The Annals of Statistics
1976, Vol. 4, No. 4, 796-806

MAXIMIZATION OF AN INTEGRAL OF A MATRIX FUNCTION AND ASYMPTOTIC EXPANSIONS OF DISTRIBUTIONS OF LATENT ROOTS OF TWO MATRICES[1]

BY A. K. CHATTOPADHYAY, K. C. S. PILLAI AND HUNG C. LI

Purdue University and University of Southern Colorado

The noncentral distribution of latent roots arising in several situations in multivariate analysis involves the integration of a hypergeometric function of matrix variates over a group of orthogonal matrices in the real case and that of unitary matrices in the complex case. In this paper the subgroup of the orthogonal group (unitary group) for which the integrand is maximized has been found under mild restrictions. The results of earlier authors (Anderson, Chang, James, Li and Pillai) follow as special cases. Further, the maximization results concerning the integrand have been used to study asymptotic expansions of the distributions of the characteristic roots of matrices arising in canonical correlation analysis and MANOVA when the corresponding parameter matrices have several multiple roots.

1. Introduction. In multivariate analysis, the distribution of characteristic roots arising in testing the equality of two covariance matrices, in MANOVA, or in the canonical correlation problem, involves the integration of a hypergeometric function of the form

$$\mathscr{I} = \int_{O(p)} {}_sF_t(a_1, \cdots, a_s, b_1, \cdots, b_t, \mathbf{AHRH}')\, d(\mathbf{H}) , \tag{1}$$

where $O(p)$ is the group of orthogonal matrices $\mathbf{H}(p \times p)$, $\mathbf{A} = \operatorname{diag}(l_1, \cdots, l_p)$, $\mathbf{R} = \operatorname{diag}(r_1, \cdots, r_p)$, $d(\mathbf{H})$ is the invariant or Haar measure over the group $O(p)$ normalized so that the measure of the whole group $O(p)$ is unity, ${}_sF_t$ is a hypergeometric function of matrix variates (James [9]) and $a_1, \cdots, a_s, b_1, \cdots, b_t$, are functions of df and are positive real numbers. In the one sample (covariance matrix) case, Anderson [1] has shown that the maximum of the integrand ($s = t = 0$) for all possible variations of $\mathbf{R}$, the sample characteristic root matrix, is attained when $\mathbf{H}$ takes a special form. Chang [2], and Li and Pillai [12], [14] found the same form for $\mathbf{H}$ when maximizing the integrand in the two sample (two covariance matrices) problem ($s = 1$, $t = 0$). In the complex analogue of both one sample and two sample cases, Li and Pillai [12], [14] obtained a similar form of the unitary matrix $\mathbf{U}$. The purpose of this paper is to generalize their results both in the complex and real situations with $a_1, \cdots, a_s, b_1, \cdots, b_t$ satisfying some suitable conditions.

Received June 1975; revised December 1975.

[1] This work is supported by the National Science Foundation Grant No. GP-11473 at Purdue University.

AMS 1970 *subject classification.* 62H10.

Key words and phrases. Maximization, integral of a matrix function, asymptotic expansions, distribution, characteristic roots, canonical correlation, MANOVA, several multiple population roots.

We develop the idea in a series of lemmas and theorems and show that the results of Anderson [1], Chang [2], James [10], and Li and Pillai [12], [14] are special cases of our results. The generalization has not only been in regard to ${}_sF_t$-hypergeometric functions but also when l_i's are equal within each of several sets. We have further proved that the integral under different forms of the matrix $\mathbf{A}$ is invariant under choices of different submatrices of $\mathbf{H}$ and our general results cover some earlier ones of the above authors.

The maximization results concerning the integrand of (1) have further been used to study asymptotic expansions of the distributions of the characteristic roots of matrices arising in I. Canonical correlation analysis and II. MANOVA when the corresponding parameter matrices have several multiple roots.

2. Maximization of some special functions. First we prove the following lemma:

LEMMA 2.1. *Let* $f(\mathbf{T})$ *be a real valued function of the elements of the matrix* $\mathbf{T}(p \times p) = (t_{ij})$. *Then*

$$df(\mathbf{T}) = \operatorname{tr}(\mathbf{Q}\, d\mathbf{T})$$

where

$$\mathbf{Q} = \begin{pmatrix} \frac{\partial f}{\partial t_{11}}, & \cdots, & \frac{\partial f}{\partial t_{1p}} \\ \vdots & & \vdots \\ \frac{\partial f}{\partial t_{p1}}, & \cdots, & \frac{\partial f}{\partial t_{pp}} \end{pmatrix} \quad \text{and} \quad d\mathbf{T} = \begin{pmatrix} dt_{11}, & \cdots, & dt_{1p} \\ \vdots & & \vdots \\ dt_{p1}, & \cdots, & dt_{pp} \end{pmatrix}.$$

Proof follows directly from the definition. We give below some special cases.

CASE 1. If $\mathbf{B}$ be a nonsingular square matrix, then

$$d|\mathbf{B}| = |\mathbf{B}| \operatorname{tr}[\mathbf{B}^{-1}(d\mathbf{B})]\,. \tag{2}$$

This is Hsu's result as reported by Deemer and Olkin [8], proved in a different way.

CASE 2. If in (2) $\mathbf{B} = \mathbf{I} + \mathbf{AHRH}'$ where

$$\begin{aligned} &\mathbf{A} = \operatorname{diag}(l_1, \cdots, l_p)\,, \qquad \mathbf{R} = \operatorname{diag}(r_1, \cdots, r_p)\,, \\ &\infty > l_1 > \cdots > l_p > 0\,, \quad \infty > r_p > r_{p-1} > \cdots > r_1 > 0 \quad \text{and} \\ &\mathbf{H} \in O(p)\,, \end{aligned} \tag{3}$$

then Lemma 1 of [2] is obtained.

CASE 3. Take $f(\mathbf{T})$ in Lemma 2.1 to be

$$f(\mathbf{AHRH}') = \exp[-\operatorname{tr}(\mathbf{AHRH}')]\,, \tag{4}$$

where $\mathbf{A}$, $\mathbf{R}$ and $\mathbf{H}$ satisfy (3), and $\mathbf{H}$ is the only variable matrix.

By Lemma 2.1

$$df(\mathbf{AHRH}') = \operatorname{tr}[\mathbf{Q}\, d(\mathbf{AHRH}')]\,.$$

But $\mathbf{Q}$ in this case is a nonzero scalar matrix. Hence,

$$df(\mathbf{AHRH'}) = 0 \Rightarrow \operatorname{tr}[d(\mathbf{AHRH'})] = 0$$

(5) $$\Rightarrow \operatorname{tr}[\mathbf{A}(d\mathbf{H})\mathbf{RH'} + \mathbf{AHR}(d\mathbf{H'})] = 0$$

$$\Rightarrow 2 \operatorname{tr}[\mathbf{RH'AH}(\mathbf{H'}\, d\mathbf{H})] = 0\,.$$

But $(\mathbf{H'}\, d\mathbf{H})$ is a skew symmetric matrix. Hence for all $\mathbf{R} > 0$

$$(5) \Rightarrow \mathbf{RH'AH} \quad \text{is symmetric}$$
$$\Rightarrow \mathbf{RH'AH} = \mathbf{H'AHR}$$
$$\Rightarrow \mathbf{H'AH} = \operatorname{diag}(\mu_1, \cdots, \mu_p)$$

and as $\mathbf{R}$ in (3) is diagonal with distinct roots implies the form for $\mathbf{H}$ as

(i) $\mathbf{H}$ has ± 1 in each row and column once and once only and zero elsewhere.

Now $\mathbf{H}$ of the form (i) after some algebra gives Anderson's result ([1], page 1158). In the above two cases although the functions are not exactly special forms of the integrand in (1), they are equivalent forms. Hence the parallel results in both cases suggest a similar approach for this general integral (1) but unfortunately attempts in this direction proved futile. Hence we give an alternative approach to handle this general problem and give special results as occasions arise.

3. Maximization of $\mathscr{I}$ when l_i's are all distinct. Let $\mathbf{S}(p \times p)$ be a symmetric matrix and $C_\kappa(\mathbf{S})$ denote the zonal polynomial of the matrix $\mathbf{S}$ corresponding to the partition κ as defined by James [9]. Let us consider the integrand in (1), i.e. let

(6) $$f(\mathbf{H}) = {}_sF_t(a_1, \cdots, a_s; b_1, \cdots, b_t, \mathbf{AHRH'})\,.$$

Also let

(7) $$a_i \geqq \tfrac{1}{2}(p-1)\,, \qquad b_j \geqq \tfrac{1}{2}(p-1)\,, \qquad i = 1, \cdots, s, j = 1, \cdots, t\,.$$

Now, by James [9],

$$f(\mathbf{H}) = \sum_{k=0}^{\infty} \sum_\kappa \frac{(a_1)_\kappa \cdots (a_s)_\kappa}{(b_1)_\kappa \cdots (b_t)_\kappa} \frac{C_\kappa(\mathbf{AHRH'})}{k!}\,,$$

where $\kappa = (k_1, \cdots, k_p)$ is a partition of k and the multivariate hypergeometric coefficient $(a)_\kappa$ is given by

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

Under (7)

(8) $$\max_{\mathbf{H} \in O(p)} f(\mathbf{H}) = \max_{\mathbf{H} \in O(p)} \sum_{k=0}^{\infty} \sum_\kappa \frac{(a_1)_\kappa \cdots (a_s)_\kappa}{(b_1)_\kappa \cdots (b_t)_\kappa} \frac{C_\kappa(\mathbf{AHRH'})}{k!}$$
$$\leqq \sum_{k=0}^{\infty} \sum_\kappa \frac{(a_1)_\kappa \cdots (a_s)_\kappa}{(b_1)_\kappa \cdots (b_t)_\kappa} \max_{\mathbf{H} \in O(p)} \frac{C_\kappa(\mathbf{AHRH'})}{k!}\,.$$

Also

$$\min_{\mathbf{H}\in O(p)} f(\mathbf{H}) = \min_{\mathbf{H}\in O(p)} \sum_{k=0}^{\infty} \sum_{\kappa} \frac{(a_1)_\kappa \cdots (a_s)_\kappa}{(b_1)_\kappa \cdots (b_t)_\kappa} \frac{C_\kappa(\mathbf{AHRH}')}{k!}$$
$$\geqq \sum_{k=0} \sum_{\kappa} \frac{(a_1)_\kappa \cdots (a_s)_\kappa}{(b_1)_\kappa \cdots (b_t)_\kappa} \min_{\mathbf{H}\in O(p)} \frac{C_\kappa(\mathbf{AHRH}')}{k!}\,.$$

Now we prove the following lemma.

LEMMA 3.1. *Let* **A** *and* **R** *be* (*unlike in* (3))

$$\mathbf{A} = \operatorname{diag}(l_1, \cdots, l_p)\,, \qquad \mathbf{R} = \operatorname{diag}(r_1, \cdots, r_p) \tag{9}$$
$$\infty > l_1 > \cdots > l_p \geqq 0 \quad \textit{and} \quad \infty > r_1 > \cdots > r_p > 0\,.$$

Then

$$\max_{\mathbf{H}\in O(p)} C_\kappa(\mathbf{AHRH}') = \max_{\mathbf{H}\in O(p)} C_\kappa(\mathbf{H}'\mathbf{AHR}) = C_\kappa(\mathbf{AR})\,, \tag{10}$$

However, if **A** *and* **R** *satisfy* (3), *then*

$$\min_{\mathbf{H}\in O(p)} C_\kappa(\mathbf{AHRH}') = \min_{\mathbf{H}\in O(p)} C_\kappa(\mathbf{H}'\mathbf{AHR}) = C_\kappa(\mathbf{AR})\,.$$

H *in both cases is of the form* (i) i.e. **H** *has* ± 1 *in each row and column once and once only and zero elsewhere.*

PROOF. We prove the first part of the lemma regarding the maximum of the zonal polynomial. For complete proof refer to [4], [5]. The following proof is due to the referee. We have from Constantine [7]

$$C_\kappa(\mathbf{ST}) = d_{k,k} t_1^{k_1} \cdots t_p^{k_p} |\mathbf{S}_1|^{k_1-k_2} |\mathbf{S}_2|^{k_2-k_3} \cdots |\mathbf{S}_p|^{k_p} + \text{“lower terms”}$$

where

$$\mathbf{S}_i = (s_{rs})\,, \qquad r, s = 1, \cdots, i\,,$$

$(k_1, \cdots, k_p)$ is a partition of κ which is lexicographically the highest and $t_i = Ch_i(\mathbf{T})$, $i = 1, \cdots, p$, $Ch_i(\mathbf{B})$ denoting the ith characteristic root of **B**. Thus

$$C_\kappa(\mathbf{H}'\mathbf{AHR}') = d_{k,k} r_1^{k_1} \cdots r_p^{k_p} |(\mathbf{H}'\mathbf{AH})_1|^{k_1-k_2} \cdots |(\mathbf{H}'\mathbf{AH})_p|^{k_p} + \text{“lower terms,”}$$

where $(\mathbf{H}'\mathbf{AH})_i$ is defined as above. Hence

$$C_\kappa(\mathbf{H}'\mathbf{AHR}) \leqq d_{k,k}(r_1 l_1)^{k_1} \cdots (r_p l_p)^{k_p} + \text{“lower terms”} = C_\kappa(\mathbf{AR})\,.$$

As a further generalization of the above we consider

$$\mathbf{R} = \operatorname{diag}(r_1, \cdots, r_p)\,, \quad \infty > r_1 > \cdots > r_p > 0\,,$$
$$\mathbf{A} = \operatorname{diag}(l_1, \cdots, l_1, \overbrace{l_2, \cdots, l_2}^{k_2}, \cdots, \overbrace{l_m, \cdots, l_m}^{k_m}, l_{k_1+\cdots+k_m+1}, \cdots, l_p)\,,$$
$$\infty > l_1 > \cdots > l_m > l_{k_1+\cdots+k_m+1} > \cdots > l_p \geqq 0\,, \tag{11}$$

or alternately

$$\infty > l_p > \cdots > l_{k_1+\cdots+k_m+1} > l_m > \cdots > l_1 \geqq 0\,, \tag{12}$$
$$\mathbf{H} = \operatorname{diag}(\mathbf{H}_1, \cdots, \mathbf{H}_m, I_0(p - k_1 - \cdots - k_m))\,, \tag{13}$$

where $\mathbf{H}_j(k_j \times k_j)$ is an orthogonal matrix of order k_j, $j = 1, \cdots, m$ and $I_0(p - k_1 - \cdots - k_m) = \text{diag}(\pm 1, \cdots, \pm 1)$ of order $(p - k_1 - \cdots - k_m) \times (p - k_1 - \cdots - k_m)$. Then we have

LEMMA 3.2. *Under* (11)

$$\max_{\mathbf{H} \in O(p)} C_\kappa(\mathbf{H}'\mathbf{AHR}) = \max_{\mathbf{H} \in O(p)} C_\kappa(\mathbf{AHRH}') = C_\kappa(\mathbf{AR}),$$

under (12)

$$\min_{\mathbf{H} \in O(p)} C_\kappa(\mathbf{H}'\mathbf{AHR}) = \min_{\mathbf{H} \in O(p)} C_\kappa(\mathbf{AHRH}') = C_\kappa(\mathbf{AR})$$

and the optimum values are attained iff **H** *has the form* (13).

For detailed proof the reader is referred to [4], [5]. Now, corresponding to the Lemmas 3.1 and 3.2 we have the following two theorems:

THEOREM 3.1. *If* **A** *and* **R** *are given by* (3) *or* (9), *the class of orthogonal matrices for which* $f(\mathbf{H})$ *in* (6) *subject to* (7) *and for all* $\mathbf{R} > 0$ *is optimum is given by* $\mathbf{H} = \text{diag}(\pm 1, \cdots, \pm 1)$. *Further, when* (9) *is satisfied*

$$\max_{\mathbf{H} \in O(p)} f(\mathbf{H}) = \sum_{k=0}^{\infty} \sum_\kappa \frac{(a_1)_\kappa \cdots (a_s)_\kappa}{(b_1)_\kappa \cdots (b_t)_\kappa} \frac{C_\kappa(\mathbf{AR})}{k!},$$

and when (9) *is replaced by* (3)

$$\min_{\mathbf{H} \in O(p)} f(\mathbf{H}) = \sum_{k=0}^{\infty} \sum_\kappa \frac{(a_1)_\kappa \cdots (a_s)_\kappa}{(b_1)_\kappa \cdots (b_t)_\kappa} \frac{C_\kappa(\mathbf{AR})}{k!}.$$

Before stating the general form of this theorem let us observe some special cases.

COROLLARY 3.1. *If* $s = t = 0$ *in* (6) *then*

$$f(\mathbf{H}) = {}_0F_0(\mathbf{AHRH}') = \exp(\text{tr}\, \mathbf{AHRH}')$$

and under (9) *we get*

$$\max_{\mathbf{H} \in O(p)} f(\mathbf{H}) = \exp(\text{tr}\, \mathbf{RA}).$$

This is Anderson's result [1] mentioned earlier as Case 3. As a further application we consider

$$g(\mathbf{H}) = |\mathbf{I} + \mathbf{AHRH}'|^{-n} = {}_1F_0(n, -\mathbf{AHRH}'),$$

where $n \geqq \frac{1}{2}(p - 1)$ and **A** and **R** are defined as

$$(14) \qquad \mathbf{A} = \text{diag}(l_1, \cdots, l_p), \qquad \mathbf{R} = \text{diag}(r_1, \cdots, r_p),$$
$$\infty > l_p > \cdots > l_1 \geqq 0 \quad \text{and} \quad \infty > r_1 > \cdots > r_p > 0.$$

As it stands, Theorem 3.1 is not direclty applicable to this function. So we write following Khatri [11],

$$|\mathbf{I} + \mathbf{AHRH}'| = |\mathbf{I} + \mathbf{R}||\mathbf{I} - (\mathbf{I} - \mathbf{A})\mathbf{HR}(\mathbf{I} + \mathbf{R})^{-1}\mathbf{H}'|.$$

We now assume $Ch_i(\mathbf{A}) < 1$, $i = 1, \cdots, p$. This is no loss of generality since for $k > 0$

$$|\mathbf{I} + k\mathbf{AHRH}'| = \prod_{i=1}^{p} (1 + k\alpha_i)$$

where $\alpha_i = Ch_i(\mathbf{AHRH}') > 0$, $i = 1, \cdots, p$.

Thus the problem of finding the maximum or minimum of $|\mathbf{I} + \mathbf{AHRH}'|$ with respect to $\mathbf{H} \in O(p)$ is the same as that of $|\mathbf{I} + k\mathbf{AHRH}'|$. Hence

$$\begin{aligned} |\mathbf{I} + \mathbf{AHRH}'|^{-n} &= |\mathbf{I} + \mathbf{R}|^{-n} |\mathbf{I} - (\mathbf{I} - \mathbf{A})\mathbf{HR}(\mathbf{I} + \mathbf{R})^{-1}\mathbf{H}'|^{-n} \\ &= |\mathbf{I} + \mathbf{R}|^{-n} {}_1F_0(n, \mathbf{BHCH}') , \end{aligned}$$

where

$$\mathbf{B} = (\mathbf{I} - \mathbf{A}) = \operatorname{diag}(b_1, \cdots, b_p)$$

and

$$\mathbf{C} = \mathbf{R}(\mathbf{I} + \mathbf{R})^{-1} = \operatorname{diag}(c_1, \cdots, c_p) .$$

Hence from (14), we get

$$1 > b_1 > \cdots > b_p \geqq 0 \qquad \text{and} \qquad \infty > c_1 > \cdots > c_p > 0 .$$

Thus

$$g(\mathbf{H}) = |\mathbf{I} + \mathbf{R}|^{-n} {}_1F_0(n, \mathbf{BHCH}') ,$$

and now we can apply Theorem 3.1 and get the following corollary:

COROLLARY 3.2. *Under the conditions stated immediately above*

$$\begin{aligned} \max_{\mathbf{H} \in O(p)} g(\mathbf{H}) &= |\mathbf{I} + \mathbf{R}|^{-n} \max_{\mathbf{H} \in O(p)} {}_1F_0(n, \mathbf{BHCH}') \\ &= |\mathbf{I} + \mathbf{R}|^{-n} {}_1F_0(n, \mathbf{BC}) \\ &= |\mathbf{I} + \mathbf{AR}|^{-n} . \end{aligned}$$

This corresponds to Chang's result [2]. We now restate the above two results in a different form.

COROLLARY 3.3. *Let* (3) *hold. Then* $\sum_{i=1}^{p} l_i r_{ij}$ *and* $\prod_{i=1}^{p} (1 + l_i r_{ij})$ *are both minimized when* $r_{ij} = r_i$, $i = 1, \cdots, p$. *They are both maximized when* $r_{ij} = r_{p-i+1}$, $i = 1, \cdots, p$.

The latter two results are implicitly assumed in Anderson [1] and Chang [2]. In fact we can go a step further and get the following: Let f be a nonnegative, nondecreasing function defined on $[0, \infty]$. Then, under (3),

$$f(\textstyle\sum_{i=1}^{p} l_i r_{ij}) \leqq f(\sum_{i=1}^{p} l_i r_i) \qquad \text{and} \qquad f(\prod_{i=1}^{p} (1 + l_i r_{ij})) \leqq f(\prod_{i=1}^{p} (1 + l_i r_i)) .$$

These results follow directly from the above discussion but are mentioned separately, since they cover a broader ground in the sense that with modification, the results apply to positive convex combinations of two symmetric matrix functions.

We now state the general form of Theorem 3.1.

THEOREM 3.2. *Let* $\mathbf{A}$ *and* $\mathbf{R}$ *satisfy* (11) *or* (12). *Then the class of orthogonal* *matrices for which* $f(\mathbf{H})$ *in* (6) *subject to* (7) *and for all* $\mathbf{R} > 0$ *is optimum is given by* (13). *Further when* (11) *is satisfied*

$$\max_{\mathbf{H}\in O(p)} f(\mathbf{H}) = \sum_{k=0}^{\infty} \sum_{\kappa} \frac{(a_1)_\kappa \cdots (a_s)_\kappa}{(b_1)_\kappa \cdots (b_t)_\kappa} \frac{C_\kappa(\mathbf{AB})}{k!}$$

and when (11) *is replaced by* (12)

$$\min_{\mathbf{H}\in O(p)} f(\mathbf{H}) = \sum_{k=0}^{\infty} \sum_{\kappa} \frac{(a_1)_\kappa \cdots (a_s)_\kappa}{(b_1)_\kappa \cdots (b_t)_\kappa} \frac{C_\kappa(\mathbf{AR})}{k!} .$$

This is the main result which will be used in the sequel to obtain asymptotic expansions of $\mathscr{I}$ in (1) under different situations.

4. Asymptotic expansion for canonical correlation-population roots all distinct. Let $X_1, \cdots, X_p, X_{p+1}, \cdots, X_{p+q}$, $p \leqq q$ be distributed $N(\mathbf{0}, \mathbf{\Sigma})$, where

$$\mathbf{\Sigma} = \begin{matrix} p \\ q \end{matrix}\begin{pmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{12}' & \mathbf{\Sigma}_{22} \end{pmatrix}_{\substack{\;p\quad\; q}} . \tag{15}$$

Let ρ_i^2, $i = 1, \cdots, p$, be the roots of

$$|\mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{21} - \rho^2\mathbf{\Sigma}_{11}| = 0 , \tag{16}$$

and $\hat{\rho}_i^2$, $i = 1, \cdots, p$, be the maximum likelihood estimate (m.l.e.) of ρ_i^2, $i = 1, \cdots, p$, from a sample of size $n \geqq p + q$. For the sake of uniformity of notation let us put $l_i = \rho_i^2$ and $r_i = \hat{\rho}_i^2$, $i = 1, \cdots, p$, and let

$$\begin{gathered} \mathbf{R} = \operatorname{diag}(r_1, \cdots, r_p) , \qquad \mathbf{A} = \operatorname{diag}(l_1, \cdots, l_p) , \\ 1 > l_1 > \cdots > l_p \geqq 0 \quad \text{and} \quad 1 > r_1 > \cdots > r_p > 0 . \end{gathered} \tag{17}$$

Then the joint density of elements of $\mathbf{R}$ is given as ([7], [9]),

$$D_1 \int_{O(p)} {}_2F_1(\tfrac{1}{2}n, \tfrac{1}{2}n, \tfrac{1}{2}q, \mathbf{AHRH}') \, d(\mathbf{H}) , \tag{18}$$

$d(\mathbf{H})$ is defined in (1) and

$$\begin{aligned} D_1 = \{&\pi^{p^2/2}\,\Gamma_p(\tfrac{1}{2}n)/\Gamma_p(\tfrac{1}{2}q)\Gamma_p(\tfrac{1}{2}(n-q))\Gamma_p(\tfrac{1}{2}p)\} \\ &\times |\mathbf{I} - \mathbf{A}|^{n/2}|\mathbf{R}|^{\frac{1}{2}(q-p-1)}|\mathbf{I} - \mathbf{R}|^{\frac{1}{2}(n-p-q-1)} \textstyle\prod_{i<j} (r_i - r_j) . \end{aligned} \tag{19}$$

The density (18) involves an integral and following Anderson [1], Chang [2], and Li and Pillai [12], [14], we try to maximize this integral. Let us now denote the integral by

$$E = \int_{O(p)} {}_2F_1(s, s, t, \mathbf{AHRH}') \, d(\mathbf{H}) , \tag{20}$$

where for notational simplicity we put $s = \frac{1}{2}n$, $t = \frac{1}{2}q$. Now with a mild restriction on s we get by Theorem 3.1 that for variations of $\mathbf{H} \in O(p)$, ${}_2F_1(s, s, t, \mathbf{AHRH}')$ is maximized when $\mathbf{H}$ has the form (i) of Case 3 and the optimum ${}_2F_1(s, s, t, \mathbf{AR})$. In order to obtain an asymptotic expansion for the density of sample roots we proceed as follows: First we use Kummer's formula and get

$${}_2F_1(s, s, t, \mathbf{AHRH}') = |\mathbf{I} - \mathbf{AHRH}'|^{-(2s-t)}\,{}_2F_1((t-s), (t-s), t, \mathbf{AHRH}') . \tag{21}$$

Now varying $\mathbf{H}$ around $N(\mathbf{I})$, a neighborhood of $\mathbf{I}(p \times p)$, i.e. varying $\mathbf{AHRH'}$ around $\mathbf{AR}$ we get

$$(22)\qquad {}_2F_1((t-s),(t-s),t,\mathbf{AHRH'}) = {}_2F_1((t-s),(t-s),t,\mathbf{AR}) + O(\varepsilon)\,.$$

We prove below a more general result.

LEMMA 4.1. *If* $\mathbf{H} \in N(\mathbf{I})$, $a_i \geqq \frac{1}{2}(p-1)$, $b_j \geqq \frac{1}{2}(p-1)$, $i = 1, \cdots, \mu$, $j = 1, \cdots, \eta$, *then*

$$ {}_\mu F_\eta(a_1, \cdots, a_\mu, b_1, \cdots, b_\eta, \mathbf{AHRH'}) = {}_\mu F_\eta(a_1, \cdots, a_\mu, b_1, \cdots, b_\eta, \mathbf{AR}) + O(\varepsilon)\,,$$

provided $t_i - \varepsilon \leqq Ch_i(\mathbf{AHRH'}) \leqq t_i + \varepsilon$, *where* $t_i = Ch_i(\mathbf{AR})$, $i = 1, \cdots, p$.

PROOF. Let $f(\mathbf{H}) = {}_\mu F_\eta(a_1, \cdots, a_\mu, b_1, \cdots, b_\eta, \mathbf{AHRH'})$. Now $f(\mathbf{H})$ is an increasing function of each of its characteristic roots. Thus varying $\mathbf{H} \in N(\mathbf{I})$, we note that the first partial derivative of $f(\mathbf{H})$ with respect to each characteristic root exists, except possibly over a set of zero measure. Agian as $f(\mathbf{H})|_{\mathbf{H}=\mathbf{I}}$ exists, the mean value theorem applies and hence the lemma.

Now an application of Lemma 4.1 in the last expression in formula (21) gives (22). Following Anderson [1], Chang [2], Li and Pillai [12], [14], and using (22) we get for large values of $(2s - t)$

$$E = 2^p \int_{N(\mathbf{I})} |\mathbf{I} - \mathbf{AHRH'}|^{-(2s-t)}\, d(\mathbf{H})\, {}_2F_1((t-s),(t-s),t,\mathbf{AR}) + O(\varepsilon)\,.$$

Further, we consider

$$(23)\qquad F = 2^p \int_{N(\mathbf{I})} |\mathbf{I} - \mathbf{AHRH'}|^{-(2s-t)}\, d(\mathbf{H})\,.$$

The integrand in (23) is quite similar to that of Chang [2] and hence using the technique given in his paper as modified by Li and Pillai [12], [14] we get Theorem 4.1 below. For details the reader is referred to [4], [5], and the references mentioned therein.

THEOREM 4.1. *For large n, an asymptotic expansion of the distribution of* $r_1, \cdots, r_p$ (m.l.e. *of the squares of population canonical correlation coefficients*) *where* $1 > r_1 > \cdots > r_p > 0$ *and the population characteristic roots from* (16) *are such that* $1 > l_1 > \cdots > l_p \geqq 0$ *is given by*

$$D_1 \prod_{i<j} \left(\frac{2\pi}{(2n-q)c_{ij}}\right)^{\frac{1}{2}} |\mathbf{I} - \mathbf{AR}|^{-\frac{1}{2}(2n-q)} \left\{1 + \frac{1}{2(2n-q)} \times \left[\sum_{i<j} c_{ij}^{-1} + \alpha(p)\right] + \cdots\right\} {}_2F_1(\tfrac{1}{2}(q-n), \tfrac{1}{2}(q-n), \tfrac{1}{2}q, \mathbf{AR}) + O(\varepsilon)$$

where $\mathbf{R}$ *and* D_1 *are given by* (17) *and* (19) *respectively*,

$$\alpha(p) = p(p-1)(2p+5)/12\,, \qquad c_{ij} = (t_{ij} - t_i t_j r_{ij}) r_{ij} = c_{ji}\,,$$
$$t_{ij} = t_i - t_j\,, \quad r_{ij} = r_i - r_j \quad \text{and} \quad t_i = l_i(1 - l_i r_i)^{-1}\,, \qquad i, j = 1, \cdots, p\,.$$

5. Asymptotic expansion for canonical correlation—several multiple population roots. Let $\mathbf{A}$ and $\mathbf{R}$ be as in (17), however, $\mathbf{A}$ being modified as

$$\text{(24)} \qquad \mathbf{A} = \operatorname{diag}(l_1, \cdots l_d, \overbrace{l_{d+1}, \cdots, l_{d+1}}^{k_1}, \cdots, \overbrace{l_{d+m}, \cdots, l_{d+m}}^{k_m}) \quad \text{and} \quad 1 > l_1 > \cdots > l_d > l_{d+1} > \cdots > l_{d+m} \geqq 0 \, ;$$

then, following Li and Pillai [13], we obtain the extension of Theorem 4.1. This result is available in [5] and in Chattopadhyay and Pillai [6].

6. Asymptotic expansion for MANOVA—population roots all distinct. Let $\mathbf{B}(p \times p)$ have a noncentral Wishart distribution with s df and matrix of noncentrality parameter $\mathbf{A}_1$ and $\mathbf{W}$ have a central Wishart distribution with t df, covariance matrix in each case being $\mathbf{\Sigma}$, and $\mathbf{A}_1 = \frac{1}{2}\boldsymbol{\mu}\boldsymbol{\mu}'\mathbf{\Sigma}^{-1}$ where $\boldsymbol{\mu}(p \times s)$ is a matrix of mean vectors. Then the probability density function of the roots of $R_1 = \mathbf{B}(\mathbf{B} + \mathbf{W})^{-1}$ is given by [7],

$$T_1 \int_{O(p)} {}_1F_1(\tfrac{1}{2}(s+t), \tfrac{1}{2}s, \mathbf{AHRH}') \, d(\mathbf{H}) \, ,$$

where

$$\text{(25)} \qquad T_1 = \pi^{p^2/2} \Gamma_p(\tfrac{1}{2}(s+t))\{\Gamma_p(\tfrac{1}{2}t)\Gamma_p(\tfrac{1}{2}s)\Gamma_p(\tfrac{1}{2}p)\}^{-1} \exp[-\operatorname{tr} \mathbf{A}] \times (\textstyle\prod_{i=1}^p r_1)^{\frac{1}{2}(s-p-1)} \prod_{i=1}^p (1-r_i)^{\frac{1}{2}(t-p-1)} \prod_{i<j} (r_i - r_j) \, ,$$

where

$$\text{(26)} \qquad \begin{aligned} &\mathbf{R} = \operatorname{diag}(r_1, \cdots, r_p) \, , \quad 1 > r_1 > \cdots > r_p > 0 \, , \\ &\mathbf{A} = \operatorname{diag}(l_1, \cdots, l_p) \, , \quad \infty > l_1 > \cdots > l_p \geqq 0 \, , \\ &l_i = Ch_i(\mathbf{A}_1) \quad \text{and} \quad r_i = Ch_i(\mathbf{R}_1) \, , \qquad i = 1, \cdots, p \, . \end{aligned}$$

As earlier, we consider

$$E_1 = \int_{O(p)} {}_1F_1(\tfrac{1}{2}(s+t), \tfrac{1}{2}s, \mathbf{AHRH}') \, d(\mathbf{H}) \, .$$

The integrand as it stands is not easy to work with, hence we apply to the integrand the confluence relation (James [9]):

$$\text{(27)} \qquad \lim_{c \to \infty} {}_2F_1(a, c, b, c^{-1}\mathbf{S}) = {}_1F_1(a, b, \mathbf{S}) \, .$$

Applying the dominated convergence theorem, since the functions involved are well defined and satisfy the conditions of the theorem, we get, using (27),

$$\begin{aligned} \lim_{a \to \infty} \int_{O(p)} {}_2F_1(\tfrac{1}{2}(s+t), a, \tfrac{1}{2}s, a^{-1}\mathbf{AHRH}') \, d(\mathbf{H}) &= \int_{O(p)} \lim_{a \to \infty} {}_2F_1(\tfrac{1}{2}(s+t), a, \tfrac{1}{2}s, a^{-1}\mathbf{AHRH}') \, d(\mathbf{H}) \\ &= \int_{O(p)} {}_1F_1(\tfrac{1}{2}(s+t), \tfrac{1}{2}s, \mathbf{AHRH}') \, d(\mathbf{H}) \, . \end{aligned}$$

Thus for evaluating E_1 we consider, for large a,

$$E_2 = \int_{O(p)} {}_2F_1(\tfrac{1}{2}(s+t), a, \tfrac{1}{2}s, a^{-1}\mathbf{AHRH}') \, d(\mathbf{H}) \, .$$

We now apply the earlier technique with slight modification which, together with other details, may be referred to [4], [5]. Thus after some algebra we obtain the following theorem:

THEOREM 6.1. *For large t (and hence for large sample size), an asymptotic expansion for the distribution of the characteristic roots of* $\mathbf{R}_1$ *with parameter matrix* $\mathbf{A}_1$ *where* $\mathbf{R}$ *and* $\mathbf{A}$ *satisfy* (26) *is given by*

$$T_1 2^p \prod_{i<j=1}^{p} \left(\frac{2\pi}{tc_{ij}}\right)^{\frac{1}{2}} \left\{1 + \frac{1}{2t}\left[\sum_{i<j} c_{ij}^{-1} + \alpha(p)\right] + \cdots\right\} \times \exp[\operatorname{tr} \mathbf{AR}]\,{}_1F_1(-\tfrac{1}{2}t, \tfrac{1}{2}s, -\mathbf{AR}) + O(\varepsilon)\,,$$

where T_1 *is given by* (25) *and other constants as defined earlier.*

7. Asymptotic expansion for MANOVA—several multiple population roots. Let $\mathbf{A}$ and $\mathbf{R}$ be as defined in (26), however, $\mathbf{A}$ being modified as

$$\mathbf{A} = \operatorname{diag}(l_1, \cdots, l_d, \overbrace{l_{d+1}, \cdots, l_{d+1}}^{k_1}, \cdots, \overbrace{l_{d+m}, \cdots, l_{d+m}}^{k_m}), \qquad \infty > l_1 > l_2 > \cdots > l_d > l_{d+1} > \cdots > l_{d+m} \geqq 0\,. \tag{28}$$

Then, following Li and Pillai [13], we have obtained the extension of Theorem 6.1 which is available in the references given above for the extension of Theorem 4.1.

8. Complex analogues of previous results. All the above results extend to the complex case as well. Corresponding to each theorem in the real case above we may obtain the complex counterpart. These results in the complex case are available in [3].

9. Remarks. In this section, we make the following remarks:

1. The method as outlined above is a generalization of Anderson's result [1] and all his comments apply here also.

2. In approximating ${}_\mu F_\eta$ by Kummer's formula we note that if $N(I)$ involved in each case be sufficiently close to $\mathbf{I}$, which is possible for large enough sample size, we can neglect $O(\varepsilon)$ in each case for good enough approximation.

3. The ordering of roots in each case is immaterial as shown in [4], and as such the only restriction is that the roots of the sample matrix and those of population matrix be ordered in the same direction.

4. From the previous remark it may be seen that the expansion for one extreme multiple population root covers the largest root case although results given in this paper are for the smallest.

5. Each formula, as displayed, gives a considerable simplification in ${}_\mu F_\eta$ function since each population root goes with its sample counterpart.

6. In the real case when a in $(a)_\kappa$ is a negative integer, the hypergeometric function involved reduces to a polynomial.

7. When all population roots are equal, we note that $O(\varepsilon)$ term is identically zero. Here we take any empty product as unity.

REFERENCES

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

[2] 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.

[3] CHATTOPADHYAY, A. K. (1971). An asymptotic distribution theory and applications in multivariate analysis. Mimeo Series No. 256, Department of Statistics, Purdue Univ.

[4] CHATTOPADHYAY, A. K. and PILLAI, K. C. S. (1970). On the maximization of an integral of a matrix function over the group of orthogonal matrices. Mimeo Series No. 248, Department of Statistics, Purdue Univ.

[5] CHATTOPADHYAY, A. K. and PILLAI, K. C. S. (1970). Asymptotic expansions of the distributions of characteristic roots in MANOVA and canonical correlation. Mimeo Series No. 249, Department of Statistics, Purdue Univ.

[6] 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 117–127 (P. R. Krishnaiah, ed.). Academic Press, New York.

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

[8] DEEMER, W. L. and OLKIN, I. (1951). The Jacobian of certain matrix transformations useful in multivariate analysis. *Biometrika* **38** 345–367.

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

[10] JAMES, A. T. (1968). Tests of equality of latent roots of the covariance matrix. In *Multivariate Analysis* II 205–218 (P. R. Krishnaiah, ed.). Academic Press, New York.

[11] KHATRI, C. G. (1967). Some distribution problems connected with the characteristic roots of $S_1S_2^{-1}$. *Ann. Math. Statist.* **38** 944–948.

[12] LI, H. C. and PILLAI, K. C. S. (1969). Asymptotic expansions for the distribution of roots of two matrices from classical and complex Gaussian populations. Mimeo Series No. 188, Department of Statistics, Purdue Univ.

[13] LI, H. C. and PILLAI, K. C. S. (1970). Asymptotic expansions of the distributions of the characteristic roots of $S_1S_2^{-1}$ when population roots are not all distinct. Mimeo Series No. 231, Department of Statistics, Purdue Univ.

[14] LI, H. C., PILLAI, K. C. S. and CHANG, T. C. (1970). Asymptotic expansions for distributions of roots of two matrics from classical and complex Gaussian populations. *Ann. Math. Statist.* **41** 1541–1556.

A. K. CHATTOPADHYAY
INDIAN STATISTICAL INSTITUTE
CALCUTTA, INDIA

K. C. S. PILLAI
DEPARTMENT OF STATISTICS
MATHEMATICAL SCIENCES BUILDING
PURDUE UNIVERSITY
WEST LAFAYETTE, INDIANA 47907

HUNG C. LI
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF SOUTHERN COLORADO
PUEBLO, COLORADO 81001