AN ASYMPTOTIC EXPANSION FOR THE DISTRIBUTION OF THE LATENT ROOTS OF THE ESTIMATED COVARIANCE MATRIX[1]

By George A. Anderson[2]

Yale University

0. Summary. The distribution of the latent roots depends on a definite integral over the group of orthogonal matrices. This integral defines a function of the latent roots of both the covariance matrix and the estimated covariance matrix. With an integration procedure involving first a substitution and then an expansion of the resulting integrand the first three terms of an expansion for the integral are found. This expansion is given in increasing powers of n^{-1}, where n is the sample number less one. A numerical example is given for the distribution of the latent roots using the expansion for the definite integral given in this paper. Improved maximum likelihood estimates for the latent roots are found and the likelihood function is considered in detail.

1. Introduction. To motivate a study of the distribution of the latent roots consider the notion of a principal component. Suppose x has the normal multivariate distribution with mean μ and covariance matrix Σ where

$$x = (x_1, x_2, \cdots, x_K)', \qquad \mu = (\mu_1, \mu_2, \cdots, \mu_K)',$$

and

$$\Sigma = \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1K} \\ \vdots & & \vdots \\ \sigma_{K1} & \cdots & \sigma_{KK} \end{bmatrix}.$$

Since Σ is positive definite symmetric, there is an orthogonal matrix H_1 such that

$$H_1'\Sigma H_1 = \begin{bmatrix} \alpha_1 & & & 0 \\ & \alpha_2 & & \\ & & \ddots & \\ 0 & & & \alpha_K \end{bmatrix} = \mathfrak{A} \tag{1.1}$$

where $\alpha_1 \geqq \alpha_2 \geqq \cdots \geqq \alpha_K > 0$. Let $z = H_1'(x - \mu)$. Then z has the normal distribution with mean zero and covariance matrix $\mathfrak{A}$. Clearly

$$x = \mu + z_1 h_1 + z_2 h_2 + \cdots + z_K h_K$$

where h_i is the ith column of H_1. If for $i > r$ the α_i are very small then the cor-

Received 31 August 1964; revised 25 January 1965.

[1] This research was partially supported by the National Science Foundation Grant No. NSF-GP-6.

[2] Now at Trinity College.

responding z_i are nearly zero and with small error we write

$$x = \mu + z_1 h_1 + \cdots + z_r h_r .$$

Thus we are interested in those principal components z_i which have large variance. The variances of the principal components are estimated by the latent roots of the estimated covariance matrix.

Consider the distribution of these latent roots. Let

$$Y = \begin{bmatrix} y_{11} & \cdots & y_{1N} \\ \vdots & & \vdots \\ y_{K1} & \cdots & y_{KN} \end{bmatrix}$$

be the sample matrix of N observations from a normal K-variate population with mean μ and covariance matrix Σ. It is well known that $W = nS$, where S, the sample covariance matrix, is given by

$$S = (s_{ij}) = n^{-1}(\textstyle\sum_{k=1}^{N} (y_{ik} - \bar{y}_i)(y_{jk} - \bar{y}_j)), \qquad n = N - 1,$$

has the Wishart distribution on n degrees of freedom. Since S is positive definite symmetric we can write

$$S = HLH' \tag{1.2}$$

where H is the group $O(K)$ of $K \times K$ orthogonal matrices and

$$L = \begin{bmatrix} l_2 & & & 0 \\ & l_2 & & \\ & & \ddots & \\ 0 & & & l_K \end{bmatrix}$$

with $l_1 \geqq l_2 \geqq \cdots \geqq l_K > 0$. Finally the (marginal) distribution of the latent roots comes to

$$dF(l_1, l_2, \cdots, l_K) = c(\textstyle\prod_{i=1}^{K} l_i)^{(n-K-1)/2} \prod_{i<j}^{K} (l_i - l_j) \prod_{i=1}^{K} dl_i |\Sigma|^{-n/2} \int_{O(K)} \exp [(-n/2) \operatorname{tr} \Sigma^{-1} HLH'](H'\, dH) \tag{1.3}$$

where

$$c = 2^{-K}(n/2)^{nK/2}/\pi^{K(K-1)/4} \prod_{i=1}^{K} \Gamma\left(\frac{n - K + i}{2}\right)$$

and $(H'\, dH) = \prod_{i<j}^{K} h_i'\, dh_j$ is the invariant measure on the group $O(K)$ with h_i and dh_j the ith and jth columns of H and dH. This measure is considered in detail in James [6]. The group $O(K)$ has volume

$$V(K) = \int_{O(K)} (H'\, dH) = 2^K \pi^{K(K+1)/4} / \textstyle\prod_{i=1}^{K} \Gamma(i/2).$$

If we replace H by $H_1 H$ with H_1 defined in (1.1) then the distribution (1.3) depends on the integral

$$I_1 = \int_{O(K)} \exp [(-n/2) \operatorname{tr} AHLH'](H'\, dH) \tag{1.4}$$

where

$$A = \begin{bmatrix} a_1 & & & 0 \\ & a_2 & & \\ & & \ddots & \\ 0 & & & a_K \end{bmatrix} = \mathfrak{A}^{-1}$$

so that $a_i = (\alpha_i)^{-1}$; $0 < a_1 \leqq a_2 \leqq \cdots \leqq a_K$.

The power series

$$I_1 = V(K) \sum_{k=0}^{\infty} (1/k!) \sum_{\kappa} [C_\kappa(-\tfrac{1}{2}A) C_\kappa(nL)/C_\kappa(I_K)]$$

has been given by James [7], [8]. With S any $K \times K$ positive definite symmetric matrix the *zonal polynomial* $C_\kappa(S)$ is defined for each partition κ of k into not more than K parts as a certain symmetric polynomial in the latent roots of S. This series converges slowly unless the latent roots of the argument matrices are small. Thus another type of expansion for I_1 appears necessary if n is large.

The main result of this paper, given in Section 2, is the expansion

$$I_1 = 2^K \exp [(-n/2) \operatorname{tr} AL](2\pi/n)^{K(K-1)/4} \left(\prod_{i<j}^{K} c_{ij}\right)^{-\frac{1}{2}} \cdot F \tag{1.5}$$

where

$$F = 1 + (1/2n) \sum_{i<j}^{K} (1/c_{ij}) + (9/8n^2) \sum_{i<j}^{K} (1/c_{ij}^2) + (1/4n^2) \sum^{K} (1/c_{ij} c_{\kappa l}) + \cdots$$

and

$$c_{ij} = (a_j - a_i)(l_i - l_j) > 0.$$

The third sum in F is simply all possible cross products of the c_{ij}^{-1} without repetition. The last two terms of F are shown for $K = 4$ and conjectured for higher values of K. Together with (1.3) this yields the distribution of the latent roots for large n. This distribution will be compared with a result given by Girshick [3] which states that for large enough n the l_i are independent normal with mean and variance α_i and $\alpha_i^2(2/n)$ respectively.

The procedure used to find the expansion (1.5) requires that

$$\alpha_1 > \alpha_2 > \cdots > \alpha_K .$$

Thus this paper considers only the case when all the population roots α_i are different. A limiting distribution of the sample roots when the population roots have arbitrary multiplicity has been given by T. W. Anderson [1]. If all the multiplicities are unity (if the roots α_i are different) then this limiting distribution reduces to Girshick's result mentioned in the preceding paragraph.

The procedure used to find the expansion (1.5) is an extension of the well known methods sketched below for the case $K = 2$. Here, as is demonstrated in the next paragraph, I_1 comes to a multiple of a Bessel function. Let

$$O^{\pm}(2) = \{H \varepsilon O(2) \mid |H| = \pm 1\}.$$

Then

$$I_1 = 2\int_{O^+(2)} \exp[(-n/2) \operatorname{tr} AHLH'](H' dH).$$

Now let

$$H = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}, \qquad -\pi < \theta \leqq \pi,$$

so that $(H' dH) = h_1' dh_2 = d\theta$ and

$$(1.6) \quad I_1 = 4 \exp[(-n/2) \operatorname{tr} AL] \int_{-\pi/2}^{\pi/2} \exp[(-nc_{12}/4)(1 - \cos 2\theta)]\, d\theta.$$

Clearly the integrand has a maximum of unity at $\theta = 0$ and then decreases to $\exp(-nc_{12}/2)$ at $\theta = \pm\pi/2$. Laplace's method (Erdélyi [2]) is used to show that for n large enough, I_1 is approximately $4 \exp[(-n/2) \operatorname{tr} AL](2\pi/nc_{12})^{\frac{1}{2}}$. Also, however, there is an expansion associated with I_1. If we expand $\cos 2\theta$ in the usual power series then

$$I_1 = 4 \exp[(-n/2) \operatorname{tr} AL] \int_{-\pi/2}^{\pi/2} \exp(-nc_{12}\theta^2/2) [\exp(nc_{12}\theta^4/6 - nc_{12}\theta^6/45 + \cdots)]\, d\theta.$$

If the exponential in the brackets is expanded and the integration performed term by term then for large n the limits can be set to $\pm\infty$ since each integration is of the form $\int_{-\pi/2}^{\pi/2} \exp(-nc_{12}\theta^2/2)\theta^{2m}\, d\theta$ and most of this integral is given in a small neighborhood of $\theta = 0$. Thus for large n, I_1 is approximately

$$(1.7) \quad 4 \exp[(-n/2) \operatorname{tr} AL](2\pi/nc_{12})^{\frac{1}{2}}(1 + 1/2nc_{12} + 9/8n^2c_{12}^2 + \cdots).$$

Section 2 also contains a statement of a lemma proved by Hsu [5] which is an extension of Laplace's method. With K arbitrary it is used to show that, for large n, I_1 is approximately (1.5) with F unity.

Clearly the change $\theta = \Psi/2$ in (1.6) yields

$$I_1 = 4\pi \exp[(-n/2) \operatorname{tr} AL] \exp(-nc_{12}/4) I_0(nc_{12}/4)$$

where $I_0(z) = (1/\pi)\int_0^\pi \exp(z \cos\Psi)\, d\Psi$ is the imaginary Bessel function of the first kind. Thus (1.7) is essentially the well known asymptotic (z large) expansion

$$I_0(z) = (e^z/(2\pi z)^{\frac{1}{2}})(1 + 1/8z + 9/128z^2 + \cdots + (1^2 \cdot 3^2 \cdots (2m-1)^2)/(m!(8z)^m) + \cdots).$$

The precise definition of asymptotic expansion is given at the end of this section. Finally it should be noted that (1.3) and the above expression for I_1 in terms of I_0 yield the distribution of the sample roots when $K = 2$. Girshick [4] gave this distribution with a power series for the integral I_1. However he does not state here that the integral is a Bessel function or that his power series is the ordinary power series for $I_0(z)$. There is also no mention of an asymptotic expansion for the integral.

In Section 3 a numerical example is given using the asymptotic distribution for the latent roots resulting from the expansion given in Section 2. This distribution comes to

$$(1.8) \quad d\left(\prod_{i=1}^K l_i\right)^{(n-K-1)/2}\left(\prod_{i<j}^K (l_i - l_j)\right)^{\frac{1}{2}}\prod_{i=1}^K dl_i \cdot \left(\prod_{i=1}^K a_i\right)^{n/2}\left(\prod_{i<j}^K (a_j - a_i)\right)^{-\frac{1}{2}} \exp\left[(-n/2)\sum_{i=1}^K a_i l_i\right]\cdot F$$

where

$$d = (n/2)^{(nK/2-K(K-1)/4)}/\prod_{i=1}^K \Gamma[(n-K+i)/2]$$

and F is given in (1.5). The likelihood function is calculated and improved maximum likelihood estimates are given. The computations here are quite lengthy and the author is indebted to Mr. Jack Alanen of the Yale Computing Center for his excellent work in programming the IBM 709 computer for this example.

The definitions below are taken from Erdélyi [2].

DEFINITION. The sequence of functions $\{\phi_n(x)\}$ is an *asymptotic sequence* as $x \to \infty$, if for each n

$$\phi_{n+1}(x) = o(\phi_n(x)) \quad \text{as } x \to \infty.$$

Let $\{\phi_n\}$ be an asymptotic sequence.

DEFINITION. The (formal) series $\sum a_n\phi_n(x)$ is an *asymptotic expansion* to N terms of $f(x)$ as $x \to \infty$ if

$$f(x) = \sum^N a_n\phi_n(x) + o(\phi_N(x)) \quad \text{as } x \to \infty.$$

This is written $f(x) \sim \sum^N a_n\phi_n(x)$.

DEFINITION. The function $\phi(x)$ is an *asymptotic representation* for $f(x)$ as $x \to \infty$ if

$$f(x) = \phi(x) + o(\phi(x)) \quad \text{as } x \to \infty.$$

This is written $f(x) \sim \phi(x)$.

2. The expansion for I_1. In this section the integral

$$I_1 = \int_{O(K)} \exp\left[(-n/2) \text{ tr } AHLH'\right](H'\,dH)$$

will be considered in detail. In the asymptotic theory it is necessary to assume $0 < a_1 < a_2 < \cdots < a_K$ and $l_1 > l_2 > \cdots > l_K > 0$.

With these assumptions it is easy to show that the integrand

$$\exp\left[(-n/2) \text{ tr AHLH}'\right] = \exp\left[(-n/2)\sum_{i,j}^K a_i l_j h_{ij}^2\right]$$

has identical maximum values of $\exp\left[(-n/2) \text{ tr } AL\right]$ at each of the 2^K matrices of the form

$$\begin{bmatrix} \pm 1 & & & 0 \\ & \pm 1 & & \\ & & \ddots & \\ 0 & & & \pm 1 \end{bmatrix}.$$

For large n the integrand is negligible except for small neighborhoods about each of these matrices and I_1 consists of identical contributions from each of these neighborhoods so that

$$\begin{aligned} I_1 &= 2^K \int_{N(I)} \exp\,[(-n/2)\,\mathrm{tr}\,AHLH'](H'\,dH) \\ &= 2^K \mathcal{I} \end{aligned} \tag{2.1}$$

where $N(I)$ is a neighborhood of the identity matrix on the orthogonal manifold and as such consists only of proper orthogonal matrices.

Since any proper orthogonal matrix can be written as the exponential of a skew symmetric matrix we transform $\mathcal{I}$ under

$$H = \exp S, \qquad S \text{ a } K \times K \text{ skew symmetric matrix,} \tag{2.2}$$

so that $N(I) \to N(S = 0)$. (Note that "S" was also used as the sample covariance matrix in the introduction). It is not necessary to go further into the nature of $N(I)$ or $N(S = 0)$ since for large n we are going to approximate $\mathcal{I}$ by integrating not over exactly $N(S = 0)$ but simply over the intervals $-\infty < s_{ij} < \infty$ for each s_{ij}. This argument is given below following Equation (2.5).

We shall now calculate the Jacobian of the transformation $H = \exp S$ given in (2.2) above. Murnaghan [9] is used here but the essential idea was suggested to the author by A. T. James. For any proper orthogonal matrix H there is an orthogonal matrix H_1 such that, with $K = 2k + 1$,

$$H = H_1' \begin{bmatrix} \cos\theta_1 & -\sin\theta_1 & & & & 0 \\ \sin\theta_1 & \cos\theta_1 & & & & \\ & & \ddots & & & \\ & & & \cos\theta_k & -\sin\theta_k & 0 \\ & & & \sin\theta_k & \cos\theta_k & 0 \\ 0 & & & 0 & 0 & 1 \end{bmatrix} H_1,$$

where $-\pi < \theta_i \leqq \pi$, $i = 1, 2, \cdots, k$. The last row and column are deleted if $K = 2k$. With

$$\Theta = \begin{bmatrix} 0 & -\theta_1 & & & & 0 \\ \theta_1 & 0 & & & & \\ & & \ddots & & & \\ & & & 0 & -\theta_k & 0 \\ & & & \theta_k & 0 & 0 \\ 0 & & & 0 & 0 & 0 \end{bmatrix}$$

it is clear that

$$H = \exp(H_1'\Theta H_1).$$

From Murnaghan [9], pp. 230–235, with $K = 2k+1$,

$$(H'\,dH) = \prod_{i=1}^{k} 4\sin^2(\theta_i/2) g_1(\Theta) \prod_{i=1}^{k} d\theta_i((H_1'\,dH_1))$$

and, when $K = 2k$,

$$(H'\,dH) = g_1(\Theta) \prod_{i=1}^{k} d\theta_i((H_1'\,dH_1))$$

where

$$g_1(\Theta) = \prod_{i<j}^{k} \{16\sin^2[(\theta_i + \theta_j)/2]\sin^2[(\theta_i - \theta_j)/2]\}$$

and

$$((H_1'\,dH_1)) = \prod_{i<j,\, i\neq j-1,\, j \text{ even}}^{K} (h_i^1)'\,dh_j^1$$

with h_i^1 and dh_j^1 the ith and jth columns of H_1 and dH_1 respectively. This procedure can also be used with the transformation

$$S = H_1'\Theta H_1$$

and easily shows that with $K = 2k + 1$

$$\prod_{i<j}^{k} ds_{ij} = \prod_{i=1}^{k} \theta_i^2 g_2(\Theta) \prod_{i=1}^{k} d\theta_i((H_1'\,dH_1))$$

and, when $K = 2k$,

$$\prod_{i<j}^{K} ds_{ij} = g_2(\Theta) \prod_{i=1}^{k} d\theta_i((H_1'\,dH_1))$$

where

$$g_2(\Theta) = \prod_{i<j}^{k} (\theta_i + \theta_j)^2(\theta_i - \theta_j)^2.$$

Thus with $K = 2k + 1$ it is clear that

$$(H'\,dH) = \prod_{i=1}^{k} [(\sin\theta_i/2)/(\theta_i/2)]^2 f(\Theta) \prod_{i<j}^{K} ds_{ij},$$

and, when $K = 2k$,

$$(H'\,dH) = f(\Theta) \prod_{i<j}^{K} ds_{ij},$$

where

$$f(\Theta) = \prod_{i<j}^{k} \{[\sin((\theta_i+\theta_j)/2)/((\theta_i+\theta_j)/2)][\sin((\theta_i-\theta_j)/2)/((\theta_i-\theta_j)/2)]\}^2.$$

Since $S = H_1'\Theta H_1$, we have $\operatorname{tr} S^{2m} = (-1)^m(2)\sum_{i=1}^{k}\theta_i^{2m}$ so that for K even or odd $(H'\,dH) \to J\prod_{i<j}^{K} ds_{ij}$ where

$$(2.3)\quad J = 1 + [(K-2)/24]\operatorname{tr} S^2 + [(8-K)/4(6!)]\operatorname{tr} S^4 + [(5K^2 - 20K + 14)/8(6!)](\operatorname{tr} S^2)^2 + \cdots.$$

Direct substitution of (2.2) yields

$$\operatorname{tr} AHLH' = \operatorname{tr} AL + \operatorname{tr}(ALS^2 - ASLS) + \operatorname{tr}(ASLS^2)$$

$$(2.4) \qquad + \operatorname{tr}(ALS^4/12 - ASLS^3/3 + AS^2LS^2/4)$$

$$+ \operatorname{tr}(ASLS^4/12 + AS^3LS^2/6)$$

$$+ \operatorname{tr}(ALS^6/360 - ASLS^5/60 + AS^2LS^4/24 - AS^3LS^3/36) + \cdots .$$

This is rewritten using curly brackets to define the expressions in parentheses so that

$$\operatorname{tr} AHLH' = \operatorname{tr} AL + \operatorname{tr}\{S^2\} + \operatorname{tr}\{S^3\} + \operatorname{tr}\{S^4\} + \operatorname{tr}\{S^5\} + \operatorname{tr}\{S^6\} + \cdots$$

Since S is skew symmetric

$$\operatorname{tr}\{S^2\} = \textstyle\sum_{i<j}^{K} c_{ij} s_{ij}^2$$

where c_{ij} is defined in (1.5).
Finally we can write

$$(2.5) \quad \mathcal{J} = \exp[(-n/2)\operatorname{tr} AL] \int \cdots \int_{N(S=0)} \exp[(-n/2) \textstyle\sum_{i<j}^{K} c_{ij} s_{ij}^2]$$

$$\cdot \exp[(-n/2)(\operatorname{tr}\{S^3\} + \operatorname{tr}\{S^4\} + \cdots)] J \textstyle\prod_{i<j}^{K} ds_{ij} .$$

If this integration is to be performed term by term on the expansion of $e^{[\;]}J$ then for large n the limits for each s_{ij} can be put to $\pm\infty$ since each integration is of the form

$$\int \cdots \int_{N(S=0)} \exp[(-n/2) \textstyle\sum_{i<j}^{K} c_{ij} s_{ij}^2] \prod_{i<j}^{K} s_{ij}^{m_{ij}} \prod_{i<j}^{K} ds_{ij} ,$$

and most of this integral is given in a small neighborhood of $S = 0$. The m_{ij} are positive even integers or zero since any term containing an odd power of an s_{ij} will integrate to zero. Since

$$(2.6a) \quad \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp[(-n/2) \textstyle\sum_{i<j}^{K} c_{ij} s_{ij}^2] \prod_{i<j}^{K} ds_{ij} = \prod_{i<j}^{K} (2\pi/nc_{ij})^{\frac{1}{2}} = C$$

and

$$(2.6b) \quad \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp[(-n/2) \textstyle\sum_{i<j}^{K} c_{ij} s_{ij}^2] s_{kl}^{2m} \prod_{i<j}^{K} ds_{ij}$$

$$= C \cdot [1 \cdot 3 \cdot 5 \cdots (2m - 1)/(nc_{kl})^m]$$

it is clear that after integration each of the terms in the expansion of $e^{[\;]}J$ is of the form $M(A, L)(C/n^m)$ where $M(A, L)$ depends on the a_i and l_j and $m = 0, 1, 2, \cdots$. Finally we expand $e^{[\;]}J$, writing the terms in groups, each group corresponding to a certain value of m. We have

$$(2.7a) \quad e^{[\;]}J = 1 + (-n/2)\operatorname{tr}\{S^4\} + (n^2/8)(\operatorname{tr}\{S^3\})^2 + ((K-2)/24)\operatorname{tr} S^2$$

$$+ (-n/2)\operatorname{tr}\{S^6\} + (n^2/8)(\operatorname{tr}\{S^4\})^2 + (n^2/4)\operatorname{tr}\{S^3\}\operatorname{tr}\{S^5\}$$

$$(2.7b) \quad (-n^3/16)(\operatorname{tr}\{S^3\})^2 \operatorname{tr}\{S^4\} + [n^4/16(4!)](\operatorname{tr}\{S^3\})^4$$

$$+ (n(2-K)/48)\operatorname{tr}\{S^4\}\operatorname{tr} S^2 + [n^2(K-2)/8(4!)](\operatorname{tr}\{S^3\})^2 \operatorname{tr} S^2$$

$$+ [(8-K)/4(6!)]\operatorname{tr} S^4 + [(5K^2 - 20K + 14)/8(6!)](\operatorname{tr} S^2)^2$$

$$+ \cdots .$$

The three theorems below yield the expansion (1.5) for $\mathcal{J}$.

THEOREM 1. *Let A and L be diagonal matrices with $0 < a_1 < a_2 < \cdots < a_K$ and $l_1 > l_2 > \cdots > l_K > 0$. Then for large n the first two terms in the expansion for $\mathcal{J}$ are given by*

$$(2.8) \quad \mathcal{J} = \exp [(-n/2) \operatorname{tr} AL] \prod_{i<j}^{K} (2\pi/nc_{ij})^{\frac{1}{2}} \cdot (1 + (1/2n) \sum_{i<j}^{K} (1/c_{ij}) + \cdots).$$

PROOF. Application of (2.6a) shows that the one in $e^{[\ ]}J$ yields C after integration. We must write out and integrate the three terms in (2.7a). We include only terms without an odd power of an s_{ij} and do not write the C which appears with each term after integration. The integrations all follow Formula (2.6b).

From (2.4)

$$12 \operatorname{tr} \{S^4\} = \sum_{i,j,k,l}^{K} s_{ij}s_{jk}s_{kl}s_{li}f(i, j, k)$$

where $f(i, j, k) = a_i(l_i - 4l_j + 3l_k)$. In detail we have

$$\begin{aligned} 12 \operatorname{tr} \{S^4\} &= \sum_{i,j,k,k\neq i}^{K} s_{ij}^2 s_{jk}^2 (f(i, j, k) + f(j, i, j)) + \sum_{i,j}^{K} s_{ij}^4 f(i, j, i) \\ &= \sum_{i<j<k}^{K} [s_{ij}^2 s_{jk}^2 (g(i, j, k) + g(k, j, i)) \\ &\quad + s_{ij}^2 s_{ik}^2 (g(k, i, j) + g(j, i, k)) \\ &\quad + s_{ik}^2 s_{jk}^2 (g(j, k, i) + g(i, k, j)] + \sum_{i<j}^{K} (-4c_{ij}) s_{ij}^4 \end{aligned}$$

where $g(i, j, k) = f(i, j, k) + f(j, i, j)$. But

$$g(i, j, k) + g(k, j, i) = -4c_{ij} - 4c_{jk} + 3c_{ik}$$

so that after term by term integration $12 \operatorname{tr} \{S^4\}$ contributes

$$\begin{aligned} &-(8/n^2) \sum_{i<j<k}^{K} (1/c_{ij} + 1/c_{ik} + 1/c_{jk}) \\ &\quad + (3/n^2) \sum_{i<j<k}^{K} (c_{ik}/c_{ij}c_{jk} + c_{jk}/c_{ij}c_{ik} + c_{ij}/c_{ik}c_{jk}) - (12/n^2) \sum_{i<j}^{K} 1/c_{ij}. \end{aligned}$$

Denote the second expression by S_{11}^1 and the third by S_1. Since the first sum can be written

$$\sum_{k<i<j}^{K} (1/c_{ij}) + \sum_{i<k<j}^{K} (1/c_{ij}) + \sum_{i<j<k}^{K} (1/c_{ij}) = (K - 2) \sum_{i<j}^{K} 1/c_{ij}$$

we have finally that $(-n/2) \operatorname{tr} \{S^4\}$ contributes

$$(2.9) \quad [(K - 2)/3n]S_1 - (1/8n)S_{11}^1 + (1/2n)S_1.$$

Again from (2.4)

$$\operatorname{tr} \{S^3\} = \sum_{i<j<k}^{K} D(i, j, k) s_{ij}s_{jk}s_{ki}$$

where

$$D(i, j, k) = a_i(l_k - l_j) + a_j(l_i - l_k) + a_k(l_j - l_i).$$

It is not difficult to show that

$$D^2(i, j, k) = c_{ij}^2 + c_{ik}^2 + c_{jk}^2 - 2(c_{ij}c_{ik} + c_{ij}c_{jk} + c_{ik}c_{jk})$$

so that after integration $(\text{tr}\,\{S^3\})^2$ yields

$$(1/n^3) \textstyle\sum_{i<j<k}^{K} (c_{ik}/c_{ij}c_{jk} + c_{jk}/c_{ij}c_{ik} + c_{ij}/c_{ik}c_{jk}) - (2/n^3) \textstyle\sum_{i<j<k}^{K} (1/c_{ij} + 1/c_{ik} + 1/c_{jk}).$$

Thus from $(n^2/8)(\text{tr}\,\{S^3\})^2$ we have

$$(1/8n)S_{11}^{1} - [(K-2)/4n]S_1. \tag{2.10}$$

Finally since

$$\text{tr}\, S^2 = -2 \textstyle\sum_{i<j}^{K} s_{ij}^2$$

it is clear that $((K-2)/24)$ tr S^2 contributes

$$-[(K-2)/12n]S_1. \tag{2.11}$$

The results in (2.9), (2.10), and (2.11) add to $S_1/2n$ so that the proof of the theorem is complete.

No proof has been given to show that (2.8) is an asymptotic expansion for $\mathcal{J}$. An extension of Laplace's method can be applied here to show that at least we have an asymptotic representation for $\mathcal{J}$. This extension is given by Hsu [5] in the lemma below.

LEMMA (Hsu). *Let $\phi(u_1, \cdots, u_m)$ and $f(u_1, \cdots, u_m)$ be real functions on an m-dimensional closed domain D such that*

1. *$f > 0$ on D.*
2. *$\phi(f)^n$ is absolutely integrable over D, $n = 0, 1, 2, \cdots$.*
3. *All partial derivatives f_{u_i} and $f_{u_iu_j}$ exist and are continuous, $i, j = 1, 2, \cdots, m$.*
4. *$f(u)$ has an absolutely maximum value at an interior point ξ of D, so that all $f_{u_i}(\xi) = 0$, and $|-f_{u_iu_j}(\xi)| > 0$.*
5. *ϕ is continuous at ξ and $\phi(\xi) \neq 0$. Then for n large*

$$\int_D \phi(f)^n\, du_1 \cdots du_m \sim [\phi(\xi)(f(\xi))^n/(\Delta(\xi_1, \cdots, \xi_m))^{\frac{1}{2}}](2\pi/n)^{m/2}$$

where

$$f(u) = e^{\psi(u)}$$

and

$$\Delta(u_1, \cdots, u_m) = |-\psi_{u_iu_j}|.$$

This lemma is used to prove that we have an asymptotic representation for $\mathcal{J}$.

THEOREM 2. *Under the conditions of Theorem* 1

$$\mathcal{J} \sim \exp\,[(-n/2)\,\text{tr}\,AL] \textstyle\prod_{i<j}^{K} (2\pi/nc_{ij})^{\frac{1}{2}}.$$

PROOF. Apply the lemma directly to $\mathcal{J}$ after making the substitution $H = \exp S$. We have

$$\mathcal{J} = \int_{N(S=0)} \{\exp\,[(-\tfrac{1}{2}) \textstyle\sum_{i,j}^{K} a_i l_j h_{ij}^2]\}^n (1 + [(K-2)/24]\,\text{tr}\, S^2 + \cdots) \textstyle\prod_{i<j}^{K} ds_{ij}$$

so that

$$f = \exp [(-\tfrac{1}{2}) \sum_{i,j}^{K} a_i l_j h_{ij}^2],$$
$$\phi = 1 + [(K-2)/24] \operatorname{tr} S^2 + \cdots,$$
$$\psi = (-\tfrac{1}{2}) \sum_{i,j} a_i l_j h_{ij}^2,$$

and

$$D = N(S = 0).$$

Also ξ corresponds to the point $S = 0$ and f^n and hence f have just the single maximum point $S = 0$ in $N(S = 0)$. From the form of ψ it is clear that the conditions of the lemma are satisfied. Also

$$-(\partial^2\psi/\partial s_{mn}^2) = \sum_{i,j}^{K} a_i l_j h_{ij}(\partial^2 h_{ij}/\partial s_{mn}^2) + \sum_{i,j}^{K} a_i l_j (\partial h_{ij}/\partial s_{mn})^2$$

and

$$-(\partial^2\psi/\partial s_{m_2n_2}\partial s_{m_1n_1}) = \sum_{i,j}^{K} a_i l_j h_{ij}(\partial^2 h_{ij}/\partial s_{m_2n_2}\partial s_{m_1n_1}) + \sum_{i,j}^{K} a_i l_j(\partial h_{ij}/\partial s_{m_2n_2})(\partial h_{ij}/\partial s_{m_1n_1})$$

so that to find $|-\psi_{u_iu_j}(S = 0)|$ we are to differentiate the elements of $H = \exp S$ at most twice and then set each s_{ij} to zero. Hence to find this determinant we can essentially use

$$H = S + S^2/2.$$

Finally, it follows easily that

$$-\partial^2\psi/\partial s_{mn}^2 = c_{mn}; \qquad \partial^2\psi/\partial s_{m_2n_2}\partial s_{m_1n_1} = \partial\psi/\partial s_{mn} = 0,$$

and the lemma shows

$$\mathcal{J} \sim \exp [(-n/2) \operatorname{tr} AL](2\pi/n)^{K(K-1)/4}(\textstyle\prod_{i<j}^{K} c_{ij})^{-\frac{1}{2}}.$$

This completes the proof of Theorem 2.

From the relative simplicity of the proof of Theorem 1 one might think that a similar proof using (2.7b) would quickly yield the next term in the expansion (2.8) for $\mathcal{J}$. However it is extremely lengthy to write out some of the expressions in (2.7b) when K is arbitrary, especially when $\operatorname{tr}\{S^3\}$ or $\operatorname{tr}\{S^4\}$ is raised to a power. The author has worked out the details of this both for $K = 3$ and $K = 4$. Of course the latter is many times more difficult. In both cases the result is

$$(9/8n^2) \sum_{i<j}^{K} (1/c_{ij}^2) + (1/4n^2) \sum^{K} (1/c_{ij}c_{kl}).$$

Certainly this indicates strongly that the result holds for arbitrary K. We now state this result precisely for $K = 4$.

THEOREM 3. *Under the conditions of Theorem* 1, *but with* $K = 4$, *the term in* n^{-2} *in the expansion* (2.8) *for* $\mathcal{J}$ *is given by*

$$(9/8n^2) \sum_{i<j}^{4} (1/c_{ij}^2) + (1/4n^2) \sum^{4} (1/c_{ij}c_{kl})$$

where the second sum contains all (fifteen) possible cross products of the $(c_{ij})^{-1}$ *without repetition.*

PROOF. Due to an enormous amount of detail necessary for this proof only an outline is given and this is put in the appendix.

There is a conjecture that can be given here from the results of Theorems 1 and 3. From (1.6) and (1.7) it is clear that when $K = 2$ the general term of the expansion for I_1 is given by $(1^2 \cdot 3^2 \cdots (2m-1)^2)/m!(2nc_{12})^m$. It appears that in the expansion (1.5) for I_1 with K arbitrary the term in $\sum_{i<j}^K (1/c_{ij}^m)$ is given by

$$[(1^2 \cdot 3^2 \cdots (2m-1)^2/(2n)^m m!] \sum_{i<j}^K (1/c_{ij}^m). \tag{2.13}$$

In the numerical example given in Section 3 the cross product term is much smaller than (2.13) with $m = 2$ so that this conjecture could be of considerable importance.

We close this section with a precise statement of a theorem concerning the distribution of the latent roots proved by Girshick [3] and with a comparison of this with the distribution (1.8). Using the notation of the introduction we have

THEOREM (Girshick). *Let* $\alpha_1, \alpha_2, \cdots, \alpha_t$ *be any set of simple nonvanishing roots of*

$$|\Sigma - \alpha I| = 0.$$

For sufficiently large samples these will be approximated by certain of the latent roots $l_1, l_2, \cdots, l_t$ *of the samples. If* $l_i - \alpha_i$ *is divided by the standard error,* $\alpha_i(2/n)^{\frac{1}{2}}$ *then the resulting variates have a distribution which, as n increases, approaches the normal distribuion of t independent variates of zero mean and unit standard deviation.*

In the introduction Σ is taken non-singular so that $t = K$. From (1.8) the distribution of the latent roots can be written

$$M(A) \prod_{i=1}^K [l_i^{(n-K-1)/2} \exp(-nl_i/2\alpha_i)] \prod_{i<j}^K (l_i - l_j)^{\frac{1}{2}} F \prod_{i=1}^K dl_i \tag{2.14}$$

where $M(A)$ depends on the a_i but not on the l_i. For n large enough (2.14) will now be shown to agree with Girshick's theorem.

Suppose we can assume that

$$T = \prod_{i<j}^K (l_i - l_j)^{\frac{1}{2}}$$

and F have little effect on the distribution (2.14). Then the l_i are independent and each l_i has the same distribution as

$$\alpha_i \chi^2 / n$$

where χ^2 has the chi-square distribution on $n - K + 1$ degrees of freedom. But for large f, χ^2 on f degrees of freedom can be approximated by a normal variate with mean f and variance $2f$. Thus for large n, l_i has the normal distribution with mean

$$(\alpha_i/n)(n - K + 1) \cong \alpha_i$$

and variance

$$(\alpha_i^2/n^2)[2(n - K + 1)] \cong \alpha_i^2(2/n).$$

Now let us expand T to find the conditions necessary in order that T not depend on the l_i. We have

$$(l_i - l_j)^{\frac{1}{2}} = (\alpha_i - \alpha_j)^{\frac{1}{2}}[1 + ((\delta l_i - \delta l_j)/(\alpha_i - \alpha_j))]^{\frac{1}{2}}$$
$$= (\alpha_i - \alpha_j)^{\frac{1}{2}}(1 + \tfrac{1}{2}(\delta l_i - \delta l_j)/(\alpha_i - \alpha_j) + \cdots)$$

where $\delta l_i = l_i - \alpha_i$. Thus

$$T = \prod_{i<j}^{K} (\alpha_i - \alpha_j)^{\frac{1}{2}}(1 + \tfrac{1}{2} \sum_{i<j}^{K} [(\delta l_i - \delta l_j)/(\alpha_i - \alpha_j)] + \cdots).$$

Using $\alpha_i(2/n)^{\frac{1}{2}}$ for the standard deviation of l_i we can assume $|\delta l_i| < 2\alpha_i(2/n)^{\frac{1}{2}}$ so that

$$|\tfrac{1}{2} \sum_{i<j}^{K} [(\delta l_i - \delta l_j)/(\alpha_i - \alpha_j)]| < (2/n)^{\frac{1}{2}} \sum_{i<j}^{K} [(\alpha_i + \alpha_j)/(\alpha_i - \alpha_j)].$$

Thus we need n large enough to make this small so that T may be taken as $\prod_{i<j}^{K} (\alpha_i - \alpha_j)^{\frac{1}{2}}$.

Also we need n large enough to take F as unity. This does not require as large an n since F has the form

$$1 + n^{-1}(\tfrac{1}{2} \sum_{i<j}^{K} (1/c_{ij})) + \cdots .$$

All of this shows that Girshick's normal approximation follows from (2.14) provided n is large enough to take F as unity, T as not depending on the l_i, and finally to use the normal approximation to the χ^2 distribution on $n - K + 1$ degrees of freedom.

3. Numerical Example. The distribution (1.8) follows from the results of the theorems in Section 2. This distribution can be written

$$\text{(3.1)}\quad \prod_{i=1}^{K} [a_i^{n/2} \exp((-n/2)a_i l_i)](\prod_{i<j}^{K} (a_j - a_i))^{-\frac{1}{2}}F \cdot d(\prod_{i=1}^{K} l_i)^{(n-K-1)/2}(\prod_{i<j}^{K} (l_i - l_j))^{\frac{1}{2}}\prod_{i=1}^{K} dl_i,$$

where d is a constant given in (1.8) and

$$F = 1 + F_1 + F_2 + F_3 + \cdots$$

with

$$F_1 = (1/2n) \sum_{i<j}^{K} (1/c_{ij}),$$
$$F_2 = (9/8n^2) \sum_{i<j}^{K} (1/c_{ij}^2),$$

and

$$F_3 = (1/4n^2) \sum^{K} (1/c_{ij}c_{kl}).$$

Here again $c_{ij} = (a_j - a_i)(l_i - l_j) > 0$ and F_3 includes all possible cross products without repetition. As mentioned just before the statement of Theorem 3 the results for F_2 and F_3 have been proved only for $K = 3, 4$ and conjectured for arbitrary K.

Since (3.1) is a marginal distribution of the sample roots it appears that the corresponding likelihood function is appropriate for inference concerning the population roots provided either there is no prior information about the characteristic vectors or such information suggests a uniform distribution for the characteristic vectors.

Suppose that the latent roots l_i have been calculated. Consider the likelihood function

$$L(a_1, a_2, \cdots, a_K) = L_1 L_2 F \tag{3.2}$$

where

$$L_1 = \prod_{i=1}^{K} [a_i^{n/2} \exp((-n/2)a_i l_i)],$$

and

$$L_2 = (\prod_{i<j}^{K} (a_j - a_i))^{-\frac{1}{2}}.$$

Assume for the moment that L_2 and F have essentially no dependence on the a_i. From L_1 we have the maximum likelihood estimates $\hat{a}_i = 1/l_i$, $i = 1, 2, \cdots, K$. Clearly the function

$$a_i^{n/2} \exp((-n/2)a_i l_i) / \int_0^\infty x^{n/2} \exp((-n/2)x l_i)\, dx$$

has the same graph as the graph for $(1/nl_i)\chi^2$ where χ^2 has the chi-square distribution on $n + 2$ degrees of freedom. Thus the likelihood function of the a_i is similar to the distribution of a variate having mean $(1/nl_i)(n + 2) \cong 1/l_i = \hat{a}_i$, and variance $(1/n^2 l_i^2)[2(n + 2)] \cong 2/nl_i^2 = \sigma_i^2$. The foregoing heuristic considerations suggest that the set of intervals

$$\Delta a_i = (\hat{a}_i - 2\sigma_i, \hat{a}_i + 2\sigma_i), \qquad i = 1, 2, \cdots, K$$

would be a suitable region over which to study the likelihood function.

The function L_2 has an effect on the maximum likelihood estimates. We have

$$\begin{aligned} ln(a_j - a_i) &= ln(\hat{a}_j - \hat{a}_i) + ln(1 + (\delta a_j - \delta a_i)/(\hat{a}_j - \hat{a}_i)) \\ &= c(i, j) + (\delta a_j - \delta a_i)/(\hat{a}_j - \hat{a}_i) + \cdots \end{aligned}$$

where $c(i, j)$ is independent of the a_i and $\delta a_i = a_i - \hat{a}_i$. Thus

$$\begin{aligned} lnL_1L_2 = {} & \sum_{i=1}^{K} [(n/2)ln\, a_i - (n/2)a_i l_i] \\ & - \tfrac{1}{2}(\sum_{i<j}^{K} (\delta a_j - \delta a_i)/(\hat{a}_j - \hat{a}_i) + \cdots) - \tfrac{1}{2} \sum_{i<j}^{K} c(i, j). \end{aligned}$$

From partial differentiation with respect to a_i the corrected maximum likelihood estimates $\check{a}_i$ have the form

$$\begin{aligned} \check{a}_i &= (1/l_i)\{1/[1 + (1/nl_i)[\textstyle\sum_{j\neq i}^{K} (1/(\hat{a}_i - \hat{a}_j))] + \cdots]\} \\ &= (1/l_i) - (1/nl_i^2)[\textstyle\sum_{j\neq i}^{K} (1/(\hat{a}_i - \hat{a}_j))] + \cdots \\ &= \hat{a}_i - (\sigma_i^2/2) \textstyle\sum_{j\neq i}^{K} (1/(\hat{a}_i - \hat{a}_j)) + \cdots . \end{aligned} \tag{3.3}$$

Thus $\hat{a}_K$, the largest of the $\hat{a}_i$, is decreased, while $\hat{a}_1$, the smallest of the $\hat{a}_i$, is increased. The effect of L_2 is to move the maximum likelihood estimates closer together.

Consider the function F. The term $(\prod_{i<j}^{K} (l_i - l_j))^{\frac{1}{2}}$ in the distribution (3.1) for the latent roots indicates that the l_i will be reasonably well spaced. At $a_i = \hat{a}_i$, $a_j = \hat{a}_j$, we have

$$c_{ij}^{-1} = l_i l_j/(l_i - l_j)^2$$

so that for large n the functions F_1, F_2 and F_3 will be small as compared to unity. In most of the region defined above by the intervals Δa_i the F_i remain small. Thus in calculating L we take F as unity and examine the F_i separately.

Suppose that we have a sample of 181 observations from a normal 3-variate population such that the latent roots of the estimated covariance matrix are

$$l_1 = .37, \qquad l_2 = .23, \qquad l_3 = .15.$$

We have

$$\hat{a}_1 = 2.70, \qquad \hat{a}_2 = 4.35, \qquad \hat{a}_3 = 6.67$$
$$\sigma_1 = .285, \qquad \sigma_2 = .459, \qquad \sigma_3 = .703.$$

From (3.3) the improved maximum likelihood estimates are

$$\check{a}_1 = \hat{a}_1 + .035 = 2.74,$$
$$\check{a}_2 = \hat{a}_2 - .018 = 4.33,$$
$$\check{a}_3 = \hat{a}_3 - .169 = 6.50.$$

The correction for $\hat{a}_3$ is about $\sigma_3/4$ and cannot be assumed trivial.

An IBM 709 computer was used to evaluate lnL_1L_2, lnL_2, F_1, F_2, and F_3 at the set of points T given by

$$(3.4) \qquad \begin{aligned} a_1 &= 2.17 + .14m_1, & m_1 &= 0, 1, 2, \cdots, 8 \\ a_2 &= 3.41 + .23m_2, & m_2 &= 0, 1, 2, \cdots, 8 \\ a_3 &= 5.09 + .35m_3, & m_3 &= 1, 2, 3, \cdots, 8. \end{aligned}$$

For each value of a_3 in T we now reproduce square arrays for the function $ln(L_1L_2) - 116$. At $(\check{a}_1, \check{a}_2, \check{a}_3)$ this function has a maximum of 5.17. Curves for constant $ln(L_1L_2) - 116$ of two units below this maximum are sketched in. These curves yield the surface of constant likelihood in 3-space. See Table I.

For $a_3 = 6.49$ we also reproduce below arrays for lnL_2, F_1, F_2, and F_3. Certainly it is clear that L_2 does contribute essential information to the likelihood function. The results for F_1, F_2, and F_3 do not change appreciably for the other values of a_3 so that F can be taken as unity over most of the points of T to within 5% accuracy. See Table II.

TABLE I

a_2 \ a_1	2.17	2.31	2.45	2.59	2.73	2.87	3.01	3.15	3.29
3.41	0.77	0.27	1.00	1.44	1.64	1.62	1.41	1.09	0.76
3.64	0.32	1.35	2.02	2.49	2.66	2.61	2.37	1.95	1.40
3.87	1.07	2.09	2.80	3.21	3.37	3.30	3.03	2.58	1.97
4.10	1.52	2.54	3.24	3.65	3.79	3.71	3.43	2.96	2.30
4.33	1.71	2.71	3.42	3.82	3.96	3.87	3.58	3.09	2.44
4.56	1.67	2.68	3.37	3.77	3.91	3.81	3.51	3.02	2.35
4.79	1.44	2.45	3.14	3.53	3.67	3.57	3.26	2.76	2.09
5.02	1.08	2.09	2.77	3.16	3.29	3.19	2.88	2.37	1.69
5.25	0.71	1.72	2.40	2.79	2.91	2.81	2.49	1.98	1.30

$ln(L_1L_2) - 116, a_3 = 5.44$

a_2 \ a_1	2.17	2.31	2.45	2.59	2.73	2.87	3.01	3.15	3.29
3.41	0.02	1.03	1.75	2.19	2.38	2.36	2.16	1.83	1.50
3.64	1.06	2.10	2.81	3.23	3.40	3.35	3.10	2.68	2.12
3.87	1.80	2.83	3.53	3.94	4.10	4.03	3.75	3.29	2.68
4.10	2.24	3.26	3.95	4.36	4.50	4.42	4.13	3.65	3.01
4.33	2.40	3.42	4.11	4.51	4.65	4.56	4.24	3.77	3.11
4.56	2.34	3.35	4.04	4.43	4.57	4.47	4.16	3.66	3.00
4.79	2.06	3.07	3.76	4.15	4.28	4.18	3.86	3.36	2.68
5.02	1.61	2.62	3.30	3.69	3.82	3.71	3.39	2.89	2.20
5.25	1.02	2.03	2.71	3.09	3.22	3.11	2.79	2.28	1.59

$ln(L_1L_2) - 116, a_3 = 5.79$

a_2 \ a_1	2.17	2.31	2.45	2.59	2.73	2.87	3.01	3.15	3.29
3.41	0.43	1.47	2.19	2.63	2.82	2.79	2.59	2.26	1.92
3.64	1.50	2.53	3.24	3.66	3.83	3.77	3.52	3.10	2.54
3.87	2.23	3.25	3.96	4.36	4.52	4.44	4.16	3.71	3.09
4.10	2.66	3.67	4.37	4.77	4.91	4.83	4.53	4.06	3.41
4.33	2.81	3.82	4.51	4.91	5.05	4.95	4.65	4.16	3.50
4.56	2.72	3.73	4.42	4.81	4.95	4.84	4.53	4.03	3.36
4.79	2.42	3.43	4.12	4.50	4.63	4.53	4.21	3.77	3.03
5.02	1.93	2.94	3.62	4.01	4.13	4.02	3.70	3.19	2.51
5.25	1.28	2.29	2.96	3.35	3.47	3.36	3.04	2.52	1.83

$ln(L_1L_2) - 116, a_3 = 6.14$

a_2 \ a_1	2.17	2.31	2.45	2.59	2.73	2.87	3.01	3.15	3.29
3.41	0.59	1.63	2.35	2.78	2.97	2.95	2.74	2.41	2.06
3.64	1.66	2.69	3.39	3.81	3.98	3.92	3.66	3.24	2.68
3.87	2.38	3.40	4.10	4.51	4.66	4.58	4.30	3.84	3.22
4.10	2.80	3.81	4.51	4.91	5.05	4.96	4.67	4.19	3.54
4.33	2.94	3.96	4.64	5.04	5.17	5.08	4.77	4.28	3.62
4.56	2.84	3.86	4.54	4.93	5.06	5.96	4.65	4.14	3.47
4.79	2.53	3.54	4.22	4.61	4.73	4.63	4.31	3.80	3.12
5.02	2.02	3.03	3.70	4.09	4.21	4.10	3.78	3.27	2.58
5.25	1.34	2.34	3.02	3.40	3.52	3.41	3.08	2.56	1.87

$ln(L_1L_2) - 116, a_3 = 6.49$

a_2 \ a_1	2.17	2.31	2.45	2.59	2.73	2.87	3.01	3.15	3.29
3.41	0.50	1.54	2.26	2.69	2.88	2.85	2.64	2.30	1.96
3.64	1.56	2.59	3.30	3.71	3.88	3.82	3.56	3.13	2.57
3.87	2.28	3.30	4.00	4.41	4.56	4.48	4.20	3.73	3.11
4.10	2.69	3.71	4.40	4.80	4.94	4.85	4.55	4.07	3.42
4.33	2.83	3.84	4.53	4.92	5.06	4.96	4.65	4.16	3.49
4.56	2.72	3.73	4.42	4.81	4.94	4.83	4.52	4.01	3.34
4.79	2.40	3.41	4.08	4.47	4.60	4.49	4.17	3.66	2.97
5.02	1.88	2.88	3.56	3.94	4.06	3.95	3.63	3.11	2.42
5.25	1.18	2.18	2.85	3.23	3.35	3.24	2.91	2.39	1.70

$ln(L_1L_2) - 116,\ a_3 = 6.84$

a_2 \ a_1	2.17	2.31	2.45	2.59	2.73	2.87	3.01	3.15	3.29
3.41	0.18	1.22	1.93	2.37	2.55	2.52	2.31	1.98	1.63
3.64	1.24	2.27	2.97	3.39	3.55	3.49	3.23	2.80	2.24
3.87	1.95	2.98	3.67	4.08	4.23	4.15	3.86	3.40	2.78
4.10	2.36	3.38	4.07	4.47	4.61	4.51	4.22	3.73	3.08
4.33	2.49	3.51	4.19	4.58	4.72	4.62	4.31	3.81	3.14
4.56	2.38	3.39	4.07	4.46	4.59	4.48	4.17	3.66	2.98
4.79	2.05	3.06	3.73	4.12	4.24	4.13	3.81	3.30	2.61
5.02	1.52	2.52	3.20	3.58	3.70	3.59	3.26	2.74	2.05
5.25	0.81	1.81	2.48	2.86	2.98	2.86	2.53	2.01	1.32

$ln(L_1L_2) - 116,\ a_3 = 7.19$

a_2 \ a_1	2.17	2.31	2.45	2.59	2.73	2.87	3.01	3.15	3.29
3.41	−0.35	0.69	1.41	1.84	2.02	1.99	1.78	1.44	1.10
3.64	0.71	1.74	2.44	2.86	3.02	2.96	2.70	2.27	1.70
3.87	1.42	2.44	3.14	3.54	3.69	3.61	3.32	2.86	2.24
4.10	1.83	2.84	3.53	3.93	4.07	3.97	3.67	3.19	2.54
4.33	1.96	2.97	3.65	4.04	4.17	4.07	3.76	3.27	2.60
4.56	1.84	2.85	3.53	3.91	4.04	3.94	3.62	3.11	2.43
4.79	1.50	2.51	3.18	3.57	3.69	3.58	3.26	2.74	2.06
5.02	0.96	1.97	2.64	3.02	3.14	3.03	2.70	2.18	1.49
5.25	0.24	1.24	1.92	2.29	2.41	2.29	1.96	1.44	0.74

$ln(L_1L_2) - 116,\ a_3 = 7.54$

a_2 \ a_1	2.17	2.31	2.45	2.59	2.73	2.87	3.01	3.15	3.29
3.41	−1.06	−0.02	0.69	1.12	1.31	1.27	1.06	0.72	0.38
3.64	−0.00	1.02	1.72	2.14	2.30	2.24	1.98	1.55	0.98
3.87	0.70	1.72	2.42	2.82	2.97	2.89	2.60	2.13	1.51
4.10	1.11	2.12	2.81	3.20	3.34	3.25	2.95	2.46	1.81
4.33	1.23	2.24	2.92	3.32	3.45	3.35	3.03	2.53	1.86
4.56	1.11	2.12	2.80	3.18	3.31	3.20	2.88	2.38	1.70
4.79	0.77	1.77	2.45	2.83	2.95	2.84	2.52	2.00	1.31
5.02	0.23	1.23	1.90	2.28	2.40	2.28	1.96	1.44	0.74
5.25	−0.50	0.50	1.17	1.55	1.66	1.55	1.21	0.69	−0.01

$ln(L_1L_2) - 116,\ a_3 = 7.89$

TABLE II

a_2 \ a_1	2.17	2.31	2.45	2.59	2.73	2.87	3.01	3.15	3.29
3.41	1.40	1.32	1.24	1.14	1.03	0.90	0.73	0.49	0.08
3.64	1.45	1.38	1.31	1.23	1.14	1.04	0.92	0.77	0.58
3.87	1.48	1.42	1.36	1.29	1.21	1.12	1.03	0.92	0.79
4.10	1.50	1.44	1.38	1.32	1.26	1.18	1.10	1.01	0.91
4.33	1.50	1.45	1.40	1.34	1.28	1.22	1.15	1.07	0.99
4.56	1.50	1.45	1.40	1.35	1.29	1.23	1.17	1.10	1.03
4.79	1.48	1.43	1.39	1.34	1.29	1.23	1.18	1.16	1.05
5.02	1.45	1.41	1.36	1.32	1.27	1.22	1.17	1.11	1.05
5.25	1.40	1.36	1.32	1.28	1.23	1.18	1.13	1.08	1.03

$-\ln L_2$, $a_3 = 6.49$

a_2 \ a_1	2.17	2.31	2.45	2.59	2.73	2.87	3.01	3.15	3.29
3.41	.030	.032	.035	.039	.044	.052	.065	.091	.181
3.63	.029	.030	.032	.034	.037	.041	.047	.056	.073
3.87	.028	.029	.030	.032	.034	.037	.040	.045	.051
4.10	.028	.029	.030	.031	.032	.034	.036	.039	.043
4.33	.028	.029	.030	.031	.032	.033	.035	.037	.039
4.56	.029	.030	.031	.031	.032	.033	.034	.036	.038
4.79	.031	.031	.032	.033	.033	.034	.035	.036	.038
5.02	.034	.034	.034	.035	.036	.036	.037	.038	.039
5.25	.037	.038	.038	.039	.039	.040	.040	.041	.042

F_1 , $a_3 = 6.49$

a_2 \ a_1	2.17	2.31	2.45	2.59	2.73	2.87	3.01	3.15	3.29
3.41	.0018	.0021	.0025	.0033	.0045	.0067	.0117	.0268	.1237
3.64	.0015	.0017	.0020	.0023	.0028	.0037	.0052	.0081	.0152
3.87	.0014	.0016	.0017	.0019	.0022	.0026	.0032	.0043	.0061
4.10	.0015	.0015	.0016	.0018	.0019	.0022	.0025	.0030	.0037
4.33	.0016	.0016	.0017	.0018	.0019	.0020	.0022	.0025	.0029
4.56	.0018	.0018	.0019	.0020	.0020	.0021	.0023	.0024	.0026
4.79	.0022	.0022	.0022	.0023	.0023	.0024	.0025	.0026	.0027
5.02	.0028	.0028	.0028	.0029	.0029	.0029	.0030	.0031	.0032
5.25	.0038	.0038	.0038	.0038	.0039	.0039	.0039	.0040	.0041

F_2 , $a_3 = 6.49$

a_2 \ a_1	2.17	2.31	2.45	2.59	2.73	2.87	3.01	3.15	3.29
3.41	.0003	.0003	.0003	.0004	.0005	.0006	.0008	.0012	.0026
3.64	.0002	.0003	.0003	.0003	.0004	.0004	.0005	.0007	.0010
3.87	.0002	.0002	.0003	.0003	.0003	.0004	.0004	.0005	.0006
4.10	.0002	.0002	.0003	.0003	.0003	.0003	.0004	.0004	.0005
4.33	.0002	.0002	.0003	.0003	.0003	.0003	.0004	.0004	.0004
4.56	.0002	.0002	.0003	.0003	.0003	.0003	.0003	.0004	.0004
4.79	.0002	.0002	.0003	.0003	.0003	.0003	.0003	.0004	.0004
5.02	.0003	.0003	.0003	.0003	.0003	.0003	.0004	.0004	.0004
5.25	.0003	.0003	.0003	.0003	.0003	.0004	.0004	.0004	.0004

F_3 , $a_3 = 6.49$

4. Acknowledgment. The author expresses his gratitude to Dr. Alan T. James of Yale University, who suggested both the problem and a general method of attack.

APPENDIX

THEOREM 3. *Under the conditions of Theorem* 1, *but with* $K = 4$, *the term in* n^{-2} *in the expansion* (2.8) *for* $\mathcal{J}$ *is given by*

$$(9/8n^2) \textstyle\sum_{i<j}^4 (1/c_{ij}^2) + (1/4n^2) \sum^4 (1/c_{ij}c_{kl})$$

where the second sum contains all (fifteen) possible cross products of the $(c_{ij})^{-1}$ *without repetition.*

OUTLINE OF PROOF: As in the proof of Theorem 1 each of the expressions in (2.7b) is written out in detail as the sum of polynomials in the s_{ij} and any term with an odd power of an s_{ij} is dropped. Then each polynomial is integrated using (2.6b) and all of these results are combined to give a linear combination of the expressions S_2, S_{11}, $\cdots$, S_{1111}^{11} defined in (2.12) below. Since an $n^{-2}C$ is common to each of the terms in (2.7b) after integration we write

$$f(S) \doteq d_2S_2 + d_{11}S_{11} + \cdots + d_{1111}^{11}S_{1111}^{11}$$

if

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp\,[(-n/2) \textstyle\sum_{i<j} c_{ij}s_{ij}^2] f(S) \prod_{i<j}^4 ds_{ij} = n^{-2}C(d_2S_2 + \cdots + d_{1111}^{11}S_{1111}^{11}).$$

Here the *d*'s are constants. We now list without proof the results for each of the expressions in (2.7b).

$$(-n/2)\ \mathrm{tr}\,\{S^6\} \doteq (1/6!)[-432S_2 - 253S_{11} - 104S_{1-1} + 45S_{21}^1 + 15S_{111}^1 + 15\,\hat{S}_{111}^1 - 10\,\check{S}_{111}^1].$$

$$\begin{aligned}(n^2/8)(\mathrm{tr}\,\{S^4\})^2 \doteq{}& (135/32)S_2 + (359/144)S_{11} + (59/36)S_{1-1} \\ &+ (-11/16)S_{21}^1 + (9/128)S_{22}^2 + (-11/48)S_{111}^1 \\ &+ (-11/48)\,\hat{S}_{111}^1 + (13/288)\,\check{S}_{111}^1 + (1/32)S_{1111}^{(1+1)^2} \\ &+ (3/64)S_{211}^{11} + (1/64)S_{1111}^2 .\end{aligned}$$

$$\begin{aligned}(n^2/4)\ \mathrm{tr}\,\{S^3\}\ \mathrm{tr}\,\{S^5\} \doteq{}& (3/4)S_2 + (13/16)S_{11} + (1/2)S_{1-1} \\ &+ (-3/16)S_{21}^1 + (-5/48)S_{111}^1 + (-5/48)\,\hat{S}_{111}^1 \\ &+ (1/12)\,\check{S}_{111}^1 .\end{aligned}$$

$$\begin{aligned}(-n^3/16)(\mathrm{tr}\,\{S^3\})^2\ \mathrm{tr}\,\{S^4\} \doteq{}& (-61/16)S_2 + (-157/48)S_{11} + (-23/12)S_{1-1} \\ &+ (7/16)S_{111}^1 + (9/16)\,\hat{S}_{111}^1 + (-7/48)\,\check{S}_{111}^1 + (19/16)S_{21}^1 \\ &+ (-9/64)S_{22}^2 + (-3/32)S_{211}^{11} + (-1/32)S_{1111}^2 + (-3/16)S_{1111}^{(1+1)^2}.\end{aligned}$$

$$(n^4/384)(\operatorname{tr}\{S^3\})^4 \doteq (33/32)S_2 + (5/6)S_{11} + (17/24)S_{1-1} + (-5/24)S^1_{111} + (-1/4)\hat{S}^1_{111} + (11/96)\check{S}^1_{111} + (-3/8)S^1_{21} + (9/128)S^2_{22} + (3/64)S^{11}_{211} + (17/192)S^2_{1111} + (7/48)S^{11}_{1111}.$$

$$(-n/24)\operatorname{tr}\{S^4\}\operatorname{tr} S^2 \doteq (-3/4)S_2 + (-23/48)S_{11} + (-7/18)S_{1-1} + (1/16)S^1_{21} + (1/48)S^1_{111} + (1/48)\hat{S}^1_{111}.$$

$$(n^2/96)(\operatorname{tr}\{S^2\})^2\operatorname{tr} S^2 \doteq (1/4)S_2 + (13/48)S_{11} + (1/6)S_{1-1} + (-1/48)S^1_{111} + (-1/48)\hat{S}^1_{111} + (-1/16)S^1_{21}.$$

$$(1/6!)(\operatorname{tr} S^4 + (7/4)(\operatorname{tr} S^2)^2) \doteq (1/6!)[27S_2 + 18S_{11} + 14S_{1-1}].$$

It is clear that the integration of (2.7b) comes to

$$(9/8)S_2 + (1/3)S_{11} + (7/12)S_{1-1} + (-1/12)S^1_{111} + (1/12)\check{S}^1_{111} + (-1/12)S^{(1+1)^2}_{1111}.$$

However from (2.12) it is not difficult to show the dependence

$$S_{11} + 4S_{1-1} - S^1_{111} + \check{S}^1_{111} - S^{(1+1)^2}_{1111} = 0.$$

Finally then integration of the expressions in (2.7b) yields

$$(9/8)S_2 + (1/4)(S_{11} + S_{1-1}).$$

We complete the outline of this proof by defining the expressions S_2, S_{11}, $\cdots$, S^{11}_{1111} which appear after the integration of (2.7b).

FORMULA (2.12).

$$S_2 = 1/c^2_{12} + 1/c^2_{13} + 1/c^2_{14} + 1/c^2_{23} + 1/c^2_{24} + 1/c^2_{34},$$

$$S_{11} = 1/c_{12}c_{13} + 1/c_{12}c_{14} + 1/c_{12}c_{23} + 1/c_{12}c_{24} + 1/c_{13}c_{14} + 1/c_{13}c_{23} + 1/c_{13}c_{34} + 1/c_{14}c_{24} + 1/c_{14}c_{34} + 1/c_{23}c_{24} + 1/c_{23}c_{34} + 1/c_{24}c_{34},$$

$$S_{1-1} = 1/c_{12}c_{34} + 1/c_{13}c_{24} + 1/c_{14}c_{23},$$

$$S^2_{22} = c^2_{12}(1/c^2_{13}c^2_{23} + 1/c^2_{14}c^2_{24}) + c^2_{13}(1/c^2_{12}c^2_{23} + 1/c^2_{14}c^2_{34}) + c^2_{14}(1/c^2_{12}c^2_{24} + 1/c^2_{13}c^2_{34}) + c^2_{23}(1/c^2_{12}c^2_{13} + 1/c^2_{24}c^2_{34}) + c^2_{24}(1/c^2_{12}c^2_{14} + 1/c^2_{23}c^2_{34}) + c^2_{34}(1/c^2_{13}c^2_{14} + 1/c^2_{23}c^2_{24}),$$

$$S^1_{21} = c_{12}(1/c_{13}c^2_{23} + 1/c^2_{13}c_{23} + 1/c_{14}c^2_{24} + 1/c^2_{14}c_{24}) + c_{13}(1/c_{12}c^2_{23} + 1/c^2_{12}c_{23} + 1/c_{14}c^2_{34} + 1/c^2_{14}c_{34}) + c_{14}(1/c_{12}c^2_{24} + 1/c^2_{12}c_{24} + 1/c_{13}c^2_{34} + 1/c^2_{13}c_{34}) + c_{23}(1/c_{12}c^2_{13} + 1/c^2_{12}c_{13} + 1/c_{24}c^2_{24} + 1/c^2_{34}c_{34})$$

$$+ c_{24}(1/c_{12}c_{14}^2 + 1/c_{12}^2c_{14} + 1/c_{23}c_{34}^2 + 1/c_{23}^2c_{34})$$

$$+ c_{34}(1/c_{13}c_{14}^2 + 1/c_{13}^2c_{14} + 1/c_{23}c_{24}^2 + 1/c_{23}^2c_{24}),$$

$$S^1_{111} = (c_{23} + c_{24} + c_{34})/c_{12}c_{13}c_{14} + (c_{13} + c_{14} + c_{34})/c_{12}c_{23}c_{24}$$

$$+ (c_{12} + c_{14} + c_{24})/c_{13}c_{23}c_{34} + (c_{12} + c_{13} + c_{23})/c_{14}c_{24}c_{34},$$

$$\hat{S}^1_{111} = (c_{14} + c_{23})[1/c_{12}c_{13}c_{34} + 1/c_{12}c_{13}c_{24} + 1/c_{12}c_{24}c_{34} + 1/c_{13}c_{24}c_{34}]$$

$$+ (c_{13} + c_{24})[1/c_{12}c_{14}c_{34} + 1/c_{12}c_{14}c_{23} + 1/c_{12}c_{23}c_{34} + 1/c_{14}c_{23}c_{34}]$$

$$+ (c_{12} + c_{34})[1/c_{13}c_{14}c_{24} + 1/c_{13}c_{14}c_{23} + 1/c_{14}c_{23}c_{24} + 1/c_{13}c_{23}c_{24}],$$

$$\check{S}^1_{111} = (c_{12}/c_{34} + c_{34}/c_{12})(1/c_{13}c_{24} + 1/c_{14}c_{23})$$

$$+ (c_{13}/c_{24} + c_{24}/c_{13})(1/c_{14}c_{23} + 1/c_{12}c_{34})$$

$$+ (c_{14}/c_{23} + c_{23}/c_{14})(1/c_{13}c_{24} + 1/c_{12}c_{34}),$$

$$S^{11}_{211} = (1/c_{12}^2 + 1/c_{34}^2)(c_{13}c_{14}/c_{23}c_{24} + c_{23}c_{24}/c_{13}c_{14} + c_{13}c_{24}/c_{14}c_{23} + c_{14}c_{23}/c_{13}c_{24})$$

$$+ (1/c_{13}^2 + 1/c_{24}^2)(c_{12}c_{14}/c_{23}c_{34} + c_{23}c_{34}/c_{12}c_{14} + c_{12}c_{34}/c_{14}c_{23} + c_{14}c_{23}/c_{12}c_{34})$$

$$+ (1/c_{14}^2 + 1/c_{23}^2)(c_{12}c_{13}/c_{24}c_{34} + c_{24}c_{34}/c_{12}c_{13} + c_{12}c_{34}/c_{13}c_{24} + c_{13}c_{24}/c_{12}c_{34}),$$

$$S^2_{1111} = (c_{12}^2 + c_{34}^2)/c_{13}c_{14}c_{23}c_{24} + (c_{13}^2 + c_{24}^2)/c_{12}c_{14}c_{23}c_{34} + (c_{14}^2 + c_{23}^2)/c_{12}c_{13}c_{24}c_{34},$$

$$S^{11}_{1111} = c_{12}c_{34}/c_{13}c_{14}c_{23}c_{24} + c_{13}c_{24}/c_{12}c_{14}c_{23}c_{34} + c_{14}c_{23}/c_{12}c_{13}c_{24}c_{34},$$

$$S^{(1+1)^2}_{1111} = S^2_{1111} + 2S^{11}_{1111}.$$

REFERENCES

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

[2] ERDÉLYI, A. (1956). *Asymptotic Expansions.* Dover, New York.

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

[4] GIRSHICK, M. A. (1941). The distribution of the ellipticity statistic L_ϵ when the hypothesis is false. *Terr. Mag.* **46** 455–457.

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

[6] JAMES, A. T. (1954). Normal multivariate analysis and the orthogonal group. *Ann. Math. Statist.* **25** 40–75.

[7] JAMES, A. T. (1960). The distribution of the latent roots of the covariance matrix. *Ann. Math. Statist.* **31** 151–158.

[8] JAMES, A. T. (1961). Zonal polynomials of the real positive definite symmetric matrices. *Ann. of Math.* **74** 456–459.

[9] MURNAGHAN, F. D. (1938). *The Theory of Group Representations.* The Johns Hopkins Press, Baltimore.