J. Japan Statist. Soc.
Vol. 35 No. 2 2005 251–272

SOME TESTS CONCERNING THE COVARIANCE MATRIX IN HIGH DIMENSIONAL DATA

Muni S. Srivastava*

In this paper, tests are developed for testing certain hypotheses on the covariance matrix Σ, when the sample size $N = n+1$ is smaller than the dimension p of the data. Under the condition that $(\operatorname{tr}\Sigma^i/p)$ exists and > 0, as $p \to \infty$, $i = 1, \ldots, 8$, tests are developed for testing the hypotheses that the covariance matrix in a normally distributed data is an identity matrix, a constant time the identity matrix (spherecity), and is a diagonal matrix. The asymptotic null and non-null distributions of these test statistics are given.

Key words and phrases: Asymptotic distributions, multivariate normal, null and non-null distributions, sample size smaller than the dimension.

1. Introduction

Recent advances in analyzing high dimensional data with fewer observations require that certain assumptions made implicitly or explicitly in analyzing them should be ascertained. For example, Dudoit *et al.* (2002) in their analysis of microarrays data on genes assume that the covariance matrices are diagonal matrices, and thus their distance function uses only the diagonal elements of the sample covariance matrix. The good performance of their procedures appear to suggest that this indeed might be the case. To ascertain these assumptions on the covariance matrix, the likelihood ratio tests cannot be used as the sample size $N = n+1$ could be smaller than the dimension p. Although, the locally best invariant (LBI) test proposed by John (1971) and considered by Suguira (1972), and Nagao (1973) for the spherecity hypothesis and the LBI test given by Nagao (1973) for testing the hypothesis that the covariance matrix Σ is an identity matrix can be computed for all sample sizes, there appears to be no theoretical justification for using them as LBI tests cannot be obtained when $n < p$. Thus, we consider a distance function between the null hypothesis and the alternative hypothesis, and propose tests based on consistent estimators of these parametric functions of the covariance matrix Σ for testing the hypothesis of spherecity of the covarinace matrix, and for testing the hypothesis that the covarinace matrix is an identity matrix. In addition, under the same set of conditions, we provide tests for testing the hypothesis that the covarinace matrix is a diagonal matrix. Asymptotic distributions of these test statistics are given under the hypothesis as well as under the alternative hypothesis. Our focus is however for the case when $n = O(p^\delta), 0 < \delta \leq 1$. Thus, it includes the case when $(n/p) \to 0$. The

Received October 6, 2004. Revised January 4, 2005. Accepted March 4, 2005.
*Department of Statistics, University of Toronto, 100 St George Street, Toronto, Canada M5S 3G3. Email: srivasta@utstat.toronto.edu

organization of the paper is as follows.

In Section 2, we present notations and some preliminary results. Sections 3 to 5 develop tests for testing that the covariance matrix $\Sigma = \sigma^2 I$, $\Sigma = I_p$, and $\Sigma = \text{diag}(\sigma_{11}, \ldots, \sigma_{pp})$ respectively. The proofs of theorems and lemmas stated in Sections 2 to 4 are given in Section 6.

2. Notations and preliminaries

Let $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N$ be independently and identically distributed (iid) as a p-dimensional random vector $\boldsymbol{x}$ which is distributed as multivariate normal with mean vector $\boldsymbol{\mu}$ and covariance matrix Σ, denoted, $\boldsymbol{x} \sim N_p(\boldsymbol{\mu}, \Sigma)$. We shall assume that $\Sigma > 0$, that is, it is a positive definite matrix. Let $\overline{\boldsymbol{x}}$ and S denote the sample mean vector and the sample covariance matrix respectively, defined as

$$\overline{\boldsymbol{x}} = N^{-1} \sum_{\alpha=1}^{N} \boldsymbol{x}_\alpha, \qquad N = n + 1, \tag{2.1}$$

and

$$S \equiv n^{-1} V = n^{-1} \sum_{\alpha=1}^{N} (\boldsymbol{x}_\alpha - \overline{\boldsymbol{x}})(\boldsymbol{x}_\alpha - \overline{\boldsymbol{x}})' \tag{2.2}$$

respectively. Let

$$a_i = (\text{tr}\, \Sigma^i / p), \qquad i = 1, \ldots, 8.$$

We make the following assumptions:

$$\text{(A)} : \text{As } p \to \infty,\ a_i \to a_i^0,\ 0 < a_i^0 < \infty,\ i = 1, \ldots, 8.$$
$$\text{(B)} : n = O(p^\delta),\ 0 < \delta \leq 1.$$

In the next lemma, we give unbiased and consistent estimators of a_1 and a_2.

Lemma 2.1. *Under the assumption* (A), *and as $n \to \infty$, an unbiased and consistent estimators of a_1 and a_2 are respectively given by*

$$\hat{a}_1 = (\text{tr}\, S)/p, \tag{2.3}$$

and

$$\hat{a}_2 = \frac{n^2}{(n-1)(n+2)} \frac{1}{p} \left[\text{tr}\, S^2 - \frac{1}{n} (\text{tr}\, S)^2 \right]. \tag{2.4}$$

For proof, *see Section* 6.

Theorem 2.1. *Let $\hat{a}_1$ and $\hat{a}_2$ be as defined in* (2.3) *and* (2.4) *respectively. Then under the assumption* (A), *asymptotically*

$$\begin{pmatrix} \hat{a}_1 \\ \hat{a}_2 \end{pmatrix} \sim N_2 \left[\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, (np)^{-1} \begin{pmatrix} 2a_2 & 4a_3 \\ 4a_3 & 8a_4 + 4(p/n)a_2^2 \end{pmatrix} \right].$$

The proof is given in Section 6.

Remark 2.1. From the definition of $\hat{a}_1$ and $\hat{a}_2$, it follows that

$$\left(\frac{\operatorname{tr} S^2}{p}\right) = \hat{a}_2 + \frac{1}{pn}(\operatorname{tr} S)^2 = \hat{a}_2 + \frac{p}{n}\hat{a}_1^2.$$

Thus, unless (p/n) goes to zero as n and $p \to \infty$, $(\operatorname{tr} S^2/p)$ is not a consistent estimator of a_2. That is, if $n = O(p)$, $(\operatorname{tr} S^2/p)$ is not a consistent estimator of $(\operatorname{tr} \Sigma^2/p)$, while $\hat{a}_2$ is always a consistent estimator of a_2 irrespective of how $n \to \infty$, provided the assumption (A) is satisfied.

COROLLARY 2.1. *Let n and $p \to \infty$ such that $(p/n) \to c$. Then, asymptotically*

$$\begin{pmatrix} p^{-1}(\operatorname{tr} S) \\ p^{-1}(\operatorname{tr} S^2) \end{pmatrix} \sim N_2\left[\begin{pmatrix} a_1 \\ a_2 + ca_1^2 \end{pmatrix}, n^{-2}c^{-1}\Delta\right],$$

where

$$\Delta = \begin{pmatrix} 2a_2 & 4(ca_1a_2 + a_3) \\ 4(ca_1a_2 + a_3) & 4(2a_4 + ca_2^2 + 4ca_1a_3 + 2c^2a_1^2a_2) \end{pmatrix}.$$

The proof is given in Section 6.

Let $A = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_q) : p \times q$. Then $\operatorname{vec}(A)$ is a $pq \times 1$ vector defined by $\operatorname{vec}(A) = (\boldsymbol{a}_1', \ldots, \boldsymbol{a}_q')'$. Now, we consider some known asymptotic results.

LEMMA 2.2. $nS = n(s_{ij}) \sim W_p(\Sigma, n)$, *where* $\Sigma = (\sigma_{ij})$. *Then, if* $\boldsymbol{\theta} = \operatorname{vec}(\Sigma)$ *and* $B = (b_{ij})$,

$$\operatorname{vec}(B) \equiv \sqrt{n}(\operatorname{vec}(S) - \boldsymbol{\theta}) \textit{ is } AN_{p^2}(\mathbf{0}, \Omega),$$

where the elements of Ω *are given by*

$$E(b_{ij}b_{kl}) = \sigma_{ik}\sigma_{jl} + \sigma_{il}\sigma_{jk},$$

see Hsu (1949), *Bilodeau and Brenner* (1999, p. 79–80) *or Srivastava and Khatri* (1979, Problem 3.32, p. 103).

LEMMA 2.3. *Let* $\boldsymbol{g} = (g_1, \ldots, g_k)'$, *where $g_i's$ are differentiable functions of* $\operatorname{vec} S : p^2 \times 1$ *at* $\boldsymbol{\theta} = \operatorname{vec} \Sigma$. *Then*

$$\sqrt{n}(\boldsymbol{g}(\operatorname{vec}(S)) - \boldsymbol{g}(\boldsymbol{\theta})) \to N_k(\mathbf{0}, D\boldsymbol{g}(\boldsymbol{\theta})\Omega(D\boldsymbol{g}(\boldsymbol{\theta}))'),$$

where

$$D\boldsymbol{g}(\boldsymbol{\theta}) = \left(\frac{\partial \boldsymbol{g}}{\partial \operatorname{vec}(S)'}\right)_{\operatorname{vec} S = \boldsymbol{\theta}} : k \times p^2.$$

For proof, see Bilodeau and Brenner (1999, p. 79).

3. A test for the spherecity

In this section, we consider the problem of testing the hypothesis

$$H : \Sigma = \sigma^2 I \qquad \text{vs} \qquad A : \Sigma \neq \sigma^2 I,$$

when a sample of size $N = n + 1$, $\boldsymbol{x}_1, \dots, \boldsymbol{x}_N$ is drawn from $N_p(\boldsymbol{\mu}, \Sigma)$. When $n > p$, the most appropriate commonly used test statistic is the likelihood ratio test which has been shown by Carter and Srivastava (1977) to have a monotone power function. However, when $n < p$ the likelihood ratio test is not available. Even the competitive locally best invariant test which can be calculated for $n < p$, cannot be justified on theoretical grounds, although Ledoit and Wolf (2002) have proposed this LBI test and have given its asymptotic distribution when $(p/n) \to c$, a constant, but the non-null distribution is not available. In this paper, we consider a test based on a consistent estimator of a parametric function of Σ which seperates the null hypothesis from the alternative hypothesis which we discuss next.

The testing problem remains invariant under the transformation $\boldsymbol{x} \to G\boldsymbol{x}$, where G belongs to the group of orthogonal marices. The problem also remains invariant under the scalar transformation $\boldsymbol{x} \to c\boldsymbol{x}$. Thus, we may assume without any loss of generality that $\Sigma = \text{diag}(\lambda_1, \dots, \lambda_p)$, a $p \times p$ diagonal matrix. From the Cauchy-Schwarz inequality, it follows that

$$\left(\sum_{i=1}^{p} \lambda_i \times 1 \right)^2 \leq p \left(\sum \lambda_i^2 \right)$$

with equality holding if and only if $\lambda_i \equiv \lambda$ for some constant λ. Thus

$$\gamma_1 \equiv \frac{(\sum_{i=1}^{p} \lambda_i^2/p)}{(\sum_{i=1}^{p} \lambda_i/p)^2} \geq 1,$$

and is equal to one if and only if $\lambda_i = \lambda$ for some constant λ. Thus, γ_1 is equal to one if and only if $\lambda_i = \lambda$ for some constant λ. Hence, we may consider testing the hypothesis

$$H : \gamma_1 - 1 = 0 \qquad \text{vs} \qquad A : \gamma_1 - 1 > 0.$$

A test for the above hypothesis H vs A can be based on a consistent estimator of γ_1.

From Lemma 2.1, it follows that under the assumptions (A) and (B), a consistent estimator of γ_1 is given by

$$\hat{\gamma}_1 = \frac{n^2}{(n-1)(n+2)} \frac{1}{p} \left[\text{tr}\, S^2 - \frac{1}{n} (\text{tr}\, S)^2 \right] / (\text{tr}\, S/p)^2 = \frac{\hat{a}_2}{\hat{a}_1^2}.$$

Thus, a test for the spherecity can be based on the statistic

$$T_1 = \hat{\gamma}_1 - 1.$$

The following theorem gives the asymptotic non-null distribution of the statistic T_1.

Theorem 3.1. *Under the assumptions* (A) *and* (B), *asymptotically*

$$\left(\frac{n}{2}\right)(T_1 - \gamma_1 + 1) \sim N(0, \tau_1^2)$$

where

$$\tau_1^2 = \frac{2n(a_4 a_1^2 - 2a_1 a_2 a_3 + a_2^3)}{p a_1^6} + \frac{a_2^2}{a_1^4}.$$

Corollary 3.1. *Under the hypothesis that* $\gamma_1 = 1$, *and under the assumptions* (A) *and* (B), *asympototically*

$$\frac{n}{2} T_1 \sim N(0, 1).$$

Remark 3.1. The reader is reminded that it is a one-sided test for testing the hypothesis that $\gamma_1 = 1$ vs $\gamma_1 > 1$.

4. A test for the covariance matrix to be an identity matrix

Let $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N$ be iid $N_p(\boldsymbol{\mu}, \Sigma)$. In this section, we consider the problem of testing the hypothesis $H : \Sigma = I$ against the alternative $A : \Sigma \neq I$.

The usual likelihood ratio test for $n > p$ which has monotone power function as shown by Nagao (1967) does not exist when $n < p$. Even, the locally best invariant test proposed by Nagao (1973) cannot be justified theoretically. However, Ledoit and Wolf (2002) have porposed a modified version of Nagao's test and have given its asymptotic null distribution when $(p/n) \to c > 0$, where c is a constant but the non-null distribution is lacking. In this paper, we consider a distance function between the hypothesis and the alternative hypothesis and propose a test based on a consistent estimator of this parametric function, which we discuss next. Since the problem remains invariant under the group of orthogonal transformations, we may assume without any loss of generality that $\Sigma = \mathrm{diag}(\lambda_1, \ldots, \lambda_p)$, a diagonal marix and test the hypothesis that $\lambda_i = 1$, $i = 1, \ldots, p$ against the alternative that $\lambda_i \neq 1$, for at least one i, $i = 1, \ldots, p$. But $\lambda_i = 1$ for all i if and only if

$$\frac{1}{p} \Sigma_{i=1}^{p} (\lambda_i - 1)^2 = 0.$$

That is,

$$\begin{aligned} \frac{1}{p}\left[\sum \lambda_i^2 - 2\sum \lambda_i + p\right] &= \frac{1}{p}\left[\mathrm{tr}\,\Sigma^2 - 2\,\mathrm{tr}\,\Sigma + p\right] \\ &= a_2 - 2a_1 + 1 = 0. \end{aligned}$$

Thus, a test for the hypothesis $H : \Sigma = I$ vs $A : \Sigma \neq I$, can be based on a consistent estimator of

$$\gamma_2 = a_2 - 2a_1. \tag{4.1}$$

From Lemma 2.1, it follows that under the assumptions (A) and (B), an unbiased and consistent estimate of γ_2 is given by

$$\hat{\gamma}_2 = \frac{n^2}{p(n-1)(n+2)}\left[\operatorname{tr} S^2 - \frac{1}{n}(\operatorname{tr} S)^2\right] - \frac{2}{p}[\operatorname{tr} S] \tag{4.2}$$

$$= \hat{a}_2 - 2\hat{a}_1, \tag{4.3}$$

and our test will be based on the test statistic

$$T_2 = \hat{\gamma}_2 + 1. \tag{4.4}$$

THEOREM 4.1. *Under the assumptions* (A) *and* (B), *asymptotically*

$$\left(\frac{n}{2}\right)(T_2 - \gamma_2 - 1) \sim N(0, \tau_2^2),$$

where

$$\tau_2^2 = \frac{2n}{p}(a_2 - 2a_3 + a_4) + a_2^2.$$

The proof is given in Section 6.

COROLLARY 4.1. *Under the hypothesis* $H : \Sigma = I$, *and the assumptions* (A) *and* (B), *the asymptotic null distribution of* T_2 *is given by*

$$\left(\frac{n}{2}\right)T_2 \to N(0, 1).$$

Remark 4.1. This is a one-sided test for testing $H : \gamma_2 + 1 = 0$ vs $A : \gamma_2 + 1 > 0$.

Remark 4.2. For testing the hypothesis that $\Sigma = \Sigma_0$, we consider the observations $\boldsymbol{y}_i = \Sigma_0^{1/2}\boldsymbol{x}_i$, and test the hypothesis that the covariance matrix of $\boldsymbol{y}_i$ is an identity matrix. Here $\Sigma = (\Sigma^{1/2})(\Sigma^{1/2})$.

5. A test for the covariance to be diagonal

Let $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N$ be iid $N_p(\boldsymbol{\mu}, \Sigma)$. We wish to test the hypothesis $H : \sigma_{ij} = 0$, $i \neq j$, $i, j = 1, \ldots, p$, against the alternative $A : \sigma_{ij} \neq 0$, for at least one pair (i, j), $i \neq j$, where $\Sigma = (\sigma_{ij})$. Without loss of generality, we may consider the hypothesis $H : \rho_{ij} = 0$ vs $A : \rho_{ij} \neq 0$ for at least one pair (i, j), $i \neq j$,

$i, j = 1, \ldots, p$ where $\rho_{ij} = \sigma_{ij}/(\sigma_{ii}\sigma_{jj})^{1/2}$, $i \neq j$. Let r_{ij} be the sample correlation coefficient defined by

$$r_{ij} = \frac{s_{ij}}{\sqrt{s_{ii}s_{jj}}} = \frac{v_{ij}}{\sqrt{v_{ii}v_{jj}}}, \qquad i \neq j, \quad i, j = 1, \ldots, p, \tag{5.1}$$

where $S = (s_{ij})$ and $nS = V = (v_{ij})$. Define

$$q = \frac{1}{2}p(p-1), \qquad \boldsymbol{r} = (r_{12}, r_{13}, \ldots, r_{p-1,p})', \tag{5.2}$$

and

$$\rho = (\rho_{12}, \rho_{13}, \ldots, \rho_{p-1,p})'.$$

Then it follows from the results of Hsu (1949), that as $n \to \infty$,

$$\sqrt{n}(\boldsymbol{r} - \boldsymbol{\rho}) \longrightarrow N_q(0, \Omega),$$

where the diagonal elements of Ω are given by $(1-\rho_{12}^2)^2, (1-\rho_{13}^2)^2, \ldots, (1-\rho_{p-1,p}^2)^2$. The off diagonal elements are given by

$$\begin{aligned} \operatorname{cov}(r_{ij}, r_{kl}) = {} & \rho_{ij}\rho_{kl} + \rho_{kl}\rho_{il} - \rho_{lj}(\rho_{ij}\rho_{kj} + \rho_{il}\rho_{kl}) \\ & - \rho_{ki}(\rho_{ij}\rho_{il} + \rho_{kj}\rho_{kl}) + \frac{1}{2}\rho_{ki}\rho_{lj}(\rho_{ij}^2 + \rho_{il}^2 + \rho_{kj}^2 + \rho_{kl}^2). \end{aligned} \tag{5.3}$$

It may be noted that when $\rho_{ij} = 0$, $\Omega = I$, and thus $\sqrt{n}\boldsymbol{\gamma} \sim N_q(\mathbf{0}, I)$ as $n \to \infty$. Since

$$\sum_{i<j}^{p} \rho_{ij}^2 = 0$$

if and only if $\rho_{ij} = 0$, a test for the hypothesis $H : \rho_{ij} = 0$ against the alternative $A : \rho_{ij} \neq 0$ for at least one pair of (i, j), $i \neq j$, can be based on the test statistic

$$T_3^* = \frac{n\sum_{i<j}^{p} r_{ij}^2 - q}{\sqrt{2q}}, \qquad q = \frac{1}{2}p(p-1). \tag{5.4}$$

Since, under the hypothesis $H : \rho_{ij} = 0$, nr_{ij}^2 are asymptotically independently distributed with mean 1 and variance 2, it follows from the central limit theorem that as $p \to \infty$,

$$\sqrt{q}\left(\frac{n\sum_{i<j}^{p} r_{ij}^2}{q} - 1\right) \to N(0, 2).$$

Thus, asymptotically

$$\frac{n\sum_{i<j}^{p} r_{ij}^2 - q}{\sqrt{2q}} \to N(0, 1).$$

Thus, we get the following theorem

THEOREM 5.1. *Under the hypothesis $H : \rho_{ij} = 0$, asymptotically as n and p go to infinity*

$$T_3^* \sim N(0,1).$$

Under the alternative $A : \rho_{ij} = n^{-1/2}c_{ij}$, $-n^{1/2} < c_{ij} < n^{1/2}$, $\Omega = I + O(n^{-1/2})$. Hence, we get the following corollary.

COROLLARY 5.1. *Under the alternative $\rho_{ij} = n^{-1/2}c_{ij}$, asymptotically*

$$\left[\frac{2q}{2(q+2\boldsymbol{c}'\boldsymbol{c})}\right]^{1/2} T_3^* \sim N\left(\frac{\boldsymbol{c}'\boldsymbol{c}}{\sqrt{2(q+2\boldsymbol{c}'\boldsymbol{c})}}, 1\right).$$

The asymptotic normality of T_3^* given in Theorem 5.1 is rather slow. Thus, as an alternative, Chen and Mudholkar (1990) considered the Fisher's z-transformation z_{ij}, defined by

$$z_{ij} = \frac{1}{2}\log_e \frac{1+r_{ij}}{1-r_{ij}} \tag{5.5}$$

and proposed a test based on the test statistic

$$T_3^{**} = (n-2)\sum_{i<j}^{p} z_{ij}^2. \tag{5.6}$$

They approximated its distribution by a $X_v^2 + b$ where v, a and b are obtained by equating the first three moments of T_3^{**} with that of $aX_v^2 + b$ where X_v^2 is a chi-square random variable with v degrees of freedom. However, neither of the two tests have been desgined for large p. For example, the values of a, v, and b depends only on n. Thus, we consider another parametric measure of the deviation of the hypothesis from the alternative and propose a test based on its consistent estimate. Let

$$a_{20} = \frac{1}{p}\sum_{i=1}^{p}\sigma_{ii}^2, \qquad \hat{a}_{20} = \frac{n}{p(n+2)}\sum_{i=1}^{p} s_{ii}^2 \tag{5.7}$$

$$a_{40} = \frac{1}{p}\sum_{i=1}^{p}\sigma_{ii}^4, \qquad \hat{a}_{40} = \frac{1}{p}\sum_{i=1}^{p} s_{ii}^4. \tag{5.8}$$

Then

$$\begin{aligned} a_2 &= \frac{1}{p}\operatorname{tr}\Sigma^2 \\ &= \frac{1}{p}\sum_{i=1}^{p}\sigma_{ii}^2 + \frac{1}{p}\sum_{i\neq j}^{p}\sigma_{ij}^2 \\ &= a_{20} + \frac{1}{p}\sum_{i\neq j}^{p}\sigma_{ij}^2. \end{aligned}$$

We consider the parametric function

$$\gamma_3 = (a_2/a_{20}). \tag{5.9}$$

Clearly $\gamma_3 = 1$ if and only if $\sigma_{ij} = 0$. And if $\sigma_{ij} \neq 0$, $\gamma_3 \geq 1$. Thus, we can base our test on a consistent estimator of γ_3 given by

$$\hat{\gamma}_3 = \left(\frac{\hat{a}_2}{\hat{a}_{20}}\right) \tag{5.10}$$

where $\hat{a}_2$ is defined in (2.4), and $\hat{a}_{20}$ is given in (5.7). We will be testing the hypothesis $H : \gamma_3 - 1 = 0$ against the alternative $A : \gamma_3 - 1 \geq 0$. Thus, it will also be a one-sided test.

We note that

$$\begin{aligned} \hat{\gamma}_3 &= \frac{n}{(n-1)} \frac{\left[\operatorname{tr} S^2 - \frac{1}{n}(\operatorname{tr} S)^2\right]}{\sum_{i=1}^p s_{ii}^2} \\ &= \frac{n}{(n-1)} \frac{\left[\sum_{i\neq j} s_{ij}^2 + \sum_{i=1}^p s_{ii}^2 - \frac{1}{n}\sum_{i=1}^p s_{ii}^2 - \frac{1}{n}\sum_{i\neq j} s_{ii}s_{jj}\right]}{\sum_{i=1}^p s_{ii}^2} \\ &= \frac{n}{(n-1)} \left[\frac{\sum_{i\neq j}\left(s_{ij}^2 - \frac{1}{n}s_{ii}s_{jj}\right)}{\sum_{i=1}^p s_{ii}^2} + \frac{n-1}{n}\right] \\ &= \frac{n}{(n-1)} \frac{\sum_{i\neq j}\left(s_{ij}^2 - \frac{1}{n}s_{ii}s_{jj}\right)}{\sum_{i=1}^p s_{ii}^2} + 1, \end{aligned} \tag{5.11}$$

which is a statistic based on the sample covariances and not on sample correlations. This appears to be a reasonable procedure as when $n < p$, the sample covariance matrix S is singular, and in this case the diagonal elements s_{ii} may be very small and may lead to larger values of r_{ij}. From the asymptotic theory given in Theorem 6.2, it follows that

$$\hat{\gamma}_3 - 1 \to N\left[\gamma_3 - 1, (a_{20}^2)^{-1}\left(\frac{4}{n^2}\right)(a^2 - p^{-1}a_4)\right].$$

Thus, we propose a test based on the test statistic

$$T_3 = \left(\frac{n}{2}\right) \frac{(\hat{\gamma}_3 - 1)}{\left[1 - \left(\frac{1}{p}\right)\left(\frac{\hat{a}_{40}}{\hat{a}_{20}^2}\right)\right]^{1/2}}. \tag{5.12}$$

In the next theorem, we give an asymptotic distribution of T_3.

Theorem 5.2. *Under the assumptions* (A) *and* (B), *the asymptotic distribution of T_3 is given by*

$$T_3 \sim N[\delta, \tau_3^2],$$

where

$$\delta = \frac{n}{2}(\gamma_3 - 1)\left[1 - \left(\frac{1}{p}\right)\left(\frac{a_{40}}{a_{20}^2}\right)\right]^{-1/2}$$

and

$$\tau_3^2 = (a_2^2 - p^{-1}a_4)/(a_{20}^2 - p^{-1}a_{40}).$$

Since under the hypothesis H, $\delta = 0$, *and* $\tau_3^2 = 1$, *we get the following corollary.*

COROLLARY 5.2. *Under the hypothesis* $H : \sigma_{ij} = 0$, *and under the assumptions* (A) *and* (B), T_3 *is asymptotically distributed as* $N(0, 1)$.

6. Proofs

In this section, we prove lemmas and theorems stated in Sections 2 to 4. But before we begin these proofs, we state some results on the moments of a chi-square random variable with n degrees of freedom, and other preliminary results.

6.1. Preliminary results

We begin with the following lemma.

LEMMA 6.1. *Let* v *be a chi-square random variable with* n *degree of freedom. Then*

$$\begin{aligned}
&E(v^r) = n(n+2)\cdots(n+2r-2), \qquad r = 1, 2, \ldots \\
&\mathrm{Var}(v) = 2n, \qquad \mathrm{Var}(v^2) = 8n(n+2)(n+3), \\
&E(v-n)^3 = 8n, \qquad E(v-n)^4 = 12n(n+4), \\
&E[v^2 - n(n+2)]^4 = 3n(n+2)[272n^4 + O(n^3)].
\end{aligned}$$

Next, we obtain expressions for $\operatorname{tr} S$, $\operatorname{tr} S^2$ and $(\operatorname{tr} S)^2$ in terms of chi-square random variables. Let $V = nS = YY' \sim W_p(\Sigma, n)$, where $Y = (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n)$ and $\boldsymbol{y}_i$ are iid $N_p(\boldsymbol{0}, \Sigma)$. Let Γ be an orthogonal matrix such that $\Gamma\Sigma\Gamma' = \wedge$, where $\wedge = \mathrm{diag}(\lambda_1, \ldots, \lambda_p)$ and λ_i are the eigenvalues of Σ. Then, if $U = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)$, where $\boldsymbol{u}_i$ are iid $N_p(\boldsymbol{0}, I)$, $Y = \Sigma^{1/2}U$, and $\Sigma^{1/2}\Sigma^{1/2} = \Sigma$,

$$\begin{aligned}
(\operatorname{tr} S) &= \frac{1}{n}\operatorname{tr} U'\Sigma U \\
&= \frac{1}{n}\operatorname{tr} W' \wedge W \\
&= \frac{1}{n}\sum_{i=1}^{p}\lambda_i \boldsymbol{w}_i'\boldsymbol{w}_i,
\end{aligned}$$

where $U'\Gamma' = W' = (\boldsymbol{w}_1, \ldots, \boldsymbol{w}_p)$, and $\boldsymbol{w}_i$ are iid $N_n(\mathbf{0}, I)$. Thus, if $v_{ii} = \boldsymbol{w}_i'\boldsymbol{w}_i$, v_{ii} are iid chi-square random variables with n degrees of freedom. Hence,

$$np\hat{a}_1 = n(\operatorname{tr} S) = \sum_{i=1}^{p} \lambda_i v_{ii}, \tag{6.1}$$

$$n^2(trS)^2 = \left[\sum_{i=1}^{p} \lambda_i^2 v_{ii}^2 + 2\sum_{i<j}^{p} \lambda_i\lambda_j v_{ii}v_{jj}\right], \tag{6.2}$$

and

$$\begin{aligned} n^2 \operatorname{tr} S^2 &= \operatorname{tr}(W' \wedge W)(W' \wedge W) \\ &= \operatorname{tr}\left[\left(\sum_{i=1}^{p} \lambda_i \boldsymbol{w}_i\boldsymbol{w}_i'\right)\left(\sum_{i=1}^{p} \lambda_i \boldsymbol{w}_i\boldsymbol{w}_i'\right)\right] \\ &= \operatorname{tr}\left[\sum_{i=1}^{p} \lambda_i^2 \boldsymbol{w}_i\boldsymbol{w}_i'\boldsymbol{w}_i\boldsymbol{w}_i' + 2\sum_{i<j}^{p} \lambda_i\lambda_j \boldsymbol{w}_i\boldsymbol{w}_i'\boldsymbol{w}_j\boldsymbol{w}_j'\right] \\ &= \left[\sum_{i=1}^{p} \lambda_i^2 (\boldsymbol{w}_i'\boldsymbol{w}_i)^2 + 2\sum_{i<j}^{p} \lambda_i\lambda_j (\boldsymbol{w}_i'\boldsymbol{w}_j)^2\right] \\ &= \sum_{i=1}^{p} \lambda_i^2 v_{ii}^2 + 2\sum_{i<j}^{p} \lambda_i\lambda_j v_{ij}^2, \end{aligned} \tag{6.3}$$

where

$$v_{ij} = \boldsymbol{w}_i'\boldsymbol{w}_j, \qquad i \neq j, \quad v_{ii} = \boldsymbol{w}_i'\boldsymbol{w}_i. \tag{6.4}$$

Thus, since $n^2/(n-1)(n+2) \simeq 1$,

$$\begin{aligned} \hat{a}_2 &\simeq \frac{1}{p}\left[\operatorname{tr} S^2 - \frac{1}{n}(\operatorname{tr} S)^2\right] \\ &= \frac{1}{n^2 p}\left[\sum_{i=1}^{p} \lambda_i^2 v_{ii}^2 + 2\sum_{i<j}^{p} \lambda_i\lambda_j v_{ij}^2 - \frac{1}{n}\left(\sum_{i=1}^{p} \lambda_i^2 v_{ii}^2 + 2\sum_{i<j}^{p} \lambda_i\lambda_j v_{ii}v_{jj}\right)\right] \\ &= \frac{1}{n^2 p}\left[\frac{n-1}{n}\sum_{i=1}^{p} \lambda_i^2 v_{ii}^2 + 2\sum_{i<j}^{b} \lambda_i\lambda_j \left(v_{ij}^2 - \frac{1}{n}v_{ii}v_{jj}\right)\right] \\ &= \left[\frac{n-1}{n^3 p}\sum_{i=1}^{p} \lambda_i^2 v_{ii}^2\right] + \left[\frac{2}{n^2 p}\sum_{i<j}^{p} \lambda_i\lambda_j z_{ij}\right] \\ &= q_1 + q_2, \end{aligned} \tag{6.5}$$

where

$$\begin{aligned} z_{ij} &= v_{ij}^2 - \frac{1}{n}v_{ii}v_{jj} \\ &= (\boldsymbol{w}_i'\boldsymbol{w}_j)^2 - \frac{1}{n}(\boldsymbol{w}_i'\boldsymbol{w}_i)(\boldsymbol{w}_j'\boldsymbol{w}_j). \end{aligned} \tag{6.6}$$

From now on, we shall consider only the asymptotic version of $\hat{a}_2$, without the sign '$\simeq$'

LEMMA 6.2. *For z_{ij} defined above, we have*

$$\begin{aligned} & E(z_{ij}) = 0, \qquad E(z_{ij}z_{ik}) = 0, \qquad \textit{for all distinct } i, j, k, \\ & \mathrm{Var}(z_{ij}) = 2(n+2)(n-1). \end{aligned} \tag{6.7}$$

PROOF. Since $\boldsymbol{w}_i$ are iid $N_n(\mathbf{0}, I_n)$, it follows that

$$\begin{aligned} E(z_{ij}) &= E\left[\boldsymbol{w}_i'\boldsymbol{w}_j\boldsymbol{w}_j'\boldsymbol{w}_i - \frac{1}{n}(\boldsymbol{w}_i'\boldsymbol{w}_i)(\boldsymbol{w}_j'\boldsymbol{w}_j)\right] \\ &= E(\boldsymbol{w}_i'\boldsymbol{w}_i) - \frac{1}{n} \times n \times n \\ &= 0, \end{aligned}$$

as $E(\boldsymbol{w}_j\boldsymbol{w}_j') = I_n$ and $E(\boldsymbol{w}_i'\boldsymbol{w}_i) = n$.

Similarly, for all distinct i, j, k,

$$\begin{aligned} E[z_{ij}z_{ik}] &= E\left[(\boldsymbol{w}_j{}'\boldsymbol{w}_i)^2 - \frac{1}{n}(\boldsymbol{w}_i{}'\boldsymbol{w}_i)(\boldsymbol{w}_j{}'\boldsymbol{w}_j)\right]\left[(\boldsymbol{w}_k{}'\boldsymbol{w}_i)^2 - \frac{1}{n}(\boldsymbol{w}_i{}'\boldsymbol{w}_i)(\boldsymbol{w}_k{}'\boldsymbol{w}_k)\right] \\ &= E\Bigg[(\boldsymbol{w}_j{}'\boldsymbol{w}_i)^2(\boldsymbol{w}_k{}'\boldsymbol{w}_i)^2 - \frac{1}{n}(\boldsymbol{w}_k{}'\boldsymbol{w}_i)^2(\boldsymbol{w}_i{}'\boldsymbol{w}_i)(\boldsymbol{w}_j{}'\boldsymbol{w}_j) \\ &\qquad - \frac{1}{n}(\boldsymbol{w}_j{}'\boldsymbol{w}_i)^2(\boldsymbol{w}_i{}'\boldsymbol{w}_i)(\boldsymbol{w}_k{}'\boldsymbol{w}_k) + \frac{1}{n^2}(\boldsymbol{w}_i{}'\boldsymbol{w}_i)^2(\boldsymbol{w}_j{}'\boldsymbol{w}_j)(\boldsymbol{w}_k{}'\boldsymbol{w}_k)\Bigg]. \end{aligned}$$

Thus,

$$\begin{aligned} E[(\boldsymbol{w}_j{}'\boldsymbol{w}_i)^2(\boldsymbol{w}_k{}'\boldsymbol{w}_i)^2] &= E[(\boldsymbol{w}_j{}'\boldsymbol{w}_i\boldsymbol{w}_i{}'\boldsymbol{w}_j)(\boldsymbol{w}_i{}'\boldsymbol{w}_k\boldsymbol{w}_k{}'\boldsymbol{w}_i)] \\ &= E[\{\boldsymbol{w}_i{}'(\boldsymbol{w}_j\boldsymbol{w}_j{}')\boldsymbol{w}_i\}\{\boldsymbol{w}_i{}'(\boldsymbol{w}_k\boldsymbol{w}_k{}')\boldsymbol{w}_i\}] \\ &= E[(\boldsymbol{w}_i{}'\boldsymbol{w}_i)(\boldsymbol{w}_i{}'\boldsymbol{w}_i)] \\ &= n(n+2), \\ E[(\boldsymbol{w}_i{}'\boldsymbol{w}_i)(\boldsymbol{w}_j{}'\boldsymbol{w}_j)(\boldsymbol{w}_i{}'\boldsymbol{w}_k\boldsymbol{w}_k{}'\boldsymbol{w}_i)] &= E[(\boldsymbol{w}_i{}'\boldsymbol{w}_i)(\boldsymbol{w}_j{}'\boldsymbol{w}_j)(\boldsymbol{w}_i{}'\boldsymbol{w}_i)] \\ &= n^2(n+2), \\ E[(\boldsymbol{w}_i{}'\boldsymbol{w}_i)(\boldsymbol{w}_k{}'\boldsymbol{w}_k)(\boldsymbol{w}_i{}'\boldsymbol{w}_j\boldsymbol{w}_j{}'\boldsymbol{w}_i)] &= nE[(\boldsymbol{w}_i{}'\boldsymbol{w}_i)^2(\boldsymbol{w}_k{}'\boldsymbol{w}_k)] \\ &= n^2(n+2), \end{aligned}$$

and,

$$E[(\boldsymbol{w}_i{}'\boldsymbol{w}_i)^2(\boldsymbol{w}_j{}'\boldsymbol{w}_j)(\boldsymbol{w}_k{}'\boldsymbol{w}_k)] = n^3(n+2).$$

Hence, for all distinct i, j, k

$$E[z_{ij}z_{ik}] = 0.$$

To calculate the variance of z_{ij}, we note that

$$z_{ij}^2 = v_{ij}^4 - \frac{2}{n}v_{ij}^2v_{ii}v_{jj} + \frac{1}{n^2}v_{ii}^2v_{jj}^2, \qquad i \neq j,$$

and

$$\begin{aligned} E(v_{ij}^4) &= E(\boldsymbol{w}_i{}'\boldsymbol{w}_j\boldsymbol{w}_j{}'\boldsymbol{w}_i)^2 \\ &= E(\boldsymbol{w}_i{}'A_j\boldsymbol{w}_i)^2, \end{aligned}$$

where $A_j = \boldsymbol{w}_j\boldsymbol{w}_j{}'$. Let G be an orthogonal matrix such that $GA_jG' = \text{diag}(\boldsymbol{w}_j{}'\boldsymbol{w}_j, 0\ldots,0)$. Since given $A_j, \boldsymbol{l}_i = G\boldsymbol{w}_i \sim N_n(\mathbf{0}, I)$, it follows that $\boldsymbol{l}_i$ is independently distributed of A_j, and

$$\begin{aligned} E(v_{ij}^4) &= E[z_{1i}^4(\boldsymbol{w}_j{}'\boldsymbol{w}_j)^2] \\ &= 3n(n+2), \end{aligned}$$

where $\boldsymbol{l}_i = (l_{1i}, \ldots, l_{ni})' = G\boldsymbol{w}_i \sim N_n(\mathbf{0}, I)$. Thus, $l_{1i} \sim N(0,1)$ and is independently distributed of $\boldsymbol{w}_j{}'\boldsymbol{w}_j$. Similarly,

$$\begin{aligned} E(v_{ij}^2 v_{ii} v_{jj}) &= E(\boldsymbol{w}_i{}'\boldsymbol{w}_j\boldsymbol{w}_j{}'\boldsymbol{w}_i)(\boldsymbol{w}_i{}'\boldsymbol{w}_i)(\boldsymbol{w}_j{}'\boldsymbol{w}_j) \\ &= E(\boldsymbol{w}_i{}'A_j\boldsymbol{w}_i)(\boldsymbol{w}_i{}'\boldsymbol{w}_i)(trA_j), \qquad A_j = \boldsymbol{w}_j\boldsymbol{w}_j{}' \\ &= E\left[l_{1i}^2\boldsymbol{w}_j{}'\boldsymbol{w}_j\left(\sum_{k=1}^n l_{ki}^2\right)(\boldsymbol{w}_j{}'\boldsymbol{w}_j)\right] \\ &= E(\boldsymbol{w}_j{}'\boldsymbol{w}_j)^2 E\left(l_{1i}^2\sum_{k=1}^n l_{ki}^2\right) \\ &= n(n+2)(n+2) \\ &= n(n+2)^2. \end{aligned}$$

Finally,

$$E(v_{ii}^2 v_{jj}^2) = n^2(n+2)^2.$$

Hence,

$$\begin{aligned} \text{Var}(z_{ij}) &= 3n(n+2) - 2(n+2)^2 + (n+2)^2 \\ &= 3n(n+2) - (n+2)^2 \\ &= 2(n+2)(n-1). \end{aligned}$$

LEMMA 6.3. *Let v_{ii} be iid as χ_n^2,*

$$q_1 = \frac{n-1}{n^3p}\sum_{i=1}^p \lambda_i^2 v_{ii}^2, \qquad \textit{and} \qquad q_2 = \frac{2}{n^2p}\sum_{i<j}^p \lambda_i\lambda_j z_{ij}. \tag{6.8}$$

Then,

$$\text{Cov}(q_1, q_2) = 0 \tag{6.9}$$

$$\text{Var}(q_1) = [8(n+2)(n+3)(n-1)^2/pn^5]a_4 \tag{6.10}$$

$$\begin{aligned} \text{Var}(q_2) &= [8(n+2)(n-1)/n^4]\left[\sum_{i<j}^p \lambda_i^2\lambda_j^2/p^2\right] \\ &\simeq (4/n^2)[a_2^2 - (a_4/p)]. \end{aligned} \tag{6.11}$$

PROOF. Since $E(q_2) = 0$, we get

$$\begin{aligned}
\left(\frac{n^5p^2}{2(n-1)}\right)\mathrm{Cov}(q_1, q_2) &= \left(\frac{n^5p^2}{2(n-1)}\right)E(q_1q_2) \\
&= E\left[\sum_{i=1}^{p}\lambda_i^2 v_{ii}^2\right]\left[\sum_{j<k}^{p}\lambda_j\lambda_k z_{jk}\right] \\
&= \lambda_1^2 E\left((v_{11}^2)\left[\sum_{j<k}^{p}\lambda_j\lambda_k z_{jk}\right]\right) \\
&\quad + \lambda_2^2 E\left((v_{22}^2)\left[\sum_{j<k}^{p}\lambda_j\lambda_k z_{jk}\right]\right) \\
&\quad + \cdots + \lambda_p^2 E\left((v_{pp}^2)\left[\sum_{j<k}^{p}\lambda_j\lambda_k z_{jk}\right]\right).
\end{aligned}$$

We note that

$$\begin{aligned}
&E[v_{ii}^2 z_{jk}, j \neq k] \\
&\quad = E\left[(\boldsymbol{w}_i{}'\boldsymbol{w}_i)^2\left\{(\boldsymbol{w}_j{}'\boldsymbol{w}_k)^2 - \frac{1}{n}(\boldsymbol{w}_j{}'\boldsymbol{w}_j)(\boldsymbol{w}_k{}'\boldsymbol{w}_k)\right\}\right] \\
&\quad = E\left[(\boldsymbol{w}_i{}'\boldsymbol{w}_i)^2(\boldsymbol{w}_j{}'\boldsymbol{w}_k\boldsymbol{w}_k{}'\boldsymbol{w}_j) - \frac{1}{n}(\boldsymbol{w}_i{}'\boldsymbol{w}_i)^2(\boldsymbol{w}_j{}'\boldsymbol{w}_j)(\boldsymbol{w}_k{}'\boldsymbol{w}_k)\right] \\
&\quad = E\left[(\boldsymbol{w}_i{}'\boldsymbol{w}_i)^3 - \frac{1}{n}(\boldsymbol{w}_i{}'\boldsymbol{w}_i)^3(\boldsymbol{w}_k{}'\boldsymbol{w}_k)\right] = 0, \qquad i = j \neq k \\
&\quad = E\left[(\boldsymbol{w}_i{}'\boldsymbol{w}_i)^2(\boldsymbol{w}_j{}'\boldsymbol{w}_j) - \frac{1}{n}(\boldsymbol{w}_i{}'\boldsymbol{w}_i)^2(\boldsymbol{w}_j{}'\boldsymbol{w}_j)(\boldsymbol{w}_k{}'\boldsymbol{w}_k)\right] = 0, \qquad i \neq j \neq k \\
&\quad = E\left[(\boldsymbol{w}_i{}'\boldsymbol{w}_i)^2(\boldsymbol{w}_i{}'\boldsymbol{w}_j\boldsymbol{w}_j{}'\boldsymbol{w}_i) - \frac{1}{n}(\boldsymbol{w}_i{}'\boldsymbol{w}_i)^3(\boldsymbol{w}_j{}'\boldsymbol{w}_j)\right] = 0, \qquad j \neq k = i.
\end{aligned}$$

The variances of q_1 and q_2 can easily be obtained.

Note that,

$$\begin{aligned}
2p^{-2}\sum_{i<j}\lambda_i^2\lambda_j^2 &= p^{-2}\left[\left(\sum_{i=1}^{p}\lambda_i^2\right)^2 - \sum_{i=1}^{p}\lambda_i^4\right] \\
&= \left(\frac{\sum\lambda_i^2}{p}\right)^2 - \frac{1}{p}\left(\sum_{i=1}^{4}\frac{\lambda_i^4}{p}\right) \\
&= a_2^2 - (a_4/p),
\end{aligned} \tag{6.12}$$

which is bonded under the assumption (A). Hence, we get the following corollary.

COROLLARY 6.1. *As* $n \to \infty$, $q_2 \to 0$ *in probability under the assumption* (A). *Thus, also under the assumptions* (A) *and* (B), $q_2 \to 0$ *in probability as* $p \to \infty$.

COROLLARY 6.2. *Let*

$$\hat{a}_1 = \frac{1}{pn}\left(\sum_{i=1}^{p} \lambda_i v_{ii}\right) = \left(\frac{\operatorname{tr} S}{p}\right).$$

Then

$$\operatorname{Cov}(q_2, \hat{a}_1) = 0.$$

LEMMA 6.4. *Let* $nS \sim W_p(\Sigma, n)$, *and*

$$\hat{a}_2 = \frac{1}{p}\left[\operatorname{tr} S^2 - \frac{1}{n}(\operatorname{tr} S)^2\right] = q_1 + q_2.$$

Then, from (6.10), *and* (6.11) *and the fact that* $\operatorname{Cov}(q_1, q_2) = 0$,

$$\begin{aligned}
\operatorname{Var}(\hat{a}_2) &= \operatorname{Var}(q_1) + \operatorname{Var}(q_2) \\
&= \frac{8(n+2)(n+3)(n-1)^2}{pn^5} a_4 + \frac{4(n+2)(n-1)}{n^4}[a_2^2 - p^{-1}a_4] \\
&\simeq \frac{8a_4}{pn} + \frac{4}{n^2}[a_2^2 - p^{-1}a_4] \\
&\simeq \frac{4}{n^2}\left[a_2^2 + \frac{2n}{p}a_4\right].
\end{aligned}$$

LEMMA 6.5. *Let* v_{ii} *be iid as* χ_n^2, $q_1 = (n-1)(pn^3)^{-1}\sum_{i=1}^{p} \lambda_i^2 v_{ii}^2$, *and* $\hat{a}_1 = (pn)^{-1}\sum_{i=1}^{p} \lambda_i v_{ii}$. *Then*

$$\operatorname{Cov}(q_1, \hat{a}_1) = 4(n-1)(n+2)a_3/pn^3 \simeq 4a_3/(pn) \quad \text{as} \quad n \to \infty.$$

PROOF. We have

$$\begin{aligned}
&\frac{p^2 n^4}{(n-1)} \operatorname{Cov}(q_1, \hat{a}_1) \\
&\quad = \left[E\left(\sum_{i=1}^{p} \lambda_i^2 v_{ii}^2\right)\left(\sum_{i=1}^{p} \lambda_i v_{ii}\right) - n^2(n+2)p^2 a_2 a_1\right] \\
&\quad = \left[E\left(\sum_{i=1}^{p} \lambda_i^3 v_{ii}^3 + \sum_{i \neq j}^{p} \lambda_i^2 v_{ii}^2 \lambda_j v_{jj}\right) - n^2(n+2)p^2 a_2 a_1\right] \\
&\quad = \left[n(n+2)(n+4)pa_3 + n^2(n+2)\sum_{i \neq j} \lambda_i^2 \lambda_j - n^2(n+2)p^2 a_2 a_1\right] \\
&\quad = [n(n+2)(n+4)pa_3 - n^2(n+2)pa_3] \\
&\quad = [4n(n+2)pa_3].
\end{aligned}$$

Next, we obtain the variance of $\hat{\gamma}_2$. We have

$$\begin{aligned}\hat{\gamma}_2 &\simeq \frac{1}{p}\left[\operatorname{tr} S^2 - \frac{1}{n}(\operatorname{tr} S)^2 - 2\operatorname{tr} S\right] \\ &= \frac{1}{p}\frac{n-1}{n^3}\sum \lambda_i^2 v_{ii}^2 + \frac{2}{pn^2}\sum_{i<j} W_{ij} - \frac{2}{pn}\sum \lambda_i v_{ii} \\ &\simeq \frac{1}{pn^2}\sum \lambda_i^2 v_{ii}^2 - \frac{2}{pn}\sum \lambda_i v_{ii} + \frac{2}{pn^2}\sum_{i<j} W_{ij} \\ &= \frac{1}{pn^2}\sum[\lambda_i^2 v_{ii}^2 - 2nv_{ii}\lambda_i] + \frac{2}{pn^2}\sum_{i<j} W_{ij},\end{aligned} \tag{6.13}$$

where

$$W_{ij} = \lambda_i \lambda_j z_{ij}. \tag{6.14}$$

Now

$$\begin{aligned}&\operatorname{Var}[\lambda_i^2 v_{ii}^2 - 2nv_{ii}\lambda_i] \\ &= \lambda_i^4 \operatorname{Var}(v_{ii}^2) - 4n\lambda_i^3 \operatorname{Cov}(v_{ii}^2, v_{ii}) + 4n^2\lambda_i^2 \operatorname{Var}(v_{ii}) \\ &= 8n(n+2)(n+3)\lambda_i^4 + 8n^3\lambda_i^2 - 4n\lambda_i^3 E[(v_{ii}-n)v_{ii}^2] \\ &= 8n(n+2)(n+3)\lambda_i^4 + 8n^3\lambda_i^2 - 4n\lambda_i^3 E[v_{ii}^3 - nv_{ii}^2] \\ &= 8n(n+2)(n+3)\lambda_i^4 + 8n^3\lambda_i^2 \\ &\quad - 4n\lambda_i^3[n(n+2)(n+4) - n^2(n+2)] \\ &= 8n(n+2)(n+3)\lambda_i^4 + 8n^3\lambda_i^2 - 16n^2(n+2)\lambda_i^3.\end{aligned} \tag{6.15}$$

Hence,

$$\begin{aligned}&\operatorname{Var}\left[\frac{1}{p}\left(\operatorname{tr} S^2 - \frac{1}{n}(\operatorname{tr} S)^2 - 2\operatorname{tr}(S)\right)\right] \\ &\simeq \operatorname{Var}\left[\frac{1}{pn^2}\left(\sum(\lambda_i^2 v_{ii}^2 - 2nv_{ii}\lambda_i)\right)\right] + \operatorname{Var}\left[\frac{2}{pn^2}\sum_{i<j}^{p} W_{ij}\right] \\ &\quad + 2\operatorname{Cov}\left[\frac{1}{pn^2}\left(\sum_{i=1}^{p}(\lambda_i^2 v_{ii}^2 - 2nv_{ii}\lambda_i)\right), \frac{2}{pn^2}\sum_{i<j}^{p} W_{ij}\right] \\ &= \frac{8n(n+2)(n+3)\sum\lambda_i^4 + 8n^3\sum\lambda_i^2 - 16n^2(n+2)\sum\lambda_i^3}{p^2n^4} \\ &\quad + \frac{4(n+2)(n-1)}{n^4p^2}\sum_{i\neq j}\lambda_i^2\lambda_j^2 + 0 \\ &= \frac{8n(n+2)(n+3)a_4 + 8n^3a_2 - 16n^2(n+2)a_3}{n^4p} \\ &\quad + \frac{4(n+2)(n-1)\left[a_2^2 - \frac{1}{p}a_4\right]}{n^4} \\ &\simeq \frac{8a_4 + 8a_2 - 16a_3}{np} + \frac{4}{n^2}\left[a_2^2 - \frac{1}{p}a_4\right].\end{aligned}$$

Hence, we get the following lemma

LEMMA 6.6. *For $\hat{\gamma}_2$ defined above,*

$$\mathrm{Var}(\hat{\gamma}_2) \simeq \frac{8a_4 - 16a_3 + 8a_2}{np} + \frac{4}{n^2}[a_2^2 - p^{-1}a_4].$$

COROLLARY 6.3. *When $\lambda_i = 1$,*

$$\mathrm{Var}(\hat{\gamma}_2) \simeq \frac{4}{n^2}, \qquad \textit{for large } n \textit{ and } p.$$

Let

$$u_{1i} = \frac{\lambda_i(v_{ii} - n)}{\sqrt{n}}, \qquad u_{2i} = \frac{\lambda_i^2[v_{ii}^2 - n(n+2)]}{\sqrt{n(n+2)(n+3)}}, \tag{6.16}$$

and

$$\varepsilon_n = [(n+2)/(n+3)]^{1/2}. \tag{6.17}$$

Then

$$\begin{aligned} &E(u_{1i}) = 0, \qquad E(u_{2i}) = 0 \\ &\mathrm{Var}(u_{1i}) = 2\lambda_i^2, \qquad \mathrm{Var}(u_{2i}) = 8\lambda_i^4, \\ &\mathrm{Cov}(u_{1i}, u_{2i}) = 4\varepsilon_n\lambda_i^3. \end{aligned} \tag{6.18}$$

Thus $\boldsymbol{u}_i = \binom{u_{1i}}{u_{2i}}$ are independently distributed random vectors, $i = 1, \ldots, p$, with mean vectors as zero vectors and the 2×2 covariance matrices M_{in} given by

$$M_{in} = \begin{pmatrix} 2\lambda_i^2 & 4\varepsilon_n\lambda_i^3 \\ 4\varepsilon_n\lambda_i^3 & 8\lambda_i^4 \end{pmatrix}, \qquad i = 1, \ldots, p. \tag{6.19}$$

Now, as $p \to \infty$

$$\begin{aligned} M_n &\equiv \frac{1}{p}(M_{1n} + \ldots + M_{pn}) \\ &= \begin{pmatrix} 2a_2 & 4\varepsilon_n a_3 \\ 4\varepsilon_n a_3 & 8a_4 \end{pmatrix} \to M_n^0 \neq 0, \qquad \text{for any} \quad n, \end{aligned} \tag{6.20}$$

where

$$M_n^0 = \begin{pmatrix} 2a_0^0 & 4\varepsilon_n a_3^0 \\ 4\varepsilon_n a_3^0 & 8a_4^0 \end{pmatrix} \to \begin{pmatrix} 2a_0^0 & 4a_3^0 \\ 4a_3^0 & 8a_4^0 \end{pmatrix} \equiv M^0 \qquad \text{as} \quad n \to \infty. \tag{6.21}$$

Also, if F_i is the distribution function of $\boldsymbol{u}_i$, then

$$\begin{aligned}\frac{1}{p}\sum_{i=1}^{p}\int_{(\boldsymbol{u}'\boldsymbol{u})>p\varepsilon^2}\boldsymbol{u}'\boldsymbol{u}dF_i &\leq \frac{1}{p}\sum_{i=1}^{p}(p\varepsilon^2)^{-1}\int(\boldsymbol{u}'\boldsymbol{u})^2dF_i\\ &= \frac{1}{p^2\varepsilon^2}\sum_{i=1}^{p}E(u_{1i}^2+u_{2i}^2)^2\\ &\leq \frac{2}{p^2\varepsilon^2}\sum_{i=1}^{p}E(u_{1i}^4+u_{2i}^4),\end{aligned}$$

from c_r-inequality, see Rao (1973, p. 149). Now

$$\begin{aligned}\frac{1}{p^2}\sum_{i=1}^{p}E(u_{1i}^4) &= \frac{1}{p^2}\sum\lambda_i^4\frac{E(v_i-n)^4}{n^2}\\ &= \frac{12n(n+4)}{n^2}\frac{a_4}{p}\to 0 \quad \text{as} \quad p\to\infty.\end{aligned}$$

Similarly

$$\begin{aligned}\frac{1}{p^2}\sum_{i=1}^{p}E(u_{2i}^4) &= \frac{1}{p^2}\sum\lambda_i^8\frac{E(v_i^2-n(n+2))^4}{n^2(n+2)^2(n+3)^2}\\ &= O(p^{-1})\to 0 \quad \text{as} \quad p\to\infty.\end{aligned}$$

Then from the multivariate central limit theorem of Liapunov type given in Rao (1973, p. 147, Problem 4.7), it follows that as $p\to\infty$, and for any n,

$$\frac{1}{\sqrt{p}}\sum_{i=1}^{p}\boldsymbol{u}_i = \frac{1}{\sqrt{np}}\begin{pmatrix}\sum_{i=1}^{p}\lambda_i(v_{ii}-n)\\ \frac{\sum_{i=1}^{p}\lambda_i^2(v_{ii}^2-n(n+2))}{\sqrt{(n+2)(n+3)}}\end{pmatrix}\sim N_2(\mathbf{0},M_n^0).$$

Thus, it follows that as $p\to\infty$ and then $n\to\infty$,

$$\frac{1}{p}\sum_{i=1}^{p}\boldsymbol{u}_i\to N_2(\mathbf{0},M^0).$$

On the otherhand, as $n\to\infty$, we get from the multivariate central limit theorem that

$$\boldsymbol{u}_i\to N_2(O,M_i), \qquad i=1,\ldots,p$$

for any p, where M_i is the limit of M_{in} given by

$$M_i=\begin{pmatrix}2\lambda_i^2 & 4\lambda_i^3\\ 4\lambda_i^3 & 8\lambda_i^4\end{pmatrix}. \tag{6.22}$$

Let

$$M=\frac{1}{p}(M_1+\ldots+M_p)=\begin{pmatrix}2a_2 & 4a_3\\ 4a_3 & 8a_4\end{pmatrix}, \tag{6.23}$$

which goes to M^0 as $p \to \infty$. Since $\boldsymbol{u}_i$ are asymptotically independently distributed, it follows from the argument given above that as $n \to \infty$, and then $p \to \infty$

$$\frac{1}{\sqrt{p}} \sum_{i=1}^{p} \boldsymbol{u}_i \to N_2(\mathbf{0}, M^0).$$

Without any loss of generality, we may replace M^0 by M. Noting that

$$\hat{a}_1 = (np)^{-1} \sum_{i=1}^{p} \lambda_i v_{ii},$$

and

$$\begin{aligned} q_1 &= \frac{n-1}{n^3 p} \sum_{i=1}^{p} \lambda_i^2 v_{ii}^2 \\ &\simeq \frac{1}{n^2 p} \sum_{i=1}^{p} \lambda_i^2 v_{ii}^2 \end{aligned}$$

we get the following theorem

Theorem 6.1. *As n and $p \to \infty$, in any manner,*

$$\begin{pmatrix} \hat{a}_1 \\ q_1 \end{pmatrix} \to N_2 \left[\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, (np)^{-1} M \right].$$

Corollary 6.4. *For any finite n, and as $p \to \infty$,*

$$\begin{pmatrix} \hat{a}_1 \\ q_1 \end{pmatrix} \to N_2 \left[\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, (np)^{-1} M_n \right].$$

Theorem 6.2. *Let z_{ij} be as defined in* (6.6), *and*

$$q_2 = \frac{2}{n^2 p} \sum_{i<j} \lambda_i \lambda_j z_{ij}.$$

Then under the assumptions (A) *and* (B), *and as n and $p \to \infty$, $q_2 \sim N[\mathbf{0}, 4n^{-2}(a_2^2 - p^{-1} a_4)]$.*

Proof. Note that

$$\begin{aligned} z_{ij} &= (\boldsymbol{w}_i{}' \boldsymbol{w}_j)^2 - \frac{1}{n} (\boldsymbol{w}_i{}' \boldsymbol{w}_i)(\boldsymbol{w}_j{}' \boldsymbol{w}_j) \\ &= v_{ij}^2 - \frac{1}{n} v_{ii} v_{jj}, \end{aligned}$$

and

$$n^{-1} z_{ij} = (n^{-1} v_{ij}^2 - 1) - n^{-2}(v_{ii} v_{jj} - n^2).$$

The second term goes to zero in probability and $\text{Cov}[n^{-1}(v_{ij}-n), n^{-1}(v_{ik}-n)] \to 0$ as $n \to \infty$ for all distinct i, j, k. Since $n^{-1/2}v_{ij} \sim N(0,1)$ as $n \to \infty$, it follows that $n^{-1}v_{ij}^2$ is a chi-square random variable with one degree of freedom and are asymptotically independently distributed for all distinct i and j. Now, when $p \to \infty$, we apply Liapunov type central limit theorem to obtain the asynptotic normality of q_2. Because of the normality assumption, the same result is obtained if we interchange the order of limit.

Next, we note that $\text{Cov}(\hat{a}_1, q_2) = 0$ and $\text{Cov}(q_1, q_2) = 0$. From the asymptotic normality of $(\hat{a}_1, q_1)$ and the fact that the covariance between $(\hat{a}_1, q_1)$ and q_2 is a zero vector, it follows that $(\hat{a}_1, q_1, q_2)$ are jointly asymptotically normally distributed as stated in the following theorem.

THEOREM 6.3. *Under the assumptions* (A) *and* (B), *asymptotically*

$$\begin{pmatrix} \hat{a}_1 \\ q_1 \\ q_2 \end{pmatrix} \sim N\left[\begin{pmatrix} a_1 \\ a_2 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{1}{np}M & 0 \\ 0 & 4n^{-2}(a_2^2 - p^{-1}a_4) \end{pmatrix}\right].$$

6.2. Proof of Lemma 2.1

From (6.5), and Lemma 6.2

$$\begin{aligned} E(\hat{a}_2) &= \frac{1}{p}E\left[\text{tr}\, S^2 - \frac{1}{n}(\text{tr}\, S)^2\right] \\ &= \frac{n-1}{n^3 p}E\left[\sum_{i=1}^p \lambda_i^2 v_i^2\right] + \frac{2}{n^2 p}E\left[\sum_{i<j}^p \lambda_i\lambda_j z_{ij}\right] \\ &= \frac{n-1}{n^3 p}n(n+2)\sum_{i=1}^p \lambda_i^2 + \frac{2}{n^2 p}\sum_{i<j}^p \lambda_i\lambda_j E(z_{ij}) \\ &= \frac{(n-1)(n+2)}{n^2}\left(\frac{\text{tr}\,\Sigma^2}{p}\right). \end{aligned}$$

Hence,

$$\frac{n^2}{(n-1)(n+2)}\frac{1}{p}\left[\text{tr}\, S^2 - \frac{1}{n}(\text{tr}\, S)^2\right] \tag{6.24}$$

is an unbiased estimator of $(\text{tr}\,\Sigma^2/p)$.

From Lemma 6.3, q_2 goes to zero in probability as $n \to \infty$. Thus it follows from Lemma 6.4 and equation (6.5) that (6.12) is under assumption (A), an unbiased and consistent estimator of a_2 as $n \to \infty$, and

$$\hat{a}_2 = \frac{1}{p}\left[\text{tr}\, S^2 - \frac{1}{n}(\text{tr}\, S)^2\right]$$

is a consistent estimator of a_2 as n goes to infinity and the assumption (A) holds.

6.3. Proof of Theorem 2.1 and Corollary 2.1

Since $\hat{a}_2 = q_1 + q_2$ and $\hat{a}_1 = [\sum_{i=1}^{p} \lambda_i v_i/pn] = (\operatorname{tr} S/p)$, it follows from Theorem 6.3 that

$$\begin{pmatrix} \hat{a}_1 \\ \hat{a}_2 \end{pmatrix} \sim N_2 \left[\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, (np)^{-1} \begin{pmatrix} 2a_2 & 4a_3 \\ 4a_3 & 8a_4 + 4\frac{p}{n}a_2^2 \end{pmatrix} \right].$$

To prove Corollary 2.1, we note that $c = p/n$,

$$g_1(\hat{a}_1, \hat{a}_2) \equiv \frac{1}{p}(\operatorname{tr} S) = \hat{a}_1,$$

and

$$g_2(\hat{a}_1, \hat{a}_2) \equiv \frac{1}{p}(\operatorname{tr} S^2) = \hat{a}_2 + c\hat{a}_1^2.$$

Hence, from Lemma 2.3, the covarinace of $(g_1, g_2)'$ is given by

$$\begin{aligned}
&(np)^{-1} \begin{pmatrix} \frac{\partial g_1}{\partial \hat{a}_1} & \frac{\partial g_1}{\partial \hat{a}_2} \\ \frac{\partial g_2}{\partial \hat{a}_1} & \frac{\partial g_2}{\partial \hat{a}_2} \end{pmatrix} \begin{pmatrix} 2a_2 & 4a_3 \\ 4a_3 & 8a_4 + 4ca_2^2 \end{pmatrix} \begin{pmatrix} \frac{\partial g_1}{\partial \hat{a}_1} & \frac{\partial g_1}{\partial \hat{a}_2} \\ \frac{\partial g_2}{\partial \hat{a}_1} & \frac{\partial g_2}{\partial \hat{a}_2} \end{pmatrix}' \\
&= (np)^{-1} \begin{pmatrix} 1 & 0 \\ 2a_1 c & 1 \end{pmatrix} \begin{pmatrix} 2a_2 & 4a_3 \\ 4a_3 & 8a_4 + 4ca_2^2 \end{pmatrix} \begin{pmatrix} 1 & 2a_1 c \\ 0 & 1 \end{pmatrix} \\
&= n^{-2}c^{-1} \begin{pmatrix} 2a_2 & 4(ca_1a_2 + a_3) \\ 4(ca_1a_2 + a_3) & 4(2c^2a_1^2a_2 + 4ca_1a_3 + ca_2^2 + 2a_4) \end{pmatrix}.
\end{aligned}$$

6.4. Proof of Theorem 3.1

We note that

$$T_1 = \frac{\hat{a}_2}{\hat{a}_1^2} - 1.$$

Hence,

$$\left(\frac{\partial T_1}{\partial \hat{a}_1}\right) = -\frac{2\hat{a}_2}{\hat{a}_1^3}, \quad \text{and} \quad \left(\frac{\partial T_1}{\partial \hat{a}_2}\right) = \frac{1}{\hat{a}_1^2}.$$

Thus, asymptotically,

$$T_1 \sim N(\gamma_1 - 1, \tau_1^2)$$

where

$$\begin{aligned}
\tau_1^2 &= \left(\frac{-2a_2}{a_1^3}, \frac{1}{a_1^2}\right) \begin{pmatrix} \frac{2a_2}{np} & \frac{4a_3}{np} \\ \frac{4a_3}{np} & \frac{8a_4}{np} + \frac{4a_2^2}{n^2} \end{pmatrix} \begin{pmatrix} -2a_2/a_1^3 \\ \frac{1}{a_1^2} \end{pmatrix} \\
&= \left[-\frac{4a_2^2}{npa_1^3} + \frac{4a_3}{npa_1^2}, \frac{-8a_2a_3}{npa_1^3} + \frac{1}{a_1^2}\left(\frac{8a_4}{np} + \frac{4a_2^2}{n^2}\right) \right] \begin{pmatrix} -2a_2/a_1^3 \\ \frac{1}{a_1^2} \end{pmatrix} \\
&= \frac{8a_2^3}{npa_1^6} - \frac{16a_2a_3}{npa_1^5} + \frac{1}{a_1^4}\left(\frac{8a_4}{np} + \frac{4a_2^2}{n^2}\right) \\
&= \frac{8a_2^3}{npa_1^6} - \frac{16a_2a_3}{npa_1^5} + \frac{8a_4}{npa_1^4} + \frac{4a_2^2}{n^2a_1^4}.
\end{aligned}$$

Thus, when $\lambda_i = \lambda$,

$$\tau_1^2 = 4/n^2.$$

6.5. Proof of Theorem 4.1

Since

$$\hat{\gamma}_2 = \hat{a}_2 - 2\hat{a}_1$$

it follows from Theorem 6.3 that asymptotically

$$T_2 \sim N(\gamma_2 + 1, \tau_2^2)$$

where

$$\tau_2^2 = \frac{8a_4 - 16a_3 + 8a_2}{np} + \frac{4a_2^2}{n^2}.$$

Thus, when $\lambda_i = 1$,

$$\tau_2^2 = \frac{4}{n^2}.$$

Acknowledgements

I wish to thank the two referees for their careful reading of the article and for their suggestions. This research was supported by the Natural Sciences and Engineering Research Council of Canada.

References

Bilodeau, M. and Brenner, D. (1999). *Theory of Multivariate Statistics*, Springer-Verlag, New York.

Carter, E. M. and Srivastava, M. S. (1977). Monotonicity of the power functions of the modified Likelihood ratio criteria for the homogeneity of variances and of the sphericity test, *J. Multiv. Anal.*, **7**, 229–233.

Chen, S. and Mudholkar, G. S. (1990). Null distribution of the sum of squared z-transformations in testing complete independence, *Ann. Inst. Statist. Math.*, **42**, 149–155.

Dudoit, S., Fridlyand, J. and Speed, T. P. (2002). Comparison of discrimination methods for the classfication of tremors using gene expression data, *J. Amer. Statist. Assoc.*, **97**, 77–87.

Hsu, P. L. (1949). The limiting distribution of functions of sample means and applications to testing hypothesis, *Proceedings of the First Berkeley Symposium of Mathematical Statistics and Probability* (ed. J. Neyman), Univ. of California Press, Berkeley and Los Angeles, 359–402.

John, S. (1971). Some optimal multivariate tests, *Biometrika*, **58**, 123–127.

Ledoit, O. and Wolf, M. (2002). Some hypothesis tests for the covariance matrix when the dimension is large copared to the sample size, *Ann. Statist.*, **30**, 1081–1102.

Nagao, H. (1967). Monotonicity of the modified likelihood ratio test for a covariance matrix, *J. Sci. Hiroshima Univ.*, **A131**, 147–150.

Nagao, H. (1973). On some test criteria for covariance matrix, *Ann. Math. Statist.*, **1**, 700–709.

Rao, C. R. (1973). *Linear Statistical Inference and Its Applications*, 2nd ed., Wiley, New York.

Srivastava, M. S. and Khatri, C. G. (1979). *An Introduction to Multivariate Statistics*, North-Holland, New York.

Sugiura, N. (1972). Locally best invariant test for spherecity and the limiting distributions, *Ann. Math. Statist.*, **43**, 1312–1316.