

Article

# Chi-Square Approximation for the Distribution of Individual Eigenvalues of a Singular Wishart Matrix

Koki Shimizu \*  and Hiroki Hashiguchi 

Department of Applied Mathematics, Tokyo University of Science, 1-3 Kagurazaka, Shinjuku-ku, Tokyo 162-8601, Japan; hiro@rs.tus.ac.jp

\* Correspondence: k-shimizu@rs.tus.ac.jp

**Abstract:** This paper discusses the approximate distributions of eigenvalues of a singular Wishart matrix. We give the approximate joint density of eigenvalues by Laplace approximation for the hypergeometric functions of matrix arguments. Furthermore, we show that the distribution of each eigenvalue can be approximated by the chi-square distribution with varying degrees of freedom when the population eigenvalues are infinitely dispersed. The derived result is applied to testing the equality of eigenvalues in two populations.

**Keywords:** hypergeometric functions; laplace approximation; spiked covariance model

**MSC:** 62E15; 62H10



**Citation:** Shimizu, K.; Hashiguchi, H. Chi-Square Approximation for the Distribution of Individual Eigenvalues of a Singular Wishart Matrix. *Mathematics* **2024**, *12*, 921. <https://doi.org/10.3390/math12060921>

Academic Editor: Debashis Paul

Received: 27 February 2024

Revised: 15 March 2024

Accepted: 19 March 2024

Published: 20 March 2024

**Correction Statement:** This article has been republished with a minor change. The change does not affect the scientific content of the article and further details are available within the backmatter of the website version of this article.



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The Wishart matrix is a symmetric random matrix defined by the sum of squares and cross-products of samples from a multivariate normal distribution. It becomes non-singular when the dimension is smaller than or equal to the number of observations; otherwise, it is singular. The distributions for the Wishart matrix and its eigenvalues have been used in many areas of science and technology, including multivariate analysis, Bayesian statistics, random matrix theory, and wireless communications. Some exact distributions of eigenvalues for a Wishart matrix are represented by the hypergeometric functions of matrix arguments. James [1] classified multivariate statistics problems into five categories based on hypergeometric functions. However, the convergence of these functions is slow, and their numerical computation is cumbersome when sample sizes or dimensions are large. Consequently, the derivation of approximate distributions of eigenvalues has received a great deal of attention. Sugiyama [2] derived the approximate distribution for the largest eigenvalue through the integral representation of the confluent hypergeometric function. Sugiura [3] showed that the asymptotic distribution of the individual eigenvalues is expressed by a normal distribution for a large sample size. The chi-square approximation was discussed when the population eigenvalues are infinitely dispersed in Kato and Hashiguchi [4] and Takemura and Sheena [5]. Approximations for hypergeometric functions have been developed and applied to the multivariate distribution theory in Butler and Wood [6–8]. Butler and Wood [6] provided the Laplace approximation for the hypergeometric functions of a single matrix argument. The numerical accuracies for that approximation were shown in the computation of noncentral moments of Wilk's lambda statistic and the likelihood ratio statistic for testing block independence. This approximation was extended to the case of two matrix arguments in Butler and Wood [7]. All the results addressed above were carried out for eigenvalue distributions for a non-singular Wishart matrix.

Recently, the distribution of eigenvalues for the non-singular case has been extended to the singular case; see Shimizu and Hashiguchi [9] and Shinozaki et al. [10]. Shimizu and Hashiguchi [9] showed the exact distribution of the largest eigenvalue for a singular case is

represented in terms of the confluent hypergeometric function as well as the non-singular case. The generalized representation for the non-singular and singular cases under the elliptical model was provided by Shinozaki et al. [10].

This paper is organized as follows. In Section 2, we apply the Laplace approximation introduced by Butler and Wood [7] to the joint density of eigenvalues of a singular Wishart matrix. Furthermore, we show that the approximation for the distribution of the individual eigenvalues can be expressed by the chi-square distribution with varying degrees of freedom when the population covariance matrix has spiked eigenvalues. Section 3 discusses the equality of the individual eigenvalues in two populations. Finally, we evaluate the precision of the chi-square approximation by comparing it to the empirical distribution through Monte Carlo simulation in Section 4.

### 2. Approximate Distributions of Eigenvalues of a Singular Wishart Matrix

Suppose that an  $m \times n$  real Gaussian random matrix  $X$  is distributed as  $X \sim N_{m,n}(O, \Sigma \otimes I_n)$ , where  $O$  is the  $m \times n$  zero matrix,  $\Sigma$  is a  $m \times m$  positive symmetric matrix, and  $\otimes$  is the Kronecker product. This means that the column vectors of  $X$  are independently and identically distributed (i.i.d.) from  $N_m(0, \Sigma)$  with sample size  $n$ , where  $0$  is the  $m$ -dimensional zero vector. The eigenvalues of  $\Sigma$  are denoted by  $\lambda_1, \lambda_2, \dots, \lambda_m$ , and  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m > 0$ . Subsequently, we define the singular Wishart matrix as  $W = XX^T$ , where  $m > n$  and its distribution is denoted by  $W(n, \Sigma)$ . The spectral decomposition of  $W$  is represented as  $W = H_1 L_1 H_1^T$ , where  $L_1 = \text{diag}(\ell_1, \dots, \ell_n)$  with  $\ell_1 > \ell_2 > \dots > \ell_n > 0$ , and the  $m \times n$  matrix  $H_1$  is satisfied by  $H_1^T H_1 = I_n$ . The set of all  $m \times n$  matrices  $H_1$  with orthonormal columns is called the Stiefel manifold, denoted by  $V_{n,m} = \{H_1 \mid H_1 H_1^T = I_n\}$ , where  $m \geq n$ . The volume of  $V_{n,m}$  is represented by

$$\text{Vol}(V_{n,m}) = \int_{H_1 \in V_{n,m}} (H_1^T dH_1) = \frac{2^n \pi^{mn/2}}{\Gamma_n(m/2)}.$$

For the definition of the above exterior product  $(H_1^T dH_1)$ , see page 63 of Muirhead [11]. If  $m = n$ , Stiefel manifold  $V_{m,m}$  coincides with the orthogonal groups  $O(m)$ . Uhlig [12] gave the density of  $W$  as

$$f(W) = \frac{\pi^{(-mn+n^2)/2}}{2^{mn/2} \Gamma_n(n/2) |\Sigma|^{n/2}} |L_1|^{(n-m-1)/2} \text{etr}(-\Sigma^{-1}W/2),$$

where  $\Gamma_m(a) = \pi^{m(m-1)/4} \prod_{i=1}^m \Gamma\{a - (i - 1)/2\}$  and  $\text{etr}(\cdot) = \exp(\text{tr}(\cdot))$ . Srivastava [13] represented the joint density of eigenvalues of  $W$  in a form that includes an integral over the Stiefel manifold;

$$f(\ell_1, \dots, \ell_n) = \frac{2^{-nm/2} \pi^{n^2/2}}{|\Sigma|^{n/2} \Gamma_n(n/2) \Gamma_n(m/2)} \prod_{i=1}^n \ell_i^{(m-n-1)/2} \prod_{i < j}^n (\ell_i - \ell_j) \times \int_{H_1 \in V_{n,m}} \text{etr}\left(-\frac{1}{2} \Sigma^{-1} H_1 L_1 H_1^T\right) (dH_1), \tag{1}$$

where  $(dH_1) = \frac{(H_1^T dH_1)}{\text{Vol}(V_{n,m})}$  and  $\int_{H_1 \in V_{n,m}} (dH_1) = 1$ .

The above integral over the Stiefel manifold was evaluated by Shimizu and Hashiguchi [9] as the hypergeometric functions of the matrix arguments. We approximate (1) by Laplace approximation for the hypergeometric functions of two matrix arguments provided by Butler and Wood [7].

For a positive integer  $k$ , let  $\kappa = (\kappa_1, \kappa_2, \dots, \kappa_m)$  denote a partition of  $k$  with  $\kappa_1 \geq \kappa_2 \geq \dots \geq \kappa_m \geq 0$  and  $\kappa_1 + \dots + \kappa_m = k$ . The set of all partitions with less than or equal to  $m$  is denoted by  $P_m^k = \{\kappa = (\kappa_1, \dots, \kappa_m) \mid \kappa_1 + \dots + \kappa_m = k, \kappa_1 \geq \kappa_2 \geq \dots \geq \kappa_m \geq 0\}$ . The Pochhammer symbol for a partition  $\kappa$  is defined as  $(\alpha)_\kappa = \prod_{i=1}^m \{\alpha - (i - 1)/2\}_{\kappa_i}$ , where  $(\alpha)_k = \alpha(\alpha + 1) \dots (\alpha + k - 1)$  and  $(\alpha)_0 = 1$ . For integers,  $p, q \geq 0$  and  $m \times m$

real symmetric matrices  $A$  and  $B$ , we define the hypergeometric function of two matrix arguments as

$${}_pF_q^{(m)}(\alpha; \beta; A, B) = \sum_{k=0}^{\infty} \sum_{\kappa \in F_m^k} \frac{(\alpha_1)_\kappa \cdots (\alpha_p)_\kappa}{(\beta_1)_\kappa \cdots (\beta_q)_\kappa} \frac{C_\kappa(A)C_\kappa(B)}{k!C_\kappa(I_m)}, \tag{2}$$

where  $\alpha = (\alpha_1, \dots, \alpha_p)^\top$ ,  $\beta = (\beta_1, \dots, \beta_q)^\top$  and  $C_\kappa(A)$  is the zonal polynomial indexed by  $\kappa$  with the symmetric matrix  $A$ ; see the details provided in Chapter 7 of Muirhead [11]. The hypergeometric functions with a single matrix are defined as

$${}_pF_q(\alpha; \beta; A) = {}_pF_q(\alpha; \beta; A, I_m). \tag{3}$$

The special cases  ${}_1F_1$  and  ${}_2F_1$  of (3) are called the confluent and Gauss hypergeometric functions, respectively. Butler and Wood [6], proposed a Laplace approximation of  ${}_1F_1$  and  ${}_2F_1$  through their integral expressions. They showed that the accuracy of that approximation is greater than the previous results. This approximation was extended to the complex case in Butler and Wood [8]. The important property of (2) is the integral representation over the orthogonal group.

$${}_pF_q^{(m)}(\alpha; \beta; A, B) = \int_{H \in O(m)} {}_pF_q(\alpha; \beta; AHBH^\top)(dH), \tag{4}$$

where  $(dH)$  is the invariant measure on the  $m \times m$  orthogonal group  $O(m)$ . Integral representations (4) are a useful tool for obtaining an approximation of  ${}_pF_q^{(m)}$ . Asymptotic expansions of  ${}_0F_0^{(m)}$  are given in Anderson [14] when both two positive definite matrix arguments are widely spaced. Constantine and Muirhead [15] gave the asymptotic behavior of  ${}_0F_0^{(m)}$  when the population eigenvalues are multiple. From the integral expression (4), Butler and Wood [7] provided Laplace approximations for  ${}_pF_q^{(m)}$ .

**Lemma 1.** Let the two diagonal matrices be  $A = \text{diag}(a_1, \dots, a_m)$  and  $B = \text{diag}(b_1, \dots, b_1, b_2, \dots, b_r, \dots, b_r)$ , where  $a_1 > a_2 > \dots > a_m > 0$ ,  $b_1 > b_2 > \dots > b_r \geq 0$ , and  $b_j$  have multiplicity  $m_j$ , in which  $m = \sum_{j=1}^r m_j$ . Let  $\Omega(m_1, \dots, m_r) = \text{Vol}(O(m))^{-1} \prod_{j=1}^r \text{Vol}(O(m_j))$ . Then the Laplace approximation of  ${}_pF_q^{(m)}$  is given as

$${}_p\hat{F}_q^{(m)}(\alpha; \beta; A, B) = (2\pi)^{\frac{s}{2}} \Omega(m_1, \dots, m_r) J^{-\frac{1}{2}} {}_pF_q(\alpha; \beta; AB),$$

where  $s = \sum_{i=1}^{r-1} \sum_{j=i+1}^r m_i m_j$  and Hessian  $J$  is defined in Butler and Wood [7].

Shimizu and Hashiguchi [9] showed the following relationship

$$\int_{H_1 \in V_{n,m}} {}_pF_q(AH_1B_1H_1^\top)(dH_1) = \int_{H \in O(m)} {}_pF_q(AHBH^\top)(dH) \tag{5}$$

for an  $m \times m$  matrix  $B = \begin{pmatrix} B_1 & O \\ O & O \end{pmatrix}$ , where  $B_1$  is an  $n \times n$  symmetric matrix and  $O$  is the zero matrix. From (5), the joint density (1) can be rewritten by

$$\begin{aligned} f(\ell_1, \dots, \ell_n) &\propto \int_{H \in O(m)} \text{etr} \left( -\frac{1}{2} \Sigma^{-1} H L H^\top \right) (dH) \\ &= {}_0F_0^{(m)} \left( -\frac{1}{2} \Sigma^{-1}, L \right), \end{aligned} \tag{6}$$

where  $L = \text{diag}(\ell_1, \dots, \ell_n, 0, \dots, 0)$  is the  $m \times m$  matrix and the symbol “ $\propto$ ” means that a constant required for scaling is removed. Applying Laplace’s method to the above joint density, we have an approximation for the joint density of eigenvalues.

**Proposition 1.** *The joint density of eigenvalues of a singular Wishart matrix by Laplace approximation is expressed by*

$$\frac{\pi^{n(n-m)/2}}{2^{nm/2} |\Sigma|^{n/2} \Gamma_n(n/2)} \prod_{i=1}^n \ell_i^{(m-n-1)/2} \prod_{i<j}^n (\ell_i - \ell_j) \exp\left(-\frac{1}{2} \sum_{i=1}^n \frac{\ell_i}{\lambda_i}\right) \prod_{i<j}^n \left(\frac{2\pi}{c_{ij}}\right)^{1/2} \prod_{i=1}^n \prod_{j=n+1}^m \left(\frac{2\pi}{d_{ij}}\right)^{1/2}, \tag{7}$$

where  $c_{ij} = \frac{(\ell_i - \ell_j)(\lambda_i - \lambda_j)}{\lambda_i \lambda_j}$ ,  $d_{ij} = \frac{\ell_i(\lambda_i - \lambda_j)}{\lambda_i \lambda_j}$ .

**Proof.** Applying Lemma 1 to the hypergeometric functions in (6), the integral over the Stiefel manifold in (1) is approximated by

$$\frac{2^n}{\text{Vol}(V_{n,m})} \exp\left(-\frac{1}{2} \sum_{i=1}^n \frac{\ell_i}{\lambda_i}\right) \prod_{i<j}^n \left(\frac{2\pi}{c_{ij}}\right)^{1/2} \prod_{i=1}^n \prod_{j=n+1}^m \left(\frac{2\pi}{d_{ij}}\right)^{1/2}. \tag{8}$$

Substituting (8) to (1), we have the desired result.  $\square$

In order to derive the approximate distributions of individual eigenvalues, we define the spiked covariance model  $\rho_k$  that implies the first  $k$ -th eigenvalues of  $\Sigma > 0$  are infinitely dispersed, namely

$$\rho_k = \max\left(\frac{\lambda_2}{\lambda_1}, \frac{\lambda_3}{\lambda_2}, \dots, \frac{\lambda_{k+1}}{\lambda_k}\right) \rightarrow 0, \tag{9}$$

where  $k \leq n$ . Under the condition of (9) when  $k = n$ , Takemura and Sheena [5] proved that the distribution of individual eigenvalues for a non-singular Wishart matrix is approximated by a chi-square distribution. The improvement for that approximation, that is, when the condition listed in (9) cannot be assumed, was discussed in Tsukada and Sugiyama [16]. The following lemma was provided by Nasuda et al. [17] and Takemura and Sheena [5] in the non-singular case and could be easily extended to the singular case.

**Lemma 2.** *Let  $W \sim W_m(n, \Sigma)$ , where  $m > n$  and  $\ell_1, \ell_2, \dots, \ell_n$  be the eigenvalues of  $W$ . If  $\rho_k \rightarrow 0$ , we have*

$$r_k = \max\left(\frac{\ell_2}{\ell_1}, \frac{\ell_3}{\ell_2}, \dots, \frac{\ell_{k+1}}{\ell_k}\right) \xrightarrow{p} 0,$$

in the sense that  $\forall \epsilon > 0, \exists \delta > 0$ ,

$$\rho_k < \delta \rightarrow \Pr(r_k > \epsilon) < \epsilon,$$

where  $k \leq n$ .

From Proposition 1 and Lemma 2, we obtain the chi-square approximation that is the main result of this paper.

**Theorem 1.** *Let  $W \sim W_m(n, \Sigma)$ , where  $m > n$  and  $\ell_1, \ell_2, \dots, \ell_n$  be the eigenvalues of  $W$ . If  $\rho_k \rightarrow 0$ , it holds that*

$$\ell_i / \lambda_i \xrightarrow{d} \chi_{n-i+1}^2, \quad 1 \leq i \leq k,$$

where  $\chi^2$  is a chi-square distribution with  $n - i + 1$  degrees of freedom and the symbol “ $\xrightarrow{d}$ ” means convergence in the distribution.

**Proof.** First, we rewrite the approximate distribution (7) as

$$f(\ell_1, \dots, \ell_n) = \frac{1}{2^{n(n+1)/4} |\Sigma|^{n/2} \prod_{i=1}^n \Gamma(\frac{n-i+1}{2})} \prod_{i=1}^n \ell_i^{(m-n-1)/2} \exp\left(-\frac{1}{2} \sum_{i=1}^n \frac{\ell_i}{\lambda_i}\right) \times \prod_{i < j}^n \left\{ (\ell_i - \ell_j) \left( \frac{\lambda_i \lambda_j}{\lambda_i - \lambda_j} \right) \right\}^{1/2} \prod_{i=1}^n \prod_{j=n+1}^m \left( \frac{1}{d_{ij}} \right)^{1/2}. \tag{10}$$

From Lemma 2, we have

$$\prod_{i < j}^n (\ell_i - \ell_j)^{1/2} = \prod_{1 \leq i \leq k, i < j}^n \ell_i^{1/2} \left(1 - \frac{\ell_j}{\ell_i}\right)^{1/2} \prod_{i=k+1 < j}^n (\ell_i - \ell_j)^{1/2} \approx \prod_{i=1}^k \ell_i^{(n-i)/2} \prod_{i=k+1 < j}^n (\ell_i - \ell_j)^{1/2}.$$

Then the approximate joint density is expressed by

$$f(\ell_1, \dots, \ell_n) \approx \frac{1}{2^{n(n+1)/4} |\Sigma|^{n/2} \prod_{i=1}^k \Gamma(\frac{n-i+1}{2})} \prod_{i=1}^k \ell_i^{(m-i+1)/2} \prod_{i=k+1}^n \ell_i^{(m-n-1)/2} \prod_{i=k+1 < j}^n (\ell_i - \ell_j)^{1/2} \times \exp\left(-\frac{1}{2} \sum_{i=1}^n \frac{\ell_i}{\lambda_i}\right) \prod_{i < j}^n \left( \frac{\lambda_i \lambda_j}{\lambda_i - \lambda_j} \right)^{1/2} \prod_{i=1}^n \prod_{j=n+1}^m \left( \frac{1}{d_{ij}} \right)^{1/2} = \frac{1}{2^{n(n+1)/4} |\Sigma|^{n/2} \prod_{i=1}^k \Gamma(\frac{n-i+1}{2})} \prod_{i=1}^k \ell_i^{(n-i+1)/2} \prod_{i=k+1}^n \ell_i^{(m-n-1)/2} \prod_{i=k+1 < j}^n (\ell_i - \ell_j)^{1/2} \times \exp\left(-\frac{1}{2} \sum_{i=1}^n \frac{\ell_i}{\lambda_i}\right) \prod_{i < j}^n \left( \frac{\lambda_i \lambda_j}{\lambda_i - \lambda_j} \right)^{1/2} \prod_{i=1}^k \prod_{j=n+1}^m \left( \frac{\lambda_j}{1 - \lambda_j/\lambda_i} \right)^{1/2} \prod_{i=k+1}^n \prod_{j=n+1}^m \left( \frac{1}{d_{ij}} \right)^{1/2}.$$

If  $\rho_k \rightarrow 0$ , we have

$$\prod_{i < j}^n \left( \frac{\lambda_i \lambda_j}{\lambda_i - \lambda_j} \right)^{1/2} = \prod_{1 \leq i \leq k, i < j}^n \left( \frac{\lambda_j}{1 - \lambda_j/\lambda_i} \right)^{1/2} \prod_{i=k+1 < j}^n \left( \frac{\lambda_i \lambda_j}{\lambda_i - \lambda_j} \right)^{1/2} \approx \prod_{i=1}^k \lambda_i^{(n-i)/2} \prod_{i=k+1 < j}^n \left( \frac{\lambda_i \lambda_j}{\lambda_i - \lambda_j} \right)^{1/2}.$$

We note  $|\Sigma|^{n/2} = \prod_{i=1}^k \lambda_i^{n/2} \prod_{i=k+1}^n \lambda_i^{n/2} \prod_{j=n+1}^m \lambda_j^{n/2}$ ; the joint density (10) is approximated as

$$f(\ell_1, \dots, \ell_n) \approx \prod_{i=1}^k \left\{ \frac{\ell_i^{(n-i-1)/2}}{(2\lambda_i)^{(n-i+1)/2} \Gamma(\frac{n-i+1}{2})} \exp\left(-\frac{\ell_i}{2\lambda_i}\right) \right\} \cdot f(\ell_{k+1}, \dots, \ell_n) \prod_{i=1}^k \lambda_i^{(n-2i+1)/2} \prod_{j=n+1}^m \lambda_j^{(k-n)/2} = \prod_{i=1}^k g_{n-i+1}(\ell_i/\lambda_i) \cdot f(\ell_{k+1}, \dots, \ell_n) \prod_{i=1}^k \lambda_i^{(n-2i+1)/2} \prod_{j=n+1}^m \lambda_j^{(k-n)/2},$$

where  $g_{n-i+1}(\cdot)$  is the density function of the chi-square distribution, and degree of freedom  $n - i + 1$  and  $f(\ell_{k+1}, \dots, \ell_n)$  is represented by

$$f(\ell_{k+1}, \dots, \ell_n) = \prod_{i=k+1}^n \frac{1}{2^{(n-i+1)/2} \lambda_i^{n/2} \Gamma(\frac{n-i+1}{2})} \prod_{i=k+1}^n \ell_i^{(m-n-1)/2} \prod_{i=k+1 < j}^n (\ell_i - \ell_j) \times \exp\left(-\frac{1}{2} \sum_{i=k+1}^n \frac{\ell_i}{\lambda_i}\right) \prod_{i=k+1 < j}^n \left(\frac{1}{c_{ij}}\right)^{1/2} \prod_{i=k+1}^n \prod_{j=n+1}^m \left(\frac{1}{d_{ij}}\right)^{1/2}.$$

□

Corollary 1 shows the chi-square approximation when all population eigenvalues are infinitely dispersed.

**Corollary 1.** Let  $W \sim W_m(n, \Sigma)$ , where  $m > n$  and  $\ell_1, \ell_2, \dots, \ell_n$  are the eigenvalues of  $W$ . If  $\rho_n \rightarrow 0$ , it holds that

$$\ell_i / \lambda_i \xrightarrow{d} \chi_{n-i+1}^2, \quad 1 \leq i \leq n.$$

**Proof.** The proof is provided in the Appendix A. □

In the context of the High Dimension-Low Sample Size (HDLSS) setting, the asymptotic behavior of the eigenvalue distribution of a sample covariance matrix was discussed in Ahn et al. [18], Bolivar-Cime and Perez-Abreu [19], Jung and Marron [20]. Jung and Marron [20] showed that the spiked sample eigenvalues are approximated by the chi-square distribution with a degree of freedom of  $n$ . In contrast, Theorem 1 provides the approximation of the distribution of individual eigenvalues by a chi-square distribution with varying degrees of freedom.

### 3. Application to Test for Equality of the Individual Eigenvalues

This section discusses testing for equality of individual eigenvalues of the covariance matrix in two populations. For testing problems, we give the approximate distribution of the statistic based on the derived results from the previous section.

Let an  $m \times n_i$  Gaussian random matrix  $X^{(i)}$  be distributed as  $X^{(i)} \sim N_{m,n}(O, \Sigma^{(i)} \otimes I_{n_i})$ , where  $\Sigma^{(i)} > 0$  and  $i = 1, 2$ . The eigenvalues of  $\Sigma^{(i)}$  are denoted by  $\lambda_1^{(i)}, \lambda_2^{(i)}, \dots, \lambda_m^{(i)}$ , where  $\lambda_1^{(i)} \geq \lambda_2^{(i)} \geq \dots \geq \lambda_m^{(i)} > 0$ . We denote the eigenvalues of  $W^{(i)} = X^{(i)} X^{(i)T}$  by  $\ell_1^{(i)}, \ell_2^{(i)}, \dots, \ell_m^{(i)}$ , where  $\ell_1^{(i)} > \ell_2^{(i)} > \dots > \ell_m^{(i)} \geq 0$ . For fixed  $j$ , we consider the test of the equality of the individual eigenvalues in two populations as

$$H_0 : \lambda_j^{(1)} = \lambda_j^{(2)}, \text{ vs. } H_1 : \lambda_j^{(1)} \neq \lambda_j^{(2)}. \tag{11}$$

Sugiyama and Ushizawa [21] reduced (11) to the equality of variance test for the principal components and proposed a testing procedure using the Ansari-Bradley test. Takeda [22] proposed the test statistic  $\ell_j^{(1)} / \ell_j^{(2)}$  with  $n \geq m$  for (11) and derived the exact distribution of  $\ell_1^{(1)} / \ell_1^{(2)}$ . Since Johnstone [23] indicated that the first few eigenvalues are very large compared to the others in the large dimensional setting, it is essential to understand how the distribution for the first few eigenvalues is constructed. We provide the exact density function of  $\ell_1^{(1)} / \ell_1^{(2)}$  with  $n < m$  in the same way as Takeda [22].

**Theorem 2.** Let  $W^{(1)}$  and  $W^{(2)}$  be two independent Wishart matrices with distribution  $W_m(n_1, \Sigma^{(1)})$  and  $W_m(n_2, \Sigma^{(2)})$ , respectively, where  $m > n_i$  ( $i = 1, 2$ ). Then we have the density of  $q = \ell_1^{(1)} / \ell_1^{(2)}$  as

$$\begin{aligned}
 f(q) = & C \sum_{k=0}^{\infty} \sum_{\kappa \in P_m^k} \sum_{t=0}^{\infty} \sum_{\tau \in P_m^t} \frac{\{(m+1)/2\}_{\kappa} C_{\kappa}(\Sigma^{(1)-1}/2)}{\{(n_1+m+1)/2\}_{\kappa} k!} \frac{\{(m+1)/2\}_{\tau} C_{\tau}(\Sigma^{(2)-1}/2)}{\{(n_2+m+1)/2\}_{\tau} t!} \\
 & \times \left\{ (mn_1/2+k)(mn_2/2+t)q^{mn_2/2+t-1}\Gamma(u)/v^u \right. \\
 & - (mn_1/2+k)(\text{tr}\Sigma^{(2)-1}/2)q^{mn_2/2+t}\Gamma(u+1)/v^{u+1} \\
 & - (mn_2/2+t)(\text{tr}\Sigma^{(1)-1}/2)q^{mn_2/2+t-1}\Gamma(u+1)/v^{u+1} \\
 & \left. + (\text{tr}\Sigma^{(1)-1}/2)(\text{tr}\Sigma^{(2)-1}/2)q^{mn_2/2+t}\Gamma(u+2)/v^{u+2} \right\}, \tag{12}
 \end{aligned}$$

where  $u = m(n_1 + n_2)/2 + k + t$ ,  $v = \text{tr}\Sigma^{(1)-1} - q\text{tr}\Sigma^{(2)-1}$  and

$$C = \frac{\Gamma_{n_1}\{(n_1+1)/2\}\Gamma_{n_2}\{(n_2+1)/2\}}{2^{m(n_1+n_2)/2}\Gamma_{n_1}\{(n_1+m+1)/2\}\Gamma_{n_2}\{(n_2+m+1)/2\}|\Sigma|^{n_1/2}|\Sigma|^{n_2/2}}.$$

**Proof.** The exact expression of  $\ell_1^{(i)}$  was provided by Shimizu and Hashiguchi [9] as

$$\Pr(\ell_1^{(i)} < x) = \frac{\Gamma_{n_i}(\frac{n_i+1}{2})(\frac{x}{2})^{mn_i/2}}{\Gamma_{n_i}(\frac{n_i+m+1}{2})|\Sigma^{(i)}|^{n_i/2}} \exp\left(-\frac{x}{2}\text{tr}\Sigma^{(i)-1}\right) {}_1F_1\left(\frac{m+1}{2}; \frac{n_i+m+1}{2}; \frac{x}{2}\Sigma^{(i)-1}\right). \tag{13}$$

The derivative of (13) is represented by

$$\begin{aligned}
 f(x) = & \frac{\Gamma_{n_i}\{(n_i+1)/2\}}{2^{mn_i/2}\Gamma_{n_i}\{(n_i+m+1)/2\}|\Sigma^{(i)}|^{n_i/2}} \sum_{k=0}^{\infty} \sum_{\kappa \in P_m^k} \frac{\{(m+1)/2\}_{\kappa} C_{\kappa}(\Sigma^{(i)-1}/2)}{\{(n_i+m+1)/2\}_{\kappa} k!} \\
 & \times \exp\left(-\frac{x}{2}\text{tr}\Sigma^{(i)-1}\right) \left\{ (n_i m/2+k)x^{mn_i/2+k-1} - (\text{tr}\Sigma^{(i)-1}/2)x^{mn_i/2+k} \right\}. \tag{14}
 \end{aligned}$$

From (14), we have the joint density of  $\ell_1^{(1)}$  and  $\ell_1^{(2)}$  as

$$\begin{aligned}
 f(x, y) = & C \sum_{k=0}^{\infty} \sum_{\kappa \in P_m^k} \sum_{t=0}^{\infty} \sum_{\tau \in P_m^t} \frac{\{(m+1)/2\}_{\kappa} C_{\kappa}(\Sigma^{(1)-1}/2)}{\{(n_1+m+1)/2\}_{\kappa} k!} \frac{\{(m+1)/2\}_{\tau} C_{\tau}(\Sigma^{(2)-1}/2)}{\{(n_2+m+1)/2\}_{\tau} t!} \\
 & \times \left\{ (mn_1/2+k)x^{mn_1/2+k-1} - (\text{tr}\Sigma^{(1)-1}/2)x^{mn_1/2+k} \right\} \\
 & \times \left\{ (mn_2/2+t)y^{mn_2/2+t-1} - (\text{tr}\Sigma^{(2)-1}/2)y^{mn_2/2+t} \right\} \exp\left(-\frac{x}{2}\text{tr}\Sigma^{(1)-1}\right) \exp\left(-\frac{y}{2}\text{tr}\Sigma^{(2)-1}\right).
 \end{aligned}$$

Translating  $x$  and  $y$  to  $q = y/x$  and  $r = x$ , we have

$$\begin{aligned}
 f(q, r) = & C \sum_{k=0}^{\infty} \sum_{\kappa \in P_m^k} \sum_{t=0}^{\infty} \sum_{\tau \in P_m^t} \frac{\{(m+1)/2\}_{\kappa} C_{\kappa}(\Sigma^{(1)-1}/2)}{\{(n_1+m+1)/2\}_{\kappa} k!} \frac{\{(m+1)/2\}_{\tau} C_{\tau}(\Sigma^{(2)-1}/2)}{\{(n_2+m+1)/2\}_{\tau} t!} \\
 & \times \left\{ (mn_1/2+k)(mn_2/2+t)q^{mn_2/2+t-1}r^{m(n_1+n_2)/2+k+t-1} \right. \\
 & - (mn_1/2+k)(\text{tr}\Sigma^{(2)-1}/2)q^{mn_2/2+t}r^{m(n_1+n_2)/2+k+t} \\
 & - (mn_2/2+t)(\text{tr}\Sigma^{(1)-1}/2)q^{mn_2/2+t-1}r^{m(n_1+n_2)/2+k+t} \\
 & \left. + (\text{tr}\Sigma^{(1)-1}/2)(\text{tr}\Sigma^{(2)-1}/2)q^{mn_2/2+t}r^{m(n_1+n_2)/2+k+t+1} \right\} \\
 & \times \exp\left(-(\text{tr}\Sigma^{(1)-1} - q\text{tr}\Sigma^{(2)-1})r\right).
 \end{aligned}$$

Noting that  $\int_0^\infty x^{\alpha-1} e^{-\beta x} dx = \Gamma(\alpha) / \beta^\alpha$ , where  $\alpha, \beta > 0$ , and integrating  $r$  with respect to  $f(q, r)$ , we have the desired result.  $\square$

As the dimension increases, it is difficult to perform the numerical computation of (12) due to the high computational complexity. From Theorem 1, we provide the approximate distribution for (12) by  $F$ -distribution.

**Corollary 2.** Let  $W^{(1)}$  and  $W^{(2)}$  be two independent Wishart matrices with distribution  $W_m(n_1, \Sigma^{(1)})$  and  $W_m(n_2, \Sigma^{(2)})$ , respectively, where  $m > n_i$  ( $i = 1, 2$ ) and  $\ell_1^{(i)}, \ell_2^{(i)}, \dots, \ell_n^{(i)}$  are the eigenvalues of  $W^{(i)}$ . If the first  $k$ -th eigenvalues of  $\Sigma^{(i)}$  are spiked, then we have

$$\frac{\ell_j^{(1)} / \{(n_1 - j + 1)\lambda_{n_1-j+1}^{(1)}\}}{\ell_j^{(2)} / \{(n_2 - j + 1)\lambda_{n_2-j+1}^{(2)}\}} \xrightarrow{d} F_{(n_1-j+1, n_2-j+1)}, \quad 1 \leq j \leq k,$$

where  $F$  is an  $F$  distribution with  $n_1$  and  $n_2$  degrees of freedom.

#### 4. Simulation Study

We investigate the accuracy of the approximation for the derived distributions. In the simulation study, we consider the following population covariance matrix:

$$\Sigma = \text{diag}(a^b, a^{b/2}, \dots, a^{b/m}), \tag{15}$$

where  $a, b > 0$ . In the large-dimensional setting, mainly the accuracy of the approximate distributions for the largest and second eigenvalues was investigated; see Imori et al. [24]. In (15), we set  $(a, b) = (200, 3)$  as Case 1 and  $(a, b) = (50, 3)$  as Case 2. These two cases imply that the population covariance matrix has two spiked eigenvalues. Parameter  $\rho_k$  in (9) is smaller in Case 1 than in Case 2. We denote  $F_1(x)$  and  $F_2(x)$  as the chi-square distributions with  $n$  and  $n - 1$  degrees of freedom, which are the approximate distributions of the largest and second eigenvalues, respectively. The empirical distribution based on  $10^6$  Monte Carlo simulations is denoted by  $F_{\text{sim}}$ . Tables 1 and 2 show the  $100\alpha$ -percentile points of the distributions of  $\ell_1$  and  $\ell_2$  for  $m = 50$  and  $n = 10$ , respectively. From the simulation study, we know that sufficient accuracy of approximation for the largest eigenvalue has already been obtained in Case 2. Case 1 is more accurate than Case 2 for the second eigenvalue. It is seen that the desired accuracy can be achieved when the parameter  $\rho_k$  is small.

**Table 1.** Percentile points of the distributions of  $\ell_1$  and  $\ell_2$  of  $W_{50}(10, \Sigma)$  (Case 1).

$\alpha$	$F_{\text{sim}}^{-1}(\alpha)$	$F_1^{-1}(\alpha)$	$\alpha$	$F_{\text{sim}}^{-1}(\alpha)$	$F_2^{-1}(\alpha)$
0.99	23.2359	23.2093	0.99	21.791	21.666
0.95	18.3026	18.307	0.95	17.0601	16.919
0.90	15.9825	15.9872	0.90	14.8377	14.6837
0.50	9.34466	9.34182	0.50	8.48676	8.34283
0.05	3.94389	3.9403	0.05	3.47796	3.32511

**Table 2.** Percentile points of the distributions of  $\ell_1$  and  $\ell_2$  of  $W_{50}(10, \Sigma)$  (Case 2).

$\alpha$	$F_{\text{sim}}^{-1}(\alpha)$	$F_1^{-1}(\alpha)$	$\alpha$	$F_{\text{sim}}^{-1}(\alpha)$	$F_2^{-1}(\alpha)$
0.99	23.239	23.2093	0.99	22.1285	21.666
0.95	18.306	18.307	0.95	17.4079	16.919
0.90	15.9857	15.9872	0.90	15.1856	14.6837
0.50	9.34844	9.34182	0.50	8.84394	8.34283
0.05	3.94744	3.9403	0.05	3.86566	3.32511



Tables 3 and 4 present the chi-square probabilities for Case 1 in 90% 95% 99% percentile points denoted as  $x_{0.90}$ ,  $x_{0.95}$ , and  $x_{0.99}$  from the empirical distribution. We denote  $G_1(x)$  and  $G_2(x)$  as the chi-square approximation of the distributions for the largest and second largest eigenvalues, respectively. In this simulation study, we set  $m = 20, 30, 40, 100$ , and  $n = 5, 15$ . It can be observed that all probabilities are close to the true theoretical probabilities.

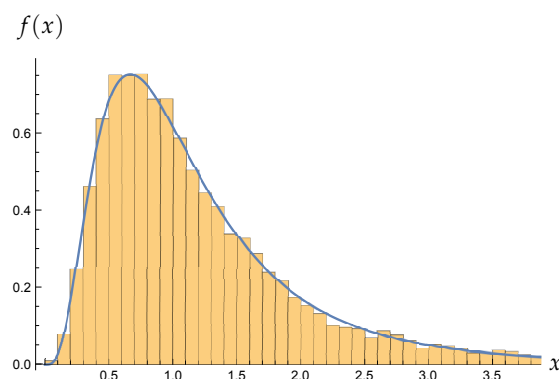
**Table 3.** Approximate probabilities of  $\ell_1$  based on the empirical percentile points (Case 1).

$n$	$m$	$G_1(x_{0.90})$	$G_1(x_{0.95})$	$G_1(x_{0.99})$
5	20	0.900047	0.950011	0.990165
	30	0.900227	0.950002	0.990018
	40	0.900173	0.950185	0.990072
	100	0.899952	0.949874	0.990048
15	20	0.900409	0.950375	0.990088
	30	0.900258	0.950232	0.990124
	40	0.900729	0.950464	0.990213
	100	0.900331	0.950039	0.990083

**Table 4.** Approximate probabilities of  $\ell_2$  based on the empirical percentile points (Case 1).

$n$	$m$	$G_2(x_{0.90})$	$G_2(x_{0.95})$	$G_2(x_{0.99})$
5	20	0.904868	0.952295	0.990482
	30	0.905512	0.953016	0.990761
	40	0.905838	0.952924	0.990609
	100	0.906393	0.95315	0.990613
15	20	0.903118	0.951587	0.990331
	30	0.903275	0.951856	0.990335
	40	0.903943	0.952062	0.990315
	100	0.904027	0.952163	0.990498

Finally, we provide the graph of the density of  $F$  distribution in Corollary 2 compared to the empirical distribution function. In Figure 1, we superimpose the graph of the  $F$  approximation with the histogram of  $\ell_1^{(1)} / \ell_1^{(2)}$  for  $n_i = 10$  ( $i = 1, 2$ ) and  $m = 30$  in Case 2. The vertical line and histograms show the empirical distribution of the  $\ell_1^{(1)} / \ell_1^{(2)}$  based on  $10^6$  iteration, respectively. The solid line is the density function of the  $F$  distribution. From the 95% points of  $F_{sim}$ , we can confirm that the approximate probability is 0.950.



**Figure 1.**  $n_i = 10$  ( $i = 1, 2$ ) and  $m = 30$ .

### 5. Concluding Remarks

In this study, we provided the approximate distribution of eigenvalues of the singular Wishart matrix, which is similar to the result of Takemura and Sheena [5] for the non-singular case. Through numerical experiments, we confirmed that the approximation accuracy is sufficient when the parameter  $\rho_k$  is small. The distribution

approximation proposed by Tsukada and Sugiyama [16] might be useful to improve the derived approximate results when  $\rho_k \rightarrow 0$  is not assumed. As a part of future work, it would be desirable to examine the robustness of the chi-square approximation to normality assumption.

**Author Contributions:** Conceptualization, K.S.; methodology, K.S. and H.H.; software, K.S.; validation, H.H.; investigation, K.S. and H.H.; writing—original draft, K.S.; writing—review and editing, H.H. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by JSPS KAKENHI Grant Numbers 23K19015 and 23K11016.

**Data Availability Statement:** No new data were created or analyzed in this study. Data sharing is not applicable to this article.

**Conflicts of Interest:** The authors declare no conflict of interest.

### Appendix A

#### Proof of Corollary 1

From Lemma 2, we have

$$\prod_{i < j}^n (\ell_i - \ell_j)^{1/2} = \prod_{i < j}^n \ell_i^{1/2} \left(1 - \frac{\ell_j}{\ell_i}\right)^{1/2} \approx \prod_{i=1}^n \ell_i^{(n-i)/2},$$

$$\prod_{i < j}^n \left(\frac{\lambda_i \lambda_j}{\lambda_i - \lambda_j}\right)^{1/2} = \prod_{i < j}^n \left(\frac{\lambda_j}{1 - \lambda_i/\lambda_j}\right)^{1/2} \approx \prod_{i=1}^n \lambda_i^{(i-1)/2}.$$

Furthermore, we note  $|\Sigma|^{n/2} = \prod_{i=1}^n \lambda_i^{n/2} \prod_{i=1}^n \prod_{j=n+1}^m \lambda_j^{1/2}$ ; the joint density (10) is represented by

$$\begin{aligned} & f(\ell_1, \dots, \ell_n) \\ & \approx \prod_{i=1}^n \left\{ \frac{\ell_i^{(n-i-1)/2}}{(2\lambda_i)^{(n-i+1)/2} \Gamma(\frac{n-i+1}{2})} \exp\left(-\frac{\ell_i}{2\lambda_i}\right) \right\} \prod_{i=1}^n \ell_i^{(m-n)/2} \prod_{i=1}^n \prod_{j=n+1}^m \left(\frac{1}{\lambda_j}\right)^{1/2} \prod_{i=1}^n \prod_{j=n+1}^m \left(\frac{1}{d_{ij}}\right)^{1/2} \\ & = \prod_{i=1}^n \left\{ \frac{\ell_i^{(n-i-1)/2}}{(2\lambda_i)^{(n-i+1)/2} \Gamma(\frac{n-i+1}{2})} \exp\left(-\frac{\ell_i}{2\lambda_i}\right) \right\} \prod_{i=1}^n \prod_{j=n+1}^m \ell_i^{1/2} \prod_{i=1}^n \prod_{j=n+1}^m \left(\frac{1}{\lambda_j}\right)^{1/2} \prod_{i=1}^n \prod_{j=n+1}^m \left(\frac{1}{d_{ij}}\right)^{1/2} \\ & = \prod_{i=1}^n \left\{ \frac{\ell_i^{(n-i-1)/2}}{(2\lambda_i)^{(n-i+1)/2} \Gamma(\frac{n-i+1}{2})} \exp\left(-\frac{\ell_i}{2\lambda_i}\right) \right\} \\ & = \prod_{i=1}^n g_{n-i+1}(\ell_i/\lambda_i). \end{aligned}$$

### References

- James, A.T. Distribution of matrix variates and latent roots derived from normal samples. *Ann. Math. Stat.* **1964**, *35*, 475–501. [CrossRef]
- Sugiyama, T. Approximation for the distribution of the largest latent root of a Wishart matrix. *Aust. J. Stat.* **1972**, *14*, 17–24. [CrossRef]
- Sugiura, N. Derivatives of the characteristic root of a symmetric or a hermitian matrix with two applications in multivariate analysis. *Commun. Stat. Theory Methods* **1973**, *1*, 393–417. [CrossRef]
- Kato, H.; Hashiguchi, H. Chi-square approximations for eigenvalue distributions and confidential interval construction on population eigenvalues. *Bull. Comput. Stat. Jpn.* **2014**, *27*, 11–28.
- Takemura, A.; Sheena, Y. 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. Multivar. Anal.* **2005**, *94*, 271–299. [CrossRef]
- Butler, R.W.; Wood, T.A. Laplace approximations for hypergeometric functions with matrix argument. *Ann. Stat.* **2002**, *30*, 1155–1177. [CrossRef]

7. Butler, R.W.; Wood, T.A. Laplace approximations to hypergeometric functions of two matrix arguments. *J. Multivar. Anal.* **2005**, *94*, 1–18. [[CrossRef](#)]
8. Butler, R.W.; Wood, T.A. Laplace approximations for hypergeometric functions with Hermitian matrix argument. *J. Multivar. Anal.* **2022**, *192*, 105087. [[CrossRef](#)]
9. Shimizu, K.; Hashiguchi, H. Heterogeneous hypergeometric functions with two matrix arguments and the exact distribution of the largest eigenvalue of a singular beta-Wishart matrix. *J. Multivar. Anal.* **2021**, *183*, 104714. [[CrossRef](#)]
10. Shinozaki, A.; Shimizu, K.; Hashiguchi, H. Generalized heterogeneous hypergeometric functions and the distribution of the largest eigenvalue of an elliptical Wishart matrix. *Random Matrices Theory Appl.* **2022**, *11*, 2250034. [[CrossRef](#)]
11. Muirhead, R.J. *Aspects of Multivariate Statistical Theory*; Wiley: New York, NY, USA, 1982.
12. Uhlig, H. On Singular Wishart and singular multivariate beta distributions. *Ann. Stat.* **1994**, *22*, 395–405. [[CrossRef](#)]
13. Srivastava, M.S. Singular Wishart and multivariate beta distributions. *Ann. Stat.* **2003**, *31*, 1537–1560. [[CrossRef](#)]
14. Anderson, G.A. An asymptotic expansion for the distribution of the latent roots of the estimated covariance matrix. *Ann. Math. Stat.* **1965**, *36*, 1153–1173. [[CrossRef](#)]
15. Constantine, A.G.; Muirhead, R.T. Asymptotic expansions for distributions of latent roots in multivariate analysis. *J. Multivar. Anal.* **1976**, *3*, 369–391. [[CrossRef](#)]
16. Tsukada, S.; Sugiyama, T. Distribution approximation of covariance matrix eigenvalues. *Comm. Statist. Simul. Comput.* **2023**, *52*, 4313–4325. [[CrossRef](#)]
17. Nasuda, R.; Shimizu, K.; Hashiguchi, H. Asymptotic behavior of the distributions of eigenvalues for beta-Wishart ensemble under the dispersed population eigenvalues. *Commun. Stat. Theory Methods* **2023**, *52*, 7840–7860. [[CrossRef](#)]
18. Ahn, J.; Marron, J.S.; Muller, K.M.; Chi, Y. The high-dimension, low-sample-size geometric representation holds under mild conditions. *Biometrika* **2007**, *94*, 760–766. [[CrossRef](#)]
19. Bolivar-Cime, A.; Perez-Abreu, V. PCA and eigen-inference for a spiked covariance model with largest eigenvalues of same asymptotic order. *Braz. J. Probab. Stat.* **2014**, *28*, 255–274. [[CrossRef](#)]
20. Jung, S.; Marron, J.S. PCA consistency in high dimension, low sample size context. *Ann. Stat.* **2009**, *37*, 4104–4130. [[CrossRef](#)]
21. Sugiyama, T.; Ushizawa, K. A non-parametric method to test equality of intermediate latent roots of two populations in a principal component analysis. *J. Jpn. Stat. Soc.* **1998**, *28*, 227–235. [[CrossRef](#)]
22. Takeda, Y. Permutation test for equality of each characteristic root in two populations. *J. Jpn. Soc. Comput. Stat.* **2001**, *14*, 1–10. [[CrossRef](#)]
23. Johnstone, I.M. On the distribution of the largest eigenvalue in principal components analysis. *Ann. Stat.* **2001**, *29*, 295–327. [[CrossRef](#)]
24. Iimori, T.; Ogura, T.; Sugiyama, T. On the distribution of the second-largest latent root for certain high dimensional Wishart matrices. *Int. J. Knowl. Eng. Soft Data Paradig.* **2013**, *4*, 187–197. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.