ACTA ET COMMENTATIONES UNIVERSITATIS TARTUENSIS DE MATHEMATICA
Volume 17, Number 1, June 2013
Available online at www.math.ut.ee/acta/

On testing the equality of mean vectors in high dimension

Muni S. Srivastava

Abstract. In this article, we review various tests that have been proposed in the literature for testing the equality of several mean vectors. In particular, it includes testing the equality of two mean vectors, the so-called two-sample problem as well as that of testing the equality of several mean vectors, the so-called multivariate analysis of variance or MANOVA problem. The total sample size, however, may be less than the dimension of the mean vectors, and so usual tests cannot be used. Powers of these tests are compared using simulation.

1. Introduction

In this article, we review various tests that have been proposed in the literature for testing the equality of several mean vectors. We begin with the comparison of the mean vectors of two groups with independently distributed p-dimensional observation vectors $\mathbf{x}_{ij}$ and the mean vectors $\boldsymbol{\mu}_i$ and the covariance matrices Σ_i, $i = 1, 2$, $j = 1, \ldots N_i$. The total number of p-dimensional observation vectors, $N = N_1 + N_2$ is less than p. Various tests have been proposed in the literature. For normally distributed observation vectors, and when $\Sigma_1 = \Sigma_2$, a test has been proposed by Dempster [3]. Bai and Saranadasa [1] proposed another test which does not require the assumption of normality but have asymptotically the same power as the one proposed by Dempster [3]. Srivastava [7] proposed a Hotelling's T^2 type test, denoted by T^{+^2}, by using Moore–Penrose inverse of S in place of S^{-1} as $n < p$, $n = N_1 + N_2 - 2$. It may be noted that all the above three tests, namely, T_D, T_{BS} and T^{+^2} are invariant under the group of orthogonal matrices. A test that is invariant under the group of non-singular $p \times p$ diagonal

Received January 12, 2012.
2010 *Mathematics Subject Classification.* 62H10, 62H15.
Key words and phrases. Equality of two mean vectors, high dimensional, multivariate analysis of variance, sample smaller than dimension, inequality of two covariance matrices.
http://dx.doi.org/10.12697/ACUTM.2013.17.03

matrices has recently been proposed by Srivastava and Du [9] and Srivastava [8]. It may be noted that this test is not invariant under the transformation by orthogonal matrices. The power comparison of these tests will be given in this article.

Bai and Saranadasa's [1] test as well as the test proposed by Srivastava and Du [9] can be generalized to the case when the covariance matrices Σ_1 and Σ_2 of the two groups are not equal; for testing the equality of two covariance matrices, see Srivastava and Yanagihara [13]. The generalized versions of these two tests have been considered by Chen and Qin [2], and Srivastava, Katayama and Kano [11].

For testing the equality of the mean vectors of several groups, the so-called multivariate analysis of variance or simply MANOVA, it is assumed that all the groups have the same covariance matrix. Under the assumption that $(p/n) \to c,\ c \in (0, \infty)$, Fujikoshi, Himeno, and Wakaki [4] and Schott [5] have given a generalized version of Dempster [3] and Bai–Saranadasa [1] two-sample tests for the MANOVA problem. Tests that do not require the above assumption have been proposed by Srivastava and Fujikoshi [10]. The above two tests considered by Fujikoshi et al. [4], Schott [5], Srivastava and Fujikoshi [10] require the assumption of normality to obtain the asymptotic distributions of these statistics. Following Srivastava [8], it can, however, be shown that these two tests are robust under a general non-normal distributions. A third test based on the Moore–Penrose Inverse of the sample covariance matrix has been proposed by Srivastava [7] under the assumption of normality. For comparison of the asymptotic powers of these three tests, see Srivastava and Fujikoshi [10].

The above three tests are, however, not invariant under the transformation by non-singular diagonal matrices. A test that has this property has been recently proposed by Yamada and Srivastava [14] under normality assumption.

The organization of this article is as follows. Since it is assumed that $N < p$, we show in Section 2 that there does not exist a test that is invariant under the transformation of the observation vector by any $p \times p$ non-singular matrix. Thus, we shall be considering tests that are invariant under smaller groups. Two such groups are the group of orthogonal matrices and the group of non-singular diagonal matrices. Such tests for the equality of two mean vectors will be given in Section 4 for the non-normal model described in Section 3. In Section 5, the problem of testing the equality of two mean vectors is considered again only under normality but the covariance matrices of the two groups are not equal. The MANOVA problem will be considered in Section 6. The paper concludes in Section 7.

2. Consequences of $N < p$, Σ non-singular

In this section, we show that no invariant test exists under the transformation of the observation vectors by a non-singular matrix. Let

$$\begin{aligned} X &= (\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_N}) : p \times N, N < p, \\ X^* &= (\mathbf{x}_1^*, \mathbf{x}_2^*, \ldots, \mathbf{x}_N^*) : p \times N, N < p \end{aligned}$$

be two sample points. Let

$$Z = (X, X_1) \text{ and } Z^* = (X^*, X_1^*),$$

where X_1 and X_1^* are $p \times (p - N)$ arbitrary matrices chosen such that Z and Z^* are non-singular (n.s.). Thus

$$\begin{aligned} I_p = (Z^*)^{-1} Z^* &= (Z^*)^{-1}(X^*, X_1^*) = [(Z^*)^{-1} X^*, (Z^*)^{-1} X_1^*] \\ &= \begin{pmatrix} I_N & 0 \\ 0 & I_{p-N} \end{pmatrix}. \end{aligned}$$

Hence

$$\begin{aligned} (Z^*)^{-1} X^* &= \begin{pmatrix} I_N \\ 0 \end{pmatrix}, Z(Z^*)^{-1} X^* = (X, X_1) \begin{pmatrix} I_N \\ 0 \end{pmatrix} = X, \\ X &= A X^*, A = Z(Z^*)^{-1}, \end{aligned}$$

where A is n.s.

Thus, for two sample points, X and X^*, on the sample space, there exists a non-singular matrix taking one point to the other. That is the whole sample space is a single orbit; group of non-singular transformations is transitive. Hence, no invariant test under the group of non-singular transformations exists. Clearly then, we need to consider smaller group of transformations, namely invariance under the group of orthogonal matrices and invariance under the group of non-singular diagonal matrices. Thus, tests have been proposed in the literature that are invariant under orthogonal group or invariant under the group of non-singular diagonal matrices. The latter tests appear to perform better than the ones that are invariant under orthogonal group.

3. A non-normal model

To show that the two-sample tests and tests in multivariate analysis of variance (MANOVA) hold when the observation vectors are not necessarily normally distributed, we consider a general model for the independently distributed p-dimensional vectors $\mathbf{x}_{ij},\ j = 1, \ldots, N_i,\ i = 1, 2$. We call this model as Model $\mathcal{M}$ which we describe next.

Model $\mathcal{M}$

In this model, we assume that the observation vectors $\mathbf{x}_{ij}$ satisfy (a), (b), and (c) given below:

(a) $\mathbf{x}_{ij} = \boldsymbol{\mu}_i + F_i \mathbf{u}_{ij},\ j = 1, \dots, N_i,\ i = 1, 2,$
$Cov(\mathbf{x}_{ij}) = \Sigma_i = F_i^2,\ i = 1, 2,$

(b) $E\left[\prod_{k=1}^{p} u_{ijk}^{\nu_k}\right] = \prod_{k=1}^{p} E(u_{ijk}^{\nu_k}),\ \mathbf{u}_{ij} = (u_{ij1}, \dots, u_{ijp})',$
$\nu_k \geq 0,\ \nu_1 + \cdots + \nu_4 \leq 4,\ \nu_i$ are integers,

(c) $E(u_{ijk}^4) = K_4 + 3,\ K_4 \geq -2.$

It may be noted that F_i is the unique $p \times p$ positive definite matrix such that $\Sigma_i = F_i F_i$. The results, however, holds for a general factorization of $\Sigma_i = C_i C_i'$ where C_i is a $p \times p$ non-singular matrix (see Srivastava [8]). For simplicity of presentation we have, however, chosen $\Sigma_i = F_i^2$. Bai and Saranadasa [1] have chosen a model in which $\Sigma_i = \Gamma_i \Gamma_i'$, Γ_i is a $p \times m$ matrix, $m \geq p$; in our model $m = p$. However, the assumptions to prove the asymptotic normality of the test statistic proposed by Bai and Saranadasa [1] under their model are much stronger than for Model $\mathcal{M}$, see Section 4.

4. Two sample tests: $\Sigma_1 = \Sigma_2$

Let $\mathbf{x}_{11}, \dots, \mathbf{x}_{1N_1}$ be independent and identically distributed (i.i.d.) vectors with p-variate normal distribution $N_p(\boldsymbol{\mu}_1, \Sigma_1)$, and $\mathbf{x}_{21}, \dots, \mathbf{x}_{2N_2}$ be i.i.d. $N_p(\boldsymbol{\mu}_2, \Sigma_2)$, where both samples are independently distributed.

The sample mean vectors are, respectively, given by

$$\bar{\mathbf{x}}_1 = \frac{1}{N_1} \sum_{j=1}^{N_1} \mathbf{x}_{ij},\ \bar{\mathbf{x}}_2 = \frac{1}{N_2} \sum_{j=1}^{N_2} \mathbf{x}_{2j},$$

The sample covariance matrices are, respectively, given by

$$S_1 = \frac{1}{n_1} \sum_{j=1}^{N_1} (\mathbf{x}_{1j} - \bar{\mathbf{x}}_{\mathbf{1}})(\mathbf{x}_{1j} - \bar{\mathbf{x}}_{\mathbf{1}})',\ n_1 = N_1 - 1,$$

$$S_2 = \frac{1}{n_2} \sum_{j=1}^{N_2} (\mathbf{x}_{2j} - \bar{\mathbf{x}}_{\mathbf{2}})(\mathbf{x}_{2j} - \bar{\mathbf{x}}_{\mathbf{2}})',\ n_2 = N_2 - 1.$$

When $\Sigma_1 = \Sigma_2 = \Sigma$, an unbiased estimator of Σ is given by

$$S = \frac{n_1 S_1 + n_2 S_2}{n},\ n = n_1 + n_2 = N_1 + N_2 - 2.$$

4.1. Dempster's test. For testing the hypothesis $H : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$ versus $A : \boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2$, four tests have been proposed in the literature by Dempster [3], Bai and Saranadasa [1], Srivastava [7] and Srivastava and Du [9]. The first three tests are invariant under orthogonal transformation and the fourth one is invariant under the transformation by any $p \times p$ non-singular diagonal matrix. We begin with the test proposed by Dempster [3] under the assumption of normality.

In the notation of Section 4, Dempster's test statistic is given by

$$T_D = \left(\frac{1}{N_1} + \frac{1}{N_2}\right)^{-1} (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)^{'} (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)/(\operatorname{tr} S). \tag{4.1}$$

Let G be an $N \times N$ orthogonal matrix given by

$$G = \left[\frac{1}{\sqrt{N}} \mathbf{1_N} | \begin{pmatrix} (N_2/N_1N)^{\frac{1}{2}} \mathbf{1}_{N_1} \\ -(N_1/N_2N)^{\frac{1}{2}} \mathbf{1}_{N_2} \end{pmatrix}, \mathbf{g}_3, \ldots, \mathbf{g}_N \right], \ N = N_1 + N_2,$$

where $GG' = G'G = I_N, \ \mathbf{g_i}^{'}\mathbf{g_i} = 1, \ i = 3, \ldots, N,$ and $\mathbf{g_i}^{'}\mathbf{g_j} = 0, \ i \neq j$. Let

$$X = (\mathbf{x}_{11}, \ldots, \mathbf{x}_{1N_1}, \mathbf{x}_{21}, \ldots, \mathbf{x}_{2N_2}) : p \times N$$

be the $p \times N$ matrix of the observation vectors from the two groups, and

$$XG = (\mathbf{y}_1, \ldots, \mathbf{y}_N).$$

Then

$$\begin{aligned} E(\mathbf{y}_1) &= N^{-\frac{1}{2}}(N_1\boldsymbol{\mu}_1 + N_2\boldsymbol{\mu}_2), \\ E(\mathbf{y}_2) &= (\frac{1}{N_1} + \frac{1}{N_2})^{-\frac{1}{2}}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2), \\ E(\mathbf{y}_j) &= 0, j = 3, \ldots N. \end{aligned}$$

Hence, we can write T_D in terms of $(N-1)$ independent random vectors $\mathbf{y}_2, ..., \mathbf{y}_N$ as

$$T_D = \frac{n\mathbf{y}_2^{'}\mathbf{y}_2}{\mathbf{y}_3^{'}\mathbf{y}_3 + \cdots + \mathbf{y}_N^{'}\mathbf{y}_N}, \quad n = N_1 + N_2 - 2. \tag{4.2}$$

If $\Sigma = \gamma^2 I_p$, then $\mathbf{y}_2, ..., \mathbf{y}_N$ are i.i.d. $N_p(0, \gamma^2 I_p)$ under $\boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$. Hence, $T_D \sim F_{p,np}$. It may be noted that when $\Sigma = \gamma^2 I_{p'}$ and under the assumption of normality Dempster's test T_D is uniformly most powerful among all tests whose power depends on $\boldsymbol{\mu}^{'}\boldsymbol{\mu}/\gamma^2$.

To obtain the distribution of T_D when $\Sigma \neq \gamma^2 I_p$, it is assumed that under the null hypothesis $\mathbf{y}_i^{'}\mathbf{y}_i, \ i = 2, \ldots, p$, are independently distributed as $m\chi_r^2$, where $m > 0, \ r > 0$ are scalar unknown quantities. Clearly, the distributions of T_D will not depend on m and thus, we need to find only r. Dempster gave two iterative methods to obtain an estimate of r. Bai and Saranadasa [1] obtained an experssion for r as follows:

$$E(\mathbf{y}_i'\mathbf{y}_i) = \operatorname{tr}\Sigma = mE(\chi_r^2) = mr$$

and

$$Var(\mathbf{y}_i'\mathbf{y}_i) = 2\operatorname{tr}\Sigma^2 = Var(m\chi_r^2) = 2m^2 r.$$

Hence,

$$r = (\operatorname{tr}\Sigma)^2/(\operatorname{tr}\Sigma^2).$$

Bai and Saranadasa [1], however, did not provide any estimator of r, as they have provided only ratio consistent estimator of $\operatorname{tr}\Sigma^2$ which cannot be used here. An estimator of r is obtained as follows. Let

$$b = (a_1^2/a_2), \text{ where } a_i = \operatorname{tr}\Sigma^i/p, \ i = 1, 2.$$

It may be noted that from Cauchy–Schwarz inequality

$$0 < (\operatorname{tr}\Sigma)^2 \leq p \operatorname{tr}\Sigma^2,$$

the equality on the right hand side holds if and only if $\Sigma = \gamma^2 I_p$, $\gamma > 0$. Hence, $0 < b \leq 1$, and $r \leq p$. Under the assumption of normality of the observation vectors and assuming that $0 < \lim_{p\to\infty} a_i < \infty$, $i = 1, \ldots, 4$, it has been shown by Srivastava [6] that

$$\hat{a}_1 = \frac{\operatorname{tr} S}{p}, \ \hat{a}_2 = \frac{1}{p}[\operatorname{tr} S^2 - \frac{1}{n}(\operatorname{tr} S)^2],$$

are consistent estimators of a_i. Thus, $\hat{b} = (\hat{a}_1^2/\hat{a}_2)$ is a consistent estimator of b giving $\hat{r} = p\hat{b}$, and

$$T_D \simeq F_{[\hat{r}],[n\hat{r}]}.$$

Thus, under the hypothesis that $\boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$, Dempster's T_D statistic for normally distributed observations is approximately distributed as an F-distribution with $[\hat{r}]$ and $[n\hat{r}]$ degrees of freedom, where $[a]$ denotes the largest integer value $\leq a$. Clearly then the hypothesis of equality of two mean vectors is rejected if $T_D > F_{[\hat{r}][n\hat{r}],1-\alpha}$, the $100(1-\alpha)\%$ point of the F-distribution with degrees of freedom mentioned above.

The asymptotic power of the T_D test has been derived by Bai and Saranadasa [1] who proposed another test and showed that the asymptotic power of the T_D test is the same as the one proposed by them. This test will be called Bai–Saranadasa test in our discussion which we describe next.

4.2. Bai–Saranadasa test. Consider an asymptotic version of Dempster's statistic. The mean of the numerator is given by

$$\begin{aligned} & \left(\frac{1}{N_1}+\frac{1}{N_2}\right)^{-1} E\left[(\bar{\mathbf{x}}_1-\bar{\mathbf{x}}_2)'(\bar{\mathbf{x}}_1-\bar{\mathbf{x}}_2)\right] \\ = & \left(\frac{1}{N_1}+\frac{1}{N_2}\right)^{-1} \left[E(\bar{\mathbf{x}}_1'\bar{\mathbf{x}}_1 - 2\bar{\mathbf{x}}_1'\bar{\mathbf{x}}_2 + \bar{\mathbf{x}}_2'\bar{\mathbf{x}}_2)\right] \\ = & \left(\frac{1}{N_1}+\frac{1}{N_2}\right)^{-1} \left[\frac{\operatorname{tr}\Sigma}{N_1} + \boldsymbol{\mu}_1'\boldsymbol{\mu}_1 - 2\boldsymbol{\mu}_1'\boldsymbol{\mu}_2 + \frac{\operatorname{tr}\Sigma}{N_2} + \boldsymbol{\mu}_2'\boldsymbol{\mu}_2\right] \\ = & \operatorname{tr}\Sigma + \tau(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)'(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2), \ \tau = N_1N_2/(N_1+N_2). \end{aligned}$$

So the numerator may be estimated by

$$\left[\left(\frac{1}{N_1}+\frac{1}{N_2}\right)^{-1} (\bar{\mathbf{x}}_1-\bar{\mathbf{x}}_2)'(\bar{\mathbf{x}}_1-\bar{\mathbf{x}}_2) - \operatorname{tr} S\right].$$

The asymptotic variance of this estimator is $2\operatorname{tr}\Sigma^2 + o(1)$. Under the assumption that $\lim(p/n) = c$, $0 < c < \infty$ and $\lambda_{max}(\Sigma) = o(p^{-\frac{1}{2}})$, Bai and Saranadasa [1] showed that as $(n,p) \to \infty$,

$$(\operatorname{tr}\Sigma^2)^{-1}\left[\operatorname{tr} S^2 - \frac{1}{n}(\operatorname{tr} S)^2\right] \xrightarrow{p} 1. \tag{4.3}$$

Thus, using the ratio consistent estimator of $(\operatorname{tr}\Sigma^2)$ given in (4.3), Bai and Saranadasa [1] proposed the statistic

$$T_{BS} = \frac{\left(\frac{1}{N_1}+\frac{1}{N_2}\right)^{-1}(\bar{\mathbf{x}}_1-\bar{\mathbf{x}}_2)'(\bar{\mathbf{x}}_1-\bar{\mathbf{x}}_2) - \operatorname{tr} S}{\sqrt{2[\operatorname{tr} S^2 - \frac{1}{n}(\operatorname{tr} S)^2]}} \tag{4.4}$$

for testing the hypothesis that $\boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$. They also showed that under the hypothesis T_{BS} is normally distributed with mean zero and variance 1 for a general model described in Section 3 that includes the normal model as a special case.

To obtain the asymptotic distribution of T_{BS} under the alternative hypothesis $A = \boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2$, it is assumed that the difference between the two mean vectors satisfy the following conditions:

$$\left(\frac{1}{N_1}+\frac{1}{N_2}\right)^{-1}(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)'\Sigma(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2) = o(\operatorname{tr}\Sigma^2), \tag{4.5}$$

$$(p/n) \to c > 0 \text{ and } \lambda_{\max}(\Sigma) = o(\sqrt{\operatorname{tr}\Sigma^2}). \tag{4.6}$$

Under the conditions (4.5) and (4.6), Bai and Saranadasa [1] showed that the asymptotic power of Dempster's test T_D and Bai and Saranadasa's test

T_{BS} are equal and is given by

$$\beta(T_D) = \beta(T_{BS}) = \Phi\left(-z_{1-\alpha} + \frac{N_1N_2}{(N_1+N_2)}\frac{(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)'(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)}{\sqrt{2pa_2}}\right), \quad (4.7)$$

where Φ is the distribution function of the standard normally distributed random variable with mean 0 and variance 1, $\Phi(z_{1-\alpha}) = 1-\alpha,\ 0<\alpha<1$.

It may be recalled that both tests T_D as well as T_{BS} are one-sided tests: Reject $H : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$, if $T_D > z_{1-\alpha}$; similarly reject H if $T_{BS} > z_{1-\alpha}$.

4.3. Srivastava's T^{+^2} test. Using Moore–Penrose of S, Srivastava [7] proposed the statistic

$$T^{+2} = \left(\frac{1}{N_1}+\frac{1}{N_2}\right)^{-1}(\bar{\mathbf{x}}_1-\bar{\mathbf{x}}_2)'S^+(\bar{\mathbf{x}}_1-\bar{\mathbf{x}}_2), \quad (4.8)$$

where S^+ is the Moore–Penrose inverse of S, which is unique and satisfies the following conditions: $S^+SS^+= S^+$, $SS^+S=S$, SS^+ and S^+S are symmetric matrices. Let $S = \frac{1}{n}V,\ n = N_1+N_2-2$. It can be shown that the T^{+^2} statistic is invariant under the transformation $\mathbf{x}_{ij} \to c\Gamma\mathbf{x}_{ij},\ c \neq 0,\ \Gamma\Gamma' = I_p$. Thus, without loss of generality, we may assume that $\Sigma = \Lambda$, a diagonal matrix. Note that

$$S^+ = nV^+ = nH'L^{-1}H, HH' = I_n,$$

where $L = \operatorname{diag}(l_1,\ldots,l_n)$, the non-zero eigenvalues of V. Thus,

$$T^{+2} = n\left(\frac{1}{N_1}+\frac{1}{N_2}\right)^{-1}(\bar{\mathbf{x}}_1-\bar{\mathbf{x}}_2)'H'L^{-1}H(\bar{\mathbf{x}}_1-\bar{\mathbf{x}}_2). \quad (4.9)$$

Let

$$\mathbf{z} = \left(\frac{1}{N_1}+\frac{1}{N_2}\right)^{-\frac{1}{2}}AH(\bar{\mathbf{x}}_1-\bar{\mathbf{x}}_2)',$$

where

$$A = (H\Lambda H')^{-\frac{1}{2}}.$$

Hence, under the hypothesis that $\boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$,

$$T^{+2} = n\mathbf{z}'(ALA)^{-1}\mathbf{z},\ \mathbf{z} \sim N_n(\mathbf{0}, I_n). \quad (4.10)$$

In order to obtain the asymptotic distribution of T^{+^2}, it is assumed that

$$0 < \lim_{p\to\infty} a_i < \infty,\ i = 1,\ldots,4,\ \text{where } a_i = (\operatorname{tr}\Sigma^i/p).$$

Under the above assumption it has been shown in Srivastava [7], that as $p\to\infty,\ p^{-1}ALA \xrightarrow{p} bI_n$, where $b = (a_1^2/a_2)$. Thus, under the hypothesis, as $p\to\infty$,

$$\frac{bpT^{+2}}{n} \xrightarrow{d} \mathbf{z}'\mathbf{z},$$

where "d" stands for "in distribution". Since, under the hypothesis $\mathbf{z}'\mathbf{z}$ has a chi-square distribution with n degrees of freedom denoted by χ^2_n $(bp/n){T^+}^2$ is asymptotically distributed as χ^2_n when $p \to \infty$. Although b is unknown, a consistent estimator of b is given by $\hat{b} = (\hat{a}_1^2/\hat{a}_2)$ as p and n go to infinity. Thus, *approximately* $(\hat{b}p/n){T^+}^2$ is distributed as χ^2_n and the hypothesis is rejected if

$$(\hat{b}p/n){T^+}^2 > \chi^2_{n,1-\alpha}, \tag{4.11}$$

where $P(\chi^2_n < \chi^2_{n,1-\alpha}) = 1 - \alpha$. Alternatively, we may consider the asymptotic distribution of $(\hat{b}p/n){T^+}^2$ as $p \to \infty$, and then $n \to \infty$. That is, we consider the standardized statistic

$$T_S^+ = \frac{(\hat{b}p/n){T^+}^2 - n}{\sqrt{2n}} \tag{4.12}$$

which is asymptotically distributed as standard normal. Thus, the hypothesis is rejected if $T_S^+ > z_{1-\alpha}$, where $\Phi(z_{1-\alpha}) = 1 - \alpha$. It may be noted that for moderate sample size, the approximate distribution in (4.11) may give a better approximation than (4.12).

To obtain the asymptotic power of T_S^+ test, we consider local alternative in which,

$$(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2) = (\tau n)^{-\frac{1}{2}}\boldsymbol{\delta}, \tag{4.13}$$

where $\boldsymbol{\delta}$ is fixed, $n = N_1 + N_2 - 2$, and $\tau = N_1N_2/(N_1+N_2)$. The asymptotic power of the T_S^+ statistic is given by

$$\beta(T_S^+) = \Phi\left[-z_{1-\alpha} + \left(\frac{n}{2}\right)^{\frac{1}{2}} \frac{\tau(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)'\Lambda(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)}{pa_2}\right]. \tag{4.14}$$

It may be noted that all the above three tests, namely, T_D, T_{BS} and ${T^+}^2$ or equivalently T_S^+ are invariant under the transformations

$$\mathbf{x}_{ij} \to a\Gamma\mathbf{x}_{ij},\ a \neq 0,\ \Gamma\Gamma' = I_p.$$

It has been shown by Bai and Saranadasa [1] that the tests T_D and T_{BS} have same asymptotic power given in (4.8). Comparing it with the power of T_S^+ given in (4.13), we find that the test T_S^+ may be preferred if

$$\left(\frac{n}{pa_2}\right)^{\frac{1}{2}} \boldsymbol{\theta}'\Lambda\boldsymbol{\theta} > \boldsymbol{\theta}'\boldsymbol{\theta},\ \boldsymbol{\theta} = (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2) \tag{4.15}$$

For example, if $\boldsymbol{\theta} \sim N_p(0, \Lambda)$, then on the average (4.15) implies that

$$\left(\frac{n}{pa_2}\right)^{\frac{1}{2}} \operatorname{tr}\Lambda^2 > \operatorname{tr}\Sigma,$$

that is

$$n > (a_1^2/a_2^2)p = bp. \tag{4.16}$$

Since $a_1^2 < a_2$, many such n exist for large p. Similarly, if $\boldsymbol{\theta} = (\sqrt{\lambda_1}, \dots, \sqrt{\lambda_p})'$, where $\lambda_1 > \cdots > \lambda_p$ are the eigenvalues of Σ, the same inequality given in (4.16) is obtained.

4.4. Srivastava–Du test. All the above three tests are invariant under the transformations

$$\mathbf{x}_{ij} \to a\Gamma\mathbf{x}_{ij},\ a \neq 0,\ \Gamma\Gamma' = I_p,$$

but not invariant under the transformations

$$\mathbf{x}_{ij} \to D\mathbf{x}_{ij},$$

where D is a $p \times p$ non-singular diagonal matrix,

$$D = \operatorname{diag}(d_1, \dots d_p).$$

This implies that change of unit of measurements will affect all the above three statistics. Srivastava and Du [9] proposed a statistic that is invariant under the transformation by any $p \times p$ non-singular diagonal matrix. This test statistic is given by

$$T_{SD} = \frac{\left(\frac{1}{N_1} + \frac{1}{N_2}\right)^{-1} (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)' D_S^{-1} (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2) - p}{\left[2\left\{\operatorname{tr}\hat{R}^2 - (\frac{p^2}{n})\right\} C_{p,n}\right]^{\frac{1}{2}}},$$

where

$$\hat{R} = D_S^{-\frac{1}{2}} S D_S^{-\frac{1}{2}},\ D_S = \operatorname{diag}(S_{11}, \dots, S_{pp}), S = (S_{ij}),$$

$$C_{p,n} = 1 + \left(\frac{\operatorname{tr}\hat{R}^2}{p^{\frac{3}{2}}}\right) \xrightarrow{p} 1,\ \text{as } (n,p) \to \infty.$$

Under normality assumptions this test was proposed by Srivastava and Du [9]. It has been shown to be robust by Srivastava [8]. Let

$$\Sigma = (\sigma_{ij}) = D_\sigma^{\frac{1}{2}} R D_\sigma^{\frac{1}{2}}, \tag{4.17}$$

where $D_\sigma^{\frac{1}{2}} = \operatorname{diag}(\sigma_{11}^{\frac{1}{2}}, \dots, \sigma_{pp}^{\frac{1}{2}})$. Then R is the correlation matrix which can be estimated by $\hat{R}$ defined above. In order to obtain the distribution of T_{SD}, Srivastava and Du [9] assumed the following:

$$\text{(i)} \quad 0 < \lim_{p\to\infty} \left(\frac{\operatorname{tr} R^i}{p}\right) < \infty,\ i = 1, \dots, 4, \tag{4.18}$$

$$\text{(ii)} \quad \lim_{p\to\infty} \max_{1\le i\le p} \frac{\lambda_{ip}}{\sqrt{p}} = 0, \tag{4.19}$$

where $\lambda_{1p}, \dots, \lambda_{pp}$ are the eigenvalues of the correlation matrix R. Under the hypothesis of equality of two mean vectors, T_{SD} has asymptotically standard normal distribution. Under local alternative defined by

$$\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2 = (\tau n)^{-\frac{1}{2}}\boldsymbol{\delta},\ \tau = N_1N_2/(N_1+N_2),\ n = N_1+N_2-2,$$

where $\boldsymbol{\delta}$ is fixed, the asymptotic distribution of the statistic T_{SD} is given by

$$\lim_{(n,p)\to\infty} P(T_{SD} > z_{1-\alpha}) = \Phi\left[-z_{1-\alpha} + \frac{\tau(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)' D_\sigma^{-1}(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)}{\sqrt{2\text{tr}\, R^2}}\right].$$

It may be noted that Yamada and Srivastava [14] have shown that the condition (4.19) is not needed in deriving the asymptotic distribution of the statistic T_{SD}.

Srivastava and Du [9] showed theoretically that under the condition

$$0 < \lim_{p\to\infty} \frac{\boldsymbol{\delta}'\boldsymbol{\delta}}{p} = \lim_{p\to\infty} \frac{\boldsymbol{\delta}' D_\sigma^{-1}\boldsymbol{\delta}}{\text{tr}\, D_\sigma^{-1}} < \infty \tag{4.20}$$

the T_{SD} test is superior to T_{BS} test. It may be noted that the condition (4.20) is satisfied for $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_p)'$ in which $\delta_i = \delta \neq 0,\ i = 1, \ldots, p$, that is, all the components of the random vector have the same mean.

In the next two sections, we compare the attained significance level and power of the two robust tests, namely T_D and T_{SD} through simulation.

4.5. Attained significance level (ASL). To compare the three tests, we need to define the attained significance levels and the empirical powers. Let $z_{1-\alpha}$ be a $100(1-\alpha)\%$ quantile of the asymptotic null distribution of the test statistic T which is N(0,1) in our case; thus $z_{1-\alpha}$ is the $100(1-\alpha)\%$ quantile of $N(0,1)$. With m replications of the data set simulated under the null hypothesis, the ASL is

$$\hat{\alpha} = \frac{(\#\text{ of } t_H \geq z_{1-\alpha})}{m},\ m = 10,000,\ \alpha = 0.05,$$

where t_H represents the values of the test statistic T based on the data sets simulated under the null hypothesis. $\hat{\alpha}$ is approximately distributed as Binomial(10000, 0.05) and has the standard deviation estimated by

$$\hat{se}(\hat{\alpha}) = \sqrt{0.05 \times 0.95/10,000} \simeq 0.0022.$$

For simplicity of presentation, we shall consider one-sample case in which we test that the mean vector is zero.

4.6. Empirical power. To compute the empirical powers, we shall use the empirical critical points. Specifically, we first simulate m replications of the data set under the null hypothesis, then select the $(m\alpha)^{th}$ largest value of the test statistic as the empirical critical point, denoted as $\hat{z}_{1-\alpha}$, that is, the $100(1-\alpha)\%$ quantile of the empirical null distribution of the test statistic obtained from the m replications. Then another m replications of the data set are simulated under the alternative with the given choice of $\boldsymbol{\mu}$.

The empirical power is calculated by

$$\hat{\beta} = \frac{(\# \text{ of } t_A \geq \hat{z}_{1-\alpha})}{m}.$$

TABLE 1. Attained significance levels of T_{SD} and T_D under the null hypothesis, when $R = I_p$ and $R = R_1$, respectively

		$D_\sigma = I_p$		$D_\sigma = D_{\sigma,1}$		$D_\sigma = D_{\sigma,2}$	
p	N	T_{SD}	T_D	T_{SD}	T_D	T_{SD}	T_D
$R = I_p$							
60	30	0.056	0.051	0.056	0.049	0.056	0.052
100	40	0.048	0.054	0.048	0.054	0.048	0.052
	60	0.050	0.053	0.050	0.053	0.050	0.051
	80	0.050	0.054	0.050	0.054	0.050	0.054
150	40	0.050	0.054	0.050	0.053	0.050	0.051
	60	0.048	0.049	0.048	0.048	0.048	0.053
	80	0.049	0.050	0.049	0.051	0.049	0.053
200	40	0.045	0.052	0.045	0.052	0.045	0.051
	60	0.048	0.052	0.045	0.052	0.045	0.051
	80	0.045	0.047	0.045	0.050	0.045	0.051
400	40	0.037	0.046	0.037	0.046	0.037	0.054
	60	0.035	0.050	0.035	0.051	0.035	0.053
	80	0.044	0.050	0.044	0.049	0.044	0.046
$R = R_1$							
60	30	0.058	0.058	0.058	0.057	0.0582	0.059
100	40	0.053	0.062	0.053	0.063	0.0526	0.061
	60	0.045	0.062	0.045	0.061	0.0450	0.057
	80	0.046	0.057	0.046	0.057	0.0463	0.060
150	40	0.049	0.065	0.049	0.065	0.0493	0.063
	60	0.050	0.064	0.050	0.063	0.0502	0.061
	80	0.044	0.059	0.044	0.059	0.0441	0.062
200	40	0.049	0.067	0.049	0.068	0.0485	0.064
	60	0.044	0.061	0.044	0.062	0.0441	0.060
	80	0.047	0.062	0.047	0.063	0.0469	0.063
400	40	0.045	0.063	0.045	0.063	0.0448	0.066
	60	0.041	0.064	0.041	0.063	0.0408	0.059
	80	0.036	0.062	0.036	0.063	0.0362	0.058

TABLE 2. Empirical powers of T_{SD} and T_D under the alternative hypothesis, when $R = I_p$ and $R = R_1$, respectively

		$D_\sigma = I_p$		$D_\sigma = D_{\sigma,1}$		$D_\sigma = D_{\sigma,2}$	
p	N	T_{SD}	T_D	T_{SD}	T_D	T_{SD}	T_D
$R = I_p$							
60	30	0.999	1.000	0.287	0.291	0.976	0.542
100	40	1.000	1.000	0.622	0.593	1.000	0.939
	60	1.000	1.000	0.880	0.861	1.000	0.998
	80	1.000	1.000	0.973	0.962	1.000	1.000
150	40	1.000	1.000	0.698	0.661	1.000	0.962
	60	1.000	1.000	0.932	0.913	1.000	1.000
	80	1.000	1.000	0.991	0.983	1.000	1.000
200	40	1.000	1.000	0.831	0.789	1.000	0.992
	60	1.000	1.000	0.979	0.967	1.000	1.000
	80	1.000	1.000	0.999	0.998	1.000	1.000
400	40	1.000	1.000	0.926	0.903	1.000	1.000
	60	1.000	1.000	0.998	0.995	1.000	1.000
	80	1.000	1.000	1.000	1.000	1.000	1.000
$R = R_1$							
60	30	0.789	0.948	0.088	0.014	0.513	0.112
100	40	0.990	1.000	0.125	0.013	0.936	0.290
	60	1.000	1.000	0.203	0.067	1.000	0.813
	80	1.000	1.000	0.303	0.261	1.000	0.973
150	40	0.980	1.000	0.115	0.003	0.974	0.143
	60	1.000	1.000	0.143	0.024	1.000	0.681
	80	1.000	1.000	0.245	0.142	1.000	0.958
200	40	0.984	1.000	0.108	0.000	0.944	0.148
	60	1.000	1.000	0.157	0.016	1.000	0.792
	80	1.000	1.000	0.249	0.096	1.000	0.985
400	40	0.933	1.000	0.085	0.000	0.814	0.018
	60	1.000	1.000	0.118	0.000	0.995	0.589
	80	1.000	1.000	0.194	0.003	1.000	0.984

4.7. Parameter selection: one-sample case. We consider both independent correlation structures $R = I_p = \text{diag}\,(1, 1, \ldots 1)$ and equal correlation structure $R = R_1 = (\rho_{ij}) : \rho_{ij} = 0.25, i \neq j$. We also consider different scalar matrix $D_\sigma = \text{diag}\,(\sigma_{11}, \ldots, \sigma_{pp})$. We select $D_\sigma = I_p, D_\sigma = D_{\sigma,1} : \sigma_{11}^{\frac{1}{2}}, \ldots, \sigma_{pp}^{\frac{1}{2}} \overset{iid}{\sim} Unif(2,3)$ and $D_\sigma = D_{\sigma,2} : \sigma_{11}, \ldots, \sigma_{pp} \overset{iid}{\sim} \chi_3^2$.

For the alternative hypothesis, we choose $\boldsymbol{\mu} = \nu = (\nu_1, \ldots \nu_p)' : \nu_{2k-1} = 0$ and $\nu_{2k} \overset{iid}{\sim} Unif(-\frac{1}{2}, \frac{1}{2}), k = 1, \ldots, \frac{p}{2}$.

5. Two sample tests: $\Sigma_1 \neq \Sigma_2$

In this section, we consider the problem of testing the equality of the mean vectors of two groups when the covariance matrices of the two groups are not equal. For normally distributed observation vectors, the equality of two covariance matrices can be ascertained using a test proposed by Srivastava and Yanagihara [13]. And if it is found that the covariance matrices of the two groups are not equal, the tests given in this section should be used. We begin with a test proposed by Chen and Qin [2].

5.1. Chen–Qin test statistic. In the notation of Section 4, this test statistic is given by

$$T_{cq} = \frac{\frac{1}{N_1 n_1}\sum_{i\neq j}^{N_1} \mathbf{x}_{1i}^{'}\mathbf{x}_{1j} + \frac{1}{N_2 n_2}\sum_{i\neq j}^{N_2} \mathbf{x}_{2i}^{'}\mathbf{x}_{2j} - 2\bar{\mathbf{x}}_{\mathbf{1}}^{'}\bar{\mathbf{x}}_{\mathbf{2}}}{\left[\frac{2}{N_1 n_1}\widehat{\operatorname{tr}\Sigma_1^2} + \frac{2}{N_2 n_2}\widehat{\operatorname{tr}\Sigma_2^2} + \frac{4}{N_1 n_2}\widehat{\operatorname{tr}\Sigma_1\Sigma_2}\right]^{\frac{1}{2}}},$$

where

$$\widehat{\operatorname{tr}\Sigma_i^2} = \frac{1}{N_i n_i}\operatorname{tr}\left\{\sum_{j\neq k}^{N_i}(\mathbf{x}_{ij} - \bar{\mathbf{x}}_{\mathbf{i(j,k)}})\mathbf{x}_{ij}^{'}(\mathbf{x}_{ik} - \bar{\mathbf{x}}_{i(jk)}\mathbf{x}_{ik}^{'})\right\},\ i = 1, 2,$$

$$\widehat{\operatorname{tr}\Sigma_1\Sigma_2} = \frac{1}{N_1 N_2}\operatorname{tr}\left\{\sum_{j=1}^{N_1}\sum_{k=1}^{N_2}(\mathbf{x}_{1j} - \bar{\mathbf{x}}_{\mathbf{1(j)}})\mathbf{x}_{1j}^{'}(\mathbf{x}_{2k} - \bar{\mathbf{x}}_{\mathbf{2(k)}})\mathbf{x}_{2k}^{'}\right\},$$

$$\bar{\mathbf{x}}_{i(j,k)} = \frac{1}{N_i - 2}(N_i\bar{\mathbf{x}}_i - \mathbf{x}_{ij} - \mathbf{x}_{ik}),\ i = 1, 2;\ j, k = 1, \ldots, N_i,$$

$$\bar{\mathbf{x}}_{i(k)} = \frac{1}{n_i}(N_i\bar{\mathbf{x}}_i - \mathbf{x}_{ik}),\ i = 1, 2;\ k = 1, \ldots . N_i.$$

In order to derive the asymptotic distribution of the statistic T_{cq}, Chen and Qin [2] made the following assumption.

Assumption (A)

A(1) $\mathbf{x}_{ij} = \Gamma_i\mathbf{z}_{ij} + \boldsymbol{\mu}_i,\ \ j = 1, \ldots, N_i,\ \ i = 1, 2,$
where each Γ_i is a $p \times m$ matrix for some $m \geq p$ such that $\Gamma_i\Gamma_i^{'} = \Sigma_i$, and $\{\mathbf{z}_{ij}\}_{j=1}^{N_i}$ are m-variate independent and identically distributed random vectors satisfying $E(\mathbf{z}_{ij}) = \mathbf{0}$, $Cov(\mathbf{z}_{ij}) = I_m$, and for $\mathbf{z}_{ij} = (z_{ij1}, \ldots, z_{ijm})^{'}$, it is assumed that $E(z_{ijk}^4) = K_4 + 3 < \infty$, and
$E(z_{ijl_1}^{\alpha_1}, z_{ijl_2}^{\alpha_2}, \ldots, z_{ijl_q}^{\alpha_q}) = \prod_{r=1}^{q} E(z_{ijl_r}^{\alpha_r}),\ \ \sum_{r=1}^{q}\alpha_r \leq 8,$
and $l_1 \neq l_2 \neq \cdots \neq l_q$,

A(2) $\lim_{p\to\infty} \frac{\mathrm{tr}\,\Sigma_i\Sigma_j\Sigma_k\Sigma_l}{\{\mathrm{tr}\,(\Sigma_1+\Sigma_2)\}^2} = 0, \quad i,j,k,l = 1 \text{ or } 2,$

A(3) $N(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)'\Sigma_i(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2) = o\left\{\mathrm{tr}\,(\Sigma_1 + \Sigma_2)^2\right\}.$

Chen and Qin (2010) proved the asymptotic normality of T_{cq} under the hypothesis that $\boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$ and under the Assumptions (A1)–(A2) given above. They also obtained the asymptotic power under Assumptions (A1)–(A3). It is given by

$$\lim_{(N_1,N_2,p)\to\infty} P_1(T_{cq} > z_{1-\alpha}) = \Phi\left(-z_{1-\alpha} + \frac{(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)'(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)}{\sigma^2_{n,p}}\right),$$

where

$$\sigma^2_{n,p} = \frac{2}{N_1 n_1}\mathrm{tr}\,(\Sigma_1^2) + \frac{2}{N_2 n_2}\mathrm{tr}\,(\Sigma_2^2) + \frac{4}{N_1 N_2}\mathrm{tr}\,(\Sigma_1\Sigma_2).$$

It may be noted that $\sigma^2_{n,p}$ is the variance of the numerator of T_{cq} which can also be estimated by using the usual consistent estimators of $\mathrm{tr}\,\Sigma_i^2/p$ and $\mathrm{tr}\,\Sigma_1\Sigma_2/p,\ i = 1,2$. Most importantly note that the numerator of T_{cq},

$$\begin{aligned} & \frac{1}{N_1 n_1}\sum_{i\neq j}^{N_1} \mathbf{x}_{1i}'\mathbf{x}_{1j} + \frac{1}{N_2 n_2}\sum_{i\neq j}^{N_2} \mathbf{x}_{2i}'\mathbf{x}_{2j} - 2\mathbf{x}_1'\mathbf{x}_2 \\ = \; & (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)'(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2) - \frac{1}{N_1}\mathrm{tr}\,S_1 - \frac{1}{N_2}\mathrm{tr}\,S_2, \end{aligned}$$

which is identical to the one obtained by generalizing Bai and Saranadasa test to the case when $\Sigma_1 \neq \Sigma_2$, and much simpler to compute. Thus, Srivastava, Katayama and Kano [11] proposed a simpler test given in the next subsection in which a different consistent estimator of the variance of the numerator is used.

5.2. A simpler test than T_{cq}. Chen–Qin test T_{cq} has rather complicated expressions and takes much longer time in computing with no apparent advantage in terms of power than the corresponding simpler test

$$T_2 = \frac{p^{-\frac{1}{2}}\left[(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)'(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2) - N_1^{-1}\mathrm{tr}\,S_1 - N_2^{-1}\mathrm{tr}\,S_2\right]}{\left[2N_1^{-2}\hat{a}_{21} + 2N_2^{-2}\hat{a}_{22} + 4(N_1N_2p)^{-1}\mathrm{tr}\,S_1S_2\right]^{\frac{1}{2}}},$$

$$\hat{a}_{2i} = \frac{1}{p}\left[\mathrm{tr}\,S_i^2 - \frac{1}{n_i}(\mathrm{tr}\,S_i)^2\right],\ n_i = N_i - 1,\ i = 1,2.$$

The test T_2 has the same distribution as T_{cq}. The power and ASL for both tests, are indistinguishable, as seen in the attached tables obtained from simulation.

5.3. Srivastava–Katayama–Kano test. It may be noted that the tests T_{cq} and T_u are both invariant under the transformation by an orthogonal matrix but not invariant under the transformation by any non-singular diagonal matrix. In this subsection we describe a test which is invariant under the group. Let

$$\begin{aligned} S_i &= (S_{ijk}), \hat{D}_i = \text{diag}\,(S_{i11}, \ldots, S_{ipp}),\ i = 1, 2, \\ \hat{D} &= N_1^{-1}\hat{D}_1 + N_2^{-1}\hat{D}_2, \\ \hat{R} &= \hat{D}^{-\frac{1}{2}}(N_1^{-1}S_1 + N_2^{-1}S_2)\hat{D}^{-\frac{1}{2}} = (r_{ij}), \\ q_n &= [(\bar{\mathbf{x}}_{\mathbf{1}} - \bar{\mathbf{x}}_{\mathbf{2}})'\hat{D}^{-1}(\bar{\mathbf{x}}_{\mathbf{1}} - \bar{\mathbf{x}}_{\mathbf{2}}) - p]. \end{aligned}$$

Let $\sigma^2(q_n) = Var(q_n) = 2trR^2$, where R is the population version of the sample correlation matrix $\hat{R}$. Then the statistic proposed by Srivastava, Katayama and Kano [11] is given by

$$T_1 = \frac{q_n}{[C_{p,n}\hat{\sigma}^2(q_n)]^{\frac{1}{2}}},\ n_i = N_i - 1,\ i = 1, 2,$$

where

$$C_{p,n} = 1 + \frac{\text{tr}\,\hat{R}^2}{p^2} \to 1,\ \text{as } (n, p) \to \infty,$$

and

$$\hat{\sigma}^2(q_n) = 2\left[(\text{tr}\,\hat{R}^2) - \frac{1}{n_1 N_1^2}(\text{tr}\,\hat{D}^{-1}S_1)^2 - \frac{1}{n_2 N_2^2}(\text{tr}\,\hat{D}^{-1}S_2)^2\right].$$

It can be shown that under the **Assumptions** (B1)–(B3) stated below, $\hat{\sigma}^2(q_n)$ is a ratio consistent estimator of $\sigma^2(q_n)$: namely $[\hat{\sigma}^2(q_n)/\sigma^2(q_n)] \xrightarrow{p} 1$.

Assumption (B)

(B1) $0 < c_1 < \min_{1 \le k \le p} \sigma_{ikp} < \max_{1 \le k \le p} \sigma_{ikk} < c_2 < \infty$ uniformly for all p,
(B2) $\lim_{p\to\infty}[\text{tr}\,R^4/(\text{tr}\,R^2)^2] = 0$,
(B3) $(N_1/N) \to k \in (0, 1)$ as $N = N_1 + N_2 \to \infty$,
(B4) $N_m = O(p^\delta),\ \delta > \frac{1}{2},\ N_m = \min(N_1, N_2)$,
(B5) $(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)'D^{-\frac{1}{2}}RD^{-\frac{1}{2}}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2) = o(\text{tr}\,R^2)$,

where

$$\begin{aligned} R &= D^{-\frac{1}{2}}(N_1^{-1}\Sigma_1 + N_2^{-1}\Sigma_2)D^{-\frac{1}{2}}, \Sigma_i = (\sigma_{ijk}),\ i = 1, 2, \\ D &= (N_1^{-1}D_1 + N_2^{-1}D_2), D_i = \operatorname{diag}(\sigma_{i11}, \ldots, \sigma_{ipp}). \end{aligned}$$

It can be shown that when $\boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$, and under the Assumptions (B1)–(B4), we get the following theorem.

Theorem 5.1. *Under the Assumptions* (B1)–(B4)*, when* $\boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$*,*

$$P_0\{T_1 \le z_{1-\alpha}\} \to \Phi(z_{1-\alpha}) \text{ as } (N_1, N_2, p) \to \infty.$$

Srivastava, Katayama and Kano [11] obtained this result assuming normality of the observation vectors. To obtain the distribution under the alternative hypothesis, we choose the local alternative given in (B5), and obtain the following theorem.

Theorem 5.2. *Under the Assumption* (B)*, the distribution of the statistic* T_1 *under the local alternative* (B5) *is given by*

$$P_1\{T_1 < -z_{1-\alpha}\} \to \Phi\left\{-z_{1-\alpha} + \frac{(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)' D^{-1}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)}{2\operatorname{tr} R^2}\right\}.$$

5.4. Comparison of T_{cq}, T_1 and T_2 tests: simulation. The parameters of the simulation are given by

$$\begin{aligned} N_1 &= N_2 = N/2, \\ \boldsymbol{\mu}_1 &= \boldsymbol{\mu}_2 = \mathbf{0} \text{ for ASL}, \\ \boldsymbol{\mu}_1 &= \mathbf{0},\ \boldsymbol{\mu}_2 = (u_1, \ldots, u_p)^T, u_i \text{ i.i.d. } \mathrm{U}(\frac{1}{2}, \frac{3}{2}), \\ \Sigma_1 &= \operatorname{diag}(d_1, \ldots, d_p)\ \mathbf{R}_1 \operatorname{diag}(d_1, \ldots, d_p), \\ \Sigma_2 &= \operatorname{diag}(\psi_1, \ldots, \psi_p)\ \mathbf{R}_2 \operatorname{diag}(\psi_1, \ldots, \psi_p), \\ d_i &= 2 + \frac{p-i+1}{p},\ \psi_i^2 \text{ i.i.d. } \chi_3^2, \\ \mathbf{R}_1 &= (r_{ij}) \text{ with } r_{ij} = (-1)^{i+j}(0.2)^{|i-j|^{0.1}}, \\ \mathbf{R}_2 &= (\rho_{ij}) \text{ with } \rho_{ij} = (-1)^{i+j}(0.4)^{|i-j|^{0.1}}. \end{aligned}$$

Table 3. ASL and power for normal distribution: 1000 simulations

				ASL			Power
p	$N_1 = N_2$	T_1	T_{cq}	T_2	T_1	T_{cq}	T_2
100	20	0.069	0.089	0.089	0.824	0.395	0.396
	30	0.067	0.075	0.075	0.979	0.727	0.731
	40	0.059	0.072	0.071	0.998	0.884	0.883
150	20	0.054	0.063	0.063	0.966	0.621	0.622
	30	0.058	0.061	0.062	0.982	0.654	0.655
	40	0.052	0.061	0.061	1.000	0.964	0.963
200	20	0.073	0.079	0.080	0.928	0.628	0.624
	30	0.061	0.073	0.074	0.997	0.852	0.854
	40	0.069	0.086	0.086	1.000	0.858	0.856

Table 4. ASL and power for χ^2_8: 10,000 simulations

p	$N_1 = N_2$	ASL_$T1$	ASL_$T2$	power_$T1$	power_$T2$
100	20	0.0905	0.0739	0.4960	0.1922
	30	0.0813	0.0787	0.8446	0.3578
	40	0.0793	0.0737	0.9496	0.5027
	60	0.0721	0.0751	0.9922	0.6144
200	20	0.0270	0.0457	0.9185	0.3138
	30	0.0261	0.0388	0.9974	0.6673
	40	0.0230	0.0401	1.0000	0.7601
	60	0.0194	0.0282	1.0000	1.0000
300	20	0.0649	0.0714	0.9319	0.4199
	30	0.0452	0.0590	0.9958	0.5699
	40	0.0457	0.0632	1.0000	0.8875
	60	0.0438	0.0625	1.0000	0.9976
500	20	0.0872	0.0887	0.8229	0.3132
	30	0.0768	0.0839	0.9956	0.5470
	40	0.0718	0.0791	0.9996	0.7511
	60	0.0649	0.0757	1.0000	0.9426

TABLE 5. ASL and power: χ^2_{32} distribution, 1,000 simulations

p	$N_1 = N_2$	ASL_$T1$	ASL_$T2$	power_$T1$	power_$T2$
100	20	0.104	0.08	0.599	0.24
	30	0.078	0.966	0.373	0.078
	40	0.081	0.078	0.972	0.35
	60	0.069	0.079	1	0.994
200	20	0.024	0.038	0.968	0.477
	30	0.024	0.038	1	0.618
	40	0.017	0.032	1	0.916
	60	0.065	0.072	1	0.768
300	20	0.06	0.072	0.908	0.406
	30	0.055	0.07	1	0.652
	40	0.056	0.061	1	0.858
	60	0.049	0.071	1	0.991
500	20	0.1	0.093	0.851	0.29
	30	0.088	0.096	0.987	0.427
	40	0.021	0.033	1	1
	60	0.063	0.065	1	0.998

6. Multivariate analysis of variance (MANOVA)

The problem of testing the equality of $(q+1)$ mean vectors, $q \geq 2$, is a special case of the multivariate regression model

$$\begin{aligned} Y &= X\Theta + U\Lambda^{\frac{1}{2}},\ Y = (\mathbf{y}_1, \ldots, \mathbf{y}_N)' : N \times p, \\ U &= (\mathbf{u}_1, \ldots, \mathbf{u}_N)' : N \times p,\ (u_{i1}, \ldots, u_{ip})' = \mathbf{u}_i, \end{aligned}$$

where $\mathbf{y}_1, \ldots, \mathbf{y}_N$ are independently distributed p-dimensional observation vectors,

$$\begin{aligned} X &: N \times k \text{ of rank k, matrix of constants,} \\ \Theta &: k \times p,\ \text{matrix of parameters ,} \end{aligned}$$

Rather than assuming normality of $\mathbf{y}_i$, or equivalently of $\mathbf{u}_i$, we shall assume that

$$E(\mathbf{u}_i) = \mathbf{0},\ Cov(\mathbf{u}_i) = I_p,\ E(u_{ik}^4) = K_4 + 3, \tag{6.1}$$

$$Cov(\mathbf{y}_i) = \Lambda = \Lambda^{\frac{1}{2}}\Lambda^{\frac{1}{2}} = (\lambda_{ij}).$$

We also assume that for $\nu_k \geq 0,\ \Sigma_{k=1}^{p}\nu_k = 4$,

$$E\left[\prod_{k=1}^{p} u_{ik}^{\nu_k}\right] = \prod_{k=1}^{p} E(u_{ik}^{\nu_k}). \tag{6.2}$$

Under the assumptions (6.1) and (6.2), we consider the problem of testing the hypothesis

$$H : C\Theta = 0 \text{ versus } A : C\Theta \neq 0,$$

where C is a $q \times k$, matrix with $q \leq k$, and $\text{rank}(C) = q$. Let

$$B = Y'GY = (C\hat{\Theta})'[C(X'X)^{-1}C']^{-1}C\hat{\Theta},$$

where

$$\begin{aligned}
&G = X(X'X)^{-1}C'[C(X'X)^{-1}C']^{-1}C(X'X)^{-1}X',\\
&\hat{\Theta} = (X'X)^{-1}X'Y,\\
&S = n^{-1}Y'(I_N - H)Y,\ H = X(X'X)^{-1}X',\ n = N - k.
\end{aligned}$$

Then, under the hypothesis $C\Theta = 0$, $E(B) = q\Sigma$, and thus B measures the departure from the hypothesis. On the other hand, $E(S) = \Sigma$, irrespective of whether the hypothesis is true or not. Thus, a test is usually constructed by comparing in some manner B with S. Since $n < p$, the inverse of S does not exist, and thus the likelihood ratio test which uses the inverse of S does not exist. Fujikoshi, Himeno and Wakaki [4] proposed a test by generalizing Dempster's two-sample test, and Srivastava [7] proposed a test using Moore–Penrose inverse of S. Srivastava and Fujikoshi [10] and Schott [5] proposed a test by generalizing Bai and Saranadasa's two sample test. These tests will be described in the following subsection.

6.1. Generalization of Dempster's test. Assume

$$(p/n) \to c \ \in \ (0, \infty), \tag{6.3}$$

$$\begin{aligned}
\tilde{T}_D &= \sqrt{p}\left[\frac{\text{tr}\, B}{\text{tr}\, S} - q\right],\ \hat{a}_2 = p^{-1}[\text{tr}\, S^2 - \frac{1}{n}(\text{tr}\, S)^2],\\
\hat{\sigma}_D^2 &= 2q\frac{\hat{a}_2}{\hat{a}_1^2},\ \hat{a}_1 = \frac{(\text{tr}\, S)}{p}.
\end{aligned}$$

Then, Fujikoshi, Himeno and Wakaki [4] showed that under the hypothesis H, (6.1), and normality,

$$(\tilde{T}_D/\hat{\sigma}_D) \to N(0, 1).$$

Distribution under the alternative hypothesis is also given.

6.2. Generalization of Bai–Saranadasa test. Let

$$T_2 = \frac{p^{-\frac{1}{2}}[\text{tr}\, B - q\text{tr}\, S]}{\sqrt{2q\hat{a}_2}}. \tag{6.4}$$

This test was proposed by Srivastava and Fujikoshi [10], and Schott [5]. Under the hypothesis that $C\Theta = 0$, the asymptotic distribution of T_2 is normal with mean 0 and variance 1. To obtain the asymptotic distribution when $C\Theta \neq 0$, we consider local alternatives. For this, let

$$(C\Theta)'[C(X'X)^{-1}C']^{-1}(C\Theta) = MM', \tag{6.5}$$

where M is a $p \times q$ matrix of rank $q < n$, when $C\Theta \neq 0$. We shall assume that q is finite and

$$\lim_{p\to\infty} \frac{\operatorname{tr} \Lambda MM'}{p} = 0. \tag{6.6}$$

Then, Srivastava and Fujikoshi [10] have shown that the power of the T_2 test is given by

$$\beta(T_2) \simeq \Phi\left(-z_{1-\alpha} + \frac{\operatorname{tr} MM'}{\sqrt{2pqa_2}}\right), \tag{6.7}$$

for local alternatives satisfying (6.6), and finite q.

6.3. Srivastava's test. Let $V = nS$, V^+ be its Moore–Penrose inverse, and $d_1, ..., d_q$ be the eigenvalues of BV^+, $q < n$. Then Srivastava [7] proposed the satistic

$$U^+ = \prod_{i=1}^{q}(1 + d_i)^{-1} = |I + BV^+|^{-1}. \tag{6.8}$$

When $\Sigma = \gamma^2 I$ or Σ is of rank $r \leq n$, Srivastava [7] obtained distributions of chi-square types, similar to the likelihood ratio test. It may be noted that the hypothesis $\Sigma = \gamma^2 I$ can be tested by a test proposed by Srivastava [6] which has been shown to be robust under some departure from normality by Srivastava, Kollo, and von Rosen [12].

For general Σ, however, he showed that under the null hypothesis $C\Theta = 0$,

$$\lim_{(n,p)\to\infty} p\left[\frac{-p\hat{b}\,\log U^+ - qn}{\sqrt{2qn}} < z_{1-\alpha}\right] = \Phi(z_{1-\alpha}). \tag{6.9}$$

Srivastava and Fujikoshi [10] considered more generalized form of $\tilde{T}_D$ and T_2 and obtained the distribution without the condition (6.3) which was required by Schott [5].

Next we give the asymptotic distribution of the statistic U^+ under local alternatives given by

$$M = n^{-\frac{1}{2}}\Delta, \tag{6.10}$$

where Δ is $O(1)$. The asymptotic power of the test based on U^+ is given by

$$\begin{aligned}\beta(U^+) &= \lim_{n\to\infty}\lim_{p\to\infty} P_1\left[\frac{p\hat{b}\,\log U^+ - qn}{\sqrt{2qn}} < z_{1-\alpha}\right] \\ &= \lim_{n\to\infty}\lim_{p\to\infty} \Phi\left[-z_{1-\alpha} + \frac{n\ \operatorname{tr}\ \Lambda MM'}{pa_2\sqrt{2qn}}\right].\end{aligned}$$

The U^+ test should be preferred over T_2 if

$$\frac{\operatorname{tr}\Lambda MM'}{\operatorname{tr} MM'} > \frac{pa_2\sqrt{2qn}}{\sqrt[n]{2pqa_2}} = \left(\frac{pa_2}{n}\right)^{\frac{1}{2}}.$$

For example, if $M = (\mathbf{m}_1, \ldots, \mathbf{m}_q)$, and $\mathbf{m}_i = (\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_p})'$ $i = 1, \ldots, q$, where $\lambda_1, \ldots, \lambda_p$ are the eigenvalues of Σ given in the same order as the matrix Λ consisting of the eigenvalues of Σ, then the test U^+ should be preferred if

$$n < p\left(\frac{a_1^2}{a_2}\right) = pb.$$

Since $0 < b \leq 1$, many such n exist for large p. It may be noted that the U^+ test assumes normality and has yet to be generalized for non-normal model such as Model $\mathcal{M}$. Thus, it is not included in the tables of comparison by simulation.

These tests are invariant under the group of orthogonal transformations but not invariant under $p \times p$ non-singular diagonal matrices. A test that is invariant under this transformation has been proposed by Yamada and Srivastava [14] assuming normality of the observation vectors. We describe this test in the next Subsection 6.4.

6.4. Invariant test statistics. As mentioned above, $\tilde{T}_D$ and T_2 tests are not invariant under non-singular diagonal matrices. Thus, Yamada and Srivastava [14] proposed the statistic

$$T_1 = \frac{\operatorname{tr} BD_S^{-1} - \left(\frac{n}{n-2}\right)pq}{[2c_{p,n}q(\operatorname{tr} R^2 - n^{-1}p^2)]^{\frac{1}{2}}}, \quad R = D_S^{-\frac{1}{2}}SD_S^{-\frac{1}{2}}$$

under normality, where $c_{p,n} = 1 + (\operatorname{tr} R^2/p^{\frac{3}{2}})$.

The asymptotic distribution under the null hypothesis is standard normal and under local alternatives similar to (6.6), the asymptotic power of the T_1-test is given by

$$\beta(T_1) \simeq \Phi\left[-z_{1-\alpha} + \frac{\operatorname{tr} MM'D_\Sigma^{-1}}{\sqrt{2q\ \operatorname{tr} R^2}}\right],$$

where $R = D_\Sigma^{-\frac{1}{2}}\Sigma D_\Sigma^{-\frac{1}{2}}$ and $D_\Sigma =$ diag $(\sigma_{11}, \ldots, \sigma_{pp})$, $\sigma_{11}, \ldots, \sigma_{pp}$ being the diagonal elements of Σ.

For details see Yamada and Srivastava [14], where a theoretical comparison between T_1 and T_2 similar to Srivastava and Du [9] is also given.

In the next section, we compare the power of T_1-test with that of T_2-test by simulation.

6.5. Power comparison by simulation. 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 testing the inequality of the three mean vectors, we write

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

The observation matrix is

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

where $\mathbf{1}_{N^*} = (1, \dots, 1)' : N^* \times 1$ 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 $\boldsymbol{\mu}_1 = \mathbf{0}$, $\boldsymbol{\mu}_2 = 3n^{-1/2}p^{-1/4}\mathbf{1}'_p$, $\boldsymbol{\mu}_3 = -\boldsymbol{\mu}_2$. To generate the $\mathbf{Y}$ 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^2_8 and χ^2_{32} with 2, 8 and 32 degrees of freedom, respectively, and centre them and scale them as

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

for $u_{ij} \sim \chi^2_m$, $m = 2, 8, 32$. Since the skewness and kurtosis $(K_4 + 3)$ of χ^2_m are, repectively, $(8/m)^{1/2}$ and $3 + 12/m$, it is noted that χ^2_2 has higher skewness and kurtosis than χ^2_8 and χ^2_{32}. Write them as

$$\mathbf{V} = (\nu_1^{(1)}, \dots, \nu_{N^*}^{(1)}; \nu_1^{(2)}, \dots, \nu_{N^*}^{(2)}; \nu_1^{(3)}, \dots, \nu_{N^*}^{(3)})',$$

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

Case 1 : $\boldsymbol{\Sigma} = \mathbf{I}_p$,

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

For the Case 1 we define

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

where under the hypothesis, $\mathbf{Y} = \mathbf{V}$, and under the alternative, $\boldsymbol{\mu}_2$ and $\boldsymbol{\mu}_3$ are replaced by the vectors mentioned above. For the Case 2 let

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

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

The simulation results under the χ^2_m distributions for $m = 2$, 8 and 32 are presented in Tables 6, 7 and 8, respectively. The critical values are computed based on 100,000 replications and the powers are obtained based on 10,000 replications. Three tables report the critical values and the power in the hypothesis of the two tests, and it is seen that the critical values are appropriate.

TABLE 6. Critical values 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		Power 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	4.94	4.97	92.04	85.02
20	80	1.5632	1.6225	4.97	4.78	90.46	85.65
30	120	1.5622	1.6440	4.69	4.72	89.94	86.21
40	200	1.5564	1.6386	5.46	5.44	90.43	87.40
				Case 2			
10	40	1.6061	1.6865	4.94	5.49	99.96	24.71
20	80	1.5632	1.6784	4.97	5.04	99.63	18.20
30	120	1.5622	1.6919	4.69	4.64	97.82	15.20
40	200	1.5564	1.6852	5.46	4.99	96.26	16.97

TABLE 7. Critical values 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		Power 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	5.13	5.00	84.92	84.48
20	80	1.6175	1.6810	5.11	5.16	86.92	86.19
30	120	1.6119	1.6812	4.78	4.76	87.04	86.49
40	200	1.5967	1.6714	4.73	4.73	87.80	87.26
				Case 2			
10	40	1.7339	1.7903	5.13	5.05	99.93	23.56
20	80	1.6175	1.7276	5.11	5.45	99.27	18.89
30	120	1.6119	1.7344	4.78	5.00	97.06	15.56
40	200	1.5967	1.7291	4.73	5.02	94.70	16.38

TABLE 8. Critical values 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		Power 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	4.78	4.77	82.13	84.72
20	80	1.6457	1.6930	4.72	4.87	84.12	84.89
30	120	1.6155	1.6812	4.54	4.57	86.09	86.20
40	200	1.6090	1.6831	5.06	4.86	86.84	87.08
				Case 2			
10	40	1.7688	1.8157	4.78	5.21	99.97	23.15
20	80	1.6457	1.7409	4.72	4.61	98.93	16.80
30	120	1.6155	1.7476	4.54	4.73	96.46	15.85
40	200	1.6090	1.7223	5.06	4.93	94.55	16.61

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 $\mathbf{\Sigma} = \mathbf{I}_p$, 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.

7. Concluding remarks

In this article we reviewed several tests for the equality of the two mean vectors including the case when the covariance matrices of the two groups may be unequal. The asymptotic distributions are given under non-normal models. Thus, the tests are robust against the departure from normality. In MANOVA, we assume that the covariance matrices are all equal. We have shown through simulation that the tests that are invariant under non-singular diagonal matrices perform better than those that are not invariant.

Acknowledgements

I would like to thank a referee and the editor handling this paper for their suggestions that led to a great improvement in the presentation. The research of this paper was supported by NSERC of Canada.

References

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

[2] S. X. Chen and Y. L. Qin, *A two-sample test for high-dimensional data with applications to gene-set testing*, Ann. Statist. **38** (2010), 808–835.

[3] A. P. Dempster, *A high dimensional two-sample significance test*, Ann. Math. Statist. **29** (1958), 995–1010.

[4] Y. Fujikoshi, T. Himeno, and H. Wakaki, *Asymptotic results of a high dimensional Manova test and power comparison when the dimension is large compared to the sample size*, J. Japan Statist. Soc. **34** (2004), 19–26.

[5] J. R. Schott, *Some high-dimensional tests for a one-way MANOVA*, J. Multivariate Anal. **98** (2007), 1825–1839.

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

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

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

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

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

[11] M. S. Srivastava, S. Katayama, and Y. Kano, *A two sample test in high dimension with fewer observations than the dimension*, J. Multivariate Anal. **114** (2013), 349–358.

[12] M. S. Srivastava, T. Kollo, and D. von Rosen, *Some tests for the covariance matrix with fewer observations than the dimension under non-normality*, J. Multivariate Anal. **102** (2011), 1090–1103.

[13] M. S. Srivastava and H. Yanagihara, *Tetsing the equality of several covariance matrices with fewer observations than the dimension*, J. Multivariate Anal. **101** (2010), 1319–1329.

[14] T. Yamada and M. S. Srivastava, *A test for the multivariate analysis of variance in high-dimension*, Commun. Statist. Theory Methods **41** (2012), 13–14, 2602–2615.

Department of Statistical Sciences, University of Toronto, 100 St George Street, Toronto, Canada

E-mail address: `srivasta@utstat.toronto.edu`