 \) for \( i \neq j \), and \( E u_i u_{k1} = 0 \) otherwise, then
\[
(4.44) \quad E u_i u'_i = I + \epsilon_{ii},
E u_i u'_j = \epsilon_{ji}
\]
where \( \epsilon_{ij} \) is a matrix with 1 in the \( i \)-th row and \( j \)-th column and 0's elsewhere. Thus
\[
(4.45) \quad E[z_i z'_j \mid W] = \frac{1}{2} W(I \delta_{ij} + \epsilon_{ii}) W' = \frac{1}{2}(I \delta_{ij} + w_i w'_j).
\]
Since \( U \) is normally distributed, the conditional distribution of \( Z \) is normal.

**Theorem 3.** Let \( D \) have the distribution \( W(I, q_2) \), \( q_2 \geq p \), and let \( A_n \) be independent of \( D_n \) and have the distribution \( W(I, n) \). Define \( X_n \) and \( \Phi_n \) by means of (4.3), (4.4) and the conditions that \( x_{ii}(n) \geq 0 \) and \( \Phi_n \) is diagonal with diagonal elements in descending order. Let \( n \Phi_n = \Theta_n \) and let \( X_n = W_n + \frac{1}{\sqrt{n}} Z_n \), where \( W_n Z_n = I \) and \( W_n Z_n = Z'_n W_n \). The limiting distribution of \( \Theta_n, W_n \) and \( Z_n \) as \( n \to \infty \) is the joint distribution of \( \Theta, W, \) and \( Z \) such that the marginal distribution of the diagonal matrix \( \Theta \) and the orthogonal matrix \( W \) is that of theorem 1 with \( m = q_2 \) and the conditional distribution of \( Z \) given \( W \) and \( \Theta \) is normal with mean 0 and covariances given by (4.45).

5. An asymptotic distribution when all population roots are equal but different from zero

Another special case that is easy to treat is the case of all roots of (2.1) being equal but different from 0, say, \( \tau_1^2(n) = \ldots = \tau_p^2(n) = \lambda_n > 0 \). Then
\[
(5.1) \quad D_n = F + \sqrt{n} E_n + n \lambda_n I,
\]
where
\[
(5.2) \quad F = \sum_{g=1}^{q_2} y_g y'_g,
\]
and \( E_n \) is composed of elements
\[
(5.3) \quad \sqrt{\lambda_n}(y_{ij} + y_{ji}).
\]
We are interested in $X_n$ and $\Phi_n$ (diagonal) defined by

\begin{equation}
\frac{1}{n} F + \frac{1}{\sqrt{n}} E_n + \lambda_n I = X'_n \Phi_n X_n,
\end{equation}

\begin{equation}
\frac{1}{n} A_n = X'_n X_n,
\end{equation}

with $x_{n(n)} \geq 0$ and the diagonal elements of $\Phi_n$ arranged in descending order.

Let

\begin{equation}
X_n = W_n + \frac{1}{\sqrt{n}} Z_n,
\end{equation}

where $W_n$ and $Z_n$ satisfy (4.9) and (4.12). Let

\begin{equation}
\Phi_n = \lambda_n I + \frac{1}{\sqrt{n}} \Theta_n,
\end{equation}

where $\Theta_n$ is diagonal. Then (5.4) and (5.5) are

\begin{equation}
\frac{1}{n} F + \frac{1}{\sqrt{n}} E_n + \lambda_n I = \left(W_n + \frac{1}{\sqrt{n}} Z_n\right)' \left(\lambda_n I + \frac{1}{\sqrt{n}} \Theta_n\right) \left(W_n + \frac{1}{\sqrt{n}} Z_n\right)
= \lambda_n I + \frac{1}{\sqrt{n}} \left[\lambda_n (Z'_n W_n + W'_n Z_n) + W'_n \Theta_n W_n\right]
+ \frac{1}{n} \left[W'_n \Theta_n Z_n + Z'_n \Theta_n W_n + \lambda_n Z'_n Z_n\right] + \frac{1}{n^{3/2}} Z'_n \Theta_n Z_n,
\end{equation}

(5.9)

\begin{equation}
\frac{1}{n} A_n = \left(W_n + \frac{1}{\sqrt{n}} Z_n\right)' \left(W_n + \frac{1}{\sqrt{n}} Z_n\right);
\end{equation}

that is,

\begin{equation}
\frac{1}{\sqrt{n}} (A_n - nI) = (W'_n Z_n + Z'_n W_n) + \frac{1}{\sqrt{n}} Z'_n Z_n.
\end{equation}

Multiply (5.10) by $\lambda_n$ and subtract from $\sqrt{n}$ times (5.8) to obtain

\begin{equation}
\frac{1}{\sqrt{n}} F + E_n - \lambda_n (A_n - nI) = W'_n \Theta_n W_n + \frac{1}{\sqrt{n}} [W'_n \Theta_n Z_n + Z'_n \Theta_n W_n]
+ \frac{1}{n} Z'_n \Theta_n Z_n.
\end{equation}

Let

\begin{equation}
U_n = \frac{1}{\sqrt{n}} (A_n - nI).
\end{equation}

Then (5.10) and (5.11) can be written as

\begin{equation}
U_n = (W'_n Z_n + Z'_n W_n) + \frac{1}{\sqrt{n}} Z'_n Z_n,
\end{equation}

\begin{equation}
\frac{1}{\sqrt{n}} F + E_n - \lambda_n U_n = W'_n \Theta_n W_n + \frac{1}{\sqrt{n}} [W'_n \Theta_n Z_n + Z'_n \Theta_n W_n]
+ \frac{1}{n} Z'_n \Theta_n Z_n,
\end{equation}

where

\begin{equation}
W'_n W_n = I,
\end{equation}

\begin{equation}
W'_n Z_n = Z'_n W_n.
\end{equation}
Thus for a given $n$, $\Theta_n$, $W_n$ and $Z_n$ are defined as functions of $U_n$, $F = F_n$ and $E_n$. The functions are unique and continuous except over a set of $U_n$, $F_n$ and $E_n$ of measure zero. The limit of the functions (as $U_n \to U$, $F_n \to F$, and $E_n \to E$) is the solution to

\begin{align}
    U &= W'Z + Z'W = 2W'Z, \\
    E - \lambda U &= W'\Theta W,
\end{align}

with $W$ satisfying (4.32) and where $\lambda = \lim_{n \to \infty} \lambda_n$. This argument is justified as in section 3. In particular, each diagonal element of $\Theta_n$ as a function of nonrandom $F_n$, $E_n$ and $U_n$ is a root of

\begin{align}
    \left( \frac{1}{\sqrt{n}} F_n + E_n + \sqrt{n}\lambda_n I - (U_n + \sqrt{n}I) \left( \lambda_n + \frac{1}{\sqrt{n}} \theta \right) \right) &= 0,
\end{align}

that is, of

\begin{align}
    \left( \frac{1}{\sqrt{n}} F_n + E_n - \lambda_n U_n \right) \left( I + \frac{1}{\sqrt{n}} U_n \right)^{-1} - \theta I &= 0.
\end{align}

Since

\begin{align}
    \left( \frac{1}{\sqrt{n}} F_n + E_n - \lambda_n U_n \right) \left( I + \frac{1}{\sqrt{n}} U_n \right)^{-1} \to E - \lambda U,
\end{align}

the ordered roots of (5.20) approach the ordered roots of

\begin{align}
    |E - \lambda U - \theta I| &= 0.
\end{align}

As in section 4 we can argue that the elements of $W_n$, $\Theta_n$ and $Z_n$ are bounded for $F_n \to F$, $E_n \to E$ and $U_n \to U$. Then the elements in (5.13) and (5.14) which are multiplied by $1/\sqrt{n}$, $1/n$ and $1/n^{3/2}$ go to zero. The remainder of the argument of section 4 applies.

We can now use Rubin's theorem since the discontinuities of the mapping occur where there is indeterminacy and the set of $E$ and $U$ where this occurs is of limiting probability 0. To find the limiting distribution of the random matrices $W_n$, $\Theta_n$, and $Z_n$ we need only consider the distribution of the random matrices $W$, $\Theta$, and $Z$ defined by (5.18), (5.19) and (4.32) (with $W$ orthogonal, $\Theta$ diagonal with diagonal elements in descending order), where the random matrices $E$ and $U$ have the limiting distribution of $E_n$ and $U_n$. Let

\begin{align}
    E - \lambda U &= V.
\end{align}

The density of $E$ and $U$ is

\begin{align}
    k e^{-\lambda^2 E^2/(4\lambda) + \lambda U^2/2} &= k e^{-\text{tr} \left( (V + \lambda U) + 2\lambda U^2 \right)/(8\lambda)} \\
    &= k e^{-\text{tr} \left(V + 2\lambda UV + 2\lambda U^2 + 2\lambda^2 U \right)/(8\lambda)} \\
    &= k e^{-\text{tr} \left(V + (2\lambda^2 + 2\lambda^2) \right) (U + (2\lambda^2 + 2\lambda^2) V) + (2\lambda^2 + 2\lambda^2) V + (2\lambda^2 + 2\lambda^2) V^2)}/(8\lambda) \\
    &= k e^{-(\lambda + 2) \text{tr} \left(U + (1/\lambda^2 + 2\lambda^2) V + (1/\lambda^2 + 2\lambda^2) V^2 \right)}/(\lambda^2 + 2\lambda^2).
\end{align}

This marginal density of $V$ is normal, and the conditional density of $U$ is normal. The distribution of $W$ and $(2\lambda^2 + 4\lambda)^{-1/2} \Theta$ is that of theorem 2. The conditional
distribution of $Z$ given $W$ and $\Theta$ is normal with mean

$$(5.26) \quad E[Z \mid W, \Theta] = \frac{1}{2} W E[U \mid W, \Theta] = \frac{1}{2} W \left(-\frac{1}{\lambda + 2} V\right) = -\frac{1}{2 (\lambda + 2)} \Theta W.$$ 

The conditional covariance between two columns of $Z$ is

$$(5.27) \quad \mathcal{E} \left[ z_i - \mathcal{E} (z_i \mid W, \Theta) \right] \left[ z_j - \mathcal{E} (z_j \mid W, \Theta) \right]' \mid W, \Theta = \frac{1}{2 (\lambda + 4)} \left(I - \frac{1}{2} W^2\right) \delta_{ij} + w_i w_j.$$ 

**Theorem 4.** Let

$$D_n = \sum_{p=1}^q \left(y_p + \sqrt{n} \sqrt{\lambda} \epsilon_p\right) \left(y_p + \sqrt{n} \sqrt{\lambda} \epsilon_p\right)'$$

where the $p$-dimensional vectors $y_1, \ldots, y_q$ ($q \geq p$) are independently distributed according to $N(0, I)$; let $A_n$ be independently distributed according to $W(I, n)$. Define $X_n$ and $\Phi_n$ by means of (4.3), (4.4), and the conditions that $x_{ij}(n) \geq 0$ and $\Phi_n$ is diagonal with diagonal elements in descending order. Let $\Theta_n = \sqrt{n}(\Phi_n - \lambda I)$ and $X_n = W_n + \frac{1}{\sqrt{n}} Z_n$, where $W_n W_n = I$, and $W_n Z_n = Z_n W_n$. The limiting distribution of $\Theta_n, W_n$ and $Z_n$ as $n \to \infty$ is the joint distribution of $\Theta, W$ and $Z$ such that the marginal distribution of $(2 \lambda^2 + 4 \lambda)^{-1/2}$ times the diagonal elements of the diagonal matrix $\Theta$ and the orthogonal matrix $W$ is that of theorem 2 and the conditional distribution of $Z$ given $W$ and $\Theta$ is normal with mean (5.26) and covariances (5.27).

6. An asymptotic distribution in the general case

Now we consider the general case of the population roots having different values. We assume that the multiplicities do not depend on $n$, but the values may. Let

$$(6.1) \quad [\tau^2(n) \delta_{ij}] = \begin{bmatrix} \lambda_1(n) I & 0 & \ldots & 0 \\ 0 & \lambda_2(n) I & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_h(n) I \\ 0 & 0 & \ldots & 0 \end{bmatrix} = \Lambda_n,$$

say, where $\lambda_i(n) I$ is of order $r_i$ and $r_{h+1} = p - \sum_{i=1}^h r_i$ is the multiplicity of $0$. Partition $X_n, F, E_n$ and $U_n$ similarly, to $[\tau^2(n) \delta_{ij}]$. Let

$$(6.2) \quad X_n = \begin{bmatrix} (W_n \tau(n) \ 0 & \ldots & 0 \\ 0 & W_n \tau(n) & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & W_n \tau(n) \end{bmatrix} + \frac{1}{\sqrt{n}} \begin{bmatrix} Z_{11}(n) & Z_{12}(n) & \ldots & Z_{1,h+1} \\ Z_{21}(n) & Z_{22}(n) & \ldots & Z_{2,h+1} \\ \vdots & \vdots & \ddots & \vdots \\ Z_{h+1,1} & Z_{h+1,2} & \ldots & Z_{h+1,h+1} \end{bmatrix} = W_n + \frac{1}{\sqrt{n}} Z_n.$$
where
\[
W_i(n)Z'_{ij}(n) = Z_{ij}(n)W'_i(n).
\]
As before, this defines \(W_i(n)\) and \(Z_{ij}(n)\) uniquely in terms of \(X_{ij}(n)\). To make \(X_n\)
unique in this case we now require that the elements of the first column of \(X_{ij}(n)\)
be nonnegative.

Let
\[
\Phi_n = \begin{pmatrix}
\lambda_1(n)I + \frac{1}{\sqrt{n}}\Theta_1(n) & 0 & \cdots & 0 & 0 \\
0 & \lambda_2(n)I + \frac{1}{\sqrt{n}}\Theta_2(n) & \cdots & 0 & 0 \\
0 & 0 & \cdots & \lambda_h(n)I + \frac{1}{\sqrt{n}}\Theta_h(n) & 0 \\
0 & 0 & \cdots & 0 & \frac{1}{\sqrt{n}}\Theta_{h+1}(n)
\end{pmatrix}.
\]
Then
\[
\frac{1}{n} D_n = \frac{1}{n} F + \frac{1}{\sqrt{n}} E_n + \Lambda = X_n'\Phi_nX_n,
\]
(6.5)
\[
\frac{1}{n} A_n = \frac{1}{\sqrt{n}} U_n + I = X_n'X_n.
\]
The submatric equations of (6.6) are
\[
\frac{1}{\sqrt{n}} U_{ij}(n) + I = \sum_i X'_{ji}(n)X_{ji}(n)
\]
\[
=W'_i(n)W_i(n) + \frac{1}{\sqrt{n}} [W'_i(n)Z_{ij}(n) + Z'_{ij}(n)W_i(n)]
\]
\[
+ \frac{1}{n} \sum_j Z'_{ji}(n)Z_{ji}(n).
\]
(6.7)
\[
\frac{1}{\sqrt{n}} U_{ij}(n) = \sum_k X_{ki}(n)X_{kj}(n)
\]
\[
= \frac{1}{\sqrt{n}} [W'_i(n)Z_{ij}(n) + Z'_{ij}(n)W_j(n)]
\]
\[
+ \frac{1}{n} \sum_k Z'_{ki}(n)Z_{kj}(n), \quad i \neq j.
\]
The submatric equations of (6.5) are
\[
\frac{1}{\sqrt{n}} F_{ii}(n) + \frac{1}{\sqrt{n}} E_{ii}(n) + \lambda_i(n)I = \sum_k X'_{ki}(n)\Phi_k(n)X_{ki}(n)
\]
\[
= \sum_{k=1}^h \left( \delta_{ik}W'_i(n) + \frac{1}{\sqrt{n}} Z'_{ki}(n) \right) \left( \lambda_k(n)I + \frac{1}{\sqrt{n}} \Theta_k(n) \right)
\]
\[
\times \left( \delta_{ik}W_i(n) + \frac{1}{\sqrt{n}} Z_{ki}(n) \right) + \frac{1}{n} Z'_{ki+1}(n)\Theta_{k+1}(n)Z_{ki+1}(n)
\]
\[
= \lambda_i(n)W'_i(n)W_i(n) + \frac{1}{\sqrt{n}} [W'_i(n)\Theta_iW_i(n) + \lambda_i(n)X'.n] \times
\]
\[ \frac{1}{n} F_{a+1,a+1} = \sum_k X'_{k,a+1}(n) \Phi_k(n) X_{k,a+1}(n) \]

\[ = \frac{1}{n} \left[ \sum_{k=1}^h Z_{k,a+1}(n) \left( \lambda_k(n) I + \frac{1}{\sqrt{n}} \Theta_k(n) \right) Z_{k,a+1}(n) \right] \]

\[ + \frac{1}{\sqrt{n}} Z_{a+1,a+1}(n) \left( W_{a+1}(n) + \frac{1}{\sqrt{n}} Z_{a+1,a+1}(n) \right) \]

\[ = \frac{1}{n} \left[ W_{a+1}(n) \Theta_{a+1}(n) W_{a+1}(n) + \sum_{k=1}^h \lambda_k(n) Z_{k,a+1}(n) Z_{k,a+1}(n) \right] \]

\[ + \frac{1}{n^{3/2}} \left[ \sum_{k=1}^h Z_{k,a+1}(n) \Theta_k(n) Z_{k,a+1}(n) + \frac{1}{x_k(1)} \right] \]

\[ + \sum_{k=1}^h Z_{k,a+1}(n) \Theta_k(n) Z_{k,a+1}(n) Z_{a+1,a+1}(n) \]

\[ + \frac{1}{n^{3/2}} \left[ \sum_{k=1}^h Z_{k,a+1}(n) \Theta_k(n) Z_{k,a+1}(n) + \frac{1}{x_k(1)} \right] \]

\[ + \sum_{k=1}^h Z_{k,a+1}(n) \Theta_k(n) Z_{k,a+1}(n) Z_{a+1,a+1}(n) \]

\[ (6.10) \]

\[ \frac{1}{n} F_{i,j} + \frac{1}{\sqrt{n}} E_{i,j}(n) = \sum_k X'_{k,i}(n) \Phi_k(n) X_{k,j}(n) \]

\[ = \sum_{k=1}^h \left( \delta_{ki} W_i(n) + \frac{1}{\sqrt{n}} Z_{ki}(n) \right) \left( \lambda_k(n) I + \frac{1}{\sqrt{n}} \Theta_k(n) \right) \]

\[ \times \left( \delta_{kj} W_j(n) + \frac{1}{\sqrt{n}} Z_{kj}(n) \right) + \frac{1}{n^{3/2}} Z_{i+1,i}(n) \Theta_{i+1}(n) Z_{i+1,j}(n) \]

\[ = \frac{1}{\sqrt{n}} \left[ \lambda_i(n) W_i(n) Z_{ij}(n) + \lambda_j(n) Z_{ij}(n) W_j(n) \right] \]

\[ + \frac{1}{n} \left[ \sum_{k=1}^h \lambda_k(n) Z_{ki}(n) Z_{kj}(n) + W_i(n) \Theta_i(n) Z_{ij}(n) + Z_{ij}(n) \Theta_j(n) \right] \]

\[ \times W_j(n) \right] + \frac{1}{n^{3/2}} \left[ \sum_{k=1}^h Z_{ki}(n) \Theta_k(n) Z_{kj}(n) \right] \]

\[ + \frac{1}{n^{3/2}} Z_{i+1,i}(n) \Theta_{i+1}(n) Z_{i+1,j}(n), \quad i \neq j ; i, j \neq h + 1, \]

\[ (6.11) \]

\[ \frac{1}{n} F_{i,a+1} + \frac{1}{\sqrt{n}} E_{i,a+1}(n) = \sum_{k=1}^{h+1} X'_{k,i}(n) \Phi_k(n) X_{k,a+1}(n) \]

\[ = \sum_{k=1}^{h+1} \left( \delta_{ki} W_i(n) + \frac{1}{\sqrt{n}} Z_{ki}(n) \right) \left( \lambda_k(n) I + \frac{1}{\sqrt{n}} \Theta_k(n) \right) \frac{1}{\sqrt{n}} Z_{k,a+1}(n) + \]

\[ (6.12) \]
\[ + \frac{1}{\sqrt{n}} z_{h+1,i}(n) \frac{1}{n} \Theta_{h+1}(n) \left( W_{h+1}(n) + \frac{1}{\sqrt{n}} z_{h+1,h+1}(n) \right) \]

\[ = \frac{1}{\sqrt{n}} \lambda_i(n) W'_i(n) Z_{i,h+1}(n) + \frac{1}{n} \left[ W'_i(n) \Theta_i(n) Z_{i,h+1}(n) + \sum_{k=1}^{h} \lambda_k(n) \right] \]

\[ \times Z_{h+1,i}(n) Z_{h,h+1}(n) + \frac{1}{n^{2/3}} \left[ \sum_{k=1}^{h} Z_{h+1,i}(n) \Theta_k(n) Z_{h+1,h+1}(n) + Z'_{h+1,i}(n) \right] \]

\[ \times \Theta_{h+1}(n) W_{h+1}(n) + \frac{1}{n^{2}} z_{h+1,i}(n) \Theta_{h+1}(n) Z_{h+1,h+1}(n), \quad i \neq h + 1. \]

For fixed \( F, E_n = E \) and \( U_n = U \) (in the proper domain) the above equations define the orthogonal \( W_i(n), Z_i(n) \) and the diagonal \( \Theta_i(n) \) uniquely (except for a set \( F, E \) and \( U \) of measure zero) under the restrictions that the elements in the first column of \( W_i(n) + \frac{1}{\sqrt{n}} Z_{ii}(n) \) are nonnegative and that the diagonal elements of the \( \Theta_i(n) \) are in descending order. Now subtract \( I \) from each side of \( (6.7) \) and multiply \( (6.7), (6.8), (6.9), (6.11) \) and \( (6.12) \) by \( \sqrt{n} \) and \( (6.10) \) by \( n \) and let \( n \to \infty \). Using the fact that \( W_i(n) \) is orthogonal and \( (6.3) \), we obtain the limiting equations [for \( \lambda_i(n) \to \lambda_i, E_n \to E, \) and \( U_n \to U \)]

\[(6.13) \quad U_{ii} = 2 W'_i Z_{ii}, \]

\[(6.14) \quad U_{ij} = W'_i Z_{ij} + Z'_j W_j, \quad i \neq j, \]

\[(6.15) \quad E_{ii} = W'_i \Theta_i W_i + 2 \lambda_i W'_i Z_{ii}, \quad i \neq h + 1, \]

\[(6.16) \quad F_{h+1,h+1} = W'_{h+1} \Theta_{h+1} W_{h+1} + \sum_{k=1}^{h} \lambda_k Z'_{k,h+1} Z_{k,h+1}, \]

\[(6.17) \quad E_{ij} = \lambda_i W'_i Z_{ij} + \lambda_j Z'_j W_j, \quad i \neq j; j, i \neq h + 1, \]

\[(6.18) \quad E_{i,h+1} = \lambda_i W'_i Z_{i,h+1}, \quad i \neq h + 1. \]

From (6.13) and (6.15) we obtain

\[(6.19) \quad E_{ii} - \lambda_i U_{ii} = W'_i \Theta_i W_i. \]

From (6.16) and (6.18) we obtain

\[(6.20) \quad F_{h+1,h+1} = \sum_{i=1}^{h} \frac{1}{\lambda_i} E'_{i,h+1} E_{i,h+1} = W'_{h+1} \Theta_{h+1} W_{h+1}. \]

Then the fact that \( W_i \) is orthogonal and the requirement that the elements of the first column of \( W_i \) be nonnegative define \( W_i, Z_{ii}, \) and \( \Theta_i \) uniquely.

To show that \( W_i, Z_{ij}, \) and \( \Theta_i \), defined by (6.13), (6.14), (6.17), (6.18), (6.19) and (6.20) are the limits of the matrices defined by (6.7)–(6.12) is more complicated than the similar demonstration in section 4. We shall only sketch this proof briefly. First we should like to prove that the diagonal elements of \( \Theta_i(n) \), for
example, converge to the characteristic roots of $E_{11} - \lambda_i U_{11}$ as $F_n \to F$, $E_n \to E$ and $U_n \to U$. From the equation

$$ (6.21) \quad \left| \frac{1}{n} F_1 + \frac{1}{\sqrt{n}} E_n + \Lambda_n - \phi \left( \frac{1}{\sqrt{n}} U_n + I \right) \right| = 0 $$

we can show that the first $r_1$ elements of $\Phi_n$ converge to $\lambda_1$. Then we need to show that the largest root of (6.21) minus $\lambda_1(n)$ times $\sqrt{n}$ converges to the largest characteristic root of $E_{11} - \lambda_i U_{11}$. That can be argued from the determinantal equation

$$ (6.22) \quad \left| \frac{1}{n} F_1 + \frac{1}{\sqrt{n}} \left[ E_n - \lambda_1(n) U_n \right] + \left( \Lambda_n - \lambda_1(n) I \right) - \frac{1}{\sqrt{n}} \theta \left( \frac{1}{\sqrt{n}} U_n + I \right) \right| = 0. $$

In the second determinant above we can factor out $\sqrt{n}$ from the first $r_1$ rows. Then as $n \to \infty$ there are $r_1$ sequences of roots each of which converges to a characteristic root of $E_{11} - \lambda_i U_{11}$. Similar arguments suffice for $\Theta_h(n)$ $(i \neq h + 1)$. For $\Theta_{h+1}(n)$ we can use a slightly more complicated demonstration.\(^3\)

Next we wish to argue that the elements of $Z_{ij}(n)$ are bounded as $n \to \infty$ (as $F_n \to F$, $E_n \to E$, and $U_n \to U$). For convenience let us take the case of $r_1 = 1$. Then the characteristic vector say $c_1(n)$ associated with the largest root $\theta_1$ satisfies

$$ (6.23) \quad \left[ \frac{1}{n} F_1 + \frac{1}{\sqrt{n}} \left[ E_n - \lambda_1(n) U_n \right] + \left( \Lambda_n - \lambda_1(n) I \right) - \frac{1}{\sqrt{n}} \theta_1 \left( \frac{1}{\sqrt{n}} U_n + I \right) \right] c = 0. $$

Then the components of $c_1(n)$ are

$$ (6.24) \quad c_{11}(n) = k(n) \left[ \prod_{i=2}^{p} [ \tau_i^2(n) - \lambda_1(n) ] + \frac{1}{\sqrt{n}} k_1(n) \right], $$

$$ (6.25) \quad c_{1i}(n) = k(n) \frac{1}{\sqrt{n}} k_i(n), \quad i \neq 1, $$

where $k(n)$ approaches a finite limit and $k_i(n)$ are bounded. Using the same reasoning for each characteristic vector (assuming $r_1 = 1$) we show that $C_n$ is a diagonal matrix with bounded elements plus $1/\sqrt{n}$ times a matrix with bounded elements. Thus $X_n = C_n^{-1}$ is of the same form. Therefore, $\sqrt{n} x_{ij}(n) = z_{ij}(n)$ $(i \neq j)$ are bounded. From

$$ (6.26) \quad \frac{1}{\sqrt{n}} u_{ii}(n) + 1 = \sum_{i=1}^{p} x_{ii}^2(n) $$

$$ = x_{ii}^2(n) + \sum_{i \neq i} x_{ii}(n) $$

we see that $\sqrt{n}[1 - x_{ii}^2(n)] = -u_{ii}(n) + \sqrt{n} x_{ii}^2(n)$ is bounded. Thus $z_{ii}(n) = \sqrt{n}[1 - x_{ii}(n)]$ is also bounded. If $r_1 \neq 1$, essentially the same argument can be carried out in terms of the partitioned matrices. Thus the norms of matrices such as $U_{ij}(n) - 2W_{ij}(n) Z_{ii}(n)$ go to zero. The argument of section 4 shows that $Z_{ii}(n)$ approaches $Z_{ii}$, etc.

\(^3\) Since these arguments are similar to Hsu's [8], it is unnecessary to go into more detail.
We are now in a position to apply Rubin's theorem. The discontinuities occur where $W$, $Z$, and $\Theta$ are not defined uniquely and the measure of $U$, $F$, and $E$ where this occurs is zero. Let the limiting distribution of the random matrices $U_n$, $F_n$, and $E_n$ be the distribution of the random matrices $U$, $F$, and $E$. Then the limiting distribution of the random matrices $W_n$ (orthogonal diagonal blocks), $Z_n$ and $\Theta_n$ (diagonal) defined by (6.7)–(6.12) is the distribution of $W$ (orthogonal diagonal blocks) $Z$ and $\Theta$ (diagonal) defined in terms of the random matrices $U$, $F$, and $E$ by (6.13), (6.14), (6.17)–(6.20). The distribution of $E_{ii} - \lambda_i U_{ii}$, $i \neq h + 1$, is that of section 5. Hence, the distribution of $W_i$ and $\Theta_i$ is that given there. Since $E_{ii} - \lambda_i U_{ii}$ is independent of $E_{ij} - \lambda_j U_{ij}$, $i \neq j$, the matrices $W_i$ and $\Theta_i$, $i = 1$, ..., $h$ are independent. The conditional distribution of $Z_{ii}$ given $W_i$ is also that of section 5 with $\lambda = \lambda_i$, $\Theta = \Theta_i$, $W = W_i$ and $p = r_i$.

Now, consider (6.20). An element of $E_{i,h+1}$ is $e_{gh} = \sqrt{\lambda_i} y_{kg} (r_1 + \ldots + r_{g-1} + 1 \leq g \leq r_1 + \ldots + r_h; \ p - r_{h+1} + 1 \leq k \leq p)$; an element of $F_{h+1,h+1}$ is

$$f_{kk'} = \sum_{g=1}^{q_1} y_{kg} y_{kg'}.$$  

Thus an element of $F_{h+1,h+1} = \sum_{i=1}^{h-1} E_{i,h+1} E_{i,h+1} \sum_{g=1}^{q_1} y_{kg} y_{kg'}$. This matrix of order $r_{h+1}$ has the distribution $W(1, q_2 - p + r_{h+1})$ and is independent of $E_{ii}$, $i \neq h + 1$, and $E_{ij}$, $i \neq j$. The distribution of $W_{h+1}$, $\Theta_{h+1}$ and $Z_{h+1,h+1}$ is that of section 4 with $W = W_{h+1}$, $\Theta = \Theta_{h+1}$, $Z = Z_{h+1,h+1}$, $p = r_{h+1}$ and $m = q_2 - p + r_{h+1}$.

The matrices $U_{ij}$ and $E_{ij}$, $i \neq j$, are independent of the other submatrices (except $E_{ii} = E_{ii}'$ and $U_{ii} = U_{ii}'$). From (6.14) and (6.17) or (6.18), we obtain

$$E_{ij} - \lambda_j U_{ij} = (\lambda_i - \lambda_j) W_{ij} Z_{ij}.$$  

The conditional distribution of $Z_{ij}$ given $W_i$ is that of

$$\frac{1}{\lambda_i - \lambda_j} W_i (E_{ij} - \lambda_j U_{ij}).$$

An element of $E_{ij}$ is

$$e_{ij} = \sqrt{\lambda_j} y_{ij} + \sqrt{\lambda_i} y_{if}, \quad r_1 + \ldots + r_{i-1} + 1 \leq g \leq r_1 + \ldots + r_i, \quad r_1 + \ldots + r_{j-1} + 1 \leq f \leq r_1 + \ldots + r_j.$$  

Thus an element of $E_{ij} - \lambda_j U_{ij}$ is normally distributed independently of the other elements with zero mean and variance $\lambda_i + \lambda_j + \lambda_j^2$. The conditional distribution of $Z_{ij}$ given $W_i$ is normal with mean 0 and covariance between two elements is

$$\mathbb{E} \left[ z_{ij} z_{ij}' | W_i \right] = \frac{1}{(\lambda_i - \lambda_j)^2} \mathbb{E} \sum_{k,k'} w_{kk'}^{(i)} (e_{kk'} - \lambda_j u_{kk'})(e_{kk'} - \lambda_j u_{kk'})$$

$$= \frac{1}{(\lambda_i - \lambda_j)^2} \sum_{k,k'} w_{kk'}^{(i)} w_{kk'}^{(i)} \delta_{kk'}(\lambda_i + \lambda_j + \lambda_j^2)$$

$$= \frac{\delta_{jj'}}{(\lambda_i - \lambda_j)^2} (\lambda_i + \lambda_j + \lambda_j^2).$$

Thus the variance of $z_{ij}$ is $(\lambda_i + \lambda_j + \lambda_j^2)/(\lambda_i - \lambda_j)^2$ and the covariances are zero.
The conditional distribution of $Z_{ji}$ is similar except that $i$ and $j$ are interchanged. $Z_{ij}$ and $Z_{ji}$ are independent of the other matrices. Now consider the conditional covariance between an element $z_{gf}$ of $Z_{ij}$ and $z_{g'f'}$ of $Z_{ji}$. It is

\[(6.31) \quad \mathbb{E} \left\{ z_{gf} z_{g'f'} \mid W_i, W_j \right\} = \frac{-1}{(\lambda_i - \lambda_j)^2} \mathbb{E} \sum_{k,k'} w_{jk}^{(i)} \left( e_{kf} - \lambda_i w_{kf} \right)
\]

\[\times \left( e_{f'k'} - \lambda_i w_{f'k'} \right)
\]

\[= -\frac{\lambda_i + \lambda_j + \lambda_i \lambda_j}{(\lambda_i - \lambda_j)^2} \sum_{k,k'} w_{jk}^{(i)} w_{g'f'} \delta_{kj} \delta_{f'k'}
\]

\[= -\frac{\lambda_i + \lambda_j + \lambda_i \lambda_j}{(\lambda_i - \lambda_j)^2} w_{gf} w_{g'f'}.
\]

We can indicate the conditional covariances between the columns $Z = (z_1, \ldots, z_p)$ in matrix form. If $g, f \leq r_1$, the conditional covariance between $z_g$ and $z_f$ is

\[(6.32) \quad \mathbb{E} \left\{ z_g z_f' \mid \Theta, W \right\} =
\]

\[
\begin{pmatrix}
\frac{1}{2\lambda_1 + 4} (I \delta_{gf} + w_{gf}^{(1)} w_{gf}^{(1)'}) & 0 & \cdots & 0 \\
0 & \frac{\lambda_1 + \lambda_2 + \lambda_1 \lambda_2}{(\lambda_1 - \lambda_2)^2} I \delta_{gf} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \frac{\lambda_1}{\lambda_2^2} I \delta_{gf}
\end{pmatrix}
\]

If $g \leq r_1$ and $r_1 + 1 \leq f \leq r_1 + r_2$, the conditional covariance is

\[(6.33) \quad \mathbb{E} \left\{ z_g z_f' \mid \Theta, W \right\} =
\]

\[
\begin{pmatrix}
0 & 0 & \cdots & 0 \\
\frac{\lambda_1 + \lambda_2 + \lambda_1 \lambda_2}{(\lambda_1 - \lambda_2)^2} w_{gf}^{(2)} w_{gf}^{(2)'}, & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 0
\end{pmatrix}
\]

**Theorem 5.** Let

\[(6.34) \quad D_n = \sum_{g=1}^{z_i} \left[ y_g + \sqrt{n} \tau_n (n) e_g \right] \left[ y_g + \sqrt{n} \tau_n (n) e_g \right]'.
\]

where the $p$-dimensional vectors $y_1, \ldots, y_{z_i}$ ($z_i \geq p$) are independently distributed according to $N(0, I)$ and $\tau_n(n) = \sqrt{\lambda_i(n)} \rightarrow \sqrt{\lambda_i}$, $r_1 + \ldots + r_i + 1 \leq g \leq r_1 + \ldots + r_i$, $\lambda_i > \lambda_j$, $i < j$ ($i, j = 1, \ldots, h + 1$); $\lambda_{h+1}(n) = 0$. Let $A_n$ be independently distributed according to $W(I, n)$. Define $X_n$ and $\Phi_n$ by means of (4.3), (4.4) and the conditions that $x_{g,r_1+1} + \ldots + x_{g,r_i+1}(n) \geq 0$, $r_1 + \ldots + r_i + 1 \leq g \leq r_1 + \ldots + r_i$, and $\Phi_n$ is diagonal with diagonal elements in descending order. Let each of $X_n$ and $\Phi_n$ be partitioned into $(h + 1)^2$ submatrices of $r_1, \ldots, r_{h+1}$ rows and
\[ r_1, \ldots, r_{h+1} \text{ columns. Let } \Theta_n \text{ be defined by } (6.4) \text{ and } W_n \text{ and } Z_n \text{ defined by } (6.2), \]
\[ (6.3) \text{ and } W_i(n)^t W_i(n) = I. \text{ The limiting distribution of } \Theta_n, W_n \text{ and } Z_n \text{ is that of } \Theta, W \text{ and } Z \text{ similarly partitioned, which can be described as follows: The matrices } \Theta_i, W_i, Z_i \text{ are independent of } \Theta_j, W_j, Z_j \text{ (} i \neq j \text{); the distribution of } \Theta_i, W_i \text{ and } Z_i, i \neq h + 1, \text{ is that of } \Theta, W \text{ and } Z \text{ given in theorem } 4 \text{ with } p = r_i; \text{ the distribution of } \Theta_{h+1}, W_{h+1} \text{ and } Z_{h+1,h+1} \text{ is that of } \Theta, W \text{ and } Z \text{ given in theorem } 3 \text{ with } p = r_{h+1}, \]
\[ m = q_2 - p + r_{h+1}; \text{ the conditional distribution of } Z \text{ given } \Theta \text{ and } W \text{ is normal; the conditional expectation of } Z_{ij} \text{ (} i \neq j \text{) is } 0; \text{ the conditional covariance matrix of } z_i, r_1 + \ldots + r_{i-1} + 1 \leq g, j \leq r_1 + \ldots + r_i; \text{ consists of nondiagonal blocks of zeros and the } i-\text{th diagonal block is} \]
\[ \frac{1}{2\lambda_i + 4} (I_{\delta_{ii}} + w_j^{(i)} w_j^{(i)^t}) \]
\[ \text{and the } j-\text{th diagonal block } (j \neq i) \text{ is} \]
\[ \frac{\lambda_i + \lambda_j + \lambda_j^2}{(\lambda_i - \lambda_j)^2} I_{\delta_{ii}}; \]
\[ \text{the covariance matrix of } z_i, r_1 + \ldots + r_{i-1} + 1 \leq g \leq r_1 + \ldots + r_i, \]
\[ r_1 + \ldots + r_{j-1} + 1 \leq j \leq r_1 + \ldots + r_j \text{ (} i \neq j \text{) consists of } 0 \text{'s except the } j-\text{th } i-\text{th block which is} \]
\[ \frac{\lambda_i + \lambda_j + \lambda_i \lambda_j}{(\lambda_i - \lambda_j)^2} w_j^{(i)} w_j^{(i)^t}. \]

7. The asymptotic distribution of characteristic roots and vectors

The columns of the matrix \( C \) defined in section 2 are the characteristic vectors of \( D_n \) in the metric of \( A_n \). We choose \( C \) so that the elements of the first row of \( C_i(n) \) are nonnegative. Let

\[ (7.1) \quad C_n = \begin{pmatrix}
C_{11}(n) & C_{12}(n) & \cdots & C_{1,h+1}(n) \\
C_{21}(n) & C_{22}(n) & \cdots & C_{2,h+1}(n) \\
\vdots & \vdots & \ddots & \vdots \\
C_{h+1,1}(n) & C_{h+1,2}(n) & \cdots & C_{h+1,h+1}(n)
\end{pmatrix}
\]
\[ = \begin{pmatrix}
V_1(n) & 0 & \cdots & 0 \\
0 & V_2(n) & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & V_{h+1}(n)
\end{pmatrix}
+ \frac{1}{\sqrt{n}} \begin{pmatrix}
Y_{11}(n) & Y_{12}(n) & \cdots & Y_{1,h+1}(n) \\
Y_{21}(n) & Y_{22}(n) & \cdots & Y_{2,h+1}(n) \\
\vdots & \vdots & \ddots & \vdots \\
Y_{h+1,1}(n) & Y_{h+1,2}(n) & \cdots & Y_{h+1,h+1}(n)
\end{pmatrix}
\]
\[ = V_n + \frac{1}{\sqrt{n}} Y_n, \]
\[ \text{where} \]
\[ (7.2) \quad V_i(n)V_i(n) = I, \]
\[ (7.3) \quad V_i(n)Y_i(n) = Y_i(n)V_i(n). \]
Let

$$\bar{C}_n = X_n^{-1}. $$

If $X_n = J_nC_n^{-1}$, then $\bar{C}_n = C_nJ_n$, where $J_n$ is a diagonal matrix with diagonal elements $+1$ and $-1$. Now define $\bar{V}_n$ and $\bar{Y}_n$ in terms of $\bar{C}_n$ as $V_n$ and $Y_n$ are in terms of $C_n$. First let us show that for nonrandom $W_n \to W$, $Z_n \to Z$ (which define $X_n \to X$), $\bar{V}_n \to W'$ and $\bar{Y}_n \to -Z'$. We have

$$\bar{C}_n = \bar{V}_n + \frac{1}{\sqrt{n}} \bar{Y}_n = \left( W_n + \frac{1}{\sqrt{n}} Z_n \right)^{-1} = \left( I + \frac{1}{\sqrt{n}} W' Z_n \right)^{-1} W'_n $$

$$= \left( I - \frac{1}{\sqrt{n}} W'_n Z_n + \frac{1}{n} (W'_n Z_n)^2 - \ldots \right) W'_n $$

$$= W'_n - \frac{1}{\sqrt{n}} W'_n Z_n W'_n + \frac{1}{n} (W'_n Z_n)^2 W'_n - \ldots $$

for $n$ sufficiently large. The $i$-th diagonal block of (7.5) is

$$\bar{V}_i(n) + \frac{1}{\sqrt{n}} Y_{ii}(n) = W'_i(n) - \frac{1}{\sqrt{n}} W'_i(n) Z_{ii}(n) W'_i(n) + \frac{1}{n} T_{ii}(n) $$

$$= W'_i(n) - \frac{1}{\sqrt{n}} Z_{ii}(n) + \frac{1}{n} T_{ii}(n), $$

where the elements of $T_{ii}(n)$ are bounded. Multiplying each side of (7.6) on the left by its transpose, we obtain

$$I + \frac{1}{\sqrt{n}} [\bar{V}_i(n) \bar{V}_{ii}(n) + \bar{Y}_{ii}(n) \bar{V}_i(n)] + \frac{1}{n} \bar{V}_{ii}(n) \bar{Y}_{ii}(n) $$

$$= I - \frac{1}{\sqrt{n}} [W'_i(n) Z_{ii}(n) + Z_{ii}(n) W'_i(n)] + \frac{1}{n} S_{ii}(n).$$

Subtracting $I$ from both sides, multiplying by $\bar{V}_i(n)$ and using (7.3) and (6.3), we find

$$Y_{ii}(n) = -\bar{V}_i(n) W'_i(n) Z_{ii}(n) + \frac{1}{\sqrt{n}} \bar{V}_i(n) R_i(n).$$

We can show (by means of an argument similar to that used in section 4) that the elements of $V_i(n)R_i(n) = S_{ii}(n) - V_i(n)Y_{ii}(n)Y_{ii}(n)$ and of $R_i(n)$ are bounded. Inserting (7.8) in (7.6) we obtain

$$\bar{V}_i(n) = \left( W'_i(n) - \frac{1}{\sqrt{n}} Z_{ii}(n) + \frac{1}{n} T_{ii}(n) \right) \left( I + \frac{1}{\sqrt{n}} W'_i(n) Z_{ii}(n) \right)^{-1} = W'_i(n) + \frac{1}{n} Q_i(n) $$

for $n$ sufficiently large. It is clear from this that $\bar{V}_i(n) \to W'_i$ and from (7.8) that $\bar{Y}_{ii}(n) \to -Z'_{ii}$ (as $W_n \to W$ and $Z_n \to Z$). A nondiagonal block of (7.5) is

$$\frac{1}{\sqrt{n}} \bar{Y}_{ij}(n) = -\frac{1}{\sqrt{n}} W'_i(n) Z_{ij}(n) W'_j(n) + \frac{1}{n} T_{ij}(n). $$

Clearly, $\bar{Y}_{ij}(n) \to W'_i Z_{ij} W'_j$. 

Theorem 6. Let

\[ D_n = \sum_{0}^{q_2} \left[ y_2 + \sqrt{n} \tau_{\theta}(n) e_2 \right] \left[ y_2 + \sqrt{n} \tau_{\theta}(n) e_2 \right]' \]

where the \( p \)-dimensional vectors \( y_1, \ldots, y_{q_2} (q_2 \geq p) \) are independently distributed according to \( N(0, I) \) and \( \tau_{\theta}(n) = \sqrt{n} \lambda_i(n) \to \sqrt{\lambda_i}, \ r_1 + \cdots + r_{i-1} + 1 \leq g \leq r_1 + \cdots + r_i; \ \lambda_i > \lambda_j, \ i < j \) \( (i, j = 1, \ldots, h + 1); \lambda_{h+1}(n) = 0. \) Let \( A_n \) be independently distributed according to \( W(I, n) \). Define \( C_n \) and \( \Phi_n \) by means of (2.11), (2.12) and the conditions that \( \sigma_{r_1 + \cdots + r_{i-1} + 1}(n) \geq 0, \ r_1 + \cdots + r_{i-1} + 1 \leq g \leq r_1 + \cdots + r_i \) \( (i = 1, \ldots, h + 1) \) and \( \Phi_n \) is diagonal with diagonal elements in descending order. Let each of \( C_n \) and \( \Phi_n \) be partitioned into \( (h + 1)^2 \) submatrices of \( r_1, \ldots, r_{h+1} \) rows and \( r_1, \ldots, r_{h+1} \) columns. Let \( \Theta_n \) be defined by (6.4) and \( V_n \) and \( Y_n \) be defined by (7.1), (7.2) and (7.3). The limiting distribution of \( \Theta_n, V_n \) and \( Y_n \) is that of \( \Theta, V \) and \( Y \) similarly partitioned, which can be described as follows: The matrices \( \Theta_i, V_i, Y_i \) are independent of \( \Theta_j, V_j, Y_j \) \( (i \neq j) \); the distribution of \( \Theta_i \) and \( V_i \) is that of \( \Theta \) and \( V \) given in theorem 4 with \( p = r_i, i \neq h + 1 \), the conditional distribution of \( Y_i \) given \( \Theta_i \) and \( V_i \) is normal with mean \( (2\lambda_i + 4)^{-1} V_i \Theta_i \); the distribution of \( \Theta_{h+1}, V_{h+1} \) is that of \( \Theta \) and \( V \) given in theorem 3 with \( p = r_{h+1}, m = q_2 - p + r_{h+1} \); the conditional distribution of \( Y_{h+1}, h+1 \) is normal with mean 0; the
conditional distribution of $Y$ given $\Theta$ and $V$ is normal; the conditional expectation of $Y_{ij}$, $i \neq j$, is 0; the conditional covariance matrix of $y_i$ and $y_j$, $r_1 + \ldots + r_i + 1 \leq g, f \leq r_1 + \ldots + r_i$, consists of nondiagonal blocks of zeros and the $i$-th diagonal block is

$$\frac{1}{2\lambda_i + 4} (I_{\delta_{gf}} + v^{(i')}_f v^{(i')}_g)$$

and the $j$-th diagonal block ($j \neq i$) is

$$\frac{\lambda_i + \lambda_j + \lambda_j^2}{(\lambda_i - \lambda_j)^2} I_{\delta_{gf}};$$

the covariance matrix of $y_i$ and $y_j$, $r_1 + \ldots + r_i + 1 \leq g \leq r_1 + \ldots + r_i, r_1 + \ldots + r_j - 1 \leq f \leq r_1 + \ldots + r_j$ ($i \neq j$) consists of 0's except the $j$-th, $i$-th block which is

$$\frac{\lambda_i + \lambda_j + \lambda_i \lambda_j}{(\lambda_i - \lambda_j)^2} v^{(j')}_f v^{(i')}_g.$$ 

Now let us consider the asymptotic distribution of $C_n = \Gamma_n C_n$. Let $\Gamma$ satisfy

$$(7.12) \quad B_n \lim_{n \to \infty} \bar{Q}_n B_n^T \Gamma = \Sigma \Gamma \Lambda.$$ 

If the diagonal elements of $\Lambda$ are all different and if $\gamma'_{ij} \neq 0$, $j = 1, \ldots, p$, then the restrictions $\gamma'_{ij} > 0$, and $\gamma'_{j} \Sigma_{j} \gamma_j = 1$ determine $\Gamma$ uniquely. If the same restrictions are placed on each $\Gamma_n$ then $\Gamma_n \to \Gamma$ as $n \to \infty$ because the set of characteristic vectors is a continuous function of $\bar{Q}_n$ (which approaches the limit, $\lim_{n \to \infty} \bar{Q}_n$). If the diagonal elements of $\Lambda$ are not all different, then another indeterminacy is involved. Partition $\Gamma$ in the manner that the matrices in section 6 were partitioned;

$$(7.13) \quad \Gamma = \begin{pmatrix} \Gamma_{11} & \cdots & \Gamma_{1,k+1} \\ \vdots & \ddots & \vdots \\ \Gamma_{h+1,1} & \cdots & \Gamma_{h+1,k+1} \end{pmatrix}.$$ 

Let $O$ be an orthogonal matrix of the form

$$(7.14) \quad O = \begin{pmatrix} O_1 & 0 & \cdots & 0 \\ 0 & O_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & O_h \end{pmatrix}.$$ 

We require that

$$(7.15) \quad \Gamma' \Sigma \Gamma = I.$$ 

Then $\Gamma O$ also satisfies (7.15); that is, there is an indeterminacy of such an orthogonal transformation. This indeterminacy can be removed by putting restrictions on $\Gamma$ (such as requiring that the first column of $\Gamma$ lie in a certain $r_1 - 1$ dimensional hyperplane, the second in a certain $r_1 - 2$ dimensional hyperplane, etc.,
with an element in each column having a specified sign). For all \( n \) greater than some particular integer, the same restrictions can be imposed on \( \Gamma_n \). With these restrictions imposed, \( \Gamma_n \to \Gamma \).

Then

\[ C_n^* = \Gamma_n C_n = \Gamma_n V_n + \frac{1}{\sqrt{n}} \Gamma_n \bar{Y}_n. \]

Let

\[ V_n^* = \Gamma_n V_n, \]
\[ Y_n^* = \Gamma_n Y_n; \]
then

\[ C_n^* = V_n^* + \frac{1}{\sqrt{n}} Y_n^*. \]

\( V_n^* \) and \( Y_n^* \) are functions of \( V_n \) and \( Y_n \). As \( n \to \infty \), the functions approach limits; that is, for fixed \( V_n = V \) and \( Y_n = Y \), \( V_n^* \to V^* \) and \( Y_n^* \to Y^* \), say. Thus by Rubin's theorem the limiting distribution of the random matrices \( V_n^* \) and \( Y_n^* \) is that of \( \Gamma V = V^* \) and \( \Gamma Y = Y^* \).

The distribution of \( V^* \) is that of \( \Gamma V \). The distribution of \( \Theta \) is given in section 6. The conditional distribution of \( Y^* \) given \( V^* \) and \( \Theta \) can be found from that of \( Y \).

We have

\[ \mathcal{E} \{ Y_{ij}^* | \Theta, V \} = \sum_{k=1}^{h+1} \Gamma_{ik} \mathcal{E} \{ Y_{kj} | \Theta, V \} = \Gamma_{ij} \frac{1}{2 (\lambda_j + 2)} V_j \Theta_j, \]
\[ j \neq h + 1; \]
\[ \mathcal{E} \{ Y_{i,h+1}^* | \Theta, V \} = 0. \]

The conditional covariances are easily obtained from theorem 6. Let \( y_a^* \) be the \( a \)-th column of \( Y^* \). Then the conditional covariance matrix of this vector for \( a \leq n \) is

\[ \begin{bmatrix} 1/2 \lambda_1 + 4 (I + v_a^{(1)} v_a^{(1)'} & 0 & \cdots & 0 \\ 0 & \lambda_1 + \lambda_2 + \lambda_2^2 I & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_1 I \end{bmatrix} \]

\[ = \frac{1}{2 \lambda_1 + 4} \Gamma_1 (I + v_a^{(1)} v_a^{(1)'}) \Gamma_1 + \sum_{j=2}^{h+1} \frac{\lambda_1 + \lambda_j + \lambda_j^2}{(\lambda_1 - \lambda_j)^2} \Gamma_j \Gamma_j', \]

where

\[ \Gamma_i = \begin{bmatrix} \Gamma_{i,i} \\ \vdots \\ \Gamma_{i,h+1,i} \end{bmatrix}. \]
If \( a \neq a^* \) and \( a, a^* \leq r_1 \), the conditional covariances between \( y_a \) and \( y_{a^*} \) are

\[
\frac{1}{2\lambda_1 + 4} \Gamma_2 v_a(1) v_a(1)' \Gamma_2'.
\]

If \( a \leq r_1 \) and \( r_1 + 1 \leq a^* \leq r_1 + r_2 \), then the covariances are

\[
\Gamma =
\begin{pmatrix}
0 & 0 & \ldots & 0 \\
0 & 0 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & 0
\end{pmatrix}
\]

\[
= \frac{-\lambda_1 + \lambda_2 + \lambda_1 \lambda_2}{(\lambda_1 - \lambda_3)^2} \Gamma_2 v_a(1) v_a(1)' \Gamma_2'.
\]

**Theorem 7.** Let the \( p \)-dimensional vectors \( x_a(a = 1, \ldots, N \geq p) \) be independently distributed according to \( N(B_1z_{1a} + B_2z_{2a}, \Sigma) \), where the matrix of vectors \( (z_{1a}, z_{2a}) \) is of rank \( p \) and \( z_{1a} \) has \( q_1 \) components and \( z_{2a} \) has \( q_2 \) components. Let \( B_n = [B_1(n), B_2(n)] \) be defined by (1.3); let \( \bar{Q}_n = \frac{1}{n} Q_n \) be defined by (1.5) where \( n = N - q_1 - q_2 \); let \( A_n^* \) be defined by (1.6). Let \( \Phi_n \) be a diagonal matrix whose diagonal elements are the roots of (1.7) arranged in descending order of size. Let \( C_n^* = [c_1^*(n), \ldots, c_p^*(n)] \) be composed of the corresponding vectors satisfying (1.8) and (1.9). Let \( \tau_i^2(n) \geq \tau_j^2(n) \) be the roots of (2.1). We assume \( \bar{Q}_n \) approaches a non-singular limit in such a way that \( \tau_i(n) = \sqrt{\lambda_i(n)} \rightarrow \sqrt{\lambda_i}, \; r_1 + \ldots + r_{i-1} + 1 \leq g \leq r_1 + \ldots + r_i, \lambda_i \leq \lambda_j, \; i < j \). Let \( \Gamma \) be a matrix satisfying (7.12) and (7.15) where \( \Lambda \) is the limit of (6.1), and satisfying other restrictions to make \( \Gamma \) uniquely defined. Let \( \Gamma_n = [\gamma_1(n), \ldots, \gamma_p(n)] \) be composed of vectors satisfying (2.2) and (2.3) and the additional restrictions on \( \Gamma \) (for \( n \) sufficiently large). Let \( \Theta_n \) be defined by (6.4). Let \( C_n^* = V_n^* + \frac{1}{\sqrt{n}} \eta_n^* \), where

\[
\Gamma_n^{-1}V_n^* = V_n = \begin{pmatrix} V_1(n) & 0 & \ldots & 0 \\
0 & V_2(n) & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & V_{h+1}(n) \end{pmatrix},
\]

\( V_i(n) = I \), the elements of the first row of \( V_i(n) \) are nonnegative and \( V_i(n)Y_i(n) = Y_i(n)V_i(n) \), where \( Y_i(n) \) is the \( i \)-th diagonal submatrix of \( \Gamma_n^{-1}Y_n = Y_n \). As \( n \rightarrow \infty \), the limiting distribution of \( \Theta_n, V_n^* \) and \( Y_n^* \) is that of \( \Theta, \Gamma V \) and \( Y^* \), similarly partitioned, which may be described as follows: The marginal distribution of \( \Theta \) and \( V \) is such that \( \Theta_i, V_i \) is independent of \( \Theta_j, V_j (i \neq j) \); the marginal distribution of \( \Theta_i, V_i \) is that of \( \Theta, W \) given in theorem 4 with \( p = r_i, i \neq h + 1 \); the distribution of \( \Theta_{h+1}, V_{h+1} \) is that of \( \Theta, W \) given in theorem 3 with \( p = r_{h+1}, m = q_2 - p + r_{h+1} \); the conditional distribution of \( Y^* \) given \( \Theta, V \) is normal; the conditional expectation of
a submatrix $Y_{ij}^*$ is given by (7.19); the conditional covariance matrix of $y^*_a$ and $y^*_{a'}$, $r_1 + \ldots + r_{i-1} + 1 \leq g, f \leq r_1 + \ldots + r_i$, is

$$(7.26) \quad \frac{1}{2\lambda_i + 4} \left[ \Gamma_i v_{ij}^{(i)} v_{ij}^{(i)'} + \delta_{ij} \Gamma_i \right] + \delta_{rf} \sum_{j=1}^{k+1} \frac{\lambda_i + \lambda_j + \lambda_j^2}{(\lambda_i - \lambda_j)^2} \Gamma_i \Gamma_j';$$

the conditional covariance between $y^*_a$ and $y^*_{a'}$, $r_1 + \ldots + r_{i-1} + 1 \leq g \leq r_1 + \ldots + r_i, r_1 + \ldots + r_{j-1} + 1 \leq f \leq r_1 + \ldots + r_j (i \neq j)$, is

$$(7.27) \quad -\frac{\lambda_i + \lambda_j + \lambda_i \lambda_j}{(\lambda_i - \lambda_j)^2} \Gamma_i v_{ij}^{(i)} v_{ij}^{(j)'} \Gamma_j;$$

where $\Gamma_i$ is defined in (7.22) and $v_{ij}^{(i)}$ is the $g$-th column of $V_i$.

A special case of considerable interest is the case of $r_i = 1, i \neq h + 1$ and $r_{h+1} = p - h$. Then $Y_{ij}$ consists of one element for $i, j \neq h + 1$. In this case $V_i = 1$ for $i \neq h + 1$. We can easily express the conditional distribution of $Y$ given $V_{h+1}$ by integrating out $\theta_1, \ldots, \theta_h$. The marginal distribution of $\theta_i$ is $N(0, 2\lambda_i^2 + 4\lambda_i)$ and the conditional distribution of $y_{ih}$ is $N(\theta_i/(2\lambda_i + 4), 1/(2\lambda_i + 2))$. $\theta_i$ and $y_{ih}$ are independent of the other variables. The marginal distribution then of $y_{ih}$ is $N(0, 1/2)$.

In this case $\mathcal{E}\{Y^*|V_{h+1}\} = 0$. The conditional covariance between $y^*_a$ and $y^*_{a'}$, $a \leq h$ is

$$(7.28) \quad \mathcal{E}\{y^*_a y^*_{a'}\} = \frac{1}{2} \Gamma_a \Gamma_{a'} + \sum_{j=1}^{k+1} \frac{\lambda_a + \lambda_j + \lambda_j^2}{(\lambda_a - \lambda_j)^2} \Gamma_j \Gamma_j'.$$

The conditional covariance between $y^*_a$ and $y^*_{a'}, a \neq a', a, a' \leq h$ is

$$(7.29) \quad \mathcal{E}\{y^*_a y^*_{a'}|V_{h+1}\} = -\frac{\lambda_a + \lambda_{a'} + \lambda_a \lambda_{a'}}{(\lambda_a - \lambda_{a'})^2} \Gamma_a \Gamma_{a'}.$$

The conditional covariance between $y^*_a$ and $y^*_{a'}, a \leq h, h + 1 \leq a' \leq p$, is

$$(7.30) \quad \mathcal{E}\{y^*_a y^*_{a'}|V_{h+1}\} = \frac{1}{\lambda_a} \Gamma_{h+1} v_{aa'}^{(h+1)} \Gamma_a.$$

The conditional covariance between $y^*_a$ and $y^*_{a'}, h + 1 \leq a, a' \leq p$, is

$$(7.31) \quad \mathcal{E}\{y^*_a y^*_{a'}|V_{h+1}\} = \frac{1}{2} \Gamma_{h+1} (I \delta_{aa'} + v_{aa'}^{(h+1)} \Gamma_{h+1}) + \sum_{j=1}^{h} \frac{1}{\lambda_j} \Gamma_j \Gamma_j' \delta_{aa'}.$$

Clearly, if $r_i = 1, i = 1, \ldots, h + 1 = p$, then $V_n = I$, and the limiting distribution of $\Phi_n$ and $Y_n^*$ is such that the marginal distribution of $Y^*$ is normal with mean 0 and covariances derived from above.

8. Remarks

8.1. Use of $N$ and $n = N - q$. In section 2 we defined $\overline{Q}_n = \frac{1}{N} Q_n$. The asymptotic distributions obtained are exactly the same if one uses $\frac{1}{N} Q_n$ for the roots of $|B_n^2 - \rho^2 \Sigma|$ are multiples by $n/N$ of the roots of (2.1). These roots converge also to $\lambda_1, \ldots, \lambda_{h+1}$, respectively, and for each $n$ the multiplicities are the same in the two cases. Using $\frac{1}{N} Q_n$ changes the definition of the sample roots again.
by a factor of \( n/N \). We might also define \( A^* \) in terms of \( N \) instead of \( n \) and normalized \( \phi^* \) in terms of \( N \) instead of \( n \). Asymptotically the effect of using \( N \) instead of \( n \) disappears. Rubin's theorem can be used to prove each such statement rigorously.

8.2. The limiting probabilities. It is interesting to see for what sequences of sets in the space of the characteristic vectors the limiting probabilities are defined. As a simple example, suppose \( p = r_1 = 2 \) and \( \Sigma = I \). We shall consider a sequence of sets for one vector \( c_1 \) and another sequence for \( c_2 \) defined in the same plane. Consider a segment on the unit circle in the right half plane. The regions for \( c_1 \) contain this segment and as \( n \to \infty \) the regions converge to the segment. The boundaries converge as \( 1/\sqrt{n} \). Consider the segment of the unit circle in the right half plane composed of the points which are 90° from the points in the other segment. There is a corresponding sequence of regions which close down on this segment as \( n \) increases. The limiting probabilities are defined for these sequences.

8.3. Cases of special interest. Two cases of the model discussed here are of particular interest. The one occurs when the "fixed variate" vectors \( z_n \) (in section 1) are composed of dummy variates; that is, variates that are 0 or 1 (see [1], for example). These can be chosen so that \( Bz_n = \mu, a = N_1 + \ldots + N_{i-1} + 1, \ldots, N_1 + \ldots + N_i \) \((N_1 + \ldots + N_r = N)\). The first \( N_1 \) \( x_n \) are observations from the first population, etc. If we require that \( N_i = k_i N \) as \( N \to \infty \), then the multiplicities of the roots of (2.1) are unchanged as \( N \to \infty \).

It can be shown that if \( \sqrt{n}[\tau^2(n) - \tau^0(n)] \to 0 \) as \( n \to \infty \), \( r_1 + \ldots + r_i + 1 \leq g, f \leq r_1 + \ldots + r_i, i = 1, \ldots, h + 1 \), then \( \Theta_s, W_s, \) and \( Z_s \) (also \( \Theta_s, V_s^*, \) and \( Y_s^* \)) can be defined in the manner of this paper and have the same limiting distribution as given here. Thus we can weaken our conditions slightly. If \( N_i = k_i N \) to within rounding error in the case mentioned above, the same theory applies.

If the fixed variate subvectors \( z_{2n} \) are not composed of dummy variates, in general the nonzero roots of (2.1) will be simple for all \( n \) and in the limit. Then the theory at the end of section 7 applies.

REFERENCES