Statistics and Probability Letters 161 (2020) 108699

Contents lists available at ScienceDirect

Statistics and Probability Letters

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

Hypothesis testing for the identity of high-dimensional covariance matrices

Manling Qian [a], Li Tao [a], Erqian Li [a], Maozai Tian [a,b,c,*]

[a] Center for Applied Statistics, School of Statistics, Renmin University of China, Beijing, 100872, China
[b] School of Statistics and Information, Xinjiang University of Finance and Economics, 830012, Xinjiang, China
[c] School of Statistics, Lanzhou University of Finance and Economics, 730101, Gansu, China

ARTICLE INFO

Article history:
Received 4 August 2019
Received in revised form 13 December 2019
Accepted 12 January 2020
Available online 10 February 2020

Keywords:
Covariance matrix
High-dimensional data
Hypothesis testing
Identity matrix

ABSTRACT

A new test statistic is proposed by utilizing the eigenvalues of the sample covariance matrix for the identity test. Under some general assumptions, asymptotic distributions of the proposed test statistic T and tests proposed in previous literature (denoted as T_s, T_1, T_2) are given. Simulations are also conducted to evaluate their performance in a finite sample.

1. Introduction

High-dimensional data are often encountered in modern statistical inferences. For example, in genomic studies, the data dimension p can be larger than the sample size n; In macro-economic data analysis, there are a large number of features of an economy. However, many classical methods fail when this happens because these conventional techniques rely on the assumption that the dimension p is fixed and the sample size n is larger than p.

In this paper, we focus on the identity test when p equals or exceeds the sample size n. Generally speaking, let $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$ be independent and identically distributed p-dimensional multivariate random variables with mean zero and a common population covariance matrix Σ_p. The hypothesis testing for an identity covariance matrix refers to:

$$H_0 : \Sigma_p = I_p \quad vs. \quad H_1 : \Sigma_p \neq I_p, \tag{1}$$

where I_p is the p-dimensional identity matrix. We are interested in studying this test based on the sample covariance matrix $S_n = \sum_{i=1}^{n} \boldsymbol{x}_i\boldsymbol{x}_i'/n$ under the framework that both p and n tend to infinity with $p/n \to c \in (0, \infty)$.

There are many research works on testing problems when $p > n$. Ledoit and Wolf (2002) modified certain statistics proposed by Hisao (1973) and John (1971) for the above hypothesis to address the situation when $p/n \to c$ for a finite constant c. Srivastava (2005) proposed a similar test but a different statistic using the first and second arithmetic means of eigenvalues of S_n. Fisher (2012) derived a statistic through utilizing the higher power of arithmetic means. However, these tests rely on an assumption that the sample is normally distributed. For non-normal situations, Chen et al. (2010) gave a new method by using the V-statistics, but it requires a heavy burden of computations. Srivastava et al. (2011) presented that the test in Srivastava (2005) is still valid when the kurtosis of the sample distribution is close to 3. For more relevant tests, one is referred to Li and Qin (2014) and Srivastava (2006).

* Corresponding author at: Center for Applied Statistics, School of Statistics, Renmin University of China, Beijing, 100872, China.
E-mail address: mztian@ruc.edu.cn (M. Tian).

https://doi.org/10.1016/j.spl.2020.108699
0167-7152/

Motivated by the precious research works, a new test statistic is proposed in this paper based on the third and sixth arithmetic means of the eigenvalues of the sample covariance matrix. It is shown that the newly proposed test is comparable to, in some cases more powerful than the tests in the current literatures.

The organization of this paper is as follows. In Section 2, we provide some preliminary results to construct the new test statistic for the identity test under moments conditions in the general asymptotic framework. Two new estimators are proposed for the fifth and sixth arithmetic mean of the eigenvalues of Σ_p. And we also investigate the asymptotic behaviors of the two new estimators. Section 3 provides the new test procedure. In Section 4, the extensive simulation is presented to show the effectiveness of the new statistic. Concluding remarks are given in Section 5. Technical proofs are deferred to the last section.

2. Preliminaries

Denote H_p and F_n as the spectral distribution of Σ_p and S_n respectively, then the kth moment of them can be expressed as:

$$\alpha_k = \int t^k dH_p(t) = \frac{1}{p} tr(\Sigma_p^k), \quad \hat{\beta}_k = \int x^k dF_n(x) = \frac{1}{p} tr(S_n^k), \quad k \in \mathbb{N}.$$

Meanwhile, the estimators of α_i, $i = 1, 2, 3, 4$ can be given as follows:

$$\hat{\alpha}_1 = \hat{\beta}_1, \qquad \hat{\alpha}_2 = \gamma_2(\hat{\beta}_2 - c_n\hat{\beta}_1^2), \qquad \hat{\alpha}_3 = \gamma_3(\hat{\beta}_3 - 3c_n\hat{\beta}_2\hat{\beta}_1 + 2c_n^2\hat{\beta}_1^3),$$
$$\hat{\alpha}_4 = \gamma_4(\hat{\beta}_4 - 4c_n\hat{\beta}_3\hat{\beta}_1 - \frac{2n^2+3n-6}{n^2+n+2}c_n\hat{\beta}_2^2 + \frac{10n^2+12n}{n^2+n+2}c_n^2\hat{\beta}_2\hat{\beta}_1{}^2 - \frac{5n^2+6n}{n^2+n+2}c_n^3\hat{\beta}_1^4),$$

where $c_n = p/n$, $\gamma_2 = n^2/[(n-1)(n+2)]$, $\gamma_3 = n^4/[(n-1)(n-2)(n+2)(n+4)]$, $\gamma_4 = n^5(n^2+n+2)/[(n+1)(n+2)(n+4)(n+6)(n-1)(n-2)(n-3)]$. When the sample data are assumed normally distributed, the estimators of α_i, $i = 1, 2, 3, 4$ were proved to be unbiased, consistent and asymptotically normal, see Srivastava (2005), Fisher (2012) and Fisher et al. (2010). When the normally distributed assumption does not hold, it was shown that these estimators were not unbiased, but the consistency and asymptotic normality can still be derived.

In this paper, we give two new estimators of α_5 and α_6, which are given as follows:

$$\hat{\alpha}_5 = \hat{\beta}_5 - 5c_n\hat{\beta}_1\hat{\beta}_4 - 5c_n\hat{\beta}_2\hat{\beta}_3 + 15c_n^2\hat{\beta}_1^2\hat{\beta}_3 + 15c_n^2\hat{\beta}_1\hat{\beta}_2^2 - 35c_n^3\hat{\beta}_1^3\hat{\beta}_2 + 14c_n^4\hat{\beta}_1^5,$$
$$\hat{\alpha}_6 = \hat{\beta}_6 - 6c_n\hat{\beta}_1\hat{\beta}_5 - 6c_n\hat{\beta}_2\hat{\beta}_4 - 3c_n\hat{\beta}_3^2 + 21c_n^2\hat{\beta}_1^2\hat{\beta}_4 + 42c_n^2\hat{\beta}_1\hat{\beta}_2\hat{\beta}_3 + 7c_n^2\hat{\beta}_2^3 - 56c_n^3\hat{\beta}_1^3\hat{\beta}_3$$
$$- 84c_n^3\hat{\beta}_1^2\hat{\beta}_2^2 + 126c_n^4\hat{\beta}_1^4\hat{\beta}_2 - 42c_n^5\hat{\beta}_1^6.$$

Under some proper assumptions, firstly we can prove the consistency and asymptotic normality of the two new estimators without the normally distributed assumption. And the joint distribution of the estimators $\hat{\alpha}_i$, $i = 1, 2, 3, 4, 5, 6$ can be derived through an elementary linear combination. Certain assumptions are as follows:

Assumption(a). Both $n, p \to \infty$, and $c_n = p/n \to c \in (0, \infty)$.

Assumption(b). There is a doubly infinite array of i.i.d. random variables (ω_{ij}), $i, j \geq 1$ with:

$$E(\omega_{11}) = 0, \quad E(\omega_{11}^2) = 1, \quad E(\omega_{11}^4) < \infty.$$

And for each p, n, let $W_n = (\omega_{i,j})_{1\leq i\leq p, 1\leq j\leq n}$, the observation vectors can be represented as $\boldsymbol{x}_j = \Sigma_p^{-1/2}\omega_{\cdot j}$, where $\boldsymbol{\omega}_{\cdot j} = (\omega_{i,j})_{1\leq i\leq p}$ denotes the jth column of W_n and $\Sigma_p^{-1/2}$ is the Hermitian square root of Σ_p.

Assumption(c). The population spectral distribution H_p of Σ_p weakly converges to a probability distribution H, as $p \to \infty$, and the sequence of spectral norms $\{\|\Sigma_p\|\}$ is uniformly bounded.

Assumptions (a)–(c) are general conditions in the central limit theorem for linear spectral statistics of sample covariance matrices, see Silverstein and Bai (2004), Bai and Silverstein (2010), Pan and Zhou (2008) and Bai et al. (2010). From the assumption (c), we can get: $\alpha_k \to \tilde{\alpha}_k = \int t^k dH(t)$ for any fixed $k \in \mathbb{N}$, as $p \to \infty$.

Lemma 1. *Suppose that the assumptions (a)–(c) are satisfied.*
(i) the estimator $\hat{\alpha}_k$ is strongly consistent, i.e.,

$$\hat{\alpha}_k - \alpha_k \xrightarrow{a.s.} 0, k = 1, 2, \ldots, 6.$$

(ii) If $E(w_{11}^4) = 3$, then

$$n(\hat{\alpha}_1 - \alpha_1, \hat{\alpha}_2 - \alpha_2, \ldots, \hat{\alpha}_6 - \alpha_6)' \xrightarrow{D} \boldsymbol{N}_6(0, \Omega), \tag{2}$$

where the covariance matrix Ω is given in the Appendix.

Lemma 1 presents the consistency and asymptotic normality of the estimators $\hat{\alpha}_k$. We can find that the convergence of the first four estimators is the same as Theorem 2 in Fisher (2012) which assumes that ω_{11} is standard normal, but our approach only needs that ω_{11} has the same first, second, and fourth moments as standard normal variables. If $E(w_{11}^4) = 3$ does not hold, we have the following result:

Lemma 2. *Suppose that the assumptions* (a)–(c) *are satisfied and* Σ_p *is diagonal for all* p *in addition. Then*

$$n(\hat{\alpha}_1 - \alpha_1, \hat{\alpha}_2 - \alpha_2, \ldots, \hat{\alpha}_6 - \alpha_6)' \xrightarrow{D} \boldsymbol{N}_6(\nu, \Gamma), \tag{3}$$

where the mean vector ν *and the covariance matrix* Γ *are respectively:*

$$\nu = \Delta \cdot (0, \tilde{\alpha}_2, 3\tilde{\alpha}_3, 6\tilde{\alpha}_4, 10\tilde{\alpha}_5, 15\tilde{\alpha}_6)',$$

$$\Gamma = \Omega + \frac{\Delta}{c}\begin{pmatrix} \tilde{\alpha}_2 & 2\tilde{\alpha}_3 & 3\tilde{\alpha}_4 & 4\tilde{\alpha}_5 & 5\tilde{\alpha}_6 & 6\tilde{\alpha}_7 \\ 2\tilde{\alpha}_3 & 4\tilde{\alpha}_4 & 6\tilde{\alpha}_5 & 8\tilde{\alpha}_6 & 10\tilde{\alpha}_7 & 12\tilde{\alpha}_8 \\ 3\tilde{\alpha}_4 & 6\tilde{\alpha}_5 & 9\tilde{\alpha}_6 & 12\tilde{\alpha}_7 & 15\tilde{\alpha}_8 & 18\tilde{\alpha}_9 \\ 4\tilde{\alpha}_5 & 8\tilde{\alpha}_6 & 12\tilde{\alpha}_7 & 16\tilde{\alpha}_8 & 20\tilde{\alpha}_9 & 24\tilde{\alpha}_{10} \\ 5\tilde{\alpha}_6 & 10\tilde{\alpha}_7 & 15\tilde{\alpha}_8 & 20\tilde{\alpha}_9 & 25\tilde{\alpha}_{10} & 30\tilde{\alpha}_{11} \\ 6\tilde{\alpha}_7 & 12\tilde{\alpha}_8 & 18\tilde{\alpha}_9 & 24\tilde{\alpha}_{10} & 30\tilde{\alpha}_{11} & 36\tilde{\alpha}_{12} \end{pmatrix},$$

where $\Delta = E(\omega_{11}^4) - 3$ *and the matrix* Ω *is defined in* (2).

Lemma 2 gives a new central limit theorem under the general fourth moment of ω_{11} and the diagonality of Σ_p. One can see that both the mean vectors and the covariance matrix have some shifts from those in (2), which indicates that the estimators $\hat{\alpha}_k, k = 2, 3, \ldots, 6$ are not unbiased. When neither $E(w_{11}^4) = 3$ nor Σ_p is diagonal, the situation will become much more complicated. Thus, we will not go into details here, if interested, one can refer to Pan and Zhou (2008).

3. Test procedure

This paper aims to analyze the existing methods in the current literature and propose a new test statistic for testing hypothesis (1). As mentioned above, Srivastava (2005) proposed a test statistic using the first and second means of the eigenvalues of S_n, denoted as:

$$T_s = \frac{n}{2}(\hat{\alpha}_2 - 2\hat{\alpha}_1 + 1). \tag{4}$$

It is shown that the statistic has the following asymptotic distribution:

$$\frac{n}{2}(\hat{\alpha}_2 - 2\hat{\alpha}_1 - \alpha_2 - 2\alpha_1) \xrightarrow{D} \boldsymbol{N}(0, \frac{c}{2}(\tilde{\alpha}_2 - 2\tilde{\alpha}_3 + \tilde{\alpha}_4 + \tilde{\alpha}_2^2))$$

under the alternative hypothesis H_1, and

$$T_s = \frac{n}{2}(\hat{\alpha}_2 - 2\hat{\alpha}_1 + 1) \xrightarrow{D} \boldsymbol{N}(0, 1)$$

under the null hypothesis.

Later, Fisher (2012) introduced two new statistics using higher powers of the arithmetic means, which are denoted as T_1 and T_2 respectively,

$$T_1 = \frac{n}{c\sqrt{8}}(\hat{\alpha}_4 - 4\hat{\alpha}_3 + 6\hat{\alpha}_1 + 1), \tag{5}$$

$$T_2 = \frac{n}{\sqrt{8(c^2 + 12c + 8)}}(\hat{\alpha}_4 - 2\hat{\alpha}_2 + 1). \tag{6}$$

Under the null hypothesis, the asymptotic distributions are given as follows:

$$T_1 = \frac{n}{c\sqrt{8}}(\hat{\alpha}_4 - 4\hat{\alpha}_3 + 6\hat{\alpha}_1 + 1) \xrightarrow{D} \boldsymbol{N}(0, 1),$$

$$T_2 = \frac{n}{\sqrt{8(c^2 + 12c + 8)}}(\hat{\alpha}_4 - 2\hat{\alpha}_2 + 1) \xrightarrow{D} \boldsymbol{N}(0, 1).$$

The asymptotic distribution under H_1 can be seen in Fisher (2012).

From the simulation, we can see that in some cases, the two statistics proposed in that paper are more powerful than other statistics. Note that under H_0, each eigenvalue of Σ_p takes on the value one. Furthermore, the equality for inequality

$$1/p\sum_{i=1}^{p}(\lambda_i^r - 1)^{2s} \geq 0 \tag{7}$$

holds if and only if each $\lambda_i = 1$. Recalling the definition of α_i and a simple algebraic expansion results in the statistics T_s, T_1, T_2 in the cases of $(r = 1, s = 1)$, $(r = 1, s = 2)$, and $(r = 2, s = 1)$ respectively, which motivates us to construct the new statistics based on the inequality (7). The case of $(r = 3, s = 1)$ is considered here. In this case, the inequality becomes $\alpha_6 - 2\alpha_3 + 1 \geq 0$. And the equality holds if and only if the hypothesis of identity stands. Performing a standardization process, we propose a new statistic:

$$T = \frac{n}{\sqrt{12(c^4 + 30c^3 + 150c^2 + 162c + 27)}}(\hat{\alpha}_6 - 2\hat{\alpha}_3 + 1).$$

Next, the asymptotic distribution of T is given under both alternative and null hypothesis:

Theorem 1. *Suppose that the assumptions (a)–(c) are satisfied.*
(i) If Σ_p is diagonal, then

$$\frac{n}{\sqrt{12(c^4 + 30c^3 + 150c^2 + 162c + 27)}}(\hat{\alpha}_6 - 2\hat{\alpha}_3 - (\alpha_6 - 2\alpha_3)) \xrightarrow{D} \boldsymbol{N}(\mu, \sigma_T^2) \tag{8}$$

as $(n, p) \to \infty$, where the mean is $\mu = \Delta \cdot (-6\tilde{\alpha}_3 + 15\tilde{\alpha}_6)/(\sqrt{12(c^4 + 30c^3 + 150c^2 + 162c + 27)})$ and the variance is

$$\begin{aligned}\sigma_T^2 =& \frac{1}{c^4 + 30c^3 + 150c^2 + 162c + 27}(6\tilde{\alpha}_{10}\tilde{\alpha}_2 + 12\tilde{\alpha}_3\tilde{\alpha}_9 + 18\tilde{\alpha}_4\tilde{\alpha}_8 + 24\tilde{\alpha}_5\tilde{\alpha}_7 + 15\tilde{\alpha}_6^2 \\ &- 12\tilde{\alpha}_2\tilde{\alpha}_7 - 24\tilde{\alpha}_3\tilde{\alpha}_6 - 24\tilde{\alpha}_4\tilde{\alpha}_5 + 6\tilde{\alpha}_3^2 + 6\tilde{\alpha}_2\tilde{\alpha}_4 + c^4\tilde{\alpha}_2^6 + c^3(6\tilde{\alpha}_2^4\tilde{\alpha}_4 + 24\tilde{\alpha}_2^3\tilde{\alpha}_3^2) \\ &+ c^2(6\tilde{\alpha}_2^3\tilde{\alpha}_6 + 36\tilde{\alpha}_2^2\tilde{\alpha}_3\tilde{\alpha}_5 + 27\tilde{\alpha}_2^2\tilde{\alpha}_4^2 + 72\tilde{\alpha}_2\tilde{\alpha}_3^2\tilde{\alpha}_4 + 9\tilde{\alpha}_3^2) + c(2\tilde{\alpha}_2^3 - 12\tilde{\alpha}_2^2\tilde{\alpha}_5 \\ &+ 6\tilde{\alpha}_2^2\tilde{\alpha}_8 - 24\tilde{\alpha}_2\tilde{\alpha}_3\tilde{\alpha}_4 + 24\tilde{\alpha}_2\tilde{\alpha}_3\tilde{\alpha}_7 + 36\tilde{\alpha}_2\tilde{\alpha}_4\tilde{\alpha}_6 + 24\tilde{\alpha}_2\tilde{\alpha}_5^2 - 4\tilde{\alpha}_3^3 + 24\tilde{\alpha}_3^2\tilde{\alpha}_6 \\ &+ 72\tilde{\alpha}_3\tilde{\alpha}_4\tilde{\alpha}_5 + 14\tilde{\alpha}_4^3) + \frac{1}{c}(6\tilde{\alpha}_{12} + 6\tilde{\alpha}_6 - 12\tilde{\alpha}_9) + \frac{\Delta}{c}(3\tilde{\alpha}_{12} + 3\tilde{\alpha}_6 - 6\tilde{\alpha}_9)).\end{aligned}$$

(ii) If $E(w_{11}^4) = 3$ then the convergence in (8) also holds with the same mean (actually zero) and variance.
(iii) Under the null hypothesis,

$$T = \frac{n}{\sqrt{12(c^4 + 30c^3 + 150c^2 + 162c + 27)}}(\hat{\alpha}_6 - 2\hat{\alpha}_3 + 1) \xrightarrow{D} \boldsymbol{N}(\mu_0, \sigma_{T0}^2),$$

as $(n, p) \to \infty$, where the mean is $\mu_0 = 9\Delta/\sqrt{12(c^4 + 30c^3 + 150c^2 + 162c + 27)}$ and the variance is $\sigma_{T0}^2 = 1$.

Considering the fact that the asymptotic distributions of T_s, T_1 and T_2 are derived under the normally distributed assumption of the sample, we still need to find the general distribution of all of them under a more relaxed condition, that is, the moments condition.

Theorem 2. *Suppose that the assumptions (a)–(c) are satisfied.*
(i) If Σ_p is diagonal, then

$$\begin{aligned}&\frac{n}{2}(\hat{\alpha}_2 - 2\hat{\alpha}_1 - (\alpha_2 - 2\alpha_1)) \xrightarrow{D} \boldsymbol{N}(\mu_s, \sigma_s^2) \\ &\frac{n}{c\sqrt{8}}(\hat{\alpha}_4 - 4\hat{\alpha}_3 + 6\hat{\alpha}_2 - 4\hat{\alpha}_1 - (\alpha_4 - 4\alpha_3 + 6\alpha_2 - 4\alpha_1)) \xrightarrow{D} \boldsymbol{N}(\mu_1, \sigma_1^2) \\ &\frac{n}{\sqrt{8(c^2 + 12c + 8)}}(\hat{\alpha}_4 - 2\hat{\alpha}_2 - (\alpha_4 - 2\alpha_2)) \xrightarrow{D} \boldsymbol{N}(\mu_2, \sigma_2^2)\end{aligned} \tag{9}$$

as $(n, p) \to \infty$, where the means are $\mu_s = \Delta \cdot \tilde{\alpha}_2/2$, $\mu_1 = \Delta \cdot (6\tilde{\alpha}_2 - 12\tilde{\alpha}_3 + 6\tilde{\alpha}_4)/(c\sqrt{8})$, $\mu_2 = \Delta \cdot (6\tilde{\alpha}_4 - 2\tilde{\alpha}_2)/\sqrt{8(c^2 + 12c + 8)}$ and the variances are

$$\begin{aligned}\sigma_s^2 =& \tilde{\alpha}_2^2 + \frac{2}{c}(\tilde{\alpha}_2 - 2\tilde{\alpha}_3 + \tilde{\alpha}_4) + \frac{\Delta}{c}(\tilde{\alpha}_2 - 2\tilde{\alpha}_3 + \tilde{\alpha}_4), \\ \sigma_1^2 =& \frac{1}{c^2}(c^2\tilde{\alpha}_2^4 + 60\tilde{\alpha}_4\tilde{\alpha}_2 + 18\tilde{\alpha}_2^2 + 48\tilde{\alpha}_3^2 + 6\tilde{\alpha}_4^2 + 4\tilde{\alpha}_2\tilde{\alpha}_6 + 8\tilde{\alpha}_3\tilde{\alpha}_5 - 72\tilde{\alpha}_2\tilde{\alpha}_3 - 24\tilde{\alpha}_2\tilde{\alpha}_5 \\ &- 48\tilde{\alpha}_3\tilde{\alpha}_4 + 12c\tilde{\alpha}_2^3 + 4c\tilde{\alpha}_4\tilde{\alpha}_2^2 + 8c\tilde{\alpha}_2\tilde{\alpha}_3^2 - 24c\tilde{\alpha}_2^2\tilde{\alpha}_3 + \frac{4}{c}(\tilde{\alpha}_8 + \tilde{\alpha}_2 + 15\tilde{\alpha}_4 + 15\tilde{\alpha}_6 - 6\tilde{\alpha}_7 \\ &- 20\tilde{\alpha}_5 - 6\tilde{\alpha}_3) + \frac{2\Delta}{c}(\tilde{\alpha}_8 + \tilde{\alpha}_2 + 15\tilde{\alpha}_4 + 15\tilde{\alpha}_6 - 6\tilde{\alpha}_7 - 20\tilde{\alpha}_5 - 6\tilde{\alpha}_3)), \\ \sigma_2^2 =& \frac{1}{c^2 + 12c + 8}(c^2\tilde{\alpha}_2^4 + 8c\tilde{\alpha}_2\tilde{\alpha}_3^2 + 4c\tilde{\alpha}_2^2\tilde{\alpha}_4 + 2\tilde{\alpha}_2^2 + 6\tilde{\alpha}_4^2 + 4\tilde{\alpha}_2\tilde{\alpha}_6 + 8\tilde{\alpha}_2\tilde{\alpha}_5 - 8\tilde{\alpha}_2\tilde{\alpha}_4 - 4\tilde{\alpha}_3^2 \\ &+ \frac{4}{c}(\tilde{\alpha}_4 + \tilde{\alpha}_8 - 2\tilde{\alpha}_6) + \frac{2\Delta}{c}(\tilde{\alpha}_4 + \tilde{\alpha}_8 - 2\tilde{\alpha}_6)).\end{aligned}$$

Table 1
Empirical sizes in percentages of four identity tests for standard normal random variables.

n	T	T_s	T_1	T_2	T	T_s	T_1	T_2
	$c = 1$				$c = 5$			
50	5.32	4.98	4.86	6.12	4.57	4.68	3.94	5.08
100	5.13	4.95	4.82	5.63	4.95	4.81	4.86	5.19
150	5.21	4.84	5.32	5.45	4.92	4.79	5.02	4.94
200	5.12	4.96	5.08	5.33	5.03	4.94	4.88	5.26
250	5.02	5.09	5.01	4.94	4.83	4.92	4.60	4.97
300	5.25	5.27	5.28	5.55	4.95	4.64	4.83	4.83
	$c = 10$				$c = 20$			
50	4.48	4.29	4.28	4.87	5.12	5.05	3.75	4.20
100	4.86	4.83	4.72	5.04	4.87	4.82	4.80	4.97
150	4.63	4.52	4.68	4.61	5.03	4.92	5.03	5.05
200	5.12	5.43	4.85	5.06	4.57	4.68	4.54	4.48
250	4.97	4.95	4.89	5.07	5.06	5.28	5.04	4.85
300	4.86	4.70	4.65	4.84	4.93	4.85	4.94	5.00

Table 2
Empirical sizes in percentages of four identity tests for uniform random variables.

n	T	T_s	T_1	T_2	T	T_s	T_1	T_2
	$c = 1$				$c = 5$			
50	6.13	4.87	3.76	3.32	6.12	5.19	3.13	4.26
100	5.93	6.19	3.93	4.26	5.76	3.65	4.32	4.57
150	4.86	5.21	4.12	4.46	5.93	4.37	4.67	4.27
200	5.54	4.93	4.68	5.13	4.87	5.03	4.72	5.15
250	4.93	5.13	5.13	5.36	4.96	4.24	4.83	4.96
300	5.03	5.05	4.96	5.07	5.04	5.91	5.03	5.06
	$c = 10$				$c = 20$			
50	5.67	4.18	3.63	3.13	5.86	3.42	3.21	3.15
100	5.43	3.72	4.53	5.13	5.67	5.23	4.53	4.34
150	4.56	4.04	5.19	4.72	4.63	5.21	4.67	4.63
200	5.36	4.73	4.57	4.96	5.27	6.21	6.21	5.26
250	5.25	5.03	4.86	5.29	4.96	5.02	4.56	4.96
300	4.98	5.31	4.97	4.98	5.06	5.32	5.14	5.12

(ii) If $E(w_{11}^4) = 3$ then the convergence in (9) also holds with the same means (actually zero) and variances.

(iii) Under the null hypothesis, $T_s = \frac{n}{2}(\hat{\alpha}_2 - 2\hat{\alpha}_1 + 1) \xrightarrow{D} \mathbf{N}(\mu_{s0}, \sigma_{s0}^2)$, $T_1 = \frac{n}{c\sqrt{8}}(\hat{\alpha}_4 - 4\hat{\alpha}_3 + 6\hat{\alpha}_1 + 1) \xrightarrow{D} \mathbf{N}(\mu_{10}, \sigma_{10}^2)$, $T_2 = \frac{n}{\sqrt{8(c^2+12c+8)}}(\hat{\alpha}_4 - 2\hat{\alpha}_2 + 1) \xrightarrow{D} \mathbf{N}(\mu_{20}, \sigma_{20}^2)$, as $(n, p) \to \infty$, where the means are $\mu_{s0} = \Delta/2, \mu_{10} = 0, \mu_{20} = 4\Delta/\sqrt{8(c^2 + 12c + 8)}$, and the variances are $\sigma_{s0}^2 = 1, \sigma_{10}^2 = 1, \sigma_{20}^2 = 1$.

4. Simulation study

The purpose of the simulation study is to (i) show the effectiveness of the newly defined test statistics T and (ii) perform a comparative study on the available statistics (denoted by T_s, T_1 and T_2) in testing the hypothesis that the covariance matrix is the identity. And we will show the empirical sizes and powers of the studied tests in four scenarios of distribution:

$$\text{(I) } \omega_{11} \sim N(0, 1), \qquad \text{(II) } \omega_{11} \sim unif(-\sqrt{(3)}, \sqrt{(3)}),$$
$$\text{(III) } \omega_{11} \sim \sqrt{3/4}t_8, \quad \text{(IV) } \omega_{11} \sim \frac{\chi^2(7) - 7}{\sqrt{14}}.$$

In the first scenario, ω_{11} is a standard normal random variable with $E(\omega_{11}^4) = 3$; the second distribution is a light tailed distribution, and ω_{11} is a uniform random variable with $E(\omega_{11}^4) = 9/5$; in the third scenario, a heavy tailed distribution is explored, and ω_{11} is a scaled Student's t-random variable with $E(\omega_{11}^4) = 9/2$; lastly, we look at a skewed distribution, and ω_{11} is a chi-squared random variable with $E(\omega_{11}^4) = 33/7$. The nominal significance level is $\alpha = 0.05$, and the number of independent replications is set to be 10 000.

Tables 1–4 show the empirical sizes of the four tests. As can be seen, in the first scenario of normal distribution, all the sizes are close to the nominal significance level. As for the non-normal distributions, there is slight size distortion for the tests when p and n are small, but this distortion fades away as p and n are increased.

Table 3
Empirical sizes in percentages of four identity tests for standardized Student's t-random variables.

n	T	T_s	T_1	T_2	T	T_s	T_1	T_2
	$c=1$				$c=5$			
50	6.23	7.27	7.63	6.03	6.34	7.03	6.21	6.32
100	5.84	5.67	6.84	5.26	6.17	5.83	4.26	5.48
150	5.94	4.69	4.79	5.46	5.75	6.54	5.47	4.86
200	5.83	4.28	5.86	5.13	5.43	5.36	5.35	5.42
250	5.63	5.08	5	5.38	5.45	5.32	4.86	5.26
300	5.13	5.23	5.13	5	5.21	4.96	5.02	4.89
	$c=10$				$c=20$			
50	6.16	6.43	4.23	5.23	6.24	6.57	4.71	6.24
100	5.66	5.92	5.63	5.06	5.98	6.16	4.59	5.53
150	5.73	6.36	4.58	4.67	5.54	4.67	5.03	4.65
200	5.53	5.43	4.96	5.63	5.62	5.37	5.21	5.04
250	5.34	5.38	5.27	4.96	5.26	5.32	4.86	4.87
300	5.12	5.26	4.98	4.94	5.2	5.08	5.06	5.06

Table 4
Empirical sizes in percentages of four identity tests for standardized chi-squared random variables.

n	T	T_s	T_1	T_2	T	T_s	T_1	T_2
	$c=1$				$c=5$			
50	6.53	6.66	7.26	8.26	5.75	5.68	5.86	6.52
100	6.21	5.9	6.51	7.07	5.28	5.26	5.39	6
150	5.46	5.47	5.9	6.56	5.25	5.31	5.54	5.28
200	5.32	5.44	5.14	6.02	5.16	5.03	4.64	5.3
250	5.36	5.48	5.25	5.44	5.05	5.06	4.77	5.11
300	5.25	5.17	5.62	5.62	4.98	4.71	5.26	5.26
	$c=10$				$c=20$			
50	5.63	5.87	4.9	5.43	5.56	5.71	4.88	5.22
100	5.36	4.83	4.93	5.37	4.76	5.45	4.5	4.91
150	4.86	4.85	5.2	5.04	5.09	5.13	5.09	5.02
200	5.27	5.08	5.34	5.2	5.03	5.16	4.87	4.98
250	4.93	5.22	5.04	5.14	5.02	4.8	4.85	4.96
300	5.03	4.84	5.25	5.25	4.98	5.23	5.06	5.06

Table 5
Empirical powers of four identity tests under Model 2 for standard normal random variables.

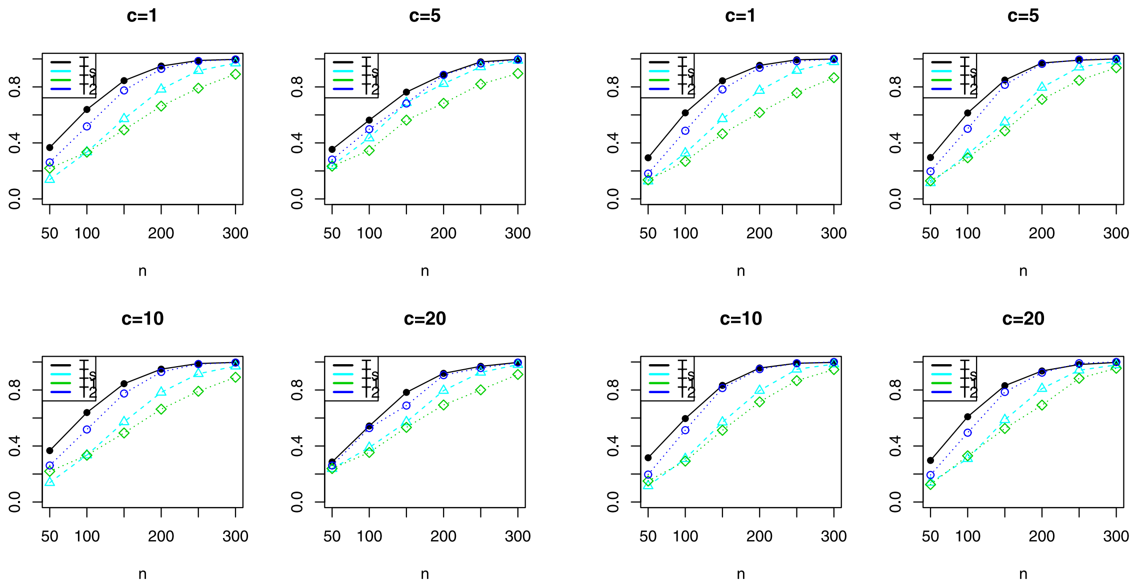
n	T	T_s	T_1	T_2	T	T_s	T_1	T_2
	$c=1$				$c=3$			
50	0.747	0.997	0.522	0.931	0.569	0.999	0.434	0.81
100	0.93	1	0.666	1	0.744	1	0.537	0.99
150	0.99	1	0.846	1	0.818	1	0.62	1
200	1	1	0.918	1	0.892	1	0.686	1
250	1	1	0.972	1	0.936	1	0.752	1
300	1	1	0.99	1	0.965	1	0.815	1

To compare the powers of the four statistics, we study two models under the alternative hypothesis, which are listed below.

Model 1: $\Sigma_p = I_p + \sqrt{(c)}diag(\underbrace{1.2, \ldots, 1.2}_{n/50}, \underbrace{0, \ldots, 0}_{p-n/50})$,

Model 2: $\Sigma_p = I_p + diag(\underbrace{0.6, \ldots, 0.6}_{p/2}, \underbrace{0, \ldots, 0}_{p/2})$($p$ even).

The difference between the two models is that there is only a small cluster of diagonal elements which deviate from the bulk in Model 1, but in Model 2 half of the diagonal elements are apart from the rest. The empirical powers of four tests in normal and uniform distributions are shown in Fig. 1. The empirical powers of four tests in Student's t and chi-squared distributions are not presented here due to limited space and are available upon request. Obviously, we can see that all the powers grow to 1 as the dimensions increase, and the shifts in data distribution do not seem to affect the powers. Comparing the performance, the newly proposed test is more powerful. Empirical powers are illustrated in Tables 5–8 which show that the test T_s is more powerful and the power function of T_1 has a slower convergence rate compared with the other three tests. Moreover, there is a slight decrease in power of tests T, T_1 and T_2 between the two concentrations, but the test T_s does not seem to be affected by concentration.



(a) standard normal distribution (b) uniform distribution

Fig. 1. Empirical powers of four tests under normal and uniform distributions in Model 1.

Table 6
Empirical powers of four identity tests under Model 2 for uniform random variables.

n	T	T_s	T_1	T_2	T	T_s	T_1	T_2
	$c = 1$				$c = 3$			
50	0.842	0.994	0.719	0.869	0.826	0.998	0.684	0.779
100	0.939	1	0.794	0.999	0.881	1	0.714	0.981
150	0.991	1	0.866	1	0.907	1	0.764	1
200	0.997	1	0.918	1	0.915	1	0.787	1
250	1	1	0.967	1	0.956	1	0.809	1
300	1	1	0.982	1	0.979	1	0.845	1

Table 7
Empirical powers of four identity tests under Model 2 for standardized Student's t-random variables.

n	T	T_s	T_1	T_2	T	T_s	T_1	T_2
	$c = 1$				$c = 3$			
50	0.903	0.994	0.817	0.953	0.878	0.999	0.766	0.865
100	0.988	1	0.889	1	0.902	1	0.78	0.99
150	0.997	1	0.936	1	0.938	1	0.794	1
200	0.999	1	0.97	1	0.958	1	0.834	1
250	1	1	0.988	1	0.973	1	0.863	1
300	1	1	0.993	1	0.989	1	0.906	1

Table 8
Empirical powers of four identity tests under Model 2 for standardized chi-squared random variables.

n	T	T_s	T_1	T_2	T	T_s	T_1	T_2
	$c = 1$				$c = 3$			
50	0.884	0.994	0.834	0.959	0.884	0.998	0.774	0.885
100	0.976	1	0.885	1	0.911	1	0.795	0.993
150	0.998	1	0.947	1	0.924	1	0.813	1
200	1	1	0.972	1	0.948	1	0.822	1
250	1	1	0.987	1	0.974	1	0.844	1
300	1	1	0.996	1	0.979	1	0.891	1

To conclude, the shifts in data distribution do not seem to affect the powers. And T performs the best when the true covariance matrix is nearly an identity matrix. The statistic T_s performs better when the true covariance greatly differs from the identity. Besides, the power function of T_1 has a slower convergence rate compared with the other three tests.

5. Conclusions and remarks

This paper proposes a new test statistic of testing hypothesis (1). Tests proposed in Srivastava (2005) and Fisher (2012) have been generalized to accommodate the situations where the underlying distribution is not normal. Simulations show that the new test performs best when just a few elements deviate from the identity matrix. Besides, the statistic T_s performs better when the true covariance matrix greatly differs from the identity. The asymptotic results derived in this paper is based on the assumption that $\boldsymbol{x}_i$ has zero mean vector. When this is not true, we need to replace S_n with S_n^*,with $S_n^* = \frac{1}{n-1}\sum_{i=1}^n(\boldsymbol{x}_i - \bar{\boldsymbol{x}})(\boldsymbol{x}_i - \bar{\boldsymbol{x}})'$, where $\bar{\boldsymbol{x}} = \sum_{i=1}^n \boldsymbol{x}_i/n$. Then if we use $c_n^* = p/(n-1)$ to replace $c_n = p/n$, all the conclusions in this paper hold, see Wang and Yao (2013) and Zheng and Bai (2013).

CRediT authorship contribution statement

Manling Qian: Writing - original draft, Conceptualization, Methodology. **Li Tao:** Methodology, Writing, review & editing, Software. **Erqian Li:** Methodology, Writing - review & editing. **Maozai Tian:** Supervision, Writing - review & editing.

Acknowledgments

The work was supported by the Fundamental Research Funds for the Central Universities, China, and the Research Funds of Renmin University of China (No. 18XNL012).

Appendix

A.1. Proof of Lemma 1

Supposing that Assumptions (a)–(c) hold, we can know that the spectral distribution F_n converges weakly to a limiting distribution $F^{c,H}$, moreover the Stieltjes transform $s_n(z)$ of F_n converges almost surely to $s(z)$, the Stieltjes transform of $F_{c,H}$, see Silverstein (1995). Let F_{c_n,H_p} be a distribution derived from $F^{c,H}$ by replacing c and H with c_n and H_p, respectively. Then the kth moments of F_{c_n,H_p} and $F_{c,H}$, can be given as follows:

$$\beta_k = \int t^k dF^{c_n,H_p}(t) \quad and \quad \tilde{\beta}_k = \int t^k dF^{c,H}(t)$$

$k = 1, 2, \ldots$. As $(n, p) \to \infty$, it is easy to derive that $\beta_k \to \tilde{\beta}$. Meanwhile, from Nica and Speicher (2006), the relationship between α_k and β_k, $k = 1, 2, \ldots, 6$ is:

$$\begin{aligned}
&\alpha_1 = \beta_1, \quad \alpha_2 = \beta_2 - c_n\beta_1^2, \quad \alpha_3 = \beta_3 - 3c_n\beta_2\beta_1 + 2c_n^2\beta_1^3,\\
&\alpha_4 = \beta_4 - 4c_n\beta_3\beta_1 - 2c_n\beta_2^2 + 10c_n^2\beta_2\beta_1^2 - 5c_n^3\beta_1^4,\\
&\alpha_5 = \beta_5 - 5c_n\beta_1\beta_4 - 5c_n\beta_2\beta_3 + 15c_n^2\beta_1^2\beta_3 + 15c_n^2\beta_1\beta_2^2 - 35c_n^3\beta_1^3\beta_2 + 14c_n^4\beta_1^5,\\
&\alpha_6 = \beta_6 - 6c_n\beta_1\beta_5 - 6c_n\beta_2\beta_4 - 3c_n\beta_3^2 + 21c_n^2\beta_1^2\beta_4 + 42c_n^2\beta_1\beta_2\beta_3 + 7c_n^2\beta_2^3 - 56c_n^3\beta_1^3\beta_3\\
&\qquad - 84c_n^3\beta_1^2\beta_2^2 + 126c_n^4\beta_1^4\beta_2 - 42c_n^5\beta_1^6.
\end{aligned}$$

When the support of H is bounded, the support of $F^{c,H}$ is also bounded. Then from Lebesgue's dominated theorem, we can get that as $(n, p) \to \infty$,

$$\hat{\beta}_k = \int x^k dF_n(x) = -\frac{1}{2\pi i}\oint_C z^k s_n(z)dz \xrightarrow{a.s.} -\frac{1}{2\pi i}\oint_C z^k s(z)dz = \tilde{\beta}_k,$$

where the contour C is simple, closed, taken in the positive direction in the complex plane, and enclosing the support of $F^{c,H}$. Thus, as $(n, p) \to \infty$, we can get

$$\hat{\alpha}_k - \alpha_k \xrightarrow{a.s.} 0, \quad k = 1, 2, \ldots, 6,$$

which is the first conclusion of this lemma.

As for the second conclusion, we firstly derive the limiting distribution of $(\hat{\beta}_1, \hat{\beta}_2, \ldots, \hat{\beta}_6)'$. From the central limit theorem in Silverstein and Bai (2004), let functions $f(x) = x^k$, $k = 1, 2, \ldots, 6$, we can get

$$n(\hat{\beta}_1 - \beta_1, \hat{\beta}_2 - \beta_2, \ldots, \hat{\beta}_6 - \beta_6)' \xrightarrow{D} N_6(\zeta, \varphi),$$

where the mean vector $\zeta = (\zeta_j)$ with

$$\zeta_j = -\frac{1}{2\pi i}\oint_{C_1} \frac{z^j \underline{s}^3 \int t^2(1 + t\underline{s}(z))^{-3}dH(t)}{(1 - c\int \underline{s}^2(z)t^2(1 + t\underline{s}(z))^{-2}dH(t))^2}dz, \tag{A.1}$$

and the covariance $\varphi = (\varphi_{ij})$ with its entries

$$\varphi_{ij} = -\frac{1}{2\pi^2 c^2}\oint_{C_2}\oint_{C_1}\frac{z_1^i z_2^j}{(\underline{s}(z_1)-\underline{s}(z_2))^2}\underline{s}'(z_1)\underline{s}'(z_2)dz_1dz_2, \tag{A.2}$$

where $\underline{s}(z) = -(1-c)/z + cs(z)$ and the contours C_1 and C_2 in (A.1) and (A.2) are simple, closed, non-overlapping, taken in the positive direction in the complex plane, and each enclosing the support of $F^{c,H}$. The contour integrals in (A.1) and (A.2) can been derived from Qin and Li (2017),

$$\begin{aligned}
&\zeta_1 = 0, \quad \zeta_2 = \tilde{\alpha}_2, \quad \zeta_3 = 3(c\tilde{\alpha}_1\tilde{\alpha}_2 + \tilde{\alpha}_3), \quad \zeta_4 = 6c^2\tilde{\alpha}_1^2\tilde{\alpha}_2 + 5c\tilde{\alpha}_2^2 + 12c\tilde{\alpha}_1\tilde{\alpha}_3 + 6\tilde{\alpha}_4,\\
&\zeta_5 = 10c^3\tilde{\alpha}_1^3\tilde{\alpha}_2 + 30c^2\tilde{\alpha}_1^2\tilde{\alpha}_3 + 25c^2\tilde{\alpha}_1\tilde{\alpha}_2^2 + 30c\tilde{\alpha}_1\tilde{\alpha}_4 + 25c\tilde{\alpha}_2\tilde{\alpha}_3 + 10\tilde{\alpha}_5,\\
&\zeta_6 = 15c^4\tilde{\alpha}_1^4\tilde{\alpha}_2 + 60c^3\tilde{\alpha}_1^3\tilde{\alpha}_3 + 75c^3\tilde{\alpha}_1^2\tilde{\alpha}_2^2 + 90c^2\tilde{\alpha}_1^2\tilde{\alpha}_4 + 150c^2\tilde{\alpha}_1\tilde{\alpha}_2\tilde{\alpha}_3 + 21c^2\tilde{\alpha}_2^3 + 60c\tilde{\alpha}_1\tilde{\alpha}_5\\
&\quad + 51\tilde{\alpha}_2\tilde{\alpha}_4 + 24c\tilde{\alpha}_3^2 + 15\tilde{\alpha}_6,
\end{aligned}$$

and

$$\varphi_{ij} = \frac{2}{c^2}\sum_{k=0}^{i-1}(i-k)q_{i,k}q_{j,i+j-k},$$

where $q_{j,k}$ is the coefficient of z^k in the Taylor expansion of $(-1-c\sum_{l=1}^{\infty}\tilde{\alpha}_l(-z)^l)^j$.

Then, we can get the final result using the Delta method and Slutsky's theorem. Let $t = (x, y, z, u, v, w)'$ and define a vector function T_n

$$\begin{aligned}
T_n(t) = {} & (x, \gamma_2(y - c_nx^2), \gamma_3(z - 3c_nxy + 2c_n^2x^3), \gamma_4(u - 4c_nxz - \frac{2n^2+3n-6}{n^2+n+2}c_ny^2\\
& + \frac{10n^2+12n}{n^2+n+2}c_n^2x^2y - \frac{5n^2+6n}{n^2+n+2}c_n^3x^4), v - 5c_nxu - 5c_nyz + 15c_n^2x^2z + 15c_n^2xy^2\\
& - 35c_n^3x^3y + 14c_n^4x^5, w - 6c_nxv - 6c_nyu = 3c_nz^2 + 21c_n^2x^2u + 42c_n^2xyz + 7c_n^2y^3\\
& - 56c_n^3x^3z - 84c_n^3x^2y^2 + 126c_n^4x^4y - 42c_n^5x^6)'.
\end{aligned}$$

Obviously, T_n has continuous partial derivative at $b = (\beta_1, \beta_2, \ldots, \beta_6)'$ and the Jacobian matrix $J_n(b) = \partial T_n(t)/\partial t|_{t=b}$ converges to a limit $J(b)$, as $(n, p) \to \infty$,

$$J_n(b) = \begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0\\
-2c\tilde{\alpha}_1 & 1 & 0 & 0 & 0 & 0\\
3c^2\tilde{\alpha}_1^2 - 3c\tilde{\alpha}_2 & -3c\tilde{\alpha}_1 & 1 & 0 & 0 & 0\\
-4c^3\tilde{\alpha}_1^3 + 8c^2\tilde{\alpha}_1\tilde{\alpha}_2 - 4c\tilde{\alpha}_3 & 6c^2\tilde{\alpha}_1^2 - 4c\tilde{\alpha}_2 & -4c\tilde{\alpha}_1 & 1 & 0 & 0\\
5c^4\tilde{\alpha}_1^4 - 15c^3\tilde{\alpha}_1^2\tilde{\alpha}_2 + 10c^2\tilde{\alpha}_1\tilde{\alpha}_3 + 5c^2\tilde{\alpha}_2^2 - 5c\tilde{\alpha}_4 & -10c^3\tilde{\alpha}_1^3 + 15c^2\tilde{\alpha}_1\tilde{\alpha}_2 - 5c\tilde{\alpha}_3 & 10c^2\tilde{\alpha}_1^2 - 5c\tilde{\alpha}_2 & -5c\tilde{\alpha}_1 & 1 & 0\\
-6c^5\tilde{\alpha}_1^5 + 24c^4\tilde{\alpha}_1^3\tilde{\alpha}_2 - 18c^3\tilde{\alpha}_1\tilde{\alpha}_2^2 - 18c^3\tilde{\alpha}_1\tilde{\alpha}_2^2 + 12c^2\tilde{\alpha}_1\tilde{\alpha}_4 + 12c^2\tilde{\alpha}_2\tilde{\alpha}_3 - 6c\tilde{\alpha}_5 & 15c^4\tilde{\alpha}_1^4 - 36c^3\tilde{\alpha}_1^2\tilde{\alpha}_2 + 18c^2\tilde{\alpha}_1\tilde{\alpha}_3 + 9c^2\tilde{\alpha}_2^2 - 6c\tilde{\alpha}_4 & -20c^3\tilde{\alpha}_1^3 + 24c^3\tilde{\alpha}_1\tilde{\alpha}_2 - 6c\tilde{\alpha}_3 & 15c^2\tilde{\alpha}_1^2 - 6c\tilde{\alpha}_2 & -6c\tilde{\alpha}_1 & 1
\end{pmatrix}.$$

Therefore, as $(n, p) \to \infty$, we can get

$$n(\tilde{\alpha}_1 - \alpha_1, \tilde{\alpha}_1 - \alpha_2, \ldots, \tilde{\alpha}_6 - \alpha_6)' + n((\alpha_1, \alpha_2, \ldots, \alpha_6)' - T_n(b)) \xrightarrow{D} N_6(J(b)\zeta, J(b)\varphi J'(b)).$$

Through some elementary calculations, we can derive that

$$n((\alpha_1, \alpha_2, \ldots, \alpha_6)' - T_n(b)) \to J(b)\zeta,$$

where $J(b)\zeta = (0, \tilde{\alpha}_2, 3\tilde{\alpha}_3, 6\tilde{\alpha}_4 + c\tilde{\alpha}_2^2, 5c\tilde{\alpha}_2\tilde{\alpha}_3 + 10\tilde{\alpha}_5, 15\tilde{\alpha}_6 + 9c\tilde{\alpha}_2\tilde{\alpha}_4 + 6c\tilde{\alpha}_3^2)'$, and $J(b)\varphi J'(b) := \Omega = (\Omega_{ij})$ with its entries

$$\Omega_{11} = 2\tilde{\alpha}_2/c, \quad \Omega_{12} = 4\tilde{\alpha}_3/c, \quad \Omega_{13} = 6\tilde{\alpha}_4/c, \quad \Omega_{14} = 8\tilde{\alpha}_5/c, \quad \Omega_{15} = 10\tilde{\alpha}_6/c, \quad \Omega_{16} = 12\tilde{\alpha}_7/c$$

$$\begin{aligned}
&\Omega_{22}=4(2\tilde{\alpha}_4/c+\tilde{\alpha}_2^2),\quad \Omega_{23}=12(\tilde{\alpha}_5/c+\tilde{\alpha}_3\tilde{\alpha}_2),\quad \Omega_{24}=8(2\tilde{\alpha}_6/c+2\tilde{\alpha}_4\tilde{\alpha}_2+\tilde{\alpha}_3^2),\\
&\Omega_{25}=20(\tilde{\alpha}_2\tilde{\alpha}_5+\tilde{\alpha}_3\tilde{\alpha}_4+\tilde{\alpha}_7/c),\quad \Omega_{26}=12(2\tilde{\alpha}_2\tilde{\alpha}_6+2\tilde{\alpha}_3\tilde{\alpha}_5+\tilde{\alpha}_4^2+\tilde{\alpha}_8/c)\\
&\Omega_{33}=6(3\tilde{\alpha}_6/c+3\tilde{\alpha}_2\tilde{\alpha}_4+3\tilde{\alpha}_3^2+c\tilde{\alpha}_2^3),\quad \Omega_{34}=24(\tilde{\alpha}_7/c+\tilde{\alpha}_2\tilde{\alpha}_5+2\tilde{\alpha}_3\tilde{\alpha}_4+c\tilde{\alpha}_2^2\tilde{\alpha}_3),\\
&\Omega_{35}=30(\tilde{\alpha}_2\tilde{\alpha}_6+2\tilde{\alpha}_3\tilde{\alpha}_5+\tilde{\alpha}_4^2+\tilde{\alpha}_8/c+c(\tilde{\alpha}_2^2\tilde{\alpha}_4+\tilde{\alpha}_2\tilde{\alpha}_3^2)),\quad \Omega_{36}=12(3\tilde{\alpha}_2\tilde{\alpha}_7+6\tilde{\alpha}_3\tilde{\alpha}_6+6\tilde{\alpha}_4\tilde{\alpha}_5\\
&\qquad+3\tilde{\alpha}_9/c+c(3\tilde{\alpha}_2^2\tilde{\alpha}_5+6\tilde{\alpha}_2\tilde{\alpha}_3\tilde{\alpha}_4+\tilde{\alpha}_3^3)),\\
&\Omega_{44}=8(4\tilde{\alpha}_8/c+6\tilde{\alpha}_4^2+4\tilde{\alpha}_2\tilde{\alpha}_6+8\tilde{\alpha}_3\tilde{\alpha}_5+4c\tilde{\alpha}_2^2\tilde{\alpha}_4+8c\tilde{\alpha}_2\tilde{\alpha}_3^2+c^2\tilde{\alpha}_2^4),\\
&\Omega_{45}=40(c^2\tilde{\alpha}_2^3\tilde{\alpha}_3+\tilde{\alpha}_2\tilde{\alpha}_7+2\tilde{\alpha}_3\tilde{\alpha}_6+3\tilde{\alpha}_4\tilde{\alpha}_5+\tilde{\alpha}_9/c+c\tilde{\alpha}_2^2\tilde{\alpha}_3+4c\tilde{\alpha}_2\tilde{\alpha}_3\tilde{\alpha}_4+c\tilde{\alpha}_3^2),\\
&\Omega_{46}=24(2\tilde{\alpha}_{10}/c+2\tilde{\alpha}_2\tilde{\alpha}_8+4\tilde{\alpha}_3\tilde{\alpha}_7+6\tilde{\alpha}_4\tilde{\alpha}_6+3\tilde{\alpha}_5^2+c^2(2\tilde{\alpha}_2^2\tilde{\alpha}_4+3\tilde{\alpha}_2^2\tilde{\alpha}_3^2)+2c\tilde{\alpha}_2^2\tilde{\alpha}_6+8c\tilde{\alpha}_2\tilde{\alpha}_3\tilde{\alpha}_5\\
&\qquad+4c\tilde{\alpha}_2\tilde{\alpha}_4^2+6c\tilde{\alpha}_3^2\tilde{\alpha}_4),\\
&\Omega_{55}=10(5\tilde{\alpha}_{10}/c+c^3\tilde{\alpha}_2^5+5\tilde{\alpha}_2\tilde{\alpha}_8+10\tilde{\alpha}_3\tilde{\alpha}_7+15\tilde{\alpha}_4\tilde{\alpha}_6+10\tilde{\alpha}_5^2+5c^2\tilde{\alpha}_2^3\tilde{\alpha}_4+15c^2\tilde{\alpha}_2^2\tilde{\alpha}_3^2+5c\tilde{\alpha}_2^2\tilde{\alpha}_6\\
&\qquad+20c\tilde{\alpha}_2\tilde{\alpha}_3\tilde{\alpha}_5+15c\tilde{\alpha}_2\tilde{\alpha}_4^2+20\tilde{\alpha}_5\tilde{\alpha}_6),\\
&\Omega_{56}=60(\tilde{\alpha}_{11}/c+c^3\tilde{\alpha}_2^4\tilde{\alpha}_3+\tilde{\alpha}_2\tilde{\alpha}_9+2\tilde{\alpha}_3\tilde{\alpha}_8+3\tilde{\alpha}_4\tilde{\alpha}_7+4\tilde{\alpha}_5\tilde{\alpha}_6+c^2\tilde{\alpha}_2^3\tilde{\alpha}_5+6c^2\tilde{\alpha}_2^2\tilde{\alpha}_3\tilde{\alpha}_4+3c^2\tilde{\alpha}_2\tilde{\alpha}_3^2\\
&\qquad+c\tilde{\alpha}_2^2\tilde{\alpha}_7+4c\tilde{\alpha}_2\tilde{\alpha}_3\tilde{\alpha}_6+6c\tilde{\alpha}_2\tilde{\alpha}_4\tilde{\alpha}_5+4c\tilde{\alpha}_3^2\tilde{\alpha}_5+5c\tilde{\alpha}_3\tilde{\alpha}_4^2),\\
&\Omega_{66}=12(6\tilde{\alpha}_2\tilde{\alpha}_{10}+6\tilde{\alpha}_{12}/c+c^4\tilde{\alpha}_2^6+12\tilde{\alpha}_3\tilde{\alpha}_9+18\tilde{\alpha}_4\tilde{\alpha}_8+24\tilde{\alpha}_5\tilde{\alpha}_7+15\tilde{\alpha}_6^2+c^3(6\tilde{\alpha}_2^4\tilde{\alpha}_4+24\tilde{\alpha}_2^3\tilde{\alpha}_3^2)\\
&\qquad+c^2(6\tilde{\alpha}_2^3\tilde{\alpha}_6+36\tilde{\alpha}_2\tilde{\alpha}_4\tilde{\alpha}_6+27\tilde{\alpha}_2^2\tilde{\alpha}_4^2+72\tilde{\alpha}_2\tilde{\alpha}_3^2\tilde{\alpha}_4+9\tilde{\alpha}_3^4)+c(6\tilde{\alpha}_2^2\tilde{\alpha}_8+24\tilde{\alpha}_2\tilde{\alpha}_3\tilde{\alpha}_7+36\tilde{\alpha}_2\tilde{\alpha}_4\tilde{\alpha}_6\\
&\qquad+24\tilde{\alpha}_2\tilde{\alpha}_5^2+24\tilde{\alpha}_3^2\tilde{\alpha}_6+72\tilde{\alpha}_3\tilde{\alpha}_4\tilde{\alpha}_5+14\tilde{\alpha}_4^3)).
\end{aligned}$$

The proof is completed.

A.2. Proof of Lemma 2

Supposing the assumptions in this lemma hold, from Theorem 1.4 in Pan and Zhou (2008), we can know that

$$n(\hat{\beta}_1-\beta_1,\hat{\beta}_2-\beta_2,\ldots,\hat{\beta}_6-\beta_6)'\xrightarrow{D}N_6(\overline{\zeta},\overline{\varphi}),$$

where the mean vector $\overline{\zeta}=(\overline{\zeta}_j)$ with

$$\begin{aligned}
\overline{\zeta}_j&=\zeta_j-\frac{E(\omega_{11}^4)-3}{2\pi i}\oint_{C_1}\frac{z^j\underline{s}^3(z)\int t^2(1+t\underline{s}(z))^{-3}dH(t)}{1-c\int\underline{s}^2(z)t^2(1+t\underline{s}(z))^{-2}dH(t)}dz\\
&:=\zeta_j+\Delta L_j
\end{aligned}\tag{A.3}$$

and the covariance $\overline{\varphi}=(\overline{\varphi}_{ij})$ with its entries

$$\begin{aligned}
\overline{\varphi}_{ij}&=\varphi_{ij}-\frac{E(\omega_{11}^4-3)}{4\pi^2c}\oint_{C_2}\oint_{C_1}z_1^iz_2^j\frac{d^2}{dz_1dz_2}\int\frac{t^2\underline{s}(z_1)\underline{s}(z_2)dH(t)}{(1+t\underline{s}(z_1))(1+t\underline{s}(z_2))}dz_1dz_2\\
&:=\varphi_{ij}+\Delta K_{ij}
\end{aligned}\tag{A.4}$$

where $\Delta=E(\omega_{11}^4-3)$, ζ_j and φ_{ij} are defined in (A.1) and (A.2), respectively.

Without loss of generality, let the contour C_2 enclose C_1 and both of them be away from the support S_F of $F^{c,H}$. Denote $\underline{s}(C_i)=\{\underline{s}(z):z\in C_i\},\quad i=1,2$. Then, similar to Qin and Li (2017), the two contours $\underline{s}(C_1)$ and $\underline{s}(C_2)$ are also simple, closed, and non-overlapping, and are taken in the negative direction. Besides, $\underline{s}(C_2)$ encloses $\underline{s}(C_1)$ and both of them enclose zero. Let

$$p(\underline{s})=-1+c\int\frac{t\underline{s}}{1+t\underline{s}}dH(t),\quad Q(\underline{s})=\int\frac{t^2}{(1+t\underline{s})^3}dH(t),$$

the integral L_j in Eq. (A.3) can be simplified

$$\begin{aligned}
L_j&=-\frac{1}{2\pi i}\oint_{C_1}z^j\underline{s}(z)\underline{s}'(z)\int t^2(1+t\underline{s}(z))^{-3}dH(t)dz\\
&=-\frac{1}{2\pi i}\oint_{\underline{s}(C_1)}\frac{P^j(\underline{s}Q(\underline{s}))}{\underline{s}^{j-1}}d\underline{s}\\
&=\begin{cases}0 & j=1\\ \frac{1}{(j-2)!}[P^j(z)Q(z)]^{(j-2)}|_{z=0} & j\geq 2,\end{cases}
\end{aligned}$$

and the integral K_{ij} in (A.4) becomes

$$K_{ij} = -\frac{1}{4\pi^2 c}\oint_{C_2}\oint_{C_1} z_1^i z_2^j \int \frac{t^2 \underline{s}'(z_1)\underline{s}'(z_2)dH(t)}{(1+t\underline{s}(z_1))^2(1+t\underline{s}(z_2))^2}dz_1 dz_2$$
$$= -\frac{1}{4\pi^2 c}\oint_{\underline{s}(C_2)}\oint_{\underline{s}(C_1)} \frac{P^i(\underline{s}_1)P^j(\underline{s}_2)}{\underline{s}_1^i \underline{s}_2^j}\int \frac{t^2 dH(t)}{(1+t\underline{s}(z_1))^2(1+t\underline{s}(z_2))^2}d\underline{s}_1\underline{s}_2$$
$$= \frac{1}{c}\int \frac{t^2}{(i-1)!(j-1)!}[\frac{P^i(z)}{(1+tz)^2}]^{(i-1)}[\frac{P^j(z)}{(1+tz)^2}]^{(j-1)}|_{z=0}dH(t),$$

where the contours integrals are calculated from the Cauchy integral theorem. Similar to the procedures in Lemma 1, we can get

$$n(\tilde{\alpha}_1-\alpha_1, \tilde{\alpha}_1-\alpha_2, \ldots, \tilde{\alpha}_6-\alpha_6)' \xrightarrow{D} N_6(J(b)(\overline{\zeta}-\zeta), J(b)\overline{\varphi}J'(b)).$$

Through some elementary calculations, we can get

$$J(b)(\overline{\zeta}-\zeta) = \Delta\cdot(0, \tilde{\alpha}_2, 3\tilde{\alpha}_3, 6\tilde{\alpha}_4, 10\tilde{\alpha}_5, 15\tilde{\alpha}_6)'$$

and

$$J(b)(\overline{\varphi}-\varphi)J'(b) = \frac{\Delta}{c}\begin{pmatrix} \tilde{\alpha}_2 & 2\tilde{\alpha}_3 & 3\tilde{\alpha}_4 & 4\tilde{\alpha}_5 & 5\tilde{\alpha}_6 & 6\tilde{\alpha}_7 \\ 2\tilde{\alpha}_3 & 4\tilde{\alpha}_4 & 6\tilde{\alpha}_5 & 8\tilde{\alpha}_6 & 10\tilde{\alpha}_7 & 12\tilde{\alpha}_8 \\ 3\tilde{\alpha}_4 & 6\tilde{\alpha}_5 & 9\tilde{\alpha}_6 & 12\tilde{\alpha}_7 & 15\tilde{\alpha}_8 & 18\tilde{\alpha}_9 \\ 4\tilde{\alpha}_5 & 8\tilde{\alpha}_6 & 12\tilde{\alpha}_7 & 16\tilde{\alpha}_8 & 20\tilde{\alpha}_9 & 24\tilde{\alpha}_{10} \\ 5\tilde{\alpha}_6 & 10\tilde{\alpha}_7 & 15\tilde{\alpha}_8 & 20\tilde{\alpha}_9 & 25\tilde{\alpha}_{10} & 30\tilde{\alpha}_{11} \\ 6\tilde{\alpha}_7 & 12\tilde{\alpha}_8 & 18\tilde{\alpha}_9 & 24\tilde{\alpha}_{10} & 30\tilde{\alpha}_{11} & 36\tilde{\alpha}_{12} \end{pmatrix},$$

which complete the proof.

A.3. Proof of Theorems 1 and 2

We first prove Theorem 1. Using a standard application of the Delta method, we can easily get this theorem from Lemmas 1 and 2. Let $t = (x, y, z, u, v, w)'$, and define a function $T = w - 2z + 1$. Obviously, T has continuous partial derivative at $a = (\alpha_1, \alpha_2, \ldots, \alpha_6)'$, and $T'(a) = \partial T/\partial t|_{t=a} = (0, 0, -2, 0, 0, 1)'$. If Σ_p is diagonal for all p large, from Lemma 2 , we can get

$$n(\hat{\alpha}_6 - 2\hat{\alpha}_3 - (\alpha_6 - 2\alpha_3)) \xrightarrow{D} N(\mu^\star, \sigma^{\star 2}).$$

where the limiting mean $\mu^\star = \Delta\cdot T'(a)(0, \tilde{\alpha}_2, 3\tilde{\alpha}_3, 6\tilde{\alpha}_4, 10\tilde{\alpha}_5, 15\tilde{\alpha}_6)' = \Delta\cdot(-6\tilde{\alpha}_3 + 15\tilde{\alpha}_6)$, and the variance $\sigma^{\star 2} = T'(a)\Gamma T'(a)^\top = 4\Gamma_{33} - 4\Gamma_{36} + \Gamma_{66}$, Γ is defined in (3). If $E(\omega_{11}^4) = 3$, from Lemma 1, we can get $\mu^\star = 0$, $\sigma^{\star 2} = T'(a)\Omega T'(a)^\top = 4\Omega_{33} - 4\Omega_{36} + \Omega_{66}$, Ω is defined in (2).

Then as $(n, p) \to \infty$, we can get

$$\frac{n}{\sqrt{12(c^4 + 30c^3 + 150c^2 + 162c + 27)}}(\hat{\alpha}_6 - 2\hat{\alpha}_3 - (\alpha_6 - 2\alpha_3)) \xrightarrow{D} N(\mu, \sigma_T^2),$$

where

$$\mu = \mu^\star/\sqrt{12(c^4 + 30c^3 + 150c^2 + 162c + 27)},$$
$$\sigma_T^2 = \sigma^{\star 2}/12(c^4 + 30c^3 + 150c^2 + 162c + 27).$$

Similarly, we can get the result of Theorem 2 by using different vector function. The proof is completed.

References

Bai, Z.D., Chen, J.Q., Yao, J.F., 2010. On estimation of the population spectral distribution from a high-dimensional sample covariance matrix. Aust. N. Z. J. Stat. 52, 423–437.
Bai, Z.D., Silverstein, J.W., 2010. Spectral Analysis of Large Dimensional Random Matrices, second ed. Springer, New York.
Chen, S.X., Zhang, L.X., Zhong, P.S., 2010. Tests for high-dimensional covariance matrices. J. Amer. Statist. Assoc. 109, 810–819.
Fisher, T.J., 2012. On testing for an identity covariance matrix when the dimensionality equals or exceeds the sample size. J. Statist. Plann. Inference 142, 312–326.
Fisher, T.J., Sun, X., Gallagher, C.M., 2010. A new test for sphericity of the covariance matrix for high dimensional data. J. Multivariate Anal. 101, 2554–2570.
Hisao, N., 1973. On some test criteria for covariance matrix. Ann. Statist. 1, 700–709.
John, S., 1971. Some optimal multivariate tests. Biometrika 58, 123–127.
Ledoit, O., Wolf, M., 2002. Some hypothesis tests for the covariance matrix when the dimension is large compared to the sample size. Ann. Statist. 30, 1081–1102.
Li, W., Qin, Y., 2014. Hypothesis testing for high-dimensional covariance matrices. J. Multivariate Anal. 128, 108–119.

Nica, A., Speicher, R., 2006. Lectures on the Combinatorics of Free Probability. Cambridge University Press, New York.
Pan, G.M., Zhou, W., 2008. Central limit theorem for signal-to-interference ratio of reduced rank linear receiver. Ann. Appl. Probab. 18, 1232–1270.
Qin, Y., Li, W., 2017. Bias-reduced moment estimators of population spectral distribution and their application. In: Big and Complex Data Analysis. Springer International Publishing.
Silverstein, J.W., 1995. Strong convergence of the empirical distribution of eigenvalues oflarge-dimensional random matrices. J. Multivariate Anal. 55, 331–339.
Silverstein, J.W., Bai, Z.D., 2004. CLT for linear spectral statistics of large-dimensional sample covariance matrices. Ann. Probab. 32, 553–605.
Srivastava, M.S., 2005. Some tests concerning the covariance matrix in high dimensional data. J. Japan Statist. Soc. 35, 251–272.
Srivastava, M.S., 2006. Some tests criteria for the covariance matrix with fewer observations than the dimension. Acta Comment. Univ. Tartu. Math. 10, 77–93.
Srivastava, M.S., Kollo, T., Rosen, D.V., 2011. Some tests for the covariance matrix with fewer observations than the dimension under non-normality. J. Multivariate Anal. 102, 1090–1103.
Wang, Q.W., Yao, J.F., 2013. On the sphericity test with large-dimensional observations. Electron. J. Stat. 7, 2164–2192.
Zheng, S.R., Bai, Z.D., 2013. A note on central limit theorems for linear spectral statistics of large dimensional F matrix. arXiv:1305.0339.