SUT Journal of Mathematics
Vol. 52, No. 2 (2016), 141–158

Approximate eigenvalue distribution for the ratio of Wishart matrices

Shusuke Matsubara and Hiroki Hashiguchi

(Received February 18, 2016; Revised September 10, 2016)

Abstract. We discuss approximations for the distribution of eigenvalues of the ratio of Wishart matrices when the population eigenvalues are infinitely dispersed. The first approximation is expressed as the F distribution with suitable parameters, and the second is expressed by the product of F distributions. Numerical examples show that the proposed approximations are more accurate than the known asymptotic expansions of the normal distribution.

AMS 2010 *Mathematics Subject Classification.* 62E20, 62H10.

Key words and phrases. Laplace's method, hypergeometric function with two matrix arguments, F distribution.

§1. Introduction

Random matrix theory originated in mathematical physics and statistics, and recently it has found a wide range of applications in the fields of science and engineering. One of the fundamental random matrices in multivariate analysis, the Wishart matrix, has important uses in estimation and in statistical tests involving the sample covariance matrix. The landmark studies on random matrix theory in statistics were Johnstone (2001, 2008, 2009). These studies focus primarily on the null case, in which the population covariance matrix is the identity matrix. Some multivariate statistics are also expressed as the function of the eigenvalues of Wishart matrices, therefore it is important to derive the distributions of these eigenvalues. The distribution of the eigenvalues of a Wishart matrix or of the ratio of Wishart matrices depends on a definite integral over the group of orthogonal matrices. This integral is expressed as a hypergeometric series involving zonal polynomials, and it is difficult to compute them numerically in a non-null case.

To approximate the distribution of the eigenvalues of a Wishart matrix, Sugiura (1973) and Muirhead and Chikuse (1975) derived asymptotic expansions with normal distributions. Approximations have also been obtained with χ^2-distributions by Sugiyama (1972), Takemura and Sheena (2005), and Kato and Hashiguchi (2014). For a ratio of Wishart matrices, Khatri (1967) derived exactly the joint probability density function (pdf) of the eigenvalues, and Li et al. (1970) derived an asymptotic expansion by evaluating an approximation of the integral over the orthogonal group. Sugiura (1976) and Chikuse (1977) derived an asymptotic expansion using the normal distribution.

In this paper, we use the F distribution to derive an approximation for the distribution of eigenvalues of the ratio of Wishart matrices when population eigenvalues are infinitely dispersed. This infinite dispersion property of population eigenvalues was introduced by Takemura and Sheena (2005). We also consider an approximation that uses the product of F distributions; we use a similar method to Kato and Hashiguchi (2014). In the remaining part of this introduction, we summarize the results of Kato and Hashiguchi (2014) for a single Wishart matrix. In Section 2, we discuss an extension of Kato and Hashiguchi (2014) for the ratio of Wishart matrices. In Section 3, numerical experiments are performed via Monte Carlo simulations.

Let W be distributed as the Wishart distribution $W_p(n, \Sigma)$, where $n \geq p$ and the covariance matrix Σ is positive definite. The eigenvalues of Σ are denoted by $\sigma_1, \ldots, \sigma_p$, and we assume that $\sigma_1 > \cdots > \sigma_p > 0$. For a Wishart matrix W, the eigenvalues are denoted by $w_1 > w_2 > \cdots > w_p$, which are random variables.

From Theorem 3.2.18 of Muirhead (1982; p. 106), the joint distribution of $w_1, w_2, \ldots, w_p$ is

$$f(w_1, \ldots, w_p) = \frac{2^{-pn/2}\, \pi^{p^2/2}}{\Gamma_p(p/2)\Gamma_p(n/2)|\Sigma|^{n/2}} \prod_{j=1}^{p} w_j^{\frac{n-p-1}{2}} \prod_{j<k}(w_j - w_k) \\ {}_0F_0^{(p)}\left(-\frac{1}{2}\Sigma^{-1}, L\right),$$

where

$${}_0F_0^{(p)}\left(-\frac{1}{2}\Sigma^{-1}, L\right) = \int_{O(p)} \mathrm{etr}\left(-\frac{1}{2}\Sigma^{-1} HLH^{\top}\right)(dH),$$

$$\Gamma_p(a) = \pi^{\frac{p(p-1)}{4}} \prod_{i=1}^{p} \Gamma\left(a - \frac{i-1}{2}\right)$$

and (dH) is the normalized Haar mesure on the orthogonal group $O(p)$. From Theorem 9.5.2 of Muirhead (1982; p.392), the integral has the following asymptotic behavior

$$(1.1)\; {}_0F_0^{(p)}\left(-\frac{1}{2}\Sigma^{-1}, L\right) \sim \frac{\Gamma_p(p/2)}{\pi^{p^2/2}} \exp\left(-\frac{1}{2}\sum_{j=1}^{p}\frac{w_j}{\sigma_j}\right)\prod_{j<k}^{p}\left(\frac{2\pi}{c_{jk}}\right)^{1/2},$$

where $c_{jk} = [(w_j - w_k)(\sigma_j - \sigma_k)]/[\sigma_j\sigma_k]$. When we say "$a \sim b$ for sufficiently large n," we mean that $a/b \to 1$ as $n \to \infty$. Furthermore, we define ρ_1 as follows:

$$\rho_1 = \max\left(\frac{\sigma_2}{\sigma_1}, \cdots, \frac{\sigma_p}{\sigma_{p-1}}\right),$$

and we consider the case for $\rho_1 \to 0$. For any random variables X and Y, we use the notation

$$X \approx Y \quad \text{or} \quad \Pr[X < x] \approx \Pr[Y < y]$$

to mean that, for sufficiently large n, X converges to Y as $\rho_1 \to 0$. By evaluating the asymptotic expansion (1.1) when $\rho_1 \to 0$ for sufficiently large n, Kato and Hashiguchi (2014) showed Propositions 1.1 and 1.2.

Proposition 1.1. *Let $w_1, \dots, w_p$ be the eigenvalues of $W \sim W_p(n, \Sigma)$, where $n \geq p$, Σ is positive definite, and $w_1 > w_2 > \cdots > w_p$. If $\rho_1 \to 0$, then for sufficiently large n, $w_1, \dots, w_p$ are mutually independent, and each w_k is asymptotically distributed as the χ^2-distribution with $n - k + 1$ degrees of freedom.*

Proposition 1.1 is almost the same as a result of Takemura and Sheena (2005) that places no assumptions on the sample size n. Considering the order of $w_1, \dots, w_p$ and their asymptotic behavior, Kato and Hashiguchi (2014) obtained the following proposition, which states that each w_k can be approximated by a product of χ^2-distributions.

Proposition 1.2. *Let $Y_1, \dots, Y_p$ be mutually independent random variables, and let each Y_k be distributed as a χ^2-distribution with $n - k + 1$ degrees of freedom. We define $\overline{U}_{(k)}$ and $\underline{U}_{(k)}$ as*

$$\begin{cases} \overline{U}_{(k)} &= \{\sigma_1 Y_1, \sigma_2 Y_2, \dots, \sigma_k Y_k\} \\ \underline{U}_{(k)} &= \{\sigma_k Y_k, \sigma_{k+1} Y_{k+1}, \dots, \sigma_m Y_m\}, \end{cases}$$

where, for convenience, we let $\overline{U}_{(0)} = \{\infty\}$ and $\underline{U}_{(m+1)} = \{0\}$. If $\rho_1 \to 0$, then for sufficiently large n, the following two equations hold.

1. $\ell_k \approx \min\{\min \overline{U}_{(k-1)}, \max \underline{U}_{(k)}\}$,

$$\begin{aligned}\Pr[w_k > x] &\approx \Pr[\min\{\min \overline{U}_{(k-1)}, \max \underline{U}_{(k)}\} > x] \\ &= \prod_{j=1}^{k-1}(1 - G_{n-j+1}(x/\sigma_j)) \times \left(1 - \prod_{j=k}^{m} G_{n-j+1}(x/\sigma_j)\right).\end{aligned}$$

2. $w_k \approx \max\{\min \overline{U}_{(k)}, \max \underline{U}_{(k+1)}\}$,

$$\begin{aligned}\Pr[w_k < x] &\approx \Pr[\max\{\min \overline{U}_{(k)}, \max \underline{U}_{(k+1)}\} < x] \\ &= \left(1 - \prod_{j=1}^{k}(1 - G_{n-j+1}(x/\sigma_j)\right) \times \prod_{j=k+1}^{m} G_{n-j+1}(x/\sigma_j).\end{aligned}$$

Corollary 1.3. *Under the same conditions as Proposition 1.2, the approximate distribution of the eigenvalue w_1 is given by*

$$w_1 \approx \max \underline{U}_{(1)} \quad \textit{and} \quad \Pr[w_1 < x] \approx \prod_{k=1}^{p} G_{n-j+1}(x/\sigma_j). \tag{1.2}$$

Similarly, we have

$$w_p \approx \min \overline{U}_{(p)} \quad \textit{and} \quad \Pr[w_p > x] \approx \prod_{k=1}^{p} \left(1 - G_{n-j+1}(x/\sigma_j)\right).$$

We note that equation (1.2) is the same as a result of Sugiyama (1972), but without assumptions on ρ and n.

§2. Main results

In this section, we consider the distribution of the eigenvalues of the ratio of Wishart matrices. Let $W_j (j = 1, 2)$ be independently distributed as $W_p(n_j, \Sigma_j)$. For $k = 1, \ldots, p$, let ℓ_k denote the eigenvalues of $W_1 W_2^{-1}$, and let λ_k denote the population eigenvalues of $\Sigma_1 \Sigma_2^{-1}$, where $\ell_1 > \cdots > \ell_p > 0$ and $\lambda_1 > \cdots > \lambda_p > 0$.

Let X and Y be $p \times p$ positive Hermitian matrices. Then the hypergeometric function ${}_1F_0^{(p)}(a; X, Y)$ with two arguments X and Y is defined by

$${}_1F_0^{(p)}(a; X, Y) = \int_{O(p)} |I - XHYH^\top|^{-a} (dH), \tag{2.1}$$

where $O(p)$ denotes the set of $p \times p$ orthogonal matrices, and (dH) is the normalized Haar measure on $O(p)$. Let $x_1, \ldots, x_p$ and $y_1, \ldots, y_p$ be the eigenvalues of X and Y, respectively, where $x_1 > x_2 > \cdots > x_p > 0$ and $y_1 > y_2 > \cdots > y_p > 0$. Using Laplace's method in a similar way to that of (1.1), the asymptotic behavior of (2.1) is given by

$${}_1F_0^{(p)}(a; X, Y) \sim \frac{\Gamma_p(p/2)}{\pi^{p^2/2}} |I - XY|^{-a} \prod_{j<k} \left(\frac{\pi}{a\, c_{jk}}\right)^{\frac{1}{2}}, \tag{2.2}$$

where $c_{jk} = [(x_j - x_k)(y_j - y_k)]/[(1-x_jy_j)(1-x_ky_k)]$. A general formula based on Laplace's method for a hypergeometric function with two matrix arguments was obtained in Butler and Wood (2005). The asymptotic properties of (1.1) and (2.2) are special cases of the results of Butler and Wood (2005). The right hand side of (2.2) is the same as the first-order term of the asymptotic expansion of ${}_1F_0^{(p)}$ given by Li et al. (1970).

James (1964) introduced the hypergeometric function for matrix arguments and gave the joint pdf of $\ell_1, \ldots, \ell_p$. Khatri (1967) provided another expression for the joint distribution, and Khatri (1972) presented the distribution of the largest and smallest eigenvalues. In Khatri (1972), the distribution of ℓ_1 and ℓ_p were expressed by a finite series of Laguerre polynomials with matrix arguments. Under the null hypothesis $\Sigma_1\Sigma_2^{-1} = I_p$, Venables (1973) proposed a method for exactly computing the distribution of ℓ_1 and ℓ_p.

Proposition 2.1. *(Joint pdf of the eigenvalues)*
Let $n = n_1 + n_2$*,* $A = \mathrm{diag}(\lambda_1, \ldots, \lambda_p)$*, and* $B = \mathrm{diag}(\ell_1, \ldots, \ell_p)$*.*

1. *(James, 1964) The joint pdf of the eigenvalues* $\ell_1, \ldots, \ell_p$ *of* $W_1W_2^{-1}$ *is given by*

(2.3)
$$f(\ell_1, \ldots, \ell_p) = \frac{\pi^{p^2/2}|A|^{-\frac{n_1}{2}}|B|^{\frac{n_1-p-1}{2}}}{B_p(n_1/2, n_2/2)\Gamma_p(p/2)} \prod_{j<k}(\ell_j - \ell_k)\,{}_1F_0^{(p)}\left(\frac{n}{2}; -A^{-1}, B\right),$$

where $B_p(n_1/2, n_2/2)$ *is the multivariate beta function with parameters* $n_1/2$ *and* $n_2/2$ *as*

$$B_p\left(n_1/2, n_2/2\right) = \frac{\Gamma_p(n_1/2)\Gamma_p(n_2/2)}{\Gamma_p(n/2)}.$$

2. *(Khatri, 1967) Another expression of* $f(\ell_1, \ldots, \ell_p)$ *is given by*

$$\frac{|A|^{-\frac{n_1}{2}}|B|^{\frac{n_1-p-1}{2}}}{B_p(n_1/2, n_2/2)} \prod_{j<k}(\ell_j - \ell_k)|I + B|^{-\frac{n}{2}}\,{}_1F_0^{(p)}\left(\frac{n}{2}; I - A^{-1}, B(I+B)^{-1}\right).$$

Applying the Laplace approximation (2.2) to Proposition 2.1, the following corollary is clearly obtained.

Corollary 2.2. *The Laplace approximation for the joint pdf* $f(\ell_1, \ldots, \ell_p)$ *of* (2.3) *in Proposition 2.1 is given by*

$$f(\ell_1, \ldots, \ell_p) \sim \frac{1}{B_p(n_1/2, n_2/2)} \prod_{j<k}\left(\frac{2\pi}{n}\right)^{\frac{1}{2}} |A|^{-\frac{n_1}{2}}|B|^{\frac{n_1-p-1}{2}} \left|I + A^{-1}B\right|^{-\frac{n}{2}} \prod_{j<k}(\ell_j - \ell_k)\, c_{jk}^{-\frac{1}{2}}, \tag{2.4}$$

where $c_{jk} = [(\lambda_j - \lambda_k)(\ell_j - \ell_k)]/[(\lambda_j + \ell_j)(\lambda_k + \ell_k)]$.

Proof. From (2.2), we have

$$ {}_1F_0\left(\frac{n}{2}; -A^{-1}, B\right) \sim \frac{\Gamma_p(p/2)}{\pi^{p^2/2}} \left|I + A^{-1}B\right|^{-\frac{n}{2}} \prod_{j<k} \left(\frac{2\pi}{n\ c_{jk}}\right)^{\frac{1}{2}}. $$

Substitute it into equation (2.3). □

For sufficiently large n, the normalizing constant on the right-hand side of (2.4) has the asymptotic property shown in the following lemma.

Lemma 2.3. *If n is sufficiently large, then*

$$ \frac{1}{B_p(n_1/2, n_2/2)} \prod_{j<k} \left(\frac{2\pi}{n}\right)^{\frac{1}{2}} \sim \prod_{j=1}^{p} \frac{1}{B(\frac{n_1-j+1}{2}, \frac{n_2-p+j}{2})}. $$

Proof. We note that the normalizing constant can be written as the product of gamma and beta functions, as follows:

$$ \begin{aligned} \frac{1}{B_p(n_1/2, n_2/2)} \prod_{j<k} \left(\frac{2\pi}{n}\right)^{\frac{1}{2}} &= \frac{\Gamma_p(n/2)}{\Gamma_p(n_1/2)\Gamma_p(n_2/2)} \left(\frac{2\pi}{n}\right)^{\frac{p(p-1)}{4}} \\ &= \left(\frac{2}{n}\right)^{\frac{p(p-1)}{4}} \prod_{j=1}^{p} \frac{\Gamma(\frac{n-j+1}{2})}{\Gamma(\frac{n-p+1}{2})} \cdot \frac{1}{B(\frac{n_1-j+1}{2}, \frac{n_2-p+j}{2})}. \end{aligned} $$

Next, we note that the following two statements hold:

$$ \left(\frac{2}{n}\right)^{\frac{p(p-1)}{4}} \prod_{j=1}^{p} \frac{\Gamma(\frac{n-j+1}{2})}{\Gamma(\frac{n-p+1}{2})} \sim 1 \quad \text{and} $$

$$ \frac{1}{B_p(n_1/2, n_2/2)} \prod_{j<k} \left(\frac{2\pi}{n}\right)^{\frac{1}{2}} \sim \prod_{j=1}^{p} \frac{1}{B(\frac{n_1-j+1}{2}, \frac{n_2-p+j}{2})}, $$

because Stirling's formula for the gamma function gives

$$\begin{aligned}
\left(\frac{2}{n}\right)^{\frac{p(p-1)}{4}} \prod_{j=1}^{p} \frac{\Gamma(\frac{n-j+1}{2})}{\Gamma(\frac{n-p+1}{2})} &= \left(\frac{2}{n}\right)^{\frac{p(p-1)}{4}} \prod_{j=1}^{p-1} \frac{\Gamma(\frac{n-j+1}{2}+1)}{\Gamma(\frac{n-p-1}{2}+1)} \\
&\sim \left(\frac{2}{n}\right)^{\frac{p(p-1)}{4}} \prod_{j=1}^{p-1} \frac{\sqrt{(n-j-1)\pi}(\frac{n-j-1}{2\mathrm{e}})^{\frac{n-j-1}{2}}}{\sqrt{(n-p-1)\pi}(\frac{n-p-1}{2\mathrm{e}})^{\frac{n-p-1}{2}}} \\
&= \left(\frac{2}{n}\right)^{\frac{p(p-1)}{4}} \left[\prod_{j=1}^{p-1} \frac{(1-\frac{j+1}{n})^{\frac{n-j}{2}}}{(1-\frac{p+1}{n})^{\frac{n-p}{2}}}\right] \left(\frac{n}{2e}\right)^{\frac{p(p-1)}{4}} \\
&\sim e^{-\frac{p(p-1)}{4}} \prod_{j=1}^{p-1} \frac{e^{-\frac{j+1}{2}}}{e^{-\frac{p+1}{2}}}. \\
&= 1.
\end{aligned}$$

□

In Proposition 2.1 and Corollary 2.2, we consider the asymptotic joint pdf of $\ell_1, \ldots, \ell_p$ in the case that the population eigenvalues are infinity dispersed when

$$\rho_2 = \max\left(\frac{\lambda_2}{\lambda_1}, \cdots, \frac{\lambda_p}{\lambda_{p-1}}\right) \to 0.$$

If $\rho_2 \to 0$, then Lemma 2.4 is obtained.

Lemma 2.4. *If n is sufficiently large and $\rho_2 \to 0$, then we have*

$$\Pr\left(\frac{\ell_k}{\ell_j} > 0\right) \to 0$$

for $1 \le j < k \le p$.

Proof. From Corollary 3.1 of Sugiura (1976), each ℓ_k is asymptotically distributed as $N(\lambda_k, 2n\lambda_k^2/n_1n_2)$, and $\ell_1, \ldots, \ell_p$ are asymptotically independent of each other. Hence, using Markov's inequality and a delta method for $E(\ell_j^{-1})$, we obtain

$$\Pr\left(\left|\frac{\ell_k}{\ell_j}\right| > \epsilon\right) \le \frac{1}{\epsilon}\mathrm{E}\left(\left|\frac{\ell_k}{\ell_j}\right|\right) = \frac{\lambda_k}{\lambda_j \epsilon}\left(1 + O\left(\frac{n}{n_1n_2}\right)\right)$$

for any $\epsilon > 0$. Set $\epsilon = \sqrt{\lambda_k/\lambda_j}$. If $\rho_2 \to 0$, then we have $\epsilon \to 0$ and

$$\lim_{\rho_2 \to 0} \Pr\left(\frac{\ell_k}{\ell_j} > \sqrt{\frac{\lambda_k}{\lambda_j}}\right) \le \lim_{\rho_2 \to 0} \sqrt{\frac{\lambda_k}{\lambda_j}}\left(1 + O\left(\frac{n}{n_1n_2}\right)\right) = 0,$$

which yields

$$\Pr\left(\frac{\ell_k}{\ell_j} > 0\right) \to 0.$$

□

The above lemma means that ℓ_k/ℓ_j asymptotically tends to zero with probability one. Furthermore, from Lemmas 2.3 and 2.4, we have the following asymptotic pdfs.

Theorem 2.5. *1. Let $x_j = \ell_j/\lambda_j$, $m_{1,j} = n_1 - j + 1$, and $m_{2,j} = n_2 - p + j$. If $\rho_2 \to 0$ and n is sufficiently large, then $x_1, \ldots, x_n$ are mutually independent, and each x_j is asymptotically distributed as the beta distribution of the second kind with parameters $m_{1,j}/2$ and $m_{2,j}/2$:*

$$f(x_1, \ldots, x_p) \approx \prod_{j=1}^{p} \frac{1}{B(m_{1,j}/2, m_{2,j}/2)} \frac{x_j^{\frac{m_{1,j}}{2}-1}}{(1+x_j)^{\frac{m_{1,j}+m_{2,j}}{2}}}.$$

2. Let $y_j = x_j/(1+x_j)$, then we also have

$$f(y_1, \ldots, y_p) \approx \prod_{j=1}^{p} \frac{y_j^{\frac{m_{1,j}}{2}-1}(1-y_j)^{\frac{m_{2,j}}{2}-1}}{B(m_{1,j}/2, m_{2,j}/2)},$$

where $0 \le y_j \le 1$. Namely, each y_j is asymptotically distributed as the beta distribution of the first kind with parameters $m_{1,j}/2$ and $m_{2,j}/2$.

3. Furthermore, if we set $z_j = m_{2,j}\, x_j/m_{1,j}$, then we also have

$$f(z_1, \ldots, z_p) \approx \prod_{j=1}^{p} \frac{(\frac{m_{1,j}}{m_{2,j}})^{\frac{m_{1,j}}{2}}}{B(m_{1,j}/2, m_{2,j}/2)} \frac{z_j^{\frac{m_{1,j}}{2}-1}}{(1+\frac{m_{1,j}}{m_{2,j}}z_j)^{\frac{m_{1,j}+m_{2,j}}{2}}}\, dz_j,$$

where $0 \le z_j < \infty$. Thus, each z_j is asymptotically distributed as the F distribution with parameters $m_{1,j}$ and $m_{2,j}$.

Proof. First, the terms other than the normalizing constant in (2.4) can be rewritten as

$$
\begin{aligned}
&|A|^{-\frac{n_1}{2}} \, |B|^{\frac{n_1-p-1}{2}} \, \left|I + A^{-1}B\right|^{-\frac{n}{2}} \prod_{j<k} (\ell_j - \ell_k) \, c_{jk}^{-\frac{1}{2}} \\
&= \prod_{j=1}^{p} \lambda_j^{-\frac{n_1}{2}} \ell_j^{\frac{n_1-p-1}{2}} \left(1 + \frac{\ell_j}{\lambda_j}\right)^{-\frac{n}{2}} \prod_{j<k} \left(\frac{1}{c_{jk}}\right)^{\frac{1}{2}} (\ell_j - \ell_k) \\
&= \prod_{j=1}^{p} \lambda_j^{-\frac{n_1}{2}} \ell_j^{\frac{n_1-p-1}{2}} \left(1 + \frac{\ell_j}{\lambda_j}\right)^{-\frac{n}{2}} \prod_{j<k} \left(\frac{(\lambda_j + \ell_j)(\lambda_k + \ell_k)(\ell_j - \ell_k)}{\lambda_j - \lambda_k}\right)^{\frac{1}{2}} \\
&= \prod_{j=1}^{p} \lambda_j^{-\frac{n_1}{2}} \ell_j^{\frac{n_1-p-1}{2}} \left(1 + \frac{\ell_j}{\lambda_j}\right)^{-\frac{n}{2}} \prod_{j<k} \left(\frac{\left(1 + \frac{\ell_j}{\lambda_j}\right)\left(1 + \frac{\ell_k}{\lambda_k}\right)\left(1 - \frac{\ell_k}{\ell_j}\right)\lambda_k \ell_j}{1 - \frac{\lambda_k}{\lambda_j}}\right)^{\frac{1}{2}} .
\end{aligned}
$$

If $\rho_2 \to 0$ and n is sufficiently large, then from Lemma 2.4, we have $\ell_k/\ell_j \approx 0$, and $\lambda_k/\lambda_j \to 0$ for $1 \le j < k \le p$. Hence, the last line of the above equation, with differential operators $d\ell_1 \cdots d\ell_p$, can be expressed as

$$
\begin{aligned}
&\approx \prod_{j=1}^{p} \lambda_j^{-\frac{n_1}{2}} \ell_j^{\frac{n_1-p-1}{2}} \left(1 + \frac{\ell_j}{\lambda_j}\right)^{-\frac{n}{2}} \prod_{j<k} \left(1 + \frac{\ell_j}{\lambda_j}\right)^{\frac{1}{2}} \left(1 + \frac{\ell_k}{\lambda_k}\right)^{\frac{1}{2}} \lambda_k^{\frac{1}{2}} \ell_j^{\frac{1}{2}} d\ell_j \\
&= \prod_{j=1}^{p} \lambda_j^{-\frac{n_1}{2}} \ell_j^{\frac{n_1-p-1}{2}} \left(1 + \frac{\ell_j}{\lambda_j}\right)^{-\frac{n}{2}} \prod_{j=1}^{p} \left\{\ell_j \left(1 + \frac{\ell_j}{\lambda_j}\right)\right\}^{\frac{p-j}{2}} \left\{\lambda_k \left(1 + \frac{\ell_k}{\lambda_k}\right)\right\}^{\frac{k-1}{2}} d\ell_j \\
&= \prod_{j=1}^{p} \left(\frac{\ell_j}{\lambda_j}\right)^{\frac{n_1-j+1}{2}-1} \frac{1}{\lambda_j} \left(1 + \frac{\ell_j}{\lambda_j}\right)^{\frac{-n+p-1}{2}} d\ell_j \\
&= \prod_{j=1}^{p} \left(\frac{\ell_j}{\lambda_j}\right)^{\frac{n_1-j+1}{2}-1} \left(1 + \frac{\ell_j}{\lambda_j}\right)^{\frac{-n+p-1}{2}} \frac{d\ell_j}{\lambda_j}.
\end{aligned}
$$

Therefore, we obtain

$$
\begin{aligned}
&\frac{1}{B_p(\frac{n_1}{2}, \frac{n_2}{2})} \prod_{j<k} \left(\frac{2\pi}{n}\right)^{\frac{1}{2}} |A|^{-\frac{n_1}{2}} \, |B|^{\frac{n_1-p-1}{2}} \, \left|I + A^{-1}B\right|^{-\frac{n}{2}} \prod_{j<k} (\ell_j - \ell_k) \, c_{jk}^{-\frac{1}{2}} \, d\ell_j \\
&\approx \prod_{j=1}^{p} \frac{1}{B(\frac{n_1-j+1}{2}, \frac{n_2-p+j}{2})} \left(\frac{\ell_j}{\lambda_j}\right)^{\frac{n_1-j+1}{2}-1} \left(1 + \frac{\ell_j}{\lambda_j}\right)^{\frac{-n+p-1}{2}} \frac{d\ell_j}{\lambda_j}. \qquad (2.5)
\end{aligned}
$$

If we set $x_j = \ell_j/\lambda_j$, $m_{1,j} = n_1 - j + 1$, and $m_{2,j} = n_2 - p + j$, then equation (2.5) becomes the product of beta distributions of the second kind with parameters

$m_{1,j}/2$ and $m_{2,j}/2$:

$$f(x_1,\dots,x_p) \approx \prod_{j=1}^{p} \frac{1}{B(m_{1,j}/2, m_{2,j}/2)} \frac{x_j^{\frac{m_{1,j}}{2}-1}}{(1+x_j)^{\frac{m_{1,j}+m_{2,j}}{2}}}. \tag{2.6}$$

If we use the transformation $y_j = x_j/(1+x_j) \quad (0 \le y_j \le 1)$, then equation (2.6) becomes the product of beta distributions of the first kind with parameters $m_{1,j}/2$ and $m_{2,j}/2$:

$$f(y_1,\dots,y_p) \approx \prod_{j=1}^{p} \frac{y_j^{\frac{m_{1,j}}{2}-1}(1-y_j)^{\frac{m_{2,j}}{2}-1}}{B(m_{1,j}/2, m_{2,j}/2)}.$$

Another transformation from x_j to $z_j = m_{2,j}\ x_j/m_{1,j}$ gives the joint pdf of $z_1,\dots,z_p$ as

$$f(z_1,\dots,z_p) \approx \prod_{j=1}^{p} \frac{(\frac{m_{1,j}}{m_{2,j}})^{\frac{m_{1,j}}{2}}}{B(m_{1,j}/2, m_{2,j}/2)} \frac{z_j^{\frac{m_{1,j}}{2}-1}}{(1+\frac{m_{1,j}}{m_{2,j}}z_j)^{\frac{m_{1,j}+m_{2,j}}{2}}},$$

and the proof of Theorem 2.5 is completed. □

Theorem 2.6 (Approximation by a product of F distributions)**.** *Let $Z_1, \dots, Z_p$ be independent random variables, where each Z_k follows the F distribution with $m_{1,j}, m_{2,j}$ degrees of freedom, where $m_{1,j}$ and $m_{2,j}$ are as defined in Theorem 2.5. Furthermore, let $\overline{V}_{(k)}$ and $\underline{V}_{(k)}$ be defined as*

$$\begin{cases} \overline{V}_{(k)} = \{\delta_1^{-1}\lambda_1 Z_1, \delta_2^{-1}\lambda_2 Z_2, \dots, \delta_k^{-1}\lambda_k Z_k\} \\ \underline{V}_{(k)} = \{\delta_k^{-1}\lambda_k Z_k, \delta_{k+1}^{-1}\lambda_{k+1} Z_{k+1}, \dots, \delta_p^{-1}\lambda_p Z_p\}, \end{cases}$$

where $\delta_k = m_{2,k}/m_{1,k}$, $\overline{V}_{(0)} = \{\infty\}$, and $\underline{V}_{(p+1)} = \{0\}$. If $\rho_2 \to 0$, we have the following two approximations for the distribution of the kth eigenvalue of $W_1W_2^{-1}$.

1. $\ell_k \approx \min\{\min \overline{V}_{(k-1)}, \max \underline{V}_{(k)}\}$,

$$\Pr[\ell_k > x] \approx \Pr[\min\{\min \overline{V}_{(k-1)}, \max \underline{V}_{(k)}\} > x] = \prod_{j=1}^{k-1}\left(1 - F_{m_{1,j},\, m_{2,j}}\left(\delta_j \frac{x}{\lambda_j}\right)\right)\left(1 - \prod_{j=k}^{p} F_{m_{1,j},\, m_{2,j}}\left(\delta_j \frac{x}{\lambda_j}\right)\right).$$

2. $\ell_k \approx \max\{\min \overline{V}_{(k)}, \max \underline{V}_{(k+1)}\}$,

$$\Pr[\ell_k < x] \approx \Pr[\max\{\min \overline{V}_{(k)}, \max \underline{V}_{(k+1)}\} < x] = \left\{1 - \prod_{j=1}^{k}\left(1 - F_{m_{1,j},\, m_{2,j}}\left(\delta_j \frac{x}{\lambda_j}\right)\right)\right\} \prod_{j=k+1}^{p} F_{m_{1,j},\, m_{2,j}}\left(\delta_j \frac{x}{\lambda_j}\right).$$

In both statements, $F_{j,k}(x)$ denotes the cumulative distribution function of an F distribution with j, k degrees of freedom.

Proof. If $\rho_2 \to 0$, $z_1, \ldots, z_p$ are approximately independent, and each z_j is distributed as the F distribution with parameters $m_{1,j}$ and $m_{2,j}$. From

$$\max \underline{V}_{(k)} \leq \frac{\lambda_k Z_k}{\delta_k} \leq \min \overline{V}_{(k-1)} \quad \text{and} \quad \ell_k \approx \frac{\lambda_k Z_k}{\delta_k} \in \underline{V}_{(k)},$$

we have $\ell_k \approx \min\{\max \underline{V}_{(k)}, \min \overline{V}_{(k-1)}\}$. Hence, the upper probability of ℓ_k can be expressed as

$$\begin{aligned}
\Pr(\ell_k > x) &\approx \Pr(\min\{\max \underline{V}_{(k)}, \min \overline{V}_{(k-1)}\} > x) \\
&= \Pr(\max \underline{V}_{(k)} > x) \Pr(\min \overline{V}_{(k-1)} > x) \\
&= \Big(1 - \Pr(\max \underline{V}_{(k)} < x)\Big) \times \prod_{j=1}^{k-1} \Pr\left(\frac{\lambda_j Z_j}{\delta_j} > x\right) \\
&= \left\{1 - \prod_{j=1}^{k-1} \Pr\left(\frac{\lambda_j Z_j}{\delta_j} < x\right)\right\} \times \prod_{j=1}^{k-1} \Pr\left(\frac{\lambda_j Z_j}{\delta_j} > x\right) \\
&= \left\{1 - \prod_{j=k}^{p} F_{m_{1,j}, m_{2,j}}\left(\delta_j \frac{x}{\lambda_j}\right)\right\} \times \prod_{j=1}^{k-1} \left\{1 - F_{m_{1,j}, m_{2,j}}\left(\delta_j \frac{x}{\lambda_j}\right)\right\}.
\end{aligned}$$

In a similar manner, we have

$$\max \underline{V}_{(k+1)} \leq \frac{\lambda_k Z_k}{\delta_k} \leq \min \overline{V}_{(k)}, \quad \ell_k \approx \frac{\lambda_k Z_k}{\delta_k} \in \overline{V}_{(k)},$$

and $\ell_k \approx \max\{\max \underline{V}_{(k+1)}, \min \overline{V}_{(k)}\}$. Hence, the probability of ℓ_k is also given by

$$\begin{aligned}
\Pr[\ell_k < x] &\approx \Pr[\max\{\min \overline{V}_{(k)}, \max \underline{V}_{(k+1)}\} < x] \\
&= \left\{1 - \prod_{j=1}^{k} \left(1 - F_{m_{1,j}, m_{2,j}}\left(\delta_j \frac{x}{\lambda_j}\right)\right)\right\} \prod_{j=k+1}^{p} F_{m_{1,j}, m_{2,j}}\left(\delta_j \frac{x}{\lambda_j}\right).
\end{aligned}$$

□

We consider an approximate distribution for the extreme eigenvalues defined in Theorem 2.6.

Corollary 2.7. *If $k = 1$ in Theorem 2.6, the approximate distribution for the largest eigenvalue ℓ_1 is given by $\ell_1 \approx \max \underline{V}_{(1)}$, and*

$$\Pr[\ell_1 < x] \approx \prod_{j=1}^{k} F_{m_{1,j}, m_{2,j}}\left(\delta_j \frac{x}{\lambda_j}\right).$$

In a similar manner, for the smallest eigenvalue ℓ_p, we have $\ell_p \approx \min \overline{V}_{(p)}$, and

$$\Pr[\ell_p > x] \approx 1 - \prod_{j=1}^{k} \left(1 - F_{m_{1,j},m_{2,j}}\left(\delta_j \frac{x}{\lambda_j}\right)\right).$$

Proof. From statement 1 in Theorem 2.6, we have

$$\ell_1 \approx \min\{\min \overline{V}_{(0)}, \max \underline{V}_{(1)}\} = \max \underline{V}_{(1)},$$

and from statement 2 in Theorem 2.6, we have

$$\ell_1 \approx \max\{\min \overline{V}_{(1)}, \max \underline{V}_{(2)}\} = \max \underline{V}_{(1)}.$$

Hence, both statements 1 and 2 in Theorem 2.6 yield the same equation. In a similar way, statements 1 and 2 in Theorem 2.6 for the smallest eigenvalue ℓ_p yield

$$\ell_p \approx \min\{\min \overline{V}_{(p-1)}, \max \underline{V}_{(p)}\} = \min \overline{V}_{(p)}$$

and

$$\ell_p \approx \max\{\min \overline{V}_{(p)}, \max \underline{V}_{(p+1)}\} = \min \overline{V}_{(p)},$$

respectively. □

§3. Numerical experiments

We perform a simulation study to evaluate the approximate accuracy of the results discussed above. We use a Monte Carlo simulation with 10^6 runs. For the jth eigenvalues of $W_1 W_2^{-1}$, we let $G_j^{(0)}(x)$ be the asymptotic distribution up to order $O(n^{-3/2})$ in Corollary 3.1 of Sugiura (1976). That is, we define $G_j^{(0)}(x)$ such that

$$\Pr\left(\left(\frac{n_1 n_2}{2n}\right)^{1/2} \frac{\ell_j - \lambda_j}{\lambda_j} < \left(\frac{n_1 n_2}{2n}\right)^{1/2} \frac{x - \lambda_j}{\lambda_j}\right) = G_j^{(0)}(x) + O(n^{-3/2})$$

which means $\Pr(\ell_j < x) = G_j^{(0)}(x) + O(n^{-3/2})$.

We also set $G_j^{(1)}(x) = F_{m_{1,j},m_{2,j}}(m_{2,j}/m_{1,j}x)$ as in Theorem 2.5 and we set $G_j^{(2)}(x) = \Pr[\max\{\min \overline{V}_{(k)}, \max \underline{V}_{(k+1)}\} < x]$ as in Theorem 2.6. In the simulation study, the matrix $\Sigma_1 \Sigma_2^{-1}$ is assumed to be a diagonal matrix without loss of generality because each eigenvalue distribution is invariant under the action of any orthogonal matrix. Therefore, we use $\Sigma_2 = I_p$ and $\Sigma_1 \Sigma_2^{-1} = \Sigma_1 = \mathrm{diag}(\lambda_1, \ldots, \lambda_p)$. In order to compare the probability $\Pr(a < X < b) = 0.95$ for some random variable X, the values of the percentiles (a

and b, $\Pr(X < a) = \Pr(X > b) = 0.025$) are obtained from $G_j^{(0)}(x)$, $G_j^{(1)}(x)$, and $G_j^{(2)}(x)$.

Tables 1 through 6 show the empirical probabilities based on percentiles calculated by $G_j^{(0)}(x)$, $G_j^{(1)}(x)$, and $G_j^{(2)}(x)$, respectively. Tables 1 through 3 show the results for the case $n_1 = n_2$, while in Tables 4 through 6, $n_1 \neq n_2$.

In Table 1 we present the results of the simulation in which $p = 3$ for various values of $\Sigma_1\Sigma_2^{-1} = \mathrm{diag}(\lambda_1, \lambda_2, \lambda_3)$. The approximations $G_1^{(0)}$ and $G_2^{(0)}$ are more sensitive than the F-type approximations when the values of λ_1, λ_2, and λ_3 are close together. These probabilities are sometimes less than 0.9. On the other hand, for the smallest eigenvalue, the approximation of $G_3^{(0)}$ is the most accurate. The approximation of $G_j^{(2)}$ tends to be better than that of $G_j^{(1)}$ for $j = 1, 2, 3$.

In Tables 2 and 3, we present the results for higher-dimensional cases for $p = 10, 20$ than those of Table 1. We see that for larger and smaller eigenvalues, $G_j^{(2)}$ is more accurate, whereas $G_j^{(0)}$ is more accurate for eigenvalues of moderate size.

In the remaining tables, we present the results of simulations when $n_1 \neq n_2$ for $p = 5, 10$, and 20. In Table 4, when $(n_1, n_2) = (20, 50)$, we find that $G_j^{(2)}$ for $j = 1, 2, 3$, and 5 is the most precise of the three approximations. When $(n_1, n_2) = (50, 20)$, $G_j^{(2)}$ for $j = 1, 2, 3$ is the best of the three, and in the other cases, all three approximations have almost the same precision. In Table 5, when $(n_1, n_2) = (10, 50)$ and $(n_1, n_2) = (50, 20)$, we see that $G_j^{(2)}$ is more precise for the larger or smaller eigenvalues. The tendencies seen in Table 6 are similar to those seen in Table 5.

§4. Concluding remarks

In this paper, we consider the approximate distribution of the eigenvalues of a ratio of Wishart matrices, where each population has a single eigenvalue. Sugiura (1976) and Butler and Wood (2005) discussed the case of multiple eigenvalues, but we leave this for future work.

The authors would like to thank the editor and the anonymous reviewer for improving this paper. This research was supported in part by the Japan Society for the Promotion Science, Grant-in-Aid for Scientific Research (C), Nos. 25330033 and 26330053.

Table 1: Approximate distribution of the jth eigenvalue when $p = 3$

	(n_1, n_2)	$(50, 50)$			$(100, 100)$		
j	$\Sigma_1\Sigma_2^{-1}$	AsN	F	F-prod	AsN	F	F-prod
1	$\mathrm{diag}(6, 5, 1)$	0.836	0.943	0.941	0.902	0.949	0.943
1	$\mathrm{diag}(6, 4, 1)$	0.934	0.958	0.949	0.949	0.958	0.952
1	$\mathrm{diag}(6, 3, 1)$	0.944	0.959	0.954	0.949	0.955	0.953
1	$\mathrm{diag}(6, 2, 1)$	0.944	0.955	0.954	0.948	0.952	0.952
2	$\mathrm{diag}(6, 5, 4)$	0.660	0.993	0.953	0.838	0.991	0.952
2	$\mathrm{diag}(7, 5, 4)$	0.866	0.991	0.954	0.932	0.987	0.956
2	$\mathrm{diag}(8, 5, 4)$	0.887	0.986	0.955	0.937	0.980	0.957
2	$\mathrm{diag}(6, 5, 3)$	0.775	0.984	0.952	0.882	0.976	0.955
2	$\mathrm{diag}(6, 5, 2)$	0.785	0.965	0.952	0.881	0.961	0.950
2	$\mathrm{diag}(6, 5, 1)$	0.788	0.952	0.946	0.883	0.954	0.946
3	$\mathrm{diag}(6, 2, 1)$	0.955	0.959	0.954	0.952	0.955	0.954
3	$\mathrm{diag}(6, 3, 1)$	0.954	0.955	0.954	0.951	0.952	0.952
3	$\mathrm{diag}(6, 4, 1)$	0.953	0.953	0.953	0.951	0.951	0.951
3	$\mathrm{diag}(6, 5, 1)$	0.953	0.952	0.952	0.951	0.951	0.951

Table 2: Approximate distribution of the jth eigenvalue when $p = 10$ and $\Sigma_1\Sigma_2^{-1} = \mathrm{diag}(2^9, 2^8, \ldots, 2^0)$

(n_1, n_2)	$(50, 50)$			$(100, 100)$		
j	AsN	F	F-prod	AsN	F	F-prod
1	0.898	0.958	0.954	0.938	0.955	0.953
2	0.938	0.978	0.964	0.947	0.964	0.961
3	0.952	0.981	0.968	0.951	0.966	0.963
4	0.959	0.981	0.968	0.953	0.966	0.963
5	0.962	0.981	0.968	0.955	0.967	0.963
6	0.965	0.981	0.969	0.956	0.967	0.963
7	0.965	0.981	0.968	0.956	0.966	0.963
8	0.961	0.981	0.967	0.954	0.966	0.962
9	0.961	0.978	0.964	0.954	0.965	0.961
10	0.966	0.958	0.954	0.954	0.954	0.953

Note. AsN: $G_j^{(0)}(x)$ F: $G_j^{(1)}(x)$ F-prod: $G_j^{(2)}(x)$

Table 3: Approximate distribution of the jth eigenvalue when $p = 20$ and $\Sigma_1\Sigma_2^{-1} = \text{diag}(2^{19}, 2^{18}, \ldots, 2^0)$

(n_1, n_2)	$(50, 50)$			$(100, 100)$		
j	AsN	F	F-prod	AsN	F	F-prod
1	0.509	0.957	0.953	0.876	0.955	0.954
2	0.684	0.980	0.964	0.917	0.965	0.961
3	0.807	0.983	0.968	0.934	0.967	0.963
4	0.876	0.984	0.969	0.938	0.967	0.964
5	0.910	0.984	0.970	0.942	0.968	0.964
6	0.927	0.984	0.970	0.947	0.968	0.964
7	0.942	0.984	0.970	0.950	0.968	0.964
8	0.953	0.984	0.970	0.953	0.968	0.964
9	0.960	0.984	0.970	0.954	0.967	0.964
10	0.964	0.984	0.970	0.955	0.967	0.964
11	0.967	0.984	0.970	0.957	0.968	0.964
12	0.967	0.984	0.967	0.957	0.968	0.964
13	0.966	0.984	0.970	0.957	0.968	0.964
14	0.966	0.984	0.970	0.956	0.968	0.964
15	0.967	0.984	0.970	0.956	0.968	0.964
16	0.971	0.984	0.970	0.957	0.968	0.964
17	0.976	0.984	0.969	0.960	0.968	0.964
18	0.980	0.983	0.968	0.963	0.967	0.963
19	0.916	0.980	0.964	0.967	0.965	0.961
20	0.875	0.957	0.953	0.969	0.955	0.954

Table 4: Approximate distribution of the jth eigenvalue when $p = 5$ and $\Sigma_1\Sigma_2^{-1} = \text{diag}(2^4, 2^3, \ldots, 2^0)$

(n_1, n_2)	$(20, 50)$			$(50, 20)$		
j	AsN	F	F-prod	AsN	F	F-prod
1	0.937	0.960	0.954	0.836	0.956	0.951
2	0.964	0.985	0.963	0.935	0.986	0.962
3	0.965	0.985	0.963	0.956	0.988	0.967
4	0.958	0.986	0.962	0.962	0.985	0.963
5	0.969	0.956	0.951	0.953	0.960	0.954

Note. AsN: $G_j^{(0)}(x)$ F: $G_j^{(1)}(x)$ F-prod: $G_j^{(2)}(x)$

Table 5: Approximate distribution of the jth eigenvalue when $p = 10$ and $\Sigma_1\Sigma_2^{-1} = \mathrm{diag}(2^9, 2^8, \ldots, 2^0)$

(n_1, n_2)	(10, 50)			(20, 50)		
j	AsN	F	F-prod	AsN	F	F-prod
1	0.908	0.955	0.952	0.903	0.959	0.954
2	0.961	0.990	0.960	0.954	0.985	0.963
3	0.968	0.996	0.967	0.966	0.989	0.968
4	0.954	0.995	0.969	0.971	0.991	0.971
5	0.972	0.996	0.971	0.965	0.992	0.971
6	0.984	0.998	0.969	0.965	0.992	0.972
7	0.981	0.998	0.973	0.973	0.993	0.971
8	0.983	0.998	0.975	0.980	0.992	0.968
9	0.991	0.997	0.975	0.956	0.990	0.960
10	0.998	0.958	0.940	0.919	0.953	0.948

Table 6: Approximate distribution of the jth eigenvalue ($j = 1, \ldots, 20$) when $p = 20$ and $\Sigma_1\Sigma_2^{-1} = \mathrm{diag}(2^{19}, 2^{18}, \ldots, 2^0)$

(n_1, n_2)	(20, 50)			(30, 50)		
j	AsN	F	F-prod	AsN	F	F-prod
1	0.660	0.960	0.954	0.581	0.958	0.953
2	0.856	0.988	0.962	0.803	0.984	0.963
3	0.910	0.990	0.965	0.883	0.988	0.969
4	0.947	0.993	0.972	0.918	0.988	0.970
5	0.961	0.993	0.971	0.942	0.989	0.971
6	0.970	0.993	0.972	0.956	0.989	0.971
7	0.976	0.994	0.975	0.964	0.990	0.972
8	0.972	0.994	0.974	0.969	0.990	0.972
9	0.972	0.995	0.973	0.971	0.990	0.972
10	0.979	0.995	0.977	0.971	0.990	0.972
11	0.985	0.997	0.976	0.970	0.991	0.972
12	0.988	0.997	0.974	0.972	0.991	0.974
13	0.970	0.996	0.971	0.977	0.991	0.972
14	0.972	0.997	0.971	0.982	0.992	0.972
15	0.982	0.997	0.972	0.969	0.992	0.972
16	0.990	0.997	0.974	0.947	0.992	0.972
17	0.996	0.998	0.972	0.944	0.993	0.971
18	0.999	0.998	0.974	0.948	0.992	0.969
19	0.999	0.997	0.972	0.954	0.990	0.961
20	1.000	0.954	0.935	0.949	0.953	0.948

Note. AsN: $G_j^{(0)}(x)$ F: $G_j^{(1)}(x)$ F-prod: $G_j^{(2)}(x)$

References

[1] Butler, R. W. and Wood, A. T. A. (2005), Laplace Approximations to Hypergeometric Functions of Two Matrix Arguments, *Journal of Multivariate Analysis*, **94**, 1–18.

[2] Chikuse, Y. (1977), Asymptotic Expansion for the Joint and Marginal Distributions of the latent roots of $S_1S_2^{-1}$, *Annals of the Institute of Statistical Mathematics*, **29**, 221–233.

[3] James, A. T. (1964), Distributions of Matrix Variates and Latent Roots Derived from Normal Samples, *The Annals of Mathematical Statistics*, **35**, 475–501.

[4] Johnstone, I. M. (2001), On the Distribution of the Largest Eigenvalue in Principal Components Analysis, *The Annals of Statistics*, **29**, 295–327.

[5] Johnstone, I. M. (2008), Multivariate Analysis and Jacobi Ensembles: Largest Eigenvalue, Tracy–Widom Limits and Rates of Convergence. *The Annals of Statistics*, **36**, 2638–2716.

[6] Johnstone, I. M. (2009), Approximate Null Distribution of the Largest Root in Multivariate Analysis, *The Annals of Applied Statistics*, **3**, 1616–1633.

[7] Kato, H. and Hashiguchi, H. (2014), Chi-Square Approximation for Eigenvalue Distributions and Confidential Interval Construction on Population Eigenvalues, *Journal of the Japanese Society of Computational Statistics*, **27**, 11–28 (in Japanese).

[8] Khatri, C. G. (1967), Some Distribution Problems Connected with the Characteristic Roots of $S_1S_2^{-1}$, *The Annals of Mathematical Statistics*, **38**, 944–948.

[9] 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.

[10] Li, H. C., Pillai, K. C. S. and Chang, T. C. (1970), Asymptotic Expansions for Distributions of the Roots of Two Matrices from Classical and Complex Gaussian Populations, *The Annals of Mathematical Statistics*, **41**, 1541–1556.

[11] Muirhead, R. J. and Chikuse, Y. (1975), Asymptotic Expansions for the Joint and Marginal Distributions of the Latent Roots of the Covariance Matrix, *The Annals of Statistics*, **3**, 1011–1017.

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

[13] Sugiura, N. (1973), Derivatives of the Characteristic Root of a Symmetric or a Hermitian Matrix with Two Applications in Multivariate Analysis, *Communications in Statistics*, **1**, 393–417.

[14] Sugiura, N. (1976), Asymptotic Expansion of the Distributions of the Latent Roots and the Latent Vector of the Wishart and Multivariate F Matrices, *Journal of Multivariate Analysis*, **6**, 500–525.

[15] Sugiyama, T. (1972), Approximation for the Distribution of the Largest Latent Root of a Wishart Matrix, *The Australian Journal of Statistics*, **14**, 17–24.

[16] Takemura, A. and Sheena, Y. (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, *Journal of Multivariate Analysis*, **94**, 271–299.

[17] Venables, W. (1973), Computation of the Null Distribution of the Largest or Smallest Latent Roots of a Beta Matrix, *Journal of Multivariate Analysis*, **3**, 125–131.

Shusuke Matsubara
Graduate School of Mathematical Information Science, Tokyo University of Science
1-3 Kagurazaka, Shinjuku-ku, Tokyo, 162-8601, Japan
E-mail: `1414624@ed.tus.ac.jp`

Hiroki Hashiguchi
Department of Mathematical Information Science, Tokyo University of Science
1-3 Kagurazaka, Shinjuku-ku, Tokyo, 162-8601, Japan
E-mail: `hiro@rs.tus.ac.jp`