

On the Maximization of an Integral of a Matrix Function
Over the Group of Orthogonal Matrices*

by

A. K. Chattopadhyay and K.C.S. Pillai

Department of Statistics

Division of Mathematical Sciences

Mimeograph Series No. 248

December, 1970

* This work is supported by the National Science Foundation Grant No. GP-11473.

In the complex analogue of both one sample and two sample cases, Li and Pillai, [8], [9], obtained a similar form of the unitary matrix U . The purpose of this paper is to generalize their results both in the complex and real situations with $a_1, \dots, a_s, b_1, \dots, b_t$, satisfying some suitable conditions.

We develop the idea in a series of lemmas and theorems and show that the results of Anderson [1], Chang [2], James [5], Li and Pillai, [8], [9], are special cases of our results. The generalization has not only been in regard to ${}_sF_t$ - hypergeometric functions but also when λ_i 's are equal within each of several sets. We have further proved that the integral under different forms of the matrix A is invariant under choice of different submatrices of H and our general results cover some earlier ones of the above authors.

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

Lemma 1. Let $f(T)$ be a real valued function of the elements of the matrix $T(p \times p) = (t_{ij})$. Then

$$d f(T) = \text{tr}(Q d T)$$

where

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

But $(H'DH)$ is a skew symmetric matrix. Hence for all $R > 0$

$$(5) \Rightarrow R H'A H \text{ is symmetric}$$

$$\Rightarrow R H'A H = H'A H R$$

$$\Rightarrow H'A H = \text{diag}(\mu_1, \dots, \mu_p)$$

as R in (3) is diagonal with distinct roots

$$\Rightarrow H \text{ has the form}$$

(i) H has $+1$ in each row and column once and once only and zero elsewhere.

Now taking H of the form (i) after some algebra gives Anderson's result [1].

In the above two corollaries, the function we have considered though not exactly special forms of the integrand in (1) but are equivalent forms. Hence the parallel results in both the cases suggest a similar approach for this general integral (1) but unfortunately attempt 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 I when λ_i 's are all distinct. At first we prove a lemma which will be used in the sequel. Let $S(p \times p)$ be a symmetric matrix, $Ch_i(S)$ denotes the i th characteristic root of S and $C_\kappa(S)$ stand for the zonal polynomial of the matrix S corresponding to the partition κ as defined by James [4]. Then we state the following lemma.

Lemma 2. Let $ch_i(S) \geq 0$, $i = 1, \dots, p$. Then $C_\kappa(S)$ is nonnegative and increasing in each characteristic root of S .

This may be shown by using the differential equation given by James [6], since a) the coefficients of all terms of a zonal polynomial when expressed

in terms of monomial symmetric functions of the characteristic roots of the matrix are positive and b) zonal polynomials are themselves symmetric functions of the characteristic roots.

Now let us consider the integrand in (4), i.e. let

$$(6) \quad f(H) = {}_sF_t(a_1, \dots, a_t; b_1, \dots, b_t, A H R H')$$

Also let

$$(7) \quad a_i \geq (1/2)(p-1), \quad b_j \geq (1/2)(p-1), \quad i = 1, \dots, s, \quad j = 1, \dots, t.$$

Now, by James [4]

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

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

$$(a)_{\kappa} = \prod_{i=1}^p (a - (1/2)(i-1))_{k_i}$$

and

$$(a)_k = a(a+1) \dots (a+k-1)$$

Under (7)

$$(8) \quad \max_{H \in 0(p)} f(H) = \max_{H \in 0(p)} \sum_{k=0}^{\infty} \sum_{\kappa} \frac{(a_1)_{\kappa} \cdots (a_s)_{\kappa}}{(b_1)_{\kappa} \cdots (b_t)_{\kappa}} \frac{C_{\kappa}(A \ H \ R \ H')}{k!}$$

$$\leq \sum_{k=0}^{\infty} \sum_{\kappa} \frac{(a_1)_{\kappa} \cdots (a_s)_{\kappa}}{(b_1)_{\kappa} \cdots (b_t)_{\kappa}} \max_{H \in 0(p)} \frac{C_{\kappa}(A \ H \ R \ H')}{k!}$$

Now to proceed any further we have to consider the maximization problem involved in (8). To this end we proceed as follows.

For A and R in (1) let us take (unlike in (3))

$$(9) \quad \infty > \lambda_1 > \lambda_2 > \dots > \lambda_p \geq 0 \quad \text{and} \quad \infty > r_1 > r_2 > \dots > r_p > 0.$$

(The ordering and labeling of λ_i 's and r_i 's will be done in different ways as may be necessary. A and R are more or less used in a generic sense in order to avoid the use of too many symbols.)

Let us consider $C_{\kappa}(H'A \ H \ R) = C_{\kappa}(H \ R \ H'A)$, where $C_{\kappa}(Z)$ is the zonal polynomial corresponding to the partition κ as defined by James [4] and $H \in 0(p)$. Let $Ch_i(Z)$ denote the i th characteristic root of Z .

Then

$$\begin{aligned} Ch_1(H \ R \ H'A) &\leq Ch_1(H \ R \ H') Ch_1(A) \\ &= Ch_1(R) Ch_1(A) \\ &= Ch_1(R \ A) \\ &= r_1 \lambda_1. \end{aligned}$$

Now if we take

$$(10) \quad \underline{H} = \begin{pmatrix} +1 & 0 \\ 0 & \underline{H}_2 \end{pmatrix},$$

where $\underline{H}_2(p-1)$ is an orthogonal matrix of order $p-1$, then

$$\text{Ch}_1(\underline{H} \underline{R} \underline{H}' \underline{A}) = \text{Ch}_1 \begin{pmatrix} r_1 \ell_1 & 0 \\ 0 & \underline{B} \end{pmatrix} = r_1 \ell_1$$

where

$$\underline{B} = \underline{H}_2 \underline{R}_2 \underline{H}'_2 \underline{A}_2$$

and

$$(11) \quad \underline{R}_2 = \text{diag}(r_2, \dots, r_p), \quad \underline{A}_2 = \text{diag}(\ell_2, \dots, \ell_p)$$

Again let

$$(12) \quad \underline{S}^2 = \underline{R} \quad \text{and} \quad \underline{S} = \text{diag}(s_1, \dots, s_p)$$

Hence we consider the matrix

$$\underline{S} \underline{H}' \underline{A} \underline{H} \underline{S} = \underline{C} = (c_{ij}) \quad (\text{say}).$$

Let

$$(13) \quad \underline{H} = (h_{ij})$$

Then

$$\underline{C} = (c_{ij}) = (s_i s_j \sum_{k=1}^p h_{ki} h_{kj} \ell_k)$$

Let us now take $e'_i = (0, \dots, 0, 1, 0, \dots, 0)$, $i = 1, \dots, p$ (in the standard notation, i.e. 1 in the i th place and zero elsewhere).

Then

$$e'_1 C e_1 = r_1 \sum_{k=1}^p h_{k1}^2 \ell_k < r_1 \ell_1, \quad (14)$$

iff $(h_{21}, \dots, h_{p1}) \neq 0$.

Also

$$e'_i C e_i = r_i \sum_{k=1}^p h_{ki}^2 \ell_k \leq r_i \ell_1, \quad i = 2, \dots, p.$$

Let

$$X' = (x_1, \dots, x_p) \rightarrow X' X = 1. \quad (15)$$

Then

$$\begin{aligned} (16) \quad |X' C X| &\leq \sum_{i,j} |x_i| |x_j| |c_{ij}| \\ &= \sum_i |x_i|^2 |c_{ii}| + \sum_{i \neq j} |x_i| |x_j| |c_{ij}| \\ &\leq \sum_i |x_i|^2 |c_{ii}| + \sum_{i \neq j} |x_i| |x_j| \sqrt{|c_{ii}| |c_{jj}|} \\ &= \left(\sum_{i=1}^p |x_i| \sqrt{|c_{ii}|} \right)^2 \\ &\leq \ell_1 \left(\sum_{i=1}^p |x_i| \sqrt{r_i} \right)^2 \\ &= r_1 \ell_1 \left(|x_1| + |x_2| \sqrt{\frac{r_2}{r_1}} + \dots + |x_p| \sqrt{\frac{r_p}{r_1}} \right)^2. \end{aligned}$$

Now for further reference we quote a theorem.

Theorem (Courant-Fischer). Let $D(p \times p)$ be a symmetric matrix with characteristic roots $\lambda_1 \geq \dots \geq \lambda_p$. Then they may be defined as

$$\lambda_1 = \max_{\tilde{X}} (\tilde{X}' D \tilde{X}) / (\tilde{X}' \tilde{X}),$$

$$\lambda_j = \min_{\substack{(\tilde{Y}'_i \tilde{Y}_i) = 1 \\ (i=1, \dots, j-1)}} \max_{\substack{(\tilde{X}' \tilde{Y}_i) = 0}} (\tilde{X}' D \tilde{X}) / (\tilde{X}' \tilde{X})$$

$$j = 2, \dots, p$$

or equivalently

$$\lambda_p = \min_{\tilde{X}} (\tilde{X}' D \tilde{X}) / (\tilde{X}' \tilde{X}),$$

$$\lambda_j = \max_{\substack{(\tilde{Y}'_i \tilde{Y}_i) = 1 \\ i=1, \dots, j-1}} \min_{\substack{(\tilde{X}' \tilde{Y}_i) = 0}} (\tilde{X}' D \tilde{X}) / (\tilde{X}' \tilde{X})$$

$$j = 1, \dots, p-1,$$

where \tilde{X}, \tilde{Y}_i are column vectors in E_p , the Euclidean space of p dimension.

Now if $R > 0$ is such that r_i/r_1 is negligibly small ($i = 2, \dots, p$) then applying the above theorem and from (16), we get for all $R > 0$

$$\max_{H \in O(p)} \text{Ch}_1(H R A H') = \lambda_1 r_1$$

iff equality holds in (14) i.e. iff

$$(17) \quad (h_{21}, \dots, h_{p1}) = 0$$

Now since \underline{H} is orthogonal

$$(17) \Rightarrow h_{11} = \pm 1$$

Thus for all $R > 0$

$$\max_{\underline{H} \in O(p)} \text{Ch}_1(\underline{H} \underline{R} \underline{H}' \underline{A}) = \lambda_1 r_1$$

iff

\underline{H} has the form (10) .

If \underline{H} has the form (10)

$$\underline{S} \underline{H} \underline{A} \underline{H}' \underline{S} = \begin{pmatrix} r_1 \lambda_1 & 0 \\ 0 & \underline{B}_1 \end{pmatrix}$$

where

$$\underline{B}_1 = \underline{S}_2 \underline{H}' \underline{A} \underline{H} \underline{S}_2 \quad \text{and} \quad \underline{S}_2 = \text{diag} (s_2, \dots, s_p)$$

Thus characteristic vector corresponding to the root $r_1 \lambda_1$ is proportional to e_1 and hence any vector belonging to E_p , Euclidean space of p dimensions, and orthogonal to e_1 is generated by e_2, \dots, e_p . Thus when H has the form (10), the problem of finding the second maximum characteristic root of $S H' A H S$ simply reduces to finding the maximum characteristic root of B_1 for $H_2 \in O(p-1)$. Thus we proceed step by step as before, only that we are now dealing with matrices of one less dimension. We also note the following

$$\begin{aligned}
 (19) \quad \text{Ch}_i(H R H' A) &\leq \text{Ch}_1(A) \text{Ch}_i(H R H') \\
 &= \text{Ch}_1(A) \text{Ch}_i(R) \\
 &= \lambda_1 r_i \quad i = 1, \dots, p
 \end{aligned}$$

Thus from the above discussion and from (19), using the fact that zonal polynomials are symmetric functions of the characteristic roots of the matrix and the monotonicity property of zonal polynomial as proved in lemma 2, we get the following lemma.

Lemma 3. When (9) holds, for all $R > 0$

$$(20) \quad \max_{H \in O(p)} C_K(H' A H R) = \max_{H \in O(p)} C_K(H R H' A) = C_K(A R)$$

If A is as in (1) and

$$(21) \quad \infty > \lambda_p > \lambda_{p-1} > \dots > \lambda_1 \geq 0, \text{ then}$$

$$(22) \quad \min_{H \in O(p)} C_K(H' A H R) = \min_{H \in O(p)} C_K(H R H' A) = C_K(A R),$$

and

$$(23) \quad \max_{\tilde{H} \in O(p)} C_K(\tilde{H}'A \tilde{H} R) = \max_{\tilde{H} \in O(p)} C_K(\tilde{H} R \tilde{H}'A) = C_K(A \tilde{E}),$$

where

$$\tilde{E} = \text{diag}(r_p, \dots, r_1).$$

The maximum in (20) and minimum in (22) are attained when and only when \tilde{H} has the form

$$(ii) \quad \tilde{H}(p \times p) = \text{diag}(\pm 1, \dots, \pm 1).$$

In (23) \tilde{H} has the form: $\tilde{H} = \tilde{H}_1 \tilde{D}$ where \tilde{H}_1 has the form (ii) and

$$\tilde{D} = (\tilde{e}_p, \dots, \tilde{e}_1).$$

In proving (22) regarding the minimum value of the zonal polynomial we used the following relations:

$$(24) \quad \begin{aligned} \text{Ch}_i(\tilde{H}'A \tilde{H} R) &= \text{Ch}_i(\tilde{H} R \tilde{H}'A) \\ &\geq \text{Ch}_i(\tilde{H} R \tilde{H}') \text{Ch}_p(A) \\ &= \text{Ch}_i(R) \text{Ch}_p(A) \\ &= r_i^{\ell_p}, \quad i = 1, \dots, p, \end{aligned}$$

and we proceed exactly as in the maximization problem but in this case using the maxi-mini characterization of Courant-Fischer theorem. In case of (23) we note that as $\tilde{H} \in O(p)$, $\tilde{H}\tilde{D} \in O(p)$ where $\tilde{D} = (\tilde{e}_p, \dots, \tilde{e}_1)$ and \tilde{e}_i 's are defined earlier.

Also the mapping $\underline{H} \rightarrow \underline{H} \underline{D}$ is one-to-one and onto and hence in (23) instead of considering $C_{\kappa}(\underline{H} \underline{R} \underline{H}' \underline{A})$ as $\underline{H} \in 0(p)$ we consider $C_{\kappa}((\underline{H} \underline{D}) \underline{R} (\underline{H} \underline{D})' \underline{A})$ as $\underline{H} \in 0(p)$ or equivalently we consider

$$C_{\kappa}(\underline{H} \underline{E} \underline{H}' \underline{A}) \quad \text{as} \quad \underline{H} \in 0(p) \quad \text{where}$$

$$\underline{E} = \underline{D} \underline{R} \underline{D}'$$

In the above discussions we note that if \underline{A} and \underline{R} has the same ordering of the element then $\underline{A} \underline{R}$ corresponds to the maximization problem and when their ordering is reversed it gives the minimum formulation.

Thus we use (20) in (8) and we get the following theorem.

Theorem 1. If \underline{A} and \underline{R} are as given in (9), the class of orthogonal matrices for which $f(\underline{H})$ in (6) subject to (7) and for all $\underline{R} > 0$ is a maximum, is given by (ii) and

$$(25) \quad \max_{\underline{H} \in 0(p)} f(\underline{H}) = \sum_{k=0}^{\infty} \sum_{\kappa} \frac{(a_1)_{\kappa} \cdots (a_s)_{\kappa}}{(b_1)_{\kappa} \cdots (b_t)_{\kappa}} \frac{C_{\kappa}(\underline{A} \underline{R})}{k!}$$

If the ordering of λ_i 's in (9) is replaced by (21)

$$(26) \quad \min_{\underline{H} \in 0(p)} f(\underline{H}) = \sum_{k=0}^{\infty} \sum_{\kappa} \frac{(a_1)_{\kappa} \cdots (a_s)_{\kappa}}{(b_1)_{\kappa} \cdots (b_t)_{\kappa}} \frac{C_{\kappa}(\underline{A} \underline{R})}{k!}$$

with \underline{H} again taking the same form (ii).

This is one of the basic results in the paper and we will subsequently generalize it to more complex cases. But first we give some special cases.

Corollary 1.1. If $s=t=0$ in (6) then

$$f(\underline{H}) = {}_0F_0(\underline{A} \underline{H} \underline{R} \underline{H}') = \exp[\text{tr } \underline{A} \underline{H} \underline{R} \underline{H}'] ,$$

and under (21) and \underline{R} as in (9), we get for all $\underline{R} > 0$

$$\max_{\underline{H} \in O(p)} f(\underline{H}) = \exp[\text{tr } \underline{E} \underline{A}] ,$$

where

$$\underline{E} = \text{diag}(r_p, \dots, r_1) .$$

This is Anderson's result [1] mentioned earlier as case 3.

As a second application of our theorem 1, let us consider the following.

Let

$$g(\underline{H}) = |\underline{I} + \underline{A} \underline{H} \underline{R} \underline{H}'|^{-n} = {}_1F_0(n; \underline{A} \underline{H} \underline{R} \underline{H}'),$$

where \underline{A} and \underline{R} are as defined in (21) and (9) respectively and $n > (1/2)(p-1)$.

As it stands, theorem 1 is not directly applicable to this function.

So we write, following Khatri, [7],

$$|\underline{I} + \underline{A} \underline{H} \underline{R} \underline{H}'| = |\underline{I} + \underline{R}| \left| \underline{I} - (\underline{I} - \underline{A}) \underline{H} \underline{R} (\underline{I} + \underline{R})^{-1} \underline{H}' \right| .$$

We now assume $\text{Ch}_i(\underline{A}) < 1$, $i = 1, \dots, p$. This is no loss of generality, since for $k > 0$, $|\underline{I} + k \underline{A} \underline{H} \underline{R} \underline{H}'| = \prod_{i=1}^p (1 + k\alpha_i)$ where

$\alpha_i = \text{Ch}_i(\underline{A} \underline{H} \underline{R} \underline{H}') > 0$, $i = 1, \dots, p$. Thus the problem of finding the maximum or minimum of $|\underline{I} + \underline{A} \underline{H} \underline{R} \underline{H}'|$ with respect to $\underline{H} \in O(p)$ is the same as that of

$|\underline{I} + \underline{k} \underline{A} \underline{H} \underline{R} \underline{H}'|$. Thus

$$\begin{aligned} |\underline{I} + \underline{A} \underline{H} \underline{R} \underline{H}'|^{-n} &= |\underline{I} + \underline{R}|^{-n} |\underline{I} - (\underline{I} - \underline{A}) \underline{H} \underline{R} (\underline{I} + \underline{R})^{-1} \underline{H}'|^{-n} \\ &= |\underline{I} + \underline{R}|^{-n} {}_1F_0(n, \underline{B} \underline{H} \underline{C} \underline{H}') \quad , \end{aligned}$$

where

$$\underline{B} = (\underline{I} - \underline{A}) = \text{diag}(b_1, \dots, b_p) \quad ,$$

and

$$\underline{C} = \underline{R}(\underline{I} + \underline{R})^{-1} = \text{diag}(c_1, \dots, c_p) \quad .$$

Hence from (21) and (9) we get

$$1 > b_1 > \dots > b_p > 0 \quad \text{and} \quad \infty > c_1 > c_2 > \dots > c_p \geq 0 \quad .$$

Thus $g(\underline{H}) = |\underline{I} + \underline{R}|^{-n} {}_1F_0(n, \underline{B} \underline{H} \underline{C} \underline{H}')$ and now we can apply the theorem 1 and get the following corollary.

Corollary 1.2. Under the conditions stated immediately above

$$\begin{aligned} \max_{\underline{H} \in \mathcal{O}(p)} g(\underline{H}) &= |\underline{I} + \underline{R}|^{-n} \max_{\underline{H} \in \mathcal{O}(p)} {}_1F_0(n, \underline{B} \underline{H} \underline{C} \underline{H}') \\ &= |\underline{I} + \underline{R}|^{-n} {}_1F_0(n, \underline{B} \underline{C}) \\ &= |\underline{I} + \underline{A} \underline{R}|^{-n} \quad . \end{aligned}$$

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

Corollary 1.3. Let (3) hold. Then $\prod_{i=1}^p l_i r_{i_j}$ and $\prod_{i=1}^p (1+l_i r_{i_j})$ are

both minimized when $r_{i_j} = r_i, i = 1, \dots, p$. They are both maximized when $r_{i_j} = r_{p-i+1}, i = 1, \dots, p$.

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 non-negative, non-decreasing function defined on $[0, \infty]$. Then

$$f\left(\sum_{i=1}^p l_i r_{i_j}\right) \leq f\left(\sum_{i=1}^p l_i r_i\right) \quad \text{and} \quad f\left(\prod_{i=1}^p (1+l_i r_{i_j})\right) \leq f\left(\prod_{i=1}^p (1+l_i r_i)\right)$$

under (3). 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.

4. Maximization of I when the l_i 's are equal in set. To this end let us consider the following form for A .

$$(27) \quad \tilde{A} = \text{diag}(l_1, \dots, l_1, l_2, \dots, l_2, \dots, l_i, \dots, l_i, l_{k_1+1}, \dots, l_{k_1+1}, \dots, l_p)$$

and $\infty > l_1 > l_2 > \dots > l_i > l_{k_1+\dots+k_i+1} > l_p \geq 0$ and \tilde{R} is as given in (9).

For the sake of simplicity in presentation we consider the case when $i = 1$, and $k_1 = 2$, i.e. let

$$(28) \quad \tilde{A} = \text{diag}(l_1, l_1, l_3, \dots, l_p), \quad \text{and} \quad \infty > l_1 > l_3 > \dots > l_p \geq 0$$

This is no loss of generality, as will be seen from our discussion — the more general case corresponding to (27) is a straight forward generalization of the same technique.

Proceeding exactly as in the earlier case of all unequal roots we get,

$$(29) \quad \tilde{e}_1' \tilde{C} \tilde{e}_1 = r_1 \sum_{k=1}^p h_{k1}^2 \ell_k < r_1 \ell_1$$

iff

$$(h_{31}, \dots, h_{p1}) \neq 0 \quad ,$$

where

$$\tilde{C} = \tilde{S} \tilde{H}' \tilde{A} \tilde{H} \tilde{S} \quad \text{and} \quad \tilde{S}, \tilde{H} \quad \text{as}$$

defined earlier and \tilde{A} as in (28)

$$\text{i.e.} \quad \tilde{S}^2 = \tilde{R} = \text{diag}(r_1, \dots, r_p) \quad .$$

Hence proceeding exactly as in the earlier case we get for all $\tilde{R} > 0$

$$\text{Ch}_1(\tilde{S} \tilde{H}' \tilde{A} \tilde{H} \tilde{S}) < r_1 \ell_1 \quad .$$

Now equality holds in (29)

$$\text{i.e.} \quad \text{Ch}_1(\tilde{S} \tilde{H}' \tilde{A} \tilde{H} \tilde{S}) = r_1 \ell_1$$

$$(30) \quad \text{iff} \quad (h_{31}, \dots, h_{p1}) = 0 \quad .$$

Now when (30) is satisfied we get by actual matrix multiplication

$$(31) \quad \begin{matrix} \underline{S} & \underline{H}' & \underline{A} & \underline{H} & \underline{S} \\ \sim & \sim & \sim & \sim & \sim \end{matrix} = \begin{pmatrix} r_1 \lambda_1 & 0 \\ 0 & G_2 \end{pmatrix},$$

where $G_2 = \underline{S}_2 \underline{H}'_2 \underline{A}_2 \underline{H}_2 \underline{S}_2$, $\underline{S}_2, \underline{H}_2, \underline{A}_2$ are defined earlier and $\underline{H}_2 \in 0(p-1)$.

As is clear from (31), the characteristic vector corresponding to the root $r_1 \lambda_1$ is proportioned to \underline{e}_1 and hence any vector belonging to E_p — the Euclidean p -space and orthogonal to \underline{e}_1 is generated by $\underline{e}_2, \dots, \underline{e}_p$.

Now the problem of finding the $\text{Ch}_2(\underline{S} \underline{H}' \underline{A} \underline{H} \underline{S})$ when (30) holds is reduced to finding $\text{Ch}_1(G_2)$.

Again

$$(32) \quad \begin{aligned} \text{Ch}_1(G_2) &= \text{Ch}_1(\underline{H}'_2 \underline{A}_2 \underline{H}_2 \underline{R}_2) \\ &= \text{Ch}_1(\underline{H}_2 \underline{R}_2 \underline{H}'_2 \underline{A}_2) \\ &\leq \text{Ch}_1(\underline{A}_2) \text{Ch}_1(\underline{H}_2 \underline{R}_2 \underline{H}'_2) \\ &= \text{Ch}_1(\underline{A}_2) \text{Ch}_1(\underline{R}_2) \\ &= \lambda_1 r_2 \end{aligned}$$

Proceeding exactly as earlier we get that equality in (32) is achieved for all variations of $\underline{R}_2 > 0$

$$(33) \quad \text{iff} \quad (h_{23}, \dots, h_{2p}) = 0$$

Again when (30) and (33) are satisfied we get by actual matrix multiplication

$$(34) \quad \underset{\sim}{S} \underset{\sim}{H}' \underset{\sim}{A} \underset{\sim}{H} \underset{\sim}{S} = \begin{pmatrix} r_1 \ell_1 & 0 & 0 \\ 0 & r_2 \ell_1 & 0 \\ 0 & 0 & G_3 \end{pmatrix}$$

where

$$\underset{\sim}{G}_3 = \underset{\sim}{S}_3 \underset{\sim}{H}'_3 \underset{\sim}{A}_3 \underset{\sim}{H}_3 \underset{\sim}{S}_3, \quad \underset{\sim}{H}_3 \in O(p-2)$$

$$\text{and} \quad \underset{\sim}{A}_3 = \text{diag}(\ell_3, \dots, \ell_p), \quad \underset{\sim}{S}_3^2 = \underset{\sim}{R}_3 = \text{diag}(r_3, \dots, r_p)$$

Now in order to find the form of $\underset{\sim}{H}_3$ so that $\text{Ch}_3(\underset{\sim}{S} \underset{\sim}{H} \underset{\sim}{A} \underset{\sim}{H}' \underset{\sim}{S})$ is maximized subject to (30) and (33), we find from (34) that we are back to the problem of all distinct roots in $\underset{\sim}{A}_3$ with the dimension of $\underset{\sim}{H}$ reduced by 2. Hence following our earlier technique step by step we get

Lemma 3.1. When (28) holds and $\underset{\sim}{R}$ is as in (9), then for all variations of $\underset{\sim}{R} > 0$, we get

$$\max_{\underset{\sim}{H} \in O(p)} C_{\underset{\sim}{K}}(\underset{\sim}{H}' \underset{\sim}{A} \underset{\sim}{H} \underset{\sim}{R}) = \max_{\underset{\sim}{H} \in O(p)} C_{\underset{\sim}{K}}(\underset{\sim}{H} \underset{\sim}{R} \underset{\sim}{H}' \underset{\sim}{A}) = C_{\underset{\sim}{K}}(\underset{\sim}{A} \underset{\sim}{R})$$

If $\underset{\sim}{R}$ is as in (3), then

$$(35) \quad \min_{\underset{\sim}{H} \in O(p)} C_{\underset{\sim}{K}}(\underset{\sim}{H}' \underset{\sim}{A} \underset{\sim}{H} \underset{\sim}{R}) = \min_{\underset{\sim}{H} \in O(p)} C_{\underset{\sim}{K}}(\underset{\sim}{H} \underset{\sim}{R} \underset{\sim}{H}' \underset{\sim}{A}) = C_{\underset{\sim}{K}}(\underset{\sim}{A} \underset{\sim}{R})$$

and this maximum or minimum value of the function in respective cases is achieved when \underline{H} has the form

$$(36) \quad \underline{H} = \begin{pmatrix} \underline{H}_1 \\ \underline{H}_2 \end{pmatrix},$$

where \underline{H}_1 ($2 \times p$) is arbitrary but otherwise satisfying the orthogonality relation of \underline{H} and

$$\underline{H}_2 = (0, \underline{H}_{22}),$$

where $0 = 0((p-2) \times 2)$ and $\underline{H}_{22}((p-2) \times (p-2))$ satisfies the condition (ii) of lemma 3.

In practice it is more frequently useful that \underline{A} instead of taking form (28) often satisfies the following

$$(37) \quad \underline{A} = \text{diag}(\lambda_1, \lambda_1, \lambda_3, \dots, \lambda_p) \quad \text{and} \quad \infty > \lambda_p > \dots > \lambda_3 > \lambda_1 \geq 0$$

and

$$\underline{R} = \text{diag}(r_1, \dots, r_p), \quad \infty > r_1 > r_2 > \dots > 0$$

The problem in this case more or less remains the same with the following changes:

Now instead of considering $C_K(\underline{H} \underline{R} \underline{H}' \underline{A})$ we consider $C_K((\underline{H} \underline{D}) \underline{R} (\underline{H} \underline{D})' \underline{A})$ as $H \in O(p)$, where $\underline{D} = (e_p, \dots, e_1)$ i.e. we consider

$$C_K(\underline{H} \underline{E} \underline{H}' \underline{A}) \quad \text{as} \quad H \in O(p),$$

where $\underline{E} = \text{diag}(r_p, \dots, r_1)$. Also as $\underline{H} \in \underline{O}(p)$, $\underline{H} \underline{D} \in \underline{O}(p)$, by earlier argument we get our results. Thus considering $C_{\kappa}(\underline{H} \underline{R} \underline{H}' \underline{A})$ as $\underline{H} \in \underline{O}(p)$ we note that the form of \underline{H} in this case is $\underline{H} \underline{D}$ where \underline{H} in this case satisfies the form (36), or more explicitly

$$\begin{aligned}
 (38) \quad \max_{\underline{H} \in \underline{O}(p)} C_{\kappa}(\underline{H} \underline{R} \underline{H}' \underline{A}) &= \max_{\underline{H} \in \underline{O}(p)} C_{\kappa}((\underline{H} \underline{D}) \underline{R} (\underline{H} \underline{D})' \underline{A}) \\
 &= \max_{\underline{H} \in \underline{O}(p)} C_{\kappa}(\underline{H} \underline{E} \underline{H}' \underline{A}) \\
 &= C_{\kappa}(\underline{E} \underline{A})
 \end{aligned}$$

and it is attained when \underline{H} has the form

$$(39) \quad \underline{H} = \begin{pmatrix} \underline{H}_1 \\ \underline{H}_2 \end{pmatrix} \underline{D} ,$$

where the left hand matrix is defined in (36). Here of course we note that zonal polynomials are symmetric functions of the characteristic roots of the defining matrices and so long as the characteristic roots of a matrix unchanged, zonal polynomials defined on them are also unchanged.

Now as a further remark we note that the above proof though stated for only one set of two equal roots, the proof is quite general, for at each step we just consider one root at a time and as can be noted that had there been three equal roots in a set, then after (34), we should have gotten a corresponding reduction and that it will work generally. Thus our earlier

$$(43) \quad \underline{H} = \begin{pmatrix} \underline{H}_1 \\ \underline{H}_2 \end{pmatrix} = \begin{pmatrix} \underline{P}_1 & \underline{P}_2 \\ \underline{P}_3 & \underline{P}_4 \end{pmatrix}, \quad \underline{R} = \begin{pmatrix} \underline{R}_1 & \underline{0} \\ \underline{0} & \underline{R}_4 \end{pmatrix},$$

where the partitions are appropriately done so that the following matrix products are defined.

$$\begin{aligned} \underline{H}' \underline{A} \underline{H} \underline{R} &= \underline{\lambda} \underline{R} + \begin{pmatrix} \underline{P}'_1 & \underline{P}'_3 \\ \underline{P}'_2 & \underline{P}'_4 \end{pmatrix} \begin{pmatrix} \underline{0} & \underline{0} \\ \underline{0} & \underline{B}_4 \end{pmatrix} \begin{pmatrix} \underline{P}_1 & \underline{P}_2 \\ \underline{P}_3 & \underline{P}_4 \end{pmatrix} \begin{pmatrix} \underline{R}_1 & \underline{0} \\ \underline{0} & \underline{R}_4 \end{pmatrix} \\ &= \underline{\lambda} \underline{R} + \begin{pmatrix} \underline{P}'_3 \underline{B}_4 \underline{P}_3 \underline{R}_1 & \underline{P}'_3 \underline{B}_4 \underline{P}_4 \underline{R}_4 \\ \underline{P}'_4 \underline{B}_4 \underline{P}_3 \underline{R}_1 & \underline{P}'_4 \underline{B}_4 \underline{P}_4 \underline{R}_4 \end{pmatrix} \\ &= \underline{\lambda} \underline{R} + \begin{pmatrix} \underline{X} \\ \underline{H}_2 \end{pmatrix}' \underline{B} \begin{pmatrix} \underline{X} \\ \underline{H}_2 \end{pmatrix} \underline{R} \end{aligned}$$

where $\underline{X}(k \times p)$ is arbitrary but otherwise is a completion of \underline{H}_2 .

Thus under (41), $f(\underline{H})$ is invariant of the choice of \underline{H}_1 in \underline{H} . This result with suitable modifications gives the results of James [5], Li and Pillai [8], [9].

5. Complex analogue of previous results. The complex analogue of the previous problems arises from the following consideration. Here instead of the problem of evaluating I in (1) in an asymptotic sense we have the parallel problem of finding an asymptotic expansion of

$$(44) \quad I_1 = \int_{U(p)} s_t^{\underline{F}}(a_1, \dots, a_s, b_1, \dots, b_t, \underline{U}' \underline{A} \underline{U} \underline{R}) d(\underline{U}),$$

where $U(p)$ is the group of $p \times p$ unitary matrices and $d(U)$ is the invariant measure on $U(p)$, normalized to make the total measure of the whole group unity, A, R as defined earlier, a_i 's and b_j 's are still functions of d.f. and hence are positive real numbers. But here considering the definition of hypergeometric functions as given by James [4] we will put the following restrictions on a_i 's and b_j 's.

$$(45) \quad a_i \geq (p-1), \quad b_j \geq (p-1) \quad i = 1, \dots, s, \quad j = 1, \dots, t,$$

$U \in U(p)$, i.e. an element of the group of $p \times p$ unitary matrix such that $U^* U = U U^* = I(p)$.

In this context we have the following lemma.

Lemma 3.3. Let A and R be as defined in (9). Then $\tilde{C}_K(U^* A U R)$ is real and for all variations of $R > 0$, we have

$$\max_{U \in U(p)} \tilde{C}_K(U^* A U R) = \max_{U \in U(p)} \tilde{C}_K(U R U^* A) = \tilde{C}_K(A R).$$

If A is defined by (21) instead of (9),

$$\min_{U \in U(p)} \tilde{C}_K(U^* A U R) = \min_{U \in U(p)} \tilde{C}_K(U^* A U R) = \tilde{C}_K(A R).$$

Proof. To prove that $\tilde{C}_K(U^* A U R)$ is real, we note

$$\text{Ch}_p(U^* A U) \text{Ch}_p(R) \leq \text{Ch}_i(U^* A U R) \leq \text{Ch}_1(U^* A U) \text{Ch}_1(R)$$

$$\text{i.e. } \text{Ch}_p(A) \text{Ch}_p(R) \leq \text{Ch}_i(U^* A U R) \leq \text{Ch}_1(A) \text{Ch}_1(R), \quad i = 1, \dots, p$$

For the rest we put as earlier $\tilde{S}^2 = \tilde{R}$ where

$$\tilde{S}^2 = \text{diag}(s_1, \dots, s_p), \quad \tilde{R} = \text{diag}(r_1, \dots, r_p) \text{ i.e. } s_i^2 = r_i, \quad i = 1, \dots, p.$$

Then let $\tilde{C} = \tilde{S} \tilde{U} \tilde{A} \tilde{U} \tilde{S} = (c_{ij})$ and $\tilde{U} = (u_{ij})$. Hence

$$c_{ij} = s_i s_j \sum_{k=1}^p \bar{u}_{ki} u_{kj} r_k, \quad (i, j = 1, \dots, p).$$

Now we proceed exactly as in the real case replacing H by \tilde{U} and get

Theorem 1.3. Let \tilde{R} be as given in (9) and \tilde{A} as in (27), then the class of unitary matrices for which $f(\tilde{U}) = \tilde{F}_t(a_1, \dots, a_s; b_1, \dots, b_t, \tilde{U}^* \tilde{A} \tilde{U} \tilde{R})$ subject to (45) and for all $\tilde{R} > 0$ is a maximum, is given by

$$(46) \quad \tilde{U} = \begin{pmatrix} U_1 \\ \vdots \\ U_i \\ U_{i+1} \end{pmatrix},$$

where

$$U_1 (k_1 \times p), \quad U_j = (0_j, U_{j1}) ,$$

$$0_j = 0(k_j \times (k_1 + \dots + k_{j-1})), \quad U_{j1} = U_{j1}(k_j \times (p - k_1 - \dots - k_{j-1})), \quad j = 2, \dots, i,$$

and

$$U_{i+1} = (0_{i+1,1}, U_{i+1,1}) ,$$

where

$$0_{i+1,1} = 0_{((p-(k_1+\dots+k_i)) \times (k_1+\dots+k_i))},$$

$$U_{i+1,1} = U_{i+1,1}^{((p-(k_1+\dots+k_i)) \times (p-(k_1+\dots+k_i)))},$$

0 stands for null matrix and $U_{i+1,1}$ is a diagonal matrix of the order indicated in the parentheses and with diagonal elements $e^{\sqrt{-1}\theta_j}$ $0 \leq \theta_j < 2\pi$, U_{k1} , $k = 2, \dots, i$, and U_1 are arbitrary matrices subject to the condition that U in (46) is unitary.

Also the maximum in this case is given by

$$(47) \quad \max_{U \in U(p)} f(U) = \max_{U \in U(p)} {}_s\tilde{F}_t(a_1, \dots, a_s; b_1, \dots, b_t, U^* A U R) \\ = {}_s\tilde{F}_t(a_1, \dots, a_s; b_1, \dots, b_t, A R),$$

where ${}_s\tilde{F}_t(a_1, \dots, a_s; b_1, \dots, b_t, Z)$ is defined in James [4],

$${}_s\tilde{F}_t(a_1, \dots, a_s, b_1, \dots, b_t, Z) = \sum_{k=0}^{\infty} \sum_{\kappa} \frac{[a_1]_{\kappa} \dots [a_s]_{\kappa}}{[b_1]_{\kappa} \dots [b_t]_{\kappa}} \frac{\tilde{C}_{\kappa}(Z)}{k!}$$

and the complex multivariate hypergeometric coefficients are as follows

$$[a]_{\kappa} = \prod_{i=1}^p (a - i + 1)_{k_i},$$

where $\kappa = (k_1, \dots, k_p)$ is a partition of the integer k , when the ordering of the elements of A is reversed a result parallel to that of (26) holds.

$U_2 = (O, D)$, $O((p-k) \times k)$ and $D = \text{diag}(e^{\sqrt{-1}\theta_1}, \dots, e^{\sqrt{-1}\theta_q})$ where $q = p-k$ and $0 \leq \theta_j < 2\pi$, $j = 1, \dots, q$, a_i 's and b_j 's are assumed to satisfy (45). This form of U in (48) immediately asserts invariance results of Li and Pillai [8], [9]. In fact obtaining the form of U as in (46) is in short the most general result obtained in this paper.