Journal of Multivariate Analysis 101 (2010) 1319–1329

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Testing the equality of several covariance matrices with fewer observations than the dimension

Muni S. Srivastava [a,*], Hirokazu Yanagihara [b]

[a] Department of Statistics, University of Toronto, 100 St George Street, Toronto, Ontario, M5S 3G3, Canada

[b] Department of Mathematics, Graduate School of Science, Hiroshima University, 1-3-1 Kagamiyama, Higashi-Hiroshima, Hiroshima 739-8626, Japan

ARTICLE INFO

Article history:
Received 18 July 2008
Available online 4 January 2010

AMS subject classifications:
primary 62H15
secondary 62H10

Keywords:
Comparison of powers
Equality of several covariance matrices
Equality of two covariances
High-dimensional data
Normality
Sample size smaller than the dimension

ABSTRACT

For normally distributed data from the k populations with $m \times m$ covariance matrices $\Sigma_1, \ldots, \Sigma_k$, we test the hypothesis $H : \Sigma_1 = \cdots = \Sigma_k$ vs the alternative $A \neq H$ when the number of observations $N_i, i = 1, \ldots, k$ from each population are less than or equal to the dimension m, $N_i \leq m, i = 1, \ldots, k$. Two tests are proposed and compared with two other tests proposed in the literature. These tests, however, do not require that $N_i \leq m$, and thus can be used in all situations, including when the likelihood ratio test is available. The asymptotic distributions of the test statistics are given, and the power compared by simulations with other test statistics proposed in the literature. The proposed tests perform well and better in several cases than the other two tests available in the literature.

1. Introduction

Let $\boldsymbol{x}_{ij}, j = 1, \ldots, N_i, i = 1, \ldots, k$, be independently distributed as multivariate normal with mean vectors $\boldsymbol{\mu}_i$ and covariance matrices Σ_i, denoted as $N_m(\boldsymbol{\mu}_i, \Sigma_i)$, where m denotes the dimension of the random vectors $\boldsymbol{x}_{ij}$ which will be assumed to be bigger than the sample sizes, $m \geq N_i, i = 1, \ldots, k$. In microarray datasets, m is usually in thousands whereas N_i are small, often much less than 50. The analysis of such datasets has often been carried out in the two-sample case under the assumption that $\Sigma_1 = \Sigma_2$ without verifying or testing this assumption. Schott [1] and Srivastava [2,3] have proposed tests for testing the equality of two or more covariance matrices. In this paper, we propose another test and compare its performance with the above tests. In order to describe these test statistics, we first consider the sufficient statistics given by

$$\bar{\boldsymbol{x}}_i = \frac{1}{N_i}\sum_{j=1}^{N_i}\boldsymbol{x}_{ij}, \qquad V_i = \sum_{j=1}^{N_i}(\boldsymbol{x}_{ij} - \bar{\boldsymbol{x}}_i)(\boldsymbol{x}_{ij} - \bar{\boldsymbol{x}}_i)', \quad i = 1, \ldots, k. \tag{1.1}$$

Let

$$V = \sum_{i=1}^{k} V_i, \qquad S_i = \frac{1}{n_i}V_i, \quad n_i = N_i - 1, \ i = 1, \ldots, k, \tag{1.2}$$

$$\hat{a}_{1i} = \frac{1}{mn_i}\operatorname{tr} V_i, \quad i = 1, \ldots, k, \tag{1.3}$$

$$\hat{a}_{2i} = \frac{1}{m(n_i - 1)(n_i + 2)}\left\{\operatorname{tr} V_i^2 - \frac{1}{n_i}(\operatorname{tr} V_i)^2\right\}, \quad i = 1, \ldots, k. \tag{1.4}$$

* Corresponding author.
E-mail address: srivasta@utstat.toronto.edu (M.S. Srivastava).

0047-259X/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmva.2009.12.010

We shall first consider the case of $k = 2$, that is, we test the hypothesis of equality of two covariance matrices:

$$H : \Sigma_1 = \Sigma_2 = \Sigma \quad \text{vs}\, A : \Sigma_1 \neq \Sigma_2.$$

Sometimes, it may also be of interest to consider only one-sided alternative such as $A_1 : \Sigma_1 > \Sigma_2$, where $\Sigma_1 > \Sigma_2$ means $\Sigma_1 - \Sigma_2$ is positive definite. For if it is known that one group has a larger covariance matrix than the other group, it may be advisable, if feasible, to take larger sample from the group that has a larger covariance matrix to offset the largeness to some degree.

For testing the equality of two covariance matrices, we note that S_1 and S_2 are unbiased estimators of Σ_1 and Σ_2, respectively. However, since $n_i < m$, these are singular matrices and hence V_i are distributed as singular Wishart, $V_i \sim W_m(\Sigma_i, n_i)$, $n_i < m$, see Srivastava [4] for the distribution of a singular Wishart matrix. When $n_i > m$, $i = 1, 2$, the likelihood ratio test is based on the eigenvalues of $V_1^{-1}V_2$. Although V_1^{-1} does not exist when $n_1 < m$ we may consider the Moore–Penrose inverse of V_1. Thus, for testing the hypothesis $H : \Sigma_1 = \Sigma_2$ against the alternative $A : \Sigma_1 \neq \Sigma_2$, Srivastava [2] proposed a test based on the statistic

$$G_2 = m\hat{b}\, \mathrm{tr}\, V_1^{+}V_2, \tag{1.5}$$

where V_1^{+} is the Moore–Penrose inverse of V_1, defined in Section 2, and $\hat{b}$ is a consistent estimator of

$$b = \frac{(\mathrm{tr}\, \Sigma/m)^2}{\mathrm{tr}\, \Sigma^2/m}, \tag{1.6}$$

Σ being the common unknown covariance matrix of the two populations under the null hypothesis that $\Sigma_1 = \Sigma_2 \equiv \Sigma$. We estimate this common covariance matrix Σ by

$$\hat{\Sigma} = \frac{1}{n}V \equiv \frac{1}{(n_1 + n_2)}(V_1 + V_2) \equiv S,$$

and b by $\hat{b} = \hat{a}_1^2/\hat{a}_2$, where $n = n_1 + n_2$ and

$$\hat{a}_1 = \frac{1}{nm}\mathrm{tr}\, V, \qquad \hat{a}_2 = \frac{1}{(n-1)(n+2)m}\left\{\mathrm{tr}\, V^2 - \frac{1}{n}(\mathrm{tr}\, V)^2\right\}. \tag{1.7}$$

It may be noted that when $\Sigma_1 = \Sigma_2 = \Sigma, V \sim W_m(\Sigma, n)$. It can be shown that for fixed n_1 and n_2, and under the hypothesis $H : \Sigma_1 = \Sigma_2$,

$$\lim_{m\to\infty} G_2 \sim \chi^2_{n_1 n_2}, \tag{1.8}$$

where χ_f^2 denotes a chi-square random variable with f degrees of freedom. It is found that this test does not perform well. Thus, we need to consider alternative tests.

We consider a measure of distance between the hypothesis and the alternative, namely the Frobenius norm. A test based on a consistent estimator of this distance has been proposed by Schott [1]. It is given by

$$J_2 = \frac{n_1 n_2}{2(n_1 + n_2)\hat{a}_2}\left(\hat{a}_{21} + \hat{a}_{22} - \frac{2}{m}\mathrm{tr}\, S_1 S_2\right), \tag{1.9}$$

which, under the null hypothesis, is distributed as $N(0, 1)$ as $(m, n_1, n_2) \to \infty$, provided $m/n_i \to c_i \in (0, \infty)$ as $(m, n_i) \to \infty, i = 1, 2$. Using a lower bound on Frobenius norm, a test based on its consistent estimator has been proposed by Srivastava [3]. It is given by

$$T_2 = \frac{\hat{a}_{21} - \hat{a}_{22}}{\sqrt{\hat{\eta}_1^2 + \hat{\eta}_2^2}}, \tag{1.10}$$

where $\hat{\eta}_i^2$ are consistent estimators of

$$\eta_i^2 = \frac{4}{n_i^2}a_2^2\left(1 + \frac{2n_i a_4}{m a_2^2}\right), \quad i = 1, 2. \tag{1.11}$$

Under the null hypothesis of equality of two covariance matrices, T_2 is asymptotically distributed on $N(0, 1)$ as $(n, m) \to \infty$. Equivalently, T_2^2 is asymptotically distributed as chi-square with one degree of freedom, χ_1^2, under the null hypothesis.

The above two tests are based on the differences of $\mathrm{tr}\, \Sigma_1^2$ and $\mathrm{tr}\, \Sigma_2^2$. But the differences between $\mathrm{tr}\, \Sigma_1$ and $\mathrm{tr}\, \Sigma_2$ may also throw a light on the differences between the two covariances. To devise a procedure that takes into account this fact, we consider a measure of distance by $\mathrm{tr}\, \Sigma_1^2/(\mathrm{tr}\, \Sigma_1)^2 - \mathrm{tr}\, \Sigma_2^2/(\mathrm{tr}\, \Sigma_2)^2$. Thus, we propose our test based on a consistent estimator of this measure of distance, namely,

$$Q_2 = \frac{\hat{\gamma}_1 - \hat{\gamma}_2}{\sqrt{\hat{\xi}_1^2 + \hat{\xi}_2^2}}, \tag{1.12}$$

where

$$\hat{\gamma}_i = \frac{\hat{a}_{2i}}{\hat{a}_{1i}^2}, \quad i = 1, 2, \tag{1.13}$$

and $\hat{\xi}_i^2$ are consistent estimators of

$$\xi_i^2 = \frac{4}{n_i^2} \left\{ \frac{a_2^2}{a_1^4} + \frac{2n_i}{m} \left(\frac{a_2^3}{a_1^6} - \frac{2a_2a_3}{a_1^5} + \frac{a_4}{a_1^4} \right) \right\}, \quad i = 1, 2. \tag{1.14}$$

Under the null hypothesis H and assumption of (A1)–(A4), which are described in Section 2, Q_2 is asymptotically normally distributed as $N(0, 1)$ as $(m, n) \to \infty$. Thus the test based on Q_2 can be used to test one-sided hypothesis such as $H : \Sigma_1 = \Sigma_2$ vs $A : \Sigma_1 > \Sigma_2$, since tr $\Sigma^2/(\text{tr } \Sigma)^2$ is a monotone increasing function of the ordered eigenvalues of Σ, see Srivastava and Khatri [5] in Corollary 10.4.2, page 317.

For testing the equality of several covariance matrices, Schott [1] proposed a test based on the statistic

$$J_k = \sum_{i<j}^{k} \frac{\text{tr } (S_i - S_j)^2}{\hat{\theta}}, \tag{1.15}$$

where $\hat{c}_i = m/n_i i = 1, \ldots, k$ and

$$\hat{\theta} = 2\hat{a}_2 \left\{ \sum_{i<j}^{k} (\hat{c}_i + \hat{c}_j)^2 + (k-1)(k-2) \sum_{i=1}^{k} \hat{c}_i^2 \right\}^{1/2}. \tag{1.16}$$

It is assumed that $\hat{c}_i \to c_i \in (0, \infty)$, as $(m, n_i) \to \infty$, $i = 1, \ldots, k$, and $0 < \lim_{m\to\infty} \text{tr } \Sigma^j/m < \infty$, $j = 1, \ldots, 8$. Under these two assumptions, Schott [1] has shown that when the null hypothesis holds, $J_k \xrightarrow{d} N(0, 1)$ as $(m, n_i) \to \infty$, where 'd' stands for 'in distribution'.

The generalized version of the T_2^2-statistic for testing the equality of k covariance matrices is given by

$$T_k^2 = \sum_{i=1}^{k} \frac{(\hat{a}_{2i} - \bar{\hat{a}}_2)^2}{\hat{\eta}_i^2}, \tag{1.17}$$

where $\bar{\hat{a}}_2$ is a weighted mean of $\hat{a}_{21}, \ldots, \hat{a}_{2k}$, i.e.,

$$\bar{\hat{a}}_2 = \frac{\sum_{i=1}^{k} \hat{a}_{2i}/\hat{\eta}_i^2}{\sum_{i=1}^{k} 1/\hat{\eta}_i^2}, \tag{1.18}$$

and $\hat{\eta}_i^2$'s are estimated values of (1.11). It is easily seen that T_k^2 with $k = 2$ coincides with a squared T_2 in (1.10). The estimators $\hat{a}_2$ and $\hat{a}_4$ use $V = V_1 + \cdots + V_k$, and $n = n_1 + \cdots + n_k$. Asymptotically T_k^2 is distributed as χ_{k-1}^2 under the null hypothesis that all covariance matrices are equal.

The generalization version of the Q_2^2-statistic for testing the equality of k covariance matrices is given by

$$Q_k^2 = \sum_{i=1}^{k} \frac{(\hat{\gamma}_i - \bar{\hat{\gamma}})^2}{\hat{\xi}_i^2}, \tag{1.19}$$

where $\bar{\hat{\gamma}}$ is a weighted mean of $\hat{\gamma}_1, \ldots, \hat{\gamma}_k$, i.e.,

$$\bar{\hat{\gamma}} = \frac{\sum_{i=1}^{k} \hat{\gamma}_i/\hat{\xi}_i^2}{\sum_{i=1}^{k} 1/\hat{\xi}_i^2}, \tag{1.20}$$

and $\hat{\xi}_i^2$'s are estimated values of (1.14). It is easily seen that Q_k^2 with $k = 2$ coincides with a squared Q_2 in (1.12). The asymptotic distribution of Q_k^2 is χ_{k-1}^2.

The organization of this article is as follows. In Section 2, we give some preliminary results. The problem of testing the equality of two covariance matrices is considered in Section 3 and that of several covariance matrices in Section 4. The comparison of power of the tests for the equality of two and three covariance matrices is done in Section 5. In Section 6, we give proofs of Theorems 2.1 and 2.2 which give important results and may be useful in future work as no such results exist in the literature. The conclusion is given in Section 7.

2. Preliminaries

The Moore–Penrose inverse of a matrix A is defined by A^+ satisfying the following four conditions : (i) $AA^+A = A$, (ii) $A^+AA^+ = A^+$, (iii) $(AA^+)' = AA^+$, (iv) $(A^+A)' = A^+A$. The Moore–Penrose inverse is unique. We shall make the following assumptions:

(A1) $n = O(m^\delta), \quad \delta > 1/2,$

(A2) $\lim_{n\to\infty} n_i/n = g_i, \quad 0 < g_i < 1,\ i = 1, \dots, k,\ n = \sum_{i=1}^{n} n_i,$

(A3) $0 < a_{i0} = \lim_{m\to\infty} \mathrm{tr}\,\Sigma^i/m < \infty, \quad i = 1, \dots, 4,$

(A4) $0 < a_{ij0} = \lim_{m\to\infty} \mathrm{tr}\,\Sigma_j^i/m < \infty, \quad i = 1, \dots, 4,\ j = 1, \dots, k.$

Lemma 2.1. *Let $V \sim W_m(\Sigma, n)$, $a_i = \mathrm{tr}\,\Sigma^i/m$, $i = 1, \dots, 4$. Then under the assumptions* (A1) *and* (A3)*, unbiased and consistent estimators of a_1 and a_2 as $(n, m) \to \infty$ are respectively given by $\hat{a}_1$ and $\hat{a}_2$ defined in (1.7).*

The following two lemmas on asymptotic normality of $\hat{a}_2$ and $\hat{\gamma} = \hat{a}_2/\hat{a}_1^2$ are given in Srivastava [6].

Lemma 2.2. *Let $V \sim W_m(\Sigma, n)$, $\hat{a}_2$ as defined in (1.7), and $a_i = tr\,\Sigma^i/m$, $i = 1, \dots, 4$. Then under the conditions* (A1) *and* (A3)

$$\lim_{(n,m)\to\infty} P\{(\hat{a}_2 - a_2)/\eta \le x\} = \Phi(x)$$

where $\Phi(x)$ denotes the cumulative distribution function of a standard normal random variable, and $\eta = \eta_i$ with $n_i = n$.

Lemma 2.3. *Let $V \sim W_m(\Sigma, n)$, $\hat{\gamma} = \hat{a}_2/\hat{a}_1^2$ with $\hat{a}_1$ and $\hat{a}_2$ as defined in (1.7), and $\xi = \xi_i$ with $n_i = n$. Then under the conditions* (A1) *and* (A3)

$$\lim_{(n,m)\to\infty} P\{(\hat{\gamma} - \gamma)/\xi \le x\} = \Phi(x),$$

where $\gamma = a_2/a_1^2$.

In order to apply Lemmas 2.2 and 2.3, we need consistent estimators of η and ξ, that is of a_i, $i = 1, \dots, 4$. While consistent estimators of a_1 and a_2 are available in Srivastava [6], and stated in Lemma 2.1, in the next two theorems we give consistent estimators of a_3 and a_4, the proofs are given in Section 6. These two theorems may be of great help in varieties of other problems where consistent estimators of the third and fourth moments are needed, as no such results are available in the literature.

Theorem 2.1. *Let $V \sim W_m(\Sigma, n)$, and $a_3 = tr\,\Sigma^3/m$. Then, under the condition* (A3)*, and as $(n, m) \to \infty$, $n = O(m^\delta)$, $\delta > 1/3$, a consistent estimator of a_3 is given by*

$$\hat{a}_3 = \frac{1}{n(n^2 + 3n + 4)} \left\{ \frac{1}{m} tr\, V^3 - 3n(n+1)m\hat{a}_2\hat{a}_1 - nm^2\hat{a}_1^3 \right\},$$

where $\hat{a}_1$ and $\hat{a}_2$ have been defined in (1.7).

It can be shown that an unbiased and consistent estimator of a_3 is given by

$$\hat{a}_{3u} = \frac{n}{(n-1)(n-2)(n+2)(n+4)} \left\{ \frac{1}{m} \mathrm{tr}\, V^3 - 3(n+2)(n-1)\hat{a}_2\hat{a}_1 - nm^2\hat{a}_1^3 \right\}.$$

Theorem 2.2. *Let $V \sim W_m(\Sigma, n)$, and $a_4 = tr\,\Sigma^4/m$. Then, under the condition* (A3)*, and as $(n, m) \to \infty$, $n = O(m^\delta)$, $\delta > 1/2$, a consistent estimator of a_4 is given by*

$$\hat{a}_4 = \frac{1}{c_0} \left(\frac{1}{m} tr\, V^4 - mc_1\hat{a}_1 - m^2c_2\hat{a}_1^2\hat{a}_2 - mc_3\hat{a}_2^2 - nm^3\hat{a}_1^4 \right),$$

where

$$\begin{aligned} &c_0 = n(n^3 + 6n^2 + 21n + 18), \qquad c_1 = 2n(2n^2 + 6n + 9), \\ &c_2 = 2n(3n + 2), \qquad c_3 = n(2n^2 + 5n + 7). \end{aligned} \tag{2.1}$$

Using these consistent estimator of $\hat{a}_3$ and $\hat{a}_4$, consistent estimators of η_i and ξ_i are obtained in this paper.

3. Testing the equality of the two covariance matrices

Let $\boldsymbol{x}_{ij}$ be independently distributed as $N_m(\boldsymbol{\mu}_i, \Sigma_i)$, $i = 1, 2, j = 1, \ldots, N_i$. When $N_i > m$ the likelihood ratio test does not depend on the sample mean vectors $\bar{\boldsymbol{x}}_i$ but only on the sample covariance matrices $n_i^{-1}V_i$, $n_i = N_i - 1$. In fact, it depends only on the eigenvalues of $V_1^{-1}V_2$. Thus, we shall consider only those tests that are based on V_i, $i = 1, 2$.

3.1. A test based on the eigenvalues

We shall assume without loss of generality that $n_1 \geq n_2$. Since $n_i < m$, the inverse of V_i does not exist. Thus, instead Srivastava [2] considered the Moore–Penrose inverse of V_1^+ and proposed a test for testing the equality of two covariance matrices based on the statistic G_2 in (1.5) with $\hat{b} = \hat{a}_1^2/\hat{a}_2$, where $\hat{a}_1$ and $\hat{a}_2$ are defined in (1.7). The asymptotic distribution of G_2 under the hypothesis, as $m \to \infty$ is $\chi^2_{n_1 n_2}$ for fixed n_1 and n_2 : $G_2 \xrightarrow{d} \chi^2_{n_1 n_2}$. Following Srivastava [2], the proof can be obtained. The G_2-statistic, however, does not perform well, and thus we shall not consider this test any further.

3.2. Two tests based on an estimator of a distance

For testing the hypothesis $H : \Sigma_1 = \Sigma_2 = \Sigma$ against the alternative $A \neq H$, that is the alternative that either $\Sigma_1 - \Sigma_2 \geq 0$ or $\Sigma_2 - \Sigma_1 \geq 0$, where for a matrix A, $A \geq 0$ denotes that A is at least positive semi-definite (p.s.d), we define the square of the distance by

$$d^2 = \frac{1}{m} \mathrm{tr}\, (\Sigma_1 - \Sigma_2)^2, \tag{3.1}$$

It can be shown that d is a distance function. It may be noted that d^2 is known as Frobenius norm whose many interesting properties have been investigated by Ledoit and Wolf [7]. Now

$$d^2 = a_{21} + a_{22} - \frac{2}{m} \mathrm{tr}\, (\Sigma_1 \Sigma_2).$$

Using consistent estimators of a_{21}, a_{22} and $\mathrm{tr}\,(\Sigma_1\Sigma_2)$, as given in Srivastava [6] and stated in Section 2, Schott [1] proposed a test based on the statistic J_2, which under the hypothesis H is asymptotically distributed as $N(0, 1)$ under the assumptions (A3)–(A4) and $m/n_i \to c_i \in (0, \infty)$, $i = 1, 2$, as $(m, n) \to \infty$. We note that

$$d^2 \geq a_{21} + a_{22} - 2(a_{11}a_{22})^{1/2} = (a_{21}^{1/2} - a_{22}^{1/2})^2 \geq 0.$$

Thus, we may consider a test statistic based on a consistent estimator of a_{21} and a_{22}, namely $\hat{a}_{21} - \hat{a}_{22}$. Thus, Srivastava [3] considered the statistic T_2, which under the hypothesis H and assumptions (A1)–(A4) is asymptotically distributed as $N(0, 1)$ for $(m, n) \to \infty$.

3.3. The proposed test

We note that when $\Sigma_1 = \Sigma_2$, then $\mathrm{tr}\ \Sigma_1 = \mathrm{tr}\ \Sigma_2$, $\mathrm{tr}\ \Sigma_1^2 = \mathrm{tr}\ \Sigma_2^2$, etc. Thus we must have, under the null hypothesis, $\gamma_1 = \gamma_2$, where

$$\gamma_i = \frac{\mathrm{tr}\ \Sigma_i^2/m}{(\mathrm{tr}\ \Sigma_i/m)^2}, \quad i = 1, 2.$$

A consistent estimator of γ_i is given by (1.13). Under the null hypothesis $\hat{\gamma}_i$ is asymptotically normally distributed with mean γ_i and variance ξ_i^2 given in (1.14), see Theorem 3.1 of Srivastava [6]. Thus we propose a test based on the test statistic Q_2 given in (1.12) where $\hat{\xi}_i^2$ are consistent estimators of ξ_i, obtained by substituting consistent estimators of $\hat{a}_i$, $i = 1, \ldots, 4$. The asymptotic distribution of Q_2 as $(n, m) \to \infty$ is $N(0, 1)$ under the null hypothesis that the two covariance matrices are equal. Equivalently Q_2^2 is asymptotically distributed as χ_1^2.

4. Testing the equality of several covariance matrices

In this section we consider the problem of testing the hypothesis

$$H : \Sigma_1 = \cdots = \Sigma_k \equiv \Sigma, \text{ say} \tag{4.1}$$

against the alternative

$$A : \Sigma_i \neq \Sigma_j \text{ for at least one pair } (i, j), \quad i \neq j,\ i, j = 1, \ldots, k,$$

based on N_i independent observations $\boldsymbol{x}_{ij}$ from the ith population, $i = 1, \dots, k, j = 1, \dots, N_i$. When $\Sigma_i = \Sigma$ for all i, we estimate Σ by

$$\hat{\Sigma} = \frac{1}{n}V = \frac{1}{n}(V_1 + \cdots + V_k), \quad n = n_1 + \cdots + n_k. \tag{4.2}$$

We now describe the generalized version of the three tests J_2, T_2, and Q_2 for testing the hypothesis

$$H : \Sigma_1 = \cdots = \Sigma_k \quad \text{vs } A \neq H.$$

For k groups, the Frobenius norm is given by

$$d^2 = \frac{1}{m}\sum_{i<j}^{k} \operatorname{tr}(\Sigma_i - \Sigma_j)^2.$$

Using a consistent estimator of d^2, Schott [1] proposed the test J_k which is normally distributed as $N(0, 1)$ under the null hypothesis, assumptions (A1)–(A4), and $\lim_{(m,n_i\to\infty)} m/n_i = c_i \in (0, \infty), i = 1, \dots, k$.

To describe the generalized version of T_2 and Q_2, or equivalently of T_2^2 and Q_2^2 as the alternative is two-sided, we remind the reader that in the definition of various terms such as γ_i, η_i, ξ_i or their estimates remain the same. Thus i runs from 1 to k, $V = V_1 + \cdots + V_k$, $n = n_1 + \cdots + n_k$ etc.

From Lemma 2.2, it follows that under the hypothesis H, $\hat{a}_{2i}$ are independently normally distributed with common mean a_2 and variances η_i^2. The common mean a_2 is estimated by weighing $\hat{a}_{2i}$ with the weights inversely proportional to its estimated variances. That is, we estimate a_2 by $\bar{\hat{a}}_2$ given in (1.18), which is also a consistent estimator of a_2. Let

$$\bar{\hat{a}}_2^* = \frac{\sum_{i=1}^{k} \hat{a}_{2i}/\eta_i^2}{\sum_{i=1}^{k} 1/\eta_i^2} = \frac{\mathbf{1}'D\hat{\boldsymbol{a}}_2}{\mathbf{1}'D\mathbf{1}}, \tag{4.3}$$

where $\mathbf{1} = (1, \dots, 1)'$, a k-vector of ones, D is a $k \times k$ diagonal matrix, $D = \operatorname{diag}(1/\eta_1^2, \dots, 1/\eta_k^2)$, and $\hat{\boldsymbol{a}}_2 = (\hat{a}_{21}, \dots, \hat{a}_{2k})'$. Note that $Cov(\hat{\boldsymbol{a}}_2) = D^{-1}$. We derive $D^{1/2}(\hat{\boldsymbol{a}}_2 - a_2\mathbf{1}) \xrightarrow{d} N_k(\mathbf{0}, I_k)$ as $(n, m) \to \infty$ under the null hypothesis, where $\mathbf{0} = (0, \dots, 0)'$, a k-vector of zeros. Moreover, we have $\{I_k - \mathbf{1}\mathbf{1}'D/(\mathbf{1}'D\mathbf{1})\}'D\{I_k - \mathbf{1}\mathbf{1}'D/(\mathbf{1}'D\mathbf{1})\} = D - D\mathbf{1}\mathbf{1}'D/(\mathbf{1}'D\mathbf{1})$. Thus, we get

$$\begin{aligned} T_k^2 &= \sum_{i=1}^{k} \frac{(\hat{a}_{2i} - \bar{\hat{a}}_2)^2}{\hat{\eta}_i^2} \overset{p}{=} \sum_{i=1}^{k} \frac{(\hat{a}_{2i} - \bar{\hat{a}}_2^*)^2}{\eta_i^2} \\ &= (\hat{\boldsymbol{a}}_2 - \bar{\hat{a}}_2\mathbf{1})'D(\hat{\boldsymbol{a}}_2 - \bar{\hat{a}}_2\mathbf{1}) = \left(\hat{\boldsymbol{a}}_2 - \frac{\mathbf{1}\mathbf{1}'D\hat{\boldsymbol{a}}_2}{\mathbf{1}'D\mathbf{1}}\right)' D \left(\hat{\boldsymbol{a}}_2 - \frac{\mathbf{1}\mathbf{1}'D\hat{\boldsymbol{a}}_2}{\mathbf{1}'D\mathbf{1}}\right) \\ &= \hat{\boldsymbol{a}}_2' \left(D - \frac{D\mathbf{1}\mathbf{1}'D}{\mathbf{1}'D\mathbf{1}}\right)\hat{\boldsymbol{a}}_2 = \hat{\boldsymbol{a}}_2' D^{1/2}\left(I_k - \frac{D^{1/2}\mathbf{1}\mathbf{1}'D^{1/2}}{\mathbf{1}'D\mathbf{1}}\right) D^{1/2}\hat{\boldsymbol{a}}_2. \end{aligned}$$

Note that $\{I_k - D^{1/2}\mathbf{1}\mathbf{1}'D^{1/2}/(\mathbf{1}'D\mathbf{1})\}(D^{1/2}a_2\mathbf{1}) = \mathbf{0}$. Hence, T_k^2 is asymptotically distributed as chi-square with $k - 1$ degrees of freedom, χ_{k-1}^2 since $I_k - D^{1/2}\mathbf{1}\mathbf{1}'D^{1/2}/(\mathbf{1}'D\mathbf{1})$ is an idempotent matrix. Thus we propose T_k^2 as a test statistic for testing the equality of k covariance matrices, that is, for testing $H : \Sigma_1 = \cdots = \Sigma_k$ against the alternative $A \neq H$. We state this result in the following theorem.

Theorem 4.1. *Let V_i be independently distributed as $W_m(\Sigma, n_i)$, $V = V_1 + \cdots + V_k$, $n = n_1 + \cdots + n_k$, $\hat{a}_{2i}$ and $\bar{\hat{a}}_2$ defined in (1.4) and (1.18) respectively, and $\hat{\eta}_i$ are consistent estimators of η_i. Then the statistic T_k^2 defined in (1.17) is distributed as χ_{k-1}^2 under the assumption* (A1)–(A4) *as $(m, n) \to \infty$.*

To obtain the generalized version of Q_2 to test the equality of k covariance matrices, we define the estimator of the common value of $\gamma_1 = \cdots = \gamma_k = \gamma$ by $\bar{\hat{\gamma}}$ given in (1.20), and propose the test statistic Q_k^2 in (1.19). Following the step of Theorem 4.1 in obtaining the distribution of T_k^2, we obtain the following theorem.

Theorem 4.2. *Let V_i be independently distributed as $W_m(\Sigma, n_i)$, $V = V_1 + \cdots + V_k$, $n = n_1 + \cdots + n_k$, $\hat{\gamma}_i$ and $\bar{\hat{\gamma}}$ defined in (1.13) and (1.20) respectively. Then under the assumptions* (A1)–(A4), *$Q_k^2 \xrightarrow{d} \chi_{k-1}^2$ as $(m, n) \to \infty$.*

5. Comparison of power of various tests and Attained Significance Level with the nominal value 0.05

To demonstrate how our tests perform, we carry out several simulations. We first consider the two-sample case for the test statistics J_2, T_2, Q_2 we define the Attained Significance Level (ASL) as

$$\hat{\alpha}_{J_2} = \frac{\#\text{ of } J_{2H} > z_\alpha}{r}, \qquad \hat{\alpha}_{T_2} = \frac{\#\text{ of } T_{2H}^2 > \chi_{1,\alpha}^2}{r}, \qquad \hat{\alpha}_{Q_2} = \frac{\#\text{ of } Q_{2H}^2 > \chi_{1,\alpha}^2}{r},$$

respectively, where r is the number of replications, z_α is the upper $100\alpha\%$ point of the $N(0, 1)$ distribution, and $\chi_{f,\alpha}^2$ is the upper $100\alpha\%$ point of the chi-square distribution with f degrees of freedom. The ASL is used to get an idea of how close the empirical distributions of J_2, T_2^2, and Q_2^2 are to their asymptotic ones. Here, J_{2H}, T_{2H}, and Q_{2H} are values of the test statistics computed from data simulated under H.

Based on the empirical distributions constructed from the above simulations, we define $\hat{z}_\alpha$ as the upper $100\alpha\%$ point of the empirical distribution of J_2 and $\hat{\chi}_{1,\alpha,i}^2$, $i = 1, 2$ as the upper $100\alpha\%$ point of the empirical distributions of T_2^2 and Q_2^2, respectively. We then define the Attained Power of J_2, T_2^2, and Q_2^2 as

$$\hat{\beta}_{J_2} = \frac{\#\text{ of } J_{2A} > \hat{z}_\alpha}{r}, \qquad \hat{\beta}_{T_2} = \frac{\#\text{ of } T_{2A}^2 > \hat{\chi}_{1,\alpha,1}^2}{r}, \qquad \hat{\beta}_{Q_2} = \frac{\#\text{ of } Q_{2A}^2 > \hat{\chi}_{1,\alpha,2}^2}{r},$$

respectively. Here, J_{2A}, T_{2A}^2, and Q_{2A}^2 are values of the test statistics computed from data simulated under A.

In our simulation, we selected $r = 1000$. Let $\Omega = \text{diag}(\omega_1, \ldots, \omega_m)$, where $\omega_1, \ldots, \omega_m \sim i.i.d.\ Unif(1, 5)$, and Δ_j $(j = 0, 1, 2)$ be a $m \times m$ matrix whose (a, b)th element are defined by $(-1)^{a+b}\{0.2 \times (j+2)\}^{|a-b|^{1/10}}$ We considered the following hypothesis testing setup:

$$H: \quad \Sigma_1 = \Sigma_2 = \Sigma = \Omega\Delta_0\Omega,$$
$$A: \quad \Sigma_1 = \Sigma \quad \text{and} \quad \Sigma_2 = \Omega\Delta_2\Omega.$$

We computed the ASL under H and Attained Power under A. Table 1 presents the ASL and Attained Power for the tests J_2, T_2^2, and Q_2^2 for our hypothesis testing setup. In Table 2, we present the results for $k = 3$. We consider the hypothesis and the alternative as follows:

$$H: \quad \Sigma_1 = \Sigma_2 = \Sigma_3 = \Sigma = \Omega\Delta_0\Omega,$$
$$A: \quad \Sigma_1 = \Sigma, \qquad \Sigma_2 = \Omega\Delta_1\Omega, \quad \text{and} \quad \Sigma_3 = \Omega\Delta_2\Omega.$$

We find that all three tests have good performance for large m. The power of T_k^2, however, is low for small sample sizes, and the power of J_k is also low when $n_1 = n_2 = 10$ and $n_1 = n_2 = n_3 = 10$. The power of Q_k^2 is higher than that of T_k^2. The reason why the power of Q_k^2 becomes better than that of T_k^2 may be that the Q_k^2-statistic measures the difference of covariances by both tr Σ_i and tr Σ_i^2 although the T_k^2-statistic measures the difference of covariances by only tr Σ_i. Moreover, in the all cases, the ASL of J_k greatly exceeds 0.05. Its difference tends to be large when the sample size increases. As for the ASL, the test based on T_k^2 will be the best among three tests. The reason why the ASL of T_k^2 is better than that of Q_k^2 may be that an estimation of η_i^2 standardizing T_k^2 is easier than that of ξ_i^2 standardizing Q_k^2, because ξ_i^2 depends on many terms than η_i^2. The difficulty in getting a fast convergent estimator of η_i^2 makes the convergence of the test statistic Q_k^2 to the chi-square distribution slow but still better than the convergence of J_k to normal. Taking into consideration the attained ASL as well as power, it appears that the test based on Q_k^2 may be preferred as the ASL for the J_k test is significantly higher than the prescribed level most of the times unless m is very large.

6. Proofs of Theorems 2.1 and 2.2

In order to prove the consistency of the estimates of $\hat{a}_3$ and $\hat{a}_4$, we need the following lemma.

Lemma 6.1. *Let $V \sim W_m(\Sigma, n)$. Then*

(a) $E(V^2) = n(n+1)\Sigma^2 + n(\text{tr}\Sigma)\Sigma,$

(b) $E(V^3) = n(n^2 + 3n + 4)\Sigma^3 + 2n(n+1)(\text{tr}\Sigma)\Sigma^2 + n(n+1)(\text{tr}\Sigma^2)\Sigma + n(\text{tr}\Sigma)^2\Sigma,$

(c) $$E(V^4) = n(n^3 + 6n^2 + 21n + 18)\Sigma^4 + n(3n^2 + 9n + 14)(\text{tr}\Sigma)\Sigma^3 + 3n(n+1)(\text{tr}\Sigma)^2\Sigma^2 + n(2n^2 + 5n + 7)(\text{tr}\Sigma^2)\Sigma^2 + n(3n+2)(\text{tr}\Sigma)(\text{tr}\Sigma^2)\Sigma + n(n^2 + 3n + 4)(\text{tr}\Sigma^3)\Sigma + n(\text{tr}\Sigma)^3\Sigma.$$

Proof of Lemma 6.1. The proof of (a) is given by Srivastava and Khatri [5] in Problem 3.2. Following in the same manner, the proofs of (b) and (c) can be obtained. □

From the above lemma, the following corollary is easily obtained.

Table 1
ASL (under H) and Attained Power (under A) of J_2, T_2^2 and Q_2^2.

m	$n_1 = n_2$	ASL			Power		
		J_2	T_2^2	Q_2^2	J_2	T_2^2	Q_2^2
20	10	0.085	0.054	0.091	0.474	0.243	0.855
	20	0.107	0.066	0.076	0.892	0.352	0.968
	40	0.082	0.050	0.052	1.000	0.633	1.000
	60	0.094	0.051	0.053	1.000	0.753	1.000
40	10	0.081	0.053	0.081	0.403	0.271	0.918
	20	0.096	0.054	0.061	0.748	0.479	0.992
	40	0.107	0.059	0.061	0.999	0.785	1.000
	60	0.094	0.053	0.058	1.000	0.908	1.000
60	10	0.072	0.050	0.060	0.424	0.342	0.910
	20	0.099	0.054	0.059	0.849	0.603	0.968
	40	0.096	0.051	0.047	1.000	0.893	1.000
	60	0.103	0.045	0.051	1.000	0.974	1.000
100	10	0.091	0.059	0.072	0.440	0.343	0.922
	20	0.084	0.045	0.046	0.946	0.619	0.997
	40	0.093	0.061	0.069	1.000	0.848	1.000
	60	0.120	0.066	0.066	1.000	0.952	1.000
200	10	0.071	0.039	0.046	0.629	0.487	0.923
	20	0.082	0.042	0.047	0.962	0.742	0.997
	40	0.092	0.048	0.043	1.000	0.959	1.000
	60	0.097	0.046	0.050	1.000	0.995	1.000

Table 2
ASL (under H) and Attained Power (under A) of J_3, T_3^2 and Q_3^2.

m	$n_1 = n_2 = n_3$	ASL			Power		
		J_3	T_3^2	Q_3^2	J_3	T_3^2	Q_3^2
20	10	0.081	0.069	0.034	0.952	0.424	0.760
	20	0.093	0.060	0.044	1.000	0.737	0.956
	40	0.102	0.053	0.047	1.000	0.961	1.000
	60	0.091	0.043	0.045	1.000	0.996	1.000
40	10	0.085	0.075	0.026	0.606	0.240	0.835
	20	0.087	0.050	0.028	1.000	0.451	0.976
	40	0.089	0.052	0.039	1.000	0.751	0.999
	60	0.103	0.059	0.048	1.000	0.876	1.000
60	10	0.083	0.073	0.022	0.464	0.182	0.869
	20	0.098	0.065	0.044	0.986	0.244	0.980
	40	0.101	0.054	0.045	1.000	0.472	1.000
	60	0.105	0.045	0.039	1.000	0.657	1.000
100	10	0.069	0.066	0.026	0.571	0.249	0.860
	20	0.111	0.066	0.046	0.963	0.408	0.984
	40	0.099	0.048	0.039	1.000	0.760	1.000
	60	0.111	0.061	0.050	1.000	0.870	1.000
200	10	0.077	0.062	0.025	0.591	0.291	0.890
	20	0.097	0.059	0.040	0.992	0.441	0.995
	40	0.092	0.058	0.043	1.000	0.813	1.000
	60	0.108	0.054	0.049	1.000	0.948	1.000

Corollary 6.1. *Let $V \sim W_m(\Sigma, n)$. Then*

(a) $E(\mathrm{tr}V^2) = n(n+1)\mathrm{tr}\Sigma^2 + n(\mathrm{tr}\Sigma)^2$,

(b) $E(\mathrm{tr}V^3) = n(n^2+3n+4)\mathrm{tr}\Sigma^3 + 3n(n+1)\mathrm{tr}\Sigma^2\mathrm{tr}\Sigma + n(\mathrm{tr}\Sigma)^3$,

(c) $E(\mathrm{tr}V^4) = c_0\mathrm{tr}\Sigma^4 + c_1(\mathrm{tr}\Sigma^3)\mathrm{tr}\Sigma + c_2\mathrm{tr}\Sigma^2(\mathrm{tr}\Sigma)^2 + c_3(\mathrm{tr}\Sigma^2)^2 + n(\mathrm{tr}\Sigma)^4$,

where c_0, c_1, c_2, and c_3 have been defined in (2.1).

It has been shown in Srivastava [6] that consistent unbiased estimators of $a_i = \mathrm{tr}\ \Sigma^i/m, i = 1, 2$ are given by $\hat{a}_i$. To obtain a consistent estimator of a_3, one may use consistent estimators of a_1 and a_2 in (b), and obtain an expression of $\hat{a}_3$, which is the same as given in Theorem 2.1. But proving the consistency of $\hat{a}_3$, that is, the proof of Theorem 2.1 still remains. To prove it, we need the following lemma which can be proved by using the results of chi-square distributions and the moments of quadratic forms.

Lemma 6.2. *Let* $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m$ *be i.i.d.* $N_n(\boldsymbol{0}, I_n)$, $w_{ij} = \boldsymbol{u}_i'\boldsymbol{u}_j$. *Then*

(a) $E(w_{ii}^r) = n(n+2)\cdots(n+2r-2)$, $r = 1, 2, \ldots,$
(b) $E(w_{ij}^2) = n$, $i \neq j$,
(c) $E(w_{ii}w_{ij}^2) = n(n+2)$, $i \neq j$,
(d) $E(w_{ij}w_{jk}w_{ki}) = n$, $i \neq j \neq k$,
(e) $Var(w_{ii}^3) = 6n(n+2)(n+4)(3n^2+30n+80)$,
(f) $E(w_{ii}^2 w_{ij}^4) = 3n(n+2)(n+4)(n+6)$, $i \neq j$,
(g) $E(w_{ii}^3 w_{jj} w_{ij}^2) = n(n+2)^2(n+4)(n+6)$, $i \neq j$,
(h) $E(w_{ii}^4 w_{jj}^2) = n^2(n+2)^2(n+4)(n+6)$, $i \neq j$,
(i) $E(w_{jk}w_{kl}w_{lj})^2 = n(n+2)(n+8)$, $j \neq k \neq l$,
(j) $E(w_{ij}^4) = 3n(n+2)$, $i \neq j$,
(k) $E(w_{ij}^3 w_{jk} w_{ki} w_{kk}) = 3n(n+2)^2$, $i \neq j \neq k$.

The following lemma gives an alternative expression of $E(\mathrm{tr}\, V^3)$.

Lemma 6.3. *An alternative expression of* $E(\mathrm{tr} V^3)$ *is given by*

$$E(\mathrm{tr} V^3) = n(n+2)(n+4)\sum_{i=1}^{m}\lambda_i^3 + 3n(n+2)\sum_{i\neq j}^{m}\lambda_i^2\lambda_j + n\sum_{i\neq j\neq k}^{m}\lambda_i\lambda_j\lambda_k, \tag{6.1}$$

where $\lambda_1 \geq \cdots \geq \lambda_m$ *are the eigenvalues of this covariance matrix* Σ.

Proof of Lemma 6.3. We have $V \sim W_m(\Sigma, n)$. Thus, we can write $V = YY'$, where $Y = (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n)$ and $\boldsymbol{y}_i$ are *i.i.d.* $N_m(\boldsymbol{0}, \Sigma)$. Let Γ be an orthogonal matrix such that $\Gamma\Sigma\Gamma' = \Lambda$, where Λ is a $m \times m$ diagonal matrix, $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_m)$ and λ_i are the eigenvalues of Σ. Then, if $\mathcal{E} = (\boldsymbol{\varepsilon}_1, \ldots, \boldsymbol{\varepsilon}_n)$, where $\boldsymbol{\varepsilon}_i$ are *i.i.d.* $N_m(\boldsymbol{0}, I_m)$, $Y = \Sigma^{1/2}\mathcal{E}$, and $\Sigma^{1/2}\Sigma^{1/2} = \Sigma$. Thus

$$V = \Sigma^{1/2}\mathcal{E}\mathcal{E}'\Sigma^{1/2} = \Gamma'\Lambda^{1/2}\Gamma\mathcal{E}\mathcal{E}'\Gamma'\Lambda^{1/2}\Gamma = \Gamma'\Lambda^{1/2}UU'\Lambda^{1/2}\Gamma,$$

where $U' = \mathcal{E}'\Gamma' = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m)$ and $\boldsymbol{u}_i$ are *i.i.d.* $N_n(\boldsymbol{0}, I_n)$. Thus $w_{ii} = \boldsymbol{u}_i'\boldsymbol{u}_i$ are *i.i.d.* χ_n^2, $i = 1, \ldots, m$. Let $w_{ij} = \boldsymbol{u}_i'\boldsymbol{u}_j$. This gives

$$\mathrm{tr}\, V^3 = \sum_{i=1}^{m}\lambda_i^3 w_{ii}^3 + 3\sum_{i\neq j}^{m}\lambda_i^2\lambda_j w_{ii}w_{ij}^2 + \sum_{i\neq j\neq k}^{m}\lambda_i\lambda_j\lambda_k w_{ij}w_{jk}w_{ki}. \tag{6.2}$$

Hence, from Lemma 6.2, we obtain the expectation in (6.1). □

6.1. *Proof of* Theorem 2.1

We note

$$\hat{a}_2 = \frac{1}{mn^2}\left\{\mathrm{tr}\, V^2 - \frac{1}{n}(\mathrm{tr}\, V)^2\right\} + O_2^*,$$

and

$$\begin{aligned}\hat{a}_3 &= \frac{1}{mn^3}\left(\mathrm{tr}\, V^3 - 3n^2\hat{a}_1\hat{a}_2 - nm\hat{a}_1^3\right) + O_3^* \\ &= \frac{1}{mn^3}\left\{\mathrm{tr}\, V^3 - \frac{3}{n}\mathrm{tr}\, V\,\mathrm{tr}\, V^2 + \frac{2}{n^2}(\mathrm{tr}\, V)^3\right\} + O_3^*,\end{aligned} \tag{6.3}$$

where O_2^* and O_3^* denote terms which tend to 0 under the conditions (A1)–(A4). Following the notation in the proof of Lemma 6.3, we can write

$$\mathrm{tr}\, V\,\mathrm{tr}\, V^2 = \sum_{i=1}^{m}\lambda_i^3 w_{ii}^3 + \sum_{i\neq j}^{m}\lambda_i^2\lambda_j(w_{ii}^2 w_{jj} + 2w_{ii}w_{ij}^2) + \sum_{i\neq j\neq k}^{m}\lambda_i\lambda_j\lambda_k w_{ii}w_{jk}^2, \tag{6.4}$$

$$(\mathrm{tr}\, V)^3 = \sum_{i=1}^{m}\lambda_i^3 w_{ii}^3 + 3\sum_{i\neq j}^{m}\lambda_i^2\lambda_j w_{ii}^2 w_{jj} + \sum_{i\neq j\neq k}^{m}\lambda_i\lambda_j\lambda_k w_{ii}w_{jj}w_{kk}. \tag{6.5}$$

By substituting (6.2), (6.4) and (6.5) into (6.3), we can write $\hat{a}_3$ as

$$\hat{a}_3 = \hat{a}_3^{(1)} + \hat{a}_3^{(2)} + \hat{a}_3^{(3)} + \hat{a}_3^{(4)} + O_3^*, \tag{6.6}$$

where

$$\hat{a}_3^{(1)} = \frac{1}{mn^3}\left(1 - \frac{3}{n} + \frac{2}{n^2}\right)\sum_{i=1}^{m}\lambda_i^3 w_{ii},$$

$$\hat{a}_3^{(2)} = \frac{3}{mn^3}\left(1 - \frac{2}{n}\right)\sum_{i\neq j}^{m}\lambda_i^2\lambda_j\left(w_{ii}w_{ij}^2 - \frac{1}{n}w_{ii}^2 w_{jj}\right),$$

$$\hat{a}_3^{(3)} = \frac{1}{mn^3}\sum_{i\neq j\neq k}^{m}\lambda_i\lambda_j\lambda_k\left(w_{ij}w_{jk}w_{ki} - \frac{1}{n^2}w_{ii}w_{jj}w_{kk}\right),$$

$$\hat{a}_3^{(4)} = -\frac{3}{mn^4}\sum_{i\neq j\neq k}^{m}\lambda_i\lambda_j\lambda_k\left(w_{ij}w_{jk}^2 - \frac{1}{n}w_{ii}w_{jj}w_{kk}\right).$$

We note that

$$E(\hat{a}_3^{(1)}) = \frac{1}{mn^3}\left(1 - \frac{3}{n} + \frac{2}{n^2}\right)E\left(\sum_{i=1}^{m}\lambda_i^3 w_{ii},\right) = a_3 + O(n^{-1}),$$

$$\begin{aligned}Var(\hat{a}_3^{(1)}) &= \frac{1}{m^2n^4}\left(1 - \frac{3}{n} + \frac{2}{n^2}\right)^2\sum_{i=1}^{m}\lambda_i^6 Var(w_{ii}^3)\\ &= \frac{1}{m}\sum_{i=1}^{m}\lambda_i^6 O(m^{-1}n^{-1}) \to 0, \quad \text{as } (n, m) \to \infty.\end{aligned}$$

Thus, in probability $\hat{a}_3^{(1)} \xrightarrow{p} a_3$ as $(n, m) \to \infty$, where 'p' stands for 'in probability'.

To prove the theorem, we need to show that the remaining terms on the right side of (6.6) go to zero as $(n, m) \to \infty$. The expected value of $\hat{a}_3^{(2)}$ is

$$E(\hat{a}_3^{(2)}) = \frac{3}{mn^3}\left(1 - \frac{2}{n}\right)\sum_{i\neq j}^{m}\lambda_i^2\lambda_j E\left(w_{ii}w_{ij}^2 - \frac{1}{n}w_{ii}^2 w_{jj}\right) = 0.$$

Similarly, $Cov(r_{ij}, r_{kl}) = 0$ for $(i \neq j) \neq (k \neq l)$, where $r_{ij} = w_{ii}(w_{ij}^2 - w_{ii}w_{jj}/n)$. Hence,

$$\begin{aligned}Var(\hat{a}_3^{(2)}) &= \frac{9}{m^2n^6}\left(1 - \frac{2}{n}\right)^2\sum_{i\neq j}^{m}\lambda_i^4\lambda_j^2 Var(r_{ij})\\ &= \frac{9}{m^2n^6}\left(1 - \frac{2}{n}\right)^2\sum_{i\neq j}^{m}\lambda_i^4\lambda_j^2(n-2)\left(1 - \frac{1}{n}\right)E(w_{ii}^4)\\ &= \frac{9}{n}\left(a_2a_4 - \frac{1}{m}a_6\right)O(1) = O(n^{-1}).\end{aligned}$$

Hence $\hat{a}_3^{(2)} \xrightarrow{p} 0$ as $(n, m) \to \infty$.

The third term $\hat{a}_3^{(3)}$ is given by $\hat{a}_3^{(3)} = \sum_{i\neq j\neq k}^{m}\lambda_i\lambda_j\lambda_k s_{ijk}/(mn^3)$, where $s_{ijk} = w_{ij}w_{jk}w_{ki} - w_{ii}w_{jj}w_{kk}/n^2$. It can easily be seen that $E(\hat{a}_3^{(3)}) = 0$ and $Cov(s_{i_1j_1k_1}, s_{i_2j_2k_2}) = 0$ for $(i_1 \neq j_1 \neq k_1) \neq (i_2 \neq j_2 \neq k_2)$. Hence

$$Var(\hat{a}_3^{(3)}) = \frac{1}{m^2n^6}\sum_{i\neq j\neq k}^{m}\lambda_i^2\lambda_j^2\lambda_k^2 Var(s_{ijk}) = \frac{1}{m^3}\sum_{i\neq j\neq k}^{m}\lambda_i^2\lambda_j^2\lambda_k^2 O(mn^{-3}),$$

from Lemma 6.2 as

$$Var(s_{ijk}) = \left(1 + \frac{2}{n}\right)(n^3 - n^2 + 4n - 4) = O(n^3).$$

Similarly $E(\hat{a}_3^{(4)}) = 0$ and $Var(\hat{a}_3^{(4)}) = O(mn^{-4})$ can be shown. Thus, all the remainder terms go to zero and $\hat{a}_3 \xrightarrow{p} a_3$ if $n = O(m^\delta)$, $\delta > 1/3$, as $(n, m) \to \infty$. □

6.2. *Proof of* Theorem 2.2

We first note that we can write

$$\begin{aligned}\operatorname{tr} V^4 &= \operatorname{tr}\left(\sum_{i=1}^{m}\lambda_i^2 w_{ii}\boldsymbol{u}_i\boldsymbol{u}_i' + \sum_{i\neq j}^{m}\lambda_i\lambda_j w_{ij}\boldsymbol{u}_i\boldsymbol{u}_j'\right)^2 \\ &= \sum_{i=1}^{m}\lambda_i^4 w_{ii}^4 + \sum_{i\neq j}^{m}\lambda_i^2\lambda_j^2 w_{ii}w_{jj}w_{ij}^2 + 4\sum_{i\neq j}^{m}\lambda_i^3\lambda_j w_{ii}^2 w_{ij}^2 + 2\sum_{i\neq j\neq k}^{m}\lambda_i\lambda_j\lambda_k^2 w_{ij}w_{jk}w_{ki}w_{kk} \\ &\quad + 2\sum_{i\neq j\neq k}^{m}\lambda_i\lambda_j^2\lambda_k w_{ij}^2 w_{kj}^2 + \sum_{i\neq j}^{m}\lambda_i^2\lambda_j^2 w_{ij}^4 + \sum_{i\neq j\neq k\neq l}^{m}\lambda_i\lambda_j\lambda_k\lambda_l w_{ij}w_{kl}w_{jk}w_{li}.\end{aligned}$$

By combining the terms from $\hat{a}_1$, $\hat{a}_2$, and $\hat{a}_3$, it can be shown as in the case of $\hat{a}_3$, that $\hat{a}_4$ is a consistent estimator of a_4 for $n = O(m^\delta)$, $\delta > 1/2$. □

7. Concluding remarks

In this article, we considered four tests for testing the equality of $k \geq 2$ covariance matrices. One of them, namely the G test does not perform well. All the remaining three tests, namely, J_k, T_k^2, and Q_k^2 tests perform well when m or n is large. For small n and m, the tests J_k and Q_k^2 perform well from the power consideration. The Average Significance Level (ASL) of the J_k test fluctuates considerably, often much larger than the specified level. Thus, taking into consideration the power as well as ASL, the Q_k^2 test may be preferred over J_k test.

Acknowledgments

This research was supported by NSERC of Canada. We wish to thank the two referees for their helpful suggestions. The computation for Tables 1 and 2 was carried out by Kelvin Yuen to whom we express our sincere thanks.

References

[1] J.R. Schott, A test for the equality of covariance matrices when the dimension is large relative to the sample size, Comput. Statist. Data Anal. 51 (2007) 6535–6542.
[2] M.S. Srivastava, Multivariate theory for analyzing high-dimensional data, J. Japan Statist. Soc. 37 (2007) 53–86.
[3] M.S. Srivastava, Testing the equality of two covariance matrices and independence of two sub-vectors with fewer observations than the dimension, International Conference on Advances in Interdisciplinary Statistics and Combinatorics, Oct 12–14, 2007, University of North Carolina at Greensboro, 2007, NC, USA.
[4] M.S. Srivastava, Singular Wishart and multivariate beta distributions, Ann. Statist. 31 (2003) 1537–1560.
[5] M.S. Srivastava, C.G. Khatri, An Introduction to Multivariate Statistics, North Holland, New York, 1979.
[6] M.S. Srivastava, Some tests concerning the covariance matrix in high-dimensional data, J. Japan Statist. Soc. 35 (2005) 251–272.
[7] O. Ledoit, M. Wolf, A well conditioned estimator for large-dimensional covariance matrices, J. Multivariate Anal. 88 (2004) 365–411.