Some tests for the covariance matrix with fewer observations than the dimension under non-normality

Muni S. Srivastava, Tõnu Kollo, and Dietrich von Rosen

Research Report
Centre of Biostochastics

Swedish University of Agricultural Sciences

Report 2009:05
ISSN 1651-8543

Some tests for the covariance matrix with fewer observations than the dimension under non-normality

Muni S. Srivastava

University of Toronto, Canada

Tõnu Kollo

University of Tartu, Estonia

Dietrich von Rosen[1]

Centre of Biostochastics
Swedish University of Agricultural Sciences, Sweden

Abstract

This article analyzes whether the existing tests for the $p \times p$ covariance matrix Σ of the N independent identically distributed observation vectors with $N \leq p$ work under non-normality. We focus on three hypotheses testing problems: (1) testing for sphericity, that is, the covariance matrix Σ is proportional to an identity matrix I_p; (2) the covariance matrix Σ is an identity matrix I_p; and (3) the covariance matrix is a diagonal matrix. It is shown that the tests proposed by Srivastava (2005) for the above three problems are robust under the non-normality assumption made in this article irrespective of whether $N \leq p$ or $N \geq p$.

Keywords: Covariance matrix, Diagonality of covariance matrix, Hypotheses tests, Identity matrix, Test of sphericity

[1]E-mail address to the correspondence author: dietrich.von.rosen@et.slu.se

1 Introduction

Quantitative measurements of thousands of genes' expressions are obtained through DNA microarrays. Since these observations on the genes are on the same subject, they are not independently distributed. Thus, if there are measurements on p genes, it has a $p \times p$ covariance matrix Σ. The number of subjects on which these measurements are obtained, say N, are often very few, that is $N \ll p$. The analysis of such data sets requires new developments of multivariate theory, many of them have recently been obtained in the literature. The analysis is, however, simplified considerably if the $p \times p$ covariance matrix Σ satisfies either of the following three hypotheses:

(1) $H_1 : \Sigma = \lambda I_p, \ \lambda > 0,$

(2) $H_2 : \Sigma = I_p,$

(3) $H_3 : \Sigma = D = diag(d_1, ..., d_p),$

where D is a $p \times p$ diagonal matrix with diagonal elements $d_1, ..., d_p$. For example, if either the hypothesis (1) or (2) holds, then most of the univariate results can be used to analyze the data. If the hypothesis H_3 holds, then a standardized version of the univariate test statistics can be used. In microarray data analysis of genes, it is invariably assumed, implicitly or explicitly that the genes are independently distributed to carry out the analysis; that is, the analysis is carried out under the hypothesis H_3. The false discovery rate (FDR) of the Benjamini and Hochberg (1995) procedure can be controlled at the specified level only if the hypothesis H_3 is true, or if the covariance matrix Σ is of the intraclass correlation structure with positive correlation provided the data is normally distributed; but so far no satisfactory test is available for testing the intraclass correlation structure when N $\leq$ p. Since $N \ll p$, it is not known how to ascertain the multivariate normality of the data. Thus, it would be desirable to have tests for which the significance levels can be controlled with or without the assumption of normality of the data; that is, to have robust tests.

When p is finite and N is large it may not be important or necessary to obtain robust tests as the level of significance can be maintained at the specified level by using the bootstrap methods of Beran and Srivastava (1985) for the covariance matrix. For this reason, most studies considered selecting a test that has better power among the available tests. For example, Chan and Srivastava (1988) compared the power of the LRT with that of LBIT defined in Section 4 for testing sphericity. Similar comparison was carried out by Nagao and Srivastava (1992) for the multivariate t-distribution with k degrees of freedom and found that LBIT is better than LRT. Purkayastha and Srivastava (1995) compared the power of LRT with a test proposal by Rao (1948) and independently by Nagao (1973) for testing that $\Sigma = I$, for the elliptical distribution. A robust and improved estimator of the covariance matrix of the elliptical model has been given by Kubokawa and Srivastava (1999).

For $N \leq p$ and both N and p going to infinity, bootstrap theory is not yet available. Thus, it is desirable to obtain robust tests for this situation. Our objective in this paper is to show that the tests proposed by Srivastava (2005) are robust for the model described below.

We assume that the p dimensional observation vectors $\mathbf{x}_1$,..., $\mathbf{x}_N$ on N subjects are independently identically distributed (iid) with mean vector $\boldsymbol{\mu}$ and covariance matrix $\Sigma = CC'$, where C is a $p \times p$ non-singular matrix, that is Σ is a positive definite (pd) matrix.

We shall assume that the N iid observation vectors $\mathbf{x}_i$ of dimension p, can be written as

$$\mathbf{x}_i = \boldsymbol{\mu} + C\mathbf{z}_i, \tag{1.1}$$

$$E(\mathbf{z}_i) = \mathbf{0}, \quad Cov(\mathbf{x}_i) = CC' = \Sigma > \mathbf{0}, \quad i = 1, ..., N.$$

For testing the hypothesis H_3 of diagonality of the covariance matrix Σ, we shall, however, assume that under H_3, $C = diag(d_1^{1/2}, ..., d_p^{1/2}) = D^{1/2}$.

Instead of normality of $\mathbf{z}_i = (z_{i1}, ..., z_{ip}), i = 1, ..., N$, we shall assume that not only that $\mathbf{z}_i$ are iid, but that z_{ij} are iid for all i and j with

$$E(z_{ij}^r) = \gamma_r, \; r = 3, ..., 8, \text{ with } \gamma_4 = \gamma. \tag{1.2}$$

Under normality, $\gamma_3 = \gamma_5 = \gamma_7 = 0, \; \gamma = 3, \; \gamma_6 = 15$, and $\gamma_8 = 105$. Unbiased estimators of $\boldsymbol{\mu}$ and Σ are respectively given by

$$\bar{\mathbf{x}} = N^{-1} \sum_{i=1}^{N} \mathbf{x}_i, \quad S = \left[\sum_{i=1}^{N} (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})' \right] / (N-1). \tag{1.3}$$

When $N \leq p$, the sample covariance matrix S is singular and no likelihood ratio test (LRT) is available for any of the three hypotheses. Thus, we consider the following tests proposed by Srivastava (2005) for the hypotheses H_1, H_2, H_3. Let

$$\hat{\delta}_1 = \text{tr}S/p, \quad \hat{\delta}_2 = [\text{tr}S^2 - N^{-1}(\text{tr}S)^2]/p, \tag{1.4}$$

$$\hat{\delta}_{20} = \sum_{i=1}^{p} s_{ii}^2/p, \text{ and } \hat{\delta}_{40} = \sum_{i=1}^{p} s_{ii}^4/p, \quad S = (s_{ij}). \tag{1.5}$$

Then for testing the hypothesis H_1, known in the literature as the 'Sphericity' hypothesis, we consider the test statistic given by

$$T_1 = \left(\frac{\hat{\delta}_2}{\hat{\delta}_1^2}\right) - 1;$$

for the hypothesis H_2, the test statistic is given by

$$T_2 = \hat{\delta}_2 - 2\hat{\delta}_1 + 1;$$

and for the hypothesis H_3, the test statistic is given by

$$T_3 = \frac{\left(\hat{\delta}_2/\hat{\delta}_{20}\right) - 1}{\left(1 - \frac{1}{p}\left(\hat{\delta}_{40}/\hat{\delta}_{20}^2\right)\right)^{1/2}}$$

Let $\delta_i = p^{-1}\text{tr}\Sigma^i,\ i = 1, ..., 4,\ \ \delta_{20} = p^{-1}\sum_{i=1}^p \sigma_{ii}^2,\ \ \delta_{40} = p^{-1}\sum_{i=1}^p \sigma_{ii}^4$. We make the following assumption for the consistency of the statistics T_1, T_2, and T_3 in the general case; this assumption, however, is not needed for their consistency or their asymptotic distributions as $(N, p) \to \infty$, under their null hypotheses:

Assumption A: As $p \to \infty,\ \delta_i \to \delta_i^o,\ 0 < \delta_i^o < \infty,\ i = 1, ..., 4.$

Under *Assumption A*, it is shown that $\hat{\delta}_1$ and $\hat{\delta}_2$ are consistent estimators of δ_1 and δ_2 as $(N, p) \to \infty$. It may be noted that $\text{tr}S^2/p$ is not a consistent estimator of δ_2 unless $p/N \to 0$.

Next, we state the asymptotic distributions of the test statistics T_1, T_2, and T_3 under the null hypotheses as $(N, p) \to \infty$. The theorems will be proved in the subsequent sections. Let $\Phi(\cdot)$ denote the cdf of a standard normal random variable, $N(0, 1)$, and P_0 denotes the distribution under the null hypothesis.

Theorem 1.1. *Under the model (1.1)-(1.2),*

$$\lim_{(N,p)\to\infty} P_0\{(N/2)T_1 \leq t_1\} = \Phi(t_1),$$

where $\Phi(\cdot)$ denotes the cdf of a standard normal random variable, $N(0, 1)$, and P_0 denotes the distribution under the hypothesis H_1.

Theorem 1.2. *Under the model (1.1)-(1.2),*

$$\lim_{(N,p)\to\infty} P_0\{(N/2)T_2 \leq t_2\} = \Phi(t_2).$$

Theorem 1.3. *Under the model (1.1)-(1.2),*

$$\lim_{(N,p)\to\infty} P_0\{(N/2)T_3 \leq t_3\} = \Phi(t_3).$$

The asymptotic distributions for $T_1 \sim T_3$ which are presented in Theorem 1.1, Theorem 1.2 and Theorem 1.3 are the same as those obtained under normality assumption in Srivastava (2005). Thus the tests based on T_1, T_3 or T_3 are robust tests.

To obtain the distribution of the test statistic T_1 and T_2 we need to obtain the joint distribution of $\hat{\delta}_1$ and $\hat{\delta}_2$ under the model (1.1)-(1.2). To prove robustness, we need only obtain the joint distribution of $\hat{\delta}_1$ and $\hat{\delta}_2$ under the null hypotheses H_1 and H_2. Since the statistic T_1 is invariant under the scalar transformation $\mathbf{x}_i \to c\mathbf{x}_i,\ c \neq 0$, we shall assume without loss of generality that $\lambda = 1$. Thus, the results of the following theorem are applicable to both the statistics T_1 and T_2.

Theorem 1.4. *Let (1.1), (1.2), and $\Sigma = I_p$ hold. Then the joint distribution of $\hat{\delta}_1$ and $\hat{\delta}_2$, displayed in (1.4), as $(N,p) \to \infty$ in any manner, is given by*

$$\begin{pmatrix} \hat{\delta}_1 \\ \hat{\delta}_2 \end{pmatrix} \xrightarrow{d} N_2\left[\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \frac{1}{Np}\Omega\right],$$

where

$$\Omega = \begin{pmatrix} \gamma - 1 & 2(\gamma - 1) \\ 2(\gamma - 1) & 4(\gamma - 1) + 4\frac{p}{N} \end{pmatrix}. \tag{1.6}$$

The organization of the paper is as follows. In Section 2, we give some preliminary results needed to prove Theorem 1.4, which is proven in Section 3. The proofs of Theorem 1.1, Theorem 1.2 and Theorem 1.3 are given in Sections 4, 5 and 6, respectively. In particular, in Section 6 some of the notion and ideas of Section 2 will be repeated but now it is focused on T_3 instead of T_1 and T_2.

2 Some preliminary results

In this section we present some preliminary results. We begin with the sample covariance matrix S. Note that in probability S is equal to

$$S^* = \frac{1}{N}\sum_{i=1}^{N}(\mathbf{x}_i - \boldsymbol{\mu})(\mathbf{x}_i - \boldsymbol{\mu})'. \tag{2.7}$$

Thus $\hat{\delta}_1$ and $\hat{\delta}_2$, given in (1.4), can be approximated in probability by

$$\hat{\delta}_1^* = (\mathrm{tr}S^*/p), \quad \text{and} \quad \hat{\delta}_2^* = p^{-1}\left[\mathrm{tr}S^{*2} - N^{-1}(\mathrm{tr}S^*)^2\right], \tag{2.8}$$

respectively. In order to prove the consistency of $\hat{\delta}_1^*$ and $\hat{\delta}_2^*$, we need some results on quadratic forms, stated in the following subsection.

2.1 Moments of quadratic forms

Lemma 2.1. *Let $\mathbf{u} = (u_1, ..., u_p)'$ where u_i are iid with mean 0, variance 1, fourth moment γ, sixth moment γ_6 and eighth moment γ_8. Then for any $A = (a_{ij})$ and $B = (b_{ij})$ symmetric matrices of size $p \times p$,*

(a) $E(\mathbf{u}'A\mathbf{u})^2 = \left[(\gamma - 3)\sum_{i=1}^{p} a_{ii}^2 + 2\mathrm{tr}A^2 + (\mathrm{tr}A)^2\right],$

(b) $Var(\mathbf{u}'A\mathbf{u}) = \left[(\gamma - 3)\sum_{i=1}^{p} a_{ii}^2 + 2\mathrm{tr}A^2\right],$

(c) $E[(\mathbf{u}'A\mathbf{u})(\mathbf{u}'B\mathbf{u})] = \left[(\gamma - 3)\sum_{i=1}^{p} a_{ii}b_{ii} + 2\mathrm{tr}(AB) + (\mathrm{tr}A)(\mathrm{tr}B)\right],$

(d) $Cov[(\mathbf{u}'A\mathbf{u}), (\mathbf{u}'B\mathbf{u})] = \left[(\gamma - 3)\sum_{i=1}^{p} a_{ii}b_{ii} + 2\mathrm{tr}(AB)\right],$

(e) $Var[(\mathbf{u}'\mathbf{u})^2] = p(\gamma_8 - \gamma^2) + 4p(p-1)(\gamma_6 - \gamma) + 4(p-1)(p-2)(p-3)(\gamma - 1),$

(f) $E(\mathbf{u}'\mathbf{u})^3 = p\gamma_6 + 3p(p-1)\gamma + p(p-1)(p-2).$

Lemma 2.2. *Let u_i and v_j be independently and identically distributed with mean 0, variance 1 and fourth moment γ, $i, j = 1, ..., p$. Then for $\mathbf{u} = (u_1, ..., u_p)'$, and $\mathbf{v} = (v_1, ..., v_p)'$, and any $p \times p$ symmetric matrix $B = (b_{ij})$,*

$$Var[\mathbf{u}'\mathbf{B}\mathbf{v}]^2 = (\gamma - 3)^2 \sum_{i=1}^{p}\sum_{j=1}^{p} b_{ij}^4 + 6(\gamma - 3)\sum_{i=1}^{p}\left(\sum_{j=1}^{p} b_{ij}^2\right)^2 + 6\mathrm{tr}B^4 + 2(\mathrm{tr}B^2)^2.$$

2.2 Consistency of $\hat{\delta}_1^*$ and $\hat{\delta}_2^*$

For the sake of convenience of presentation, we shall not distinguish between δ_i and $\delta_i^o = \lim_{p\to\infty} \delta_i, i = 1, ..., 4$. From (1.1), $S^* = N^{-1}\sum_{i=1}^N C\mathbf{z}_i\mathbf{z}_i'C'$. Let $B = C'C = (b_{ij})$. Then

$$E(\hat{\delta}_1^*) = \frac{N}{Np}E(\mathbf{z}_i'B\mathbf{z}_i) = \frac{\text{tr}B}{p} = \delta_1,$$

$$Var(\hat{\delta}_1^*) = \frac{N}{N^2p^2}Var(\mathbf{z}_i'B\mathbf{z}_i) = \frac{1}{Np}\left[(\gamma-3)\sum_{i=1}^p \frac{b_{ii}^2}{p} + 2\frac{\text{tr}B^2}{p}\right].$$

Thus, under Assumption A, $Var(\hat{\delta}_1^*) = O((Np)^{-1})$, and $\hat{\delta}_1^*$ is a consistent estimator of δ_1. Now $\hat{\delta}_2^* = p^{-1}\left[\text{tr}S^{*2} - N^{-1}(\text{tr}S^*)^2\right] = N(N-1)N^{-2}\left(\text{tr}B^2/p\right) + a_1 + a_2 + a_3$, where

$$a_1 = \frac{1}{N^2p}\sum_{i=1}^N(\mathbf{z}_i'B\mathbf{z}_i - \text{tr}B)^2, \quad a_2 = -\left[\frac{1}{N^3p}\sum_{i=1}^N(\mathbf{z}_i'B\mathbf{z}_i - \text{tr}B)\right]^2,$$

$$a_3 = \frac{2}{N^2p}\sum_{i<j}^N[(\mathbf{z}_i'B\mathbf{z}_j)^2 - \text{tr}B^2].$$

We have,

$$E(a_1) = \frac{1}{Np}Var(\mathbf{z}_i'B\mathbf{z}_i) = \frac{1}{N}\left[(\gamma-3)\sum_{i=1}^p \frac{b_{ii}^2}{p} + \frac{2\text{tr}B^2}{p}\right],$$

$$E(-a_2) = \frac{1}{N^2}\left[\frac{Var(\mathbf{z}_i'B\mathbf{z}_i)}{p}\right].$$

Thus, from Markov's inequality, both a_1 and a_2 go to zero in probability as $(N,p) \to \infty$. Similarly, from Lemma 2.2, it can be shown that $Var(a_3) \to 0$ as $(N,p) \to \infty$. Hence $\hat{\delta}_2^*$ is a consistent estimator of δ_2 under the Assumption A.

2.3 Variance of $\hat{\delta}_2^*$ under the hypotheses H_1 and H_2

The proposed statistic T_1 is invariant under the scalar transformations $\mathbf{x}_i \to c\mathbf{x}_i$, $c \neq 0$. Thus we may assume without any loss of generality that $\Sigma = I$ under the hypothesis H_1, the same as for the hypothesis H_2. Hence all the results in this subsection are obtained under the assumption that $\Sigma = I_p$. When $\Sigma = I_p$, the observation matrix can be expressed in two ways:

$$Z = (z_{ij}) = (\mathbf{z}_1, ..., \mathbf{z}_N) = (\mathbf{w}_1, ..., \mathbf{w}_p)' = (w_{ij}). \tag{2.9}$$

Under H_1 and H_2 all the elements z_{ij} or w_{ij} are *iid* with mean 0 and variance 1. Thus,

$$E(\mathbf{w}_i) = \mathbf{0}, \quad Cov(\mathbf{w}_i) = I_N,$$

since $\mathbf{w}_i$ is an N-dimensional random vector. We shall now express $\hat{\delta}_2^*$ in terms of $\mathbf{w}_i$ as $B = I$ under H_1 and H_2. Thus under H_1 or H_2,

$$S^* = \frac{1}{N} ZZ' = \frac{1}{N}(\mathbf{w}_1, ..., \mathbf{w}_p)'(\mathbf{w}_1, ..., \mathbf{w}_p) \tag{2.10}$$

To evaluate the variance of $\hat{\delta}_2^*$, we rewrite $\hat{\delta}_2^*$ in terms of random vectors $\mathbf{w}_i,\ i = 1, ..., p$. That is, ($\approx$ stands for approximately equal to)

$$\hat{\delta}_2^* \approx q_1 + q_2, \tag{2.11}$$

where

$$q_1 = \frac{1}{N^2 p} \sum_{i=1}^{p} v_{ii}^2, \ v_{ii} = (\mathbf{w}_i' \mathbf{w}_i), \tag{2.12}$$

$$q_2 = \frac{2}{N^2 p} \left[\sum_{i<j}^{p} \left(v_{ij}^2 - \frac{1}{N} v_{ii} v_{jj} \right) \right], \ v_{ij} = \mathbf{w}_i' \mathbf{w}_j. \tag{2.13}$$

Let $\mathbf{w}$ be a random vector having the same distribution as $\mathbf{w}_i$, and $v = \mathbf{w}'\mathbf{w}$. Then, from Lemma 2.1(a)

$$E(q_1) = \frac{1}{N^2} E(v^2) = \frac{1}{N}(N + \gamma - 1).$$

Let

$$u_{ij} = v_{ij}^2 - \frac{1}{N} v_{ii} v_{jj} = (\mathbf{w}_i' \mathbf{w}_j \mathbf{w}_j' \mathbf{w}_i) - \frac{1}{N} (\mathbf{w}_i' \mathbf{w}_i)(\mathbf{w}_j' \mathbf{w}_j). \tag{2.14}$$

Then

$$q_2 = \frac{2}{Np} \sum_{i<j}^{p} u_{ij}, \text{ and } E(q_2) = 0. \tag{2.15}$$

From Lemma 2.1(e), we get the following theorem.

Theorem 2.3. *Let q_1 be given in (2.12). Then,*

$$Var(q_1) = 4(\gamma - 1)(Np)^{-1}[1 + O(N^{-1}p^{-1})].$$

To calculate the variance of q_2, we first evaluate

$$Cov(u_{ij}, u_{ik}) = \\ E\left[((\mathbf{w}_j'\mathbf{w}_i)^2 - N^{-1}(\mathbf{w}_i'\mathbf{w}_i)(\mathbf{w}_j'\mathbf{w}_j))((\mathbf{w}_k'\mathbf{w}_i)^2 - N^{-1}(\mathbf{w}_i'\mathbf{w}_i)(\mathbf{w}_k'\mathbf{w}_k))\right],$$

for $i \neq j \neq k$. Since,

$$\begin{aligned} E[(\mathbf{w}_i'\mathbf{w}_j\mathbf{w}_j'\mathbf{w}_i)(\mathbf{w}_i'\mathbf{w}_k\mathbf{w}_k'\mathbf{w}_i)] &= E(\mathbf{w}_i'\mathbf{w}_i)^2, \\ -\frac{1}{N}E[(\mathbf{w}_i'\mathbf{w}_j\mathbf{w}_j'\mathbf{w}_i)(\mathbf{w}_i'\mathbf{w}_i)(\mathbf{w}_k'\mathbf{w}_k)] &= -E(\mathbf{w}_i'\mathbf{w}_i)^2, \\ -\frac{1}{N}E[(\mathbf{w}_i'\mathbf{w}_i)(\mathbf{w}_j'\mathbf{w}_j)(\mathbf{w}_i'\mathbf{w}_k\mathbf{w}_k'\mathbf{w}_i)] &= -E(\mathbf{w}_i'\mathbf{w}_i)^2, \\ \frac{1}{N^2}E[(\mathbf{w}_i'\mathbf{w}_i)(\mathbf{w}_j'\mathbf{w}_j)(\mathbf{w}_i'\mathbf{w}_i)(\mathbf{w}_k'\mathbf{w}_k)] &= E(\mathbf{w}_i'\mathbf{w}_i)^2, \end{aligned}$$

it follows that

$$Cov(u_{ij}, u_{ik}) = 0, \ i \neq j \neq k. \tag{2.16}$$

Hence,

$$Var(q_2) = \frac{4}{N^4p^2}\sum_{i<j}^{p} Var(u_{ij}) = \frac{2p(p-1)}{N^4p^2} Var(u_{ij}).$$

Thus, we need to evaluate $Var(u_{ij}) = E(u_{ij}^2)$, since $E(u_{ij}) = 0$. Let $A_j = (a_{ik}(j)) = \mathbf{w}_j\mathbf{w}_j'$, $\mathbf{w}_j = (w_{j1}, ..., w_{jN})'$. Then, for $i \neq j$,

$$u_{ij}^2 = v_{ij}^4 - \frac{2}{N}v_{ij}^2 v_{ii}v_{jj} + \frac{1}{N^2}v_{ii}^2 v_{jj}^2, \quad \text{and} \quad v_{ij}^4 = (\mathbf{w}_i'\mathbf{w}_j\mathbf{w}_j'\mathbf{w}_i)^2 = (\mathbf{w}_i'A_j\mathbf{w}_i)^2.$$

Hence, for $i \neq j$

$$E(v_{ij}^4) = E[E(\mathbf{w}_i'A_j\mathbf{w}_i)|A_j]^2 = N[3N + (\gamma^2 - 3)].$$

Next, we evaluate

$$E(v_{ij}^2 v_{ii}v_{jj}) = E[\mathbf{w}_i'A_j\mathbf{w}_i\mathbf{w}_i'\mathbf{w}_i trA_j] = N(N + \gamma - 1)^2.$$

Finally,

$$E(v_{ii}^2 v_{jj}^2) = E(\mathbf{w}_i'\mathbf{w}_i)^2 E(\mathbf{w}_j'\mathbf{w}_j)^2 = N^2[N + \gamma - 1]^2.$$

Hence,

$$Var(u_{ij}) = (N-1)[(\gamma - 1)^2 + 2N],$$

and we get the following theorem.

Theorem 2.4. *Let $\mathbf{w}_1, ..., \mathbf{w}_p$ be iid with mean $\mathbf{0}$ and covariance I_N, and fourth moment γ. Then the variance of q_2 in (2.15) is given by*

$$Var(q_2) = \frac{4}{N^4 p^2} \frac{p(p-1)}{2}(N-1)[(\gamma-1)^2 + 2N] \approx \frac{4}{N^2}\left[1 + \frac{(\gamma-1)^2}{2N}\right].$$

We may also prove

Theorem 2.5. *Let q_1 and q_2 be given by (2.12) and (2.15), respectively. Then, $Cov(q_1, q_2) = 0$.*

Theorem 2.6. *Let $\hat{\delta}_1^*$ and q_2 be given by (2.8) and (2.15), respectively. Then, $Cov(\hat{\delta}_1^*, q_2) = 0$.*

3 Proof of Theorem 1.4

To establish the joint asymptotic normality of k statistics

$$t_{i,p}^{(n)} = \sum_{j=1}^{p} x_{ij}^{(n)}, \quad i = 1, ..., k$$

we consider an arbitrary linear combination

$$t_p^{(n)} = c_1 t_{1,p}^{(n)} + ... + c_k t_{k,p}^{(n)} = \sum_{j=1}^{p}\sum_{i=1}^{k} c_i x_{ij}^{(n)} \equiv \sum_{j=1}^{p} y_j^{(n)}$$

where without any loss of generality $c_1^2 + ... + c_k^2 = 1$, and $y_j^{(n)} = \sum_{i=1}^{k} c_i x_{ij}^{(n)}$. From the definition of multivariate normality, see Srivastava and Khatri (1979), the joint normality for all $c_1, .., c_k$ will follow if the normality of $t_p^{(n)}$ is established. Let $F_l^{(n)}$ be the σ-algebra generated by the random variables $(x_{1j}^{(n)}, ..., x_{kj}^{(n)},\ j = 1, ..., l),\quad l = 1, ..., p$. Then $F_0 \subset F_1^{(n)} \subset ... \subset F_p^{(n)} \subset F$, where $(\emptyset, F_0, \Lambda)$ is the probability space and $\emptyset$ being the null set.

Lemma 3.1. *Let $x_{ij}^{(n)}$ be a sequence of random variables, and $y_j^{(n)} = \sum_{i=1}^{k} c_i x_{ij}^{(n)}$, $j = 1, ..., p$, and $n = O(p^\delta)$, $\delta > 0$. We assume that*

(i) $E(y_j^{(n)} | F_{j-1}^{(n)}) = 0$,

(ii) $\lim_{(N,p)\to\infty} E[(y_j^{(n)})^2] < \infty$,

(iii) $\sum_{j=0}^{p} E[(y_j^{(n)})^2|F_{j-1}^{(n)}] \xrightarrow{p} \sigma_0^2$, *as* $(n,p) \to \infty$,

(iv) $L \equiv \sum_{j=0}^{p} E[(y_j^{(n)})^2 \ I(|y_j^{(n)}| > \epsilon)|F_{j-1}^{(n)}] \xrightarrow{p} 0$, *as* $(n,p) \to \infty$,

Then $t_p^{(n)} = \sum_{j=1}^{p} y_j^{(n)} \xrightarrow{d} N(0,\sigma_0^2)$, *as* $(n,p) \to \infty$.

The proof of this lemma follows from Theorem 4 of Shiryayev (1984, p. 511), since the first two conditions imply that $\{x_j^{(n)}, F_j^{(n)}\}$ forms a sequence of integrable martingale differences. The condition (iv) is known as Lindeberg's condition. To verify this condition, we note that from Markov's and Cauchy-Schwarz inequalities

$$P[L > \delta] \leq \sum_{j=0}^{p} E[(y_j^{(n)})^4]/\delta\epsilon^2.$$

Thus,

$$E[(y_j^{(n)})^4] \quad \leq \quad k^3 \sum_{i=1}^{k} c_i^4 E[(x_{ij}^{(n)})^4] \leq k^3 \sum_{i=1}^{k} E[(x_{ij}^{(n)})^4].$$

Hence, if

$$\sum_{j=1}^{p} E[(x_{ij}^{(n)})^4] \to 0,$$

for all $i = 1, ..., k$, the Lindeberg condition is satisfied. It is rather simple to evaluate σ_0^2 in most cases.

Because of the invariance of the statistic T_1 under a scalar transformation, we shall assume as before that $\Sigma = I$ and thus $B = I$ in both the hypotheses H_1 and H_2. We first consider the joint distribution of $\hat{\delta}_1^*$ and q_1 defined in (2.2) and (2.6) respectively, under $\Sigma = I_p$.

Let $\boldsymbol{\xi}_i = (\xi_{1i}, \xi_{2i})'$ where $\xi_{1i} = N^{-\frac{1}{2}}(\mathbf{w}_i'\mathbf{w}_i - N)$, $\xi_{2i} = N^{-\frac{3}{2}}[(\mathbf{w}_i'\mathbf{w}_i)^2 - N^2 - N(\gamma-1)]$, $i = 1, ..., p$ and $\mathbf{w}_i$ is as in Section 2. Then the vectors $\boldsymbol{\xi}_1$,...,$\boldsymbol{\xi}_p$ are *iid* with mean $\mathbf{0}$ and covariance matrix Ω_1 given by

$$\Omega_1 = \begin{pmatrix} \gamma - 1 & 2(\gamma-1) \\ 2(\gamma-1) & 4(\gamma-1) \end{pmatrix}.$$

Hence, from the multivariate central limit theorem

$$(1/\sqrt{p}) \sum_{i=1}^{p} \boldsymbol{\xi}_i \longrightarrow N_2(\mathbf{0}, \Omega_1),$$

irrespective of whether N goes to infinity and then p goes to infinity or p goes to infinity and then N goes to infinity. Since

$$\hat{\delta}_1^* = \frac{1}{p\sqrt{N}}\sum_{i=1}^{p}\xi_{1i} + 1, \text{ and } q_1 = \frac{1}{p\sqrt{N}}\sum_{i=1}^{p}\xi_{2i} + 1 + \frac{\gamma - 1}{N},$$

we get the following Lemma.

Lemma 3.2. *The asymptotic distribution of* $\hat{\delta}_1^*$ *and* q_1 *is bivariate normal given by*

$$\begin{pmatrix} \hat{\delta}_1^* \\ q_1 \end{pmatrix} \xrightarrow{d} N_2\left[\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \frac{1}{Np}\Omega_1\right]$$

as $(N, p) \to \infty$ *in any manner.*

It remains to find the distribution of q_2, to obtain the joint distribution of $\hat{\delta}_1^*$ and $\hat{\delta}_2^*$. Note that from (2.15),

$$Nq_2 = \sum_{j=2}^{p}\eta_j = \frac{2}{Np}\sum_{j=2}^{p}\left(\sum_{i=1}^{j-1}u_{ij}\right).$$

Let F_j be the σ-algebra generated by the random vectors $\mathbf{w}_1$,...,$\mathbf{w}_j$. Letting $\mathbf{w}_0 = 0$, and $F_0 = (\emptyset, \Lambda) = F_{-1}$, where $\emptyset$ is the empty set and Λ is the whole space, we find that $F_0 \subset F_1 \subset ... \subset F_p \subset F$. Also,

$$E(\eta_j|F_{j-1}) = 0.$$

Then

$$\begin{aligned} E(\eta_j^2|F_{j-1}) &= \frac{4}{N^2p^2}\left[\sum_{i=1}^{j-1}E(u_{ij}^2|F_{j-1}) + 2\sum_{k<l}^{j-1}E(u_{kj}u_{lj}|F_{j-1})\right] \\ &\equiv \frac{4}{N^2p^2}\left[\sum_{i=1}^{j-1}b_{iN} + 2\sum_{k<l}^{j-1}c_{klN}\right] \end{aligned}$$

and

$$E(\eta_j^2) = \frac{4}{N^2p^2}[(j-1)b_N + (\delta-1)(\delta-2)c_N], \; j \le p$$

where

$$b_N = E(b_{iN}) = E(u_{ij}^2) = N(N-1)[2 + \frac{(\gamma-1)^2}{N}],$$

giving

$$E(\eta_j^2) = \frac{4N(N-1)}{N^2p^2}(j-1)[2 + \frac{(\gamma-1)^2}{N}] < \infty,\ j \le p.$$

From the definition, it follows that (η_k, F_k) is a sequence of integrable martingale differences. To prove the normality of Nq_2, we apply Lemma 3.1. We note that

$$E\left[\sum_{j=0}^{p} E(\eta_j^2|F_{j-1})\right] = \sum_{j=0}^{p} E(\eta_j^2) = \frac{2N(N-1)}{N^2p^2}p(p-1)[2 + \frac{(\gamma-1)^2}{N}].$$

Thus

$$\lim_{p\to\infty} E\left[\sum_{j=0}^{p} E(\eta_j^2|F_{j-1})\right] = 4,$$

and $\sigma_0^2 = 4$. If we show that $\upsilon^2 = Var\left[\sum_{j=2}^{p} E(\eta_j^2|F_{j-1})\right] \to 0$, as $(N,p) \to \infty$, we find that

$$\upsilon^2 = Var\left\{\frac{4}{N^2p^2}\left[\sum_{j=2}^{p}\left(\sum_{i=1}^{j-1} b_{iN} + 2\sum_{k<l}^{j-1} c_{klN}\right)\right]\right\},$$

where

$$\begin{aligned} b_{iN} &= E(u_{ij}^2|F_{j-1}),\ i<j \\ &= E\left[(\mathbf{w}_j' A_i \mathbf{w}_j)^2 - \frac{2}{N}(\mathbf{w}_j' A_i \mathbf{w}_j)v_{jj}v_{ii} + \frac{1}{N^2}v_{ii}^2(\mathbf{w}_j'\mathbf{w}_j)^2|F_{j-1}\right], \end{aligned}$$

with $A_i = \mathbf{w}_i\mathbf{w}_i' = (a_{rl}(i)) : N \times N$. Using Lemma 2.1, yields

$$\begin{aligned} b_{iN} &= (\gamma-3)\sum_{r=1}^{N} a_{rr}^2(i) + 3(\mathbf{w}_i'\mathbf{w}_i)^2 \\ &\quad -\frac{2}{N}\left[(\gamma-3)\sum_{l=1}^{N} a_{ll}(i) + 2\mathbf{w}_i'\mathbf{w}_i + N\mathbf{w}_i'\mathbf{w}_i\right](\mathbf{w}_i'\mathbf{w}_i) \\ &\quad +\frac{1}{N^2}[(\gamma-3)N + 2N + N^2](\mathbf{w}_i'\mathbf{w}_i)^2 \\ &= d(\mathbf{w}_i'\mathbf{w}_i)^2 + (\gamma-3)\left(\sum_{k=1}^{N} w_{ik}^4\right),\ \ d = \left(2 - \frac{\gamma-1}{N}\right). \end{aligned}$$

Thus, to show that the variance of $4(N^2p^2)^{-1}\left(\sum_{j=2}^{p}\sum_{i=1}^{j-1} b_{iN}\right)$ goes to zero, it will be sufficient to show that the variance of $4d(N^2p^2)^{-1}\sum_{j=2}^{p}\sum_{i=1}^{j-1}\mathbf{w}_i'\mathbf{w}_i$, as well as the variance of $4(\gamma-3)(N^2p^2)^{-1}\sum_{j=2}^{p}\sum_{i=1}^{j-1}\left(\sum_{k=1}^{N} w_{ik}^4\right)$ go to zero. Clearly,

$$\begin{aligned} Var\left[\frac{4d}{N^2p^2}\sum_{j=2}^{p}\left(\sum_{i=1}^{j-1}\mathbf{w}_i'\mathbf{w}_i\right)\right] &= \frac{16d^2}{N^4p}Var\sum_{j=1}^{p-1}(p-j)(\mathbf{w}_j'\mathbf{w}_j) \\ &\leq \frac{16d^2}{N^4p}[(\gamma-3)N+N^2] \to 0 \text{ as } (N,p)\to\infty. \end{aligned}$$

Similarly, we need to show that

$$Var\left[\frac{8}{N^2p^2}\sum_{j=2}^{p}\sum_{k<l}^{j-1} c_{klN}\right] = \frac{8^2}{N^4p^4}Var\left[\sum_{1\leq k<l}^{p-1}(p-l-1)c_{klN}\right] \to 0.$$

For this, we need to calculate c_{klN} which after some calculations can be shown to equal

$$c_{klN} = E[u_{kj}u_{lj}|F_{j-1}] = (\gamma-3)\sum_{r=1}^{N} w_{rr}^2(k)w_{rr}^2(l) + 2\left[v_{kl}^2 - \frac{\gamma-1}{N}v_{kk}v_{ll}\right], \quad k<l<j.$$

Thus,

$$\begin{aligned} &\frac{64}{N^4p^4}Var\left[\sum_{1\leq k<l}^{p-1}(p-l-1)c_{klN}\right] \leq \frac{64}{N^4p^2}Var\left[\sum_{1\leq k<l}^{p-1} c_{klN}\right] \\ &= \frac{64}{N^4p^2}Var\left[\sum_{1\leq k<l}^{p-1}\left\{(\gamma-3)\sum_{r=1}^{N} w_{rr}^2(k)w_{rr}^2(l) + 2\left(v_{kl}^2 - \frac{\gamma-1}{N}v_{kk}v_{ll}\right)\right\}\right]. \end{aligned}$$

We need to show that the variance of each of the terms goes to zero. Clearly, the first term is of the order $O(N^{-3})$. Similarly, from the results of Section 2, the second term is of the order $O(N^{-2})$ and the third term is of the order $O(N^{-3})$. Hence, we have shown that condition (iii) is satisfied.

Next, we show that

$$\sum_{k=0}^{p} E(\eta_k^4) \to 0 \text{ as } (N,p)\to\infty.$$

For this, we note that $\eta_j = 2(Np)^{-1}\sum_{i=1}^{j-1} u_{ij}$, and hence,

$$\begin{aligned}
N^4p^4\sum_{j=0}^{p} E(\eta_j^4) &= 16E\sum_{j=2}^{p}\left[\sum_{i=1}^{j-1} u_{ij}^4 + 6\sum_{k<l}^{j-1} u_{kj}^2 u_{lj}^2\right]. \\
&= 16E\left[\sum_{j=2}^{p}\sum_{i=1}^{j-1} E(u_{ij}^4|F_{j-1}) + 6\sum_{k<l}^{j-1} E(u_{kj}^2u_{lj}^2|F_{j-1})\right].
\end{aligned}$$

Now ($A_i = \mathbf{w}_i\mathbf{w}_i'$)

$$\begin{aligned}
u_{ij}^4 &= \left[(\mathbf{w}_j'A_i\mathbf{w}_j)^2 - \frac{2}{N}(\mathbf{w}_j'A_i\mathbf{w}_j)v_{jj}v_{ii} + \frac{1}{N^2}v_{ii}^2(\mathbf{w}_j'\mathbf{w}_j)^2\right]^2 \\
&= (\mathbf{w}_j'A_i\mathbf{w}_j)^4 + \frac{4}{N^2}(\mathbf{w}_j'A_i\mathbf{w}_j)^2v_{jj}^2v_{ii}^2 + \frac{1}{N^4}v_{ii}^4(\mathbf{w}_j'\mathbf{w}_j)^4 \\
&\quad -\frac{4}{N}(\mathbf{w}_j'A_i\mathbf{w}_j)^3v_{jj}v_{ii} \\
&\quad +\frac{2}{N^2}(\mathbf{w}_j'A_i\mathbf{w}_j)^2(\mathbf{w}_j'\mathbf{w}_j)^2v_{ii}^2 - \frac{4}{N^3}(\mathbf{w}_j'A_i\mathbf{w}_j)(\mathbf{w}_j'\mathbf{w}_j)^2v_{ii}^3v_{jj}.
\end{aligned}$$

Let

$$g_i = E(u_{ij}^4|F_{j-1}),\ i<j,$$

and

$$h_{kl} = E(u_{kj}^2u_{lj}^2|F_{j-1}).$$

Then,

$$\begin{aligned}
\sum_{j=2}^{p} E(\eta_j^4) &= \frac{16}{N^4p^4}\left[\sum_{j=1}^{p-1}(p-j)E(g_j) + 6\sum_{1\le k<l}^{p-1}(p-l-1)h_{kl}\right] \\
&\le \frac{16}{N^4p^3}\left[\sum_{j=1}^{p-1}E(g_j) + 6\sum_{1\le k<l}^{p-1}E(h_{kl})\right] \\
&= O(p^{-2}) + O(p^{-1}),
\end{aligned}$$

from Lemma 2.1. Thus, the Lindeberg condition is also satisfied. Hence, as $(N,p)\to\infty$,

$$Nq_2 \to N(0,4)$$

or equivalently q_2 is asymptotically normally distributed as normal with mean 0 and variance $4/N^2$ under the hypothesis H.

We shall now apply Lemma 3.1 again to obtain the joint normality of $\hat{\delta}_1^*, q_1$, and q_2. In the notation of Lemma 3.1, let

$$t_{1,p}^{(n)} = \sum_{i=1}^{p} \left(\frac{\xi_{1i}}{\sqrt{p}} \right), \quad t_{2,p}^{(n)} = \sum_{i=1}^{p} \left(\frac{\xi_{2i}}{\sqrt{p}} \right), \quad t_{3,p}^{(n)} = \sum_{i=1}^{p} \eta_i.$$

It is easy to check that

$$\sum_{i=1}^{p} E\left[\left(\frac{\xi_{1i}}{\sqrt{p}} \right)^4 \right] \text{ and } \sum_{i=1}^{p} E\left[\left(\frac{\xi_{2i}}{\sqrt{p}} \right)^4 \right]$$

go to zero as $(N, p) \to \infty$ while we have already shown that $\sum_{i=1}^{p} E(\eta_i^4) \to 0$ as $(N, p) \to \infty$. Similarly, the convergence can be satisfied. Hence, we have

$$\begin{pmatrix} \hat{\delta}_1^* \\ q_1 \\ q_2 \end{pmatrix} \sim N_3 \left[\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} (Np)^{-1}\Omega_1 & \mathbf{0} \\ \mathbf{0} & 4/N^2 \end{pmatrix} \right]$$

Hence

$$\begin{pmatrix} \hat{\delta}_1^* \\ \hat{\delta}_2^* \end{pmatrix} \sim N_2 \left[\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \frac{1}{Np}\Omega \right],$$

where Ω is defined in (1.7). This proves Theorem 1.4.

4 Robustness of the sphericity test: proof of Theorem 1.1

In this section, we first discuss various tests available for testing the hypothesis of 'sphericity' H_1. When $N > p$, the likelihood ratio test (LRT) is based on the ratio of the arithmetic mean to the geometric mean of the eigenvalues of the sample covariance matrix S. The power of the LRT is a monotonically increasing function of the ratio of the eigenvalues of Σ, see Carter and Srivastava (1977). Another test, known in the literature as the locally best invariant test (LBIT) was originally proposed by Nagao (1970) but it was John (1971) and Sugiura (1972) who showed that it is the LBIT. It is based on the statistic

$$U = \left[\frac{\text{tr}S^2}{p\hat{\delta}_1^2} \right] - 1$$

It may be noted that $\left(\frac{\text{tr}S^2}{p}\right)$ is a consistent estimator of $\left(\frac{\text{tr}\Sigma^2}{p}\right)$, if $\left(\frac{p}{N}\right) \to 0$. Thus, when $\frac{p}{N} \to c \neq 0$, Ledoit and Wolf considered the statistic $U - c$ and using the asymptotic result of Jonsson (1982) gave its (N, p) asymptotic null-distribution under the assumption A and the assumption that $\left(\frac{p}{N}\right) \to c$ as $(N, p) \to \infty$. The (N, p) asymptotic non-null distribution of $U - c$ can be obtained from Corollary 2.1 of Srivastava (2005).

It may be noted that the statistic U exists irrespective of whether $N \leq p$ or $N > p$. Next, we define a measure of sphericity which differs from the one given by Ledoit and Wolf (2002). From Cauchy-Schwarz inequality, we have for a $p \times p$ positive definite matrix Σ,

$$\frac{\delta_2}{\delta_1^2} = \frac{(\text{tr}(\Sigma^2)/p)}{(\text{tr}\Sigma/p)^2} \geq 1. \tag{4.17}$$

The equality holds if and only if (iff) all the eigenvalues of Σ are equal to some unknown constant, say λ. That is, iff $\Sigma = \lambda I_p$. Thus, as in Srivastava (2005), a measure of sphericity may be defined by

$$m_s = \left[\frac{(\text{tr}(\Sigma^2)/p)}{(\text{tr}\Sigma/p)^2} - 1\right], \tag{4.18}$$

which takes the value zero iff $\Sigma = \lambda I$, the sphericity hypothesis. The statistic T_1 defined in Section 1 is a consistent estimator of m_s. It may be noted that the statistic T_1 is invariant under the scalar transformation $\mathbf{x}_i \to a\mathbf{x}_i,\ a \neq 0$. Thus, without any loss of generality, we may assume that $\lambda = 1$ in obtaining the distribution of T_1.

We use Theorem 1.4 to obtain the distribution of T_1 under the hypothesis H_1 as $(N, p) \to \infty$. Under H_1, $\hat{\delta}_1$ and $\hat{\delta}_2$ are consistent estimators of $\delta_1 = 1$, and $\delta_2 = 1$ respectively. Now

$$\frac{\partial T_1}{\partial \hat{\delta}_1} = -2\frac{\hat{\delta}_2}{\hat{\delta}_1^3}, \quad \frac{\partial T_1}{\partial \hat{\delta}_2} = \frac{1}{\hat{\delta}_1^2}$$

Thus $(Np)^{-1}(-2, 1)\Omega(-2, 1)' = 4N^{-2}$.

Hence, under H_1, $T_1 \xrightarrow{d} N(0, 4N^{-2})$ as $(N, p) \to \infty$, proving Theorem 1.1, as well as showing that the test statistic T_1 for testing sphericity is robust.

5 A robust test for testing that Σ is an identity matrix: proof of Theorem 1.2

Despite the monotonicity property of the power function of the LRT for this problem established by Nagao (1967) and DasGupta (1969), it cannot be considered since $N \leq p$. Thus, we consider a test based on a consistent estimator of the distance function that measures the departure of the hypothesis from the alternative, namely,

$$m_I = \frac{1}{p}\text{tr}(\Sigma - I)^2 = \delta_2 - 2\delta_1 + 1.$$

Thus, Rao (1948), and independently Nagao (1973) proposed a test statistic

$$V = \frac{1}{p}\text{tr}S^2 - 2\hat{\delta}_1 + 1,$$

for testing the hypothesis that $\Sigma = I_p$. Ledoit and Wolf (2002) modified it to

$$W = V - \frac{p}{n}[\hat{\delta}_1^2 - 1],$$

and obtained its null distribution under the condition that

$$\lim_{(N,p)\to\infty} \frac{p}{N} = c > 0.$$

Using consistent estimators of δ_1 and δ_2, Srivastava (2005) proposed a test based on the statistic

$$T_2 = \hat{\hat{\delta_2}} - 2\hat{\hat{\delta_1}} + 1,$$

and obtained its null as well as non-null distribution as $(N, p) \to \infty$. In this article we show that T_2 is a robust test under the non-normality model (1.1)-(1.2). To obtain the distribution T_2, we use Theorem 1.4. Since

$$\frac{\partial T_2}{\partial \hat{\hat{\delta}}_1} = -2, \quad \frac{\partial T_2}{\partial \hat{\hat{\delta}}_2} = 1,$$

we have

$$(Np)^{-1}(-2, 1)'\Omega(-2, 1)' = 4N^{-2}$$

Thus as $(N, p) \to \infty$, $T_2 \xrightarrow{d} N(0, \frac{4}{N^2})$, and hence proving Theorem 1.2 and the robustness of the test statistic T_2 as it does not depend on $\gamma, \gamma_3, \gamma_5 - \gamma_8$, it is the same distribution as given by Srivastava (2005) under the assumption of normality.

6 Robustness of the diagonality test T_3: proof of Theorem 1.3

When the observations are normally distributed, the LRT is based on the determinant of the sample correlation matrix.

$$R = (r_{ij}),\ r_{ii} = 1,\ r_{ij} = \frac{s_{ij}}{(s_{ii}s_{jj})^{1/2}},$$

provided $N > p$. When $N \leq p$, the determinant of R does not exist. By defining the distance function as the sum of squared correlations $\rho_{ij}^2 \frac{\sigma_{ij}^2}{\sigma_{ii}\sigma_{jj}}$, $\sum\limits_{i<j} \rho_{ij}^2$ which is zero iff $\rho_{ij} = 0$, Srivastava (2005, 2006) proposed a test based on the normalized version of its consistent estimator. Schott (2005) also gave its distribution under the condition that $\frac{p}{N} \to c$. However, since the convergence to normality is slow, Srivastava (2005, 2006) proposed a test based on Fisher's transformation, and obtain its (N, p) asymptotic distribution. Srivastava (2005) defined another distance function to measure the departure from the hypothesis H_3. It is given by

$$m_d = \left[\left(\frac{\text{tr}\Sigma^2}{\sum\limits_{i=1}^{p} \sigma_{ii}^2} \right) - 1 \right],\ \Sigma = (\sigma_{ij}),$$

which is zero if and only if $\rho_{ij} = 0$. Under normality, a test based on its consistent estimator is given by the test statistic T_3 defined in Section 1. The (N,p) asymptotic distribution is given in Srivastava (2005) and its power compared in Srivastava (2006) with the test based on Fisher's transformation and shown to be at least as good as based on the Fisher's transformation. In this section, we show that this test T_3 defined in Section 1 is robust under the model (1.1)-(1.2). As in Section 2, we can for the asymptotic distribution purposes, consider $\hat{\delta}_2^*$ based on S^* instead of S, and N in place of $N-1$ and may show that

$$\hat{\delta}_2^* \quad \approx \quad \hat{\delta}_{20}^* + 2\sum_{i<j}^{p} ({s_{ij}^*}^2 - \frac{1}{N} s_{ii}^* s_{jj}^*),$$

where $\hat{\delta}_{20}^* = p^{-1} \sum_{i=1}^{p} {s_{ii}^*}^2$.

Under the hypothesis H_3, $\Sigma = D$ with $C = D^{1/2}$. Hence, if $\mathbf{w}_i$ are *iid* with mean $\mathbf{0}$, covariance I_n, with fourth moment γ and the existence of eight moments, we can write

$$s_{ij}^* = d_i d_j \mathbf{w}_i' \mathbf{w}_j \text{ for all } i, j = 1, ..., p.$$

Let

$$q_3^* = \frac{2}{p} \sum_{i<j}^{p} \left({s_{ij}^*}^2 - \frac{1}{N} s_{ii}^* s_{jj}^* \right) \equiv \frac{2}{N^2 p} \sum_{i<j}^{p} d_i d_j u_{ij},$$

with $E(u_{ij}) = 0$, and $Cov(u_{ij}, u_{ik}) = 0, i \neq j \neq k$. Hence

$$Var(q_3^*) = 4 \sum_{i<j}^{p} d_i^2 d_j^2 Var(u_{ij}) = \frac{4}{N^2} [\delta_{20}^2 - p^{-1} \delta_{40}] + O(N^{-3}).$$

We now show that $\hat{\delta}_{20}^*$ and $\hat{\delta}_{40}^*$ are consistent estimators of $\delta_{20} = p^{-1} \sum_{i=1}^{p} \sigma_{ii}^2$ and $\delta_{40} = p^{-1} \sum_{i=1}^{p} \sigma_{ii}^4$, respectively under the hypothesis H_3 when $C = D^{1/2} = diag(d_1^{\frac{1}{2}}, ..., d_p^{\frac{1}{2}})$. In terms of the *iid* random vector $\mathbf{w}_i$,

$$\hat{\delta}_{20} = \frac{1}{pN^2} \sum_{i=1}^{p} d_i^2 (\mathbf{w}_i' \mathbf{w}_i)^2,$$

and its variance is given by

$$Var(\hat{\delta}_{20}) = \frac{1}{pN^4} Var(\mathbf{w}_i' \mathbf{w}_i)^2 \left(\sum_{i=1}^{p} \frac{d_i^4}{p} \right) = O(N^{-1} p^{-1})$$

from Assumption A and Lemma 2.1(e). Since $E(\hat{\delta}_{20}) = \delta_{20}[1 + O(N^{-1})]$, $\hat{\delta}_{20}$ is a consistent estimator of δ_{20}. Similarly, it can be shown that $\hat{\delta}_{40}$ is a consistent estimator of δ_{40}. Let

$$\eta_k^* = \frac{2}{Np} d_k \sum_{i=1}^{k-1} d_i u_{ik}$$

Then following the steps of Section 3, it can be shown that $\{\eta_k^*, F_k\}$ is a sequence of integrable martingale difference satisfying the convergence condition and Lindeberg's condition, i.e. Lemma 3.1, (iii), (iv). Thus, Theorem 1.3 follows and thus the test statistic T_3 is shown to be robust.

References

[1] Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing, *J. Royal Statist. Soc.*, **B57**, 289-300

[2] Beran, R. and Srivastava, M.S. (1985). Bootstrap tests and confidence regions for functions of a covariance matrix. *Ann. Statist.*, **13**, 95-115.

[3] Carter, E.M. 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. Multivariate Anal.*, **7**, 229-233.

[4] Chan, Y.M. and Srivastava, M.S. (1988). Comparison of powers for the sphericity test using both the asymptotic distribution and the bootstrap. *Comm. Statist.*, **17**, 671-690.

[5] DasGupta, S. (1969). Properties of power functions of some tests concerning dispersion matrices of multivariate normal distributions. *Ann. Math. Statist.*, **40**, 697-701.

[6] John, S. (1971). Some optimal multivariate tests, *Biometrika*, **58**, 123-127.

[7] John, S. (1972). The distribution of a statistic used for testing sphericity of normal distributions. *Biometrika*, **67**, 31-43.

[8] Jonsson, D. (1982). Some limit theorems for the eigenvalues of a sample covariance matrix. *J. Multivariate Anal.*, **12**, 1-38.

[9] Kubokawa, T. and Srivastava, M.S. (1999). Robust improvement in estimation of a covariance matrix in an elliptically contoured distribution. *Ann. Statist.*, **27**, 600-609.

[10] Ledoit, O. and 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.

[11] Nagao, H. (1967). Monotonicity of the modified likelihood ratio test for a covariance matrix. *J. Sci. Hiroshima Univ., Ser. A-I* **31**, 147-150.

[12] Nagao, H. (1970). Asymptotic expansions of some test criteria for homogeneity of variances and covariance matrices from normal populations. *J. Sci. Hiroshima Univ., Ser. A-I*, **34**, 153-247.

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

[14] Nagao, H. and Srivastava, M.S. (1992). On the distribution of some test criteria for a covariance matrix under local alternatives and bootstrap approximations. *J. Multivariate Anal.*, **43**, 331-350.

[15] Purkayastha, S. and Srivastava, M.S. (1995). Asymptotic distributions of some test criteria for the covariance matrix in elliptical distributions under local alternatives, *J. Multivariate Anal.*, **55**, 165-186.

[16] Rao, C.R. (1948). Large sample tests of statistical hypotheses concerning several parameters with applications to problems of estimation. *Mathematical Proc. Cambridge Philo. Soc.*, **44**, 50-57.

[17] Schott, J.R. (2005). Testing for complete independence in high dimensions. *Biometrika*, **92**, 951-956.

[18] Shiryayev, A.N. (1984). *Probability*, Springer-Verlag, New York.

[19] Srivastava, M.S. (2005). Some tests concerning the covariance matrix in high-dimensional data, *J. Japan Stat. Soc.*, **35**, 251-272.

[20] Srivastava, M.S. (2006). Some tests criteria for the covariance matrix with fewer observations than the dimension, *Acta et Commentationes Universitatis Tartuensis de Mathematica*, **10**, 77-93.

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

[22] Sugiura, N. (1972). Locally best invariant test for sphericity and the limiting distributions, *Ann. Math. Statist.*, **43**, 1312-1316.