中 央 大 学 理 工 学 研 究 所 論 文 集　第 14 号　2008 年
Journal of the Institute of Science and Engineering. Chuo University

Behavior of Latent Vector of Trivariate Wishart Matrix

Masanori Ishizaki*, **Yuichi Takeda**†,
Takayuki Yamada‡ **and Takakazu Sugiyama***

abstract

This paper is concerned with the probability density function of the latent vector corresponding to the largest latent root of Wishart matrix. The latent vector may be expressed by the polar coordinates. Sugiyama (1966) give the exact expression of the probability density function of the polar coordinates. The function contained the alternating series, thus the function may not be converged on the domain of definition, numerically. In this paper we derived an improved expression of the function to be the positive series, for which we provide graphs of a population latent vector and latent roots.

1 Introduction

Let $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N$ be a random sample from the p-variate normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$, which is denoted by $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Then, the unbiased estimator of $\boldsymbol{\Sigma}$ is given by

$$\boldsymbol{S} = \frac{1}{N-1}\sum_{\alpha=1}^{N}(\boldsymbol{x}_\alpha - \bar{\boldsymbol{x}})(\boldsymbol{x}_\alpha - \bar{\boldsymbol{x}})'$$

where $\bar{\boldsymbol{x}} = (1/n)\sum_{\alpha=1}^{N}\boldsymbol{x}_\alpha$. In this case, $n\boldsymbol{S} = (N-1)\boldsymbol{S}$ is distributed according to the Wishart distribution with n degrees of freedom and the covariance matrix $\boldsymbol{\Sigma}$, denoted by $W_p(n, \boldsymbol{\Sigma})$. The latent roots and vectors of $\boldsymbol{S}$ have important roles for statistical inference. For example, the principal components analysis results in the linear combination of latent vectors of $\boldsymbol{\Sigma}$ and of the original variable vector. The variances of the principal components are given by the latent roots of $\boldsymbol{\Sigma}$. Generally these parameters are unknown, thus we must estimate them using the latent roots and vectors of $\boldsymbol{S}$. The statistical inferences containing the asymptotic distribution have been studied intensely. However, few papers have been published on latent vectors and how they may be expressed by their polar coordinates. Sugiyama (1965) gave the probability density function of the polar coordinates when $p = 2$. Later, Sugiyama (1966) extended this initial result to the general p. The main aim of this paper is to examine the stability of variation from the graph when $p = 3$. According to Sugiyama (1966), the probability density function is expressed by the alternating series, thus the

*Department of Mathematics, Graduate School of Sciences and Engineering, Chuo University, 1-13-27 Kasuga, Bunkyo-ku, Tokyo 112-8551, Japan
†Center for Basic Education and Integrated Learning, Kanagawa Institute of Technology, 1030 Shimo-Ogino, Atsugi-shi, Kanagawa 243-0292, Japan
‡Department of Clinical Medicine, School of Pharmacy, Kitasato University, 5-9-1, Shirokane, minato-ku, Tokyo, 108-8641, Japan

function may not be converged on the domain of definition numerically. Initially in this paper, we shall derive an expression by a positive term series, which is useful for numerical computation. We then drew the graphs of the probability density function when $p = 3$ and $n = 4, 10$ and 50. These graphs are unimodal on the region of definition.

The related works of this paper are as follows: Anderson (1963) gave an asymptotic distribution of the latent vectors of a Wishart matrix. Sugiura (1976) gave an asymptotic expansion of the distribution of the latent vector corresponding to the simple root of $\boldsymbol{\Sigma}$. Khatri and Pillai (1969) derived an exact distribution of the latent vectors corresponding to the largest latent roots in one and two sample cases. Takemura and Sheena (2007) derived an asymptotic normality of the latent vectors for the normalized sample latent roots when the population eigenvalues were infinitely dispersed.

This paper is organized as follows: in Section 2, we provide an improved expression of the probability density function of the latent vector. In Section 3, we show the graph of the probability density function for the case of $p = 3$. The conclusion of this paper follows in Section 4.

2 Improvement of the Density Function

In this section, we improve the expression of the probability density function of the latent vector corresponding to the largest latent root of the Wishart matrix that was given originally by Sugiyama (1966), modifying his argument.

Let $\boldsymbol{U}$ be distributed as $W_p(n, \boldsymbol{\Sigma})$. Then, it is well known that the probability density function of $\boldsymbol{U}$ is given by

$$K|\boldsymbol{U}|^{\frac{n-p-1}{2}} \exp\left(-\frac{1}{2}\mathrm{tr}\boldsymbol{\Sigma}^{-1}\boldsymbol{U}\right),$$

where $K = |\boldsymbol{\Sigma}|^{-\frac{n}{2}}/(2^{\frac{np}{2}}\Gamma_p(\frac{n}{2}))$ and $\Gamma_p(u) = \pi^{\frac{p(p-1)}{4}}\prod_{i=1}^{p}\Gamma(u-(i-1)/2)$. Consider the following spectral decomposition:

$$\boldsymbol{U} = \boldsymbol{H}\boldsymbol{D}_\ell\boldsymbol{H}', \tag{1}$$

where $\boldsymbol{D}_\ell$ is the diagonal matrix with diagonal elements $\ell_1 > \ell_2 > \cdots > \ell_p > 0$ and

$$\boldsymbol{H} = (\ \boldsymbol{h}_1,\ \boldsymbol{h}_2,\ \ldots,\ \boldsymbol{h}_p\ )$$

is $p \times p$ orthogonal matrix. Let $\boldsymbol{R}_\nu(t)$ be the single rotation matrix defined by

$$\boldsymbol{R}_\nu(t) = \begin{pmatrix} \boldsymbol{I}_{\nu-1} & 0 & 0 & 0 \\ 0 & \cos t & -\sin t & 0 \\ 0 & \sin t & \cos t & 0 \\ 0 & 0 & 0 & \boldsymbol{I}_{p-\nu-1} \end{pmatrix},$$

where $\boldsymbol{I}_\nu$ is the $\nu \times \nu$ identity matrix. Then, $\boldsymbol{H}$ may be expressed as

$$\boldsymbol{L}_1(t_{1\cdot})\cdots\boldsymbol{L}_{p-1}(t_{p-1,\cdot})\begin{pmatrix} 1 & \mathbf{0}' & 0 \\ \mathbf{0} & \boldsymbol{I}_{p-2} & \mathbf{0} \\ 0 & \mathbf{0}' & \varepsilon \end{pmatrix}, \tag{2}$$

where ε is 1 or -1, $t_{\nu\cdot} = (t_{\nu,p-1}, t_{\nu,p-2}, \ldots, t_{\nu,\nu})$ and

$$\boldsymbol{L}_\nu(t_{\nu\cdot}) = \boldsymbol{R}_{p-1}(t_{\nu,p-i})\boldsymbol{R}_{p-2}(t_{\nu,p-2})\cdots\boldsymbol{R}_\nu(t_{\nu,\nu}).$$

for $0 \le t_{ij} \le \pi$ with $1 \le j \le p-1$ and $0 \le t_{i,p-1} < 2\pi$. In this case, we may write

$$\boldsymbol{H} = \boldsymbol{H}(t_{..}) = \boldsymbol{L}_1(t_{1.}) \begin{pmatrix} 1 & \boldsymbol{0}' \\ \boldsymbol{0} & \boldsymbol{H}_{p-1}(\bar{t}_{..}) \end{pmatrix},$$

where $t_{..} = (t_{1.}, t_{2.}, \dots, t_{p-1,.})$, $\bar{t}_{..} = (t_{2.}, \dots, t_{p-1,.})$ and $\boldsymbol{H}_{p-1}(\bar{t}_{..})$ is an orthogonal matrix of degree $p-1$. The jacobian of the transformation $\boldsymbol{U} = \boldsymbol{H}\boldsymbol{D}_\ell \boldsymbol{H}'$ is given by

$$J = \prod_{i<j}^{p} (\ell_i - \ell_j) \prod_{i=1}^{p-2} \prod_{j=i}^{p-2} \sin^{p-i-1} t_{ij}. \tag{3}$$

Let $\boldsymbol{L} = \mathrm{diag}(\ell_2, \ell_3, \dots, \ell_p)$. Then, we can express $\mathrm{tr}\boldsymbol{\Sigma}^{-1}\boldsymbol{U}$ as

$$\mathrm{tr}\left[\boldsymbol{L}_1(t_{1.})'\boldsymbol{\Sigma}^{-1}\boldsymbol{L}_1(t_{1.}) \begin{pmatrix} 1 & \boldsymbol{0}' \\ \boldsymbol{0} & \boldsymbol{H}_{p-1}(\bar{t}_{..}) \end{pmatrix} \begin{pmatrix} 0 & \boldsymbol{0}' \\ \boldsymbol{0} & \boldsymbol{L} \end{pmatrix} \begin{pmatrix} 1 & \boldsymbol{0}' \\ \boldsymbol{0} & \boldsymbol{H}_{p-1}(\bar{t}_{..}) \end{pmatrix}'\right] + \ell_1 \boldsymbol{h}_1' \boldsymbol{\Sigma}^{-1} \boldsymbol{h}_1. \tag{4}$$

From (3) and (4) the joint probability density function of t_{ij} ($i = 1, \dots, p-1$ and $j = i, \dots, p-1$) and $(\ell_1, \dots, \ell_p)$ is

$$K(\ell_1|\boldsymbol{L}|)^{\frac{n-p-1}{2}} \exp\left(-\frac{1}{2}\ell_1 \boldsymbol{h}_1'\boldsymbol{\Sigma}^{-1}\boldsymbol{h}_1\right) \prod_{i<j}^{p} (\ell_i - \ell_j)$$
$$\cdot \exp\left\{\mathrm{tr}\left(-\frac{1}{2}\boldsymbol{\Sigma}_{p-1}(t_{1.})\boldsymbol{H}_{p-1}(\bar{t}_{..})\boldsymbol{L}\boldsymbol{H}_{p-1}(\bar{t}_{..})'\right)\right\} \prod_{i=1}^{p-2}\prod_{j=i}^{p-2} \sin^{p-i-1} t_{ij},$$

where $\boldsymbol{\Sigma}_{p-1}(t_{1.})$ is $(p-1) \times (p-1)$ matrix obtained from deleting the first row and column of $\boldsymbol{L}_1(t_{1.})'\boldsymbol{\Sigma}^{-1}\boldsymbol{L}_1(t_{1.})$. We use the following lemma (see e.g., Muirhead [1982]) to find the joint probability density function of $\boldsymbol{t}_1 = (t_{11}, \dots, t_{1,p-1})$ and $(\ell_1, \dots, \ell_p)$.

Lemma 2.1 *Let $\boldsymbol{S}$ and $\boldsymbol{T}$ be $p \times p$ positive definite matrices. Then,*

$$\frac{1}{c}\int (\mathrm{tr}\boldsymbol{H}'\boldsymbol{S}\boldsymbol{H}\boldsymbol{T})^k \prod_{i=1}^{p-2}\prod_{j=i}^{p-2} \sin^{p-j-1} t_{ij} \prod_{i=1}^{p-1}\prod_{j=i}^{p-1} dt_{ij} = \sum_\kappa \frac{C_\kappa(\boldsymbol{S})C_\kappa(\boldsymbol{T})}{C_\kappa(\boldsymbol{I})},$$

where $\boldsymbol{\Sigma}_\kappa$ stands for the sum of all possible partition $\kappa = \{k_1, \dots, k_{p-1}\}$ of nonnegative integer k satisfying $k_1 \ge \cdots \ge k_{p-1} \ge 0; C_\kappa(\boldsymbol{X})$ stands for the zonal polynomial corresponding to κ and $c = \pi^{\frac{p^2}{2}}/\Gamma_p(\frac{p}{2})$.

From Lemma 2.1 the joint probability density function of $\boldsymbol{t}_1$ and $(\ell_1, \dots, \ell_p)$ is given by

$$K(\ell_1|\boldsymbol{L}|)^{\frac{n-p-1}{2}} \left(\frac{2^{p-1}\pi^{\frac{(p-1)^2}{2}}}{\Gamma_{p-1}(\frac{p-1}{2})}\right) \exp\left(-\frac{1}{2}\ell_1\boldsymbol{h}_1'\boldsymbol{\Sigma}^{-1}\boldsymbol{h}_1\right) \prod_{j=1}^{p-2} \sin^{p-2} t_{1j}$$
$$\cdot \prod_{i<j}^{p}(\ell_i - \ell_j) \sum_{k=0}^{\infty}\sum_\kappa \frac{C_\kappa(-\frac{1}{2}\boldsymbol{\Sigma}_{p-1}(t_{1.}))C_\kappa(\boldsymbol{L})}{k! C_\kappa(\boldsymbol{I}_{p-1})}, \tag{5}$$

Let $\ell_i = \ell_1 x_i$ for $i = 2, \dots, p$, the joint probability density function of $\boldsymbol{t_1}$, $(x_2, \dots, x_p)$ and ℓ_1 is given by

$$K\left(\frac{\pi^{\frac{(p-1)^2}{2}}}{\Gamma_{p-1}(\frac{p-1}{2})}\right)\exp\left(-\frac{1}{2}\ell_1\boldsymbol{h}_1'\boldsymbol{\Sigma}^{-1}\boldsymbol{h}_1\right)\prod_{j=1}^{p-2}\sin^{p-2}t_{1j}$$
$$\cdot\sum_{k=0}^{\infty}\sum_{\kappa}\ell_1^{\frac{np+2k}{2}-1}|\boldsymbol{L_x}|^{\frac{n-p-1}{2}}\prod_{i=2}^{p}(1-\ell_i)\prod_{i<j=2}^{p}(\ell_i-\ell_j)\cdot\frac{C_\kappa(-\frac{1}{2}\boldsymbol{\Sigma}_{p-1}(t_{1\cdot}))C_\kappa(\boldsymbol{L_x})}{k!C_\kappa(\boldsymbol{I}_{p-1})} \tag{6}$$

where $\boldsymbol{L_x}=\mathrm{diag}(x_2,\ldots,x_p)$. In order to derive the joint probability density function of $\boldsymbol{t}_1$ and ℓ_1, we must integrate (6) with respect to $(x_2,\ldots,x_p)$. The following lemma (see e.g., Muirhead [1982]) is used to integrate.

Lemma 2.2 *Let $\boldsymbol{\Lambda}$ be a diagonal matrix with diagonal elements $1>\lambda_1>\cdots>\lambda_m>0$ .Then,*

$$\int_{1>\lambda_1>\cdots>\lambda_m>0}|\boldsymbol{\Lambda}|^{t-\frac{m+1}{2}}\prod_{i=1}^{m}(1-\lambda_i)^{u-\frac{m+1}{2}}\prod_{i<j}^{m}(\lambda_i-\lambda_j)C_\kappa(\boldsymbol{\Lambda})\prod_{i=1}^{m}d\lambda_i$$
$$=\frac{\Gamma_m(\frac{m}{2})}{\pi^{\frac{m^2}{2}}}\cdot\frac{(t)_\kappa\Gamma_m(t)\Gamma_m(u)}{(t+u)_\kappa\Gamma_m(t+u)}C_\kappa(\boldsymbol{I}_{p-1}).$$

where

$$(a)_\kappa=\prod_{i=1}^{p}(a-\frac{1}{2}(i-1))_{k_i},\ (b)_k=\prod_{j=1}^{k}(b+(j-1)).$$

Integrating (6) with respect to $(x_2,\ldots,x_p)$ the joint probability density function of $\boldsymbol{t}_1$ and ℓ_1 is given by

$$\frac{|\boldsymbol{\Sigma}|^{-\frac{n}{2}}}{2^{\frac{np}{2}}\Gamma_p(\frac{n}{2})}\frac{\Gamma_{p-1}(\frac{p+2}{2})\Gamma_{p-1}(\frac{n-1}{2})}{\Gamma_{p-1}(\frac{n+p+1}{2})}\ell_1^{\frac{np}{2}-1}\exp\left(-\frac{1}{2}\ell_1\boldsymbol{h}_1'\boldsymbol{\Sigma}^{-1}\boldsymbol{h}_1\right)$$
$$\cdot\sum_{k=0}^{\infty}\sum_{\kappa}\frac{(\frac{n-1}{2})_\kappa}{(\frac{n+p+1}{2})_\kappa}\frac{C_\kappa(-\frac{1}{2}\ell_1\boldsymbol{\Sigma}_{p-1}(t_{1\cdot}))}{k!}\prod_{j=1}^{p-2}\sin^{p-2}t_{1j}$$
$$=\frac{|\boldsymbol{\Sigma}|^{-\frac{n}{2}}}{2^{\frac{np}{2}}\Gamma_p(\frac{n}{2})}\frac{\Gamma_{p-1}(\frac{p+2}{2})\Gamma_{p-1}(\frac{n-1}{2})}{\Gamma_{p-1}(\frac{n+p+1}{2})}\ell_1^{\frac{np}{2}-1}\exp\left(-\frac{1}{2}\ell_1\boldsymbol{h}_1'\boldsymbol{\Sigma}^{-1}\boldsymbol{h}_1\right)$$
$$\cdot{}_1F_1(\tfrac{n-1}{2};\tfrac{n+p+1}{2};-\tfrac{1}{2}\ell_1\boldsymbol{\Sigma}_{p-1}(t_{1\cdot}))\prod_{j=1}^{p-2}\sin^{p-2}t_{1j}. \tag{7}$$

where ${}_1F_1(a;c;\boldsymbol{X})$ denotes the hypergeometric function with the matrix argument defined by

$${}_1F_1(a;c;\boldsymbol{X})=\sum_{k=0}^{\infty}\sum_{\kappa}\frac{(a)_\kappa}{(c)_\kappa}\frac{C_\kappa(\boldsymbol{X})}{k!}.$$

These calculations are the same as those of Sugiyama (1966). The positive term series of expression of the joint probability density function is obtained by the Kummer transformation (see e.g., Muirhead [1982]):

$${}_1F_1(a;c;\boldsymbol{X})=\mathrm{etr}(\boldsymbol{X})\cdot{}_1F_1(c-a;c;-\boldsymbol{X}).$$

Applying the Kummer transformation and using the fact that $h_1'\boldsymbol{\Sigma}^{-1}h_1+\mathrm{tr}\boldsymbol{\Sigma}_{p-1}=\mathrm{tr}\boldsymbol{\Sigma}^{-1}$, the joint probability density function in (7) can be expressed by

$$\frac{|\boldsymbol{\Sigma}|^{-\frac{n}{2}}}{2^{\frac{np}{2}}\Gamma_p(\frac{n}{2})}\frac{\Gamma_{p-1}(\frac{p+2}{2})\Gamma_{p-1}(\frac{n-1}{2})}{\Gamma_{p-1}(\frac{n+p+1}{2})}\sum_{k=0}^{\infty}\sum_{\kappa}\frac{(\frac{p+2}{2})_\kappa}{(\frac{n+p+1}{2})_\kappa}\frac{C_\kappa(\frac{1}{2}\boldsymbol{\Sigma}_{p-1}(t_{1.}))}{k!}$$
$$\cdot\ell_1^{\frac{np+2k}{2}-1}\exp\left(-\frac{1}{2}\ell_1\mathrm{tr}\boldsymbol{\Sigma}^{-1}\right)\prod_{j=1}^{p-2}\sin^{p-2}t_{1j}.$$

Integrating with respect to ℓ_1, the joint probability density function of $\boldsymbol{t}_1$ is given by

$$\frac{|\boldsymbol{\Sigma}|^{-\frac{n}{2}}\Gamma(\frac{np}{2})}{\Gamma_p(\frac{n}{2})(\mathrm{tr}\boldsymbol{\Sigma}^{-1})^{\frac{np}{2}}}B_{p-1}\left(\frac{p+2}{2},\frac{n-1}{2}\right)$$
$$\cdot\sum_{k=0}^{\infty}\frac{(\frac{np}{2})_k}{k!}\sum_{\kappa}\frac{(\frac{p+2}{2})_\kappa}{(\frac{n+p+1}{2})_\kappa}C_\kappa\left(\frac{1}{\mathrm{tr}\boldsymbol{\Sigma}^{-1}}\boldsymbol{\Sigma}_{p-1}(t_{1.})\right)\prod_{j=1}^{p-2}\sin^{p-2}t_{1j},\qquad(8)$$

where $B_p(\alpha,\beta)=\Gamma_p(\alpha)\Gamma_p(\beta)/\Gamma_p(\alpha+\beta)$. We must confirm the uniform convergence of the positive series. The following lemma is needed to show that

Lemma 2.3 (Poincaré separation theorem) *Let $\boldsymbol{A}$ be an $p\times p$ symmetric matrix and $\boldsymbol{B}$ be an $p\times h$ matrix satisfying $\boldsymbol{B}'\boldsymbol{B}=\boldsymbol{I}_h$.Then, for $i=1,\ \ldots,\ h$, it follows that:*

$$\lambda_{p-h+i}(\boldsymbol{A})\le\lambda_i(\boldsymbol{B}'\boldsymbol{A}\boldsymbol{B})\le\lambda_i(\boldsymbol{A}),$$

where $\lambda_j(\boldsymbol{A})$ denotes the j-th largest latent root of $\boldsymbol{A}$.

Let $\tilde{\boldsymbol{\Sigma}}=\mathrm{diag}(1/\lambda_p,\ldots,1/\lambda_2)$, where λ_j denotes the j-th largest latent root of $\boldsymbol{\Sigma}$. Using Lemma 2.3,

$$C_\kappa(\boldsymbol{\Sigma}_{p-1}(t_{1.}))\le C_\kappa(\tilde{\boldsymbol{\Sigma}}),$$

for any $\boldsymbol{t}_1$ and thus

$$\sum_{k=0}^{\infty}\frac{(\frac{np}{2})_k}{k!}\sum_{\kappa}\frac{(\frac{p+2}{2})_\kappa}{(\frac{n+p+1}{2})_\kappa}C_\kappa\left(\frac{1}{\mathrm{tr}\boldsymbol{\Sigma}^{-1}}\boldsymbol{\Sigma}_{p-1}(t_{1.})\right)$$
$$\le\sum_{k=0}^{\infty}\frac{(\frac{np}{2})_k}{k!}\sum_{\kappa}\frac{(\frac{p+2}{2})_\kappa}{(\frac{n+p+1}{2})_\kappa}C_\kappa\left(\frac{1}{\mathrm{tr}\boldsymbol{\Sigma}^{-1}}\tilde{\boldsymbol{\Sigma}}\right).\qquad(9)$$

which is bounded by ${}_1F_0(np/2;\mathrm{tr}\tilde{\boldsymbol{\Sigma}}/\mathrm{tr}\boldsymbol{\Sigma}^{-1})$. Noting that $\mathrm{tr}\tilde{\boldsymbol{\Sigma}}/\mathrm{tr}\boldsymbol{\Sigma}^{-1}$ is less than 1, infinite series in (8) is uniformly convergent in the wider sense.

Theorem 2.1 *The joint probability density function of $t_{11},\ldots,t_{1p}$ is given by (8).*

When $p=2$, it can be verified after some calculations that (8) is identical to the probability density function given in Sugiyama (1965), stating that:

$$\frac{1}{\pi(n+1)}\left\{\frac{4|\boldsymbol{\Sigma}|}{(\mathrm{tr}\boldsymbol{\Sigma})^2}\right\}^{\frac{n}{2}}\left\{\frac{n+1}{2}\,{}_2F_1\left(1,n;\frac{n+1}{2};x(t)\right)-\frac{n+3}{2}\,{}_2F_1\left(1,n;\frac{n+3}{2};x(t)\right)\right\},$$

where $x(t)=(\lambda_1\cos^2 t+\lambda_2\sin^2 t)/\mathrm{tr}\boldsymbol{\Sigma}$.

3 Numerical Result

In this section we drew the graphs of the joint probability density function of $\boldsymbol{t}_1$ when $p = 3$. From (8), the joint probability density function of t_{11} and t_{12} is given by

$$\frac{|\boldsymbol{\Sigma}|^{-\frac{n}{2}}\Gamma(\frac{3}{2}n)}{\Gamma_3(\frac{n}{2})(\mathrm{tr}\boldsymbol{\Sigma}^{-1})^{\frac{3}{2}n}} B_2\left(\frac{5}{2}, \frac{n-1}{2}\right) \cdot \sum_{k=0}^{\infty} \frac{(\frac{3}{2}n)_k}{k!} \sum_{\kappa} \frac{(\frac{5}{2})_\kappa}{(\frac{n+4}{2})_\kappa} C_\kappa\left(\frac{1}{\mathrm{tr}\boldsymbol{\Sigma}^{-1}} \boldsymbol{\Sigma}_2(t_{11}, t_{12})\right) \sin t_{11}. \tag{10}$$

From James (1968), it is known that for any 2×2 positive definite matrix $\boldsymbol{A}$,

$$\frac{C_\kappa(\boldsymbol{A})}{C_\kappa(\boldsymbol{I}_2)} = (a_1 a_2)^{\frac{1}{2}} P_{k_1 - k_2}\left(\frac{1}{2}(a_1 + a_2)(a_1 a_2)^{-\frac{1}{2}}\right),$$

where $\boldsymbol{I}_2$ is the 2×2 identity matrix, a_1 and a_2 are the ordered latent roots of $\boldsymbol{A}$, and $P_{k_1-k_2}(u)$ is the Legendre polynomials defined by

$$P_{2q}(u) = (-1)^q (2q)!(2^{2q}(q!)^2)^{-1} \cdot {}_2F_1\left(-q, q + \frac{1}{2}; \frac{1}{2}; u^2\right)$$

for $k_1 - k_2 = 2q$ and

$$P_{2q+1}(u) = (-1)^q (2q+1)!(2^{2q}(q!)^2)^{-1} u \cdot {}_2F_1\left(-q, q + \frac{3}{2}; \frac{3}{2}; u^2\right)$$

for $k_1 - k_2 = 2q + 1$. It is known that

$$C_\kappa(\boldsymbol{I}_2) = 2^{2k} k!(1)_\kappa \frac{\prod_{i<j}^{m}(2k_i - 2k_j - i + j)}{\prod_{i=1}^{m}(2k_i + m - i)!},$$

where m denotes the number of non-zero parts of κ (see e.g.,Muirhead[1982]). We can construct a spectrum decomposition, whereby: $\boldsymbol{\Sigma} = \boldsymbol{\Gamma\Lambda\Gamma}'$, where $\boldsymbol{\Gamma}$ is a 3×3 orthogonal matrix and $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \lambda_2, \lambda_3)$. From (2), we can write

$$\begin{aligned}\boldsymbol{\Gamma} &= \left(\ \boldsymbol{\gamma}_1,\ \boldsymbol{\gamma}_2,\ \boldsymbol{\gamma}_3\ \right) \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_{12} & -\sin\theta_{12} \\ 0 & \sin\theta_{12} & \cos\theta_{12} \end{pmatrix} \begin{pmatrix} \cos\theta_{11} & -\sin\theta_{11} & 0 \\ \sin\theta_{11} & \cos\theta_{11} & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &\qquad \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_{22} & -\sin\theta_{22} \\ 0 & \sin\theta_{22} & \cos\theta_{22} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix},\end{aligned}$$

where $0 \le \theta_{11},\ \theta_{12} \le \pi,\ 0 \le \theta_{22} < 2\pi$. It follows that

$$\boldsymbol{\gamma}_1 = \left(\ \cos\theta_{11},\quad \sin\theta_{11}\cos\theta_{12},\quad \sin\theta_{11}\sin\theta_{12}\ \right)'. \tag{11}$$

We can choose θ_{11}, θ_{12} to characterize γ_1. In this paper, we drew the density function for the following cases (A to D). Multiple root ($\Lambda = \mathrm{diag}(3,2,2)$):

$$\begin{aligned}
&\text{Case A}_1\text{: } \boldsymbol{\gamma}_1 = (0,0,1)', \\
&\text{Case B}_1\text{: } \boldsymbol{\gamma}_1 = \begin{pmatrix} -1/\sqrt{2}, & 0, & 1/\sqrt{2} \end{pmatrix}', \\
&\text{Case C}_1\text{: } \boldsymbol{\gamma}_1 = \begin{pmatrix} 1/\sqrt{2}, & 0, & 1/\sqrt{2} \end{pmatrix}', \\
&\text{Case D}_1\text{: } \boldsymbol{\gamma}_1 = \begin{pmatrix} 1/\sqrt{3}, & 1/\sqrt{3}, & 1/\sqrt{3} \end{pmatrix}';
\end{aligned}$$

Simple root ($\Lambda = \mathrm{diag}(3,2,1)$):

$$\begin{aligned}
&\text{Case A}_2\text{: } \boldsymbol{\gamma}_1 = (0,0,1)', \\
&\text{Case B}_2\text{: } \boldsymbol{\gamma}_1 = \begin{pmatrix} -1/\sqrt{2}, & 0, & 1/\sqrt{2} \end{pmatrix}', \\
&\text{Case C}_2\text{: } \boldsymbol{\gamma}_1 = \begin{pmatrix} 1/\sqrt{2}, & 0, & 1/\sqrt{2} \end{pmatrix}', \\
&\text{Case D}_2\text{: } \boldsymbol{\gamma}_1 = \begin{pmatrix} 1/\sqrt{3}, & 1/\sqrt{3}, & 1/\sqrt{3} \end{pmatrix}';
\end{aligned}$$

Case A_1 represents the transformation of the first element of the original variable to the first principal component. Cases B_1 and C_1 represent the subtraction and sum of the first and third variables, respectively. Case D_1 represents the total variation in the matrix. Cases A_2, B_2, C_2, and D_2 represent the same transformations to principal components and variability as in Cases A_1, B_1, C_1, D_1, respectively, when $\boldsymbol{\Sigma}$ is simple. We drew the graphs when $p = 3$ and $n = 4, 10, 50$.

3.1 Multiple roots

We checked the numerical convergence of the infinite series of (8). Table 1 shows that the sum up to $k = 70$ is sufficient to obtain 3 digits of accuracy for $n = 4, k = 200$ for $n = 10$, and $k = 210$ for $n = 50$.

3.1.1 Case $\mathbf{A}_1$: $\boldsymbol{\gamma}_1 = (0, 0, 1)'$

The probability density function in (8) is shown in Figures 1-3 when $\gamma_1 = (1, 0, 0)$ for $n = 4, 10, 50$. The graphs are symmetric with respect to $(\pi/2, \pi/2)$.

3.1.2 Case $\mathbf{B}_1$: $\boldsymbol{\gamma}_1 = (-1/\sqrt{2},\ 0,\ 1/\sqrt{2})'$

The probability density function in (8) is shown in Figures 4-6 when $(\theta_{11}, \theta_{12}, \theta_{22}) = (3\pi/4, \pi/2, \pi/4)$ for $n = 4, 10, 50$.

3.1.3 Case $\mathbf{C}_1$: $\boldsymbol{\gamma}_1 = (1/\sqrt{2},\ 0,\ 1/\sqrt{2})'$

The probability density function in (8) is shown in Figures 7-9 when $(\theta_{11}, \theta_{12}, \theta_{22}) = (\pi/4, \pi/2, \pi/4)$ for $n = 4, 10, 50$.

3.1.4 Case $\mathbf{D}_1$: $\boldsymbol{\gamma}_1 = (1/\sqrt{3},\ 1/\sqrt{3},\ 1/\sqrt{3})'$

The probability density function in (8) is shown in Figures 10-12 when $(\theta_{11}, \theta_{12}, \theta_{22}) = (\arccos(1/\sqrt{3}), \pi/4, \pi/4)$ for $n = 4, 10, 50$.

Table 1: Right-hand side of (9) for $n = 4, 10, 50, p = 3$ and $\Lambda = \mathrm{diag}(3, 2, 2)$.

	$n = 4$	$n = 10$	$n = 50$
$0 \sim 10$	73.1265	398.2075	3144.8170
$10 \sim 20$	35.3206	638.4612	34827.9060
$20 \sim 30$	5.5627	254.3336	93050.2335
$30 \sim 40$	0.6156	57.4889	124883.2652
$40 \sim 50$	0.0574	9.5579	109915.2269
$50 \sim 60$	0.0048	1.3080	71552.3132
$60 \sim 70$	0.0004	0.1564	36875.6823
$70 \sim 80$	0.0000	0.0169	15730.9644
$80 \sim 90$	0.0000	0.0017	5733.9068
$90 \sim 100$	0.0000	0.0002	1828.7816
$100 \sim 110$	0.0000	0.0000	519.8977
⋮	⋮	⋮	⋮
$190 \sim 200$	0.0000	0.0000	0.0002
$200 \sim 210$	0.0000	0.0000	0.0000
Total	114.6880	1359.5323	498236.7287

· Case A_1 : $\gamma_1 = (0, 0, 1)'$

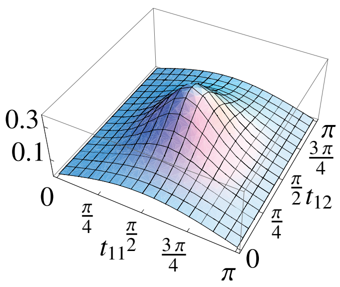

Figure1: $n = 4$

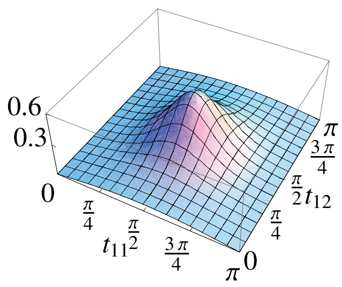

Figure2: $n = 10$

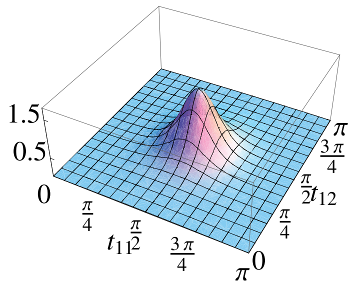

Figure3: $n = 50$

· Case B_1 : $\gamma_1 = (-1/\sqrt{2}, 0, 1/\sqrt{2})'$

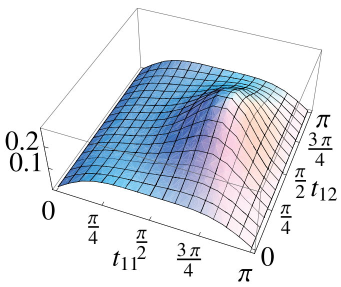

Figure4: $n = 4$

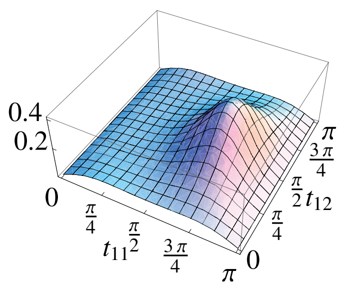

Figure5: $n = 10$

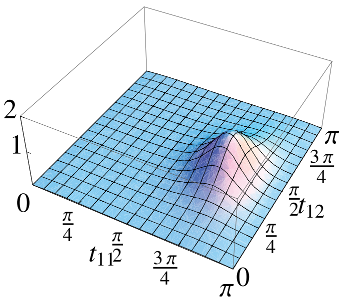

Figure6: $n = 50$

· Case C_1 : $\gamma_1 = (1/\sqrt{2}, 0, 1/\sqrt{2})'$

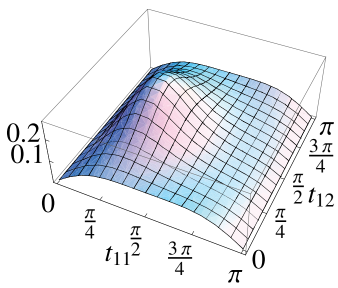

Figure7: $n = 4$

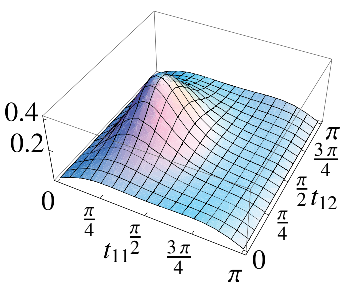

Figure8: $n = 10$

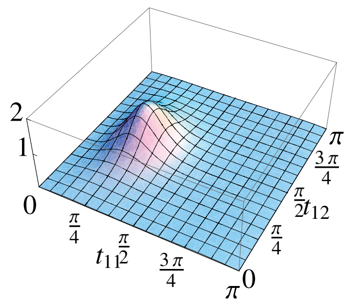

Figure9: $n = 50$

· Case D_1 : $\gamma_1 = (1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})'$

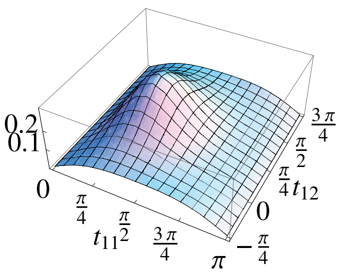

Figure10: $n = 4$

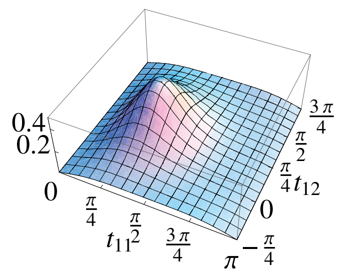

Figure11: $n = 10$

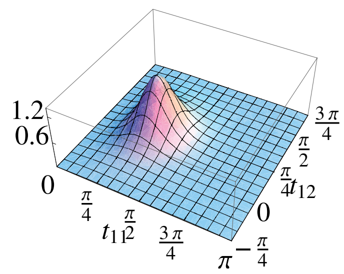

Figure12: $n = 50$

3.2 Simple root

We checked the numerical convergence of the infinite series of (8) when $\Lambda = \mathrm{diag}(3, 2, 1)$. Table 2 shows that the sum up to $k = 110$ is sufficient to obtain 3 digits of accuracy for $n = 4$, $k = 170$ for $n = 10$, and $k = 370$ for $n = 50$.

3.2.1 Case $\mathbf{A}_2$: $\boldsymbol{\gamma}_1 = (0, 0, 1)'$

We studied the case in which Figures 13-15 show the probability density function (8) when $\boldsymbol{\Sigma} = \mathrm{diag}(3, 2, 1)$ for $n = 4, 10, 50$. The graph shows that large amount of probability mass is concentrated on the line $t_{11} = \pi/2$.

3.2.2 Case $\mathbf{B}_2$: $\boldsymbol{\gamma}_1 = (-1/\sqrt{2},\ 0,\ 1/\sqrt{2})'$

The probability density function in (8) is shown in Figures 16-18 when $(\theta_{11}, \theta_{12}, \theta_{22}) = (3\pi/4, \pi/2, \pi/4)$ for $n = 4, 10, 50$. These figures show that the curve of the points of the highly concentrated probability density is not linear. This is the case in all subsequent Figures.

3.2.3 Case $\mathbf{C}_2$: $\boldsymbol{\gamma}_1 = (1/\sqrt{2},\ 0,\ 1/\sqrt{2})'$

The probability density function in (8) is shown in Figures 19-21 when $(\theta_{11}, \theta_{12}, \theta_{22}) = (\pi/4, \pi/2, \pi/4)$ for $n = 4, 10, 50$.

3.2.4 Case $\mathbf{D}_2$: $\boldsymbol{\gamma}_1 = (1/\sqrt{3},\ 1/\sqrt{3},\ 1/\sqrt{3})'$

The probability density function in (8) is shown in Figures 22-24 when $(\theta_{11}, \theta_{12}, \theta_{22}) = (\arccos(1/\sqrt{3}), \pi/4, \pi/4)$ for $n = 4, 10, 50$.

Table 2: Right-hand side of (9) for $n = 4, 10, 50, p = 3$ and $\Lambda = \mathrm{diag}(3, 2, 1)$.

	$n = 4$	$n = 10$	$n = 50$
$0 \sim 10$	133.3410	894.7891	9866.7591
$10 \sim 20$	141.1833	3411.4462	478120.5372
$20 \sim 30$	52.6759	3338.7759	5212000.6417
$30 \sim 40$	13.8721	1820.6691	23858193.0181
$40 \sim 50$	3.0813	725.3823	61828279.8142
$50 \sim 60$	0.6186	237.3524	108276386.5542
$60 \sim 70$	0.1161	67.8066	142851150.8319
$70 \sim 80$	0.0208	17.5365	151898163.2695
$80 \sim 90$	0.0036	4.2029	135975058.7892
$90 \sim 100$	0.0006	0.9484	105579380.5341
$100 \sim 110$	0.0001	0.2038	72670899.3408
$110 \sim 120$	0.0000	0.0421	45083190.6042
⋮	⋮	⋮	⋮
$150 \sim 160$	0.0000	0.0001	2968699.9101
$160 \sim 170$	0.0000	0.0000	1279348.1454
⋮	⋮	⋮	⋮
$350 \sim 360$	0.0000	0.0000	0.0002
$360 \sim 370$	0.0000	0.0000	0.0000
Total	344.9133	10519.1656	904210197.6290

· Case A_2 : $\gamma_1 = (0, 0, 1)'$

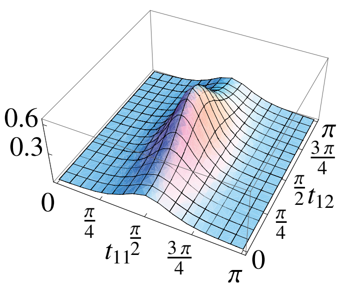

Figure13: $n = 4$

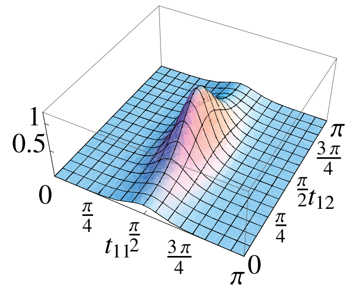

Figure14: $n = 10$

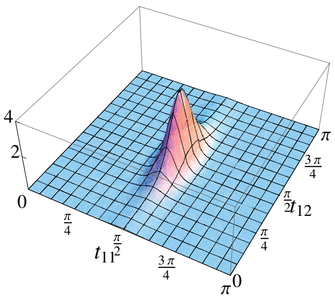

Figure15: $n = 50$

· Case B_2 : $\gamma_1 = (-1/\sqrt{2}, 0, 1/\sqrt{2})'$

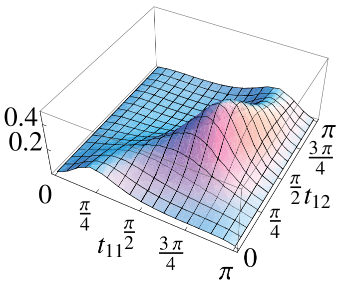

Figure16: $n = 4$

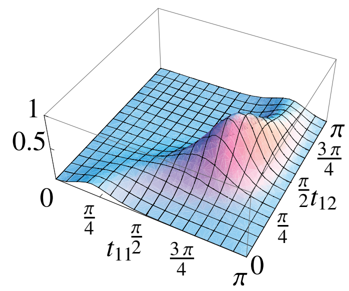

Figure17: $n = 10$

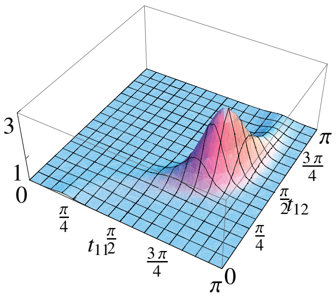

Figure18: $n = 50$

· Case C_2 : $\gamma_1 = (1/\sqrt{2}, 0, 1/\sqrt{2})'$

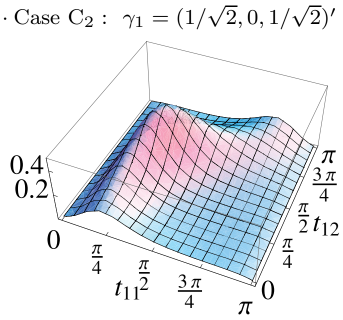

Figure19: $n = 4$

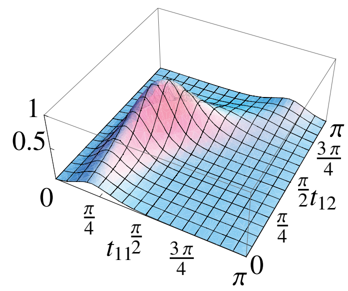

Figure20: $n = 10$

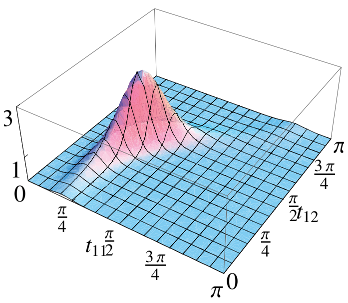

Figure21: $n = 50$

· Case D_2 : $\gamma_1 = (1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})'$

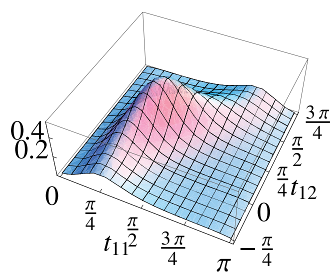

Figure22: $n = 4$

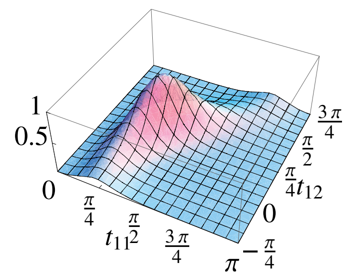

Figure23: $n = 10$

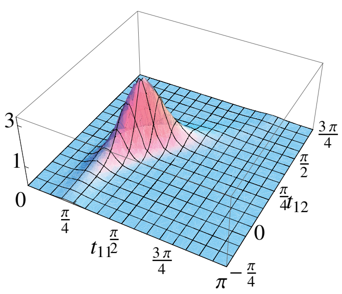

Figure24: $n = 50$

4 Conclusion

In this paper, we have derived an improved expression of the probability density function of the latent vector corresponding to the largest latent root of Wishart matrix given originally in Sugiyama(1966). To illustrate our improvement, we drew the graph of the probability density function of the polar coordinates of the latent vectors when $p = 3$ and confirmed that the concentration of the distribution of the latent vector increases as n increases. As a future study, we will provide the exact confidence region of the vector. This study is currently underway in our laboratory.

Acknowledgment

This work is supported in past by the Institute of Science and Engineering, Chuo University.

References

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

[2] James, A. T. (1968). Calculation of zonal polynomial coefficients by use of the Laplace-Beltrami operator. *Ann.Math.Statist.*, **39**, 1711-1719.

[3] Khatri, C. G. and Pillai, K. C. S. (1969). Distributions of vectors corresponding to the largest roots of three matrices. In *Multivariate Analysis-II*(P. R. Krishnaiah, Ed.), Academic Press, New York.

[4] Muirhead, R. J. (1982). Aspects of Multivariate Statistical Analysis. Wiley, New York.

[5] Sugiura, N. (1976). Asymptotic expansions of the distributions of the latent roots and the latent vector of the Wishart and multivariate F matrices. *J. Multivariate Anal.*, **6**, no. 4, 500-525.

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

[7] 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.

[8] Takemura, Akimichi; Sheena, Yo. (2005). Distribution of eigenvalues and eigenvectors of Wishart matrix when the population eigenvalues are infinitely dispersed and its application to minimax estimation of covariance matrix. *J. Multivariate Anal.*, **94**, no. 2, 271-299.