CIRJE-F-831

Tests for Multivariate Analysis of Variance in High Dimension Under Non-Normality

Muni S. Srivastava
University of Toronto

Tatsuya Kubokawa
University of Tokyo

December 2011; Revised in January 2012

Tests for Multivariate Analysis of Variance in High Dimension Under Non-Normality

Muni S. Srivastava, and Tatsuya Kubokawa
University of Toronto, and University of Tokyo

ABSTRACT

In this article, we consider the problem of testing the equality of mean vectors of dimension p of several groups with a common unknown non-singular covariance matrix $\mathbf{\Sigma}$, based on N independent observation vectors where N may be less than the dimension p. This problem, known in the literature as the Multivariate Analysis of variance (MANOVA) in high-dimension has recently been considered in the statistical literature by Srivastava and Fujikoshi[7], Srivastava [5] and Schott[3]. All these tests are not invariant under the change of units of measurements. On the lines of Srivastava and Du[8] and Srivastava[6], we propose a test that has the above invariance property. The null and the non-null distributions are derived under the assumption that $(N, p) \to \infty$ and N may be less than p and the observation vectors follow a general non-normal model.

Keywords and phrases: Asymptotic distributions, high dimension, MANOVA, multivariate linear model, non-normal model, sample size smaller than dimension.

AMS 1991 subject classification: primary 62H15, Secondary 62F05

1 Introduction

The problem of testing the equality of mean vectors of several groups with common unknown nonsingular covariance matrix, the so called MANOVA or multivariate analysis of variance has been considered many times in the statistical literature. For normally distributed observation vectors when the total sample size N is considerably larger than the dimension p of the vector, Wilks[9] likelihood ratio test is commonly used with Box's[2] approximation for the distribution of the test statistic. For dimension p larger than the sample size N, this testing problem has also been recently considered in the literature by Srivastava and Fujikoshi[7], Srivastava[5], and Schott[3] for normally distributed observation vectors.

In this article, we consider a general model which includes normal distributions and propose a test that is invariant under the change of units of measurements. That is, the test statistic is invariant under the transformation by non singular diagonal matrices. Thus, without any loss of generality, we assume that the covariance matrix is a correlation matrix $\mathbf{\Lambda} = \mathbf{\Lambda}^{1/2}\mathbf{\Lambda}^{1/2}$, where $\mathbf{\Lambda}^{1/2}$ is the unique positive definite matrix. Since the MANOVA problem is a special case of the multivariate regression model, we assume that the $N \times p$ matrix of observations follow the model

$$\boldsymbol{Y} = \boldsymbol{X}\mathbf{\Theta} + \boldsymbol{U}\mathbf{\Lambda}^{1/2} \tag{1.1}$$

where $\boldsymbol{X}$ is an $N \times k$ matrix of known constants of rank k, $\mathbf{\Theta}$ is a $k \times p$ matrix of unknown parameters, $k \leq p$,

$$\boldsymbol{U} = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_N)',$$

and $\boldsymbol{u}_i = (u_{i1}, \ldots, u_{ip})'$ are independent and identically distributed with

$$E(\boldsymbol{u}_i) = \mathbf{0}, \ \mathbf{Cov}\,(\boldsymbol{u}_i) = \boldsymbol{I}_p, \ E(u_{ik}^4) = K_4 + 3, \tag{1.2}$$

and for $\nu_k \geq 0, \sum_{k=1}^{p} \nu_k \leq 4, i = 1, \ldots, N$,

$$E\Big[\prod_{k=1}^{p} u_{ik}^{\nu_k}\Big] = \prod_{k=1}^{p} E(u_{ik}^{\nu_k}). \tag{1.3}$$

Here $\mathbf{\Lambda} = (\lambda_{ij}) = \mathbf{\Lambda}^{1/2}\mathbf{\Lambda}^{1/2}$ is the non-singular correlation matrix. For normally distributed $\boldsymbol{u}_i$ with zero mean vector and identity covariance matrix, the conditions (1.2)-(1.3) are satisfied with $K_4 = 0$.

The problem of testing in the model (1.1) is that of testing the hypothesis

$$H : \boldsymbol{C}\mathbf{\Theta} = \mathbf{0} \text{ vs } A : \boldsymbol{C}\mathbf{\Theta} \neq \mathbf{0},$$

where $\boldsymbol{C}$ is a $q \times k$ known matrix of rank $q \leq k$. For example, in testing the equality of $k = (q+1)$ mean vectors, the observation matrix $\boldsymbol{Y}$ is of the form given by

$$\boldsymbol{Y} = (\boldsymbol{y}_{11}, \ldots, \boldsymbol{y}_{1N}; \ldots; \boldsymbol{y}_{k1}, \ldots, \boldsymbol{y}_{kN_k})', \tag{1.4}$$

where N_i independent vectors are obtained from the ith group with mean vector $\boldsymbol{\mu}_i$, $i = 1, \ldots, q+1$, and $N = N_1 + \cdots + N_{q+1}$. All the observation vectors have the same covariance matrix which we have assumed in this article as non singular correlation matrix $\boldsymbol{\Lambda}$. To write the problem of testing the equality of $k = (q+1)$ mean vectors as a regression model, we define a vector $\mathbf{1}_r = (1, \ldots, 1)'$ as an r-vector with all the elements equal to one,

$$\boldsymbol{X} = \begin{pmatrix} \mathbf{1}_{N_1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{1}_{N_2} & \mathbf{0} \\ \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{1}_{N_k} \end{pmatrix} : N \times k \tag{1.5}$$

and

$$\boldsymbol{\Theta} = (\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_k)' : k \times p, \;\; k = q+1. \tag{1.6}$$

Thus, the regression model representing the mean vectors of $k = (q+1)$ groups is given by (1.1) with $\boldsymbol{Y}$, $\boldsymbol{X}$ and $\boldsymbol{\Theta}$ defined respectively in (1.4)-(1.6). The problem of testing the equality of $k = (q+1)$ mean vectors is given by $H : \boldsymbol{C\Theta} = \mathbf{0}$ against the alternative $A : \boldsymbol{C\Theta} \neq \mathbf{0}$ where $\boldsymbol{C}$ is now given by $q \times (q+1)$ matrix.

$$\boldsymbol{C} = (\boldsymbol{I}_q, -\mathbf{1}_q) : q \times k, \;\; k = q+1. \tag{1.7}$$

In general, for testing the hypothesis $H : \boldsymbol{C\Theta} = \mathbf{0}$, we consider the variation due to the hypothesis given by

$$\boldsymbol{B} = \boldsymbol{Y}'\boldsymbol{G}\boldsymbol{Y}, \tag{1.8}$$

where

$$\boldsymbol{G} = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{C}'[\boldsymbol{C}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{C}']^{-1}\boldsymbol{C}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}', \tag{1.9}$$

is an $N \times N$ matrix of rank $q < N$. The matrix $\boldsymbol{G}$ is an idempotent matrix of rank q, $\boldsymbol{G}^m = \boldsymbol{G}$ for a positive integer m. That is, there are q eigenvalues that are equal to 1 and the remaining $N - q$ eigenvalues are zero. Also $\boldsymbol{G}$ is symmetric and positive semi-definite. That is, if $\boldsymbol{G} = (g_{ij})$, then we have

$$g_{ii} \geq 0, \;\; \sum_{i=1}^{N} g_{ii} = q, \;\; \sum_{i=1}^{N} g_{ii}^2 + \sum_{i \neq j}^{N} g_{ij}^2 = q.$$

The last equality implies that $\sum_{i=1}^{N} g_{ii}^2 \leq q$ and $\sum_{i\neq j}^{N} g_{ij}^2 \leq q$. In fact for $\boldsymbol{X}$ and $\boldsymbol{C}$ defined by (1.5) and (1.7) respectively in testing the equality of $(q+1)$ mean vectors, we have

$$\sum_{i=1}^{N} g_{ii}^2 = \sum_{i=1}^{q+1} N_i^{-1} - (2q+1)N^{-1} = O(N^{-1}),$$

under the assumption that $N_i = O(N)$, $N = N_1 + \cdots + N_{(q+1)}$. Thus, in this article for the general multivariate regression model, we shall make the following assumptions:

Assumption (A).

A(1) $\sum_{i=1}^{N} g_{ii}^2/q = o(1)$, $N = O(p^\delta)$, $\delta > 1/2$,

A(2) $\lim_{p\to\infty}(\mathrm{tr}\,[\boldsymbol{\Lambda}^2]/p) < \infty$,

A(3) $\lim_{p\to\infty}(\mathrm{tr}\,[\boldsymbol{\Lambda}^4]/p^2) = 0$,

A(4) $\lim_{N\to\infty}(\mathrm{tr}\,[\boldsymbol{G}_+^4]/q^2) = a$, $0 \leq a < \infty$,

A(5) $\lim_{(n,p)\to\infty}\{(pq)^{-1}\mathrm{tr}\,[\boldsymbol{\Lambda}\boldsymbol{M}\boldsymbol{M}']\} = 0$,

where

$$\boldsymbol{M} = \boldsymbol{\Theta}'\boldsymbol{C}'[\boldsymbol{C}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{C}']^{-1/2},\ \boldsymbol{G}_+ = (g_{ij+}), \tag{1.10}$$

and $g_{ij+} = |g_{ij}|$, $i \neq j$, $i, j = 1, \ldots, N$, $g_{ii} \geq 0$.

The matrix $\boldsymbol{G}$ is a positive semi-definite matrix and hence $g_{ii} \geq 0$. Also $\mathrm{tr}\,[\boldsymbol{G}^4] = q$. So the condition A(4) is not a strong condition. The Assumption A(5) gives the local alternative under which the non-null distribution of the statistic will be obtained.

The variation due to the error which can be used to estimate the correlation matrix $\boldsymbol{\Lambda}$ with or without the hypothesis H being true is given by

$$\boldsymbol{S} = n^{-1}\boldsymbol{Y}'(\boldsymbol{I}_N - \boldsymbol{H})\boldsymbol{Y},\ \boldsymbol{H} = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}',\ n = N - k, \tag{1.11}$$

where $\boldsymbol{I}_N - \boldsymbol{H}$ is also an idempotent matrix of rank $N - q - 1 = N - k = n$, $(\boldsymbol{I}_N - \boldsymbol{H})\boldsymbol{G} = 0$, and hence, under normality assumption, it implies that $\boldsymbol{B}$ and $\boldsymbol{S}$ are independently distributed but we do not have normality. The sample correlation matrix $\boldsymbol{R}$ is defined by

$$\boldsymbol{R} = \boldsymbol{D}_{\boldsymbol{S}}^{-1'/2}\boldsymbol{S}\boldsymbol{D}_{\boldsymbol{S}}^{-1/2}, \tag{1.12}$$

where $\boldsymbol{D_S} = diag(\boldsymbol{S})$ is a diagonal matrix with the same diagonal elements as the diagonal elements of $\boldsymbol{S}$. In this paper, we propose the test statistic

$$T_1 = \frac{\text{tr}\,[\boldsymbol{B}\boldsymbol{D}_{\boldsymbol{S}}^{-1}] - npq(n-2)^{-1}}{[2c_{p,n}q(\text{tr}\,[\boldsymbol{R}^2] - n^{-1}p^2)]^{1/2}}, \; n = N - k, \tag{1.13}$$

where

$$c_{p,n} = 1 + (\text{tr}\,[\boldsymbol{R}^2]/p^{3/2}) \tag{1.14}$$

is a correction factor to speed up the convergence of the statistic T_1 to normal which goes to one for $n = O(p^\delta)$, $\delta > 1/2$, as given in Srivastava and Du [8]. Under the assumption of normality, Yamada and Srivastava[10] have shown that as $(n, p) \to \infty$, T_1 is asymptotically normally distributed. In this article we show that this result holds under the general distributions described above in (1.2)-(1.3).

The organization of this paper is as follows. In Section 2, we derive the asymptotic distribution of T_1 under the general distribution described in (1.2)-(1.3) when the hypothesis H holds. The asymptotic non-null distribution of this statistic under local alternative is given in Section 3. The asymptotic distribution of another statistics proposed in the literature is considered in Section 4. In Section 5, the power of the proposed test is compared with some existing tests through simulation. The results on moments are given in Sectionsec:moment. The paper concludes in Section 7.

2 Asymptotic Null Distribution of T_1

We first note that the diagonal elements of the sample covariance matrix $\boldsymbol{S}$ goes in probability to the corresponding diagonal elements of the covariance matrix which in the case of this paper is $\boldsymbol{\Lambda}$. Thus, $\boldsymbol{D_S} \to \boldsymbol{I}_p$ in probability as $n \to \infty$. It also follows from Srivastava and Du [8] and Srivastava [6] that for $n = O(p^\delta)$, $\delta > 0$, $N = O(p^\delta)$, $\delta > 1/2$

$$\frac{1}{p}[\text{tr}\,[\boldsymbol{R}^2] - n^{-1}p^2] \to (\text{tr}\,[\boldsymbol{\Lambda}^2]/p) \tag{2.1}$$

in probability. From the Assumption A(2), it is finite. Thus, we need only to find the asymptotic distribution of

$$
\begin{aligned}
T_1 &\stackrel{p}{=} \frac{1}{\sqrt{pq}}\big\{\mathrm{tr}\,[\boldsymbol{B}] - pq\big\}/(2\mathrm{tr}\,[\boldsymbol{\Lambda}^2]/p)^{1/2} \\
&= \frac{1}{\sqrt{pq}}\big\{\mathrm{tr}\,[\boldsymbol{Y}'\boldsymbol{G}\boldsymbol{Y}] - pq\big\}/(2\mathrm{tr}\,[\boldsymbol{\Lambda}^2]/p)^{1/2} \qquad (2.2)
\end{aligned}
$$

Under the hypothesis $H : \boldsymbol{C\Theta} = 0$, and hence $\boldsymbol{GX\Theta} = 0$. Thus, under H, T_1 becomes

$$
T_1 = \frac{1}{\sqrt{pq}}\big\{\mathrm{tr}\,[\boldsymbol{\Lambda}\boldsymbol{U}'\boldsymbol{G}\boldsymbol{U}] - pq\big\}/(2\mathrm{tr}\,[\boldsymbol{\Lambda}^2]/p)^{1/2}, \qquad (2.3)
$$

where $\boldsymbol{U} = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_N)'$, and $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_N$ are independent and identically distributed p-vectors with mean vector 0 and covariance matrix $\boldsymbol{I}_p$. The fourth moment of each component $\boldsymbol{u}_i = (u_{i1}, \ldots, u_{ip})'$ is the same, namely $E(u_{ik}^4) = K_4 + 3$, $k = 1, \ldots, p$ as the model satisfies the conditions (1.2)-(1.3). Alternatively, we may assume that $u_{i1}, \ldots, u_{ip}$ are independently distributed as is done in Srivastava [6] which results in somewhat simpler algebraic manipulations. But we will continue with the assumptions (1.2) - (1.3). Writing $\boldsymbol{G} = (g_{ij})$, we find that the numerator of T_1 in (2.3) is given by

$$
\begin{aligned}
q_{n,p} &= \frac{1}{\sqrt{pq}}\big\{\mathrm{tr}\,[\boldsymbol{G}\boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}'] - pq\big\} \\
&= \frac{1}{\sqrt{pq}}[\sum_{i=1}^{N}\sum_{j=1}^{N} g_{ij}\boldsymbol{u}_i'\boldsymbol{\Lambda}\boldsymbol{u}_j - pq] \\
&= \frac{1}{\sqrt{pq}}[\sum_{i=1}^{N} g_{ii}\boldsymbol{u}_i'\boldsymbol{\Lambda}\boldsymbol{u}_i - pq] + \frac{1}{\sqrt{pq}}\sum_{i\neq j}^{N} g_{ij}\boldsymbol{u}_i'\boldsymbol{\Lambda}\boldsymbol{u}_j \\
&= J_1 + J_2. \qquad (2.4)
\end{aligned}
$$

We note that

$$
E(J_1) = \frac{1}{\sqrt{pq}}\Big\{\sum_{i=1}^{N} g_{ii}(\mathrm{tr}\,\boldsymbol{\Lambda}) - pq\Big\} = 0,
$$

since $\mathrm{tr}\,(\boldsymbol{\Lambda}) = p$ and $\sum_{i=1}^{N} g_{ii} = q$. Using Lemma 6.1 given in Section 6, we find that the variance of J_1 is given by

$$\begin{aligned} Var(J_1) &= \frac{1}{pq}\sum_{i=1}^{N} g_{ii}^2(K_4 p + 2\mathrm{tr}\,[\boldsymbol{\Lambda}^2]) \\ &= [K_4 + (2\mathrm{tr}\,[\boldsymbol{\Lambda}^2]/p)](\sum_{i=1}^{N} g_{ii}^2/q) \\ &= o(1). \end{aligned} \tag{2.5}$$

Hence, the first term goes to zero in probability. Thus, in probability

$$q_{n,p} \stackrel{p}{=} \frac{1}{\sqrt{pq}}\sum_{i\neq j}^{N} g_{ij}\boldsymbol{u}_i'\boldsymbol{\Lambda}\boldsymbol{u}_j = \frac{2}{\sqrt{pq}}\sum_{j=2}^{N}\sum_{i=1}^{j-1} g_{ij}\boldsymbol{u}_i'\boldsymbol{\Lambda}\boldsymbol{u}_j \tag{2.6}$$

with $E(q_{n,p}) = 0$, and

$$\begin{aligned} Var(q_{n,p}) &= \frac{4}{pq}\sum_{j=2}^{N}\sum_{i=1}^{j-1} g_{ij}^2\mathrm{tr}\,[\boldsymbol{\Lambda}^2] = \frac{2}{pq}\sum_{i\neq j}^{N} g_{ij}^2\mathrm{tr}\,[\boldsymbol{\Lambda}^2] \\ &\cong 2\mathrm{tr}\,[\boldsymbol{\Lambda}^2]/p < \infty, \end{aligned} \tag{2.7}$$

from the Assumption (A). Let

$$\eta_j = \frac{2}{\sqrt{pq}}\sum_{i=1}^{j-1} g_{ij}\boldsymbol{u}_i'\boldsymbol{\Lambda}\boldsymbol{u}_j, \tag{2.8}$$

and let $\Im_j$ be the σ- algebra generated by the random vectors $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_j$. Letting $\boldsymbol{u}_0 = \boldsymbol{0}$, and $\Im_0 = (\phi, \Omega) = \Im_{-1}$, where ϕ is the empty set and Ω the whole space, we find that $\Im_0 \subset \Im_1 \subset \cdots \subset \Im_N \subset \Im_1$, and

$$\begin{aligned} E(\eta_j|\Im_{j-1}) &= 0,\ E(\eta_j) = 0, \\ E(\eta_j^2|\Im_{j-1}) &= \frac{4}{pq}\sum_{i=1}^{j-1} g_{ij}^2 E[\boldsymbol{u}_i'\boldsymbol{\Lambda}\boldsymbol{u}_j\boldsymbol{u}_j'\boldsymbol{\Lambda}\boldsymbol{u}_i|\Im_{j-1}] \\ &\quad + \frac{4}{pq}\sum_{i\neq k}^{j-1} g_{ij}g_{kj}E[\boldsymbol{u}_i'\boldsymbol{\Lambda}\boldsymbol{u}_j\boldsymbol{u}_j'\boldsymbol{\Lambda}\boldsymbol{u}_k|\Im_{j-1}] \\ &= \frac{4}{pq}\sum_{i=1}^{j-1} g_{ij}^2\boldsymbol{u}_i'\boldsymbol{\Lambda}^2\boldsymbol{u}_i + \frac{4}{pq}\sum_{i\neq k}^{j-1} g_{ij}g_{kj}\boldsymbol{u}_i'\boldsymbol{\Lambda}^2\boldsymbol{u}_k, \end{aligned} \tag{2.9}$$

$$E(\eta_j^2) = 4(\sum_{i=1}^{j-1} g_{ij}^2)(\mathrm{tr}\,[\mathbf{\Lambda}^2]/pq) < \infty. \tag{2.10}$$

Hence, the sequence $\{\eta_k, \Im_k\}$ is a sequence of integrable martingale difference. Thus, to establish the asymptotic normality of the random variable $q_{n,p}$ given in (2.5), we may use the Theorem 4 from Shirayev[4]. This requires establishing the Lindberg condition.

For $\varepsilon > 0$,

(I) $L = \sum_{k=2}^{N} E[\eta_k^2 I(|\eta_k| > \varepsilon)]|\Im_{k-1}] \xrightarrow{p} 0$ in probability.

And showing that

(II) $C = \sum_{k=2}^{N} E(\eta_k^2|\Im_{k-1}) \xrightarrow{p} \sigma_0^2$ for some constant σ_0^2.

We first show (II). From (2.10) we find that

$$\begin{aligned} \sum_{j=2}^{N} E(\eta_j^2) &= 4(\sum_{j=2}^{N}\sum_{i=1}^{j-1} g_{ij}^2)(\mathrm{tr}\,[\mathbf{\Lambda}^2]/pq) \\ &= 2(\sum_{i\neq j}^{N} g_{ij}^2)(\mathrm{tr}\,[\mathbf{\Lambda}^2]/pq) \\ &\to 2(\mathrm{tr}\,[\mathbf{\Lambda}^2]/p) = \sigma_0^2 < \infty. \end{aligned}$$

Thus, to show that the convergence condition (II) is satisfied, we need to show that the variance of the random variable C goes to zero. The variance of C is given by

$$Var(C) = \frac{4}{q^2p^2} Var\Big[\sum_{j=2}^{N}(\sum_{i=1}^{j-1} g_{ij}^2 \boldsymbol{u}_i'\mathbf{\Lambda}^2\boldsymbol{u}_i + 2\sum_{i<k}^{j-1} g_{ij}g_{kj}\boldsymbol{u}_i'\mathbf{\Lambda}^2\boldsymbol{u}_k)\Big].$$

We will show that the variance of each term in the right side goes to zero which will imply that $Var(C)$ goes to zero. The variance of the first term is

$$\begin{aligned}
&\frac{4}{q^2p^2}Var\Big[\sum_{j=2}^{N}(\sum_{i=1}^{j-1} g_{ij}^2 \boldsymbol{u}_i'\boldsymbol{\Lambda}^2\boldsymbol{u}_i)\Big] \\
&=\frac{4}{q^2p^2}Var\Big[\sum_{i=1}^{N-1}(\boldsymbol{u}_i'\boldsymbol{\Lambda}^2\boldsymbol{u}_i)(\sum_{j=i+1}^{N} g_{ij}^2)^2\Big] \\
&=\frac{4}{p^2q^2}\Big\{K_4\sum_{i=1}^{p}(\boldsymbol{\Lambda}^2)_{ii}^2 + 2\text{tr}\,[\boldsymbol{\Lambda}^4]\Big\}(\sum_{i\neq j} g_{ij}^2)^2
\end{aligned}$$

where $(\boldsymbol{\Lambda}^2)_{ii}$ is the $(i,i)th$ term of $\boldsymbol{\Lambda}^2$, $i = 1, \ldots, p$. Since

$$\left(\sum_{i=1}^{p}(\boldsymbol{\Lambda}^2)_{ii}^2/p^2\right) \leq (\text{tr}\,[\boldsymbol{\Lambda}^4]/p^2) \to 0,$$

and $(\sum_{i\neq j} g_{ij}^2)^2/q^2 \leq 1$, the variance of the first term goes to zero. Next, we show that under the Assumption (A), the variance of the second term goes to zero. That is

$$\begin{aligned}
&\frac{4}{p^2q^2}Var\Big[2\sum_{j=2}^{N}\sum_{i<k}^{j-1} g_{ij}g_{kj}\boldsymbol{u}_i'\boldsymbol{\Lambda}^2\boldsymbol{u}_k\Big] \\
&=\frac{16}{p^2q^2}Var\Big[\sum_{i\leq k<l}^{N-1}(\sum_{j=l+1}^{N} g_{jk}g_{jl})\boldsymbol{u}_k'\boldsymbol{\Lambda}^2\boldsymbol{u}_l\Big] \\
&=\frac{16}{p^2q^2}(\text{tr}\,[\boldsymbol{\Lambda}^4])\sum_{i\leq k<l}^{N-1}\left(\sum_{j=l+1}^{N} g_{jk}g_{jl}\right)^2 \\
&\leq 16(\text{tr}\,[\boldsymbol{\Lambda}^4]/p^2)\sum_{i\leq k<l}^{N-1}\left(\sum_{j=l+1}^{N} |g_{jk}|\ |g_{jl}|\right)^2 \\
&\leq 16(\text{tr}\,[\boldsymbol{\Lambda}^4]/p^2)(\text{tr}\,[\boldsymbol{G}_+^4]/q^2),
\end{aligned}$$

which goes to zero under Assumptions A(3) and A(4). Then,

$$\sum_{k=2}^{N} E(\eta_k^2|\Im_{k-1}) \xrightarrow{p} \sigma_0^2 = 2(\text{tr}\,[\boldsymbol{\Lambda}^2]/p).$$

To show that Lindberg's condition (L) is satisfied, we need to only show that

$$\sum_{j=2}^{N} E(\eta_j^4) \to 0 \text{ as } N \to \infty,$$

See Srivastava [6]. That is,

$$\begin{aligned}
&\frac{16}{p^2q^2}\sum_{j=2}^{N} E(\sum_{i=1}^{j-1} g_{ij}\boldsymbol{u}_i'\boldsymbol{\Lambda}\boldsymbol{u}_j)^4 \\
&=\frac{16}{p^2q^2}\sum_{j=2}^{N} E[\sum_{i=1}^{j-1} g_{ij}^2(\boldsymbol{u}_i'\boldsymbol{\Lambda}\boldsymbol{u}_j)^2 + 2\sum_{i<l}^{j-1} g_{ij}g_{lj}(\boldsymbol{u}_i'\boldsymbol{\Lambda}\boldsymbol{u}_j)(\boldsymbol{u}_l'\boldsymbol{\Lambda}\boldsymbol{u}_j)]^2 \\
&\leq\frac{16}{p^2q^2}\sum_{j=2}^{N}\Big[(K_4+3)^2(\text{tr}\,[\boldsymbol{\Lambda}^4])\sum_{i=1}^{j-1} g_{ij}^4 + 4E\Big\{\sum_{i<l}^{j-1} g_{ij}g_{lj}(\boldsymbol{u}_i'\boldsymbol{\Lambda}\boldsymbol{u}_j)(\boldsymbol{u}_l'\boldsymbol{\Lambda}\boldsymbol{u}_j)\Big\}^2\Big] \\
&\quad+\frac{16}{p^2q^2}\sum_{j=2}^{N}\Big[2E\Big\{\sum_{i<l}^{j-1} g_{ij}^2g_{lj}^2(\boldsymbol{u}_i'\boldsymbol{\Lambda}\boldsymbol{u}_j)^2(\boldsymbol{u}_l'\boldsymbol{\Lambda}\boldsymbol{u}_j)^2\Big\}\Big] \\
&=\frac{16}{p^2q^2}\sum_{j=2}^{N}(K_4+3)^2\Big\{\sum_{i=1}^{j-1} g_{ij}^4 + 6(K_4+3)\sum_{i<l}^{j-1} g_{ij}^2g_{lj}^2\Big\}(\text{tr}\,[\boldsymbol{\Lambda}^4]) \\
&\leq\frac{16}{p^2q^2}(K_4+3)^2\sum_{j=2}^{N}\Big[\sum_{i<1}^{j-1} g_{ij}^4 + 2\sum_{i<l}^{j-1} g_{ij}^2g_{lj}^2\Big](\text{tr}\,[\boldsymbol{\Lambda}^4]) \\
&=\frac{16}{p^2q^2}(K_4+3)^2\sum_{j=2}^{N}(\sum_{i<1}^{j-1} g_{ij}^2)^2(\text{tr}\,[\boldsymbol{\Lambda}^4]) \\
&\leq\frac{16}{p^2q^2}(K_4+3)^2(\text{tr}\,[\boldsymbol{G}^4])(\text{tr}\,[\boldsymbol{\Lambda}^4]) \\
&=\frac{16}{q}(K_4+3)^2(\text{tr}\,[\boldsymbol{\Lambda}^4]/p^2) \to 0,
\end{aligned}$$

from Assumption A(3). Thus, we have proved the following theorem.

Theorem 2.1 *Consider the model* (1.1) *satisfying* (1.2) *and* (1.3). *Then under the hypothesis* $H : \boldsymbol{C\Theta} = \boldsymbol{0}$, *the statistic* T_1 *defined in* (1.12) *is asymptotically normally distributed with mean* 0 *and variance* 1, *namely*

$$\lim_{(N,p)\to\infty} P_0\left\{T_1 < z_{1-\alpha}\right\} = \Phi(z_{1-\alpha})$$

where Φ denotes a standard normal distribution function, and P_0 denotes that the probability has been computed under the hypothesis H

Corollary 2.1 *As $(N,p) \to \infty$,*

$$T_1 \stackrel{p}{=} \sum_{i\neq j}^{N} g_{ij}\boldsymbol{y}_i'\boldsymbol{y}_j / \left\{2c_{p,n}q(\mathrm{tr}\,[\boldsymbol{R}^2] - n^{-1}p^2)\right\}^{1/2},$$

where $\boldsymbol{G} = (g_{ij}) = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{C}'[\boldsymbol{C}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{C}']^{-1}\boldsymbol{C}(\boldsymbol{X}\boldsymbol{X})^{-1}\boldsymbol{X}'$.

3 Asymptotic Non-Null Distribution of T_1

In this section, we derive the asymptotic distribution of the statistic T_1 under local alternative given by the Assumption A(5), namely

$$\lim_{(N,p)\to\infty} (pq)^{-1}\mathrm{tr}\,[\boldsymbol{\Lambda}\boldsymbol{M}\boldsymbol{M}'] = 0, \tag{3.1}$$

where

$$\boldsymbol{M} = \boldsymbol{\Theta}'\boldsymbol{C}'[\boldsymbol{C}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{C}']^{-1/2} \tag{3.2}$$

From Theorem 2.1, it follows that in probability the statistic

$$T_1 \stackrel{p}{=} (pq)^{-1/2}\left\{\mathrm{tr}\,\boldsymbol{Y}'\boldsymbol{G}\boldsymbol{Y} - pq\right\}/(2\mathrm{tr}\,[\boldsymbol{\Lambda}^2]/p)^{1/2}$$

goes to $\mathcal{N}(0,1)$ under the hypothesis H. This implies that irrespective of any hypothesis, the random variable

$$\begin{aligned}
T_1^* &= (pq)^{-1/2}\left\{\mathrm{tr}\,[(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\Theta})'\boldsymbol{G}(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\Theta})] - pq\right\}/(2\mathrm{tr}\,[\boldsymbol{\Lambda}^2]/p)^{1/2} \\
&= \frac{(pq)^{-1/2}}{(2\mathrm{tr}\,[\boldsymbol{\Lambda}^2]/p)^{1/2}}\left\{\mathrm{tr}\,[\boldsymbol{Y}'\boldsymbol{G}\boldsymbol{Y}] - 2\mathrm{tr}\,[\boldsymbol{\Theta}'\boldsymbol{X}'\boldsymbol{G}\boldsymbol{Y}] + \mathrm{tr}\,[\boldsymbol{\Theta}'\boldsymbol{X}'\boldsymbol{G}\boldsymbol{X}\boldsymbol{\Theta}] - pq\right\} \\
&= T_1 + \frac{(pq)^{-1/2}}{(2\mathrm{tr}\,[\boldsymbol{\Lambda}^2]/p)^{1/2}}\left\{-2\mathrm{tr}\,[\boldsymbol{\Theta}'\boldsymbol{X}'\boldsymbol{G}\boldsymbol{Y}] + \mathrm{tr}\,[\boldsymbol{\Theta}'\boldsymbol{X}'\boldsymbol{G}\boldsymbol{X}\boldsymbol{\Theta}]\right\} \\
&\to \mathcal{N}(0,1) \text{ as } (N,p) \to \infty
\end{aligned}$$

It may be noted that the random variable T_1^* depends on unknown parameters $\boldsymbol{\Theta}$. We now show that under the assumption A(5)

$$(pq)^{-1/2}\mathrm{tr}\,[\boldsymbol{\Theta}'\boldsymbol{X}'\boldsymbol{G}\boldsymbol{Y}] \to (pq)^{-1/2}\mathrm{tr}\,[\boldsymbol{\Theta}'\boldsymbol{X}'\boldsymbol{G}\boldsymbol{X}\boldsymbol{\Theta}].$$

Let

$$\boldsymbol{A} = \boldsymbol{\Theta}'\boldsymbol{X}'\boldsymbol{G} = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n), \ \boldsymbol{Y} = (\boldsymbol{y}_1, \ldots.\boldsymbol{y}_N)'.$$

Then

$$E[(pq)^{-1/2}\mathrm{tr}\,[\boldsymbol{AY}]] = (pq)^{-1/2}\mathrm{tr}\,[\boldsymbol{A\Theta X}] = (pq)^{-1/2}\mathrm{tr}\,[\boldsymbol{\Theta}'\boldsymbol{X}'\boldsymbol{GX\Theta}]$$

and since $\boldsymbol{G}^2 = \boldsymbol{G}$,

$$\begin{aligned}
&Var[(pq)^{-1/2}\mathrm{tr}\,[\boldsymbol{AY}]] \\
=&(pq)^{-1}Var[\mathrm{tr}\sum_{i=1}^{N}\boldsymbol{a}_i\boldsymbol{y}_i'] = (pq)^{-1}Var[\sum_{i=1}^{N}\boldsymbol{a}_i'\boldsymbol{y}_i] \\
=&(pq)^{-1}\sum_{i=1}^{N}\boldsymbol{a}_i'\boldsymbol{\Lambda}\boldsymbol{a}_i = (pq)^{-1}\mathrm{tr}\,[\boldsymbol{\Lambda}(\sum_{i=1}^{N}\boldsymbol{a}_i\boldsymbol{a}_i')] \\
=&(pq)^{-1}\mathrm{tr}\,[\boldsymbol{\Lambda AA}'] = (pq)^{-1}\mathrm{tr}\,[\boldsymbol{\Lambda\Theta}'\boldsymbol{X}'\boldsymbol{GX\Theta}'] \\
=&(pq)^{-1}\mathrm{tr}\,[\boldsymbol{\Lambda MM}'],
\end{aligned}$$

which goes to zero under the Assumption A(5). Thus,

$$\begin{aligned}
(pq)^{-1/2}\mathrm{tr}\,[\boldsymbol{\Theta}'\boldsymbol{X}'\boldsymbol{GY}] \xrightarrow{p} &(pq)^{-1/2}\mathrm{tr}\,[\boldsymbol{\Theta}'\boldsymbol{X}'\boldsymbol{GX\Theta}] \\
=&(pq)^{-1/2}\mathrm{tr}\,\boldsymbol{MM}'
\end{aligned}$$

and

$$T_1^* \stackrel{p}{=} T_1 - (pq)^{-1/2}\mathrm{tr}\,[\boldsymbol{MM}']/\sqrt{2\mathrm{tr}\,\boldsymbol{\Lambda}^2/p}$$

Hence,

$$\begin{aligned}
&P_1\Big\{T_1 > z_{1-\alpha}|\text{ under A(5)}\Big\} \\
=&P_1\Big\{T_1 - \frac{\mathrm{tr}\,[\boldsymbol{MM}']}{\sqrt{2q\mathrm{tr}\,[\boldsymbol{\Lambda}^2]}} > z_{1-\alpha} - \frac{\mathrm{tr}\,[\boldsymbol{MM}']}{\sqrt{2q\mathrm{tr}\,[\boldsymbol{\Lambda}^2]}}\Big\} \\
=&P_1\Big\{T_1^* > z_{1-\alpha} + \frac{\mathrm{tr}\,[\boldsymbol{MM}']}{\sqrt{2q\mathrm{tr}\,[\boldsymbol{\Lambda}^2]}}\Big\} \\
=&\Phi\Big(-z_{1-\alpha} + \frac{\mathrm{tr}\,[\boldsymbol{MM}']}{\sqrt{2q\mathrm{tr}\,[\boldsymbol{\Lambda}^2]}}\Big),
\end{aligned}$$

where P_1 denotes that the probability has been computed under the local alternative hypothesis given in A(5). It may be noted that if the assumption that $\boldsymbol{D_\Sigma} = \boldsymbol{I}_p$, where $\boldsymbol{D_\Sigma} = diag(\sigma_{11}, ..., \sigma_{pp})$, $\boldsymbol{\Sigma} = (\sigma_{ij})$ is dropped, then the power can be written as

$$P_1\left\{T_1 > z_\alpha | \text{ under A(5)}\right\} = \Phi\left(-z_\alpha + \frac{\mathrm{tr}\,[\boldsymbol{D}_{\boldsymbol{\Sigma}}^{-1}\boldsymbol{M}\boldsymbol{M}']}{\sqrt{2q\mathrm{tr}\,[\boldsymbol{\Lambda}^2]}}\right),$$

for the model $\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\Theta} + \boldsymbol{\Sigma}^{1/2}\boldsymbol{\Lambda}^{1/2}\boldsymbol{U}$. Hence, we get the following theorem.

Theorem 3.1 *Under the model* $\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\Theta} + \boldsymbol{U}\boldsymbol{\Lambda}^{1/2}\boldsymbol{D}_{\boldsymbol{\Sigma}}^{1/2}$, *where the elements of* $\boldsymbol{U}$ *satisfies conditions* (1.2) - (1.3)

$$P_1\left\{T_1 > z_\alpha\right\} = \Phi\left(-z_\alpha + \frac{\mathrm{tr}\,[\boldsymbol{D}_{\boldsymbol{\Sigma}}^{-1}\boldsymbol{M}\boldsymbol{M}']}{\sqrt{2q\mathrm{tr}\,[\boldsymbol{\Lambda}^2]}}\right).$$

The Assumption A(5) *becomes* $\lim_{(N,p)\to\infty}(pq)^{-1}\mathrm{tr}\,[\boldsymbol{\Lambda}\boldsymbol{D}_{\boldsymbol{\Sigma}}^{-1/2}\boldsymbol{M}\boldsymbol{M}'\boldsymbol{D}_{\boldsymbol{\Sigma}}^{-1/2}] = 0$.

4 Other Tests

Bai and Saranadasa [1] proposed a two-sample test for testing the equality of two mean vectors. A generalized version of this test for the MANOVA problem was given by Srivastava and Fujikoshi [7] for normally distributed observation vectors. It is given by

$$T_2 = [2pq\hat{a}_2(1 + n^{-1}q)]^{-1/2}\left\{\mathrm{tr}\,[\boldsymbol{B}] - q\mathrm{tr}\,[\boldsymbol{S}]\right\},$$

where

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

Under the hypothesis $H : \boldsymbol{C}\boldsymbol{\Theta} = 0$, T_2 is asymptotically normally distributed as $\mathcal{N}(0, 1)$. That is,

$$\lim_{(N,p)\to\infty} P_0\left\{T_2 < z_\alpha\right\} = \Phi(z_\alpha).$$

By following the methods given in Section 2 of this article, it can be shown that the asymptotic normality of T_2 under the hypothesis still holds for the non-normal model considered in this paper under the corresponding modified assumptions on the covariance matrix $\boldsymbol{\Sigma}$ in place of the correlation matrix

$\mathbf{\Lambda}$. Similarly it can be shown that under the alternative hypothesis A(5), the asymptotic distribution is given by

$$\lim_{(n,p)\to\infty} P_1\left\{T_2 > z_\alpha\right\} = \Phi\Big(-z_\alpha + \frac{\mathrm{tr}\left[\boldsymbol{M}\boldsymbol{M}'\right]}{\sqrt{2q\mathrm{tr}\left[\boldsymbol{\Sigma}^2\right]}}\Big).$$

The test T_2 for normally distributed observation vectors was also considered by Schott [3] who obtained its distribution under the condition that (n/p) goes to a constant as $(n,p) \to \infty$. It has been shown in Srivastava and Du [8] that T_1 performs better than T_2. The test proposed by Srivastava [5], which has been shown to perform better than T_2 in Srivastava and Fujikoshi [7] is not considered in this paper as its distribution under non-normal model has yet to be derived.

5 Power and Attained Significance Level

In this section we compare the power of the statistics T_1 and T_2 in finite samples by simulation. We first examine the attained significance level to the nominal value $\alpha = 0.05$.

The attained significance level (ASL) is $\hat{\alpha}_T = \#(T_{1H} > z_{1-\alpha})/r$ where T_{1H} are values of the test statistic T_1 (or T_2) computed from data simulated under H, r is the number of replications and $z_{1-\alpha}$ is the $100(1-\alpha)\%$ point of the standard normal distribution. The ASL assesses how close the null distribution of T_1 (or T_2) is to its limiting null distribution. From the same simulation, we also obtain $\hat{z}_{1-\alpha}$ as the $100(1-\alpha)\%$ point of the empirical null distribution, and define the attained power by $\hat{\beta}_T = \#(T_{1A} > \hat{z}_{1-\alpha})/r$, where T_{1A} are values of the T_1 (or T_2) computed from data simulated under A.

Through the simulation, we compare the proposed test T_1 with T_2. It may be noted that irrespective of the ASL of any statistic, the power has been computed when all the statistics in the comparison have the same specified significance level as the cut off points have been obtained by simulation. The ASL gives an idea as to how close it is to the specified significance level. If it is not close, the only choice left is to obtain it from simulation, not from the asymptotic distribution. It is common in practice, although not recommended, to depend on the asymptotic distribution, rather than relying on simulations to determine the ASL.

For simulation, we consider the problem of testing the equality of 3 mean vectors, that is, $k = q + 1 = 3$ and $q = 2$, where $N_1 = N_2 = N_3 = N^*$, and the cases of $(N^*, p) = (10, 40), (20, 80), (30, 120)$ and $(40, 200)$ are treated.. Note that $n = N_1 + N_2 + N_3 - k = 3(N^* - 1)$. For the three mean vectors will

$$\begin{aligned} \boldsymbol{\Theta} &= (\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \boldsymbol{\mu}_3)' : 3 \times p, \\ \boldsymbol{C} &= \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix}, \ \boldsymbol{C\Theta} = \begin{pmatrix} \boldsymbol{\mu}_1' & -\boldsymbol{\mu}_3' \\ \boldsymbol{\mu}_2 & -\boldsymbol{\mu}_3' \end{pmatrix}. \end{aligned}$$

The observation matrix is

$$\begin{aligned} \boldsymbol{Y} &= (\boldsymbol{y}_1^{(1)}, \ldots, \boldsymbol{y}_{N^*}^{(1)}; \boldsymbol{y}_1^{(2)}, \ldots, \boldsymbol{y}_{N^*}^{(2)}; \boldsymbol{y}_1^{(3)}, \ldots, \boldsymbol{y}_{N^*}^{(3)})' \\ \underset{N \times 3}{\boldsymbol{X}} &= \begin{pmatrix} \mathbf{1}_{N^*} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{1}_{N^*} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1}_{N^*} \end{pmatrix}, \end{aligned}$$

where $\mathbf{1}_{N^*} = (1, \ldots, 1)' : N^* \times 1$ for $N = 3N^*$. For the hypothesis, without loss of generality we choose $\boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = \boldsymbol{\mu}_3 = \mathbf{0}$. For the alternative hypothesis, we choose $\boldsymbol{\mu}_1 = 0$, $\boldsymbol{\mu}_2 = 3n^{-1/2}p^{-1/4}\mathbf{1}_p'$, $\boldsymbol{\mu}_3 = -\boldsymbol{\mu}_2$.

To generate the $\boldsymbol{Y}$ matrix from a non-normal distribution, we generate $3N^*p$ i.i.d. random variables u_{ij} from three kinds of chi-square distributions, namely, χ_2^2, χ_8^2 and χ_{32}^2 with 2, 8 and 32 degrees of freedom, respectively, and centre them and scale them as

$$\nu_{ij} = (u_{ij} - m)/\sqrt{2m},$$

for $u_{ij} \sim \chi_m^2$, $m = 2, 8, 32$. Since the skewness and kurtosis $(K_4 + 3)$ of χ_m^2 is, repectively, $(8/m)^{1/2}$ and $3 + 12/m$, it is noted that χ_2^2 has higher skewness and kurtosis than χ_8^2 and χ_{32}^2. Write them as

$$\boldsymbol{V} = (\boldsymbol{\nu}_1^{(1)}, \ldots, \boldsymbol{\nu}_{N^*}^{(1)}; \boldsymbol{\nu}_1^{(2)}, \ldots, \boldsymbol{\nu}_{N^*}^{(2)}; \boldsymbol{\nu}_1^{(3)}, \ldots, \boldsymbol{\nu}_{N^*}^{(3)})'$$

where $\boldsymbol{\nu}_j^{(i)}$ vectors are p-vectors, $j = 1, \ldots, N^*$, $i = 1, 2, 3$. For the covariance matrix, we consider two cases

(Case 1) $\boldsymbol{\Sigma} = \boldsymbol{I}_p$,

(Case 2) $\boldsymbol{\Sigma} = \boldsymbol{D}_a = \mathrm{diag}\,(a_1^2, \ldots, a_p^2)$, where a_i are i.i.d. as chi-square with 3 degrees of freedom.

For the first case, we define

(Case 1) $\quad \boldsymbol{Y} = \boldsymbol{V} + \boldsymbol{X}(\boldsymbol{0}, \boldsymbol{\mu}_2, \boldsymbol{\mu}_3)'$,

where under the hypothesis, $\boldsymbol{Y} = \boldsymbol{V}$, and under the alternative, $\boldsymbol{\mu}_2$ and $\boldsymbol{\mu}_3$ are replaced by the vectors mentioned above.

For the second case

(Case 2) $\quad \boldsymbol{Y} = \boldsymbol{V}\boldsymbol{D}_a^{1/2} + \boldsymbol{X}(\boldsymbol{0}, \boldsymbol{\mu}_2, \boldsymbol{\mu}_3)'$,

where under the hypothesis, $\boldsymbol{Y} = \boldsymbol{V}\boldsymbol{D}_a^{1/2}$, and in the alternative, $\boldsymbol{\mu}_2$ and $\boldsymbol{\mu}_3$ are replaced by the vectors mentioned above.

The simulation results under the χ_m^2 distributions for $m = 2$, 8 and 32 are presented in Tables 1, 2 and 3, respectively. The critical values are computed based on 100,000 replications and the ASL and the powers are obtained based on 10,000 replications. It is noted that the 95% point of the standard normal distribution is 1.64485. Three tables report the critical values and the power in the hypothesis of the two tests, and it is seen that the values of the ASL are appropriate. As reported in the tables, the powers of the two tests perform similarly in Case 1, but the proposed test T_1 has much higher powers than T_2 in Case 2. For the χ_2^2-distribution, which has higher skewness and kurtosis, T_1 has slightly higher power than T_2 in Case 1. Clearly, when $\boldsymbol{\Sigma} = \boldsymbol{I}_p$, all the components have the same unit of measurements and hence both tests perform equally well but when the unit of measurements are not the same, as in Case 2, the proposed test performs much better than the test based on T_2.

6 Results on moments

We here provide results on moments.

Lemma 6.1 *Let* $\boldsymbol{u} = (u_1, \dots, u_p)'$ *be a* p*-dimensional random vector such that* $E[\boldsymbol{u}] = \boldsymbol{0}$, $\mathbf{Cov}\,[\boldsymbol{u}] = \boldsymbol{I}_p$, $E[u_i^4] = K_4 + 3$, $i = 1, \dots, p$, *and*

$$E[u_i^a u_j^b u_k^c u_l^d] = E[u_i^a]E[u_j^b]E[u_k^c]E[u_l^d], \tag{6.1}$$

$0 \leq a + b + c + d \leq 4$ *for all* i, j, k, l. *Then for any* $p \times p$ *symmetric matrices* $\boldsymbol{A} = (a_{ij})$ *and* $\boldsymbol{B} = (b_{ij})$ *of constants, we have*

Table 1: Critical values, ASL and powers of the tests T_1 and T_2 in the case of χ^2_2-distribution with skewness 2 and kurtosis 9

		Critical Value		**ASL in H**		**Power in A**	
N^*	p	T_1	T_2	T_1	T_2	T_1	T_2
				Case 1			
10	40	1.6061	1.5966	4.72	4.37	92.04	85.02
20	80	1.5632	1.6225	4.18	4.81	90.46	85.65
30	120	1.5622	1.6440	4.54	5.28	89.94	86.21
40	200	1.5564	1.6386	4.05	4.76	90.43	87.40
				Case 2			
10	40	1.6061	1.6865	4.72	5.57	99.96	24.71
20	80	1.5632	1.6784	4.18	5.21	99.63	18.20
30	120	1.5622	1.6919	4.54	5.73	97.82	15.20
40	200	1.5564	1.6852	4.05	5.36	96.26	16.97

Table 2: Critical values, ASL and powers of the tests T_1 and T_2 in the case of χ^2_8-distribution with skewness 1 and kurtosis 4.5

		Critical Value		**ASL in H**		**Power in A**	
N^*	p	T_1	T_2	T_1	T_2	T_1	T_2
				Case 1			
10	40	1.7339	1.7029	5.95	5.64	84.92	84.48
20	80	1.6175	1.6810	4.69	5.36	86.92	86.19
30	120	1.6119	1.6812	4.42	5.07	87.04	86.49
40	200	1.5967	1.6714	4.29	5.03	87.80	87.26
				Case 2			
10	40	1.7339	1.7903	5.95	6.21	99.93	23.56
20	80	1.6175	1.7276	4.69	6.19	99.27	18.89
30	120	1.6119	1.7344	4.42	5.60	97.06	15.56
40	200	1.5967	1.7291	4.29	5.83	94.70	16.38

Table 3: Critical values, ASL and powers of the tests T_1 and T_2 in the case of χ^2_{32}-distribution with skewness 0.5 and kurtosis 3.375

		Critical Value		**ASL in** H		**Power in** A	
N^*	p	T_1	T_2	T_1	T_2	T_1	T_2
				Case 1			
10	40	1.7688	1.7184	5.99	5.56	82.13	84.72
20	80	1.6457	1.6930	4.92	5.16	84.12	84.89
30	120	1.6155	1.6812	4.83	5.29	86.09	86.20
40	200	1.6090	1.6831	4.29	5.01	86.84	87.08
				Case 2			
10	40	1.7688	1.8157	5.99	6.28	99.97	23.15
20	80	1.6457	1.7409	4.92	5.34	98.93	16.80
30	120	1.6155	1.7476	4.83	6.33	96.46	15.85
40	200	1.6090	1.7223	4.29	5.59	94.55	16.61

$$\text{(a)} \quad E[(\boldsymbol{u}'\boldsymbol{A}\boldsymbol{u})^2] = K_4 \sum_{i=1}^{p} a_{ii}^2 + 2\mathrm{tr}\,[\boldsymbol{A}^2] + (\mathrm{tr}\,[\boldsymbol{A}])^2,$$

$$\text{(b)} \quad Var[\boldsymbol{u}'\boldsymbol{A}\boldsymbol{u}] = K_4 \sum_{i=1}^{p} a_{ii}^2 + 2\mathrm{tr}\,[\boldsymbol{A}^2],$$

$$\text{(c)} \quad E[\boldsymbol{u}'\boldsymbol{A}\boldsymbol{u}\boldsymbol{u}'\boldsymbol{B}\boldsymbol{u}] = K_4 \sum_{i=1}^{p} a_{ii}b_{ii} + 2\mathrm{tr}\,[\boldsymbol{A}\boldsymbol{B}] + \mathrm{tr}\,[\boldsymbol{A}]\mathrm{tr}\,[\boldsymbol{B}].$$

Proof. For (a), note that $a_{ij} = a_{ji}$. Then under condition (6.1),

$$\begin{aligned} E[(\boldsymbol{u}'\boldsymbol{A}\boldsymbol{u})^2] =& E\Big[\Big(\sum_{i=1}^p a_{ii}u_i^2 + 2\sum_{j<k}^p a_{jk}u_ju_k\Big)^2\Big] \\ =& E\Big[\sum_{i=1}^p a_{ii}^2 u_i^4 + \sum_{i\neq j}^p a_{ii}a_{jj}u_i^2u_j^2 + 4\big(\sum_{j<k}^p a_{jk}u_ju_k\big)^2 \\ &+ 8\sum_{j<k} a_{jj}a_{jk}u_j^3u_k + 8\sum_{i\neq j, j<k} a_{ii}a_{jk}u_i^2u_ju_k\Big] \\ =& (K_4+3)\sum_{i=1}^p a_{ii}^2 + \sum_{i\neq j}^p a_{ii}a_{jj} + 4\sum_{j<k}^p (a_{jk})^2 \\ =& K_4\sum_{i=1}^p a_{ii}^2 + (\sum_{i=1}^p a_{ii}^2 + \sum_{i\neq j}^p a_{ii}a_{jj}) + 2\sum_{i=1}^p a_{ii}^2 + 2\sum_{i\neq j}^p (a_{ij})^2 \\ =& K_4\sum_{i=1}^p a_{ii}^2 + (\mathrm{tr}\,[\boldsymbol{A}])^2 + 2\mathrm{tr}\,[\boldsymbol{A}^2]. \end{aligned}$$

For (b), from condition (6.1), it follows that

$$E[\boldsymbol{u}'\boldsymbol{A}\boldsymbol{u}] = E\Big[\sum_{i=1}^p a_{ii}u_i^2 + \sum_{i\neq j} a_{ii}a_{jj}u_iu_j\Big] = \sum_{i=1}^p a_{ii} = \mathrm{tr}\,\boldsymbol{A},$$

which, together with the equality in (a), yields the equality in (b).

For (c), it is seen that

$$\begin{aligned} E[\boldsymbol{u}'\boldsymbol{A}\boldsymbol{u}\boldsymbol{u}'\boldsymbol{B}\boldsymbol{u}] =& E\Big[\Big(\sum_{i=1}^p a_{ii}u_i^2 + 2\sum_{i<j}^p a_{ij}u_iu_j\Big)\Big(\sum_{i=1}^p b_{ii}u_i^2 + 2\sum_{i<j}^p b_{ij}u_iu_j\Big)\Big] \\ =& \gamma\sum_{i=1}^p a_{ii}b_{ii} + \sum_{i\neq j} a_{ii}b_{jj} + 4\sum_{i<j} a_{ij}b_{ij}, \end{aligned}$$

for $\gamma = K_4 + 3$. Noting that $\mathrm{tr}\,[\boldsymbol{A}\boldsymbol{B}] = \sum_{i=1}^p a_{ii}b_{ii} + 2\sum_{i<j}^p a_{ij}b_{ij}$ and $\mathrm{tr}\,[\boldsymbol{A}]\mathrm{tr}\,[\boldsymbol{B}] = \sum_{i=1}^p a_{ii}b_{ii} + \sum_{i\neq j}^p a_{ii}b_{jj}$, we can get the equality in (c). ■

Corollary 6.1 *Let* $\bar{\boldsymbol{u}} = N^{-1}\sum_{i=1}^N \boldsymbol{u}_i$*, where* $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_N$ *are independently and identically distributed. Then*

$$Var(\bar{\boldsymbol{u}}'\boldsymbol{A}\bar{\boldsymbol{u}}) = \frac{K_4}{N^3}\sum_{i=1}^p a_{ii}^2 + \frac{2}{N^2}\mathrm{tr}\,[\boldsymbol{A}^2].$$

Proof. This corrolary is shown as follows:

$$\begin{aligned}
N^2 Var(\bar{\boldsymbol{u}}'\boldsymbol{A}\bar{\boldsymbol{u}}) =& \frac{1}{N^2} Var\Big[\Big(\sum_{i=1}^{N}\boldsymbol{u}_i\Big)'\boldsymbol{A}\Big(\sum_{i=1}^{N}\boldsymbol{u}_i\Big)\Big] \\
=& \frac{1}{N^2} Var\Big[\sum_{i=1}^{N}\boldsymbol{u}_i'\boldsymbol{A}\boldsymbol{u}_i + 2\sum_{i<k}^{N}\boldsymbol{u}_j'\boldsymbol{A}\boldsymbol{u}_k\Big] \\
=& \frac{1}{N^2} Var\Big[\sum_{i=1}^{N}\boldsymbol{u}_i'\boldsymbol{A}\boldsymbol{u}_i\Big] + \frac{4}{N^2} Var\Big[\sum_{j<k}^{N}\boldsymbol{u}_j'\boldsymbol{A}\boldsymbol{u}_k\Big] \\
&+ \frac{4}{N^2} Cov\Big[\sum_{i=1}^{N}\boldsymbol{u}_i'\boldsymbol{A}\boldsymbol{u}_i, \sum_{j<k}^{N}\boldsymbol{u}_j'\boldsymbol{A}\boldsymbol{u}_k\Big] \\
=& \frac{1}{N} Var[\boldsymbol{u}_1'\boldsymbol{A}\boldsymbol{u}_1] + \frac{2N(N-1)}{N^2}\mathrm{tr}\,[\boldsymbol{A}^2] \\
=& \frac{1}{N}\Big\{K_4\sum_{j=1}^{p} a_{jj}^2 + 2\mathrm{tr}\,[\boldsymbol{A}^2]\Big\} + \frac{2(N-1)}{N}\mathrm{tr}\,[\boldsymbol{A}^2] \\
=& \frac{1}{N} K_4 \sum_{j=1}^{p} a_{jj}^2 + 2\mathrm{tr}\,[\boldsymbol{A}^2].
\end{aligned}$$

■

Lemma 6.2 *Let* $\boldsymbol{u}$ *and* $\boldsymbol{v}$ *be independently and identically distributed random vectors with zero mean vector and covariance matrix* $\boldsymbol{I}_p$. *Then under condition* (6.1) *for any* $p \times p$ *symmetric matrix* $\boldsymbol{A} = (a_{ij})$,

$$Var[(\boldsymbol{u}'\boldsymbol{A}\boldsymbol{v})^2] = K_4^2 \sum_{i,j}^{p} a_{ij}^4 + 6K_4 \sum_{i,j,k}^{p} a_{ij}^2 a_{ik}^2 + 6\mathrm{tr}\,[\boldsymbol{A}^4] + 2(\mathrm{tr}\,[\boldsymbol{A}^2])^2.$$

Proof. Since $E[(\boldsymbol{u}'\boldsymbol{A}\boldsymbol{v})^2] = E[\boldsymbol{u}'\boldsymbol{A}\boldsymbol{v}\boldsymbol{v}'\boldsymbol{A}\boldsymbol{u}] = \mathrm{tr}\,[\boldsymbol{A}^2]$, we have $Var[(\boldsymbol{u}'\boldsymbol{A}\boldsymbol{v})^2] = E[(\boldsymbol{u}'\boldsymbol{A}\boldsymbol{v})^4] - (\mathrm{tr}\,[\boldsymbol{A}^2])^2$. Let $\boldsymbol{C} = (c_{ij}) = \boldsymbol{A}\boldsymbol{v}\boldsymbol{v}'\boldsymbol{A}$. Then, $\mathrm{tr}\,[\boldsymbol{C}] = \boldsymbol{v}'\boldsymbol{A}^2\boldsymbol{v}$ and $\mathrm{tr}\,[\boldsymbol{C}^2] = \mathrm{tr}\,[\boldsymbol{A}\boldsymbol{v}\boldsymbol{v}'\boldsymbol{A}\boldsymbol{A}\boldsymbol{v}\boldsymbol{v}'\boldsymbol{A}] = (\boldsymbol{v}'\boldsymbol{A}^2\boldsymbol{v})^2 = (\mathrm{tr}\,[\boldsymbol{C}])^2$. Since $(\boldsymbol{u}'\boldsymbol{A}\boldsymbol{v})^4 =$

$(\boldsymbol{u}'\boldsymbol{C}\boldsymbol{u})^2$, from (a) and (c) in Lemma 6.1, it follows that

$$
\begin{aligned}
E[(\boldsymbol{u}'\boldsymbol{A}\boldsymbol{v})^4] =& E[(\boldsymbol{u}'\boldsymbol{C}\boldsymbol{u})^2] = E\big[E[(\boldsymbol{u}'\boldsymbol{C}\boldsymbol{u})^2|\boldsymbol{C}]\big] \\
=& E\Big[K_4\sum_{i=1}^p c_{ii}^2 + 2\mathrm{tr}\,[\boldsymbol{C}^2] + (\mathrm{tr}\,[\boldsymbol{C}])^2\Big] \\
=& E\Big[K_4\sum_{i=1}^p c_{ii}^2 + 3(\boldsymbol{v}'\boldsymbol{A}^2\boldsymbol{v})^2\Big] \\
=& K_4\sum_{i=1}^p E[c_{ii}^2] + 3\Big\{K_4\sum_{i=1}^p\{(\boldsymbol{A}^2)_{ii}\}^2 + 2\mathrm{tr}\,[\boldsymbol{A}^4] + (\mathrm{tr}\,[\boldsymbol{A}^2])^2\Big\}.
\end{aligned}
$$

Let $\boldsymbol{A}' = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_p)$ for column vectors $\boldsymbol{a}_i$'s. Since $\boldsymbol{C} = \boldsymbol{A}\boldsymbol{v}\boldsymbol{v}'\boldsymbol{A}'$ and $\boldsymbol{v}'\boldsymbol{A}' = (\boldsymbol{v}'\boldsymbol{a}_1, \ldots, \boldsymbol{v}'\boldsymbol{a}_p)$, it is seen that $c_{ii} = \boldsymbol{a}_i'\boldsymbol{v}\boldsymbol{v}'\boldsymbol{a}_i = \boldsymbol{v}'\boldsymbol{a}_i\boldsymbol{a}_i'\boldsymbol{v}$ and $c_{ii}^2 = (\boldsymbol{v}'\boldsymbol{a}_i\boldsymbol{a}_i'\boldsymbol{v})^2 = (\boldsymbol{v}'\boldsymbol{G}_i\boldsymbol{v})^2$ for $\boldsymbol{G}_i = \boldsymbol{a}_i\boldsymbol{a}_i'$. Hence, from (a) in Lemma 6.1,

$$
\begin{aligned}
E[c_{ii}^2] =& E[(\boldsymbol{v}'\boldsymbol{G}_i\boldsymbol{v})^2] \\
=& K_4\sum_{j=1}^p\{(\boldsymbol{G}_i)_{jj}\}^2 + 2\mathrm{tr}\,[\boldsymbol{G}_i^2] + (\mathrm{tr}\,[\boldsymbol{G}_i])^2 \\
=& K_4\sum_{j=1}^p\{(\boldsymbol{G}_i)_{jj}\}^2 + 3(\boldsymbol{a}_i'\boldsymbol{a}_i)^2.
\end{aligned}
$$

Since $\boldsymbol{G}_i = \boldsymbol{a}_i\boldsymbol{a}_i'$, it is noted that $(\boldsymbol{G}_i)_{jj} = a_{ij}^2$. Since $\boldsymbol{A}^2 = \boldsymbol{A}\boldsymbol{A}' = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_p)'(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_p)$, it is seen that $(\boldsymbol{A}^2)_{ii} = \boldsymbol{a}_i'\boldsymbol{a}_i = \sum_{j=1}^p a_{ij}^2$. Hence, we get

$$
\begin{aligned}
E[(\boldsymbol{u}'\boldsymbol{A}\boldsymbol{v})^4] =& K_4^2\sum_{i=1}^p\sum_{j=1}^p a_{ij}^4 + 3K_4\sum_{i=1}^p(\boldsymbol{a}_i'\boldsymbol{a}_i)^2 \\
&+ 3K_4\sum_{i=1}^p(\boldsymbol{a}_i'\boldsymbol{a}_i)^2 + 6\mathrm{tr}\,[\boldsymbol{A}^4] + 3(\mathrm{tr}\,[\boldsymbol{A}^2])^2.
\end{aligned}
$$

Thus,

$$
Var[(\boldsymbol{u}'\boldsymbol{A}\boldsymbol{v})^2] = K_4^2\sum_{i=1}^p\sum_{j=1}^p a_{ij}^4 + 6K_4\sum_{i=1}^p(\boldsymbol{a}_i'\boldsymbol{a}_i)^2 + 6\mathrm{tr}\,[\boldsymbol{A}^4] + 2(\mathrm{tr}\,[\boldsymbol{A}^2])^2.
$$

Noting that $\sum_{i=1}^p(\boldsymbol{a}_i'\boldsymbol{a}_i)^2 = \sum_{i=1}^p(\sum_{j=1}^p a_{ij}^2)^2 = \sum_{i,j,k}^p a_{ij}^2 a_{ik}^2$, we get the equality in Lemma 6.2. ∎

7 Concluding Remarks

In this article, we have proposed a test which is invariant under the change of unit of measurements. It has been shown to perform better than the test proposed by Srivastava and Fujikoshi[7] and Schott[3] unless $\boldsymbol{\Sigma} = \sigma^2 \boldsymbol{I}_p$ in which case both tests are equally good. Our simulation results show that both tests are robust and the assumptions of normality is not needed to carry out any of the two tests.

Acknowledgments. The research of the first author was supported by NSERC and that of the second author was supported in part by Grant-in-Aid for Scientific Research Nos. 19200020 and 21540114 from Japan Society for the Promotion of Science.

References

[1] Z. Bai and H. Saranadasa. Effect of high dimension: by an example of a two sample problem. *Statistics Sinice*, 6:311–329, 1996.

[2] G.E.P Box. A general distribution theory for a class of likelihood criteria. *Biometrika*, 36:317–346, 1949.

[3] J.R. Schott. Some high-dimensional tests for a one-way manova. *J. Multivariate Anal.*, 98:1825–1839., 2007.

[4] A.N. Shiryaev. *Probability.* Springer-Verlag, 1984.

[5] M.S. Srivastava. Multivariate theory for analyzing high dimensional data. *J. Japan Statist. Soc.*, 37(1):53–86, 2007.

[6] M.S. Srivastava. A test of the mean vector with fewer observations than the dimension under non-normality. *J. Multivariate Analysis*, (100):518–532, 2009.

[7] M.S. Srivastava and Y. Fujikoshi. Multivariate analysis of variance with fewer observations than the dimension. *J. Multivariate Anal.*, 97:1927–1940, 2006.

[8] M.S. Srivastava and Du. M. A test for the mean vector with fewer observations than the dimension. *J. Multivariate Analysis*, 99:386–402, 2008.

[9] S.S. Wilks. Certain generalizations in the analysis of variance. *Biometrika*, 24:471–494, 1932.

[10] T. Yamada and M.S. Srivastava. A test for the multivariate analysis of variance in high-dimension. *(Unpublished)*, 2011.