The Annals of Statistics
2003, Vol. 31, No. 5, 1517–1536

TESTING HOMOGENEITY OF MULTIVARIATE NORMAL MEAN VECTORS UNDER AN ORDER RESTRICTION WHEN THE COVARIANCE MATRICES ARE COMMON BUT UNKNOWN

BY SHOICHI SASABUCHI, KOJI TANAKA AND TAKESHI TSUKAMOTO

Kyushu Institute of Design, Tosu High School and Hitachi Software Corporation

Suppose that an order restriction is imposed among several p-variate normal mean vectors. We are interested in testing the homogeneity of these mean vectors under this restriction. This problem is a multivariate extension of Bartholomew's [*Biometrika* **46** (1959) 36–48]. When the covariance matrices are known, this problem has been studied by Sasabuchi, Inutsuka and Kulatunga [*Hiroshima Math. J.* **22** (1992) 551–560], Sasabuchi, Kulatunga and Saito [*Amer. J. Math. Management Sci.* **18** (1998) 131–158] and some others. In the present paper, we consider the case when the covariance matrices are common but unknown. We propose a test statistic, study its upper tail probability under the null hypothesis and estimate its critical points.

1. Introduction. Bartholomew (1959) considered the problem of testing the homogeneity of several univariate normal means against an order restricted alternative hypothesis. He derived the likelihood ratio test statistic, $\bar{\chi}_k^2$, and its null distribution under the assumption that the variances are known. Since then, an extensive literature concerning this problem has appeared and most of it has been summarized by Barlow, Bartholomew, Bremner and Brunk (1972) and Robertson, Wright and Dykstra (1988). They have shown by numerical computation that, for the order restricted alternative hypothesis, the $\bar{\chi}_k^2$ test is more powerful than the usual χ^2 test, which is the likelihood ratio test for testing homogeneity against the unrestricted alternative hypothesis.

Sasabuchi, Inutsuka and Kulatunga (1983) generalized Bartholomew's (1959) problem to that of several multivariate normal mean vectors. Their theory enables us to study, for example, statistical inference in the case where the effects of several factors increase (or decrease) simultaneously. [See Sasabuchi, Inutsuka and Kulatunga (1983).]

Consider p-variate normal distributions $N_p(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$, $i = 1, 2, \dots, k$. We are interested in the problem of testing $\mathrm{H}_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = \cdots = \boldsymbol{\mu}_k$ versus $\mathrm{H}_1 : \boldsymbol{\mu}_1 \leq \boldsymbol{\mu}_2 \leq \cdots \leq \boldsymbol{\mu}_k$, where "$\boldsymbol{\mu}_i \leq \boldsymbol{\mu}_j$" means that all the elements of $\boldsymbol{\mu}_j - \boldsymbol{\mu}_i$ are nonnegative. This problem is a multivariate extension of Bartholomew's (1959) one

Received April 2002; revised October 2002.
AMS 2000 subject classifications. Primary 62F30; secondary 62F03, 62H15.
Key words and phrases. Common but unknown covariance matrices, multivariate isotonic regression, multivariate normal distribution, order restriction, testing homogeneity of mean vectors, upper tail probability.

and may arise in the situation where the values of several parameters increase simultaneously. For example, if we want to know whether both average height and average weight of the children of an area increase simultaneously year by year, we could apply our test to the bivariate data sets obtained by random sampling from the population.

When the covariance matrices are known, this problem has been studied to some extent. Sasabuchi, Inutsuka and Kulatunga (1983) derived the likelihood ratio test statistic, $\bar{\chi}^2_{k,p}$, Kulatunga and Sasabuchi (1984) studied its null distribution, and Nomakuchi and Shi (1988) proposed a new test whose null distribution was easier to calculate than that of the $\bar{\chi}^2_{k,p}$ test. Sasabuchi, Kulatunga and Saito (1998) made some power comparisons by simulation in the bivariate case and showed that over H_1 the $\bar{\chi}^2_{k,p}$ test is more powerful than the usual χ^2 test, which is the likelihood ratio test for H_0 against the unrestricted alternative hypothesis.

In the present paper, we assume that the covariance matrices are common but unknown. When $p = 1$, this case has been studied by Bartholomew (1961) and many others, and the likelihood ratio test statistic is well known as $\bar{E}^2_k$. We shall study its multivariate extension. For the multivariate case, the likelihood ratio test statistic for H_0 against the unrestricted alternative hypothesis was given by Anderson [(1984), Section 8.8]. But, in our problem, the alternative has an order restriction; thus we should take this restriction into consideration.

To the authors' knowledge, the likelihood ratio test for H_0 versus H_1 in the case when the covariance matrices are common but unknown has not been obtained yet. Perlman (1969) studied a multivariate one-sided testing problem with an unknown covariance matrix and derived its likelihood ratio test. It may seem that our problem may be reduced to his, but the structure of our model is different from his, so we cannot apply his methods or results directly to our problem.

The purposes of the present paper are to propose a test statistic and study its upper tail probability under H_0. In Section 2 we describe the problem and propose a test statistic. Main theorems about its null distribution and upper tail probability under H_0 are presented in Section 3. In Section 4 preliminary definitions and results on a convex cone and projection are given. Proofs of the main theorems are given in Section 5. In Section 6 a table of the critical points estimated by simulation is presented. In Section 7 a justification of the plausibility of our test statistic is discussed. Proofs of some of the lemmas are given in the Appendix.

2. The problem and proposed test statistic. Suppose that $\mathbf{X}_{i1}, \dots, \mathbf{X}_{iN_i}$ are random samples from a p-variate normal distribution $N_p(\boldsymbol{\mu}_i, \boldsymbol{\Sigma})$, $i = 1, 2, \dots, k$. We assume that $\boldsymbol{\Sigma}$ is unknown and $N_1 + \cdots + N_k > p + k$.

Throughout this paper, any vectors are column vectors as a rule. As usual, for any vector $\mathbf{x}$ and matrix A, $\mathbf{x}'$ and A' denote their transposed vector and transposed matrix, respectively.

Consider the problem of testing

$$\mathrm{H}_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = \cdots = \boldsymbol{\mu}_k \quad \text{versus} \quad \mathrm{H}_1 : \boldsymbol{\mu}_1 \le \boldsymbol{\mu}_2 \le \cdots \le \boldsymbol{\mu}_k,$$

where "$\boldsymbol{\mu}_i \le \boldsymbol{\mu}_j$" means that all the elements of $\boldsymbol{\mu}_j - \boldsymbol{\mu}_i$ are non-negative. The restriction imposed on H_1 means that all the components of the p-dimensional mean vector $\boldsymbol{\mu}_i$ increase simultaneously as i increases.

In order to discuss the order restricted problem, we prepare some definitions.

DEFINITION 2.1 [Barlow, Bartholomew, Bremner and Brunk (1972)]. Given real numbers $x_1, \dots, x_k$ and positive numbers $w_1, \dots, w_k$, a k-dimensional real row vector $(\hat{\theta}_1, \dots, \hat{\theta}_k)$ is said to be the isotonic regression (IR) of $x_1, \dots, x_k$ with weights $w_1, \dots, w_k$ if $\hat{\theta}_1 \le \hat{\theta}_2 \le \cdots \le \hat{\theta}_k$ and $(\hat{\theta}_1, \dots, \hat{\theta}_k)$ satisfies

$$\min_{\theta_1 \le \theta_2 \le \cdots \le \theta_k} \sum_{i=1}^{k} (x_i - \theta_i)^2 w_i = \sum_{i=1}^{k} (x_i - \hat{\theta}_i)^2 w_i.$$

The isotonic regression can be computed easily by the well-known method, Pool–Adjacent–Violators algorithm. [See Barlow, Bartholomew, Bremner and Brunk (1972).]

DEFINITION 2.2 [Sasabuchi, Inutsuka and Kulatunga (1983)]. Given p-dimensional real vectors $\mathbf{x}_1, \dots, \mathbf{x}_k$ and $p \times p$ positive definite matrices $\boldsymbol{\Lambda}_1, \dots, \boldsymbol{\Lambda}_k$, a $p \times k$ real matrix $(\hat{\boldsymbol{\theta}}_1, \dots, \hat{\boldsymbol{\theta}}_k)$ is said to be the multivariate isotonic regression (MIR), in fact, p-variate isotonic regression, of $\mathbf{x}_1, \dots, \mathbf{x}_k$ with weights $\boldsymbol{\Lambda}_1^{-1}, \dots, \boldsymbol{\Lambda}_k^{-1}$ if $\hat{\boldsymbol{\theta}}_1 \le \hat{\boldsymbol{\theta}}_2 \le \cdots \le \hat{\boldsymbol{\theta}}_k$ and $(\hat{\boldsymbol{\theta}}_1, \dots, \hat{\boldsymbol{\theta}}_k)$ satisfies

$$\min_{\boldsymbol{\theta}_1 \le \boldsymbol{\theta}_2 \le \cdots \le \boldsymbol{\theta}_k} \sum_{i=1}^{k} (\mathbf{x}_i - \boldsymbol{\theta}_i)' \boldsymbol{\Lambda}_i^{-1} (\mathbf{x}_i - \boldsymbol{\theta}_i) = \sum_{i=1}^{k} (\mathbf{x}_i - \hat{\boldsymbol{\theta}}_i)' \boldsymbol{\Lambda}_i^{-1} (\mathbf{x}_i - \hat{\boldsymbol{\theta}}_i),$$

where "$\boldsymbol{\theta}_i \le \boldsymbol{\theta}_j$" means that all the elements of $\boldsymbol{\theta}_j - \boldsymbol{\theta}_i$ are non-negative.

Sasabuchi, Inutsuka and Kulatunga (1983) proposed an iterative algorithm for the computation of MIR and studied its convergence in the case when $p = 2$. This algorithm was extended to the general multivariate case by Sasabuchi, Inutsuka and Kulatunga (1992).

When $\boldsymbol{\Sigma}$ is known, the likelihood ratio test for H_0 versus H_1 is given by the following theorem.

THEOREM 2.1 [Sasabuchi, Inutsuka and Kulatunga (1983)]. *When $\boldsymbol{\Sigma}$ is known, the critical region of the likelihood ratio test for* H_0 *versus* H_1 *is given by*

$$\bar{\chi}^2_{k,p} = \sum_{i=1}^{k} N_i (\widehat{\boldsymbol{\mu}}_i - \overline{\mathbf{X}})' \boldsymbol{\Sigma}^{-1} (\widehat{\boldsymbol{\mu}}_i - \overline{\mathbf{X}}) \ge c_1,$$

where $(\widehat{\boldsymbol{\mu}}_1, \dots, \widehat{\boldsymbol{\mu}}_k)$ is the MIR of $\overline{\mathbf{X}}_1, \dots, \overline{\mathbf{X}}_k$ with weights $N_1 \boldsymbol{\Sigma}^{-1}, \dots, N_k \boldsymbol{\Sigma}^{-1}$, c_1 is a positive constant depending on the significance level, $\overline{\mathbf{X}}_i = N_i^{-1} \sum_{j=1}^{N_i} \mathbf{X}_{ij}$, $i = 1, \dots, k$, and $\overline{\mathbf{X}} = (\sum_{i=1}^{k} N_i)^{-1} \sum_{i=1}^{k} N_i \overline{\mathbf{X}}_i$.

The $\bar{\chi}^2_{k,p}$ statistic is the multivariate extension of Bartholomew's (1959) $\bar{\chi}^2_k$ statistic. This test has been studied by Kulatunga and Sasabuchi (1984), Kulatunga (1984), Sasabuchi, Kulatunga and Saito (1998) and some others.

In this paper, we assume that $\boldsymbol{\Sigma}$ is unknown. In this case, to the authors' knowledge, the likelihood ratio test for H_0 versus H_1 has not been obtained yet.

Replacing the unknown covariance matrix in the $\bar{\chi}^2_{k,p}$ by its estimator, we propose the following test:

$$\overline{T}^2 = \sum_{i=1}^{k} N_i(\widehat{\boldsymbol{\mu}}_i - \overline{\mathbf{X}})' S^{-1}(\widehat{\boldsymbol{\mu}}_i - \overline{\mathbf{X}}) \geq c_2 \quad \Longrightarrow \quad \text{reject } \mathrm{H}_0,$$

where $S = \sum_{i=1}^{k} \sum_{j=1}^{N_i} (\mathbf{X}_{ij} - \overline{\mathbf{X}}_i)(\mathbf{X}_{ij} - \overline{\mathbf{X}}_i)'$, $(\widehat{\boldsymbol{\mu}}_1, \ldots, \widehat{\boldsymbol{\mu}}_k)$ is the MIR of $\overline{\mathbf{X}}_1, \ldots, \overline{\mathbf{X}}_k$ with weights $N_1 S^{-1}, \ldots, N_k S^{-1}$, and c_2 is a positive constant depending on the significance level. Note that S has the Wishart distribution $W_p(N-k; \boldsymbol{\Sigma})$, where $N = N_1 + \cdots + N_k$, and is statistically independent of $\overline{\mathbf{X}}_1, \ldots, \overline{\mathbf{X}}_k$.

3. Main theorems. When we use $\overline{T}^2$ to test H_0 versus H_1, we need to compute the supremum of its upper tail probability under H_0; that is,

$$\sup_{\boldsymbol{\Sigma}} \sup_{\mathrm{H}_0} \mathrm{P}_{\boldsymbol{\mu},\boldsymbol{\Sigma}}\big(\overline{T}^2 \geq c\big) \tag{1}$$

for arbitrary constant c. Here $\mathrm{P}_{\boldsymbol{\mu},\boldsymbol{\Sigma}}$ denotes the probability measure corresponding to the parameters $\boldsymbol{\mu} = (\boldsymbol{\mu}'_1, \ldots, \boldsymbol{\mu}'_k)'$ and $\boldsymbol{\Sigma}$, $\sup_{\mathrm{H}_0}$ denotes the supremum for $\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_k$ with $\boldsymbol{\mu}_1 = \cdots = \boldsymbol{\mu}_k$, and $\sup_{\boldsymbol{\Sigma}}$ denotes the supremum for all the $p \times p$ positive definite real matrices.

Proofs of the theorems presented in this section are given in Section 5.

First we get the following theorem.

THEOREM 3.1. *Under* H_0, *the distribution of* $\overline{T}^2$ *is independent of* $\boldsymbol{\mu}_0$, *where* $\boldsymbol{\mu}_0$ *is the common value of* $\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_k$.

According to this theorem, we can assume that $\boldsymbol{\mu} = \mathbf{0}$ in computing the upper tail probability of $\overline{T}^2$ under H_0. But it still depends on the unknown $\boldsymbol{\Sigma}$.

Now we introduce the following statistic:

$$T^* = \sum_{i=1}^{k} N_i(\overline{\mathbf{X}}_i - \overline{\mathbf{X}})' S^{-1}(\overline{\mathbf{X}}_i - \overline{\mathbf{X}}) - \frac{1}{s_{11}} \sum_{i=1}^{k} N_i(\overline{X}_{i1} - \widehat{\mu}_{i1})^2,$$

where $\overline{X}_{i1}$ is the first component of $\overline{\mathbf{X}}_i$, $i = 1, 2, \ldots, k$, s_{11} is the $(1,1)$th element of S, and $(\widehat{\mu}_{11}, \ldots, \widehat{\mu}_{k1})$ is the IR of $\overline{X}_{11}, \ldots, \overline{X}_{k1}$ with weights $N_1, \ldots, N_k$.

We can show the following theorem.

THEOREM 3.2. *Under* H_0, *the distribution of* T^* *is independent of* $\boldsymbol{\mu}_0$ *and* $\boldsymbol{\Sigma}$, *where* $\boldsymbol{\mu}_0$ *is the common value of* $\boldsymbol{\mu}_1, \dots, \boldsymbol{\mu}_k$.

According to this theorem, we can assume that $\boldsymbol{\mu} = \mathbf{0}$ and $\boldsymbol{\Sigma} = \mathrm{I}_p$, the unit matrix, in computing the upper tail probability of T^* under H_0.

The supremum of the upper tail probability of $\overline{T}^2$ under H_0 is given by the following theorem.

THEOREM 3.3. *For any real number* c,

$$\sup_{\boldsymbol{\Sigma}} \sup_{\mathrm{H}_0} \mathrm{P}_{\boldsymbol{\mu},\boldsymbol{\Sigma}}(\overline{T}^2 \geq c) = \sup_{\boldsymbol{\Sigma}} \mathrm{P}_{\mathbf{0},\boldsymbol{\Sigma}}(\overline{T}^2 \geq c) = \mathrm{P}_{\mathbf{0},\mathrm{I}_p}(T^* \geq c).$$

According to this theorem, we only need to compute $\mathrm{P}_{\mathbf{0},\mathrm{I}_p}(T^* \geq c)$ in order to compute (1). We do not know the exact distribution of T^*, but we can easily get the approximate value of $\mathrm{P}_{\mathbf{0},\mathrm{I}_p}(T^* \geq c)$ by using Monte Carlo simulation generating standard normal random numbers, because T^* is easy to calculate and the distribution is that under $\boldsymbol{\mu} = \mathbf{0}$ and $\boldsymbol{\Sigma} = \mathrm{I}_p$.

4. Preliminary definitions and results on a convex cone and projection. In this section, we prepare some definitions and basic results about a convex cone and projection. They are modifications of those of Perlman (1969) in accordance with our situation. Proofs of the lemmas in this section are omitted.

Let $\mathbb{R}^p$ and $\mathbb{R}^{pk}$ denote the p-dimensional and pk-dimensional real Euclidean spaces, respectively. For a point $\mathbf{x}$ in a real Euclidean space, we write "$\mathbf{x} \geq \mathbf{0}$ ($\mathbf{x} > \mathbf{0}$)" to indicate that all the elements of $\mathbf{x}$ are non-negative (positive).

Note that all vectors are column vectors as mentioned in Section 2.

Let $\mathcal{C}$ be a nonempty subset in a real Euclidean space. We call $\mathcal{C}$ a *convex cone* if

$$\mathbf{x}, \mathbf{y} \in \mathcal{C},\ \beta \geq 0,\ \gamma \geq 0 \quad \Longrightarrow \quad \beta\mathbf{x} + \gamma\mathbf{y} \in \mathcal{C}.$$

Further, we call $\mathcal{C}$ a *closed convex cone* if it is a convex cone and closed set.

LEMMA 4.1. *Let*

$$\mathcal{A}_n = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \middle| \begin{array}{l} x \in \mathbb{R}^1, y \in \mathbb{R}^1 \\ x \geq 0, y \geq -nx \end{array} \right\}, \qquad n = 1, 2, \dots.$$

Then

(i) $\mathcal{A}_n$ *is a closed convex cone in* $\mathbb{R}^2$, $n = 1, 2, \dots$,

(ii) $\mathcal{A}_n \subset \mathcal{A}_{n+1}$, $n = 1, 2, \dots$,

(iii) $\bigcup_{n=1}^{\infty} \mathcal{A}_n = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \middle| \begin{array}{l} x \geq 0 \\ y \geq 0 \end{array} \right\} \cup \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \middle| \begin{array}{l} x > 0 \\ y \leq 0 \end{array} \right\}$.

For a $q \times r$ matrix $A = (a_{ij})$ and an $s \times t$ matrix $B = (b_{kl})$, as usual, $A \otimes B$ denotes their *Kronecker product*, that is,

$$A \otimes B = \begin{pmatrix} a_{11}B & \dots & a_{1r}B \\ \vdots & & \vdots \\ a_{q1}B & \dots & a_{qr}B \end{pmatrix}.$$

Let $\mathbf{\Lambda}$ be a $p \times p$ positive definite real matrix. Let $\mathbf{x}_1, \dots, \mathbf{x}_k, \mathbf{y}_1, \dots, \mathbf{y}_k$ be p-dimensional real vectors and put

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_k \end{pmatrix}, \qquad \mathbf{y} = \begin{pmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_k \end{pmatrix}.$$

Define an *inner product* $\langle \cdot, \cdot \rangle_{\mathbf{\Lambda}}$ in $\mathbb{R}^{pk}$ by

$$\begin{aligned} \langle \mathbf{x}, \mathbf{y} \rangle_{\mathbf{\Lambda}} &= \sum_{i=1}^{k} N_i \mathbf{x}_i' \mathbf{\Lambda}^{-1} \mathbf{y}_i \\ &= (\mathbf{x}_1', \dots, \mathbf{x}_k') \begin{pmatrix} N_1 \mathbf{\Lambda}^{-1} & & 0 \\ & \ddots & \\ 0 & & N_k \mathbf{\Lambda}^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_k \end{pmatrix} \\ &= \mathbf{x}'(D \otimes \mathbf{\Lambda}^{-1})\mathbf{y}, \end{aligned}$$

where

$$D = \begin{pmatrix} N_1 & & 0 \\ & \ddots & \\ 0 & & N_k \end{pmatrix}.$$

Further, define a *norm* $\| \cdot \|_{\mathbf{\Lambda}}$ in $\mathbb{R}^{pk}$ by

$$\|\mathbf{x}\|_{\mathbf{\Lambda}} = \langle \mathbf{x}, \mathbf{x} \rangle_{\mathbf{\Lambda}}^{1/2}.$$

Let $\mathcal{C}$ be a closed convex cone in $\mathbb{R}^{pk}$.

For $\mathbf{x} \in \mathbb{R}^{pk}$, the *orthogonal projection of* $\mathbf{x}$ *onto* $\mathcal{C}$ *with respect to* $\langle \cdot, \cdot \rangle_{\mathbf{\Lambda}}$, denoted by $\pi_{\mathbf{\Lambda}}(\mathbf{x}; \mathcal{C})$, is defined by the point which minimizes $\|\mathbf{x} - \mathbf{z}\|_{\mathbf{\Lambda}}$ under the restriction that $\mathbf{z} \in \mathcal{C}$.

Note that $\pi_{\mathbf{\Lambda}}(\mathbf{x}; \mathcal{C})$ is determined uniquely since $\mathcal{C}$ is a closed convex cone. If $\mathbf{x} \in \mathcal{C}$, then $\pi_{\mathbf{\Lambda}}(\mathbf{x}; \mathcal{C}) = \mathbf{x}$.

For any set $\mathcal{A} \subset \mathbb{R}^{pk}$ and any $p \times p$ real nonsingular matrix B, $(\mathrm{I}_k \otimes B)\mathcal{A}$ denotes the *image of* $\mathcal{A}$ *by* $\mathrm{I}_k \otimes B$; that is, $(\mathrm{I}_k \otimes B)\mathcal{A} = \{(\mathrm{I}_k \otimes B)\mathbf{x} \mid \mathbf{x} \in \mathcal{A}\}$.

LEMMA 4.2. *For any point* $\mathbf{x} \in \mathbb{R}^{pk}$ *and any* $p \times p$ *real nonsingular matrix* B,

$$\|\mathbf{x} - \pi_{\mathbf{\Lambda}}(\mathbf{x}; \mathcal{C})\|_{\mathbf{\Lambda}} = \big\|(\mathrm{I}_k \otimes B)\mathbf{x} - \pi_{B\mathbf{\Lambda}B'}\big((\mathrm{I}_k \otimes B)\mathbf{x}; (\mathrm{I}_k \otimes B)\mathcal{C}\big)\big\|_{B\mathbf{\Lambda}B'}.$$

LEMMA 4.3. *Let $\{\mathcal{C}_n\}_{n=1,2,\dots}$ be a sequence of closed convex cones in $\mathbb{R}^{pk}$. If $\mathcal{C}_1 \subset \mathcal{C}_2 \subset \cdots \subset \mathcal{C}_n \subset \cdots$, then*

$$\lim_{n\to\infty} \|\mathbf{x} - \pi_{\mathbf{\Lambda}}(\mathbf{x}; \mathcal{C}_n)\|_{\mathbf{\Lambda}} = \|\mathbf{x} - \pi_{\mathbf{\Lambda}}(\mathbf{x}; \overline{\mathcal{C}_\infty})\|_{\mathbf{\Lambda}},$$

where $\overline{\mathcal{C}_\infty}$ denotes the closure of $\mathcal{C}_\infty = \bigcup_{n=1}^{\infty} \mathcal{C}_n$.

5. Proofs of the main theorems. Now we can represent our test statistic in terms of the projection. Recall that

$$\overline{T}^2 = \sum_{i=1}^{k} N_i (\widehat{\boldsymbol{\mu}}_i - \overline{\mathbf{X}})' S^{-1} (\widehat{\boldsymbol{\mu}}_i - \overline{\mathbf{X}}).$$

By using Theorem 2.2 of Sasabuchi, Inutsuka and Kulatunga (1983), $\overline{T}^2$ is rewritten as

$$\overline{T}^2 = \sum_{i=1}^{k} N_i (\overline{\mathbf{X}}_i - \overline{\mathbf{X}})' S^{-1} (\overline{\mathbf{X}}_i - \overline{\mathbf{X}}) - \sum_{i=1}^{k} N_i (\overline{\mathbf{X}}_i - \widehat{\boldsymbol{\mu}}_i)' S^{-1} (\overline{\mathbf{X}}_i - \widehat{\boldsymbol{\mu}}_i).$$

Put

$$\mathbf{X} = \begin{pmatrix} \overline{\mathbf{X}}_1 \\ \vdots \\ \overline{\mathbf{X}}_k \end{pmatrix}, \qquad \widetilde{\mathbf{X}} = \begin{pmatrix} \overline{\mathbf{X}} \\ \vdots \\ \overline{\mathbf{X}} \end{pmatrix}, \qquad \widehat{\boldsymbol{\mu}} = \begin{pmatrix} \widehat{\boldsymbol{\mu}}_1 \\ \vdots \\ \widehat{\boldsymbol{\mu}}_k \end{pmatrix}.$$

Then, by using the norm defined in Section 4, we can write

$$\overline{T}^2 = \|\widehat{\boldsymbol{\mu}} - \widetilde{\mathbf{X}}\|_S^2 = \|\mathbf{X} - \widetilde{\mathbf{X}}\|_S^2 - \|\mathbf{X} - \widehat{\boldsymbol{\mu}}\|_S^2.$$

Now it is important to note that $(\overline{\mathbf{X}}, \dots, \overline{\mathbf{X}})$ and $(\widehat{\boldsymbol{\mu}}_1, \dots, \widehat{\boldsymbol{\mu}}_k)$ are the values of $(\boldsymbol{\mu}_1, \dots, \boldsymbol{\mu}_k)$ which minimize $\sum_{i=1}^{k} N_i (\overline{\mathbf{X}}_i - \boldsymbol{\mu}_i)' S^{-1} (\overline{\mathbf{X}}_i - \boldsymbol{\mu}_i)$ under the restrictions that $\boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = \cdots = \boldsymbol{\mu}_k$ and $\boldsymbol{\mu}_1 \le \boldsymbol{\mu}_2 \le \cdots \le \boldsymbol{\mu}_k$, respectively.

Define two closed convex cones $\mathcal{C}_0$ and $\mathcal{C}_1$ in $\mathbb{R}^{pk}$ by

$$\mathcal{C}_0 = \left\{ \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \\ \vdots \\ \boldsymbol{\mu}_k \end{pmatrix} \middle| \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = \cdots = \boldsymbol{\mu}_k, \boldsymbol{\mu}_i \in \mathbb{R}^p, i = 1, \dots, k \right\},$$

$$\mathcal{C}_1 = \left\{ \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \\ \vdots \\ \boldsymbol{\mu}_k \end{pmatrix} \middle| \boldsymbol{\mu}_1 \le \boldsymbol{\mu}_2 \le \cdots \le \boldsymbol{\mu}_k, \boldsymbol{\mu}_i \in \mathbb{R}^p, i = 1, \dots, k \right\}.$$

Then, from the definition of projection in Section 4, $\widetilde{\mathbf{X}} = \pi_S(\mathbf{X}; \mathcal{C}_0)$ and $\widehat{\boldsymbol{\mu}} = \pi_S(\mathbf{X}; \mathcal{C}_1)$; hence we can write

$$\overline{T}^2 = \|\pi_S(\mathbf{X}; \mathcal{C}_1) - \pi_S(\mathbf{X}; \mathcal{C}_0)\|_S^2 = \|\mathbf{X} - \pi_S(\mathbf{X}; \mathcal{C}_0)\|_S^2 - \|\mathbf{X} - \pi_S(\mathbf{X}; \mathcal{C}_1)\|_S^2.$$

PROOF OF THEOREM 3.1. The proof is based on the location invariance of $\overline{T}^2$. Recall that $\boldsymbol{\mu}_0$ is the common unknown value of $\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_k$ under H_0. Put

$$\mathbf{Y} = \begin{pmatrix} \mathbf{Y}_1 \\ \vdots \\ \mathbf{Y}_k \end{pmatrix} = \begin{pmatrix} \overline{\mathbf{X}}_1 - \boldsymbol{\mu}_0 \\ \vdots \\ \overline{\mathbf{X}}_k - \boldsymbol{\mu}_0 \end{pmatrix} = \mathbf{X} - \begin{pmatrix} \boldsymbol{\mu}_0 \\ \vdots \\ \boldsymbol{\mu}_0 \end{pmatrix}.$$

Then we can easily show that

$$\|\mathbf{X} - \pi_S(\mathbf{X}; \mathcal{C}_0)\|_S^2 = \|\mathbf{Y} - \pi_S(\mathbf{Y}; \mathcal{C}_0)\|_S^2,$$
$$\|\mathbf{X} - \pi_S(\mathbf{X}; \mathcal{C}_1)\|_S^2 = \|\mathbf{Y} - \pi_S(\mathbf{Y}; \mathcal{C}_1)\|_S^2$$

and hence

$$\begin{aligned}\overline{T}^2 &= \|\mathbf{X} - \pi_S(\mathbf{X}; \mathcal{C}_0)\|_S^2 - \|\mathbf{X} - \pi_S(\mathbf{X}; \mathcal{C}_1)\|_S^2 \\ &= \|\mathbf{Y} - \pi_S(\mathbf{Y}; \mathcal{C}_0)\|_S^2 - \|\mathbf{Y} - \pi_S(\mathbf{Y}; \mathcal{C}_1)\|_S^2.\end{aligned}$$

Under H_0, the distribution of $\|\mathbf{Y} - \pi_S(\mathbf{Y}; \mathcal{C}_0)\|_S^2 - \|\mathbf{Y} - \pi_S(\mathbf{Y}; \mathcal{C}_1)\|_S^2$ is independent of $\boldsymbol{\mu}_0$ because $\mathbf{Y}_i$ is distributed as $N_p(\mathbf{0}, N_i^{-1}\boldsymbol{\Sigma})$, $i = 1, \ldots, k$. This completes the proof. □

Now we define another closed convex cone $\mathcal{C}_2$ in $\mathbb{R}^{pk}$ by

$$\mathcal{C}_2 = \left\{ \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \\ \vdots \\ \boldsymbol{\mu}_k \end{pmatrix} \middle| \mu_{11} \le \mu_{21} \le \cdots \le \mu_{k1},\ \boldsymbol{\mu}_i \in \mathbb{R}^p,\ i = 1, \ldots, k \right\},$$

where μ_{i1} denotes the first component of $\boldsymbol{\mu}_i$, $i = 1, \ldots, k$. Further we define another statistic T^{**} by

$$T^{**} = \|\mathbf{X} - \pi_S(\mathbf{X}; \mathcal{C}_0)\|_S^2 - \|\mathbf{X} - \pi_S(\mathbf{X}; \mathcal{C}_2)\|_S^2.$$

Note that $\mathcal{C}_0 \subset \mathcal{C}_1 \subset \mathcal{C}_2$, so we have

$$\|\mathbf{X} - \pi_S(\mathbf{X}; \mathcal{C}_0)\|_S^2 \ge \|\mathbf{X} - \pi_S(\mathbf{X}; \mathcal{C}_2)\|_S^2$$

and thus $T^{**} \ge 0$.

LEMMA 5.1. *Let $\boldsymbol{\Sigma}$ be a $p \times p$ positive definite real matrix. Then we have*:

(i) *for any $p \times p$ orthogonal matrix U*,

$$\big(\mathrm{I}_k \otimes (U\boldsymbol{\Sigma}^{-1/2})\big)\mathcal{C}_0 = \mathcal{C}_0;$$

(ii) *there exists a $p \times p$ orthogonal matrix U which satisfies*

$$\big(\mathrm{I}_k \otimes (U\boldsymbol{\Sigma}^{-1/2})\big)\mathcal{C}_2 = \mathcal{C}_2.$$

Note that $\mathbf{\Sigma}^{1/2}$ is the positive definite real (symmetric) matrix such that $\mathbf{\Sigma}^{1/2}\mathbf{\Sigma}^{1/2} = \mathbf{\Sigma}$ and $\mathbf{\Sigma}^{-1/2} = (\mathbf{\Sigma}^{1/2})^{-1}$.

The proof will be given in the Appendix.

LEMMA 5.2. *Under* H_0, *the distribution of* T^{**} *is independent of* $\boldsymbol{\mu}_0$ *and* $\mathbf{\Sigma}$, *where* $\boldsymbol{\mu}_0$ *is the common value of* $\boldsymbol{\mu}_1, \dots, \boldsymbol{\mu}_k$.

PROOF. Let $\mathbf{Y}$ be the same as that defined in the proof of Theorem 3.1. We can easily show that

$$\begin{aligned}\|\mathbf{X} - \pi_S(\mathbf{X}; \mathcal{C}_0)\|_S^2 &= \|\mathbf{Y} - \pi_S(\mathbf{Y}; \mathcal{C}_0)\|_S^2,\\ \|\mathbf{X} - \pi_S(\mathbf{X}; \mathcal{C}_2)\|_S^2 &= \|\mathbf{Y} - \pi_S(\mathbf{Y}; \mathcal{C}_2)\|_S^2\end{aligned}$$

and hence

$$\begin{aligned}T^{**} &= \|\mathbf{X} - \pi_S(\mathbf{X}; \mathcal{C}_0)\|_S^2 - \|\mathbf{X} - \pi_S(\mathbf{X}; \mathcal{C}_2)\|_S^2\\ &= \|\mathbf{Y} - \pi_S(\mathbf{Y}; \mathcal{C}_0)\|_S^2 - \|\mathbf{Y} - \pi_S(\mathbf{Y}; \mathcal{C}_2)\|_S^2.\end{aligned}$$

Let U be the $p \times p$ orthogonal matrix which satisfies (ii) of Lemma 5.1 and consider the linear transformation given by $(\mathrm{I}_k \otimes (U\mathbf{\Sigma}^{-1/2}))$; then, by Lemma 4.2, T^{**} can be rewritten as

$$\begin{aligned}T^{**} &= \|\mathbf{Y} - \pi_S(\mathbf{Y}; \mathcal{C}_0)\|_S^2 - \|\mathbf{Y} - \pi_S(\mathbf{Y}; \mathcal{C}_2)\|_S^2\\ &= \big\|(\mathrm{I}_k \otimes (U\mathbf{\Sigma}^{-1/2}))\mathbf{Y}\\ &\qquad - \pi_{U\mathbf{\Sigma}^{-1/2}S\mathbf{\Sigma}^{-1/2}U'}\big((\mathrm{I}_k \otimes (U\mathbf{\Sigma}^{-1/2}))\mathbf{Y};\\ &\qquad\qquad (\mathrm{I}_k \otimes (U\mathbf{\Sigma}^{-1/2}))\mathcal{C}_0\big)\big\|_{U\mathbf{\Sigma}^{-1/2}S\mathbf{\Sigma}^{-1/2}U'}^2\\ &\quad - \big\|(\mathrm{I}_k \otimes (U\mathbf{\Sigma}^{-1/2}))\mathbf{Y}\\ &\qquad - \pi_{U\mathbf{\Sigma}^{-1/2}S\mathbf{\Sigma}^{-1/2}U'}\big((\mathrm{I}_k \otimes (U\mathbf{\Sigma}^{-1/2}))\mathbf{Y};\\ &\qquad\qquad (\mathrm{I}_k \otimes (U\mathbf{\Sigma}^{-1/2}))\mathcal{C}_2\big)\big\|_{U\mathbf{\Sigma}^{-1/2}S\mathbf{\Sigma}^{-1/2}U'}^2.\end{aligned}$$

Put

$$\begin{aligned}\mathbf{Z} &= \begin{pmatrix}\mathbf{Z}_1\\ \vdots\\ \mathbf{Z}_k\end{pmatrix} = \begin{pmatrix}U\mathbf{\Sigma}^{-1/2}\mathbf{Y}_1\\ \vdots\\ U\mathbf{\Sigma}^{-1/2}\mathbf{Y}_k\end{pmatrix} = \big(\mathrm{I}_k \otimes (U\mathbf{\Sigma}^{-1/2})\big)\mathbf{Y},\\ S^* &= U\mathbf{\Sigma}^{-1/2}S\mathbf{\Sigma}^{-1/2}U';\end{aligned}$$

then, by Lemma 5.1, we have

$$T^{**} = \|\mathbf{Z} - \pi_{S^*}(\mathbf{Z}; \mathcal{C}_0)\|_{S^*}^2 - \|\mathbf{Z} - \pi_{S^*}(\mathbf{Z}; \mathcal{C}_2)\|_{S^*}^2.$$

By their definition, S^* and $\mathbf{Z}_1, \dots, \mathbf{Z}_k$ are mutually independent, S^* has the Wishart distribution $W_p(N-k; \mathrm{I}_p)$ and $\mathbf{Z}_i$ is distributed as $N_p(\mathbf{0}, N_i^{-1}\mathrm{I}_p)$, $i = 1, \dots, k$, hence the joint distribution of $(S^*, \mathbf{Z}_1, \dots, \mathbf{Z}_k)$ does not depend on $\boldsymbol{\mu}_0$ or $\boldsymbol{\Sigma}$. This completes the proof. □

LEMMA 5.3. *For any real number c,*

$$\sup_{\boldsymbol{\Sigma}} \sup_{\mathrm{H}_0} \mathrm{P}_{\boldsymbol{\mu}, \boldsymbol{\Sigma}}\big(\overline{T}^2 \geq c\big) = \sup_{\boldsymbol{\Sigma}} \mathrm{P}_{\mathbf{0}, \boldsymbol{\Sigma}}\big(\overline{T}^2 \geq c\big) \leq \mathrm{P}_{\mathbf{0}, \mathrm{I}_p}(T^{**} \geq c).$$

PROOF. Since $\mathcal{C}_1 \subset \mathcal{C}_2$, we have $\|\mathbf{X} - \pi_S(\mathbf{X}; \mathcal{C}_1)\|_S^2 \geq \|\mathbf{X} - \pi_S(\mathbf{X}; \mathcal{C}_2)\|_S^2$ and hence $\overline{T}^2 \leq T^{**}$. Thus, by Theorem 3.1 and Lemma 5.2, we have

$$\begin{aligned} \sup_{\boldsymbol{\Sigma}} \sup_{\mathrm{H}_0} \mathrm{P}_{\boldsymbol{\mu}, \boldsymbol{\Sigma}}\big(\overline{T}^2 \geq c\big) &= \sup_{\boldsymbol{\Sigma}} \mathrm{P}_{\mathbf{0}, \boldsymbol{\Sigma}}\big(\overline{T}^2 \geq c\big) \\ &\leq \sup_{\boldsymbol{\Sigma}} \mathrm{P}_{\mathbf{0}, \boldsymbol{\Sigma}}(T^{**} \geq c) \\ &= \mathrm{P}_{\mathbf{0}, \mathrm{I}_p}(T^{**} \geq c). \end{aligned}$$ □

In order to prove the inverse inequality of that in the above lemma, define a sequence of $p \times p$ nonsingular real matrices $\{B_n\}$ and a sequence of $p \times p$ positive definite real matrices $\{\boldsymbol{\Sigma}_n\}$ by

$$B_n = \begin{pmatrix} 1 & & & 0 \\ -n & 1 & & \\ \vdots & & \ddots & \\ -n & 0 & & 1 \end{pmatrix}, \qquad n = 1, 2, \dots,$$

and

$$\boldsymbol{\Sigma}_n = (B_n' B_n)^{-1}, \qquad n = 1, 2, \dots,$$

respectively.

LEMMA 5.4. *For the closed convex cones $\mathcal{C}_0$, $\mathcal{C}_1$ and $\mathcal{C}_2$, $\{B_n\}$ satisfies*:

(i) $(\mathrm{I}_k \otimes B_n)\mathcal{C}_0 = \mathcal{C}_0$, $n = 1, 2, \dots$,
(ii) $\mathcal{C}_1 \subset (\mathrm{I}_k \otimes B_1)\mathcal{C}_1 \subset (\mathrm{I}_k \otimes B_2)\mathcal{C}_1 \subset \cdots \subset (\mathrm{I}_k \otimes B_n)\mathcal{C}_1 \subset \cdots$,
(iii) $\overline{\bigcup_{n=1}^{\infty} (\mathrm{I}_k \otimes B_n)\mathcal{C}_1} = \mathcal{C}_2$.

LEMMA 5.5. *For any real number c and $n = 1, 2, \dots$,*

$$\begin{aligned} &\mathrm{P}_{\mathbf{0}, \boldsymbol{\Sigma}_n}\big(\overline{T}^2 \geq c\big) \\ &\quad = \mathrm{P}_{\mathbf{0}, \mathrm{I}_p}\Big(\|\mathbf{X} - \pi_S(\mathbf{X}; \mathcal{C}_0)\|_S^2 - \big\|\mathbf{X} - \pi_S\big(\mathbf{X}; (\mathrm{I}_k \otimes B_n)\mathcal{C}_1\big)\big\|_S^2 \geq c\Big). \end{aligned}$$

Proofs of the above two lemmas will be given in the Appendix.

LEMMA 5.6. *For any real number* c,

$$\sup_{\boldsymbol{\Sigma}} \sup_{\mathrm{H}_0} \mathrm{P}_{\boldsymbol{\mu},\boldsymbol{\Sigma}}\big(\overline{T}^2 \geq c\big) = \sup_{\boldsymbol{\Sigma}} \mathrm{P}_{\mathbf{0},\boldsymbol{\Sigma}}\big(\overline{T}^2 \geq c\big) \geq \mathrm{P}_{\mathbf{0},\mathrm{I}_p}(T^{**} \geq c).$$

PROOF. By Lemmas 5.5, 4.3 and 5.4,

$$\begin{aligned}
&\lim_{n\to\infty} \mathrm{P}_{\mathbf{0},\boldsymbol{\Sigma}_n}\big(\overline{T}^2 \geq c\big)\\
&\quad= \lim_{n\to\infty} \mathrm{P}_{\mathbf{0},\mathrm{I}_p}\Big(\|\mathbf{X} - \pi_S(\mathbf{X}; \mathcal{C}_0)\|_S^2 - \|\mathbf{X} - \pi_S\big(\mathbf{X}; (\mathrm{I}_k \otimes B_n)\mathcal{C}_1\big)\|_S^2 \geq c\Big)\\
&\quad= \mathrm{P}_{\mathbf{0},\mathrm{I}_p}\Big(\|\mathbf{X} - \pi_S(\mathbf{X}; \mathcal{C}_0)\|_S^2 - \lim_{n\to\infty} \|\mathbf{X} - \pi_S\big(\mathbf{X}; (\mathrm{I}_k \otimes B_n)\mathcal{C}_1\big)\|_S^2 \geq c\Big)\\
&\quad= \mathrm{P}_{\mathbf{0},\mathrm{I}_p}\left(\|\mathbf{X} - \pi_S(\mathbf{X}; \mathcal{C}_0)\|_S^2 - \left\|\mathbf{X} - \pi_S\left(\mathbf{X}; \overline{\bigcup_{n=1}^{\infty}(\mathrm{I}_k \otimes B_n)\mathcal{C}_1}\right)\right\|_S^2 \geq c\right)\\
&\quad= \mathrm{P}_{\mathbf{0},\mathrm{I}_p}\big(\|\mathbf{X} - \pi_S(\mathbf{X}; \mathcal{C}_0)\|_S^2 - \|\mathbf{X} - \pi_S(\mathbf{X}; \mathcal{C}_2)\|_S^2 \geq c\big)\\
&\quad= \mathrm{P}_{\mathbf{0},\mathrm{I}_p}(T^{**} \geq c).
\end{aligned}$$

Thus, by Theorem 3.1, we get

$$\begin{aligned}
&\sup_{\boldsymbol{\Sigma}} \sup_{\mathrm{H}_0} \mathrm{P}_{\boldsymbol{\mu},\boldsymbol{\Sigma}}\big(\overline{T}^2 \geq c\big)\\
&\qquad= \sup_{\boldsymbol{\Sigma}} \mathrm{P}_{\mathbf{0},\boldsymbol{\Sigma}}\big(\overline{T}^2 \geq c\big) \geq \lim_{n\to\infty} \mathrm{P}_{\mathbf{0},\boldsymbol{\Sigma}_n}\big(\overline{T}^2 \geq c\big) = \mathrm{P}_{\mathbf{0},\mathrm{I}_p}(T^{**} \geq c). \qquad \square
\end{aligned}$$

PROOF OF THEOREM 3.2. By Lemma 5.2, it is sufficient to show that $T^* = T^{**}$. Recall that

$$T^{**} = \|\mathbf{X} - \pi_S(\mathbf{X}; \mathcal{C}_0)\|_S^2 - \|\mathbf{X} - \pi_S(\mathbf{X}; \mathcal{C}_2)\|_S^2.$$

For the first term of T^{**}, as we have seen at the beginning of this section,

$$\|\mathbf{X} - \pi_S(\mathbf{X}; \mathcal{C}_0)\|_S^2 = \sum_{i=1}^{k} N_i(\overline{\mathbf{X}}_i - \overline{\mathbf{X}})' S^{-1}(\overline{\mathbf{X}}_i - \overline{\mathbf{X}}).$$

For the second term of T^{**},

$$\begin{aligned}
\|\mathbf{X} - \pi_S(\mathbf{X}; \mathcal{C}_2)\|_S^2 &= \min_{\boldsymbol{\mu}\in\mathcal{C}_2} \|\mathbf{X} - \boldsymbol{\mu}\|_S^2\\
&= \min_{\boldsymbol{\mu}\in\mathcal{C}_2} \sum_{i=1}^{k} N_i(\overline{\mathbf{X}}_i - \boldsymbol{\mu}_i)' S^{-1}(\overline{\mathbf{X}}_i - \boldsymbol{\mu}_i).
\end{aligned}$$

Let $\boldsymbol{\mu}_i, \overline{\mathbf{X}}_i$ and S be partitioned as

$$\boldsymbol{\mu}_i = \begin{pmatrix} \mu_{i1} \\ \boldsymbol{\mu}_{i2} \end{pmatrix}, \qquad \overline{\mathbf{X}}_i = \begin{pmatrix} \overline{X}_{i1} \\ \overline{\mathbf{X}}_{i2} \end{pmatrix}, \qquad S = \left(\begin{array}{c|c} s_{11} & S_{12} \\ \hline S_{21} & S_{22} \end{array}\right),$$

where μ_{i1}, $\overline{X}_{i1}$ and s_{11} are scalars, $i = 1, 2, \ldots k$. Then, by a well-known result on a partitioned matrix [e.g., see Anderson (1984), Appendix A],

$$
\begin{aligned}
&(\overline{\mathbf{X}}_i - \boldsymbol{\mu}_i)' S^{-1} (\overline{\mathbf{X}}_i - \boldsymbol{\mu}_i) \\
&\quad = \begin{pmatrix} \overline{X}_{i1} - \mu_{i1} \\ \overline{\mathbf{X}}_{i2} - \boldsymbol{\mu}_{i2} \end{pmatrix}' \begin{pmatrix} s_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}^{-1} \begin{pmatrix} \overline{X}_{i1} - \mu_{i1} \\ \overline{\mathbf{X}}_{i2} - \boldsymbol{\mu}_{i2} \end{pmatrix} \\
&\quad = \frac{1}{s_{11}} (\overline{X}_{i1} - \mu_{i1})^2 \\
&\qquad + \left\{ \overline{\mathbf{X}}_{i2} - \frac{S_{21}}{s_{11}} (\overline{X}_{i1} - \mu_{i1}) - \boldsymbol{\mu}_{i2} \right\}' S_{22.1}^{-1} \left\{ \overline{\mathbf{X}}_{i2} - \frac{S_{21}}{s_{11}} (\overline{X}_{i1} - \mu_{i1}) - \boldsymbol{\mu}_{i2} \right\},
\end{aligned}
$$

where $S_{22.1} = S_{22} - S_{21}^2 / s_{11}$. Hence

$$
\begin{aligned}
&\|\mathbf{X} - \pi_S(\mathbf{X}; \mathcal{C}_2)\|_S^2 \\
&\quad = \min_{\boldsymbol{\mu} \in \mathcal{C}_2} \sum_{i=1}^{k} N_i \left[\frac{1}{s_{11}} (\overline{X}_{i1} - \mu_{i1})^2 + \left\{ \overline{\mathbf{X}}_{i2} - \frac{S_{21}}{s_{11}} (\overline{X}_{i1} - \mu_{i1}) - \boldsymbol{\mu}_{i2} \right\}' \right. \\
&\qquad\qquad \left. \times S_{22.1}^{-1} \left\{ \overline{\mathbf{X}}_{i2} - \frac{S_{21}}{s_{11}} (\overline{X}_{i1} - \mu_{i1}) - \boldsymbol{\mu}_{i2} \right\} \right] \\
&\quad = \min_{\mu_{11} \le \cdots \le \mu_{k1}} \min_{\boldsymbol{\mu}_{12}, \ldots, \boldsymbol{\mu}_{k2}} \left[\frac{1}{s_{11}} \sum_{i=1}^{k} N_i (\overline{X}_{i1} - \mu_{i1})^2 \right. \\
&\qquad\qquad + \sum_{i=1}^{k} N_i \left\{ \overline{\mathbf{X}}_{i2} - \frac{S_{21}}{s_{11}} (\overline{X}_{i1} - \mu_{i1}) - \boldsymbol{\mu}_{i2} \right\}' \\
&\qquad\qquad \left. \times S_{22.1}^{-1} \left\{ \overline{\mathbf{X}}_{i2} - \frac{S_{21}}{s_{11}} (\overline{X}_{i1} - \mu_{i1}) - \boldsymbol{\mu}_{i2} \right\} \right] \\
&\quad = \min_{\mu_{11} \le \cdots \le \mu_{k1}} \left[\frac{1}{s_{11}} \sum_{i=1}^{k} N_i (\overline{X}_{i1} - \mu_{i1})^2 \right. \\
&\qquad\qquad + \min_{\boldsymbol{\mu}_{12}, \ldots, \boldsymbol{\mu}_{k2}} \sum_{i=1}^{k} N_i \left\{ \overline{\mathbf{X}}_{i2} - \frac{S_{21}}{s_{11}} (\overline{X}_{i1} - \mu_{i1}) - \boldsymbol{\mu}_{i2} \right\}' \\
&\qquad\qquad \left. \times S_{22.1}^{-1} \left\{ \overline{\mathbf{X}}_{i2} - \frac{S_{21}}{s_{11}} (\overline{X}_{i1} - \mu_{i1}) - \boldsymbol{\mu}_{i2} \right\} \right] \\
&\quad = \min_{\mu_{11} \le \cdots \le \mu_{k1}} \frac{1}{s_{11}} \sum_{i=1}^{k} N_i (\overline{X}_{i1} - \mu_{i1})^2 = \frac{1}{s_{11}} \sum_{i=1}^{k} N_i (\overline{X}_{i1} - \widehat{\mu}_{i1})^2,
\end{aligned}
$$

where $(\widehat{\mu}_{11}, \ldots, \widehat{\mu}_{k1})$ is the IR of $\overline{X}_{11}, \ldots, \overline{X}_{k1}$ with weights $N_1, \ldots, N_k$. Thus we have obtained

$$T^{**} = \sum_{i=1}^{k} N_i(\overline{\mathbf{X}}_i - \overline{\mathbf{X}})' S^{-1}(\overline{\mathbf{X}}_i - \overline{\mathbf{X}}) - \frac{1}{s_{11}} \sum_{i=1}^{k} N_i(\overline{X}_{i1} - \widehat{\mu}_{i1})^2 = T^*. \qquad \square$$

PROOF OF THEOREM 3.3. By Lemmas 5.3 and 5.6, we have

$$\sup_{\boldsymbol{\Sigma}} \sup_{\mathrm{H}_0} \mathrm{P}_{\boldsymbol{\mu},\boldsymbol{\Sigma}}\big(\overline{T}^2 \geq c\big) = \sup_{\boldsymbol{\Sigma}} \mathrm{P}_{\mathbf{0},\boldsymbol{\Sigma}}\big(\overline{T}^2 \geq c\big) = \mathrm{P}_{\mathbf{0},\mathrm{I}_p}(T^{**} \geq c).$$

As we have shown in the proof of Theorem 3.2, $T^* = T^{**}$, thus we complete the proof. $\square$

6. Critical points estimated by simulation. By Theorem 3.3, in order to get the upper α point of $\overline{T}^2$ under H_0, we only need to obtain that of T^* when $\boldsymbol{\mu} = \mathbf{0}$ and $\boldsymbol{\Sigma} = \mathrm{I}_p$.

We generated $N_1 + N_2 + \cdots + N_k$ sets of p-variate normal random vectors from $N_p(\mathbf{0}, \mathrm{I}_p)$ and computed T^*. We repeated this computation 20,000 times to get an estimated upper α point of T^*. We further repeated this process 10 times and computed the average of the 10 estimated upper α points for $\alpha = 0.01, 0.05$, $p = 2, 5, 10$, $k = 2, 5, 10$, and $N_i = 5, 10, 20, 30$, $i = 1, 2, \ldots, k$, respectively, and we list them in Table 1.

TABLE 1
Upper α points estimated by simulation

			$N_1(= N_2 = \cdots = N_k)$			
α	p	k	5	10	20	30
0.01	2	2	2.4623	0.6471	0.2511	0.1572
		5	1.1673	0.4157	0.1810	0.1151
		10	0.7778	0.3053	0.1377	0.0889
	5	2	18.9569	1.6017	0.5083	0.2990
		5	3.2160	0.9683	0.3988	0.2526
		10	2.0458	0.7675	0.3419	0.2196
	10	2	—	5.8425	1.0101	0.5411
		5	10.5996	2.0200	0.7565	0.4645
		10	4.6477	1.5503	0.6623	0.4218
0.05	2	2	1.1673	0.3639	0.1488	0.0953
		5	0.7761	0.2936	0.1310	0.0836
		10	0.5847	0.2359	0.1079	0.0699
	5	2	7.6463	1.0071	0.3472	0.2081
		5	2.3794	0.7688	0.3247	0.2056
		10	1.7130	0.6587	0.2952	0.1901
	10	2	—	3.4425	0.7310	0.4018
		5	7.6404	1.6830	0.6483	0.4008
		10	4.0074	1.3824	0.5985	0.3812

7. Discussion. Our proposed test statistic, $\overline{\overline{T}}^2$, is obtained by replacing the unknown covariance matrix in the $\bar{\chi}^2_{k,p}$ statistic by its estimator. This method seems somewhat ad hoc. In this section, we discuss a justification of the plausibility of our test statistic.

Using Lemma 3.2.2 of Anderson (1984) in the way similar to that of Anderson [(1984), Section 8.8], in order to get the likelihood ratio test for our problem we need to minimize the determinant

$$\left| S + \sum_{i=1}^{k} N_i (\overline{\mathbf{X}}_i - \boldsymbol{\mu}_i)(\overline{\mathbf{X}}_i - \boldsymbol{\mu}_i)' \right|$$

under the restriction that $\boldsymbol{\mu}_1 \le \boldsymbol{\mu}_2 \le \cdots \le \boldsymbol{\mu}_k$. We have

$$\begin{aligned} &\left| S + \sum_{i=1}^{k} N_i (\overline{\mathbf{X}}_i - \boldsymbol{\mu}_i)(\overline{\mathbf{X}}_i - \boldsymbol{\mu}_i)' \right| \\ &\qquad = |S| \cdot \left| \mathrm{I}_p + S^{-1/2} \sum_{i=1}^{k} N_i (\overline{\mathbf{X}}_i - \boldsymbol{\mu}_i)(\overline{\mathbf{X}}_i - \boldsymbol{\mu}_i)' S^{-1/2} \right|, \end{aligned}$$

so we need to minimize

$$\left| \mathrm{I}_p + S^{-1/2} \sum_{i=1}^{k} N_i (\overline{\mathbf{X}}_i - \boldsymbol{\mu}_i)(\overline{\mathbf{X}}_i - \boldsymbol{\mu}_i)' S^{-1/2} \right|$$

under the above restriction. The difficulty in deriving the likelihood ratio test arises in this minimization problem.

Now suppose that A is a $p \times p$ non-negative definite real (symmetric) matrix, $\lambda_1, \lambda_2, \ldots, \lambda_p$ are the characteristic roots of A, and ε is a positive number. Then

$$|\mathrm{I}_p + \varepsilon A| = \prod_{i=1}^{p} (1 + \varepsilon \lambda_i) = 1 + \sum_{i=1}^{p} \varepsilon \lambda_i + O(\varepsilon^2) = 1 + \mathrm{tr}(\varepsilon A) + O(\varepsilon^2).$$

Taking this expansion into account, instead of minimizing the determinant, consider the problem of minimizing

$$\begin{aligned} &1 + \mathrm{tr}\left[S^{-1/2} \sum_{i=1}^{k} N_i (\overline{\mathbf{X}}_i - \boldsymbol{\mu}_i)(\overline{\mathbf{X}}_i - \boldsymbol{\mu}_i)' S^{-1/2} \right] \\ &\qquad = 1 + \sum_{i=1}^{k} N_i (\overline{\mathbf{X}}_i - \boldsymbol{\mu}_i)' S^{-1} (\overline{\mathbf{X}}_i - \boldsymbol{\mu}_i) \end{aligned}$$

under the restriction that $\boldsymbol{\mu}_1 \le \boldsymbol{\mu}_2 \le \cdots \le \boldsymbol{\mu}_k$. The solution of this minimization problem is nothing but the MIR of $\overline{\mathbf{X}}_1, \ldots, \overline{\mathbf{X}}_k$ with weights $N_1 S^{-1}, \ldots, N_k S^{-1}$. This consideration might add some justification to the plausibility of our test statistic.

APPENDIX

A. Proofs of some of the lemmas.

PROOF OF LEMMA 5.1.

(i)

$$\begin{aligned}(\mathrm{I}_k \otimes (U\boldsymbol{\Sigma}^{-1/2}))\mathcal{C}_0 &= \left\{ (\mathrm{I}_k \otimes (U\boldsymbol{\Sigma}^{-1/2})) \begin{pmatrix} \boldsymbol{\mu}_1 \\ \vdots \\ \boldsymbol{\mu}_k \end{pmatrix} \middle| \boldsymbol{\mu}_1 = \cdots = \boldsymbol{\mu}_k \right\} \\ &= \left\{ \begin{pmatrix} U\boldsymbol{\Sigma}^{-1/2}\boldsymbol{\mu}_1 \\ \vdots \\ U\boldsymbol{\Sigma}^{-1/2}\boldsymbol{\mu}_k \end{pmatrix} \middle| \boldsymbol{\mu}_1 = \cdots = \boldsymbol{\mu}_k \right\} \\ &= \mathcal{C}_0.\end{aligned}$$

(ii) $\mathcal{C}_2$ is rewritten as

$$\begin{aligned}\mathcal{C}_2 &= \left\{ \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \vdots \\ \boldsymbol{\mu}_k \end{pmatrix} \middle| \mu_{11} \le \cdots \le \mu_{k1} \right\} \\ &= \bigcap_{i=1}^{k-1} \left\{ \begin{pmatrix} \boldsymbol{\mu}_1 \\ \vdots \\ \boldsymbol{\mu}_k \end{pmatrix} \middle| \mu_{i1} \le \mu_{(i+1)1} \right\} \\ &= \bigcap_{i=1}^{k-1} \left\{ \begin{pmatrix} \boldsymbol{\mu}_1 \\ \vdots \\ \boldsymbol{\mu}_k \end{pmatrix} \middle| \mathbf{e}_1'(\boldsymbol{\mu}_{i+1} - \boldsymbol{\mu}_i) \ge 0 \right\},\end{aligned}$$

where $\mathbf{e}_1 = (1, 0, 0, \ldots, 0)'$. Hence

$$(\mathrm{I}_k \otimes \boldsymbol{\Sigma}^{-1/2})\mathcal{C}_2 = \bigcap_{i=1}^{k-1} \left\{ \begin{pmatrix} \boldsymbol{\Sigma}^{-1/2}\boldsymbol{\mu}_1 \\ \vdots \\ \boldsymbol{\Sigma}^{-1/2}\boldsymbol{\mu}_k \end{pmatrix} \middle| \mathbf{e}_1'(\boldsymbol{\mu}_{i+1} - \boldsymbol{\mu}_i) \ge 0 \right\}.$$

Put $\boldsymbol{\Sigma}^{-1/2}\boldsymbol{\mu}_i = \boldsymbol{\nu}_i$. Then $\boldsymbol{\mu}_i = \boldsymbol{\Sigma}^{1/2}\boldsymbol{\nu}_i$ and hence

$$(\mathrm{I}_k \otimes \boldsymbol{\Sigma}^{-1/2})\mathcal{C}_2 = \bigcap_{i=1}^{k-1} \left\{ \begin{pmatrix} \boldsymbol{\nu}_1 \\ \vdots \\ \boldsymbol{\nu}_k \end{pmatrix} \middle| \mathbf{e}_1'\boldsymbol{\Sigma}^{1/2}(\boldsymbol{\nu}_{i+1} - \boldsymbol{\nu}_i) \ge 0 \right\}.$$

Put

$$\mathbf{a}_1 = (\mathbf{e}_1'\boldsymbol{\Sigma}\mathbf{e}_1)^{-1/2}\boldsymbol{\Sigma}^{1/2}\mathbf{e}_1 = (\sigma_{11})^{-1/2}\boldsymbol{\Sigma}^{1/2}\mathbf{e}_1,$$

where σ_{11} is the $(1,1)$th element of $\mathbf{\Sigma}$. Then $\mathbf{a}_1'\mathbf{a}_1 = 1$ and hence there exists a $p \times p$ orthogonal matrix U such that $U\mathbf{a}_1 = \mathbf{e}_1$ since $\mathbf{e}_1'\mathbf{e}_1 = 1$. For this U, we have

$$\bigl(\mathrm{I}_k \otimes (U\mathbf{\Sigma}^{-1/2})\bigr)\mathcal{C}_2 = \bigcap_{i=1}^{k-1} \left\{ \begin{pmatrix} U\boldsymbol{\nu}_1 \\ \vdots \\ U\boldsymbol{\nu}_k \end{pmatrix} \middle| \mathbf{e}_1'\mathbf{\Sigma}^{1/2}(\boldsymbol{\nu}_{i+1} - \boldsymbol{\nu}_i) \geq 0 \right\}.$$

Put $U\boldsymbol{\nu}_i = \boldsymbol{\eta}_i$. Then $\boldsymbol{\nu}_i = U'\boldsymbol{\eta}_i$ and hence

$$\bigl(\mathrm{I}_k \otimes (U\mathbf{\Sigma}^{-1/2})\bigr)\mathcal{C}_2 = \bigcap_{i=1}^{k-1} \left\{ \begin{pmatrix} \boldsymbol{\eta}_1 \\ \vdots \\ \boldsymbol{\eta}_k \end{pmatrix} \middle| \mathbf{e}_1'\mathbf{\Sigma}^{1/2}U'(\boldsymbol{\eta}_{i+1} - \boldsymbol{\eta}_i) \geq 0 \right\}.$$

By the definition of U, $U\mathbf{\Sigma}^{1/2}\mathbf{e}_1 = (\sigma_{11})^{1/2}\mathbf{e}_1$, and hence $\mathbf{e}_1'\mathbf{\Sigma}^{1/2}U' = (\sigma_{11})^{1/2}\mathbf{e}_1'$ since $\mathbf{\Sigma}^{1/2}$ is symmetric. Thus we obtain

$$\begin{aligned} \bigl(\mathrm{I}_k \otimes (U\mathbf{\Sigma}^{-1/2})\bigr)\mathcal{C}_2 &= \bigcap_{i=1}^{k-1} \left\{ \begin{pmatrix} \boldsymbol{\eta}_1 \\ \vdots \\ \boldsymbol{\eta}_k \end{pmatrix} \middle| (\sigma_{11})^{1/2}\mathbf{e}_1'(\boldsymbol{\eta}_{i+1} - \boldsymbol{\eta}_i) \geq 0 \right\} \\ &= \bigcap_{i=1}^{k-1} \left\{ \begin{pmatrix} \boldsymbol{\eta}_1 \\ \vdots \\ \boldsymbol{\eta}_k \end{pmatrix} \middle| \mathbf{e}_1'(\boldsymbol{\eta}_{i+1} - \boldsymbol{\eta}_i) \geq 0 \right\} \\ &= \mathcal{C}_2. \end{aligned}$$

□

PROOF OF LEMMA 5.4.
(i)

$$(\mathrm{I}_k \otimes B_n)\mathcal{C}_0 = \left\{ \begin{pmatrix} B_n\boldsymbol{\mu}_1 \\ \vdots \\ B_n\boldsymbol{\mu}_k \end{pmatrix} \middle| \boldsymbol{\mu}_1 = \cdots = \boldsymbol{\mu}_k \right\} = \mathcal{C}_0.$$

(ii) Put $\boldsymbol{\nu}_i = B_n\boldsymbol{\mu}_i$, $i = 1, 2, \ldots, k$; then

$$\begin{aligned} (\mathrm{I}_k \otimes B_n)\mathcal{C}_1 &= \left\{ \begin{pmatrix} B_n\boldsymbol{\mu}_1 \\ \vdots \\ B_n\boldsymbol{\mu}_k \end{pmatrix} \middle| \boldsymbol{\mu}_1 \leq \cdots \leq \boldsymbol{\mu}_k \right\} \\ &= \left\{ \begin{pmatrix} \boldsymbol{\nu}_1 \\ \vdots \\ \boldsymbol{\nu}_k \end{pmatrix} \middle| B_n^{-1}\boldsymbol{\nu}_1 \leq \cdots \leq B_n^{-1}\boldsymbol{\nu}_k \right\}. \end{aligned}$$

By the definition of B_n,

$$B_n^{-1} = \begin{pmatrix} 1 & & & 0 \\ n & 1 & & \\ \vdots & & \ddots & \\ n & 0 & & 1 \end{pmatrix}, \qquad n = 1, 2, \ldots.$$

Hence put

$$\boldsymbol{\nu}_i = \begin{pmatrix} \nu_{i1} \\ \nu_{i2} \\ \vdots \\ \nu_{ip} \end{pmatrix}, \qquad i = 1, 2, \dots, k.$$

Then

$$(\mathrm{I}_k \otimes B_n)\mathcal{C}_1 = \left\{ \begin{pmatrix} \boldsymbol{\nu}_1 \\ \vdots \\ \boldsymbol{\nu}_k \end{pmatrix} \middle| \begin{array}{c} \nu_{11} \le \cdots \le \nu_{k1} \\ n\nu_{11} + \nu_{12} \le \cdots \le n\nu_{k1} + \nu_{k2} \\ \vdots \\ n\nu_{11} + \nu_{1p} \le \cdots \le n\nu_{k1} + \nu_{kp} \end{array} \right\}.$$

From the above expression, it is easy to see that

$$\mathcal{C}_1 \subset (\mathrm{I}_k \otimes B_1)\mathcal{C}_1 \subset (\mathrm{I}_k \otimes B_2)\mathcal{C}_1 \subset \cdots \subset (\mathrm{I}_k \otimes B_n)\mathcal{C}_1 \subset \cdots.$$

(iii)

$$(\mathrm{I}_k \otimes B_n)\mathcal{C}_1 = \left\{ \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \vdots \\ \boldsymbol{\mu}_k \end{pmatrix} \middle| \begin{array}{c} \mu_{11} \le \cdots \le \mu_{k1} \\ n\mu_{11} + \mu_{12} \le \cdots \le n\mu_{k1} + \mu_{k2} \\ \vdots \\ n\mu_{11} + \mu_{1p} \le \cdots \le n\mu_{k1} + \mu_{kp} \end{array} \right\}$$

$$= \bigcap_{i=1}^{k-1} \left\{ \boldsymbol{\mu} \middle| \begin{array}{l} \mu_{(i+1)1} - \mu_{i1} \ge 0 \\ \mu_{(i+1)2} - \mu_{i2} \ge -n(\mu_{(i+1)1} - \mu_{i1}) \\ \vdots \\ \mu_{(i+1)p} - \mu_{ip} \ge -n(\mu_{(i+1)1} - \mu_{i1}) \end{array} \right\}$$

$$= \bigcap_{i=1}^{k-1} \bigcap_{j=2}^{p} \left\{ \boldsymbol{\mu} \middle| \begin{array}{l} \mu_{(i+1)1} - \mu_{i1} \ge 0 \\ \mu_{(i+1)j} - \mu_{ij} \ge -n(\mu_{(i+1)1} - \mu_{i1}) \end{array} \right\}.$$

Put

$$\mathcal{C}_{n,ij} = \left\{ \boldsymbol{\mu} \middle| \begin{array}{l} \mu_{(i+1)1} - \mu_{i1} \ge 0 \\ \mu_{(i+1)j} - \mu_{ij} \ge -n(\mu_{(i+1)1} - \mu_{i1}) \end{array} \right\}.$$

Then

$$\bigcup_{n=1}^{\infty} (\mathrm{I}_k \otimes B_n)\mathcal{C}_1 = \bigcup_{n=1}^{\infty} \bigcap_{i=1}^{k-1} \bigcap_{j=2}^{p} \mathcal{C}_{n,ij} = \bigcap_{i=1}^{k-1} \bigcap_{j=2}^{p} \bigcup_{n=1}^{\infty} \mathcal{C}_{n,ij}.$$

It is easy to see that the second equality holds true since $\mathcal{C}_{n,ij} \subset \mathcal{C}_{n+1,ij}$ for $n = 1, 2, \dots;\ i = 1, 2, \dots, k-1;\ j = 2, 3, \dots, p$. Put

$$\mathcal{C}_{ij} = \left\{ \boldsymbol{\mu} \middle| \begin{array}{l} \mu_{(i+1)1} - \mu_{i1} \ge 0 \\ \mu_{(i+1)j} - \mu_{ij} \ge 0 \end{array} \right\} \cup \left\{ \boldsymbol{\mu} \middle| \begin{array}{l} \mu_{(i+1)1} - \mu_{i1} > 0 \\ \mu_{(i+1)j} - \mu_{ij} \le 0 \end{array} \right\}.$$

Then, by Lemma 4.1,

$$\bigcup_{n=1}^{\infty} \mathcal{C}_{n,ij} = \mathcal{C}_{ij},$$

hence

$$\bigcup_{n=1}^{\infty} (\mathrm{I}_k \otimes B_n)\mathcal{C}_1 = \bigcap_{i=1}^{k-1} \bigcap_{j=2}^{p} \mathcal{C}_{ij}.$$

Now

$$\begin{aligned}\mathcal{C}_{ij} &\subset \left[\left\{\boldsymbol{\mu} \,\middle|\, \begin{matrix}\mu_{(i+1)1} - \mu_{i1} \geq 0 \\ \mu_{(i+1)j} - \mu_{ij} \geq 0\end{matrix}\right\} \cup \left\{\boldsymbol{\mu} \,\middle|\, \begin{matrix}\mu_{(i+1)1} - \mu_{i1} \geq 0 \\ \mu_{(i+1)j} - \mu_{ij} \leq 0\end{matrix}\right\}\right] \\ &= \{\boldsymbol{\mu} | \mu_{(i+1)1} - \mu_{i1} \geq 0\}\end{aligned}$$

and

$$\begin{aligned}\mathcal{C}_{ij} &\supset \left[\left\{\boldsymbol{\mu} \,\middle|\, \begin{matrix}\mu_{(i+1)1} - \mu_{i1} > 0 \\ \mu_{(i+1)j} - \mu_{ij} \geq 0\end{matrix}\right\} \cup \left\{\boldsymbol{\mu} \,\middle|\, \begin{matrix}\mu_{(i+1)1} - \mu_{i1} > 0 \\ \mu_{(i+1)j} - \mu_{ij} \leq 0\end{matrix}\right\}\right] \\ &= \{\boldsymbol{\mu} | \mu_{(i+1)1} - \mu_{i1} > 0\},\end{aligned}$$

hence

$$\bigcup_{n=1}^{\infty} (\mathrm{I}_k \otimes B_n)\mathcal{C}_1 \subset \bigcap_{i=1}^{k-1} \{\boldsymbol{\mu} | \mu_{(i+1)1} - \mu_{i1} \geq 0\} = \{\boldsymbol{\mu} | \mu_{11} \leq \cdots \leq \mu_{k1}\},$$

$$\bigcup_{n=1}^{\infty} (\mathrm{I}_k \otimes B_n)\mathcal{C}_1 \supset \bigcap_{i=1}^{k-1} \{\boldsymbol{\mu} | \mu_{(i+1)1} - \mu_{i1} > 0\} = \{\boldsymbol{\mu} | \mu_{11} < \cdots < \mu_{k1}\}.$$

Thus we obtain

$$\overline{\bigcup_{n=1}^{\infty} (\mathrm{I}_k \otimes B_n)\mathcal{C}_1} = \{\boldsymbol{\mu} | \mu_{11} \leq \cdots \leq \mu_{k1}\} = \mathcal{C}_2.$$

□

PROOF OF LEMMA 5.5. Put

$$\mathbf{X}^* = \begin{pmatrix}\mathbf{X}_1^* \\ \vdots \\ \mathbf{X}_k^*\end{pmatrix} = \begin{pmatrix}B_n\overline{\mathbf{X}}_1 \\ \vdots \\ B_n\overline{\mathbf{X}}_k\end{pmatrix} = (\mathrm{I}_k \otimes B_n)\mathbf{X}, \qquad S^* = B_n S B_n'.$$

By Lemmas 4.2 and 5.4,

$$\begin{aligned}&\|\mathbf{X} - \pi_S(\mathbf{X}; \mathcal{C}_0)\|_S^2 \\ &\quad = \big\|(\mathrm{I}_k \otimes B_n)\mathbf{X} - \pi_{B_n S B_n'}\big((\mathrm{I}_k \otimes B_n)\mathbf{X}; (\mathrm{I}_k \otimes B_n)\mathcal{C}_0\big)\big\|_{B_n S B_n'}^2 \\ &\quad = \|\mathbf{X}^* - \pi_{S^*}(\mathbf{X}^*; \mathcal{C}_0)\|_{S^*}^2.\end{aligned}$$

Similarly, by Lemma 4.2,

$$\|\mathbf{X} - \pi_S(\mathbf{X}; \mathcal{C}_1)\|_S^2 = \big\|\mathbf{X}^* - \pi_{S^*}\big(\mathbf{X}^*; (\mathrm{I}_k \otimes B_n)\mathcal{C}_1\big)\big\|_{S^*}^2.$$

Now $\mathrm{P}_{\mathbf{0},\boldsymbol{\Sigma}_n}(\overline{T}^2 \geq c)$ is the probability that

$$\overline{T}^2 = \|\mathbf{X} - \pi_S(\mathbf{X}; \mathcal{C}_0)\|_S^2 - \|\mathbf{X} - \pi_S(\mathbf{X}; \mathcal{C}_1)\|_S^2 \geq c$$

when $\overline{\mathbf{X}}_i \sim N_p(\mathbf{0}, N_i^{-1}\boldsymbol{\Sigma}_n)$, $i = 1, 2, \dots, k$, and $S \sim W_p(N - k; \boldsymbol{\Sigma}_n)$. This probability can be rewritten as the probability that

$$\|\mathbf{X}^* - \pi_{S^*}(\mathbf{X}^*; \mathcal{C}_0)\|_{S^*}^2 - \big\|\mathbf{X}^* - \pi_{S^*}\big(\mathbf{X}^*; (\mathrm{I}_k \otimes B_n)\mathcal{C}_1\big)\big\|_{S^*}^2 \geq c$$

when $\mathbf{X}_i^* \sim N_p(\mathbf{0}, N_i^{-1}\mathrm{I}_p)$, $i = 1, 2, \dots, k$, and $S^* \sim W_p(N - k; \mathrm{I}_p)$, since $\boldsymbol{\Sigma}_n = (B_n' B_n)^{-1}$, while $\mathrm{P}_{\mathbf{0},\mathrm{I}_p}(\|\mathbf{X} - \pi_S(\mathbf{X}; \mathcal{C}_0)\|_S^2 - \|(\mathbf{X} - \pi_S(\mathbf{X}; (\mathrm{I}_k \otimes B_n)\mathcal{C}_1)\|_S^2 \geq c)$ is the probability that

$$\|\mathbf{X} - \pi_S(\mathbf{X}; \mathcal{C}_0)\|_S^2 - \big\|(\mathbf{X} - \pi_S\big(\mathbf{X}; (\mathrm{I}_k \otimes B_n)\mathcal{C}_1\big)\big\|_S^2 \geq c$$

when $\overline{\mathbf{X}}_i \sim N_p(\mathbf{0}, N_i^{-1}\mathrm{I}_p)$, $i = 1, 2, \dots, k$, and $S \sim W_p(N - k; \mathrm{I}_p)$. Thus the proof is complete. $\square$

Acknowledgments. The authors are deeply grateful to the referees for their valuable comments and suggestions. This paper is dedicated to the memory of the late Professor Akio Kudô, who had been a supervisor of the first author and who passed away in February, 2003.

REFERENCES

ANDERSON, T. W. (1984). *An Introduction to Multivariate Statistical Analysis*, 2nd ed. Wiley, New York.

BARLOW, R. E., BARTHOLOMEW, D. J., BREMNER, J. M. and BRUNK, H. D. (1972). *Statistical Inference under Order Restrictions*: *The Theory and Application of Isotonic Regression*. Wiley, New York.

BARTHOLOMEW, D. J. (1959). A test of homogeneity for ordered alternatives. *Biometrika* **46** 36–48.

BARTHOLOMEW, D. J. (1961). Ordered tests in the analysis of variance. *Biometrika* **48** 325–332.

KULATUNGA, D. D. S. (1984). Convolutions of the probabilities $P(l, k)$ used in order restricted inference. *Mem. Fac. Sci. Kyushu Univ. Ser. A Math.* **38** 9–15.

KULATUNGA, D. D. S. and SASABUCHI, S. (1984). A test of homogeneity of mean vectors against multivariate isotonic alternatives. *Mem. Fac. Sci. Kyushu Univ. Ser. A Math.* **38** 151–161.

NOMAKUCHI, K. and SHI, N.-Z. (1988). A test for a multiple isotonic regression problem. *Biometrika* **75** 181–184.

PERLMAN, M. D. (1969). One-sided testing problems in multivariate analysis. *Ann. Math. Statist.* **40** 549–567.

ROBERTSON, T., WRIGHT, F. T. and DYKSTRA, R. L. (1988). *Order Restricted Statistical Inference*. Wiley, New York.

SASABUCHI, S., INUTSUKA, M. and KULATUNGA, D. D. S. (1983). A multivariate version of isotonic regression. *Biometrika* **70** 465–472.

SASABUCHI, S., INUTSUKA, M. and KULATUNGA, D. D. S. (1992). An algorithm for computing multivariate isotonic regression. *Hiroshima Math. J.* **22** 551–560.

SASABUCHI, S., KULATUNGA, D. D. S. and SAITO, H. (1998). Comparison of powers of some tests for testing homogeneity under order restrictions in multivariate normal means. *Amer. J. Math. Management Sci.* **18** 131–158.

S. SASABUCHI
COMMON TECHNICAL COURSES
KYUSHU INSTITUTE OF DESIGN
SHIOBARU, MINAMI-KU
FUKUOKA-SHI, 815-8540
JAPAN
E-MAIL: sasabuti@kyushu-id.ac.jp

K. TANAKA
TOSU HIGH SCHOOL
SAGA
JAPAN

T. TSUKAMOTO
HITACHI SOFTWARE CORPORATION
TOKYO
JAPAN