ON THE DISTRIBUTION OF THE LARGEST LATENT ROOT AND THE CORRESPONDING LATENT VECTOR FOR PRINCIPAL COMPONENT ANALYSIS

By T. Sugiyama

Aoyama Gakuin University

1. Summary. The distribution of the latent vectors of a sample covariance matrix was found by T. W. Anderson [1] in 1951 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 [3] in 1963. The exact distribution of the latent vectors of a sample covariance matrix has been described by the author [10] in 1965 when the observations are obtained from a bi-variate normal distribution. The elements of each latent vector are the coefficients of a principal component (with sum of squares of coefficients being unity), and the corresponding latent root is the variance of the principal component. In this paper, the exact distribution of the latent vector corresponding to the largest latent root of a sample covariance matrix is given when the observations are from a multivariate normal distribution whose population covariance matrix is arbitrary Σ, and the distribution of the largest latent root is given when the population covariance matrix is a scalar matrix, $\Sigma = \sigma^2 I$.

2. Introduction. Let χ be a p-component random vector with mean vector μ and covariance matrix $\mathcal{E}(\chi - \mu)(\chi - \mu)' = \Sigma$ where

$$\Sigma = \begin{pmatrix} \sigma_{11} & \cdots & \sigma_{1p} \\ \vdots & & \vdots \\ \sigma_{1p} & \cdots & \sigma_{pp} \end{pmatrix}.$$

The variance of a linear combination $\gamma_1'\chi$ is

$$\mathcal{E}(\gamma_1'\chi - \mathcal{E}\gamma_1'\chi)^2 = \mathcal{E}\gamma_1'(\chi - \mu)(\chi - \mu)'\gamma_1 = \gamma_1'\Sigma\gamma_1 . \tag{2.1}$$

The linear combination normalized by $\gamma_1'\gamma_1 = 1$ which has maximum variance may be called the first principal component of χ and the coefficients γ_1 the latent vector corresponding to the first principal component. The linear combination uncorrelated with the first principal component and similarly normalized which has maximum variance may be called the second principal component and the coefficients the latent vector corresponding to the second principal component. The other $p - 2$ principal components and the latent vectors corresponding to each principal component are similarly defined. The variances of principal components and the corresponding latent vectors are estimated by the latent roots and the corresponding latent vectors of the estimated covariance matrix.

Let

Received 13 December 1965.

$$X = \begin{pmatrix} \chi_{11} & \cdots & \chi_{1N} \\ \vdots & & \vdots \\ \chi_{p1} & \cdots & \chi_{pN} \end{pmatrix}$$

be the sample matrix of N observations from $N(\mu, \Sigma)$. It is well known that

$$(2.2) \qquad nS = (ns_{ij}) = \left(\textstyle\sum_{\xi=1}^{N}(\chi_{i\xi} - \bar{\chi}_i)(\chi_{j\xi} - \bar{\chi}_j)\right), \qquad n = N - 1,$$

has the Wishart distribution on n degrees of freedom. Since S is positive symmetric matrix we can write $S = HDH'$, where H is the $p \times p$ orthogonal matrix and

$$D = \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdot & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & d_p \end{pmatrix}$$

with $d_1 > d_2 > \cdots > d_p > 0$. The first column of the orthogonal matrix H, that is, the latent vector corresponding to the largest latent root d_1, is represented by the $p - 1$ independent elements. Explicitly it may be written the following way:

$$\begin{aligned} h_1' &= (h_{11}h_{12} \cdots h_{1i} \cdots h_{1p-1}\, h_{1p}) \\ (2.3) \qquad &= (\cos\theta_{11} \sin\theta_{11} \cos\theta_{12} \cdots \textstyle\prod_{\nu=1}^{i-1} \sin\theta_{1\nu} \cos\theta_{1i} \\ &\quad \cdots \textstyle\prod_{\nu=1}^{p-2} \sin\theta_{1\nu} \cos\theta_{1p-1} \prod_{\nu=1}^{p-2} \sin\theta_{1\nu} \sin\theta_{1p-1}). \end{aligned}$$

The purpose of this paper is to find the distribution of the latent vector corresponding to the largest latent root d_1 for non-null case, and to find the distribution of the largest latent root d_1 for null case.

3. Notation and preliminary results.

3.1. *On the Jacobian of the orthogonal transformation.* In research on distribution problems in multivariate analysis, Tumura's theorem for the Jacobian of the orthogonal transformation is very useful. The orthogonal matrix is represented in terms of rotation angles. The $p \times p$ orthogonal matrix has only $p(p - 1)/2$ independent elements, and every rotation in the p-dimensional space consists of $p(p - 1)/2$ single rotations which is such a rotation in the two-dimensional plane. Let $R_p^{\nu}(\theta)$ be a single rotation matrix defined by

$$(3.1) \qquad R_p^{\nu}(\theta) = \begin{pmatrix} I_{\nu-1} & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & I_{p-\nu-1} \end{pmatrix},$$

where I_ν is the identity matrix ($\nu \times \nu$). Let

$$(3.2) \qquad H_p^{\nu}(\theta_j) = R_p^{p-1}(\theta_{p-1}) R_p^{p-2}(\theta_{p-2}) \cdots R_p^{\nu}(\theta_\nu),$$

and

$$H_p(\theta_{ij}) = H_p^1(\theta_{1j})H_p^2(\theta_{2j}) \cdots H_p^{p-1}(\theta_{p-1p-1}). \tag{3.3}$$

The matrix $H_p(\theta_{ij})$ defined by (3.3) is the general form of the orthogonal matrix. Therefore, explicitly we can write

$$H_p(\theta_{ij}) = H_p^1(\theta_{1j}) \begin{pmatrix} 1 & 1 \\ 0 & H_{p-1}(\theta_{ij}) \end{pmatrix}, \tag{3.4}$$

where $H_{p-1}(\theta_{ij})$ is arbitrary orthogonal matrix of the $p - 1$ dimensional space.

LEMMA 3.1. (Tumura's theorem). *The Jacobian of the transformation* $U = H\Lambda H'$ *is*

$$J(U; \lambda, \theta) = \text{mod} \prod_{i<j}^{p} (\lambda_i - \lambda_j) \prod_{i=1}^{p-2} \prod_{j=i}^{p-2} \sin^{p-i-1} \theta_{ij}, \tag{3.5}$$

where U is $p \times p$ *symmetric matrix,* Λ, *is a* $p \times p$ *diagonal matrix with diagonal elements in descending order,* $H = H_p(\theta_{ij})$, *that is, the orthogonal matrix defined by* (3.3), *and* θ*'s range in the intervals:* $0 \leqq \theta_{ij} \leqq \pi$ $(j \leqq p - 2)$, $0 \leqq \theta_{ip-1} \leqq 2\pi$.

The proof has been described in detail in (11).

3.2. *On a property of the zonal polynomials.* With S any positive definite symmetric $p \times p$ matrix the zonal polynomials $Z_\kappa(S)$ are defined for each partition $\kappa = (k_1, \cdots, k_p)$, $k_1 \geqq \cdots \geqq k_p \geqq 0$, of k into not more than p parts, as certain symmetric polynomials in the latent roots of S. A detailed discussion of zonal polynomials may be found in A. T. James [7], and [8]. The notation with respect to the zonal polynomial given by A. G. Constantine [4] is used here, that is,

$$C_\kappa(S) = c(\kappa)Z_\kappa(S)/1\cdot 3 \cdots (2k - 1). \tag{3.6}$$

The fundamental property of the zonal polynomials is given by the following integral, proved in [7]:

$$\int_{O(m)} C_\kappa(H'SHT)\, d(H) = C_\kappa(S)C_\kappa(T)/C_\kappa(I), \tag{3.7}$$

where I is the identity matrix, and $d(H)$ is the invariant Haar measure on the orthogonal group, normalized to make the volume of the group manifold unity. Another derivation of this property has been given by Y. Tumura by representing in terms of rotation angles the orthogonal matrix.

LEMMA 3.2. *Let S and T be positive definite* $p \times p$ *matrices.*

$$(1/C) \int (\text{tr } H'SHT)^k \prod_{i=1}^{p-2} \prod_{j=i}^{p-2} \sin^{p-j-1} \theta_{ij} \prod_{i=1}^{p-1} \prod_{j=i}^{p-1} d\theta_{ij} = \sum_\kappa C_\kappa(S)C_\kappa(T)/C_\kappa(I), \tag{3.8}$$

where κ *is a partition of k, and* θ*'s range in the integral is* $0 \leqq \theta_{ij} \leqq \pi$ $(j \leqq p - 2)$, $0 \leqq \theta_{ip-1} \leqq 2\pi$. *The constant C, also, is given by*

$$C = \int \prod_{i=1}^{p-2} \prod_{j=i}^{p-2} \sin^{p-j-1} \theta_{ij} \prod_{i=1}^{p-1} \prod_{j=i}^{p-1} d\theta_{ij} = \pi^{p^2/2}/\Gamma_p(p/2), \tag{3.9}$$

where $\Gamma_p(u) = \pi^{p(p-1)/4} \prod_{i=1}^{p} \Gamma(u - (i - 1)/2)$.

3.3. *On a beta-function integral.* Let S be a positive definite symmetric $m \times m$ matrix, and R a positive definite $m \times m$ matrix, then

$$\int_0^I |S|^{t-(m+1)/2}|I - S|^{u-(m+1)/2} C_\kappa(RS)\, dS = \{\Gamma_m(t, \kappa)\Gamma_m(u)/\Gamma_m(t+u, \kappa)\} C_\kappa(R), \tag{3.10}$$

where κ is a partition $(k_1, \cdots, k_m)$ of kth degree, valid for all complex numbers t and u satisfying real $(t) > (m-1)/2$ and real $(u) > (m-1)/2$, and

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

The proof has been described in detail in (4). Let $R = I$ in (3.10) and consider the following transformation: $S = H\Lambda H'$, where H is an orthogonal matrix, and Λ is a diagonal matrix. By use of Lemma 2.1. on the left hand side of (3.10) we can write it as follows:

$$\begin{aligned} &\int_{1>\lambda_1>\cdots>\lambda_m>0} |\Lambda|^{t-(m+1)/2}|I - \Lambda|^{u-(m+1)/2} C_\kappa(\Lambda) \prod_{i<j}^m (\lambda_i - \lambda_j)\, d\Lambda \\ &\cdot \int \prod \sin^{m-j-1}\theta_{ij} \prod d\theta_{ij} = (\pi^{m^2/2}/\Gamma_m(m/2)) \\ &\cdot \int_{1>\lambda_1>\cdots>\lambda_m>0} |\Lambda|^{t-(m+1)/2}|I - \Lambda|^{u-(m+1)/2} C_\kappa(\Lambda) \prod_{i<j}^m (\lambda_i - \lambda_j)\, d\Lambda. \end{aligned} \tag{3.11}$$

Therefore we obtain the following lemma.

LEMMA 3.3. *Let* Λ *be a diagonal matrix with diagonal elements* $1 > \lambda_1 > \cdots > \lambda_m > 0$. Then,

$$\begin{aligned} &\int_{1>\lambda_1>\cdots>\lambda_m>0} |\Lambda|^{t-(m+1)/2} \prod_{i=1}^m (1-\lambda_i)^{u-(m+1)/2} \prod_{i<j}^m (\lambda_i - \lambda_j) \\ &\qquad \cdot C_\kappa(\Lambda) \prod_{i=1}^m d\lambda_i = (\Gamma_m(m/2)/\pi^{m^2/2}) \\ &\qquad \cdot (\Gamma_m(t, \kappa)\Gamma_m(u)/\Gamma_m(t+u, \kappa)) C\kappa(I). \end{aligned} \tag{3.12}$$

4. On the distribution of the latent vector corresponding to the largest latent root of sample covariance matrix. Let U have the Wishart distribution $W(p, n, \Sigma)$, where U corresponds to nS defined by Section 2. The probability elements of U are

$$K\, |U|^{(n-p-1)/2} \exp\left(-\tfrac{1}{2} \operatorname{tr} \Sigma^{-1}U\right) dU, \tag{4.1}$$

where $K = |\Sigma|^{-n/2}/2^{np/2}\Gamma_p(n/2)$. We shall make the following transformation:

$$U = HD_\lambda H', \tag{4.2}$$

where D_λ is diagonal with diagonal elements $\lambda_1 > \lambda_2 > \cdots > \lambda_p > 0$, and H is orthogonal matrix. By use of (3.4) we can write

$$\begin{aligned} \operatorname{tr} \Sigma^{-1}U &= \operatorname{tr} \Sigma^{-1}H_1 \begin{pmatrix} 1 & 0 \\ 0 & H_{p-1} \end{pmatrix} \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_p \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & H_{p-1} \end{pmatrix}' H_1' = h_1'\Sigma^{-1}h_1\lambda_1 \\ &+ \operatorname{tr} H_1'\Sigma^{-1}H_1 \begin{pmatrix} 1 & 0 \\ 0 & H_{p-1} \end{pmatrix} \begin{bmatrix} 0 & \cdots & & 0 \\ & \lambda_2 & & \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & & \lambda_p \end{bmatrix} \begin{pmatrix} 1 & 0 \\ 0 & H_{p-1} \end{pmatrix}', \end{aligned} \tag{4.3}$$

where $\lambda_1 > \lambda_2 > \cdots > \lambda_p > 0$, the first column of H_1 is the latent vector h_1, H_1 is the orthogonal matrix $H_p^1(\theta_{1j})$ with $p - 1$ independent elements, $\theta_{11}, \theta_{12}, \cdots, \theta_{1p-1}$, H_{p-1} is the orthogonal matrix of the $p - 1$ dimensional space with $(p - 1)(p - 2)/2$ independent elements, θ_{ij}, $i = 2, \cdots, p - 1$, $j = i, \cdots, p - 1$, and h_1 is given by (2.3). We note that H_1 depends only on the same $p - 1$ parameters as h_1. Σ_{p-1} is the $(p - 1) \times (p - 1)$ matrix obtained from $H_1'\Sigma^{-1}H_1$ by deleting the first row and column, and

$$\Lambda = \begin{pmatrix} \lambda_2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_p \end{pmatrix}.$$

Then, by use of Lemma 3.1 we have from (4.1), the probability element of θ_{ij}, $i = 1, \cdots, p - 1, j = i, \cdots, p - 1$, and λ_i, $i = 1, \cdots, p$,

$$\begin{aligned} &K \cdot (\lambda_1|\Lambda|)^{(n-p-1)/2} \exp\left(-\tfrac{1}{2} h_1' \Sigma^{-1} h_1\lambda_1\right) \\ (4.4)\quad &\cdot \exp \operatorname{tr}\left(-\tfrac{1}{2}\Sigma_{p-1}H_{p-1}\Lambda H_{p-1}'\right) \textstyle\prod_{i=2}^{p-2} \prod_{j=i}^{p-2} \sin^{p-j-1} \theta_{ij} \prod_{i=2}^{p-1} \prod_{j=i}^{p-1} d\theta_{ij} \\ &\cdot \textstyle\prod_{i<j}^{p} (\lambda_i - \lambda_j) \prod_{j=1}^{p-2} \sin^{p-j-1} \theta_{1j} \prod_{j=1}^{p-1} d\theta_{1j} \prod_{i=1}^{p} d\lambda_i . \end{aligned}$$

Integrating (4.4) with respect to $(p - 1)(p - 2)/2$ independent elements of H_{p-1} we have from Lemma 3.2 the probability element of θ_{1j}, $j = 1, \cdots, p - 1$, and λ_i, $i = 1, \cdots, p$,

$$\begin{aligned} &K \cdot \left(\pi^{(p-1)^2/2}/\Gamma_{p-1}((p - 1)/2)\right) \cdot (\lambda_1 |\Lambda|)^{(n-p-1)/2} \\ (4.5)\quad &\cdot \exp\left(-\tfrac{1}{2}h_1' \Sigma^{-1} h_1\lambda_1\right) \textstyle\prod_{j=1}^{p-2} \sin^{p-j-1} \theta_{1j} \prod_{j=1}^{p-1} d\theta_{ij} \\ &\cdot \textstyle\prod_{i<j}^{p} (\lambda_i - \lambda_j) \sum_{k=0}^{\infty} \sum_{\kappa} \left(C_\kappa(-\tfrac{1}{2}\Sigma_{p-1})C_\kappa(\Lambda)/k!\; C_\kappa(I_{p-1})\right) \prod_{i=1}^{p} d\lambda_i . \end{aligned}$$

We shall integrate (4.5) with respect to $\lambda_2, \cdots, \lambda_p$, where the range of the integral is $\lambda_1 > \lambda_2 > \cdots > \lambda_p > 0$. Let $\lambda_i = \lambda_1 l_i$, then the range of the integral result in $1 > l_2 > \cdots > l_p > 0$. Since we can write

$$\begin{aligned} &\textstyle\int_{\lambda_1>\lambda_2>\cdots>\lambda_p>0} (\lambda_1 |\Lambda|)^{(n-p-1)/2} \prod_{i<j}^{p} (\lambda_i - \lambda_j)C_\kappa(\Lambda) \prod_{i=2}^{p} d\lambda_i \\ &\qquad = \lambda_1^{(pn+2k)/2-1} \cdot \textstyle\int_{1>l_2>\cdots>l_p>0} |\Lambda_l|^{(n-p-1)/2} \\ &\qquad\qquad \cdot \textstyle\prod_{i=2}^{p} (1 - l_i) \prod_{i<j=2}^{p} (l_i - l_j)C_\kappa(\Lambda_l) \prod_{i=2}^{p} dl_i , \end{aligned}$$

by letting $m = p - 1$, $t = (n - 1)/2$, and $u = (p + 2)/2$ in Lemma 3.3 we have the probability element of θ_{1j}, $j = 1, \cdots, p - 1$, and λ_1,

$$\begin{aligned} &K \cdot \Gamma_{p-1}((p + 2)/2) \\ (4.6)\quad &\cdot \textstyle\sum_{k=0}^{\infty} \sum_{\kappa}\{C_\kappa(-\tfrac{1}{2}\Sigma_{p-1})\Gamma_{p-1}((n - 1)/2, \kappa)/k!\Gamma_{p-1}((n + p + 1)/2, \kappa)\} \\ &\cdot \lambda_1^{(pn+2k)/2-1} \exp\left(-\tfrac{1}{2}h_1' \Sigma^{-1} h_1\lambda_1\right) d\lambda_1 \textstyle\prod_{j=1}^{p-2} \sin^{p-j-1} \theta_{1j} \prod_{j=1}^{p-1} d\theta_{1j} . \end{aligned}$$

Integrating (4.6) with respect to λ_1 we obtain the probability element of

θ_{ij}, $j = 1, \cdots, p-1$,

$$(4.7)\quad \begin{aligned}&(|\Sigma|^{-n/2}\Gamma_{p-1}((p+2)/2)/\Gamma_p(n/2))\\ &\cdot \textstyle\sum_{k=0}^{\infty}\sum_{\kappa}\{\Gamma_{p-1}((n-1)/2,\kappa)\Gamma(pn/2+k)/k!\Gamma_{p-1}((n+p+1)/2,\kappa)\}\\ &\cdot C_\kappa(-\Sigma_{p-1}/h_1'\Sigma^{-1}h_1)(h_1'\Sigma^{-1}h_1)^{-pn/2}\textstyle\prod_{j=1}^{p-2}\sin^{p-j-1}\theta_{1j}\prod_{j=1}^{p-1}d\theta_{1j}.\end{aligned}$$

Let $(a)_\kappa = \prod_{i=1}^{p}(a-(i-1)/2)_{k_i}$, $\kappa = (k_1, \cdots, k_p)$ and, as usual, if a is such that the gamma functions are defined, then $(a)_\kappa = \Gamma_p(a,\kappa)/\Gamma_p(a)$. Therefore, (4.7) may be rewritten as

$$(4.8)\quad \begin{aligned}&(|\Sigma|^{-n/2}\Gamma(pn/2)/\Gamma_p(n/2))B_{p-1}((p+2)/2,\ (n-1)/2)\\ &\cdot \textstyle\sum_{k=0}^{\infty}\{(pn/2)_k/k!\}\sum_{\kappa}\{((n-1)/2)_\kappa/((n+p+1)/2)_\kappa\}\\ &\cdot C_\kappa(-\Sigma_{p-1}/h_1'\Sigma^{-1}h_1)(h_1'\Sigma^{-1}h_1)^{-pn/2}\textstyle\prod_{j=1}^{p-2}\sin^{p-j-1}\theta_{1j}\prod_{j=1}^{p-1}d\theta_{1j},\end{aligned}$$

where $B_{p-1}(\alpha,\beta) = (\Gamma_{p-1}(\alpha)\Gamma_{p-1}(\beta)/\Gamma_{p-1}(\alpha+\beta))$, Σ_{p-1} is the $(p-1)\times(p-1)$ matrix obtained from $H_1'\Sigma^{-1}H_1$ by deleting the first row and column, the first column of H_1 is the latent vector h_1, H_1 depends only on the same $p-1$ parameters as h_1, and $0 \leqq \theta_{1i} \leqq \pi$ $(i \leqq p-2)$, $0 \leqq \theta_{1p-1} \leqq 2\pi$. Therefore we obtain the following theorem:

THEOREM 1. *Let U have the Wishart distribution $W(p, n, \Sigma)$, then the distribution of the latent vector corresponding to the largest latent root of the positive definite symmetric matrix U is given by* (4.8).

Let $\Sigma = I$ in (4.8), then (4.8) can be written as follows:

$$(4.9)\quad \text{Const}\cdot\textstyle\prod_{j=1}^{p-2}\sin^{p-j-1}\theta_{1j}\prod_{j=1}^{p-1}d\theta_{1j}.$$

Integrating (4.9) with respect to θ_{1j} we obtain Const $= \Gamma(p/2)/2\pi^{p/2}$. Therefore we have the probability element of θ_{1j}, $j = 1, \cdots, p-1$, in the null case;

$$(4.10)\quad (\Gamma(p/2)/2\pi^{p/2})\cdot\textstyle\prod_{j=1}^{p-2}\sin^{p-j-1}\theta_{1j}\prod_{j=1}^{p-1}d\theta_{1j}.$$

Also, let $p = 2$ in (4.8), then using the Kummer transformation formula we find the same result as given in (10).

5. On the distribution of the largest latent root of sample covariance matrix when the population covariance matrix is a scalar matrix. Let $\Sigma = I$ in (4.6). Then we have the probability element of θ_{1j}, $j = 1, \cdots, p-1$, and λ_1,

$$(5.1)\quad \begin{aligned}&K\cdot\Gamma_{p-1}((p+2)/2)\\ &\cdot\textstyle\sum_{k=0}^{\infty}\sum_{\kappa}\{C_\kappa(-\frac{1}{2}I_{p-1})\Gamma_{p-1}((n-1)/2,\kappa)/k!\Gamma_{p-1}((n+p+1)/2,\kappa)\}\\ &\cdot\lambda_1^{(pn+2k)/2-1}\exp(-\tfrac{1}{2}\lambda_1)\,d\lambda_1\textstyle\prod_{j=1}^{p-2}\sin^{p-j-1}\theta_{1j}\prod_{j=1}^{p-1}d\theta_{1j}.\end{aligned}$$

Integrating (5.1) with respect to θ_{1j} we obtain the distribution of the largest latent root

$$(5.2)\quad \begin{aligned}&(\pi^{p/2}/2^{np/2}\Gamma(p/2)\Gamma_p(n/2))\Gamma_{p-1}((p+2)/2)\\ &\cdot\textstyle\sum_{k=0}^{\infty}\sum_{\kappa}\{C_\kappa(-\frac{1}{2}I_{p-1})\Gamma_{p-1}((n-1)/2,\kappa)/k!\Gamma_{p-1}((n+p+1)/2,\kappa)\}\\ &\cdot\lambda_1^{(pn+2k)/2-1}\exp(-\tfrac{1}{2}\lambda_1)\,d\lambda_1.\end{aligned}$$

This result will be, also, easily derived from the formula that is obtained by letting $\Sigma = I$ in (4.1) by use of Lemma 3.3 and

$$\exp(\operatorname{tr} Z) = \sum_{k=0}^{\infty} \sum_{\kappa} C_\kappa(Z)/k!, \tag{5.3}$$

where Z is a complex symmetric matrix. A. G. Constantine [4] has defined as the series of zonal polynomials the hypergeometric functions ${}_pF_q(Z)$ of a complex symmetric matrix Z defined by C. S. Herz [5] by means of a multidimensional form of the Laplace transform. Namely

$$\begin{aligned} {}_pF_q(a_1, \cdots, a_p; b_1, \cdots, b_q; Z) \\ = \sum_{k=0}^{\infty} \sum_{\kappa} [(a_1)_\kappa \cdots (a_p)_\kappa C_\kappa(Z)/(b_1)_\kappa \cdots (b_q)_\kappa k!]. \end{aligned}$$

By this definition, (5.2) may be rewritten as

$$\begin{aligned} &\{\pi^{p/2}/2^{np/2}\Gamma(p/2)\Gamma_p(n/2)\}B_{p-1}((p+2)/2, (n-1)/2) \\ &\cdot {}_1F_1((n-1)/2; (n+p+1)/2; -(\lambda_1/2)I_{p-1})\lambda_1^{pn/2-1} \exp(-\tfrac{1}{2}\lambda_1)\, d\lambda_1 . \end{aligned} \tag{5.4}$$

Therefore we obtain the following theorem:

THEOREM 2. *Let U have the Wishart distribution $W(p, n, I)$, then the distribution of the largest latent root of the positive definite symmetric matrix U is given by* (5.4).

6. Acknowledgments. The author wishes to express very sincere thanks to Professor Y. Tumura and Dr. K. Fukutomi of the College of Tokyo Science for their instructive suggestions throughout this study.

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. (1958). *An Introduction to Multivariate Statistical Analysis.* Wiley, New York.

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

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

[5] HERZ, C. S. (1955). Bessel functions of matrix argument. *Ann. Math.* **61** 474–523.

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

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

[8] JAMES, A. T. (1961). Zonal polynomials of the real positive definite symmetric matrices. *Ann. Math.* **74** 456–469.

[9] ROY, S. N. (1957). *Some Aspects of Multivariate Analysis.* Wiley, New York.

[10] SUGIYAMA, T. (1965). On the distribution of the latent vectors for principal component analysis. *Ann. Math. Statist.* **36** 1875–1876.

[11] TUMURA, Y. (1965). The distribution of latent roots and vectors. *Tokyo Rika Univ. Math.* **1** 1–16.