TESTING HOMOGENEITY OF A PARAMETER MATRIX SOME OF THE ROWS OF WHICH ARE UNDER SYNCHRONIZED ORDER RESTRICTIONS

A Dissertation by

Arijit Banerjee

Master of Science, University of Calcutta, 2004
Bachelor of Science, University of Calcutta, 2002

Submitted to the Department of Mathematics, Statistics and Physics and the faculty of the Graduate School of Wichita State University in partial fulfillment of the requirements for the degree of Doctor of Philosophy

December 2011
The following faculty members have examined the final copy of this dissertation for form and content, and recommend that it be accepted in partial fulfillment of the requirement for the degree of Doctor of Philosophy with a major in Applied Mathematics.

________________________
Xiaomi Hu, Committee Chair

________________________
Hari Mukerjee, Committee Member

________________________
Lop-Hing Ho, Committee Member

________________________
Tianshi Lu, Committee Member

________________________
K Suresh Raju, Committee Member

 Accepted for the College of Liberal Arts and Sciences

________________________
William Bischoff, Dean

Accepted for the Graduate School

________________________
J. David McDonald, Dean
DEDICATION

To my parents, my teachers, my friends, and my wife
Get wisdom . . . though it costs all your possessions,
get understanding.
ACKNOWLEDGEMENTS

I would like to thank my advisor Dr. Xiaomi Hu, without whose help and guidance, I would not have done this dissertation. I am really grateful to him for all his time and efforts in teaching me various concepts in the subject. I am also grateful to Dr. Hari Mukerjee for his valuable suggestions and help. I would also like to thank my other committee members for their comments and help. I will be grateful to Dr. Kenneth Miller and Dr. Buma Fridman for supporting me throughout this time. I want to thank our department secretaries Terry, Deana and Janise for their co-operation and help. I am also indebted to Mark and Tom for assisting me with numerous computer related issues. Finally, I would like to thank my friends in and outside the department, and especially Dr. Deepak Aralumallige and Dr. Nanhee Kim for their valuable suggestions and the good time we spent together as office mates.

I am indebted to my parents for their everlasting support throughout this long journey. I would like to thank my wife for being with me here, thereby making life easier and happier for me.
ABSTRACT

This research considers a multiple multivariate linear regression model with a parameter matrix some of the rows of which are constrained by synchronized order restrictions. The test on the homogeneity of the parameter matrix is considered. Under the assumption that the common variance covariance matrix is unknown, an ad-hoc test statistic is proposed by replacing the unknown covariance matrix with its estimate in a likelihood ratio test statistic. The deterministic and probabilistic properties of the test statistic are studied. It is shown that the family of ad hoc tests share the same alpha level critical values and follow the same distribution for computing their p values. A sufficient condition is established for other tests to enjoy these properties and to be more powerful. Two examples of such more powerful tests are provided.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Introduction</td>
</tr>
<tr>
<td>1.1</td>
<td>Motivation</td>
</tr>
<tr>
<td>1.2</td>
<td>Literature review</td>
</tr>
<tr>
<td>1.3</td>
<td>Current Work</td>
</tr>
<tr>
<td>2</td>
<td>Preliminary definitions and results</td>
</tr>
<tr>
<td>2.1</td>
<td>Closed convex set in Hilbert space</td>
</tr>
<tr>
<td>2.2</td>
<td>Projections</td>
</tr>
<tr>
<td>2.2.1</td>
<td>Projection onto a closed convex set</td>
</tr>
<tr>
<td>2.2.2</td>
<td>Projection onto a subspace</td>
</tr>
<tr>
<td>2.2.3</td>
<td>Projection onto a closed convex cone</td>
</tr>
<tr>
<td>3</td>
<td>Multivariate isotonic regression model</td>
</tr>
<tr>
<td>3.1</td>
<td>A Multivariate model</td>
</tr>
<tr>
<td>3.1.1</td>
<td>Estimation of $\beta$ and $\Sigma$</td>
</tr>
<tr>
<td>3.2</td>
<td>Order restrictions on $\beta$</td>
</tr>
<tr>
<td>3.3</td>
<td>Multivariate isotonic regression</td>
</tr>
<tr>
<td>4</td>
<td>The testing problem and test statistic</td>
</tr>
<tr>
<td>4.1</td>
<td>Inner products, norms and projections</td>
</tr>
<tr>
<td>4.2</td>
<td>The testing problem</td>
</tr>
<tr>
<td>4.3</td>
<td>Construction of the likelihood ratio test statistic</td>
</tr>
<tr>
<td>5</td>
<td>Properties of $T_H(Y)$</td>
</tr>
<tr>
<td>5.1</td>
<td>$T_H$ of transformed $Y$</td>
</tr>
<tr>
<td>5.2</td>
<td>$T_H(Y)$ with different $H$</td>
</tr>
<tr>
<td>5.3</td>
<td>A limit of $T_H$ of transformed $Y$</td>
</tr>
<tr>
<td>6</td>
<td>The distributions of $T_H(Y)$</td>
</tr>
<tr>
<td>6.1</td>
<td>Condition for non-dependence on $\beta$</td>
</tr>
<tr>
<td>6.2</td>
<td>Condition for non-dependence on $\Sigma$</td>
</tr>
<tr>
<td>6.3</td>
<td>Two theorems</td>
</tr>
<tr>
<td>7</td>
<td>Main results</td>
</tr>
<tr>
<td>7.1</td>
<td>$\alpha$-level critical values</td>
</tr>
<tr>
<td>7.2</td>
<td>p-values</td>
</tr>
<tr>
<td>7.3</td>
<td>More powerful tests</td>
</tr>
</tbody>
</table>
TABLE OF CONTENTS (continued)

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>8 Conclusions</td>
<td>54</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>56</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION

This chapter is organized in the following way. In the next section, an example is provided which gave us the motivation behind this work. In section 1.2, a brief literature review is done for the topic of Order Restricted Statistical Inference. In section 1.3, we provide a synopsis of our work.

1.1 Motivation

Suppose a study is conducted among students of different grades with respect to age, family income, height and play hours. We represent the mean values of the four variables row-wise and the grades columnwise in the following table.

<table>
<thead>
<tr>
<th></th>
<th>4th grade</th>
<th>5th grade</th>
<th>6th grade</th>
<th>7th Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>$\beta_{11}$</td>
<td>$\beta_{12}$</td>
<td>$\beta_{13}$</td>
<td>$\beta_{14}$</td>
</tr>
<tr>
<td>Income</td>
<td>$\beta_{21}$</td>
<td>$\beta_{22}$</td>
<td>$\beta_{23}$</td>
<td>$\beta_{24}$</td>
</tr>
<tr>
<td>Height</td>
<td>$\beta_{31}$</td>
<td>$\beta_{32}$</td>
<td>$\beta_{33}$</td>
<td>$\beta_{34}$</td>
</tr>
<tr>
<td>Play hours</td>
<td>$\beta_{41}$</td>
<td>$\beta_{42}$</td>
<td>$\beta_{43}$</td>
<td>$\beta_{44}$</td>
</tr>
</tbody>
</table>

Intuitively, from the above table, one can assume

$$\beta_{11} \leq \beta_{12} \leq \beta_{13} \leq \beta_{14}$$

that is, the age of a student is positively correlated with the grade. We can also assume the height of a student to be positively correlated with grade, while the play hours to be negatively correlated with grade. The mean household income, however, is uncorrelated with the grade of a student, and hence will not be under any ordering.
Then, we construct our parameter matrix $\beta$ as

$$
\beta = \begin{pmatrix}
\beta_{11} & \beta_{12} & \beta_{13} & \beta_{14} \\
\beta_{21} & \beta_{22} & \beta_{23} & \beta_{24} \\
\beta_{31} & \beta_{32} & \beta_{33} & \beta_{34} \\
\beta_{41} & \beta_{42} & \beta_{43} & \beta_{44}
\end{pmatrix}
$$

We have $\beta_{31} \leq \beta_{32} \leq \beta_{33} \leq \beta_{34}$ and $\beta_{41} \geq \beta_{42} \geq \beta_{43} \geq \beta_{44}$, while the second row is free of any ordering.

Then, this matrix has four rows, three of which are constrained by known synchronized orderings. Problems like these are often encountered in practice, and we would like to test the homogeneity of such parameter matrices against the known synchronized orderings.

### 1.2 Literature review

A frequent problem which arises in statistical analysis is to test whether a number of normal variates have the same mean. Let $(X_1, X_2, \ldots, X_k)$ be independent observations where $X_i \sim N(\mu_i, \sigma_i), i = 1, 2, \ldots, k$, where $\sigma_i$’s are known. Then the null hypothesis of interest is given by

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_k \quad (1.1)$$

The most general form of the alternative hypothesis for the above testing problem is

$$H_1 : \mu_i \neq \mu_j \quad (1.2)$$

for some $i \neq j$ which imposes the minimum restriction on the $\mu_i, i = 1, 2, \ldots, k$.

In the standard ANOVA setting, the $\chi^2$-tests can be used for testing $H_0$ against $H_1$ using the likelihood-ratio test procedure.
Bartholomew (1959) considered a simple ordering over the \( \mu_i \)'s, and considered the alternative as

\[
H_2 : \mu_1 \leq \mu_2 \leq \cdots \leq \mu_k, \tag{1.3}
\]

not all \( \mu_i \) are equal. Here, at least one of the inequalities must be strict so that \( H_2 \) is not included in \( H_0 \).

He provided a general solution for testing \( H_0 \) against \( H_2 \). He used the same likelihood-ratio principle for the problem, which gave the following test criterion.

\[
\chi^2_k = \sum_{i=1}^{k} a_i (x_i - \bar{x})^2 - \sum_{i=1}^{k} a_i (x_i - \hat{\mu}_i)^2 \tag{1.4}
\]

where \( a_i = \frac{1}{\sigma_i^2}, \bar{x} = \frac{\sum_{i=1}^{k} a_i x_i}{\sum_{i=1}^{k} a_i} \) and \((\hat{\mu}_1, \hat{\mu}_2, \ldots, \hat{\mu}_k)\) are the values of \((\mu_1, \mu_2, \ldots, \mu_k)\) which maximize (1.4) subject to the condition \( \hat{\mu}_1 \leq \hat{\mu}_2 \leq \cdots \leq \hat{\mu}_k \).

Since then, different types of restrictions have been imposed upon the means, some of them are

\[
\begin{align*}
\mu_1 &\leq \mu_2 \leq \cdots \leq \mu_k \text{ (monotone increasing)} \\
\mu_1 &\geq \mu_2 \geq \cdots \geq \mu_k \text{ (monotone decreasing)} \\
\mu_1 &\leq \mu_i \text{ for } i = 2, 3, \ldots, k \text{ (simple tree ordering)} \\
\mu_1 &\leq \mu_2 \leq \cdots \leq \mu_i \geq \mu_{i+1} \geq \cdots \geq \mu_k \text{ (umbrella ordering)}
\end{align*}
\]


They considered $X_i(i = 1, 2, \ldots, k)$ as independent p-variate Normal random vectors with mean vectors $\mu_i$ and covariance matrices $\Sigma_i, i = 1, 2, \ldots, k$. They considered the problem of testing

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_k \text{ against } H_1 : (\mu_1, \mu_2, \ldots, \mu_k) \text{ is isotonic.}$$

The test statistic $\bar{\chi}^2_{k,p}$ was obtained under the assumption that the $\Sigma_i$'s are known. This test statistic is the multivariate extension of the $\chi^2$ statistic as defined by Bartholomew (1959, 1961b). They also gave an algorithm for computing the isotonic regression under a bivariate set-up.

Sasabuchi et al (2003) considered p-variate Normal distribution $N_p(\mu_i, \Sigma), i = 1, 2, \ldots, k$, and they tested the problem

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_k \text{ against } H_1 : \mu_1 \leq \mu_2 \leq \cdots \leq \mu_k$$

where $\mu_i \leq \mu_j$ meant that all the elements of $\mu_j - \mu_i$ are non-negative. They derived the likelihood ratio test (LRT) statistic for the testing problem using the LRT procedure. Under the assumption that the common covariance matrix $\Sigma$ is unknown, they proposed an ad-hoc test statistic by replacing $\Sigma$ with its estimate. They also studied the null distribution and the upper-tail probability of the test statistic under $H_0$. The simple expression of the upper tail probability is useful for computing the p-values of the test.

From component-wise simple ordering, Hu (2009) generalized the model restriction to a general vector quasi-ordering $\preceq$, i.e., he considered the alternative hypothesis as

$$H_3 : \mu_i \preceq \mu_j \text{ for } i, j \in B$$

Here, $\preceq$ is a special reflexive and binary relation of vectors in $R^p$ and $B$ is a subset of the column index space $\{(i, j) : i, j = 1, 2, \ldots, q\}$. 

4
Sasabuchi (2007) provided some tests for the testing problem which are more powerful than the tests proposed by Sasabuchi et al (2003). He also derived some theorems about the null distribution of the test statistics and their powers.

1.3 Current Work

The work presented here is the continuation of the research in this direction. We have extended and generalized the results in Sasabuchi et al (2003) and Hu (2009). While Sasabuchi et al (2003) considered a component-wise simple ordering and Hu (2009) considered a general vector quasi-ordering over the rows of the unknown parameter matrix, we have considered a synchronized order restriction over the rows of the parameter matrix. Also, the restrictions considered by Sasabuchi et al (2003) and Hu (2009) were imposed over all the rows, while in our case, only some of the rows of the parameter matrix are under synchronized order restrictions.

This paper considers a multivariate linear regression model

\[ Y \sim N(\beta M', \Sigma) \]  

\[ (1.5) \]

where \( \beta \) is a \( p \times q \) unknown parameter matrix and \( M \) is a \( n \times q \) known matrix of full column rank. Some of the rows of the unknown parameter matrix \( \beta \) is assumed to be under synchronized order restrictions. We have considered the problem of testing the homogeneity of the unknown parameter matrix against these synchronized order restrictions. We have assumed that the covariance matrix \( \Sigma \) is unknown. For this model, we propose an ad-hoc test statistic obtained by replacing the unknown \( \Sigma \) in a LRT statistic with its estimator. The row indices of the restricted rows of \( \beta \) form a non-empty set \( H \subset \{1, 2, \ldots, p\} \). With all possible \( H \), an ad hoc test family is produced.

We have shown here that all members of the ad hoc test family share the same \( \alpha \)-level critical values. Thus, all \( \alpha \)-level tests have the same rejection regions. It is also shown that a
unique distribution can be employed for the computation of p-values for all members of the family. A sufficient condition is established for other tests to enjoy the above two properties, and moreover, to be more powerful. Two such more powerful tests are examined.

The framework of the paper is given as follows.

Chapter 2 discusses the preliminary definitions and results which are used throughout the later chapters. The notion of a closed convex set is given and the concept of projections have also been introduced. We have discussed the necessary and sufficient conditions of a projection onto a closed convex set and a closed convex cone.

In Chapter 3, various multivariate models are discussed and a multivariate multiple regression model is used for our purpose. For model (1.5), we derive the Maximum Likelihood Estimate (MLE) and Unbiased Estimate (UE) for $\beta$ and $\Sigma$. In the next section, the order restriction on the unknown $\beta$ is defined and the multivariate isotonic regression model used in our discussion is introduced.

In Chapter 4, the testing problem is defined and using the LRT procedure, the LRT statistic is derived when $\Sigma$ is known. Under the assumption that $\Sigma$ is unknown, an ad-hoc test statistic is constructed on replacing the unknown $\Sigma$ by its estimate. The test statistic is denoted by $T_H(Y)$.

Chapter 5 deals with certain deterministic properties of $T_H(Y)$. A linear transformation of the data matrix $Y$ is considered and it is shown that $T_H(Y)$ is invariant under such transformation. We also discuss the limit of $T_H(Y)$ under a sequence of transformed $Y$.

Chapter 6 deals with the distributional and probabilistic properties of $T_H(Y)$. Two
theorems are stated and proved discussing the probabilistic properties of our test statistic.

In Chapter 7, we provide the main results of our tests, which are based on the two theorems presented in Chapter 6. Chapter 8 provides the summary of our findings.

To facilitate the discussion and avoid the operations such as vectorization and Kronecker product, the study is conducted in the space $\mathbb{R}^{p \times q}$ rather than $\mathbb{R}^{pq}$, using the concept of projections and the notation for random matrices with independent normal columns.
CHAPTER 2
PRELIMINARY DEFINITIONS AND RESULTS

In this chapter, the notion of Hilbert space and projections are presented. We are particularly interested in closed convex cones in a Hilbert space and projections onto those closed convex cones. Section 2.1 discusses about closed convex sets in a Hilbert space along with a proposition. Section 2.2 is about projections onto a closed convex set, subspace and closed convex cones.

2.1 Closed convex set in Hilbert space

We present some definitions to begin with.

**Definition 1**: Let $V$ be a linear space. A set $A \subset V$ is said to be convex if for $x, y \in A$, $\alpha x + (1 - \alpha) y \in A$, for all $\alpha \in (0, 1)$.

**Definition 2**: Let $M$ be a metric space. A set $A \subset M$ is said to be closed if for a sequence $x_n \in A$ where $x_n \to x$ as $n \to \infty$, then $x \in A$, that is, a closed set is one which contains all its limit points.

**Definition 3**: A sequence $X = (x_n)$ in a metric space is said to be a Cauchy sequence if for all $\epsilon > 0$, there exists a $K(\epsilon) > 0$ such that for all $m, n \geq K(\epsilon)$,

$$\|x_m - x_n\| < \epsilon.$$

**Definition 4**: A metric space $M$ is said to be complete (or Cauchy) if every Cauchy sequence in $M$ converges in $M$. 
**Definition 5:** Let $I$ be an inner product space. Then, $I$ is a Hilbert space if it is complete with respect to the metric defined by the norm $\|x - y\|$ induced from the inner product for $x, y \in I$.

Let $I$ be an inner product space. Then, for $x, y \in I$,

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2) \quad (2.1)$$

The equation (2.1) is known as the **Parallelogram equation**.

Now, we have the following proposition.

**Proposition 1:** Suppose $A$ is a closed convex set in a Hilbert space $\mathcal{H}$. Then, we have the following.

1. For all $x \in \mathcal{H}$, there exists $x^* \in A$ such that

   $$\|x - x^*\| = \inf\{\|x - y\| : y \in A\}$$

2. If $x \in \mathcal{H}$ and $x^* \in A$ such that $\|x - x^*\| = \inf\{\|x - y\| : y \in A\}$, then

   $$\langle x - x^*, x^* - y \rangle \geq 0 \text{ for all } y \in A.$$

3. If $x^* \in A$ and $\langle x - x^*, x^* - y \rangle \geq 0$ for all $y \in A$, then

   $$\|x - y\|^2 \geq \|x - x^*\|^2 + \|x^* - y\|^2 \text{ for all } y \in A.$$ 

4. If $x^* \in A$ and $\|x - y\|^2 \geq \|x - x^*\|^2 + \|x^* - y\|^2$ for all $y \in A$, then

   $$\|x - x^*\| = \inf\{\|x - y\| : y \in A\}$$

5. If $x_1^* \in A, x_2^* \in A, \|x - y\|^2 \geq \|x - x_1^*\|^2 + \|x_1^* - y\|^2$ and
\[\|x - y\|^2 \geq \|x - x^*_2\|^2 + \|x^*_2 - y\|^2 \text{ for all } y \in A\]

then \( x^*_1 = x^*_2 \).

**Proof:**

(1) Let \( d = \inf\{\|x - y\| : y \in A\} \). Then, by definiton of infimum, there exists a sequence \( \{x_n\} \) in \( A \) such that \( d \leq \|x - x_n\| \to d \) as \( n \to \infty \).

We claim that \( \{x_n\} \) is a Cauchy sequence.

Replacing \( x \) by \( x_s - x \) and \( y \) by \( x_t - x \) in (2.1), we get

\[
\|(x_s + x_t) - 2x\|^2 + \|x_s - x_t\|^2 = 2(\|x_s - x\|^2 + \|x_t - x\|^2)
\]

i.e.

\[
\|x_s - x_t\|^2 + \|2(x_s + x_t) - x\|^2 = 2(\|x_s - x\|^2 + \|x_t - x\|^2)
\]

Since \( A \) is convex, hence \( \frac{x_s + x_t}{2} \in A \) and \( d \leq \|\frac{x_s + x_t}{2} - x\| \).

Hence,

\[
\|x_s - x_t\|^2 \\
= 2(\|x_s - x\|^2) + 2(\|x_t - x\|^2) - 4\|\frac{x_s + x_t}{2} - x\|^2 \\
\leq 2\|x_s - x\|^2 + 2\|x_t - x\|^2 - 4d^2 \\
= 2(\|x_s - x\|^2 - d^2) + 2(\|x_t - x\|^2 - d^2).
\]

As \( s, t \to \infty, \|x_s - x_t\|^2 \to 0 \Rightarrow \{x_n\} \) is Cauchy, proving the claim.

Again, since \( \mathcal{H} \) is a Hilbert space, every Cauchy sequence in \( \mathcal{H} \) is convergent in \( \mathcal{H} \), that is, there exists \( x^* \in \mathcal{H} \) such that \( x_n \to x^* \). But, since \( A \) is closed, \( x^* \in A \).

Now, \( x^* \in A \) and \( x_n \to x^* \Rightarrow \|x - x_n\| \to \|x - x^*\| \). But

\[
\|x - x_n\| \to d \Rightarrow d = \|x - x^*\|. \text{ This proves (1).}
\]
(2) We define \( z = (1 - \alpha)x^* + \alpha y, \alpha \in (0, 1) \). Then, \( z \in A \), since \( A \) is convex. Then,

\[
\|x - x^*\|^2 \leq \|x - z\|^2 \\
= \|x - (1 - \alpha)x^* - \alpha y\|^2 \\
= \|x - x^* + \alpha(x^* - y)\|^2
\]

\[
\Rightarrow \|x - x^*\|^2 \leq \|x - x^*\|^2 + \alpha^2\|x^* - y\|^2 + 2\alpha \langle x - x^*, x^* - y \rangle
\]

\[
\Rightarrow 2\alpha \langle x - x^*, x^* - y \rangle \geq -\alpha^2\|x^* - y\|^2
\]

\[
\Rightarrow \langle x - x^*, x^* - y \rangle \geq -\frac{\alpha}{2}\|x^* - y\|^2 \text{ for all } \alpha \in (0, 1).
\]

We claim that \( \langle x - x^*, x^* - y \rangle \geq 0 \). This is true by the following argument. Suppose that \( \langle x - x^*, x^* - y \rangle \geq 0 \) is not true. Then, \( \langle x - x^*, x^* - y \rangle < 0 \), that is, \( \langle x - x^*, x^* - y \rangle < -\delta < 0 \) for some \( \delta > 0 \). But for this \( \delta \), we can find an \( \alpha \in (0, 1) \) such that \( 0 < \frac{\alpha}{2}\|x^* - y\|^2 < \delta \), i.e., \( -\delta < -\frac{\alpha}{2}\|x^* - y\|^2 < 0 \).

Hence, \( \langle x - x^*, x^* - y \rangle < -\delta < -\frac{\alpha}{2}\|x^* - y\|^2 \). But this contradicts

\[
\langle x - x^*, x^* - y \rangle \geq -\frac{\alpha}{2}\|x^* - y\|^2 \text{ for all } \alpha \in (0, 1)
\]

Hence, \( \langle x - x^*, x^* - y \rangle \geq 0 \).

(3) \( \|x - y\|^2 = \langle x - y, x - y \rangle = \langle (x - x^*) + (x^* - y), (x - x^*) + (x^* - y) \rangle \\
= \|x - x^*\|^2 + 2\langle x - x^*, x^* - y \rangle + \|x^* - y\|^2. \)

Now, since \( \langle x - x^*, x^* - y \rangle \geq 0 \), we get (3).

(4) \( \|x - y\|^2 \geq \|x - x^*\|^2 + \|x^* - y\|^2 \) for all \( y \in A \).

\[
\Rightarrow \|x - x^*\|^2 \leq \|x - y\|^2 - \|x^* - y\|^2 \text{ for all } y \in A.
\]

\[
\Rightarrow \|x - x^*\|^2 \leq \|x - y\|^2, \text{ since } \|x^* - y\|^2 \geq 0 \text{ for all } y \in A.
\]
\[ \Rightarrow \|x - x^*\| \leq \|x - y\| \text{ for all } y \in A. \]

But \( x^* \in A \). So, \( \|x - x^*\| = \inf\{\|x - y\| : y \in A\} \).

(5) Replacing \( y \) by \( x^*_2 \) in the first inequality, we have

\[ \|x - x^*_2\|^2 \geq \|x - x^*_1\|^2 + \|x^*_1 - x^*_2\|^2 \]  

(2.2)

Replacing \( y \) by \( x^*_1 \) in the second inequality, we have

\[ \|x - x^*_1\|^2 \geq \|x - x^*_2\|^2 + \|x^*_1 - x^*_2\|^2 \]  

(2.3)

Adding (2.2) and (2.3), we get

\[ \|x - x^*_2\|^2 + \|x - x^*_1\|^2 \geq \|x - x^*_1\|^2 + \|x - x^*_2\|^2 + 2\|x^*_1 - x^*_2\|^2 \]

\[ \Rightarrow \|x^*_1 - x^*_2\|^2 = 0 \Rightarrow x^*_1 = x^*_2. \]

\[ \blacksquare \]

2.2 Projections

2.2.1 Projection onto a closed convex set

Definition 6 : Let \( A \) be a closed convex set in a Hilbert space \( \mathcal{H} \). Then, by (1) of Proposition 1, there exists a \( x^* \in A \) such that

\[ \|x - x^*\| = \inf\{\|x - y\| : y \in A\} \]  

(2.4)

This \( x^* \) is defined as the projection of \( x \) onto \( A \) and is denoted by \( P(x|A) \).

We have the following lemma.

Lemma 1 : Suppose \( A \) is a closed convex set in a Hilbert space \( \mathcal{H} \). Then \( x^* \), the projection of \( x \) onto \( A \), denoted by \( P(x|A) \) for all \( x \in \mathcal{H} \), is unique.
Proof: Suppose for all \( x \in \mathcal{H} \), there exist \( x_1^*, x_2^* \in A \) such that
\[
\|x - x_1^*\| = \inf\{\|x - y\| : y \in A\} \quad \text{and} \quad \|x - x_2^*\| = \inf\{\|x - y\| : y \in A\}.
\]
Then, by (2) of Proposition 1, \( \langle x - x_1^*, x_1^* - y \rangle \geq 0 \) for all \( y \in A \) and \( \langle x - x_2^*, x_2^* - y \rangle \geq 0 \) for all \( y \in A \). Again, using (3) of Proposition 1, we get
\[
\|x - y\|^2 \geq \|x - x_1^*\|^2 + \|x_1^* - y\|^2 \quad \text{for all} \quad y \in A.
\]
Finally, using (5) of Proposition 1, we get \( x_1^* = x_2^* \), thereby implying that the projection \( x^* \) of \( x \) onto \( A \), is unique for all \( x \in \mathcal{H} \). □.

The following lemma gives a necessary and sufficient condition for \( x^* \) to be a projection onto a closed convex set.

Lemma 2 : \( x^* = P(x|A) \Leftrightarrow x^* \in A \) and \( \langle x - x^*, x^* - y \rangle \geq 0 \) for all \( y \in A \).

Proof: Suppose \( x^* = P_A(x) \), i.e., \( x^* \in A \) and \( \|x - x^*\| = \inf\{\|x - y\| : y \in A\} \). Then, by (2) of Proposition 1, \( x^* \in A \) and \( \langle x - x^*, x^* - y \rangle \geq 0 \) for all \( y \in A \).

Conversely, suppose \( x^* \in A \) and \( \langle x - x^*, x^* - y \rangle \geq 0 \) for all \( y \in A \). Then, by (3) of Proposition 1, we have
\[
x^* \in A \quad \text{and} \quad \|x - y\|^2 \geq \|x - x^*\|^2 + \|x^* - y\|^2 \quad \text{for all} \quad y \in A.
\]

Applying (4) of Proposition 1, we get \( x^* \in A \) and
\[
\|x - x^*\| = \inf\{\|x - y\| : y \in A\}
\]  (2.5)

Hence, by definition of projection, \( x^* = P(x|A) \). □

2.2.2 Projection onto a subspace

Let \( S \) be a closed subspace in a Hilbert space \( \mathcal{H} \). Then, we have the following lemma.

Lemma 3 : \( x^* = P_S(x) \Leftrightarrow x^* \in S \) and \( \langle x - x^*, y \rangle = 0 \) for all \( y \in S \)

Proof: \( S \) is closed, and being a subspace, \( S \) is convex. So, \( S \) is a closed, convex set in a Hilbert space \( \mathcal{H} \).
Thus, \( x^* = P(x|S) \Leftrightarrow x^* \in S \) and \( \langle x - x^*, x^* - y \rangle \geq 0 \) for all \( y \in S \) is true from Lemma 1.

Now, we need to show that
\[
\begin{align*}
x^* \in S & \text{ and } \langle x - x^*, x^* - y \rangle \geq 0 \text{ for all } y \in S \\
\Leftrightarrow x^* \in S & \text{ and } \langle x - x^*, y \rangle = 0 \text{ for all } y \in S \\
\Rightarrow \langle x - x^*, y \rangle &= \langle x - x^*, x^* - (x^* - y) \rangle. \text{ Since } x^* \in S, y \in S \text{ and } S \text{ is a subspace, hence } x^* - y \in S. \\
\end{align*}
\]

Thus,
\[
\langle x - x^*, y \rangle = \langle x - x^*, x^* - (x^* - y) \rangle \geq 0 \tag{2.6}
\]
from the given condition.

Again,
\[
\begin{align*}
\langle x - x^*, y \rangle &= \langle x - x^*, -(x^* - y) \rangle \\
&= -\langle x - x^*, x^* - (x^* + y) \rangle \\
\text{But, since } \langle x - x^*, x^* - y \rangle &\geq 0 \text{ for all } y \in S \text{ and } x^* + y \in S, \text{ hence } \\
\langle x - x^*, y \rangle &= -\langle x - x^*, x^* - (x^* + y) \rangle \leq 0 \tag{2.7}
\end{align*}
\]
Combining (2.6) and (2.7), we get \( \langle x - x^*, y \rangle = 0 \) for all \( y \in S \).

\( \Leftarrow \)

Since \( x^* \in S \) and \( y \in S \), hence \( x^* - y \in S \), thereby \( \langle x - x^*, x^* - y \rangle = 0 \), which is a special case of \( \langle x - x^*, x^* - y \rangle \geq 0 \). □

### 2.2.3 Projection onto a closed convex cone

**Definition 7**: A set \( C \) is said to be a cone if for \( x \in C \) and \( \alpha \geq 0 \), \( \alpha x \in C \).

**Definition 8**: A subset \( C \) of a Hilbert space \( H \) is said to be a closed convex cone if \( C \) is closed, \( C \) is convex and \( C \) is a cone.

The following lemma gives a necessary and sufficient condition for \( x^* \) to be a projection onto a closed convex cone \( C \).
Lemma 4: Let $C \subset \mathcal{H}$, where $C$ is a closed convex cone and $\mathcal{H}$ is a Hilbert space. Then, $x^* = P(x|C) \iff x^* \in C, \langle x - x^*, x^* \rangle = 0$ and $\langle x - x^*, y \rangle \leq 0$ for all $y \in C$.

Proof: $x^* = P(x|C) \iff x^* \in C$ and $\langle x - x^*, x^* - y \rangle \geq 0$ for all $y \in C$ is true from Lemma 2.

Now, to show that $x^* \in C$ and $\langle x - x^*, x^* - y \rangle \geq 0$ for all $y \in C$ $\iff x^* \in C, \langle x - x^*, x^* \rangle = 0$ and $\langle x - x^*, y \rangle \leq 0$ for all $y \in C$.

$\Rightarrow$ part: Suppose $x^* \in C$ and $\langle x - x^*, x^* - y \rangle \geq 0$ for all $y \in C$. Let $y = \alpha x^*$ for $\alpha > 0$. Then, since $C$ is a cone, $\alpha x^* \in C$.

Now, $\langle x - x^*, x^* - y \rangle = \langle x - x^*, x^* - \alpha x^* \rangle \geq 0 \Rightarrow \langle x - x^*, x^* \rangle - \alpha \langle x - x^*, x^* \rangle \geq 0$.

If $\alpha \in (0, 1)$, then $\langle x - x^*, x^* \rangle \geq 0$ and if $\alpha > 1$, $\langle x - x^*, x^* \rangle \leq 0$, thus combining, we have $\langle x - x^*, x^* \rangle = 0$.

Again, $\langle x - x^*, x^* - y \rangle = \langle x - x^*, x^* \rangle - \langle x - x^*, y \rangle$.

Since $\langle x - x^*, x^* - y \rangle \geq 0$ and $\langle x - x^*, x^* \rangle = 0$, $\langle x - x^*, y \rangle \leq 0$ for all $y$.

$\Leftarrow$ part: Suppose $\langle x - x^*, x^* \rangle = 0$ and $\langle x - x^*, y \rangle \leq 0$ for all $y \in C$.

Then $\langle x - x^*, x^* - y \rangle = \langle x - x^*, x^* \rangle - \langle x - x^*, y \rangle \geq 0$ for all $y \in C$.

□

Lemma 5: Suppose $L \subset C \subset \mathcal{H}$, where $L$ is a linear space, $C$ is a closed convex cone and $\mathcal{H}$ is a Hilbert space. Then $\langle x - P(x|C), y \rangle = 0$ for all $y \in L$, i.e., $\langle x, y \rangle = \langle P(x|C), y \rangle$ for all $x \in \mathcal{H}, y \in L$.

Proof: From Lemma 4, we see that if $x^* = P(x|C)$, then $\langle x - x^*, y \rangle \leq 0$ for all $y \in C$.

Now, $y \in L$ and $L$ is a linear space, i.e., if $y \in L$, then $-y \in L$.

Hence, $\langle x - P(x|C), y \rangle \leq 0 \Rightarrow \langle x, y \rangle \leq \langle P(x|C), y \rangle$ for all $y \in L$. 

15
Replacing $y$ by $-y$, we get $\langle x, y \rangle \geq \langle P(x|C), y \rangle$ for all $y \in L$.

Combining, we have the result. □
CHAPTER 3
MULTIVARIATE ISOTONIC REGRESSION MODEL

3.1 A Multivariate model

Let $Y_1, Y_2, \ldots, Y_n$ be a set of $p$-variate random observations. We have a data matrix given by $Y = (Y_1, Y_2, \ldots, Y_n)$, where $Y \in \mathbb{R}^{p \times n}$. It is assumed that $Y \sim N(E(Y), \Sigma)$. This means that the columns of the data matrix $Y$ are independent Normal vectors with common covariance matrix $\Sigma$ and the expectation of $Y$ is $E(Y)$. This notation, proposed by Hu (2009), has many convenient properties. Here, we will need that

$$Y \sim N(E(Y), \Sigma) \Rightarrow AY + B \sim N(AE(Y) + B, A\Sigma A')$$

By specifying different structures of $E(Y)$, one can have different models for the data matrix $Y$. A few are discussed below.

(1) One sample model: Suppose $Y_1, Y_2, \ldots, Y_n$ is a random sample from a Normal distribution with mean vector $\mu$ and common variance-covariance matrix $\Sigma$. Then, the data matrix is given by

$$Y = (Y_1, Y_2, \ldots, Y_n) \sim N(\mu 1_n', \Sigma)$$

where $1_n = (1, 1, \ldots, 1)' \in \mathbb{R}^n$.

(2) Two sample model: Let $E(Y) = \beta B'$, where $\beta = (\mu_1, \mu_2) \in \mathbb{R}^{p \times 2}$, $B = \begin{pmatrix} 1_{n_1} & 0 \\ 0 & 1_{n_2} \end{pmatrix} \in \mathbb{R}^{n \times 2}$, with $n = n_1 + n_2$. Then, $Y \sim N(\beta B', \Sigma)$ means that $Y_1, Y_2, \ldots, Y_{n_1}$ is a random sample from $N(\mu_1, \Sigma)$ and $Y_{n_1+1}, Y_{n_1+2}, \ldots, Y_n$ is a random sample from $N(\mu_2, \Sigma)$. 
(3) **MANOVA model**: With $\beta = (\mu_1, \mu_2, \ldots, \mu_q) \in R^{p \times q}$ and $B = \begin{pmatrix} 1_{n_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1_{n_q} \end{pmatrix}$ with $n = n_1 + n_2 + \cdots + n_q$, $Y \sim N(\beta B', \Sigma)$ means that $Y_1, Y_2, \ldots, Y_n$ are $p$-variate random vectors from $q$ populations with mean vectors $\mu_1, \mu_2, \ldots, \mu_q$ respectively and common covariance matrix $\Sigma$. The sample size for the $i$th population is $n_i$.

(4) **A multivariate linear regression model**: In this paper, we consider the data matrix $Y \in R^{p \times n}$ as $Y \sim N(\beta M', \Sigma)$, where $\beta$ is a $(p \times q)$ parameter matrix, $M$ is a $(n \times q)$ matrix of full column rank which is known, and has the structure similar to the matrix $B$ in model 3.

Then this model is the generalization of all the above three models.

### 3.1.1 Estimation of $\beta$ and $\Sigma$

We consider model (4) in our discussion. From the joint probability density function of $Y_1, Y_2, \ldots, Y_n$, we obtain the MLE and UE of $\beta$ and $\Sigma$.

For $Y = (Y_1, Y_2, \ldots, Y_n) \sim N(\beta M', \Sigma)$ with $\mu = \beta M' = (\mu_1, \mu_2, \ldots, \mu_n) \in R^{p \times n}$, the joint p.d.f. of $Y_1, Y_2, \ldots, Y_n$ is

\[
\prod_{i=1}^{n} \frac{1}{(2\pi)^{p/2}|\Sigma|^{1/2}} \exp\{-\frac{1}{2}[(Y_i - \mu_i)^\prime \Sigma^{-1} (Y_i - \mu_i)]\}
\]

\[
= \frac{1}{(2\pi)^{np/2}|\Sigma|^{n/2}} \exp\{-\frac{1}{2} \sum_{i=1}^{n} [(Y_i - \mu_i)^\prime \Sigma^{-1} (Y_i - \mu_i)]\}
\]

\[
= \frac{1}{(2\pi)^{np/2}|\Sigma|^{n/2}} \exp\{-\frac{1}{2} \text{tr}(Y - \mu)^\prime \Sigma^{-1} (Y - \mu)\}
\]

\[
= \frac{1}{(2\pi)^{np/2}|\Sigma|^{n/2}} \exp\{-\frac{1}{2} \text{tr}[\Sigma^{-1/2} (Y - \mu)(Y - \mu)^\prime \Sigma^{-1/2}]\}
\]
Thus the likelihood function of $\beta$ and $\Sigma$ is

$$L(\beta, \Sigma) = \frac{1}{(2\pi)^{np/2}|\Sigma|^{n/2}} \exp\{-\frac{1}{2} \text{tr}[\Sigma^{-1/2}(Y - \beta M')(Y - \beta M')\Sigma^{-1/2}]\}.$$ 

Now, we define

$$m(Y) = YM(M'M)^{-1}$$  \hspace{1cm} (3.1) 

$$v(Y) = Y[I - M(M'M)^{-1}M']Y'$$  \hspace{1cm} (3.2) 

The following lemma gives us the MLE and UE of $\beta$ and $\Sigma$.

**Lemma 6**: Suppose $Y \sim N(\beta M', \Sigma)$.

(1) $m(Y)$ is the MLE for $\beta$ and it is also an UE for $\beta$.

(2) $\frac{n(Y)}{n}$ is the MLE for $\Sigma$ and $\frac{v(Y)}{n-q}$ is an UE for $\Sigma$.

**Proof:**

(1) Substituting $m(Y) = YM(M'M)^{-1}$, we get

$$[Y - m(Y)M'][m(Y)M' - \beta M']' = 0$$

and

$$[m(Y)M' - \beta M'][Y - m(Y)M']' = 0$$

Then,

$$(Y - \beta M')(Y - \beta M')' = [Y - m(Y)M'][Y - m(Y)M']' + [m(Y)M' - \beta M'][m(Y)M' - \beta M']'$$
Premultiplying by $-\frac{1}{2} \Sigma^{-1/2}$ and post-multiplying by $\Sigma^{-1/2}$ and taking trace on both sides, we get

$$\text{tr}\{-\frac{1}{2} \Sigma^{-1/2}(Y-\beta M')(Y-\beta M')' \Sigma^{-1/2}\} \leq \text{tr}\{-\frac{1}{2} \Sigma^{-1/2}(Y-m(Y)M')(Y-m(Y)M')' \Sigma^{-1/2}\}$$

since $\text{tr}\{\Sigma^{-1/2}[m(Y) - \beta M'][m(Y) - \beta M']' \Sigma^{-1/2}\} \geq 0$.

Taking exponents on both sides, we get $L(m(Y), \Sigma) \geq L(\beta, \Sigma)$ for all $\beta, \Sigma$, thus proving that $m(Y)$ is the MLE for $\beta$.

Again, $E(m(Y)) = E[YM(M'M)^{-1}] = E(Y)M(M'M)^{-1} = \beta M'M(M'M)^{-1} = \beta$, thereby proving that $m(Y)$ is the UE for $\beta$.

(2) Let $A = \Sigma^{-1/2}(Y - m(Y)M')(Y - m(Y)M')' \Sigma^{-1/2}$.

By spectral decomposition,

$$A = P \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_p) P'.$$

Then, we have

$$\text{tr}[A] = \sum_{i=1}^{n} \lambda_i$$

and

$$|A| = \prod_{i=1}^{n} \lambda_i.$$
Thus,

\[
\frac{1}{|\Sigma|^{\frac{n}{2}}} = |\Sigma^{-1}|^{\frac{n}{2}}
\]
\[
= |\Sigma^{-1/2} [Y - m(Y)M'] [Y - m(Y)M'] (\Sigma^{-1/2})^{\frac{n}{2}}
\]
\[
= \frac{(\lambda_1, \lambda_2, \ldots, \lambda_p)^{\frac{n}{2}}}{|\Sigma^{-1/2} [Y - m(Y)M'] [Y - m(Y)M'] (\Sigma^{-1/2})^{\frac{n}{2}}}
\]

Hence,

\[
L(m(Y), \Sigma) = \frac{1}{(2\pi)^{np/2}} \cdot \exp\left\{ -\frac{\lambda_1 + \lambda_2 + \cdots + \lambda_p}{2} \right\} \cdot \frac{(\lambda_1, \lambda_2, \ldots, \lambda_p)^{\frac{n}{2}}}{|\Sigma^{-1/2} [Y - m(Y)M'] [Y - m(Y)M'] (\Sigma^{-1/2})^{\frac{n}{2}}}
\]

where

\[
c = \frac{1}{(2\pi)^{np/2}} \cdot \frac{1}{|\Sigma^{-1/2} [Y - m(Y)M'] [Y - m(Y)M'] (\Sigma^{-1/2})^{\frac{n}{2}}}
\]

Now, we define \( g(\lambda_i) = \lambda_i^{n/2} e^{-\lambda_i/2} \).

Then, \( \ln g(\lambda_i) = \frac{n}{2} \ln \lambda_i - \frac{\lambda_i}{2} \)

and \( [\ln g(\lambda_i)]' = \frac{n}{2} \left( \frac{1}{\lambda_i} - \frac{1}{n} \right) \).

Solving \( [\ln g(\lambda_i)]' = 0 \) for \( \lambda_i \) gives \( \lambda_i = n \).

Also, \( [\ln g(\lambda_i)]' < 0 \) for \( \lambda_i > n \) and \( [\ln g(\lambda_i)]' > 0 \) for \( \lambda_i < n \), implying that \( \ln g(\lambda_i) \), and hence \( g(\lambda_i) \) is maximized for \( \lambda_i = n \).

Thus,

\[
\Sigma^{-\frac{1}{2}} [Y - m(Y)M'] [Y - m(Y)M'] \Sigma^{-\frac{1}{2}} = PP'n = nI
\]

Pre-multiplying and post-multiplying both sides by \( \Sigma^{\frac{1}{2}} \), we get

\[
\hat{\Sigma} = \frac{(Y - m(Y)M')}n = \frac{\hat{v}(Y)}n,
\]

using the definition of \( m(Y) \) which maximizes the likelihood function \( L(m(Y), \Sigma) \). Thus, we get
\[ L(\beta, \Sigma) \leq L(m(Y), \Sigma) \leq L(m(Y), \frac{v(Y)}{n}) \]

thereby implying that \( \frac{v(Y)}{n} \) is the MLE for \( \Sigma \).

Let \( A = I - M(M'M)^{-1}M' \). Then

\[
E\left[\frac{v(Y)}{n-q}\right] = \frac{1}{n-q} E(YAY')
\]

\[
= \frac{1}{n-q} [(E(Y))A(E(Y))' + tr(A)\Sigma]
\]

\[
= \frac{1}{n-q} [(\beta M')(I - M(M'M)^{-1}M')\beta M')' + tr(A)\Sigma]
\]

But, \( tr(A) = n - q \) and \( (\beta M')[I - M(M'M)^{-1}M'](\beta M')' = 0 \)

thereby implying \( E\left[\frac{v(Y)}{n-q}\right] = \Sigma \).

Thus, \( \frac{v(Y)}{n-q} \) is an UE of \( \Sigma \). □

3.2 Order restrictions on \( \beta \)

Let \( \ll \) be a partial order in \( \Omega = \{x_1, x_2, \ldots, x_q\} \), i.e., \( \ll \) is a binary relation of elements of \( \Omega \), which is reflexive (\( x_i \ll x_i \) for all \( x_i \in \Omega \)), transitive (\( x_i \ll x_j \) and \( x_j \ll x_k \Rightarrow x_i \ll x_k \)) and antisymmetric (\( x_i \ll x_j \) and \( x_j \ll x_i \Rightarrow x_i = x_j \)). Let \( f = (f_1, f_2, \ldots, f_q)' \) be a vector in \( R^q \) with \( f_i = f(x_i), i = 1, 2, \ldots, q \). Then this relation defines an order restriction on the elements of \( f \in R^q \) as \( x_i \ll x_j \Rightarrow f_i \leq f_j \). Denote the collection of all vectors \( f \) satisfying the order restriction as \( C_+ \). It is now shown that \( C_+ \) is a convex cone in \( R^q \).

Note that \( f \in C_+ \Rightarrow f_i \leq f_j \) for \( x_i \ll x_j \). Similarly, \( g \in C_+ \Rightarrow g_i \leq g_j \) for \( x_i \ll x_j \). For \( \alpha, \beta \geq 0, \alpha f_i \leq \alpha f_j \) and \( \beta g_i \leq \beta g_j \).

Thus, for \( x_i \ll x_j \),

\[
\alpha f_i + \beta g_i \leq \alpha f_j + \beta g_j
\]
implying that \( \alpha f + \beta g \in C_+ \). Hence, \( C_+ \) is a convex cone in \( \mathbb{R}^q \).

The partial order \( \ll \) also defines an opposite order restriction as
\( x_i \ll x_j \Rightarrow f_i \geq f_j \). The collection of all vectors \( f \) satisfying this order restriction is denoted by \( C_- \). By similar argument as \( C_+ \), \( C_- \) is also a convex cone in \( \mathbb{R}^q \), and clearly, \( C_+ = -C_- \).

The rows of the parameter matrix \( \beta = (\beta(1), \beta(2), \ldots, \beta(p))' \in \mathbb{R}^{p \times q} \) are constrained by synchronized order restrictions with respect to \( \ll \) if
\[
\beta(i) \in C_{(i)} \text{ for all } i
\]
where \( C_{(i)} \) is either \( C_+ \) or \( C_- \).
These order restrictions are said to be synchronized since they are based on the same \( \ll \).
Suppose \( H \) is a non-empty subset of row-index space \( \{1, 2, \ldots, p\} \). Then
\[
C_H = \{ (\beta(1), \beta(2), \ldots, \beta(p))' \in \mathbb{R}^{p \times q} : \beta(i) \in C_{(i)}, i \in H \} \tag{3.3}
\]
is the collection of all matrices with some rows, the rows with indices in \( H \), constrained by synchronized orderings. Then, \( C_H \) is a convex cone in \( \mathbb{R}^{p \times q} \). Set \( H \) is partitioned by \( H_+ \) and \( H_- \) such that
\[
i \in H_+ \Rightarrow C_{(i)} = C_+, \ i \in H_- \Rightarrow C_{(i)} = C_-
\]
and without loss of generality, we assume that the set \( H_+ \) is non-empty.

### 3.3 Multivariate isotonic regression

The multivariate linear regression model on page 18, with \( \beta \in C_H \), is called a multivariate isotonic regression model. A model like this covers many interesting practical problems.

Referring to the example which we provided in section 1.1, we can see that our unknown parameter matrix \( \beta \) has four rows, three of which are under synchronized order...
restrictions, that is, $\beta_{(1)} \in C_+, \beta_{(3)} \in C_+$ while $\beta_{(4)} \in C_-$, where $\beta_{(i)}$ is the $i$th row of $\beta$. $\beta_{(2)}$, however, is free of any restriction.


CHAPTER 4
THE TESTING PROBLEM AND TEST STATISTIC

In this chapter, we formulate the test and derive the likelihood-ratio test statistic using the likelihood-ratio test procedure under the assumption that the common covariance matrix is known. Then, when the common covariance matrix is unknown, we derive an ad-hoc test statistic for the testing problem.

4.1 Inner products, norms and projections

Here, we define an inner product for matrices $A$ and $B$ in $R^{p \times q}$. Let $A$ and $B$ be matrices in $R^{p \times q}$. Let $V \in R^{p \times p}$ be a positive-definite matrix. Define

$$
\langle A, B \rangle_V = \text{tr}(A'V^{-1}BM'M).
$$

We have the following lemma.

Lemma 7 : The expression in (4.1) defines an inner product for matrices $A$ and $B$ in $R^{p \times q}$.

Proof: It is enough to show that (4.1) follows all the properties that define an inner product.

(a) We have

$$
\langle A, A \rangle_V = \text{tr}(A'V^{-1}AM'M)
= \text{tr}(MA'V^{-1}AM')
= \text{tr}(MA'V^{-1/2}V^{-1/2}AM')
= \text{tr}(DD')
$$

where $D = (d_{ij}) = MA'V^{-1/2} \in R^{m \times p}$. Therefore, $\langle A, A \rangle_V = \sum_{i=1}^{n} \sum_{j=1}^{p} d_{ij}^2 \geq 0$.

Now, to show that $\langle A, A \rangle_V = 0$ if and only if $A = 0$

Suppose $A = 0$. Then, $D = 0$. Consequently, $\langle A, A \rangle_V = \text{tr}(DD') = 0$.

Conversely, let $\langle A, A \rangle_V = 0$, i.e., $\langle A, A \rangle_V = \text{tr}(DD') = \sum_{i=1}^{n} \sum_{j=1}^{p} d_{ij}^2 = 0$. Then,
\[ d_{ij} = 0 \text{ for all } i, j. \] So, \( D = 0 \), i.e., \( MA'V^{-1/2} = 0 \).

Now, \( MA'V^{-1/2} = 0 \Rightarrow MA' = 0 \Rightarrow (M'M)^{-1}M'MA' = 0 \Rightarrow A' = 0 \Rightarrow A = 0 \).

(b) \( \langle A, B \rangle_V = \text{tr}(A'V^{-1}BM'M) = \text{tr}[(M'MB'V^{-1}A)'] = \text{tr}[(BM'M)'V^{-1}A]' = \langle B, A \rangle_V \).

(c) \( \langle A+C, B \rangle_V = \text{tr}[(A+C)'V^{-1}BM'M] = \text{tr}[(C'+A')V^{-1}BM'M] = \text{tr}(C'V^{-1}BM'M) + \text{tr}(A'V^{-1}BM'M) = \langle A, B \rangle_V + \langle C, B \rangle_V \).

(d) For any scalar \( \alpha \),
\[
\langle \alpha A, B \rangle_V = \text{tr}[(\alpha A)'V^{-1}BM'M] = \alpha \text{tr}(A'V^{-1}BM'M) = \alpha \langle A, B \rangle_V .
\]

Thus, (4.1) follows all the properties of an inner product, thereby proving the lemma.

\[\square\]

Then, \( \|A\| = \langle A, A \rangle^{1/2} \) induces a norm in \( R^{p \times q} \) with respect to the inner product.

We have the following lemma.

**Lemma 8**: \( C_H \) is a closed convex cone in \( R^{p \times q} \) with respect to the norm induced from the inner product as defined.

**Proof**: We have defined \( C_H \) as
\[
C_H = \{ (\beta(1), \beta(2), \ldots, \beta(p))' \in R^{p \times q} : \beta(i) \in C_{(i)} \text{ for all } i \in H \}
\]
which is a convex cone in \( R^{p \times q} \). We need to show that \( C_H \) is closed with respect to the norm induced from the inner product as defined in equation (4.1). By definition of a closed set, we need to show the following:
If \( \beta^{[n]} \in C_H \) and \( \beta^{[n]} \rightarrow \beta^* \), i.e., \( \|\beta^{[n]} - \beta^*\| \rightarrow 0 \), then \( \beta^* \in C_H \).
Now,
\[ \beta_n \in C_H \]
\[ \iff \beta_{(i)}^n \in C_{(i)} \text{ for all } i \in H \]
\[ \iff \beta_{(i)}^n \in C_+ \text{ when } C_{(i)} = C_+ \text{ and } \beta_{(i)}^n \in C_- \text{ when } C_{(i)} = C_- \]
\[ \iff \text{ if } x_s \ll x_t, \beta_{is}^n \leq \beta_{it}^n \text{ when } C_{(i)} = C_+ \text{ and } \beta_{is}^n \geq \beta_{it}^n \text{ when } C_{(i)} = C_- \]

In a finite-dimensional space, the convergence with respect to a metric induced from an inner product is always equivalent to the component-wise convergence. Thus,
\[ \beta^n \rightarrow \beta^* \iff \beta_{is}^n \rightarrow \beta_{is}^* \text{ and } \beta_{it}^n \rightarrow \beta_{it}^* \]

Thus, we get
\[ \beta_{is}^* \leq \beta_{it}^* \text{ for } C_{(i)} = C_+ \text{ and } x_s \ll x_t, \beta_{is}^* \geq \beta_{it}^* \text{ for } C_{(i)} = C_- \text{ and } x_s \ll x_t \]
\[ \iff \beta_{(i)}^* \in C_+ \text{ if } C_{(i)} = C_+ \text{ and } \beta_{(i)}^* \in C_- \text{ if } C_{(i)} = C_- \]
\[ \iff \beta_{(i)}^* \in C_{(i)} \text{ for all } i \in H \]
\[ \iff \beta^* \in C_H. \]

Thus, for any matrix \( A \in \mathbb{R}^{p \times q} \), there exists a unique matrix \( A^* \in C_H \) such that
\[ \|A - A^*\| = \inf\{\|A - B\| : B \in C_H\} \quad (4.2) \]

The matrix \( A^* \) is called the projection of \( A \) onto \( C_H \) and is denoted by \( P_V(A|C_H) \).

**Definition 9** : A matrix \( \beta = (\beta_1, \beta_2, \ldots, \beta_q) \in \mathbb{R}^{p \times q} \) is a homogeneous matrix if \( \beta_1 = \beta_2 = \cdots = \beta_q \).

The collection of all homogeneous matrices, \( \mathcal{L} = \{x_1' : x \in \mathbb{R}^p\} \), where
\[ 1_q = (1, 1, \ldots, 1)' \in \mathbb{R}^q, \]
is a special convex cone, a linear space in \( \mathbb{R}^{p \times q} \), and \( \mathcal{L} \subset C_H \).

Since the linear space \( \mathcal{L} \) is also closed, \( P_V(A|\mathcal{L}) \) can be defined analogously.

### 4.2 The testing problem

Let \( Y = (Y_1, Y_2, \ldots, Y_n) \) be the data matrix in \( \mathbb{R}^{p \times n} \). For model \( Y \sim N(\beta M', \Sigma) \), where \( \beta \in \mathbb{R}^{p \times q} \) is the unknown parameter matrix and \( M \in \mathbb{R}^{n \times q} \) is a known matrix of full column rank, we consider a testing problem.
\[ H_0 : \beta \in \mathcal{L} \text{ against } H_1 : \beta \in \mathcal{C}_H. \]

### 4.3 Construction of the likelihood ratio test statistic

For our testing problem, we use likelihood-ratio test procedure to derive our test statistic. We consider two cases: one where \( \Sigma \) is known, and one where \( \Sigma \) is unknown.

(i) \( \Sigma \) is known: Under the testing scheme, the likelihood function is given by

\[
L(\beta) = \frac{1}{(2\pi)^{np/2}|\Sigma|^{1/2}} \exp\left\{-\frac{1}{2}\text{tr}[\Sigma^{-1/2}(Y - \beta M')(Y - \beta M')\Sigma^{-1/2}]\right\}.
\]

Let \( m(Y) = YM(M'M)^{-1} = \hat{\beta} \). Then, considering the exponential part, we have

\[
\text{tr}[\Sigma^{-1/2}(Y - \beta M')(Y - \beta M')\Sigma^{-1/2}]
= \text{tr}[\Sigma^{-1/2}(Y - \hat{\beta} M' + \hat{\beta} M' - \beta M')(Y - \hat{\beta} M' + \hat{\beta} M' - \beta M')\Sigma^{-1/2}]
= \text{tr}[\Sigma^{-1/2}(Y - \hat{\beta} M')(Y - \hat{\beta} M') + (\hat{\beta} M' - \beta M')(\hat{\beta} M' - \beta M')]
= \text{tr}[\Sigma^{-1/2}(\hat{\beta} M' - \beta M')\Sigma^{-1/2}]
\]

Now, \( \text{tr}[\Sigma^{-1/2}(Y - \hat{\beta} M')(\hat{\beta} M' - \beta M')\Sigma^{-1/2}] = 0 \).

Also, \( \text{tr}[\Sigma^{-1/2}(\hat{\beta} M' - \beta M')(Y - \hat{\beta} M')\Sigma^{-1/2}] = 0 \).

Thus,

\[
\text{tr}[\Sigma^{-1/2}(Y - \beta M')(Y - \beta M')\Sigma^{-1/2}]
= \text{tr}[\Sigma^{-1/2}(Y - \hat{\beta} M')(Y - \hat{\beta} M')\Sigma^{-1/2}] + \text{tr}[\Sigma^{-1/2}(\hat{\beta} M' - \beta M')(\hat{\beta} M' - \beta M')\Sigma^{-1/2}]
= 0 + \text{tr}[\Sigma^{-1/2}(\hat{\beta} M' - \beta M')\Sigma^{-1/2}]
\]

But, \( Q = \text{tr}[\Sigma^{-1/2}(Y - \hat{\beta} M')(Y - \hat{\beta} M')\Sigma^{-1/2}] \) is a statistic and

\[
\text{tr}[\Sigma^{-1/2}(\hat{\beta} M' - \beta M')(\hat{\beta} M' - \beta M')\Sigma^{-1/2}]
= \text{tr}[\Sigma^{-1/2}(\hat{\beta} - \beta)M'M(\hat{\beta} - \beta)\Sigma^{-1/2}]
= \text{tr}[(\hat{\beta} - \beta)^2\Sigma^{-1/2}\Sigma^{-1/2}(\hat{\beta} - \beta)M'M]
= \text{tr}[(\hat{\beta} - \beta)^2]\Sigma^{-1}(\hat{\beta} - \beta)M'M
= \|\hat{\beta} - \beta\|^2_S
\]
Thus
\[ L(\beta) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{n/2}} \exp\{-\frac{1}{2}Q - \frac{1}{2}\|\hat{\beta} - \beta\|^2_\Sigma\}. \]

Clearly,
\[ \max\{L(\beta) : \beta \in C_H\} = \frac{1}{(2\pi)^{n/2}|\Sigma|^{n/2}} \exp\{-\frac{1}{2}Q - \frac{1}{2}\|\hat{\beta} - P_\Sigma(\hat{\beta}|C_H)\|^2_\Sigma\}, \]
\[ \max\{L(\beta) : \beta \in L\} = \frac{1}{(2\pi)^{n/2}|\Sigma|^{n/2}} \exp\{-\frac{1}{2}Q - \frac{1}{2}\|\hat{\beta} - P_\Sigma(\hat{\beta}|L)\|^2_\Sigma\}. \]

So, \( \Lambda = \frac{\max\{L(\beta) : \beta \in C_H\}}{\max\{L(\beta) : \beta \in L\}} = e^{-1/2T} \)
where \( T = \|\hat{\beta} - P_\Sigma(\hat{\beta}|L)\|^2_\Sigma - \|\hat{\beta} - P_\Sigma(\hat{\beta}|C_H)\|^2_\Sigma \)

Since \( \Lambda \) is an increasing function in \( T \), hence, substituting \( m(Y) \) for \( \hat{\beta} \), the likelihood ratio test statistic is given by
\[ T(Y) = \|m(Y) - P_\Sigma(m(Y)|L)\|^2_\Sigma - \|m(Y) - P_\Sigma(m(Y)|C_H)\|^2_\Sigma. \]

This test statistic can be simplified as that in the next lemma.

**Lemma 9**: The test statistic \( T(Y) \) can be written as
\[ T(Y) = \|P_\Sigma(m(Y)|L) - P_\Sigma(m(Y)|C_H)\|^2_\Sigma. \]

**Proof**: We have
\[ \|m(Y) - P_\Sigma(m(Y)|L)\|^2_\Sigma \]
\[ = \|m(Y) - P_\Sigma(m(Y)|C_H) + P_\Sigma(m(Y)|C_H) - P_\Sigma(m(Y)|L)\|^2_\Sigma \]
\[ = \|m(Y) - P_\Sigma(m(Y)|C_H)\|^2_\Sigma + \|P_\Sigma(m(Y)|L) - P_\Sigma(m(Y)|C_H)\|^2_\Sigma \]
\[ -2\langle P_\Sigma(m(Y)|L) - P_\Sigma(m(Y)|C_H), m(Y) - P_\Sigma(m(Y)|C_H) \rangle \]
Now,
\[ \langle m(Y) - P_\Sigma(m(Y)|C_H), P_\Sigma(m(Y)|L) - P_\Sigma(m(Y)|C_H) \rangle \]
\[ = \langle m(Y) - P_\Sigma(m(Y)|C_H), P_\Sigma(m(Y)|L) \rangle - \langle m(Y) - P_\Sigma(m(Y)|C_H), P_\Sigma(m(Y)|C_H) \rangle \]
\[ = \langle m(Y) - P_\Sigma(m(Y)|C_H), P_\Sigma(m(Y)|L) \rangle \]

29
since \( (m(Y) - P_\Sigma(m(Y)|C^H), P_\Sigma(m(Y)|C^H)) = 0 \) by Lemma 4.

Now, \( P_\Sigma(m(Y)|\mathcal{L}) \in \mathcal{L} \) and \( \mathcal{L} \subset C^H \). Hence, by Lemma 5,

\[
(m(Y) - P_\Sigma(m(Y)|C^H), P_\Sigma(m(Y)|\mathcal{L})) = 0.
\]

Thus,

\[
\|m(Y) - P_\Sigma(m(Y)|\mathcal{L})\|_\Sigma^2 = \|m(Y) - P_\Sigma(m(Y)|C^H)\|_\Sigma^2 + \|P_\Sigma(m(Y)|\mathcal{L}) - P_\Sigma(m(Y)|C^H)\|_\Sigma^2
\]

\( \Rightarrow T(Y) = \|P_\Sigma(m(Y)|\mathcal{L}) - P_\Sigma(m(Y)|C^H)\|_\Sigma^2 \)

This proves the lemma. \( \square \).

Thus, when \( \Sigma \) is known, the likelihood ratio test statistic is given by

\[
T(Y) = \|P_\Sigma(m(Y)|\mathcal{L}) - P_\Sigma(m(Y)|C^H)\|_\Sigma^2
\]

(4.3)

and \( H_0 \) is rejected for large values of \( T(Y) \).

(ii) \( \Sigma \) is unknown: When \( \Sigma \) is unknown, we can obtain three different test statistics as follows:

(1) We replace \( \Sigma \) by its MLE \( \hat{\Sigma}(Y) \) and obtain a test statistic as

\[
T_1(Y) = \|P_{\hat{\Sigma}(Y)}(m(Y)|\mathcal{L}) - P_{\hat{\Sigma}(Y)}(m(Y)|C^H)\|_{\hat{\Sigma}(Y)}^2
\]

(2) We replace \( \Sigma \) by its UE \( \hat{\Sigma}(Y) \) and obtain a test statistic as

\[
T_2(Y) = \|P_{\hat{\Sigma}(Y)}(m(Y)|\mathcal{L}) - P_{\hat{\Sigma}(Y)}(m(Y)|C^H)\|_{\hat{\Sigma}(Y)}^2
\]

(3) We replace \( \Sigma \) simply by \( v(Y) \) and obtain a test statistic as

\[
T_H(Y) = \|P_{v(Y)}(m(Y)|\mathcal{L}) - P_{v(Y)}(m(Y)|C^H)\|_{v(Y)}^2
\]

The following lemma relates the test statistics \( T_1(Y) \), \( T_2(Y) \) and \( T_H(Y) \).

Lemma 10 : The test statistic \( T_H(Y) \) is proportional to \( T_1(Y) \) and \( T_2(Y) \)
Proof: For $k > 0$, we have

$$
\langle A, B \rangle_{kV} = \text{tr}[A'(kV)^{-1}BM'M]
= \frac{1}{k}\text{tr}[A'V^{-1}BM'M]
= \frac{1}{k}\langle A, B \rangle_V
$$

Thus,

$$
\langle A, B \rangle_V = k\langle A, B \rangle_{kV} \text{ and } \|A\|_V^2 = k\|A\|_{kV}^2.
$$

(4.4)

Clearly $P_{kV}(A|C_H) \in C_H$. By the first equation in (4.4) and Lemma 4,

$$
\langle A - P_{kV}(A|C_H), P_{kV}(A|C_H) \rangle_V
= k\langle A - P_{kV}(A|C_H), P_{kV}(A|C_H)_{kV} \rangle
= 0.
$$

Also with $B \in C_H$,

$$
\langle A - P_{kV}(A|C_H), B \rangle_V
= k\langle A - P_{kV}(A|C_H), B \rangle_{kV}
\leq 0.
$$

Hence, by Lemma 4,

$$
P_V(A|C_H) = P_{kV}(A|C_H)
$$

(4.5)

for any $k > 0$ and matrix $A \in C_H$.

Then, using (4.5) and (4.4) we have

$$
T_1(Y) = \|P_{v(Y)}(m(Y)|C_H) - P_{v(Y)}(m(Y)|C_H)\|_{v(Y)}^2
= \|P_{v(Y)}(m(Y)|L) - P_{v(Y)}(m(Y)|C_H)\|_{v(Y)}^2
= n\|P_{v(Y)}(m(Y)|L) - P_{v(Y)}(m(Y)|C_H)\|_{v(Y)}^2
= nT_H(Y)
$$

31
By the same argument, we have $T_2(Y) = (n - q)T_H(Y)$.

Thus, our lemma is proved. □.

Thus, our proposed test statistic is given by

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

(4.6)

and $H_0$ is rejected for large values of this statistic.
CHAPTER 5

PROPERTIES OF $T_H(Y)$

In this section, we study the deterministic properties of our proposed test statistic $T_H(Y)$. Here, $T_H(Y)$ is treated as a function of $Y \in \mathbb{R}^{p \times n}$ without assuming the randomness of $Y$.

5.1 $T_H$ of transformed $Y$

Before studying the deterministic properties of $T_H(Y)$, the following lemma is stated and proved.

Lemma 11:

1. $m(Y - \beta M') = m(Y) - \beta$
2. $v(Y - \beta M') = v(Y)$
3. $m(DY) = Dm(Y)$ where $D$ is a $p \times p$ non-singular matrix.
4. $v(DY) = Dv(Y)D'$ where $D$ is a $p \times p$ non-singular matrix.
5. $\langle A, B \rangle_{v(DY)} = \langle D^{-1}A, D^{-1}B \rangle_{v(Y)}$

Proof:

1. By the definition of $m(Y)$ in (3.2),

   $$m(Y - \beta M') = (Y - \beta M')M(M'M)^{-1}$$
   $$= YM(M'M)^{-1} - \beta (M'M)(M'M)^{-1}$$
   $$= m(Y) - \beta$$
(2) By the definition of $v(Y)$ in (3.3),

$$v(Y - \beta M')$$

$$= (Y - \beta M')[I - M(M'M)^{-1}M'](Y - \beta M')'$$

$$= Y[I - M(M'M)^{-1}M']Y'$$

$$= v(Y)$$

(3) $m(DY) = DY M(M'M)^{-1} = Dm(Y)$

(4) $v(DY) = DY[I - M(M'M)^{-1}M']Y'D' = Dv(Y)D'$.

(5) By our definition of norm and using (4) of this lemma, we have

$$\langle A, B \rangle_{v(DY)}$$

$$= \text{tr}[A'(v(DY))^{-1}BM'M]$$

$$= \text{tr}[A'(Dv(Y)D')^{-1}BM'M]$$

$$= \text{tr}[A'(D^{-1})'(v(Y))^{-1}D^{-1}BM'M]$$

$$= \text{tr}[(D^{-1}A)'v^{-1}(Y)(D^{-1}B)M'M]$$

$$= \langle D^{-1}A, D^{-1}B \rangle_{v(Y)}. \square$$

Lemma 12 : Let $C_H$ and $L$ be defined as in (3.3) and page 27 respectively. Then,

(1) When $\beta \in L$, 

(a) $P_{v(Y)}(m(Y) - \beta | C_H) = P_{v(Y)}(m(Y) | C_H) - \beta$

(b) $P_{v(Y)}(m(Y) - \beta | L) = P_{v(Y)}(m(Y) | L) - \beta$

(2) For any non-singular matrix $D$,

(a) $D^{-1}P_{v(DY)}(m(DY) | C_H) = P_{v(Y)}(m(Y) | D^{-1}C_H)$

(b) $D^{-1}P_{v(DY)}(m(DY) | L) = P_{v(Y)}(m(Y) | L)$
Proof:

(1) Only (a) is proved, because (b) can be proven analogously. In order to prove (a), by Lemma 4, one needs to show

\[ (P_v(Y)(m(Y)|C_H) - \beta) \in C_H \]
\[ (m(Y) - \beta) - (P_v(Y)(m(Y)|C_H) - \beta) \leq 0 \]
\[ (m(Y) - \beta) - (P_v(Y)(m(Y)|C_H) - \beta) \leq 0 \]

(i): \( P_v(Y)(m(Y)|C_H) \in C_H \). If \( \beta \in L \), then \(-\beta \in L \) since \( L \) is a linear space. But \( L \subset C_H \). So \(-\beta \in L \). It is to be noted that \( C_H \) is a convex cone. Hence \( P_v(Y)(m(Y)|C_H) - \beta \in C_H \).

(ii): For (ii), we have

\[ ((m(Y) - \beta) - (P_v(Y)(m(Y)|C_H) - \beta), P_v(Y)(m(Y)|C_H) - \beta) \in (v(Y) \]
\[ = (m(Y) - P_v(Y)(m(Y)|C_H), P_v(Y)(m(Y)|C_H) - \beta) \]
\[ = (m(Y) - P_v(Y)(m(Y)|C_H), P_v(Y)(m(Y)|C_H) - \beta) \]

But, \( (m(Y) - P_v(Y)(m(Y)|C_H), P_v(Y)(m(Y)|C_H) - \beta) \in (v(Y) \]
and \( (m(Y) - P_v(Y)(m(Y)|C_H), \beta) \in (v(Y) \]
Thus, \( ((m(Y) - \beta) - (P_v(Y)(m(Y)|C_H) - \beta), P_v(Y)(m(Y)|C_H) - \beta) \in (v(Y) \]

(iii): For any matrix \( B \in C_H \), by Lemma 4,

\[ ((m(Y) - \beta) - (P_v(Y)(m(Y)|C_H) - \beta), B) \]
\[ = (m(Y) - (P_v(Y)(m(Y)|C_H), B) \]
\[ \leq 0 \]

Thus, claim (a) is proved.

(2) Only (a) is proved, because (b) can be proven analogously.

Note that \( D^{-1}C_H \) is a closed convex cone. In order to prove (a), one needs to show

\[ i) D^{-1}P_{v(DY)}(m(DY)|C_H) \in D^{-1}C_H \]
(ii) $\langle m(Y) - D^{-1}P_v(DY)(m(DY)|C_H), D^{-1}P_v(DY)(m(DY)|C_H) \rangle_{v(Y)} = 0$

(iii) $\langle m(Y) - D^{-1}P_v(DY)(m(DY)|C_H), B \rangle_{v(Y)} \leq 0$ for all $B \in D^{-1}C_H$.

(i): $P_v(DY)(m(DY)|C_H) \in C_H$ implies that $\Rightarrow D^{-1}P_v(DY)(m(DY)|C_H) \in D^{-1}C_H$

(ii): By (5) of Lemma 11, (3) of Lemma 11 and Lemma 4,

\[
\langle m(Y) - D^{-1}P_v(DY)(m(DY)|C_H), D^{-1}P_v(DY)(m(DY)|C_H) \rangle_{v(Y)} \\
= \langle Dm(Y) - P_v(DY)(m(DY)|C_H), P_v(DY)(m(DY)|C_H) \rangle_{v(DY)} \\
= \langle m(DY) - P_v(DY)(m(DY)|C_H), P_v(DY)(m(DY)|C_H) \rangle_{v(DY)} \\
= 0
\]

(iii): For any matrix $B \in D^{-1}C_H$, according to (5) of Lemma 11, (3) of Lemma 11 and Lemma 4,

\[
\langle m(Y) - D^{-1}P_v(DY)(m(DY)|C_H), B \rangle_{v(Y)} \\
= \langle Dm(Y) - P_v(DY)(m(DY)|C_H), DB \rangle_{v(DY)} \\
= \langle m(DY) - P_v(DY)(m(DY)|C_H), G \rangle_{v(DY)} \\
\leq 0.
\]

Hence, (a) is proved. \qed

The deterministic properties of $T_H(Y)$ are established in the following two lemma.

**Lemma 13**: If $\beta \in L$, then $T_H(Y - \beta M') = T_H(Y)$.

**Proof**: By the definition of $T_H(Y)$ in equation (4.6), (1) of Lemma 11, (2) of Lemma 11 and (1) of Lemma 12,
\[ T_H(Y - \beta M') \]
\[ = \| P_v(Y - \beta M') (m(Y - \beta M')|\mathcal{L}) - P_v(Y - \beta M') (m(Y - \beta M')|\mathcal{C}_H) \|^2 \]
\[ = \| P_v(Y - \beta |\mathcal{L}) - P_v(Y - \beta |\mathcal{C}_H) \|^2 \]
\[ = \| P_v(Y)|\mathcal{L}) - P_v(Y)m(Y)|\mathcal{C}_H \|^2 \]
\[ = \| P_v(Y)|\mathcal{L}) - P_v(Y)m(Y)|\mathcal{C}_H \|^2 \]
\[ = T_H(Y). \square \]

**Lemma 14**: If \( D^{-1}C_H = C_H \), then \( T_H(DY) = T_H(Y) \).

**Proof**: According to (5) of Lemma 11,
\[ \| A \|^2_{v(DY)} = \| D^{-1}A \|^2_{v(Y)} \] (5.1)

Then, by the definition of \( T_H(Y) \) in (4.6), equation (5.1), (2) of Lemma 12 and the given condition,
\[ T_H(DY) \]
\[ = \| P_v(DY)(m(DY)|\mathcal{L}) - P_v(DY)(m(DY)|\mathcal{C}_H) \|^2 \]
\[ = \| D^{-1}P_v(DY)(m(DY)|\mathcal{L}) - D^{-1}P_v(DY)(m(DY)|\mathcal{C}_H) \|^2 \]
\[ = \| P_v(Y)(m(Y)|\mathcal{L}) - P_v(Y)(m(Y)|\mathcal{C}_H) \|^2 \]
\[ = T_H(Y). \square \]

**5.2 \( T_H(Y) \) with different \( H \)**

We consider a non-empty subset \( H_* \) of \( H \) and establish a relationship between the two test statistics \( T_H(Y) \) and \( T_{H*}(Y) \). Before that, we prove the following proposition.

**Proposition 2**: Let \( H_* \) be a non-empty subset of \( H \). Define
\[ s = \langle P_v(Y)(m(Y)|\mathcal{L}) - P_v(Y)(m(Y)|\mathcal{C}_H), P_v(Y)(m(Y)|\mathcal{C}_H) - P_v(Y)(m(Y)|\mathcal{C}_{H*}) \rangle_{v(Y)}, \]
Then, \( s \geq 0. \)
Proof: From the definition of $s,$

$$\langle P_v(Y) (m(Y)| L) - P_v(Y) (m(Y)| H) \rangle = \langle P_v(Y) (m(Y)| L) - P_v(Y) (m(Y)| H) \rangle = \langle P_v(Y) (m(Y)| L) - P_v(Y) (m(Y)| H) \rangle = \langle P_v(Y) (m(Y)| L) - P_v(Y) (m(Y)| H) \rangle.$$ 

Here $P_v(Y)(m(Y)| L) \in L$ and $L \subset C_H.$ Hence, by lemma 5, the first term is zero. By lemma 4, the second term is zero. Again, $P_v(Y)(m(Y)| L) \in L$ and $L \subset C_{H^*}.$ Hence, by lemma 5, the third term is zero. Thus

$$s = -\langle P_v(Y) (m(Y)| C_H), m(Y) - P_v(Y) (m(Y)| C_{H^*}) \rangle.$$ 

But, $H^* \subset H$ implies $C_H \subset C_{H^*}$ and hence $P_v(Y)(m(Y)| C_H) \in C_{H^*}.$ By Lemma 4

$$\langle P_v(Y)(m(Y)| C_H), m(Y) - P_v(Y)(m(Y)| C_{H^*}) \rangle \leq 0.$$

So $s \geq 0.$

The following lemma establishes a relationship between $T_H(Y)$ and $T_{H^*}(Y)$.

Lemma 15: For any non-empty subset $H$ of $H, T_H(Y) \leq T_{H^*}(Y)$ holds.

Proof: Using the definition of $T_H(Y)$ in equation (4.6),

$$T_{H^*}(Y) = \| P_v(Y)(m(Y)| L) - P_v(Y)(m(Y)| C_{H^*}) \|^2_{v(Y)}$$

$$= \| P_v(Y)(m(Y)| L) - P_v(Y)(m(Y)| C_H) + P_v(Y)(m(Y)| C_H) - P_v(Y)(m(Y)| C_{H^*}) \|^2_{v(Y)}$$

$$= T_H(Y) + \| P_v(Y)(m(Y)| C_H) - P_v(Y)(m(Y)| C_{H^*}) \|^2_{v(Y)} + 2s,$$

where $s$ is as in proposition 2. Using the conclusion from proposition 2, $T_H(Y) \leq T_{H^*}(Y),$ thereby the lemma is established. □.
5.3 A limit of $T_H$ of transformed Y

In this section, we construct a sequence of non-singular matrices $D^{[N]}$ in $\mathbb{R}^{p \times p}$ and study the asymptotic properties of the sequence of projections of $m(Y)$ onto $D^{[N]}C_H$ denoted by $\{P_{e(Y)}(m(Y)|D^{[N]}C_H)\}$. Finally, it is proved that when $N$ goes to infinity, the sequence of test statistics $T_H(D^{-N}Y)$ converges to $T_i(Y)$ for $i_0 \in H$.

For $i_0 \in H$ let $u = (u_1, u_2, \ldots, u_p)' \in \mathbb{R}^p$, where

$$u_i = \begin{cases} 
1, i \neq i_0, i \in H_+ \\
0, i = i_0 \\
-1, i \neq i_0, i \in H_-
\end{cases}$$

Define

$$D^{[N]} = \begin{cases} 
I_p - N \cdot u e'_{i_0}, i_0 \in H_+ \\
I_p + N \cdot u e'_{i_0}, i_0 \in H_-
\end{cases}$$

where $N$ is an integer and $e_{i_0}$ is the $i_0$th column of $I_p$. Then

$$D^{[m]}D^{[N]} = \begin{cases} 
(I_p - m \cdot u e'_{i_0})(I_p - N \cdot u e'_{i_0}), i_0 \in H_+ \\
(I_p + m \cdot u e'_{i_0})(I_p + N \cdot u e'_{i_0}), i_0 \in H_-
\end{cases}$$

$$= \begin{cases} 
I_p - (m + N) \cdot u e'_{i_0} + mN(u e'_{i_0})(u e'_{i_0}), i_0 \in H_+ \\
I_p + (m + N) \cdot u e'_{i_0} + mN(u e'_{i_0})(u e'_{i_0}), i_0 \in H_-
\end{cases}$$

Now, $ue'_{i_0}ue'_{i_0} = 0$, since $e'_{i_0}u = 0$. Thus,

$$D^{[m]}D^{[N]} = \begin{cases} 
I_p - (m + N) \cdot u e'_{i_0}, i_0 \in H_+ \\
I_p + (m + N) \cdot u e'_{i_0}, i_0 \in H_-
\end{cases}$$

Clearly, $D^{[m]}D^{[N]} = D^{[m+N]}$ and $D^{[0]} = I_p$. Hence, $(D^{[N]})^{-1} = D^{-N}$. 

39
When $H$ contains only one element, say $i_0$, $C_H$ is simply denoted by $C_{i_0}$. Clearly, by definition of $C_H$, $B = (B(1), B(2), \ldots, B(p))' \in C_{i_0}$ if and only if $B(i_0) \in C(i_0)$. We now present a proposition where a sequence of matrices $B^{[m]} \in R^{p \times q}$ is constructed and it is shown that as $m \to \infty$, the sequence $B^{[m]}$ converges to a matrix $B \in C_{i_0}$ for $i_0 \in H$. It is also shown that for all $N > N_m$, $B^{[m]} \in D^{[N]} C_H$.

Proposition 3: For $i_0 \in H$ and $B \in C_{i_0}$, there exists $B^{[m]} \in R^{p \times q}$ and positive integers $N_m, m = 1, 2, \ldots$ such that

1. $B^{[m]} \to B$ as $m \to \infty$.
2. $B^{[m]} \in D^{[N]} C_H$ for all $N > N_m$.

Proof: Without loss of generality, assume $i_0 \in H_+$.

(1) Let $v = (v_1, v_2, \ldots, v_q)'$ be an interior point of $C_{(i_0)}$, i.e., $v \in C_{(i_0)}$, $v_j \neq v_k$ when $j \neq k$.

Let $B = (B(1), \ldots, B(p))' \in C_{i_0}$, i.e., $B(i_0) \in C(i_0)$. Define

$B^{[m]} = B + \frac{e_{i_0}v'}{m}$. Then, $B^{[m]} \to B$ as $m \to \infty$.

(2) Write $B^{[m]} = D^{[N]}(D^{-N} B^{[m]})$. Then, for (2), it is sufficient to show that

$A = (A(1), A(2), \ldots, A(p))' = D^{-N} B^{[m]} \in C_H$ for sufficiently large $N$, or equivalently, $A(i) \in C(i)$ for all $i$ for sufficiently large $N$.

By the construction of $D^{[N]}$ and the definition of $B^{[m]}$,

$$(A(1), A(2), \ldots, A(p))' = D^{-N} B^{[m]}$$

$$= (I_p + N \cdot u e_{i_0}')(B + \frac{e_{i_0}v'}{m})$$

$$= B + \frac{e_{i_0}v'}{m} + N \cdot u e_{i_0}'(B + \frac{e_{i_0}v'}{m})$$

$$= B + \frac{e_{i_0}v'}{m} + N \cdot u(B(i_0) + \frac{v}{m})'$$

Thus for $i \neq i_0$ and $i \in H_+$,
Lemma 16:

\[ A_{(i)} = B_{(i)} + N \cdot (B_{(i_0)} + \frac{\omega}{m}) \]

But \( B_{(i_0)} + \frac{\omega}{m} \in \mathcal{C}_+ = C_{(i)} \). For sufficiently large \( N \), \( N \cdot (B_{(i_0)} + \frac{\omega}{m}) \) dominates \( B_{(i)} \), and hence \( A_{(i)} \in \mathcal{C}_+ = C_{(i)} \). Now, for \( i \neq i_0 \) and \( i \in H_\cdot \),

\[ A_{(i)} = B_{(i)} + N \cdot (-B_{(i_0)} - \frac{\omega}{m}) \]

But, \( -B_{(i_0)} - \frac{\omega}{m} = -(B_{(i_0)} + \frac{\omega}{m}) \in \mathcal{C}_- = C_{(i)} \). Again, for large \( N \), \( N \cdot (B_{(i_0)} + \frac{\omega}{m}) \) dominates \( B_{(i)} \), and hence \( A_{(i)} \in \mathcal{C}_- = C_{(i)} \). \( \square \)

The following lemma talks about the boundedness and convergence of the sequence of projections \( \{ P_{v(Y)}(m(Y)|D^{[N]}C_H) \} \).

**Lemma 16**: 

1. The sequence \( \{ P_{v(Y)}(m(Y)|D^{[N]}C_H) \} \) lies in a bounded set.

2. Every convergent subsequence of \( \{ P_{v(Y)}(m(Y)|D^{[N]}C_H) \} \) converges to \( P_{v(Y)}(m(Y)|C_{i_0}) \).

**Proof:**

1. Note that \( \| m(Y) \|^2_{v(Y)} = \| m(Y) - P_{v(Y)}(m(Y)|D^{[N]}C_H) \|^2_{v(Y)} + \| P_{v(Y)}(m(Y)|D^{[N]}C_H) \|^2_{v(Y)} \) since \( \langle m(Y) - P_{v(Y)}(m(Y)|D^{[N]}C_H), P_{v(Y)}(m(Y)|D^{[N]}C_H) \rangle = 0 \). Therefore

\[ \| P_{v(Y)}(m(Y)|D^{[N]}C_H) \|^2_{v(Y)} \leq \| m(Y) \|^2_{v(Y)} \]

Thus \( \{ P_{v(Y)}(m(Y)|D^{[N]}C_H) \} \) lies in a bounded set.

2. Every bounded sequence has a convergent subsequence. So, we assume \( P_{v(Y)}(m(Y)|D^{[N_k]}C_H) \to A^* \) as \( k \to \infty \). Then, by the sufficient and necessary conditions for projections onto closed convex cones, we need to show that

   (a) \( A^* \in C_{i_0} \), (b) \( \langle m(Y) - A^*, A^* \rangle_{v(Y)} = 0 \) and (c) \( \langle m(Y) - A^*, B \rangle_{v(Y)} \leq 0 \) for all \( B \in C_{i_0} \).

   (a) Since \( D^{[N]}C_H \subset C_i \), \( P_{v(Y)}(m(Y)|D^{[N_k]}C_H) \in C_{i_0} \). But \( C_{i_0} \) is closed. So \( A^* \in C_{i_0} \).

41
(b) Note that $\langle m(Y) - P_{v(Y)}(m(Y)|D^{[N_k]}C_H), P_{v(Y)}(m(Y)|D^{[N_k]}C_H)\rangle_{v(Y)} = 0$. Then, as $k \to \infty$, $P_{v(Y)}(m(Y)|D^{[N_k]}C_H) \to A^*$. Hence, (b) is true.

(c) For $B \in \mathcal{C}_{i_0}$, select $B^{[m]}$ and $N_m$ as in proposition 3. When $N_k > N_m$, by (b) of proposition 3, $\langle m(Y) - P_{v(Y)}(m(Y)|D^{[N_k]}C_H), B^{[m]}\rangle_{v(Y)} \leq 0$. As $k \to \infty, \langle m(Y) - A^*, B^{[m]}\rangle_{v(Y)} \leq 0$. Then, letting $m \to \infty$ and applying (a) of proposition 3, we get (c). □

Using the above lemma, we can establish a limit of the test statistic $T_H(Y)$.

**Lemma 17**: If $i_0 \in \mathcal{H}$, then $T_H(D^{[\infty]}Y) \to T_{i_0}(Y)$ as $N \to \infty$.

**Proof**: By Lemma 14,

$$T_H(D^{[\infty]}Y) = \|P_{v(Y)}(m(Y)|\mathcal{L}) - P_{v(Y)}(m(Y)|D^{[N]}C_H)\|^2_{v(Y)}.$$  

According to (2) of Lemma 16, as $N \to \infty$, $P_{v(Y)}(m(Y)|D^{[N]}C_H) \to P_{v(Y)}(m(Y)|\mathcal{C}_{i_0})$. Then, as $N \to \infty$,

$$T_H(D^{[\infty]}Y) \to \|P_{v(Y)}(m(Y)|\mathcal{L}) - P_{v(Y)}(m(Y)|\mathcal{C}_{i_0})\|^2_{v(Y)}$$

i.e., $T_H(D^{[\infty]}Y) \to T_{i_0}(Y)$ as $N \to \infty$. □
CHAPTER 6
THE DISTRIBUTIONS OF $T_H(Y)$

In this chapter, the distribution of our proposed test statistic $T_H(Y)$ is explored using the results obtained from the previous chapter. In the next section, we show that the distribution of $T_H(Y)$ is invariant to $\beta$ when $\beta \in \mathcal{L}$.

6.1 Condition for non-dependence on $\beta$

Before discussing about the distribution of $T_H(Y)$, we use the following notation in our discussion. With $Y \sim N(\beta M', \Sigma)$, the distribution of $T_H(Y)$ depends on $\beta$ and $\Sigma$, and we denote this by $T_H(\beta, \Sigma)$, i.e., $Y \sim N(\beta M', \Sigma) \Rightarrow T_H(Y) \sim T_H(\beta, \Sigma)$.

We now present the following lemma.

Lemma 18 : Let $T_H(\beta, \Sigma)$ be defined as our notation. Then

(1) If $\beta \in \mathcal{L}$, then $T_H(\beta, \Sigma) = T_H(0, \Sigma)$.

(2) $T_H(\beta, \Sigma) = T_H(D\beta, D\Sigma D')$ for all $\beta \in R^{p \times q}$ and all $\Sigma > 0$ if $D^{-1}C_H = C_H$.

(3) If $\phi \neq H_+ \subset H$, then $T_H(\beta, \Sigma)$ is stochastically less than or equal to $T_{H_+}(\beta, \Sigma)$

(4) For $i_0 \in H$, there exists $T_H(0, \Sigma^{[N]})$ and $T_{i_0}(0, I)$ such that $T_H(0, \Sigma^{[N]}) \rightarrow T_{i_0}(0, I)$ in distribution.
Proof:

(1) \( Y \sim N(\beta M', \Sigma) \) implies \( Y - \beta M' \sim N(0M', \Sigma) \). Thus, by our notation, \( T_H(Y) \sim T_H(\beta, \Sigma) \) and \( T_H(Y - \beta M') \sim T_H(0, \Sigma) \). But, by Lemma 13, \( T_H(Y - \beta M') = T_H(Y) \) if \( \beta \in \mathcal{L}. \) Thus, under this condition, the distributions \( T_H(\beta, \Sigma) \) and \( T_H(0, \Sigma) \) are identical.

(2) \( Y \sim N(\beta M', \Sigma) \) implies \( DY \sim N(D\beta M', D\Sigma D') \). Thus, by our notation, we have \( T_H(Y) \sim T_H(\beta, \Sigma) \) and \( T_H(DY) \sim T_H(D\beta, D\Sigma D') \). But, by Lemma 14, it is seen that \( T_H(Y) = T_H(DY) \) if \( D^{-1}C_H = C_H \). Thus, under this condition, the distributions \( T_H(\beta, \Sigma) \) and \( T_H(D\beta, D\Sigma D') \) are identical.

(3) \( Y \sim N(\beta M', \Sigma) \) implies that \( T_{H^*}(Y) \sim T_{H^*}(\beta, \Sigma) \) and \( T_H(Y) \sim T_H(\beta, \Sigma) \). But by Lemma 15, \( T_H(Y) \leq T_{H^*}(Y) \) if \( \phi \neq H_* \subset H \). Thus, under this condition, the distribution \( T_H(\beta, \Sigma) \) is stochastically less than or equal to the distribution \( T_{H^*}(\beta, \Sigma) \), that is,

\[
P(T_H(\beta, \Sigma) > t) \leq P(T_{H^*}(\beta, \Sigma) > t)
\]

for all \( t \geq 0 \).

(4) For \( i_0 \in H \), using the definition of \( D^{[N]} \) and by Lemma 17, \( T_H(D^{-[N]}Y) \rightarrow T_{i_0}(Y) \) with probability 1. Thus, it converges in distributions. But when \( Y \sim N(0M', I) \), \( D^{-[N]}Y \sim N(0M', \Sigma^{[N]}) \) with \( \Sigma^{[N]} = D^{-[N]}(D^{-[N]})' \).

Thus, \( T_H(D^{-[N]}Y) \sim T_H(0, \Sigma^{[N]}) \) and \( T_{i_0}(Y) \sim T_{i_0}(0, I) \). Consequently, \( T_H(0, \Sigma^{[N]}) \rightarrow T_{i_0}(0, I) \). \( \square \)

When \( \beta \in \mathcal{L} \), by (1) of Lemma 18, the distribution \( T_H(\beta, \Sigma) \) is invariant to \( \beta \), but still depends on \( H \) and \( \Sigma \). In the next section, we show that if \( H \) contains only one element,
then the distribution of \( T_H(Y) \) will no longer be \( \Sigma \) specific.

### 6.2 Condition for non-dependence on \( \Sigma \)

Suppose \( H = \{i\} \). Then, the following lemma is stated and proved.

**Lemma 19**: For a given \( i \in \{1, 2, \ldots, p\} \), there exists a non-singular \((p \times p)\) matrix \( D \) such that

1. \( D^{-1}C_i = C_i \)
2. \( D\Sigma D' = I \).

**Proof:**

1. Let \( G \) be the elementary matrix obtained by interchanging the \( i \)th row and \( p \)th row of \( I_p \). Then \( G \) is symmetric, non-singular and \( G^{-1} = G \). So, \( G\Sigma G' \) is a positive definite matrix and hence \((G\Sigma G')^{-1/2}\) exists. Then, by QR decomposition, we have

\[
(G\Sigma G')^{-1/2} = QR
\]

where \( Q \) is an orthogonal matrix and \( R = (r_{ij})_{p \times p} \) is an upper-triangular matrix with positive diagonal elements.

Let \( D = GRG \), where \( R \) is the matrix obtained by the QR decomposition of \((G\Sigma G')^{-1/2}\).

Then,

\[
D^{-1} = GR^{-1}G
\]

where \( R^{-1} \) is also an upper triangular matrix with positive diagonal elements.

Suppose \( Y = D^{-1}X \), where \( X = (X_{(1)}, X_{(2)}, \ldots, X_{(p)})' \), \( Y = (Y_{(1)}, Y_{(2)}, \ldots, Y_{(p)})' \).
Then

\[ Y = D^{-1} X \]

\[ = GR^{-1} GX \]

\[ = GR^{-1}(X_{(1)}, X_{(2)}, \ldots, X_{(p)}, \ldots, X_{(i)})' \]

\[ = G \left( \frac{1}{r_{11}} X_{(1)} \frac{1}{r_{22}} X_{(2)} \cdots \frac{1}{r_{ii}} X_{(i)} \cdots \frac{1}{r_{pp}} X_{(p)} \right)' \]

\[ = \left( \frac{1}{r_{11}} X_{(1)} \frac{1}{r_{22}} X_{(2)} \cdots \frac{1}{r_{pp}} X_{(i)} \cdots \frac{1}{r_{ii}} X_{(p)} \right)' \]

So, we have

\[ Y_{(i)} = \frac{1}{r_{pp}} X_{(i)} \]  \hspace{1cm} (6.1)

Now, using the definition of \( Y \) and equation (6.1),

\[ Y \in C_i \]

\[ \iff Y_{(i)} \in C_{(i)} \]

\[ \iff \frac{1}{r_{pp}} X_{(i)} \in C_{(i)} \]

\[ \iff X_{(i)} \in C_{(i)} \]

\[ \iff X \in C_i \]

Thus \( D^{-1} X \in C_i \iff X \in C_i \). Hence \( D^{-1} C_i = C_i \).

(2) With \( D \) defined in the proof of (1) of this lemma and by the fact that \( Q \) is orthogonal,
\[ D \Sigma D' = (GRG) \Sigma (GRG)' \]
\[ = (GR)(G \Sigma G)(R'G) \]
\[ = (GQ')(QR)(G \Sigma G)(QR)'QG \]
\[ = (GQ')(G \Sigma G)^{-1/2}(G \Sigma G)(G \Sigma G)^{-1/2}QG. \]
\[ = GQ'QG = I_p \]

Thus, the lemma is proved. □

In the next lemma, it is shown that \( T_i(0, \Sigma) \) is invariant to \( \Sigma \).

Lemma 20 : \( T_i(0, \Sigma) = T_i(0, I) \) for all \( \Sigma > 0 \) and all \( i = 1, 2, \ldots, p \).

Proof : Using (2) of lemma 18, we have \( T_H(\beta, \Sigma) = T_H(D\beta, D \Sigma D') \) for all \( \beta \in \mathbb{R}^{p \times q} \) and all \( \Sigma > 0 \) if \( D^{-1}C_H = C_H \). But from Lemma 19, there exists a matrix \( D \) such that (i) \( D \Sigma D' = I \) and (ii) \( D^{-1}C_i = C_i \). Therefore, \( T_i(\beta, \Sigma) = T_i(D\beta, D \Sigma D') = T_i(D\beta, I) \) for all \( \beta \in \mathbb{R}^{p \times q} \). Let \( \beta = 0 \). Then the conclusion follows. □.

6.3 Two theorems

We now establish two theorems directly related to the properties of the tests that will be stated in the next chapter.

Theorem 1 : \( \sup\{P(T_H(\beta, \Sigma) > t) : \beta \in \mathcal{L}, \Sigma > 0\} = P(T_i(0, I) > t) \) for all \( i \) in \( \{1, 2, \ldots, p\} \), all non-empty \( H \), and all \( t \geq 0 \).

Proof: If the theorem is true for \( i \in H \), then with \( H = \{1, 2, \ldots, p\} \) it can be seen that \( T_i(0, I) = T_j(0, I) \) for all \( i = 1, 2, \ldots, p \) and \( j = 1, 2, \ldots, p \). Consequently, the theorem holds for all \( i \).
Suppose $i \in H$. Then, by (1) of Lemma 18, (3) of Lemma 18 and Lemma 20,

$$
\sup \left[ P(T_H(\beta, \Sigma) > t) : \beta \in \mathcal{L}, \Sigma > 0 \right]
= \sup \left[ P(T_H(0, \Sigma) > t) : \Sigma > 0 \right]
\leq \sup \left[ P(T_i(0, \Sigma) > t) : \Sigma > 0 \right]
= P(T_i(0, I) > t)
$$

On the other hand, using (1) of Lemma 18, we get

$$
\sup \left[ P(T_H(\beta, \Sigma) > t) : \beta \in \mathcal{L}, \Sigma > 0 \right]
= \sup \left[ P(T_H(0, \Sigma) > t) : \Sigma > 0 \right]
\geq P(T_H(0, \Sigma^{[N]}) > t)
$$

where $\Sigma^{[N]}$ are given as (4) of Lemma 18, and also, from (4) of Lemma 18,

$$
P(T_H(0, \Sigma^{[N]}) > t) \rightarrow P(T_i(0, I) > t)
$$

Hence

$$
\sup \left[ P(T_H(\beta, \Sigma) > t) : \beta \in \mathcal{L}, \Sigma > 0 \right] \geq P(T_i(0, I) > t)
$$

Combining the results, the theorem is proved. $\square$.

Let $T_*(Y)$ be a statistic with the distribution denoted by $T_*(\beta, \Sigma)$ when $Y \sim N(\beta M', \Sigma)$. For this distribution, we have the following theorem.

**Theorem 2** : If $T_H(\beta, \Sigma) \leq T_*(\beta, \Sigma) \leq T_{i_0}(\beta, \Sigma)$ stochastically for some $i_0 \in \{1, 2, \ldots, p\}$ and some non-empty $H$, then

$$
\sup \left[ P(T_*(\beta, \Sigma) > t) : \beta \in \mathcal{L}, \Sigma > 0 \right] = P(T_i(0, I) > t)
$$
for all \( i \in \{1, 2, \ldots, p\} \) and all \( t \geq 0 \).

**Proof:** The condition of this theorem implies that

\[
P(T_H(\beta, \Sigma) > t) \leq P(T_*(\beta, \Sigma) > t) \leq P(T_0(\beta, \Sigma) > t)
\]

for all \( t \geq 0 \). Thus

\[
\sup\{P(T_H(\beta, \Sigma) > t) : \beta \in \mathcal{L}, \Sigma > 0\} \\
\leq \sup\{P(T_*(\beta, \Sigma) > t) : \beta \in \mathcal{L}, \Sigma > 0\} \\
\leq \sup\{P(T_0(\beta, \Sigma) > t) : \beta \in \mathcal{L}, \Sigma > 0\}
\]

But, by Theorem 1, for all \( i \),

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

and

\[
\sup\{P(T_0(\beta, \Sigma) > t) : \beta \in \mathcal{L}, \Sigma > 0\} = P(T_i(0, I) > t)
\]

Thus,

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

\( \Box \).
CHAPTER 7
MAIN RESULTS

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

7.1 $\alpha$-level critical values

With a given $H$, the test has $\alpha$-level critical value $t_\alpha$ determined by

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

where $\alpha$ is the significance level of the test. By Theorem 1, the above 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 $\alpha$-level critical values, and consequently, all $\alpha$-level tests share the same rejection regions.

The complexity of the computation for the $\alpha$-level critical values is greatly reduced by using equation (7.2) since it does not involve unknown $\beta \in \mathcal{L}$ and unknown $\Sigma > 0$.

7.2 p-values

Let $t_{ob}$ be the observed value of $T_H(Y)$. Then, the p-value, or the observed significance level is, by definition, given as

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

By Theorem 1, this equation is equivalent to

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

50
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 $T_i(0, I)$.

7.3 More powerful tests

Let $T_\ast(Y)$ be a test statistic satisfying the conditions in Theorem 2 for a given $H$ for testing $H_0 : \beta \in \mathcal{L}$ against $H_1 : \beta \in \mathcal{C}_H$. We reject the null hypothesis when $T_\ast(Y) > t$. For convenience, this test scheme is denoted as Test II and the one which uses the test statistic $T_H(Y)$ as Test I. When $Y \sim N(\beta M', \Sigma)$, the distribution of $T_\ast(Y)$ is denoted by $T_\ast(\beta, \Sigma)$.

The next theorem gives a sufficient condition for Test II to have the same properties as that of Test I in terms of $\alpha$-level critical values and p-values.

**Theorem 3**: If $T_H(\beta, \Sigma) \leq T_\ast(\beta, \Sigma) \leq T_i(\beta, \Sigma)$ stochastically for some $i$ in \{1, 2, \ldots, $p$\}, then the followings are true.

1. Test I and Test II have the same $\alpha$-level critical values.

2. Let $t_\ast$ be the observed value of $T_\ast(Y)$. Then the p-value of Test II is

   $$P(T_i(0, I) > t_\ast) \text{ for } i \in \{1, 2, \ldots, p\}$$

3. At the same $\alpha$-level, Test II is more powerful than Test I.
Proof:

(1) From Theorem 2, we have

\[ \sup \left[ P(T_H(\beta, \Sigma) > t) : \beta \in L, \Sigma > 0 \right] = P(T_i(0, I) > t). \]

But, by Theorem 1, \( \sup \left[ P(T_H(\beta, \Sigma) > t) : \beta \in L, \Sigma > 0 \right] = P(T_i(0, I) > t) \). Also, by combining equations (7.1) and (7.2),

\[ \alpha = \sup \left[ P(T_H(\beta, \Sigma) > t) : \beta \in L, \Sigma > 0 \right] = P(T_i(0, I) > t_\alpha). \]

Hence, \( \alpha = \sup \left[ P(T_i(\beta, \Sigma) > t_\alpha) : \beta \in L, \Sigma > 0 \right] \) is true if and only if \( \alpha = \sup \left[ P(T_H(\beta, \Sigma) > t_\alpha) : \beta \in L, \Sigma > 0 \right] \) is true, i.e., Test I and Test II have the same \( \alpha \)-level critical value \( t_\alpha \).

(2) When \( t \) is replaced with \( t_* \), we have

\[ \sup \left[ P(T_*(\beta, \Sigma) > t_*) : \beta \in L, \Sigma > 0 \right] = P(T_i(0, I) > t_*) \]

which is the p-value for Test II.

(3) Under the condition of this lemma, we have

\[ \sup \left[ P(T_H(\beta, \Sigma) > t_\alpha) : \Sigma > 0 \right] \leq \sup \left[ P(T_*(\beta, \Sigma) > t_\alpha) : \Sigma > 0 \right] \]

i.e., the power of Test II at \( \beta \) is no less than the power of Test I at \( \beta \). Hence, by definition of More Powerful tests, we can say that at the same \( \alpha \) level, Test II is more powerful than Test I. \( \square \)

We now give two examples of such \( T_*(Y) \) as follows.

**Example 1**: Let \( i_0 \in H \). Define
\[ T_*(Y) = T_{i_0}(Y) \]

From Lemma 15, for any non-empty subset \( H_* \) of \( H \), \( T_H(Y) \leq T_{H_*}(Y) \). Then, one can choose \( T_*(Y) = T_{H_*}(Y) = T_{i_0}(Y) \), then, \( T_{i_0}(Y) \) satisfies the conditions in Theorem 2. Thus, \( \alpha \)-level test using \( T_{i_0}(Y) \) is more powerful than that using \( T_H(Y) \).

In constructing \( T_H(Y) \), each rows under the restriction made their contributions, which, intuitively, is good. But Example 1 uncovers that this test is less powerful than the one that uses a statistic constructed with the input from only one row. That row may not even be under the restriction according to the model specification. So, a question can be raised if there is a test statistic that takes input from each rows under the restriction, the test using this test statistic has properties for \( \alpha \)-level critical values and observed significance levels just as each members in \( T_H \) family, and the test is more powerful than that using \( T_H \). The question is answered in the next example.

**Example 2:** Define

\[ T_*(Y) = \min\{T_i(Y) : i \in H\}. \tag{7.3} \]

Then, by Lemma 15, \( T_H(Y) \leq T_*(Y) \). Again

\[ \min\{T_i(Y) : i \in H\} \leq T_i(Y) \text{ for } i \in H \]

\[ \Rightarrow T_*(Y) \leq T_i(Y) \text{ for } i \in H \]

Thus, \( T_H(Y) \leq T_*(Y) \leq T_i(Y) \) for \( i \in H \) and hence \( T_H(\beta, \Sigma) \leq T_*(\beta, \Sigma) \leq T_i(\beta, \Sigma) \) for \( i \in H \). Then, the condition in Theorem 2 is satisfied for \( i \in H \). Thus \( \alpha \)-level test using equation (7.3) is more powerful than that using \( T_H(Y) \).
CHAPTER 8

CONCLUSIONS

In this research, a parameter matrix $\beta \in \mathbb{R}^{p \times q}$ is considered. Some rows of the parameter matrix are under an ordering while some other rows of the matrix are under an opposite ordering. Also, $\beta$ may have some rows which are unrestricted, that is, these rows are free of any ordering. A set $H$ is constructed which is a subset of the row index space of the parameter matrix. $H$ is the index of those rows of $\beta$ which are under any ordering. Thus, with $H = \{1, 3, 4\}$, it means that the first, third and fourth row of $\beta$ are under some order restrictions. Since the order restrictions are based on the same partial order $ll$, they are termed as synchronized order restrictions. Under the null hypothesis that the parameter matrix $\beta$ is homogeneous, a testing problem is considered where the alternative hypothesis states that some of the rows of $\beta$ are under synchronized orderings.

When the common variance-covariance matrix $\Sigma$ is known, a LRT statistic can be obtained using the LRT procedure. But when $\Sigma$ is unknown, an ad-hoc test statistic is obtained by replacing $\Sigma$ with its estimator in the LRT statistic. This ad hoc test statistic is denoted by $T_H(Y)$, where $Y$ is the data matrix in $\mathbb{R}^{p \times n}$. Considering different possibilities of $H$, a family of ad-hoc test statistics can be obtained.

The deterministic and probabilistic properties of $T_H(Y)$ have been studied. It has been shown that $T_H(Y)$ is invariant under a linear transformation of $Y$. A non-empty subset $H_*$ of $H$ has been considered and it has been shown that the values of $T_H(Y)$ is less than or equal to that of $T_{H_*}(Y)$. Studying the probabilistic properties, it has been proved that the null distribution of $T_H(Y)$ is free of $\beta$, and when $H$ contains only one element, the distribution of $T_H(Y)$ does not depend on $\Sigma$ as well.
The main results about the test has been obtained from Theorem 1 and Theorem 2. From Theorem 1, it has been shown that all tests in the ad hoc test family share the same $\alpha$-level critical values. The expression in equation (7.2) does not involve any unknown parameters and hence the computation for the $\alpha$-level critical values is greatly reduced by using equation (7.2).

It has also been shown that all tests in the test family follow the same distribution $T_i(0, I)$. Thus, the computation of p-value is easier due to the fact that the distribution $T_i(0, I)$ does not involve any unknown parameters. Hence estimation of p-values can be done using the Monte Carlo method by simulating observations from $T_i(0, I)$.

A sufficient condition has been given in Theorem 2 for a test scheme to be more powerful than the test scheme that uses the proposed test statistic $T_H(Y)$ at the same $\alpha$-level. A test statistic which satisfies the condition of Theorem 2 is denoted by $T_*(Y)$. Two examples of $T_*(Y)$ has been provided which are more powerful than $T_H(Y)$. Between them, the second example in equation (7.3) is intuitively better than the first example as it takes input from all the restricted rows of the parameter matrix.
REFERENCES
LIST OF REFERENCES


