



On the test for the homogeneity of a parameter matrix with some rows constrained by synchronized order restrictions

Xiaomi Hu^{a,*}, Arijit Banerjee^b

^a Department of Mathematics and Statistics, Wichita State University, Wichita, KS 67260-0033, USA

^b Wichita State University, Wichita, KS 67260-0033, USA

ARTICLE INFO

Article history:

Received 11 June 2010

Available online 10 January 2012

AMS 2000 subject classifications:

62F30

62H15

Keywords:

Closed convex cone

Hypothesis testing

Multivariate linear regression

Power function

Projection

ABSTRACT

The tests on the homogeneity of the columns of the coefficient matrix in a multiple multivariate linear regression with some rows of the matrix constrained by synchronized orderings, using the test statistics obtained by replacing the unknown variance–covariance matrix with its estimator in likelihood ratio test statistics, form a family of ad hoc tests. It is shown in this paper that the tests in the family share the same alpha-level critical values and follow the same distributions for computing their p -values. A sufficient condition is established for other tests to enjoy these properties, and to be more powerful. Two such more powerful tests are examined.

© 2012 Elsevier Inc. All rights reserved.

1. Introduction

Let $\beta \in R^{p \times q}$ and positive definite $\Sigma \in R^{p \times p}$ ($\Sigma > 0$) be the coefficient matrix and the variance–covariance matrix of a multivariate multiple linear regression model. Suppose that the elements of some of the rows of β are constrained by synchronized orderings, the monotone non-decreasing or monotone non-increasing orderings with respect to a common partial order of column indices. For this model we consider the test of the hypothesis that the columns of β are all equal using an ad hoc test statistic obtained by replacing the unknown Σ in a likelihood ratio test (LRT) statistic with its estimator. The row indices of the restricted rows of β form a non-empty set $H \subset \{1, \dots, p\}$. With all possible H , a family of ad hoc tests is produced.

It is shown in this paper that all members of the ad hoc test family share the same α -level critical values, and a unique distribution can be employed for computing the p -values for all members of the family. A sufficient condition is established for other tests to enjoy the above two properties, and to be more powerful. Two such more powerful tests are examined.

The study of the tests on the homogeneity of multiple parameters with an alternative specifying an ordering dates back to the late 50s. Bartholomew [2] derived an LRT on the homogeneity of independent normal means against a simple ordering. Since then researchers have studied many cases of hypothesis testing under constraints. Mukerjee and Tu [5] investigated the inferences for a polynomial regression with the regression function restricted to be monotonic. For the testing on the restricted components of a multivariate normal mean vector Silvapulle [9] proposed a Hotelling's T^2 -type test statistic by replacing the unknown variance–covariance matrix with its estimator. Many results of the statistical inferences on order restricted parameters are summarized in [1,6,10].

* Corresponding author.

E-mail address: xiaomi.hu@wichita.edu (X. Hu).

Sasabuchi et al. [8] extend the test in [2] to a test on vector homogeneity against vector simple ordering in which the relation of $\mu_i \leq \mu_j$ for vectors μ_i and μ_j is interpreted as “componentwise less than or equal to”. The model restriction in [8] is later generalized by Hu [4] from componentwise simple ordering to a general vector quasi ordering. Recently, Sasabuchi [7] explored a more powerful test than that in [8]. The work in this paper is the continuation of the research in this direction, and extended and generalized the results in [7,4].

In the next section we give preliminaries on the problem and explain the notations. Two lemmas and two theorems are presented in Section 3. The proofs of the lemmas, however, are placed in the appendices. In Section 4 we give the main results on the tests. These results are directly based on the conclusions of the two theorems in Section 3.

2. Preliminaries

2.1. Order restrictions

Let $\ll_{(i)}$ be a partial order in $\Omega = \{1, \dots, q\}$, and $\mathcal{C}_{(i)}$ be either $\{v \in R^q : v_s \leq v_t \text{ for } s \ll_{(i)} t\}$ or $\{v \in R^q : v_s \geq v_t \text{ for } s \ll_{(i)} t\}$, $i = 1, \dots, p$. The rows matrix $(\beta_{(1)}, \dots, \beta_{(p)})' \in R^{p \times q}$ are under separate order restrictions if $\beta_{(i)} \in \mathcal{C}_{(i)}$ for all $i = 1, \dots, p$. These separate restrictions are said to be synchronized if $\ll_{(i)}$, $i = 1, \dots, p$, are identical. Synchronized order restrictions are essentially isotonic restrictions and antitonic restrictions with respect to a common partial order \ll . For this \ll , let \mathcal{C}_+ and \mathcal{C}_- be the collections of all isotonic and antitonic vectors, and $\mathcal{C}_{(i)}$ be either \mathcal{C}_+ or \mathcal{C}_- . With H , a non-empty subset of $\{1, \dots, p\}$,

$$\mathcal{C}_H = \{(\beta_{(1)}, \dots, \beta_{(p)})' \in R^{p \times q} : \beta_{(i)} \in \mathcal{C}_{(i)} \text{ for all } i \in H\} \tag{1}$$

is the collection of all matrices with some rows, the rows specified by H , constrained by synchronized orderings. Here \mathcal{C}_+ and \mathcal{C}_- are convex cones in R^q , and \mathcal{C}_H is a convex cone in $R^{p \times q}$. When $H = \{i\}$, the notation \mathcal{C}_H is simplified as \mathcal{C}_i . Suppose that H is partitioned by H_+ and H_- such that $\mathcal{C}_{(i)} = \mathcal{C}_+$ for $i \in H_+$, and $\mathcal{C}_{(i)} = \mathcal{C}_-$ for $i \in H_-$. For $f, g \in R^p$ define $f \leq g$ if $f_i \leq g_i$ for all $i \in H_+$ and $f_i \geq g_i$ for all $i \in H_-$. Then \leq is a quasi order for vectors in R^p , and

$$\mathcal{C}_H = \{(\beta_{(1)}, \dots, \beta_{(p)})' \in R^{p \times q} : \beta_{(i)} \leq \beta_{(j)} \text{ for all } i \ll j\}.$$

Thus the restriction $\beta \in \mathcal{C}_H$ is a special case of multivariate isotonic restriction.

Restrictions $\beta \in \mathcal{C}_H$ are frequently encountered in statistics. For example, suppose a survey is conducted among the students in 4th and 5th grades in districts I and II. The means of the age, the household income, the height and the time for non-academic activities (NAA) in school are listed in matrix $\beta = (\beta_{ij})_{4 \times 4} = (\beta_{(1)}, \dots, \beta_{(4)})' \in R^{4 \times 4}$.

	4th grade District I	5th grade District I	4th grade District II	5th grade District II
Age	β_{11}	β_{12}	β_{13}	β_{14}
Household income	β_{21}	β_{22}	β_{23}	β_{24}
Height	β_{31}	β_{32}	β_{33}	β_{34}
Time for NAA	β_{41}	β_{42}	β_{43}	β_{44}

It is reasonable to assume that $\beta_{11} \leq \beta_{12}$, $\beta_{11} \leq \beta_{14}$, $\beta_{13} \leq \beta_{12}$ and $\beta_{13} \leq \beta_{14}$. Define \ll in $\Omega = \{1, 2, 3, 4\}$ by $1 \ll 2$, $1 \ll 4$, $3 \ll 2$ and $3 \ll 4$, and let \mathcal{C}_+ be $\{v \in R^4 : v_s \leq v_t \text{ for } s \ll t\}$. The assumption becomes $\beta_{(1)} \in \mathcal{C}_+$. Since the height and the age are positively correlated, the time for non-academic activities and the age are negatively correlated, and the household income and the age are uncorrelated, we may also impose the restrictions $\beta_{(3)} \in \mathcal{C}_+$ and $\beta_{(4)} \in \mathcal{C}_-$ while no restriction is imposed on $\beta_{(2)}$. Thus the restriction is expressed as $\beta \in \mathcal{C}_H$ where $H = \{1, 3, 4\}$.

One can change the signs of some of the variables to make synchronized orderings identical. But to avoid variables such as negative age, negative income, or negative time for non-academic activities, we kept two opposite cone structures for the restrictions in this paper, which did not create any difficulties in mathematical processing.

2.2. A multivariate model

Let the columns of data matrix $Y = (Y_1, \dots, Y_n) \in R^{p \times n}$ be p -variate independent normal vectors with a common variance-covariance matrix Σ . The distribution of Y is denoted by $Y \sim N_{p \times n}(E(Y), \Sigma)$. This notation, proposed by Hu [3], has many convenient properties. For example, with $A \in R^{q \times p}$ and $B \in R^{q \times n}$,

$$Y \sim N_{p \times n}(E(Y), \Sigma) \Rightarrow AY + B \sim N_{q \times n}(AE(Y) + B, A\Sigma A')$$

The structure of $E(Y)$ is dictated by the specifications of the model for data. For the example described in Section 2.1, suppose data matrix $Y \in R^{4 \times 1000}$ is obtained from the observations on the four variables, the age, the household income, the height and the non-academic time in school, of 1000 students. According to multivariate analysis of variance (MANOVA),

$$Y \sim N_{4 \times 1000}(\beta X', \Sigma), \quad X = \begin{pmatrix} 1_{n_1} & 0 & 0 & 0 \\ 0 & 1_{n_2} & 0 & 0 \\ 0 & 0 & 1_{n_3} & 0 \\ 0 & 0 & 0 & 1_{n_4} \end{pmatrix} \quad \text{and} \quad \sum_{i=1}^4 n_i = 1000.$$

In this paper we consider multivariate multiple linear regression

$$Y \sim N_{p \times n}(\beta X', \Sigma) \quad (2)$$

where the known matrix $X \in R^{n \times q}$ has full column rank. Let

$$m(Y) = YX(X'X)^{-1} \quad \text{and} \quad v(Y) = Y[I - X(X'X)^{-1}X']Y'. \quad (3)$$

It can be shown that $m(Y)$ is the maximum likelihood estimator (MLE) and an unbiased estimator (UE) for β ; $v(Y)/n$ is the MLE for Σ , and $v(Y)/(n - q)$ is an UE for Σ .

2.3. Projections

Let $V \in R^{p \times p}$ be a positive definite matrix ($V > 0$). For A and B in $R^{p \times q}$

$$\langle A, B \rangle_V = \text{tr}[(AX')'V^{-1}BX'] \quad (4)$$

defines an inner product of A and B . With respect to the induced norm $\|\cdot\|_V$, \mathcal{C}_H in (1) is closed. Thus for $A \in R^{p \times q}$ there exists a unique matrix in \mathcal{C}_H that minimizes the distance to A over all matrices in \mathcal{C}_H . This matrix, called the projection of A onto \mathcal{C}_H , is denoted by $P_V(A|\mathcal{C}_H)$. It is known, see Lemma 1.1 in [11], that $A^* = P_V(A|\mathcal{C}_H)$ if and only if

$$A^* \in \mathcal{C}_H, \quad \langle A - A^*, A^* \rangle_V = 0 \quad \text{and} \quad \langle A - A^*, B \rangle_V \leq 0 \quad \text{for all } B \in \mathcal{C}_H. \quad (5)$$

This projection plays an important role in restricted statistical inference. For example, if Σ in (2) is known, then $P_\Sigma(m(Y)|\mathcal{C}_H)$ is the restricted MLE of β under $\beta \in \mathcal{C}_H$. We say that $\beta = (\beta_1, \dots, \beta_q) \in R^{p \times q}$ is homogeneous if $\beta_1 = \dots = \beta_q$, i.e., $\beta \in \mathcal{L}$ where

$$\mathcal{L} = \{x1'_q : x \in R^p\}.$$

This \mathcal{L} is a linear space inside \mathcal{C}_H . The restricted MLE of β under $\beta \in \mathcal{L}$, if Σ is known, is $P_\Sigma(m(Y)|\mathcal{L})$.

2.4. An ad hoc test

Model (2) under $\beta \in \mathcal{C}_H$ generalized the settings in [7] in three aspects. MANOVA is generalized to be multivariate regression; identical simple ordering is generalized to be synchronized quasi ordering; restrictions on all rows are generalized to be that on some rows. It can be shown that with model (2) under $\beta \in \mathcal{C}_H$ for the test on $H_0 : \beta \in \mathcal{L}$,

$$\|P_\Sigma(m(Y)|\mathcal{L}) - P_\Sigma(m(Y)|\mathcal{C}_H)\|_\Sigma^2$$

is an LRT statistic and the null hypothesis is rejected for large values of the statistic if Σ is known. When Σ is unknown, conventionally we replace Σ in the LRT statistic with its estimator to get an ad hoc test statistic. But the test statistics obtained by replacing Σ with its MLE and UE are proportional to that by replacing Σ with $v(Y)$ in (3). Thus

$$T_H(Y) = \|P_{v(Y)}(m(Y)|\mathcal{L}) - P_{v(Y)}(m(Y)|\mathcal{C}_H)\|_{v(Y)}^2 \quad (6)$$

is proposed as a test statistic. The notation $T_H(Y)$ is simplified as $T_i(Y)$ when $H = \{i\}$. Let D be a non-singular $p \times p$ matrix. One can show that with transformed data DY

$$T_H(DY) = \|P_{v(Y)}(m(Y)|\mathcal{L}) - P_{v(Y)}(m(Y)|D^{-1}\mathcal{C}_H)\|_{v(Y)}^2. \quad (7)$$

3. On the distributions of $T_H(Y)$

Denote the distribution of $T_H(Y)$, when $Y \sim N_{p \times n}(\beta X', \Sigma)$, by $T_H(\beta, \Sigma)$. On this distribution we list two lemmas and place the proofs in the appendices.

Lemma 1. *The following statements are true.*

- (a) If $\beta \in \mathcal{L}$, then $T_H(\beta, \Sigma) = T_H(0, \Sigma)$.
- (b) If $\emptyset \neq H_* \subset H$, then $T_H(\beta, \Sigma) \leq T_{H_*}(\beta, \Sigma)$ stochastically.
- (c) $T_i(0, \Sigma) = T_i(0, I)$ for all $\Sigma > 0$ and all $i = 1, \dots, p$.

Proof. See Appendix A. \square

Lemma 2. For $i_0 \in H$ there exists $\{\Sigma^{[N]} > 0 : N\}$ such that $T_H(0, \Sigma^{[N]}) \rightarrow T_{i_0}(0, I)$ in distributions as $N \rightarrow \infty$.

Proof. See Appendix B. \square

Based on the results of two lemmas we establish two theorems directly related to the properties of the tests.

Theorem 1. For all non-empty H , all $t \geq 0$, and all $i = 1, \dots, p$,

$$\sup\{P(T_H(\beta, \Sigma) > t) : \beta \in \mathcal{L}, \Sigma > 0\} = P(T_i(0, I) > t).$$

Proof. If the theorem is true for $i \in H$, then with $H = \{1, \dots, p\}$ we see that $T_i(0, I) = T_j(0, I)$ for all i and j . Consequently, the theorem holds for all i . Assume that $i \in H$. By (a)–(c) of Lemma 1,

$$\begin{aligned} \sup\{P(T_H(\beta, \Sigma) > t) : \beta \in \mathcal{L}, \Sigma > 0\} &= \sup\{P(T_H(0, \Sigma) > t) : \Sigma > 0\} \\ &\leq \sup\{P(T_i(0, \Sigma) > t) : \Sigma > 0\} \\ &= P(T_i(0, I) > t). \end{aligned}$$

Let $\{\Sigma^{[N]} > 0 : N\}$ be produced in Lemma 2 for $i \in H$. Then

$$\begin{aligned} \sup\{P(T_H(\beta, \Sigma) > t) : \beta \in \mathcal{L}, \Sigma > 0\} &= \sup\{P(T_H(0, \Sigma) > t) : \Sigma > 0\} \\ &\geq P(T_H(0, \Sigma^{[N]}) > t) \\ &\rightarrow P(T_i(0, I) > t). \end{aligned}$$

Conclusion $\sup\{P(T_H(\beta, \Sigma) > t) : \beta \in \mathcal{L}, \Sigma > 0\} = P(T_i(0, I) > t)$ follows. \square

Theorem 2. Let $T_*(Y)$ be a statistic with the distribution denoted by $T_*(\beta, \Sigma)$ when $Y \sim N_{p \times n}(\beta X', \Sigma)$. If $T_H(\beta, \Sigma) \leq T_*(\beta, \Sigma) \leq T_{i_0}(\beta, \Sigma)$ stochastically for some $i_0 = 1, \dots, p$, then for all $t \geq 0$ and all $i = 1, \dots, p$,

$$\sup\{P(T_*(\beta, \Sigma) > t) : \beta \in \mathcal{L}, \Sigma > 0\} = P(T_i(0, I) > t).$$

Proof. Under the condition of this theorem,

$$\begin{aligned} \sup\{P(T_H(\beta, \Sigma) > t) : \beta \in \mathcal{L}, \Sigma > 0\} &\leq \sup\{P(T_*(\beta, \Sigma) > t) : \beta \in \mathcal{L}, \Sigma > 0\} \\ &\leq \sup\{P(T_{i_0}(\beta, \Sigma) > t) : \beta \in \mathcal{L}, \Sigma > 0\}. \end{aligned}$$

But by Theorem 1, $\sup\{P(T_H(\beta, \Sigma) > t) : \beta \in \mathcal{L}, \Sigma > 0\} = P(T_i(0, I) > t)$ and $\sup\{P(T_{i_0}(\beta, \Sigma) > t) : \beta \in \mathcal{L}, \Sigma > 0\} = P(T_{i_0}(0, I) > t)$. Therefore, $\sup\{P(T_*(\beta, \Sigma) > t) : \beta \in \mathcal{L}, \Sigma > 0\} = P(T_i(0, I) > t)$. \square

4. Main results on the tests

The tests on $H_0 : \beta \in \mathcal{L}$ versus $H_1 : \beta \in \mathcal{C}_H$ using the proposed test statistics $T_H(Y)$ form a family of ad hoc tests. The members of the family are indexed by H .

4.1. α -level critical values

The test has α -level critical value t_α determined by

$$\alpha = \sup\{P(T_H(\beta, \Sigma) > t_\alpha) : \beta \in \mathcal{L}, \Sigma > 0\}.$$

By Theorem 1, this equation is equivalent to

$$\alpha = P(T_i(0, I) > t_\alpha)$$

which is free of H . Therefore, all tests in the family share the same α -level critical values. The complexity of the computation for the α -level critical values is greatly reduced by using the second equation since it does not involve unknown $\beta \in \mathcal{L}$ and unknown $\Sigma > 0$.

4.2. p -values

Let t_{ob} be the observed value of the test statistic. The observed significance level, or p -value, is given by

$$p\text{-value} = \sup\{P(T_H(\beta, \Sigma) > t_{ob}) : \beta \in \mathcal{L}, \Sigma > 0\}.$$

By Theorem 1, this equation is equivalent to

$$p\text{-value} = P(T_i(0, I) > t_{ob})$$

which is free of H . Therefore, all tests in the family follow the same distribution $T_i(0, I)$ for computing observed significance levels. The fact that this distribution does not involve unknown parameters allows us to estimate the p -values using the Monte Carlo method by simulating the observations from $T_i(0, I)$.

4.3. More powerful tests

Let $T_*(Y)$ be a statistic satisfying the conditions in [Theorem 2](#). Then the test using $T_*(Y)$ as a test statistic has the same α -level critical values as that for the members of the ad hoc test family since

$$\alpha = \sup\{P(T_*(\beta, \Sigma) > t_\alpha) : \beta \in \mathcal{L}, \Sigma > \mathbf{0}\},$$

by [Theorem 2](#), is equivalent to

$$\alpha = P(T_i(\mathbf{0}, I) > t_\alpha).$$

This test also follows the distribution $T_i(\mathbf{0}, I)$ for computing p -values since

$$p\text{-value} = \sup\{P(T_*(\beta, \Sigma) > t_{ob}) : \beta \in \mathcal{L}, \Sigma > \mathbf{0}\},$$

by [Theorem 2](#), is equivalent to

$$p\text{-value} = P(T_i(\mathbf{0}, I) > t_{ob}).$$

Moreover, the test using $T_*(Y)$ is more powerful than that of using $T_H(Y)$ since

$$P(T_H(\beta, \Sigma) > t_\alpha) \leq P(T_*(\beta, \Sigma) > t_\alpha)$$

for all $\beta \in \mathcal{C}_H$ and all $\Sigma > \mathbf{0}$ by the condition of [Theorem 2](#).

We now give two examples of such $T_*(Y)$. The first one is $T_{i_0}(Y)$ where $i_0 \in H$. The second one is $\min\{T_i(Y) : i \in H\}$. Both statistics satisfy the conditions for $T_*(Y)$ in [Theorem 2](#), and hence both produce the tests more powerful than that of using $T_H(Y)$. It is interesting to notice that the first statistic only takes one restricted row into the consideration while the second one does consider every rows under the restriction. Therefore the second statistic, by our intuition, is more appropriate. The power of this test, however, is less than or equal to that of the test using the first test statistic.

Acknowledgments

The authors wish to thank four anonymous reviewers for their comments and suggestions that improved this paper.

Appendix A. Proof of Lemma 1

Proof. (a): From (3), $m(Y - \beta X') = m(Y) - \beta$ and $v(Y - \beta X') = v(Y)$. By (5), $\beta \in \mathcal{L}$ implies that $P_{v(Y)}(m(Y) - \beta | \mathcal{L}) = P_{v(Y)}(m(Y) | \mathcal{L}) - \beta$ and $P_{v(Y)}(m(Y) - \beta | \mathcal{C}_H) = P_{v(Y)}(m(Y) | \mathcal{C}_H) - \beta$. So,

$$\begin{aligned} T_H(Y - \beta X') &= \|P_{v(Y - \beta X')} (m(Y - \beta X') | \mathcal{L}) - P_{v(Y - \beta X')} (m(Y - \beta X') | \mathcal{C}_H)\|_{v(Y - \beta X')}^2 \\ &= \|P_{v(Y)} (m(Y) - \beta | \mathcal{L}) - P_{v(Y)} (m(Y) - \beta | \mathcal{C}_H)\|_{v(Y)}^2 \\ &= \|P_{v(Y)} (m(Y) | \mathcal{L}) - P_{v(Y)} (m(Y) | \mathcal{C}_H)\|_{v(Y)}^2 \\ &= T_H(Y). \end{aligned}$$

But $T_H(Y - \beta X') \sim T_H(\mathbf{0}, \Sigma)$ and $T_H(Y) \sim T_H(\beta, \Sigma)$, (a) follows.

(b): Note that $\mathcal{L} \subset \mathcal{C}_H \subset \mathcal{C}_{H^*}$ and \mathcal{L} is a linear space. Let S be the inner product $\langle P_{v(Y)}(m(Y) | \mathcal{L}) - P_{v(Y)}(m(Y) | \mathcal{C}_H), P_{v(Y)}(m(Y) | \mathcal{C}_H) - P_{v(Y)}(m(Y) | \mathcal{C}_{H^*}) \rangle_{v(Y)}$. By (5),

$$\begin{aligned} S &= \langle P_{v(Y)}(m(Y) | \mathcal{L}) - P_{v(Y)}(m(Y) | \mathcal{C}_H), P_{v(Y)}(m(Y) | \mathcal{C}_H) - m(Y) \rangle_{v(Y)} \\ &\quad + \langle P_{v(Y)}(m(Y) | \mathcal{L}), m(Y) - P_{v(Y)}(m(Y) | \mathcal{C}_{H^*}) \rangle_{v(Y)} - \langle P_{v(Y)}(m(Y) | \mathcal{C}_H), m(Y) - P_{v(Y)}(m(Y) | \mathcal{C}_{H^*}) \rangle_{v(Y)} \\ &= \mathbf{0} + \mathbf{0} - \langle P_{v(Y)}(m(Y) | \mathcal{C}_H), m(Y) - P_{v(Y)}(m(Y) | \mathcal{C}_{H^*}) \rangle_{v(Y)} \\ &\geq \mathbf{0}. \end{aligned}$$

Then

$$\begin{aligned} T_{H^*}(Y) &= \|P_{v(Y)}(m(Y) | \mathcal{L}) - P_{v(Y)}(m(Y) | \mathcal{C}_{H^*})\|_{v(Y)}^2 \\ &= T_H(Y) + \|P_{v(Y)}(m(Y) | \mathcal{C}_H) - P_{v(Y)}(m(Y) | \mathcal{C}_{H^*})\|_{v(Y)}^2 + 2S \\ &\geq T_H(Y). \end{aligned}$$

But $T_H(Y) \sim T_H(\beta, \Sigma)$ and $T_{H^*}(Y) \sim T_{H^*}(\beta, \Sigma)$. (b) follows.

(c): For $Y \sim N_{p \times n}(0X', \Sigma)$, let G be the elementary matrix obtained by interchanging the i th row and the p th row of I_p , and $(G\Sigma G)^{-1/2} = QR$ be the QR decomposition such that Q is orthogonal and R is upper-triangular with all diagonal elements being positive numbers. Define $D = GRG$. For $A \in R^{p \times q}$, the i th row of $D^{-1}A = (GRG)^{-1}A = GR^{-1}GA$ is the i th row of A divided by the last diagonal element of R . Therefore $A \in \mathcal{C}_i$ if and only if $D^{-1}A \in \mathcal{C}_i$. By (7), $T_i(DY) = T_i(Y)$ where $T_i(Y) \sim T_i(\mathbf{0}, \Sigma)$

and $T_i(DY) \sim T_i(0, D\Sigma D')$ with

$$\begin{aligned} D\Sigma D' &= (GR)(G\Sigma G)(R'G) \\ &= GQ'(QR)(G\Sigma G)(QR)'QG \\ &= GQ'(G\Sigma G)^{-1/2}(G\Sigma G)(G\Sigma G)^{-1/2}QG \\ &= I. \end{aligned}$$

(c) follows. \square

Appendix B. Proof of Lemma 2

Proof. Without loss of generality assume that $\mathcal{C}_{(i_0)} = \mathcal{C}_+$. Let $u = (u_1, \dots, u_p)'$ with $u_i = \begin{cases} 1, & i \neq i_0 \text{ and } \mathcal{C}_{(i)} = \mathcal{C}_+ \\ 0, & i = i_0 \\ -1, & i \neq i_0 \text{ and } \mathcal{C}_{(i)} = \mathcal{C}_- \end{cases}$ and e_{i_0} be the i_0 th column of I_p . Define $D^{[N]} = I_p - N \cdot ue'_{i_0}$. Clearly, $D^{[0]} = I_p$, $D^{[m]}D^{[N]} = D^{[m+N]}$, and $(D^{[N]})^{-1} = D^{[-N]}$. With $\Sigma^{[N]} = D^{[-N]}(D^{[-N]})'$ and $Z \sim N_{p \times n}(0, I)$, $T_H(D^{[-N]}Z) \sim T_H(0, \Sigma^{[N]})$ and $T_{i_0}(Z) \sim T_{i_0}(0, I)$. It suffices to show

$$T_H(D^{[-N]}Z) \rightarrow T_{i_0}(Z) \quad \text{with probability 1.} \tag{B.1}$$

By (7), for (B.1) we only need

$$P_{v(Y)}(m(Y)|D^{[N]}\mathcal{C}_H) \rightarrow P_{v(Y)}(m(Y)|\mathcal{C}_{i_0}) \quad \text{with probability 1.} \tag{B.2}$$

But the sequence $\{P_{v(Y)}(m(Y)|D^{[N]}\mathcal{C}_H) : N\}$ lies in a bounded set since by (5)

$$\begin{aligned} \|P_{v(Y)}(m(Y)|D^{[N]}\mathcal{C}_H)\|_{v(Y)}^2 &= \|m(Y)\|_{v(Y)}^2 - \|m(Y) - P_{v(Y)}(m(Y)|D^{[N]}\mathcal{C}_H)\|_{v(Y)}^2 \\ &\leq \|m(Y)\|_{v(Y)}^2. \end{aligned}$$

Thus (B.2) is equivalent to that every convergent subsequence converges to $P_{v(Y)}(m(Y)|\mathcal{C}_{i_0})$, i.e.,

$$P_{v(Y)}(m(Y)|D^{[N_k]}\mathcal{C}_H) \rightarrow A^* \Rightarrow A^* = P_{v(Y)}(m(Y)|\mathcal{C}_{i_0}). \tag{B.3}$$

Note that $P_{v(Y)}(m(Y)|D^{[N_k]}\mathcal{C}_H) \in \mathcal{C}_{i_0}$ since $D^{[N]}A$ does not change the i_0 th row of A , and that $\langle m(Y) - P_{v(Y)}(m(Y)|D^{[N_k]}\mathcal{C}_H), P_{v(Y)}(m(Y)|D^{[N_k]}\mathcal{C}_H) \rangle_{v(Y)} = 0$. By letting $k \rightarrow \infty$ we have $A^* \in \mathcal{C}_{i_0}$ and $\langle m(Y) - A^*, A^* \rangle_{v(Y)} = 0$. By (5) for (B.3) we need to show that

$$\langle m(Y) - A^*, B \rangle_{v(Y)} \leq 0 \quad \text{for all } B \in \mathcal{C}_{i_0}. \tag{B.4}$$

For $B = (B_{(1)}, \dots, B_{(p)})' \in \mathcal{C}_{i_0}$ let $B^{[m]} = B + e_{i_0}v'/m$ where $v = (v_1, \dots, v_q)'$ is an interior point of $\mathcal{C}_{(i_0)}$, i.e., $v \in \mathcal{C}_{(i_0)}$ and $v_i \neq v_j$ for $i \neq j$ and $i \ll j$. Then $B^{[m]} \rightarrow B$ as $m \rightarrow \infty$. If

$$\exists \{N_{k_m}\} \text{ such that } B^{[m]} \in D^{[N_{k_m}]\mathcal{C}_H} \text{ for all } m = 1, 2, \dots, \tag{B.5}$$

then by (5)

$$\langle m(Y) - P_{v(Y)}(m(Y)|D^{[N_{k_m}]\mathcal{C}_H}), B^{[m]} \rangle_{v(Y)} \leq 0.$$

Let $m \rightarrow \infty$. We obtain (B.4). To establish the lemma we prove (B.5). For each fixed positive integer m write $B^{[m]} = D^{[N]} \cdot (D^{[-N]}B^{[m]})$ where

$$D^{[-N]}B^{[m]} = (I_p + Nue'_{i_0})(B + e_{i_0}v'/m) = B + e_{i_0}v'/m + Nu(B_{(i_0)} + v/m)'$$

denoted by $(A_{(1)}, \dots, A_{(p)})'$ has $A_{(i_0)} = B_{(i_0)} + v/m \in \mathcal{C}_{(i_0)}$. For $i \neq i_0$ and $\mathcal{C}_{(i)} = \mathcal{C}_+$, $A_{(i)} = B_{(i)} + N(B_{(i_0)} + v/m)$ where $N(B_{(i_0)} + v/m)$ is an interior point in $\mathcal{C}_+ = \mathcal{C}_{(i)}$ that dominates $B_{(i)}$ when N is sufficiently large. Thus $A_{(i)} \in \mathcal{C}_{(i)}$ when N is sufficiently large. For $i \neq i_0$ and $\mathcal{C}_{(i)} = \mathcal{C}_-$, $A_{(i)} = B_{(i)} + N(-B_{(i_0)} - v/m)$ where $N(-B_{(i_0)} - v/m)$ is an interior point in $\mathcal{C}_- = \mathcal{C}_{(i)}$ that dominates $B_{(i)}$ when N is sufficiently large. Thus $A_{(i)} \in \mathcal{C}_{(i)}$ when N is sufficiently large. Therefore there exists N_{k_m} such that $D^{[-N_{k_m}]}B^{[m]} \in \mathcal{C}_H$. So $B^{[m]} = D^{[N_{k_m}]}(D^{[-N_{k_m]}B^{[m]}) \in D^{[N_{k_m}]\mathcal{C}_H}$. (B.5) is established. \square

References

[1] R.E. Barlow, E. Bartholomew, J.M. Bremner, H.H. Brunk, Statistical Inference Under Order Restrictions, John Wiley and Sons, New York, 1972.
 [2] D.J. Bartholomew, A test of homogeneity for ordered alternatives, Biometrika 46 (1959) 36–48.
 [3] X. Hu, Multivariate analysis without vec and \otimes , Amer. Statist. 61 (2007) 64–66.
 [4] X. Hu, p -values of a test on homogeneous means in a multivariate isotonic regression, Statist. Probab. Lett. 79 (2009) 2005–2011.
 [5] H. Mukerjee, R. Tu, Order-restricted inferences in linear regression, J. Amer. Statist. Assoc. 90 (1995) 717–728.
 [6] T. Robertson, F.T. Wright, R.L. Dykstra, Order Restricted Statistical Inferences, Wiley, New York, 1988.

- [7] S. Sasabuchi, More powerful tests for homogeneity of multivariate normal mean vectors under an order restriction, *Sankhyā: The Indian Journal of Statistics* 69 (2007) 700–716.
- [8] S. Sasabuchi, K. Tanaka, T. Tsukamoto, Testing homogeneity of multivariate normal mean vectors under an order restriction when the covariance matrices are common but unknown, *Ann. Statist.* 31 (2003) 1517–1536.
- [9] M.J. Silvapulle, A Hotelling's T^2 -type statistic for testing against one-sided hypotheses, *J. Multivariate Anal.* 55 (1995) 312–319.
- [10] M.J. Silvapulle, P.K. Sen, *Constrained Statistical Inference, Inequality, Order, and Shape Restrictions*, Wiley, New York, 2005.
- [11] E.H. Zarantonello, Projection on convex sets in Hilbert space and spectral theory, in: E.H. Zarantonello (Ed.), *Contributions to Nonlinear Functional Analysis*, Academic Press, New York, 1971, pp. 237–424.