J. Jpn. Soc. Comp. Statist., **30.2**(2018), 1–12 DOI:10.5183/jjscs.1708001_244

DISTRIBUTION OF THE LARGEST EIGENVALUE OF AN ELLIPTICAL WISHART MATRIX AND ITS SIMULATION

Aya Shinozaki*, Hiroki Hashiguchi† and Toshiya Iwashita‡

ABSTRACT

This paper provides an alternative proof of the derivation of the distribution of the largest eigenvalue of an elliptical Wishart matrix in contrast to the result of Caro-Lopera et al. (2016). We show the relation between multivariate and matrix-variate t distributions. From this relation, we can generate random numbers drawn from the matrix-variate t distribution. A Monte Carlo simulation is conducted to evaluate the accuracy for the truncated distribution function of the largest eigenvalue of the elliptical Wishart matrix. Exact computation of the distribution of the smallest eigenvalue is also presented.

1. Introduction

The study of the eigenvalue distribution under normal population is a fundamentally important component of multivariate analysis. This study originated from Wishart (1928), James (1960, 1964), and Constantine (1963). It is summarized in standard textbook such as Muirhead (1982). Furthermore, it is developed as a matrix-variate distribution, which is described in Gupta and Nagar (1999). Recently, discussion of the exact distribution theory under the elliptical model is specifically examined, for example, Caro-Lopera et al. (2010) and Caro-Lopera et al. (2014a, 2016). Caro-Lopera et al. (2010) introduced a family of matrix-variate elliptically contoured distributions including a matrix-variate normal distribution, Pearson type VII configuration distribution, Kotz type configuration distribution, Bessel configuration distribution, and Jensen-logistic configuration distribution. As mentioned by Sutradhar and Ali (1989), the elliptical models are generalization of the multivariate normal distribution, accordingly the models have robust statistical characteristics. Caro-Lopera et al. (2014a) presented the generalized Wishart distribution and discussed some properties of covariance matrices. They first derived a class of generalized Wishart distributions under the elliptical model, which includes the classical Wishart distribution as a special case. As an application of this result, they obtained the distribution of the eigenvalues of an elliptical Wishart matrix; then they deduced the corresponding known results for the ordinal Wishart case. Caro-Lopera et al. (2014b) also deduced the sphericity test on a trace-type matrix-variate elliptical distribution, and derived an exact distribution of the statistics. Caro-Lopera et al. (2016) provided the density of an elliptical Wishart

*Department of Applied Mathematics, Graduate School of Science, Tokyo University of Science, 1-3, Kagurazaka, Shinjuku, Tokyo 162-8601, Japan E-mail: 1413704@ed.tus.ac.jp

†Faculty of Science Division 1, Tokyo University of Science, 1-3, Kagurazaka, Shinjuku, Tokyo 162-8601, Japan

‡Faculty of Science and Technology, Tokyo University of Science, 2641, Yamazaki, Noda, Chiba 278-8510, Japan

Key words: Eigenvalue distribution; Elliptically contoured distribution; Hypergeometric function of a matrix argument; Matrix-variate t distribution; Percentile point

distribution, the joint density of the eigenvalues, and the largest and smallest eigenvalues of an elliptical Wishart matrix. However, numerical computation of exact distributions under an elliptical model is not carried out because it should calculate an infinite series with zonal polynomials. This is another reason why a random generation procedure is not established in order to evaluate accuracy by Monte Carlo simulation.

This paper provides an alternative proof of the derivation of the largest-eigenvalue distribution of an elliptical Wishart matrix in contrast to the result of Caro-Lopera et al. (2016). Our derivation is based on the method of Sugiyama (1967). On the other hand, their result followed the method of Constantine (1963) and Theorem 9.7.1 of Muirhead (1982). We also discuss the largest-eigenvalue distribution under a matrix-variate t distribution, which is a special case of the elliptical case. Section 2 presents a summary of the theory of eigenvalues of an elliptical Wishart matrix as given by Caro-Lopera et al. (2014a, 2016). In Section 3, we establish a relation between the matrix-variate and multivariate t distributions. This relation can be used to generate random numbers drawn from the matrix-variate t distribution because the random generation for the multivariate t distribution is known. In Section 4, under the matrix-variate t distribution, we compare the empirical and the approximate distributions of the largest eigenvalues. The empirical distribution is generated by the relation of the multivariate t distribution described in Section 3 and the approximate distribution is a truncated distribution with a finite series of zonal polynomials. Calculation of zonal polynomials is based on the algorithm of Hashiguchi et al. (2000). Finally, conclusion and future work are described in Section 5.

2. The distribution of the largest eigenvalue of an elliptical Wishart matrix

We introduce the vec operator $\mathrm{vec}(\boldsymbol{X})$ for an $m \times n$ matrix $\boldsymbol{X} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$ and the Kronecker product $\boldsymbol{A} \otimes \boldsymbol{B}$ for $p \times q$ matrix $\boldsymbol{A} = (a_{ij})$ and $r \times s$ matrix $\boldsymbol{B}$. They are defined respectively as

$$\mathrm{vec}(\boldsymbol{X}) = \begin{pmatrix} \boldsymbol{x}_1 \\ \vdots \\ \boldsymbol{x}_n \end{pmatrix} \text{ and } \boldsymbol{A} \otimes \boldsymbol{B} = \begin{pmatrix} a_{11}B & \cdots & a_{1q}B \\ \vdots & \ddots & \vdots \\ a_{p1}B & \cdots & a_{pq}B \end{pmatrix}.$$

An $n \times m$ random matrix $\boldsymbol{X}$ is said to have a matrix-variate elliptically contoured distribution, denoted by $\mathcal{E}_{n\times m}(\boldsymbol{M}, \boldsymbol{\Omega} \otimes \boldsymbol{\Sigma}, h)$, if its density function is given as

$$g_{\boldsymbol{X}}(\boldsymbol{X}) = \frac{1}{|\boldsymbol{\Sigma}|^{n/2}|\boldsymbol{\Omega}|^{m/2}} h\left(\mathrm{tr}\boldsymbol{\Omega}^{-1}(\boldsymbol{X}-\boldsymbol{M})\boldsymbol{\Sigma}^{-1}(\boldsymbol{X}-\boldsymbol{M})^\top\right), \tag{1}$$

where $\boldsymbol{\Sigma} : m \times m$, $\boldsymbol{\Omega} : n \times n$, $\boldsymbol{\Sigma} > 0$, $\boldsymbol{\Omega} > 0$, and the generator function $h : \mathbb{R} \to [0, \infty)$, satisfies $h(u) \in C^\infty$. The elliptical Wishart matrix $\boldsymbol{A}$ is defined as $\boldsymbol{A} = \boldsymbol{X}^\top \boldsymbol{X}$ if $\boldsymbol{X}$ is distributed as $\mathcal{E}_{n\times m}(\boldsymbol{O}, \boldsymbol{I}_n \otimes \boldsymbol{\Sigma}, h)$. That is, $\boldsymbol{M} = \boldsymbol{O}$, $\boldsymbol{\Omega} = \boldsymbol{I}_n$ and $n \geq m$. The elliptical Wishart distribution is written as $\mathcal{EW}_m(n, \boldsymbol{\Sigma}, h)$. Its density function of $\boldsymbol{A}$ is given as

$$\frac{\pi^{mn/2}}{\Gamma_m\left(\frac{1}{2}n\right)\ |\boldsymbol{\Sigma}|^{n/2}} |\boldsymbol{A}|^{(n-m-1)/2}\ h(\mathrm{tr}\boldsymbol{\Sigma}^{-1}\boldsymbol{A}), \tag{2}$$

where $\Gamma_m(a) = \pi^{m(m-1)/4} \prod_{i=1}^m \Gamma(a - (i-1)/2)$, with $\Re(a) > (m-1)/2$ and $\Gamma(a)$ is the ordinary gamma function. The distribution function of $\boldsymbol{A}$ is given as

$$\begin{aligned}\Pr(\boldsymbol{A} < \boldsymbol{\Omega}) = &\frac{\pi^{mn/2}\Gamma_m((m+1)/2)}{\Gamma_m((n+m+1)/2)}|\boldsymbol{\Sigma}^{-1}\boldsymbol{\Omega}|^{n/2} \\ &\times \sum_{k=0}^{\infty} \frac{h^{(k)}(0)}{k!} \sum_{\kappa} \frac{(n/2)_\kappa}{((n+m+1)/2)_\kappa} \mathcal{C}_\kappa(\boldsymbol{\Sigma}^{-1}\boldsymbol{\Omega}),\end{aligned} \tag{3}$$

where $\boldsymbol{A} < \boldsymbol{\Omega}$ represents that $\boldsymbol{\Omega} - \boldsymbol{A}$ is positive definite, $h^{(k)}(a) = d^k h(u)/du^k|_{u=a}$. In (3), $\mathcal{C}_\kappa(\boldsymbol{X})$ is the zonal polynomial with a symmetric matrix $\boldsymbol{X}$ indexed by a partition of k into not more than m parts, $\kappa = (\kappa_1, \ldots, \kappa_m)$ where $\kappa_1 \geq \cdots \geq \kappa_m \geq 0$ and $\sum_{i=1}^m \kappa_i = k$. The sum $\sum_\kappa$ in (3) runs over all partitions of k. The Pochhammer symbol for a partition $\kappa = (\kappa_1, \ldots, \kappa_m)$ is defined as $(\alpha)_\kappa = \prod_{i=1}^m (\alpha - (i-1)/2)_{\kappa_i}$, where $(\alpha)_t = \alpha(\alpha+1)\cdots(\alpha+t-1)$ and $(\alpha)_0 = 1$. James (1964) presented the definition and properties of the zonal polynomial. The equations (2) and (3) were given by Caro-Lopera et al. (2014) and Caro-Lopera et al. (2016), respectively. We note that the distribution $\mathcal{E}_{n\times m}(\boldsymbol{M}, \boldsymbol{\Omega} \otimes \boldsymbol{\Sigma}, h)$ includes the matrix-variate normal distribution $\mathcal{N}_{n\times m}(\boldsymbol{M}, \boldsymbol{\Omega} \otimes \boldsymbol{\Sigma})$, where $n \geq m$, if $h(y) = \exp(-y/2)/(2\pi)^{mn/2}$. Thereby, if $\boldsymbol{X} \sim \mathcal{N}_{n\times m}(\boldsymbol{O}, \boldsymbol{I}_n \otimes \boldsymbol{\Sigma})$, namely, row vectors of $\boldsymbol{X}$ are mutually independent and identically distributed as $\mathcal{N}_m(\boldsymbol{0}, \boldsymbol{\Sigma})$, then $\boldsymbol{A} = \boldsymbol{X}^\top \boldsymbol{X}$ is the classical Wishart matrix.

Let $\boldsymbol{A} \sim \mathcal{EW}_m(n, \boldsymbol{\Sigma}, h)$. We consider the distribution of the eigenvalues of $\boldsymbol{A}$ denoted by $\ell_1, \ldots, \ell_m$ where $\ell_1 > \cdots > \ell_m > 0$. Their joint density function is given by Caro-Lopera et al. (2016) as

$$\frac{\pi^{m(n+m)/2}}{\Gamma_m(n/2)|\boldsymbol{\Sigma}|^{n/2}\Gamma_m(m/2)}|\boldsymbol{L}|^{(n-m-1)/2}\prod_{i<j}^m(\ell_i - \ell_j)\sum_{k=0}^\infty \frac{h^{(k)}(0)}{k!}\sum_\kappa \frac{\mathcal{C}_\kappa(\boldsymbol{\Sigma}^{-1})\mathcal{C}_\kappa(\boldsymbol{L})}{\mathcal{C}_\kappa(\boldsymbol{I}_m)}, \tag{4}$$

where $\boldsymbol{L} = \mathrm{diag}(\ell_1, \ldots, \ell_m)$. They also derived the distributions of the largest and smallest eigenvalues using (4).

We also give the distribution function of ℓ_1 of $\boldsymbol{A}$ by the procedures of Sugiyama (1967) and Khatri (1967). The key procedure of Khatri (1967) is the following multiple integral:

$$\begin{aligned}&\int_{1>\ell_1>\ell_2>\cdots>\ell_m>0} |\boldsymbol{L}|^{t-(m+1)/2}|\boldsymbol{I}_m - \boldsymbol{L}|^{u-(m+1)/2}\mathcal{C}_\kappa(\boldsymbol{L})\prod_{i<j}^m(\ell_i - \ell_j)\prod_{i=1}^m d\ell_i \\ &= \frac{\Gamma_m(m/2)\Gamma_m(t,\kappa)\Gamma_m(u)\mathcal{C}_\kappa(\boldsymbol{I}_m)}{\pi^{m^2/2}\Gamma_m(t+u,\kappa)},\end{aligned} \tag{5}$$

where $\Re(t) > (m-1)/2$, $\Re(u) > (m-1)/2$, $\Gamma_m(\alpha,\kappa) = (\alpha)_\kappa\Gamma_m(\alpha)$. Let $\boldsymbol{X}_1 = \mathrm{diag}(1, x_2, \ldots, x_m)$, $\boldsymbol{X}_2 = \mathrm{diag}(x_2, \ldots, x_m)$, where $x_2 > \cdots > x_m > 0$. Using (5), the other procedure of Sugiyama (1967) is given as

$$\begin{aligned}&\int_{1>x_2>\cdots>x_m>0} |\boldsymbol{X}_2|^{t-(m+1)/2}\mathcal{C}_\kappa(\boldsymbol{X}_1)\prod_{i=2}^m(1-x_i)\prod_{i<j}^m(x_i - x_j)\prod_{i=2}^m dx_i \\ &=(mt+k)(\Gamma_m(m/2)/\pi^{m^2/2})\frac{\Gamma_m(t,\kappa)\Gamma_m((m+1)/2)}{\Gamma_m(t+(m+1)/2,\kappa)}\mathcal{C}_\kappa(\boldsymbol{I}_m),\end{aligned} \tag{6}$$

which yields the following theorem.

Theorem 1 *Let $\boldsymbol{A} \sim \mathcal{EW}_m(n, \boldsymbol{\Sigma}, h)$. Then the distribution function of the largest eigenvalue ℓ_1 of $\boldsymbol{A}$ is given as*

$$\Pr(\ell_1 < x) = \frac{\pi^{mn/2}\Gamma_m((m+1)/2)|x\boldsymbol{\Sigma}^{-1}|^{n/2}}{\Gamma_m((n+m+1)/2)} \sum_{k=0}^{\infty} \frac{h^{(k)}(0)}{k!} \sum_{\kappa} \frac{(n/2)_\kappa}{((n+m+1)/2)_\kappa} \mathcal{C}_\kappa(x\boldsymbol{\Sigma}^{-1}). \quad (7)$$

Proof. The translation of $x_i = \ell_i/\ell_1$ for $i = 2, \ldots, m$ and the integration of $x_2, \ldots, x_m$ in (4) give the density function of ℓ_1 as

$$\begin{aligned}
f(\ell_1) &= \frac{\pi^{m(n+m)/2}}{\Gamma_m(n/2)|\boldsymbol{\Sigma}|^{n/2}\Gamma_m(m/2)} \ell_1^{mn/2+k-1} \int_{1>x_2>\cdots>x_m>0} |\boldsymbol{X}_2|^{(n-m-1)/2} \\
&\quad \times \prod_{i=2}^{m}(1-x_i) \prod_{2\le i<j}^{m} (x_i - x_j)\mathcal{C}_\kappa(\boldsymbol{X}_2) \sum_{k=0}^{\infty} \frac{h^{(k)}(0)}{k!} \sum_{\kappa} \frac{\mathcal{C}_\kappa(\boldsymbol{\Sigma}^{-1})}{\mathcal{C}_\kappa(\boldsymbol{I}_m)} \prod_{i=2}^{m} dx_i \\
&= \frac{\pi^{m(n+m)/2}}{\Gamma_m(n/2)|\boldsymbol{\Sigma}|^{n/2}\Gamma_m(m/2)} \ell_1^{mn/2+k-1} \frac{(mn/2+k)\Gamma_m(m/2)}{\pi^{m^2/2}} \\
&\quad \times \frac{\Gamma_m(n/2)(n/2)_\kappa\Gamma_m((m+1)/2)\mathcal{C}_\kappa(\boldsymbol{I}_m)}{\Gamma_m((n+m+1)/2)((n+m+1)/2)_\kappa} \sum_{k=0}^{\infty} \frac{h^{(k)}(0)}{k!} \sum_{\kappa} \frac{\mathcal{C}_\kappa(\boldsymbol{\Sigma}^{-1})}{\mathcal{C}_\kappa(\boldsymbol{I}_m)} \\
&= \frac{\pi^{mn/2}\Gamma_m(\frac{m+1}{2})}{\Gamma_m(\frac{n+m+1}{2})} |\boldsymbol{\Sigma}^{-1}|^{n/2} \\
&\quad \times \sum_{k=0}^{\infty} (mn/2+k)\ell_1^{mn/2+k-1} \frac{h^{(k)}(0)}{k!} \sum_{\kappa} \frac{(n/2)_\kappa}{((n+m+1)/2)_\kappa} \mathcal{C}_\kappa(\boldsymbol{\Sigma}^{-1}).
\end{aligned}$$

Moreover, integrating it with respect to ℓ_1, we have the distribution function of ℓ_1 as

$$\begin{aligned}
\Pr(\ell_1 < x) &= \int_0^x f(\ell_1)d\ell_1 \\
&= \frac{\pi^{mn/2}\Gamma_m(\frac{m+1}{2})}{\Gamma_m(\frac{n+m+1}{2})} |x\boldsymbol{\Sigma}^{-1}|^{n/2} \sum_{k=0}^{\infty} \frac{h^{(k)}(0)}{k!} \sum_{\kappa} \frac{(n/2)_\kappa}{((n+m+1)/2)_\kappa} \mathcal{C}_\kappa(x\boldsymbol{\Sigma}^{-1}).
\end{aligned}$$

□

In fact, the probability of $\ell_1 < x$ is equivalent to the one of $\boldsymbol{A} < x\boldsymbol{I}_m$. Therefore, the result (7) in Theorem 1 follows by putting $\boldsymbol{\Omega} = x\boldsymbol{I}_m$ in (3), as provided in Caro-Lopera et al. (2016).

The hypergeometric generalized function ${}_1P_1$ of a matrix argument $\boldsymbol{X}$ is defined as

$${}_1P_1\left(h^{(k)}(0) : a; c; \boldsymbol{X}\right) = \sum_{k=0}^{\infty} \frac{h^{(k)}(0)}{k!} \sum_{\kappa} \frac{(a)_\kappa}{(c)_\kappa} \mathcal{C}_\kappa(\boldsymbol{X}), \quad (8)$$

where the function $h^{(k)}(0)$ is independent of κ. This function ${}_1P_1$ was defined by Díaz-García and Caro-Lopera (2008). We have ${}_1P_1(1 : a; c; \boldsymbol{X}) = {}_1F_1(a; c; \boldsymbol{X})$ where ${}_1F_1(a; c; \boldsymbol{X})$ is the confluent hypergeometric function of a matrix argument defined by

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

Analogous to the Kummer relation to ${}_1F_1$ of Theorem 7.4.3 presented by Muirhead (1982), the generalized Kummer relation is given by Díaz-García and Caro-Lopera (2008) as

$$ {}_1P_1\left(h^{(k)}(0) : a; c; \boldsymbol{X}\right) = {}_1P_1\left(h^{(k)}(\mathrm{tr}\boldsymbol{X}) : c-a; c; -\boldsymbol{X}\right). \tag{9} $$

Corollary 1 *Using (8), the equation (7) can be written as*

$$ \Pr(\ell_1 < x) = \frac{\pi^{\frac{mn}{2}}\Gamma_m((m+1)/2)|x\boldsymbol{\Sigma}^{-1}|^{n/2}}{\Gamma_m((n+m+1)/2)}\,{}_1P_1\left(h^{(k)}(0) : \frac{n}{2}; \frac{n+m+1}{2}; x\boldsymbol{\Sigma}^{-1}\right) \tag{10} $$

and, for $h(y) = \exp(-y/2)/(2\pi)^{mn/2}$ *and* $h^{(k)}(0) = (-1/2)^k/(2\pi)^{mn/2}$*, as*

$$ \Pr(\ell_1 < x) = \frac{\Gamma_m((m+1)/2)|x\boldsymbol{\Sigma}^{-1}/2|^{n/2}}{\Gamma_m((n+m+1)/2)}\,{}_1F_1\left(\frac{n}{2}; \frac{n+m+1}{2}; -\frac{x}{2}\boldsymbol{\Sigma}^{-1}\right). \tag{11} $$

Applying the generalized Kummer relation (9) to (10), (7) is also written as

$$ \Pr(\ell_1 < x) = \frac{\pi^{\frac{mn}{2}}\Gamma_m((m+1)/2)|x\boldsymbol{\Sigma}^{-1}|^{n/2}}{\Gamma_m((n+m+1)/2)}\,{}_1P_1\left(h^{(k)}(\mathrm{tr}x\boldsymbol{\Sigma}^{-1}) : \frac{m+1}{2}; \frac{n+m+1}{2}; -x\boldsymbol{\Sigma}^{-1}\right) $$

as well as, for $h(y) = \exp(-y/2)/(2\pi)^{mn/2}$ *and* $h^{(k)}(0) = (-1/2)^k/(2\pi)^{mn/2}$*,*

$$ \Pr(\ell_1 < x) = \frac{\Gamma_m((m+1)/2)|x\boldsymbol{\Sigma}^{-1}/2|^{n/2}}{\Gamma_m((n+m+1)/2)}\mathrm{etr}(-x\boldsymbol{\Sigma}^{-1}/2)\,{}_1F_1\left(\frac{m+1}{2}; \frac{n+m+1}{2}; \frac{x}{2}\boldsymbol{\Sigma}^{-1}\right). $$

3. Relation between multivariate and matrix-variate t distributions

This section presents consideration of the matrix-variate t distribution with covariance matrix $\boldsymbol{\Sigma}$ and ρ degrees of freedom, as denoted by $T_{n\times m}(\rho, \boldsymbol{\Sigma})$. Its generator function and k-th derivative are given as

$$ h(y) = \frac{\Gamma\left(\frac{mn+\rho}{2}\right)}{(\pi\rho)^{mn/2}\Gamma\left(\frac{\rho}{2}\right)}\left(1+\frac{y}{\rho}\right)^{-(mn+\rho)/2}, \tag{12} $$

$$ h^{(k)}(y) = \frac{\Gamma((mn+\rho)/2)(-1)^k((mn+\rho)/2)_k}{(\pi\rho)^{mn/2}\Gamma(\rho/2)\rho^k}(1+y/\rho)^{-((mn+\rho)/2+k)}, \tag{13} $$

respectively. In fact, the matrix-variate t distribution has definitions of two kinds. The ordinary definition has a density function given as

$$ \frac{\Gamma_m(\frac{\rho+n+m-1}{2})}{\pi^{mn/2}\Gamma_m(\frac{\rho+m-1}{2})|\boldsymbol{\Sigma}|^{n/2}|\boldsymbol{\Theta}|^{m/2}}|\boldsymbol{I}_m + \boldsymbol{\Sigma}^{-1}(\boldsymbol{Y}-\boldsymbol{M})^\top\boldsymbol{\Theta}^{-1}(\boldsymbol{Y}-\boldsymbol{M})|^{-(m+n+\rho-1)/2}, $$

where $\boldsymbol{Y} : n\times m$, $\boldsymbol{M} : n\times m$, $\boldsymbol{\Theta}(n\times n) > 0$, $\boldsymbol{\Sigma}(m\times m) > 0$ and $\rho > 0$. This distribution is similar to a matrix-variate beta distribution. Chapter 4 of Gupta and Nagar (1999) presents related details. The other definition is given, using (12), as

$$ g_{\boldsymbol{X}}(\boldsymbol{X}) = \frac{\Gamma((mn+\rho)/2)}{(\pi\rho)^{mn/2}\Gamma(\rho/2)|\boldsymbol{\Sigma}|^{n/2}}\left(1+\mathrm{tr}(\boldsymbol{X}\boldsymbol{\Sigma}^{-1}\boldsymbol{X}^\top)/\rho\right)^{-(mn+\rho)/2}. \tag{14} $$

The common point of the two kinds of matrix-variate t distributions shares a multivariate t distribution when $n = 1$. In this case of $n = 1$, a random vector $\boldsymbol{x} = \boldsymbol{X}^\top$ where $\boldsymbol{X} \sim$

$T_{1\times m}(\rho, \boldsymbol{\Sigma})$ is said to be distributed as a multivariate t distribution. Its density function is given as

$$g(\boldsymbol{x}) = \frac{\Gamma((m+\rho)/2)}{(\pi\rho)^{m/2}\Gamma(\rho/2)|\boldsymbol{\Sigma}|^{1/2}} \left(1 + \frac{1}{\rho}\boldsymbol{x}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{x}\right)^{-(m+\rho)/2}.$$

Then the relation between multivariate and matrix-variate t distributions is satisfied in the following theorem.

Theorem 2 *The statement* $\boldsymbol{X} \sim T_{n\times m}(\rho, \boldsymbol{\Sigma})$ *is equivalent to* $\mathrm{vec}(\boldsymbol{X}^\top) \sim T_{1\times nm}(\rho, \boldsymbol{I}_n \otimes \boldsymbol{\Sigma})$.

Proof. From Lemma 2.2.3 of Muirhead (1982), we have

$$\mathrm{tr}(\boldsymbol{X}\boldsymbol{\Sigma}^{-1}\boldsymbol{X}^\top) = (\mathrm{vec}(\boldsymbol{X}^\top))^\top(\boldsymbol{I}_m \otimes \boldsymbol{\Sigma})^{-1}(\mathrm{vec}(\boldsymbol{X})) = \boldsymbol{x}^\top(\boldsymbol{I}_m \otimes \boldsymbol{\Sigma})^{-1}\boldsymbol{x}$$

for $\boldsymbol{x} = \mathrm{vec}(\boldsymbol{X}^\top)$ and $|\boldsymbol{\Sigma}|^{n/2} = |\boldsymbol{I}_n \otimes \boldsymbol{\Sigma}|^{1/2}$. Therefore, the density function of $\boldsymbol{x}$ is given as

$$\frac{\Gamma((mn+\rho)/2)}{(\pi\rho)^{mn/2}\Gamma(\rho/2)|\boldsymbol{I}_n \otimes \boldsymbol{\Sigma}|^{1/2}} \left(1 + \frac{1}{\rho}\boldsymbol{x}^\top \left(\boldsymbol{I}_n \otimes \boldsymbol{\Sigma}\right)^{-1} \boldsymbol{x}\right)^{-(mn+\rho)/2} \tag{15}$$

which yields that $\boldsymbol{x} = \mathrm{vec}(\boldsymbol{X}^\top) \sim T_{1\times nm}(\rho, \boldsymbol{I}_n \otimes \boldsymbol{\Sigma})$. The converse condition is clearly obtained from the proof above. □

Theorem 3 *Assume that* $\boldsymbol{Z} \sim N_{n\times m}(\boldsymbol{O}, \boldsymbol{I}_n \otimes \boldsymbol{I}_m)$ *and* $R \sim \chi^2_\rho$ *where* $\boldsymbol{Z}$ *and* R *are independent. Then the random matrix*

$$\boldsymbol{X} = \frac{\rho^{1/2}\boldsymbol{Z}\boldsymbol{\Sigma}^{1/2}}{R^{1/2}} \tag{16}$$

is distributed as $\boldsymbol{X} \sim T_{n\times m}(\rho, \boldsymbol{\Sigma})$.

Proof. We put $\boldsymbol{z} = \mathrm{vec}(\boldsymbol{Z}^\top)$ and

$$\boldsymbol{x} = \mathrm{vec}(\boldsymbol{X}^\top) = \rho^{1/2}(\boldsymbol{I}_n \otimes \boldsymbol{\Sigma}^{1/2})\mathrm{vec}(\boldsymbol{Z}^\top)/R^{1/2} = \left(\frac{\rho}{R}\right)^{1/2}(\boldsymbol{I}_n \otimes \boldsymbol{\Sigma}^{1/2})\boldsymbol{z}.$$

The joint density function of $\boldsymbol{z} = \mathrm{vec}(\boldsymbol{Z}^\top)$ and R is given as

$$f(\boldsymbol{z}, R) = \frac{1}{(2\pi)^{mn/2}} \exp\left(-\frac{1}{2}\boldsymbol{z}^\top\boldsymbol{z}\right) \frac{1}{\Gamma\left(\frac{\rho}{2}\right)2^{\rho/2}} R^{\rho/2-1}e^{-R/2}. \tag{17}$$

Substituting $\boldsymbol{z} = (R/\rho)^{1/2}(\boldsymbol{I}_n \otimes \boldsymbol{\Sigma}^{1/2})^{-1}\boldsymbol{x}$ and $d\boldsymbol{z} = (R/\rho)^{mn/2}|\boldsymbol{\Sigma}|^{-n/2}d\boldsymbol{x}$ for (17), we have

$$\begin{aligned}
f(\boldsymbol{x})d\boldsymbol{x} &= \int_{R>0} f((R/\rho)^{1/2}(\boldsymbol{I}_n \otimes \boldsymbol{\Sigma}^{1/2})^{-1}\boldsymbol{x}, R)(R/\rho)^{mn/2}|\boldsymbol{\Sigma}|^{-n/2}d\boldsymbol{x}dR \\
&= \frac{\rho^{-mn/2}|\boldsymbol{\Sigma}|^{-n/2}}{(2\pi)^{mn/2}\Gamma\left(\frac{\rho}{2}\right)2^{\rho/2}} \int_0^\infty R^{(mn+\rho)/2-1} \exp\left\{-\frac{R}{2}\left(1 + \frac{1}{\rho}\boldsymbol{x}^\top(\boldsymbol{I} \otimes \boldsymbol{\Sigma})^{-1}\boldsymbol{x}\right)\right\} dR\, d\boldsymbol{x} \\
&= \frac{\Gamma((mn+\rho)/2)}{(\pi\rho)^{mn/2}\Gamma(\rho/2)|\boldsymbol{\Sigma}|^{n/2}} \left(1 + \frac{1}{\rho}\boldsymbol{x}^\top(\boldsymbol{I} \otimes \boldsymbol{\Sigma})^{-1}\boldsymbol{x}\right)^{-(mn+\rho)/2} d\boldsymbol{x},
\end{aligned}$$

where $d\boldsymbol{x} = dx_1 \cdots dx_{mn}$ for $\boldsymbol{x} = (x_1, \ldots, x_{mn})^\top$. Therefore, it is readily apparent that $\boldsymbol{x} \sim T_{1\times mn}(\rho, \boldsymbol{\Sigma})$. From Theorem 2, we have $\boldsymbol{X} \sim T_{n\times m}(\rho, \boldsymbol{\Sigma})$. □

From Theorem 3 we are able to generate random numbers of $\boldsymbol{X} \sim T_{n\times m}(\rho, \boldsymbol{\Sigma})$ and study numerical simulations in Section 4. We note that a similar result of Theorem 3 can be found in the Problem 4.6. of Gupta and Nagar (1999).

4. Numerical experiments

In this section, we describe calculation of the distributions of the largest and smallest eigenvalues of the elliptical Wishart matrix under the matrix-variable t distribution. A Monte Carlo simulation study under a matrix-variate t distribution can be carried out from Theorem 3. Throughout this section, the empirical distribution based on 10^6 times Monte Carlo simulation is represented by F_{sim}.

The distribution function of ℓ_1 of $\boldsymbol{A} = \boldsymbol{X}^\top \boldsymbol{X}$ where $\boldsymbol{X} \sim T_{n\times m}(\rho, \boldsymbol{\Sigma})$ is given as

$$\begin{aligned}&\Pr(\ell_1 < x)\\ &= \frac{\Gamma_m(\frac{m+1}{2})\,\Gamma(\frac{mn+\rho}{2})}{\Gamma_m(\frac{n+m+1}{2})\,\Gamma(\frac{\rho}{2})}|x\boldsymbol{\Sigma}^{-1}|^{n/2}\sum_{k=0}^{\infty}\frac{(-1)^k(\frac{nm+\rho}{2})_k}{\rho^{mn/2+k}}\sum_{\kappa}\frac{(\frac{n}{2})_\kappa}{(\frac{n+m+1}{2})_\kappa}\frac{\mathcal{C}_\kappa(x\boldsymbol{\Sigma}^{-1})}{k!}\end{aligned} \tag{18}$$

for substitution of (8) and (13) for (10). This distribution function is an alternating series. Its translation with positive terms can be obtained as

$$\begin{aligned}\Pr(\ell_1 < x) =& \frac{\Gamma_m(\frac{m+1}{2})\,\Gamma(\frac{mn+\rho}{2})}{\Gamma_m(\frac{n+m+1}{2})\,\Gamma(\frac{\rho}{2})}|x\boldsymbol{\Sigma}^{-1}|^{n/2}\\ &\times\sum_{k=0}^{\infty}\frac{(\frac{nm+\rho}{2})_k\rho^{\frac{\rho}{2}}}{(\rho + x\mathrm{tr}\boldsymbol{\Sigma}^{-1})^{\frac{mn}{2}+k+\frac{\rho}{2}}}\sum_{\kappa}\frac{(\frac{m+1}{2})_\kappa}{(\frac{n+m+1}{2})_\kappa}\frac{\mathcal{C}_\kappa(x\boldsymbol{\Sigma}^{-1})}{k!}\end{aligned} \tag{19}$$

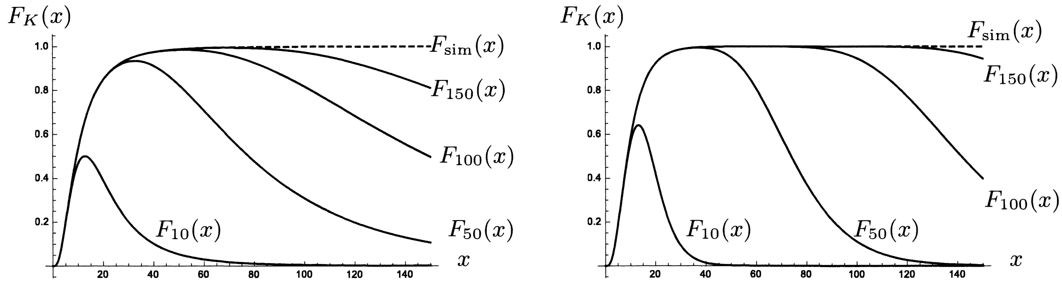
by using the Kummer relation (9). The truncated distribution up to the Kth degree of (19) is denoted by

$$\begin{aligned}F_K(x) =& \frac{\Gamma_m(\frac{m+1}{2})\,\Gamma(\frac{mn+\rho}{2})}{\Gamma_m(\frac{n+m+1}{2})\,\Gamma(\frac{\rho}{2})}|x\boldsymbol{\Sigma}^{-1}|^{n/2}\\ &\times\sum_{k=0}^{K}\frac{(\frac{nm+\rho}{2})_k\,\rho^{\frac{\rho}{2}}}{(\rho + x\mathrm{tr}\boldsymbol{\Sigma}^{-1})^{\frac{mn}{2}+k+\frac{\rho}{2}}}\sum_{\kappa}\frac{(\frac{m+1}{2})_\kappa}{(\frac{n+m+1}{2})_\kappa}\frac{\mathcal{C}_\kappa(x\boldsymbol{\Sigma}^{-1})}{k!}.\end{aligned} \tag{20}$$

For example, the function $F_{10}(x)$ is calculated as

$$\begin{aligned}F_{10}(x) = \frac{7x^4}{1250\left(\frac{3x}{20}+1\right)^9}\Bigg(&1 + \frac{81x}{14\left(\frac{3x}{2}+10\right)} + \frac{85x^2}{4\left(\frac{3x}{2}+10\right)^2} + \frac{3525x^3}{56\left(\frac{3x}{2}+10\right)^3}\\ &+\frac{238725x^4}{1456\left(\frac{3x}{2}+10\right)^4} + \frac{12537x^5}{32\left(\frac{3x}{2}+10\right)^5} + \frac{957299x^6}{1088\left(\frac{3x}{2}+10\right)^6} + \frac{77968575x^7}{41344\left(\frac{3x}{2}+10\right)^7}\\ &+\frac{1127986815x^8}{289408\left(\frac{3x}{2}+10\right)^8} + \frac{875329785x^9}{111872\left(\frac{3x}{2}+10\right)^9}\frac{17160151239x^{10}}{1118720\left(\frac{3x}{2}+10\right)^{10}}\Bigg)\end{aligned}$$

when $m = 2, n = 4, \rho = 10, \boldsymbol{\Sigma} = \mathrm{diag}(1, 2)$ and the graph of $F_{10}(x)$ are as presented on the left-hand-side of Fig. 1. We also present graphs for F_{50}, F_{100} and F_{150}, which show that F_{150} might give high accuracy by comparison with F_{sim}. On the right-hand-side of Fig. 1 we take $\rho = 50$ greater than $\rho = 10$, which shows that the truncated distributions in the right



(a) $\rho = 10$ (b) $\rho = 50$

Fig. 1: $m = 2$, $n = 4$, $\mathbf{\Sigma} = \mathrm{diag}(1, 2)$.

panel converge more rapidly to 1 than the ones in the left panel. Several percentage points are presented in Table 1 and have three-decimal-place precision for $\rho = 10, 20, 30$, and 50. Table 2 shows the vaules of u and $F_K(u)$ such that $F_K'(u) = 0$. For $\rho = 10$, $m = 2$ and $\mathbf{\Sigma} = \mathrm{diag}(1, 2)$, Table 3 shows that F_{150}s have poor accuracy, especially for 95% and 99%. When $n = 50$, the 90%, 95%, and 99% points fail to achieve the desired accuracy. As shown in Table 3, when n is large, more terms are required because F_{150} does not reach to 0.90, 0.95, or 0.99. Numerical calculations of F_{150} for $m = 3$ are much more cumbersome than

Table 1: Percentile points of truncated function ($m = 2$, $n = 4$).

ρ	% point	5%	10%	50%	90%	95%	99%
10	$F_{\rm sim}$	2.65	3.54	9.34	24.0	31.5	52.8
	F_{150}	2.65	3.54	9.35	24.0	31.5	53.1
20	$F_{\rm sim}$	2.81	3.69	9.00	20.4	25.5	38.3
	F_{150}	2.81	3.69	9.01	20.4	25.4	38.3
30	$F_{\rm sim}$	2.87	3.75	8.88	19.3	23.7	34.5
	F_{150}	2.87	3.75	8.89	19.3	23.7	34.5
50	$F_{\rm sim}$	2.91	3.80	8.80	18.4	22.4	31.7
	F_{150}	2.92	3.80	8.80	18.5	22.4	31.8

Table 2: Values of $(u, F_K(u))$ such that $F_K'(u) = 0$ for $m = 2$, $n = 4$, $\mathbf{\Sigma} = \mathrm{diag}(1, 2)$.

ρ		u	$F_K(u)$
10	F_{10}	12.7	0.501
	F_{50}	32.6	0.935
	F_{100}	52.5	0.985
	F_{150}	69.6	0.995
50	F_{10}	13.2	0.641
	F_{50}	37.3	0.994
	F_{100}	62.1	1.00
	F_{150}	84.3	1.00

those for $m = 2$. It took less than 4 seconds to calculate F_{150} for $m = 2$, but more than 6 hours for $m = 3$. Calculation times were measured using `Mathematica 10` on a computer (MacBookAir OS X Yosemite, ver. 10.10.5; Apple Computer, Inc. with a 1.6GHz Intel Core i5 processor; Intel Corp.) with 4 GB memory. Table 4 shows the behavior similarly to the case for $m = 2$, where the cases for larger values of ρ have a rapid convergence property. In Tables 3 and 4, the symbol "-" signifies that F_{150} does not reach to the probability.

Table 3: Percentile points of truncated function ($\rho = 10$, $m = 2$).

n	% point	5%	10%	50%	90%	95%	99%
3	F_{sim}	1.72	2.40	7.06	19.3	25.7	44.0
	F_{150}	1.71	2.39	7.06	19.3	25.7	44.0
5	F_{sim}	3.63	4.72	11.6	28.5	37.0	61.7
	F_{150}	3.62	4.71	11.6	28.5	37.1	62.4
10	F_{sim}	8.62	10.6	22.5	50.2	63.9	103
	F_{150}	8.61	10.6	22.5	50.2	64.1	-
20	F_{sim}	18.9	22.5	43.9	92.0	115	182
	F_{150}	18.9	22.6	43.9	95.2	-	-
30	F_{sim}	29.4	34.7	65.2	133	167	261
	F_{150}	29.4	34.8	65.2	-	-	-
50	F_{sim}	50.8	59.4	108	216	269	421
	F_{150}	50.8	59.4	109	-	-	-

Table 4: Percentile points of truncated function ($m = 3$, $n = 4$).

ρ	% point	5%	10%	50%	90%	95%	99%
10	F_{sim}	5.60	7.18	17.1	41.3	53.4	88.6
	F_{150}	5.59	7.17	17.1	41.3	53.8	-
20	F_{sim}	5.98	7.54	16.5	34.7	42.7	63.1
	F_{150}	5.99	7.54	16.5	34.7	42.7	63.2
30	F_{sim}	6.16	7.70	16.2	32.7	39.7	56.3
	F_{150}	6.15	7.69	16.3	32.7	39.6	56.3
50	F_{sim}	6.30	7.83	16.1	31.2	37.3	51.6
	F_{150}	6.30	7.83	16.1	31.2	37.3	51.5

Finally, we present a discussion of numerical computation of the distribution of the smallest eigenvalue under $\boldsymbol{X} \sim T_{n\times m}(\rho, \boldsymbol{\Sigma})$ and $n \geq m$. In the general case of $\boldsymbol{A} \sim \mathcal{EW}(n, \boldsymbol{\Sigma}, h)$, the upper-tail probability of the smallest eigenvalue was given by Caro-Lopera et al. (2016) as well as the case of $T_{n\times m}(\rho, \boldsymbol{\Sigma})$. If $r = (n-m-1)/2$ is a positive integer, then the upper-tail probability of the smallest eigenvalue ℓ_m of $\boldsymbol{A} = \boldsymbol{X}^\top \boldsymbol{X}$ under $T_{n\times m}(\rho, \boldsymbol{\Sigma})$ is given as

$$\Pr(\ell_m > x) = \sum_{k=0}^{mr} \frac{(\rho/2)_k \; x^k}{k! \; \rho^k} \left(1 + \frac{x}{\rho}\mathrm{tr}\boldsymbol{\Sigma}^{-1}\right)^{-\rho/2-k} \sum_{\kappa^*} \mathcal{C}_{\kappa^*}(\boldsymbol{\Sigma}^{-1}), \tag{21}$$

where the summation runs over partitions $\kappa^* = (\kappa_1, \ldots, \kappa_m)$ of κ with restriction of $\kappa_1 \le r$. It is easier to calculate (21) numerically than (19) because (21) is a finite series up to mr. We check the limit of ρ. If $\rho \to \infty$, then the matrix-variate t distribution with a trace type converges in law to a matrix-variate normal distribution $\mathcal{N}_{n\times m}(\mathbf{0}, \boldsymbol{I}_n \otimes \boldsymbol{\Sigma})$ because of

$$\begin{aligned}\lim_{\rho\to\infty} h(y) &= \lim_{\rho\to\infty} \frac{\Gamma((mn+\rho)/2)}{(2\pi)^{mn/2}(\rho/2)^{mn/2}\Gamma(\rho/2)} \left(1+\frac{y}{\rho}\right)^{-mn/2} \left\{\left(1+\frac{y}{\rho}\right)^{\rho}\right\}^{-1/2} \\ &= \frac{1}{(2\pi)^{mn/2}} \exp(-y/2).\end{aligned}$$

Therefore, if $\rho \to \infty$, then the upper-tail probability of ℓ_m is given as

$$\Pr(l_m > x) = \operatorname{etr}\left(-\frac{1}{2}x\boldsymbol{\Sigma}^{-1}\right) \sum_{k=0}^{mr} \frac{1}{k!} \sum_{\kappa^*} \mathcal{C}_{\kappa^*}\left(\frac{1}{2}x\boldsymbol{\Sigma}^{-1}\right)$$

which coincides with the result of Khatri (1972). Fig. 2 presents graphs of (21) when $\rho = 5, 40$ and $\rho \to \infty$ under $n = 10$ and $\boldsymbol{\Sigma} = \operatorname{diag}(1, 2, 4)$. Their percentages are shown in Table 5.

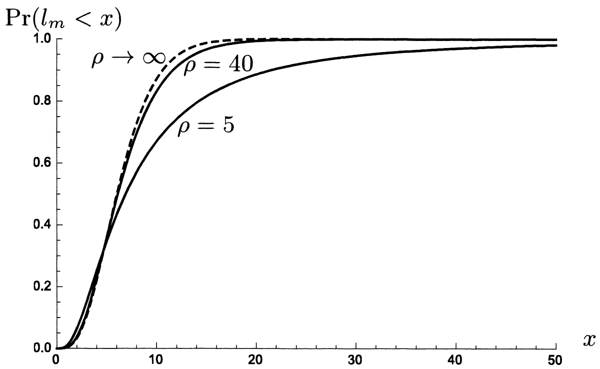


Fig. 2: $m = 3$, $n = 10$, $\boldsymbol{\Sigma} = \operatorname{diag}(1, 2, 4)$.

Table 5: Percentile points of the smallest eigenvalue ($n = 10$, $\boldsymbol{\Sigma} = \operatorname{diag}(1, 2, 4)$).

ρ	5%	10%	50%	90%	95%	99%
5	1.82	2.45	6.90	21.6	31.0	66.0
10	2.01	2.64	6.44	15.3	19.7	32.1
20	2.14	2.77	6.23	12.8	15.6	22.4
40	2.22	2.84	6.13	11.7	13.8	18.7
∞	2.31	2.93	6.03	10.6	12.2	15.5

5. Conclusion

As described herein, we provide an alternative proof of the derivation of the largest-eigenvalue distribution under an elliptically contoured distribution. We also discuss numerical computation of the largest and smallest eigenvalues' distributions under the matrix-variate t distribution. The relation between the multivariate and matrix-variate t distributions is useful in conducting Monte Carlo simulations. Numerical computation is limited to cases in which the dimension m is small. Therefore, we must investigate the asymptotic properties for m or n, which remains as task for future work. Other future work includes numerical computation and simulation for other elliptically contoured distributions.

Acknowledgements

The authors would like to thank the editors and the two anonymous referees for their helpful comments and suggestions.

REFERENCES

Caro-Lopera, F. J., Díaz-García, J. A. and González-Farías G. (2010). Noncentral elliptical configuration density. *Journal of Multivariate Analysis,* **101** (1), 32–43.

Caro-Lopera, F. J., González-Farías, G. and Balakrishnan N. (2014a). On generalized Wishart distributions – I: Likelihood ratio test for homogeneity of covariance matrices. *Sankhyā: The Indian Journal of Statistics,* **76-A** (2), 179–194.

Caro-Lopera, F. J., González-Farías, G. and Balakrishnan N. (2014b). On generalized Wishart distributions – II: Sphericity test. *Sankhyā: The Indian Journal of Statistics,* **76-A** (2), 195–218.

Caro-Lopera, F. J., González-Farías, G. and Balakrishnan N. (2016). Matrix-Variate distribution theory under elliptical models – 4: Joint distribution of latent roots of covariance matrix and the largest and smallest latent roots. *Journal of Multivariate Analysis,* **145**, 224–235.

Constantine, A. G. (1963). Some non-central distribution problems in multivariate analysis. *Annals of Mathematical Statistics,* **34**, 1270–1285.

Díaz-García, J. A. and Caro-Lopera, F. J. (2008). Matrix Kummer – general relation. Comunicaión del CIMAT No I-08-16/25-09-2008.

Gupta, A. K. and Nagar, D. K. (1999). Matrix Variate Distributions. *Chapman & Hall / CRC.*

Hashiguchi, H., Nakagawa, S. and Niki, N. (2000). Simplification of the Laplace–Beltrami operator. *Mathematics and Computers in Simulation,* **51** (5), 489–496.

James, A. T. (1960). The distribution of the latent roots of the covariance matrix. *Annals of Mathematical Statistics,* **31**, 151–158.

James, A. T. (1964). Distribution of matrix variates and latent roots derived from normal samples. *Annals of Mathematical Statistics,* **35**, 475–501.

Khatri, C. G. (1967). Some distribution problems connected with the characteristic roots of $S_1S_2^{-1}$. *Annals of Mathematical Statistics,* **38**, 944–948.

Khatri, C. G. (1972). On the exact finite series distribution of the smallest or the largest root of matrices in three situations. *Journal of Multivariate Analysis,* **2**, 201–207.

Muirhead, R. J. (1982). Aspects of Multivariate Statistical Theory. *John Wiley & Sons, New York.*

Sugiyama, T. (1967). On the distribution of the largest latent root of the covariance matrix. *Annals of Mathematical Statistics,* **38**, 1148–1151.

Sutradhar, B. C. and Ali, M. M. (1989). A generalization of the Wishart distribution for the elliptical model and its moments for the multivariate t model. *Journal of Multivariate Analysis,* **29** (1), 155–162.

Wishart, J. (1928). The generalised product moment distribution in samples from a normal multivariate population. *Biometrika,* **20A** (1), 32–52.

Wolfram Research, Inc., Mathematica, Ver. 10.3, Champaign, IL (2015).

(Received: August 16, 2017, Accepted: February 6, 2018)