The Annals of Statistics
1981, Vol. 9, No. 4, 725–736

ASYMPTOTIC INFERENCE FOR EIGENVECTORS

BY DAVID E. TYLER

Old Dominion University

Asymptotic procedures are given for testing certain hypotheses concerning eigenvectors and for constructing confidence regions for eigenvectors. These asymptotic procedures are derived under fairly general conditions on the estimates of the matrix whose eigenvectors are of interest. Applications of the general results to principal components analysis and canonical variate analysis are given.

1. Introduction and summary. Let M be a $p \times p$ matrix which is symmetric in the metric of the positive definite symmetric matrix Γ, i.e., ΓM is symmetric. Let the eigenvalues of M be represented by $\lambda_1 \geq \lambda_2 \cdots \geq \lambda_p$. Also, let M_n be a sequence of estimates of M such that $a_n(M_n - M)$ converges in distribution to a multivariate normal distribution, where a_n is an increasing sequence of real numbers, and let A be a $p \times r$ matrix with $\operatorname{rank}(A) = r$.

In this paper, under the assumption that $\lambda_{i-1} \neq \lambda_i$ and $\lambda_{i+m-1} \neq \lambda_{i+m}$, the following two null hypotheses are considered. For $r \leq m$, we consider

(1.1) H_0: the columns of A lie in the subspace generated by the set of eigenvectors of M associated with the roots $\lambda_i, \lambda_{i+1}, \ldots, \lambda_{i+m-1}$.

For $r \geq m$, we consider

(1.2) H_0^*: the eigenvectors of M associated with the roots $\lambda_i, \lambda_{i+1}, \ldots, \lambda_{i+m-1}$ lie in the subspace generated by the columns of A.

The assumption on the eigenvalues is to be interpreted as $\lambda_{i+m-1} \neq \lambda_{i+m}$ when $i = 1$, and $\lambda_{i-1} \neq \lambda_i$ when $i + m - 1 = p$.

Under fairly general condition on M_n, a consistent asymptotic chi-square test of H_0 is given. This test is based upon the asymptotic normality of the "orthogonal" projection of the columns of A onto the subspace generated by the eigenvectors of M_n associated with the ith to $(i + m - 1)$th roots of M_n. For H_0^*, an asymptotic chi-square test is constructed by relating this hypothesis to a hypothesis of the form given in (1.1).

An asymptotic confidence region for the subspace generated by the eigenvectors of M associated with the roots $\lambda_i, \lambda_{i+1}, \ldots, \lambda_{i+m-1}$ is then given. This confidence region is based upon the asymptotic chi-square test of H_0 for the special case when $r = m$.

Anderson (1963) gives an asymptotic chi-square test of H_0 for the special case $m = 1$ and when M_n is the sample covariance matrix from a multivariate normal sample with population covariance matrix M. This paper is thus a generalization of Anderson's results.

James (1977) gives exact tests for a hypothesis similar to (1.1) when M_n is the sample covariance matrix from a multivariate normal sample with population covariance matrix M. James considers the hypothesis that the columns of A generate an invariant subspace of M. His hypothesis does not state with which eigenvalues of M the invariant space is associated, whereas the hypothesis considered in this paper does. The approach used by

Received October, 1979; revised August, 1980.

Research supported in part by ONR Grants N0014-67-0151-0017 and N00014-75-C-0453, ARO Grant DAHC-04-74-G0178 and ERDA Grant E-111-2310 while the author was a graduate student in the Department of Statistics, Princeton University.

AMS 1970 *subject classifications*. Primary 62H15; secondary 62H20, 62H25, 62E20.

Key words and phrases. Eigenvectors, eigenprojections, generalized inverses and asymptotic chi-square statistics, principal components analysis, canonical variate analysis, elliptical distributions.

James uses special properties of the sample covariance matrix from a normal sample and does not readily generalize to other matrices.

For other related works on the distributional and inferential theory for eigenvectors, the reader is referred to Anderson (1951), Mallows (1961), Chambers (1967), Hirakawa (1976a, 1976b), Izenman (1976) and Sugiura (1976).

Applications of the general results in this paper are illustrated through the following two examples: the principal component vectors for the covariance matrix of an elliptical distribution, and the canonical vectors associated with two random vectors which jointly have an elliptical distribution.

2. Preliminaries. *Spectral Theory.* Let S be a $q \times q$ real matrix which is symmetric in the metric of a real positive definite symmetric matrix T. In order to establish notation and vocabulary, the spectral theory for S is briefly reviewed below. A more detailed review can be found in Kato (1966) or Nerring (1970).

If $S\mathbf{x} = \lambda\mathbf{x}$ for some $\mathbf{x} \neq 0$, then λ is an eigenvalue of S and $\mathbf{x}$ is an eigenvector of S associated with λ. All eigenvalues of S are real. The spectral set of S, denoted $\mathscr{S}$, is the set of all eigenvalues of S.

The eigenspace of S associated with λ is $V(\lambda) = \{\mathbf{x} \in R^q \mid S\mathbf{x} = \lambda\mathbf{x}\}$, where R^q is the set of all q-dimensional real vectors. The dimension of $V(\lambda)$ is the multiplicity of λ, say $m(\lambda)$. If λ and μ are two distinct eigenvalues of S, then $V(\lambda)$ and $V(\mu)$ are orthogonal subspaces in the metric of T. That is, if $\mathbf{x} \in V(\lambda)$ and $\mathbf{y} \in V(\mu)$, then $\mathbf{x}'T\mathbf{y} = 0$.

Since S is symmetric in the metric of T, we have the decomposition, $R^q = \sum_{\lambda \in \mathscr{S}} V(\lambda)$. The eigenprojection of S associated with λ, denoted $P(\lambda)$, is the projection operator onto $V(\lambda)$ with respect to this decomposition of R^q. The spectral decomposition of S is $S = \sum_{\lambda \in \mathscr{S}} \lambda P(\lambda)$. If v is any subset of the spectral set $\mathscr{S}$, then the total eigenprojection for S associated with the eigenvalues in v is defined to be $\sum_{\lambda \in v} P(\lambda)$. For any set of vectors $\{\mathbf{x}_j\}$ in $V(\lambda)$ such that $\mathbf{x}_j'T\mathbf{x}_k = \delta_{jk}$, where δ_{jk} denotes the Kronecker delta, $P(\lambda)$ has the representation $P(\lambda) = \sum_{j=1}^{m(\lambda)} \mathbf{x}_j\mathbf{x}_j'T$. Thus, $P(\lambda)$ is symmetric in the metric of T.

The eigenvalues and eigenprojections of "symmetric" matrices have the following important continuity property.

LEMMA 2.1. *Let S_k be a $q \times q$ matrix symmetric in the metric of T_k and with eigenvalues $\lambda_1(S_k) \geq \lambda_2(S_k) \geq \cdots \geq \lambda_q(S_k)$. Let $P_{j,t}(S_k)$ represent the total eigenprojection for S_k associated with $\lambda_j(S_k), \ldots, \lambda_t(S_k)$ for $t \geq j$. If $S_k \to S$ as $k \to \infty$, then*

(i) *$\lambda_j(S_k) \to \lambda_j(S)$, and*

(ii) *$P_{j,t}(S_k) \to P_{j,t}(S)$ provided $\lambda_{j-1}(S) \neq \lambda_j(S)$ and $\lambda_t(S) \neq \lambda_{t+1}(S)$.*

Generalized Inverses and Quadratic Forms. A generalized inverse of S is any S^- such that $SS^-S = S$. The Moore-Penrose generalized inverse of S, denoted by S^+, can be represented by $S^+ = \sum_{\lambda \in \mathscr{S}, \lambda \neq 0} \lambda^{-1} P(\lambda)$. In this section, some basic results concerning generalized inverses and quadratic forms involving generalized inverses are presented.

LEMMA 2.2. *Let S_n be a $q \times q$ random matrix symmetric in the metric of T_n, where T_n is random. If $S_n \to S$ in probability with* $\operatorname{rank}(S) = s$ *and* $\operatorname{Prob}[\operatorname{rank}(S_n) = s] \to 1$, *then $S_n^+ \to S^+$ in probability.*

PROOF. Let $P_s(S_n)$ be the eigenprojection of S_n associated with its zero eigenvalue. It then follows from Lemma 2.1.(ii) that $S_n^+ = \{[S_n + P_s(S_n)]^{-1} - P_s(S_n)\} \to \{[S + P_s(S)]^{-1} - P_s(S)\} = S^+$ in probability.

For a $q \times k$ matrix C, let $\mathscr{M}(C)$ represent the manifold of C, that is

$$\mathscr{M}(C) = \{\mathbf{v} \in R^q \mid \mathbf{v} = C\mathbf{w} \text{ for some } \mathbf{w} \in R^k\}. \tag{2.1}$$

LEMMA 2.3. *Let S be a positive semidefinite symmetric matrix of order $q \times q$, $\mathbf{x} \in \mathscr{M}(S)$, and B a $q \times k$ matrix with rank$(B) = k$, then*

(i) *$\mathbf{x}'B(B'SB)^- B'\mathbf{x}$ is invariant with respect to the choice of the generalized inverse for $B'SB$.*

(ii) *$\mathbf{x}'B(B'SB)^- B'\mathbf{x} \le \mathbf{x}'S^- \mathbf{x}$ with equality if $k = q$.*

(iii) *$\mathbf{x}'S^- \mathbf{x} \ge (\mathbf{x}'\mathbf{x})^2(\mathbf{x}'S\mathbf{x})^{-1}$.*

PROOF. For $B = I$, part (i) follows since $\mathbf{x} = S\mathbf{y}$ for some $\mathbf{y}$, and so $\mathbf{x}'S^- \mathbf{x} = \mathbf{y}'SS^- S\mathbf{y} = \mathbf{y}'S\mathbf{y}$ for any generalized inverse. For general B, part (i) follows by noting that $B'\mathbf{x} \in \mathscr{M}(B'SB)$.

For part (ii), note that $S^{1/2}B(B'SB)^- B'S^{1/2}$ is idempotent, where $S^{1/2} = \sum_{\lambda \in \mathscr{S}} \lambda^{1/2} P(\lambda)$, and thus $\mathbf{x}'B(B'SB)^- B'\mathbf{x} = \mathbf{y}'SB(B'SB)^- B'S\mathbf{y} \le \mathbf{y}'S\mathbf{y} = \mathbf{x}'S^- \mathbf{x}$.

To prove part (iii), the Cauchy-Schwarz inequality can be used to obtain $(\mathbf{x}'\mathbf{x})^2 = [\mathbf{x}'S^{1/2}(S^+)^{1/2}\mathbf{x}] \le \mathbf{x}'S\mathbf{x}\ \mathbf{x}'S^+ \mathbf{x} = \mathbf{x}'S\mathbf{x}\ \mathbf{x}'S^- \mathbf{x}$.

Let $\chi_k^2(\delta)$ represent a chi-square distribution on k degrees of freedom and with noncentrality parameter δ. That is, if $\mathbf{X} \sim \text{Normal}(\boldsymbol{\mu}, I_{k \times k})$ then $\mathbf{X}'\mathbf{X} \sim \chi_k^2(\boldsymbol{\mu}'\boldsymbol{\mu})$. The next lemma concerning asymptotic chi-square variables is given by Moore (1977).

LEMMA 2.4. *If $\mathbf{X}_n \to_d \text{Normal}(\boldsymbol{\mu}, \Sigma_X)$ with rank$(\Sigma_X) = v$, and if $B_n \to B$ in probability where B is any generalized inverse of Σ_X, then*

(i) *$\mathbf{X}_n' B_n \mathbf{X}_n \to_d \chi_v^2$ for $\boldsymbol{\mu} = \mathbf{0}$, and*

(ii) *$\mathbf{X}_n' B_n \mathbf{X}_n \to_d \chi_v^2(\boldsymbol{\mu}' \Sigma_X^+ \boldsymbol{\mu})$ provided $\boldsymbol{\mu} \in \mathscr{M}(\Sigma_X)$.*

Kronecker Products and the "vec" *Transformation.* For random matrices, it is convenient to introduce the following notation. If B is a $b \times t$ matrix, then vec(B) is the transformation of B into a bt-dimensional vector in the following fashion. Let $B = [\mathbf{b}_1\ \mathbf{b}_2 \cdots \mathbf{b}_t]$ where $\mathbf{b}_j$ is the jth column of B, then

$$\text{vec}(B) = \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \\ \vdots \\ \mathbf{b}_t \end{bmatrix}. \tag{2.2}$$

If B is a $b \times t$ matrix and C is a $c \times u$ matrix, then the Kronecker product of B and C is the $bc \times tu$ partitioned matrix $B \otimes C = [b_{jk}C]$, $j = 1, 2, \ldots, b$ and $k = 1, 2, \ldots, t$ with j varying over rows of matrices and k varying over columns of matrices.

An important property relating the "vec" transformation and the Kronecker product is

$$\text{vec}(BCD) = (D' \otimes B)\text{vec}(C), \tag{2.3}$$

where the dimensions of the matrices B, C, and D are such that the multiplications are properly defined. Other properties of the "vec" transformation and the Kronecker product can be found in Neudecker (1968).

The commutation matrix or permuted identity matrix is the $ab \times ab$ matrix $I_{(a,b)} = \sum_{i=1}^a \sum_{j=1}^b E_{ij} \otimes E_{ij}'$, where E_{ij} is an $a \times b$ matrix with a one in the (i, j) position and zeroes elsewhere. The commutation matrix has been extensively investigated recently by Magnus and Neudecker (1979). Two important properties of the commutation matrix are

$$I_{(a,b)}\ \text{vec}(B) = \text{vec}(B'), \tag{2.4}$$

and

$$I_{(a,b)}(C \otimes D) = (D \otimes C)I_{(c,d)}, \tag{2.5}$$

where B is $b \times a$, C is $b \times d$, and D is $a \times c$.

Finally, the following two lemmas are needed. The first lemma is a special case of the theorem given by Okamoto (1973).

LEMMA 2.5. *If* B *is a* $k_1 \times k_2$ *random matrix such that* $\text{vec}(B) \sim \text{Normal}(\mathbf{0}, S)$ *with* $\text{rank}(S) = k_1 k_2$, *then* $\text{rank}(B) = \min(k_1, k_2)$ *almost surely.*

LEMMA 2.6. *If* B_n *and* B *are* $k_1 \times k_2$ *random matrices such that* $\text{rank}(B) = b$ *almost surely,* $\text{Prob}[\text{rank}(B_n) \le b] \to 1$ *and* $B_n \to_d B$, *then* $\text{Prob}[\text{rank}(B_n) = b] \to 1$.

PROOF. By Lemma 2.1(i), it follows that $\lambda_b(B_n' B_n) \to_d \lambda_b(B'B)$ which is almost surely nonzero. The lemma follows since $\text{rank}(B_n) = \text{rank}(B_n' B_n)$.

3. Assumptions. In order to form an asymptotic test for (1.1), a sequence of estimators M_n for M are needed which satisfies the following assumptions. The implications of Assumption 3.1(iii) are to be discussed in Section 8.

ASSUMPTION 3.1.
(i) M_n is symmetric in the metric of Γ_n, a positive definite symmetric matrix, with $\Gamma_n \to \Gamma$ in probability.
(ii) $a_n(M_n - M) \to_d N$ where a_n is an increasing sequence of positive numbers such that $a_n \to \infty$ as $n \to \infty$, and $\text{vec}(N)$ is multivariate normal, mean zero, and covariance matrix Σ.
(iii) For B which is $p \times p$, $\Sigma\, \text{vec}(\Gamma B) = 0$ implies $M(B + B') = 0$.

It is also necessary to have a sequence of estimators Σ_n for Σ which satisfies the following properties.

ASSUMPTION 3.2
(i) Σ_n is symmetric and positive semidefinite.
(ii) $\Sigma_n \to \Sigma$ in probability.
(iii) Let $\Omega_n = \{\Sigma_n\, \text{vec}(\Gamma_n B) = 0 \text{ implies } M_n(B + B') = 0\}$ then $\text{Prob}(\Omega_n) \to 1$.

It is to be understood that the asymptotic procedures given in this paper are only defined on the intersection of Ω_n with

$$C_n = \{\hat{\lambda}_{i-1} \ne \hat{\lambda}_i \quad \text{and} \quad \hat{\lambda}_{i+m-1} \ne \hat{\lambda}_{i+m}\}, \tag{3.1}$$

where $\hat{\lambda}_1 \ge \hat{\lambda}_2 \ge \cdots \ge \hat{\lambda}_p$ are the eigenvalues of M_n. It is irrelevant to the asymptotic properties of the procedures what action is taken otherwise, since by the continuity of the eigenvalues of "symmetric" matrices, that is Lemma 2.1 (i), $\text{Prob}(C_n) \to 1$.

4. Asymptotic Distribution of the Eigenprojection. Let $w = \{\lambda_i, \lambda_{i+1}, \ldots, \lambda_{i+m-1}\}$ and let $\hat{w} = \{\hat{\lambda}_i, \hat{\lambda}_{i+1}, \ldots, \hat{\lambda}_{i+m-1}\}$. Also, for λ an eigenvalue of M, let P_λ represent the eigenprojection of M associated with λ. For λ an eigenvalue of M_n, let $\hat{P}_\lambda$ represent the eigenprojection of M_n associated with λ. For convenience, define $P_0 = \sum_{\lambda \in w} P_\lambda$ and $\hat{P}_0 = \sum_{\lambda \in \hat{w}} \hat{P}_\lambda$. P_0 represents the total eigenprojection of M associated with the eigenvalues of M in w, and $\hat{P}_0$ represents the total eigenprojection of M_n associated with the eigenvalues of M_n in $\hat{w}$.

The null hypothesis (1.1) can thus be rephrased as

$$H_0: P_0 A = A \tag{4.1}$$

where A is $p \times r$ with $\text{rank}(A) = r \le m$.

A natural statistic to consider in testing H_0 is the difference between A and its "orthogonal" projection onto the eigenspace of M_n associated with the eigenvalues in $\hat{w}$, that is $(A - \hat{P}_0 A)$. In obtaining the asymptotic distribution of this statistic, the Taylor series expansion of $\hat{P}_0$ about P_0 is to be used. This expansion is given in the following lemma. The lemma is a simplified version of more general results given in Chapter 2 of Kato (1966). A proof of this simplified version can be found in Appendix B of the author's dissertation.

LEMMA 4.1. *Let* $d_0 = \min\{\lambda_{i-1} - \lambda_i, \lambda_{i+m-1} - \lambda_{i+m}\}$, *and* $d_1 = (\lambda_i - \lambda_{i+m-1})$. *Also define the norm* $\|B\| = [\max$ *eigenvalue* $(\Gamma^{-1}B'\Gamma B)]^{1/2}$. *If* $\|M_n - M\| \le d_0/2$, *then*

$$\hat{P}_0 = P_0 - \textstyle\sum_{\lambda\in w}[P_\lambda(M_n - M)(M - \lambda I)^+ + (M - \lambda I)^+(M_n - M)P_\lambda] + E_n,$$

where $\|E_n\| \le (1 + d_1/d_0)(2\|M_n - M\|/d_0)^2(1 - 2\|M_n - M\|/d_0)^{-1}$.

THEOREM 4.1. *If* $P_0A = A$, *then* $\mathrm{vec}[a_n(I - \hat{P}_0)A] \to_d \mathrm{Normal}[\mathbf{0}, \Sigma_0(A)]$, *with* $\Sigma_0(A) = (A' \otimes I)C_w'\Sigma_n C_w(A \otimes I)$ *and where* $C_w = \sum_{\lambda\in w}\sum_{\mu\notin w}(\lambda - \mu)^{-1}P_\lambda \otimes P_\mu'$.

PROOF. Lemma 4.1 immediately yields the following limiting normal distribution

$$a_n(\hat{P}_0 - P_0) \to_d N_0 = -\textstyle\sum_{\lambda\in w}[P_\lambda N(M - \lambda I)^+ + (M - \lambda I)^+ NP_\lambda]. \tag{4.2}$$

Thus, under H_0, $a_n(A - \hat{P}_0A) = a_n(I - \hat{P}_0)A \to_d -N_0A$, which has a multivariate normal distribution with zero mean. The form of the covariance matrix follows by noting that

$$\begin{aligned}\mathrm{vec}[-N_0A] &= (A' \otimes I)\textstyle\sum_{\lambda\in w}[(M' - \lambda I)^+ \otimes P_\lambda + P_\lambda' \otimes (M - \lambda I)^+]\mathrm{vec}(N)\\ &= -(A' \otimes I)C_w'\mathrm{vec}(N).\end{aligned}$$

An estimate of $\Sigma_0(A)$ which is consistent under H_0 is obtained by defining

$$\hat{\Sigma}_0(A) = (A' \otimes I)\hat{C}_w'\Sigma_n\hat{C}_w(A \otimes I), \tag{4.3}$$

where

$$\hat{C}_w = \textstyle\sum_{\lambda\in\hat{w}}\sum_{\mu\notin\hat{w}}(\lambda - \mu)^{-1}\hat{P}_\lambda \otimes \hat{P}_\mu'.$$

THEOREM 4.2. $\hat{\Sigma}_0(A) \to \Sigma_0(P_0A)$ *in probability.*

PROOF. Since $\Sigma_n \to \Sigma$ in probability by Assumption 3.2(ii), it only needs to be shown that

$$\hat{C}_w \to C_w \quad \text{in probability.} \tag{4.4}$$

To show this, note that $\hat{C}_w$ and C_w are symmetric in the metric of $\hat{\Gamma}_n \otimes \hat{\Gamma}_n^{-1}$ and $\Gamma \otimes \Gamma^{-1}$ respectively, and that $\hat{C}_w = [M_n\hat{P}_0 \otimes (I - \hat{P}_0') - \hat{P}_0 \otimes M_n'(I - \hat{P}_0')]^+$ and $C_w = [MP_0 \otimes (I - P_0') - P_0 \otimes M'(I - P_0')]^+$. Statement (4.4) then follows by Lemma 2.2, since $\mathrm{rank}(\hat{C}_w) = m(p - m) = \mathrm{rank}(C_w)$.

5. An Asymptotic Chi-square Test. In this section, an asymptotic chi-square test for H_0 is given based upon the asymptotic normality of $a_n(I - \hat{P}_0)A$. Before introducing the test statistic, the following theorem and corollary are needed.

THEOREM 5.1. *If* $\mathrm{rank}[P_0A] = r$, *then*
(i) $\mathscr{M}[\Sigma_0(A)] = \mathscr{M}[I_{r\times r} \otimes (I - P_0)]$ *and*
(ii) $\mathrm{rank}[\Sigma_0(A)] = r(p - m)$.

PROOF. The proof consists of determining the null space of $\Sigma_0(A)$. For G which is $p \times r$, $[\mathrm{vec}(\Gamma G)]'\ \Sigma_0(A)\mathrm{vec}(\Gamma G) = [\mathrm{vec}(\Gamma B)]'\ \Sigma\ \mathrm{vec}(\Gamma B)$, where $B = \sum_{\lambda\in w}\sum_{\mu\notin w}(\lambda - \mu)^{-1}P_\mu GA'P_\lambda'$. So by Assumption 3.1(iii), $\Sigma_0(A)\mathrm{vec}(\Gamma G) = 0$ implies $M(B + B') = 0$, which implies $P_\mu G = 0$ for $\mu \notin w$.

The last implication is justified by the following contrapositive argument. Suppose $\mu \notin w$ and $P_\mu G \neq 0$.

Case I: $0 \notin w$. $M(B + B')P'_\mu = \{\sum_{\lambda \in w} \lambda(\lambda - \mu)^{-1} P_\lambda\} A G' P'_\mu \neq 0$, since $\operatorname{rank}[\{\sum_{\lambda \in w} \lambda \cdot (\lambda - \mu)^{-1} P_\lambda\} A] = r$.

Case II: $0 \in w$. $P_\mu M(B + B') = \mu P_\mu G A' \{\sum_{\lambda \in w} (\lambda - \mu)^{-1} P'_\lambda\} \neq 0$.

The converse, that is $P_\mu G = 0$ for all $\mu \notin w$ implies $\Sigma_0(A)\operatorname{vec}(\Gamma G) = 0$, is obviously true. Thus, the null space of $\Sigma_0(A)$ is $\eta = \{\operatorname{vec}(\Gamma G) \mid (I - P_0)G = 0\}$, and the theorem immediately follows.

COROLLARY 5.1. *If* $\operatorname{rank}(\hat{P}_0 A) = r$, *then*

(i) $\mathscr{M}[\hat{\Sigma}_0(A)] = \mathscr{M}[I_{r \times r} \otimes (I - \hat{P}_0)]$,

(ii) $\operatorname{rank}[\hat{\Sigma}_0(A)] = r(p - m)$, *and*

(iii) $\operatorname{vec}[(I - \hat{P}_0)A] \in \mathscr{M}[\hat{\Sigma}_0(A)]$.

The test statistic is now introduced. Define

$$T_n(A) = a_n^2 \{\operatorname{vec}[(I - \hat{P}_0)A]\}' [\hat{\Sigma}_0(A)]^- \operatorname{vec}[(I - \hat{P}_0)A]. \tag{5.1}$$

The value of $T_n(A)$ does not depend upon the choice of the generalized inverse for $\hat{\Sigma}_0(A)$, at least asymptotically. Another property of the statistic $T_n(A)$ is that asymptotically it is invariant under post-multiplication of A by a nonsingular matrix. This property is important since the hypothesis H_0 is invariant under postmultiplication of A by a nonsingular matrix. More specifically,

THEOREM 5.2.

(i) *On the set* $\{\operatorname{rank}(\hat{P}_0 A) = r\}$, $T_n(A)$ *is invariant under different choices of a generalized inverse for* $\hat{\Sigma}_0(A)$.

(ii) *On the set* $\{\operatorname{rank}(\hat{P}_0 A) = r\}$, $T_n(A) = T_n(AB)$ *for any nonsingular matrix* B.

(iii) *Whether or not* H_0 *is true*, $\operatorname{Prob}\{\operatorname{rank}(\hat{P}_0 A) = r\} \to 1$.

PROOF.

(i) By corollary 5.1(iii), this result is obtained by application of Lemma 2.3(i) to $T_n(A)$.

(ii) This result follows immediately from part (i) by noting that $\operatorname{vec}[(I - \hat{P}_0)AB] = (B' \otimes I)\operatorname{vec}[(I - \hat{P}_0)A]$, and that $(B' \otimes I)[\hat{\Sigma}_0(AB)]^-(B \otimes I)$ is a generalized inverse of $\hat{\Sigma}_0(A)$.

(iii) *Case* I: $\operatorname{rank}(P_0 A) = r$. Since $\hat{P}_0 A \to P_0 A$ in probability, and $\operatorname{rank}(\hat{P}_0 A) \leq r$, the theorem follows from Lemma 2.6.

Case II: $\operatorname{rank}(P_0 A) = r_0 < r$. Let C be a nonsingular matrix such that $AC = [A_1\ A_2]$ and where $P_0 A_1 = A_1$ and $P_0 A_2 = 0$. The order of A_1 and A_2 are $p \times r_0$ and $p \times (r - r_0)$ respectively. Also, let $A_* = [A_1\ A_3]$ where A_3 has order $p \times (m - r_0)$ and chosen such that $P_0 A_3 = A_3$ and $A'_1 \Gamma A_3 = 0$.

Since $r \geq \operatorname{rank}(\hat{P}_0 A) \geq \operatorname{rank}(A'_* \Gamma \hat{P}_0 A) = \operatorname{rank}[A'_* \Gamma \hat{P}_0 A_1\ a_n A'_* \Gamma \hat{P}_0 A_2]$, and $[A'_* \Gamma \hat{P}_0 A_1\ a_n A'_* \Gamma \hat{P}_0 A_2] \to_d [A'_* \Gamma A_1\ A'_* \Gamma N_0 A_2]$, where N_0 is defined in (4.2), the theorem then follows from Lemma 2.6 provided it is shown that

$$\operatorname{rank}[A'_* \Gamma A_1\ A'_* \Gamma N_0 A_2] = r \qquad \text{almost surely.} \tag{5.2}$$

To show this, note that $\operatorname{vec}(A'_* \Gamma N_0 A_2) \sim \operatorname{Normal}(\mathbf{0}, \Sigma_*)$, where $\Sigma_* = C'_* \Sigma C_*$ with $C_* = \sum_{\lambda \in w} \sum_{\mu \notin w} (\lambda - \mu)^{-1} P_\mu A_2 \otimes \Gamma P_\lambda A_*$. If G is an $m \times (r - r_0)$ matrix, then $\operatorname{vec}(G)' \Sigma_* \operatorname{vec}(G) = \operatorname{vec}(\Gamma B)' \Sigma_* \operatorname{vec}(\Gamma B)$, where $B = \sum_{\lambda \in w} \sum_{\mu \notin w} (\lambda - \mu)^{-1} P_\lambda A_* G A'_2 P'_\mu$. It is then easy to verify that if $G \neq 0$, then $M(B + B') \neq 0$. So, by Assumption 3.1(iii), Σ_* is nonsingular.

Since Σ_* is nonsingular, $\operatorname{rank}[A'_* \Gamma N_0 A_2] = r - r_0$ almost surely by Lemma 2.5, and $A'_* \Gamma N_0 A_0$ is almost surely linearly independent of $A'_* \Gamma A_1 = [I, 0]'$. Statement (5.2) follows by noting that $\operatorname{rank}[A'_* \Gamma A_1] = r_0$. The proof is thus complete.

By the previous theorem, $T_n(A)$ is unique on the set $\{\operatorname{rank}(\hat{P}_0 A) = r\}$, and the Moore-

Penrose inverse for $\hat{\Sigma}_0(A)$ can thus be used on this set. In addition, if M_n is symmetric, then $T_n(A)$ has the representation

(5.3) $$T_n(A) = a_n^2[\text{vec}(A)]'[\hat{\Sigma}_0(A)]^+\text{vec}(A)$$

on the set $\{\text{rank}(\hat{P}_0A) = r\}$. This statement is justified by noting that for a symmetric matrix B, $B\mathbf{x} = \mathbf{0}$ if and only if $B^+\mathbf{x} = \mathbf{0}$. It is easy to verify that $\hat{\Sigma}_0(A)\text{vec}(\hat{P}_0A) = \mathbf{0}$.

The asymptotic distribution of $T_n(A)$ is given in the next theorem.

THEOREM 5.3.
(i) *If* $P_0A = A$, *then* $T_n(A) \to_d \chi^2_{r(p-m)}$.
(ii) *If* $A_n = A + a_n^{-1}B$ *with* $P_0A = A$ *and* $P_0B = 0$, *then* $T_n(A_n) \to_d \chi^2_{r(p-m)}[\delta(A, B)]$, *where* $\delta(A, B) = \text{vec}(B)'[\Sigma_0(A)]^+\text{vec}(B)$.
(iii) *If* $P_0A \neq A$, *then for any fixed* x, $\text{Prob}[T_n(A) > x] \to 1$.

PROOF. (i) By Lemma 2.2, Theorems 4.2 and 5.1(ii), and Corollary 5.1(ii), it follows that $[\hat{\Sigma}_0(A)]^+ \to [\Sigma_0(A)]^+$ in probability. The theorem is then obtained by application of Lemma 2.4(i) and Theorem 4.1.

(ii) By Theorem 4.1, $\text{vec}[a_n(I - \hat{P}_0)A_n] \to_d \text{Normal}[\text{vec}(B), \Sigma_0(A)]$. Also, by expanding $\hat{\Sigma}_0(A_n)$ and applying (4.4) it follows that $\hat{\Sigma}_0(A_n) = \hat{\Sigma}_0(A) + o_p(a_n^{-1+\epsilon})$ for any $\epsilon > 0$. So, by Theorem 4.2, $\hat{\Sigma}_0(A_n) \to \Sigma_0(A)$ in probability and by Theorem 5.1(i), $\text{vec}(B) \in \mathscr{M}[\Sigma_0(A)]$. The result thus follows from Lemma 2.4(ii).

(iii) By Corollary 5.2(iii) and Lemma 2.3(iii), it follows that $T_n(A) \geq a_n^2(\mathbf{c}_n'\mathbf{c}_n)^2 \cdot [\mathbf{c}_n'\hat{\Sigma}_0(A)\mathbf{c}_n]^{-1}$, where $\mathbf{c}_n = \text{vec}[(I - \hat{P}_0)A]$. The theorem follows since $\mathbf{c}_n \to \text{vec}[(I - P_0)A]$ in probability, which is nonzero.

In summary, consider the following test for H_0.

(5.4) Reject H_0 if either
(i) $\text{rank}(\hat{P}_0A) < r$, or
(ii) $\text{rank}(\hat{P}_0A) = r$ and $T_n(A) > \chi^2_{r(p-m),\hat{\alpha}}$

where $\chi^2_{k,\hat{\alpha}}$ is the $(1 - \alpha)$ percentile of a χ^2_k distribution. By Theorems 5.2 and 5.3, this test is a well-defined consistent asymptotic α level test for H_0. Its local power function is given by Theorem 5.3(ii) and it is invariant under the transformation $A \to AB$ for any nonsingular matrix B.

By theorem 5.2(ii), it is irrelevant to the asymptotic properties of a test of H_0 what action is taken on the set $\{\text{rank}(\hat{P}_0A) < r\}$. However, rejecting H_0 for this case enables the rejection region to be "continuous" in the sense given in the following theorem. This property is important when using the test defined by (5.4) for constructing confidence regions for the range of P_0, as is done in Section 6.

THEOREM 5.4. *If* $\{A_k\}$ *is any sequence such that* $\text{rank}(\hat{P}_0A_k) = r$, $A_k \to A$, *and* $\text{rank}(\hat{P}_0A) < r$, *then* $T_n(A_k) \to \infty$, *as* $k \to \infty$.

PROOF. Let $r_0 = \text{rank}(\hat{P}_0A)$, and let B be an $r \times (r - r_0)$ matrix with $\text{rank}(B) = r - r_0$ and such that $\hat{P}_0AB = 0$. By Corollary 5.1(iii), we can apply Lemmas 2.3(ii) and 2.3(iii) to obtain

(5.5) $$T_n(A_k) \geq T_n(A_kB) \geq a_n^2(\mathbf{b}_k'\mathbf{b}_k)^2[\mathbf{b}_k'\hat{\Sigma}_0(A_kB)\mathbf{b}_k]^{-1},$$

where $\mathbf{b}_k = \text{vec}[(I - \hat{P}_0)A_kB]$. As $k \to \infty$, $\hat{\Sigma}_0(A_kB) \to 0$ and $\mathbf{b}_k \to \text{vec}(AB)$, which is nonzero. Thus, the right-hand side of (5.5) goes to infinity.

REMARK 1. If the assumption $\lambda_{i-1} \neq \lambda_i$ or the assumption $\lambda_{i+m-1} \neq \lambda_{i+m}$ is false, then the asymptotic chi-square test given by (5.4) is not generally valid. If the assumptions on the eigenvalues are true, then by Lemma 4.1, the "sample" size n necessary to insure that the asymptotic chi-square test is a "good" approximation is in general inversely related to the quantity $\min(\lambda_{i-1} - \lambda_i, \lambda_{i+m-1} - \lambda_{i+m})$. In addition, if λ_{i-1} is "close" to λ_i, one may not wish to study the eigenspace associated with λ_{i-1} separately from the eigenspace associated with λ_i. So, in practice, before determining which eigenspaces are of interest, a study of the eigenvalues is necessary.

REMARK 2. Let $v = \{\lambda_j, j \in I\}$, where I is some index set. Under the assumption $\lambda_j \neq \lambda_k$ for all $j \in I$ and $k \notin I$, consider the hypothesis

(5.6) H_0: the columns of A lie in the subspace generated by the eigenvectors of M associated with $\{\lambda_j, j \in I\}$,

where A is $p \times r$ with $\operatorname{rank}(A) = r \leq m = \operatorname{rank}(\sum_{\lambda \in v} P_\lambda)$. This hypothesis can be tested by using the test given by (5.4) provided w is replaced by v and $\hat{w}$ is replaced by $\hat{v} = \{\hat{\lambda}_j, j \in I\}$.

REMARK 3. For $r \geq m$, the hypothesis H_0^* given by (1.2) can be tested by using the following approach. Let B be $p \times (p - r)$ with $\operatorname{rank}(B) = p - r$ and such that $A'B = 0$. The hypothesis (1.2) can then be rephrased as

(5.7) H_0^*: the columns of B lie in the subspace generated by the eigenvectors of M' associated with the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_{i-1}, \lambda_{i+m}, \ldots, \lambda_p$.

Note that if M is symmetric in the metric of Γ, then M' is symmetric in the metric of Γ^{-1}. It is easy to verify that if the conditions on M, Γ, M_n, Γ_n, Σ and Σ_n given by Assumptions 3.1 and 3.2 are satisfied, then the conditions are satisfied when M, Γ, M_n, Γ_n, Σ and Σ_n are replaced by M', Γ^{-1}, M_n', Γ_n^{-1}, $\operatorname{var}(N') = I_{(p,p)} \Sigma I_{(p,p)}$ and $I_{(p,p)} \Sigma_n I_{(p,p)}$ respectively. So, by Remark 2, the results of this section apply to testing the hypothesis (5.7).

Note that if $r = m$, then the hypothesis (1.1) and (1.2) are equivalent. For this case, the test given by (5.4) when applied to the hypothesis (1.1) is the same as the test for (1.2) suggested in this remark.

6. Asymptotic confidence regions. The test of H_0 given by (5.4) yields the following asymptotic $(1 - \alpha)$ confidence region for the range of P_0,

(6.1) $$\{\mathscr{M}(A) \mid A \text{ is } p \times m, \operatorname{rank}(A) = m, \quad \text{and} \quad T_n(A) < \chi^2_{m(p-m), \hat{\alpha}}\}.$$

One "undesirable" aspect of this confidence region is that $T_n(A)$ involves a generalized inverse of $\hat{\Sigma}_0(A)$, which must be recalculated for each A. However, this problem can be alleviated and the confidence region can be given a simpler representation.

To make the simplification, let

(6.2) $$X_n = [\hat{\mathbf{x}}_i \, \hat{\mathbf{x}}_{i+1} \cdots \hat{\mathbf{x}}_{i+m-1}],$$

where $\{\hat{\mathbf{x}}_j\}$ is defined such that $M_n \hat{\mathbf{x}}_j = \hat{\lambda}_j \hat{\mathbf{x}}_j$, and $\hat{\mathbf{x}}_j' \Gamma_n \hat{\mathbf{x}}_k = \delta_{jk}$. By noting that $\hat{P}_0 = X_n X_n' \Gamma_n$, it can then be easily verified that

$$\hat{\Sigma}_0(A) = (A'\Gamma_n X_n \otimes I) \, \hat{\Sigma}_0(X_n)(X_n'\Gamma_n A \otimes I).$$

So, by Theorem 5.2(ii), if $\operatorname{rank}(\hat{P}_0 A) = m$, then

$$T_n(A) = T_n[A(X_n'\Gamma_n A)^{-1}] = a_n^2 \{\operatorname{vec}[A(X_n'\Gamma_n A)^{-1} - X_n]\}' \, \hat{\Sigma}_0(X_n)^+ \operatorname{vec}[A(X_n'\Gamma_n A)^{-1} - X_n].$$

Thus, (6.1) can be rewritten as

(6.3) $$\{\mathscr{M}(A) \mid X_n'\Gamma_n A = I \quad \text{and} \quad a_n^2[\operatorname{vec}(A - X_n)]'\hat{\Sigma}_0(X_n)^+ \operatorname{vec}(A - X_n) < \chi^2_{m(p-m), \hat{\alpha}}\},$$

For the special case $m = 1$, (6.3) reduces to

$$\{c\mathbf{a} \mid \hat{\mathbf{x}}_i'\Gamma_n\mathbf{a} = I, \quad \text{and} \quad a_n^2(\mathbf{a} - \hat{\mathbf{x}}_i)'\Lambda_n^+(\mathbf{a} - \hat{\mathbf{x}}_i) < \chi^2_{p-1,\hat{\alpha}}\}, \tag{6.4}$$

where $\Lambda_n = [\hat{\mathbf{x}}_i' \otimes (M_n - \hat{\lambda}_i I)^+]\Sigma_n[\hat{\mathbf{x}}_i \otimes (M_n' - \hat{\lambda}_i I)^+]$.

If M_n and M are symmetric, (6.3) and (6.4) respectively reduce to

$$\{\mathcal{M}(A) \mid X_n'A = I \quad \text{and} \quad a_n^2[\text{vec}(A)]'\hat{\Sigma}_0(X_n)^+[\text{vec}(A)] < \chi^2_{m(p-m),\hat{\alpha}}\}, \tag{6.5}$$

and

$$\{c\mathbf{a} \mid \hat{\mathbf{x}}_i'\mathbf{a} = 1, \quad \text{and} \quad a_n^2\mathbf{a}'\Lambda_n^+\mathbf{a} < \chi^2_{p-1}, \hat{\lambda}\}, \tag{6.6}$$

where $\Lambda_n = [\hat{\mathbf{x}}_i' \otimes (M_n - \hat{\lambda}_i I)^+]\Sigma_n[\hat{\mathbf{x}}_i \otimes (M_n - \hat{\lambda}_i I)^+]$.

7. Applications.

7.1. *Principal Components Analysis.* One of the most common uses of eigenvectors in statistics is in the principal components analysis of a covariance matrix. For this case, let M_n be the sample covariance matrix from a sample of size n from an elliptical distribution with nonsingular covariance matrix M. That is, $M_n = (1/n)\sum_{j=1}^n (\mathbf{Y}_j - \bar{\mathbf{Y}})(\mathbf{Y}_j - \bar{\mathbf{Y}})'$, where $\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_n$ are independent and identically distributed as $\mathbf{Y}$, where $\mathbf{Y}$ has density function of the form $f(\mathbf{y}) = c|M|^{-1/2}g[(\mathbf{y} - \boldsymbol{\mu})'M^{-1}(\mathbf{y} - \boldsymbol{\mu})]$ for some constant c and nonnegative function g. Also, assume that g is defined such that the fourth moments of $\mathbf{Y}$ exist.

It is well-known that $\sqrt{n}(M_n - M) \to_d N$, where N is multivariate normal, mean zero, and with covariances given by

$$\text{Cov}(n_{ij}, n_{k\ell}) = m_{ik}m_{j\ell} + m_{i\ell}m_{jk} + \kappa(m_{ij}m_{k\ell} + m_{ik}m_{j\ell} + m_{i\ell}m_{jk}), \tag{7.1}$$

where κ is a "kurtosis" parameter defined such that 3κ is the kurtosis of any marginal distribution of $\mathbf{Y}$. (For example, see Muirhead and Waternaux (1980).) If $\mathbf{Y}$ is multivariate normal, then $\kappa = 0$.

If Σ represents the covariance matrix of vec(N), then (7.1) can be reexpressed as

$$\Sigma = (1 + \kappa)(I + I_{(p,p)})(M \otimes M) + \kappa \,\text{vec}(M)[\text{vec}(M)]'. \tag{7.2}$$

Let $\hat{\kappa}$ be some consistent estimate of κ, and let $\Sigma_n = (1 + \hat{\kappa})(I + I_{(p,p)})(M_n \otimes M_n) + \hat{\kappa}\,\text{vec}(M_n)[\text{vec}(M_n)]'$.

It can be verified that Assumptions 3.1 and 3.2 are satisfied, and so the results of this paper apply to this example. In particular,

$$\Sigma_0(A) = (1 + \kappa)\sum_{\lambda\in\omega}\sum_{\mu\notin\omega}\mu\lambda/(\lambda - \mu)^2A'P_\lambda A \otimes P_\mu. \tag{7.3}$$

On the set $\{\text{rank}(\hat{P}_0A) = r\}$,

$$T_n(A) = n(1 + \hat{\kappa})\sum_{\mu\notin\hat{\omega}}\mu^{-1}\text{Trace}\{A'\hat{P}_\mu A[A'X_nD_n(\mu)X_n'A]^{-1}\}, \tag{7.4}$$

where X_n is defined in (6.2) and $D_n(\mu)$ is an $m \times m$ diagonal matrix with entries $\hat{\lambda}_j/(\hat{\lambda}_j - \mu)^2, j = i, i+1, \ldots, i+m-1$. For $r = m$, by Theorem 5.2 (ii), A can be normalized such that $A'X_n = I$ and (7.4) can be expressed as

$$T_n(A) = n(1 + \hat{\kappa})\text{Trace}[A'M_n^{-1}A\Delta_n + A'M_nA\Delta_n^{-1} - 2A'A] \tag{7.5}$$

where Δ_n is an $m \times m$ diagonal matrix with entries $\hat{\lambda}_i, \hat{\lambda}_{i+1}, \ldots, \hat{\lambda}_{i+m-1}$.

Under the additional assumption $\lambda_i = \lambda_{i+1} = \cdots = \lambda_{i+m-1}$, (7.4) is asymptotically equivalent to the statistic

$$t_n(A) = n(1 + \hat{\kappa})\text{Trace}[\hat{\lambda}A'M_n^{-1}A + \hat{\lambda}^{-1}A'M_nA - 2A'A], \tag{7.6}$$

where $\hat{\lambda} = m^{-1} \sum_{j=i}^{i+m-1} \hat{\lambda}_i$ and A is normalized so that $A'\hat{P}_0 A = I$. Statement (7.6) is not valid if the assumption $\lambda_i = \lambda_{i+1} = \cdots = \lambda_{i+m-1}$ does not hold.

For $r = m = 1$ and $\mathbf{Y}$ multivariate normal, $T_n(A)$ is asymptotically equivalent under H_0 to the chi-square statistics given by Anderson (1963). Anderson shows that

$$n[\hat{\lambda}_i \mathbf{a}' M_n^{-1} \mathbf{a} + \hat{\lambda}_i^{-1} \mathbf{a}' M_n \mathbf{a} - 2] \to_d \chi^2_{p-1}, \tag{7.7}$$

when λ_i is distinct, $M\mathbf{a} = \lambda_i \mathbf{a}$ and with $\mathbf{a}$ normalized so that $\mathbf{a}'\mathbf{a} = 1$.

7.2 *Canonical Analysis.* Another common use of eigenvectors in statistics is in canonical analysis. Let C_n represent the sample covariance matrix for a sample of size n from a $(p + q)$ random vector with an elliptical distribution, nonsingular covariance matrix C, and finite fourth moments. For this case, $M_n = \hat{C}_{11}^{-1}\hat{C}_{12}\hat{C}_{22}^{-1}\hat{C}_{21}$ and $M = C_{11}^{-1}C_{12}C_{22}^{-1}C_{21}$, where

$$C_n = \begin{bmatrix} \hat{C}_{11} & \hat{C}_{12} \\ \hat{C}_{21} & \hat{C}_{22} \end{bmatrix} \quad \text{and} \quad C = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}.$$

In this example, $\Gamma = C_{11}$ and $\Gamma_n = \hat{C}_{11}$.

By expanding M_n in a Taylor series about C and defining N_{ij}^* such that $\sqrt{n}(\hat{C}_{ij} - C_{ij}) \to_d N_{ij}^*$, we obtain $\sqrt{n}(M_n - M) \to_d N$ where

$$N = C_{11}^{-1}N_{12}^*C_{22}^{-1}C_{21} + C_{11}^{-1}C_{12}C_{22}^{-1}N_{21}^* - C_{11}^{-1}N_{11}^*C_{11}^{-1}C_{12}C_{22}^{-1}C_{21} - C_{11}^{-1}C_{12}C_{22}^{-1}N_{22}^*C_{22}^{-1}C_{21}. \tag{7.8}$$

The random matrix N is multivariate normal with mean zero. The covariance matrix for $\text{vec}(N)$ is obtained from (7.8) by using the form of the covariance matrix for the limiting normal distribution of $\text{vec}[\sqrt{n}(C_n - C)]$ given in (7.2). The covariance matrix for $\text{vec}(N)$ is thus

$$\begin{aligned} \Sigma = (1 + \kappa)\{&\Gamma(I - M) \otimes \Gamma^{-1}M' + \Gamma(M - M^2) \otimes \Gamma^{-1}(I - 2M') \\ &+ I_{(p,p)}[(I - M) \otimes (M - M^2)' + (M - M^2) \otimes (I - M')]\}. \end{aligned} \tag{7.9}$$

Let $\hat{\kappa}$ be a consistent estimate of κ, and let Σ_n have the same form as Σ with Γ_n, M_n, and $\hat{\kappa}$ replacing Γ, M and κ respectively. For $C_{12} \neq 0$, it can be verified that Assumptions 3.1 and 3.2 are satisfied, and so the results of this paper apply to this example. In particular,

$$\Sigma_0(A) = (1 + \kappa) \sum_{\lambda \in w} \sum_{\mu \notin w} (1 - \lambda)(\mu + \lambda - 2\mu\lambda)/(\lambda - \mu)^2 A'\Gamma P_\lambda A \otimes \Gamma^{-1}P_\mu'. \tag{7.10}$$

On the set $\text{rank}(\hat{P}_0 A) = r$, we have the representation

$$T_n(A) = n(1 + \hat{\kappa}) \sum_{\mu \notin \hat{w}} \text{Trace}\{A'\Gamma_n\hat{P}_\mu A[A'\Gamma_n X_n D_n(\mu) X_n' \Gamma_n A]^{-1}\}, \tag{7.11}$$

where $D_n(\mu)$ is an $m \times m$ diagonal matrix with entries

$$(1 - \hat{\lambda}_j)(\mu + \hat{\lambda}_j - 2\mu\hat{\lambda}_j)/(\hat{\lambda}_j - \mu)^2, \quad j = i, i + 1, \ldots, i + m - 1.$$

In particular, for $m = 1$ and $\hat{\lambda}_i \neq 0$, (7.11) becomes

$$T_n(\mathbf{a}) = n(1 + \hat{\kappa})\mathbf{a}'\Gamma_n(M_n - \hat{\lambda}_i I)^2[(1 - 2\hat{\lambda}_i)M_n + \hat{\lambda}_i I]^{-1}\mathbf{a}/[(1 - \hat{\lambda}_i)(\mathbf{a}'\Gamma_n\hat{\mathbf{x}}_i)^2]. \tag{7.12}$$

For the special case $\text{rank}(M) = i - 1 < p$, (7.11) is asymptotically equivalent under H_0 to

$$t_n(A) = n(1 + \hat{\kappa})\text{Trace}[A'\hat{C}_{12}\hat{C}_{22}^{-1}\hat{C}_{21}A][A'(\hat{C}_{11} - \hat{C}_{12}\hat{C}_{22}^{-1}\hat{C}_{21})A]^{-1}. \tag{7.13}$$

If the elliptical distribution $\mathbf{Y}' = (\mathbf{Y}_1' \; \mathbf{Y}_2')$ is multivariate normal and we choose $\hat{\kappa} = 0$, then (7.13) is the Lawley-Hotelling trace statistic for testing independence between $A'\mathbf{Y}_1$ and $\mathbf{Y}_2$.

8. Relaxation of Assumptions. The results of Section 4 do not depend upon Assumptions 3.1 (iii) and 3.2 (iii). However, the results of Section 5 do. If these two assumptions do not hold, then for the case $r \le m$ the quadratic form

$$t_n(A) = a_n^2\{\text{vec}[(I - \hat{P}_0)A]\}'[\hat{\Sigma}_0(A)]^+ \text{vec}[(I - \hat{P}_0)A] \tag{8.1}$$

has the following limiting distribution.

THEOREM 8.1. *Let* $\nu(A) = \text{rank}[\Sigma_0(A)]$.

(i) *If* $P_0A = A$ *and* $\text{Prob}\{\text{rank}[\hat{\Sigma}_0(A)] = \nu(A)\} \to 1$, *then* $t_n(A) \to_d \chi^2_{\nu(A)}$.

(ii) *If* $A_n = A + a_n^{-1}B$ *with* $P_0A = A$, $P_0B = 0$, $\text{vec}(B) \in M[\Sigma_0(A)]$, *and* $\text{Prob}\{\text{rank}[\hat{\Sigma}_0(A_n)] = \nu(A)\} \to 1$, *then* $t_n(A_n) \to_d \chi^2_{\nu(A)}[\delta(A, B)]$, *where* $\delta(A, B) = [\text{vec}(B)]'\Sigma_0(A)^+\text{vec}(B)$.

PROOF. The proof for this theorem is analogous to the proof for Theorem 5.3(i) and 5.3(ii). The results follow from Theorems 4.1 and 4.2 by applying Lemmas 2.2 and 2.4.

In using $t_n(A)$ as a test statistic for H_0, it is important to note the following shortcomings. In general, the value of $\nu(A)$ may depend upon the hypothesized value of A, and the property $\text{Prob}\{\text{rank}[\hat{\Sigma}_0(A)] = \nu(A)\} \to 1$ may hold for some hypothesized value of A, and may not hold for others. The definition of $t_n(A)$ is dependent on the use of the Moore-Penrose generalized inverse, and the invariance property $t_n(A) = t_n(AB)$ for any nonsingular B may not hold. Finally, the consistency of the test statistic, that is $\text{Prob}[t_n(A) > x] \to 1$ for all x whenever $P_0A \neq A$, does not necessarily hold in general.

If Assumptions 3.1(iii) and 3.2(iii) do not hold, then the above conditions must be investigated for the specific problem. In such cases, a general chi-square test may not be desirable. Other tests could be constructed using Theorem 4.1 and the specific form of $\Sigma_0(A)$.

As an aid in understanding Assumption 3.1(iii), consider the case when $\Gamma = \Gamma_n = I$ and M is nonsingular. Assumption 3.1(iii) states that if $\Sigma\,\text{vec}(B) = 0$, then $B + B' = 0$. This is equivalent to stating that the covariance matrix of the $p(p + 1)/2$ diagonal and upper triangular entries of N must be nonsingular. An important example when this condition does not hold is when M_n and M are the sample and population correlation matrices respectively.

REFERENCES

ANDERSON, T. W. (1951). The asymptotic distribution of certain characteristic roots and vectors. *Proc. Second Berkeley Symp. Math. Stat. Probability* 103–130. Univ. California Press.

ANDERSON, T. W. (1963). Asymptotic theory for principal components analysis. *Ann. Math. Stat.* **34** 122–148.

CHAMBERS, J. M. (1967). On methods of asymptotic approximations for multivariate distributions. *Biometrika* **54** 367–383.

HIRAKAWA, F. (1976a). On a test for the equality of the characteristic spaces of two population covariance matrices. *TRU Math.* **12** 75–81.

HIRAKAWA, F. (1976b). Asymptotic distributions of test criteria for the identities of the characteristic spaces of two population covariance matrices. *TRU Math.* **12** 83–92.

IZENMAN, A. J. (1976). Reduced-rank regression for the multivariate linear model. *J. Multivariate Anal.* **5** 248–264.

JAMES, A. T. (1977). Test for a prescribed subspace of principal components. In *Multivariate Analysis—IV* (P. R. Krishnaiah, ed.) 73–77. North Holland, Amsterdam.

KATO, T. (1966). *Perturbation Theory for Linear Operators.* Springer, Berlin.

MAGNUS, J. R. and NEUDECKER, H. (1979). The commutation matrix: some properties and applications. *Ann. Statist.* **7** 381–394.

MALLOWS, C. L. (1961). Latent vectors of random symmetric matrices. *Biometrika* **48** 133–149.

MOORE, D. S. (1977). Generalized inverses, Wald's method, and the construction of chi-squared tests of fit. *J. Amer. Statist. Assoc.* **72** 131–137.

MUIRHEAD, R. J. and WATERNAUX, C. M. (1980). Asymptotic distributions in canonical correlation analysis and other multivariate procedures for non-normal populations. *Biometrika* **67** 31–43.

NERRING, E. D. (1970). *Linear Algebra and Matrix Theory*, 2nd Ed. Wiley, New York.

NEUDECKER, H. (1968). The Kronecker matrix product and some of its applications. *Statistica*

Neerlandica **22** 69–82.

OKAMOTO, M. (1973). Distinctness of the eigenvalues of a quadratic form in a multivariate normal sample. *Ann. Statist.* **1** 763–765.

RAO, C. R. and MITRA, S. K. (1971) *Generalized Inverses of Matrices and its Application.* Wiley, New York.

SUGIURA, N. (1976). Asymptotic expansions of the distributions of the latent roots and latent vectors of the Wishart and multivariate F matrices. *J. Multivariate Anal.* **6** 500–525.

TYLER, D. E. (1979). Redundancy analysis and associated asymptotic distribution theory. Ph.D. dissertation, Dept. Statist., Princeton Univ.

DEPARTMENT OF MATHEMATICAL SCIENCES
OLD DOMINION UNIVERSITY
NORFOLK, VIRGINIA 23508