A test for the mean vector with fewer observations than the dimension under non-normality

Muni S. Srivastava

University of Toronto, Department of Statistics, 100 St. George Street, M5S 3G3 Toronto, Ontario, Canada
Bowling Green State University, Bowling Green, OH, 43403, USA

ARTICLE INFO

Article history:
Received 14 May 2007
Available online 17 June 2008

AMS subject classifications:
62H10
62H15

Keywords:
Asymptotic null and non-null distribution
Fewer observations
High dimension
Non-normality
Testing mean vector

ABSTRACT

In this article, we consider the problem of testing that the mean vector \( \mu = 0 \) in the model

\[
\mathbf{x}_j = \mu + C \mathbf{z}_j, \quad j = 1, \ldots, N,
\]

where \( \mathbf{z}_j \) are random \( p \)-vectors, \( \mathbf{z}_j = (z_{ij}, \ldots, z_{pj})' \) and \( z_{ij} \) are independently and identically distributed with finite four moments, \( i = 1, \ldots, p, \quad j = 1, \ldots, N \); that is, \( \mathbf{x}_j \) need not be normally distributed. We shall assume that \( C \) is a \( p \times p \) non-singular matrix, and there are fewer observations than the dimension, \( N \leq p \). We consider the test statistic

\[
T = \left[ \frac{N \mathbf{x} \mathbf{D}^{-1} \mathbf{x} - np(n - 2)}{2tr \mathbf{R}^2 - p^2/n} \right]^{1/2},
\]

where \( \mathbf{x} \) is the sample mean vector, \( S = (s_{ij}) \) is the sample covariance matrix, \( D_S = \text{diag}(s_{11}, \ldots, s_{pp}) \), \( R = D_S^{-1/2}D_S^{-1} \) and \( n = N - 1 \). The asymptotic null and non-null distributions of the test statistic \( T \) are derived.

© 2008 Published by Elsevier Inc.

1. Introduction

Let \( \mathbf{x}_1, \ldots, \mathbf{x}_N \) be \( N \) independent and identically distributed (hereafter, iid) random \( p \)-vectors such that

\[
\mathbf{x}_j = \mu + C \mathbf{z}_j,
\]

where

\[
\mathbf{z}_j = (z_{1j}, \ldots, z_{pj})',
\]

\( j = 1, \ldots, N, \quad N \leq p \), and \( C \) is a \( p \times p \) non-singular matrix, and thus \( \Sigma = CC' \) is a positive definite matrix, written as \( \Sigma > 0 \). It is assumed that \( z_{ij} \) are iid with first, second and fourth moments given by

\[
E(z_{ij}) = 0, \quad E(z_{ij}^2) = 1, \quad E(z_{ij}^4) = \gamma < \infty,
\]

\( i = 1, \ldots, p, \quad j = 1, \ldots, N, \quad N \leq p \). There is however, no restriction on the third moment. Let the sample mean vector and the sample covariance matrix of the \( N \) observation vectors \( \mathbf{x}_1, \ldots, \mathbf{x}_N \) be respectively given by

\[
\bar{\mathbf{x}} = N^{-1} \sum_{j=1}^{N} \mathbf{x}_j, \quad \mathbf{S} = n^{-1}V = n^{-1} \sum_{j=1}^{N} (\mathbf{x}_j - \bar{\mathbf{x}})(\mathbf{x}_j - \bar{\mathbf{x}})', \quad n = N - 1.
\]

* Corresponding address: University of Toronto, Department of Statistics, 100 St. George Street, M5S 3G3 Toronto, Ontario, Canada.
E-mail address: srivasta@utstat.toronto.edu.

0047-259X/$ – see front matter © 2008 Published by Elsevier Inc.
doi:10.1016/j.jmva.2008.06.006
We consider testing the hypothesis
\[ H : \mu = 0 \text{ vs } A : \mu \neq 0 \] (1.5)
with the test statistics
\[ T = \frac{[\mathbf{N} \mathbf{X} D^{-1}_s \mathbf{X} - np/(n - 2)]}{[2tr R^2 - p^2/n]^{1/2}}, \] (1.6)
where
\[ R = D^{-1}_s SD^{-1}_s, \]
\[ D_s = \text{diag}(s_{11}, \ldots, s_{pp}), \quad S = (s_{ij}). \] (1.7)

The statistic \( T \) uses the information from the diagonal elements of the singular sample covariance matrix \( S \) as compared to the tests proposed by Dempster \[2\] and Bai and Saranadasa \[1\] where the information from the \( s_{ii} \) has not been used. It may be noted that \( \frac{1}{n} \) is the sample variance of \( X_i \), where \( X_i = (X_{i1}, \ldots, X_{ip})' \). In this article, we derive the asymptotic distribution of the test statistic \( T \) when model (1.1) holds but without assuming that the random vectors \( X_i \) are iid multivariate normal. Srivastava and Du \[8\] obtained the distribution of \( T \) under the normality assumption on the random vectors \( X_j, j = 1, \ldots, N \) where they showed by simulation that the statistics \( T \) divided by \( q_{p,n} \) defined in (1.9) leads to a faster convergence to normality and thus assumed that \( n = O(p^3), \frac{1}{2} < \delta \leq 1 \) so that \( q_{p,n} \) converges to one. The quantity \( q_{p,n} \) is given by
\[ q_{p,n} = [1 + (\text{tr} R^2)/p^2]^{1/2}. \] (1.9)

Srivastava and Du \[8\] also showed that the power of the test \( T \) in (1.6) is superior to the tests proposed by Dempster \[2\] and Bai and Saranadasa \[1\], theoretically as well as numerically. Thus, it is important to know how robust the test based on the statistic \( T \) is. This goal is achieved in this article. The results of this article can also be used to show the robustness of Dempster, and Bai and Saranadasa tests.

In the course of establishing the robustness of the \( T \)-test described above under the distributional assumptions given in (1.1)–(1.3), we also obtain the consistency of \( \text{tr} V/np \) and \((n - 1)(n + 2)p^{-1}[\text{tr} V^2 - 2(\text{tr} V)^2] \) as estimators of \( (\text{tr} \Sigma/p) \) and \( (\text{tr} \Sigma^2/p) \) respectively, without assuming normality. In addition, the asymptotic normality of certain quadratic forms is obtained which may be of great many other problems.

The organization of the article is as follows. In Section 2, some technical results are presented which may be of general interest. The asymptotic distribution of the test statistic \( T \) is given in Section 3 when the hypothesis is true and in Section 4 when the alternative hypothesis holds. The asymptotic distributions are obtained when \( n = O(p^3), 0 < \delta \leq 1 \). The technical proofs are given in Section 5. For a review of some tests in high dimensional data, see \[3\], where it is required that \( n > p \) and \((n/p) \to C > 1 \).

2. Some preliminary results and notations

In this section, we state some results, which though needed in the development of technical results of this article, may be of general interest to the reader. The proofs will be given in Section 5. Throughout this paper, we shall use the notation that \( Z_n \xrightarrow{p} a \) whenever, as \( n \to \infty \),
\[ P \{|Z_n - a| > \epsilon\} \to 0 \]
for \( \epsilon > 0 \), which can be established by using Markov’s inequality or Chebyshev’s inequality. When \( Z_n \) is an estimator, then we may also say that \( Z_n \) is a consistent estimator of \( a \). In this article, most estimators depend on \( n \) and \( p \). But under the stated assumptions, the variances or the expected values depend only on \( n \) as the parts depending on \( p \) are bounded. Some limiting distributions are given when both \( n \) and \( p \) go to infinity and the notation \((n, p) \to \infty \) or \( \lim_{(n, p) \to \infty} \) means that the results hold irrespective of how they go to infinity.

2.1. Moments of quadratic forms

We state some results on the quadratic forms in the following three lemmas. Their proofs can be obtained from the author.

**Lemma 2.1.** Let \( \mathbf{u} = (u_1, \ldots, u_p)' \) where \( u_i \) are iid with mean 0, variance \( \sigma^2 \) and fourth moment \( \sigma^4 \gamma \). Then, for any \( A = (a_{ij}) \) and \( B = (b_{ij}) \) symmetric matrices of order \( p \times p \),
\[ (a) \quad E(\mathbf{u}' A \mathbf{u})^2 = \sigma^4 \left( \gamma^2 - 3 \right) \sum_{i=1}^{p} a_{ii}^2 + 2 \text{tr} A^2 + (\text{tr} A)^2, \]
\[ (b) \quad \text{Var}(\mathbf{u}' A \mathbf{u}) = \sigma^4 \left( \gamma^2 - 3 \right) \sum_{i=1}^{p} a_{ii}^2 + 2 \text{tr} A^2, \]
Let $u_i$ and $v_j$ be independently and identically distributed with mean 0, variance 1 and fourth moment $\gamma$, $i, j = 1, \ldots, p$. Then for $u = (u_1, \ldots, u_p)'$ and $v = (v_1, \ldots, v_p)'$, and any $p \times p$ symmetric matrix $B = (b_{ij})$, we have

$$\text{Var}[u'Bv]^2 = (\gamma - 3)^2 \sum_{i=1}^p \sum_{j=1}^p b_{ij}^4 + 6(\gamma - 3) \sum_{i=1}^p \left( \sum_{j=1}^p b_{ij}^2 \right)^2 + 6trB^4 + 2(trB^2)^2.$$ 

Lemma 2.3. Let $z_{ij}$ be iid with mean 0, variance 1, fourth moment $\gamma$ and $\mathbf{Z} = (z_1, \ldots, z_p)'$, where $z_i = (1/N) \sum_{j=1}^N z_{ij}$, $i = 1, \ldots, p, j = 1, \ldots, N$. Then for any $p \times p$ symmetric matrix $A = (a_{ij})$,

$$\text{Var}(N\mathbf{Z}^t A \mathbf{Z}) = \frac{\gamma - 3}{N^2} \sum_{i=1}^p a_{ii}^2 + 2trA^2.$$ 

Corollary 2.1.

$$\text{Var} \left( N \sum_{i=1}^p a_{ii} z_i^2 \right) = \frac{\gamma - 3}{N} \sum_{i=1}^p a_{ii}^2 + 2 \sum_{i=1}^p a_{ii}^2.$$ 

Corollary 2.2.

$$\text{Var} \left( N \sum_{j=k} a_{jk} \bar{z}_j \bar{z}_k \right) = \frac{1}{2} \left( \text{tr}A^2 - \sum_{i=1}^p a_{ii}^2 \right).$$

Corollary 2.3.

$$\text{Var} (N\bar{z}_i^2) = \frac{\gamma - 3}{N} + 2.$$ 

2.2. Some results on asymptotic normality

We first state a lemma on asymptotic normality of a linear combination of iid random variables, due to Srivastava [7].

Lemma 2.4. Let $z_1, \ldots, z_p$ be iid random variables with mean 0 and variance 1. Then for a sequence of constants $a_{ip}$ satisfying

$$\lim_{p \to \infty} \max_{1 \leq i \leq p} a_{ip}^2 = 0 \quad \text{and} \quad \lim_{p \to \infty} \sum_{i=1}^p a_{ip}^2 = 1,$$

$$\lim_{p \to \infty} P \left( \sum_{i=1}^p a_{ip} z_i \leq x \right) = \Phi(x),$$

where $\Phi$ is the cumulative distribution function (cdf) of a standard normal random variable.

Let $A_{++} = (a_{ii}+)$ be a $p \times p$ symmetric matrix defined by $a_{ii}+ = a_{ii}$ and $a_{ij}+ = |a_{ij}|$, where $A = (a_{ij})$ is a $p \times p$ symmetric matrix. We shall assume that

$$\lim_{p \to \infty} \frac{\text{tr}A_{++}^i}{p} < \infty, \quad i = 1, 2, 4. \tag{2.1}$$

It may be noted that the assumption in (2.1) implies that we also have

$$\lim_{p \to \infty} \frac{\text{tr}A_{i}^i}{p} < \infty, \quad i = 1, 2, 4.$$ 

Next, we state a theorem regarding the conditions under which the asymptotic normality of a quadratic form is obtained.
Theorem 2.1. Let \( z_{ij} \) be iid random variables with mean 0, variance 1, fourth moment \( \gamma \) and \( \bar{z} = (\bar{z}_1, \ldots, \bar{z}_p)' \) where \( \bar{z}_i = (1/N) \sum_{j=1}^{N} z_{ij}, i = 1, \ldots, p \). Then for any \( p \times p \) symmetric matrix, \( A = (a_{ij}) \) for which

\[
\begin{align*}
(i) \quad & \lim_{p \to \infty} \max_{1 \leq j \leq p} \left( \frac{a_{jj}^2}{p} \right) = 0, \\
(ii) \quad & \lim_{p \to \infty} (\text{tr } A_i / p) < \infty, \quad i = 1, 2, 4,
\end{align*}
\]

hold,

\[
\lim_{(N,p) \to \infty} P \left[ \left( N \frac{\bar{z} A \bar{z}}{\sqrt{2p^2}} - \frac{\text{tr } A}{\sqrt{p}} \right)^2 \leq x \right] = \Phi(x),
\]

where \( \Phi(x) \) is the cdf of a standard normal random variable, and \( \tau^2 = \frac{\nu A^2}{p} \).

Corollary 2.4. Let \( z_1, z_2, \ldots \) be a sequence of iid random variables on the probability space \((\Omega, F, P)\) with mean 0, variance one and fourth moment \( \gamma \). Let \( A = (a_{ij}) \) be a \( p \times p \) symmetric matrix satisfying (i) and (ii) of Theorem 2.1. Then, for \( z = (z_1, \ldots, z_p)' \), a random \( p \)-vector,

\[
\lim_{p \to \infty} P \left[ \frac{z' A z - \text{tr } A}{\sqrt{p}} \leq x \right] = \Phi(x),
\]

where

\[
\tau^2 = \lim_{p \to \infty} \frac{1}{p} \left[ (\gamma - 3) \sum_{i=1}^{p} a_{ii}^2 + 2 \text{tr } A^2 \right].
\]

Corollary 2.5. Let \( A = I \). Then \( a_{ii} = 1 \), and \( a_{ij} = 0 \), \( i \neq j \). Hence,

\[
\lim_{p \to \infty} P \left[ \frac{z' z - p}{\sqrt{p(\gamma - 1)}} < x \right] = \Phi(x).
\]

2.3. Estimators based on the sample covariance matrix

For model (1.1), the \( p \times p \) covariance matrix \( \Sigma = CC' \), is assumed to be positive definite.

Define

\[
\delta_j = \frac{1}{p} \text{tr } (\Sigma^j), \quad j = 1, 2, 3, 4.
\]

We shall assume that

\[
0 < \lim_{p \to \infty} \delta_j = \delta_{j0} < \infty. \tag{2.2}
\]

Let

\[
\hat{\delta}_1 = \frac{1}{p} \text{tr } (S) \tag{2.3}
\]

and

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

Theorem 2.2. Under the assumption (2.2), \( \hat{\delta}_1 \) and \( \hat{\delta}_2 \) are consistent estimators of \( \delta_1 \) and \( \delta_2 \) as \( N \to \infty \). If \( N = O(p^\epsilon) \), \( 0 < \epsilon \leq 1 \), \( \hat{\delta}_1 \) and \( \hat{\delta}_2 \) are consistent estimators of \( \delta_1 \) and \( \delta_2 \) as \( (N, p) \to \infty \).

In order to simplify the derivation of the asymptotic distribution and consistency of the estimators, we assume that for large \( N \), \( S \) can be replaced by \( S^* \) given by

\[
S^* = \frac{1}{N} \sum_{i=1}^{N} (x_i - \mu)(x_i - \mu)' = (S^*_y), \tag{2.5}
\]
throughout the theoretical derivations of this article. This follows from the fact that

\[
S = \frac{1}{n} \sum_{j=1}^{N} (x_j - \bar{x})(x_j - \bar{x})'
\]

\[
= \frac{1}{n} \left[ \sum_{j=1}^{N} (x_j - \mu)(x_j - \mu)' - N(\bar{x} - \mu)(\bar{x} - \mu)' \right]
\]

\[
= \frac{N}{n} S^* - \frac{N}{n} (\bar{x} - \mu)(\bar{x} - \mu)'
\]

where the second term goes to zero in probability for large \( N \), and \( (N/n) \approx 1 \). Thus, from model (1.1), for large \( N \), \( S \) is approximated by

\[
S^* = \frac{1}{N} \sum_{j=1}^{N} (x_i - \mu)(x_i - \mu)'
\]

\[
= \frac{1}{N} C \left( \sum_{j=1}^{N} z_j z_j' \right) C',
\]

where \( z_j = (z_{1j}, \ldots, z_{pj})' \) and \( z_j \) are iid with mean 0 and variance 1 etc. as given in (1.3). Let \( C' = (c_1, \ldots, c_p) \), and \( A_i = c_i c_i' = (a_{ij}(i)) \). Then

\[
s_{11}^* = \frac{1}{N} c_1 \left( \sum_{j=1}^{N} z_j z_j' \right) c_1 = \frac{1}{N} \sum_{j=1}^{N} z_j A_1 z_j
\]

and

\[
s_{22}^* = \frac{1}{N} c_2 \left( \sum_{j=1}^{N} z_j z_j' \right) c_2 = \frac{1}{N} \sum_{j=1}^{N} z_j A_2 z_j.
\]

Hence

\[
E(s_{ii}^*) = \text{tr} A_i = c_i' c_i = \sigma_{ii},
\]

since

\[
\Sigma = (\sigma_{ij}) = CC'.
\]

From Lemma 2.1,

\[
\text{Var}(s_{ii}^*) = \frac{1}{N} \left[ (\gamma - 3) \sum_{j=1}^{p} \sigma_{jj}(i) + 2\sigma_{ii}^2 \right].
\]

Note that

\[
A_i = c_i c_i' = \begin{pmatrix} c_{11} \\ \vdots \\ c_{pi} \end{pmatrix} (c_{11}, \ldots, c_{pi}).
\]

Hence

\[
\sum_{j=1}^{p} \sigma_{jj}(1) = \sum_{j=1}^{p} \sigma_{jj}^A = \left( \sum_{j=1}^{p} \sigma_{jj}^A \right)^2 = \sigma_{11}^2.
\]

Thus,

\[
\text{Var}(s_{ii}^*) = O(N^{-1}).
\]

Again, from Lemma 2.1(d),

\[
\text{Cov}(s_{11}^*, s_{22}^*) = \frac{1}{N} \left[ (\gamma - 3) \sum_{j=1}^{p} \sigma_{jj}(1) \sigma_{jj}(2) + 2\text{tr} A_1 A_2 \right]
\]

\[
= \frac{1}{N} \left[ (\gamma - 3) \sum_{j=1}^{p} \sigma_{jj}(1) \sigma_{jj}(2) + 2\sigma_{12}^2 \right].
\]
Let

\[ \sum_{j=1}^{p} a_{ij}(1) a_{ij}(2) = \sum_{j=1}^{p} c_{ij}^2 c_{j2}^2 \leq \left( \sum_{j=1}^{p} c_{ij}^2 \right) \left( \sum_{j=1}^{p} c_{j2}^2 \right) \leq \sigma_{11} \sigma_{22}, \]

it follows that

\[ \text{Cov}(s_{11}^*, s_{22}^*) = O(N^{-1}). \]

Thus, we get the following theorem.

**Theorem 2.3.** Let

\[ S^* = (s_{ij}^*) = \frac{1}{N} C \left( \sum_{j=1}^{N} z_j z'_j \right) C', \]

where \( C = (c_1, \ldots, c_i), \quad C' = (c_1, \ldots, c_i), \quad \Sigma = (\sigma_{ij}) > 0, \) and \( z_j = (z_{ij}, \ldots, z_{ip}), \) \( z_j \) iid with mean 0, variance 1 and fourth moment \( \gamma. \) Then with \( A_i = (a_{id}(i)) = c_i c'_i, \)

(a) \( E(s_{ii}^*) = \sigma_{ii}. \)

(b) \( \text{Var}(s_{ii}^*) = \frac{1}{N} \left[ \nu - 3 \sum_{j=1}^{p} a_{ij}^2(1) + 2 \sigma_{ii}^2 \right] = O(N^{-1}). \)

(c) \( \text{Cov}(s_{11}^*, s_{22}^*) = \frac{1}{N} \left[ (\nu - 3) \sum_{j=1}^{p} a_{ij}(1) a_{ij}(2) + 2 \sigma_{12}^2 \right] = O(N^{-1}). \)

Next, we study the properties of \( (s_{ii}^*)^{-1} \). From Taylor’s expansion, it follows that

\[ (s_{11}^*)^{-1} = \frac{1}{\sigma_{11}} - \frac{s_{11}^* - \sigma_{11}}{\sigma_{11}^2} + \frac{(s_{11}^* - \sigma_{11})^2}{(\sigma_{11})^3} + O_p \left( N^{-1} \right). \]

Hence

\[ E(s_{11}^{-1}) = \frac{1}{\sigma_{11}} + O \left( N^{-1} \right), \]

\[ \text{Var}(s_{11}^{-1}) = \frac{\text{Var}(s_{11}^*)}{\sigma_{11}^4} + O \left( N^{-2} \right), \]

and

\[ \text{Cov}(s_{11}^{-1}, s_{22}^{-1}) = \frac{1}{\sigma_{11}^2 \sigma_{22}} \text{Cov}(s_{11}^*, s_{22}^*) + O \left( N^{-2} \right). \]

Thus, we have the following corollary.

**Corollary 2.6.** In the notation of **Theorem 2.3**

\[ E \left[ (s_{11}^*)^{-1} \right] = \sigma_{11}^{-1} + O \left( N^{-1} \right), \]

\[ \text{Var} \left[ (s_{11}^*)^{-1} \right] = \frac{\text{Var}(s_{11}^*)}{\sigma_{11}^4} + O \left( N^{-2} \right), \]

and

\[ \text{Cov}(s_{11}^{-1}, s_{22}^{-1}) = \frac{1}{\sigma_{11}^2 \sigma_{22}} \text{Cov}(s_{11}^*, s_{22}^*) + O \left( N^{-2} \right). \]

In **Theorem 2.2**, the consistency of the estimator \( \hat{\delta}_2 \) of the quantity \( \delta_2 \) is stated. We shall now consider the corresponding result for the correlation matrix

\[ \mathcal{R} = D_{1/2} \Sigma D_{1/2}, \]

where \( \Sigma = (\sigma_{ij}) \) and \( D_{1/2} = \text{diag}(\sigma_{11}^{1/2}, \ldots, \sigma_{pp}^{1/2}). \) The sample correlation matrix is given by

\[ R = D_{1/2}^{-1} S D_{1/2}^{-1/2}. \]
where \( S = (s_{ij}) \) and \( D_1^2 = \text{diag}(s_{11}, \ldots, s_{pp}) \). Following Srivastava and Du [8], it can be shown that

\[
\frac{1}{p} \left[ \text{tr} R^2 - \frac{p^2}{n} \right]
\]

is a consistent estimator of \( \frac{1}{p} \text{tr} \mathcal{R}^2 \) for model (1.1) since from Theorem 2.2, \( \hat{\delta}_1 \) and \( \hat{\delta}_2 \) are consistent estimators of \( \delta_1 \) and \( \delta_2 \) respectively. Then, we have the following lemma.

**Lemma 2.5.** As \((n, p) \to \infty\), \( \frac{1}{p} [\text{tr} R^2 - \left( \frac{p^2}{n} \right)] \) is a consistent estimator of \( \frac{1}{p} \text{tr} \mathcal{R}^2 \) under the condition (2.2), and \( n = O(p^\epsilon), \ \frac{1}{2} < \epsilon \leq 1 \).

### 2.4. Some algebraic inequalities

**Lemma 2.6.** Let \( A = C'C \) be a \( p \times p \) symmetric and positive semi-definite matrix, where \( C' = (c_{ij}) \). Then

(a) \( a_{ii} = \sum_{j=1}^{p} c_{ij}^2, \quad i = 1, \ldots, p \).

(b) \( \text{tr} (C'C)^2 = \sum_{k=1}^{p} \sum_{i=1}^{p} \left( \sum_{j=1}^{p} c_{kj} c_{ij} \right)^2 \).

(c) \( \sum_{i=1}^{p} a_{ii}^2 = \sum_{i=1}^{p} \left( \sum_{j=1}^{p} c_{ij}^2 \right)^2 \leq \text{tr} (C'C)^2 \).

**Corollary 2.7.** Let \( B = (b_{ij}) \) be a \( p \times p \) symmetric matrix. Then

(a) \( \text{tr} B^4 = \sum_{k=1}^{p} \sum_{i=1}^{p} \left( \sum_{j=1}^{p} b_{kj} b_{ij} \right)^2 = \sum_{k=1}^{p} \sum_{i=1}^{p} \left( \sum_{j=1}^{p} b_{kj} b_{ji} \right)^2 \).

(b) \( \sum_{i \neq j} b_{ij}^4 \leq \sum_{i=1}^{p} \sum_{j=1}^{p} b_{ij}^4 \leq \sum_{i=1}^{p} \left( \sum_{j=1}^{p} b_{ij}^2 \right)^2 \)

\[ \leq \sum_{i=1}^{p} \left( \sum_{j=1}^{p} b_{ij}^2 \right)^2 + \sum_{k=1}^{p} \sum_{i=1}^{p} \sum_{j=1}^{p} b_{ij} b_{kj}^2 \]

\[ = \sum_{k=1}^{p} \sum_{i=1}^{p} \left( \sum_{j=1}^{p} b_{kj} b_{ij} \right) = \text{tr} B^4. \]

### 3. Distribution of the test statistic \( T \) under the hypothesis that \( \mu = 0 \)

In this section, we derive the distribution of the test statistic \( T \) defined in (1.6) under the hypothesis \( H : \mu = 0 \). We shall assume that

\[
n = O(p^\epsilon), \quad 0 < \epsilon \leq 1,
\]

and

\[
\lim_{p \to \infty} \tau_i \left( \frac{\text{tr} \mathcal{R}_i}{p} \right) = \tau_{0i} < \infty, \quad i = 1, \ldots, 4.
\]

From Lemma 2.5, it follows that under the condition (3.2)

\[
\frac{1}{p} \left[ \text{tr} R^2 - \frac{1}{n} p^2 \right] \to \tau_{20}.
\]
and \( n/(n - 2) \to 1 \) as \((n, p) \to \infty \). Also, from Corollary 2.6, \( s_{ii}^{-1} \) are consistent estimators of \( \sigma_{ii}^{-1} \). Thus, we need only consider the asymptotic distribution of the test statistic

\[
T = \frac{N\bar{X}D_{\sigma}^{-1}\bar{X} - p}{\sqrt{2p\tau_2}}
= \frac{N\bar{X}D_{\sigma}^{-1}\bar{X} - p}{\sqrt{2p\tau_2}}
\]

where \( B = C'D_{\sigma}^{-1}C, D_{\sigma} = \text{diag}(\sigma_{11}, \ldots, \sigma_{pp}) \) and \( \bar{X} = C\bar{X} \). From Theorem 2.1, it follows that it has a normal distribution. Thus, we have the following theorem.

**Theorem 3.1.** Under the condition (3.1) and (3.2) and in the notation of (1.4)–(1.9),

\[
\lim_{(n, p) \to \infty} \left[ P_0 \left[ \left\{ \frac{N\bar{X}D_{\sigma}^{-1}\bar{X} - \left( \frac{p}{n-2} \right) p}{2(\text{tr} R^2 - \frac{1}{n}p^2)} \right\} \leq \zeta \right] = \Phi(\zeta),
\]

where \( \Phi(\zeta) \) denotes the standard normal cdf, and \( P_0 \) denotes that it has been computed under the null hypothesis.

4. Distribution of the statistic \( T \) under local alternatives

For the asymptotic distribution of the test statistic \( T \) when \( \mu \neq 0 \), we shall consider the local alternative under which

\[
\mu = \left( \frac{1}{nN} \right)^{\frac{1}{2}} \delta,
\]

where \( \delta \) is a vector of constants. We shall assume that

\[
\frac{\delta'D_{\sigma}^{-1}\delta}{p} \leq M,
\]

for every \( p \), where \( M \) does not depend on \( p \). From Theorem 2.1, we know that asymptotically as \((n, p) \to \infty \),

\[
T = \frac{N(\bar{X} - \mu)'D_{\sigma}^{-1}(\bar{X} - \mu) - p}{\sqrt{2p\text{tr} R^2}}
\]

has a standard normal distribution, \( N(0, 1) \). For the local alternative \( \mu = (nN)^{-\frac{1}{2}} \delta \),

\[
\frac{1}{\sqrt{p}} \left[ N(\bar{X} - \mu)'D_{\sigma}^{-1}(\bar{X} - \mu) \right] \sim \frac{1}{\sqrt{p}} \left[ \frac{n\bar{X}D_{\sigma}^{-1}\bar{X} - \delta'D_{\sigma}^{-1}\delta}{n\sqrt{2p}} \right]
\]

and under conditions (4.1) and (4.2)

\[
\left[ \frac{n\bar{X}D_{\sigma}^{-1}\bar{X}}{n\sqrt{p}} \right] \overset{p}{\to} \frac{\delta'D_{\sigma}^{-1}\delta}{n\sqrt{p}}.
\]

Hence,

\[
\frac{1}{\sqrt{p}} \left[ N(\bar{X} - \mu)'D_{\sigma}^{-1}(\bar{X} - \mu) \right] \overset{p}{\to} \frac{1}{\sqrt{p}} \left[ \frac{n\bar{X}D_{\sigma}^{-1}\bar{X} - \delta'D_{\sigma}^{-1}\delta}{n} \right].
\]

Thus, we get the following theorem.

**Theorem 4.1.** Under the conditions of Theorem 3.1, and (4.1) and (4.2)

\[
\lim_{(n, p) \to \infty} \left[ P_1(T < z) = \Phi(z + \frac{\delta'D_{\sigma}^{-1}\delta}{n\sqrt{2p\text{tr} R^2}}) \right] = 0.
\]

5. Proofs of lemmas and theorems in Section 2

In this section, we give proofs of some of the results stated in Section 2.
5.1. Proof of Theorem 2.1

Let
\[ q = \frac{1}{\sqrt{p}} \left[ N \z^2 A \z - \text{tr} A \right] \]
\[ = \frac{1}{\sqrt{p}} \sum_{i=1}^{p} a_{ii} (Nz_i^2 - 1) + \frac{2N}{\sqrt{p}} \sum_{i<j} a_{ij} \z_i \z_j. \]

Since, from Assumption (i)
\[ \lim_{p \to \infty} \max_{1 \leq i \leq p} \left( \frac{a_{ii}^2}{p} \right) = 0, \]
it follows from Lemma 2.4 and Corollary 2.1 that
\[ q_1 = \frac{1}{\sqrt{p}} \sum_{i=1}^{p} a_{ii} (Nz_i^2 - 1) \to N(0, \sigma_1^2), \]
where
\[ \sigma_1^2 = 2 \lim_{p \to \infty} \left( \frac{\sum_{i=1}^{p} a_{ii}^2}{p} \right). \]

It can be verified that
\[ \text{Cov} \left[ \frac{1}{\sqrt{p}} \sum_{i=1}^{p} a_{ii} (Nz_i^2 - 1), \frac{N}{\sqrt{p}} \sum_{i<j} a_{ij} \z_i \z_j \right] = 0. \]

Thus, it remains to prove the asymptotic normality of the second term
\[ q_2 = \frac{2N}{\sqrt{p}} \sum_{i<j} a_{ij} \z_i \z_j. \]

Note that
\[ 2N \sum_{i<j} a_{ij} \z_i \z_j / \sqrt{p} = 2N \sum_{j=2}^{p} z_j \left( \sum_{i=1}^{j-1} a_{ij} \z_i \right) / \sqrt{p}. \]

Let
\[ \eta_j = Nz_j \left( \sum_{i=1}^{j-1} a_{ij} \z_i \right) / \sqrt{p}, \]
and let \( \mathcal{F}_j \) be the \( \sigma \)-algebra generated by the random variables \( \z_1, \ldots, \z_j \). Letting \( \mathcal{Z}_0 = 0 \), and \( \mathcal{F}_0 = (\phi, \Omega) = \mathcal{F}_{-1} \), where \( \phi \) is the empty set, we find that \( \mathcal{F}_0 \subset \mathcal{F}_1 \subset \cdots \subset \mathcal{F}_p \subset \mathcal{F} \), and
\[ E(\eta_j|\mathcal{F}_{j-1}) = 0, \quad \text{and} \quad E(\eta_j^2|\mathcal{F}_{j-1}) = \frac{N}{p} \left( \sum_{i=1}^{j-1} a_{ij} \z_i \right)^2. \]

Also,
\[ E(\eta_j^2) = \frac{N}{p} E \left( \sum_{i=1}^{j-1} a_{ij} \z_i \right)^2 \]
\[ = \frac{1}{p} \sum_{i=1}^{j-1} a_{ij}^2 < \infty. \]

Hence, the sequence \( \{\eta_k, \mathcal{F}_k\} \) is a sequence of square integrable martingale difference. Thus, to establish the asymptotic normality of the random variable \( q_2 \), we may use Theorem 4 from [6], pp. 511. This requires establishing the Lindeberg condition: for \( \epsilon > 0 \),

\[(i) L = \sum_{k=0}^{p} E \left[ \eta_k^2 I(|\eta_k| > \epsilon)|\mathcal{F}_{k-1} \right] \overset{p}{\to} 0,\]
and showing that

\[(ii) \quad G = \sum_{k=0}^{p} E \left( \eta_{k}^{2} | \mathcal{F}_{k-1} \right) \xrightarrow{p} \sigma_{20}^{2},\]

for some finite constant $\sigma_{20}^{2}$. We first show (ii). We have

\[
\sum_{j=0}^{p} E(\eta_{j}^{2} | \mathcal{F}_{j-1}) = N \frac{1}{p} \sum_{j=2}^{p} \left( \sum_{i=1}^{j-1} a_{j,i} \bar{z}_{i} \right)^{2} \equiv N \frac{1}{p} \sum_{j=2}^{p} \left[ \sum_{i=1}^{j-1} a_{j,i}^{2} \bar{z}_{i}^{2} + 2 \sum_{k<i}^{j-1} a_{j,k} a_{j,i} \bar{z}_{k} \bar{z}_{i} \right].
\]

Hence,

\[
E \left[ \sum_{j=0}^{p} E(\eta_{j}^{2} | \mathcal{F}_{j-1}) \right] = \frac{1}{p} \sum_{j=2}^{p} \sum_{i=1}^{j-1} a_{j,i}^{2} = \frac{1}{p} \sum_{i=1}^{p} a_{j,i}^{2} = \sigma_{2}^{2}.
\]

Note that

\[
\sigma_{2}^{2} = \frac{1}{2} \left[ \frac{1}{p} \sum_{i \neq j} a_{j,i}^{2} \right] = \frac{1}{2} \left[ \frac{\text{tr} A^{2}}{p} - \frac{\sum a_{i}^{2}}{p} \right] \leq \frac{1}{2} a_{2}.
\]

Thus, from Assumption (ii)

\[
\lim_{p \to \infty} \sigma_{2}^{2} = \sigma_{20}^{2} < \infty.
\]

In order to show that

\[
\sum_{j=0}^{p} E \left( \eta_{j}^{2} | \mathcal{F}_{j-1} \right) \xrightarrow{p} \sigma_{20}^{2},
\]

we need to show that its variance goes to zero. That is

\[
\nu^{2} = \text{Var} \left[ \frac{N}{p} \sum_{j=2}^{p} \left( \sum_{i=1}^{j-1} a_{j,i}^{2} \bar{z}_{i}^{2} + 2 \sum_{k<i}^{j-1} a_{j,k} a_{j,i} \bar{z}_{k} \bar{z}_{i} \right) \right]
\]

goes to zero. That is the variance of each of the two terms goes to zero. We find that the variance of the first term

\[
= \text{Var} \left[ \frac{N}{p} \sum_{j=2}^{p} \sum_{i=1}^{j-1} a_{j,i}^{2} \bar{z}_{i}^{2} / p \right]
\]

\[
= \sum_{i=1}^{p-1} \left( \sum_{j=i+1}^{p} a_{j,i}^{2} \right) \left[ \frac{\nu - 3}{N} + 2 \right] / p^{2}
\]

\[
\leq \left( \text{tr} A^{2} / p \right) \left( 2 + \frac{\nu - 3}{N} \right) / p \to 0 \quad \text{as} \quad (N, p) \to \infty.
\]

In order to show that the variance of the second term goes to zero, we need to show that

\[
\text{Var} \left[ \frac{N}{p} \sum_{j=2}^{p} \sum_{1 \leq k < l}^{j-1} a_{j,k} a_{j,l} \bar{z}_{k} \bar{z}_{l} \right] = (N/p)^{2} \text{Var} \left[ \sum_{1 \leq k < l}^{p-1} \left( \sum_{j=l+1}^{p} a_{j,k} a_{j,l} \right) \bar{z}_{k} \bar{z}_{l} \right]
\]

\[
= p^{-2} \sum_{1 \leq k < l}^{p-1} \left( \sum_{j=l+1}^{p} a_{j,k} a_{j,l} \right)^{2}
\]

\[
\leq p^{-2} \sum_{1 \leq k < l}^{p} \left( \sum_{j=l+1}^{p} |a_{j,k}| |a_{j,l}| \right)^{2}
\]

\[
\leq p^{-2} \sum_{1 \leq k < l}^{p} \left( \sum_{j=1}^{p} |a_{j,k}| |a_{j,l}| \right)^{2}
\]

\[
\leq p^{-1} (\text{tr} A^{2} / p),
\]

\[
\]

527
from Corollary 2.7. Since, it has been assumed that \( \text{tr} A^4 / p \) is bounded, the variance of the second term also goes to zero. Hence, \( v^2 \to 0 \) as \((N, p) \to \infty\), and thus from Chebyshev’s inequality

\[
\sum_{j=0}^{\infty} E \left[ \eta_j^2 | \mathcal{F}_{j-1} \right] \overset{p}{\to} \sigma^2_{20}.
\]

Next, we verify the Lindeberg condition:

\[
L = \sum_{k=0}^{p} E \left[ \eta_k^2 I(|\eta_k| > \varepsilon) | \mathcal{F}_{k-1} \right] \overset{p}{\to} 0.
\]

From Chebyshev’s inequality

\[
P[L > \xi] \leq \frac{E(L^2)}{\xi^2} = \frac{\text{Var}(L) + E^2(L)}{\xi^2}
\]

\[
\leq \frac{v^2 + \left( \sum_{k=0}^{p} E \left[ \eta_k^2 I(|\eta_k| > \varepsilon) \right] \right)^2}{\xi^2}
\]

\[
\leq \frac{v^2 + \left[ \sum_{k=0}^{p} E \left[ (\eta_k^2)^2 (P(\eta_k^2 > \varepsilon^2)) \right] \right]^2}{\xi^2}
\]

\[
\leq \frac{\varepsilon^4 v^2 + \left[ \sum_{k=0}^{p} E(\eta_k^4) \right]^2}{\varepsilon^4 \xi^2}
\]

\[
= \frac{\varepsilon^4 v^2}{\varepsilon^4 \xi^2}.
\]

Since \( v^2 \to 0 \), we need only show that

\[
\sum_{k=0}^{p} E(\eta_k^4) \to 0 \quad \text{as} \quad p \to \infty
\]

for the Lindeberg condition to hold. Now

\[
\eta_k = N z_k \left( \sum_{j=1}^{k-1} a_{kj} z_j \right) / \sqrt{p},
\]

\[
\eta_k^4 = N^4 p^{-2} z_k^4 \left[ \sum_{j=1}^{k-1} a_{kj}^2 z_j^2 + 2 \sum_{j<l} a_{kj} a_{kl} z_j z_l \right]^2
\]

\[
= N^4 p^{-2} z_k^4 \left[ \left( \sum_{j=1}^{k-1} a_{kj}^2 z_j^2 \right)^2 + 4 \left( \sum_{j=1}^{k-1} a_{kj}^2 z_j^2 \right) \left( \sum_{j<m} a_{kj} a_{km} z_j z_m \right) + \left( \sum_{j<l} a_{kj} a_{kl} z_j z_l \right) \right].
\]

Hence, for \( \xi_{0,N} = \frac{\varepsilon^2}{N} + 3 \)

\[
E(\eta_k^4) = \left( \frac{1}{p} \right)^2 \left[ \frac{\gamma - 3}{N} + 3 \right] \left[ \sum_{j=1}^{k-1} \xi_{0,N} a_{kj}^4 + 6 \sum_{j<l} a_{kj}^2 a_{kl}^2 \right]
\]

\[
E(\eta_k^4) \leq p^{-2} \xi_{0,N} \left[ \sum_{j=1}^{k-1} a_{kj}^4 + 2 \sum_{j<l} a_{kj}^2 a_{kl}^2 \right]
\]

\[
\leq p^{-2} \xi_{0,N} \left( \sum_{j=1}^{k-1} a_{kj}^2 \right)^2
\]

and

\[
\sum_{k=0}^{p} E(\eta_k^4) \leq p^{-2} \xi_{0,N} \sum_{k=0}^{p} \left( \sum_{j=0}^{k-1} a_{kj}^2 \right)^2
\]

\[
= (\xi_{0,N} / p) \left( \text{tr} A^4 / p \right),
\]
which goes to zero as $(N, p) \to \infty$. Thus, the Lindeberg condition is satisfied, and
\[
\frac{2N}{\sqrt{p}} \sum_{i<j} a_{ij} \bar{z}_i \bar{z}_j \to N(0, 4\sigma^2).
\]
Combining the two terms, we get
\[
q \overset{d}{\to} N(0, \sigma^2),
\]
where
\[
\sigma^2 = \lim_{p \to \infty} \frac{1}{p} \left[ 2 \sum_{i=1}^{p} a_{ii}^2 + 2 \sum_{i \neq j} a_{ij}^2 \right] = \lim_{p \to \infty} (2 \text{tr} A^2 / p).
\]

5.2. Proof of Theorem 2.2

Let
\[
\mathbf{u}_i = \mathbf{x}_i - \mathbf{\mu}, \quad \text{and} \quad \bar{\mathbf{u}} = N^{-1} \sum_{i=1}^{N} \mathbf{u}_i.
\]
Then
\[
\mathbf{u}_i = C\mathbf{z}_i,
\]
where $\mathbf{z}_1, \ldots, \mathbf{z}_N$ are iid with mean vector $\mathbf{0}$ and covariance matrix $I_p$. It is assumed that the $p$-components of $\mathbf{z}_i$ are independently distributed. Since $N/n \to 1$ as $N \to \infty$, and $\bar{\mathbf{u}} \bar{\mathbf{u}}' \to 0$ in probability as $N \to \infty$ for all values of $p$, we may replace
\[
S = \frac{1}{n} \sum_{i=1}^{N} (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})'
\]
by
\[
S^* = \frac{1}{N} \sum_{i=1}^{N} \mathbf{u}_i \mathbf{u}_i' = \frac{1}{N} \sum_{i=1}^{N} C\mathbf{z}_i \mathbf{z}_i' C'
\]
for theoretical purposes. Thus,
\[
\hat{\delta}_1 = \frac{1}{Np} \sum_{i=1}^{N} \text{tr} C\mathbf{z}_i \mathbf{z}_i' C' = \frac{1}{Np} \sum_{i=1}^{N} \mathbf{z}_i' B \mathbf{z}_i,
\]
where
\[
B = (b_{ij}) = C'C : p \times p.
\]
Hence,
\[
E(\hat{\delta}_1) = \delta_1
\]
and, from Lemma 2.1
\[
\text{Var}(\hat{\delta}_1) = \frac{1}{Np^2} \left[ (\gamma - 3) \sum_{i=1}^{p} b_{ii}^2 + 2\text{tr} B^2 \right].
\]
Thus, as $(n, p) \to \infty$, $\text{Var}(\hat{\delta}_1) \to 0$ and $\hat{\delta}_1 \overset{p}{\to} \delta_1$. 

Next, with $B = (b_{ij}) = C'C$ and $NS^* = \sum_{i=1}^{p}(Cz_i'z_i')$, we have

$$
\hat{\delta}_2^2 = \frac{1}{p} \left[ \text{tr} S' - \frac{1}{N} (\text{tr} S')^2 \right]
$$

$$
= \frac{1}{N^2p} \left[ \text{tr} \left( \sum_{i=1}^{N} Cz_i'z_i' \right)^2 - \frac{1}{N} \left( \text{tr} \sum_{i=1}^{N} Cz_i'z_i' \right)^2 \right]
$$

$$
= \frac{1}{N^2p} \left[ \sum_{i,j} (z_i'Bz_i - \text{tr} B)^2 - \frac{1}{N} \left( \sum_{i=1}^{N} z_i'Bz_i - \text{tr} B \right)^2 \right]
$$

$$
= \frac{1}{N^2p} \left[ \sum_{i=1}^{N} (z_i'Bz_i - \text{tr} B)^2 - \frac{1}{N} \left( \sum_{i=1}^{N} z_i'Bz_i - \text{tr} B \right)^2 \right]
$$

$$
= \frac{1}{N^2p} \left[ \sum_{i=1}^{N} (z_i'Bz_i - \text{tr} B)^2 - \frac{1}{N} \left( \sum_{i=1}^{N} (z_i'Bz_i) - \text{tr} B \right)^2 \right]
$$

$$
= \frac{N(N-1)}{N^2} \left( \frac{\text{tr} B^2}{p} \right) + q_1 + q_2 + q_3,
$$

where, for example,

$$
q_1 = \frac{1}{N^2p} \sum_{i=1}^{N} (z_i'Bz_i - \text{tr} B)^2
$$

with mean (see Lemma 2.1)

$$
E(q_1) = \frac{1}{Np} E \left[ (z_i'Bz_i - \text{tr} B)^2 \right]
$$

$$
= \frac{1}{N} \left[ (\gamma - 3) \left( \sum_{i=1}^{p} b_i^2 \right) \right] + 2(\text{tr} B^2/p)
$$

which goes to zero as $(N, p) \to \infty$. Hence, from Markov's inequality $q_1 \to 0$ in probability as $(N, p) \to \infty$. Next, we consider

$$
-q_2 = \frac{1}{N^3} \left[ \sum_{i=1}^{N} (z_i'Bz_i - \text{tr} B)/\sqrt{p} \right]^2
$$

$$
= \frac{1}{N^3} \left( \sum_{i=1}^{N} u_i \right)^2,
$$

where $u_i = (z_i'Bz_i - \text{tr} B)/\sqrt{p}$ are iid random variables with mean 0 and finite variance (see Lemma 2.1),

$$
\xi_0 = \left[ (\gamma - 3) \left( \sum_{i=1}^{p} b_i^2 \right) \right] + 2(\text{tr} B^2/p) < \infty.
$$

Hence, from the Marcinkiewicz and Zygmund [4] inequality, as given in [5, p. 304],

$$
E \left[ \sum_{i=1}^{N} u_i \right]^2 = O(N).
$$

Thus, from Markov's inequality

$$
P \left\{ \frac{1}{N^3p} \left[ \sum_{i=1}^{N} (z_i'Bz_i - \text{tr} B)^2 \right] > \epsilon \right\} = O(N^{-2})
which goes to zero as \((N, p) \to \infty\). Hence \(q_2 \to 0\) in probability as \((N, p) \to \infty\). Finally, we consider the third term

\[
q_3 = \frac{2}{N^2p} \sum_{i<j} \left( (z_i^t B z_j)^2 - \text{tr} B^2 \right).
\]

From Lemma 2.2

\[
\text{Var}(q_3) = \frac{2N(N-1)}{N^4p^2} \left[ (\gamma - 3)^2 \sum_{i=1}^{p} \sum_{j=1}^{p} b_{ij}^4 + 6(\gamma - 3) \left( \sum_{i=1}^{p} b_{ii}^2 \right)^2 + 6 \text{tr} B^4 + 2(\text{tr} B^2)^2 \right]
\leq \frac{2(N-1)}{N^3} \left[ (\gamma - 3)^2 \frac{\text{tr} B^4}{p^2} + 6(\gamma - 3) \frac{\text{tr} B^4}{p^2} + 6 \frac{\text{tr} B^4}{p^2} + 2 \left( \frac{\text{tr} B^2}{p} \right)^2 \right]
\to \frac{4}{N^2} \left( \frac{\text{tr} B^2}{p} \right)^2
\]

which goes to zero as \((N, p) \to \infty\). Hence \(q_3 \to 0\) in probability as \((N, p) \to \infty\). Thus, \(\delta_2 \xrightarrow{p} (\text{tr} B^2/p)\).

5.3. Proof of Lemma 2.6

Since

\[
A = (a_{ij}) = C'C = \begin{pmatrix} C_{11} & C_{12} & \cdots & C_{1p} \\ C_{21} & C_{22} & \cdots & C_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ C_{p1} & C_{p2} & \cdots & C_{pp} \end{pmatrix},
\]

\[
\text{tr} (C'C)^2 = \text{tr} \left( \sum_{i=1}^{p} \sum_{j=1}^{p} C_{ij}^2 \right) = \sum_{i=1}^{p} \sum_{j=1}^{p} C_{ij}^2,
\]

The diagonal elements of \((C'C)^2\) are given by

\[
\left( \sum_{j=1}^{p} C_{ij}^2 \right)^2 + \sum_{i=2}^{p} \left( \sum_{j=1}^{p} C_{ij}^2 \right)^2 + \left( \sum_{j=1}^{p} C_{ij}^2 \right)^2 + \sum_{i=1}^{p} \left( \sum_{j=1}^{p} C_{ij}^2 \right)^2,
\]

and \(\sum_{i=1}^{p} \left( \sum_{j=1}^{p} C_{ij}^2 \right)^2\) respectively. Hence

\[
\text{tr} (C'C)^2 = \sum_{i=1}^{p} \left( \sum_{j=1}^{p} C_{ij}^2 \right)^2 + \sum_{i=1}^{p} \left( \sum_{j=1}^{p} C_{ij}^2 \right)^2 + \cdots + \sum_{i=1}^{p} \left( \sum_{j=1}^{p} C_{ij}^2 \right)^2
\]

\[
= \sum_{k=1}^{p} \sum_{i=1}^{p} \left( \sum_{j=1}^{p} C_{ij}^2 \right)^2.
\]

Also,

\[
a_{11} = \sum_{j=1}^{p} C_{ij}^2, \ldots, a_{pp} = \sum_{i=1}^{p} C_{ij}^2.
\]

Hence,

\[
\sum_{i=1}^{p} a_{ii} = \sum_{i=1}^{p} \left( \sum_{j=1}^{p} C_{ij}^2 \right)^2 \leq \sum_{i=1}^{p} \left( \sum_{j=1}^{p} C_{ij}^2 \right)^2 + \sum_{i=1}^{p} \sum_{k=1}^{p} \left( \sum_{j=1}^{p} C_{ij}^2 \right)^2 = \text{tr} (C'C)^2.
\]
Acknowledgments

I wish to sincerely thank Dr. Murari Singh and the two reviewers for their careful reading of this article and offering suggestions. This research was supported by the Natural Sciences and Engineering Research Council of Canada, NSERC.

References