On An Asymptotic Representation of the Distribution of the Characteristic Roots of $S_1S_2^{-1}$ When Roots Are Not All Distinct

Ъy

Tseng C. Chang*

Department of Statistics

Division of Mathematical Sciences

Mimeograph Series No. 180

January, 1969

This research was supported by the National Science Foundation, Grant No. GP-7663.

On An Asymptotic Representation of the Distribution of the Characteristic Roots of $S_1S_2^{-1}$ When Roots Are Not All Distinct

bу

Tseng C. Chang*

Purdue University

Introduction and Summary

In the study of the distribution of the characteristic roots of $\Sigma_1\Sigma_2^{-1}$ where Σ_i : p x p (i = 1,2) are independently distributed as Wishart (n_i,p, Σ_i). We encounter the difficulty of evaluating the following integral [9]

(1.1)
$$E = \int_{O(p)} \left| \frac{1}{n} + \underbrace{AHLH}' \right|^{-\frac{n}{1+n}2} \left(\underbrace{H'dH} \right)$$

and characteristic roots of $\sum_{l}\sum_{2}^{-l}$ and $\{\sum_{l}\sum_{2}^{-l}\}^{-l}$ respectively, and $(\underline{H}'d\underline{H})$ is an invariant measure on the group O(p) of $p \times p$ orthogonal matrices as discussed in detail in [6], [7] and [8].

An asymptotic form of (1.1) when $n = n_1 + n_2$ is large and all a_i 's (i = 1, 2, ..., p) are distinct has been derived in [3]. If A = aI, i.e., $a_1 = a_2 = ... = a_p = a$ it can be shown easily that the integrand of (1.1) is

^{*} This research was supported by the National Science Foundation Grant No. GP-7663.

independent of H and has the form

(1.2)
$$E = \left\{ 2^{p} \prod_{i=1}^{\frac{1}{4}p(p+1)} / \prod_{i=1}^{p} \Gamma(\frac{i}{2}) \right\} \prod_{i=1}^{p} \left(1 + a \ell_{i} \right)^{-\frac{n}{2}}$$

In particular, when A = I, i.e., $a_1 = a_2 = \dots = a_p = 1$, we can obtain the so-called central distribution of the characteristic roots of $S_1 S_2^{-1}$ [10]. Note (1.2) is an exact form where we assume no asymptotic condition.

However, the roots need not all be equal. And when we are interested in the likelihood of equality of population roots, the asymptotic formula of the distinct roots case blows up (see [3]). A reasonable method to conjecture a new formula for that case is to remove the factor in the asymptotic formula which blows up and to insert the exact formula for the group containing equal roots in such a way that when the exact formula for the group is replaced by its asymptotic expansion we shall recover the original asymptotic expansion of all the variables, at least up to the first order.

With this motivation an attempt is made in this paper to give an asymptotic distribution of the characteristic roots of $\sum_{l}\sum_{2}^{-l}$ under the hypothesis, $0 < a_1 < \ldots < a_k < a_{k+1} = a_{k+2} = \ldots = a_p (1 \le k \le p-1)$ while ℓ_i 's (i = 1,2,...,p) are assumed to be distinct.

2. Asymptotic Representation When All a 's Are Equal Except an Extreme One

The procedure used to find the asymptotic representation of (1.1) when $0 < a_1 < a_2 < \ldots < a_k < a_{k+1} = \ldots = a_p$ and $\ell_1 > \ell_2 > \ldots > \ell_p > 0$ is an extension of the elementary method sketched below for the case when k = 1. We first derive in the following a Jacobian which enables us to study through a random orthogonal matrix instead of an orthogonal group.

Lemma (2.1). If f(H) is any integrable function of $H \in O(p)$ and dH is the normalized Haar measure, i.e., the integral will be one when taken over the whole group, and A and H are defined as in (1.1) then

(2.1)
$$\int_{O(p)} f(H)(dH) = \frac{2^{p} \prod_{i=1}^{p} \Gamma(\frac{k}{2})}{\prod_{i=1}^{p} (p+1)} \int_{\{HH: i=1\}} f(H)|J| \prod_{i>j} dh_{i,j}$$

where

(2.2)
$$J = \left\{ h_{pp} \cdot \begin{vmatrix} h_{p-1,p-1} & h_{p,p-1} \\ h_{p-1,p} & h_{p,p} \end{vmatrix} \cdot \cdot \cdot \begin{vmatrix} h_{2p} & h_{3p} & \cdots & h_{pp} \\ h_{2p} & h_{3p} & \cdots & h_{pp} \end{vmatrix} \right\}^{-1}$$

(Note that the choice of the set of p(p-1)/2 random variables on the right hand side of (2.1) is completely arbitrary.)

Proof: Let f(S) be the Wishart distribution with n degrees of freedom

(2.3)
$$\mathbf{f}(\mathbf{S}) = \frac{\left(\frac{\mathbf{n}}{2}\right)^{\frac{\mathbf{n}}{2}} \exp\left\{-\frac{\mathbf{n}}{2} \sum^{-1} \mathbf{S}\right\} \left|\mathbf{S}\right|^{\frac{\mathbf{n}-\mathbf{p}-1}{2}}}{\prod^{\frac{1}{\mathbf{n}}\mathbf{p}(\mathbf{p}-1)} \left|\sum^{\frac{\mathbf{n}}{2}} \prod^{\mathbf{p}} \Gamma\left(\frac{\mathbf{n}-\mathbf{i}+1}{2}\right)\right|}$$

As usual our objective is to find the joint distribution of the characteristic roots ℓ_i 's $(i=1,\ldots,p)$ of S, $(\ell_1 \geq \ell_2 \geq \ldots \geq \ell_p \geq 0)$. Without loss of generality we may write the exponential part in (2.3) as $\exp\{-\frac{n}{2} \text{ tr BS}\}$ since $\Sigma^{-1} = H_1 \stackrel{B}{\sim} H_1'$ where H_1 is a fixed orthogonal matrix and the characteristic roots of S is invariant under any orthogonal transformation. Now consider H as a complete random matrix with random variables $h_{ij}(i > j)$ such that

$$\begin{array}{ccc} S &=& H L H' \\ \sim & \sim \sim \sim \end{array}$$

For economy of notation let

$$\overset{S*}{\sim} = \overset{dS}{\sim} = (\overset{d}{s_{ij}}),$$

the differential of S. Then from (2.4) we have

or

$$(2.6) \qquad \qquad \underset{\sim}{\text{H'}} \text{ S* H} = \underset{\sim}{\text{H'}} \text{H* L} + \underset{\sim}{\text{L*}} + \underset{\sim}{\text{L}} \text{H'** H} .$$

It is known that H'H* is a p x p skew symmetric matrix (denote it as T) \sim then

$$H'H^{\pm} = T$$

or

$$(2.8) \qquad \qquad \underset{\sim}{\text{H*}^*\text{H}} = -\text{T} \quad .$$

Application of (2.7) to (2.6) yields

$$(2.9) \qquad \qquad H^{\dagger}S*H = TL + L* - LT \qquad .$$

Let
$$W = H'S*H$$

then

$$(2.10) \qquad \qquad \underset{\sim}{\mathbb{W}} = \underset{\sim}{\mathbb{TL}} + \underset{\sim}{\mathbb{L}} * - \underset{\sim}{\mathbb{L}} T \qquad .$$

Now by the property of conditional probability we can show that

$$(2.11) J(S;H,L) = J(S*;H*,L*) = J(S*;W)J(W;TL*)J(T;H*)$$

We know $J(\underbrace{S^*;W}) = 1$ since it is an orthogonal transformation. Also $J(\underbrace{W}; \underbrace{T,L^*}) = \underbrace{\Pi} (\ell_i - \ell_j)$. By substitution of the above results (2.11) can be reduced to

(2.12)
$$J(S; H L) = \prod_{i < j} (\ell_i - \ell_j) J(T, H^*) .$$

To compute $J(T,H^*)$ we equate corresponding elements on both sides of $H^* = HT$ which gives the following set of equations:

(2.13)
$$h_{il}^* = \sum_{k=2}^{p} h_{ik} t_{kl}$$
 $i = 2,...,p$

(2.14)
$$h_{i2}^* = -h_{i1} t_{21} + \sum_{k=3}^{p} h_{ik} t_{k2}$$
 $i = 3,...,p$

(2.15)
$$h_{i,p-2}^* = -\sum_{k=1}^{p-3} h_{ik} t_{p-2,k} + h_{i,p-1} t_{p-1,p-2} + h_{ip} t_{p,p-2} i = p-1,p$$

(2.16)
$$h_{p,p-1}^* = -\sum_{k=1}^{p-2} h_{pk} t_{p-1,k} + h_{pp} t_{p,p-1}$$

where $t_{ij}(i,j=1,2,...,p)$ are elements of T. It is easy to see that

(2.17)
$$J(H^*; T) = J(h^*_{i1}; t_{i1}; i = 2,...,p) J(h^*_{i2}; t_{i2}; i = 3,...,p) ...$$

$$... J(h^*_{p,p-1}; t_{p,p-1}) ...$$

In detail we can show

(2.18)
$$J(h_{11}^{*}; t_{11}^{*}; i = 2,...,p) = mod \begin{vmatrix} h_{22} & h_{32} & \cdots & h_{p2} \\ \vdots & & & \vdots \\ h_{2p} & h_{3p} & \cdots & h_{pp} \end{vmatrix},$$

(2.19)
$$J(h_{i2}^*; t_{i2}^*; i = 3,...,p) = mod \begin{vmatrix} h_{33} & h_{43} & \cdots & h_{p3} \\ \vdots & & & \vdots \\ h_{3p} & h_{4p} & \cdots & h_{pp} \end{vmatrix}$$
,

(2.20)
$$J(h_{p-1,p-2,p,p-2}^{*}; t_{p-1,p-2}, t_{p,p-2}) = mod \begin{vmatrix} h_{p-1,p-1} & h_{p,p-1} \\ h_{p-1,p} & h_{p,p} \end{vmatrix},$$

and

(2.21)
$$J(h_{p,p-1}^*; t_{p,p-1}) = |h_{pp}|$$

But

$$J(\underline{T}; \underline{H}^*) = 1/J(\underline{H}^*; \underline{T}),$$

hence using the above results (2.3) has the form

(2.23)
$$dF(\ell_1, \dots, \ell_p) = \frac{\left(\frac{n}{2}\right)^p \prod_{i=1}^{n} (\ell_i)^{\frac{n-p-1}{2}}}{\prod_{i=1}^{\frac{1}{n}p(p-1)} \left|\sum_{i=1}^{\frac{n}{2}}\prod_{i=1}^{p} \Gamma(\frac{n-i+1}{2})\right|}$$

$$\int_{\{HH'=I\}} \exp\left(\frac{n}{2} \operatorname{tr} AHLH'\right) |J| \prod_{i>j} dh_{i,j}$$

Comparing (2.23) with the result we already know (see [7], [10]) which is given by

(2.24)
$$dF(\ell_1, \ell_2, \dots, \ell_p) = \frac{\left(\frac{n}{2}\right)^{\frac{np}{2}} \frac{p}{n} (\ell_i)^{\frac{n-p-1}{2}} \frac{p}{n^2}}{2^p \left|\sum_{i=1}^{\frac{n}{2}} \left(\sum_{i=1}^{p} \left\{\Gamma(\frac{n-i+1}{2}) \Gamma(\frac{i}{2})\right\}\right\}}$$

$$\frac{1}{i < j} (\ell_i - \ell_j) \int_{o(p)} \exp(-\frac{n}{2} \operatorname{tr} AHLH^{\dagger}) (dH)$$

the proof of the lemma is completed.

Note that we may also write the Jacobian in (2.2) as

(2.25)
$$J = \left\{ h_{11} \begin{vmatrix} h_{11} & h_{21} \\ h_{12} & h_{22} \end{vmatrix} \dots \begin{vmatrix} h_{11} \cdots h_{p-1,1} \\ h_{1,p-1} \cdots h_{p-1,p-1} \end{vmatrix} \right\}^{-1}$$

since the choice of random elements in computing $J(H^*;T)$ is arbitrary.

Lemma (2.2). Given matrices $A(p \times p)$, $B(p \times q)$, $C(q \times p)$, $D(q \times q)$ where D is non-singular then

(2.26)
$$\begin{vmatrix} A & B \\ C & D \\ C & D \end{vmatrix} = \begin{vmatrix} D \\ C \end{vmatrix} \begin{vmatrix} A - BD^{-1} & C \\ C \end{vmatrix}$$

Proof: See Anderson [3] .

We now partition H as

$$(2.27) \qquad \qquad \underbrace{H}_{\infty} = \begin{bmatrix} A & Y & M \\ \sim & \sim & \sim \\ B & X & N \end{bmatrix}$$

where the dimensions of these submatrices are A_{∞} : $(\alpha+1) \times \alpha$, B_{∞} : $t \times \alpha, Y_{\infty}$: $(\alpha+1) \times 1$, X_{∞} : $t \times 1, M_{\infty}$: $(\alpha+1) \times t$, X_{∞} : $t \times 1, M_{\infty}$: $(\alpha+1) \times t$, X_{∞} : X_{∞} :

$$(2.28) \qquad \qquad A' A + B'B = I$$

$$(2.29) \qquad \qquad \underset{\sim}{A_{0}} \stackrel{\mathsf{T}}{\times} \qquad + \qquad \underset{\sim}{B_{1}} \stackrel{\mathsf{T}}{\times} \qquad = \qquad 0$$

$$(2.30) \qquad \qquad \underset{\sim}{x^{i}x} + \underset{\sim}{y^{i}y} = 1 .$$

 $\underbrace{\text{Lemma }(2.3)}. \quad \text{Let } \overset{X}{\approx} = (x_1, x_2, \dots, x_t)' \quad \text{and } \overset{A}{\approx} = (\overset{A}{\approx} \overset{Y}{\approx}) \quad \text{then}$

(2.31)
$$\int_{\mathbb{D}} \frac{\mathrm{d}x_{1}, \dots, \mathrm{d}x_{t}}{\left|\frac{A}{C}\right|} = \frac{\frac{t+1}{2}}{\Gamma(\frac{t+1}{2})}$$

where $D = \{(x_1, ..., x_t) \text{ such that } H'H = I \text{ and } |A| > 0\}$.

Proof:
$$\left| A \right| = \left| A'A \right|^{\frac{1}{2}}$$

or

$$|A| = \begin{vmatrix} A' & A & A' & Y \\ Y' & A & Y' & Y \end{vmatrix}^{\frac{1}{2}}$$

$$(2.32)$$

Applying Lemma (2.2) to the right hand side of (2.32) we have

(2.33)
$$|A' A_0|^{\frac{1}{2}} \left\{1 - X' \left[1 + B(A' A_0)^{-1} B' \right] X\right\}^{\frac{1}{2}} .$$

Let

(2.34)
$$Z = \left[I + B(A', A)^{-1} \right]^{\frac{1}{2}} X$$

and the Jacobian of this transformation is simply

$$J(X; Z) = \left| I + B(A' A_o)^{-1} B' \right|^{-\frac{1}{2}} .$$

In application of lemma (2.2) the right hand side of (2.35) can be shown to be equal to

(2.36)
$$|\mathbf{I} + (\mathbf{A}' \ \mathbf{A})^{-1} \ \mathbf{B'B}|^{-\frac{1}{2}} = |\mathbf{A}' \ \mathbf{A}|^{\frac{1}{2}}$$

since $\begin{vmatrix} A' & A + B'B \end{vmatrix} = 1$ by (2.28). Finally by substitution of the result we have

$$(2.37) \qquad \int_{D} \frac{dx_{1}, \dots, dx_{t}}{\begin{vmatrix} A \\ \sum \\ i=1 \end{vmatrix}} = \int_{t} \dots \int_{1-\frac{ZZ}{2}} \frac{dz_{1}, dz_{2}, \dots, dz_{t}}{\left[1-\frac{ZZ}{2}\right]^{\frac{1}{2}}}$$

$$= \frac{\frac{t+1}{2}}{\Gamma(\frac{t+1}{2})}.$$

Lemma (2.4). If $H = \begin{pmatrix} H_1 \\ H_2 \end{pmatrix}$ where H_2 consists of only the pth row of H_2 and $f(H_2)$ is an integrable function which depends only on H_2 then

(2.38)
$$\int_{O(p)} f(H_2)(dH) = \frac{2^p \Gamma(\frac{p}{2})}{\frac{p}{n^2}} \int_{\substack{p-1 \ i=1}}^{\infty} \dots \int_{\substack{p'=1 \ i=1}}^{\infty} f(H_2)(1 - \sum_{i=1}^{p-1} h_{pi}^2) \prod_{i=1}^{2} h_{pi}^{-1} .$$

Proof: We first prove below the lemma when p = 4 by an elementary method then apply lemma (2.3) in proving the case for the general p. When p = 4, the left hand side of (2.38) becomes

(2.39)
$$\int_{O(4)} f(h_{41}, h_{42}, h_{43}) (dH) \sim$$

Application of lemma (2.1) to (2.39) yields

$$(2.40) \qquad \frac{2^{3}}{\prod^{l_{1}}} \int_{\mathbb{R}^{n}} \int_{\mathbb{R}^{n}} \int_{\mathbb{R}^{n}} f(h_{l_{1}}, h_{l_{2}}, h_{l_{3}}) \mod \left\{ \frac{\prod_{i \geq j} dh_{ij}}{\left|h_{11} h_{12} \right| \left|h_{11} h_{12} h_{13} \right|} \right\} .$$

We know by the orthogonality of H

(2.41)
$$|h_{11}|^{-1} = (1-h_{21}^2 - h_{31}^2 - h_{41}^2)^{-\frac{1}{2}} ,$$

and

(2.42)
$$\mod \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix}^{-1} = |h_{11}|^{-1} = (1 - h_{11}^{2} - h_{12}^{2} - h_{12}^{2})^{-\frac{1}{2}} .$$

Moreover,

$$(2.43) \mod \binom{h_{11} \ h_{12}}{h_{21} \ h_{22}}^{-1} = \{ [1 - (h_{31}^2 + h_{41}^2)][1 - (h_{32}^2 + h_{42}^2)] - (h_{31} \ h_{32}^2 + h_{41}h_{42})^2 \}^{-\frac{1}{2}}$$

$$= [(1 - h_{41}^2 - h_{42}^2) + h_{31}^2 h_{42}^2 - h_{31}^2 + h_{32}^2 h_{41}^2 - h_{32}^2 - 2h_{31}h_{32}h_{41}h_{42}]^{-\frac{1}{2}}$$

$$= [1 - \frac{(1 - h_{42}^2)}{(1 - h_{41}^2 - h_{42}^2)} \ h_{31}^2 - \frac{(1 - h_{41}^2)}{(1 - h_{41}^2 - h_{42}^2)} h_{32}^2 - \frac{2h_{42}h_{41}h_{31}h_{32}}{(1 - h_{41}^2 - h_{42}^2)}]^{-\frac{1}{2}} \frac{1}{\int_{1 - h_{41}^2 - h_{42}^2}}$$

Since f is independent of h_{21} so we make the following transformation into (2.41)

$$\mu = \frac{h_{21}}{\sqrt{1 - h_{31}^2 - h_{41}^2}}$$

then

$$(2.45) \qquad (1-h_{21}^2-h_{31}^2-h_{41}^2)^{-\frac{1}{2}} dh_{21} dh_{31} dh_{41} = (1-\mu^2)^{-\frac{1}{2}} d\mu dh_{31} dh_{41} .$$

So we can integrate with respect to μ from -1 to 1 easily. Next let

$$(2.46) x = \sqrt{\frac{1-h_{42}^2}{(1-h_{41}^2-h_{42}^2)}} h_{31} + \sqrt{\frac{h_{41}h_{42}}{(1-h_{42}^2)(1-h_{41}^2-h_{42}^2)}} h_{32} ,$$

(2.47)
$$y = \sqrt{\frac{1}{(1-h_{42}^2)}} h_{32}$$

then

(2.48)
$$J(h_{31}, h_{32}; x, y) = \sqrt{1 - h_{41}^2 - h_{42}^2} .$$

By substitution of the above results (2.40) reduces to

$$(2.49) \qquad \frac{2^{3}}{\Pi^{4}} \int \int \int \frac{f(h_{41}, h_{42}, h_{43})}{3 \frac{3}{1 + 1}} dh_{41} dh_{42} dh_{43}$$

$$\left\{ \sum_{i=1}^{3} h_{41}^{2} \leq 1 \right\} \int \frac{f(h_{41}, h_{42}, h_{43})}{1 + 2 \frac{3}{2}} dh_{41} dh_{42} dh_{43}$$

$$\int \int \frac{dx \, dy}{\sqrt{1-x^2-y^2}} \int_{-1}^{+1} \frac{1}{\sqrt{1-\mu^2}} \, du = \frac{2^{l_1}}{l_1^2} \int \int \int \frac{f(h_{l_1}, h_{l_2}, h_{l_3})}{\left(\frac{3}{1-\sum_{i=1}^{\infty} h_{l_i}^2}\right)^{\frac{1}{2}}} \left\{ \frac{3}{1-\sum_{i=1}^{\infty} h_{l_i}^2} \right\} \frac{1}{\left(\frac{3}{1-\sum_{i=1}^{\infty} h_{l_i}^2}\right)^{\frac{1}{2}}}$$

 $dh_{41} dh_{42} dh_{43}$

For general p repeated applications of lemma (2.3) to E yields the following

(2.50)
$$\int_{O(p)} f(H_2)(dH) = \frac{2^p \prod_{k=1}^p \Gamma(\frac{k}{2})}{\prod_{k=1}^{\frac{k}{p}} (p+1)} \cdot \frac{\frac{p-1}{2}}{\Gamma(\frac{p-1}{2})} \cdot \frac{\frac{p-2}{2}}{\Gamma(\frac{p-2}{2})} \cdot \frac{\frac{3}{2}}{\Gamma(\frac{3}{2})} \cdot \frac{\Pi}{\Gamma(\frac{2}{2})}$$

$$\int \dots \int \frac{f(H_2)}{\sum_{i=1}^{p-1} \prod_{j=1}^{p-1} dh_{pi}} \frac{f(H_2)}{\prod_{i=1}^{p-1} dh_{pi}} = \frac{2^p \Gamma(\frac{p}{2})}{\prod^{p/2}} \int \dots \int \frac{f(H_2) \prod_{i=1}^{p-1} dh_{pi}}{\sum_{i=1}^{p-1} \prod_{j=1}^{p-1} dh_{pi}}$$

which completes the proof.

<u>Lemma (2.5)</u>. If $H = \begin{pmatrix} H_1 \\ H_2 \end{pmatrix}$ where H_1 consists of the 1st row of H and $f(H_1)$ is an integrable function which depends only on H_1 then

$$(2.51) \qquad \int_{O(p)} f(H_{1})(dH) = \frac{2^{p}\Gamma(\frac{p}{2})}{\Pi^{p/2}} \int_{\prod_{i=2}^{p} h_{1i} \leq 1} \int_{\substack{1-\sum h_{1i}^{2} \\ i=2}}^{f(H_{1})} \prod_{i=2}^{p} dh_{1i}.$$

Proof: We choose h_{ij} , i < j, (i,j = 1,2,...,p) as random variables and the rest of the proof is analagous to the proof of lemma (2.4).

Theorem (2.1). Let A and L be defined as in (1.1) with

$$0 < a_1 = a_2 = \dots = a_{p+1} = a, a_p > a, and l_1 > l_2 > \dots > l_p > 0$$

then, for large degrees of freedom n, the first term of the asymptotic expansion of E is given by

(2.52)
$$E \simeq \frac{2^{p} \Gamma(\frac{p}{2})}{\frac{p}{n^{2}}} (1 + a_{p} \ell_{p})^{-\frac{n}{2}} \prod_{i=1}^{p-1} (1 + a \ell_{i})^{-\frac{n}{2}} \prod_{i=1}^{p-1} (\frac{2\Pi}{nC_{ip}})^{\frac{1}{2}} \left[1 + O(\frac{1}{n})\right]$$

where

(2.53)
$$C_{ip} = \frac{(a_p - a)(\ell_i - \ell_p)}{(1 + a_\ell)(1 + a_p \ell_p)} \qquad (i = 1, 2, ..., p).$$

Proof:
$$A = aI + (a_p-a)$$

then

(2.54)
$$|\underline{\mathbf{I}} + \underline{\mathbf{A}} + \underline{\mathbf{H}} \underline{\mathbf{L}} + \underline{\mathbf{H}'}| - \frac{\underline{\mathbf{n}}}{2} = |\underline{\mathbf{I}} + \underline{\mathbf{H}'}\underline{\mathbf{A}} + \underline{\mathbf{L}}| - \frac{\underline{\mathbf{n}}}{2}$$

$$= |\underline{\mathbf{I}} + \underline{\mathbf{a}} \underline{\mathbf{L}} + (\underline{\mathbf{a}}_{p} - \underline{\mathbf{a}})\underline{\mathbf{h}} \underline{\mathbf{h}'} \underline{\mathbf{L}}| - \frac{\underline{\mathbf{n}}}{2}$$

where

(2.55)
$$h_{p}^{i} = (h_{p1}, ..., h_{pp})$$

So

(2.56)
$$|\underline{\mathbf{I}} + \underline{\mathbf{A}} + \underline{\mathbf{H}} + \underline{\mathbf{L}} + \underline{\mathbf{I}}|^{-\frac{n}{2}} = |\underline{\mathbf{I}} + \underline{\mathbf{a}} + \underline{\mathbf{L}}|^{-\frac{n}{2}} |\underline{\mathbf{I}} + (\underline{\mathbf{a}}_{p} - \underline{\mathbf{a}})(\underline{\mathbf{I}} + \underline{\mathbf{a}} + \underline{\mathbf{L}})^{-1} |\underline{\mathbf{h}}_{p} \underline{\mathbf{h}}_{p}^{\dagger} \underline{\mathbf{L}}|^{-\frac{n}{2}}$$

$$= |\underline{\mathbf{I}} + \underline{\mathbf{a}} + \underline{\mathbf{L}}|^{-\frac{n}{2}} (1 + (\underline{\mathbf{a}}_{p} - \underline{\mathbf{a}}) \underline{\mathbf{h}}_{p}^{\dagger} \underline{\mathbf{L}}(\underline{\mathbf{I}} + \underline{\mathbf{a}} + \underline{\mathbf{L}})^{-1} |\underline{\mathbf{h}}_{p}|^{-\frac{n}{2}}$$

$$= |\underline{\mathbf{I}} + \underline{\mathbf{a}} + \underline{\mathbf{L}}|^{-\frac{n}{2}} (1 + \sum_{i=1}^{p} z_{i} \underline{\mathbf{h}}_{pi}^{2})^{-\frac{n}{2}}$$

where

(2.57)
$$z_{i} = \frac{\ell_{i}(a_{p}-a)}{1+a\ell_{i}} \qquad (i = 1,2,...,p) .$$

We see that (2.54) depends only on h_p^* so that application of lemma (2.4) expands E into the following form

(2.58)
$$E \approx \frac{2^{p}\Gamma(\frac{p}{2})}{\Pi^{p/2}} \prod_{i=1}^{p} (1+a\ell_{i})^{-\frac{n}{2}} \int \dots \int \frac{\left\{1+\frac{1}{2}\sum_{i=1}^{p-1}h_{pi}^{2}+\dots\right\}}{\prod_{i=1}^{p-1}h_{pi}^{2}} \prod_{i=1}^{p-1}dh_{pi}$$

$$\left\{\sum_{i=1}^{p-1}h_{pi}^{2} \leq 1\right\} \left(1+\sum_{i=1}^{p}z_{i}h_{pi}^{2}\right)^{\frac{n}{2}} \prod_{i=1}^{p-1}dh_{pi}$$

Since $h_{pp}^2 = 1 - \sum_{i=1}^{p-1} h_{pi}^2$, it is easy to show that in a small neighborhood of origin

$$(2.59) \qquad \left(1 + \sum_{i=1}^{p} z_{i}h_{pi}^{2}\right)^{-\frac{n}{2}} = \left(\frac{1+a_{p}p_{p}}{1+al_{p}}\right)^{-\frac{n}{2}} \left(1 + \sum_{i=1}^{p-1} c_{ip}h_{pi}^{2}\right)^{-\frac{n}{2}}$$

$$= \left(\frac{1+a_{p}p_{p}}{1+al_{p}}\right)^{-\frac{n}{2}} \exp\left\{-\frac{n}{2}\sum_{i=1}^{p-1} c_{ip}h_{pi}^{2}\right\} \left[1 + \frac{n}{4}\left(\sum_{i=1}^{p-1} c_{ip}h_{pi}^{2}\right)^{2} + \ldots\right]$$

Direct substitution of (2.59) into (2.58) yields

(2.60)
$$E \approx \frac{2^{p}\Gamma(\frac{p}{2})}{n^{p/2}} (1 + a_{p}\ell_{p})^{-\frac{n}{2}} \prod_{i=1}^{p-1} (1 + a\ell_{i})^{-\frac{n}{2}}$$

$$\int \dots \int \exp\left(-\frac{n}{2} \sum_{i=1}^{p-1} c_{ip}h_{pi}^{2}\right) \prod_{i=1}^{p-1} dh_{pi} \left\{1 + o\left(\frac{1}{n}\right)\right\} .$$

$$\left\{\sum_{j=1}^{p-1} h_{pj}^{2} \leq 1\right\}$$

Since the integrand is of exponential form it attains its maximum value at origin and most of this integral is concentrated in a small neighborhood of origin. Therefore, for large n, all limits can be set to $\pm \infty$ [4], and we have for the first term of the expansion of E

(2.61)
$$E \simeq \frac{2^{p} \Gamma(\frac{p}{2})}{\prod^{p/2}} \frac{\prod_{i=1}^{p-1} (1+a\ell_{i})^{-\frac{n}{2}}}{\prod_{i=1}^{n} (1+a\ell_{i})^{\frac{n}{2}}} \prod_{i=1}^{p-1} \left(\frac{2\prod}{c_{ip}^{n}}\right)^{\frac{1}{2}} \left[1+o(\frac{1}{n})\right] .$$

Theorem (2.2). Let $\underset{\sim}{A}$ and $\underset{\sim}{L}$ be defined as (1.1) with

$$0 < a_1 < a_2 = a_3 = \dots = a_{p-1} = a_p = a$$
 and $\ell_1 > \ell_2 > \dots > \ell_p > 0$

then, for large degrees of freedom n, the first term of the asymptotic expansion of E is given by

(2.62)
$$E \simeq \frac{2^{p} \Gamma(\frac{p}{2})}{\frac{p}{n^{2}}} \frac{\prod_{i=2}^{p} (1+a\ell_{i})^{-\frac{n}{2}}}{\prod_{i=2}^{n} (1+a\ell_{i})^{\frac{n}{2}}} \prod_{i=2}^{p} \left(\frac{2\pi}{C_{1i}n}\right)^{\frac{1}{2}} \left[1+O(\frac{1}{n})\right]$$

Proof: We first note that

(2.63)
$$|\underbrace{\text{I+A}}_{\sim} \text{H} \underset{\sim}{\text{L}} \text{H'}|^{-\frac{n}{2}} = |\underbrace{\text{I}}_{\sim} + \underset{\sim}{\text{aL}} + (a_1 - a) \underset{\sim}{\text{h}}_{\sim} \underset{\sim}{\text{h'}} \underset{\sim}{\text{L'}}|^{-\frac{n}{2}}$$

which depends only on $h'_1 = (h_{11}, h_{12}, \dots, h_{1p})$. Hence we can apply lemma (2.5) and the rest of the proof is analogous to the proof of theorem (2.1).

Corollary. In one sample case under the same condition of theorem (2.1), i.e., $0 < a_1 = \dots = a_{p-1} < a_p$, $\ell_1 > \ell_2 > \dots > \ell_p > 0$ then

(2.64)
$$\int_{o(p)} \exp \left(-\frac{n}{2} \operatorname{tr} A H L H'\right) (dH)$$

$$= 2^{p} \frac{\Gamma(\frac{p}{2})}{\prod^{\frac{1}{2}}} \exp \left(-\frac{n}{2} \operatorname{tr} \underbrace{A}_{i} \underbrace{L}\right) \sum_{j=1}^{p-1} \left(\frac{2 \prod i}{n C_{ip}^{*}}\right) \left[1 + O\left(\frac{1}{n}\right)\right]$$

where

(2.65)
$$C_{ip}^* = (a_p - a_i)(\ell_i - \ell_p)$$
 (i = 1,2,..., p-1)

Proof: For the proof, note that

(2.66)
$$\exp(-\frac{n}{2} \operatorname{tr} A + L + 1) = \exp(-\frac{n}{2} \operatorname{tr} A + L) = \exp(-\frac{n}{2} \sum_{i=1}^{p-1} C_{ip}^* h_{pi}^2)$$
.

The rest follows easily.

3. Asymptotic Representation When Not All a 's Are Equal

When $0 < a_1 < a_2 < \ldots < a_k < a_{k+1} = \ldots = a_p$ (1 < k \le p-1) the method in the previous section is too troublesome to be applied. However, we notice that if we partition the matrix $\frac{H}{L}$ into the submatrices $\frac{H}{L}$ and $\frac{H}{L}$ consisting of its first k and (p-k) rows of $\frac{H}{L}$ and if the integrand does not depend upon $\frac{H}{L}$ then we can integrate over $\frac{H}{L}$ for fixed $\frac{H}{L}$ by the formula

$$(3.1) \qquad \qquad \int_{\mathbb{H}_2} (d\mathbb{H}) = (d\mathbb{H}_1)$$

where the symbol (dH₁) stands for the invariant volume element on the Stiefel manifold of orthonormal k-frames in p space normalized to make its integral unity. Note that (3.1) can be easily verified by using lemma (2.4) of the last section.

We then apply the transformation

$$(3.2) \qquad \qquad \underset{\sim}{\text{H}} = \exp S,$$

S a p x p skew symmetric matrix and a parameterization of H may be obtained by writing

where S_{11} is a k x k skew symmetric matrix and S_{12} is a k x (p-k) rectangular matrix.

The Jacobian of (3.3) has been computed in [1] as follows

(3.4)
$$J = 1 + \frac{p-2}{24} \operatorname{tr} s^2 + \frac{8-p}{4 \times 6!} \operatorname{tr} s^4 + \dots$$

It can be easily shown from (3.4) that

(3.5)
$$\frac{\prod_{i=1}^{\frac{1}{2}kp}}{\Gamma_{k(\frac{1}{2}p)}} \quad (dH_{1}) = (dS_{11})(dS_{12}) \quad \{1+0 \quad (squares of s_{ij}'s)\}$$

where the symbols (dS_{11}) and (dS_{12}) stand for dS_{11} and dS_{11} and dS_{11} and dS_{11} and dS_{11} respectively and $\Gamma_p(t) = \Pi^{p(p-1)/4} \prod_{j=1}^p \Gamma(t-\frac{1}{2}j+\frac{1}{2})$.

<u>Lemma (3.1)</u>. If $\stackrel{A}{\sim}$ and $\stackrel{L}{\sim}$ are defined as in (1.1) with

$$0 < a_1 < a_2 < \dots < a_k < a_{k+1} = \dots = a_p = a$$
 and $\ell_1 > \ell_2 > \dots > \ell_p > 0$

and if we partition $H = \begin{pmatrix} H_1 \\ H_2 \end{pmatrix}$ such that H_2 consisting of last (p-k) rows of H then $|I + A H L H^*|$ does not depend upon H_2 .

Proof:

then

(3.7)
$$|\mathbf{I} + \mathbf{A} + \mathbf{H} + \mathbf{H}'| = |\mathbf{I} + \mathbf{H}' + \mathbf{A} + \mathbf{H}|$$

$$= |\mathbf{I} + \mathbf{A} + \mathbf{H} + \mathbf{H}'|$$

$$= |\mathbf{I} + \mathbf{A} + \mathbf{H} + \mathbf{H}'|$$

$$\circ \quad \mathbf{a_k - a}$$

which is independent of H_2 .

It has been shown in [3]

(3.8)
$$E = \int_{\mathcal{O}(\mathbf{p})} \left| \underbrace{\mathbf{l} + \mathbf{A}}_{\sim} + \underbrace{\mathbf{L}}_{\sim} + \underbrace{\mathbf{l} \cdot \mathbf{l}}_{\sim} \right|^{-\frac{n}{2}} (dH) \simeq 2^{\mathbf{p}} \int_{\mathbb{N}(\mathbf{l})} \left| \underbrace{\mathbf{l} + \mathbf{A}}_{\sim} + \underbrace{\mathbf{L}}_{\sim} + \underbrace{\mathbf{l} \cdot \mathbf{l}}_{\sim} \right|^{-\frac{n}{2}} (dH)$$

where N(I) is the neighborhood of the identity element of the orthogonal manifold. If $0 < a_1 < a_2 < \dots < a_k < a_{k+1} = \dots = a_p = a$ and $\ell_1 > \ell_2 > \dots > \ell_p > 0$ by lemma (3.1) the integrand does not depend upon H₂. Therefore, we can use (3.1) to integrate over H₂ for fixed H₁

(3.9)
$$\mathbb{E} \sim 2^{p} \int_{\mathbb{N}(\mathbb{I})} \left| \mathbb{I} + \mathbb{A} + \mathbb{I} + \mathbb{I} + \mathbb{I} \right|^{-\frac{n}{2}} (dH_{1}) .$$

Now under the transformation (3.3), $N(I) \rightarrow N(S=0)$ and in the small neighborhood of N(S=0) its Jacobian can be assumed to be equal to one. Moreover, the integrand of (3.9) can be shown to be equal to

(3.10)
$$\prod_{i=1}^{k} (1+a_{i}l_{i})^{-\frac{n}{2}} \prod_{i=k+1}^{p} (1+al_{i})^{-\frac{n}{2}} \prod_{i < j}^{k} \exp\left\{-\left(\frac{n}{2}\right) \frac{(a_{j}-a_{i})(l_{i}-l_{j})}{(1+a_{i}l_{i})(1+a_{j}l_{j})} s_{i,j}^{2}\right\}$$

Substitution of (3.10) into (3.9) yields

(3.11)
$$E \simeq 2^{p} \frac{\Gamma_{k}(\frac{p}{2})}{\prod^{\frac{1}{2}kp}} \prod_{i=1}^{k} (1+a_{i}\ell_{i})^{-\frac{n}{2}} \prod_{i=k+1}^{p} (1+a\ell_{i})^{-\frac{n}{2}}$$

$$\int_{S_{11}} \int_{S_{12}} \prod_{i < j}^{k} \exp \left\{ \frac{(-\frac{n}{2})(a_{j} - a_{i})(\ell_{i} - \ell_{j})s_{i,j}^{2}}{(1 + a_{i}\ell_{i})(1 + a_{j}\ell_{j})} \right\}$$

For large n and a_i 's (i = 1,2,...,k) and ℓ_i 's (i = 1,...,k) well spaced most of the integral in (3.11) will be obtained from small values of the elements of s_{ij} and s_{ij} . Hence, to obtain an asymptotic series, we can replace the finite range of s_{ij} by the range of all real values of s_{ij}

(3.12)
$$E \approx 2^{p} \frac{\Gamma_{k}(\frac{p}{2})}{\prod^{\frac{1}{2}kp}} \stackrel{k}{\underset{i=1}{\eta}} (1+a_{i}\ell_{i})^{-\frac{n}{2}} \stackrel{p}{\underset{i=k+1}{\eta}} (1+a\ell_{i})^{-\frac{n}{2}}$$

$$\prod_{\mathbf{i} < \mathbf{j}}^{k} \int_{-\infty}^{\infty} \exp \left\{ \frac{-(\frac{n}{2})(\mathbf{a}_{\mathbf{j}} - \mathbf{a}_{\mathbf{i}})(\mathbf{l}_{\mathbf{i}} - \mathbf{l}_{\mathbf{j}}) \mathbf{s}_{\mathbf{i}\mathbf{j}}^{2}}{(1 + \mathbf{a}_{\mathbf{i}} \mathbf{l}_{\mathbf{i}})(1 + \mathbf{a}_{\mathbf{j}} \mathbf{l}_{\mathbf{j}})} \right\} \ \mathrm{d}\mathbf{s}_{\mathbf{i}\mathbf{j}}$$

$$\prod_{i=1}^{k} \prod_{j=k+1}^{p} \int_{-\infty}^{\infty} \exp \left\{ \frac{-(\frac{n}{2})(a-a_{i})(\ell_{i}-\ell_{j})s_{i,j}^{2}}{(1+a_{i}\ell_{i})(1+a\ell_{j})} \right\} ds_{i,j} \left\{ 1+o(\frac{1}{n}) \right\} .$$

We thus have the following theorem.

Theorem (3.1). The asymptotic distribution of the roots, $\ell_1 > \ell_2 > \dots > \ell_p > 0$ of $\sum_{k=2}^{n-1} s_k$, for large degrees of freedom $n = n_1 + n_2$, when roots of $(\sum_{k=2}^{n-1} s_k)^{-1}$ are $0 < a_1 < a_2 < \dots < a_k < a_{k+1} = \dots = a_p$ is given by

(3.13)
$$C* \prod_{i < j}^{p} (\ell_{i} - \ell_{j}) \prod_{i=1}^{p} \left\{ (\ell_{i})^{\frac{n_{1} - p - 1}{2}} (a_{i})^{\frac{n_{1}}{2}} \right\} \prod_{i=k+1}^{p} \left\{ (1 + a_{i} \ell_{i})^{-\frac{(n_{1} + n_{2})}{2}} \right\}$$

$$\prod_{i=1}^{k} \left\{ (1 + a_{i} \ell_{i})^{\frac{(n_{1} + n_{2})}{2}} \right\} \prod_{i < j}^{k} \left\{ \frac{2 \pi}{(n_{1} + n_{2}) c_{i,j}} \right\}^{\frac{1}{2}} \prod_{i=1}^{k} \prod_{j=k+1}^{p} \left\{ \frac{2 \pi}{(n_{1} + n_{2}) c_{i,j}} \right\}^{\frac{1}{2}}$$

$$\left[1+0\left(\frac{1}{n_1+n_2}\right)\right]$$

where

(3.14)
$$c_{ij} = \frac{(a_j - a_i)(\ell_i - \ell_j)}{(1 + a_i \ell_i)(1 + a_j \ell_j)}$$

and

(3.15)
$$C* = \frac{\prod_{i=1}^{\frac{1}{4}p(p-1)-\frac{1}{2}kp}}{\left[\Gamma_{k}(\frac{p}{2})\right]^{-1}} \prod_{i=1}^{p} \Gamma_{p}(\frac{i}{2})\Gamma_{p}(\frac{n_{1}+n_{2}}{2}) \left\{\Gamma_{p}(\frac{p}{2})\Gamma_{p}(\frac{n_{1}}{2})\Gamma_{p}(\frac{n_{2}}{2})\right\}^{-1}.$$

(3.13) can be rewritten as the following form

(3.16)
$$C * \prod_{i < j} (\ell_i - \ell_j) \prod_{i=1}^{p} \left\{ \ell_i - \ell_j \right\} \prod_{i=1}^{n_1 - p - 1} a_i^{\frac{n_1}{2}} G_1 \cdot G_2 \cdot G_3 \left[1 + O(\frac{1}{n_1 + n_2}) \right]$$

where

(3.17)
$$G_{1} = \prod_{i=1}^{k} \left\{ (1 + a_{i} \ell_{i})^{\frac{-(n_{1} + n_{2})}{2}} \right\} \prod_{i < j}^{k} \left\{ \frac{2 \pi}{(n_{1} + n_{2})^{c}} \right\}^{\frac{1}{2}}$$

(3.18)
$$G_{2} = \prod_{i=k+1}^{p} \left\{ (1+a_{i} \ell_{i})^{-\frac{(n_{1}+n_{2})}{2}} \right\},$$

and

(3.19)
$$G_{3} = \prod_{i=1}^{k} \prod_{j=k+1}^{p} \left\{ \frac{2\pi}{(n_{1}+n_{2})C_{i,j}} \right\}^{\frac{1}{2}}.$$

Note that except differing by a constant G_1 is actually an asymptotic representation of the first k distinct roots, $(2_i's, a_i's; i = 1, 2, ..., k)$ [c.f. Theorem (1.1)] and G_2 is related to the next (p-k) equal roots

(ℓ_i 's, a_i 's; i = k+1,...,p) [c.f. Eq. (1.2)] while G_3 can be considered as cross effects between the roots of G_1 and G_2 .

Acknowledgement

The author wishes to express his sincere thanks to Professor K.C.S. Pillai for introducing me to this and related problems of research and his guidance throughout. Special thanks are also given to Professor C.G. Khatri for his generous discussions during his visit at Purdue University.

REFERENCES

- [1] Anderson, G.A. (1965). An Asymptotic Expansion for the Distribution of the Latent Root of the Estimated Covariance Matrix. Ann. Math. Statist. 36, 1153-1173.
- Anderson, T.W. (1948). The Asymptotic Distribution of the Roots of Certain Determinantal Equations. Roy. Stat. Soc. 10, 132-139.
- [3] Chang, T.C. (1968). On an Asymptotic Representation of the Distribution of the Characteristic Roots of SlS2. Mimeo Series No. 165, Department of Statistics, Purdue University.
- [4] Erdélyi, A. (1956). Asymptotic Expansions. Dover, New York.
- [5] Hsu, L.C. (1948). A Theorem on the Asymptotic Behavior of a Multiple Integral. <u>Duke Math</u>. J. 15, 623-632.
- [6] James, A.T. (1954). Normal Multivariate Analysis and the Orthogonal Group. Ann. Math. Statist. 25, 40-75.
- [7] (1960). The Distribution of the Latent Roots of the Covariance Matrix. Ann. Math. Statist. 31, 151-158.
- [8] (1964). Distributions of Matrix Variates and Latent Roots Derived from Normal Samples. Ann.Math. Statist. 35, 475-501.
- [9] Khatri, C.G. (1967). Some Distribution Problems Connected with the Characteristic Roots of S₁S₂¹. Ann. Math. Statist. 38, 944-948.
- [10] Roy, S.N. (1957). Some Aspects of Multivariate Analysis. John Wiley and Sons, Inc., New York.