Journal of Multivariate Analysis 171 (2019) 412–420

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

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

High-dimensional testing for proportional covariance matrices

Koji Tsukuda [a,*], Shun Matsuura [b]

[a] Graduate School of Arts and Sciences, The University of Tokyo, 3-8-1 Komaba, Meguro-ku, Tokyo 153-8902, Japan
[b] Faculty of Science and Technology, Keio University, 3-14-1 Hiyoshi, Kohoku-ku, Yokohama, Kanagawa 223-8522, Japan

ARTICLE INFO

Article history:
Received 31 May 2018
Available online 2 February 2019

AMS 2010 subject classifications:
primary 62H15
secondary 62F03
62F05

Keywords:
Asymptotic test
High-dimension
Multivariate normal distribution
Proportional covariance model

ABSTRACT

Hypothesis testing for the proportionality of covariance matrices is a classical statistical problem and has been widely studied in the literature. However, there have been few treatments of this test in high-dimensional settings, especially for the case where the number of variables is larger than the sample size, despite high-dimensional statistical inference having recently received considerable attention. This paper studies hypothesis testing for the proportionality of two covariance matrices in the high-dimensional setting: $m, n \asymp p^{\delta}$ for some $\delta \in (1/2, 1)$, where m and n denote the sample sizes and p denotes the number of variables. A test statistic is proposed and its asymptotic distribution is derived under multivariate normality. The non-asymptotic performance of the proposed test procedure is numerically examined.

1. Introduction

Let $\boldsymbol{x} \sim \mathcal{N}_p(\boldsymbol{\mu}_x, \boldsymbol{\Sigma}_x)$ and $\boldsymbol{y} \sim \mathcal{N}_p(\boldsymbol{\mu}_y, \boldsymbol{\Sigma}_y)$ be independent p-dimensional random vectors with $\mathrm{rank}(\boldsymbol{\Sigma}_x) = \mathrm{rank}(\boldsymbol{\Sigma}_y) = p$, and let $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$ with $m \geq 3$ and $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n$ with $n \geq 3$ be independent and identically distributed copies of $\boldsymbol{x}$ and $\boldsymbol{y}$, respectively. This paper discusses hypothesis testing for the proportionality of $\boldsymbol{\Sigma}_x$ and $\boldsymbol{\Sigma}_y$: the null hypothesis $\mathcal{H}_0$ and the alternative hypothesis $\mathcal{H}_1$ are expressed as

$$\mathcal{H}_0 : \ \boldsymbol{\Sigma}_x = k\boldsymbol{\Sigma}_y \text{ for some } k \in (0, \infty), \tag{1}$$

$$\mathcal{H}_1 : \ \boldsymbol{\Sigma}_x \neq k\boldsymbol{\Sigma}_y \text{ for any } k \in (0, \infty). \tag{2}$$

Without loss of generality, $m \geq n$ is assumed throughout this paper.

In the theory of multivariate statistical analysis, numerous hypothesis testing techniques for covariance matrices have been developed. The hypothesis (1), which is the target of this paper, is a typical problem in two-sample hypothesis testing; see, e.g., Manly and Rayner [16]. When quadratic discriminant analysis is used, it is known that (1), which is called the proportional covariance model or proportional scatter model, provides good performances in many cases [7,8]. Moreover, (1) is sometimes a natural assumption in the quantitative genetics domain [6,10]. Hence, estimating covariance matrices under the constraint (1) has been studied extensively; see, e.g., [2,4,6,9,16]. To judge whether (1) can be assumed, it is important to test the hypothesis (1). We will provide a brief review of this test problem in the next paragraph, and in Remark 6 after presenting our proposed approach. For other hypothesis testing regarding covariance matrices, readers are referred to [2,16].

* Corresponding author.
E-mail address: ctsukuda@g.ecc.u-tokyo.ac.jp (K. Tsukuda).

https://doi.org/10.1016/j.jmva.2019.01.011

Testing (1) against (2) is one of the classical problems in multivariate hypothesis testing, and was first studied by Federer [5] using an approach based on a likelihood ratio test. The likelihood ratio test requires the condition $p < \min(m, n) = n$, and under $\mathcal{H}_0$ twice the logarithm of its statistic converges in distribution to a chi-squared variable as

$$m, n \to \infty \text{ with fixed } p. \tag{3}$$

Following the work of Federer [5], this problem has frequently been considered [6]. Schott [18] proposed a Wald statistic under (3) without the normality assumption. However, the tests derived under (3) do not necessarily work well as the dimensionality (the number of variables) p becomes larger. To overcome this problem, Xu et al. [24] and Liu et al. [15] constructed tests under the setting where m, n, p increase simultaneously. In particular, Liu et al. [15] proposed a test procedure that was established under

$$m, n, p \to \infty \quad \text{with} \quad p/m \to c_1 \in (0, 1), \quad p/n \to c_2 \in (0, \infty). \tag{4}$$

Alternatively, we consider the asymptotic setting

$$p \to \infty, \quad m = m(p) \asymp p^\delta, \quad n = n(p) \asymp p^\delta, \quad \delta \in (0, 1), \tag{5}$$

where $m \asymp p^\delta$ means that there exist positive constants C_1 and C_2 such that $C_1 p^\delta \le m \le C_2 p^\delta$ for all sufficiently large p. Obviously, (5) implies that $p/m, p/n \to \infty$ which clearly differs from (3) and (4). Several settings implying that $p/m, p/n \to \infty$ have been considered in other hypothesis testing. Specifically, in [20–22], several tests for covariance matrices have been considered under (5). Under the setting (5) with $\delta \in (1/2, 1)$, we will propose a test procedure for the proportionality of two covariance matrices assuming multivariate normality. The proposed procedure is designed for the case where the dimensionality p is larger than the sample sizes m, n. Under our assumptions that observations follow the normal distribution and the eigenvalues are asymptotically bounded (see Remark 3), the situations to which our procedure is applicable are limited. The relaxation of these assumptions is left for future research.

1.1. Notations

The indicator function is denoted by $\mathbf{1}(\cdot)$. The (i, j)-element of a matrix $\boldsymbol{A}$ is denoted by $(\boldsymbol{A})_{i,j}$. For a matrix $\boldsymbol{A}$, $\boldsymbol{A}^2$ means $\boldsymbol{A}\boldsymbol{A}$. Throughout this paper, the asymptotic setting that m, n and p simultaneously tend to infinity is considered, and the limit operation is denoted by $p \to \infty$. For a sequence of random variables (or random vectors) $\boldsymbol{X} = \boldsymbol{X}_p$ with $p \in \{1, 2, \ldots\}$ and a random variable (or a random vector) $\boldsymbol{Y}$, $\boldsymbol{X} \rightsquigarrow \boldsymbol{Y}$ means that $\boldsymbol{X}$ converges to $\boldsymbol{Y}$ in distribution. Moreover, for a constant $\boldsymbol{b}$, $\boldsymbol{X} \to^p \boldsymbol{b}$ means that $\boldsymbol{X}$ converges to $\boldsymbol{b}$ in probability. As for the stochastic order symbols $O_P(\cdot)$ and $o_P(\cdot)$, see Section 2.2 of van der Vaart [23].

1.2. Organization of this paper

In Section 2, several assumptions on the covariance matrices are introduced. Section 3 presents preliminary results that will be used in subsequent sections. Following these preparations, a procedure to test the proportionality of two covariance matrices is proposed in Section 4. The performance of the proposed procedure is verified in Section 5 through a simulation study. Proofs of two theoretical results, Lemma 1 and Theorem 2, are included in Appendix A. Some technical details of the proof of Lemma 1 are presented in the Online Supplement.

2. Assumptions on covariance matrices

For $i \in \{1, \ldots, 8\}$, define the following quantities:

$$a_{x,i}(p) = \frac{1}{p}\operatorname{tr}(\boldsymbol{\Sigma}_x^i), \quad a_{y,i}(p) = \frac{1}{p}\operatorname{tr}(\boldsymbol{\Sigma}_y^i), \quad a_{xy}(p) = \frac{1}{p}\operatorname{tr}(\boldsymbol{\Sigma}_x\boldsymbol{\Sigma}_y).$$

Remark 1. The null hypothesis (1) holds if and only if

$$\frac{\boldsymbol{\Sigma}_x}{a_{x,1}(p)} = \frac{\boldsymbol{\Sigma}_y}{a_{y,1}(p)}, \tag{6}$$

which is further equivalent to

$$\frac{a_{x,2}(p)}{\{a_{x,1}(p)\}^2} + \frac{a_{y,2}(p)}{\{a_{y,1}(p)\}^2} - \frac{2a_{xy}(p)}{a_{x,1}(p)a_{y,1}(p)} = 0.$$

We will construct a test statistic based on the above facts.

To derive the asymptotic distribution of a test statistic, it is assumed that $\boldsymbol{\Sigma}_x$ and $\boldsymbol{\Sigma}_y$ satisfy

$$\lim_{p\to\infty} a_{x,i}(p) = a_{x,i}, \quad \lim_{p\to\infty} a_{y,i}(p) = a_{y,i}, \quad 0 < a_{x,i}, a_{y,i} < \infty \tag{7}$$

for $i \in \{1, \ldots, 8\}$, and

$$\lim_{p\to\infty} a_{xy}(p) = a_{xy}, \quad 0 < a_{xy} < \infty. \tag{8}$$

Remark 2. Under $\mathcal{H}_0$, as $a_{x,2}(p) = a_{y,2}(p) = a_{xy}(p)$, (8) is equivalent to (7) for $i = 2$. It will be proved that a test statistic converges in distribution under $\mathcal{H}_0$, so (7) is sufficient to show our theoretical result.

Remark 3. Assumption (7) is frequently used in the context of high-dimensional statistics; see, e.g., [19,20]. This assumption means that the moments of the eigenvalues of $\boldsymbol{\Sigma}_x$ and $\boldsymbol{\Sigma}_y$ are uniformly bounded in p up to eighth order. In some actual data analyses, this assumption is too strong. To address such cases, the spiked structures proposed by Johnstone [11] have been applied in different asymptotic regimes; see, e.g., Section 2 of [12]. This new regime requires totally different asymptotic discussions.

3. Preliminary results

The unbiased sample covariance matrices of $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$ and $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n$ are denoted by $\boldsymbol{S}_x$ and $\boldsymbol{S}_y$, respectively, that is

$$\boldsymbol{S}_x = \frac{1}{m-1}\sum_{i=1}^{m}(\boldsymbol{x}_i - \bar{\boldsymbol{x}})(\boldsymbol{x}_i - \bar{\boldsymbol{x}})^\top, \quad \boldsymbol{S}_y = \frac{1}{n-1}\sum_{i=1}^{n}(\boldsymbol{y}_i - \bar{\boldsymbol{y}})(\boldsymbol{y}_i - \bar{\boldsymbol{y}})^\top,$$

where $\bar{\boldsymbol{x}} = (\boldsymbol{x}_1 + \cdots + \boldsymbol{x}_m)/m$ and $\bar{\boldsymbol{y}} = (\boldsymbol{y}_1 + \cdots + \boldsymbol{y}_n)/n$.

Let us introduce the following statistics:

$$\hat{a}_{x,1} = \frac{1}{p}\,\mathrm{tr}(\boldsymbol{S}_x), \quad \hat{a}_{y,1} = \frac{1}{p}\,\mathrm{tr}(\boldsymbol{S}_y), \quad \hat{a}_{xy} = \frac{1}{p}\,\mathrm{tr}(\boldsymbol{S}_x\boldsymbol{S}_y),$$
$$\hat{a}_{x,2} = \frac{(m-1)^2}{p(m-2)(m+1)}\left[\mathrm{tr}(\boldsymbol{S}_x^2) - \frac{\{\mathrm{tr}(\boldsymbol{S}_x)\}^2}{m-1}\right], \quad \hat{a}_{y,2} = \frac{(n-1)^2}{p(n-2)(n+1)}\left[\mathrm{tr}(\boldsymbol{S}_y^2) - \frac{\{\mathrm{tr}(\boldsymbol{S}_y)\}^2}{n-1}\right].$$

Note that the so-called trace estimators $\hat{a}_{x,1}, \hat{a}_{x,2}, \hat{a}_{y,1}, \hat{a}_{y,2}$ were originally introduced by Bai and Saranadasa [1] in the context of the two-sample high-dimensional mean testing problem. Other unbiased trace estimators that do not require the normality assumption have been discussed in [3,14,22,25].

Remark 4. Under our assumptions, Lemma 2.1 and Theorem 2.1 of Srivastava [20] imply that the following properties hold for $\hat{a}_{x,i}$ and $\hat{a}_{y,i}$ with $i \in \{1, 2\}$. For $i \in \{1, 2\}$, $\mathrm{E}(\hat{a}_{x,i}) = a_{x,i}(p)$ and $\mathrm{E}(\hat{a}_{y,i}) = a_{y,i}(p)$. Moreover, for $i \in \{1, 2\}$,

$$\hat{a}_{x,i} \to^p a_{x,i}, \quad \hat{a}_{y,i} \to^p a_{y,i} \tag{9}$$

and

$$p^{(1+\delta)/2}\{\hat{a}_{x,1} - a_{x,1}(p)\} = O_P(1), \quad p^{(1+\delta)/2}\{\hat{a}_{y,1} - a_{y,1}(p)\} = O_P(1), \tag{10}$$
$$p^{\delta}\{\hat{a}_{y,2} - a_{y,2}(p)\} = O_P(1), \quad p^{\delta}\{\hat{a}_{x,2} - a_{x,2}(p)\} = O_P(1)$$

as $p \to \infty$ under assumptions (5) and (7). Note that these properties do not generally hold when our assumptions fail.

Remark 5. As $\mathrm{E}(\hat{a}_{xy}) = a_{xy}(p)$ and $\mathrm{var}(\hat{a}_{xy}) = O(p^{-2\delta})$, if (8) holds, then $\hat{a}_{xy} \to^p a_{xy}$ as $p \to \infty$.

In the next section, a test procedure will be proposed. To prove the convergence in distribution of our test statistic, we will use the following lemma, which corresponds to the two-sample version of Theorem 1 of Schott [19] in which the asymptotic setting $p/(m+n) \not\to \infty$ is considered. The relation between Z, which will be defined in (11), and $t_{(m+n)p}$ in [19] is

$$t_{(m+n)p} = p(m+n)Z/(mn).$$

Our proof employs a similar strategy to Schott [19].

Lemma 1. *Assume that (5) with $\delta > 1/2$ and (7) hold. If $\boldsymbol{\Sigma}_x = \boldsymbol{\Sigma}_y$ then*

$$Z = \frac{mn}{m+n}(\hat{a}_{x,2} + \hat{a}_{y,2} - 2\hat{a}_{xy}) \tag{11}$$

converges in distribution to $\mathcal{N}(0, 4a_2^2)$ as $p \to \infty$, where $a_2 = a_{x,2} = a_{y,2}$.

4. Test statistic

Let us consider the following statistic:

$$T=\frac{mn}{m+n}\left\{\frac{\hat{a}_{x,2}}{(\hat{a}_{x,1})^2}+\frac{\hat{a}_{y,2}}{(\hat{a}_{y,1})^2}-\frac{2\hat{a}_{xy}}{\hat{a}_{x,1}\hat{a}_{y,1}}\right\}.$$

Then, we have the following theorem.

Theorem 2. *Assume that (5) with $\delta > 1/2$ and (7) hold. Under $\mathcal{H}_0$, $T \rightsquigarrow \mathcal{N}(0, 4b_2^2)$ as $p \to \infty$, where $b_2 = a_{x,2}/(a_{x,1})^2 = a_{y,2}/(a_{y,1})^2$.*

As b_2 is unknown, we use a consistent estimator. Define

$$\hat{b}_2=\frac{m^2}{m^2+n^2}\frac{\hat{a}_{x,2}}{(\hat{a}_{x,1})^2}+\frac{n^2}{m^2+n^2}\frac{\hat{a}_{y,2}}{(\hat{a}_{y,1})^2}.$$

The ratio of the weights is m^2/n^2 because the ratio of the leading terms of the asymptotic variances of $\hat{a}_{x,2}/(\hat{a}_{x,1})^2$ and $\hat{a}_{y,2}/(\hat{a}_{x,1})^2$ is n^2/m^2 under $\mathcal{H}_0$. For this quantity, Lemma 3 follows from (9).

Lemma 3. *Assume that* (5) and (7) *hold. Under $\mathcal{H}_0$, $\hat{b}_2 \to^p b_2$ as $p \to \infty$.*

We now propose the test statistic $\tilde{T} = T/(2\hat{b}_2)$ and a one-sided test procedure for the proportionality of the covariance matrices by regarding $\tilde{T}$ as a standard normal variable, i.e., (1) is rejected when $\tilde{T} > \Phi^{-1}(1-\alpha)$, where Φ^{-1} is the quantile function of the standard normal distribution, $\mathcal{N}(0, 1)$, and α is an assigned significance level. This procedure is based on the fact that if $\mathcal{H}_0$ is true, $\tilde{T} \rightsquigarrow \mathcal{N}(0, 1)$ as $p \to \infty$ under assumptions (5) with $\delta > 1/2$ and (7), which follows from Theorem 2 and Lemma 3 combined with Slutsky's Lemma.

Remark 6. In the literature, there are two approaches for testing (1) against (2): using the likelihood ratio statistics or using the F-matrices $\boldsymbol{S}_x^{-1}\boldsymbol{S}_y$. Federer [5] used the former approach, for which a Bartlett correction was derived by Eriksen [4]. As $\ln|\boldsymbol{S}_x|$ and $\ln|\boldsymbol{S}_y|$ are used for their test procedures, the condition $p < \min(m, n) = n$ is required. The F-matrices approach is useful because (1) is equivalent to the sphericity of $\boldsymbol{\Sigma}_x^{-1}\boldsymbol{\Sigma}_y$. The properties of the F-matrices have been investigated by Khatri [13], Pillai et al. [17] and Zheng [26]. The test developed by Liu et al. [15] was constructed using a closely related approach. For this approach, $\boldsymbol{S}_x$ must be full rank and hence the condition $p < \max(m, n) = m$ is necessary. Moreover, Xu et al. [24] applied a pseudo-likelihood method for sphericity, assuming that

$$m, n, p \to \infty \quad \text{with} \quad p/m \to c_1 \in (0, 1), \quad p/n \to c_2 \in (0, 1).$$

In contrast, our approach is designed for the case $m < p < n^2$. Furthermore, a statistic proposed by Srivastava and Yanagihara [21] (Q_2 in their notation) can be used to test the proportionality hypothesis although this perspective is not mentioned in their paper.

5. Simulation study

In this section, the non-asymptotic performance of the proposed procedure is verified by conducting a simulation study. Let us describe the simulation setting. The number of iterations is 10,000, the significance level α is 0.05, and the number of variables p and the sample sizes are set as follows:

(I) $p = d^3$ and $m = n = d^2(= p^{2/3})$ for $d \in \{3, 4, 5, 6, 7, 8, 9, 10\}$;

(II) $p = d^5$ and $m = n = d^3(= p^{3/5})$ for $d \in \{2, 3, 4\}$.

The covariance matrix $\boldsymbol{\Sigma}_x$ is always the identity matrix $\boldsymbol{I}_p$ of size p, and $\boldsymbol{\Sigma}_y$ is set as follows:

(A) $\boldsymbol{\Sigma}_y = \boldsymbol{I}_p$;

(B) $\boldsymbol{\Sigma}_y = 5\boldsymbol{I}_p$;

(C) $(\boldsymbol{\Sigma}_y)_{i,j} = \{1 + 4(i-1)/(p-1)\} \times \mathbf{1}(i = j)$ for $i, j \in \{1, \ldots, p\}$;

(D) $(\boldsymbol{\Sigma}_y)_{i,j} = \mathbf{1}(i = j) + 0.4 \times \mathbf{1}(|i-j| = 1)$ for $i, j \in \{1, \ldots, p\}$;

(E) $(\boldsymbol{\Sigma}_y)_{i,j} = \mathbf{1}(i = j) + 0.5 \times \mathbf{1}(|i-j| = 1)$ for $i, j \in \{1, \ldots, p\}$;

(F) $(\boldsymbol{\Sigma}_y)_{i,j} = 0.4^{|i-j|}$ for $i, j \in \{1, \ldots, p\}$;

(G) $(\boldsymbol{\Sigma}_y)_{i,j} = 0.5^{|i-j|}$ for $i, j \in \{1, \ldots, p\}$.

The rejection rates of $\mathcal{H}_0$ obtained under the above settings are summarized in Tables 1 and 2. For cases (A) and (B), the numbers in the table represent the type I error rates, and for the other cases they represent the powers of the test. The results can be summarized as follows:

Table 1
Results (rejection rates of $\mathcal{H}_0$) of simulation study (I) $p = d^3$, $m = n = d^2$.

d	Size		Power				
	(A)	(B)	(C)	(D)	(E)	(F)	(G)
3	0.0353	0.0355	0.0675	0.1094	0.1698	0.1258	0.2246
4	0.0401	0.0384	0.1169	0.2639	0.4603	0.3252	0.6150
5	0.0488	0.0451	0.2102	0.5229	0.8160	0.6236	0.9421
6	0.0514	0.0478	0.3387	0.8123	0.9859	0.9001	0.9991
7	0.0489	0.0477	0.5325	0.9662	0.9996	0.9927	1.0000
8	0.0464	0.0475	0.7198	0.9995	1.0000	1.0000	1.0000
9	0.0469	0.0456	0.8872	1.0000	1.0000	1.0000	1.0000
10	0.0486	0.0490	0.9686	1.0000	1.0000	1.0000	1.0000

Table 2
Results (rejection rates of $\mathcal{H}_0$) of simulation study (II) $p = d^5$, $m = n = d^3$.

d	Size		Power				
	(A)	(B)	(C)	(D)	(E)	(F)	(G)
2	0.0372	0.0331	0.0569	0.0917	0.1368	0.1022	0.1835
3	0.0420	0.0470	0.2288	0.5923	0.8795	0.7001	0.9745
4	0.0506	0.0493	0.7290	0.9981	1.0000	0.9999	1.0000

(1) The proposed test is conservative for almost all settings.
(2) The Type I error rate approaches α $(= 0.05)$ as d increases, i.e., m, n and p increase.
(3) The power monotonically increases as d increases.
(4) By comparing settings (I) $d = 4$ and (II) $d = 2$, both of which have $p/m = p/n = 4$, and by comparing (I) $d = 9$ and (II) $d = 3$, both of which have $p/m = p/n = 9$, we see that the power of (I) is greater than that of (II) when p/m is the same.

Let us discuss the result of (C)–(G) in more detail. For case (C), in which the non-diagonal elements of $\boldsymbol{\Sigma}_y$ (the covariances of $\boldsymbol{y}$) are all 0 (which equal the non-diagonal elements of $\boldsymbol{\Sigma}_x$), but the diagonal elements of $\boldsymbol{\Sigma}_y$ (the variances of $\boldsymbol{y}$) are different from those of $\boldsymbol{\Sigma}_x$, the power becomes greater as d increases. The power gets closer to 1 as d increases in cases (D) and (E), too, where the variances of $\boldsymbol{y}$ are all 1 and two consecutive variables are correlated. The power in case (E), with correlation coefficients 0.5, is greater than that in case (D), with correlation coefficients 0.4, for the same values of d. Similar results were obtained in cases (F) and (G), although their powers go to 1 more rapidly than in cases (D) and (E). This may be because all p variables of $\boldsymbol{y}$ are correlated in cases (F) and (G), unlike cases (D) and (E).

Acknowledgments

This work was partly supported by the Japan Society for the Promotion of Science KAKENHI, Japan Grant Number 18K13454 (KT). Shun Matsuura would like to thank Mr. Shinya Ogasawara of National Statistics Center in Japan for his suggestion of the topic of this paper. The authors would like to express their gratitude to the Editor-in-Chief, Christian Genest, and the referees who provided us valuable comments.

Appendix A. Proofs

Proof of Lemma 1. Assume that $\boldsymbol{\Sigma}_x = \boldsymbol{\Sigma}_y$ and let us denote them by $\boldsymbol{\Sigma}$. As Z is invariant if $\boldsymbol{S}_x$ and $\boldsymbol{S}_y$ are transformed to $\boldsymbol{P}^\top \boldsymbol{S}_x \boldsymbol{P}$ and $\boldsymbol{P}^\top \boldsymbol{S}_y \boldsymbol{P}$ with any orthogonal matrix $\boldsymbol{P}$, we assume that $\boldsymbol{\Sigma}$ is diagonal without loss of generality. Let us denote the (i, j)-element of $\boldsymbol{S}_w$ by $s_{ij,w}$ for $w = x, y$, and the (i, i)-element of $\boldsymbol{\Sigma}$ by σ_{ii}. Then, we can write

$$s_{ij,x} = \frac{1}{m-1}\sigma_{ii}^{1/2}\sigma_{jj}^{1/2}\boldsymbol{z}_{i,x}^\top \boldsymbol{z}_{j,x}, \quad s_{ij,y} = \frac{1}{n-1}\sigma_{ii}^{1/2}\sigma_{jj}^{1/2}\boldsymbol{z}_{i,y}^\top \boldsymbol{z}_{j,y},$$

with two independent iid sequences $\boldsymbol{z}_{1,x}, \ldots, \boldsymbol{z}_{p,x}$ and $\boldsymbol{z}_{1,y}, \ldots, \boldsymbol{z}_{p,y}$, where for each $h \in \{1, \ldots, p\}$, $\boldsymbol{z}_{h,x}$ is an $(m-1)$-dimensional standard normal variable and $\boldsymbol{z}_{h,y}$ is an $(n-1)$-dimensional standard normal variable.

Letting

$$\rho_{1,h} = \frac{mn}{p(m+n)}\frac{(m-1)^2}{(m+1)(m-2)}\left(2\sum_{\ell=1}^{h-1} s_{\ell h,x}^2 + s_{hh,x}^2\right), \quad \rho_{2,h} = \frac{mn}{p(m+n)}\frac{(n-1)^2}{(n+1)(n-2)}\left(2\sum_{\ell=1}^{h-1} s_{\ell h,y}^2 + s_{hh,y}^2\right),$$

$$\rho_{3,h} = 2\frac{mn}{p(m+n)}\left(2\sum_{\ell=1}^{h-1} s_{\ell h,x}s_{\ell h,y} + s_{hh,x}s_{hh,y}\right), \quad \rho_{4,h} = \frac{mn}{p(m+n)}\frac{m-1}{(m+1)(m-2)}\left(2s_{hh,x}\sum_{\ell=1}^{h-1} s_{\ell\ell,x} + s_{hh,x}^2\right)$$

and

$$\rho_{5,h} = \frac{mn}{p(m+n)} \frac{n-1}{(n+1)(n-2)} \left(2s_{hh,y} \sum_{\ell=1}^{h-1} s_{\ell\ell,y} + s_{hh,y}^2 \right),$$

we define a sequence $\tau_1, \ldots, \tau_p$ of random variables by $\tau_h = \rho_{1,h} + \rho_{2,h} - \rho_{3,h} - \rho_{4,h} - \rho_{5,h}$ for each $h \in \{1, \ldots, p\}$. Then, $Z = \tau_1 + \cdots + \tau_p$. Moreover, let $\rho_{i,h}^* = \rho_{i,h} - \mathrm{E}(\rho_{i,h})$ for each $i \in \{1, \ldots, 5\}$. For each $h \in \{1, \ldots, p\}$, let us introduce a sigma-field $\mathcal{F}_h^p$ generated by $\{\boldsymbol{z}_{1,x}, \ldots, \boldsymbol{z}_{h,x}\}$ and $\{\boldsymbol{z}_{1,y}, \ldots, \boldsymbol{z}_{h,y}\}$. Then, $\tau_1, \ldots, \tau_p$ is a martingale difference sequence with respect to the filtration $\{\mathcal{F}_1^p, \ldots, \mathcal{F}_p^p\}$; see the proof of Theorem 1 in [19]. As we will show that

$$\sum_{h=1}^{p} \mathrm{E}(\tau_h^2 | \mathcal{F}_{h-1}^p) \to^p 4a_2^2 \tag{A.1}$$

and

$$\sum_{h=1}^{p} \mathrm{E}(\tau_h^4) \to 0, \tag{A.2}$$

the Lyapunov type Martingale Central Limit Theorem yields the conclusion.

Proof of (A.1). It is sufficient to show

$$\sum_{h=1}^{p} \mathrm{E}(\tau_h^2) \to 4a_2^2 \tag{A.3}$$

and

$$\mathrm{var}\left\{ \sum_{h=1}^{p} \mathrm{E}(\tau_h^2 | \mathcal{F}_{h-1}^p) \right\} \to 0. \tag{A.4}$$

Proof of (A.3). As $\tau_1, \ldots, \tau_p$ is a martingale difference sequence, one has

$$\mathrm{E}(\tau_1^2) + \cdots + \mathrm{E}(\tau_p^2) = \mathrm{E}\{(\tau_1 + \cdots + \tau_h)^2\} = \mathrm{var}(Z).$$

It also holds that

$$\mathrm{var}(Z) = \frac{m^2 n^2}{(m+n)^2} \{\mathrm{var}(\hat{a}_{x,2}) + \mathrm{var}(\hat{a}_{y,2}) + 4\mathrm{var}(\hat{a}_{xy}) - 4\mathrm{cov}(\hat{a}_{x,2}, \hat{a}_{xy}) - 4\mathrm{cov}(\hat{a}_{y,2}, \hat{a}_{xy})\}.$$

From $\mathrm{cov}(\hat{a}_{x,2}, \hat{a}_{xy}) = o(p^{-1}m^{-2}n^{-1})$ and $\mathrm{cov}(\hat{a}_{y,2}, \hat{a}_{xy}) = o(p^{-1}m^{-1}n^{-2})$ (see Subsection 4.2 of [22]), it follows that

$$\mathrm{var}(Z) = \frac{m^2 n^2}{(m+n)^2} \left(\frac{4a_2^2}{m^2} + \frac{4a_2^2}{n^2} + 4\frac{2a_2^2}{mn} \right) + o(1) \to 4a_2^2.$$

Proof of (A.4). From $\tau_h = (\rho_{1,h}^* - \rho_{4,h}^*) + (\rho_{2,h}^* - \rho_{5,h}^*) - \rho_{3,h}^*$, it follows that

$$\tau_h^2 = (\rho_{1,h}^* - \rho_{4,h}^*)^2 + (\rho_{2,h}^* - \rho_{5,h}^*)^2 + (\rho_{3,h}^*)^2 + 2(\rho_{1,h}^* - \rho_{4,h}^*)(\rho_{2,h}^* - \rho_{5,h}^*) - 2(\rho_{1,h}^* - \rho_{4,h}^*)\rho_3^* - 2(\rho_{2,h}^* - \rho_{5,h}^*)\rho_{3,h}^*.$$

Hence, to see $\mathrm{var}[\sum_{h=1}^p \mathrm{E}(\tau_h^2 | \mathcal{F}_{h-1}^p)\} \to 0$, it is sufficient to prove

$$\mathrm{var}\left[\sum_{h=1}^{p} \mathrm{E}\{(\rho_{1,h}^* - \rho_{4,h}^*)^2 | \mathcal{F}_{h-1}^p\} \right] \to 0, \tag{A.5}$$

$$\mathrm{var}\left[\sum_{h=1}^{p} \mathrm{E}\{(\rho_{2,h}^* - \rho_{5,h}^*)^2 | \mathcal{F}_{h-1}^p\} \right] \to 0, \tag{A.6}$$

$$\mathrm{var}\left[\sum_{h=1}^{p} \mathrm{E}\{(\rho_{3,h}^*)^2 | \mathcal{F}_{h-1}^p\} \right] \to 0, \tag{A.7}$$

$$\mathrm{var}\left[\sum_{h=1}^{p} \mathrm{E}\{(\rho_{1,h}^* - \rho_{4,h}^*)(\rho_{2,h}^* - \rho_{5,h}^*) | \mathcal{F}_{h-1}^p\} \right] = 0, \tag{A.8}$$

$$\mathrm{var}\left[\sum_{h=1}^{p} \mathrm{E}\{(\rho_{1,h}^* - \rho_{4,h}^*)\rho_3^* | \mathcal{F}_{h-1}^p\} \right] = 0, \tag{A.9}$$

and

$$\mathrm{var}\left[\sum_{h=1}^{p}\mathrm{E}\{(\rho_{2,h}^{*}-\rho_{5,h}^{*})\rho_{3,h}^{*}|\mathcal{F}_{h-1}^{p}\}\right]=0. \tag{A.10}$$

Here, let us briefly explain these assertions; the detailed calculations are presented in the Online Supplement. For (A.5), one has

$$\begin{aligned}
&\mathrm{var}\left[\sum_{h=1}^{p}\mathrm{E}\{(\rho_{1,h}^{*}-\rho_{4,h}^{*})^2|\mathcal{F}_{h-1}^{p}\}\right]\\
&=\left\{\frac{mn}{p(m+n)}\frac{m-1}{(m+1)(m-2)}\right\}^4\mathrm{var}\left[\sum_{h=1}^{p}\left\{8\sum_{\ell=1}^{h-1}\frac{(m-2)\sigma_{\ell\ell}^2\sigma_{hh}^2}{(m-1)^3}(\boldsymbol{z}_{\ell,x}^{\top}\boldsymbol{z}_{\ell,x})^2\right.\right.\\
&\quad\left.\left.+8\sum_{\ell=1}^{h-1}\sum_{\ell'\neq\ell}^{h-1}\frac{\sigma_{\ell\ell}\sigma_{\ell'\ell'}\sigma_{hh}^2}{(m-1)^2}(\boldsymbol{z}_{\ell,x}^{\top}\boldsymbol{z}_{\ell',x})^2-8\sum_{\ell=1}^{h-1}\sum_{\ell'\neq\ell}^{h-1}\frac{\sigma_{\ell\ell}\sigma_{\ell'\ell'}\sigma_{hh}^2}{(m-1)^3}(\boldsymbol{z}_{\ell,x}^{\top}\boldsymbol{z}_{\ell,x})(\boldsymbol{z}_{\ell',x}^{\top}\boldsymbol{z}_{\ell',x})\right\}\right]\\
&\leq 3\left\{\frac{mn}{p(m+n)}\frac{m-1}{(m+1)(m-2)}\right\}^4\left[\mathrm{var}\left\{\sum_{h=1}^{p}8\sum_{\ell=1}^{h-1}\frac{(m-2)\sigma_{\ell\ell}^2\sigma_{hh}^2}{(m-1)^3}(\boldsymbol{z}_{\ell,x}^{\top}\boldsymbol{z}_{\ell,x})^2\right\}\right.\\
&\quad\left.+\mathrm{var}\left\{8\sum_{\ell=1}^{h-1}\sum_{\ell'\neq\ell}^{h-1}\frac{\sigma_{\ell\ell}\sigma_{\ell'\ell'}\sigma_{hh}^2}{(m-1)^2}(\boldsymbol{z}_{\ell,x}^{\top}\boldsymbol{z}_{\ell',x})^2\right\}+\mathrm{var}\left\{8\sum_{\ell=1}^{h-1}\sum_{\ell'\neq\ell}^{h-1}\frac{\sigma_{\ell\ell}\sigma_{\ell'\ell'}\sigma_{hh}^2}{(m-1)^3}(\boldsymbol{z}_{\ell,x}^{\top}\boldsymbol{z}_{\ell,x})(\boldsymbol{z}_{\ell',x}^{\top}\boldsymbol{z}_{\ell',x})\right\}\right].
\end{aligned}$$

We can also write

$$\begin{aligned}
&\left\{\frac{mn}{p(m+n)}\frac{m-1}{(m+1)(m-2)}\right\}^4\mathrm{var}\left[\sum_{h=1}^{p}\left\{8\sum_{\ell=1}^{h-1}\frac{(m-2)\sigma_{\ell\ell}^2\sigma_{hh}^2}{(m-1)^3}(\boldsymbol{z}_{\ell,x}^{\top}\boldsymbol{z}_{\ell,x})^2\right\}\right]\\
&=512\,\frac{m^4n^4(m+2)}{p^4(m+n)^4(m-2)^2(m-1)(m+1)^3}\sum_{\ell=1}^{p-1}\left(\sum_{h=\ell+1}^{p}\sigma_{hh}^2\right)^2\sigma_{\ell\ell}^4\\
&\leq 512\,\frac{m^4n^4(m+2)}{p(m+n)^4(m-2)^2(m-1)(m+1)^3}\frac{\mathrm{tr}(\boldsymbol{\Sigma}^4)}{p}\left\{\frac{\mathrm{tr}(\boldsymbol{\Sigma}^2)}{p}\right\}^2\to 0.
\end{aligned}$$

Moreover, one has

$$\begin{aligned}
&\left\{\frac{mn}{p(m+n)}\frac{m-1}{(m+1)(m-2)}\right\}^4\mathrm{var}\left[\sum_{h=1}^{p}\left\{8\sum_{\ell=1}^{h-1}\sum_{\ell'\neq\ell}^{h-1}\frac{\sigma_{\ell\ell}\sigma_{\ell'\ell'}\sigma_{hh}^2}{(m-1)^2}(\boldsymbol{z}_{\ell,x}^{\top}\boldsymbol{z}_{\ell',x})^2\right\}\right]\\
&=64\,\frac{m^4n^4}{p^4(m+n)^4(m-2)^4(m+1)^4}\left\{2\sum_{\ell=1}^{p-1}\sum_{\ell'\neq\ell}^{p-1}\sum_{h=\max\{\ell,\ell'\}+1}^{p}\sum_{h'=\max\{\ell,\ell'\}+1}^{p}\sigma_{\ell\ell}^2\sigma_{\ell'\ell'}^2\sigma_{hh}^2\sigma_{h'h'}^2 2(m-1)(m+2)\right.\\
&\quad\left.+4\sum_{\ell=1}^{p-1}\sum_{\ell'\neq\ell}^{p-1}\sum_{k\neq\ell,\ell'}^{p-1}\sum_{h=\max\{\ell,\ell'\}+1}^{p}\sum_{h'=\max\{\ell,k\}+1}^{p}\sigma_{\ell\ell}^2\sigma_{\ell'\ell'}\sigma_{kk}\sigma_{hh}^2\sigma_{h'h'}^2 2(m-1)\right\}\\
&\leq 64\,\frac{m^4n^4}{(m+n)^4(m-2)^4}\left[4\frac{(m-1)(m+2)}{(m+1)^4}\left\{\frac{\mathrm{tr}(\boldsymbol{\Sigma}^2)}{p}\right\}^4+8\frac{(m-1)p}{(m+1)^4}\left\{\frac{\mathrm{tr}(\boldsymbol{\Sigma})}{p}\right\}^2\left\{\frac{\mathrm{tr}(\boldsymbol{\Sigma}^2)}{p}\right\}^3\right]
\end{aligned}$$

and

$$\begin{aligned}
&\left\{\frac{mn}{p(m+n)}\frac{m-1}{(m+1)(m-2)}\right\}^4\mathrm{var}\left[\sum_{h=1}^{p}\left\{8\sum_{\ell=1}^{h-1}\sum_{\ell'\neq\ell}^{h-1}\frac{\sigma_{\ell\ell}\sigma_{\ell'\ell'}\sigma_{hh}^2}{(m-1)^3}(\boldsymbol{z}_{\ell,x}^{\top}\boldsymbol{z}_{\ell,x})(\boldsymbol{z}_{\ell',x}^{\top}\boldsymbol{z}_{\ell',x})\right\}\right]\\
&=512\,\frac{m^4n^4}{p^4(m+n)^4(m-2)^4(m+1)^4}\left[\left\{\sum_{\ell=1}^{p-1}\sum_{\ell'\neq\ell}^{p-1}\sum_{h=\max\{\ell,\ell'\}+1}^{p}\sum_{h'=\max\{\ell,\ell'\}+1}^{p}\sigma_{\ell\ell}^2\sigma_{\ell'\ell'}^2\sigma_{hh}^2\sigma_{h'h'}^2 m\right\}\right.
\end{aligned}$$

$$+\left\{\sum_{\ell=1}^{p-1}\sum_{\ell'\neq\ell}^{p-1}\sum_{k\neq\ell,\ell'}^{p-1}\sum_{h=\max\{\ell,\ell'\}+1}^{p}\sum_{h'=\max\{\ell,k\}+1}^{p}\sigma_{\ell\ell}^2\sigma_{\ell'\ell'}\sigma_{kk}\sigma_{hh}^2\sigma_{h'h'}^2(m-1)\right\}\Bigg]$$
$$\leq 512\,\frac{m^4n^4}{(m+n)^4(m-2)^4}\left[\frac{m}{(m+1)^4}\left\{\frac{\mathrm{tr}(\boldsymbol{\Sigma}^2)}{p}\right\}^4+\frac{p(m-1)}{(m+1)^4}\left\{\frac{\mathrm{tr}(\boldsymbol{\Sigma})}{p}\right\}^2\left\{\frac{\mathrm{tr}(\boldsymbol{\Sigma}^2)}{p}\right\}^3\right].$$

These right-hand sides converge to 0 because $\delta > 1/3$. In the same way as for (A.5), we can show (A.6). As for (A.7), it holds that

$$\mathrm{var}\left[\sum_{h=1}^{p}\mathrm{E}\{(\rho_{3,h}^*)^2|\mathcal{F}_{h-1}^p\}\right]=512\,\frac{m^4n^4}{p^4(m+n)^3(m-1)^3(n-1)^3}\sum_{\ell=1}^{p-1}\sigma_{\ell\ell}^4\left(\sum_{h=\ell+1}^{p}\sigma_{hh}^2\right)^2$$
$$\leq 512\,\frac{m^4n^4}{p(m+n)^3(m-1)^3(n-1)^3}\left\{\frac{\mathrm{tr}(\boldsymbol{\Sigma}^4)}{p}\right\}\left\{\frac{\mathrm{tr}(\boldsymbol{\Sigma}^2)}{p}\right\}^2\to 0.$$

For (A.8), (A.9) and (A.10), see the Online Supplement.

Proof of (A.2). It follows from Jensen's inequality that

$$\mathrm{E}(\tau_h^4)\leq 5^3[\mathrm{E}\{(\rho_{1,h}^*)^4\}+\mathrm{E}\{(\rho_{2,h}^*)^4\}+\mathrm{E}\{(\rho_{3,h}^*)^4\}+\mathrm{E}\{(\rho_{4,h}^*)^4\}+\mathrm{E}\{(\rho_{5,h}^*)^4\}].$$

The right-hand side of the above expression is $O(p^{-2\delta})$, because some simple calculations give

$$\mathrm{E}\{(\rho_{1,h}^*)^4\}\leq\frac{m^4n^4}{p^4(m+n)^4}\frac{(m-1)^8}{(m+1)^4(m-2)^4}\Big[\sigma_{hh}^4[O(m^{-4})\mathrm{tr}(\boldsymbol{\Sigma}^4)+O(m^{-5})\mathrm{tr}(\boldsymbol{\Sigma}^3)\mathrm{tr}(\boldsymbol{\Sigma})$$
$$+O(m^{-5})\mathrm{tr}(\boldsymbol{\Sigma}^2)\{\mathrm{tr}(\boldsymbol{\Sigma})\}^2+O(m^{-4})\{\mathrm{tr}(\boldsymbol{\Sigma}^2)\}^2+O(m^{-6})\{\mathrm{tr}(\boldsymbol{\Sigma})\}^4]+\sigma_{hh}^8O(m^{-2})\Big]$$
$$=\frac{m^4n^4}{p^4(m+n)^4}O(p^{-6\delta+4})=O(p^{-2\delta}),$$
$$\mathrm{E}\{(\rho_{3,h}^*)^4\}=\frac{16m^4n^4}{p^4(m+n)^4}\Big[\sigma_{hh}^4[O(m^{-2}n^{-2})\mathrm{tr}(\boldsymbol{\Sigma}^4)+O(m^{-2}n^{-2})\{\mathrm{tr}(\boldsymbol{\Sigma}^2)\}^2]$$
$$+\sigma_{hh}^6O\{(m+n)m^{-2}n^{-2})\mathrm{tr}(\boldsymbol{\Sigma}^2)+\sigma_{hh}^8O(m^{-1}n^{-1})\Big]$$
$$=\frac{m^4n^4}{p^4(m+n)^4}O(p^{-4\delta+2})=O(p^{-2}),$$
$$\mathrm{E}\{(\rho_{4,h}^*)^4\}\leq\frac{m^4n^4}{p^4(m+n)^4}\frac{(m-1)^8}{(m+1)^4(m-2)^4}\Big[\sigma_{hh}^4[O(m^{-7})\mathrm{tr}(\boldsymbol{\Sigma}^4)+O(m^{-7})\mathrm{tr}(\boldsymbol{\Sigma}^3)\mathrm{tr}(\boldsymbol{\Sigma})$$
$$+O(m^{-6})\mathrm{tr}(\boldsymbol{\Sigma}^2)\{\mathrm{tr}(\boldsymbol{\Sigma})\}^2+O(m^{-6})\{\mathrm{tr}(\boldsymbol{\Sigma}^2)\}^2+O(m^{-6})\{\mathrm{tr}(\boldsymbol{\Sigma})\}^4]+\sigma_{hh}^8O(m^{-6})\Big]$$
$$=O(p^{-2\delta}),$$

Now $\mathrm{E}\{(\rho_{2,h}^*)^4\}=O(p^{-2\delta})$, and $\mathrm{E}\{(\rho_{5,h}^*)^4\}=O(p^{-2\delta})$; see also the evaluation on pp. 6541–6542 of [19]. Thus, we conclude from $\delta > 1/2$ and (7) for $i\in\{1,\ldots,8\}$ that $\mathrm{E}(\tau_1^4)+\cdots+\mathrm{E}(\tau_p^4)=O(p^{1-2\delta})\to 0$. This completes the argument. □

Proof of Theorem 2. Write $T=T_1+T_2+T_3+T_4+T_5$, where

$$T_1=\frac{mn}{m+n}\left[\frac{\hat{a}_{x,2}}{\{a_{x,1}(p)\}^2}+\frac{\hat{a}_{y,2}}{\{a_{y,1}(p)\}^2}-\frac{2\hat{a}_{xy}}{a_{x,1}(p)a_{y,1}(p)}\right],$$
$$T_2=\hat{a}_{x,2}\left\{\frac{1}{\hat{a}_{x,1}}+\frac{1}{a_{x,1}(p)}\right\}\frac{mn}{m+n}\left\{\frac{1}{\hat{a}_{x,1}}-\frac{1}{a_{x,1}(p)}\right\},\quad T_3=\hat{a}_{y,2}\left\{\frac{1}{\hat{a}_{y,1}}+\frac{1}{a_{y,1}(p)}\right\}\frac{mn}{m+n}\left\{\frac{1}{\hat{a}_{y,1}}-\frac{1}{a_{y,1}(p)}\right\},$$
$$T_4=\frac{\hat{a}_{xy}}{\hat{a}_{x,1}}\frac{2mn}{m+n}\left\{\frac{1}{\hat{a}_{y,1}}-\frac{1}{a_{y,1}(p)}\right\},\quad T_5=\frac{\hat{a}_{xy}}{a_{y,1}(p)}\frac{2mn}{m+n}\left\{\frac{1}{\hat{a}_{x,1}}-\frac{1}{a_{x,1}(p)}\right\}.$$

First, we see that $T=T_1+o_P(1)$. It follows from (10) that

$$\frac{mn}{m+n}\left\{\frac{1}{\hat{a}_{x,1}}-\frac{1}{a_{x,1}(p)}\right\}=\frac{mn}{p^{(1+\delta)/2}(m+n)}\frac{p^{(1+\delta)/2}\{a_{x,1}(p)-\hat{a}_{x,1}\}}{\hat{a}_{x,1}a_{x,1}(p)}\to^p 0$$

and

$$\frac{mn}{m+n}\left\{\frac{1}{\hat{a}_{y,1}}-\frac{1}{a_{y,1}(p)}\right\}\to^p 0,$$

which imply $T_2, T_3, T_4, T_5 \to^p 0$. It remains to show that

$$T_1 \rightsquigarrow \mathcal{N}(0, 4b_2^2). \tag{A.11}$$

Letting $\tilde{\boldsymbol{x}} = \boldsymbol{x}/\sqrt{a_{x,1}(p)}$ and $\tilde{\boldsymbol{y}} = \boldsymbol{y}/\sqrt{a_{y,1}(p)}$, one has $\mathrm{var}(\tilde{\boldsymbol{x}}) = \boldsymbol{\Sigma}_x/a_{x,1}(p)$ and $\mathrm{var}(\tilde{\boldsymbol{y}}) = \boldsymbol{\Sigma}_y/a_{y,1}(p)$. If we regard these $\tilde{\boldsymbol{x}}$ and $\tilde{\boldsymbol{y}}$ as the original $\boldsymbol{x}$ and $\boldsymbol{y}$, respectively, Lemma 1 implies (A.11), recalling that $\mathcal{H}_0$ is equivalent to (6). This completes the argument. $\square$

Appendix B. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jmva.2019.01.011.

References

[1] Z. Bai, H. Saranadasa, Effect of high dimension: By an example of a two sample problem, Statist. Sinica 6 (1996) 311–329.
[2] G. Boente, L. Orellana, Robust plug-in estimators in proportional scatter models, J. Statist. Plann. Inference 122 (2004) 95–110.
[3] S.X. Chen, L.-X. Zhang, P.-S. Zhong, Tests for high-dimensional covariance matrices, J. Amer. Statist. Assoc. 105 (2010) 810–819.
[4] P.S. Eriksen, Proportionality of covariance matrices, Ann. Statist. 15 (1987) 732–748.
[5] W.T. Federer, Testing proportionality of covariance matrices, Ann. Math. Stat. 22 (1951) 102–106.
[6] B.K. Flury, Proportionality of k covariance matrices, Statist. Probab. Lett. 4 (1986) 29–33.
[7] B.W. Flury, M.J. Schmid, Quadratic discriminant functions with constraints on the covariance matrices: Some asymptotic results, J. Multivariate Anal. 40 (1992) 244–261.
[8] B.W. Flury, M.J. Schmid, A. Narayanan, Error rates in quadratic discrimination with constraints on the covariance matrices, J. Classification 11 (1994) 101–120.
[9] S.T. Jensen, S. Johansen, Estimation of proportional covariances, Statist. Probab. Lett. 6 (1987) 83–85.
[10] S.T. Jensen, J. Madsen, Estimation of proportional covariances in the presence of certain linear restrictions, Ann. Statist. 32 (2004) 219–232.
[11] I.M. Johnstone, On the distribution of the largest eigenvalue in principal components analysis, Ann. Statist. 29 (2001) 295–327.
[12] S. Jung, J.S. Marron, PCA consistency in high dimension, low sample size context, Ann. Statist. 37 (2009) 4104–4130.
[13] C.G. Khatri, Some distribution problems connected with the characteristic roots of $S_1S_2^{-1}$, Ann. Math. Stat. 38 (1967) 944–948.
[14] J. Li, S.X. Chen, Two sample tests for high-dimensional covariance matrices, Ann. Statist. 40 (2012) 908–940.
[15] B. Liu, L. Xu, S. Zheng, G.-L. Tian, A new test for the proportionality of two large-dimensional covariance matrices, J. Multivariate Anal. 131 (2014) 293–308.
[16] B.F.J. Manly, J.C.W. Rayner, The comparison of sample covariance matrices using likelihood ratio tests, Biometrika 74 (1987) 841–847.
[17] K.C.S. Pillai, S. Al-Ani, G.M. Jouris, On the distributions of the ratios of the roots of a covariance matrix and Wilks' criterion for tests of three hypotheses, Ann. Math. Stat. 40 (1969) 2033–2040.
[18] J.R. Schott, A test for proportional covariance matrices, Comput. Statist. Data Anal. 32 (1999) 135–146.
[19] J.R. Schott, A test for the equality of covariance matrices when the dimension is large relative to the sample sizes, Comput. Statist. Data Anal. 51 (2007) 6535–6542.
[20] M.S. Srivastava, Some tests concerning the covariance matrix in high dimensional data, J. Japan Statist. Soc. 35 (2005) 251–272.
[21] M.S. Srivastava, H. Yanagihara, Testing the equality of several covariance matrices with fewer observations than the dimension, J. Multivariate Anal. 101 (2010) 1319–1329.
[22] M.S. Srivastava, H. Yanagihara, T. Kubokawa, Tests for covariance matrices in high dimension with less sample size, J. Multivariate Anal. 130 (2014) 289–309.
[23] A.W. van der Vaart, Asymptotic Statistics, Cambridge University Press, Cambridge, 1998.
[24] L. Xu, B. Liu, S. Zheng, S. Bao, Testing proportionality of two large-dimensional covariance matrices, Comput. Statist. Data Anal. 78 (2014) 43–55.
[25] K. Yata, M. Aoshima, Correlation tests for high-dimensional data using extended cross-data-matrix methodology, J. Multivariate Anal. 117 (2013) 313–331.
[26] S. Zheng, Central limit theorems for linear spectral statistics of large dimensional F-matrices, Ann. Inst. Henri Poincaré Probab. Stat. 48 (2012) 444–476.