

RESEARCH ARTICLE

On the powers of tests for homogeneity of regression coefficient vectors under synchronised order restrictions

Shoichi Sasabuchi¹ and D.D. Sarath Kulatunga^{2*}

¹Department of Environment Design, Faculty of Design, Kyushu University, 4-9-1, Shiobaru, Minami-ku, Fukuoka-shi, 815-8540, Japan.

²Department of Mathematics, Faculty of Science, University of Kelaniya, Kelaniya.

Revised: 27 July 2015; Accepted: 26 August 2015

Abstract: We consider a multivariate multiple linear regression model and study the problem of testing homogeneity of regression coefficient vectors under synchronised order restrictions when the covariance matrices are common but unknown. Synchronised order restrictions are the generalisation of the multivariate isotonic order restrictions. Synchronised order restricted test could be applied to a situation where the values of some parameters increase, those of some other parameters decrease, and those of the rest of the parameters have no restriction, simultaneously. For this problem, some test statistics were proposed and some inequalities among their powers were obtained in the past. This showed that the proposed test statistics may equally be good in terms of their powers. In the present paper, we mathematically prove that the strict inequalities hold among the powers of the test statistics. Thus we attain an exact comparison among the powers of the test statistics indicating more accurate and stronger results than those obtained in the past.

Keywords: Homogeneity of regression coefficient vectors, multivariate multiple linear regression model, statistical testing hypotheses, strictly more powerful test, synchronised order restrictions, unknown covariance matrix.

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 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 them have been

summarised by Barlow *et al.* (1972), Robertson *et al.* (1988), and Silvapulle and Sen (2005).

Sasabuchi *et al.* (1983) considered the problem of testing the homogeneity of several p -variate normal mean vectors against the alternative hypothesis determined by a multivariate isotonic order restriction. This is a multivariate extension of Bartholomew's (1959) problem. When the covariance matrices are known, this problem has been studied by Sasabuchi *et al.* (1983; 1998; 2003a), Kulatunga and Sasabuchi (1984), Nomakuchi and Shi (1988) and some others.

When the covariance matrices are common but unknown, Sasabuchi *et al.* (2003b) studied the above problem under component-wise simple ordering. Suppose that $\mu_1, \mu_2, \dots, \mu_k$ are mean vectors of several p -variate normal distributions with common but unknown covariance matrices. Sasabuchi *et al.* (2003b) considered the problem of testing the null hypothesis $H_0: \mu_1 = \mu_2 = \dots = \mu_k$ against the alternative hypothesis $H_1: \mu_1 \leq \mu_2 \leq \dots \leq \mu_k$ where " $\mu_i \leq \mu_j$ " means that all the elements of $\mu_j - \mu_i$ are non-negative, proposed a test statistic, and studied its null distribution. Sasabuchi (2007) considered the same problem and provided some tests, which are more powerful than the test proposed in Sasabuchi *et al.* (2003b). Hu (2009) generalised the model considered in Sasabuchi *et al.* (2003b) from component-wise simple ordering to a general vector quasi ordering.

These problems may arise in situations where the values of several parameters increase simultaneously.

* Corresponding author (kul@kln.ac.lk)

For example, if we wish to know whether both the average height and the average weight of the children of an area increase simultaneously year by year, then we could apply our test to the bivariate datasets obtained by random sampling from the population.

Hu and Banerjee (2012) considered a multivariate multiple linear regression model and studied the problem of testing homogeneity of regression coefficient vectors under synchronised order restrictions when the covariance matrices are common but unknown. For this problem, Hu and Banerjee (2012) proposed some test statistics, studied their distributions, and obtained some inequalities among their powers.

Synchronised order restrictions are the generalisation of the multivariate isotonic order restrictions. Synchronised order restricted test could be applied to a situation where the values of some parameters increase, those of some other parameters decrease, and those of the rest of the parameters have no restriction, simultaneously. An example of such situation has been given in section 2.1 of Hu and Banerjee (2012). Besides, they (Hu & Banerjee, 2012) have considered general partial orderings including the simple ordering considered by Sasabuchi *et al.* (2003b). Furthermore, multivariate multiple linear regression model is also a generalisation of several p -variate normal distributions model. Thus the problem of Hu and Banerjee (2012) is considered to be the generalisation of that of Sasabuchi *et al.* (2003b) in three senses.

The main objective of this paper is to prove mathematically that the strict inequalities hold among the powers of the tests given in Hu and Banerjee (2012). In order to prove them some careful geometrical discussion is needed and they are elucidated in the text. The exact results among the powers of the test statistics obtained in this paper are stronger than those obtained in Hu and Banerjee (2012).

In the next section, we review the problem and summarise Hu and Banerjee's (2012) findings. The main theorem and its proof are given later. Proofs of some lemmas are given in the Appendix.

PRELIMINARIES

In this section, we summarise Hu and Banerjee's (2012) findings.

Throughout this paper, column vectors are considered as vectors. In the usual notation, x' and A' denote the transpose of vector x and the transpose of matrix A , respectively.

Let R^q denote the q -dimensional real Euclidean space. Let $R^{p \times q}$ denote the set of all $p \times q$ real matrices, which is essentially equivalent to the pq -dimensional real Euclidean space.

Let \ll be a partial order in $\Omega = \{1, 2, \dots, q\}$. Let us define $C_+ = \{v = (v_1, v_2, \dots, v_q)' \in R^q : v_s \leq v_t \text{ for } s \ll t\}$, the collection of all isotonic q -dimensional column vectors, and $C_- = \{v = (v_1, v_2, \dots, v_q)' \in R^q : v_s \geq v_t \text{ for } s \ll t\}$, the collection of all antitonic q -dimensional column vectors. Let $C_{(i)}$ be either C_+ or C_- , $i = 1, 2, \dots, p$, and H be a non-empty subset of $\{1, 2, \dots, p\}$.

Let $\beta = (\beta_{(1)}, \beta_{(2)}, \dots, \beta_{(p)})' = (\beta_1, \beta_2, \dots, \beta_q)$ be a $p \times q$ real matrix. We say that $\beta = (\beta_{(1)}, \beta_{(2)}, \dots, \beta_{(p)})'$ is constrained by synchronised orderings (with respect to \ll and H) if and only if $\beta_{(i)} \in C_{(i)}$ for all $i \in H$.

For example, suppose that \ll is the simple order : $1 \ll 2 \ll \dots \ll q$, $p = 3$, $H = \{1, 2\}$, $C_{(1)} = C_+$, $C_{(2)} = C_-$. Then the synchronised orderings mean that $\beta_{11} \leq \beta_{12} \leq \dots \leq \beta_{1q}$ (isotonic), $\beta_{21} \geq \beta_{22} \geq \dots \geq \beta_{2q}$ (antitonic), and $\beta_{31}, \beta_{32}, \dots, \beta_{3q}$ have no restriction, where β_{ij} is the (i, j) -th element of β , $i = 1, 2, 3$; $j = 1, 2, \dots, q$.

Specially when $H = \{1, 2, \dots, p\}$ and $C_{(i)} = C_+$, $i = 1, 2, \dots, p$, the synchronised orderings coincide with the multivariate isotonic orderings defined in Sasabuchi *et al.* (1983). Thus the synchronised order restrictions are the generalisation of the multivariate isotonic order restrictions.

Let us consider the following multivariate multiple linear regression model. Let Y_1, Y_2, \dots, Y_n be independent p -variate normal random vectors with a common covariance matrix Σ . Assume that Σ is unknown. Put $Y = (Y_1, Y_2, \dots, Y_n) : p \times n$ and $E(Y) = \beta X'$, where β is a $p \times q$ unknown regression coefficient matrix and X is an $n \times q$ known matrix with $\text{rank}(X) = q < n$.

Define

$$\mathcal{L} = \{ \beta = (\beta_1, \beta_2, \dots, \beta_q) \in R^{p \times q} : \beta_1 = \beta_2 = \dots = \beta_q \},$$

$$C_H = \{ \beta = (\beta_{(1)}, \beta_{(2)}, \dots, \beta_{(p)})' \in R^{p \times q} : \beta_{(i)} \in C_{(i)} \text{ for all } i \in H \}.$$

Hu and Banerjee (2012) studied the following problem of testing the null hypothesis

$$H_0 : \beta \in \mathcal{L}, \text{ i.e., } \beta_1 = \beta_2 = \dots = \beta_q$$

against the alternative hypothesis

$H_1 : \beta \in C_H$, i.e., $\beta = (\beta_{(1)}, \beta_{(2)}, \dots, \beta_{(p)})'$ is constrained by synchronised orderings (with respect to \ll and H).

This problem is a generalisation of that considered by Sasabuchi *et al.* (2003b). As stated in the introduction, the problem considered by Sasabuchi *et al.* (2003b) may arise in the situation where the values of several parameters increase simultaneously. While the above problem could be applied to the situation where the values of some parameters increase, those of some other parameters decrease, and those of the rest of the parameters have no restriction, simultaneously. An example of such situation was given in section 2.1 of Hu and Banerjee (2012). Besides, now we consider general partial order: \ll including the simple order considered by Sasabuchi *et al.* (2003b). Furthermore, multivariate multiple linear regression model is also a generalisation of several p -variate normal distributions model. Thus the above problem is considered to be the generalisation of that of Sasabuchi *et al.* (2003b) in three senses.

Now we define the inner product, norm and projection in $R^{p \times q}$ as follows.

Let V be a $p \times p$ positive definite symmetric matrix. For A and B in $R^{p \times q}$,

$$\langle A, B \rangle_V = \text{tr}(XA'V^{-1}BX') = \text{trace of } XA'V^{-1}BX'$$

and

$$\|A\|_V = \left[\langle A, A \rangle_V \right]^{\frac{1}{2}},$$

define an inner product of A and B , and a norm of A , respectively.

Let C be a closed convex cone in $R^{p \times q}$. Then, for any $A \in R^{p \times q}$, there exists a unique matrix B in C , which minimises $\|B - A\|_V$ under the restriction that $B \in C$. This matrix is called the orthogonal projection of A onto C with respect to the norm $\|\cdot\|_V$, and denoted by $P_V(A | C)$.

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

Hu and Banerjee (2012) proposed the following test statistic:

$$T_H(Y) = \left\| P_{v(Y)}(m(Y) | \mathcal{L}) - P_{v(Y)}(m(Y) | C_H) \right\|_{v(Y)}^2.$$

Note that \mathcal{L} and C_H are closed convex cones in $R^{p \times q}$, $m(Y) \in R^{p \times q}$ and $v(Y)$ is a $p \times p$ positive definite symmetric matrix (with probability one), thus the above $T_H(Y)$ is well-defined.

The notation $T_H(Y)$ is simplified as $T_i(Y)$ when $H = \{i\}$.

$$\text{Put } T_{MIN}(Y) = \min\{T_i(Y) : i \in H\}.$$

We consider the test statistics $T_H(Y)$, $T_{MIN}(Y)$ and $T_{i_0}(Y)$, where $i_0 \in H$.

Note that, in general, when we use T as a test statistic, we reject the null hypothesis H_0 if $T > c$, where c is a positive constant depending on the significance level. Thus it is important to evaluate the upper tail probability of each test statistic.

Hu and Banerjee (2012) derived the following results.

(a) For any $i_0 \in H$ and all $t \geq 0$,

$$\begin{aligned} \sup\{P_{\beta, \Sigma}(T_H(Y) > t) : \beta \in \mathcal{L}, \Sigma > 0\} \\ &= \sup\{P_{\beta, \Sigma}(T_{MIN}(Y) > t) : \beta \in \mathcal{L}, \Sigma > 0\} \\ &= \sup\{P_{\beta, \Sigma}(T_{i_0}(Y) > t) : \beta \in \mathcal{L}, \Sigma > 0\} \\ &= P_{0, I_p}(T_{i_0}(Y) > t). \end{aligned}$$

(b) For any $i_0 \in H$, all $t \geq 0$, all β , and all Σ ,

$$\begin{aligned} P_{\beta, \Sigma}(T_H(Y) > t) &\leq P_{\beta, \Sigma}(T_{MIN}(Y) > t) \\ &\leq P_{\beta, \Sigma}(T_{i_0}(Y) > t). \end{aligned}$$

Here $P_{\beta, \Sigma}$ denotes the probability measure corresponding to the parameters β and Σ , " $\Sigma > 0$ " means that Σ is any $p \times p$ positive definite symmetric matrix, and I_p is the $p \times p$ identity matrix. Note that we have changed the notation given in Hu and Banerjee (2012).

The reason for the above change is as follows. In the present paper, we need to emphasise that the test statistics are functions of Y . Besides, it is important to note that $T_H(Y) \leq T_{MIN}(Y) \leq T_{i_0}(Y)$ for each value of Y ,

where $i_0 \in H$, as given in the next section. This is a fact stronger than the one that $T_H(\beta, \Sigma) \leq T_{MIN}(\beta, \Sigma) \leq T_{i_0}(\beta, \Sigma)$ stochastically, according to the notation in Hu and Banerjee (2012), which is equivalent to the one that $P_{\beta, \Sigma}(T_H(Y) > t) \leq P_{\beta, \Sigma}(T_{MIN}(Y) > t) \leq P_{\beta, \Sigma}(T_{i_0}(Y) > t)$ for any real number t .

By (a), we can use the same critical values for $T_H(Y)$, $T_{MIN}(Y)$ and $T_{i_0}(Y)$ for a given significance level.

From (b), we can see that the power of $T_H(Y)$ is less than or equal to that of $T_{MIN}(Y)$, and the power of $T_{MIN}(Y)$ is less than or equal to that of $T_{i_0}(Y)$, as stated in Hu and Banerjee (2012).

In the next section, we show that some results, which are more accurate and stronger than (b), hold true.

THE MAIN THEOREM

The main objective of this paper is to prove that the strict inequalities hold among the powers of the three test statistics $T_H(Y)$, $T_{MIN}(Y)$ and $T_{i_0}(Y)$ given in the previous section. That is equivalent to proving the result in the following theorem:

Theorem 1.

If the number of elements of H is larger than one, then, for any $i_0 \in H$, all $t > 0$, all β , and all Σ ,

$$P_{\beta, \Sigma}(T_H(Y) > t) < P_{\beta, \Sigma}(T_{MIN}(Y) > t) < P_{\beta, \Sigma}(T_{i_0}(Y) > t),$$

that is, the power of $T_{MIN}(Y)$ is strictly higher than that of $T_H(Y)$, and the power of $T_{i_0}(Y)$ is strictly higher than that of $T_{MIN}(Y)$.

Note that if the number of elements of H is one, then $T_H(Y) = T_{MIN}(Y) = T_{i_0}(Y)$. Thus the assumption of the theorem is quite natural.

The results of the theorem may seem to be intuitively obvious. But some careful geometrical discussions are needed to prove that those strict inequalities hold for all $t > 0$.

In order to prove the theorem, we state three lemmas. The proofs of these lemmas need some geometrical discussions and they are given in the Appendix.

Lemma 1.

$$T_H(Y) \leq T_{MIN}(Y) \leq T_{i_0}(Y) \text{ for any } i_0 \in H \text{ and all } Y.$$

Note that $T_H(Y)$, $T_{MIN}(Y)$ and $T_{i_0}(Y)$ are functions of Y only through $m(Y)$ when $v(Y)$ is fixed.

Lemma 2.

Assume that the number of elements of H is larger than one. Let $v(Y)$ be fixed. Let \mathcal{A} be the set of all the sample points of $m(Y)$ which satisfy $T_{MIN}(Y) > T_H(Y)$ in $R^{p \times q}$. Then \mathcal{A} is not empty.

Lemma 3.

Assume that the number of elements of H is larger than one and $i_0 \in H$. Let $v(Y)$ be fixed. Let \mathcal{B} be the set of all the sample points of $m(Y)$ which satisfy $T_{i_0}(Y) > T_{MIN}(Y)$ in $R^{p \times q}$. Then \mathcal{B} is not empty.

Proof of Theorem 1.

First we prove the first inequality of the theorem.

Let $v(Y)$ and any positive number t be fixed. Let \mathcal{A} be the non-empty set in $R^{p \times q}$ defined in Lemma 2. Both $T_{MIN}(Y)$ and $T_H(Y)$ are continuous functions of $m(Y)$, hence \mathcal{A} is an open set, and thus the Lebesgue measure of \mathcal{A} in $R^{p \times q}$ is positive.

For any $\epsilon > 0$, let $\mathcal{A}(\epsilon)$ denote the set of all the sample points of $m(Y)$ which satisfy $T_{MIN}(Y) > \epsilon > T_H(Y)$. Then, there exists a positive number ϵ_0 for which the Lebesgue measure of $\mathcal{A}(\epsilon_0)$ in $R^{p \times q}$ is also positive.

For any $\delta > 0$, define the set $\delta\mathcal{A}(\epsilon_0)$ in $R^{p \times q}$ by $\delta\mathcal{A}(\epsilon_0) = \{\delta x : x \in \mathcal{A}(\epsilon_0)\}$. It is trivial that, for any $\delta > 0$, the Lebesgue measure of $\delta\mathcal{A}(\epsilon_0)$ in $R^{p \times q}$ is also positive. Further, we can easily see that $\delta\mathcal{A}(\epsilon_0) = \mathcal{A}(\delta^2\epsilon_0)$ from the definitions of $T_{MIN}(Y)$ and $T_H(Y)$.

Hence, if we put $\delta_0 = (t/\epsilon_0)^{1/2}$, then $\mathcal{A}(t) = \mathcal{A}(\delta_0^2\epsilon_0) = \delta_0\mathcal{A}(\epsilon_0)$, and thus the Lebesgue measure

of the set of all the sample points of $m(Y)$ which satisfy $T_{MIN}(Y) > t > T_H(Y)$ in $R^{p \times q}$ is also positive.

Since $m(Y)$ is statistically independent of $\nu(Y)$ and the support of the probability distribution of $m(Y)$ is $R^{p \times q}$, we have

$$P_{\beta, \Sigma}(T_{MIN}(Y) > t > T_H(Y)) > 0,$$

for any β and Σ . Therefore, by using Lemma 1, we finally get

$$\begin{aligned} & P_{\beta, \Sigma}(T_{MIN}(Y) > t) \\ &= P_{\beta, \Sigma}(T_{MIN}(Y) > t, T_H(Y) > t) \\ &\quad + P_{\beta, \Sigma}(T_{MIN}(Y) > t, T_H(Y) \leq t) \\ &= P_{\beta, \Sigma}(T_H(Y) > t) \\ &\quad + P_{\beta, \Sigma}(T_{MIN}(Y) > t \geq T_H(Y)) \\ &\geq P_{\beta, \Sigma}(T_H(Y) > t) \\ &\quad + P_{\beta, \Sigma}(T_{MIN}(Y) > t > T_H(Y)) \\ &> P_{\beta, \Sigma}(T_H(Y) > t), \end{aligned}$$

for any β and Σ . Thus we complete the proof of the first inequality.

Next we prove the second inequality of the theorem. Let $\nu(Y)$ and any positive number t be fixed again. Let \mathcal{B} be the non-empty set in $R^{p \times q}$ defined in Lemma 3. Then we can apply exactly the same technique as that in the proof of the first inequality to complete the proof. \square

CONCLUSION

We considered the problem of testing homogeneity of regression coefficient vectors under synchronised order restrictions when the covariance matrices are common but unknown. For this problem, Hu and Banerjee (2012) proposed some test statistics, studied their distributions, and obtained some inequalities among their powers. In the present paper, we mathematically proved that the strict inequalities hold among the powers of the test statistics through a careful geometrical discussion. Thus we have attained an exact comparison among the powers of the test statistics indicating more accurate and stronger results than those obtained in Hu and Banerjee (2012).

REFERENCES

1. Barlow R.E., Bartholomew D.J., Bremner J.M. & Brunk H.D. (1972). *Statistical Inference under Order Restrictions: The Theory and Application of Isotonic Regression*. John Wiley & Sons, New York, USA.
2. Bartholomew D.J. (1959). A test of homogeneity for ordered alternatives. *Biometrika* **46**: 36 – 48.
DOI: <http://dx.doi.org/10.1093/biomet/46.1-2.36>
3. Hu X. (2009). p-values of a test on homogeneous means in a multivariate isotonic regression. *Statistics and Probability Letters* **79**: 2005 – 2011.
DOI: <http://dx.doi.org/10.1016/j.spl.2009.06.021>
4. Hu X. & Banerjee A. (2012). On the test for the homogeneity of a parameter matrix with some rows constrained by synchronized order restrictions. *Journal of Multivariate Analysis* **107**: 64 – 70.
DOI: <http://dx.doi.org/10.1016/j.jmva.2012.01.006>
5. Kulatunga D.D.S. & Sasabuchi S. (1984). A test of homogeneity of mean vectors against multivariate isotonic alternatives. *Memoirs of the Faculty of Science, Kyushu University, Series A (Mathematics)* **38**: 151 – 161.
DOI: <http://dx.doi.org/10.2206/kyushumfs.38.151>
6. Nomakuchi K. & Shi N.Z. (1988). A test for a multiple isotonic regression problem. *Biometrika* **75**: 181 – 184.
DOI: <http://dx.doi.org/10.1093/biomet/75.1.181>
7. Robertson T., Wright F.T. & Dykstra R.L. (1988). *Order Restricted Statistical Inference*. John Wiley & Sons, New York, USA.
8. Sasabuchi S. (2007). More powerful tests for homogeneity of multivariate normal mean vectors under an order restriction. *Sankhyā* **69**: 700 – 716.
9. Sasabuchi S., Inutsuka M. & Kulatunga D.D.S. (1983). A multivariate version of isotonic regression. *Biometrika* **70**: 465 – 472.
DOI: <http://dx.doi.org/10.1093/biomet/70.2.465>
10. Sasabuchi S., Kulatunga D.D.S. & Saito H. (1998). Comparison of powers of some tests for testing homogeneity under order restrictions in multivariate normal means. *American Journal of Mathematical and Management Sciences* **18**: 131 – 158.
DOI: <http://dx.doi.org/10.1080/01966324.1998.10737457>
11. Sasabuchi S., Miura T. & Oda H. (2003a). Estimation and test of several multivariate normal means under an order restriction when the dimension is larger than two. *Journal of Statistical Computation and Simulation* **73**: 619 – 641.
12. Sasabuchi S., Tanaka K. & Tsukamoto T. (2003b). Testing homogeneity of multivariate normal mean vectors under an order restriction when the covariance matrices are common but unknown. *Annals of Statistics* **31**: 1517 – 1536.
DOI: <http://dx.doi.org/10.1214/aos/1065705117>
13. Silvapulle M.J. & Sen P.K. (2005). *Constrained Statistical Inference: Inequality, Order, and Shape Restrictions*. John Wiley & Sons, New York, USA.

APPENDIX: PROOFS OF LEMMAS

Here we state the following lemma in order to prove Lemmas 1, 2 and 3.

Lemma 4.

For any H , a non-empty subset of $\{1, 2, \dots, p\}$, the following statements are true.

- (a) $\|m(Y) - P_{v(Y)}(m(Y) | \mathcal{L})\|_{v(Y)}^2 = \|m(Y)\|_{v(Y)}^2 - \|P_{v(Y)}(m(Y) | \mathcal{L})\|_{v(Y)}^2.$
- (b) $\|m(Y) - P_{v(Y)}(m(Y) | \mathcal{C}_H)\|_{v(Y)}^2 = \|m(Y)\|_{v(Y)}^2 - \|P_{v(Y)}(m(Y) | \mathcal{C}_H)\|_{v(Y)}^2.$
- (c) $T_H(Y) = \|P_{v(Y)}(m(Y) | \mathcal{L}) - P_{v(Y)}(m(Y) | \mathcal{C}_H)\|_{v(Y)}^2$
 $= \|m(Y) - P_{v(Y)}(m(Y) | \mathcal{L})\|_{v(Y)}^2 - \|m(Y) - P_{v(Y)}(m(Y) | \mathcal{C}_H)\|_{v(Y)}^2$
 $= \|P_{v(Y)}(m(Y) | \mathcal{C}_H)\|_{v(Y)}^2 - \|P_{v(Y)}(m(Y) | \mathcal{L})\|_{v(Y)}^2.$
- (d) If $P_{v(Y)}(m(Y) | \mathcal{C}_H) = 0$, then $P_{v(Y)}(m(Y) | \mathcal{L}) = 0$.

Proof.

(a): This equality follows from the usual orthogonal property of the projection because \mathcal{L} is a linear subspace in $R^{p \times q}$.

(b): This equality follows from equation (5) in Hu and Banerjee (2012).

$$\begin{aligned} \text{(c): } & \|m(Y) - P_{v(Y)}(m(Y) | \mathcal{L})\|_{v(Y)}^2 \\ &= \|m(Y) - P_{v(Y)}(m(Y) | \mathcal{C}_H) + P_{v(Y)}(m(Y) | \mathcal{C}_H) - P_{v(Y)}(m(Y) | \mathcal{L})\|_{v(Y)}^2 \\ &= \|m(Y) - P_{v(Y)}(m(Y) | \mathcal{C}_H)\|_{v(Y)}^2 + \|P_{v(Y)}(m(Y) | \mathcal{C}_H) - P_{v(Y)}(m(Y) | \mathcal{L})\|_{v(Y)}^2 \\ &\quad + 2\langle m(Y) - P_{v(Y)}(m(Y) | \mathcal{C}_H), P_{v(Y)}(m(Y) | \mathcal{C}_H) - P_{v(Y)}(m(Y) | \mathcal{L}) \rangle_{v(Y)}, \end{aligned}$$

by using the relation between the norm and inner product defined in Preliminaries. By equation (5) in Hu and Banerjee (2012),

$$\langle m(Y) - P_{v(Y)}(m(Y) | \mathcal{C}_H), P_{v(Y)}(m(Y) | \mathcal{C}_H) \rangle_{v(Y)} = 0.$$

Besides, since both $P_{v(Y)}(m(Y) | \mathcal{L})$ and $-P_{v(Y)}(m(Y) | \mathcal{L})$ belong to \mathcal{C}_H ,

$$\langle m(Y) - P_{v(Y)}(m(Y) | \mathcal{C}_H), P_{v(Y)}(m(Y) | \mathcal{L}) \rangle_{v(Y)} \leq 0$$

and

$$\langle m(Y) - P_{v(Y)}(m(Y) | \mathcal{C}_H), P_{v(Y)}(m(Y) | \mathcal{L}) \rangle_{v(Y)} \geq 0,$$

also by equation (5) in Hu and Banerjee (2012). Hence

$$\langle m(Y) - P_{v(Y)}(m(Y) | \mathcal{C}_H), P_{v(Y)}(m(Y) | \mathcal{C}_H) - P_{v(Y)}(m(Y) | \mathcal{L}) \rangle_{v(Y)} = 0,$$

and thus

$$\begin{aligned} & \left\| m(Y) - P_{\nu(Y)}(m(Y) | \mathcal{L}) \right\|_{\nu(Y)}^2 \\ &= \left\| m(Y) - P_{\nu(Y)}(m(Y) | \mathcal{C}_H) \right\|_{\nu(Y)}^2 + \left\| P_{\nu(Y)}(m(Y) | \mathcal{C}_H) - P_{\nu(Y)}(m(Y) | \mathcal{L}) \right\|_{\nu(Y)}^2, \end{aligned}$$

so the second equality of (c) follows. Further, by (a) and (b), the last equality of (c) follows.

(d): If $\left\| P_{\nu(Y)}(m(Y) | \mathcal{C}_H) \right\|_{\nu(Y)} = 0$, then, $\left\| P_{\nu(Y)}(m(Y) | \mathcal{L}) \right\|_{\nu(Y)} = 0$ because of (c) and the fact that $T_H(Y) \geq 0$. Thus (d) follows. \square

Proof of Lemma 1.

By (c) of Lemma 4,

$$T_H(Y) = \left\| m(Y) - P_{\nu(Y)}(m(Y) | \mathcal{L}) \right\|_{\nu(Y)}^2 - \left\| m(Y) - P_{\nu(Y)}(m(Y) | \mathcal{C}_H) \right\|_{\nu(Y)}^2$$

and

$$T_i(Y) = \left\| m(Y) - P_{\nu(Y)}(m(Y) | \mathcal{L}) \right\|_{\nu(Y)}^2 - \left\| m(Y) - P_{\nu(Y)}(m(Y) | \mathcal{C}_{\{i\}}) \right\|_{\nu(Y)}^2.$$

For $i \in H$, since $\mathcal{C}_H \subset \mathcal{C}_{\{i\}}$,

$$\left\| m(Y) - P_{\nu(Y)}(m(Y) | \mathcal{C}_H) \right\|_{\nu(Y)}^2 \geq \left\| m(Y) - P_{\nu(Y)}(m(Y) | \mathcal{C}_{\{i\}}) \right\|_{\nu(Y)}^2,$$

thus $T_H(Y) \leq T_i(Y)$. Hence $T_H(Y) \leq T_{MIN}(Y) \leq T_{i_0}(Y)$ for any $i_0 \in H$ and all Y . \square

Proof of Lemma 2.

For the geometrical discussions used here, readers are referred to, for example, Silvapulle and Sen (2005).

Assume that the number of the elements of H is larger than one. Let $H = \{i_1, i_2, \dots, i_k\}$, where $k > 1$.

$$\text{Then } \mathcal{C}_H = \mathcal{C}_{\{i_1\}} \cap \mathcal{C}_{\{i_2\}} \cap \dots \cap \mathcal{C}_{\{i_k\}}.$$

Let $\nu(Y)$ be fixed.

For any closed convex cone C in $R^{p \times q}$, let C^o denote the negative dual cone or polar cone of C , that is,

$$C^o = \{x \in R^{p \times q} : \langle x, y \rangle_{\nu(Y)} \leq 0 \text{ for all } y \in C\}.$$

Then, it is important to note that $P_{\nu(Y)}(x | C) = 0$ if and only if $x \in C^o$.

Since \mathcal{C}_H and $\mathcal{C}_{\{i_j\}}$, $j=1,2,\dots,k$, are (closed convex) polyhedral cones defined by some linear inequality constraints, \mathcal{C}_H^o and $\mathcal{C}_{\{i_j\}}^o$, $j=1,2,\dots,k$, are (closed convex) polyhedral cones generated by some half lines.

Then we have

$$\mathcal{C}_H^o = (\mathcal{C}_{\{i_1\}} \cap \mathcal{C}_{\{i_2\}} \cap \dots \cap \mathcal{C}_{\{i_k\}})^o = \mathcal{C}_{\{i_1\}}^o + \mathcal{C}_{\{i_2\}}^o + \dots + \mathcal{C}_{\{i_k\}}^o,$$

where $C_{\{i_1\}}^o + C_{\{i_2\}}^o + \dots + C_{\{i_k\}}^o$ is the direct sum of $C_{\{i_1\}}^o, C_{\{i_2\}}^o, \dots, C_{\{i_k\}}^o$.

The relative interior of C_H^o , denoted by $ri(C_H^o)$, is defined as the interior of C_H^o with respect to the relative topology induced in the linear subspace spanned by C_H^o .

There exists a sample point of $m(Y)$ such that $m(Y) \in ri(C_H^o)$, since the support of the probability distribution of $m(Y)$ is $R^{p \times q}$. For this point, since $m(Y) \in C_H^o$, $P_{v(Y)}(m(Y) | C_H) = 0$ and hence $P_{v(Y)}(m(Y) | \mathcal{L}) = 0$ by (d) of Lemma 4, thus $T_H(Y) = 0$. While, for $j = 1, 2, \dots, k$, since $C_{\{i_j\}}^o$ is a subset of the boundary (with respect to the relative topology) of C_H^o , $m(Y) \notin C_{\{i_j\}}^o$, hence $P_{v(Y)}(m(Y) | C_{\{i_j\}}) \neq 0$, and thus $T_{i_j}(Y) = \left\| P_{v(Y)}(m(Y) | C_{\{i_j\}}) \right\|_{v(Y)}^2 > 0$. Therefore, $T_{MIN}(Y) = \min\{T_{i_j}(Y) : j = 1, 2, \dots, k\} > 0 = T_H(Y)$. This implies that \mathcal{A} is not empty. \square

Proof of Lemma 3.

Assume that the number of the elements of H is larger than one and $i_0 \in H$.

Let $v(Y)$ be fixed.

There exists a number i such that $i \in H$ and $i \neq i_0$. For this i , there exists a sample point of $m(Y)$ such that $m(Y) \in C_{\{i_0\}}^o$ and $m(Y) \notin C_{\{i\}}^o$, since the support of the probability distribution of $m(Y)$ is $R^{p \times q}$. For this point, since $P_{v(Y)}(m(Y) | C_{\{i_0\}}) = m(Y)$ and $P_{v(Y)}(m(Y) | C_{\{i\}}) \neq m(Y)$, by (c) of Lemma 4,

$$\begin{aligned} T_{i_0}(Y) &= \left\| P_{v(Y)}(m(Y) | \mathcal{L}) - P_{v(Y)}(m(Y) | C_{\{i_0\}}) \right\|_{v(Y)}^2 \\ &= \left\| P_{v(Y)}(m(Y) | \mathcal{L}) - m(Y) \right\|_{v(Y)}^2 \end{aligned}$$

and

$$\begin{aligned} T_i(Y) &= \left\| m(Y) - P_{v(Y)}(m(Y) | \mathcal{L}) \right\|_{v(Y)}^2 - \left\| m(Y) - P_{v(Y)}(m(Y) | C_{\{i\}}) \right\|_{v(Y)}^2 \\ &< \left\| m(Y) - P_{v(Y)}(m(Y) | \mathcal{L}) \right\|_{v(Y)}^2, \end{aligned}$$

thus $T_{i_0}(Y) > T_i(Y)$, and hence $T_{i_0}(Y) > T_{MIN}(Y)$. This implies that \mathcal{B} is not empty. \square