THE DISTRIBUTION OF CHARACTERISTIC VECTORS CORRESPONDING TO THE TWO LARGEST ROOTS OF A MATRIX*

by

Hung C. Li

Department of Statistics

Division of Mathematical Sciences

Mimeograph Series No. 183

February, 1969

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

THE DISTRIBUTION OF CHARACTERISTIC VECTORS CORRESPONDING TO THE TWO LARGEST ROOTS OF A MATRIX*

by

Hung C. Li

Purdue University

1. Summary. The distribution of the characteristic (ch.) vectors of a sample covariance matrix was found by T. W. Anderson [1], when the population covariance matrix is a scalar matrix $\Sigma = \sigma^2 I$. The asymptotic distribution for arbitrary Σ also was obtained by T. W. Anderson [2]. For unknown Σ, the distribution of the ch. vector corresponding to the largest root of a covariance matrix was found by T. Sugiyama [10] and C. G. Khatri and K. C. S. Pillai [8]. In this paper, for arbitrary Σ, we obtain the joint distribution of the vectors corresponding to the two largest roots for the non-central linear case, i.e. when the rank of the mean matrix is one.

2. Notations and some useful results. Matrices will be denoted by bold face capital letters, and their dimensions will be indicated parenthetically. The mxm identity matrix will be denoted by I_m, and in particular, I denotes I_{p-2} throughout this paper. $|\alpha|$ denotes the absolute value of α, and $|X|$ denotes the determinant of X. $O(n)$ denotes the group of all orthogonal nxn matrices.

Let $S(m \times m)$ be any symmetric positive definite matrix. The zonal polynomials $Z_\kappa(S)$ are defined for each partition $\kappa = (k_1, k_2, \ldots, k_m)$, $k_1 \geq k_2 \geq \cdots \geq k_m \geq 0$ of k into not more than m parts, as certain symmetric polynomials in the ch. roots of S, (see A. T. James [5], [6], [7] and A. G. Constantine [3]). Further (see A. G. Constantine [3])

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

$$(2.1) \qquad \int_{O(m)} C_\kappa(H'SHT)\,dH = C_\kappa(S)C_\kappa(T)/C_\kappa(I_m)$$

where dH is the invariant Haar measure on the orthogonal group $O(m)$, normalized to make the volume of the group manifold unity. Also note that (see [3])

$$(2.2) \qquad \int_{0}^{I_m} |S|^{t-\frac{1}{2}(m+1)} |I_m - S|^{u-\frac{1}{2}(m+1)} C_\kappa(TS)\,dS = \frac{\Gamma_m(t,\kappa)\Gamma_m(u)}{\Gamma_m(t+u,\kappa)} C_\kappa(T)$$

where T is a positive definite matrix,

$$\Gamma_m(u) = \pi^{\frac{1}{4}m(m-1)} \prod_{i=0}^{m-1} \Gamma(u-\tfrac{1}{2}i)$$

and

$$\Gamma_m(t,\kappa) = \pi^{\frac{1}{4}m(m-1)} \prod_{i=0}^{m-1} \Gamma(t+k_i-\tfrac{1}{2}i) .$$

Let R (nxn) be an orthogonal matrix such that the first $r(\leq n)$ columns have random elements and the remaining $(n-r)$ columns depend on these random elements. We will denote $dR^{(n,r)}$ a normalized measure over this space, i.e.

$$\int_{O(n)} dR^{(n,r)} = 1$$

In terms of Roy's notation [9], let

$$J(R) = 2^n / \left| \frac{\partial(RR')}{\partial(R_D)} \right|_{R_I}$$

Thus $J(\underset{\sim}{R})$ is a function of random elements of $\underset{\sim}{R}$. We will write

$$(2.3) \qquad d\underset{\sim}{R}^{(n,r)} = \pi^{-\frac{1}{2}rn}\Gamma_r(\tfrac{n}{2})J(\underset{\sim}{R}) .$$

<u>Lemma 1.</u> Let $\underset{\sim}{U}(p\text{x}n)$ and $\underset{\sim}{V}((p-r)\text{ x }(n-r))$ be random matrices, $(p \leq n)$ and let

$$(2.4) \qquad \underset{\sim}{U} = \underset{\sim}{H}\begin{pmatrix} \alpha_1 & & & 0 \\ & \ddots & & \\ & & \alpha_r & \\ 0 & & & \underset{\sim}{V} \end{pmatrix}\underset{\sim}{G}'$$

be a transformation such that the first $r(\leq p)$ column vectors of orthogonal matrices $\underset{\sim}{H}$ (pxp) and $\underset{\sim}{G}$ (nxn) contain random elements, and $\alpha_i \neq 0$ $(i=1,\ldots,r)$ and $\alpha_1^2 > \alpha_2^2 > \ldots > \alpha_r^2 > 0$ be the first r non-zero largest ordered ch. roots of $\underset{\sim}{U}\underset{\sim}{U}'$. Then the jacobian of the transformation is given by

$$(2.5) \qquad J(\underset{\sim}{U} : \underset{\sim}{H}\ \alpha_1, \alpha_2, \ldots, \alpha_r, \underset{\sim}{V}, \underset{\sim}{G}) =$$

$$C \prod_{i=1}^{r} |\alpha_i|^{n-p}|\alpha_i^2\underset{\sim}{I} - \underset{\sim}{V}\underset{\sim}{V}'| \prod_{i<j} (\alpha_i^2 - \alpha_j^2)d\underset{\sim}{H}^{(p,r)}d\underset{\sim}{G}^{(n,r)}$$

where

$$C = \pi^{\frac{1}{2}r(p+n)} \left\{\Gamma_r(\tfrac{p}{2})\Gamma_r(\tfrac{n}{2})\right\}^{-1} .$$

<u>Proof</u>: Taking differentials of (2.4)

$$d\underset{\sim}{U} = d\underset{\sim}{H}\begin{pmatrix} \alpha_1 & & & 0 \\ & \ddots & & \\ & & \alpha_r & \\ 0 & & & \underset{\sim}{V} \end{pmatrix}\underset{\sim}{G}' + \underset{\sim}{H}\begin{pmatrix} d\alpha_1 & & & 0 \\ & \ddots & & \\ & & d\alpha_r & \\ 0 & & & d\underset{\sim}{V} \end{pmatrix}\underset{\sim}{G}' + \underset{\sim}{H}\begin{pmatrix} \alpha_1 & & & 0 \\ & \ddots & & \\ & & \alpha_r & \\ 0 & & & \underset{\sim}{V} \end{pmatrix}d\underset{\sim}{G}'$$

Both sides, pre- and post- multiply $\underset{\sim}{H}'$ and $\underset{\sim}{G}$, we obtain

$$\underset{\sim}{H}'\,(d\underset{\sim}{U})\,\underset{\sim}{G} =$$

$$\underset{\sim}{H}'d\underset{\sim}{H}\begin{pmatrix}\alpha_1 & & & 0\\ & \ddots & & \\ & & \alpha_r & \\ 0 & & & \underset{\sim}{V}\end{pmatrix} + \begin{pmatrix}d\alpha_1 & & & 0\\ & \ddots & & \\ & & d\alpha_r & \\ 0 & & & d\underset{\sim}{V}\end{pmatrix} + \begin{pmatrix}\alpha_1 & & & 0\\ & \ddots & & \\ & & \alpha_r & \\ 0 & & & \underset{\sim}{V}\end{pmatrix}(d\underset{\sim}{G}')\underset{\sim}{G}$$

Let

$$\underset{\sim}{H}'(d\underset{\sim}{U})\underset{\sim}{G} = \underset{\sim}{W} = \begin{pmatrix}w_{11} & \cdots & w_{1n}\\ \cdots & \cdots & \cdots\\ w_{p1} & \cdots & w_{pn}\end{pmatrix}$$

Since $\underset{\sim}{H}'d\underset{\sim}{H}$ and $(d\underset{\sim}{G}')\underset{\sim}{G}$ are pxp and nxn skew symmetric matrices, hence we can put

$$\underset{\sim}{H}'d\underset{\sim}{H} = \underset{\sim}{A} = \begin{pmatrix}0 & a_{12} & \cdots & a_{1p}\\ -a_{12} & 0 & \cdots & a_{2p}\\ \cdots & \cdots & \cdots & \cdots\\ -a_{1p} & -a_{2p} & \cdots & 0\end{pmatrix}$$

and

$$(d\underset{\sim}{G}')\underset{\sim}{G} = \underset{\sim}{B} = \begin{pmatrix}0 & b_{12} & \cdots & b_{1n}\\ -b_{12} & 0 & \cdots & b_{2n}\\ \cdots & \cdots & \cdots & \cdots\\ -b_{1n} & -b_{2n} & \cdots & 0\end{pmatrix}$$

Denote $\underset{\sim}{W}_{(r,r)}$ to be the matrix from $\underset{\sim}{W}$ by deleting the first r rows and the first r columns of $\underset{\sim}{W}$. Similarly for $\underset{\sim}{A}_{(r,r)}$ and $\underset{\sim}{B}_{(r,r)}$.

Then

$$(w_{ij}, w_{ji}) = (a_{ij}, b_{ij}) \begin{pmatrix} \alpha_j - \alpha_i \\ \alpha_i - \alpha_j \end{pmatrix},$$

and $$(w_{i,r+1}, \ldots, w_{in}, w_{r+1,i}, \ldots, w_{pi}) =$$

$$(a_{i,r+1}, \ldots, a_{ip}, b_{i,r+1}, \ldots, b_{in}) \begin{pmatrix} \underset{\sim}{V} & -\alpha_i \underset{\sim}{I} \\ \alpha_i \underset{\sim}{I}_{n-r} & -\underset{\sim}{V}' \end{pmatrix}$$

imply

$$J_{ij} = J(w_{ij}, w_{ji} : a_{ij}, b_{ij}) = \alpha_i^2 - \alpha_j^2 \qquad i, j = 1, \ldots, r \text{ and } i < j,$$

and

$$J_i = J(w_{ik}, W_{\ell i} : A_{i\ell}, b_{ik}; \qquad k = r+1, \ldots, n; \quad \ell = r+1, \ldots, p)$$

$$= |\alpha_i|^{n-p} |\alpha_i^2 \underset{\sim}{I} - \underset{\sim}{V}\underset{\sim}{V}'| \qquad i = 1, \ldots, r.$$

Moreover $J(d\underset{\sim}{U} : \underset{\sim}{W}) = 1$

$$J_i^* = J(w_{ii} : d\alpha_i) = 1, \qquad i = 1, \ldots, r$$

and $$\underset{\sim}{W}_{(r,r)} = \underset{\sim}{A}_{(r,r)}\underset{\sim}{V} + d\underset{\sim}{V} + \underset{\sim}{V}\underset{\sim}{B}_{(r,r)}$$

implies $$J(\underset{\sim}{W}_{(r,r)} : d\underset{\sim}{V}) = 1 .$$

Finally,

$$J_i^{**} = J(a_{i\ell}, b_{ik} : dh_{i\ell}, dg_{ik}; \quad k = r+1, \ldots, n, \ \ell = r+1, \ldots, p)$$
$$= 1 \qquad\qquad i = 1, \ldots, r.$$

where h_{ij} and g_{ij} are the ith row and jth column elements of $\underset{\sim}{H}$ and $\underset{\sim}{G}$ respectively. Therefore

$$J(\underset{\sim}{U} : \underset{\sim}{H}, \alpha_1, \ldots, \alpha_r, \underset{\sim}{V}, \underset{\sim}{G})$$

$$= J(d\underset{\sim}{U} : d\underset{\sim}{H}, d\alpha_1, \ldots, d\alpha_r, d\underset{\sim}{V}, d\underset{\sim}{G})$$

$$= J(d\underset{\sim}{U} : \underset{\sim}{W})J(\underset{\sim}{W} : dh_{1\ell}, \ldots, dh_{r\ell}, d\alpha_1, \ldots, d\alpha_r, d\underset{\sim}{V}, dg_{1k},$$
$$\ldots, dg_{rk}; \quad k = r+1, \ldots, n; \quad \ell = r+1, \ldots, p)J(\underset{\sim}{H})J(\underset{\sim}{G})$$

$$= \prod_{i=1}^{r} J_i \, J_i^{*} \, J_i^{**} \prod_{i<j} J_{ij} \; J(d\underset{\sim}{U} : \underset{\sim}{W})J(\underset{\sim}{W}_{(r,r)} : d\underset{\sim}{V})J(\underset{\sim}{H})J(\underset{\sim}{G}) .$$

Using (2.3), we obtain (2.5) .

In (2.5) if we put $\lambda_i = \alpha_i^2 \quad i = 1, \ldots, r$ and notice that each λ_i corresponding α_i and $-\alpha_i$, then (2.5) can be written

(2.6)
$$J(\underset{\sim}{U} : \underset{\sim}{H}, \lambda_1, \lambda_2, \ldots, \lambda_r \underset{\sim}{V}, \underset{\sim}{G})$$
$$= C \prod_{i=1}^{r} \lambda_i^{\frac{1}{2}(n-p-1)} |\lambda_i \underset{\sim}{I} - \underset{\sim}{V}\underset{\sim}{V}'| \prod_{i<j} (\lambda_i - \lambda_j) d\underset{\sim}{H}^{(p,r)} d\underset{\sim}{G}^{(n,r)} \quad .$$

Lemma 2. Let $\underset{\sim}{V}(m\times t)$ be a random matrix and $\underset{\sim}{A}(m\times m)$ be a symmetric matrix. For definiteness, assume $m \le t$. Then

$$\int_{\mathcal{D}} |a\underset{\sim}{I}_m - \underset{\sim}{V}\underset{\sim}{V}'|^{\alpha} |b\underset{\sim}{I}_m - \underset{\sim}{V}\underset{\sim}{V}'|^{\beta} C_{\kappa}(\underset{\sim}{A}\underset{\sim}{V}\underset{\sim}{V}')d\underset{\sim}{V} \tag{2.7}$$

$$= \pi^{\frac{1}{2}mt}\Gamma_m(\beta + \tfrac{m+1}{2})C_{\kappa}(\underset{\sim}{A}) \left\{\Gamma_m(\beta + \tfrac{t+m+1}{2})C_{\kappa}(\underset{\sim}{I}_m)\right\}^{-1}$$

$$\cdot \sum_{q=0}^{\infty} \sum_{\eta} \sum_{\delta} \frac{(-1)_{\eta}}{q!} g^{\delta}_{\kappa,\eta}\, C_{\delta}(\underset{\sim}{I}_m)(\tfrac{t}{2})_{\delta} \left\{(\beta + \tfrac{t+m+1}{2})_{\delta}\right\}^{-1} a^{m\alpha - q} b^{m(\beta+\frac{1}{2}t)+d}$$

where $\alpha > \frac{1}{2}(m-1)$, $\beta > \frac{1}{2}(m-1)$, $a > b > 0$,

$$\mathcal{D} = \mathcal{D}\,\{\underset{\sim}{V} \text{ such that } b\underset{\sim}{I}_m - \underset{\sim}{V}\underset{\sim}{V}' \text{ is positive definite}\}$$

$$\kappa = \{k_1, k_2, \ldots, k_m\}\ , \quad \sum_{i=1}^{m} k_i = k\ ,$$

$(x)_{\kappa} = \Gamma_m(x,\kappa)/\Gamma_m(x)$ if x is such that the gamma functions are defined, and

$$\left\{\begin{array}{l} \delta = (\delta_1,\delta_2,\ldots,\delta_m), \quad \delta_1 \ge \delta_2 \ge \cdots \ge \delta_m \ge 0, \quad \sum_{i=1}^{m} \delta_i = q+k = d \\ g^{\delta}_{\kappa,\eta} \text{ is the coefficient of } C_{\delta}(\underset{\sim}{B}) \text{ in the} \\ \text{product of } C_{\kappa}(\underset{\sim}{B})C_{\eta}(\underset{\sim}{B}) \quad . \end{array}\right. \tag{2.8}$$

Proof: Let us write

$$h = \int_{\mathcal{D}} |a\underset{\sim}{I}_m - \underset{\sim}{V}\underset{\sim}{V}'|^{\alpha} |b\underset{\sim}{I}_m - \underset{\sim}{V}\underset{\sim}{V}'|^{\beta} C_{\kappa}(\underset{\sim}{A}\underset{\sim}{V}\underset{\sim}{V}')\, d\underset{\sim}{V}$$

Then

$$h = \int_{\mathcal{D}} |a\underset{\sim}{I}_m - \underset{\sim}{V}\underset{\sim}{V}'|^{\alpha} |b\underset{\sim}{I}_m - \underset{\sim}{V}\underset{\sim}{V}'|^{\beta} C_{\kappa}(\underset{\sim}{A}\underset{\sim}{V}\underset{\sim}{V}')d\underset{\sim}{V} \int_{O(m)} d\underset{\sim}{H} \quad .$$

Making transformation $\underset{\sim}{V} \to \underset{\sim}{H}\underset{\sim}{V}$ and notice

$$C_{\kappa}(\underset{\sim}{H}'\underset{\sim}{A}\underset{\sim}{H}\underset{\sim}{V}\underset{\sim}{V}') = C_{\kappa}(\underset{\sim}{A}\underset{\sim}{H}\underset{\sim}{V}\underset{\sim}{V}'\underset{\sim}{H}')$$

then

$$h = \frac{C_{\kappa}(\underset{\sim}{A})}{C_{\kappa}(\underset{\sim}{I}_m)} \int_{\mathcal{D}} |a\underset{\sim}{I}_m - \underset{\sim}{V}\underset{\sim}{V}'|^{\alpha} |b\underset{\sim}{I}_m - \underset{\sim}{V}\underset{\sim}{V}'|^{\beta} C_{\kappa}(\underset{\sim}{V}\underset{\sim}{V}')d\underset{\sim}{V}$$

by (2.1). Let $\underset{\sim}{V}\underset{\sim}{V}' = \underset{\sim}{S}$, then

$$h = C^* \int_{\underset{\sim}{0}}^{b\underset{\sim}{I}_m} |\underset{\sim}{S}|^{\frac{1}{2}t-\frac{1}{2}(m+1)} |a\underset{\sim}{I}_m - \underset{\sim}{S}|^{\alpha} |b\underset{\sim}{I}_m - \underset{\sim}{S}|^{\beta} C_{\kappa}(\underset{\sim}{S})d\underset{\sim}{S}$$

where

$$C^* = \pi^{\frac{1}{2}mt} \, C_{\kappa}(\underset{\sim}{A}) \, \{\Gamma_m(\tfrac{t}{2}) C_{\kappa}(\underset{\sim}{I}_m)\}^{-1} \quad .$$

Next, put $\underset{\sim}{S} = a\underset{\sim}{S}^*$, we get

$$h = C^* \, a^{m(\alpha+\beta+\frac{1}{2}t)+k}$$

$$\cdot \int_0^{\frac{b}{a}\underset{\sim}{I}_m} |\underset{\sim}{S}^*|^{\frac{1}{2}t-\frac{1}{2}(m+1)} |\underset{\sim}{I}_m - \underset{\sim}{S}^*|^{\alpha} |\tfrac{b}{a}\underset{\sim}{I}_m - \underset{\sim}{S}^*|^{\beta} C_{\kappa}(\underset{\sim}{S}^*)d\underset{\sim}{S}^* \quad .$$

By A. G. Constantine [4], notice that

$$\sum_{q=0}^{\infty} \sum_{\eta} \frac{(z)_\eta}{q!} C_\eta(\underset{\sim}{S}) = |\underset{\sim}{I}_m - \underset{\sim}{S}|^{-z}$$

h can be written as

$$h = C^* A^{m(\alpha+\beta+\frac{1}{2}t)+k} \sum_{q=0}^{\infty} \sum_{\eta} \sum_{\delta} \frac{(-\alpha)_\eta}{q!} g_{\kappa,\eta}^{\delta}$$

$$\cdot \int_{\underset{\sim}{0}}^{\frac{b}{a}\underset{\sim}{I}_m} |\underset{\sim}{S}^*|^{\frac{1}{2}t-\frac{1}{2}(m+1)} |\tfrac{b}{a}\underset{\sim}{I}_m - \underset{\sim}{S}^*|^{\beta} C_\delta(\underset{\sim}{S}^*) d\underset{\sim}{S}^* ,$$

where δ, δ_i and $g_{\kappa,\eta}^{\delta}$ are defined by (2.8).

Finally, put $\underset{\sim}{S}^* = \frac{b}{a}\underset{\sim}{T}$ then by (2.2) and

$$\Gamma_m(\tfrac{t}{2}, \delta) = \Gamma_m(\tfrac{t}{2})(\tfrac{t}{2})_\delta$$

we have

$$h = C^* \sum_{q=0}^{\infty} \sum_{\eta} \sum_{\delta} a^{m\alpha-q} b^{m(\beta+\frac{1}{2}t)+q+k} \frac{(-\alpha)_\eta}{q!} g_{\kappa,\eta}^{\delta}$$

$$\cdot \int_{\underset{\sim}{0}}^{\underset{\sim}{I}_m} |\underset{\sim}{T}|^{\frac{1}{2}t-\frac{1}{2}(m+1)} |\underset{\sim}{I}_m - \underset{\sim}{T}|^{\beta} C_\delta(\underset{\sim}{T}) d\underset{\sim}{T}$$

$$= \pi^{\frac{1}{2}mt} C_\kappa(\underset{\sim}{A}) \left\{C_\kappa(\underset{\sim}{I}_m)\right\}^{-1} \sum_{q=0}^{\infty} \sum_{\eta} \sum_{\delta} \frac{(-\alpha)_\eta}{q!} g_{\kappa,\eta}^{\delta} C_\delta(\underset{\sim}{I}_m)(\tfrac{t}{2})_\delta$$

$$\cdot \Gamma_m(\beta + \tfrac{m+1}{2}) \left\{\Gamma_m(\beta + \tfrac{t+m+1}{2}, \delta)\right\}^{-1} a^{m\alpha-q} b^{m(\beta+\frac{1}{2}t)+q+k} .$$

After rearranging, we obtain (2.7).

3. <u>Distribution of the characteristic vectors corresponding to the two largest roots of a matrix in the non-central case</u>. Let the matrix $\underset{\sim}{X}(pxn)$ be distributed as

$$(2\pi)^{-\frac{1}{2}pn}|\underset{\sim}{\Sigma}|^{-\frac{1}{2}n} \quad \exp\left\{-\tfrac{1}{2}\ \mathrm{tr}\ \underset{\sim}{\Sigma}^{-1}(\underset{\sim}{X} - \underset{\sim}{M})(\underset{\sim}{X} - \underset{\sim}{M})'\right\} \quad \cdot \cdot \cdot$$

i.e.

$$(3.1) \qquad (2\pi)^{-\frac{1}{2}pn}|\underset{\sim}{\Sigma}|^{-\frac{1}{2}n} \quad \exp\left\{-\tfrac{1}{2}\ \mathrm{tr}\ \underset{\sim}{\Sigma}^{-1}\underset{\sim\sim}{MM'} + \mathrm{tr}\ \underset{\sim}{\Sigma}^{-1}\underset{\sim\sim}{MX'} - \tfrac{1}{2}\ \mathrm{tr}\ \underset{\sim}{\Sigma}^{-1}\underset{\sim\sim}{XX'}\right\}$$

where $E[\underset{\sim}{X}] = \underset{\sim}{M}$.

Making transformation

$$\underset{\sim}{X} = \underset{\sim}{L}\begin{pmatrix} \alpha_1 & & & 0 \\ & \ddots & & \\ & & \alpha_r & \\ 0 & & & \underset{\sim}{Y} \end{pmatrix}\underset{\sim}{Q}'$$

where $r \leq p$, $\underset{\sim}{L}(pxp)$ and $\underset{\sim}{Q}(nxn)$ are orthogonal matrices, $\underset{\sim}{Y}$ is an $(p-r)X(n-r)$ matrix and $\alpha_1^2 > \alpha_2^2 > \ldots > \alpha_r^2 > 0$ are the first r largest ordered ch. roots of $\underset{\sim\sim}{XX'}$. Using Lemma 1, the joint density function of $\underset{\sim}{L}$, α_1, α_2, ..., α_r, $\underset{\sim}{Y}$ and $\underset{\sim}{Q}$ is given by

$$(3.2) \qquad C_1|\Sigma|^{-\frac{1}{2}n}\prod_{i=1}^{r}|\alpha_i|^{n-p}|\alpha_i^2 \underset{\sim}{I} - \underset{\sim\sim}{YY'}|\prod_{i<j}(\alpha_i^2 - \alpha_j^2)$$

$$\cdot \exp\left\{-\tfrac{1}{2}\operatorname{tr}\underset{\sim}{\Sigma}^{-1}\underset{\sim\sim}{MM'} + \operatorname{tr}\underset{\sim}{\Sigma}^{-1}\underset{\sim\sim}{MQ}\begin{pmatrix}\alpha_1 & & 0\\ & \ddots\alpha_r & \\ 0 & & \underset{\sim}{Y'}\end{pmatrix}\underset{\sim}{L'}\right.$$

$$\left.-\tfrac{1}{2}\operatorname{tr}\underset{\sim}{\Sigma}^{-1}\underset{\sim}{L}\begin{pmatrix}\alpha_1^2 & & 0\\ & \ddots\alpha_r^2 & \\ 0 & & \underset{\sim\sim}{YY'}\end{pmatrix}\underset{\sim}{L'}\right\} d\underset{\sim}{L}^{(p,r)}d\underset{\sim}{Q}^{(n,r)}$$

where

$$(3.3) \qquad C_1 = C(2\pi)^{-\frac{1}{2}pn} = \pi^{\frac{1}{2}r(p+n)}\left\{(2\pi)^{\frac{1}{2}pn}\Gamma_r(\tfrac{p}{2})\Gamma_r(\tfrac{n}{2})\right\}^{-1} .$$

In integrating α_i $(i = 1,\ldots, r)$, $\underset{\sim}{Y}$, $\underset{\sim}{Q}$, or $\underset{\sim}{L}$, we only consider the non-central linear case, i.e. when the rank of the mean matrix $\underset{\sim}{M}$ is one, because the general problem is extremely difficult.

For $r = 1$, we get the same result as given by C. G. Khatri and K. C. S. Pillai [8].

For $r = 2$, let

$$\underset{\sim}{L} = (\underset{\sim}{\ell_1}, \underset{\sim}{\ell_2}, \underset{\sim}{L_2}) ,$$

where $\underset{\sim}{\ell_1}$ and $\underset{\sim}{\ell_2}$ are the first two columns of $\underset{\sim}{L}$, corresponding to the two largest ordered ch. roots λ_1 and λ_2 of $\underset{\sim\sim}{XX'}$, having random elements and the others $\underset{\sim}{L_2}$ depend on these random elements. Then (3.2) becomes

$$(3.4)\qquad C_2|\alpha_1\alpha_2|^{n-p}(\alpha_1^2 - \alpha_2^2)|\alpha_1^2 \underset{\sim}{I} - \underset{\sim}{Y}\underset{\sim}{Y}'|\alpha_2^2\underset{\sim}{I} - \underset{\sim}{Y}\underset{\sim}{Y}'|$$

$$\cdot \exp\left\{\text{tr}\ \underset{\sim}{\Sigma}^{-1}\underset{\sim}{M}\underset{\sim}{Q}\begin{pmatrix}\alpha_1 & 0 & \underset{\sim}{0}\\ 0 & \alpha_2 & \\ \underset{\sim}{0} & & \underset{\sim}{Y}'\end{pmatrix}\underset{\sim}{L}' - \tfrac{1}{2}\ \text{tr}\ \underset{\sim}{\Sigma}^{-1}\underset{\sim}{L}\begin{pmatrix}\alpha_1^2 & 0 & \underset{\sim}{0}\\ 0 & \alpha_2^2 & \\ \underset{\sim}{0} & & \underset{\sim}{Y}\underset{\sim}{Y}'\end{pmatrix}\underset{\sim}{L}'\right\}$$

$$\cdot\ d\ \underset{\sim}{L}^{(p,2)} d\underset{\sim}{Q}^{(n,2)}\ ,$$

where

$$C_2 = C_1|\underset{\sim}{\Sigma}|^{-\frac{1}{2}n}\ e^{-\frac{1}{2}\text{tr}\underset{\sim}{\Sigma}^{-1}\underset{\sim}{M}\underset{\sim}{M}'}$$

Integrating (3.4) with respect to $\underset{\sim}{Q}$, we obtain

$$(3.5)\qquad C_2|\alpha_1\alpha_2|^{n-p}(\alpha_1^2 - \alpha_2^2)\sum_{k=0}^{\infty}\left\{k!\ (\tfrac{n}{2})_k\right\}^{-1} d\underset{\sim}{L}^{(p,2)} f_k(\alpha_1,\ \alpha_2,\ \underset{\sim}{L})$$

or

$$(3.5')\qquad C_2(\lambda_1\lambda_2)^{\frac{1}{2}(n-p-1)}(\lambda_1-\lambda_2)\sum_{k=0}^{\infty}\left\{k!\ (\tfrac{n}{2})_k\right\}^{-1} d\underset{\sim}{L}^{(p,2)} f_k(\lambda_1,\ \lambda_2,\ \underset{\sim}{L}),$$

where

$$(3.6)\qquad f_k(\lambda_1,\ \lambda_2,\ \underset{\sim}{L}) =$$

$$\int_{\mathcal{D}}|\lambda_1\underset{\sim}{I} - \underset{\sim}{Y}\underset{\sim}{Y}'|\cdot|\lambda_2\underset{\sim}{I} - \underset{\sim}{Y}\underset{\sim}{Y}'|\cdot\left\{\text{tr}\left[\tfrac{1}{4}\ \underset{\sim}{\Sigma}^{-1}\underset{\sim}{L}\begin{pmatrix}\lambda_1 & 0 & \underset{\sim}{0}\\ 0 & \lambda_2 & \\ \underset{\sim}{0} & & \underset{\sim}{Y}\underset{\sim}{Y}'\end{pmatrix}\underset{\sim}{L}'\ \underset{\sim}{\Sigma}^{-1}\underset{\sim}{M}\underset{\sim}{M}'\right]\right\}^k$$

$$\cdot \exp\left\{-\tfrac{1}{2}\ \text{tr}\ \underset{\sim}{\Sigma}^{-1}\underset{\sim}{L}\begin{pmatrix}\lambda_1 & 0 & \underset{\sim}{0}\\ 0 & \lambda_2 & \\ \underset{\sim}{0} & & \underset{\sim}{Y}\underset{\sim}{Y}'\end{pmatrix}\underset{\sim}{L}'\right\} d\underset{\sim}{Y}\ ,$$

where

$$\mathcal{D} = \mathcal{D}\,\{\underset{\sim}{Y} \text{ such that } \lambda_2 \underset{\sim}{I} - \underset{\sim\sim}{YY'} \text{ is a positive definite}\}$$

or

$$(3.7) \qquad f(\theta, \lambda_1, \lambda_2, \underset{\sim}{L}) = \sum_{k=0}^{\infty} \frac{\theta^k}{k!} f_k(\lambda_1, \lambda_2, \underset{\sim}{L})$$

$$= \int_{\mathcal{D}} |\lambda_1 \underset{\sim}{I} - \underset{\sim\sim}{YY'}| \cdot |\lambda_2 \underset{\sim}{I} - \underset{\sim\sim}{YY'}|$$

$$\cdot \exp\left\{-\tfrac{1}{2} \operatorname{tr} \underset{\sim}{\Sigma}^{-1} \underset{\sim}{L} \begin{pmatrix} \lambda_1 & 0 & \\ 0 & \lambda_2 & \underset{\sim}{0} \\ \underset{\sim}{0} & & \underset{\sim\sim}{YY'} \end{pmatrix} \underset{\sim}{L'}(\underset{\sim}{I} - \tfrac{1}{2}\theta \underset{\sim}{\Sigma}^{-1} \underset{\sim\sim}{MM'})\right\} d\underset{\sim}{Y}$$

$$= \exp\{-\tfrac{1}{2}\lambda_1 \underset{\sim}{\ell_1'} \underset{\sim}{\Delta} \underset{\sim}{\ell_1} - \tfrac{1}{2}\lambda_2 \underset{\sim}{\ell_2'} \underset{\sim}{\Delta} \underset{\sim}{\ell_2}\}$$

$$\cdot \int_{\mathcal{D}} \lambda_1 \underset{\sim}{I} - \underset{\sim\sim}{YY'}| \cdot |\lambda_2 \underset{\sim}{I} - \underset{\sim\sim}{YY'}| \exp\{-\tfrac{1}{2} \operatorname{tr} \underset{\sim}{L_2'} \underset{\sim}{\Delta} \underset{\sim}{L_2} \underset{\sim\sim}{YY'}\} d\underset{\sim}{Y}$$

where

$$(3.8) \qquad \underset{\sim}{\Delta} = \underset{\sim}{\Sigma}^{-1} - \tfrac{1}{2}\theta \underset{\sim}{\Sigma}^{-1} \underset{\sim\sim}{MM'} \underset{\sim}{\Sigma}^{-1} .$$

Let $\underset{\sim}{H} \in O(p-2)$ such that

$$\int_{O(p-2)} d\underset{\sim}{H} = 1$$

Making transformation $\underset{\sim}{Y} \to \underset{\sim\sim}{HY}$, and notice that

$$\int_{O(p-2)} \exp\{-\tfrac{1}{2} \underset{\sim}{L_2'} \underset{\sim}{\Delta} \underset{\sim}{L_2} \underset{\sim\sim\sim}{HYY'} \underset{\sim}{H'}\} d\underset{\sim}{H}$$

$$= {}_0F_0(-\tfrac{1}{2} \underset{\sim}{L_2'} \underset{\sim}{\Delta} \underset{\sim}{L_2}, \underset{\sim\sim}{YY'})$$

$$= \sum_{k=0}^{\infty} \sum_{\kappa} (-1)^k C_\kappa(\tfrac{1}{2} \underset{\sim}{L_2'} \underset{\sim}{\Delta} \underset{\sim}{L_2}) C_\kappa(\underset{\sim\sim}{YY'}) \{k!\, C_\kappa(\underset{\sim}{I})\}^{-1}$$

and put

$$(3.9) \qquad \omega_i = \tfrac{1}{2}\, \underset{\sim}{\ell}_i'\, \underset{\sim}{\Delta}\, \underset{\sim}{\ell}_i \qquad i = 1,\ 2.$$

Then (3.7) can be written

$$(3.7') \qquad f(\theta,\ \lambda_1,\ \lambda_2,\ \underset{\sim}{L}) =$$

$$e^{-\omega_1\lambda_1-\omega_2\lambda_2} \sum_{k=0}^{\infty} \sum_{\kappa} (-1)^k C_\kappa(\tfrac{1}{2}\, \underset{\sim}{L}_2'\, \underset{\sim}{\Delta}\, \underset{\sim}{L}_2)\ \{k!\ C_\kappa(\underset{\sim}{I})\}^{-1} h_k(\lambda_1,\ \lambda_2)$$

where

$$h_k(\lambda_1,\ \lambda_2) = \int_{\mathcal{D}} |\lambda_1 \underset{\sim}{I} - \underset{\sim}{YY'}| \cdot |\lambda_2 \underset{\sim}{I} - \underset{\sim}{YY'}|\, C_\kappa(\underset{\sim}{YY'})\ d\underset{\sim}{Y} \ .$$

Using Lemma 2, and since $\alpha = 1$, then for $q \geq (p-2)+1$, all coefficients in (2.7) vanish, so that the function reduces to a polynomial of degree $p-2$. Hence

$$(3.10) \qquad h_k(\lambda_1,\ \lambda_2) = \pi^{\frac{1}{2}(p-2)(n-2)} \Gamma_{p-2}\ (\tfrac{p+1}{2}) \left\{\Gamma_{p-2}\ (\tfrac{p+n-1}{2})\right\}^{-1}$$

$$\cdot \sum_{q=0}^{p-2} \sum_{\eta} \sum_{\delta} \frac{(-1)_\eta}{q!}\ g^{\delta}_{\kappa,\eta} C_\delta(\underset{\sim}{I}) (\tfrac{n-2}{2})_\delta$$

$$\cdot \left\{(\tfrac{p+n-1}{2})_\delta\right\}^{-1} \lambda_1^{\,p-q-2}\ \lambda_2^{\frac{1}{2}n(p-2)+q+k} \ .$$

At this stage, we integrate

$$(\lambda_1\lambda_2)^{\frac{1}{2}(n-p-1)}(\lambda_1 - \lambda_2)\ e^{-\omega_1\lambda_1-\omega_2\lambda_2}\lambda_1^{p-q-2}\lambda_2^{\frac{1}{2}n(p-2)+q+k}$$

i.e. $$\lambda_1^{\xi-q}\ \lambda_2^{\zeta+q+k}\ (\lambda_1 - \lambda_2)\ e^{-\omega_1\lambda_1-\omega_2\lambda_2}$$

with respect to λ_1 and λ_2, where

$$\xi = \tfrac{1}{2}(p+n-5)\ , \quad \zeta = \tfrac{1}{2}(pn-n-p-1).$$

First, integrating

$$\lambda_1^{\xi-q}\ \lambda_2^{\zeta+q+k}\ (\lambda_1 - \lambda_2)\ e^{-\omega_2\lambda_2}$$

with respect to λ_2 from 0 to λ_1 and using formula

$$\int_0^a x^{b-1}\ e^{-cx}\ dx = e^{-ca} \sum_{i=0}^{\infty} \frac{c^i\ a^{b+i}\ \Gamma(b)}{\Gamma(b+1+i)}\ .$$

we obtain

$$\int_0^{\lambda_1} \lambda_1^{\xi-q}\ \lambda_2^{\zeta+q+k}\ (\lambda_1 - \lambda_2)\ e^{-\omega_2\lambda_2}\ d\lambda_2$$

$$= \sum_{i=0}^{\infty} \left\{\frac{\Gamma(v)}{\Gamma(v+1+i)} - \frac{\Gamma(v+1)}{\Gamma(v+2+i)}\right\} \omega_2^i\ \lambda_1^{s+i-1}\ e^{-\omega_2\lambda_1}$$

where

(3.11) $\quad v = \zeta + q + k + 1, \quad s = \xi + \zeta + k + 3$.

Next, integrating $\lambda_1^{s+i-1} e^{-(\omega_i+\omega_2)\lambda_1}$ with respect to λ_1 from 0 to ∞, and notice that

$$\sum_{i=1}^{\infty} \frac{(1+i)\Gamma(v)}{\Gamma(v+2+i)} \Gamma(s+i) \left(\frac{\omega_2}{\omega_1+\omega_2}\right)^i$$

$$= \frac{\Gamma(v)\Gamma(s)}{\Gamma(v+2)} F\left(2, s, v+2; \frac{\omega_2}{\omega_1+\omega_2}\right) ,$$

where $F(\alpha, \beta, \gamma; x)$ is hypergeometric function defined as in [7]. Therefore,

$$\int_0^\infty \int_0^{\lambda_1} \lambda_1^{\xi-q} \lambda_2^{\zeta+q+k} (\lambda_1 - \lambda_2) e^{-\omega_1\lambda_1-\omega_2\lambda_2} d\lambda_2 d\lambda_1$$

$$= \Gamma(v) \Gamma(s) \{(\omega_1+\omega_1)^s \Gamma(v+2)\}^{-1} F\left(2, s, v+2; \frac{\omega_2}{\omega_1+\omega_2}\right) .$$

Substituting (3.10) into (3.7') we get $f(\theta, \lambda_1, \lambda_2, \underset{\sim}{L})$ and the coefficient of $\theta^k/k!$ from $f(\theta, \lambda_1, \lambda_2, \underset{\sim}{L})$ gives $f_k(\lambda_1, \lambda_2, \underset{\sim}{L})$. Then using this value in (3.5'), we get the joint density function of $\underset{\sim}{\ell_1}, \underset{\sim}{\ell_2}, \lambda_1$, and λ_2. Integrating λ_1 and λ_2 we obtain the joint density function of $\underset{\sim}{\ell_1}$ and $\underset{\sim}{\ell_2}$. Hence we have the following theorem:

<u>Theorem</u>. Let the matrix $\underset{\sim}{X}(pxn)$ be distributed as (3.1), and $\lambda_1 > \lambda_2 > 0$ be the two largest ordered ch. roots of $\underset{\sim}{XX'}$ and let

$$\underset{\sim}{L} = (\underset{\sim}{\ell_1}, \underset{\sim}{\ell_2}, \underset{\sim}{L_2})$$

where $\underset{\sim}{\ell}_1$ and $\underset{\sim}{\ell}_2$ are the two columns of $\underset{\sim}{L}$, corresponding to the two largest ordered ch. roots λ_1 and λ_2 of $\underset{\sim}{X}\underset{\sim}{X}'$, having random elements and the others $\underset{\sim}{L}_2$ depend on these random elements. Let the rank of $\underset{\sim}{M}$ be one. Then the joint density function of $\underset{\sim}{\ell}_1$ and $\underset{\sim}{\ell}_2$ is given by

$$|\Sigma|^{-\frac{1}{2}n} e^{-\frac{1}{2}\mathrm{tr}\underset{\sim}{\Sigma}^{-1}\underset{\sim}{M}\underset{\sim}{M}'} \sum_{k=0}^{\infty} \left\{k!\ (\tfrac{n}{2})_k\right\}^{-1} \ d\underset{\sim}{L}^{(p,2)}\ f_k\ (\underset{\sim}{L})$$

where $f_k\ (\underset{\sim}{L})$ satisfying

$$f(\theta,\underset{\sim}{L}) = \sum_{k=0}^{\infty} \frac{\theta^k}{k!}\ f_k\ (\underset{\sim}{L})$$

$$= c_3 \sum_{k=0}^{\infty} \sum_{\kappa} (-1)^k\ C_\kappa(\tfrac{1}{2}\ \underset{\sim}{L}_2'\ \underset{\sim}{\Delta}\ \underset{\sim}{L}_2)\ \{k!\ C_\kappa(\underset{\sim}{I})\}^{-1}$$

$$\cdot \sum_{q=0}^{p-2} \sum_{\eta} \sum_{\delta} (-1)_\eta\ g_{\kappa,\eta}^{\delta}\ C_\delta(\underset{\sim}{I})(\tfrac{n-2}{2})_\delta \Gamma(v)\Gamma(s)$$

$$\cdot \left\{q!(\tfrac{p+n-1}{2})_\delta(\omega_1+\omega_2)^s\Gamma(v+2)\right\}^{-1} F(2,s,v+2;\frac{\omega_2}{\omega_1+\omega_2})\ ,$$

where

$$(3.12) \qquad c_3 = \pi^2\Gamma_{p-2}(\tfrac{p+1}{2})\left\{2^{\frac{1}{2}pn}\Gamma_{p-2}(\tfrac{p+n-1}{2})\Gamma_2(\tfrac{p}{2})\Gamma_2(\tfrac{n}{2})\right\}^{-1}$$

and v and s, $\underset{\sim}{\Delta}$ and ω_i are defined by (3.11), (3.8) and (3.9).

Note that the explicit expression will be obtained by evaluating the coefficient of $\theta^k/k!$ from $f(\theta,\underset{\sim}{L})$.

4. Remarks: (I). If $\underset{\sim}{M} = \underset{\sim}{0}$, then $\underset{\sim}{\Delta} = \underset{\sim}{\Sigma}^{-1}$ and $\underset{\sim}{Q}$ and $(\lambda_i = \alpha_i^2, \underset{\sim}{Y}, \underset{\sim}{L})$ are independently distributed and their respective density functions are given by

$$d\underset{\sim}{Q}^{(n,r)}$$

and

$$(4.1)\qquad C_1|\underset{\sim}{\Sigma}|^{-\frac{1}{2}n} \prod_{i=1}^{r} \lambda_i^{\frac{1}{2}(n-p-1)} |\lambda_i \underset{\sim}{I} - \underset{\sim}{Y}\underset{\sim}{Y}'| \prod_{i<j} (\lambda_i - \lambda_j)$$

$$\cdot \exp\left\{-\tfrac{1}{2} \operatorname{tr} \underset{\sim}{\Sigma}^{-1} \underset{\sim}{L} \begin{pmatrix} \lambda_1 & & & 0 \\ & \ddots & & \\ & & \lambda_r & \\ 0 & & & \underset{\sim}{Y}\underset{\sim}{Y}' \end{pmatrix} \underset{\sim}{L}'\right\} d\underset{\sim}{L}^{(p,r)}$$

where C_1 is defined by (3.3).

For $r = 1$, integrating (4.1) with respect to $\lambda\ (=\lambda_1)$ and $\underset{\sim}{Y}$, we get the same density function of $\underset{\sim}{L}$ as given by Sugiyama [10] and Khatri and Pillai [8].

For $r = 2$, we obtain the joint density function of $\underset{\sim}{\ell}_1$ and $\underset{\sim}{\ell}_2$ is given by

$$(4.2)\qquad C_3|\underset{\sim}{\Sigma}|^{-\frac{1}{2}n} \sum_{k=0}^{\infty} \sum_{\kappa} (-1)^k 2^{-k} C_\kappa(\underset{\sim}{L}_2' \underset{\sim}{\Sigma}^{-1} \underset{\sim}{L}_2)\{k!\, C_\kappa(\underset{\sim}{I})\}^{-1}$$

$$\cdot \sum_{q=0}^{p-2} \sum_{\eta} \sum_{\delta} (-1)_\eta\, g_{\kappa,\eta}^{\delta}\, C_\delta(\underset{\sim}{I}) \left(\tfrac{n-2}{2}\right)_\delta \Gamma(v)\Gamma(s)$$

$$\cdot \left\{q!\, \left(\tfrac{p+n-1}{2}\right)_\delta\, (\omega_1+\omega_2)^s \Gamma(v+2)\right\}^{-1} F\left(2,\ s,\ v+2;\ \frac{\omega_2}{\omega_1+\omega_2}\right) d\underset{\sim}{L}^{(p,2)}$$

where C_3 is defined by (3.12). (4.2) is a special case of the Theorem.

(II). Put $r = p$, integrating (4.1) with respect to $\underset{\sim}{L}$, we get the same distribution of ch. roots $\lambda_1, \ldots, \lambda_p$ of $\underset{\sim}{XX'}$ as given by A. T. James [7].

(III). If $n \leq p$, then in the all adequate formulas change the roles of p and n.

Acknowledgment. The author wishes to express his very sincere thanks to Professor K. C. S. Pillai of Purdue University for suggesting the problem and guidance.

References

[1] Anderson, T. W. (1951). The asymptotic distribution of certain characteristic roots and vectors. Proc. Second Berkeley Symp. Math. Statist. Prob. 103-130. Univ. of California Press.

[2] Anderson, T. W. (1963). Asymptotic theory for principal component analysis. Ann. Math. Statist. 34, 122-148.

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

[4] Constantine, A. G. (1966). The distribution of Hotelling's generalised To^2. Ann. Math. Statist. 37, 215-225.

[5] James, A. T. (1961). The distribution of non-central means with known covariance. Ann. Math. Statist. 32, 874-882.

[6] James, A. T. (1961). Zonal polynomial of the real positive definite symmetric matrices. Ann. of Math. 74, 456-469.

[7] James, A. T. (1964). Distributions of matrix variates and latent roots derived from normal samples. Ann. Math. Statist. 35, 475-501.

[8] Khatri, C. G. and Pillai, K. C. S. (1968). Distributions of vectors corresponding to the largest roots of three matrices. Mimeograph Series No. 160., Department of Statistics, Purdue University.

[9] Roy, S. N. (1958). Some Aspects of Multivariate Analysis, Wiley and Sons, New York.

[10] Sugiyama, T. (1966). On the distribution of the largest latent root and the corresponding latent vector for principal component analysis. Ann. Math. Statist. 37, 995-1001.