The Annals of Statistics
1978, Vol. 6, No. 1, 5–33

LATENT ROOTS AND MATRIX VARIATES: A REVIEW OF SOME ASYMPTOTIC RESULTS

BY ROBB J. MUIRHEAD[1]

Yale University

The exact noncentral distributions of matrix variates and latent roots derived from normal samples involve hypergeometric functions of matrix argument. These functions can be defined as power series, by integral representations, or as solutions of differential equations, and there is no doubt that these mathematical characterizations have been a unifying influence in multivariate noncentral distribution theory, at least from an analytic point of view. From a computational and inference point of view, however, the hypergeometric functions are themselves of very limited value due primarily to the many difficulties involved in evaluating them numerically and consequently in studying the effects of population parameters on the distributions. Asymptotic results for large sample sizes or large population latent roots have so far proved to be much more useful for such problems. The purpose of this paper is to review some of the recent results obtained in these areas.

1. Introduction and summary. The classic 1964 paper of A. T. James provides a survey of exact noncentral distributions of matrix variates and latent roots derived from normal samples. These distributions, and consequently the likelihood functions of noncentrality parameters and population roots, all involve hypergeometric functions of matrix argument, functions which have power series representations in terms of zonal polynomials. These series, however, tend to converge extremely slowly for cases of particular interest (for example, large sample sizes, large population roots) and it is very difficult to obtain from them any feeling for the behavior of the density and likelihood functions. In particular, two problems are generally of interest:

(a) Where do the regions of appreciable likelihood occur? What are the parameter values which maximize the likelihood?

(b) In making inferences about a subset of population latent roots the other roots are nuisance parameters. What effect do these have on such inferences?

The first problem is primarily a mathematical one while the second is statistical; unfortunately the power series expansions for the hypergeometric functions shed little or no light on either.

In the years since James' paper appeared a number of authors have worked

Received December 1976; Revised July 1977.

[1] The preparation of this paper was supported by National Science Foundation Grant MCS 76-14876.

AMS 1970 *subject classifications*. Primary 62H10, 62E20.

Key words and phrases. Latent roots, matrix variates, asymptotic distributions, hypergeometric functions, principal components, noncentral means, discriminant analysis, canonical correlations.

on the problems of approximating the distributions and likelihood functions by finding asymptotic representations for the hypergeometric functions involved. Such an approach leads to solutions (or at least to asymptotic solutions) of the problems posed above since these asymptotic representations can be written in terms of more elementary functions or in terms of functions which are comparatively easy to compute, and as such permit an examination of the way in which sample and population latent roots interact with each other. The purpose of this paper is to review some of the work done in this area; the paper is mainly expository but a few new results are included.

The distributions reviewed here follow those in James (1964); the latent root distributions are (i) the roots of a covariance matrix (Section 3), (ii) roots when $\Sigma_1 \neq \Sigma_2$ (Section 4), (iii) noncentral means with known covariance (Section 6), (iv) noncentral roots in multiple discriminant analysis (Section 8) and (v) canonical correlation coefficients (Section 9), while the matrix variate distributions are the noncentral Wishart (Section 5) and the noncentral multivariate F (Section 7). Section 3 is the longest; this is not only because more has been written about the distribution of the latent roots of a Wishart matrix but also because many of the comments made there, particularly with respect to problem areas, are applicable also to the other root distributions and are not repeated elsewhere.

It should perhaps be said that the main emphasis of this paper is in the area of *asymptotic representations* for distributions of sample roots and matrix variates, rather than the area of *asymptotic distributions* of suitably standardized variables. These asymptotic representations for latent root distributions involve "linkage factors" of the form $\alpha_i - \alpha_j$ corresponding to *distinct* population roots α_i and α_j (an advantage in an investigation of likelihood functions since the effects of the population roots become obvious) whereas the asymptotic distributions do not preserve such linkage factors. This is not to say the asymptotic distributions are not important; they are indeed but they do not form the primary motivation for this work although some results from this area are also included. One of the most interesting facets about asymptotic representations for joint distributions of latent roots is that they yield asymptotic representations for conditional distributions of subsets of roots given the remainder which do not depend on the population roots corresponding to the conditioned sample roots. This suggests, for example, testing equality of a subset of population roots using such a conditional distribution since it is then possible, in an asymptotic sense, to eliminate the effects of nuisance parameters. Some work in this area is also reviewed. Finally it should be noted that in most cases no attempt has been made to present results in complete generality. In the latent root distributions different asymptotic results can be obtained by varying the multiplicities of the population roots. In practice, however, it is generally of interest to test equality of a subset of roots, usually the smallest. Most of the asymptotic results given here cover the situation where the smallest population root is a multiple one while the remaining ones are simple.

2. Preliminaries. There are essentially three ways of defining the hypergeometric functions of matrix argument which occur in noncentral multivariate distributions. These are briefly reviewed, with some comments on their uses and limitations.

(i) *Power series.* The matrix variate distributions involve hypergeometric functions ${}_pF_q$, with an $m \times m$ matrix R as argument, defined by the series (Constantine (1963), James (1964))

$$(2.1)\qquad {}_pF_q(a_1, \cdots, a_p; b_1, \cdots, b_q; R) = \sum_{k=0}^{\infty} \sum_{\kappa} \frac{(a_1)_\kappa \cdots (a_p)_\kappa}{(b_1)_\kappa \cdots (b_q)_\kappa} \frac{C_\kappa(R)}{k!}$$

where $C_\kappa(R)$ is the zonal polynomial of R (a homogeneous symmetric polynomial of degree k in the latent roots of R) corresponding to the partition $\kappa = (k_1, \cdots, k_m)$, $k_1 \geqq \cdots \geqq k_m \geqq 0$, of k and

$$(a)_\kappa = \prod_{i=1}^{m} (a - \tfrac{1}{2}(i-1))_{k_i}, \qquad (x)_k = x(x+1)\cdots(x+k-1).$$

The series (2.1) is a generalization of the classical (generalized) hypergeometric function, to which it reduces when $m = 1$. The latent root distributions involve hypergeometric functions ${}_pF_q^{(m)}$, with two $m \times m$ matrices R and S as arguments, defined as

$$(2.2)\qquad {}_pF_q^{(m)}(a_1, \cdots, a_p; b_1, \cdots, b_q; R, S) = \sum_{k=0}^{\infty} \sum_{\kappa} \frac{(a_1)_\kappa \cdots (a_p)_\kappa}{(b_1)_\kappa \cdots (b_q)_\kappa} \frac{C_\kappa(R) C_\kappa(S)}{k!\, C_\kappa(I_m)}.$$

Although no explicit formula in general is known for the zonal polynomials, tables and algorithms are available for their computation (see James (1966, 1968), Parkhurst and James (1974), McLaren (1976)) and in principle the series (2.1) and (2.2) could be used for numerical work. Unfortunately, however, they tend to converge very slowly if even one of the latent roots of the argument matrix or matrices is large. Another type of convergence problem arises when the roots in the argument are not large (as in the distribution of canonical correlation coefficients where the roots lie between 0 and 1) so that as k increases the zonal polynomials quickly become numerically very small. At the same time however, the other terms $(a_i)_\kappa$ in the numerator rapidly become very large, the net effect being that convergence of the series is extremely slow. A discussion of the computational problems involved in computing zonal polynomials has been given by McLaren (1976). A review of the literature on zonal polynomials has been given by Subrahmaniam (1974), and Farrell (1976) has given a detailed discussion of their group theoretic construction.

It is worth pointing out that when the argument matrices are of size two it is possible to express some of these functions in terms of series of classical hypergeometric functions. For example, the ${}_2F_1$ function in this case has the expansion

$$(2.3)\qquad {}_2F_1(a, b; c; R) = \sum_{k=0}^{\infty} \frac{(a)_k (c-a)_k (b)_k (c-b)_k}{(c)_{2k} (c-\frac{1}{2})_k} \frac{(r_1 r_2)^k}{k!} \times {}_2F_1(a+k, b+k; c+2k; r_1 + r_2 - r_1 r_2) \quad \mathrm{Re}\,(c) > \tfrac{1}{2}$$

where r_1 and r_2 are the latent roots of R. Similar types of expansions for the ${}_1F_1$ and ${}_0F_1$ functions given by Herz (1955) and Muirhead (1975) follow from (2.3) via the confluence relations

$$ {}_1F_1(a; c; R) = \lim_{b\to\infty} {}_2F_1(a, b; c; b^{-1}R) \tag{2.4} $$

and

$$ {}_0F_1(c; R) = \lim_{a\to\infty} {}_1F_1(a; c; a^{-1}R) \,. \tag{2.5} $$

Expansions for some of the two matrix functions are also given in Muirhead (1975).

(ii). *Integral representations.* Starting with the function

$$ {}_0F_0(R) = \exp(\operatorname{tr} R) \,, $$

Herz (1955) defined the general system of hypergeometric functions by means of the Laplace and inverse Laplace transforms

$$ \begin{aligned} &{}_{p+1}F_q(a_1, \cdots, a_p, a; b_1, \cdots, b_q; -R^{-1})(\det R)^{-1} \\ &\quad = \frac{1}{\Gamma_m(a)} \int_{X>0} \exp(\operatorname{tr} -XR)\, {}_pF_q(a_1, \cdots, a_p; b_1, \cdots, b_q; -X) \\ &\qquad \times (\det X)^{a-\frac{1}{2}(m+1)}\, dX \end{aligned} \tag{2.6} $$

and

$$ \begin{aligned} &{}_pF_{q+1}(a_1, \cdots, a_p; b_1, \cdots, b_q, b; -R)(\det R)^{b-\frac{1}{2}(m+1)} \\ &\quad = \Gamma_m(b) \frac{2^{\frac{1}{2}m(m-1)}}{(2\pi i)^{\frac{1}{2}m(m+1)}} \int_{R(Z)=X_0>0} \exp(\operatorname{tr} RZ) \\ &\qquad \times {}_pF_q(a_1, \cdots, a_p; b_1, \cdots, b_q; -Z^{-1})(\det Z)^{-b}\, dZ \end{aligned} \tag{2.7} $$

where

$$ \Gamma_m(a) = \pi^{\frac{1}{4}m(m-1)} \prod_{i=1}^m \Gamma(a - \tfrac{1}{2}(i-1)) $$

and the integral (2.7) is taken over all matrices $Z = X_0 + iY$ for fixed positive definite X_0 and Y arbitrary real symmetric. The equivalence of these functions with the ones defined by the zonal polynomial series (2.1) was established by Constantine (1963). A large number of other integral representations for particular functions (primarily ${}_0F_1$, ${}_1F_1$, ${}_2F_1$) were derived from (2.6) and (2.7) by Herz (1955). The hypergeometric functions having two matrices as arguments follow from the one matrix functions via (James (1964))

$$ \begin{aligned} &{}_pF_q^{(m)}(a_1, \cdots, a_p; b_1, \cdots, b_q; R, S) \\ &\quad = \int_{O(m)} {}_pF_q(a_1, \cdots, a_p; b_1, \cdots, b_q; RH'SH)(dH) \end{aligned} \tag{2.8} $$

where (dH) is the invariant measure on the group $O(m)$ of orthogonal $m \times m$ matrices normalized so that the volume of $O(m)$ is unity.

Multiple integrals of the form (2.6), (2.7) and (2.8) do not lend themselves easily to exact numerical work; however, integral representations have proved the most useful tool in obtaining the asymptotic behaviors of hypergeometric functions, as indeed they are in investigations of the asymptotic behaviors of a large number of classical special functions. A multivariate extension of Laplace's method for integrals developed by Hsu (1948) has been widely used; this states that if a function $f(x) = f(x_1, \cdots, x_m)$ has an absolute maximum at an interior point ξ of a closed domain $\mathscr{D}$ in real m-dimensional space, then, under suitable smoothness conditions, as $n \to \infty$

$$\int_{\mathscr{D}} [f(x)]^n \phi(x)\, dx \sim \left(\frac{2\pi}{n}\right)^{\frac{1}{2}m} [f(\xi)]^n \phi(\xi) [\Delta(\xi)]^{-\frac{1}{2}} \tag{2.9}$$

where $a \sim b$ means that $a/b \to 1$ as $n \to \infty$, and $\Delta(\xi)$ denotes the Hessian of $-\log f$, namely

$$\Delta(\xi) = \det\left(-\frac{\partial^2 \log f(\xi)}{\partial \xi_i\, \partial \xi_j}\right).$$

In most applications of this result in the area of hypergeometric functions of matrix argument the space $\mathscr{D}$ is the orthogonal group $O(m)$, the Stiefel manifold $V(k, m)$ of $m \times k$ matrices with orthonormal columns, or products of such spaces.

(iii) *Differential equations.* Extending earlier work by James (1955) on the ${}_0F_1$ function, Muirhead (1970a) has shown that the commonly occurring hypergeometric functions of one matrix argument (${}_2F_1$, ${}_1F_1$, ${}_0F_1$) can be defined as solutions of systems of second order partial differential equations in the latent roots of the argument matrix. For example the function ${}_2F_1(a, b; c; R)$, where R has latent roots $r_1, \cdots, r_m$, is the unique solution of each of the m partial differential equations

$$\begin{aligned} r_i(1 - r_i)\frac{\partial^2 F}{\partial r_i^2} + \Big\{c - \tfrac{1}{2}(m-1) - [a + b + 1 - \tfrac{1}{2}(m-1)]r_i \\ + \tfrac{1}{2}\textstyle\sum_{j=1; j\neq i}^m \frac{r_i(1 - r_i)}{r_i - r_j}\Big\}\frac{\partial F}{\partial r_i} - \tfrac{1}{2}\textstyle\sum_{j=1; j\neq i}^m \frac{r_j(1 - r_j)}{r_i - r_j}\frac{\partial F}{\partial r_j} \\ = abF \qquad\qquad i = 1, 2, \cdots, m \end{aligned} \tag{2.10}$$

subject to the conditions that $F(R)$ be a symmetric function of $r_1, \cdots, r_m$, analytic at $R = 0$ with $F(0) = 1$. When $m = 1$ (2.10) reduces to the classical hypergeometric differential equation. Systems of differential equations for the functions ${}_1F_1$ and ${}_0F_1$ follow from (2.10) using the confluence relations (2.4) and (2.5). The two matrix functions have been studied by Constantine and Muirhead (1972); the function ${}_2F_1^{(m)}(a, b; c; R, S)$ is the unique solution of the partial differential equation

$$\begin{aligned} \textstyle\sum_{i=1}^m r_i \frac{\partial^2 F}{\partial r_i^2} + \sum_{i=1}^m \sum_{j=1; j\neq i}^m \frac{r_i}{r_i - r_j}\frac{\partial F}{\partial r_i} + [c - \tfrac{1}{2}(m-1)] \sum_{i=1}^m \frac{\partial F}{\partial r_i} \\ - (a + b + 2 - m) \textstyle\sum_{i=1}^m s_i^2 \frac{\partial F}{\partial s_i} - \sum_{i=1}^m s_i^3 \frac{\partial^2 F}{\partial s_i^2} \\ - \textstyle\sum_{i=1}^m \sum_{j=1; j\neq i}^m \frac{s_i^3}{s_i - s_j}\frac{\partial F}{\partial s_i} = ab \sum_{i=1}^m s_i F \end{aligned} \tag{2.11}$$

subject to the conditon that F has the series expansion

$$F(R, S) = \sum_{k=0}^{\infty} \sum_{\kappa} \alpha_{\kappa} \frac{C_{\kappa}(R)C_{\kappa}(S)}{C_{\kappa}(I_m)}$$

with $F(0, 0) = 1$. Differential equations for the corresponding ${}_1F_1^{(m)}$ and ${}_0F_1^{(m)}$ functions follow from (2.11) using obvious extensions of the confluences (2.4) and (2.5); a nontrivial differential equation for ${}_0F_0^{(m)}(R, S)$ is obtained by putting $a = c = \frac{1}{2}(m - 1)$ in the differential equation for ${}_1F_1^{(m)}(a; c; R, S)$. Further results in this area have been given by Constantine and Muirhead (1976), Chikuse (1976), Fujikoshi (1975) and Glynn (1977). Although the differential equations do not appear useful for obtaining the actual asymptotic behavior of these hypergeometric functions they have proved instrumental in the derivation of further terms in asymptotic series, see e.g., Muirhead (1970b, 1972a, b), Sugiura (1972, 1974), Muirhead and Chikuse (1975), Constantine and Muirhead (1976), Glynn (1977) and Chikuse (1976).

The hypergeometric functions are not the only classical functions which have been generalized for matrix arguments and it seems appropriate to mention in this section a few references where extensions of some other special functions may be found:

Bessel functions of the second kind (Herz (1955), Muirhead (1972b));

The second confluent hypergeometric function (Muirhead (1970b), Muirhead and Chikuse (1975));

Laguerre polynomials (Herz (1955), Constantine (1966), James and Constantine (1974), James (1976));

Hermite polynomials (Herz (1955), Hayakawa (1969), James (1976));

Jacobi polynomials (James and Constantine (1974), James (1976));

Gegenbauer polynomials (Herz (1955), James (1976)).

The paper by Herz is concerned almost entirely with multiple integral representations; most of the other papers referenced deal primarily either with differential equations or zonal polynomial expansions, or both.

A number of other review papers in the general area of multivariate distribution theory have been written in recent years and it is appropriate to conclude this section by referencing some which are particularly relevant to this work; these include papers by Crowther and Young (1974), Subrahmaniam (1974), Pillai (1976) and Krishnaiah (1977).

3. Latent roots of a covariance matrix. Of the latent root distributions surveyed in this paper those associated with principal component analysis have been the most widely studied, presumably because they are the least complex. Here some of the work done in the areas of asymptotics and inference is reviewed and some problems which appear to warrant further investigation are indicated.

The exact joint density function of the latent roots $l_1, l_2, \cdots, l_m$ of a sample covariance matrix S based on a sample of size $n + 1$ from an m-variate normal distribution with covariance matrix Σ is (James (1964))

$$(3.1)\qquad \frac{(\frac{1}{2}n)^{\frac{1}{2}mn}\pi^{\frac{1}{2}m^2}}{\Gamma_m(\frac{1}{2}n)\Gamma_m(\frac{1}{2}m)} \prod_{i=1}^m \alpha_i^{\frac{1}{2}n} \prod_{i=1}^m l_i^{\frac{1}{2}(n-m-1)} \prod_{i<j}^m (l_i - l_j)\, {}_0F_0^{(m)}(-\tfrac{1}{2}nL, A)\,,$$

$$l_1 > l_2 > \cdots > l_m > 0$$

where $0 < \alpha_1 \leqq \alpha_2 \leqq \cdots \leqq \alpha_m$ are the latent roots of the "information" matrix Σ^{-1}, $L = \operatorname{diag}(l_1, \cdots, l_m)$, $A = \operatorname{diag}(\alpha_1, \cdots, \alpha_m)$. James (1966) has argued that L is sufficient for A in the absence of knowledge about the latent vectors of Σ (this definition of partial sufficiency being a group invariant one due to Barnard (1963)) and suggests that, in such a situation, the distribution (3.1) of L be used as a basis for inference on the population roots. The marginal likelihood function of the population roots is, then,

$$(3.2)\qquad \prod_{i=1}^m \alpha_i^{\frac{1}{2}n}\, {}_0F_0^{(m)}(-\tfrac{1}{2}nL, A)\,.$$

A number of authors have studied the problem of approximating the ${}_0F_0$ function for large degrees of freedom n; to obtain its asymptotic behavior all have used Laplace's method (see Section 2) applied to the integral representation

$$(3.3)\qquad {}_0F_0^{(m)}(-\tfrac{1}{2}nL, A) = \int_{O(m)} \exp(\operatorname{tr} -\tfrac{1}{2}nLH'AH)(dH)$$

where (dH) is the invariant measure on the group $O(m)$ of orthogonal $m \times m$ matrices, normalized so that the volume of $O(m)$ is unity. The asymptotic behavior is basically determined by the maximum value of the integrand in (3.3) and the sharpness of the peaks of the integrand at its maxima; this sharpness depends fundamentally on the spread of the sample and population roots. Although the sample roots are distinct (with probability one) it is of course possible that all of the population roots $\alpha_1, \cdots, \alpha_m$ are widely spaced or that some of them are widely spaced while others are close together (and possibly equal). Asymptotic expansions for the ${}_0F_0$ function when $\alpha_1, \cdots, \alpha_m$ are widely spaced have been obtained by G. A. Anderson (1965), and by Bingham (1972) when $m = 3$. Asymptotic results in the case of one or more multiple population roots have been derived by James (1969), Chattopadhyay and Pillai (1973), Chikuse (1976) and Constantine and Muirhead (1976).

One of the most important and commonly used tests in principal component analysis is the likelihood ratio test of the null hypothesis that the q-smallest latent roots of Σ are all equal. If they are, then the variation in the last q dimensions is spherical, and, if their common value is small compared with the other $m - q$ roots, then most of the variation in the sample is explained by the first $m - q$ principal components, and a reduction in dimensionality is achieved. It seems reasonable here to concentrate on the case of one multiple population root (which can be chosen as the smallest) rather than on the more general situation of an arbitrary number of multiple roots.

The case when a group of the population roots are not widely spaced (and are possibly equal) has been studied by Constantine and Muirhead (1976). It is assumed that the first k roots of Σ are widely spaced but that the smallest $q = m - k$ roots are not. This assumption is expressed more precisely by requiring that

$$\alpha_1 < \alpha_2 < \cdots < \alpha_k < \alpha_{k+1} \leqq \alpha_{k+2} \leqq \cdots \leqq \alpha_m \tag{3.4}$$

where

$$\alpha_i - \alpha_j = O(n^{-1}) \qquad \text{for} \quad i, j = k+1, \cdots, m\,.$$

This, of course, includes the case when the q-smallest roots of Σ are equal and, when $k = m$, it gives the case when all the roots are widely spaced. Under the assumption (3.4) it is shown in Constantine and Muirhead (1976), using Laplace's method to obtain the limiting behavior and partial differential equations satisfied by the ${}_0F_0$ function to obtain correction terms, that

$$\begin{aligned} &{}_0F_0^{(m)}(-\tfrac{1}{2}nL, A) \\ &\sim \frac{\Gamma_k(\frac{1}{2}m)}{\pi^{\frac{1}{2}km}} \exp(-\tfrac{1}{2}n \textstyle\sum_{i=1}^{k} l_i\alpha_i) \prod_{i=1}^{k} \prod_{j=1;i<j}^{m} \left(\dfrac{2\pi}{nc_{ij}}\right)^{\frac{1}{2}} \\ &\quad \times {}_0F_0^{(m-k)}(-\tfrac{1}{2}nL_2, A_2)\left\{1 + \frac{1}{n}P_1 + \frac{1}{n^2}P_2 + O(n^{-3})\right\}, \end{aligned} \tag{3.5}$$

where

$$\begin{aligned} c_{ij} &= (l_i - l_j)(\alpha_j - \alpha_i)\,, \\ A_2 &= \mathrm{diag}\,(\alpha_{k+1}, \cdots, \alpha_m)\,, \qquad L_2 = \mathrm{diag}\,(l_{k+1}, \cdots, l_m)\,, \\ P_1 &= \tfrac{1}{2} \textstyle\sum_{i=1}^{k} \sum_{j=1;i<j}^{m} c_{ij}^{-1} \end{aligned}$$

and

$$P_2 = \textstyle\sum_{i=1}^{k} \sum_{j=1;i<j}^{m} c_{ij}^{-2} + \tfrac{1}{8}(\sum_{i=1}^{k} \sum_{j=1;i<j}^{m} c_{ij}^{-1})^2\,.$$

The term of order n^{-3} is also implicit in Constantine and Muirhead (1976). The exact power series expansion (2.2) for the function ${}_0F_0^{(m-k)}(-\frac{1}{2}nL_2, A_2)$ could be used for computational purposes. James (1966) has used such a combined asymptotic and power series expansion to numerically investigate the likelihood function (3.2). When $k = m$ (all population roots widely spaced) the ${}_0F_0^{(m-k)}$ function is taken to be unity; when $\alpha_{k+1} = \cdots = \alpha_m = \alpha$ (i.e., $A_2 = \alpha I_{m-k}$ and the smallest root of Σ has multiplicity q)

$${}_0F_0^{(m-k)}(-\tfrac{1}{2}nL_2, A_2) = \exp(-\tfrac{1}{2}n\alpha \textstyle\sum_{i=k+1}^{m} l_i)$$

and the right side of (3.5) is hence expressed entirely in terms of elementary functions.

It is perhaps worth noting that when $m = k = 2$ a complete asymptotic series is available, namely

$${}_0F_0^{(2)}(-\tfrac{1}{2}nL, A) \sim \exp(-\tfrac{1}{2}n \textstyle\sum_{i=1}^{2} l_i\alpha_i) \left(\dfrac{2}{nc_{12}\pi}\right)^{\frac{1}{2}} {}_2F_0\left(\tfrac{1}{2}, \tfrac{1}{2}; \dfrac{2}{nc_{12}}\right).$$

This follows directly from a representation of ${}_0F_0^{(2)}$ in terms of a classical confluent hypergeometric function (see Muirhead (1975)) and has been obtained by G. A. Anderson (1965) and Bingham (1972).

Asymptotic results such as those described above have proved useful in inference problems. For example, when all the population roots $\alpha_1, \cdots, \alpha_m$ are well spaced and n is large the likelihood function (3.2) can be approximated as

$$(3.6) \qquad \prod_{i=1}^m \alpha_i^{\frac{1}{2}n} {}_0F_0^{(m)}(-\tfrac{1}{2}nL, A)$$
$$\approx K \cdot (\prod_{i=1}^m \alpha_i^{\frac{1}{2}n}) \exp(-\tfrac{1}{2}n \sum_{i=1}^m l_i\alpha_i) \prod_{i<j}^m (\alpha_j - \alpha_i)^{-\frac{1}{2}}$$

where K is a constant (depending on n, $l_1, \cdots, l_m$ but not on $\alpha_1, \cdots, \alpha_m$ and hence irrelevant for likelihood purposes). This approximation has been investigated numerically by G. A. Anderson (1965) and Bingham (1972). As noted by G. A. Anderson (1965) and James (1966), the asymptotic likelihood is a product of likelihoods of independent variance estimates

$$\alpha_i^{\frac{1}{2}n} \exp(-\tfrac{1}{2}nl_i\alpha_i)$$

multiplied by linkage factors

$$(\alpha_j - \alpha_i)^{-\frac{1}{2}}$$

which show the dependence of the likelihood function on the effects of interactions between roots. It is well-known (see e.g., Lawley (1956)) that the sample roots l_i are biased estimates of the corresponding population roots α_i^{-1}, the bias term being of order n^{-1}. G. A. Anderson (1965) showed that, in estimating the population roots, a correction for bias is obtained by considering maximum marginal likelihood estimates (i.e., the values of the α_i^{-1} which maximize the right side of (3.6)), namely

$$\hat{\alpha}_i^{-1} = l_i - \frac{1}{n} l_i \sum_{j\neq i} \frac{l_j}{l_i - l_j} + O(n^{-2}) .$$

These estimates utilize information from other sample roots, adjacent ones of course having the most effect, and their bias terms are of order n^{-2}. An illuminating discussion on the estimation of latent roots has been provided by Dempster (1966).

Substitution of the asymptotic expansion (3.5) for ${}_0F_0^{(m)}$ in (3.1) yields an expansion for the density function of $l_1, \cdots, l_m$ when the q-smallest roots of Σ satisfy assumption (3.4). This expansion can be written, to order n^{-1}, as

$$k_1 \prod_{i=1}^k l_i^{\frac{1}{2}(n-m-1)} \exp(-\tfrac{1}{2}n \sum_{i=1}^k l_i\alpha_i) \prod_{i<j}^k \left(\frac{l_i - l_j}{\alpha_j - \alpha_i}\right)^{\frac{1}{2}} \left[1 + \frac{1}{2n} \sum_{i<j}^k c_{ij}^{-1}\right]$$
$$(3.7) \qquad \times \prod_{i=1}^k \prod_{j=k+1}^m \left(\frac{l_i - l_j}{\alpha_j - \alpha_i}\right)^{\frac{1}{2}} \left[1 + \frac{1}{2n} \sum_{i=1}^k \sum_{j=k+1}^m c_{ij}^{-1}\right]$$
$$\times \prod_{i=k+1}^m l_i^{\frac{1}{2}(n-m-1)} \prod_{k+1;i<j}^m (l_i - l_j) {}_0F_0^{(m-k)}(-\tfrac{1}{2}nL_2, A_2)$$

where

$$k_1 = \frac{(\frac{1}{2}n)^{\frac{1}{2}mn-\frac{1}{4}k(2m-k-1)}\pi^{\frac{1}{2}m^2-\frac{1}{4}k(k+1)}\Gamma_k(\frac{1}{2}m)}{\Gamma_m(\frac{1}{2}n)\Gamma_m(\frac{1}{2}m)} \prod_{i=1}^m \alpha_i^{\frac{1}{2}n} .$$

The last two lines in (3.7), multiplied by an appropriate constant, give an expansion to order n^{-1} of the conditional density function of the smallest $q = m - k$ roots $l_{k+1}, \cdots, l_m$ given the largest k roots $l_1, \cdots, l_k$. From (3.7) two points can be noted:

(i) In the asymptotic conditional distribution of the last q roots l_j the influence of the first k roots l_i is approximately via linkage factors of the form $(l_i - l_j)^{\frac{1}{2}}$. If, as a first approximation these factors are ignored with $l_{k+1}, \cdots, l_m$ regarded as satisfying $\infty > l_{k+1} > \cdots > l_m > 0$, then this asymptotic distribution in the case when $A_2 = \alpha I_q$ would just be the distribution of the roots of a $q \times q$ sample covariance matrix with n replaced by $n - k$, i.e., one degree of freedom is lost for each conditioned root.

(ii) The asymptotic conditional distribution of the last q roots given the first k does not depend on $\alpha_1, \cdots, \alpha_k$ (the first k sample roots $l_1, \cdots, l_k$ are asymptotically sufficient for the corresponding population roots $\alpha_1, \cdots, \alpha_k$). In a test of the null hypothesis that the smallest q roots of Σ are all equal, namely

$$H_q: \alpha_{k+1} = \cdots = \alpha_m \quad (= \alpha, \text{ say}),$$

$\alpha_1, \cdots, \alpha_k$ are nuisance parameters and James (1969) has suggested that the effects of these should be eliminated by drawing inferences about the smallest roots using this asymptotic conditional distribution.

The likelihood ratio statistic for testing the hypothesis H_q is (T. W. Anderson (1963))

$$V_q = \prod_{i=k+1}^m (l_i/\bar{l}_q)$$

where $\bar{l}_q = (1/q) \sum_{i=k+1}^m l_i$, the average of the last q roots and, when H_q is true (i.e., $A_2 = \alpha I_q$), the asymptotic conditional distribution of $l_{k+1}, \cdots, l_m$ given $l_1, \cdots, l_k$ reduces to

$$\text{(3.8)} \qquad \text{const.} \prod_{i=1}^k \prod_{j=k+1}^m (l_i - l_j)^{\frac{1}{2}} \prod_{i=k+1}^m l_i^{\frac{1}{2}(n-k-q-1)} \times \exp(-\tfrac{1}{2}n\alpha \textstyle\sum_{i=k+1}^m l_i) \prod_{k+1;i<j}^m (l_i - l_j).$$

The hypothesis $H_m: \alpha_1 = \cdots = \alpha_m$ (the "sphericity test") and its associated likelihood ratio statistic V_m have been widely studied. It is shown in Anderson (1958, page 263) that the statistic

$$T_m = -\left(n - \frac{2m^2 + m + 2}{6m}\right) \log V_m$$

has an asymptotic χ^2 distribution with $\frac{1}{2}(m + 2)(m - 1)$ degrees of freedom when H_m is true and that

$$E(T_m) = \tfrac{1}{2}(m + 2)(m - 1) + O(n^{-2}).$$

From testing H_q Bartlett (1954) suggested the statistic

$$-\left(n - k - \frac{2q^2 + q + 2}{6q}\right) \log V_q$$

which has an asymptotic χ^2 distribution with $\frac{1}{2}(q + 2)(q - 1)$ degrees of freedom when H_q is true. This statistic is suggested by the null conditional distribution (3.8) when the linkage factors

$$\prod_{i=1}^{k} \prod_{j=k+1}^{m} (l_i - l_j)^{\frac{1}{2}}$$

are ignored. A further refinement which takes these factors into account was obtained by Lawley (1956) and James (1969). The method of James also provides some information on the accuracy of the approximation and is based directly on the null conditional distribution (3.8). James shows that when H_q is true the statistic

$$T_q = -\left(n - k - \frac{2q^2 + q + 2}{6q} + \sum_{i=1}^{k} \frac{\bar{l}_q^2}{(l_i - \bar{l}_q)^2}\right) \log V_q$$

has an asymptotic χ^2 distribution with $\frac{1}{2}(q+2)(q-1)$ degrees of freedom and that

$$E(T_q) = \tfrac{1}{2}(q+2)(q-1) + O(n^{-2}),$$

the expectation being taken with respect to the distribution (3.8).

In connection with the likelihood ratio test of H_q there are (at least) two problem areas which need further investigation.

(a) *How good is the asymptotic χ^2 approximation?* If n is small or moderate the χ^2 approximation to the distribution of T_q may not be accurate enough for practical purposes. Little appears known about the accuracy of the approximation. The case $q = m$ has been studied and correction terms are available which are useful for small or moderate sample sizes. For example, it is shown in Anderson (1958, page 263) that, when H_m is true,

$$P(T_m < x) = P(\chi_f^2 < x) + \frac{(m^2 - 4)(m - 1)(2m^3 + 6m^2 + 3m + 2)}{288m^2n^2\rho^2} \times [P(\chi_{f+4}^2 < x) - P(\chi_f^2 < x)] + O(n^{-3})$$

where $f = \frac{1}{2}(m+2)(m-1)$ and

$$\rho = 1 - \frac{2m^2 + m + 2}{6mn}.$$

The correction term, which is of order n^{-2}, can be used to assess the accuracy of the asymptotic χ^2 distribution, and to correct it if necessary. It would be of interest to know how such correction terms should be modified when dealing with the null distribution of T_q for $q \neq m$. Some recent work by Fujikoshi (1976c) is related to this problem although the statistic under investigation is not precisely T_q.

(b) *How sensitive is the test T_q?* Very little appears known about the power of the test of H_q based on the likelihood ratio statistic T_q, except when $q = m$. The asymptotic nonnull distribution of T_m depends in a fundamental way on the alternatives being considered. For example, under the alternative $\Sigma \neq \alpha^{-1}I_m$, Sugiura (1969) has shown that

$$\frac{T_m - n\log\left[\left(\frac{1}{m}\operatorname{tr}\Sigma\right)^m \Big/ |\Sigma|\right]}{\left\{2mn\left[\frac{m\operatorname{tr}\Sigma^2}{(\operatorname{tr}\Sigma)^2} - 1\right]\right\}^{\frac{1}{2}}}$$

is asymptotically standard normal, and has obtained an asymptotic expansion up to terms of order $n^{-\frac{3}{2}}$ for its distribution function in terms of the standard normal distribution and density functions. This asymptotic result breaks down (or "blows up") for alternatives that are too close to the null hypothesis and various authors have obtained asymptotic nonnull distributions of T_m for sequences of alternative hypotheses approaching the null hypothesis. For example, under the sequence of alternatives

$$\Sigma = \alpha^{-1}(I_m + n^{-\frac{1}{2}}\Omega)$$

whose Ω is a fixed matrix, Nagao (1970) has shown that the asymptotic distribution of T_m is noncentral χ^2 on $\frac{1}{2}(m+2)(m-1)$ degrees of freedom with noncentrality parameter $\frac{1}{4}(\operatorname{tr}\Omega^2 - m^{-1}(\operatorname{tr}\Omega)^2)$ and has also obtained the term of order $n^{-\frac{1}{2}}$ in an asymptotic expansion for the distribution function of T_m. By changing the sequence of alternatives it is possible to go from a limiting noncentral χ^2 distribution to a central χ^2 distribution. Under the sequence of alternatives

$$\Sigma = \alpha^{-1}(I_m + n^{-1}\Omega),$$

considered by Nagao (1970) and Khatri and Srivastava (1974), the asymptotic distribution of T_m is (central) χ^2 on $\frac{1}{2}(m+2)(m-1)$ degrees of freedom, with Ω entering only in later correction terms.

Very little attention appears to have been paid to the numerical evaluation and comparison of these (and other) asymptotic results. Problems of some interest are: How close can one get to the null hypothesis before Sugiura's normal approximation becomes too inaccurate for power calculations? For alternatives close to the null hypothesis, which approximation is more accurate, the central or noncentral χ^2?

The entire area dealing with the nonnull distribution of T_q, for $q \neq m$, is open to study. One might suppose that the asymptotic distributions are similar, with appropriate modifications in means, variances, degrees of freedom, noncentrality parameters, and so on, but this remains a conjecture. It certainly appears that the asymptotic distribution (3.7) should be useful for investigating the nonnull distribution of T_q under the sequence of alternatives

$$A_2 = \alpha^{-1}I_{m-k} + n^{-1}\Omega$$

for fixed (diagonal) Ω.

Before moving on to discuss other types of asymptotic distributions for the sample roots it is worth noting that an asymptotic representation for the marginal distribution of l_i, valid for $l_i \in (\lambda_{i+1}, \lambda_{i-1})$, can be obtained easily (from (3.7), for example) when α_i is a simple population root as

$$(3.9) \qquad \tfrac{1}{2}\left(\frac{n}{\pi}\right)^{\frac{1}{2}} e^{\frac{1}{2}n} \lambda_i^{-\frac{1}{2}(n-m+1)} \exp(-nl_i/2\lambda_i) l_i^{\frac{1}{2}(n-m-1)} \prod_{j=1}^{i-1} \left(\frac{\lambda_j - l_i}{\lambda_j - \lambda_i}\right)^{\frac{1}{2}} \times \prod_{j=i+1}^{m} \left(\frac{l_i - \lambda_j}{\lambda_i - \lambda_j}\right)^{\frac{1}{2}}$$

where $\lambda_i = \alpha_i^{-1}$, $i = 1, \cdots, m$, are the latent roots of Σ. If the term

$$\prod_{j=1}^{i-1} (\lambda_j - l_i)^{\frac{1}{2}} \prod_{j=i+1}^{m} (l_i - \lambda_j)^{\frac{1}{2}}$$

is neglected, (3.9) suggests that nl_i/λ_i is approximately distributed as χ^2 with n degrees of freedom (G. A. Anderson (1965)).

The asymptotic distributions discussed above in (3.7) and (3.8) resemble products of gamma distributions linked by factors of the form $l_i - l_j$. By making a suitable transformation it is possible to obtain a more highly asymptotic "normal type" of distribution which no longer preserves linkage factors corresponding to distinct population roots. Putting

$$x_i = \left(\frac{n}{2}\right)^{\frac{1}{2}} \left(\frac{l_i - \lambda_i}{\lambda_i}\right) \qquad i = 1, \cdots, m,$$

Girshick (1939) showed, using the asymptotic theory of maximum likelihood estimates, that if λ_i is simple then x_i is asymptotically independent of x_j for $j \neq i$ and the limiting distribution of x_i is standard normal. When Σ has multiple latent roots the theory is rather more complicated; the definitive paper in this area is due to T. W. Anderson (1963) who investigated the asymptotic distributions of the roots and vectors of S in some detail, together with a large number of inference problems in principal component analysis. Suppose that Σ has r different roots with multiplicities $q_1, \cdots, q_r$ ($\sum_{\alpha=1}^r q_\alpha = m$), namely

$$\begin{aligned} \lambda_1 &= \cdots = \lambda_{q_1} &= \delta_1 \\ \lambda_{q_1+1} &= \cdots = \lambda_{q_1+q_2} &= \delta_2 \\ &\vdots \\ \lambda_{m-q_r+1} &= \cdots = \lambda_m &= \delta_r, \end{aligned}$$

with $\delta_1 > \delta_2 > \cdots > \delta_r > 0$, and put

$$(3.10) \qquad x_i = \left(\frac{n}{2}\right)^{\frac{1}{2}} \left(\frac{l_i - \delta_\alpha}{\delta_\alpha}\right) \quad \text{for} \quad i \in J_\alpha \qquad \alpha = 1, \cdots, r,$$

where J_α denotes the set of integers

$$\sum_{i=1}^{\alpha-1} q_\alpha + 1, \cdots, \sum_{i=1}^{\alpha} q_\alpha \quad q_0 \equiv 0.$$

Using the fact that $n^{\frac{1}{2}}(S - \Sigma)$ is asymptotically normal T. W. Anderson (1963) showed that the x_i's corresponding to different roots of Σ are asymptotically independent and that the x_i's corresponding to the same multiple root of Σ are, asymptotically, the roots of a symmetric matrix whose elements have limiting independent normal distributions. The limiting distribution of the x_i's can then be obtained from the asymptotic distribution of this matrix. An extension of Anderson's argument was used by Fujikoshi (1976a) to obtain an asymptotic expansion for the joint distribution of $x_1, \cdots, x_m$ up to terms of order n^{-1}, namely

$$\prod_{\alpha=1}^{r} \left\{ \frac{\pi^{\frac{1}{4}q_\alpha(q_\alpha-1)}}{2^{\frac{1}{2}q_\alpha}\Gamma_{q_\alpha}(\frac{1}{2}q_\alpha)} \exp(-\tfrac{1}{4}\textstyle\sum_{i\in J_\alpha} x_i^2) \prod_{i<j;i,j\in J_\alpha}(x_i - x_j)\right\} \tag{3.11}$$
$$\times \{1 + n^{-\frac{1}{2}}Q_1 + n^{-1}Q_2 + O(n^{-\frac{3}{2}})\}$$

where

$$Q_1 = 2^{\frac{1}{2}} \left\{ \textstyle\sum_{i=1}^m (\frac{1}{3}x_i^3 - \frac{1}{2}(m+1)x_i) + \frac{1}{2} \sum_{\alpha<\beta} \sum_{i\in J_\alpha} \sum_{j\in J_\beta} \left(\frac{\delta_\alpha x_i - \delta_\beta x_j}{\delta_\alpha - \delta_\beta} \right) \right\}$$

and

$$Q_2 = \tfrac{1}{2}Q_1^2 - \tfrac{1}{24}m(2m^2 + 3m - 1) - \textstyle\sum_{i=1}^m (\frac{1}{2}x_i^4 - \frac{1}{2}(m+1)x_i^2)$$
$$- \tfrac{1}{2} \textstyle\sum_{\alpha<\beta} \sum_{i\in J_\alpha} \sum_{j\in J_\beta} \frac{(\delta_\alpha x_i - \delta_\beta x_j)^2 - \delta_\alpha\delta_\beta}{(\delta_\alpha - \delta_\beta)^2} .$$

This expansion can also be obtained by making the change of variables (3.10) in expansions for the joint distribution of $l_1, \cdots, l_m$ of the type discussed previously, obtained by expanding the ${}_0F_0^{(m)}$ function. For example, (3.11) has also been derived by Sugiura (1976a) from an expansion for the ${}_0F_0^{(m)}$ function in the multiple root case given by Chattopadhyay and Pillai (1973). The result (3.11) encompasses all cases of interest; when $r = 1$, $q_1 = m$ (all roots of Σ are equal), the limiting term in (3.11) has been given by T. W. Anderson (1963); when $r = m$ (each root is simple), the limiting term is a product of standard normal density functions and the correction terms Q_1 and Q_2 were also derived in this case by Muirhead and Chikuse (1975). An expansion for the marginal distribution of x_i when λ_i is simple is then easily obtained and has been given by Muirhead and Chikuse (1975) and Sugiura (1973). This expansion has been examined numerically for the largest root x_1 by Muirhead (1974).

4. Latent roots when $\Sigma_1 \neq \Sigma_2$. Suppose that the $m \times m$ matrix variates S_1 and S_2 have independent Wishart distributions $W_m(n_1, \Sigma_1)$ and $W_m(n_2, \Sigma_2)$ respectively and let $l_1, \cdots, l_m$ be the latent roots of $S_1S_2^{-1}$. Various functions of $l_1, \cdots, l_m$ have been proposed as statistics suitable for testing the null hypothesis $\Sigma_1 = \Sigma_2$. The exact joint density function of these roots is (see James (1964))

$$\frac{\pi^{\frac{1}{2}m^2}\Gamma_m(\frac{1}{2}n)}{\Gamma_m(\frac{1}{2}n_1)\Gamma_m(\frac{1}{2}n_2)\Gamma_m(\frac{1}{2}m)} \prod_{i=1}^m \alpha_i^{\frac{1}{2}n_1} \prod_{i=1}^m l_i^{\frac{1}{2}(n_1-m-1)} \prod_{i<j}^m (l_i - l_j) \tag{4.1}$$
$$\times\ {}_1F_0^{(m)}(\tfrac{1}{2}n; -L, A) \qquad l_1 > l_2 > \cdots > l_m > 0$$

where $n = n_1 + n_2$, $0 < \alpha_1 \leqq \alpha_2 \leqq \cdots \leqq \alpha_m$ are the latent roots of $(\Sigma_1\Sigma_2^{-1})^{-1}$, $L = \operatorname{diag}(l_1, \cdots, l_m)$, $A = \operatorname{diag}(\alpha_1, \cdots, \alpha_m)$. The marginal likelihood function of the population roots is thus

$$\prod_{i=1}^m \alpha_i^{\frac{1}{2}n_1}\ {}_1F_0^{(m)}(\tfrac{1}{2}n; -L, A) . \tag{4.2}$$

To approximate the ${}_1F_0^{(m)}$ function for large n Laplace's method is applied to the integral representation

$$ {}_1F_0^{(m)}(\tfrac{1}{2}n; -L, A) = \int_{O(m)} \det (I_m + LH'AH)^{-\frac{1}{2}n}(dH) . $$

The asymptotic behavior depends on the spread of the population roots $\alpha_1, \cdots, \alpha_m$. When these are all distinct Chang (1970) derived the limiting term and further terms in an asymptotic series were obtained by Li, Pillai and Chang (1970) and Constantine and Muirhead (1976). Asymptotic results in the case of one or more multiple population roots have been derived by Chattopadhyay and Pillai (1973), Sugiura (1976a), Chang (1973) and Li, Pillai and Chang (1970).

A situation which has not thus far been dealt with explicitly is when some of the population roots are close, parallelling the discussion in the previous section. Assume that $\alpha_1, \cdots, \alpha_m$ satisfy

$$ \alpha_1 < \cdots < \alpha_k < \alpha_{k+1} \leqq \alpha_{k+2} \leqq \cdots \leqq \alpha_m \tag{4.3} $$

with

$$ \alpha_i - \alpha_j = O(n^{-1}) \qquad \text{for} \quad i, j = k+1, \cdots, m . $$

The method used by Constantine and Muirhead (1976) to obtain the asymptotic expansion (3.5) for ${}_0F_0^{(m)}$ can be used here to show that

$$ {}_1F_0^{(m)}(\tfrac{1}{2}n; -L, A) \sim \frac{\Gamma_k(\frac{1}{2}m)}{\pi^{\frac{1}{2}mk}} \prod_{i=1}^{k} (1 + l_i\alpha_i)^{-\frac{1}{2}n} \prod_{i=1}^{k} \prod_{j=1;i<j}^{m} \left(\frac{2\pi}{nc_{ij}}\right)^{\frac{1}{2}} \times {}_1F_0^{(m-k)}(\tfrac{1}{2}n; -L_2, A_2) \left\{1 + \frac{1}{n} P_1 + O(n^{-2})\right\} \tag{4.4} $$

where

$$ c_{ij} = \frac{(l_i - l_j)(\alpha_j - \alpha_i)}{(1 + l_i\alpha_i)(1 + l_j\alpha_j)} , $$

$$ L_2 = \operatorname{diag}(l_{k+1}, \cdots, l_m) , \qquad A_2 = \operatorname{diag}(\alpha_{k+1}, \cdots, \alpha_m) , $$

$$ P_1 = \tfrac{1}{2} \sum_{i=1}^{k} \sum_{j=1;i<j}^{m} c_{ij}^{-1} + \frac{k}{24} [(k-1)(4k+1) + 6(m-k)(2m-1)] . $$

The zonal polynomial series for the function ${}_1F_0^{(m-k)}(\frac{1}{2}n; -L_2, A_2)$ could be used for computational purposes. When $k = m$ (all population roots well spaced) the ${}_1F_0^{(m-k)}$ function is taken to be unity, and when $\alpha_{k+1} = \cdots = \alpha_m = \alpha$ (i.e., $A_2 = \alpha I_{m-k}$),

$$ {}_1F_0^{(m-k)}(\tfrac{1}{2}n; -L_2, A_2) = \prod_{i=k+1}^{m} (1 + l_i\alpha)^{-\frac{1}{2}n} . $$

When all the population roots are well spaced and n is large the likelihood function (4.2) can be approximated as

$$ \prod_{i=1}^{m} \alpha_i^{\frac{1}{2}n_1} {}_1F_0^{(m)}(\tfrac{1}{2}n; -L, A) \approx K \cdot \prod_{i=1}^{m} \frac{\alpha_i^{\frac{1}{2}n_1}}{(1 + l_i\alpha_i)^{\frac{1}{2}(n-m+1)}} \prod_{i<j}^{m} (\alpha_j - \alpha_i)^{-\frac{1}{2}} \tag{4.5} $$

where K does not depend on $\alpha_1, \cdots, \alpha_m$. Essentially the asymptotic likelihood can be interpreted as a product of likelihoods of independent variance ratio estimates

$$\alpha_i^{\frac{1}{2}n_1}(1 + l_i\alpha_i)^{-\frac{1}{2}(n-m+1)}$$

multiplied by linkage factors $(\alpha_j - \alpha_i)^{-\frac{1}{2}}$. The maximum marginal likelihood estimates of $\alpha_1^{-1}, \cdots, \alpha_m^{-1}$, obtained by maximizing the right side of (4.5), are, with $n_1 = k_1 n$, $n_2 = k_2 n$, $k_1 + k_2 = 1$,

$$\hat{\alpha}_i^{-1} = \frac{k_2 l_i}{k_1} - \frac{(m+1)l_i}{nk_1} - \frac{1}{nk_1^2}\sum_{j\neq i}\frac{l_i l_j}{l_i - l_j} + O(n^{-2}) .$$

Comparing this with (Chikuse (1974))

$$E(l_i) = \frac{k_1}{k_2}\alpha_i^{-1} + \frac{k_1}{nk_2^2}(m+1)\alpha_i^{-1} + \frac{1}{nk_2^2}\sum_{j\neq i}\frac{\alpha_i^{-1}\alpha_j^{-1}}{\alpha_i^{-1} - \alpha_j^{-1}} + O(n^{-2}) ,$$

it is seen that $E(\hat{\alpha}_i^{-1}) = \alpha_i^{-1} + O(n^{-2})$ and hence the estimates $\hat{\alpha}_i^{-1}$ provide a correction for bias.

Substitution of (4.4) in (4.1) gives an asymptotic representation for the density function of $l_1, \cdots, l_m$ under the assumption (4.3). Ignoring terms of $O(n^{-1})$ this representation is

$$\begin{aligned} &k_2[\textstyle\prod_{i=1}^k l_i^{\frac{1}{2}(n_1-m-1)}(1 + l_i\alpha_i)^{-\frac{1}{2}(n-k+1)}] \prod_{i<j}^k \left(\frac{l_i - l_j}{\alpha_j - \alpha_i}\right)^{\frac{1}{2}} \\ (4.6)\qquad &\times \textstyle\prod_{i=1}^k \prod_{j=k+1}^m \left(\frac{l_i - l_j}{\alpha_j - \alpha_i}\right)^{\frac{1}{2}} \prod_{i=k+1}^m [l_i^{\frac{1}{2}(n_1-m-1)}(1 + l_i\alpha_i)^{\frac{1}{2}k}] \\ &\times \textstyle\prod_{k+1;i<j}^m (l_i - l_j)\, {}_1F_0^{(m-k)}(\frac{1}{2}n; -L_2, A_2) \end{aligned}$$

where

$$k_2 = \pi^{\frac{1}{2}m^2 - \frac{1}{4}k(k+1)}\frac{\Gamma_m(\frac{1}{2}n)\Gamma_k(\frac{1}{2}m)(\frac{1}{2}n)^{-\frac{1}{4}k(2m-k-1)}}{\Gamma_m(\frac{1}{2}n_1)\Gamma_m(\frac{1}{2}n_2)\Gamma_m(\frac{1}{2}m)}\prod_{i=1}^m \alpha_i^{\frac{1}{2}n_1} .$$

The last two lines in (4.6), multiplied by an appropriate constant, give an asymptotic representation for the conditional density function of the smallest $q = m - k$ roots $l_{k+1}, \cdots, l_m$ given the largest k roots $l_1, \cdots, l_k$. In a test of equality of $\alpha_{k+1}^{-1}, \cdots, \alpha_m^{-1}$, the q-smallest roots of $\Sigma_1\Sigma_2^{-1}$, the effects of the nuisance parameters $\alpha_1, \cdots, \alpha_k$ can be eliminated asymptotically by basing inferences on the conditional distribution of $l_{k+1}, \cdots, l_m$ given $l_1, \cdots, l_k$. When the null hypothesis $H_q : \alpha_{k+1} = \cdots = \alpha_m\ (= \alpha)$ is true, it is seen from (4.6) that this conditional distribution has the asymptotic representation

$$\begin{aligned} \text{const.} &\textstyle\prod_{i=1}^k \prod_{j=k+1}^m (l_i - l_j)^{\frac{1}{2}} \prod_{i=k+1}^m [l_i^{\frac{1}{2}(n_1-k-q-1)}(1 + \alpha l_i)^{-\frac{1}{2}(n_1+n_2-k)}] \\ &\times \textstyle\prod_{k+1;i<j}^m (l_i - l_j) \end{aligned}$$

which does not depend on $\alpha_1, \cdots, \alpha_k$. If the linkage factors

$$\prod_{i=1}^k \prod_{j=k+1}^m (l_i - l_j)^{\frac{1}{2}}$$

are ignored then this asymptotic distribution is just the distribution of the roots of $S_1 S_2^{-1}$ where S_1 and S_2 have independent Wishart distributions $W_q(n_1 - k, \Sigma_1)$ and $W_q(n_2, \alpha\Sigma_1)$ respectively.

The asymptotic distributions obtained by approximating the ${}_1F_0^{(m)}$ function resemble products of F-distributions linked by factors of the form $l_i - l_j$. By a suitable standardization of the roots asymptotic distributions can be obtained which are of a normal type. Assume that $\Sigma_1\Sigma_2^{-1}$ has r different roots with multiplicities $q_1, \cdots, q_r$ ($\sum_{\alpha=1}^r q_\alpha = m$), namely

$$\begin{aligned} \lambda_1 &= \cdots = \lambda_{q_1} &&= \delta_1 \\ \lambda_{q_1+1} &= \cdots = \lambda_{q_1+q_2} &&= \delta_2 \\ &\vdots \\ \lambda_{m-q_r+1} &= \cdots = \lambda_m &&= \delta_r \end{aligned}$$

with $\delta_1 > \delta_2 > \cdots > \delta_r > 0$. Write $n_1 = k_1 n$, $n_2 = k_2 n$ where $k_1 + k_2 = 1$ and put

$$x_i = \left(\frac{k_1 k_2 n}{2}\right)^{\frac{1}{2}} \left(\frac{k_2 l_i}{k_1 \delta_\alpha} - 1\right) \qquad \text{for} \quad i \in J_\alpha \qquad \alpha = 1, \cdots, r$$

where J_α is the set of integers

$$\textstyle\sum_{i=1}^{\alpha-1} q_i + 1, \cdots, \sum_{i=1}^{\alpha} q_i \,.$$

Making this transformation in an expansion of the density function of $l_1, \cdots, l_m$ given by Chattopadhyay and Pillai (1973) in the multiple root case Sugiura (1976a) obtained an expansion for the joint density function of $x_1, \cdots, x_m$ as

$$\begin{aligned} &\textstyle\prod_{\alpha=1}^r \left\{ \dfrac{\pi^{\frac{1}{4}q_\alpha(q_\alpha-1)}}{\alpha^{\frac{1}{2}q_\alpha}\Gamma_{q_\alpha}(\frac{1}{2}q_\alpha)} \exp(-\frac{1}{2} \sum_{i\in J_\alpha} x_i^2) \prod_{i<j;i,j\in J_\alpha} (x_i - x_j) \right\} \\ &\qquad \times \{1 + n^{-\frac{1}{2}} Q_1 + O(n^{-1})\} \end{aligned}$$

where

$$\begin{aligned} Q_1 = \left(\frac{2}{k_1 k_2}\right)^{\frac{1}{2}} &\Big\{ \tfrac{1}{3}(1 + k_1) \textstyle\sum_{i=1}^m x_i^3 + [\frac{1}{2}k_1(m-1) - \frac{1}{2}(m+1)] \sum_{i=1}^m x_i \\ &+ \tfrac{1}{2} \textstyle\sum_{\alpha<\beta} \sum_{i\in J_\alpha} \sum_{j\in J_\beta} \dfrac{\delta_\alpha x_i - \delta_\beta x_j}{\delta_\alpha - \delta_\beta} \Big\} \,. \end{aligned}$$

The term of order n^{-1} has also been given by Sugiura. In the special case $r = k + 1$, $q_\alpha = 1$, $\alpha = 1, \cdots, k$, $q_{k+1} = m - k$ (i.e., all roots simple except the smallest, which has multiplicity $m - k$), this expansion was also obtained by Chikuse (1974); when $r = m$ (each root is simple) the limiting term is a product of standard normal density functions.

5. Noncentral Wishart distribution. Let X be an $m \times n$ matrix variate whose columns are independently normally distributed with common covariance matrix Σ and $E(X) = M$; the density function of XX' is (see James (1964))

$$\text{(5.1)} \qquad \begin{aligned} &\frac{\det \Sigma^{-\frac{1}{2}n}}{\Gamma_m(\frac{1}{2}n)2^{\frac{1}{2}mn}} \exp(-\tfrac{1}{2} \operatorname{tr} \Sigma^{-1}XX') \det (XX')^{\frac{1}{2}(n-m-1)} \\ &\qquad \times \exp(-\tfrac{1}{2} \operatorname{tr} \Sigma^{-1}MM')\,{}_0F_1(\tfrac{1}{2}n; \tfrac{1}{4}X'\Sigma^{-1}MM'\Sigma^{-1}X) \,. \end{aligned}$$

The ${}_0F_1$ function in (5.1) is a symmetric function of the nonzero latent roots of $A'A$ where $A = M'\Sigma^{-1}X$, having the integral representation (James (1964), Herz (1955))

$$(5.2) \qquad {}_0F_1(\tfrac{1}{2}n; \tfrac{1}{4}A'A) = \int_{O(n)} \exp(\operatorname{tr} AH)(dH) .$$

If $A'A$ has rank r it can be assumed without loss of generality that $A = \operatorname{diag}(a_1, \cdots, a_r, 0, \cdots, 0)$ where $a_1 \geqq a_2 \geqq \cdots \geqq a_r > 0$ with $a_1^2, \cdots, a_r^2$ being the nonzero latent roots of $A'A$. When these roots are large, a situation which would arise for example if M has the form

$$M = \begin{bmatrix} M_1 \\ \cdots \\ 0 \end{bmatrix},$$

with the mean vectors in M_1 being large, the zonal polynomial series for ${}_0F_1$ converges slowly. Laplace's method can be applied to the integral (5.2) in order to derive its asymptotic behavior for large A (i.e., large $a_1, \cdots, a_r$) and partial differential equations satisfied by ${}_0F_1$ can be used to obtain further terms in an asymptotic series. The result is

$$\begin{aligned} {}_0F_1(\tfrac{1}{2}n; \tfrac{1}{4}A'A) \sim \frac{\Gamma_r(\frac{1}{2}n)}{2^r\pi^{\frac{1}{2}rn}} \exp(\textstyle\sum_{i=1}^r a_i) \prod_{i=1}^r \left(\frac{2\pi}{a_i}\right)^{\frac{1}{2}(n-r)} \prod_{i<j}^r \left(\frac{2\pi}{a_i + a_j}\right)^{\frac{1}{2}} \\ \times [1 + P_1 + O(A^{-2})] \end{aligned}$$

where the terms of order A^{-1} (i.e., a_i^{-1}, $i = 1, \cdots, r$) are given by

$$P_1 = \tfrac{1}{8} \textstyle\sum_{i<j}^r (a_i + a_j)^{-1} - \tfrac{1}{8}(n - r)(n - r - 2) \sum_{i=1}^r a_i^{-1} .$$

The asymptotic behavior of ${}_0F_1$ was derived when $r = n$ by G. A. Anderson (1970). For $r \neq n$ the above asymptotic behavior was conjectured by Anderson and verified by the author. Anderson also gives the term of order A^{-2}.

6. Noncentral means with known covariance. Let X be an $m \times n$ matrix variate whose columns are independently normally distributed with common covariance matrix Σ and $E(X) = M$. The joint density function of $w_1, \cdots, w_m$, the latent roots of $\Sigma^{-1}XX'$, is (James (1964))

$$(6.1) \qquad \begin{aligned} &\frac{\pi^{\frac{1}{2}m^2}}{2^{\frac{1}{2}mn}\Gamma_m(\frac{1}{2}n)\Gamma_m(\frac{1}{2}m)} \exp(-\tfrac{1}{2} \textstyle\sum_{i=1}^m w_i) \prod_{i=1}^m w_i^{\frac{1}{2}(n-m-1)} \prod_{i<j}^m (w_i - w_j) \\ &\times \exp(-\tfrac{1}{2} \textstyle\sum_{i=1}^m \omega_i)\, {}_0F_1^{(m)}(\tfrac{1}{2}n; \tfrac{1}{4}\Omega, W) \qquad w_1 > w_2 > \cdots > w_m > 0 \end{aligned}$$

where $\omega_1 \geqq \omega_2 \geqq \cdots \geqq \omega_m \geqq 0$ are the latent roots of $\Sigma^{-1}MM'$, $\Omega = \operatorname{diag}(\omega_1, \cdots, \omega_m)$, $W = \operatorname{diag}(w_1, \cdots, w_m)$. The marginal likelihood function of the noncentrality parameters $\omega_1, \cdots, \omega_m$ is then

$$(6.2) \qquad \exp(-\tfrac{1}{2} \textstyle\sum_{i=1}^m \omega_i)\, {}_0F_1^{(m)}(\tfrac{1}{2}n; \tfrac{1}{4}\Omega, W) .$$

For large Ω the asymptotic behavior of the ${}_0F_1^{(m)}$ function can be obtained from the integral representation

$${}_0F_1^{(m)}(\tfrac{1}{2}n; \tfrac{1}{4}\Omega, W) = \int_{O(m)} {}_0F_1(\tfrac{1}{2}n; \tfrac{1}{4}\Omega H'WH)(dH)$$

and the asymptotic behavior of the ${}_0F_1$ function in the integrand given in the previous section. Under the assumption that

$$(6.3) \qquad \omega_1 > \cdots > \omega_k > \omega_{k+1} = \cdots = \omega_m = 0$$

where $\omega_1, \cdots, \omega_k$ are large, Leach (1969) has shown that

$$(6.4) \qquad {}_0F_1^{(m)}(\tfrac{1}{2}n; \tfrac{1}{4}\Omega, W) \sim 2^{\frac{1}{2}k(m+n-k-3)}\pi^{-\frac{1}{2}k(k+1)}\Gamma_k(\tfrac{1}{2}n)\Gamma_k(\tfrac{1}{2}m)\exp[\textstyle\sum_{i=1}^k (w_i\omega_i)^{\frac{1}{2}}] \times \prod_{i=1}^k (\omega_i w_i)^{\frac{1}{4}(m-n)} \prod_{i=1}^k \prod_{j=1;i<j}^m c_{ij}^{-\frac{1}{2}}$$

where

$$c_{ij} = (\omega_i - \omega_j)(w_i - w_j) \qquad i, j = 1, \cdots, k$$
$$= \omega_i(w_i - w_j) \qquad i = 1, \cdots, k; \quad j = k+1, \cdots, m.$$

The asymptotic behavior when $\omega_{k+1}, \cdots, \omega_m$ are small compared with $\omega_1, \cdots, \omega_k$ but nonzero is not known. When $k = m$ (i.e., all noncentrality parameters large and distinct) further terms in an asymptotic series can be obtained using a partial differential equation for ${}_0F_1^{(m)}$ given in Constantine and Muirhead (1972). To order $\Omega^{-\frac{1}{2}}$ the expansion is

$${}_0F_1^{(m)}(\tfrac{1}{2}n; \tfrac{1}{4}\Omega, W) \sim G_0[1 + P_1 + O(\Omega^{-1})]$$

where G_0 denotes the right side of (6.4) with $k = m$ and

$$P_1 = \tfrac{1}{2}\textstyle\sum_{i<j}^m \dfrac{(\omega_i w_i)^{\frac{1}{2}} + (\omega_j w_j)^{\frac{1}{2}}}{(\omega_i - \omega_j)(w_i - w_j)} - \tfrac{1}{8}(n-m)(n-m-2)\textstyle\sum_{i=1}^m (\omega_i w_i)^{-\frac{1}{2}}.$$

Substitution of (6.4) in (6.1) gives an asymptotic representation for the density function of $w_1, \cdots, w_m$ under the assumption (6.3) with $\omega_1, \cdots, \omega_k$ large. Ignoring terms of order $\omega_i^{-\frac{1}{2}}$, $i = 1, \cdots, k$, this representation is

$$(6.5) \qquad k_3 \exp[-\tfrac{1}{2}\textstyle\sum_{i=1}^k w_i + \sum_{i=1}^k (\omega_i w_i)^{\frac{1}{2}}] \prod_{i=1}^k w_i^{\frac{1}{4}(n-m-2)} \prod_{i<j}^k \left(\dfrac{w_i - w_j}{\omega_i - \omega_j}\right)^{\frac{1}{2}} \times \prod_{i=1}^k \prod_{j=k+1}^m (w_i - w_j)^{\frac{1}{2}} \exp(-\tfrac{1}{2}\textstyle\sum_{i=k+1}^m w_i) \prod_{i=k+1}^m w_i^{\frac{1}{2}(n-m-1)} \times \prod_{i<j;k+1}^m (w_i - w_j)$$

where

$$k_3 = \frac{\pi^{\frac{1}{2}m^2 - \frac{1}{2}k(k+1)}\Gamma_k(\frac{1}{2}n)\Gamma_k(\frac{1}{2}m)}{\Gamma_m(\frac{1}{2}n)\Gamma_m(\frac{1}{2}m)2^{\frac{1}{2}mn - \frac{1}{2}k(m+n-k-3)}} \exp(-\tfrac{1}{2}\textstyle\sum_{i=1}^m \omega_i) \prod_{i=1}^k \omega_i^{-\frac{1}{4}(m+n-2k)}.$$

The last line in (6.5), multiplied by an appropriate constant, is an asymptotic representation for the conditional density function of the $q = m - k$ smallest roots $w_{k+1}, \cdots, w_m$ given the k largest roots $w_1, \cdots, w_k$ when the null hypothesis H_q: $\omega_{k+1} = \cdots = \omega_m = 0$ is true, with $\omega_1, \cdots, \omega_k$ large. Note that it does not depend on $\omega_1, \cdots, \omega_k$, the nuisance parameters in a test of H_q (cf. Section 3). If the linkage factors

$$\prod_{i=1}^k \prod_{j=k+1}^m (w_i - w_j)^{\frac{1}{2}}$$

are ignored, then the asymptotic conditional distribution is just the distribution of the roots of a matrix having the $W_q(n - k, I_q)$ distribution. This has been exploited by Leach (1969) in a study of the likelihood ratio test of the rank of the matrix of means M.

7. Noncentral multivariate F. Let X and Y be $m \times n_1$ and $m \times n_2$ matrix variates respectively, $n_1 \leqq m \leqq n_2$, with columns all independently normally distributed with covariance matrix Σ and with $E(X) = M$, $E(Y) = 0$. The density function of the $n_1 \times n_1$ matrix variate $F = X'(YY')^{-1}X$ is (James (1964))

$$(7.1) \qquad \frac{\Gamma_{n_1}(\frac{1}{2}(n_1+n_2))}{\Gamma_{n_1}(\frac{1}{2}m)\Gamma_{n_1}(\frac{1}{2}(n_1+n_2-m))} \det F^{\frac{1}{2}(m-n_1-1)} \det(I+F)^{-\frac{1}{2}(n_1+n_2)} \times \exp(-\tfrac{1}{2}\operatorname{tr}\Omega)\,{}_1F_1(\tfrac{1}{2}(n_1+n_2); \tfrac{1}{2}m; \tfrac{1}{2}\Omega(I+F^{-1})^{-1})$$

where $\Omega = M'\Sigma^{-1}M$. If the transformation $n_1 \to m$, $m \to n_1$, $n_2 \to n_1 + n_2 - m$ is made then (7.1) becomes the density function of $F = S_1^{\frac{1}{2}}S_2^{-1}S_1^{\frac{1}{2}}$ where S_1 is noncentral Wishart $W_m(n_1, \Sigma, \Omega)$ and S_2 is central Wishart $W_m(n_2, \Sigma)$, with S_1 and S_2 independent.

The ${}_1F_1$ function in (7.1) is a symmetric function of the latent roots of the argument matrix $R = \frac{1}{2}\Omega(I+F^{-1})^{-1}$. The problem of approximating this function when some of these roots are large has been studied by Constantine and Muirhead (1976). Let $r_1 \geqq r_2 \geqq \cdots \geqq r_k > r_{k+1} \geqq \cdots \geqq r_{n_1} \geqq 0$ be the latent roots of R; the asymptotic behavior of the ${}_1F_1$ function for large $r_1, \cdots, r_k$ can be obtained from an integral representation due to Herz (1955, Equation (2.9)) as

$$(7.2) \qquad {}_1F_1(\tfrac{1}{2}(n_1+n_2); \tfrac{1}{2}m; R) \sim \frac{\Gamma_k(\frac{1}{2}m)}{\Gamma_k(\frac{1}{2}(n_1+n_2))} \exp(\textstyle\sum_{i=1}^k r_i) \prod_{i=1}^k r_i^{\frac{1}{2}(n_1+n_2-m)} \times {}_1F_1(\tfrac{1}{2}(n_1+n_2-k); \tfrac{1}{2}(m-k); R_2)(1 + P_1 + O(R_1^{-2}))$$

where $R_1 = \operatorname{diag}(r_1, \cdots, r_k)$, $R_2 = \operatorname{diag}(r_{k+1}, \cdots, r_m)$ and

$$P_1 = \tfrac{1}{4}(1-n_2)(m-n_1-n_2)\textstyle\sum_{i=1}^k r_i^{-1}.$$

Since R_2 contains the small roots of R the zonal polynomial series for the ${}_1F_1$ function on the right side of (7.2) could be used for computational purposes; when $R_2 = 0$ this function is identically equal to one.

When $k = n_1$ (i.e., all roots of R large) it is possible to obtain a complete asymptotic series, namely (see Constantine and Muirhead (1976))

$${}_1F_1(\tfrac{1}{2}(n_1+n_2); \tfrac{1}{2}m; R) \sim \frac{\Gamma_{n_1}(\frac{1}{2}m)}{\Gamma_{n_1}(\frac{1}{2}(n_1+n_2))} \exp(\textstyle\sum_{i=1}^{n_1} r_i) \prod_{i=1}^{n_1} r_i^{\frac{1}{2}(n_1+n_2-m)} \times {}_2F_0(\tfrac{1}{2}(1-n_2), \tfrac{1}{2}(m-n_1-n_2); R^{-1}).$$

This approximation has also been obtained, up to terms of order r_i^{-1}, $i = 1, \cdots, n_1$, by Sargan (1976).

8. Noncentral latent roots in discriminant analysis. Suppose that the $m \times m$ matrix variates S_1 and S_2 have independent Wishart distributions $W_m(n_1, \Sigma, \Omega)$ and $W_m(n_2, \Sigma)$ respectively and let $l_1, \cdots, l_m$ be the latent roots of $S_1(S_1+S_2)^{-1}$.

The exact joint density function of these roots is, for $n_1 \geqq m$ (Constantine (1963))

$$\frac{\pi^{\frac{1}{2}m^2}\Gamma_m(\frac{1}{2}(n_1+n_2))}{\Gamma_m(\frac{1}{2}n_1)\Gamma_m(\frac{1}{2}n_2)\Gamma_m(\frac{1}{2}m)} \prod_{i=1}^m [l_i^{\frac{1}{2}(n_1-m-1)}(1-l_i)^{\frac{1}{2}(n_2-m-1)}] \prod_{i<j}^m (l_i - l_j)$$

$$(8.1) \qquad \times \exp(-\tfrac{1}{2}\textstyle\sum_{i=1}^m \omega_i)\, {}_1F_1^{(m)}(\tfrac{1}{2}(n_1+n_2); \tfrac{1}{2}n_1; \tfrac{1}{2}\Omega, L)$$

$$1 > l_1 > l_2 > \cdots > l_m > 0$$

where $\omega_1 \geqq \omega_2 \geqq \cdots \geqq \omega_m \geqq 0$ are the latent roots of the noncentrality matrix Ω, $L = \operatorname{diag}(l_1, \cdots, l_m)$ and, without loss of generality, $\Omega = \operatorname{diag}(\omega_1, \cdots, \omega_m)$. For $n_1 \leqq m$ the distribution of the nonzero roots $l_1, \cdots, l_{n_1}$ is obtained from (8.1) via the transformation

$$m \to n_1, \qquad n_1 \to m, \qquad n_2 \to n_1 + n_2 - m.$$

In a multivariate analysis of variance situation the matrices S_1 and S_2 are respectively the "between groups" and "within groups" matrices of sums of squares and sums of products. In a typical multivariate analysis of variance it is usually of interest to test whether $\Omega = 0$, at least as a first step. If this is rejected (and it is concluded that there exist real differences between the groups) interest centers on the problem of testing whether the last few noncentrality parameters ω_i are zero, in which case the corresponding discriminant functions are not useful for discriminating between the groups (see e.g., Kshirsagar (1972, Chapter 9)). From (8.1) it is seen that the marginal likelihood function of the noncentrality parameters is

$$(8.2) \qquad \exp(-\tfrac{1}{2}\textstyle\sum_{i=1}^m \omega_i)\, {}_1F_1^{(m)}(\tfrac{1}{2}(n_1+n_2); \tfrac{1}{2}n_1; \tfrac{1}{2}\Omega, L).$$

When some or all of the noncentrality parameters are large the asymptotic behavior of the ${}_1F_1^{(m)}$ function can be obtained from the integral representation

$${}_1F_1^{(m)}(\tfrac{1}{2}(n_1+n_2); \tfrac{1}{2}n_1; \tfrac{1}{2}\Omega, L) = \int_{O(m)} {}_1F_1(\tfrac{1}{2}(n_1+n_2); \tfrac{1}{2}n_1; \tfrac{1}{2}\Omega H'LH)(dH)$$

and the asymptotic behavior of the ${}_1F_1$ function in the integrand given in the previous section. Under the assumption that

$$(8.3) \qquad \omega_1 > \cdots > \omega_k > \omega_{k+1} \geqq \omega_{k+2} \geqq \cdots \geqq \omega_m \geqq 0$$

where $\omega_1, \cdots, \omega_k$ are large. Constantine and Muirhead (1976) have shown that

$${}_1F_1^{(m)}(\tfrac{1}{2}(n_1+n_2); \tfrac{1}{2}n_1; \tfrac{1}{2}\Omega, L)$$

$$(8.4) \qquad \sim \frac{\Gamma_k(\frac{1}{2}n_1)\Gamma_k(\frac{1}{2}m)}{\Gamma_k(\frac{1}{2}(n_1+n_2))} \frac{2^{\frac{1}{2}k(2m-k-2n_2-1)}}{\pi^{\frac{1}{2}k(k+1)}} \exp(\tfrac{1}{2}\textstyle\sum_{i=1}^k \omega_i l_i) \prod_{i=1}^k (\omega_i l_i)^{\frac{1}{2}n_2}$$

$$\times \prod_{i=1}^k \prod_{j=1; i<j}^m c_{ij}^{-\frac{1}{2}}\, {}_1F_1^{(m-k)}(\tfrac{1}{2}(n_1+n_2-k); \tfrac{1}{2}(n_1-k); \tfrac{1}{2}\Omega_2, L_2)$$

$$\times [1 + P_1 + O(\Omega_1^{-2})]$$

where

$$c_{ij} = (\omega_i - \omega_j)(l_i - l_j),$$

$$\Omega_1 = \operatorname{diag}(\omega_1, \cdots, \omega_k), \quad \Omega_2 = \operatorname{diag}(\omega_{k+1}, \cdots, \omega_m),$$

$$L_2 = \operatorname{diag}(l_{k+1}, \cdots, l_m)$$

and

$$P_1 = \tfrac{1}{2}n_2(n_1 + n_2 - m - 1) \sum_{i=1}^k (\omega_i l_i)^{-1} + \tfrac{1}{2} \sum_{i=1}^k \sum_{j=1;i<j}^m c_{ij}^{-1} .$$

When $k = m$ (all noncentrality parameters large and distinct) the ${}_1F_1^{(m-k)}$ function is taken to be unity, as it is also when $\omega_{k+1} = \cdots = \omega_m = 0$ (i.e., $\Omega_2 = 0$).

Substitution of (8.4) in (8.1) gives an asymptotic representation for the density function of $l_1, \cdots, l_m$ under the assumption (8.3) with $\omega_1, \cdots, \omega_k$ large. Ignoring terms of order ω_i^{-1}, $i = 1, \cdots, k$, this representation is

$$\begin{aligned} &k_4 \exp(\tfrac{1}{2} \textstyle\sum_{i=1}^k l_i \omega_i) \prod_{i=1}^k [l_i^{\frac{1}{2}(n_1+n_2-m-1)}(1 - l_i)^{\frac{1}{2}(n_2-m-1)}] \prod_{i<j}^k \left(\dfrac{l_i - l_j}{\omega_i - \omega_j}\right)^{\frac{1}{2}} \\ (8.5) \qquad &\times \textstyle\prod_{i=1}^k \prod_{j=k+1}^m (l_i - l_j)^{\frac{1}{2}} \prod_{i=k+1}^m [l_i^{\frac{1}{2}(n_1-m-1)}(1 - l_i)^{\frac{1}{2}(n_2-m-1)}] \\ &\times \textstyle\prod_{k+1;i<j}^m (l_i - l_j)\, {}_1F_1^{(m-k)}(\tfrac{1}{2}(n_1 + n_2 - k); \tfrac{1}{2}(n_1 - k); \tfrac{1}{2}\Omega_2, L_2) \end{aligned}$$

where

$$\begin{aligned} k_4 &= \pi^{\frac{1}{2}m^2 - \frac{1}{4}k(k+1)} 2^{\frac{1}{4}k(2m-k-2n_2-1)} \frac{\Gamma_m(\frac{1}{2}(n_1 + n_2))\Gamma_k(\frac{1}{2}n_1)\Gamma_k(\frac{1}{2}m)}{\Gamma_m(\frac{1}{2}n_1)\Gamma_m(\frac{1}{2}n_2)\Gamma_m(\frac{1}{2}m)\Gamma_k(\frac{1}{2}(n_1 + n_2))} \\ &\quad \times \exp(-\tfrac{1}{2} \textstyle\sum_{i=1}^m \omega_i) \prod_{i=1}^k \omega_i^{\frac{1}{4}(n_2-m+k)} . \end{aligned}$$

The last two lines in (8.5), multiplied by an appropriate constant, give an asymptotic representation for the conditional density function of the $q = m - k$ smallest roots $l_{k+1}, \cdots, l_m$ given the k-largest roots $l_1, \cdots, l_k$. Note that it does not depend on $\omega_1, \cdots, \omega_k$. When the null hypothesis $H_q : \omega_{k+1} = \cdots = \omega_m = 0$ is true the asymptotic conditional distribution becomes

$$\begin{aligned} (8.6) \qquad &\text{const.} \textstyle\prod_{i=1}^k \prod_{j=k+1}^m (l_i - l_j)^{\frac{1}{2}} \prod_{i=k+1}^m [l_i^{\frac{1}{2}(n_1-m-1)}(1 - l_i)^{\frac{1}{2}(n_2-m-1)}] \\ &\times \textstyle\prod_{k+1;i<j}^m (l_i - l_j) . \end{aligned}$$

If the linkage factors

$$\prod_{i=1}^k \prod_{j=k+1}^m (l_i - l_j)^{\frac{1}{2}}$$

are ignored, this is just the distribution of the latent roots of $S_1(S_1 + S_2)^{-1}$ where S_1 and S_2 have independent Wishart distributions $W_q(n_1 - k, \Sigma)$ and $W_q(n_2 - k, \Sigma)$ respectively.

Another situation which can arise in practice is when both the error degrees of freedom and the noncentrality matrix are large. Here it is assumed that $\Omega = n_2\Delta$ where Δ is a fixed matrix with latent roots $\delta_1 \geqq \cdots \geqq \delta_m \geqq 0$ and that n_2 is large. W. Glynn (1977) in a recent Yale Ph. D. thesis, has derived the asymptotic behavior of the function ${}_1F_1^{(m)}(\frac{1}{2}(n_1 + n_2); \frac{1}{2}n_1; \frac{1}{2}n_2\Delta, L)$ as $n_2 \to \infty$ under the assumption that

$$\delta_1 > \cdots > \delta_k > \delta_{k+1} = \cdots = \delta_m = 0 .$$

Glynn's results will undoubtably be published at a later date. What is particularly interesting is that although the asymptotic representation of the joint density function of $l_1, \cdots, l_m$ in this case is markedly different from that given by (8.5), the asymptotic representation for the conditional distribution of $l_{k+1}, \cdots, l_m$ given $l_1, \cdots, l_k$ is the same as that given by (8.6); that is, (8.6) serves as both the asymptotic conditional distribution when $\omega_1, \cdots, \omega_k$ are large (independently of n_2) and the asymptotic conditional distribution when $\omega_1, \cdots, \omega_k$ are large in a way depending on n_2 via $\omega_i = n_2 \delta_i$.

A third possibility that can arise is when n_2 is large and the noncentrality matrix Ω remains fixed. Some asymptotic work in this direction has been done by Chattopadhyay and Pillai (1973) and Chattopadhyay, Pillai and Li (1976); however the asymptotic behavior of the ${}_1F_1^{(m)}$ function given by these authors involves the one-matrix function ${}_1F_1(\frac{1}{2}(n_1 + n_2); \frac{1}{2}n_1; \frac{1}{2}\Omega L)$ and as it stands does not appear particularly amenable to numerical or inference work. It appears that this situation requires further analysis.

To conclude this section it is worth noting that in the second case discussed above (i.e., $\Omega = n_2 \Delta$) it is possible to obtain asymptotic distributions of suitably standardized roots which are of a normal type and do not preserve linkage factors corresponding to distinct noncentrality parameters. Such asymptotic distributions have been studied by P. L. Hsu (1941 b) and T. W. Anderson (1951). Assume that Δ has r different roots with multiplicities $q_1, \cdots, q_r$ ($\sum_{\alpha=1}^{r} q_\alpha = m$), namely

$$\begin{aligned} \delta_1 &= \cdots = \delta_{q_1} &&= \gamma_1 \\ \delta_{q_1+1} &= \cdots = \delta_{q_1+q_2} &&= \gamma_2 \\ &\vdots \\ \delta_{m-q_r+1} &= \cdots = \delta_m &&= \gamma_r \end{aligned}$$

with $\gamma_1 > \cdots > \gamma_{r-1} > \gamma_r = 0$ and define new variables $x_1, \cdots, x_m$ by putting

$$x_i = \left(\frac{n_2}{4\gamma_\alpha + 2\gamma_\alpha^2}\right)^{\frac{1}{2}} \left(\frac{l_i}{1 - l_i} - \gamma_\alpha\right) \qquad \text{for} \quad i \in J_\alpha , \qquad \alpha = 1, \cdots, r - 1$$

and

$$x_i = n_2 \frac{l_i}{1 - l_i} \qquad \text{for} \quad i \in J_r$$

where J_α is the set of integers $\sum_{i=1}^{\alpha-1} q_i + 1, \cdots, \sum_{i=1}^{\alpha} q_i$. The asymptotic joint density function of $x_1, \cdots, x_m$ as $n_2 \to \infty$ is (Hsu (1941 b), Anderson (1951))

$$\prod_{\alpha=1}^{r-1} \frac{\pi^{\frac{1}{4}q_\alpha(q_\alpha-1)}}{2^{\frac{1}{2}q_\alpha}\Gamma_{q_\alpha}(\frac{1}{2}q_\alpha)} \exp(-\tfrac{1}{2} \textstyle\sum_{i\in J_\alpha} x_i^2) \prod_{i<j;i,j\in J_\alpha} (x_i - x_j) \dfrac{\pi^{\frac{1}{2}s_1^2}}{2^{\frac{1}{2}s_1 s_2}\Gamma_{s_1}(\frac{1}{2}s_1)\Gamma_{s_1}(\frac{1}{2}s_2)}$$
$$\times \exp(-\tfrac{1}{2} \textstyle\sum_{i\in J_r} x_i) \prod_{i\in J_r} x_i^{\frac{1}{2}(s_2-s_1-1)} \prod_{i<j;i,j\in J_r} (x_i - x_j)$$

where

$$s_1 = m - \textstyle\sum_{i=1}^{r-1} q_i , \qquad s_2 = n_1 - \sum_{i=1}^{r-1} q_i .$$

More recently Fujikoshi (1976 b) has obtained the term of order $n_2^{-\frac{1}{2}}$ in an asymptotic expansion for the joint density function. The x_i's corresponding to the multiple zero population root have, asymptotically, the same distribution as the roots of a matrix with the $W_{s_1}(s_2, I)$ distribution while the asymptotic distribution of any x_i corresponding to a simple nonzero population root is standard normal.

9. Canonical correlation coefficients. The joint density function of the squares $r_1^2, \cdots, r_p^2$ of the sample canonical correlation coefficients between variates $y_1, \cdots, y_p$ and $x_1, \cdots, x_q$, $p \leqq q$, calculated from a sample of size $n+1$ from a $(p+q)$-variate normal distribution is (Constantine (1963))

$$(9.1)\qquad \begin{aligned} &\frac{\Gamma_p(\frac{1}{2}n)\pi^{\frac{1}{2}p^2}}{\Gamma_p(\frac{1}{2}(n-q))\Gamma_p(\frac{1}{2}q)\Gamma_p(\frac{1}{2}p)} \prod_{i=1}^p [(r_i^2)^{\frac{1}{2}(q-p-1)}(1-r_i^2)^{\frac{1}{2}(n-q-p-1)}] \\ &\quad \times \prod_{i<j}^p (r_i^2 - r_j^2) \prod_{i=1}^p (1-\rho_i^2)^{\frac{1}{2}n} {}_2F_1^{(p)}(\tfrac{1}{2}n, \tfrac{1}{2}n; \tfrac{1}{2}q; P^2, R^2) \\ &\qquad\qquad 1 > r_1^2 > r_2^2 > \cdots > r_p^2 > 0 \end{aligned}$$

where $1 > \rho_1 \geqq \rho_2 \geqq \cdots \geqq \rho_p \geqq 0$ are the population canonical correlation coefficients, $P \doteq \operatorname{diag}(\rho_1, \cdots, \rho_p)$, $R = \operatorname{diag}(r_1, \cdots, r_p)$. The likelihood function of the population coefficients is then

$$(9.2)\qquad \prod_{i=1}^p (1-\rho_i^2)^{\frac{1}{2}n} {}_2F_1^{(p)}(\tfrac{1}{2}n, \tfrac{1}{2}n; \tfrac{1}{2}q; P^2, R^2) .$$

An often used test in canonical correlation analysis is the likelihood ratio test of the null hypothesis that the $p-k$ smallest population canonical correlation coefficients are zero when the first k population coefficients, corresponding to real relationships between the two sets of variates, have been removed. Hence it is of particular interest to know the asymptotic behavior for large n of the ${}_2F_1^{(p)}$ function in (9.1) under the assumption that

$$(9.3)\qquad 1 > \rho_1 > \rho_2 > \cdots > \rho_k > \rho_{k+1} = \rho_{k+2} = \cdots = \rho_p = 0 .$$

This has been obtained by Glynn and Muirhead (1977) using Laplace's method applied to a complicated multiple integral representation for ${}_2F_1^{(p)}$. For large n,

$$(9.4)\qquad \begin{aligned} &{}_2F_1^{(p)}(\tfrac{1}{2}n, \tfrac{1}{2}n; \tfrac{1}{2}q; P^2, R^2) \\ &\quad \sim (\tfrac{1}{2}n)^{-\frac{1}{2}k(p+q-k-1)}\pi^{-\frac{1}{2}k(k+1)}\Gamma_k(\tfrac{1}{2}q)\Gamma_k(\tfrac{1}{2}p)2^{-k} \\ &\qquad \times \prod_{i=1}^k \{(1-r_i\rho_i)^{-n+\frac{1}{2}(p+q-1)}(r_i\rho_i)^{\frac{1}{2}(p-q)}\} \prod_{i=1}^k \prod_{j=k+1; i<j}^p c_{ij}^{-\frac{1}{2}} \end{aligned}$$

where

$$\begin{aligned} c_{ij} &= (r_i^2 - r_j^2)(\rho_i^2 - \rho_j^2) & i, j = 1, \cdots, k \\ &= (r_i^2 - r_j^2)\rho_i^2 & i = 1, \cdots, k;\ j = k+1, \cdots, p . \end{aligned}$$

An alternative asymptotic result has been given by Chattopadhyay and Pillai (1973) and Chattopadhyay, Pillai and Li (1976); however the asymptotic behavior given in these papers involves a ${}_2F_1$ function with the matrix P^2R^2 as argument and as it stands does not appear amenable to numerical or inference work. It would be of some interest to have an asymptotic result for ${}_2F_1^{(p)}$ when the last $p-k$ population coefficients are small (and not necessarily zero), but this remains an unsolved problem.

When the population coefficients are all distinct $(k = p)$ and n is large the likelihood function (9.2) can be approximated as

$$(9.5)\qquad \begin{aligned} &\prod_{i=1}^p (1-\rho_i^2)^{\frac{1}{2}n} {}_2F_1^{(p)}(\tfrac{1}{2}n, \tfrac{1}{2}n; \tfrac{1}{2}q; P^2, R^2) \\ &\quad \approx K \cdot \prod_{i=1}^p [(1-\rho_i^2)^{\frac{1}{2}n}(1-r_i\rho_i)^{-n+\frac{1}{2}(p+q-1)}\rho_i^{\frac{1}{2}(p-q)}] \prod_{i<j}^p (\rho_i^2 - \rho_j^2)^{-\frac{1}{2}} \\ &\qquad \times [1 + n^{-1}P_1 + O(n^{-2})] \end{aligned}$$

where K is a function of $n, r_1, \cdots, r_p$ but not of $\rho_1, \cdots, \rho_p$ and

$$P_1 = \tfrac{1}{2} \sum_{i<j}^{p} \frac{(1 - r_i\rho_i)(1 - r_j\rho_j)(r_i\rho_i + r_j\rho_j)}{(r_i^2 - r_j^2)(\rho_i^2 - \rho_j^2)}$$
$$- \tfrac{1}{8}(q - p)(q - p - 2) \sum_{i=1}^{p} (r_i\rho_i)^{-1} + \tfrac{1}{8} \sum_{i=1}^{p} r_i\rho_i \, .$$

As estimates of the parameters $\xi_i = \tanh^{-1}\rho_i$ the statistics $z_i = \tanh^{-1} r_i$ have bias terms of order n^{-1}; the maximum marginal likelihood estimates of the ξ_i, obtained by maximizing the right side of (9.5), are

$$\hat{\xi}_i = z_i - \frac{1}{2nr_i}\left[p + q - 2 + r_i^2 + 2(1 - r_i^2) \sum_{j\neq i} \frac{r_j^2}{r_i^2 - r_j^2}\right] + O(n^{-2})$$

having as their first two moments

$$E(\hat{\xi}_i) = \xi_i + O(n^{-2})$$
$$\mathrm{Var}\,(\hat{\xi}_i) = \frac{1}{n} + O(n^{-2}) \, .$$

Thus these corrected estimates $\hat{\xi}_i$ not only stabilize the variance to order n^{-1} but also provide a correction for bias.

Substitution of (9.4) in (9.1) gives an asymptotic representation, for large n, of the density function of $r_1^2, \cdots, r_p^2$ under the assumption (9.3). Ignoring terms of order n^{-1} this representation is

$$k_5 \prod_{i=1}^{k} [r_i^{\frac{1}{2}(q-p)-1}(1 - r_i\rho_i)^{-n+\frac{1}{2}(p+q-1)}(1 - r_i^2)^{\frac{1}{2}(n-q-p-1)}] \prod_{i<j}^{k} \left(\frac{r_i^2 - r_j^2}{\rho_i^2 - \rho_j^2}\right)^{\frac{1}{2}}$$

(9.6)
$$\times \prod_{i=1}^{k} \prod_{j=k+1}^{p} (r_i^2 - r_j^2)^{\frac{1}{2}} \prod_{i=k+1}^{p} [(r_i^2)^{\frac{1}{2}(q-p-1)}(1 - r_i^2)^{\frac{1}{2}(n-q-p-1)}]$$
$$\times \prod_{k+1;i<j}^{p} (r_i^2 - r_j^2)$$

where

$$k_5 = \frac{\pi^{\frac{1}{2}p^2 - \frac{1}{2}k(k+1)}}{2^k}\left(\frac{2}{n}\right)^{\frac{1}{2}k(p+q-k-1)} \frac{\Gamma_p(\frac{1}{2}n)\Gamma_k(\frac{1}{2}q)\Gamma_k(\frac{1}{2}p)}{\Gamma_p(\frac{1}{2}(n-q))\Gamma_p(\frac{1}{2}q)\Gamma_p(\frac{1}{2}p)}$$
$$\times \prod_{i=1}^{p} (1 - \rho_i^2)^{\frac{1}{2}n} \prod_{i=1}^{k} \rho_i^{k-\frac{1}{2}(p+q)} \, .$$

The second line in (9.6) is proportional to an asymptotic representation for the conditional density function of the $p - k$ smallest coefficients $r_{k+1}^2, \cdots, r_p^2$ given the largest k coefficients $r_1^2, \cdots, r_k^2$ when the null hypothesis that $\rho_{k+1} = \cdots = \rho_p = 0$ with $\rho_1 > \cdots > \rho_k > 0$ is true. Note that it does not depend on $\rho_1, \cdots, \rho_k$. If the linkage factors

$$\prod_{i=1}^{k} \prod_{j=k+1}^{p} (r_i^2 - r_j^2)^{\frac{1}{2}}$$

are neglected then this asymptotic conditional distribution is just the null distribution of the canonical correlation coefficients between variates $y_1, \cdots, y_{p'}$ and $x_1, \cdots, x_{q'}$ calculated from a sample of size $n' + 1$ observations from a $(p' + q')$-variate normal distribution, where $n' = n - 2k$, $p' = p - k$ and $q' = q - k$.

The likelihood ratio statistic for testing that the last $p - k$ population canonical correlation coefficients are zero when the first k population coefficients have been removed is

$$V_{p-k} = \prod_{i=k+1}^{p} (1 - r_i^2) .$$

Glynn and Muirhead (1977) have shown (see also Lawley (1959), Bartlett (1947)) that, under the null hypothesis, the statistic

$$T_{p-k} = -[n - k - \tfrac{1}{2}(p + q + 1) + \sum_{i=1}^{k} r_i^{-2}] \log V_{p-k}$$

has an asymptotic χ^2 distribution with $(p - k)(q - k)$ degrees of freedom and

$$E(T_{p-k}) = (p - k)(q - k) + O(n^{-2}) ,$$

the expectation being taken with respect to the conditional distribution of $r_{k+1}^2, \cdots, r_p^2$ given $r_1^2, \cdots, r_k^2$. Comments similar to those made in Section 3 regarding remaining problems in connection with the likelihood ratio test are also applicable here.

From the asymptotic joint distribution of $r_1^2, \cdots, r_p^2$ it is a simple matter to obtain marginal distributions. When ρ_i is distinct from the other population coefficients the asymptotic marginal density function of r_i^2 is

$$\frac{1}{2}\left(\frac{n}{2\pi}\right)^{\frac{1}{2}} (1 - \rho_i^2)^{\frac{1}{2}n} \rho_i^{\frac{1}{2}(p-q)} (1 - r_i \rho_i)^{-n+\frac{1}{2}(p+q-1)} (r_i^2)^{\frac{1}{2}(q-p)-\frac{1}{2}}$$
$$\times (1 - r_i^2)^{\frac{1}{2}(n-p-q-1)} \prod_{j=1}^{i-1} \left(\frac{\rho_j^2 - r_i^2}{\rho_j^2 - \rho_i^2}\right)^{\frac{1}{2}} \prod_{j=i+1}^{p} \left(\frac{r_i^2 - \rho_j^2}{\rho_i^2 - \rho_j^2}\right)^{\frac{1}{2}} ,$$

valid for $r_i^2 \in (\rho_{i+1}^2, \rho_{i-1}^2)$.

The asymptotic distributions discussed so far resemble products of beta distributions linked by factors of the form $r_i^2 - r_j^2$. A suitable standardization of the coefficients yields asymptotic distributions which are of a normal type. Putting

$$x_i = \frac{n^{\frac{1}{2}}(r_i^2 - \rho_i^2)}{2\rho_i(1 - \rho_i^2)} \qquad i = 1, \cdots, k$$
$$x_j = nr_j^2 \qquad j = k + 1, \cdots, p$$

in the asymptotic distribution (9.6) of $r_1^2, \cdots, r_p^2$ under the assumption (9.2) gives an expansion for the joint density function of $x_1, \cdots, x_m$, the limiting term of which is

$$(9.7) \qquad \prod_{i=1}^{k} \phi(x_i) \cdot \frac{\pi^{\frac{1}{2}p'^2}}{2^{\frac{1}{2}p'q'} \Gamma_{p'}(\frac{1}{2}q') \Gamma_{p'}(\frac{1}{2}p')} \exp(-\tfrac{1}{2} \sum_{j=k+1}^{p} x_j) \prod_{j=k+1}^{p} x_j^{\frac{1}{2}(q'-p'-1)}$$
$$\times \prod_{k+1; i<j}^{p} (x_i - x_j)$$

where $\phi(\cdot)$ denotes the standard normal density function and $p' = p - k$, $q' = q - k$. This limiting result was first given by P. L. Hsu (1941a). From (9.7) it is seen that asymptotically the x_i's corresponding to distinct nonzero ρ_i's are marginally standard normal, independent of all x_j, $j \neq i$, while the x_j's corresponding to zero population coefficients are nonnormal and dependent, and their asymptotic distribution is the same as the distribution of the latent roots of a $p' \times p'$ matrix having the $W_{p'}(q', I_{p'})$ distribution.

Acknowledgments. A number of people made useful suggestions and criticisms; for their help and advice the author would like to thank the referees, the associate editor, J. A. Hartigan, A. T. James, I. R. Savage, N. Sugiura and C. Waternaux.

REFERENCES

[1] ANDERSON, G. A. (1965). An asymptotic expansion for the distribution of the latent roots of the estimated covariance matrix. *Ann. Math. Statist.* **36** 1153–1173.

[2] ANDERSON, G. A. (1970). An asymptotic expansion for the noncentral Wishart distribution. *Ann. Math. Statist.* **41** 1700–1707.

[3] ANDERSON, T. W. (1951). The asymptotic distribution of certain characteristic roots and vectors. *Proc. Second Berkeley Symp. Math. Statist. Prob.* 103–130. Univ. of Calif. Press.

[4] ANDERSON, T. W. (1958). *An Introduction to Multivariate Statistical Analysis.* Wiley, New York.

[5] ANDERSON, T. W. (1963). Asymptotic theory for principal component analysis. *Ann. Math. Statist.* **34** 122–148.

[6] BARNARD, G. A. (1963). Some logical aspects of the fiducial argument. *J. Roy. Statist. Soc. Ser. B* **25** 111–114.

[7] BARTLETT, M. S. (1947). Multivariate analysis. *J. Roy. Statist. Soc. (Suppl.)* **9** 176–190.

[8] BARTLETT, M. S. (1954). A note on multiplying factors for various χ^2 approximations. *J. Roy. Statist. Soc. Ser. B* **16** 296–298.

[9] BINGHAM, C. (1972). An asymptotic expansion for the distribution of the eigenvalues of a 3 by 3 Wishart matrix. *Ann. Math. Statist.* **43** 1498–1506.

[10] CHANG, T. C. (1970). On an asymptotic representation of the distribution of the characteristic roots of $S_1S_2^{-1}$. *Ann. Math. Statist.* **41** 440–445.

[11] CHANG, T. C. (1973). On an asymptotic distribution of the characteristic roots of $S_1S_2^{-1}$ when roots are not all distinct. *Ann. Inst. Statist. Math.* **25** 447–452.

[12] CHATTOPADHYAY, A. K. and PILLAI, K. C. S. (1973). Asymptotic expansions for the distributions of characteristic roots when the parameter matrix has several multiple roots. In *Multivariate Analysis III* (P. R. Krishnaiah, Ed.), Academic Press, New York.

[13] CHATTOPADHYAY, A. K., PILLAI, K. C. S. and LI, H. C. (1976). Maximization of an integral of a matrix function and asymptotic expansions of distributions of latent roots of two matrices. *Ann. Statist.* **4** 796–806.

[14] CHIKUSE, Y. (1974). Asymptotic expansions for the distributions of the latent roots of two matrices in multivariate analysis. Ph. D. Thesis, Yale University.

[15] CHIKUSE, Y. (1976). Asymptotic distributions of the latent roots of the covariance matrix with multiple population roots. *J. Multivariate Anal.* **6** 237–249.

[16] CONSTANTINE, A. G. (1963). Some noncentral distribution problems in multivariate analysis. *Ann. Math. Statist.* **34** 1270–1285.

[17] CONSTANTINE, A. G. (1966). The distribution of Hotelling's generalized T_0^2. *Ann. Math. Statist.* **37** 215–225.

[18] CONSTANTINE, A. G. and MUIRHEAD, R. J. (1972). Partial differential equations for hypergeometric functions of two argument matrices. *J. Multivariate Anal.* **3** 332–338.

[19] CONSTANTINE, A. G. and MUIRHEAD, R. J. (1976). Asymptotic expansions for distributions of latent roots in multivariate analysis. *J. Multivariate Anal.* **6** 369–391.

[20] CROWTHER, N. A. S. and YOUNG, D. L. (1974). Notes on the distributions of characteristic roots and functions of characteristic roots of certain matrices in multivariate analysis. Tech. Rep. No. 92, Stanford Univ.

[21] DEMPSTER, A. (1966). Estimation in multivariate analysis. In *Multivariate Analysis* (P. R. Krishnaiah, Ed.), 315-334. Academic Press, New York.

[22] FARRELL, R. H. (1976). *Techniques of Multivariate Calculation.* Springer, New York.

[23] FUJIKOSHI, Y. (1975). Partial differential equations for hypergeometric functions ${}_3F_2$ of matrix argument. *Canad. J. Statist.* **3** 153–163.

[24] FUJIKOSHI, Y. (1976a). An asymptotic expansion for the distributions of the latent roots of the Wishart matrix with multiple population roots. Unpublished manuscript.

[25] FUJIKOSHI, Y. (1976b). Asymptotic expansions for the distribution of the latent roots in MANOVA and canonical correlations. Unpublished manuscript.

[26] FUJIKOSHI, Y. (1976c). Asymptotic expansions for the distributions of some multivariate tests. In *Multivariate Analysis—IV* (P. R. Krishnaish, Ed.), 55–71. North-Holland Publishing Co.

[27] GIRSHICK, M. A. (1939). On the sampling theory of roots of determinantal equations. *Ann. Math. Statist.* **10** 203–224.

[28] GLYNN, W. J. (1977). Asymptotic distributions of latent roots in canonical correlation analysis and in discriminant analysis with applications to testing and estimation. Ph. D. Thesis, Yale University.

[29] GLYNN, W. J. and MUIRHEAD, R. J. (1977). Inference in canonical correlation analysis. To appear in *J. Multivariate Anal.*

[30] HAYAKAWA, T. (1969). On the distribution of the latent roots of a positive definite random symmetric matrix I. *Ann. Inst. Statist. Math.* **21** 1–21.

[31] HERZ, C. S. (1955). Bessel functions of matrix argument. *Ann. of Math.* **61** 474–523.

[32] HSU, L. C. (1948). A theorem on the asymptotic behavior of a multiple integral. *Duke Math. J.* **15** 623–632.

[33] HSU, P. L. (1941a). On the limiting distribution of the canonical correlations. *Biometrika* **32** 38–45.

[34] HSU, P. L. (1941b). On the limiting distribution of roots of a determinantal equation. *J. London Math. Soc.* **16** 183–194.

[35] JAMES, A. T. (1955). A generating function for averages over the orthogonal group. *Proc. Roy. Soc. Ser. A* **229** 367–375.

[36] JAMES, A. T. (1964). Distributions of matrix variates and latent roots derived from normal samples. *Ann. Math. Statist.* **35** 475–501.

[37] JAMES, A. T. (1966). Inference on latent roots by calculation of hypergeometric functions of matrix argument. In *Multivariate Analysis* (P. R. Krishnaiah, Ed.), 209–235. Academic Press, New York.

[38] JAMES, A. T. (1968). Calculation of zonal polynomial coefficients by use of the Laplace-Beltrami operator. *Ann. Math. Statist.* **39** 1711–1718.

[39] JAMES, A. T. (1969). Test of equality of the latent roots of the covariance matrix. In *Multivariate Analysis II* (P. R. Krishnaiah, Ed.), 205–218. Academic Press, New York.

[40] JAMES, A. T. (1976). Special functions of matrix and single argument in statistics. In *Theory and Applications of Special Functions* (R. A. Askey, Ed.), 497–520. Academic Press, New York.

[41] JAMES, A. T. and CONSTANTINE, A. G. (1974). Generalized Jacobi polynomials as spherical functions of the Grassmann manifold. *Proc. London Math. Soc.* **29** 174–192.

[42] KHATRI, C. G. and SRIVASTAVA, M. S. (1974). Asymptotic expansions of the nonnull distributions of likelihood ratio criteria for covariance matrices. *Ann. Statist.* **2** 109–117.

[43] KRISHNAIAH, P. R. (1977). Some recent developments on real multivariate distributions. Unpublished.

[44] KSHIRSAGAR, A. M. (1972). *Multivariate Analysis.* Marcel Dekker, New York.

[45] LAWLEY, D. N. (1956). Test of significance for the latent roots of covariance and correlation matrices. *Biometrika* **43** 128–136.

[46] LAWLEY, D. N. (1959). Tests of significance in canonical analysis. *Biometrika* **46** 59–66.

[47] LEACH, B. G. (1969). Bessel functions of matrix argument with statistical applications. Ph. D. Thesis, Univ. of Adelaide.

[48] Li, H. C., Pillai, K. C. S. and Chang, T. C. (1970). Asymptotic expansions for distributions of the roots of two matrices from classical and complex Gaussian populations. *Ann. Math. Statist.* **41** 1541–1556.

[49] McLaren, M. L. (1976). Coefficients of the zonal polynomials. *Appl. Statist.* **25** 82–87.

[50] Muirhead, R. J. (1970a). Partial differential equations for hypergeometric functions of matrix argument. *Ann. Math. Statist.* **41** 991–1001.

[51] Muirhead, R. J. (1970b). Asymptotic distributions of some multivariate tests. *Ann. Math. Statist.* **41** 1002–1010.

[52] Muirhead, R. J. (1972a). On the test of independence between two sets of variates. *Ann. Math. Statist.* **43** 1491–1497.

[53] Muirhead, R. J. (1972b). The asymptotic noncentral distribution of Hotelling's generalized T_0^2. *Ann. Math. Statist.* **43** 1671–1677.

[54] Muirhead, R. J. (1974). Powers of the largest latent root test of $\Sigma = I$. *Comm. Statist.* **3** 513–524.

[55] Muirhead, R. J. (1975). Expressions for some hypergeometric functions of matrix argument with applications. *J. Multivariate Anal.* **5** 283–293.

[56] Muirhead, R. J. and Chikuse, Y. (1975). Asymptotic expansions for the joint and marginal distributions of the latent roots of the covariance matrix. *Ann. Statist.* **3** 1011–1017.

[57] Nagao, H. (1970). Asymptotic expansions of some test criteria for homogeneity of variances and covariance matrices from normal populations. *J. Sci. Hiroshima Univ. Ser. A—I Math.* **34** 153–247.

[58] Parkhurst, A. M. and James, A. T. (1974). Zonal polynomials of order 1 through 12. In *Selected Tables in Mathematical Statistics* (H. L. Harter and D. B. Owen, Eds.), 199–388. American Mathematical Society, Providence.

[59] Pillai, K. C. S. (1976). Distributions of characteristic roots. Mimeo Series No. 462, Purdue Univ.

[60] Sargan, J. D. (1976). Econometric estimators and the Edgeworth approximation. *Econometrica* **44** 421–448.

[61] Subrahmaniam, K. (1974). Recent trends in multivariate distribution theory: On the zonal polynomials and other functions of matrix argument. Part I: zonal polynomials. Tech. Rep. No. 69, Univ. of Manitoba.

[62] Sugiura, N. (1969). Asymptotic expansions for the distributions of the likelihood ratio criteria for covariance matrix. *Ann. Math. Statist.* **40** 2051–2063.

[63] Sugiura, N. (1972). Asymptotic solutions of the hypergeometric functions ${}_1F_1$ of matrix argument, useful in multivariate analysis. *Ann. Inst. Statist. Math.* **24** 517–524.

[64] Sugiura, N. (1973). Derivatives of the characteristic root of a symmetric or Hermitian matrix with two applications in multivariate analysis. *Comm. Statist.* **1** 393–417.

[65] Sugiura, N. (1974). Asymptotic formulas for the hypergeometric function ${}_2F_1$ of matrix argument, useful in multivariate analysis. *Ann. Inst. Statist. Math.* **26** 117–126.

[66] Sugiura, N. (1976a). Asymptotic expansions of the distributions of the latent roots and the latent vector of the Wishart and multivariate F matrices. *J. Multivariate Anal.* **6** 500–525.

[67] Sugiura, N. (1976b). Asymptotic expansions of the distributions of the canonical correlations. Essay in memory of establishment of faculty and integrated science, Hiroshima University.

[68] Sugiura, N. (1976c). Asymptotic nonnull distributions of the likelihood ratio criteria for the equality of several characteristic roots of a Wishart matrix. In *Essays in Probability and Statistics* (S. Ikeda et al., Eds.), 253–264. Shinko Tsusho Co. Ltd., Tokyo.

Department of Statistics
Yale University
New Haven, Connecticut 06520