A PSEUDO RESTRICTED MAXIMUM LIKELIHOOD ESTIMATOR UNDER MULTIVARIATE SIMPLE TREE ORDER RESTRICTION AND AN ALGORITHM

A Dissertation by

Huruy Debessay Asfha

Master of Science, Anadolu University, 2017

Bachelor of Arts, University of Asmara, 2007

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

July 2021

© Copyright 2021 by Huruy Debessay Asfha

All Rights Reserved

A PSEUDO RESTRICTED MAXIMUM LIKELIHOOD ESTIMATOR UNDER MULTIVARIATE SIMPLE TREE ORDER RESTRICTION AND AN ALGORITHM

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

Chunsheng Ma , Committee Member

Adam Jaeger , Committee Member

Ziqi Sun , Committee Member

Hyuck Kwon , Committee Member

Accepted for the College of Liberal Arts and Sciences

Andrew Hippisley, Dean

Accepted for the Graduate School

Coleen Pugh, Dean

DEDICATION

This dissertation is dedicated to my father Debessay Asfha, my mother Hidat Berhe, my grandmother Demet Bairu and my lovely wife Hermon Ghebresilassie Belay who have always been there for me.

ACKNOWLEDGEMENTS

It is an authentic delight to offer my profound gratitude to my advisor, Dr. Xiaomi Hu, for his advice, patience, and guidance toward my Ph.D. His commitment and staggering disposition to help his students have been solely and essentially responsible for finishing my work. His practical exhortation, diligent scrutiny, and informed advice have helped me to an incredible degree to achieve this task. I will forever be indebted to him.

I owe a profound feeling of appreciation to Dr.Chunsheng Ma for his guidance and advice during the time I had classes with him. He has been an inspirational professor to me. Moreover, I would like to express my deepest gratitude to Dr. Thalia Jeffres for her motherly and scholarly advice and encouragement during the difficult times I had during my graduate studies.

I abundantly thank my committee members Dr. Ziqi Sun, Dr. Adam Jaeger, and Dr. Hyuck Kwon from the department of Electrical Engineering, for their constructive suggestions and timely response throughout my examination period.

Furthermore, I would like to extend my gratitude to all the faculty and administrative staff of the Department of Mathematics and Statistics for making my time at Wichita State University a wonderful experience. I am also grateful to all my colleagues and friends in WSU, especially Sirvan Rahmati, his son Hebey, and Nam Nguyen, for their intellectual ideas and discussions.

I am incredibly appreciative of my parents, grandmother, and siblings for their unconditional love and endless support, without whom this journey would have been impossible. Their love and support have been the driving source of energy for me to complete this journey. I also would like to acknowledge my uncle, Dr. Tesfa Mehari of the University of Greenwich, UK, who supported me during the admission process to WSU.

Last but not least, I am privileged to thank my beautiful wife Mrs. Hermon Ghebreslassie Belay for her understanding, patience, and continuous support. Despite the hardship she has been through, she constantly encouraged me to strive to achieve my dreams. Her unconditional love never faded away, notwithstanding the physical distance between us.

ABSTRACT

The minimum distance projection of a given matrix $X \in R^{p\times q}$ onto the order restricted cone in an appropriately defined inner product system, $\pi(X|C_{p\times q})$, plays an important role in order restricted statistical inference since in many cases the restricted maximum likelihood estimator (RMLE) for a parameter matrix under an order restriction is the projection of the maximum likelihood estimator (MLE) without any restrictions onto the order restricted cone. The RMLE plays an important part in the maximum likelihood ratio tests. The computation for $\pi(X|_{p\times q})$ however is currently a great challenge to researchers.

It is known that the order relation $\preceq$ in R^p is a multivariate order relation if and only if it is generated from a closed convex cone $C \in R^p$, called an order generating cone. The collection of all matrices $\mu = (\mu_1, ..., \mu_q) \in R^{p\times q}$ whose columns satisfy the multivariate order restriction $\mu_i \preceq \mu_j$ for all (i, j) in a specified set $H \subset \{1, ..., q\} \times \{1, ..., q\}$ is a closed convex cone $C_{p\times q}$ in $R^{p\times q}$, called an order restricted cone. For $C_{p\times q}$ created by multivariate simple-tree order restriction and a given matrix $X \in R^{p\times q}$, in this dissertation, a closed convex subset $D(X)_{p\times q} \subset C_{p\times q}$ is defined. The projection of X onto this subset, $\pi(X|D(X)_{p\times q})$, is studied. In addition, an algorithm for computing $\pi(X|D(X)_{p\times q})$ is proposed and proved.

The proposed algorithm for $\pi(X|D(X)_{p\times q})$ only depends on projections of vectors onto the order generating cone. Thus, it converts the relatively difficult matrix projection problem to a much easier vector projection problems. It is also revealed that when $q = 2$, $\pi(X|D(X)_{p\times q}) = \pi(X|C_{p\times q})$, and if $X \in C_{p\times q}$, then $\pi(X|D(X)_{p\times q}) = \pi(X|C_{p\times q})$. With all these good properties we could treat the projection onto $D(X)_{p\times q}$ as the approximation of the projection onto $C_{p\times q}$.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Order restrictions on model parameters appear in many statistical problems. Statistical tests that do not use the available information regarding the order restriction usually fail to be powerful. On the other hand, considering any additional information on the parameter of interest improves the power of the test. When comparing the means from two independent normal populations with the same variance, if information is available regarding the order of the two means, a one-sided t-test is uniformly more powerful than a two-sided t-test.

The area of order restricted statistical inference date back to the early 1950s. It was developed rapidly during the 1960s and the early 1970s. In testing homogeneity of several univariate normal means, Bartholomew (1959) and Bartholomew (1961) considered an alternative hypothesis where all μ_i's or some are under order restriction, assuming that the variances are known. They showed that the generalized test, $\bar{\chi}^2$ and $\bar{E}^2$, they proposed happens to be more potent than the ordinary χ^2 and $\bar{E}^2$ test which do not assume prior information on the order of the means. In the literature, many have discussed that taking into consideration preliminary information often results in a more robust test. Nevertheless, in practice it is common that the population variances may not be known. So, Bartholomew (1961) extended the work in (Bartholomew, 1959) to testing homogeneity of normal population means against an order restricted alternative hypothesis when the variances are unknown.

Summary of the developments in the 1960s and 1970s is well documented in Brunk et al. (1972) and is used as a basis for researchers in the field of order restricted statistical inference since then. The first conference on the area of order restricted statistical inference was held in 1981 and the second four years later. Fourteen of the presentation from the second conference were compiled and published in Dykstra et al. (2012).

Test statistics such as Bartlett (M), Hartley (F_{max}) and Cochran (G) have been already investigated in the 1950s and before to test homogeneity of several variances of normal populations against unordered alternative hypothesis. Fujino (1979) introduced a generalization of test for homogeneity of several variances of normal populations against order restricted alternative hypothesis. As expected, their investigation shows that taking into consideration information available about the order of the variances produces a superior test.

For testing the hypothesis $H_0 : \mu_1 = ... = \mu_q$ vs $H_1 : \mu_1 \leq ... \leq \mu_q$ where $\mu = (\mu_1, ..., \mu_q) = (\mu_{ij})_{p\times q}$ when the covariance matrices are known, Sasabuchi et al. (1983) provided an extension of the work in Bartholomew (1959). Sasabuchi et al. (2003) generalized these methods to include cases when the covariance matrices are unknown but common. The restriction $\mu_j \leq \mu_{j+1}$ means $\mu_{ij} \leq \mu_{ij+1}$ for all $i = 1, ..., p$ and $j = 1, ..., q-1$. Hu (2012) extended the study by introducing a vector quasi order " $\preceq$ " which is defined as $\mu_{ij} \leq \mu_{ij+1}$ for $i \in D_1$, $\mu_{ij} \geq \mu_{ij+1}$ for $i \in D_2$ and $\mu_{ij} = \mu_{ij+1}$ for $i \in D_3$ where D_1, D_2 and D_3 are prior defined disjoint subsets of $\{1, ..., p\}$. In practice, it is also of interest to test a hypothesis when an order restriction is involved in the null hypothesis. Silvapulle and Sen (2005) presented a brief detail of such tests.

In the study of order restricted statistical inference, one of the main challenges is computing the isotonic regression which means computing an estimator of a parameter under order restriction. In the univariate case, there are numerous algorithms developed through the years. The pool-adjacent-violator algorithm (PAVA) for example is well-known method mainly for computing isotonic regression associated with simple ordering. The Merge and Chop Algorithm (MCA) is also an alternative method for computing a univariate isotonic regression. Sasabuchi et al. (1992) introduced an algorithm to compute the isotonic regression in a univariate cases, and presented a multivariate extension in Sasabuchi et al. (2003). Furthermore, Geng and Shi (1991) proposed two algorithms to compute an isotonic regression under umbrella ordering in two independent variables.

The most widely used is the restricted maximum likelihood method. In an attempt to compute multivariate isotonic regression, Hu (2020) proposed an algorithm for obtaining a pseudo restricted maximum likelihood estimator when the mean matrix is restricted under multivariate simple ordering.

The choice of weights and order restriction corresponds to different isotonic regressions (Silvapulle and Sen, 2005) . Hence, the availability of different order restrictions makes the computation of isotonic regression more challenging as compared to the ordinary maximum likelihood estimation method.

This dissertation is organized as follows. In chapter 2, the concept of closed convex cone and projections is presented. In chapter 3, a multivariate order restriction is introduced. In addition, an order restricted cone, and an order induced cones are discussed. In chapter 4, a restricted maximum likelihood estimator (RMLE) in an order restricted MANOVA model is presented. In chapter 5, we present the main work of this dissertation; a pseudo RMLE is drived and an algorithm is proposed. Besides, we will discuss the conclusions and future work in chapter 6 and chapter 7 respectively.

CHAPTER 2

Closed convex cone and projections

Usually optimization is about maximization or minimization. In economics, minimizing a cost function and maximizing a profit function, and in statistics, minimizing a loss function and maximizing a likelihood function are examples of optimization objectives. Convex optimization which can be considered as a generalization of linear programming as discussed in Boyd et al. (2004) , has wide range of applications since many practical problems can be expressed in such form.

In this chapter, we present some important concepts of order restriction in relation to closed convex cone and projection onto closed convex cones in R^p.

2.1 Order relation

For a given set X, the binary relation "$\preceq$" on the elements of X is called a quasi order if it is

(1). reflexive: $x \preceq x$ for all $x \in X$, and

(2). transitive: for $x, y, z \in X$, $x \preceq y$, and $y \preceq z \Rightarrow x \preceq z$.

The relations " $\leq$ ", " $\geq$ ", " $\geq$ or $\leq$ ", and " $=$ " are all quasi orders on the set of real numbers. Without loss of generality, in this dissertation we will only use "$\leq$" to represent a quasi order on the elements of the set of real numbers.

Other two important properties of "$\leq$" are:

(1). The quasi order "$\leq$" is closed under linear combinations with non-negative coefficients, i.e.

for $x_1, x_2, y_1, y_2 \in R$ and $\alpha, \beta \geq 0$,

$$x_1 \leq y_1 \text{ and } x_2 \leq y_2 \Rightarrow \alpha x_1 + \beta x_2 \leq \alpha y_1 + \beta y_2.$$

(2). "$\leq$" is closed under limits i.e for $x_n, y_n, x, y \in R$,

$$x_n \leq y_n,\;\; x_n \to x \text{ and } y_n \to y \Rightarrow x \leq y.$$

A vector $x = (x_1, ..., x_p)' \in R^p$, is said to be order restricted if $x_i \leq x_j$ for some $(i, j) \in H$ where $H \subset \{1, ..., p\} \times \{1, ..., p\}$, and a function that takes such vector as an argument is said to be under order restriction.

An order restriction often appears in comparing parameters from two or more populations. Consider a test of homogeneity of means from k normal populations.

$$H_0 : \mu_1 = ... = \mu_k \text{ versus } H_1 : \mu_1 \leq ... \leq \mu_k.$$

Under H_1, $\mu = (\mu_1, ..., \mu_k)'$ is under an order restriction.

2.2 Closed convex cone

Definition 2.2.1.

1. A set C in a linear space $\mathcal{V}$ is said to be convex if

$$x_1, x_2 \in C \Rightarrow \alpha x_1 + (1 - \alpha)x_2 \in C \text{ for all } \alpha \in (0, 1). \tag{2.1}$$

2. A set $C \subset \mathcal{V}$, where $\mathcal{V}$ is a finite dimensional linear space $\mathcal{V}$, is said to be closed with respect to a norm induced from an inner product if

$$x_n \in C \text{ and } x_n \to x \Rightarrow x \in C. \tag{2.2}$$

3. A set C in a linear space $\mathcal{V}$ is called a cone if

$$x \in C \Rightarrow \alpha x \in C \text{ for all } \alpha > 0. \tag{2.3}$$

A set that satisfies all three is called a closed convex cone.

Figure 2.1 is a geometrical representation of a convex cone in R^2.

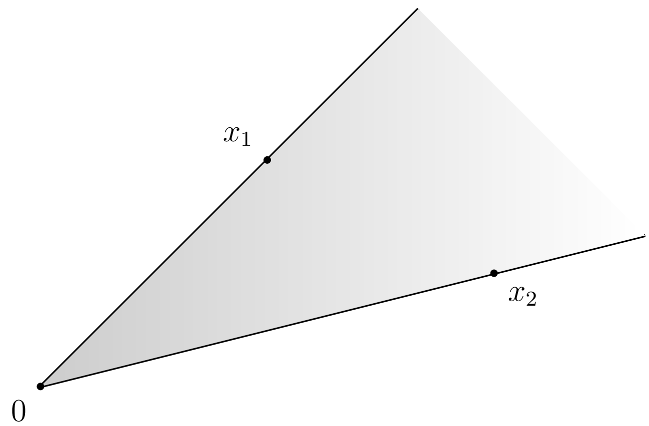

Figure 2.1: Geometrical representation of a closed convex cone in R^2

Lemma 2.2.2. *A set C in a linear space $\mathcal{V}$ is convex cone if and only if*

$$x_1, x_2 \in C \Rightarrow \alpha x_1 + \beta x_2 \in C \textit{ for all } \alpha, \beta > 0. \tag{2.4}$$

Proof. Suppose C is a convex cone. Then, by definition of a cone we have

$$x_1, x_2 \in C \Rightarrow 2\alpha x_1, 2\beta x_2 \in C \text{ for all } \alpha, \beta > 0$$

and by definition of a convex set we have

$$\frac{1}{2}(2\alpha x_1) + \frac{1}{2}(2\beta x_2) = \alpha x_1 + \beta x_2 \in C.$$

Suppose,

$$x_1, x_2 \in C \Rightarrow \alpha x_1 + \beta x_2 \in C \text{ for all } \alpha, \beta > 0.$$

Then, for $x \in C$ and $\gamma > 0$,

$$\gamma x = \frac{\gamma}{2}x + \frac{\gamma}{2}x \in C.$$

So, C is a cone.

Moreover, for $x_1, x_2 \in C$ and $\alpha \in (0,1)$ let $\beta = (1-\alpha) > 0$. Then,

$$\alpha x_1 + (1-\alpha)x_2 = \alpha x1 + \beta x_2 \in C.$$

So, C is a convex set. □

Clearly, for a closed convex cone C, when $x_1, x_2 \in C$, $\alpha x_1 + \beta x_2 \in C$ for all $\alpha \geq 0$ and $\beta \geq 0$.

2.3 Projection onto a closed convex cone

Let D be a closed convex set in a Hilbert space $\mathcal{H}$, $z \in \mathcal{H}$ be a given vector. Then a function defined as $f(x) = \|x - z\|^2$, where $x \in D$, is said to be under the restriction of $x \in D$. Under such restrictions, the function $f(x)$ is minimized at $z^* \in \mathcal{H}$.

Definition 2.3.1. For $z \in \mathcal{H}$, there exists a unique $z^* \in D$ such that $\|z^* - z\| \leq \|x - z\|$ for all $x \in D$. This z^* is called the minimum distance projection of z onto D, or simply a projection of z onto D denoted by $\pi(z|D)$.

The following lemma presents a sufficient and necessary condition for the projection onto a closed convex set.

Lemma 2.3.2. *Suppose $D \subset \mathcal{H}$ is a closed convex set and z is a given vector in $\mathcal{H}$. Then,*

$$z^* = \pi(x|D) \Leftrightarrow z^* \in D \text{ and } \langle z - z^*, z^* - y\rangle \geq 0 \text{ for all } y \in D \tag{2.5}$$

Proof. Suppose $z^* = \pi(z|D)$. Then, $z^* \in D$. For $y \in D$, $\alpha y + (1-\alpha)z^* \in D$, and

$$\|z - z^*\|^2 \leq \|z - [\alpha y + (1-\alpha)z^*]\|^2 = \|z - z^* + \alpha(z^* - y)\|^2 \ \forall y \in D \text{ and } \forall \alpha \in (0,1).$$

So,

$$0 \leq \alpha^2\|z^* - y\|^2 + 2\alpha\langle z - z^*, z^* - y\rangle$$

and hence,

$$\langle z - z^*, z^* - y \rangle \geq -\frac{\alpha}{2} \|z^* - y\|^2.$$

Since, $\alpha \in (0, 1)$, by letting $\alpha \to 0$, we have

$$\langle z - z^*, z^* - y \rangle \geq 0.$$

To show the "if" part, let $z^* \in D$ and $\langle z - z^*, z^* - y \rangle \geq 0$ for all $y \in D$. Then,

$$\begin{aligned} \|z - y\|^2 &= \|(z - z^*) + (z^* - y)\|^2 \\ &= \|z - z^*\|^2 + \|z^* - y\|^2 + 2\langle z - z^*, z^* - y \rangle \\ &\geq \|z - z^*\|^2 \ \forall y \in D. \end{aligned}$$

Thus, by definition of projection, $z^* = \pi(z|D)$. □

Since a cone is a special set, lemma 2.3.2 can be extended into that for a closed convex cone.

Lemma 2.3.3. *Let C be a closed convex cone. The projection of z onto C, denoted by $\pi(z|C)$, exists and is unique. Moreover,*

$$z^* = \pi(z|C) \Leftrightarrow z^* \in C, \langle z - z^*, z^* \rangle = 0 \textit{ and } \langle z - z^*, y \rangle \leq 0 \textit{ for all } y \in C.$$

Proof. Suppose $z^* = \pi(z|C)$. Then $z^* \in C$. With $y = 0 \in C$, by lemma 2.3.2,

$$\langle z - z^*, z^* - 0 \rangle \geq 0 \tag{2.6}$$

and with $y = 2z^* \in C$, by lemma 2.3.2,

$$0 \leq \langle z - z^*, z^* - 2z^* \rangle = -\langle z - z^*, z^* \rangle \tag{2.7}$$

So, by combining (2.6) and (2.7), we have $\langle z - z^*, z^* \rangle = 0$. Consequently,

$$0 \leq \langle z - z^*, z^* - y \rangle = \langle z - z^*, z^* \rangle - \langle z - z^*, y \rangle \text{ for all } y \in C$$

$$= -\langle z - z^*, y \rangle \text{ for all } y \in C$$

Thus, $\langle z - z^*, y \rangle \leq 0$ for all $y \in C$.

Now suppose $z^* \in C$, $\langle z - z^*, z^* \rangle = 0$ and $\langle z - z^*, y \rangle \leq 0$ for all $y \in C$.

$$\langle z - z^*, z^* - y \rangle = \langle z - z^*, z^* \rangle - \langle z - z^*, y \rangle$$

$$= -\langle z - z^*, y \rangle \geq 0.$$

So, by lemma 2.3.2, $z^* = \pi(z|C)$. □

CHAPTER 3

Multivariate order restriction and order restricted cone

3.1 Multivariate order restriction

In many applications, there is an encounter of large data with multiple variables. In such cases, parameters are represented in vector form. There has been efforts to describe comparison of two vectors. For example, Sasabuchi et al. (2003) investigated a test on the homogeneity of mean vectors against $H_1 : \mu_1 \preceq ... \preceq \mu_q$ where $\mu_i \in R^p$ for all $i = 1, ..., q$ and $\mu_i \preceq \mu_j$ means all the components of $\mu_j - \mu_i$ are non-negative. Here, $\preceq$ is an order on vectors.

Definition 3.1.1. With respect to a properly defined inner product induced norm, the relation "$\preceq$" of vectors in R^p is called a multivariate order if it is

(1). reflexive: for $x \in R^p$, $x \preceq x$,

(2). transitive: for $x, y, z \in R^p$, $x \preceq y$ and $y \preceq z \Rightarrow x \preceq z$,

(3). preserved under linear combinations with non-negative coefficients:

for $x_1, y_1, x_2, y_2 \in R^p$ and $\alpha, \beta \geq 0$

$$x_1 \preceq y_1 \text{ and } x_2 \preceq y_2 \Rightarrow \alpha x_1 + \beta x_2 \preceq \alpha y_1 + \beta y_2,$$

(4). closed under limits:

for a sequences $x_n, y_n \in R^p$ and $x, y \in R^p$, $x_n \preceq y_n$, $x_n \to x$ and $y_n \to y \Rightarrow x \preceq y$. Here, the convergence is with respect to a norm induced from an inner product and hence, it is componentwise.

A multivariate order relation covers a diversified situations in the literature. For example, Hu and Banerjee (2012) defined a multivariate order for vectors $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ and $y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$ as $x \preceq y$ if $x_1 \le y_1$, $x_2 = y_2$ and $x_3 \ge y_3$.

The following two lemmas present the relationship between a multivariate order and a closed convex cone.

Lemma 3.1.2. *Let C be a closed convex cone in Hilbert space R^p. For $x, y \in R^p$ define a relation $x \preceq y$ if $y - x \in C$. Then, "$\preceq$" is a multivariate order.*

Proof. We need to show that "$\preceq$" satisfies the four properties of a multivariate order.

(1). $x \in R^p \Rightarrow x - x = 0 \in C \Rightarrow x \preceq x$. So, $\preceq$ is reflexive.

(2). For $x, y, z \in R^p$, let $x \preceq y$ and $y \preceq z$. Then,

$$\begin{aligned} y - x, z - y \in C &\Rightarrow (z - y) + (y - x) = z - x \in C \\ &\Rightarrow x \preceq z. \end{aligned}$$

Hence, $\preceq$ is transitive.

(3). For $x_1, x_2, y_1, y_2 \in R^p$, let $x_1 \preceq y_1$ and $x_2 \preceq y_2$. Then, by definition of "$\preceq$", we have $y_1 - x_1 \in C$ and $y_2 - x_2 \in C$. But C is a closed convex cone, hence with $\alpha \ge 0$ and $\beta \ge 0$, by lemma 2.2.2 it follows that

$$\alpha(y_1 - x_1) + \beta(y_2 - x_2) \in C.$$

So,

$$(\alpha y_1 + \beta y_2) - (\alpha x_1 + \beta x_2) \in C, \text{ i.e.}$$

$$\alpha x_1 + \beta x_2 \preceq \alpha y_1 + \beta y_2.$$

So, "$\preceq$" is closed under linear combinations with non-negative coefficients

(4). Suppose $x_n \preceq y_n,\ x_n \to x,$ and $y_n \to y$. Then,

$$\begin{aligned} y_n - x_n \in C \text{ and } y_n - x_n \to y - x &\Rightarrow y - x \in C \\ &\Rightarrow x \preceq y \end{aligned}$$

So, "$\preceq$" is closed under limits. Hence, "$\preceq$" is a multivariate order.

□

Such an order is called a closed convex cone C induced multivariate order.

Lemma 3.1.3. *Let $\preceq$ be a multivariate order in a Hilbert space R^p. Then there is a closed convex cone $C \subset R^p$ such that $x \preceq y \Leftrightarrow y - x \in C$.*

Proof. Define $C = \{x \in R^p : 0 \preceq x\}$. Suppose $x, y \in C$. Then, $0 \preceq x$ and $0 \preceq y$. By property (3) of a multivariate order, we have $0 \preceq \alpha x + \beta y, \forall \alpha, \beta > 0$. Thus, $\alpha x + \beta y \in C$ and hence C is a convex cone.

To show that C is closed, let $x_n \in C$ and $x_n \to x$. Then, $0 \preceq x_n$ and $x_n \to x$. It follows by property (4) of a multivariate order that $0 \preceq x$. So $x \in C$. Therefore, C is closed and hence it is a closed convex cone.

Next we need to show that $x \preceq y \Leftrightarrow y - x \in C$.

"$\Rightarrow$" : $x \preceq y \Rightarrow x \preceq y$ and $-x \preceq -x$, by proporty (1) of a multivariate order

$\Rightarrow 0 \preceq y - x$, by proporty (3) of a multivariate order

$\Rightarrow y - x \in C$

"$\Leftarrow$" : $y - x \in C \Rightarrow 0 \preceq y - x$ and $x \preceq x$

$\Rightarrow 0 + x \preceq y - x + x$ by proporty (3) of a multivariate order

$\Rightarrow x \preceq y$

□

Such a closed convex cone is called an order generating cone.

Table 3.1 presents four closed convex cones in R and the corresponding induced orders.

Convex cone	Induced order
$\{x \in R : x \geq 0\}$	$\leq$
$\{x \in R : x \leq 0\}$	$\geq$
$\{0\}$	$=$
$\{x : x \in R\}$	$\geq$ or $\leq$

Table 3.1: Closed convex cones in R and respective induced orders

In the literature, there are convex cones which are useful in different fields. Next, we present two examples of order generating cones in R^p.

Example 3.1.4. A polyhedral cone which is represented by

$$C[A] = \{x \in R^p : Ax \geq 0 \text{ (componentwise)}\}$$

where $A \in R^{k \times p}$, is a closed convex cone in R^p. As it will be discussed in the forthcoming sections, an order restricted cone C is a polyhedral cone with $k < p$. For example, let C be the collection of all $x \in R^4$ such that $x_1 \leq x_2$, $x_1 \leq x_3$ and $x_1 \leq x_4$, then $C = C[A]$ is a polyhedral cone where

$$A = \begin{pmatrix} -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix}.$$

The multivariate order "$\preceq$" generated from this closed convex cone $C[A]$ is

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \preceq \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = y \Leftrightarrow y_2 - y_1 \geq x_2 - x_1,\ y_3 - y_1 \geq x_3 - x_1 \text{ and } y_4 - y_1 \geq x_4 - x_1.$$

Example 3.1.5. Given a cone C, the set

$$C^* = \{x \in R^p | \langle x, y\rangle \leq 0 \quad \text{for all } y \in C\}$$

where $\langle x, y\rangle$ is a defined inner product in R^p, is said to be a dual cone of C. A dual cone is always a convex cone regardless of whether the original cone is convex or not.

3.1.1 Multivariate order restricted cone

Definition 3.1.6. For $A = (A_1, ..., A_q) \in R^{p\times q}$, the restriction $A_i \preceq A_j$ for all $(i, j) \in H \subset \{1, ..., q\} \times \{1, ..., q\}$ on A is called a multivariate order restriction.

For a given matrix $A = (A_1, ..., A_q) \in R^{p\times q}$, some common multivariate order restrictions on A are,

(1). multivariate simple order restriction: $A_1 \preceq ... \preceq A_q$,

(2). multivariate simple-tree order restriction: $A_1 \preceq A_2$, $A_1 \preceq A_3$,...,$A_1 \preceq A_q$,

(3). multivariate umbrella order restriction: $A_1 \preceq ... \preceq A_i \succeq ... \succeq A_q$ where $1 < i < q$.

Let $C_{p\times q}$ be the collection of all matrices $\mu = (\mu_1, ..., \mu_q) \in R^{p\times q}$ under a multivariate order restriction

$$\mu_i \preceq \mu_j \text{ for } (i, j) \in H \subset \Omega \times \Omega \text{ where } \Omega = \{1, ..., q\}.$$

Then, $C_{p\times q}$ can take of the form

$$C_{p\times q} = \{\mu = (\mu_1, ..., \mu_q) \in R^{p\times q} : \mu_i \preceq \mu_j, \quad (i, j) \in H\}. \tag{3.1}$$

Depending on the choice of the multivariate order considered, $C_{p\times q}$ can have different forms.

The following theorem establishes that $C_{p\times q}$ defined in (3.1) is a closed convex cone.

Theorem 3.1.7. *Suppose $C_{p\times q}$ be the collection of all $p \times q$ matrices in $R^{p\times q}$ constrained by a multivariate order restriction. Then $C_{p\times q}$ is a closed convex cone.*

Proof. Suppose $A = (A_1, ..., A_q) \in C_{p\times q}$ and $B = (B_1, ..., B_q) \in C_{p\times q}$. Then $A_i \preceq A_j$ and $B_i \preceq B_j$ for all $(i, j) \in H$. For $\alpha, \beta > 0$,

$$\alpha A + \beta B = (\alpha A_1 + \beta B_1, ..., \alpha A_q + \beta B_q).$$

Using the fact that $\preceq$ is preservable under linear combinations with positive coefficients, it can be noted that $\alpha A_i + \beta B_i \preceq \alpha A_j + \beta B_j$ for all $\alpha, \beta > 0$ and $(i, j) \in H$. Thus, $\alpha A + \beta B \in C_{p\times q}$ and hence by lemma 2.2.2, $C_{p\times q}$ is a convex cone.

To show the closedness under limits, let $A^{[n]} = \left(A_1^{[n]}, ..., A_q^{[n]}\right) \in C_{p\times q}$, and $A^{[n]} \to A = (A_1, ..., A_q)$. Then $A_i^{[n]} \preceq A_j^{[n]}$ for all $(i, j) \in H$, $A_i^{[n]} \to A_i$ and $A_j^{[n]} \to A_j$. Consequently, since $\preceq$ is preservable under limits with respect to a norm induced from an inner product, we have $A_i \preceq A_j$ for all $(i, j) \in H$. So, $A \in C_{p\times q}$ and hence $C_{p\times q}$ is a closed cone. $\square$

CHAPTER 4

Restricted maximum likelihood estimator in order restricted MANOVA

In statistical inference problems where a parameter matrix $\mu = (\mu_1, ..., \mu_q) \in R^{p\times q}$ is known to be under a given multivariate order restriction i.e. $\mu \in C_{p\times q}$, quite often with the maximum likelihood estimator (MLE) $\hat{\mu}$, the restricted maximum likelihood estimator (RMLE) under the restriction $\mu \in C_{p\times q}$ is $\tilde{\mu} = \pi(\hat{\mu}|C_{p\times q})$ with an appropriately defined inner product system. In this chapter we discuss this concept.

4.1 An order restricted MANOVA model

Consider an MANOVA model with q p-dimensional normal populations $N_p(\mu_i, \Sigma)$, $i = 1, ..., q$, where the positive definite matrix $\Sigma \in R^{p\times p}$ is known, and $\mu = (\mu_1, ..., \mu_q) \in R^{p\times q}$ is an unknown parameter matrix.

With respect to the multivariate order "$\preceq$" generated from the closed convex cone $C \subset R^q$, μ is under the multivariate order restriction $\mu_i \preceq \mu_j$ for $(i, j) \in H$ i.e. $\mu \in C_{p\times q}$ where $H \subset \{1, ..., q\} \times \{1, ..., q\}$. Here, $C_{p\times q}$ is the order restricted cone defined in (3.1).

In order to obtain the estimator for μ, a random sample $X_{i1}, ..., X_{i,n_i}$ is taken from the ith population with distribution $N_p(\mu_i, \Sigma)$, sample size n_i, sample mean $\bar{X}_i = \frac{\sum_{i=1}^{n_i} X_{ij}}{n_i}$ and corrected sum of squares and cross product (CSSCP)

$$CSSCP_i = \sum_{j=1}^{n_i}(X_{ij} - \bar{X}_i)(X_{ij} - \bar{X}_i)' = \big(X_i - \bar{X}_i 1'_{n_i}\big)\big(X_i - \bar{X}_i 1'_{n_i})'.$$

The data matrix from the ith population can be written in one matrix as $X_i = (X_{i1}, ..., X_{in_i}) \in R^{p\times n_i}$ with a distribution $X_i \sim N_p(\mu_i 1'_{n_i}, \Sigma, I_{n_i})$. Then, the sample mean is

$$\bar{X}_i = X_i 1_{n_i}(1'_{n_i} 1_{n_i})^{-1} = \frac{X_i 1_{n_i}}{n_i},$$

and the corrected sum of squares and cross product is given by

$$
\begin{aligned}
CSSCP_i &= \sum_{j=1}^{n_i}(X_{ij}-\bar{X}_i)(X_{ij}-\bar{X}_i)' = \big(X_i-\bar{X}_i1'_{n_i}\big)\big(X_i-\bar{X}_i1'_{n_i}\big)' \\
&= \big(X_i-\frac{X_i1_{n_i}1'_{n_i}}{n_i}\big)\big(X_i-\frac{X_i1_{n_i}1'_{n_i}}{n_i}\big)' \\
&= \big[X_i\big(I_{n_i}-\frac{1_{n_i}1'_{n_i}}{n_i}\big)\big]\big[X_i\big(I_{n_i}-\frac{1_{n_i}1'_{n_i}}{n_i}\big)\big]' \\
&= X_i\big(I_{n_i}-\frac{1_{n_i}1'_{n_i}}{n_i}\big)\big(I_{n_i}-\frac{1_{n_i}1'_{n_i}}{n_i}\big)'X'_i \\
&= X_i\big(I_{n_i}-\frac{1_{n_i}1'_{n_i}}{n_i}\big)X'_i.
\end{aligned}
$$

Notice that the last equality is obtained since the matrix $\big(I_{n_i}-\frac{1_{n_i}1'_{n_i}}{n_i}\big)$ is idempotent.

Furthermore, from the pooled data matrix $X=(X_1,...,X_q)\sim N_{p\times n}(\mu J',\Sigma,I_n)$ where $n=n_1+...+n_q$ and

$$
J=\begin{pmatrix} 1_{n_1} & ... & 0 \\ \vdots & \ddots & \vdots \\ 0 & ... & 1_{n_q} \end{pmatrix}\in R^{n\times q}
$$

we have the statistical matrices

$$
\bar{X}=(\bar{X}_1,...,\bar{X}_q)\sim N_{p\times q}(\mu,\Sigma,(J'J)^{-1})
$$

and

$$
CSSCP=CSSCP_1+...+CSSCP_q=X\big[I_n-J(J'J)^{-1}J'\big]X'.
$$

Based on the pooled sample, the likelihood function is

$$
\begin{aligned}
L(\mu) &= \Pi_{i=1}^{q}\Pi_{j=1}^{n_i}\frac{1}{(2\pi)^{p/2}|\Sigma|^{1/2}}\exp\big[-\frac{1}{2}(X_{ij}-\mu_i)'\Sigma^{-1}(X_{ij}-\mu_i)\big] \\
&= \frac{1}{(2\pi)^{(np)/2}|\Sigma|^{n/2}}\exp\Bigg[-\frac{1}{2}\sum_{i=1}^{q}\sum_{j=1}^{n_i}(X_{ij}-\mu_i)'\Sigma^{-1}(X_{ij}-\mu_i)\Bigg] \\
&= \frac{1}{(2\pi)^{(np)/2}|\Sigma|^{n/2}}\exp\Bigg[-\frac{1}{2}\sum_{i=1}^{q}\sum_{j=1}^{n_i}[(X_{ij}-\bar{X}_i)+(\bar{X}_i-\mu_i)]'\Sigma^{-1}[(X_{ij}-\bar{X}_i)+(\bar{X}_i-\mu_i)]\Bigg].
\end{aligned}
\tag{4.1}
$$

Notice that the exponent term in (4.1) can is

$$\sum_{i=1}^{q}\sum_{j=1}^{n_i}[(X_{ij}-\bar{X}_i)+(\bar{X}_i-\mu_i)]'\Sigma^{-1}[(X_{ij}-\bar{X}_i)+(\bar{X}_i-\mu_i)] = \sum_{i=1}^{q}\sum_{j=1}^{n_i}(X_{ij}-\bar{X}_i)'\Sigma^{-1}(X_{ij}-\bar{X}_i)$$
$$+\sum_{i=1}^{q}\sum_{j=1}^{n_i}(X_{ij}-\mu_i)'\Sigma^{-1}(X_{ij}-\mu_i)$$
$$+\sum_{i=1}^{q}\sum_{j=1}^{n_i}(X_{ij}-\bar{X}_i)'\Sigma^{-1}(\bar{X}_i-\mu_i)$$
$$+\sum_{i=1}^{q}\sum_{j=1}^{n_i}(\bar{X}_i-\mu_i)'\Sigma^{-1}(X_{ij}-\bar{X}_i). \tag{4.2}$$

But, the last two terms in (4.2) are

$$\begin{aligned}\sum_{i=1}^{q}\sum_{j=1}^{n_i}(X_{ij}-\bar{X}_i)'\Sigma^{-1}(\bar{X}_i-\mu_i) &= \sum_{i=1}^{q}\left[\sum_{j=1}^{n_i}(X_{ij}-\bar{X}_i)'\right]\Sigma^{-1}(\bar{X}_i-\mu_i)\\ &= \sum_{i=1}^{q}[0]\Sigma^{-1}\bar{X}_i-\mu_i)\\ &= 0,\end{aligned}$$

and

$$\begin{aligned}\sum_{i=1}^{q}\sum_{j=1}^{n_i}(\bar{X}_i-\mu_i)'\Sigma^{-1}(x_{ij}-\bar{X}_i) &= \sum_{i=1}^{q}(\bar{X}_i-\mu_i)'\Sigma^{-1}\left[\sum_{j=1}^{n_i}(X_{ij}-\bar{X}_i)\right]\\ &= \sum_{i=1}^{q}(\bar{X}_i-\mu_i)'\Sigma^{-1}[0]\\ &= 0.\end{aligned}$$

So, (4.1) is expressed as

$$\begin{aligned}L(\mu) = \frac{1}{(2\pi)^{(np)/2}|\Sigma|^{n/2}}\exp\bigg\{ &-\frac{1}{2}\sum_{i=1}^{q}\sum_{j=1}^{n_i}(X_{ij}-\bar{X}_i)'\Sigma^{-1}(X_{ij}-\bar{X}_i)\\ &-\frac{1}{2}\sum_{i=1}^{q}\sum_{j=1}^{n_i}(\bar{X}_i-\mu_i)'\Sigma^{-1}(\bar{X}_i-\mu_i)\bigg\}.\end{aligned}$$

Moreover,

$$\begin{aligned}\sum_{i=1}^{q}\sum_{j=1}^{n_i}(X_{ij}-\bar{X}_i)'\Sigma^{-1}(X_{ij}-\bar{X}_i) &= \mathrm{tr}\Big[\sum_{i=1}^{q}\sum_{j=1}^{n_i}(X_{ij}-\bar{X}_i)'\Sigma^{-1}(X_{ij}-\bar{X}_i)\Big]\\ &= \sum_{i=1}^{q}\sum_{j=1}^{n_i}\mathrm{tr}\Big[(X_{ij}-\bar{X}_i)'\Sigma^{-1}(X_{ij}-\bar{X}_i)\Big]\\ &= \sum_{i=1}^{q}\sum_{j=1}^{n_i}\mathrm{tr}\Big[\Sigma^{-1}(X_{ij}-\bar{X}_i)(X_{ij}-\bar{X}_i)'\Big]\\ &= \mathrm{tr}\Big[\Sigma^{-1}\sum_{i=1}^{q}\sum_{j=1}^{n_i}(X_{ij}-\bar{X}_i)(X_{ij}-\bar{X}_i)'\Big]\\ &= \mathrm{tr}\big[\Sigma^{-1}\sum_{i=1}^{q}\big(\mathrm{CSSCP}_i\big)\big]\\ &= \mathrm{tr}\big[\Sigma^{-1}\big(\mathrm{CSSCP}\big)\big].\end{aligned}$$

So, the likelihood function is expressed as

$$L(\mu)=\frac{1}{(2\pi)^{(np)/2}|\Sigma|^{n/2}}\exp\Big\{-\frac{1}{2}\mathrm{tr}\big[\Sigma^{-1}(\mathrm{CSSCP})\big]\Big\}\exp\Big\{-\frac{1}{2}\sum_{i=1}^{q}n_i(\bar{X}_i-\mu_i)'\Sigma^{-1}(\bar{X}_i-\mu)\Big\}. \tag{4.3}$$

Next, we define a general inner product in $R^{p\times q}$.

4.2 An inner product in $R^{p\times q}$

For $x, y \in R^p$ and a positive definite matrix $V \in R^{p\times p}$, define an inner product by

$$\langle x, y\rangle_V = y'Vx. \tag{4.4}$$

Moreover, $\|.\|_V$ is the norm induced from the inner product in (4.4).

With $w_i > 0$, $i = 1, ..., q$ as weight of column i, and matrices $A = (A_1, ..., A_q) \in R^{p\times q}$ and $B = (B_1, ..., B_q) \in R^{p\times q}$ define $\langle A, B\rangle_{p\times q}$ by

$$\langle A, B\rangle_{p\times q} = \sum_{i=1}^{q} w_i\langle A_i, B_i\rangle_V. \tag{4.5}$$

Then, $\langle ., .\rangle_{p\times q}$ satisfies the following

(1). $\langle A, A\rangle_{p\times q} \geq 0$ for all $A \in R^{p\times q}$ and $\langle A, A\rangle_{p\times q} = 0 \Leftrightarrow A = 0$.

(2). $\langle A, B\rangle_{p\times q} = \langle B, A\rangle_{p\times q}$.

(3). For $D \in R^{p\times q}$, $\langle \alpha A + \beta B, D\rangle_{p\times q} = \alpha\langle A, D\rangle_{p\times q} + \beta\langle B, D\rangle_{p\times q}$.

and hence, it is a proper inner product in $R^{p\times q}$, and $\|.\|_{p\times q}$ is the norm induced from this inner product.

Next, we discuss a maximum likelihood estimator and restricted maximum likelihood estimator for μ.

4.3 Maximum likelihood and restricted maximum likelihood estimators for μ

Replacing V by Σ^{-1} in (4.4), we have

$$\langle x, y\rangle_{\Sigma^{-1}} = y'\Sigma^{-1}x \tag{4.6}$$

and $\|.\|_{\Sigma^{-1}}$ is the induced norm.

Furthermore, with $w_i = n_i$ and making use of (4.6), the inner product defined given by (4.5) can be expressed as

$$\langle A, B\rangle_{p\times q} = \sum_{i=1}^{q} n_i\langle A_i, B_i\rangle_{\Sigma^{-1}}. \tag{4.7}$$

So, making use of this specific inner product given in (4.7), we have

$$\begin{aligned}\sum_{i=1}^{q}\sum_{j=1}^{n_i}(\bar{X}_i - \mu_i)'\Sigma^{-1}(\bar{X}_i - \mu_i) &= \sum_{i=1}^{q} n_i(\bar{X}_i - \mu_i)'\Sigma^{-1}(\bar{X}_i - \mu_i)\\ &= \|\bar{X} - \mu\|^2_{p\times q}. \end{aligned}\tag{4.8}$$

Therefore, making use of the expressions in (4.8), the likelihood function in (4.3) can further be expressed as

$$L(\mu) = \frac{1}{(2\pi)^{(np)/2}|\Sigma|^{n/2}}\exp\left\{-\frac{1}{2}\mathrm{tr}\left[\Sigma^{-1}(\mathrm{CSSCP})\right]\right\}\exp\left\{-\frac{1}{2}\|\bar{X} - \mu\|^2_{p\times q}\right\}. \tag{4.9}$$

Note that the first term in the exponent of (4.9) is free of μ. Moreover, $L(\mu)$ is a decreasing function of $\|\bar{X}-\mu\|^2_{p\times q}$. So, $L(\mu)$ is maximized when $\|\bar{X}-\mu\|^2_{p\times q}$ is minimized. When there is no known multivariate order restriction on the columns of μ i.e. $\mu \in R^{p\times q}$, $\|\bar{X}-\mu\|^2_{p\times q}$ is minimized at $\mu=\bar{X}$. Thus, $\bar{X}$ is the maximum likelihood estimator (MLE) for $\mu \in R^{p\times q}$. Recall that $\bar{X}$ is an unbiased estimator for μ.

Now, suppose μ is under multivariate order restriction i.e., $\mu \in C_{p\times q}$. Then, by lemma 2.3.3, $\|\bar{X}-\mu\|^2_{p\times q}$ is minimized when $\mu=\pi(\bar{X}|C_{p\times q})$. $\pi(\bar{X}|C_{p\times q})$, is called the restricted maximum likelihood estimator (RMLE) for $\mu \in C_{p\times q}$. Clearly, finding RMLE for $\mu \in C_{p\times q}$ is a problem of finding a projection of $\bar{X}$ onto a closed convex cone $C_{p\times q}$, $\pi(\bar{X}|C_{p\times q})$, with respect to a properly defined inner product.

The computation of $\pi(\bar{X}|C_{p\times q})$ is a great challenge. For $q=2$, however, $\pi(\bar{X}|C_{p\times 2})$ can be obtained through a vector projection with respect to an inner product in R^p.

The following lemma provides a technique to find the projection of a matrix $X \in R^{p\times 2}$ onto $C_{p\times 2}$.

Lemma 4.3.1. *For $X=(X_1,X_2)\in R^{p\times 2}$, let $\bar{X}_* = \frac{w_1X_1+w_2X_2}{w_1+w_2}$ and $P_C=\pi(X_2-X_1|C)$. Define $\hat{X}=(\hat{X}_1,\hat{X}_2)$ by*

$$\hat{X}_1 = \bar{X}_* - \frac{w_2P_C}{w_1+w_2} \text{ and}$$
$$\hat{X}_2 = \bar{X}_* + \frac{w_1P_c}{w_1+w_2}.$$

Then $\hat{X}=\pi(X|C_{p\times 2})$.

Proof. By definition of $\hat{X}_1$ and $\hat{X}_2$, we have

$$\hat{X}_2-\hat{X}_1 = \bar{X}_* + \frac{w_1P_c}{w_1+w_2} - \big(\bar{X}_* - \frac{w_1P_C}{w_1+w_2}\big)$$

and it follows that

$$P_c = \pi(X_2 - X_1|C) \in C.$$

So, by lemma 3.1.3, $\hat{X}_2 - \hat{X}_1 \in C \Leftrightarrow \hat{X}_1 \preceq \hat{X}_2$. Thus, $\hat{X} \in C_{p\times 2}$.

Let $Y = (Y_1, Y_2) \in C_{p\times 2}$ where $Y_2 - Y_1 \in C$. Then, since $P_C = \pi(X_2 - X_1|C)$, by lemma 2.5 we have

$$\langle X_2 - X_1 - P_C, P_C - (Y_2 - Y_1)\rangle \geq 0.$$

Note that,

$$\begin{aligned} X_1 - \hat{X}_1 &= X_1 - \frac{w_1X_1 + w_2X_2}{w_1 + w_2} + \frac{w_2P_c}{w_1 + w_2} \\ &= -\frac{w_2}{w_1 + w_2}(X_2 - X_1 - P_C) \end{aligned}$$

and

$$\begin{aligned} X_2 - \hat{X}_2 &= X_2 - \frac{w_1X_1 + w_2X_2}{w_1 + w_2} - \frac{w_1P_c}{w_1 + w_2} \\ &= \frac{w_1}{w_1 + w_2}(X_2 - X_1 - P_c). \end{aligned}$$

So,

$$\begin{aligned} \langle X - \hat{X}, \hat{X} - Y\rangle_{p\times 2} &= w_1\langle X_1 - \hat{X}_1, \hat{X}_1 - Y_1\rangle + w_2\langle X_2 - \hat{X}_2, \hat{X}_2 - Y_2\rangle \\ &= w_1\langle -\frac{w_2}{w_1 + w_2}(X_2 - X_1 - P_C), \hat{X}_1 - Y_1\rangle \\ &+ w_2\langle \frac{w_1}{w_1 + w_2}(X_2 - X_1 - P_c), \hat{X}_2 - Y_2\rangle \\ &= -\frac{w_1w_2}{w_1 + w_2}\langle X_2 - X_1 - P_C, \hat{X}_1 - Y_1\rangle + \\ &\quad \frac{w_1w_2}{w_1 + w_2}\langle X_2 - X_1 - P_C, \hat{X}_2 - Y_2\rangle \\ &= \frac{w_1w_2}{w_1 + w_2}\langle X_2 - X_1 - P_C, (\hat{X}_2 - \hat{X}_1) - (Y_2 - Y_1)\rangle \\ &= \frac{w_1w_2}{w_1 + w_2}\langle X_2 - X_1 - P_C, P_C - (Y_2 - Y_1)\rangle \\ &\geq 0 \end{aligned}$$

Hence, $\hat{X} = \pi(X|C_{p\times 2})$. □

In an order restricted MANOVA problem with $q = 2$, lemma 4.3.1 gives the projection of $\bar{X}$ onto $C_{p\times 2}$, $\hat{\mu} = \pi(\bar{X}|C_{p\times 2})$, where $\hat{\mu} = (\hat{\mu}_1, \hat{\mu}_2)$ and

$$\hat{\mu}_1 = \frac{n_1\bar{X}_1 + n_2\bar{X}_2}{n_1 + n_2} - \frac{n_2}{n_1 + n_2}\pi(\bar{X}_2 - \bar{X}_1|C)$$

$$\hat{\mu}_2 = \frac{n_1\bar{X}_1 + n_2\bar{X}_2}{n_1 + n + 2} + \frac{n_1}{n_1 + n_2}\pi(\bar{X}_2 - \bar{X}_1|C).$$

Therefore, $\hat{\mu}$ obtained through the procedure in lemma 4.3.1 is in fact an RMLE for $\mu \in C_{p\times 2}$.

CHAPTER 5

A proposed pseudo RMLE under simple-tree order restriction

With a multivariate order $\preceq$ generated from the closed convex cone $C \subset R^p$, for $\mu_i \in R^p$, $i = 1, ..., q$,

$$\mu_1 \preceq \mu_2,\ \mu_1 \preceq \mu_3,\ \ldots,\ \mu_1 \preceq \mu_q \tag{5.1}$$

is called a simple-tree ordering. The collection of all matrices $\mu = (\mu_1, ..., \mu_q) \in R^{p\times q}$ whose columns satisfy the simple tree ordering from a closed convex cone

$$C_{p\times q} = \{\mu = (\mu_1, ..., \mu_q) \in R^{p\times q} : \mu_1 \preceq \mu_i \text{ for all } i = 2, ..., q\}. \tag{5.2}$$

The restriction $\mu \in C_{p\times q}$ often occurs in the experiments where μ_1 is a parameter vector from the response to a control group and μ_i, $i = 2, ..., q$, are the parameter vectors from the response to treatment groups.

For a given $X \in R^{p\times q}$, let $D(X)_{p\times q}$ be the collection of matrices $Y = (Y_1, ..., Y_2) \in R^{p\times q}$ such that $Y_i - Y_1 = \pi(X_i - X_1|C)$ with respect to the inner product $\langle ., . \rangle_V$ in R^p, i.e.,

$$D(X)_{p\times q} = \{Y = (Y_1, ..., Y_q) \in R^{p\times q} : Y_i - Y_1 = \pi(X_i - X_1|C) \text{ for all } i = 2, ..., q\}. \tag{5.3}$$

Next we show that $D(X)_{p\times q}$ is a closed convex subset of $C_{p\times q}$.

Lemma 5.0.1. *For $X \in R^{p\times q}$, $D(X)_{p\times q}$ defined in (5.3) is closed convex set.*

Proof. Suppose $Y, Z \in D(X)_{p\times q}$. Then, by definition of $D(X)_{p\times q}$, we have

$$Y_j - Y_1 = Z_j - Z_1 = \pi(X_j - X_1|C) \text{ for all } j = 2, ..., q.$$

For $\alpha \in (0, 1)$,

$$\alpha Y + (1-\alpha)Z = \left[\alpha Y_1 + (1-\alpha)Z_1, ..., \alpha Y_q + (1-\alpha)Z_q\right].$$

But,

$$\begin{aligned}\big[\alpha Y_j + (1-\alpha)Z_j\big] - \big[\alpha Y_1 + (1-\alpha Z_1)\big] &= \alpha\big(Y_j - Y_1\big) + (1-\alpha)\big(Z_j - Z_1\big)\\ &= \alpha\pi(X_j - X_1|C) + (1-\alpha)\pi(X_j - X_1|C)\\ &= \pi(X_j - X_1|C) \in C \text{ for all } j = 2, ..., q.\end{aligned}$$

Thus, $\alpha Y + (1-\alpha)Z \in D(X)_{p\times q}$ and hence, $D(X)_{p\times q}$ is a convex set.

Suppose $Y^{(n)} \in D(X)_{p\times q}$ and $Y^{(n)} \to Y$. Then,

$$\begin{aligned}&Y_j^{(n)} - Y_1^{(n)} = \pi(X_j - X_1|C) \text{ and}\\ &Y_j^{(n)} - Y_1^{(n)} \to Y_j - Y_1 = \pi(X_j - X_1|C) \text{ for all } j = 2, ..., q.\end{aligned}$$

So, $Y \in D(X)_{p\times q}$, and hence, $D(X)_{p\times q}$ is a closed. □

Lemma 5.0.2. *For $X \in R^{p\times q}$, $D(X)_{p\times q}$ defined in (5.3) is a subset of $C_{p\times q}$.*

Proof. Let $Z = (Z_1, ..., Z_q) \in D(X)_{p\times q}$. Then, by definition of $D(X)_{p\times q}$, $Z_j - Z_1 = \pi(X_j - X_1|C)$ for all $j = 2, ..., q$. So, $Z_j - Z_1 \in C$ for all $j = 2, ..., q$. By lemma 3.1.2, we have

$$Z_j - Z_1 \in C \Rightarrow Z_1 \preceq Z_j \text{ for all } j = 2, ..., q.$$

So, $Z \in C_{p\times q}$ and hence, $D(X)_{p\times q} \subset C_{p\times q}$. □

Thus by lemma 2.3.2, $\pi(Y|D(X)_{p\times q}))$ exists and is unique for all $Y \in R^{p\times q}$. Specifically, $\pi(X|D(X)_{p\times q})$ exists and is unique.

Example 5.0.3. When $q = 2$, $\pi(X|D(X)_{p\times q}) = \pi(X|C_{p\times q})$.

Let $\hat{X} = \pi(X|C_{p\times q})$. By lemma 4.3.1, $\hat{X}_1 = \bar{X}_* - \frac{w_2 P_C}{w_1+w_2}$ and $\hat{X}_2 = \bar{X}_* + \frac{w_1 P_c}{w_1+w_2}$ where $\bar{X}_* = \frac{w_1X_1+w_2X_2}{w_1+w_2}$. Hence, $\hat{X}_2 - \hat{X}_1 = \pi(X_2 - X_1|C)$. Therefore, $\hat{X} \in D(X)_{p\times 2}$. Thus, $\pi(X|D(X)_{p\times 2}) = \hat{X}$.

Example 5.0.4. When $X \in C_{p\times q}$, $\pi(X|D(X)_{p\times q}) = \pi(X|C_{p\times q})$.

Let $X \in C_{p\times q}$. Then, $\pi(X|C_{p\times q}) = X$, and $X_i - X_1 = \pi(X_i - X_1|C)$ for all $i = 2, ..., q$. So, $X \in D(X)_{p\times q}$.

For all $Y \in D(X)_{p\times q}$, $\|X - Y\|_{p\times q} \geq \|X - X\|_{p\times q} = 0$. Therefore, $\pi(X|D(X)_{p\times q}) = X$ and hence, $\pi(X|D(X)_{p\times q}) = \pi(X|C_{p\times q})$.

Generally, $\pi(X|D(X)_{p\times q})$ could be utilized as an approximation of $\pi(X|C_{p\times q})$. When this approximation is used to the simple tree order restricted MANOVA model introduced in chapter 4, $\pi(\bar{X}|D(\bar{X})_{p\times q})$ replaces $\pi(\bar{X}|C_{p\times q})$ and becomes an estimator for μ under $\mu \in C_{p\times q}$. This estimator is obtained by maximizing the likelihood function over modified domain $D(X)_{p\times q}$ and hence is our proposed pseudo RMLE for $\mu \in C_{p\times q}$.

For theoretical and/or computational simplicity, researchers often modify the likelihood function or restricted domain to obtain a pseudo restricted maximum likelihood estimator. Hu (2020) considered the case where the components of μ are constrained by a multivariate simple order restriction and proposed an algorithm for computing a pseudo maximum likelihood estimator for μ. In this work, we considered the case where the components of μ are under multivariate simple tree ordering i.e. $\mu \in C_{p\times q}$ where $C_{p\times q}$ is as defined in (3.1).

5.1 A proposed algorithm

The computation for the proposed pseudo RMLE is a computation for $\pi(X|D(X)_{p\times q})$. Here, $D(X)_{p\times q}$ is a one column index matrix set since assuming $d_i = \pi(X_i - X_1|C)$, $i = 2, ..., q$, are computable and hence are available, then

$$Y = (Y_1, ..., Y_q) \in D(X)_{p\times q} \Leftrightarrow Y = (Y_1, Y_2 + d_2, ..., Y_q + d_q).$$

So, each Y in $D(X)_{p\times q}$ is identified by its first column Y_1. Now, consider the minimizing the function defined by $f(Y_1) = \|X - Y\|^2_{p\times q}$ over $Y \in D(X)_{p\times q}$. For convenience, let $d_1 = 0$.

Then,

$$\begin{aligned}
f(Y_1) &= \|X - Y\|^2_{p\times q} \\
&= \sum_{i=1}^{q} w_i \|Y_i - X_i\|^2_V \\
&= \sum_{i=1}^{q} w_i \|Y_1 + d_i - X_i\|^2_V \\
&= \sum_{i=1}^{q} w_i \|Y_1\|^2_V + 2\sum_{i=1}^{q} w_i \langle Y_1, d_i - X_i\rangle_V + \sum_{i=1}^{q} w_i \|d_i - X_i\|^2_V \\
&= Y_1' V Y_1 \sum_{i=1}^{q} w_i + 2\sum_{i=1}^{q} w_i (d_i - X_i)' V Y_1 + \sum_{i=1}^{q} w_i \|d_i - X_i\|^2_V \qquad (5.4)
\end{aligned}$$

The last term of (5.4) is free of Y_1, the gradient vector and the Hessian matrix for $f(Y_1)$ are respectively given by

$$\bigtriangledown f = \frac{\partial}{\partial Y_1} f(Y_1) = 2VY_1 \sum_{i=1}^{q} w_i + 2\sum_{i=1}^{q} w_i V(d_i - X_i) \qquad (5.5)$$

and

$$H_f = \frac{\partial}{\partial Y_1'} \bigtriangledown f = 2V \sum_{i=1}^{q} w_i \geq 0. \qquad (5.6)$$

H_f is a positive definite matrix and hence, $f(Y_1)$ is minimized at the stationary points which can be obtained by setting $\bigtriangledown f$ to zero and solving the equation.

$$2VY_1 \sum_{i=1}^{q} w_i + 2\sum_{i=1}^{q} w_i V(d_i - X_i) = 0.$$

It follows that

$$Y_1 = \frac{1}{w_*} \sum_{i=1}^{q} w_i (X_i - d_i) = \bar{X}_* - \bar{d}$$

where $\bar{X}_* = \frac{1}{w_*}\sum_{i=1}^{q} w_i X_i$, $w_* = \sum_{i=1}^{q} w_i$, and $\bar{d} = \frac{1}{w_*}\sum_{i=1}^{q} d_i$.

Therefore, Y can be written as,

$$Y = (\bar{X}_* - \bar{d}, \bar{X}_* - \bar{d} + d_2, ..., \bar{X}_* - \bar{d} + d_q). \qquad (5.7)$$

It is important to note that the expression in (5.7) involves only vector projection. We therefore come to a proposed algorithm that produces $\hat{X}$.

5.1.1 Algorithm

For a given matrix $X \in R^{p\times q}$, weight w_i, $i = 1, ..., q$, and a properly defined inner product $\langle ., . \rangle_V$, the following algorithm is proposed.

Step 1. With $w_* = w_1 + ... + w_q$, compute

$\bar{X}_* = \frac{1}{w_*}\sum_{i=1}^{q} w_i X_i$ and $\pi(X_i - X_1|C)$ for all $i = 2, ..., q$.

Step 2. Calculate $\hat{X}_1$ by

$\hat{X}_1 = \bar{X}_* - \frac{1}{w_*}\sum_{i=2}^{q} w_i \pi(X_i - X_1|C)$.

Step 3. Compute the remaining columns of $\hat{X}$ as

$\hat{X}_i = \hat{X}_1 + \pi(X_i - X_1|C)$ for $i = 2, ..., q$.

Step 4. Output, $\hat{X} = (\hat{X}_1, ..., \hat{X}_q)$.

5.1.2 Proof of the algorithm

Let $\hat{X}$ be the matrix produced by the algorithm. We show that $\hat{X} = \pi(X|D(X)_{p\times q})$, i.e.,

$$\hat{X} \in D(X)_{p\times q} \text{ and } \|X - \hat{X}\|^2_{p\times q} \leq \|X - Y\|^2_{p\times q} \text{ for all } Y \in D(X)_{p\times q}.$$

Lemma 5.1.1. *Let $\hat{X}$ be a matrix produced by the algorithm. Then, $\hat{X} \in D(X)_{p\times q}$.*

Proof. By step 3 of the algorithm, we have

$$\hat{X}_i = \hat{X}_1 + \pi(X_i - X_1|C) \text{ for } i = 2, ..., q$$

which is equivalent to

$$\hat{X}_i - \hat{X}_1 = \pi(X_i - X_1|C) \text{ for all } i = 2, ..., q. \tag{5.8}$$

So, by definition of $D(X)_{p\times q}$, it follows that $\hat{X} \in D(X)_{p\times q}$. □

Corollary 5.1.2. *Let* $\hat{X} = (\hat{X}_1, ..., \hat{X}_q)$ *be a matrix produced by the algorithm, and* $Y = (Y_1, ..., Y_q) \in D(X)_{p\times q}$. *Then,* $\hat{X}_i - Y_i = \hat{X}_1 - Y_1$ *for all* $i = 2, ..., q$.

Proof. By definition of $D(X)_{p\times q}$, we have

$$Y_i - Y_1 = \pi(X_i - X_1|C) \text{ for all } i = 2, ..., q.$$

and, by (5.8),

$$\hat{X}_i - \hat{X}_1 = \pi(X_i - X_1|C) \text{ for all } i = 2, ..., q.$$

Therefore, $\hat{X}_i - \hat{X}_1 = Y_i - Y_1$ for all $i = 2, ..., q$, which is equivalent to

$$\hat{X}_i - Y_i = \hat{X}_1 - Y_1 \text{ for all } i = 2, ..., q.$$

□

Lemma 5.1.3. *Let* $X = (X_1, ..., X_q) \in R^{p\times q}$, *and let* $\hat{X} = (\hat{X}_1, ..., \hat{X}_q) \in D(X)_{p\times q}$ *be produced by the algorithm. Then,*

$$\sum_{i=1}^{q} w_i \hat{X}_i = \sum_{i=1}^{q} w_i X_i.$$

Proof. Let $k = \sum_{i=2}^{q} w_i \pi(X_i - X_1|C)$.

By step 3 of the algorithm,

$$\hat{X}_i = \hat{X}_1 + \pi(X_i - X_1|C) \text{ for all } i = 2, ..., q.$$

With w_* defined in step 1 of the algorithm, we have

$$\begin{aligned}\sum_{i=1}^{q} w_i \hat{X}_i &= w_1 \hat{X}_1 + \sum_{i=2}^{q} w_i \hat{X}_i \\ &= w_1 \hat{X}_1 + \sum_{i=2}^{q} w_i \big[\hat{X}_1 + \pi(X_i - X_1|C)\big] \\ &= w_* \hat{X}_1 + k,\end{aligned}$$

and by step 2 of the algorithm, $\hat{X}_1 = \bar{X}_* - \frac{k}{w_*}$. So,

$$\sum_{i=1}^{q} w_i \hat{X}_i = w_* \big(\bar{X}_* - \frac{k}{w_*}\big) + k$$
$$= w_* \bar{X}_*.$$

But, by step 1 of the algorithm, $w_* \bar{X}_* = \sum_{i=1}^{q} w_i X_i$, and hence, the result

$$\sum_{i=1}^{q} w_i \hat{X}_i = \sum_{i=1}^{q} w_i X_i$$

follows. □

Theorem 5.1.4. *For $X = (X_1, ..., X_q) \in R^{p \times q}$, let $\hat{X} = (\hat{X}_1, ..., \hat{X}_q)$ be the matrix produced by the algorithm. Then, $\hat{X} = \pi(X|D(X)_{p\times q})$.*

Proof. For a given matrix $X = (X_1, ..., X_q) \in R^{p \times q}$, suppose $\hat{X} = (X_1, ..., X_q)$ is obtained using the algorithm. Then, we need to show that $\hat{X} \in D(X)_{p\times q}$ and $\|X - \hat{X}\|^2_{p\times q} \leq \|X - Y\|^2_{p\times q}$ for all $Y \in D(X)_{p\times q}$. But, by lemma 5.1.1 $\hat{X} \in D(X)_{p\times q}$.

Note that by definition of $\langle ., . \rangle_{p\times q}$ given in (4.5), for all $Y \in D(X)_{p\times q}$ we have

$$\langle X - \hat{X}, \hat{X} - Y \rangle_{p\times q} = \sum_{i=1}^{q} w_i \langle X_i - \hat{X}_i, \hat{X}_i - Y_i \rangle_V.$$

Furthermore, by corollary 5.1.2 $\hat{X}_i - \hat{X}_i = \hat{X}_1 - Y_1$, so,

$$\langle X - \hat{X}, \hat{X} - Y \rangle_{p\times q} = \sum_{i=1}^{q} w_i \langle X_i - \hat{X}_i, \hat{X}_1 - Y_1 \rangle_V$$
$$= \Big\langle \sum_{i=1}^{q} w_i X_i - \sum_{i=1}^{q} w_i \hat{X}_i, \hat{X}_1 - Y_1 \Big\rangle_V$$
$$= \langle 0, \hat{X}_1 - Y_1 \rangle_V$$
$$= 0.$$

So, by Pythagorean theorem,

$$\begin{aligned}\|X - Y\|^2_{p\times q} &= \|X - \hat{X} + \hat{X} - Y\|^2_{p\times q} \\ &= \|X - \hat{X}\|^2_{p\times q} + \|\hat{X} - Y\|^2_{p\times q} \\ &\geq \|X - \hat{X}\|^2_{p\times q}\end{aligned}$$

Therefore, by definition of projection clearly $\hat{X} = \pi(X|D(X)_{p\times q})$. □

Since $D(X)_{p\times q} \subset C_{p\times q}$, theorem 5.1.4 guarantees that $\hat{X}$, the output from the algorithm, is always in $C_{p\times q}$ which means that $\hat{X}$ satisfies the multivariate simple-tree order generated from $C_{p\times q}$. When $q = 2$, each step in the algorithm becomes as follows

Step 1. $\bar{X}_* = \frac{1}{w_*}\sum_{i=1}^{q} w_i X_i = \dfrac{w_1 X_1 + w_2 X_2}{w_1 + w_2}$, and $\pi(X_j - X_1|C) = \pi(X_2 - X_1) = P_c$.

Step 2. $\hat{X}_1 = \bar{X}_* - \frac{1}{w_*}\sum_{j=2}^{q} w_j \pi(X_j - X_1|C) = \bar{X}_* - \dfrac{w_2 P_c}{w_1 + w_2}$

Step 3. $\hat{X}_j = \hat{X}_1 + \pi(X_2 - X_1|C) = \hat{X}_1 + P_c = \bar{X}_* - \dfrac{w_2 P_c}{w_1 + w_2} + P_c = \bar{X}_* + \dfrac{w_1 P_c}{w_1 + w_2}$

Step 4. $\hat{X} = (\hat{X}_1, \hat{X}_2)$.

This means that for $q = 2$, the algorithm reduces to lemma 4.3.1, and hence, it can be regarded as an extension of the lemma.

Moreover, it was shown in example 5.0.3 that in an order restricted MANOVA problem, when $q = 2$, the pseudo RMLE, $\pi(\bar{X}|D(X)_{p\times 2})$, obtained using the algorithm is in fact the RMLE of the parameter matrix $\mu \in C_{p\times 2}$. Moreover, in example 5.0.4 it was discussed that when $\bar{X} \in C_{p\times q}$, the pseudo RMLE $\pi(\bar{X}|D(X)_{p\times q})$ is the RMLE for $\mu \in C_{p\times q}$.

When, $q > 2$ and $\bar{X} \notin C_{p\times q}$, then the pseudo RMLE, $\pi(\bar{X}|D(\bar{X})_{p\times q})$, produced by the algorithm is an approximation for the restricted maximum likelihood estimator, $\pi(\bar{X}|C_{p\times q})$, for $\mu \in C_{p\times q}$.

Next we present a numerical example for computing $\hat{X}$ using the algorithm.

5.1.3 Numerical example

This experimental data is adopted from Jacqueline Dietz (1989). One of six dosages of vinylidene fluoride was administered to groups of ten male rats. Different response variables were measured on the rats three of which were levels of serum enzymes, SDH, SGOT and SGPT. An increase in levels of the enzymes is an indication of liver damage (Jacqueline Dietz, 1989). The data is given in table 5.1 below.

		Rat within dosage									
Dosage	Enzyme	1	2	3	4	5	6	7	8	9	10
0	SDH	18	27	16	21	26	22	17	27	26	27
	SGPOT	101	103	90	98	101	92	123	105	92	88
	SGPT	65	67	52	58	64	60	66	63	68	56
1500	SDH	25	21	24	19	21	22	20	25	24	27
	SGPOT	113	99	102	144	109	135	100	95	89	98
	SGPT	65	63	70	73	67	66	58	53	58	65
5000	SDH	22	21	22	30	25	21	29	22	24	21
	SGPOT	88	95	104	92	103	96	100	122	102	107
	SGPT	54	56	71	59	61	57	61	59	63	61
15000	SDH	31	26	28	24	33	23	27	24	28	29
	SGPOT	104	123	105	98	167	111	130	93	99	99
	SGPT	57	61	54	56	45	49	57	51	51	48

Table 5.1: Serum enzyme level

The mean response matrix is

$$\bar{X} = \begin{pmatrix} 22.7 & 22.8 & 23.7 & 27.3 \\ 99.3 & 108.4 & 100.9 & 112.9 \\ 61.9 & 63.8 & 60.2 & 52.9 \end{pmatrix}$$

where the columns and rows represent the treatment groups (dosage) and response variables (enzyme level) respectively.

Consider testing the hypothesis

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4 \quad \text{vs} \quad H_1 : \mu_1 \preceq \mu_2, \mu_1 \preceq \mu_3 \text{ and } \mu_1 \preceq \mu_4.$$

In R^p, consider

$$C = \{x \in R^3 : x \geq 0 \text{ (component wise)}\}.$$

Clearly, under H_1, $\mu \in C_{p\times q}$, but $\bar{X} \notin C_{p\times q}$.

With $V = I_{3\times 3}$ (a 3x3 identity matrix), and $w_i = 10$, $i = 1, 2, 3, 4$, in calculating the pseudo RMLE for $\mu \in C_{p\times q}$ using the algorithm proposed in this dissertation,

Step 1 produces:

$$\bar{X}_* = \begin{pmatrix} 24.125 \\ 105.375 \\ 59.7 \end{pmatrix},$$

$$\pi(\bar{X}_2 - \bar{X}_1|C) = \begin{pmatrix} 0.1 \\ 9.1 \\ 1.9 \end{pmatrix}, \ \pi(\bar{X}_3 - \bar{X}_1|C) = \begin{pmatrix} 1 \\ 1.6 \\ 0 \end{pmatrix}, \ \pi(\bar{X}_4 - \bar{X}_1|C) = \begin{pmatrix} 4.6 \\ 13.6 \\ 0 \end{pmatrix}.$$

Step 2 yields,

$$\hat{X}_1 = \begin{pmatrix} 22.7 \\ 99.3 \\ 59.225 \end{pmatrix},$$

Step 3 yields,

$$\hat{X}_2 = \begin{pmatrix} 22.8 \\ 108.34 \\ 61.125 \end{pmatrix}$$

$$\hat{X}_3 = \begin{pmatrix} 23.7 \\ 91.8 \\ 59.225 \end{pmatrix}$$

$$\hat{X}_4 = \begin{pmatrix} 27.3 \\ 112.9 \\ 59.225 \end{pmatrix}.$$

So,

$$\hat{X} = \begin{pmatrix} 22.7 & 22.8 & 23.7 & 27.3 \\ 99.3 & 108.34 & 91.8 & 112.9 \\ 59.225 & 61.125 & 59.225 & 59.225 \end{pmatrix}$$

is a pseudo RMLE. Clearly, $\hat{X}$ is under multivariate simple-tree order restriction, i.e. $\hat{X} \in C_{p\times q}$.

In the next section, we study the behavior of the squared norm of $\hat{X} - \mu$ with varying sample sizes.

5.2 Simulation

One way to measure the performance of the algorithm is to look at how close the output is to the parameter with respect to the norm defined in (4.5). We used R statistical software for the simulation and following is the discussion of the results.

Here, we used two mean matrices, μ_a and μ_b. Each μ is chosen in a way that the components of the first column are at most equal to the corresponding components of the remaining columns. The MLE for μ_a is expected more often to not be in $C_{4\times 5}$ whereas the MLE for μ_b will most of the time be in $C_{4\times 5}$.

$$\mu_a = \begin{pmatrix} 10 & 10 & 10 & 10 & 10 \\ 4 & 4 & 4 & 4 & 4 \\ 84 & 84 & 84 & 84 & 84 \\ 22 & 22 & 22 & 22 & 22 \end{pmatrix}, \text{ and}$$

$$\mu_b = \begin{pmatrix} 8 & 13.0 & 11.5 & 15 & 14.5 \\ 4 & 8.0 & 18.0 & 12 & 11.0 \\ 70 & 84.5 & 84.1 & 89 & 86.0 \\ 10 & 26.0 & 24.9 & 25 & 25.4 \end{pmatrix}.$$

With respect the two mean matrices given above, here the following notations are used:

$\bar{X}_a$: sample mean from generated data with $\mu = \mu_a$

$\hat{X}_a$: a pseudo RMLE for $\mu_a \in C_{p\times q}$

$\bar{X}_b$: sample mean from generated data with $\mu = \mu_b$

$\hat{X}_b$: Pseudo RMLE for $\mu_b \in C_{p\times q}$.

For a given mean matrix $\mu = (\mu_1, ..., \mu_5) \in C_{4\times 5}$, a variance-covariance matrix $\Sigma > 0$ and a sample size n, a random multivariate data was generated from $N_4(\mu_i, \Sigma)$ for all $i = 1, ..., 5$, and the sample mean $\bar{X} \in R^{4\times 5}$ was computed. Using the algorithm proposed, we then calculate the pseudo RMLE $\hat{X}$, $\|\bar{X} - \mu\|_{4\times 5}$ and $\|\hat{X} - \mu\|_{4\times 5}$. This process was repeated

1000 times and the average values were stored. For computing the pseudo RMLE, $C \subset R^5$ is defined as

$$C = \{x = (x_1, ..., x_5) \in R^5 : x_i \geq 0 \text{ for all } i = 1, 2, 3, 4, 5\} \tag{5.9}$$

and in $\langle ., . \rangle_V$, $V = I_{4\times5}$ is used.

n	$\|\hat{X}_a - \mu_a\|_{4\times5}$	$\|\bar{X}_a - \mu_a\|_{4\times5}$
10	7.6873780	8.4061438
50	4.7452446	5.9936577
100	2.5947777	3.5047222
200	2.0646409	2.4548309
500	1.2942564	1.6495736
1000	0.6709823	0.8587125

Table 5.2: Norm of difference matrices for selected n when $\mu = \mu_a$.

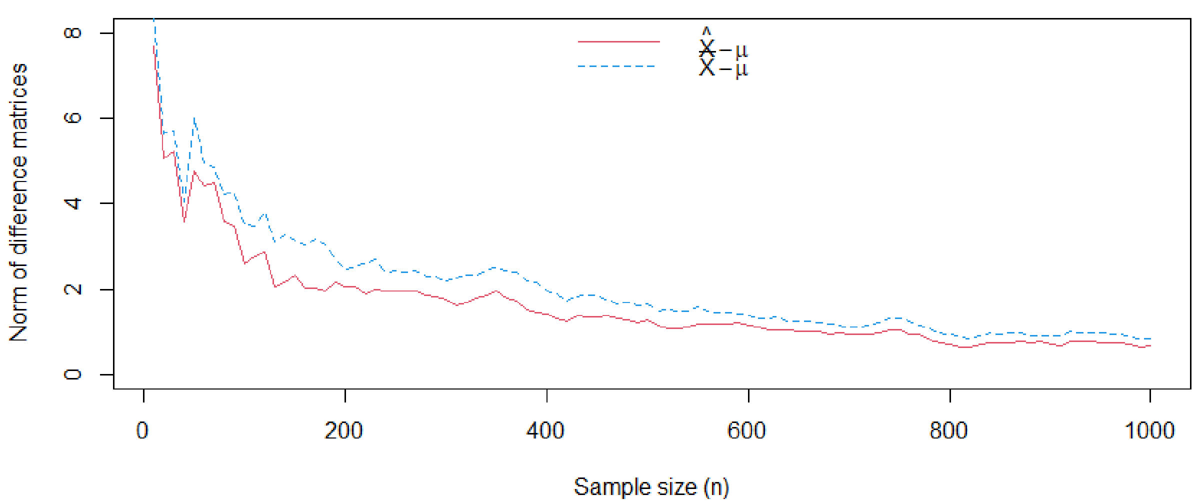

Figure 5.1: Norm of difference matrices for selected n when $\mu = \mu_a$

Table 5.2 presents values of $\|\hat{X}_a - \mu_a\|_{4\times5}$ and $\|\bar{X}_a - \mu_a\|_{4\times5}$ for selected sample sizes, n. Clearly, for each n in the table $\|\hat{X}_a - \mu_a\|_{4\times5} < \|\bar{X}_a - \mu_a\|_{4\times5}$. This comparison is also presented in figure 5.1. From the figure, it can be noted that the norm for $\hat{X} - \mu_a$ is at most equals to that of $\bar{X} - \mu_a$.

n	$\|\hat{X}_b - \mu_b\|_{4\times5}$	$\|\bar{X}_b - \mu_b\|_{4\times5}$
10	8.4061438	8.4061438
50	5.9936577	5.9936577
100	3.5047222	3.5047222
200	2.4548309	2.4548309
500	1.6495736	1.6495736
1000	0.8587125	0.8587125

Table 5.3: Norm of difference matrices for selected n when $\mu = \mu_b$.

Table 5.3 presents $\|\hat{X}_b - \mu\|_{4\times5}$ and $\|\bar{X}_b - \mu_b\|_{4\times5}$ for selected sample sizes. Clearly, all the values are identical and this is due to the fact that $\bar{X}_b$ is always in $C_{4\times5}$ and hence $\hat{X}_b = \bar{X}_b$. This is further shown in figure 5.2 as the two lines are superimposed over each other.

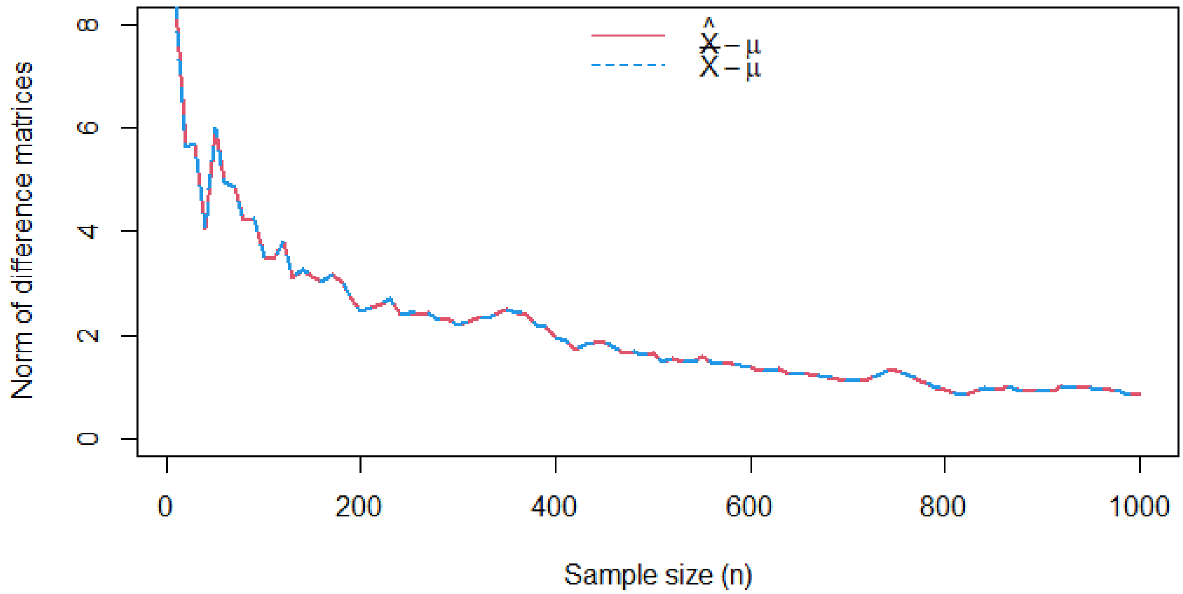

Figure 5.2: Norm of difference matrices for selected n when $\mu = \mu_b$

In both cases, as n increases both $\|\bar{X} - \mu\|_{4\times5}$ and $\|\hat{X} - \mu\|_{4\times5}$ are decreasing. This is because the sample mean, $\bar{X}$ is consistent i.e. it converges to the population mean matrix μ in probability.

CHAPTER 6

CONCLUSIONS

In MANOVA, it is shown that a parameter matrix constrained with a multivariate simple-tree order restriction belongs to a closed convex cone, i.e. $\mu \in C_{p\times q}$. The maximum likelihood estimator of such a parameter, called the restricted maximum likelihood estimator (RMLE), is a projection of its maximum likelihood estimator (MLE) under no order restriction, $\bar{X}$, onto the closed convex cone, $C_{p\times q}$, denoted by $\pi(\bar{X}|C_{p\times q})$. The computation of the RMLE could be cumbersome and numerically impossible in some cases.

For computational simplicity, we modified the domain by defining a closed convex subset $D(X)_{p\times q}$ of $C_{p\times q}$ and proposed an algorithm to compute the projection of $X \in R^{p\times q}$ onto $D(X)_{p\times q}$. This computation of matrix projection utilizes a relatively easy vector projection. In an MANOVA model, given sample mean matrix $\bar{X} \in R^{p\times q}$, $\pi(\bar{X}|D(\bar{X})_{p\times q})$ is introduced in this dissertation as a pseudo RMLE for $\mu \in C_{p\times q}$. When $q = 2$ or $\bar{X} \in C_{p\times q}$, the algorithm produces the RMLE for $\mu \in C_{p\times q}$, $\pi(\bar{X}|C_{p\times q})$. This means that the pseudo RMLE obtained using the algorithm is identical with the RMLE for $\mu \in C_{p\times q}$. However, when $q > 2$ and $\bar{X} \notin C_{p\times q}$, the pseudo RMLE produced by the algorithm is approximation of the RMLE for $\mu \in C_{p\times q}$. Furthermore, the result from the simulation shows that the pseudo RMLE is always a better estimate for $\mu \in C_{p\times q}$ as compared to its MLE i.e.
$\|\pi(\bar{X}|D(\bar{X})_{p\times q}) - \mu\|^2_{p\times q} \leq \|\bar{X} - \mu\|^2_{p\times q}$.

The pseudo RMLE proposed in this dissertation is relatively easy to compute since the algorithm utilizes only a vector projection. Besides, it has advantages as in most cases it gives the closed form of the RMLE, $\pi(\bar{X}|C_{p\times q})$, and it is a better estimator for $\mu \in C_{p\times q}$ compared to the MLE, $\bar{X}$.

Therefore for restricted statistical inference problems where the maximum likelihood estimator for a parameter matrix under multivariate simple-tree order restriction is the projection of unrestricted MLE onto the order restricted cone, the projection of MLE onto the subset of the order restricted cone can be utilized as a pseudo RMLE. MANOVA with q independent p-dimensional normal populations is such an example.

CHAPTER 7

FUTURE WORK

Many potential works are awaiting to be investigated in this area. However, as a continuation of the work in this dissertation, few extensions can be done immediately.

1. A computer program of the proposed algorithm can be developed and make it available for use in the statistical and/or mathematical softwares.

2. It could be of interest to study the properties of the pseudo RMLE. For this, it is essential to investigate the distribution function of the pseudo RMLE.

REFERENCES

REFERENCES

[1] Abelson, R. P. and Tukey, J. W. (1963). Efficient utilization of non-numerical information in quantitative analysis general theory and the case of simple order. *The Annals of Mathematical Statistics*, pages 1347-1369.

[2] Bartholomew, D. (1959). A test of homogeneity for ordered alternatives. *Biometrika*, 46(1/2): 36-48.

[3] Bartholomew, D. (1961a). Ordered tests in the analysis of variance. *Biometrika* 48(3/4): 325-332.

[4] Bartholomew, D. (1961a). A test of homogeneity of means under restricted alternatives. *Journal of the Royal Statistical Society: Series B (Methodological)* 23(2): 239-272.

[5] Boyd, S., Boyd, S. P., and Vandenberghe, L. (2004). *Convex optimization*. Cambridge university press.

[6] Brunk, H., Barlow, R. E., Bartholomew, D. J., and Bremner, J. M., (1972). Statistical inference under order restrictions.(the theory and application of isotonic regression). Technical report, Missouri Univ Columbia Dept of Statistics.

[7] Cohen,A., and Sackrowitz, H. (1996). Cone order association and stochastic cone ordering with applications to order-restricted testing. *The Annals of statistics*, 24(5): 2036-2048.

[8] Conaway, M., Pillers, C., Robertson, T., and Sconing, J. A circular-cone test for testing homogeneity against a simple tree order. *Canadian Journal of Statistics*, 19(3):283-296.

[9] Dykstra, R., Robertson, T., and Wright, F. T. (2012). *Advances in Order Restricted Statistical Inference: Proceedings of the Symposium on Order Restricted Statistical Inference Held in Iowa City, Iowa, September 11–13, 1985*, volume 37. Springer Science & Business Media.

[10] Fujino, Y. (1979). Tests for the homogeneity of a set of variances against ordered alternatives.*Biometrika*, 66(1):133-139.

[11] Geng, Z. and Shi, N.-Z. (1991). Isotonic regression in two independent variables under umbrella orderings. *Journal of the Japanese Society of Computational Statistics*, 4(1):49-61.

[12] Hamdan, M., and Hu, X. (2019). Unbiasedness of LRT on the homogeneity of means in MANOVA under general order restrictions. *Statistics & Probability Letters*, 148: 59-65.

REFERENCES

[13] Hu, X., and 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.

[14] Hu, X. (2020). A pseudo restricted MLE under multivariate order restrictions and its algorithm. *Communications in Statistics-Theory and Methods*, 49(1):169-177.

[15] Jacquelin Dietz, E. (1989). Multivariate generalizations of jonckheere's test for ordered alternatives. *Communications in Statistics-Theory and Methods*, 18(10):3763-3783.

[16] Mukerjee, H., Robertson, T., and Wright, F. (1986). A probability inequality for elliptically contoured densities with applications in order restricted inference. *The Annals of statistics*, pages 1544-1554.

[17] Peypouquet, J. (2015). *Convex optimization in normed spaces: theory, methods and examples*. Springer.

[18] Robertson, T. (1988). Order restricted statistical inference. No. 04; QA278. 7, R6.

[19] Sasabuchi, S., Inutsuka, M., and Kulatunga, D. (1983). A multivariate version of isotonic regression. *Biometrika*, 70(2):465-472.

[20] Sasabuchi, S., Inutsuka, M., and Kulatunga, D. S. (1992). An algorithm for computing multivariate isotonic regression. *Hiroshima Mathematical Journal*, 22(3):551-560.

[21] Sasabuchi, S., Tanaka, K., and Tsukamoto, T. (2003). Testing homogeneity of multivariate normal mean vectors under an order restriction when the covariance matrices are common but unknown. *The Annals of statistics*, 31(5):1517-1536.

[22] Silvapulle, M. J., and Sen, P. K. (2005).*Constrained statistical inference: Inequality, order and shape restrictions*. John Wiley & Sons.

[23] Zarantonello, E. H., (1972). Projections on convex sets in hilbert space and spectral theory: Part i. projections on convex sets: Part ii. spectral theory. In *Contributions to nonlinear functional analysis*, pages 237-424. Elsevier.