ÇUKUROVA UNIVERSITESI
INSTITUTE OF NATURAL AND APPLIED SCIENCES

PhD THESIS

Buatikan MIREZI

ADMISSIBILITY IN LINEAR MODELS

DEPARTMENT OF STATISTICS

ADANA-2023

ABSTRACT

PhD THESIS

ADMISSIBILITY IN LINEAR MODELS

Buatikan MIREZI

ÇUKUROVA UNIVERSITY
INSTITUTE OF NATURAL AND APPLIED SCIENCES
DEPARTMENT OF STATISTICS

Supervisor : Prof. Dr. Selahattin KAÇIRANLAR
Year: 2023, Pages: 202
Jury : Prof. Dr. Selahattin KAÇIRANLAR
: Prof. Dr. Sadullah SAKALLIOĞLU
: Prof. Dr. Hüseyin GÜLER
: Prof.Dr. Aşır GENÇ
: Dr. Öğr. Üyesi. Fikriye KURTOĞLU

In a statistical decision problem, the risk function is an indicator that reflects the merits of the decision function. Admissibility is one of the best decision criteria to minimize risk, which is to analyze the results of estimation and plays an important role to select a better decision function. Therefore, there are many different loss functions and admissibility studies have been done in the literature.

The key point of this article is to use the extended balanced loss functions (EBLF) which is more flexible compared to other loss functions to discuss the admissibility. Our study will be carried out by following steps, firstly, characterized the admissibility of linear estimators (AOLE) in different types of linear regression model (LRM)' s. Moreover, the admissibility of some known estimator and their properties under the EBLF are discussed. In the next step, a new more comprehensive loss function is proposed and under it, the linear admissibility of different types of regression coefficients (RC) is investigated. After that, some estimators are compared under the mean squares error (mse). Finally, a minimum matrix-valued risk estimator is proposed that combines the restricted least squares (RLS) and ordinary least squares (OLS) estimators. The results are supported by numerical examples and Monte Carlo simulation.

Key Words: Admissibility, Loss function, Linear model, Linear estimators, Risk.

ÖZ

DOKTORA TEZİ

LİNEER MODELLERDE KABUL EDİLEBİLİRLİK

Buatikan MIREZI

ÇUKUROVA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
İSTATİSTİK ANABİLİM DALI

Danışman : Prof. Dr. Selahattin KAÇIRANLAR
Yıl: 2023, Sayfa: 202
Jüri : Prof. Dr. Selahattin KAÇIRANLAR
: Prof. Dr. Sahdullah SAKALLIOĞLU
: Prof. Dr. Hüseyin GÜLER
: Prof.Dr. Aşır Genç
: Dr. Öğr. Üyesi. Fikriye KURTOĞLU

İstatistiksel bir karar probleminde risk fonksiyonu, karar fonksiyonunun performansını yansıtan bir göstergedir. Kabul edilebilirlik, tahmin sonuçlarını analiz etme riskini en aza indiren en iyi karar kriterlerinden biridir ve daha iyi bir karar fonksiyonu seçmek için önemli bir rol oynar. Bu nedenle literatürde birçok farklı kayıp fonksiyonu bulunmakta ve kabul edilebilirlik çalışmaları yapılmıştır.

Bu tezin temel noktası, kabul edilebilirliğin tartışılması için diğer kayıp fonksiyonlarına göre daha esnek olan genişletilmiş dengeli kayıp fonksiyonunun kullanılmasıdır. Çalışmamızda ilk olarak çeşitli lineer regresyon modellerinde tahmin edicilerin kabul edilebilirliği karakterize edilmiştir. Ayrıca, bu kayıp fonksiyonu altında bilinen bazı tahmin edicilerin kabul edilebilirliği ve özellikleri tartışılmıştır. Bir sonraki adımda, daha kapsamlı yeni bir kayıp fonksiyonu önerilir ve bu kayıp fonksiyonuna göre stokastik olan (olmayan) regresyon katsayılarının lineer kabul edilebilirliği araştırılmıştır. Daha sonra, birkaç kayıp fonksiyonu altında bazı tahmin ediciler karşılaştırılmıştır. Son olarak, kısıtlı ve sıradan en küçük kareler tahmin edicilerini birleştiren bir minimum matris değerli risk tahmin edici önerilmiştir. Bulunan sonuçlar Monte Carlo simülasyonu ve sayısal örnek ile desteklenmiştir.

Anahtar Kelimeler: Kabul edilebilirlik, Kayıp fonksiyonu, Lineer model, Lineer tahmin ediciler, Risk.

EXTENDED ABSTRACT

In the statistical linear model, which is one of the most important branches of Mathematical Statistics, the main problem is parameter estimation and then further statistical analysis depending on the estimation results. Although there are various types of parameter estimators in the literature, finding the best estimator and measuring the performance of these different estimators in terms of statistical decisions has always been an important research topic.

Famous statistician Abraham Wald proposed a unified mathematical model of statistics in 1950 and founded the statistical decision theory. This theory has made a great contribution to the development of modern statistics and has given rise to many new studies. In a statistical decision problem, choosing a better decision function is important in determining the goodness of the estimator. For this, it is necessary to create an indicator that reflects the character of the decision function. The risk function is one such indicator, the smaller the risk function, the better the decision is made. Under this principle, Admissibility criteria, minimization maximum criteria, bayesian criteria, and optimal homomorphism criteria are more specific and applicable criteria. Therefore, the selection of the loss function is very important. Discussing the superiority of different parameter estimation methods under different loss functions has significant theoretical value and practical application value, as acceptability is one of the best decision criteria to minimize its risk.

The AOLE has been extensively studied in the literature under different loss functions such as vector loss, quadratic loss, and matrix loss. The general loss function EBLF properties proposed by Shalabh (2009) and the AOLE with respect to the EBLF in the linear model are mainly investigated for the first time in this thesis. The content of this study is presented in seven chapters arranged as follows:

Chapter 1, the first chapter is the part of the introduction, where the recent developments of admissible estimators and the main work done in this thesis are introduced.

In Chapter 2, some information in the second part, including information about the LRM, and the basic concepts, theorems, and criteria used to determine the goodness of the estimators needed in the next chapters are given.

In Chapter 3, some biased estimators commonly used in statistics and econometrics are described and their statistical properties are examined.

In Chapter 4, the admissibility of the linear estimator and the structure of the admissible linear estimator (ALE) in different LRMs such as unrestricted model, restricted model, and elliptical restricted models are characterized under EBLF, which are more flexible than other loss functions. Sufficient and necessary conditions (SNC) for the linear estimator to be admissible are discussed in homogeneous and non-homogeneous classes (HHC), respectively. In addition, the admissibility of some known estimators such as OLS, RLS, general ridge (GR) estimator, and Bayes estimator according to EBLF was examined. In addition, the relationship between Obenchain class (OBC) and admissible estimators, and the shrinkage properties of admissible estimators were also investigated. Finally, ALEs on linear functions of the RC are considered in singular models.

In Chapter 5, in light of the above work, a more comprehensive new generalized extended balanced loss function (GEBLF) is proposed. The properties of the new loss function and the admissibility and the special case of linear estimators in the General Gauss-Markov (GGM) model are discussed. Considering the special condition that the RC have a distribution, under the proposed new loss function, the admissibility of stochastic regression coefficients (SRC) with random effects is discussed under several models.

In Chapter 6, firstly, Farebrother's proposed estimator with the ridge estimator according to the Matrix Mean Square Error (MSE) criterion under three different restrictions is compared. Secondly, a minimum matrix value risk estimator

is proposed that combines the RLS and OLS estimators. Finally, the above theoretical results are supported by numerical examples and Monte Carlo simulation.

In Chapter 7, the current fruits and the innovative points from the previous parts of the thesis are summarized.

GENİŞLETİLMİŞ ÖZET

Matematiksel istatistiğin en önemli dallarından biri olan istatistiksel lineer modelde, temel sorun ise parametre tahminidir ve ardından tahmin sonuçlarına bağlı olan daha ileri istatistiksel analiz yapılmasıdır. Literatürde çeşitli türde parametre tahmin edicileri olmasına rağmen, en iyi tahmin ediciyi bulmak ve bu farklı tahmin edicilerin performanslarını istatistiksel karar açısından ölçmek önemli bir araştırma konusudur.

Ünlü istatistikçi Abraham Wald, 1950'de birleşik bir matematiksel istatistik modeli önerdi ve istatistiksel karar teorisini kurdu. Bu teori, modern istatistiğin gelişimi üzerinde büyük bir katkı sağlamış ve birçok yeni araştırmaya yol açmıştır. İstatistiksel bir karar verme probleminde, daha iyi bir karar fonksiyonu seçmek tahmin edicinin iyiliğini belirlemekte önemlidir. Bunun için karar fonksiyonunun özelliğini yansıtan bir gösterge oluşturmak gerekir. Risk fonksiyonu böyle bir göstergedir. Kabul edilebilirlik kriteri, min-maks kriteri, Bayes kriteri ve optimal homomorfizm kriteri yukarıdaki ilke kapsamında daha spesifik ve uygulanabilir kriterlerdir. Bu nedenle, kayıp fonksiyonunun seçimi çok önemlidir. Kabul edilebilirlik, riski en az yapmak için en iyi karar kriterlerinden biri olduğu için, farklı parametre tahmin yöntemlerinin farklı kayıp fonksiyonları altında üstünlüğünü araştırmak önemli teorik değere ve pratik uygulama değerine sahiptir.

Lineer tahmin edicilerin kabul edilebilirliği, literatürde vektör kaybı, ikinci dereceden kayıp ve matris kaybı gibi farklı kayıp fonksiyonları altında kapsamlı bir şekilde incelenmiştir. Shalabh (2009) tarafından önerilen genel kayıp fonksiyonu EBLF'nun özellikleri ve EBLF'ye göre tahmin edicilerin lineer modelde kabul edilebilirliği esas olarak ilk kez bu tezde incelenmiştir. Bu çalışmanın içeriği aşağıdaki şekilde düzenlenmiş yedi bölüm halinde sunulmuştur:

1. Bölümde, kabul edilebilir tahmin edicilerdeki son gelişmeleri ve bu yazıda yapılan ana çalışmayı tanıttığımız giriş bölümüdür.

2. Bölümde, lineer regresyon modeli ile ilgili bilgiler ve sonraki bölümlerde ihtiyaç duyulan tahmin edicilerin iyiliğini belirlemek için kullanılan temel kavramlar, teorem ve kriterler verilmiştir.

Bölüm 3'te, ekonometri ve istatistikte kullanılan bazı yanlı tahminciler açıklanmakta ve istatistiksel özellikleri incelenmektedir.

Bölüm 4'te, diğer kayıp fonksiyonlarına göre daha esnek olan genişletilmiş dengeli kayıp fonksiyonları altında, kısıtsız model, kısıtlı model ve eliptik kısıtlı modeller gibi farklı lineer regresyon modellerinde lineer tahmin edicilerin kabul edilebilirliği ve kabul edilebilir lineer tahmin edicilerin yapısı karakterize edilmiştir. Lineer tahmin edicilerin kabul edilebilmesi için yeterli ve gerekli koşul sırasıyla homojen ve homojen olmayan sınıflarda ele alınmıştır. Ayrıca OLS, RLS, genel ridge tahmin edicisi ve bayes tahmin edicisi gibi bilinen bazı tahmin edicilerin EBLF'ye göre kabul edilebilirliği incelenmiştir. Ayrıca Obenchain sınıfı tahmin ediciler ile kabul edilebilir tahmin ediciler arasındaki ilişki, kabul edilebilir tahmin edicilerin büzülme özellikleri de incelenmiştir. Sonraki kısımda, regresyon katsayısının lineer fonksiyonları üzerinde kabul edilebilir lineer tahmin ediciler singular modellerde ele alınmıştır.

Bölüm 5'te, yukarıdaki çalışmanın ışığında, daha kapsamlı yeni bir genelleştirilmiş genişletilmiş dengeli kayıp fonksiyonu önerilmiştir. Yeni kayıp fonksiyonunun özellikleri ve GGM modelindeki lineer tahmin edicilerin kabul edilebilirliği ve özel durumu tartışılmıştır. Regresyon katsayılarının bir dağılıma sahip olması özel koşulu göz önüne alınarak, önerilen yeni kayıp fonksiyonu altında, rastgele etkilere sahip stokastik regresyon katsayılarının kabul edilebilirliği tartışılmıştır.

Bölüm 6'da, ilk olarak Farebrother'ın önerdiği tahmin edici, üç farklı kısıtlama altında hata kareleri ortalaması matrisi kriterine göre ridge tahmin edici ile karşılaştırılmıştır. Daha sonra, en küçük kareler tahmin edici ile sınırlı en küçük kareler tahmin edicisini birleştiren bir minimum matris değeri risk tahmin edicisi

önerilmiştir. Son olarak, yukarıdaki teorik sonuçlar, sayısal örnek ve Monte Carlo simülasyonu ile desteklenmiştir.

7. Bölüm'de tezin önceki bölümlerinde elde edilen sonuçlar ve yenilikçi noktaları özetlenmiştir.

ACKNOWLEDGEMENTS

My sincere gratitude goes to my supervisor, Prof. Dr. Selahattin KAÇIRANLAR. During my Ph.D., his invaluable knowledge, academic attitude, and expertise have deeply infected and inspired me, and will benefit me endlessly! Here, I am grateful and sincerely respect him for his careful guidance and opportunities for me!

At the same time, I would like to thank my thesis committee: Prof. Dr. Sadullah SAKALLIOĞLU and Prof. Dr. Hüseyin GÜLER, who share their useful information and opinions during my thesis work for their contribution to this work.

I would like to express my gratitude to all my teachers and friends in Turkey for their moral support and effort.

I would like to express my sincere thanks to my dear husband Abdurrahman MASUM, who was with me by providing peace and happiness at home during the completion of my thesis work.

Last but not least, I wish my express my heartfelt thanks to my parents and all of my family members too. Although I have not been able to communicate with my parents for five years, I feel their longing and prayers in my heart, they brought me to this day, who never spared her love and support, and taught me to take a courageous and determined stance against life's difficulties.

CONTENTS **PAGES**

ABSTRACT I
ÖZ II
EXTENDED ABSTRACT III
GENİŞLETİLMİŞ ÖZET VII
ACKNOWLEDGEMENTS XI
CONTENTS XII
LIST OF TABLES XVI
LIST OF FIGURES XVIII
SYMBOLS AND ABBREVIATIONS XX
1. INTRODUCTION 1
2. FUNDAMENTALS INFORMATION 5
2.1. LRM 5
2.2. Decision Theory and Decision Rule 9
2.2.1. Decision Rule 10
2.2.2. Loss Function and Risk 10
2.2.3. Selection Of Decision Rule 11
2.3. Admissibility 13
2.3.1 Linear estimators 14
2.3.2 The Parameter Space 15
2.4. Criteria Used to Determine Goodness of Estimators 16
2.4.1 Matrix Valued Quadratic Loss Function (QLF) and Mean Squared Error 16
2.4.2 Scalar Valued QLF and Mean Squared Error 18
2.4.3 Vector Valued QLF And Its Risk 19
2.4.4 Weighted QLF And Its Risk : 20
2.4.5 Mahalanobis Loss Function 20
2.4.6. Fisherian Loss Function 21

2.4.7. p' -norm Loss Function 21
2.4.8. Asymmetric Loss Function 22
2.4.9. PLF and Its Risk 23
2.4.10. Balanced Loss Function 23
2.4.11. PMSE Criterion Under the Target Functions 24
2.4.12. EBLF 26
2.5. Preliminary Knowledge 28
2.5.1. Some Concepts Necessary for Our Theses 28
2.5.2. Some Important Theorems About Matrix 28
2.5.3. Some Important Theorems Necessary for Our Thesis 31
3. SOME BIASED ESTIMATORS 37
3.1. Generalized Least Square Estimator 37
3.2. Weighted Least Square Estimator 38
3.3. Ridge estimator 38
3.4. GR Estimator 39
3.5. RLS Estimation 41
3.6. Farebrother Estimator 43
3.7. Bayes Estimator 46
4. CHARACTERIZATION OF ALEs UNDER EBLF 55
4.1. AOLE under EBLF 55
4.1.1. Explicit Characterization 55
4.1.2. Implicit Characterization 66
4.2. Characterization of ALE under the EBLF 67
4.2.1. OLS and ALE 68
4.2.2. Linear Transforms of OLS 71
4.2.3. About the Admissibility of Some Known Estimators 78
4.3. ALEs On Linear Functions of RC under EBLF 96
4.4. Linear Admissibility Under Linear Restrictions 103

4.5. Linear Admissibility Under Elliptical Restrictions 115
5. AOLE IN THE GGM MODEL UNDER GEBLF 119
5.1. A New Loss Function 119
5.2. Admissibility under the GEBLF 121
5.2.1. Example. 135
5.3. AOLE for the stocastic RC under the GEBLF 136
5.3.1. Example. 151
6. CC OF SOME ESTIMATORS 155
6.1. Comparision of Farebrother and Ridge Estimator 155
6.1.1. Case 1 (Under the unbiased restrictions) 155
6.1.2. Case 2 (Under the biased restrictions) 157
6.1.3. Case 3: r is fixed and $r = R\beta + \delta_1$ 159
6.1.2. The optimal choice of k 162
6.2. A Minimum Matrix Valued Risk Estimator Combining RLS and OLS Estimators 164
6.2.1. Introduction 164
6.2.2. Estimation Methods and Risk Functions 165
6.2.3. A Minimum mse – Matrix Estimator 166
6.2.4. A Numerical Example And Monte Carlo Simulation 175
7. RESULTS AND SUMMARY 179
REFERENCES 183
CURRICULUM VITAE 191
APPENDICES 193

LIST OF TABLES **PAGES**

Table 3.1. Linear ridge type estimators 198
Table 4.1. Some known estimators from the obenchain class 199
Table 6.1. The results of the Monte Carlo simulation for $n = 30$ 200
Table 6.2. The results of the Monte Carlo simulation for $n = 60$ 201
Table 7.1. Admissible estimators under the GEBLF and its special different types of criteria 202

LIST OF FIGURES **PAGES**

Figure 4.1. $\hat{\beta}^{(1)}$, $\hat{\beta}_{1/5}$ and true estimator when n= 30 78

Figure 4.2. n=30 ridge estimates 89

Figure.4.3. 30 combined ridge and principal components estimates

$0.5\hat{\beta}^{(1)} + 0.5\hat{\beta}_{1/5}$ 90

SYMBOLS AND ABBREVIATIONS

ALE : Admissible Linear Estimator
AOLE Admissibility of Linear Estimator
CC : Convex Combination
EBLF : Extended Balanced Loss Function
GBLF : Generalized Balanced Loss Function
GEBLF : Generalized Extended Balanced Loss Function
GGM : General Gauss – Markov
GLS : Generalized Least Squares
GR : General ridge
HHC : Homogeneous and non-Homogeneous Classes
HL : Homogeneous linear
LRM : Linear Regression Model
MDE : Mean Dispersion Error Or Matrix Valued Risk
MSE : Matrix Mean Square Error
mse : Scalar Mean Square Error
n.n.d : Non-Negative Definite
OBC : Obenchain class
OLS : Ordinary Least Squares
p.d : Positive Definite
PLF : Predictive Loss Function
QLF : Quadratic Loss Function
RC : Regression Coefficients
RLS : Restricted Least Squares
ShP : Shrinkage Property
SNC : Sufficient and necessary conditions
SRC : Stochastic Regression Coefficients
UNW : Uniformly not worse

WBLF : Weighted Balanced Loss Function

WBTLF : Weighted Balanced-Type Loss Function

ZBLF : Zellner Balanced Loss Function

1. INTRODUCTION

Linear models play a central role in modern statistical methods, as the most advanced tools are generalizations of the linear model, and the statistical methods associated with it are surprisingly versatile and robust. and a series of statistical models have been established in specific applications, such as analysis of variance models, LRMs, analysis of covariance models, etc. In this thesis, we assume that the underlying model is a LRM.

The LRM is a special type of linear model that is widely used. Which is a statistical tool used to estimate the relationship between the response (dependent) variable and explanatory (independent) variables. As to the linear model, the main problem is the prediction accuracy and interpretability of the model. These problems are evaluated with respect to the model coefficients. It means that the basic problem is parameter estimation followed by further statistical analysis which depends on the estimation results. Though there exist various kinds of biased and unbiased parameter estimators such as OLS and RLS in the literatures, it has always been important research subject to measure the performance of these different estimators from the view of the statistical decision.

In a statistical decision problem, to select a better decision function, it is necessary to establish an index reflecting the quality of the decision function. The risk function is such an indicator, the smaller the risk function, the better decision. Under this principle, admissibility is one of the more specific and feasible criteria. and the admissibility of parameter estimation is just one of the important statistical tools to solve such a problem.

The AOLE has been extensively examined in the literature under different loss functions such as vector loss, quadratic loss, and matrix loss. Some of these are as

follows: Cohen (1966), Rao (1976), L.R. LaMotte (1982), Mathew et al. (1984), Chen et al. (1985), Baksalary and Markiewicz (1985,1986,1988), Wu (1986), Klonecki and Zontek (1988), Stępniak (1989), Hoffmann (1995), Markiewicz (1996), Lu and Shi (2002), Groß and Markiewicz (2004).

In statistical decision theory, the most important problem is the selection of the appropriate loss function. On the basis of this, Zellner (1994) proposed a balanced loss function (ZBLF). ZBLF has received considerable attention in the literature. Some of these are as follows: Rodrigues and Zellner (1994), Wan (1994), Giles et al. (1996), Ohtani et al. (1997), Ohtani (1998, 1999), Dey et al. (1999), Gruber (2004), Akdeniz et al. (2005). In addition to these studies, Xu and Wu (2000), Cao (2009), Hu and Peng (2010), Cao and Kong (2013), Zhang and Gui (2014), Cao (2014), Cao and Shen (2016), Cao and He (2017) investigated the AOLE under ZBLF.

Shalabh (1995) has presented a predictive loss function (PLF), also called the weighted balanced loss function (WBLF). Liu and Yin (2020) investigated the admissibility in the GGM model with respect to an ellipsoidal constraint under WBLF.

Further, Cao (2014) has extended the ZBLF, combining Zellner's idea of balanced loss and Rao's idea of a unified theory of least squares. This loss function is called the generalized balanced loss function (GBLF). Cao (2014) investigated the admissibility of the SRC in the GGM model under GBLF.

Jafari Jozani et al. (2006) have defined a loss function which is called a weighted balanced-type loss function (WBTLF). Cao and He (2019) investigated the admissibility of the RC under WBTLF.

Shalabh et al. (2009) proposed the more general loss function (EBLF), an extended version of the PLF. EBLF is more flexible than ZBLF and extends ZBLF. On account of these advantages of the EBLF, Shalabh et al. (2009) have also compared the

risk performance of the OLS estimator and the Stein-rule estimator (SRE) under the EBLF.

EBLF has received considerable attention in the literature. Some recent studies are as follows: Özbay and Kaçıranlar (2019) studied the risk performance of some shrinkage estimators. Kaçıranlar and Dawoud (2019) derived the optimal EBLF estimators in a LRM under the EBLF. Özbay and Kaçıranlar (2017) discussed the performance of the adaptive optimal estimator of Farebrother (1975) under the EBLF. Zinodiny et al. (2017) considered the minimax estimation of the mean of the multivariate normal distribution with respect to EBLF. Chaturvedi and Shalabh (2014) Bayesian estimation of RC is discussed and obtained estimator is compared with another estimator with respect to the EBLF.

The main purpose of this thesis is to examine the properties of a new risk loss function in the field of linear model parameter estimation and to give some related research methods and results. Under the EBLF, the Admissibility of the linear estimator is to discuss and explore the property of acceptable estimators by linking with other loss functions to perform better predictive analysis that minimizes new risk.

2. FUNDAMENTALS INFORMATION

In this chapter, for the LRM, some preliminary information to be used in the later parts of the study and the criteria used to determine the goodness of the estimators are included.

2.1. LRM

In the real world, we encounter the plenty of situations where two variables or more variables such as $X=(x_1,...,x_p)$ and y have some dependencies. Regression model is one of the types of linear model that variables $x_1,...,x_p$ are quantitative and a powerful tool for studying this relationship between the variables.

According to whether the parameter or parameters in the model are included in the equation as a first-order polynomial, Regression model is divided into linear and non-linear models. It is also defined as simple or multiple linear regression according to the number of independent variables in the model.

Consider the multiple LRM of the form

$$y = X\beta + \varepsilon \tag{2.1.}$$

where n is the number of observations and p is the number of variables, y is an n-dimensional vector of observations on the dependent variable vector of responses, X is an $n \times p$ observation matrix of p independent variables with full column rank, $\beta \in \Re^p$ is a $p \times 1$ vector of unknown parameters, ε is an n-dimensional vector of disturbances. A multiple linear model is used in many fields for purposes such as estimation of RC, prediction of new observations and forecasting.

Fundamental assumptions of model (2.1.) can be listed as follows:

(i) $X \in \Re^{n \times p}$ is a non-stocastic matrix with $p < n$. Means that number of independent variable is strictly smaller than the number of observations. It is essencial for X has full column rank.

(ii) The matrix X has full column rank ($rk(X) = rk(X'X) = p$), means that there is no linear or close to linear relationship between the independent variables. If $rk(X'X) \neq p$, it may be cause the multicollinarity problem. In this case, parameter estimations cannot be made or may be inconsistent and estimation may be has the large variances.

(iii) $E(\varepsilon) = 0$ or equivalently, $E(y) = X\beta$. The elements of y are observable random vectors and depends only on X.

(iv) $Cov(\varepsilon) = \sigma^2 I_n$ with $\sigma^2 > 0$ or equivalently, $Cov(y) = \sigma^2 I_n$. Which includes both previous assumptions $Cov(\varepsilon_i) = \sigma^2$ and $Cov(\varepsilon_i, \varepsilon_j) = 0$. This means that variance of *y* is constant and the variables y and ε are uncorrelated with each other .

Assumptions (iii) and (iv) will write $\varepsilon \sim (0, \sigma^2 I_n)$ for short.

If the model satisfied these fundamental assumptions, then the model also called classic linear regression. Some additional assumptions can be added when necessary as follows:

(v) Random error with n- variate ε has normal distribution. Under which the maximum likelihood estimators have excellent properties. (v) is

usually introduced when a linear hypothesis test on β is required. Under normality $\varepsilon \sim N(0, \sigma^2 I_n)$ (equivalently $y \sim N(X\beta, \sigma^2 I_n)$), then the condition (vi) and (v) both satisfied and vice versa.

(vi) The inequality $p \geq 3$ is true; this assumption correlated with Stein estimator, it will often be true.

(vii) The identity $X'X = I_p$ is true. Which often not satisfied, but can always be obtained by an appropriate reparameterization of the model.

(viii) The parameter vector β satisfies $R\beta = r$ for a given matrix $R \in \Re^{m \times p}$ and a given $r \in \Re^m$. We will consider this in the next chapters.

In the (2.1) model, unknown parameters are called RC. It must be estimate. The aim in estimating the RC is to find an estimator as close as possible to the true points. The OLS method is common used to estimate unknown regression parameters. OLS obtained by minimizing the sum of squares of errors.

$$\begin{aligned} \varepsilon'\varepsilon &= \sum_{i=1}^{n} \varepsilon_i^2 \\ &= (y - X\beta)'(y - X\beta) \\ &= y'y - 2\beta' X'y + \beta' X'X\beta \end{aligned} \tag{2.2.}$$

We can find the normal equations

$$X'X\beta = X'y \tag{2.3.}$$

which minimize $\varepsilon'\varepsilon$ by diferentiating (2.2) with respect to the β and setting the result equal to zero. If X is matrix of full column rank ($rk(X) = rk(X'X) = p$), we can get the OLS estimator of β solving the normal equation

$$\hat{\beta} = (X'X)^{-1} X'y . \tag{2.4.}$$

Guass-markov theorem asserted that under the assumption $\varepsilon \sim (0, \sigma^2 I_n)$, OLS is the best linear unbiased estimator with the minimum variance. In the other words,

unbiased estimator, since $E(\varepsilon) = 0$

$$\begin{aligned} E(\hat{\beta}) &= E\left[(X'X)^{-1} X'y\right] = E\left[(X'X)^{-1} X'(X\beta + \varepsilon)\right] \\ &= \beta + (X'X)^{-1} X'E(\varepsilon) = \beta. \end{aligned} \tag{2.5.}$$

and

$$Var(\hat{\beta}) = Var\left[(X'X)^{-1} X'y\right] = \sigma^2 (X'X)^{-1} . \tag{2.6.}$$

The most important properties of the Gauss-Markov theorem is its distributional generalizability. If the assumption that the errors are normally distributed for the (2.1.) model, the maximum likelihood estimator of β is again given $\hat{\beta}$ as (2.4.). The Maximum likelihood estimator $\hat{\beta}$ has the minimum variance among the unbiased estimators of β .

Unfortunately, in practice, we may encounter the problem like that $X'X$ is singular or is not invertable under the some condition such as multicollinarity and high dimentionality. Which causes OLS coefficients to be unstable due to larger variance.

Alternative estimators to the OLS are required to overcome the above mentioned problem.

In the sence of above mentioned, the alternative estimators proposed need to be smaller variance than OLS. All of this arise some questions as follows:

- Is there any rules guarantees that the considered alternative linear estimator is better than the OLS estimator for at least one possible value of the unknown β ?
- Do there exist any standart to define “better” ?
- Can we find estimators which are better compared to any other estimator? under what conditions which is better than others?

We present a brief introduction to some important conception before seeking the answer the above questions.

2.2. Decision Theory and Decision Rule

In light of the related work of Abraham Wald, Ferguson (1967) and Berger (1985) made further contributions to Statistical decision theory.

Under the LRM, we are interested in estimates and analysis about the unknown parameters. As can be seen from the assumption (iii), since observable random vector y depends on unknown parameters, then it will be used estimates the unknowns. So that the decision rules in question are estimators $\beta_*(y)$ and $\sigma^2(y)$ of β and σ^2.

2.2.1. Decision Rule

Suppose we have to reach a decision β_d depending on p unknown quantities $\beta_1,...,\beta_p$. These quantities are combined in the parameter vector $\beta = (\beta_1,...,\beta_p)'$. Then:

- the set of all possible parameter vectors β is called parameter space and is denoted by Θ ;
- the set of all possible β_d is called decision space and is denoted by D .

Definition 2.1 (Decision Rule) . A mapping

$$\beta_* : \Re^n \to D$$
$$y \to \beta_*(y) = \beta_d$$

which assigns each observed value y exactly one decision β_d is called decision rule. In point estimation theory, the decision simply consists in accepting one of the possible values from Θ . Thus $D = \Theta$ and decision rules β_* are point estimators of β .

2.2.2. Loss Function and Risk

Loss function is a value of determines the extent of loss. If the decision is correct, then the loss is zero means that we have to admit when β is the true parameter and the decision $\beta_d = \beta_*(y)$ is taken. And it will take the larger value with an increasing level of incorrectness.

Definition 2.2 (Loss function). A mapping

$$L:\Theta\times D\to\square \tag{2.7.}$$

where $L(\beta,\beta_d)$ gives the loss when β is the true parameter taken from the parameter space Θ and the decision β_d is taken, is called loss function. It satisfies the following properties:

i. $\forall\beta\in\Theta$ and $L(\beta,\beta_d)>0$ for $\forall\beta_d$

ii. $L(\beta,\beta_d)=0$ for $\beta=\beta_d$.

Let β_* be an estimator of β, and $L\big(\beta,\beta_*(y)\big)$ shows the loss that desition $\beta_*(y)$ is taken for an observation y is given. Then if we consider the average loss with respect to all possible realizations of Y. Risk of this estimator under the corresponding loss function defined as.

Definition 2.3 (Risk). The expected loss

$$R(\beta,\beta_*)=E\big[L(\beta,\beta_*(Y))\big] \tag{2.8.}$$

of a decision rule $\beta_*(Y)$ is called the risk of $\beta_*(Y)$ when β is the true parameter.

2.2.3. Selection Of Decision Rule

Point estimation of β has only one decision rule to come to a decision. There are many different decision rules, and finding the most appropriate one is the key to

solving a particular decision problem. From the above definitions, we know that decision rules in question are estimators $\beta_*(y)$ of β . In this case, it is very important make a good decision by choosing the most suitable one from the alternative estimators. Compare the risks of the respective estimators for all possible values from the parameter space is a possible way.

Definition 2.4 . Let (Θ, D, L, Y) denote a decision problem and let $\beta_{*1}(Y)$ and $\beta_{*2}(Y)$ be two estimators of β . Then :

(i) The rule $\beta_{*1}(Y)$ is called uniformly not worse (UNW) than the rule $\beta_{*2}(Y)$ if

$$R(\beta, \beta_{*1}(Y)) \leq R(\beta, \beta_{*2}(Y)) \tag{2.9.}$$

is satisfied for all $\beta \in \Theta$.

(ii) The rule $\beta_{*1}(Y)$ is called uniformly better than the rule $\beta_{*2}(Y)$ if $\beta_{*1}(Y)$ is UNW than $\beta_{*2}(Y)$ and in addition

$$R(\beta, \beta_{*1}(Y)) < R(\beta, \beta_{*2}(Y)) \tag{2.10.}$$

for at least one . $\beta \in \Theta$.

One may call $\beta_{*1}(Y)$ uniformly strictly better than $\beta_{*2}(Y)$ if

$$R(\beta, \beta_{*1}(Y)) < R(\beta, \beta_{*2}(Y)) \tag{2.11.}$$

for all $\beta \in \Theta$.

OLS $\hat{\beta}$ is the uniformly strictly better than $\beta_{*2}(Y)$, where $\beta_{*2}(Y)$ belongs to the set of unbiased linear estimators.

The term admissibility, which will be introduced in the next section, is used as widely in decision theory as 'loss' and 'risk'.

2.3. Admissibility

A number of point estimators for β have already been introduced in the literature. In applications, some may be more useful than others under certain conditions. This sparks the curiosity to examine which estimators perform well relative to each other or to a class, and under what conditions. Which is the main issue of the admissibility.

Admissibility is one best decision criterion which is to measure the performance of these different estimators from the view of statistic decision and plays an important role in order to select a better decision function.

Definition 2.5 Let $\beta_{*1}(y)$, $\beta_{*2}(y)$ and $\beta_{*}(y)$ be three estimators of β, then $\beta_{*1}(y)$ is said to be better than $\beta_{*2}(y)$ if $R\left(\beta_{*1}(y);\beta,\sigma^2\right) \le R\left(\beta_{*2}(y);\beta,\sigma^2\right)$ for all (β,σ^2) with strict inequality holding for at least some point. If there does not exist estimators which are better than $\beta_{*}(y)$ in the class of some estimators L, then $\beta_{*}(y)$ will be said to be admissible in L.

Admissibility is not a criterion for choosing a decision rule (estimator of β). An admissible estimator is not necessarily a reasonable estimator. Its means that different estimator may be better in different regions or different estimator may be better according to different criteria. From the above statements, we can see that admissibility is a property which depending on specific context as follows.

1) The set of point estimators that we consider and from which we take the possible competitors of a specific point estimator.
2) The loss function and the corresponding risk function we apply for the assessment of a point estimator.
3) The parameter space we assume the unknown parameter vector to lie in.

2.3.1 Linear estimators

Linear estimators of vector $\beta \in \Re^p$ are of the form

$$\beta = Ly + a,$$

where L is a $p \times n$ real matrix and a is a p - dimensional vector. If $a = 0$ is zero vector, then the linear estimator is called homogeneous linear (HL) estimator , while in case $a \neq 0$ the linear estimator is called non- homogeneous.

The OLS estimator $\hat{\beta}$ is a homogeneous estimator with $L = (X'X)^{-1}X$ and $a = 0$.

We will consider the admissibility in the class of homogeneous and non-HL estimators, respectively

$$\mathrm{L}^h(\beta) = \left\{ Ly : L \in \Re^{p \times n} \right\}$$

and for the non-HL estimators

$$\mathrm{L}(\beta) = \left\{ Ly + a : L \in \Re^{p \times n}, a \in \Re^p \right\}.$$

2.3.2 The Parameter Space

As for the parameter space in which the unknown parameter vector to lie in, there are three different parameter spaces that reflect different states of prior knowledge about the β that need to be considered.

Case 1. When we consider the LRM without any further assumption on the parameter vector β except fundamental assumptions (i) and (iv).

We assume that true parameter $\beta \in \Re^p$, then parameter space Θ_β for β is given by $\Theta_\beta = \Re^p$. If an estimator has not greater risk for all values $\beta \in \Theta_\beta$ and all values $\sigma^2 \in (0,\infty)$, then it is uniformly better than another. Hence , the whole parameter space in this case defined as

$$\Theta(\beta,\sigma^2) = \Theta_\beta \times (0,\infty) = \Re^p \times (0,\infty)$$

Case 2. When we consider the parameter vector β with linear restrictions $R\beta = r$ for some known matrix R and some known vector r. In this case, the whole parameter space is given by

$$\Theta(\beta,\sigma^2) = \Theta_\beta \times (0,\infty) = \{\beta \in \Re^p : R\beta = r\} \times (0,\infty)$$

Case 3. When we assumed the unknown parameters β and σ^2 are satisfy the relationship $\beta' T\beta \le \sigma^2$ for some unknown symmetric positive definite (p.d) matrix $T \in \Re_s^>$, the whole parameter space is

$$\Theta = \Theta_{\beta} = \left\{ \beta \in \Re^{p} : \beta' T \beta \le \sigma^{2} \right\}.$$

2.4. Criteria Used to Determine Goodness of Estimators

Before giving some criteria, let's take a look at the following definitions that we will use frequently from now on.

Definition 2.6. An $n \times n$ symmetric matrix $A \in \Re^{n \times n}$ has the property

$$x'Ax > 0$$

for all $n \times 1$ vector x except $x = 0$, then A is called positive definite (p.d) matrix.

Definition 2.7. For an $n \times n$ symmetric matrix $A \in \Re^{n \times n}$ and $n \times 1$ vector $\forall x \neq 0$, if

$$x'Ax \ge 0$$

then A is called non negative definite (n.n.d) matrix.

2.4.1 Matrix Valued Quadratic Loss Function (QLF) and Mean Squared Error

According to the Definition 2.2, for estimating a unknown parameter $\beta \in \Re^{p}$, if $\beta_* = \beta_*(y)$ be an estimator of β the matrix – valued quaratic loss function as

$$L(\beta, \beta_*) = (\beta_* - \beta)(\beta_* - \beta)'$$

The corresponding risk i.e. MSE is given by

$$\begin{aligned} MSE(\beta,\beta_*) &= E\left[(\beta_*-\beta)(\beta_*-\beta)'\right] \\ &= \mathrm{cov}(\beta_*) + bias(\beta_*)bias(\beta_*)' \end{aligned} \tag{2.12.}$$

where $\mathrm{cov}(\beta_*) = E\left[(\beta_* - E(\beta_*))(\beta_* - E(\beta_*))'\right]$, $bias(\beta_*) = E(\beta_*) - \beta$.

According to Definition 2.4, if we want to compare estimators with respect to MSE, we have modified the previous definition accordingly.

Definition 2.8. Let (Θ, L, Y) denote a estimation problem and let $\beta_{*1} = \beta_{*1}(Y)$ and $\beta_{*2} = \beta_{*2}(Y)$ be two estimators of β. Then :

(i) The estimator β_{*1} is called UNW than the rule β_{*2}, if the symmetric matrix

$$MSE(\beta,\beta_{*2}) - MSE(\beta,\beta_{*1}) \geq 0 \tag{2.13.}$$

is n.n.d for all $\beta \in \Theta$.

(ii) The estimator β_{*1} is called uniformly better than β_{*2}, if β_{*1} is UNW than β_{*2} and in addition

$$MSE(\beta,\beta_{*2}) \neq MSE(\beta,\beta_{*1}) \tag{2.14.}$$

for at least one $\beta \in \Theta$.

Some authors define β_{*1} is called uniformly better than β_{*2} , if the $MSE(\beta,\beta_{*2})-MSE(\beta,\beta_{*1})$ is p.d for all $\beta\in\Theta$. Correspondingly, Groß (2003) call β_{*1} uniformly strictly better than β_{*2} if

$$MSE(\beta,\beta_{*2})-MSE(\beta,\beta_{*1})>0 \tag{2.15.}$$

for some $\beta\in\Theta$.

He also stated that it is essential not to mix up both definitions since they are not equivalent. Clearly, if the difference $MSE(\beta,\beta_{*2})-MSE(\beta,\beta_{*1})$ is p.d for some $\beta\in\Theta$, then this difference is also n.n.d and in addition nonzero , but the converse is not true, i.e, from 'uniformly better ' in our sense does not necessarily follow 'uniformly strictly better '.

2.4.2 Scalar Valued QLF and Mean Squared Error

The most commonly used loss function to evaluate the performance of a estimator of β is the scalar valued QLF defined as

$$L\left(\beta,\beta_{*}\right)=(\beta_{*}-\beta)'(\beta_{*}-\beta) \tag{2.16.}$$

the corresponding risk is given by

$$mse\left(\beta,\beta_{*}\right)=E\left[(\beta_{*}-\beta)'(\beta_{*}-\beta)\right] \tag{2.17.}$$

We can see relation between the MSE and mse .

$$mse(\beta,\beta_*) = trE\left[(\beta_* - \beta)(\beta_* - \beta)'\right] = tr\left[MSE(\beta,\beta_*)\right]$$

As we know from the definition above, MSE is a better criterion than mse because mse is a special case of MSE when we ignore non-diagonal elements of MSE and weight all $(\beta_{*i} - \beta_i)$ s equally. However, since MSE is more difficult to calculate in practice, mse is generally preferred as the main criterion.

2.4.3 Vector Valued QLF And Its Risk

The MSE might be replaced by the vector -valued quadratic loss such as

$$L(\beta,\beta_*) = dg\left[(\beta_* - \beta)(\beta_* - \beta)'\right]$$

where $\beta_* = \beta_*(Y)$, $dg(A)$ denotes the vector formed by i-th element of the main diagonal of the matrix A .

For arbitrary $\beta \in \Theta$

$$MSE(\beta,\beta_{*1}) \le MSE(\beta,\beta_{*2})$$
$$\Rightarrow dg\left[MSE(\beta,\beta_{*1})\right] \le dg\left[MSE(\beta,\beta_{*2})\right]$$
$$\Rightarrow tr\left[MSE(\beta,\beta_{*1})\right] \le tr\left[MSE(\beta,\beta_{*2})\right]$$

Similary

$$MSE(\beta,\beta_{*1}) \neq MSE(\beta,\beta_{*2})$$
$$\Rightarrow dg\left[MSE(\beta,\beta_{*1})\right] \neq dg\left[MSE(\beta,\beta_{*2})\right]$$
$$\Rightarrow tr\left[MSE(\beta,\beta_{*1})\right] \neq tr\left[MSE(\beta,\beta_{*2})\right]$$

showing that this criterion is stronger than mse but weaker than MSE.

2.4.4 Weighted QLF And Its Risk :

Let $\beta_* = \beta_*(Y)$ be an estimator of β, for any $\beta \in \Theta$, weighted quadratic loss is defined as

$$L_W(\beta,\beta_*) = (\beta_* - \beta)'W(\beta_* - \beta) \tag{2.18.}$$

for symmetric p.d matrix W. If $W = I$, the scalar-valued QLF is obtained. Using a $W \neq I$ weighted matrix can sometimes be useful for estimation problems. For example, when diagonal elements of W are different weighted diagonal matrix, the individual distance squares $(\beta_{*i} - \beta_i)^2$ will greater or lesser effects on the total loss. Corresponding risk is given by

$$wmse(\beta,\beta_*) = trE\left[(\beta_* - \beta)W(\beta_* - \beta)'\right] = trWMSE(\beta,\beta_*) \tag{2.19.}$$

2.4.5 Mahalanobis Loss Function

Let $\beta_* = \beta_*(y)$ be an estimator of β, $\mathrm{var}(\beta_*)$ is variance-covariance matrix of β_*. Mahalanobis loss function (MLF) is defined as

$$L_M\left(\beta,\beta_*\right)=(\beta_*-\beta)'\,\mathrm{var}^{-1}(\beta_*)(\beta_*-\beta) \tag{2.20.}$$

MLF was recommended by Mahalanobis. It is also called the minimum distance rule in the study of classification of the normal population. It is more appropriate to use it for comparison of estimators, at least one of which is a biased estimator. If both are unbiased estimators, MLF is insensitive, unable to distinguish either estimator.

2.4.6. Fisherian Loss Function

Let $\beta_*=\beta_*(y)$ be an estimator of β, $\sum$ is variance-covariance matrix of β_*. Fisherian loss function (FLF) is given by

$$L_F\left(\beta,\beta_*\right)=(\beta_*-\beta)'\Sigma^{-1}(\beta_*-\beta) \tag{2.21.}$$

2.4.7. p' -norm Loss Function

Let $\beta_*=\beta_*(y)$ be an estimator of β, and β_{*_i}, β_i denote the i-th row elements of β_* and β respectively. Then

$$L_\beta^{p'}(\beta_*)=\left\{\sum_i\left|\beta_{*_i}-\beta_i\right|^{p'}\right\}^{1/p'} \tag{2.22.}$$

it called p'-norm loss function. If we take $p=1,2\ ve\ \infty$, then p'-norm loss functions as follows rexpectively

$$L_{\beta}^{1}(\beta_{*}) = \sum_{i} \left| \beta_{*_i} - \beta_i \right| \tag{2.23.}$$

$$L_{\beta}^{2}(\beta_{*}) = \left\{ \sum_{i} \left| \beta_{*_i} - \beta_i \right|^2 \right\}^{1/2} \tag{2.24.}$$

$$L_{\beta}^{\infty}(\beta_{*}) = \max \left\{ \left| \beta_{*_i} - \beta_i \right|, i = 1,..., p \right\} \tag{2.25.}$$

2.4.8. Asymmetric Loss Function

LINEX means linear exponential loss function which used in the analysis of statistical estimation and prediction problem which containes the term of exponentially and almost linearly. It is used in both overestimation and underestimation problems. Asymmetric LINEX loss function is proposed by Varian(1975), which defined as

$$L_{L}(\beta_{*}, \beta) = b\left[\mathrm{e}^{a\Delta} - a\Delta - 1 \right], \quad a \neq 0, b > 0, \Delta = \beta_{*} - \beta \tag{2.26.}$$

where b serving to scale the loss function and a is determine its shape. The sign of the shape parameter a represents the direction of the asymmetry and the magnitude of a represents the degree of asymmetry. Generally, $a > 0$ ($a < 0$) is taken if the overestimation is more important (less important) than the underestimation. For small values of $|a|$, the LINEX loss function is almost symmetric and yields similar results as a corresponding QLF. The properties of LINEX have been investigated in detail by Zellner (1986).

2.4.9. PLF and Its Risk

Let β_* be an estimator of β . PLF is defined as

$$L_P(\beta,\beta_*)=(y-X\beta_*)'(y-X\beta_*).$$

Usually, the estimation of y for future X values is made by using the $X\beta_*$ estimator. Then corresponding risk

$$pmse(\beta,\beta_*)=E\left[(y-X\beta_*)'(y-X\beta_*)\right] \tag{2.27.}$$

is called scalar prediction mean square error (pmse), if matrix-valued prediction mean square error (PMSE) is as follows

$$PMSE(\beta,\beta_*)=E\left[(y-X\beta_*)(y-X\beta_*)'\right] \tag{2.28.}$$

2.4.10. Balanced Loss Function

Quadratic and LINEX loss functions only emphasize measuring the goodness of the prediction. Zellner (1994) proposed a balanced loss function in the evaluation of an estimator, combining both goodness of fit and precision of the estimation. For any estimator β_* of β :

$$ZBLF(\beta_*,\beta)=\lambda(y-X\beta_*)'(y-X\beta_*)+(1-\lambda)(\beta_*-\beta)'X'X(\beta_*-\beta) \tag{2.29.}$$

where λ is a scalar lying between 0 and 1 which provides assigned to the goodness of fit of model. The components $\left(y - X\beta_*\right)'\left(y - X\beta_*\right)$ and $(\beta_* - \beta)'X'X(\beta_* - \beta)$ reflect the goodness of fit and precision of estimation, respectively of β_*. Corresponding risk of ZBLF is

$$ZBLF(\beta_*, \beta) = \lambda E\left[\left(y - X\beta_*\right)'\left(y - X\beta_*\right)\right] + (1-\lambda)E\left[(\beta_* - \beta)'X'X(\beta_* - \beta)\right].$$

Zeller introduced the ZBLF which concentrates of estimates around true parameter and goodness of fit of model. Since its definition, many researchers have studied on the GR regression estimator. Some of these are as follows: Rodrigues and Zellner (1994), Wan (1994), Giles et al. (1996), Ohtani (1998) studied the risk function of some specific estimators. Shalabh (1995), Gruber (2005), Akdeniz et al. [8] did some work on the application of the ZBLF. In addition to these studies, Xu and Wu (2000), Cao (2009), Hu and Peng (2010), Cao and Kong (2013), Zhang and Gui (2014), Cao (2014), Cao and Shena (2016), Cao and He (2017) investigated the AOLE under ZBLF.

2.4.11. PMSE Criterion Under the Target Functions

Let β_* be an estimator of β. Generally, in a LRM, prediction are made for the actual values or average values $E(y) = X\beta$ of the observations y at a given time and $\hat{T}_l = X\beta_*$ is using for this. However, in some cases it may be necessary to consider the predictions of the actual and average values simultaneously. For this reason Shalabh (1995) has defined a target function for the purpose of simultaneous prediction of actual and average values of y as

$$T_l = \lambda y + (1-\lambda) E(y)\,, \tag{2.30.}$$

where λ is a scalar between 0 and 1. Loss function under target function for $\hat{T}$ predictor of T .

$$L_T(\hat{T}_l) = \left(\hat{T}_l - T_l\right)\left(\hat{T}_l - T_l\right)'$$

Then corresponding PMSE matrix

$$PMSE(\hat{T}_l) = E\left[\left(\hat{T}_l - T_l\right)\left(\hat{T}_l - T_l\right)'\right]$$

Scalar PMSE is

$$pmse(\hat{T}_l) = E\left[\left(\hat{T}_l - T_l\right)'\left(\hat{T}_l - T_l\right)\right]$$

The following PLF arises when we use the predictor $X\beta_*$ for simultaneous prediction of actual and average values of y through the target function (2.30.)

$$\begin{aligned} PLF(\beta_*, \beta, \lambda) &= (X\beta_* - T_l)'(X\beta_* - T_l) \\ &= \lambda^2 \left(y - X\beta_*\right)'\left(y - X\beta_*\right) + (1-\lambda)^2 (\beta_* - \beta)' X'X (\beta_* - \beta) \\ &\quad + 2\lambda(1-\lambda)(X\beta_* - y)' X (\beta_* - \beta) \end{aligned} \tag{2.31.}$$

PLF not only incorporates the ZBLF as its particular case but also measures the correlation between the goodness of fit of model and concentration of estimates around the true parameter.

2.4.12. EBLF

Appereciating the popularity of the ZBLF, Shalabh.et.al (2009) proposed the weighted loss function:

$$EBLF(\beta_*, \beta, \lambda_1, \lambda_2) = \lambda_1 \left(y - X\beta_* \right)' \left(y - X\beta_* \right) + \lambda_2 (\beta_* - \beta)' X'X (\beta_* - \beta)$$
$$+(1-\lambda_1-\lambda_2)(X\beta_* - y)' X (\beta_* - \beta) \qquad (2.32.)$$

where $\lambda_1, \lambda_2 \in [0,1]$ are characterizing the loss functions. The weighted loss function in Equation (2.32.) can be called the EBLF. EBLF encompass the following particular cases:

i. When $\lambda_1 = 1$ and $\lambda_2 = 0$, the EBLF in Equation (2.32.) reduces to the mse which is the criterion of goodness of fit of model.

ii. If we put $\lambda_1 = 0$ and $\lambda_2 = 1$, the EBLF in Equation (2.32.) reduces to the weighted QLF of (2.18.) which is the criterion of precision of estimation.

iii. If $\lambda_1 = 0$ and $\lambda_2 = 0$, the EBLF in Equation (2.32.) reduces to the component $(X\beta_* - y)' X (\beta_* - \beta)$ which is the criterion of interaction or covariation between the precision of estimation and goodness of fit.

iv. When we take $\lambda_1 = w$ and $\lambda_2 = 1 - w$, where $0 < w < 1$, we obtain the ZBLF in Equation (2.29.).

Then , from Equations (2.1.) and (2.30), Shalabh et al (2009) observed that

$$EBLF(\tilde{\beta},\beta,\lambda_1,\lambda_2)=\lambda_1\varepsilon'\varepsilon+(1-\lambda_1-\lambda_2)\varepsilon'X(\tilde{\beta}-\beta)+(\tilde{\beta}-\beta)'X'X(\tilde{\beta}-\beta) \tag{2.33.}$$

and the risk of any estimator β is the expected value of the EBLF which is calculated by

$$\mathrm{R}(\tilde{\beta},\beta,\lambda_1,\lambda_2)=E(EBLF(\tilde{\beta},\beta,\lambda_1,\lambda_2)). \tag{2.34.}$$

The main advantage of the EBLF is that it is more flexible in comparision to (2.29.) and (2.31.). The ZBLF in (2.29.) takes care only either the precision of estimation and the goodness of fit whereas (2.31.) extends it by considering the interaction or covariation between the precision of estimation and goodness of fit. Ignoring this covariation in the formulation of loss function may lead to wrong statistical inferences. The weights assigned in (2.31.) for precision of estimation, goodness of fit and their covariation depends only on one factor λ assigned to one of the criterion only. The EBLF in (2.32.) provides a more flexible options of choosing different weights for precision of estimation, goodness of fit and covariation terms and, in the presence of covariation, it will obviously lead to improved statistical inferences (Charturvedi and Shalabh, 2014).

2.5. Preliminary Knowledge

2.5.1. Some Concepts Necessary for Our Theses

Definition 2.9. Let $A \in \Re^{n\times n}$ and $B \in \Re^{n\times n}$ be any two matrics. Then the roots $\lambda_i = \lambda_i^B(A)$ of the equation

$$|A - \lambda B| = 0$$

are called the eigenvalues of A in the metric of B.

Definition 2.10. In the model $y = X\beta + \varepsilon$, where $E(y) = X\beta$ and X is $n \times p$ of rank $k < p < n$. A linear function of parameters $C\beta$ is said to be estimable if all columns of C' are in the columns space of X'.

Definition 2.11. Let F be a class of estimators. For arbitrary $d(y) \notin \mathrm{F}$, if there exists an estimator $d_1(y) \in \mathrm{F}$ such that $d_1(y)$ is better than $d(y)$, then F is called a complete class (Lehmann and Casella, 2005).

Thus from this definition, it is known that we only need to search admissible estimators in the complete class.

2.5.2. Some Important Theorems About Matrix

Theorem 2.1. Let $A \in \Re^{m\times m}$ be a matrix, $G \in \Re^{n\times m}$ a matrix with full column rank. In this case, the necessary and sufficient condition for A to be symmetric n.n.d is that the GAG' matrix is symmetric n.n.d (Rao and Toutenburg, 1995).

Theorem 2.2. Let $A \in \Re^{m\times m}$ be a symmetric p.d matrix, $a \in \Re^{m}$ a m - dimensional vector, and $\alpha > 0$ an arbitrary scalar. The necessary and sufficient condition for $\alpha A - aa' \geq 0$ is that $a'A^{-1}a \leq \alpha$ (Farebrother, 1976).

Theorem 2.3. Let y be an $n\times 1$ random vector with $E(y) = \mu$ and $Cov(y) = V$ and let W be an $n\times n$ non- stochastic matrix. Then the identity

$$E(y'Wy) = tr(WV) + \mu'W\mu$$

holds true. Note that the above result holds for any not necessarily symmetric square matix W .

Theorem 2.4. Let A and B be $n\times n$ non-negative matrices; then

(a) $tr(AB) \geq 0$,

(b) $tr(AB) = 0$ if and only if $AB = 0$.

Theorem 2.5. Let A be an $m\times n$ matrix and let B be an $n\times m$ matrix. Then the nonzero eigenvalues of AB coincide with the non zero eigenvalues of BA, counting multiplicities.

Theorem 2.6. Let A and B be two symmetric $m\times m$ matrices satisfying $AB = BA$. Then there exists an $m\times m$ orthogonal matrix U such that

$$A=U\Lambda U' \text{ and } B = U\Gamma U' ,$$

where Λ and Γ are two $m \times m$ diagonal matrices.

The above result ensures that two commuting symmetric matrices are simultaneously diagonalizable via a single orthogonal transformation.

Theorem 2.7. Let A and B be two symmetric p.d. $m \times m$ matrices. Then $A \leq_L B$ if and only if $B^{-1} \leq_L A^{-1}$.

Theorem 2.8. Let $A > 0$ (or $A \geq 0$) and $B > 0$. Then $B - A > 0$ if and only if $\lambda_i^B(A) < 1$.

Theorem 2.9. Let A and B be two symmetric n.n.d. $m \times m$ matrices with $A \leq_L B$. Then

$$\lambda_i \leq \mu_i, \ \forall i \in \{1,...,m\},$$

where $\lambda_1 \leq ... \leq \lambda_m$ are the ordered eigenvalues of A and $\mu_1 \leq ... \leq \mu_m$ are the ordered eigenvalues of B.

Theorem 2.10. Let A and B be two symmetric n.n.d. $m \times m$ matrices. Then the relation $A \leq_L B$ holds true if and only if the conditions

$$C(A) \subseteq C(B) \text{ and } AB^{+}A \leq_L A,$$

are satisfied.

Theorem 2.11. The p eigenvalues of symmetric $p \times p$ matrix A are real.

Theorem 2.12. If P, Q and R are rectangular matrices such that the product PQR is defined. Then

$$rank(PQR) + rank(Q) \geq rank(PQ) + rank(QR).$$

which is called the Frobenius inequality in the literature (Banerjee and Roy, 2014).

2.5.3. Some Important Theorems Necessary for Our Thesis

The following Theorem 2.13 is essentially due to Rao (1976) as extended by Chen et al. (1985).

Theorem 2.13. Let $\mathrm{L}^* = \left\{ Lz : L \in \Re^{k \times n} \right\}$ and

$$\begin{cases} z = X_0\beta + u; \\ E(u) = 0, E(uu') = \sigma^2 V, \end{cases} \tag{2.35.}$$

where, $X_0 : n \times p$ and $V > 0$ is known, $\beta \in \Re^p$ and $\sigma^2 > 0$ is unknown. If C is $k \times p$ matrix and $C\beta$ is estimable, then $Lz \overset{\mathrm{L}^*}{\square} C\beta$ holds under loss function

$$LF_I(\beta_*, C\beta) = (\beta_* - C\beta)'(\beta_* - C\beta),$$

if and only if the following statements are valid:

$(a)\ L = LX_0T^{-}X_0'V^{-1}$

$(b)\ LX_0T^{-}X_0'L' \le LX_0T^{-}C'$,

where $LX_0T^{-}C'$ is symmetric, $T = X_0'V^{-1}X_0$ and T^{-} is g - inverse of T.

We can see that if $k = p$, the set L^{*} turns in to L^{h}.

Theorem 2.14. In the linear model (2.35.), if $\hat{F}z \overset{\tilde{\mathrm{L}}}{\sqcup} C\beta$ holds under loss function $LF_G(\hat{F}z, C\beta)$ is unrelated to G only if $G > 0$ (Dong and Wu, 1988).

Which means that $\hat{F}z \overset{\tilde{\mathrm{L}}}{\sqcup} C\beta$ holds under loss $LF_G(\hat{F}z, C\beta)$ is equivalent to the $\hat{F}z \overset{\tilde{\mathrm{L}}}{\sqcup} C\beta$ holds under loss function $LF_I(\hat{F}z, C\beta)$.

Theorem 2.15. In linear model

$$\begin{cases} z = X'T^{-}X\beta + \varepsilon; \\ E(\varepsilon) = 0, E(\varepsilon\varepsilon') = \sigma^2 X'T^{-}VT^{-}X, \end{cases}$$

the SNC for $\hat{F}z \overset{\mathrm{L}^{h}}{\sqcup} C\beta$ holds with respect to the QLF

$$LF(Ly, C\beta) = (Ly - C\beta)'(Ly - C\beta),$$

are as follows:

(*a*) $\hat{F}V_0 = \hat{F}X_0(X_0'W^-X_0)^-X_0'W^-V_0$;

(*b*) $\hat{F}X_0\left[(X_0'W^-X_0)^- - I_p\right]X_0'\hat{F}' \le \hat{F}X_0\left[(X_0'W^-X_0)^- - I_p\right]C'$;

(*c*) $rk\left\{(\hat{F}X_0 - C)\left[(X_0'W^-X_0)^- - I_p\right]X_0'\right\} = rk(\hat{F}X_0 - C)$,

where $V \ge 0$ (not necessarily $V > 0$), $X_0 = X'T^-X$, $W = V_0 + X_0X_0'$ and $V_0 = X'T^-VT^-X$ (Wu (1986)).

Theorem 2.16. In GGM model, if $C\beta$ is estimable, then SNC for $F^*y \sqsubset C\beta$ under the QLF

$$LF(F^*y, C\beta) = (F^*y - C\beta)'(F^*y - C\beta),$$

is as follows:

(*a*) $F^*V = F^*P_{XT}V$

(*b*) $F^*XWC' - F^*XWX'F^{*'} \ge 0$, F^*XWC' is symmetric;

(*c*) $rk\left\{(F^*X - C)WX'\right\} = rk(F^*X - C)$,

where $V \ge 0$ (not necessarily $V > 0$), $T_* = V + XX'$, $P_{XT} = XS_{T_*}^-X'T_*^+$, $W = S_{T_*}^- - I$, $S_{T_*} = X'T_*^+X$. (Wu (1986)).

Theorem 2.17 (Gauss-Markov Theorem) Under the linear regression model with assumptions (i) to (iv) , the least squares estimator

$$\hat{\beta} = \left(X'X\right)^{-1} X'y$$

is UNW than any estimator unbiased estimator β_* with respect to $MSE(\beta, \beta_*)$, as well as with respect to

$$R(\beta, \beta_*) = E\left[(\beta - \beta_*)'W(\beta - \beta_*)\right]$$

where W is an arbitrary $p \times p$ symmetric n.n.d weight matrix.

Theorem 2.18. In model $y = X\beta + \varepsilon,\ E(\varepsilon) = 0,\ E(\varepsilon\varepsilon') = \sigma^2 V_n$, $BY \sim \beta\left[\Theta_T, \mathrm{D}_L\right]$ iff $B = PX'V^{-1}$ with $\left(X'V^{-1}X\right)^{1/2} P \left(X'V^{-1}X\right)^{1/2} \in \Gamma$, and $tr\left(X'V^{-1}X\right)^{-1} T\left[(I - PX'V^{-1}X)^{-1} - I\right] \le 1$, where $\Gamma = \left\{L \in \Re^{\geq}, (I - L) \in \Re^{>}\right\}$, $\Theta_T = \left\{\beta \in \Re^p, \beta' T \beta \le \sigma^2\right\}$, (Kurt Hoffmann , 1977).

Theorem 2.19. For ellipsodial restriction, we have

$$\tilde{L}Z + a \overset{mse}{\sim} C\beta[\Theta] \ \Leftrightarrow\ C^{-1}\tilde{L}Z + C^{-1}a \overset{mse}{\sim} \beta[\Theta] .$$

Theorem 2.20. Consider the linear model

$$\begin{cases} Z = \tilde{X}_0\xi + \varepsilon; \\ E(\varepsilon) = 0,\ \mathrm{cov}(\varepsilon) = \sigma^2\tilde{\Lambda}, \end{cases}$$

where $\xi \in \Re^p$ and $\varepsilon \in \Re^n$ are random vectors, respectively, with $E(\xi) = K\gamma$, $\mathrm{cov}(\xi) = \sigma^2\tilde{\Phi}$ and $E(\xi\varepsilon') = 0$. X_0, K, $\tilde{\Phi} \geq 0$ and $\tilde{\Lambda} \geq 0$ are known, whereas $\gamma \in \Re^p$ and $\sigma^2 > 0$ are unknown parameters. Let $\mathrm{L}\hat{I} = \{AZ : A \in \Re^{k\times n}\}$. If $Q\xi$ is linearly estimable, then $AZ \overset{\mathrm{L}^*}{\sqcup} Q\xi$ holds under the QLF $LF(AZ, Q\xi) = (AZ, Q\xi)'(AZ, Q\xi)$ if and only if

$$(i)\ A_*\tilde{\Sigma} = A_*\tilde{X}_*\tilde{R}(\tilde{R}'\tilde{X}_*'\tilde{W}^+\tilde{X}_*\tilde{R})^-\tilde{R}\tilde{X}_*'\tilde{W}^+\tilde{\Sigma} + Q\tilde{\Phi}\tilde{X}_*'\tilde{W}^+\left[I - \tilde{X}_*\tilde{R}(\tilde{R}'\tilde{X}_*'\tilde{W}^+\tilde{X}_*\tilde{R})^-\tilde{R}\tilde{X}_*'\tilde{W}^+\right]\tilde{\Sigma}$$

$$(ii)\ A_*\tilde{X}_*\tilde{R}\tilde{M}\tilde{R}'\tilde{X}_*'A_*' \leq (A_*\tilde{X}_* - Q)\tilde{R}(\tilde{R}'\tilde{X}_*'\tilde{W}^+\tilde{X}_*\tilde{R})^-\tilde{R}\tilde{X}_*'\tilde{W}^+\tilde{X}_*\tilde{\Phi}Q' + Q\tilde{R}\tilde{M}K'\tilde{X}_*'A_*',$$

$$(iii)\ C\left[(A_*\tilde{X}_* - Q)\tilde{R}\right] = C(\tilde{N}),$$

holds simultaneously, where

$$\tilde{M} = (\tilde{R}'\tilde{X}_*'\tilde{W}^+\tilde{X}_*\tilde{R})^-\tilde{R}'\tilde{X}_*'\tilde{W}^+\Sigma\tilde{W}^+\tilde{X}_*\tilde{R}(\tilde{R}'\tilde{X}_*'\tilde{W}^+\tilde{X}_*\tilde{R})^-,$$

$$\tilde{N} = (A_*\tilde{X}_* - Q)\tilde{R}\left[(\tilde{R}'\tilde{X}_*'W^+\tilde{X}_*\tilde{R})^-\tilde{R}\tilde{X}_*'\tilde{W}^+\tilde{X}_*\tilde{\Phi}Q' - M\tilde{R}'\tilde{X}_*'A_*'\right] + (A_*\tilde{X}_* - Q)\tilde{R}M\tilde{R}'(A_*\tilde{X}_* - Q)'$$

$$\tilde{N} = (A\tilde{X}_0 - Q)K\left[(K'\tilde{X}_0'W^+\tilde{X}_0K)^-K\tilde{X}_0'\tilde{W}^+\tilde{X}_0\tilde{\Phi}Q' - MK'X_0'A'\right] + (A\tilde{X}_0 - Q)KMK'(A\tilde{X}_0 - Q)'$$

$\tilde{\Sigma} = \tilde{X}_* \tilde{\Phi} \tilde{X}'_* + \tilde{\Lambda}$ and $\tilde{W} = \tilde{\Sigma} + \tilde{X}_* \tilde{R} \tilde{R}' \tilde{X}'_*$ (Wu, 1988).

3. SOME BIASED ESTIMATORS

As mentioned in the last section, OLS is the best linear unbiased estimator, but when multicollinearity is present it tends to be unstable due to its larger variances. To overcome the above-mentioned problem, a need arises for alternative estimators with variance smaller than OLS.

In this section, some biased estimators proposed as an alternative to the OLS estimator will be discussed. Some biased estimators commonly used in statistics and econometrics are described and their statistical properties are examined.

3.1. Generalized Least Square Estimator

Aitken (1936) first described generalized least squares (GLS) as an alternative technique for estimating unknown parameters in a LRM when there is a correlation between residuals in the model (2.1).

Consider the following GGM model,

$$y = X\beta + \varepsilon, \;\; E(\varepsilon) = 0, \;\; E(\varepsilon\varepsilon') = \sigma^2 V, \tag{3.1.}$$

where y is an n- dimensional vector of observations on the dependent variables such that $E(y) = X\beta$, $\mathrm{Var}(y) = \sigma^2 V$, V is a known n.n.d matrix ($V \geq 0$), X, β and ε are same as one in model (2.1).

According to Rao's unified theory of least squares, the GLS estimator of β parameter in model (3.1.) is given by

$$\hat{\beta}_{GLSE} = S_{T_*}^{-} X' T_*^{+} y, \tag{3.2.}$$

where $S_{T_*} = X'T_*^{+}X$, $T_* = V + XX'$.

3.2. Weighted Least Square Estimator

A special case of GLS called weighted least squares (WLS) is a generalization of OLS with weight matrix is a diagonal matrix, which situation arises when the variance of observations are unequal but where no correlations exist among the observed variances. Aitken showed that when the observational errors are uncorrelated and the weight matrix is a diagonal, these may be written as

$$X'V^{-1}X\hat{\beta} = X'V^{-1}y$$

If the errors are correlated, the resulting estimator is the BLUE if the weight matrix is equal to the inverse of the variance-covariance matrix of the observations.

3.3. Ridge estimator

Ridge is one of the alternative estimators to OLS to overcome the problem such as multicollinearity problem which causes OLS coefficients to be unstable due to larger variance. The ridge regression estimator

$$\hat{\beta}(k) = (X'X + kI_p)^{-1}X'y, \qquad k > 0 \tag{3.3.}$$

proposed by Hoerl and Kennard (1970) and biased for $k > 0$. The ridge regression estimator is derived from the idea of making the matrix more stable by adding positive small k values to the diagonal elements of the $X'X$ matrix. Ridge estimator is related to the parameter k and selection of the k is affected the performans of ridge.

The ridge estimator depends on the parameter k , and selection of k affects the performance of the estimator. When $k = 0$, the ridge estimator reduce to the OLS estimator. Also, the ridge estimator can be written as a linear transformation of the OLS estimator. The k value should be large enough that the ridge estimator's variance covariance matrix is smaller than the OLS estimator, and small enough that the bias is acceptable.

$$\hat{\beta}(k) = (X'X + kI_p)^{-1} X'X\hat{\beta}.$$

3.4. GR Estimator

Hoerl and Kennard (1970) defined the GR estimator as for any such non-negative symmetric matrix K

$$\hat{\beta}(K) = (X'X + K)^{-1} X'y,$$

And also a non-homogeneous GR estimator is given as the form of

$$\hat{\beta}_{K,\beta_0} = \beta_0 + (X'X + K)X'(y - X\beta_0).$$

When the $p \times 1$ vector β_0 and the $p \times p$ symmetric nonnegative definite matrix K are chosen non-stochastically, then $\hat{\beta}_{K,\beta_0}$ is also called a linear estimator of ridge type (see Groß , 2003, p.227).

It is quite obvious that the estimators $\hat{\beta}$ and $\hat{\beta}(k)$ are special case of $\hat{\beta}_{K,\beta_0}$ with $K = 0$ and $K = kI$ respectively. In addition, Table 3.1. (see Appendix 4) displays

linear ridge type estimator. The Marquard, Farebrother, Almost unbiased ridge, Iteration, Shrinkage are also of the form of a GR estimator for special choices of K and β_0

Many studies have been done on the ridge regression estimator. Some of the studies using various loss functions for comparison are as follows: Theobald (1974) compared the OLS estimator and the ridge regression estimator according to the generalized squares of error (gmse) and showed in which case the ridge regression estimator is superior to the OLS. In addition to Theobald's result, Gunst and Mason (1976) compared OLS with principal component regression estimators and ridge regression with principal component estimators. Lin and Kmenta (1982) used the p'-norm loss measure as an alternative loss criterion to compare the OLS estimator and the ridge regression estimator. They obtained the loss functions by taking $p = 1, 2 \; and \; \infty$ and showed that in all three cases the ridge regression estimator was no worse than the OLS estimator.

Chawla and Jain (1988) compared the ridge regression estimator with the generalized ridge regression estimator according to the risk of mse. Peddada ve ark. (1989) compared the OLS and the generalized ridge regression estimator under the Mahalanobis loss function and obtained that generalized ridge estimator is inadmissible with respect to the OLS. Ohtani (1995) analyzed the risk performance of a feasible generalized ridge regression estimator under asymmetric linex loss function; obtained the necessary condition for the generalized ridge regression estimator to be superior to the OLS estimator and the results supported by numerically example. Wan (1999) using the linex loss function, showed that the almost unbiased applicable generalized ridge regression estimator proposed by Ohtani (1986) is not unequivocally superior to the OLS estimator. Akdeniz and Namba (2003) defined a new applicable generalized ridge regression estimator and examined its risk performance under the asymmetric linex loss

function. Wan (2002) investigated the risk performances of nearly unbiased applicable generalized ridge regression and applicable generalized ridge regression estimators using the balanced loss function. Groß (2003) also included the admissibility of the GR estimator under the mse in his book.

3.5. RLS Estimation

In some cases, there are unknown parameter vector β constraints in the classical LRM given by (2.1.) which can be expressed in the form of a linear equation

$$R\beta = r \tag{3.1.}$$

where $R \in \Re^{m\times n}$ and $rank(R) = m$, $r \in \Re^m$ is known vector.

A classic and obvious method for obtaining a suitable restricted estimator for β under additional constraints is to find the vector that minimizes mse, same as the method of obtain the OLS. In other words, the restricted estimator assumed to satisfied the $R\beta_R = r$ and in addition

$$(y - X\hat{\beta}_R)'(y - X\hat{\beta}_R) = \min_{\beta_*: R\beta_* = r} (y - X\beta_*)'(y - X\beta_*) .$$

To obtain the restricted estimator $\hat{\beta}_R$, we using method of Lagrange multipliers, for this we consider the function

$$S(\beta, \lambda) = (y - X\beta_*)'(y - X\beta_*) - 2\lambda'(R\beta_* - r),$$

where $\lambda \in \Re^p$ is the vector of Lagrange multipliers. If we take a derivative with respect to β_* and λ and set it equal to zero, we get normal equation as follows

$$\frac{\partial S(\beta_*, \lambda)}{\partial \beta_*} = -2Xy + 2X'X\beta - 2R'\lambda = 0$$

$$\frac{\partial S(\beta_*, \lambda)}{\partial \lambda} = R\beta_* - r = 0 \qquad (3.2.)$$

$S(\beta_*, \lambda)$ is minimized if $\hat{\beta}_R$ is the solution to

$$\hat{\beta}_R = \hat{\beta} + (X'X)^{-1} R'\lambda \qquad (3.3.)$$

where $\hat{\beta} = (X'X)^{-1} Xy$. From this identity, we can also get the

$$R\hat{\beta}_R = r = R\hat{\beta} + R(X'X)^{-1} R'\hat{\lambda}_R . \qquad (3.4.)$$

Since $R(X'X)^{-1} R'$ is p.d, then

$$\hat{\lambda}_R = \left(R(X'X)^{-1} R'\right)^{-1} \left(r - R\hat{\beta}\right) . \qquad (3.5.)$$

Substituting (3.3.) and (3.5.), it easily get

$$\hat{\beta}_R = \hat{\beta} - (X'X)^{-1} R'\left(R(X'X)^{-1} R'\right)\left(R\hat{\beta} - r\right) \qquad (3.6.)$$

It seems clear that this estimator is biased under constraint, otherwise it is an unbiased estimator.

$$E\left(\hat{\beta}_R\right) = \beta - \left(X'X\right)^{-1} R'\left(R\left(X'X\right)^{-1} R'\right)\left(R\beta - r\right) \tag{3.7.}$$

Although it is biased, the reason for the constraint is to reduce the variance.

$$Var\left(\hat{\beta}_R\right) = \sigma^2\left(X'X\right)^{-1} - \sigma^2\left(X'X\right)^{-1} R'\left(R\left(X'X\right)^{-1} R'\right)^{-1} R\left(X'X\right)^{-1} \tag{3.8.}$$

From (2.6.) and (3.8.)

$$Var\left(\hat{\beta}\right) - Var\left(\hat{\beta}_R\right) = \sigma^2\left(X'X\right)^{-1} R'\left(R\left(X'X\right)^{-1} R'\right)^{-1} R\left(X'X\right)^{-1} \ . \tag{3.9.}$$

Since the right-hand side of (3.9.) is n.n.d, we can say that the use of linear constraints caused a decrease in variance.

3.6. Farebrother Estimator

If uncertainty exists about the prior information specifications, one alternatives is to make use of stochastic linear restrictions of the following form:

$$r = R\beta + \phi \tag{3.10.}$$

where r is a known $m \times 1$ random vector , R is a known $m \times p$ prior information of full row rank $m < p$, and ϕ is a random vector, independent of ε , with $E(\phi) = 0$ and

$\text{Cov}(\phi) = \sigma^2 W$. Theil and Goldberger (1961) and Theil (1963) introduced the mixed estimator by unifying (3.1.) and (3.10.) in a common model as

$$\hat{\beta}_M = (X'V^{-1}X + R'W^{-1}R)^{-1}(X'V^{-1}y + R'W^{-1}r)\,. \tag{3.11.}$$

The stochastically restricted estimator has a smaller covariance matrix than the GLS estimator and has desirable properties in the context of autocorrelation. The comparisons of $\hat{\beta}_M$ and GLS estimator have been given by Rao and Toutenburg (1995) with respect to the different mse criterions under the biased restrictions (see Rao and Toutenburg, 1995, p. 139-147).

An application of mixed estimation to combat autocorrelation and multicollinearity is in Belsley et al.(2004). To reduce the effects of autocorrelation and multicollinearity problems, Trenkler (1984) defined the ridge estimator in model (3.1.). This estimator was given by

$$\hat{\beta}_V(k) = (X'V^{-1}X + kI_p)^{-1}X'V^{-1}y\,. \tag{3.12.}$$

The comparisons of $\hat{\beta}_M$ and $\hat{\beta}_V(k)$ have been given by Güler and Kaçıranlar (2009) with respect to the variance-covariance matrix and the MSE under the unbiased restrictions.

Farebrother (1984) introduced the following generalized ridge estimator GRE: Combining the information in (3.1.) with $V = I$ and the prior information contained in (3.10.) under the special conditions $E(\phi) = 0$, $\text{Cov}(\phi) = \sigma^2 / k\, I_m$, and following GRE obtained as

$$\hat{\beta}_F(k) = (X'X + k\,R'R)^{-1}(X'y + kR'r)\ ,\ k \geq 0$$
$$= \hat{\beta} - k(X'X + k\,R'R)^{-1}R'(R\hat{\beta} - r)$$
$$= \hat{\beta} - kT_k^{-1}R'(R\hat{\beta} - r)\ , \qquad (3.13.)$$

where $T_k = \mathrm{S} + k\,R'R$. In fact, this estimator was firstly described in Terasvirta (1979a) and was applied to distributed models in Terasvirta (1979b) (see Toutenburg (1982), p.60-61).

When r and β are fixed, Farebrother pointed out that this derivation of $\hat{\beta}_F(k)$, and the equivalent bayesian derivation are invalid. These ideas also led him to introduce a new GRE for β by replacing r with $R\hat{\beta}_R$. So, we have the following estimator:

$$\hat{\beta}_F(k,r) = (X'X + k\,R'R)^{-1}(X'y + kR'R\hat{\beta}_R)\ ,\ k \geq 0\ . \qquad (3.14.)$$

We call $\hat{\beta}_F(k)$ as the Farebrother's estimator. $\hat{\beta}_F(k)$ has many interesting features as we will mention below:

- $\hat{\beta}_F(k)$ is a special case of the linear mixed estimator.
- $\hat{\beta}_F(k,r)$ is a weighted average of the OLS estimator and the RLS estimator.
- $\hat{\beta}_F(k,r) = H\hat{\beta} + (I - H)\hat{\beta}_R$, where $H = (S + k\,R'R)^{-1}S = T_k^{-1}S$.
- $\hat{\beta}_F(0) = \hat{\beta},\ \hat{\beta}_F(\infty) = \hat{\beta}_R$ (see also Groß (2003), p.135)
- $\hat{\beta}_F(k) = \hat{\beta} - kS^{-1}R'(I_m + k\,RS^{-1}R')^{-1}(R\hat{\beta} - r)$

  where $(X'X + k\,R'R)S^{-1}R' = R'(I_m + kRS^{-1}R')$.

- $\hat{\beta}_F(k)$ is the minimax linear estimator of β subject to $((R\beta - r)'(R\beta - r) \leq \sigma^2/k$ (see Kuks and Olman (1972)).
- $\hat{\beta}_F(k)$ has lower MSE (MSEM) than OLS estimator under the certain conditions as r fixed (see details in Farebrother (1984), Groß (2003), p.137-138).
- RLS has lower MSEM than $\hat{\beta}_F(k)$ estimator under the certain conditions as r fixed (see details in Farebrother (1984)).
- When, $r = 0$. Farebrother estimator $\hat{\beta}_F(k)$ can be written following as

$$\hat{\beta}^{(r)}(k) = (X'X + kI_p - kU_1U_1')^{-1}X'y,$$

This estimator can be seen as a weighted average of the ordinary least squares and principal components estimator. (see details in Groß (2003), p.138-141).

- If we choose matrix $R = I_p$ but $r \neq 0$, we can obtained the estimator from Swindel (1976).

$$\hat{\beta}_{k,r} = (X'X + kI_p)^{-1}(X'y + kr),$$

it may also called direction modified ridge estimator.

3.7. Bayes Estimator

This topic is has deep roots in decision theory. A Bayesian estimator is an estimator of an unknown parameter β that minimizes the expected loss for all observations. In other words, bayesian methods are a convenient method proposed to

generate appropriate estimators for β, assuming some random vector b is performed for which the unknown parameter vector β is not fixed but a-priority information about its distribution is available.

Under the above assumption, it is not necessary to have knowledge of the distribution of b as a whole to use bayesian methods, but it is sufficient to have basic knowledge of the unknown parameter vector β in classical linear regression, such as the expectation vector and the covariance matrix model given in (2.1.)

$$E(b) = \mu \text{ and } Cov(b) = \sigma^2 \Phi$$

where μ is a known $p \times 1$ vector and $\Phi \in \Re_s^{\geq}$ is a $p \times p$ matrix. How the knowledge of μ and Φ can be employed to estimate β is a fairly natural question. To answer this, let us remind the usual unweighted mse of a linear estimator $\beta_* = Ay + a$ for β, given as

$$\begin{aligned} mse(\beta, \beta_*) &= E\left[(Ay + a - \beta)'(Ay + a - \beta)\right] \\ &= \sigma^2 tr(AA') + tr\left[\left((AX - I_p)\beta + a\right)\left((AX - I_p)\beta + a\right)'\right] \\ &= R_{\beta,\sigma^2}(A, a) \end{aligned}$$

This is seen as a risk function of A and a that depends on unknown parameters β and σ^2. Here we use bayes risk

$$R^B(\beta,\ Ay + a) = E\left[R_{b,\sigma^2}(A, a)\right]$$

instead of the usual risk $mse(\beta,\beta_*)=R_{\beta,\sigma^2}(A,a)$ employing the prior information to remove this function's dependency on β. It can be easily known that Bayesian risk can be interpreted as the average of the usual risk over all possible realizations β of the random vector b and evaluated $E\left[R_{b,\sigma^2}(A,a)\right]$ with knowledge of $E(b)=\mu$ and $Cov(b)=\sigma^2\Phi$. Then a linear estimator A_*y+a_* is called for bayes estimator for β, if its risk $R^B(\beta, A_*y+a_*)$ doesn't exceed the risk $R^B(\beta, Ay+a)$ of any linear estimator $Ay+a$ of β.

Now, we try to find the linear bayes estimator. Under the following assumption

$$E(\mu/b)=0$$
$$Var(\mu/b)=\sigma_u^2 I \qquad (3.15.)$$

To find the linear bayes estimator, we calculate the linear equation $al'+y$ which satisfied the following

$$E(p'b-a-l'y)=0 \qquad (3.16.)$$

and minimized the variance (Rao, 1971)

$$Var(p'b-a-l'y) \qquad (3.17.)$$

By solving Equation (3.16) and (3.17.) with piror knowledge, we get

$$E(p'b - a - l'y) = 0$$

$$\Rightarrow\ p'E(b) - a - l'EX(b) = 0$$

$$\Rightarrow a = (p' - l'X)\mu \qquad (3.18.)$$

and

$$Var(p'b - a - l'y) = Var(p'b) + Var(l'y) - 2\mathrm{cov}(p'b, l'y) \qquad (3.19.)$$

Since

$$\begin{aligned} Var(l'y) &= l'Var(y)l \\ &= l'(X\Phi X' + \sigma_\varepsilon^2 I)l \end{aligned} \qquad (3.20.)$$

and

$$\begin{aligned} Cov(p'b, l'y) &= p'Cov(b, y)l \\ &= p'\{E[(b - E(b))(y - E(y))]\}l \\ &= p'\{E[(b - u)(y - Xu)]\}l \\ &= p'\{E[bb'X' + b\varepsilon' - bu'X' - ub'X' - u\varepsilon' + uu'X']\}l \\ &= p'\{E[(b - u)(b - u)']X'\}l \\ &= p'\Phi X'l \end{aligned} \qquad (3.21.)$$

then (3.21.) eaqual to

$$Var(p'b - a - l'y) = p'\Phi p + l'(X\Phi X' + \sigma_\varepsilon^2 I)l - 2p'\Phi X'l \tag{3.22.}$$

If we take a derivative with respect to l and set it equal to zero,

$$l' = p'\Phi X'(X\Phi X' + \sigma_\varepsilon^2 I)^{-1} \tag{3.23.}$$

Then using (3.18.) and (3.23.) we obtain the linear bayes estimator

$$p'\hat{\beta}^B = p'u + p'\Phi X'(X\Phi X' + \sigma_\varepsilon^2 I)^{-1}(y - Xu). \tag{3.24.}$$

If the matrix X has full column rank, linear bayes estimator for β as follows

$$\hat{\beta}^B = u + \Phi X'(X\Phi X' + \sigma_\varepsilon^2 I)^{-1}(y - Xu). \tag{3.25.}$$

Alternative forms of linear Bayesian estimator as follows (Gruber, 1998, 2010).

$$\hat{\beta}^B = u + (X'X + \sigma_\varepsilon^2 \Phi^{-1})^{-1} X'(y - Xu) \tag{3.26.}$$

$$= (X'X + \sigma_\varepsilon^2 \Phi^{-1})^{-1}(X'y - \sigma_\varepsilon^2 \Phi^{-1} u) \tag{3.27.}$$

$$= u + \Phi\left(\sigma_\varepsilon^2 (X'X)^{-1} + \Phi\right)^{-1}(\hat{\beta} - u) \tag{3.28.}$$

$$= (X'X)^{-1}\left(\Phi + \sigma_\varepsilon^2 (X'X)^{-1}\right)^{-1} b\sigma_\varepsilon^2 + \Phi\left(\Phi + \sigma_\varepsilon^2 (X'X)^{-1}\right)^{-1}\hat{\beta} \tag{3.29.}$$

$$= \hat{\beta} - (X'X)^{-1}\left(\Phi + \sigma_\varepsilon^2 (X'X)^{-1}\right)^{-1}(\hat{\beta} - u)\sigma_\varepsilon^2 \tag{3.30.}$$

From (3.27) bayes estimator can be written as

$$\begin{aligned}\hat{\beta}^{B} &= (X'X + \sigma_{\varepsilon}^{2}\Phi^{-1})^{-1}(X'y - \sigma_{\varepsilon}^{2}\Phi^{-1}u) \\ &= (X'X + \sigma_{\varepsilon}^{2}\Phi^{-1})^{-1}X'X\hat{\beta} + \sigma_{\varepsilon}^{2}(X'X + \sigma_{\varepsilon}^{2}\Phi^{-1})^{-1}\Phi^{-1}u \\ &= (\frac{X'X}{n} + \frac{\sigma_{\varepsilon}^{2}\Phi^{-1}}{n})^{-1}\frac{X'X}{n}\hat{\beta} + \sigma_{\varepsilon}^{2}(\frac{X'X}{n} + \frac{\sigma_{\varepsilon}^{2}\Phi^{-1}}{n})^{-1}\frac{\Phi^{-1}u}{n}.\end{aligned} \tag{3.31.}$$

In this case

$$\begin{aligned}plim(\hat{\beta}^{B}) = \lim_{n\to\infty}(\frac{X'X}{n} + \frac{\sigma_{\varepsilon}^{2}\Phi^{-1}}{n})^{-1}\frac{X'X}{n}\,plim(\hat{\beta}) \\ + \lim_{n\to\infty}\sigma_{\varepsilon}^{2}(\frac{X'X}{n} + \frac{\sigma_{\varepsilon}^{2}\Phi^{-1}}{n})^{-1}\frac{\Phi^{-1}u}{n}\end{aligned}$$

Since

$$\lim_{n\to\infty}(\frac{X'X}{n} + \frac{\sigma_{\varepsilon}^{2}\Phi^{-1}}{n})^{-1}\frac{X'X}{n} = I_{n} \text{ and } \lim_{n\to\infty}(\frac{X'X}{n} + \frac{\sigma_{\varepsilon}^{2}\Phi^{-1}}{n})^{-1}\frac{1}{n} = 0$$

$$plim(\hat{\beta}^{B}) = plim(\hat{\beta}) = \beta \tag{3.32.}$$

$\hat{\beta}^{B}$ is consistent and asymptotically unbiased. It can also be easily seen that $\hat{\beta}^{B}$ is equal to $\hat{\beta}$ for samples larger than the spelling in equation (3.31).

Then the expected value of the linear bayesian estimator with $\Phi^{*} = (X'X + \sigma_{\varepsilon}^{2}\Phi^{-1})^{-1}$ and $\beta^{*} = \hat{\beta} - u$ as follows

$$\begin{aligned} E(\hat{\beta}^B) &= u + \Phi^* X'X\beta - \Phi^* X'Xu \\ &= u + \Phi^* X'X\beta - (I - \sigma_\varepsilon^2 \Phi^* \Phi^{-1})u \\ &= \Phi^* X'X\beta + \sigma_\varepsilon^2 \Phi^* \Phi^{-1} u \end{aligned} \tag{3.33.}$$

and

$$Bias(\hat{\beta}^B) = E(\hat{\beta}^B) - \beta = -\sigma_\varepsilon^2 \Phi^* \Phi^{-1} \beta^* \tag{3.34.}$$

Also we get variance - covariance matrix from (3.27.) as follows

$$\begin{aligned} Var(\hat{\beta}^B) &= Var\left(\Phi^* X'y + \sigma_\varepsilon^2 \Phi^* \Phi^{-1} u\right) \\ &= Var\left(\Phi^* X'y\right) \\ &= \sigma_\varepsilon^2 \Phi^* X'X\Phi^* \end{aligned} \tag{3.35.}$$

Using (3.34.) and (3.35.) we have MSE of the linear bayes estimator

$$\begin{aligned} MSE(\hat{\beta}^B) &= Var(\hat{\beta}^B) + Bias(\hat{\beta}^B)Bias(\hat{\beta}^B)' \\ &= \sigma_\varepsilon^2 \left(\Phi^* X'X\Phi^* + \sigma_\varepsilon^2 \Phi^* \Phi^{-1} \beta^* \beta^{*\prime} \Phi^{-1} \Phi^*\right) \end{aligned} \tag{3.36.}$$

and compare the linear bayes estimator and OLS with respect to the MSE

$$\begin{aligned}\Delta &= MSE(\hat{\beta}) - MSE(\hat{\beta}^B) \\ &= \sigma_\varepsilon^2\left((X'X)^{-1} - \Phi^* X'X\Phi^* - \sigma_\varepsilon^2\Phi^*\Phi^{-1}\beta^*\beta^{*\prime}\Phi^{-1}\Phi^*\right) \\ &= \sigma_\varepsilon^2\Phi^*\left(\Phi^{*-1}(X'X)^{-1}\Phi^{*-1} - X'X - \sigma_\varepsilon^2\Phi^{-1}\beta^*\beta^{*\prime}\Phi^{-1}\right)\Phi^* \\ &= \sigma_\varepsilon^4\Phi^*\Phi^{-1}\left(2\Phi + \sigma_\varepsilon^2(X'X)^{-1} - \beta^*\beta^{*\prime}\right)\Phi^{-1}\Phi^*\end{aligned} \tag{3.37.}$$

Since $\Phi^*\Phi^{-1}$ full colomun matrix; from Theorem 2.1, we have

$$\Delta \geq 0 \Leftrightarrow 2\Phi + \sigma_\varepsilon^2(X'X)^{-1} - \beta^*\beta^{*\prime} \geq 0$$

and using Theorem 2.2, if the following condition is satisfied,

$$\beta^{*\prime}\left(\frac{2\Phi}{\sigma_\varepsilon^2} + (X'X)^{-1}\right)\beta^* \leq \sigma_\varepsilon^2 \tag{3.38.}$$

We can say that linear bayes estimator is superior to the OLS estimator with respect to the MSE criterion.

4. CHARACTERIZATION OF ALEs UNDER EBLF

In this chapter, the admissibility of the linear estimator, which has deep roots in decision theory, will be discussed, in different types of LRMs. Moreover, under the EBLF, we will investigate the admissibility of some known estimator and their properties.

4.1. AOLE under EBLF

Consider the linear model

$$y = X\beta + \varepsilon,\ \ E(\varepsilon) = 0,\ \ E(\varepsilon\varepsilon') = \sigma^2 I_n \tag{4.1.}$$

where y is an $n \times 1$ vector of observations on the dependent variable, $X \in \Re^{n\times p}$ is an model matrix of observations on the p regressors and $rk(X) = p$ has full of column rank, $\beta \in \Re^p$ and $\sigma^2 > 0$ are unknown parameters, $\varepsilon \in \Re^n$ is an vector of disturbances.

4.1.1. Explicit Characterization

It is known that EBLF is more sensitive than the quadratic loss and reduces to the quadratic loss when $t_1 = 0$ and $t_2 = 1$. Thus, there exists some relationship between the EBLF and the quadratic loss. It is presented in Theorem 4.1 which plays a crucial role in proving the main results.

Theorem 4.1. Let $\mathrm{L}_0^h = \left\{ L_0 X'y : L_0 \in \Re^{p\times p} \right\}$. Then L_0^h is a complete class of L^h.

Proof: If $\beta_* = Ly \in \mathrm{L}^h$, using some properties of trace and Theorem 2.3, from Equation (2.32) and (2.34), we get

$$
\begin{aligned}
R\left(Ly;\beta,\sigma^2\right) &= E\left[t_1\left(y-XLy\right)'\left(y-XLy\right)+t_2\left(Ly-\beta\right)'X'X\left(Ly-\beta\right)\right] \\
&+\left[(1-t_1-t_2)\mathrm{E}\left(XLy-y\right)'X\left(Ly-\beta\right)\right] \\
&=\sigma^2 tr\left[t_1\left(I_n-XL\right)'\left(I_n-XL\right)+t_2\left(L'X'XL\right)\right] \\
&+\sigma^2 tr\left[\left(1-t_1-t_2\right)\left(\left(XL-I_n\right)'XL\right)\right] \qquad (4.2.)\\
&+\beta' X'\left(I_n-XL\right)'\left(I_n-XL\right)X\beta
\end{aligned}
$$

On the other hand, if $Ly\in \mathrm{L}^h$, then, $LP_X y = L_0 X'y \in \mathrm{L}_0^h$, where $P_X = X\left(X'X\right)^{-1}X'$ and $L_0 = LX(X'X)^{-1}$. So, from Theorem 2.4 we have

$$R\left(Ly;\beta,\sigma^2\right)-R\left(LP_X y;\beta,\sigma^2\right)=\sigma^2 tr\left[L'X'XL\left(I_n-P_X\right)\right]\geq 0,$$

Moreover, the equality holds if and only if $L=LP_X$. The proof is given in detail in the Appendix 1.

Theorem 4.2. Given linear model

$$
\begin{cases} z = X'X\beta+u \\ E(u)=0, E(uu')=\sigma^2 X'X, \end{cases} \qquad (4.3.)
$$

where $\beta\in\Re^p$ and $\sigma^2>0$ is unknown and u is a p - dimensional error vector. Let $C=\left(1-w\right)I_p$, $\tilde{\mathrm{L}}=\left\{\tilde{L}z:\tilde{L}\in\Re^{p\times p}\right\}$, $\mathrm{L}_0^h=\left\{L_0X'y:L_0\in\Re^{p\times p}\right\}$, the loss function

$$LF_B(\beta_*,C\beta)=(\beta_*-C\beta)'B(\beta_*-C\beta),$$

$\tilde{L} = L_0 - w\left(X'X\right)^{-1}$, and $C\beta$ is linearly estimable in model (4.3.). Then, under model (4.3.) and loss $LF_B(\beta_*, C\beta)$, $\tilde{L}z \in C\beta$ holds if and only if $L_0X'y \in \mathrm{L}_0^h\beta$ holds under model (4.1.) and the EBLF, where $w = \left(1 + t_1 - t_2\right)/2$ and $0 \le w \le 1$.

Proof : Let $L_0X'y \in \mathrm{L}_0^h$. We get from Equation (4.2.)

$$R\left(L_0X'y;\, \beta,\, \sigma^2\right) = \sigma^2 tr(X\tilde{L}'B\tilde{L}X') + \sigma^2 tr\left(t_1 I_n - w^2 X\left(X'X\right)^{-1}X'\right)$$
$$+ \beta'\left[\tilde{L}X'X - C\right]' B\left[\tilde{L}X'X - C\right]\beta,$$

where $w = \left(1 + t_1 - t_2\right)/2$, $0 \le w \le 1$, $B = X'X$, $\tilde{L} = L_0 - w\left(X'X\right)^{-1}$ and $C = \left(1 - w\right)I_p$.

(The proof of above Equation is given in detail in the Appendix 2).

Note that

$$ELF_B(\tilde{L}z, C\beta) = \sigma^2\, \mathrm{tr}(X\tilde{L}'B\tilde{L}X') + \beta'\left[\left(\tilde{L}X'X - C\right)' B\left(\tilde{L}X'X - C\right)\right]\beta. \quad (4.4.)$$

If $L_0X'y \in \mathrm{L}_0^h\beta$, then for an arbitrary $M_0X'y \in \mathrm{L}_0^h$, we have

$$R\left(L_0X'y;\, \beta,\, \sigma^2\right) \le R\left(M_0X'y;\, \beta,\, \sigma^2\right), \quad (4.5.)$$

for all $\beta \in R^p$ and $\sigma^2 > 0$, where $\tilde{M} = M_0 - w\left(X'X\right)^{-1}$. We can see from (4.3.) and (4.4.),

$$R\left(L_0X'y;\, \beta,\, \sigma^2\right) - R\left(M_0X'y;\, \beta,\, \sigma^2\right) = ELF_B(\tilde{L}z, C\beta) - ELF_B(\tilde{M}z, C\beta).$$

Then, we get that (4.5.) holds if and only if for any $\tilde{M}z \in \tilde{\mathrm{L}}$

$$ELF_B(\tilde{L}z, C\beta) \le ELF_B(\tilde{M}z, C\beta), \tag{4.6.}$$

holds for all $\beta \in \Re^p$ and $\sigma^2 > 0$, where $\tilde{M} = M_0 - w\left(X'X\right)^{-1}$. Thus, from the definition of admissibility, (4.5.) holds if and only if $\tilde{L}z \overset{\tilde{\mathrm{L}}}{\sqsubset} C\beta$ in model (4.3.). So, proof of Theorem 4.2 is completed.

Theorem 4.3. Under model (4.1.) and the EBLF, $Ly \overset{\mathrm{L}^h}{\sqsubset} \beta$ holds if and only if L can be written as

$$L = \left[w\left(X'X\right)^{-1} + (1-w)\left(X'X\right)^{-1/2} D \left(X'X\right)^{-1/2} \right] X',$$

where $w = \left(1 + t_1 - t_2\right)/2$, $0 \le w \le 1$ and $D = \left(X'X\right)^{1/2} LX \left(X'X\right)^{-1/2}$ is symmetric and $\lambda(D) \subset [0,1]$.

Proof: In this proof, first of all, we demonstrate that $\tilde{L}$ which is given in Theorem 4.2. Then, using Theorem 4.1 and relationship between the $\tilde{L}, L_0$ and L, we will get the desired result.

Let $k = n = p$. If $X_0 = V = X'X$, model (2.35.) reduces to the model (4.3) and the set L^* turns in to $\tilde{\mathrm{L}}$. Then, using Theorem 2.13, Theorem 2.14 and Theorem 4.2, under the model (4.3) when $B = X'X$, we get the necessary and sufficient condition for the $\tilde{L}z \overset{\tilde{\mathrm{L}}}{\sqsubset} C\beta$ under

$$LF_{X'X}(\beta_*, C\beta) = (\beta_* - C\beta)' X'X (\beta_* - C\beta),$$

are as follows:

$$(a)\ \tilde{L} = \tilde{L}X'X\left(X'X\left(X'X\right)^{-1}X'X\right)^{-}X'X\left(X'X\right)^{-1} = \tilde{L}$$

$$(b)\ \tilde{L}X'X\tilde{L}' \le (1-w)\tilde{L}, \qquad (4.7.)$$

where $(1-w)\tilde{L}$ is symmetric and $T = V = X'X$.

If we take $D = (1-w)^{-1}\left(X'X\right)^{1/2}\tilde{L}\left(X'X\right)^{1/2}$, using (4.7.), we get

$$\tilde{L}X'X\tilde{L}' \le (1-w)\tilde{L} \ \Leftrightarrow\ DD' \le D,$$

where

$$\begin{aligned} DD' &= (1-w)^{-2}\left(X'X\right)^{1/2}\tilde{L}X'X\tilde{L}'\left(X'X\right)^{1/2} \\ &\le (1-w)^{-1}\left(X'X\right)^{1/2}\tilde{L}\left(X'X\right)^{1/2} = D, \end{aligned}$$

then $D' = D$ and $\lambda(D) \subset [0,1]$. So, from $D = (1-w)^{-1}\left(X'X\right)^{1/2}\tilde{L}\left(X'X\right)^{1/2}$ we can obtain the $\tilde{L} = (1-w)\left(X'X\right)^{-1/2}D\left(X'X\right)^{-1/2}$. Since (4.4.) and Theorem 4.2, we know

$$L_0 = w\left(X'X\right)^{-1} + \tilde{L} \text{ and } \tilde{L}z \overset{\tilde{L}}{\sqsubset} C\beta \ \Leftrightarrow\ L_0X'y \overset{L_0^h}{\sqsubset} \beta .$$

Then, under the model (4.1.) and the EBLF, $L_0X'y \overset{L_0^h}{\sqsubset} \beta$ holds if and only if

$$L_0 = w\left(X'X\right)^{-1} + (1-w)\left(X'X\right)^{-1/2}D\left(X'X\right)^{-1/2}. \qquad (4.8.)$$

As $Ly \in \mathrm{L}^h$ and $L_0 X'y \in \mathrm{L}_0^h$, then, $L_0 X'y = Ly \in \mathrm{L}^h$ and using Theorem 4.1, we can obtain the desired result that $Ly \overset{\mathrm{L}^h}{\sqcup} \beta$ under the EBLF holds if and only if L can be written as

$$L = \left[w\left(X'X\right)^{-1} + (1-w)\left(X'X\right)^{-1/2} D\left(X'X\right)^{-1/2} \right] X', \tag{4.9.}$$

so, the proof of Theorem 4.3 is completed.

According to the Theorem 4.3, we can obtain the folowing corollary.

Corollary 4.1. In linear model (4.1.), if A is the set of ALE of β in L^h under the $LF_1(\beta_*, \beta) = (\beta_* - \beta)'(\beta_* - \beta),$ and let A^* is the set of ALE of β under the EBLF. Then, A^* can be demonstrated as follows

$$\begin{aligned} A^* &= \left\{ w\hat{\beta} + (1-w)\tilde{\beta}_{mse} : \tilde{\beta}_{mse} \in A \right\} \\ &= \left\{ \hat{\beta} - (1-w)\left(X'X\right)^{-1/2} D\left(X'X\right)^{-1/2} X'y : \tilde{\beta}_{mse} \in A \right\}, \end{aligned}$$

Proof: If we consider the special case of Theorem 2.13 when $k = p$, $X_0 = X, V = I_n$ and $C = I_p$, then , we get the sufficient and necessary condition for $Lz \overset{\mathrm{L}^h}{\sqcup} \beta$ under the loss function $LF_1(\beta_*, \beta)$ are as follows:

a) $L = LX\left(X'X\right)^{-1} X' = LP_X,$ (4.10.)

b) $LP_x L' \le LX\left(X'X\right)^{-1}$, where $LX\left(X'X\right)^{-1}$ is symmetric and $T = X'X$. (4.11.)

If we take $D = \left(X'X\right)^{1/2} LX\left(X'X\right)^{-1/2}$ and using (4.11.), we get

$$LP_X L' \le LX\left(X'X\right)^{-1} \;\Leftrightarrow\; DD' \le D,$$

where $D' = D$ and $\lambda(D) \subset [0,1]$. From (4.10.), $L = LP_X$, we get

$$L = LX\left(X'X\right)^{-1} X' = \left(X'X\right)^{-1/2} D\left(X'X\right)^{-1/2} X'.$$

For model (4.1.), let

$$\begin{aligned} A &= \left\{ Ly : L = LP_X \text{ and } LP_X L' \le LX\left(X'X\right)^{-1} \right\} \\ &= \left\{ \left(X'X\right)^{-1/2} D\left(X'X\right)^{-1/2} X'y : D' = D \text{ and } 0 \le D \le I_p \right\}. \end{aligned} \tag{4.12.}$$

If $\tilde{\beta}_{mse} \in A$, then

$$\tilde{\beta}_{mse} = Ly = \left(X'X\right)^{-1/2} D\left(X'X\right)^{-1/2} X'y.$$

This is same as the results in Groß (2003) when $a = 0$ (see pp.221).

As A^* is the set of ALEs of β in L^h under the EBLF, from (4.10.), we have

$$\begin{aligned} A^* &= \left\{ Ly = \left[w\left(X'X\right)^{-1} + (1-w)\left(X'X\right)^{-1/2} D\left(X'X\right)^{-1/2} \right] X'y \right. \\ &\qquad \left. where\ D = D' \text{ and } 0 \le D \le I_p \right\} \\ &= \left\{ w\hat{\beta} + (1-w)\tilde{\beta}_{mse} : \tilde{\beta}_{mse} \in \mathrm{A} \right\}, \end{aligned} \tag{4.13.}$$

for any $\beta^*_{EBLF} \in A^*$ and $L \in \mathrm{L}^h$. Then we get

$$\beta^*_{EBLF} = Ly = w\hat{\beta} + (1-w)\tilde{\beta}_{mse}.$$

We have the following conclusion from the above statements:

The expression of A^* shows that the ALE under the EBLF is a convex combination (CC) of the ALEs under the unweighted squared error loss (SEL)

$$\left(y - X\beta_*\right)' \left(y - X\beta_*\right)$$

and loss (2.18.) when $W = X'X$.

Since $D = D'$ and $0 \le D \le I_p$, then $D_1 = I_p - D$ also symmetric and $0 \le D_1 \le I_p$.

So,

$$\begin{aligned} A &= \left\{ \left(X'X\right)^{-1/2} D \left(X'X\right)^{-1/2} X'y : D = D' \text{ and } 0 \le D \le I_p \right\} \\ &= \left\{ \hat{\beta} - \left(X'X\right)^{-1/2} D_1 \left(X'X\right)^{-1/2} X'y : D_1 = D_1' \text{and } 0 \le D_1 \le I_p \right\}. \end{aligned} \tag{4.14.}$$

Combining (4.13.) and (4.14.), A^* can be rewritten as follows

$$\begin{aligned} A^* &= w\hat{\beta} + (1-w)\tilde{\beta} = w\hat{\beta} + (1-w)\left(\hat{\beta} - \left(X'X\right)^{-1/2} D_1 \left(X'X\right)^{-1/2} X'y\right) \\ &= \hat{\beta} - (1-w)\tilde{\beta}_{mse}. \end{aligned} \tag{4.15.}$$

Theorem 4.4. Under model (4.1.) and the EBLF, $Ly + a \overset{\mathrm{L}}{\sqsubset} \beta$ holds if and only if $Ly \overset{\mathrm{L}^h}{\sqsubset} \beta$ and $a \in C\left(LX - I_p\right)$ holds simultaneously.

Proof: If a linear estimator $\beta_* = Ly + a \in \mathrm{L}$, then, using Equation (2.32.) and (2.34.), we get the risk of its with respect to the EBLF.

$$\begin{aligned} R\left(Ly + a; \beta, \sigma^2\right) &= E\left[t_1 \left(y - X(Ly + a)\right)' \left(y - X(Ly + a)\right) \right] \\ &\quad + t_2 E\left[\left((Ly + a) - \beta\right)' X'X \left((Ly + a) - \beta\right) \right] \\ &\quad + \left[(1 - t_1 - t_2)\mathrm{E}\left(X(Ly + a) - y\right)' X \left((Ly + a) - \beta\right) \right] \end{aligned}$$

$$
\begin{aligned}
&= \sigma^2 tr\left[t_1 \left(I_n - XL \right)' \left(I_n - XL \right) + t_2 \left(L'X'XL \right) \right] \\
&+ \sigma^2 tr\left[\left(1 - t_1 - t_2 \right) \left(\left(XL - I_n \right)' XL \right) \right] \\
&+ \left[\left(LX - I_p \right) \beta + a \right]' X'X \left[\left(LX - I_p \right) \beta + a \right].
\end{aligned} \tag{4.16.}
$$

(The proof of Equation (4.16.) is given in detail in the Appendix 3).

As $a \in C\left(LX - I_p \right)$, there exists $a_0 \in R^p$ such that $a = \left(LX - I_p \right) a_0$. Hence, using (4.15.) we get

$$
R\left(Ly + a;\, \beta, \sigma^2 \right) = R\left(Ly;\, \beta + a_0, \sigma^2 \right). \tag{4.17.}
$$

Suppose $Ly + a$ ∅ β, then there exists an estimator $Qy + q$ such that

$$
R\left(Qy + q;\, \beta, \sigma^2 \right) \le R\left(Ly + a;\, \beta, \sigma^2 \right), \tag{4.18.}
$$

holds for all $\beta \in R^p$ and $\sigma^2 > 0$.

Let $\beta = -a_0$ in (4.18.), then by (4.16.) and (4.17.)

$$
\begin{aligned}
&\left[q - \left(QX - I_p \right) a_0 \right]' X'X \left[q - \left(QX - I_p \right) a_0 \right] \\
&\qquad = \lim_{\sigma^2 \to 0^+} R\left(Qy + q;\, -a_0, \sigma^2 \right) \\
&\qquad \le \lim_{\sigma^2 \to 0^+} R\left(Ly + a;\, \beta, \sigma^2 \right) \\
&\qquad = \lim_{\sigma^2 \to 0^+} R\left(Ly;\, 0, \sigma^2 \right).
\end{aligned} \tag{4.19.}
$$

Since $X'X > 0$ and the right side of inequality (4.19.) is approaching zero. Therefore,

$$q = \left(QX - I_p\right)a_0 . \tag{4.20.}$$

From (4.17.), (4.18.) and (4.20.),

$$\begin{aligned} R\left(Qy;\, \beta + a_0,\, \sigma^2\right) &= R\left(Qy + a_0;\, \beta,\, \sigma^2\right) \\ &\le R\left(Ly + a_0;\, \beta,\, \sigma^2\right) = R\left(Ly;\, \beta + a_0,\, \sigma^2\right), \end{aligned}$$

and we get

$$R\left(Qy;\, \beta_0,\, \sigma^2\right) \le R\left(Ly;\, \beta_0,\, \sigma^2\right),$$

for all $\beta_0 \in R^p$ and $\sigma^2 > 0$, where $\beta_0 = \beta + a_0$. Which means that Qy is better than Ly, this is contradicting $Ly \in \beta$. Hence, the sufficiency is proved.

Necessity. Firstly, it is proved that $a \in C\left(LX - I_p\right)$ when $Ly + a \in \beta$.

Denote S as the orthogonal matrix of projection onto $C\left\{\left(X'X\right)^{1/2}\left(LX - I_p\right)\right\}$ and $s = \left(X'X\right)^{-1/2} S \left(X'X\right)^{1/2} a$. Let $\eta = \left(X'X\right)^{1/2}(LX - I_p)$, using the knowledge of projection matrix in Groß (2003) (see pp.242-243), we get

$$\begin{aligned} S &= \eta(\eta'\eta)^{+}\eta' \\ &= \left(X'X\right)^{1/2}(LX - I_p)\left[(LX - I_p)'\left(X'X\right)(LX - I_p)\right]^{+}, \\ &\times (LX - I_p)'\left(X'X\right)^{1/2} \end{aligned} \tag{4.21.}$$

and

$$s=(LX-I_p)\left[(LX-I_p)'\left(X'X\right)(LX-I_p)\right]^{+}(LX-I_p)'\left(X'X\right)a\,. \qquad (4.22.)$$

So, $s\in C\left(LX-I_p\right)$ holds. Using Moore - Penrose properties such as $\eta'\eta\left(\eta'\eta\right)^{+}\eta'=\eta'$, we have

$$\begin{aligned}&R\left(Ly+a;\,\beta,\,\sigma^2\right)-R\left(Ly+s;\,\beta,\,\sigma^2\right)\\&=a'X'Xa-s'X'Xs=a'\left(X'X\right)^{1/2}\left(I_p-S\right)\left(X'X\right)^{1/2}a\geq 0.\end{aligned}$$

From which we get

$$R\left(Ly+a;\,\beta,\,\sigma^2\right)\geq R\left(Ly+s;\,\beta,\,\sigma^2\right), \qquad (4.23.)$$

for all $\beta\in R^p$ and $\sigma^2>0$. The inequality holds if and only if $(X'X)^{1/2}a=S(X'X)^{1/2}a$ holds, which is $a=(X'X)^{-1/2}S(X'X)^{1/2}a=s$. Inequality (4.23.) and the admissibility of $Ly+a$ together result in $a=s$. Thus $a\in C\left(LX-I_p\right)$ is right.

Next, we prove that $Ly\in\beta$, when $Ly+a\in\beta$ holds.

If $Ly\notin\beta$, there exists an estimator Ky such that

$$R\left(Ky;\,\beta,\,\sigma^2\right)\leq R\left(Ly;\,\beta,\,\sigma^2\right), \qquad (4.24.)$$

is right for all (β,σ^2) with the strict inequality holding at some point (β_1,σ_1^2). As $a\in C\left(LX-I_p\right)$, there exists $a_0\in R^p$ such that $a=\left(LX-I_p\right)a_0$.

So, by (4.24.),

$$R\left(Ly+a;\,\beta,\,\sigma^2\right)=R\left(Ly;\,\beta+a_0,\,\sigma^2\right)$$
$$\geq R\left(Ky;\,\beta+a_0,\,\sigma^2\right)=R\left(Ky+(KX-I_p)a_0;\,\beta,\,\sigma^2\right),$$

holds for all (β,σ^2) and the strict inequality is true at point $\left(\beta_1-a_0,\sigma_1^2\right)$, which means that $Ky+(KX-I_p)a_0$ is better than $Ly+a$, this is contradictory to the admissibility of $Ly+a$. Hence necessity is proved. Therefore, the proof of Theorem 4.4 is completed.

4.1.2. Implicit Characterization

Theorem 4.3 and Theorem 4.4 provide the explicit form of L and a which is necessary and sufficient for $Ly+a\overset{L}{\square}\beta$. From the following theorem, the implicit characterizations of ALEs $Ly+a$ by giving necessary and sufficient algebraic conditions on L and a can be deduced.

Theorem 4.5. Under model (4.1.) and the EBLF, $Ly+a\overset{L}{\square}\beta$ holds if and only if

i. $XL=L'X'$,

ii. $\lambda(XL)\subset[0,1]$,

iii. $a\in C\left(LX-I_p\right)$.

Proof: Using (4.9.) in Theorem 4.3, we have

$$XL=wX\left(X'X\right)^{-1}X'+(1-w)X\left(X'X\right)^{-1/2}D\left(X'X\right)^{-1/2}X'. \qquad (4.25.)$$

If we set

$$H=X\left(X'X\right)^{-1}X',\ G=X\left(X'X\right)^{-1/2}D\left(X'X\right)^{-1/2}X',$$

since D and H is symmetric, then $XL = wH + (1-w)G$ is symmetric, which means condition (i) satisfied. Using Theorem 2.6 we know that, the non-zero eigenvalues of G coincide with the nonzero eigenvalues of $D(X'X)^{-1/2} X'X (X'X)^{-1/2} = D$, where $\lambda(D) \subset [0,1]$. This shows that, all eigenvalues of the matrix G lie in [0,1], which denoted by $\lambda(G) \subset [0,1]$.

Let H and G be two symmetric $n \times n$ matrices satisfying $GH = HG$. Then from Theorem 2.6, there exists an $n \times n$ orthogonal matrix U such that

$$H = U \Lambda U' \text{ and } G = U \Gamma U',$$

where Λ and Γ are two $n \times n$ diagonal matrices. Then, XL can be written as

$$XL = U(w\Lambda + (1-w)\Gamma)U' .$$

Since H is idempotent, H has only eigenvalues 0 or 1 and $\lambda(G) \subset [0,1]$ means that diagonal element $\gamma_i, i = 1,..., p$ lies in the interval $[0,1]$. Then, we get

$$\lambda(w\Lambda + (1-w)\Gamma) \subset [0,1] \Leftrightarrow \lambda(XL) \subset [0,1] ,$$

where $0 \le w \le 1$ (see Graybill, 1983, p.43). So, condition (ii) also satisfied.

Using results of Theorem 4.4, we have $a \in C(LX - I_p)$. Thus, the proof of Theorem 4.5 is completed.

4.2. Characterization of ALE under the EBLF

In this section, the new results about the characterization of ALE with respect to the EBLF are obtained.

4.2.1. OLS and ALE

In this section, we will consider the relationship between the OLS and ALE with respect to the EBLF. Under the condition $a \in C\left(I_p - LX\right)$ in Theorem 4.3 and Theorem 4.4, there existence of some vector d with

$$a = \left(I_p - LX\right)d = (1-w)S^{-1/2}(I_p - D)S^{1/2}d \ .$$

Using Theorem 4.4, for any ALE for β can be written in the form

$$\begin{aligned} \beta_* &= Ly + a \\ &= \left(wI_p + (1-w)S^{-1/2}DS^{1/2}\right)\hat{\beta} + (1-w)S^{-1/2}\left(I_p - D\right)S^{1/2}d, \end{aligned} \tag{4.26.}$$

where $S = X'X$. This shows that any ALE is any linear transformation of the $\hat{\beta}$.

Firstly, when we examine their covariance matrices, we see from the following theorem that the covariance matrix of an ALE for β under the EBLF is always comparable to the covariance matrix of the OLS with respect to the Löwner partial ordering.

Theorem 4.6. Under LRM (4.1) and the EBLF, let $Ly + a \overset{L}{\sqsubset} \beta$ holds. Then the inequality

$$Cov(Ly + a) \leq_L Cov(\hat{\beta}) \ ,$$

holds true.

Proof: By using Theorem 4.5, we have XL is symmetric and $XL = U\Lambda U'$, where U is an orthogonal matrix and Λ is a diagonal matrix containing the eigenvalues of XL on its main diagonal. Since $\lambda(XL) \subset [0,1]$, then

$$\lambda(\Lambda^2) \subset [0,1] \Leftrightarrow U(I_n - \Lambda^2)U' = I_n - XLL'X' \geq 0$$

is n.n.d. Multiplying this matrix from the left by $(X'X)^{-1}X'$ and from the right by $X(X'X)^{-1}$ yields the following n.n.d. matrix

$$(X'X)^{-1}X'X(X'X)^{-1} - (X'X)^{-1}X'XLL'X'X(X'X)^{-1},$$

being identical to

$$(X'X)^{-1} - LL' \geq 0.$$

This means that $Cov(\hat{\beta}) - Cov(Ly + a) \geq 0$.

From the above theorem it follows that for any estimator $Ly + a \in \mathrm{L}$, if $Ly + a \sqsubset \beta$ holds, then $Cov(Ly + a) \leq_L Cov(\hat{\beta})$ holds simultaneously. That means, of course, also the variance of each element of $Ly + a$ is smaller than or equal to the variance of the corresponding element of $\hat{\beta}$.

However, as we have known from the Gauss-Markov Theorem 2.17 that the OLS is an unbiased estimator with minimum variance. The above result also shows that ALEs cannot be unbiased for β. As matter of fact, the following holds true.

Theorem 4.7. In LRM (4.1) and the EBLF, let $Ly+a \overset{L}{\sqsubset} \beta$ holds. Then $Ly+a$ is unbiased for β if and only if

$$Ly+a=\hat{\beta},$$

holds true.

Proof: Since $E(Ly+a)=\beta$ holds if and only if $a=0$ and $LX=I_p$. In view of the above identity can equivalently be written as $XLX=X$. When $Ly+a \sqsubset \beta$, then $XL=L'X'$, implying that the identity

$$XLX=X \Leftrightarrow L'X'X=X,$$

i.e. $L'=X\left(X'X\right)^{-1}$, so that $L=\left(X'X\right)^{-1}X'$.

Now, we have seen that $Ly+a \sqsubset \beta$ which does not coincide with β is necessarily a biased estimator and has a $Cov(Ly+a) \leq_L Cov(\hat{\beta})$. Following Example 4.1 shows that the converse is however not true, that is, there may be biased estimators for β that have a smaller covariance matrix than the OLS estimator, but nonetheless being not linear admissible for β. Let's give the shrinkage estimator as an example below.

Example 4.1. Consider the linear shrinkage estimator

$$\hat{\beta}^{-}(\rho)=-\frac{1}{1+\rho}\hat{\beta}, \quad \rho \geq 0,$$

with the negative shrinkage parameter $-1/(1+\rho)$. Then this estimator $\hat{\beta}^{-}(\rho)$ has the same covariance matrix as the usual shrinkage estimator $\hat{\beta}(\rho)=\frac{1}{1+\rho}\hat{\beta}$, i.e.

$$Cov\left(\hat{\beta}^{-}(\rho)\right)=\frac{\sigma^2}{(1+\rho)^2}(X'X)^{-1}.$$

Clearly, the difference

$$Cov\left(\hat{\beta}\right)-Cov\left(\hat{\beta}^{-}(\rho)\right)=\frac{\rho(2+\rho)\sigma^2}{(1+\rho)^2}(X'X)^{-1},$$

is n.n.d. whenever $\rho\geq 0$. Moreover, $E\left(\hat{\beta}^{-}(\rho)\right)=-\frac{1}{1+\rho}\beta\neq\beta$ for any vector $\beta\neq 0$, so that $\hat{\beta}^{-}(\rho)$ is not an unbiased estimator for β.

To examine whether $\hat{\beta}^{-}(\rho)$ is linear admissible for β or not, we write

$$\hat{\beta}^{-}(\rho)=Ly+a,\quad L=-\frac{1}{1+\rho}(X'X)X',\quad a=0.$$

Then the eigenvalues of LX are all equal to $-1/(1+\rho)$ and thus dot not lie in the interval [0,1]. Hence, from Theorem 4.4 it follows that $\hat{\beta}^{-}(\rho)$ ▯ β.

4.2.2. Linear Transforms of OLS

Theorem 4.3 shows that the ALE under the EBLF is a CC of the ALEs $\hat{\beta}$ and $\tilde{\beta}_{mse}$, where $\tilde{\beta}_{mse}$ is ALE with respect to the QLF. So, ALE is equal to

$$\begin{aligned}\beta_{EBLF} &= w\hat{\beta} + (1-w)\tilde{\beta}_{mse} \\ &= \left[wI_p + (1-w)L_{mse} \right]\hat{\beta} \\ &= \tilde{L}\hat{\beta},\end{aligned} \tag{4.27.}$$

where $\tilde{\beta}_{mse} = Ay = \left(X'X\right)^{-1/2} D\left(X'X\right)^{-1/2} X'y = L_{mse}\hat{\beta}$, $L_{mse} = \left(X'X\right)^{-1/2} D\left(X'X\right)^{1/2}$

We can come to the conclusion that any HL estimator being admissible for β within the set of linear (not necessarily homogeneous) estimators can be written as a linear transformation $\tilde{L}\hat{\beta}$ of the OLS estimator for some specific matrix $\tilde{L}$. On the other hand, the following holds true.

Theorem 4.8. Given some matrix $\tilde{L}$, the linear transformation $\tilde{L}\hat{\beta}$ □ β holds with respect to the EBLF if and only if the conditions

$$X'X\tilde{L} = \tilde{L}'X'X \quad \text{and} \quad \lambda(\tilde{L}) \subset [0,1],$$

are satisfied.

Proof: Necessity.

If $\tilde{L}\hat{\beta}$ □ β, from (4.27)

$$\beta_{EBLF} = \tilde{L}\hat{\beta} = \tilde{L}(X'X)^{-1}X'y = Ly,$$

and Theorem 4.5, we have

$$XL = L'X' \;\Leftrightarrow\; X\tilde{L}\left(X'X\right)^{-1}X' = X\left(X'X\right)^{-1}\tilde{L}'X' \Leftrightarrow$$

$$X'X\tilde{L}\left(X'X\right)^{-1}X'X = X'X\left(X'X\right)^{-1}\tilde{L}'X'X \Leftrightarrow$$
$$X'X\tilde{L} = \tilde{L}'X'X\ .$$

We can get the same results in another way by using $X'XL_{mse} = L'_{mse}X'X$,

$$\begin{aligned} X'X\tilde{L} &= wX'X + (1-w)X'XL_{mse} \\ &= wX'X + (1-w)L'_{mse}X'X \\ &= \left[wI_p + (1-w)L'_{mse}\right]X'X = \tilde{L}X'X. \end{aligned}$$

Since $\lambda(XL) \subset [0,1]$, using Theorem 2.5. we have

$$\lambda(XL) = \lambda(LX) \subset [0,1],$$

Then

$$\lambda(X\tilde{L}\left(X'X\right)^{-1}X') = \lambda(\tilde{L}\left(X'X\right)^{-1}X'X) = \lambda(\tilde{L}) \subset [0,1].$$

Sufficiency.

If $X'X\tilde{L} = \tilde{L}'X'X$ and $\lambda(\tilde{L}) \subset [0,1]$ are satisfied, then from

$$X'X\tilde{L} = \tilde{L}'X'X \ \Leftrightarrow \ XL = L'X',$$

and Theorem 4.5 we can easily get $\tilde{L}\hat{\beta}$ □ β . Thus, the proof of Theorem 4.7 is completed.

If $\tilde{L}$ is a symmetric matrix, then the condition $X'X\tilde{L} = \tilde{L}X'X$ is fulfilled if and only if there exists an orthogonal matrix U and diagonal matrices Λ and Γ such that

$$X'X = U\Lambda U' \quad \text{and} \quad \tilde{L} = U\Gamma U' ,$$

(see also Theorem 2.6). Therefore, the condition $\lambda(\tilde{L}) \subset [0,1]$ means that every diagonal element γ_i, $i = 1,..., p$, lies in the interval $[0,1]$.

The Theorem 4.8 shows that such linear transformations $\tilde{L}\hat{\beta}$ are related to the set of estimators introduced by Obenchain [3], which can be written as

$$\mathcal{L}^{Ob}(\beta) = U\Gamma U'\hat{\beta}, \quad \Gamma = diag(\gamma_1,...,\gamma_p), \quad 0 \le \gamma_i \le 1,$$

where the orthogonal matrix U stems from a spectral decomposition $U\Lambda U'$ of $X'X$. In view of the above considerations we can also write

$$\mathcal{L}^{Ob}(\beta) = \left\{ \tilde{L}\hat{\beta} : \tilde{L} = \tilde{L}',\ X'X\tilde{L} = \tilde{L}X'X,\ \lambda(\tilde{L}) \subset [0,1] \right\},$$

which we will call the OBC of estimators for β.

Any $\tilde{L}\hat{\beta} \in \mathcal{L}^{Ob}(\beta)$ is linear admissible for β and in addition satisfies

$$\left\| \tilde{L}\hat{\beta} \right\|^2 = \hat{\beta}'\tilde{L}'\tilde{L}\hat{\beta} \le \left\| \hat{\beta} \right\|^2 . \qquad (4.28.)$$

The OBC is sometimes referred to as the class of generalized ridge estimators in the literature. But an estimator for the OBC is not necessarily of the form of a GR

estimator $\hat{\beta}_K = (X'X + K)^{-1} X'y$, where K is a symmetric n.n.d. matrix. Now we will investigate conditions under which an element from $\mathcal{L}^{Ob}(\beta)$ can be written as $\hat{\beta}_K$ for some K.

Theorem 4.9. Under the LRM (4.1), an estimator $\tilde{L}\hat{\beta} \in \mathcal{L}^{Ob}(\beta)$ can be written as $\tilde{L}\hat{\beta} = \hat{\beta}_K$ for some $K \in \Re_s^{\geq}$ irrespective of the realization $y \in \Re^n$ if and only if $\lambda(\tilde{L}) \subset (0,1]$.

Proof: If we suppose that $\tilde{L}\hat{\beta} \in \mathcal{L}^{Ob}(\beta)$ can be written as $\tilde{L}\hat{\beta} = \hat{\beta}_K$ for some $K \in \Re_s^{\geq}$ and every $y \in \Re^n$, then

$$\tilde{L}(X'X)^{-1}X' = (X'X + K)^{-1}X',$$

and therefore

$$\tilde{L} = (X'X + K)^{-1}X'X \,. \tag{4.29.}$$

Since $\tilde{L} = \tilde{L}'$ and $R\left((X'X + K)^{-1}X'X\right) = p$, using Theorem 2.11 $\tilde{L}$ has p non-zero real eigenvalues, so that $\lambda(\tilde{L}) \subset (0,1]$.

Now, let $X'X = U\Lambda U'$. if $\tilde{L}\hat{\beta} \in \mathcal{L}^{Ob}(\beta)$ with $\lambda(\tilde{L}) \subset (0,1]$, then the matrix $\tilde{L}$ can be written as $\tilde{L} = U\Gamma U'$, where Γ is a diagonal matrix whose every diagonal element lies in the interval (0,1]. Therefore, from (4.29.), we have

$$(X'X + K) = X'X\tilde{L}^{-1} \;\Leftrightarrow\; U\Lambda U' + K = U(\Lambda\Gamma^{-1})U' \Leftrightarrow$$

$$K = U(\Lambda\Gamma^{-1} - \Lambda)U',$$

is symmetric n.n.d., and

$$\begin{aligned}(X'X + K)^{-1} X'y &= U\Gamma\Lambda^{-1}U'X'y \\ &= U\Gamma U'U\Lambda^{-1}U'X'y \\ &= \tilde{L}\hat{\beta},\end{aligned}$$

thus concluding the proof.

In view of the identity $\tilde{L} = wI_p + (1-w)L_{mse}$, the $L_{mse} = (X'X + K)^{-1}X'X$ is satisfied when $w = 0$. This is same as the results in (Groß ,2003 .p225).

Table 4.1. Displays some well-known estimators (see Appendix 5) from the $\mathcal{L}^{Ob}(\beta)$ with respect to a given spectral decomposition $X'X = U\Lambda U' = U_1\Lambda_1U_1' + U_2\Lambda_2U_2'$.

We know that the idea behind the ridge estimator is quite different from the idea behind the principal components estimator. In the following example, we show the different behavior of both estimators under the criteria such as EBLF and mse.

Example 4.2. Consider the LRM $y = X\beta + \varepsilon$ with

$$X' = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 2.9 & 3.6 & 3.9 & 3.7 & 4.4 & 4.1 & 4.7 & 3.7 & 4.5 & 5.1 & 5.7 & 5.5 \end{pmatrix}, \quad \beta = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix},$$

$$y' = (4.0; 4.5; 5.1; 5.2; 6.0; 5.3; 6.3; 7.5; 6.0; 5.8; 6.9; 7.0),$$

True parameter vector β given as $\beta' = (0.5, 1.2)$. While the ridge estimate $\hat{\beta}_k = (X'X + kI_2)^{-1} X'y$ for the choice $k = 1/5$ is $\hat{\beta}_{1/5} = \begin{pmatrix} 1.5208 \\ 0.9855 \end{pmatrix}$.

Since the spectral decomposition above, the principal components estimator $\hat{\beta}^{(1)}$ comes up to

$$\hat{\beta}^{(1)} = u_1 (u_1' X'X u_1)^{-1} u_1' X'y = \begin{pmatrix} 0.2837 \\ 1.2641 \end{pmatrix},$$

where $u_1' = (0.2189, 0.9757)$ is first column of the matrix U corresponding to the max eigenvalue $\lambda_1 = 242.8436$. Then ridge estimator and principal component estimator is distinctly closer to the true parameter $\beta' = (0.5, 1.2)$ than the value

$$\hat{\beta} = (X'X)^{-1} X'y = \begin{pmatrix} 2.1785 \\ 0.8389 \end{pmatrix},$$

of the OLS estimator.

Following Fig 4.1 shows the 30 estimates $\hat{\beta}^{(1)}$ when $r = 1$ and $\hat{\beta}_{1/5}$ for the 30 realizations of y which we have generated using Monte Carlo simulation.

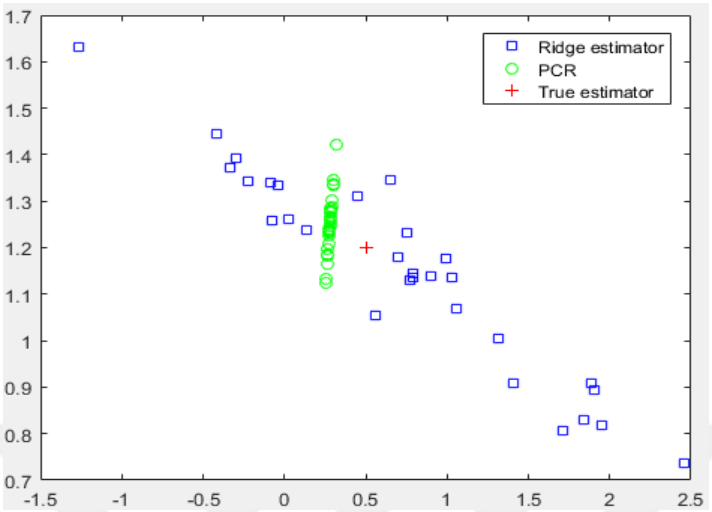

Figure 4.1. $\hat{\beta}^{(1)}$, $\hat{\beta}_{1/5}$ and true estimator when n= 30.

Note that although the loss values of $\hat{\beta}^{(1)}$ and $\hat{\beta}_{1/5}$ are rather similar, and the estimates do in fact scatter around these values, the behavior of the estimators is quite different.

Unweighted squared error loss corresponding the both estimators are $mse(\hat{\beta}) = 2.9478$, $mse(\hat{\beta}_{1/5}) = 1.0880$, $mse(\hat{\beta}^{(1)}) = 0.0509$. EBLF values corresponding the estimators are $EBLF(\hat{\beta}) = 3.0819$, $EBLF(\hat{\beta}^{(1)}) = 3.2465$, $EBLF(\hat{\beta}_{1/5}) = 2.8137$.

4.2.3. About the Admissibility of Some Known Estimators

As mentioned in last section, if some HL estimator from the OBC, then it is also linear admissible for β within the set of all linear estimators. Also Table 4.1 implies that quite a few of the known estimators in fact belong to $\mathcal{L}^{Ob}(\beta)$. Nonetheless, there exist (possibly non-homogeneous) some estimators such as GR

type and restricted estimators, which do not belong to this class and thus in this section we will discuss the admissibility of such estimators.

4.2.3.1. GR Estimator

A non-homogeneous GR estimator is of the form

$$\hat{\beta}_{K,\beta_0} = \beta_0 + (X'X + K)X'(y - X\beta_0) .$$

When the $p \times 1$ vector β_0 and the $p \times p$ symmetric nonnegative definite matrix K are chosen non-stochastically, then $\hat{\beta}_{K,\beta_0}$ is also called a linear estimator of ridge type (see Groß, 2003, p.227).

Theorem 4.10. Under the model (4.1), let $K \in \Re_s^{\geq}$ be a given $p \times p$ matrix and let $\beta_0 \in \Re^p$. Then the GR estimator

$$\hat{\beta}_{K,\beta_0} = \beta_0 + (X'X + K)^{-1}X'(y - X\beta_0),$$

is linear admissible for β with respect to the EBLF.

Proof: We can write $\hat{\beta}_{K,\beta_0} = Ly + a$ with

$$L = (X'X + K)^{-1}X',$$

and

$$a = \beta_0 - (X'X + K)^{-1} X'X\beta_0 = (X'X + K)^{-1} K\beta_0 .$$

It can easily be checked that

$$I_p - LX = (X'X + K)^{-1} K ,$$

and

$$XL = L'X' ,$$

then the condition $a \in C(I_p - LX)$ is satisfied. Moreover, according to Theorem 4.5 we can easily get $\lambda(XL) \in [0,1]$.

Since $K \in \Re_s^{\geq}$ and $XL \in \Re_s^{\geq}$, which showing that all eigenvalues of XL are real nonnegative. In addition,

$$X'X \leq_L X'X + K ,$$

implying that

$$\left(X'X + K\right)^{-1} \leq_L \left(X'X\right)^{-1}.$$

Multiplication of this inequality from the left by X and from the right by X' shows

$$X\left(X'X + K\right)^{-1} X' \leq_L X\left(X'X\right)^{-1} X' .$$

Let

$$H_K = X\left(X'X + K\right)^{-1} X',\ H = X\left(X'X\right)^{-1} X'.$$

Then, From Theorem 2.9, any eigenvalues of H_K is not greater than the largest eigenvalue of H . Since $H_K = XL$ and H symmetric and idempotent, $\lambda(H)$ are only 0 or 1 (with at least one eigenvalue 1 when X is non-zero). We can easily get $\lambda\left(H_k\right) = \lambda\left(XL\right) \subset [0,1]$. Thus the proof is completed.

Table 3.1 is given in chapter 3, displays linear ridge type estimators, some of them also belonging to $\mathcal{L}^{Ob}(\beta)$. From the above results, any of these estimators is linear admissible for β .

4.2.3.2. Restricted Estimator

A linear estimator $\beta_* \in \mathrm{L}$ satisfying the identity

$$R\beta_* = r$$

for some given $R \neq 0$, some given r , and any possible realization $y \in \Re^n$ can be called a restricted linear estimator (RLS).

A linear ridge type estimator $\hat{\beta}_{K,\beta_0}$ cannot be a restricted estimator in this sense when the trivial case $(R, r) = (0, 0)$ is excluded. On the other hand, the RLS $\hat{\beta}_R$ is of course a linear restricted estimator. The admissibility of $\hat{\beta}_R$ within the set of linear estimators is guaranteed by the following result.

Theorem 4.11. Under the model (4.1), let $R \in \Re^{p\times n}$ be a given matrix of full row rank and let r be a given $m \times 1$ vector. Then the RLS estimator

$$\hat{\beta}_R = \hat{\beta} - S^{-1}R'\left[RS^{-1}R'\right]^{-1}(R\hat{\beta} - r),$$

is linear admissible for β, where $S = X'X$.

Proof: We can write $\hat{\beta}_R = Ly + a$ with

$$L = S^{-1}X' - S^{-1}R'\left[RS^{-1}R'\right]^{-1}RS^{-1}X',$$

$$a = S^{-1}R'\left[RS^{-1}R'\right]^{-1}r.$$

Then, the matrix XL is obviously symmetric and it can be written as $XL = U\Lambda U'$, where U is an orthogonal matrix and Λ is a diagonal matrix containing the eigenvalues $\lambda_1,...,\lambda_n$ of XL on its main diagonal. On the other hand, it can easily be implied that $XLXL = XL$ is satisfied.
Then this equality can be written as

$$XL - XLL'X' = U\left(\Lambda - \Lambda^2\right)U' = 0,$$

this matrix is n.n.d. if and only if $\Lambda - \Lambda^2$ is n.n.d., which in turn is equivalent to

$$\lambda_i - \lambda_i^2 = 0, \quad i = 1,...,n.$$

Since XL symmetric and idempotent, then the eigenvalues of XL are only 0 or 1 (with at least one eigenvalue 1 when X is non-zero), it is shown that $\lambda(XL) \subset [0,1]$.

Since $a=(I_p-LX)R'(RR')^{-1}r$, according to the Theorem 4.5, $\hat{\beta}_R$ is linear admissible for β .

As we have known from corresponding sources, Principal components estimator $\hat{\beta}^{(r)}$ is a specific RLS estimator which satisfy the condition

$$U_2'\hat{\beta}^{(r)}=0,$$

where $r<p$, and $U_1\Lambda_1U_1'+U_2\Lambda_2U_2'$ is a spectral decomposition of the matrix $X'X$. We can obtain that Principal components estimator $\hat{\beta}^{(r)}$ is admissible for β in two ways, one can be deducted from the above theorem and the other can be obtained since $\hat{\beta}^{(r)}$ belongs to the OBC.

There are some estimator such as $\hat{\beta}_k^{(r)}$ from Baye and Parker (1984), which is also a restricted estimator satisfying $U_2'\hat{\beta}_k^{(r)}=0$ for any realization $y\in\Re^n$. At the same time, it is not a specific RLS estimator. The linear admissibility of this estimator follows from the fact that $\hat{\beta}^{(r)}$ belongs to the OBC.

4.2.3.3. Shrinkage Property (ShP) and Linear Admissibility

From Section 4.2.1, we have already known that any estimator $\tilde{L}\hat{\beta}\in\mathcal{L}^{Ob}(\beta)$ satisfies the ShP

$$\left\|\tilde{L}\hat{\beta}\right\|^2\le\left\|\hat{\beta}\right\|^2 \text{ for every realization } y\in\Re^n .$$

Then, of course, also

$$E\left(\left\|\tilde{L}\hat{\beta}\right\|^2\right) \leq E\left(\left\|\hat{\beta}\right\|^2\right),$$

holds true.

In the light of the above information, it has been a curiosity to examine whether any element from the set of ALEs has this ShP. Now, we can try to find some answers by examining the set of all these HL estimators which satisfying the shrinkage condition.

Theorem 4.11. Under the model (4.1) and EBLF, a HL estimator Ly of β satisfies the condition

$$\|Ly\|^2 \leq \left\|\hat{\beta}\right\|^2 \text{ for every realization } y \in \Re^n,$$

if and only if the matrix L can be written as

$$L = \tilde{L}(X'X)^{-1}X',$$

for some $p \times p$ matrix $\tilde{L}$ such that $\lambda(\tilde{L}'\tilde{L}) \subset [0,1]$.

Proof: Let us notice in advance that

$$\lambda(\tilde{L}'\tilde{L}) \subset [0,1] \Leftrightarrow \tilde{L}'\tilde{L} \leq_L I_p.$$

We can proof this using same method as Groß (2003) p.230, here we omit it.

The above theorem shows that any HL estimator satisfying the shrinkage condition is of the form $L\hat{\beta}$ with $\lambda(\tilde{L}'\tilde{L}) \subset [0,1]$.

The remaining question is any ALE satisfies the ShP. For this, we will also examine whether the HL estimator satisfying the shrinkage condition is linearly admissible for β . The following statements show that the answer is negative.

As we know from Sect. 4.2.1, any ALE can be written as $L\hat{\beta}$ for some matrix which fulfills the conditions $X'X\tilde{L} = \tilde{L}'X'X$ and $\lambda(\tilde{L}) \subset [0,1]$. Moreover, since it is easy to find matrices $\tilde{L}$ which satisfy the condition $\lambda(\tilde{L}'\tilde{L}) \subset [0,1]$ but which do not satisfy the condition $X'X\tilde{L} = \tilde{L}'X'X$. From above of them, HL estimator satisfying the shrinkage condition is not necessarily linearly admissible for β . Above mentioned also leads to the question whether the latter two conditions also imply the condition $\lambda(\tilde{L}'\tilde{L}) \subset [0,1]$. From the following example we will try to answer for this.

Example 4.3. Consider a LRM where the matrix $X'X$ is given as

$$X'X = \begin{pmatrix} 4.6 & 10.2 \\ 10.2 & 23.4 \end{pmatrix}.$$

The matrix

$$\tilde{L} = \begin{pmatrix} 0.7 & -0.6 \\ 0.1 & 1.2 \end{pmatrix},$$

has eigenvalues $\lambda_1 = 0.9$ and $\lambda_2 = 1$, so that all eigenvalues of $\tilde{L}$ lie in the closed interval $[0,1]$. Moreover, the matrix

$$X'X\tilde{L} = \begin{pmatrix} 4.24 & 9.48 \\ 9.48 & 21.96 \end{pmatrix}$$

is symmetric, which means that $\tilde{L}\hat{\beta}$ β. On the other hand the $\lambda(\tilde{L}'\tilde{L})$ given as $\lambda_1 = 0.4341$, $\lambda_2 = 1.8659$, so that one of $\lambda(\tilde{L}'\tilde{L})$ is strictly greater than 1 and the estimator $\tilde{L}\hat{\beta}$ has not the above described ShP.

Now, by the above example we can conclude that there can exist ALEs for which the condition $\lambda(\tilde{L}'\tilde{L}) \subset [0,1]$ does not hold true.

However, if we confine to symmetric matrices $\tilde{L}$ (implying that $\tilde{L}\hat{\beta}$ is an element from the OBC), then the condition implies $\lambda(\tilde{L}'\tilde{L}) \subset [0,1]$. A natural generalization of the OBC of estimators is therefore

$$\left\{ \tilde{L}\hat{\beta} :\ X'X\tilde{L} = \tilde{L}X'X,\ \lambda(\tilde{L}) \subset [0,1], \lambda(\tilde{L}'\tilde{L}) \subset [0,1] \right\},$$

which describes the set of all HL estimators with the ShP which are also linear admissible for β.

The matrix $\tilde{L}$ from Example 4.3 can also be written in the form $\tilde{L} = (X'X + K)^{-1} X'X$ for

$$K = \begin{pmatrix} 0.4 & 0.8 \\ 0.8 & 1.6 \end{pmatrix},$$

where the matrix K is symmetric n.n.d. From this, we can also conclude that not any estimator

$$\hat{\beta}_K = (X'X + K)^{-1} X'y = (X'X + K)^{-1} X'X\hat{\beta},$$

has the ShP.

Remark. If $\hat{\beta}_K = (X'X + K)^{-1} X'y$ for some $K \in \Re_s^{\geq}$, then it does not necessarily follow that $\left\|\hat{\beta}_K\right\|^2 \leq \left\|\hat{\beta}\right\|^2$.

4.2.3.4. CC of Estimators

If we combine the ALEs, it is natural to ask whether CC is again linear admissible. In this section we will try to answer this question. The following theorem considers the CC of two HL estimators.

Theorem 4.12. Under the linear model (4.1), let $L_1 y \in \mathrm{L}^h$ and $L_2 y \in \mathrm{L}^h$ be linear admissible for β with respect to the EBLF. Then under the some special condition like $w_1 = w_2 = w$, for an arbitrary non-stochastic number $0 \leq \alpha \leq 1$ the estimator $\alpha L_1 y + (1-\alpha) L_2 y$ is linear admissible for β.

Proof: Let $S = X'X$. Using Theorem 4.3, we can write

$$L_1 = w_1 S^{-1} X' + (1 - w_1) S^{-1/2} D_1 S^{-1/2} X'$$

and

$$L_2 = w_2 S^{-1} X' + (1 - w_2) S^{-1/2} D_2 S^{-1/2} X',$$

for symmetric matrices D_1 and D_2 whose eigenvalues all belong to $[0,1]$. Then for the estimator

$L_3 y = \alpha L_1 y + (1-\alpha) L_2 y$ the matrix L_3 is given as

$$\begin{aligned} L_3 y &= \left(\alpha L_1 + (1-\alpha) L_2\right) y \\ &= \alpha\left(w_1 S^{-1} + (1-w_1) S^{-1/2} D_1 S^{-1/2}\right) X'y \\ &+ (1-\alpha)\left(w_2 S^{-1} + (1-w_2) S^{-1/2} D_2 S^{-1/2}\right) X'y \\ &= \left[\left(\alpha w_1 + (1-\alpha) w_2\right) S^{-1} + S^{-1/2}\left(\alpha(1-w_1) D_1 + (1-\alpha)(1-w_2) D_2\right) S^{-1/2}\right] X'y, \end{aligned}$$

if let $w_1 = w_2 = w$;

$$L_3 = \left[wS^{-1} + (1-w) S^{-1/2} D_3 S^{-1/2}\right] X', \quad \text{where } D_3 = \alpha D_1 + (1-\alpha) D_2 .$$

Since αD_1 and $(1-\alpha) D_2$ are symmetric n.n.d., also the matrix $D_3 \in \Re_s^{\geq}$ too and all $\lambda(D_3)$ are thus real nonnegative numbers. If $\|\cdot\|$ denotes the spectral norm of a matrix, then $\|D_1\| \leq 1$ and $\|D_2\| \leq 1$, and therefore

$$\|D_3\| = \|\alpha D_1 + (1-\alpha) D_2\| \leq \alpha \|D_1\| + (1-\alpha)\|D_2\| \leq 1.$$

this shows that no eigenvalue of D_3 can be strictly greater than 1, from Theorem 4.3 the estimator $L_3 y$ □ β. Thus, proof is complete.

From Theorem 4.12 above, it concludes that when $w_1 = w_2 = w$, a CC of HL estimators being linear admissible for β under certain condition, e.g. for elements from the OBC, the admissibility property is maintained.

Example 4.4. Consider the LRM $y = X\beta + \varepsilon$, where

$$X' = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 2.9 & 3.6 & 3.9 & 3.7 & 4.4 & 4.1 & 4.7 & 3.7 & 4.5 & 5.1 & 5.7 & 5.5 \end{pmatrix}.$$

Let us consider the simple CC $\beta_{new} = 0.5\hat{\beta}^{(1)} + 0.5\hat{\beta}_{1/5}$.

We generate 30 realizations of the vector y and compute the corresponding principal components estimates $\hat{\beta}^{(1)}$ as well as the corresponding ridge estimates $\hat{\beta}_{1/5}$. Risk value of these estimates with respect to the EBLF is respectively $R_{EBLF}(\hat{\beta}_{1/5}) = 4.4258$, $R_{EBLF}(\hat{\beta}^{(1)}) = 4.6566$.

Figures 4.2 and Figure 4.3 show a comparison between the ridge and the combined estimates. We can see that combined estimator more approaches the true value, that means that combined estimator performs better than ridge estimator. The average observed EBLF loss of $0.5\hat{\beta}^{(1)} + 0.5\hat{\beta}_{1/5}$ is 4.4506.

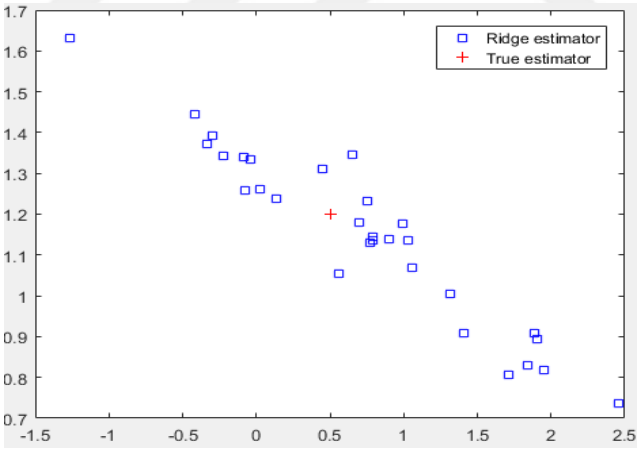

Figure.4.2. n=30 ridge estimates

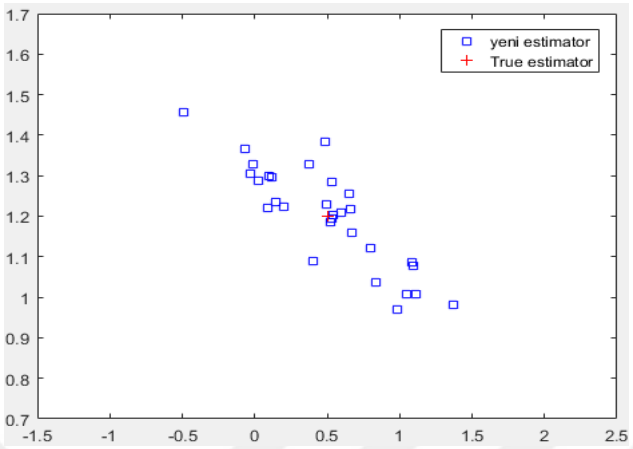

Figure.4.3. 30 combined ridge and principal components estimates $0.5\hat{\beta}^{(1)} + 0.5\hat{\beta}_{1/5}$

Now, we will consider that this property holds for non-HL estimators or not. The following theorem shows that the answer is affirmative.

Theorem 4.13. Under the LRM (4.1), let $L_1 y + a_1 \in \mathrm{L}$ and $L_2 y + a_2 \in \mathrm{L}$ be linear admissible for β with respect to the EBLF. Then for an arbitrary non-stochastic number $0 \le \alpha \le 1$ the estimator

$$\alpha(L_1 y + a_1) + (1-\alpha)(L_2 y + a_2),$$

is linear admissible for β.

Proof: Let $S = X'X$. Using Theorem 4.4, we can write

$$L_1 = w_1 S^{-1} X' + (1-w_1) S^{-1/2} D_1 S^{-1/2} X', \qquad a_1 = (1-w_1) S^{-1/2} (I_p - D_1) S^{1/2} d_1,$$

and

$$L_2 = w_2 S^{-1} X' + (1-w_2) S^{-1/2} D_2 S^{-1/2} X', \qquad a_2 = (1-w_2) S^{-1/2} (I_p - D_2) S^{1/2} d_2$$

Here D_1 and D_2 are symmetric matrices, and both $\lambda(D_1) \in [0,1]$ and $\lambda(D_2) \in [0,1]$, while d_1 and d_2 are $p \times 1$ vectors.

Now, from Theorem 4.4 and Theorem 4.13, it follows that the CC $\alpha(L_1 y + a_1) + (1-\alpha)(L_2 y + a_2) \square \beta$ holds if and only if there exists some $p \times 1$ vector d_3 such that

$a_3 = \alpha a_1 + (1-\alpha) a_2$, can be written as

$$a_3 = \left(\alpha(1-w_1) I_p + (1-\alpha)(1-w_2) I_p - \alpha(1-w_1) S^{-1/2} D_1 S^{1/2}\right) d_3 \\ - \left((1-\alpha)(1-w_2) S^{-1/2} D_1 S^{1/2}\right) d_3 ,$$

if let $w_1 = w_2 = w$ and $D_3 = \alpha D_1 + (1-\alpha) D_2$; above identity equal to

$$a_3 = (1-w) S^{-1/2} \left(I_p - D_3\right) S^{1/2} d_3 .$$

To see that such a vector d_3 does in fact exist, let

$$H_1 = a(I_p - D_1) \quad \text{and} \quad H_2 = (1-a)(I_p - D_2) ,$$

where the both matrices $H_1 \in \Re_s^{\geq}$ and $H_2 \in \Re_s^{\geq}$. Then $I_p - D_3 = H_1 + H_2$ and the vector in question d_3 exists if and only if d_3 satisfied the equality

$$H_1 S^{1/2} d_1 + H_2 S^{1/2} d_2 = (H_1 + H_2) S^{1/2} d_3 .$$

But such a vector d_3 can always be found, since for any two symmetric n.n.d. matrices A and B the identity

$$C(A)+C(B)=C(A+B),$$

holds true. Now, since $H_1S^{1/2}d_1 \in C(H_1)$ and $H_2S^{1/2}d_2 \in C(H_2)$, there must exist some vector z with

$$H_1S^{1/2}d_1 + H_2S^{1/2}d_2 = (H_1 + H_2)z .$$

The choice $d_3 = S^{-1/2}z$ shows the required existence.

Note: if $w=0$, it will be reduced to the results of CC of estimator with QLF given in Groß [1].p235, which is the special case of above theorems.

We will give the Farebrother estimator as a further example of the application of a CC of a non-homogeneous and a HL estimator consider

$$\hat{\beta}_F = (1-\alpha)\hat{\beta}_R + \alpha\hat{\beta}$$

However, Farebrother's estimator is not really a CC estimator of OLS and RLS . Further research on it will be given next section.

4.2.3.5. Linear Bayes Estimator

As we know from the information about the Bayes estimator, assuming that the unknown parameter vector β is not fixed, but the realization of some random vector b for which a-priori information about its distribution is available, then Bayes methods suggest themselves for creating appropriate estimators for β .

Then, linear Bayes estimator

$$\hat{\beta}^{B} = \mu + \Phi\left[\Phi + \left(X'X\right)^{-1}\right]^{-1}\left(\hat{\beta} - \mu\right),$$

which can be applied under the assumption that β is the realization of a random vector b satisfying

$$E(b) = \mu \text{ and } Cov(b) = \sigma^2\Phi .$$

Here μ is a known $p\times 1$ vector and $\Phi \in \Re_s^{\geq}$ is a $p \times p$ matrix. It has already given in Groß (2003) that any Bayes estimator is linearly admissible for β under the QLF. Now, it is rather natural to wonder if the Bayes estimator is an ALE for β under the EBLF. For this, we can try to find some answer by the following theorem.

Theorem 4.14. Under the model (4.1) and EBLF, the estimator $Ly + a$ is a linear Bayes estimator for β if and only if the matrix L and the vector a can be written as

$$L = \left[w\left(X'X\right)^{-1} + (1-w)\left(X'X\right)^{-1/2} D\left(X'X\right)^{-1/2}\right]X' ,$$

and

$$a = (1-w)\left(X'X\right)^{-1/2}(I_p - D)\left(X'X\right)^{1/2} d ,$$

for some symmetric $p \times p$ matrix D such that $\lambda\left(D\right) \subset [0,1)$ and some $p\times 1$ vector d .

Proof: By letting

$$D = I_p - (1-w)^{-1} S^{-1/2} (\Phi + S^{-1})^{-1} S^{-1/2}, \quad S = X'X,$$

and noting

$$S^{-1/2}\left(wI_p + (1-w)D\right)S^{1/2} = I_p - S^{-1}\left(\Phi + S^{-1}\right)^{-1} = \Phi\left(\Phi + S^{-1}\right)^{-1},$$

the linear Bayes estimator can also be written as

$$\hat{\beta}^B = \mu + S^{-1/2}\left(wI_p + (1-w)D\right)S^{1/2}(\hat{\beta} - \mu).$$

It is obvious that the above matrix D is symmetric. Since Φ and S is n.n.d., matrix D is also n.n.d., and all eigenvalues of D are strictly smaller than 1, means D has all eigenvalues in the interval $[0,1)$.

Hence, with $d := \mu$ it follows that a linear Bayes estimator can always be written as

$$\hat{\beta}^B = \left(wI_p + (1-w)S^{-1/2}DS^{1/2}\right)\hat{\beta} + (1-w)S^{-1/2}(I_P - D)S^{1/2}d,$$

for some symmetric matrix D such that $\lambda(D) \subset [0,1)$. Then, Theorem 4.3, Theorem 4.4, and Equation (4.26) show that any linear Bayes estimator is linearly admissible for β.

On the other hand, let $Ly+a \in \beta$, where L and a can be written in the asserted form. Let $S = X'X$. Since D is a symmetric matrix with $\lambda(D) \subset [0,1)$, the matrix $(1-w)^{-1} S^{1/2}(I_P - D)S^{1/2}$ is symmetric p.d. Moreover, the matrix

$$0 < (1-w)^{-1} S^{1/2}(I_P - D)S^{1/2} \le S^{1/2}(I_P - D)S^{1/2} \le S,$$
$$\Phi = (1-w)^{-1}\left[S^{1/2}(I_P - D)S^{1/2}\right]^{-1} - S^{-1},$$

is symmetric n.n.d. For this matrix Φ the identity

$$D = I_P - (1-w)^{-1} S^{-1/2}\left(\Phi + S^{-1}\right)^{-1} S^{-1/2},$$

holds true, showing that the linear estimator $Ly+a$ can be written as

$$\begin{aligned}\hat{\beta}^B &= Ly + a \\ &= \left[wS^{-1} + (1-w)S^{-1/2} D S^{-1/2}\right] X'y + (1-w)S^{-1/2}(I_P - D)S^{1/2} d \\ &= S^{-1/2}\left(wI + (1-w)D\right) S^{1/2}\hat{\beta} + (1-w)S^{-1/2}(I_P - D)S^{1/2} d \\ &= d + \Phi\left[\Phi + \left(X'X\right)^{-1}\right]^{-1}\left(\hat{\beta} - d\right).\end{aligned}$$

But this is in fact a linear Bayes estimator for β when β is viewed as the realization of a random vector b satisfying

$$E(b) = d \text{ and } Cov(b) = \sigma^2\left((1-w)^{-1}\left[S^{1/2}(I_P - D)S^{1/2}\right]^{-1} - S^{-1}\right).$$

This concludes the proof.

On the basis of above theorem and Theorem 4.3, we can easily see that there are strong similarity between the set of linear Bayes estimators and the set of ALEs for β except one point. The only difference is that the symmetric matrix D can have all eigenvalues in $[0,1]$ when the estimator is linearly admissible, while it can only have all eigenvalues in $[0,1)$ when the estimator is linear Bayes.

If for example, we choose $D = I_p$ and $d = 0$, then the corresponding OLS estimator $\hat{\beta}$ is linearly admissible for β but cannot be a linear Bayes estimator.

Note: if $w = 0$, it will be reduced to the results of CC of estimator with QLF given in Groß (2003).p237, which is the special case of above theorems.

4.3. ALEs On Linear Functions of RC under EBLF

Consider the singular linear model

$$y = X\beta + \varepsilon, \tag{4.30.}$$

where y is an observable random vector with mean $X\beta$ and covariance $\sigma^2 V$, $V \geq 0$ is known and $\beta, \sigma^2, \varepsilon$ are same as given in model (4.1.), while $X \in \Re^{n\times p}$, here we assume that $rk(X) \neq p$, in this case, which means the RC β is not estimable. In this paper, we concern the ALEs of $G\beta$, where $G \in \Re^{m\times p}$ is known and $\mu(G') \subset \mu(X')$ which is to ensure that $G\beta$ is linear estimable.

For estimating an unknown parameter θ, Jozani et al (2006) introduced and motivated the use of the balanced-type loss function

$$L_{\omega,\delta_0}(\theta,\delta) = \omega q(\theta)(\delta - \delta_0)^2 + (1-\omega)q(\theta)(\delta - \theta)^2$$

where $0 \le \omega \le 1$, $q(\theta)$ is a positive weight function, and δ_0 is a general "target" estimator, and given the analysis and various examples with regards to the issues of admissibility, dominance, bayesianity, and minimaxity.

Based on the Jozani et al (2006), under the condition of singular model (4.30.) and ZBLF, Cao and He investigated the AOLE on linear functions of RC.

Inspired by the balanced-type loss function in Jozani et al. (2006) and Cao and He (2019), we here use the following EBLF

$$\begin{aligned} L(g;\beta) &= t_1(g - Gg_0)'(g - Gg_0) + t_2(g - G\beta)'S(g - G\beta) \\ &\quad + (1 - t_1 - t_2)(g - Gg_0)'(g - G\beta), \end{aligned} \tag{4.31.}$$

where $t_1, t_2 \in [0,1]$ and $S > 0$ are known, $g = g(y)$ is an estimator of $G\beta$ and $Gg_0 = Gg_0(y)$ is a "target" estimator of $G\beta$. Jozani et al. (2006) used such a balanced loss function, but in this research the author did not study admissibility of linear functions of unknown parameters. Just as what Jozani et al. (2006) said that loss function in loss (4.31.) depends on the observed value of target estimator, reflects a desire of closeness of g both to:

(i) $G\beta$ in terms of weighted squared error loss,

(ii) the target estimator Gg_0 in terms of squared distance.

The weight $0 \le t_1, t_2 \le 1$ calibrates the relative importance of these two criteria. Here, we select the target estimator Gg_0 in loss (4.31.) as the unified least square estimator of $G\beta$, namely $G\hat{\beta} = G(X'T^{-}X)^{-}X'T^{-}y$ where $T = V + XUX'$ with $U \ge 0$ such that $rk(T) = rk(X:V)$. Without loss of generality, we assume $U = I_p$ here.

Theorem 4.15. Let $\mathrm{L}_0^h = \left\{ L_0 X'T^- y : L_0 \in \Re^{m\times p} \right\}$. Then L_0^h is a complete class of L^h .

Proof: If $\beta_* = Ly \in \mathrm{L}^h$, using some properties of trace and Theorem 2.3, from loss (4.31), we get

$$\begin{aligned} R\left(Ly;\beta,\sigma^2\right) &= t_1 E\left[\left(Ly - G(X'T^-X)^- X'T^- y\right)'\left(Ly - G(X'T^-X)^- X'T^- y\right)\right] \\ &\quad + t_2 E\left[\left(Ly - G\beta\right)' S\left(Ly - G\beta\right)\right] \\ &\quad + (1-t_1-t_2)\left[\mathrm{E}\left(Ly - G(X'T^-X)^- X'T^- y\right)'\left(Ly - G\beta\right)\right] \\ &= \sigma^2 tr\left[LVL'B - 2wLVT^- X(X'T^-X)^- G'\right] \\ &\quad + \beta'\left(LX - G\right)' B\left(LX - G\right)\beta \\ &\quad + t_1\sigma^2 tr\left[G(X'T^-X)^- X'T^- VT^- X(X'T^-X)^- G'\right], \end{aligned} \tag{4.32.}$$

where $B = t_2 S + (1-t_2) I_m$, $w = (1+t_1-t_2)/2$.

On the other hand, $LP_X y = L_0 X'T^- y \in \mathrm{L}_0^h$ holds from $Ly \in \mathrm{L}^h$, where $P_{XT} = X\left(X'T^-X\right)^- X'T^-$ and $L_0 = LX\left(X'T^-X\right)^-$. So, using Theorem 2.4, we have

$$R\left(Ly;\beta,\sigma^2\right) - R\left(LP_X y;\beta,\sigma^2\right) = \sigma^2 tr\left[L\left(I_n - P_X\right)V\left(I_n - P_X\right)' L'B\right] \geq 0,$$

Moreover, the equality holds if and only if $LV = LP_{XT}V$. It is showed that L_0^h is a complete class of L^h .

As we know, EBLF was more flexible than QLF and could reduce to the QLF when $t_2 = 0$. The special relationship between them is also valid when the design matrix has not full column rank. It presented as Theorem 4.16.

Theorem 4.16. Given linear model

$$\begin{cases} z = X'T^{-}X\beta + \varepsilon; \\ E(\varepsilon) = 0, E(\varepsilon\varepsilon') = \sigma^2 X'T^{-}VT^{-}X, \end{cases} \tag{4.33.}$$

where z is a p dimensional observation vector, $\varepsilon \in \Re^n$ is a random error vector with mean 0 and covariance $\sigma^2 X'T^{-}VT^{-}X$. Let $C = (I_m - wB^{-1})G$ and $\tilde{\mathrm{L}} = \left\{ \hat{F}z : \hat{F} \in \Re^{m\times p} \right\}$, then $\hat{F}z \underset{\tilde{\mathrm{L}}}{\sqcup} C\beta$ holds under weighted QLF

$$LF_I(\beta_*, C\beta) = (d - G\beta)'B(d - G\beta),$$

if and only if $FX'T^{-}y \underset{\mathrm{L}_0}{\sqcup} G\beta$ holds under the model (4.30) and loss (4.31), where $\hat{F} = F - wB^{-1}G(X'T^{-}X)^{-}$.

Proof: Firstly, we get from (4.31)

$$\begin{aligned} R\left(FX'T^{-}y;\ \beta,\ \sigma^2\right) = {} & \sigma^2 tr\left[FX'T^{-}VT^{-}XF'B - 2wFX'T^{-}VT^{-}X(X'T^{-}X)^{-}G'\right] \\ & + t_1\sigma^2 tr\left[G(X'T^{-}X)^{-}X'T^{-}VT^{-}X(X'T^{-}X)^{-}G'\right] \\ & + \beta'\left(FX'T^{-}X - G\right)' B\left(FX'T^{-}X - G\right)\beta. \end{aligned} \tag{4.34.}$$

It is note that

$$E\left[(\hat{F}z - C\beta)'B(\hat{F}z - C\beta)\right] = \sigma^2 tr\left(FX'T^{-}VT^{-}XF'B\right)$$
$$- 2w\sigma^2 tr\left(FX'T^{-}VT^{-}X(X'T^{-}X)^{-}G'\right)$$
$$+ w^2\sigma^2 tr\left[G(X'T^{-}X)^{-}X'T^{-}VT^{-}X(X'T^{-}X)^{-}G'B^{-1}\right]$$
$$+ \beta'\left(FX'T^{-}X - G\right)' B\left(FX'T^{-}X - G\right)\beta. \tag{4.35.}$$

Thus, by Equation (4.34) and (4.35), it is easy to know that the conclusion of Theorem 4.16 holds.

Theorem 4.17. Under model (4.30) and loss (4.31), $Ly \in L^h G\beta$ holds if and only if following three propositions hold simultaneously:

(a) $LV = LX(X'T^{-}X)^{-}X'T^{-}V$;

(b) $\tilde{L}W_T\tilde{L}' \le \tilde{L}W_T G'(I_m - wB^{-1})$;

(c) $rk\{(LX - G)W_T X'\} = rk(LX - G)$,

where $\tilde{L} = LX - wB^{-1}G$, $W_T = (X'T^{-}X)^{-} - I_p$.

Proof: According to the Theorem 4.14, Theorem 4.15 and Theorem 2.14, $FX'T^{-}y \in G\beta$ if and only if

(a') $\hat{F}V_0 = \hat{F}X_0(X_0'W^{-}X_0)^{-}X_0'W^{-}V_0$;

(b') $\hat{F}X_0\left[(X_0'WX_0)^{-} - I_p\right]X_0'\hat{F}' \le \hat{F}X_0\left[(X_0'WX_0)^{-} - I_p\right]C'$;

(c') $rk\left\{(\hat{F}X_0 - C)\left[(X_0'W^{-}X_0)^{-} - I_p\right]X_0'\right\} = rk(FX_0 - C)$,

where $\hat{F} = F - wB^{-1}G(XT^{-}X)^{-}$, $X_0 = XT^{-}X$, $W = V_0 + X_0X_0'$, and $V_0 = XT^{-}VT^{-}X$.

By calculations, (a') is identical equation, (b') is equivalent to

$$\begin{aligned}&(FXT^{-}X - wB^{-1}G)W_T(FXT^{-}X - wB^{-1}G)'\\ &\leq (FXT^{-}X - wB^{-1}G)W_TG'(I_m - wB^{-1}),\end{aligned} \tag{4.36.}$$

which means that Equation (4.36) is equivalent to (b) . (c') is equivalent to

$$rk\left\{(FXT^{-}X - G)W_TXT^{-}X\right\} = rk(FXT^{-}X - G). \tag{4.37.}$$

$$\Leftrightarrow$$

$$rk\left\{(LX - G)W_TXT^{-}X\right\} = rk(LX - G). \tag{4.38.}$$

If we let $N = (LX - G)W_T$. Firstly, $rk(NXT^{-}X) \leq rk(NX')$ holds evidently. On the other hand, by Frobenius inequality

$$rk(NXT^{-}X) \geq rk(NX') + rk(XT^{-}X) - rk(X'),$$

and

$$rk(XT^{-}X) = rk(X'),$$

we have

$$rk(NXT^{-}X) \geq rk(NX').$$

Then

$$rk(NX'T^{-}X) = rk(NX').$$

Which means Equation (4.38) is equivalent to (c). Thus, this completes the proof of theorem.

Corollary 4.2. When model (4.30) is nonsingular ($V > 0$), from the Theorem 3.1 we can obtain that $Ly \in G\beta$ holds with respect to loss (4.31) if and only if

(a'') $L = LX(X'V^{-1}X)^{-}X'V^{-1}$;

(b'') $\tilde{L}(X'V^{-1}X)^{-}\tilde{L}' \leq \tilde{L}(X'V^{-1}X)^{-}G'(I_m - wB^{-1})$,

hold simultaneously.

Theorem 4.18. Under the model (4.30) and loss function (4.31), $Ly + a \in G\beta$ holds if and only if $Ly \in G\beta$ is admissible and $\alpha \in C(LX - G)$.

Proof: If a linear estimator $\tilde{\beta} = Ly + a \in \mathrm{L}$, then, using loss function (4.31), we get the risk of its with respect to the GEBLF as follows:

$$\begin{aligned} R(Ly + a; \beta, \sigma^2) &= \sigma^2 tr\left[LVL'B - 2t_1 LVT^{-}X(X'T^{-}X)^{-}G'\right] \\ &\quad + \left[(LX - G)\beta + a\right]' B\left[(LX - G)\beta + a\right] \\ &\quad + t_1\sigma^2 tr\left[G(X'T^{-}X)^{-}X'T^{-}VT^{-}X(X'T^{-}X)^{-}G'\right] \end{aligned} \qquad (4.39.)$$

where $B = t_2 S + (1 - t_2) I_m$. The proof is obtained in a similar way to the proof of Theorem 4.4.

4.4. Linear Admissibility Under Linear Restrictions

Consider the general LRM

$$y = X\beta + \varepsilon, \tag{4.40.}$$

where y is an observable random vector. $X \in \Re^{n \times p}$ is an model matrix of observations on p non-stochastic explanatory variables $\varepsilon \in \Re^n$ is error vector with mean $E(\varepsilon) = X\beta$ and covariance $\text{cov}(\varepsilon) = \sigma^2 V$, where $\beta \in \Re^p$ and σ^2 are unknown parameters, $V > 0$ is known. It is well known that the best linear unbiased estimators of β are shown as

$$\hat{\beta}_{GLS} = S_V^{-1} X' V^{-1} y ,$$

where $S_V^{-1} = X'V^{-1}X$. $\hat{\beta}_{GLS}$ is obtained using the least squares and maximum likelihood principle to drive the estimator of the unknown parameters. While using these principles, we considered that there is no a priori information about the parameters of the model, assuming that our level information is just sample information. In many models of practical interest, there are supplementary information in addition to the sample information of the matrix (y, X). We now focus only on precise preliminary information about the β coefficients.

In general, this previous information on the coefficients can be expressed as follows:

$$\begin{cases} y = X\beta + \varepsilon, \\ r = R\beta, \\ E(\varepsilon) = 0, E(\varepsilon\varepsilon') = \sigma^2 V, \end{cases} \tag{4.41.}$$

it is denote a restricted linear model obtained from model (4.40) by imposing on β consistent linear restrictions $r = R\beta$. It results in a reduction of the parameter space $\Theta = \Re^p \times (0,\infty)$ corresponding to the model (4.40) to $\Theta_r = \left\{R^+ r + N(R)\right\} \times (0,\infty)$ corresponding to the model (4.41). As we known, the unbiased estimator corresponding to the constrained model (4.41) is given as

$$\hat{\beta}_R = \hat{\beta}_{GLS} - S_V^{-1} R'(RS_V^{-1}R')^+ (R\hat{\beta}_{GLS} - r).$$

There are two types of equality restriction, one is $R\beta = r$, the other is stochastic restriction of type $r = R\beta + \phi$.

In this section, we examine the admissibility of the linear estimators for two types of equality constraints. First of all, in case 1, we will conclude by the following preliminary preparations that the restricted model can be transformed to the unrestricted model.

Case 1. When $R\beta = r$:

It is known that general solution to a consistent set of linear equations $R\beta = r$, with $r \in C(R)$, may be written in the form

$$\beta = R^+ r + (I - R^+ R)\eta. \tag{4.42.}$$

If we let

$$Q = I - R^{+}R,$$

Equation (4.42) rewriting as

$$\beta = R^{+}r + Q\eta,$$

where Q is any matrix such that $C(Q) = N(R)$. The model (4.41) is equivalently expressible as

$$\{y_r, XQ\eta, \sigma^2 V\}, \tag{4.43.}$$

where

$$y_r = y - XR^{+}r, \tag{4.44.}$$

η is a new vector of unknown parameters, and XQ is not full column rank, since

$$r(XQ) \le \min\{r(X), r(Q)\} \ne p.$$

Further, it can easily be verified with the use of (4.42) and (4.44) that, for any $Ly + a \in \mathrm{L}$,

$$\begin{aligned}\rho(Ly + a; \beta) &= \rho(Ly_r + LXR^{+}r + a;\ R^{+}r + Q\eta) \\ &= \rho(Ly_r + a - (I - LX)R^{+}r;\ Q\eta).\end{aligned}$$

Hence it follows that $Ly + a \overset{L}{\square} \beta$ holds under the model (4.41) if and only if $Ly_r + a - (I - LX)R^+ r \overset{L}{\square} Q\eta$ under the model (4.41). Consequently, Theorem 4.14 leads to the following.

Theorem 4.19. Under the model (4.41) and loss (4.31.), $Ly + a \overset{L}{\square} \beta$ if and only if

(i) $LV = LP_{XQT}V$;

(ii) $\tilde{L}QW_Q Q'\tilde{L}' \le \tilde{L}W_Q Q'(I_m - wB^{-1})$;

(iii) $rk\{(LX - I)QW_Q Q'X'\} = rk(LX - I)Q$;

$(iiii)$ $a - (I - LX)R^+ r \in C\left[\left(LX - I\right)Q\right]$,

where $\tilde{L} = LX - t_1 B^{-1}$, $W_Q = (Q'X'T^- XQ)^- - I_p$, $P_{XT} = X(X'T^- X)^- X'T^-$.

Proof: On account of Theorem 4.14, $Ly_r + a - (I - LX)R^+ r \overset{L}{\square} Q\eta$ under the model (4.30) and loss (4.31) if and only if

$$(a)\quad LV = LXQ(Q'X'T^- XQ)^- Q'X'T^- V \ ; \tag{4.45.}$$

$$(b)\quad \tilde{L}_1\left[(Q'X'T^- XQ)^- - I_p\right]\tilde{L}_1' \le \tilde{L}_1\left[(Q'X'T^- XQ)^- - I_p\right]Q'(I - wB^{-1}) \ ; \tag{4.46.}$$

$$(c)\quad rk\{(LX - I)Q\left[(Q'X'T^- XQ)^- - I_p\right]Q'X'\} = rk(LX - I)Q \ , \tag{4.47.}$$

and

$$a - (I - LX)R^+ r \in C\left[\left(LX - I\right)Q\right] ,$$

where $\tilde{L}_1 = (LX - t_1 B^{-1})Q$ and $Q = I - R^+ R$.

Since Q is symmetric and idempotent matrix satisfying the $C(Q)=N(R)$, we have according to the above, we can see that (4.45), (4.46), (4.48) are equivalent to the conditions $(i),(ii),(iii)$ in Theorem 4.19.

For the restricted model

$$\begin{cases} y = X\beta + \varepsilon, \\ R\beta = 0, \\ E(\varepsilon) = 0, E(\varepsilon\varepsilon') = \sigma^2 I_n, \end{cases} \tag{4.49.}$$

satisfied the following conditions

$$\mu(R') \subset \mu(X'), r(R) = m, r(X) \neq p,$$

and if we let

$$H = (X'X)^- - (X'X)^- R'\left(R(X'X)^- R'\right)^- R(X'X)^-,$$

then

$$HR' = 0,\ (XHX')^2 = XHX'(XHX')' = XHX'.$$

Using the Theorem 4.17, we will give the following theorem.

Theorem 4.20. Under the model (4.49.) and EBLF, if the assumptions $\mu(R') \subset \mu(X')$, $r(R) = m$, $r(X) \neq p$ are satisfied. Then sufficient and necessary condition of $Ly \in \beta$ is

(*a*) $\tilde{L} = \tilde{L}X'XH$,

(*b*) $\tilde{L}X'XHX'X\tilde{L}' \leq \tilde{L}X'XHC'$.

where ,

$$B = X'X,\ \tilde{L} = L_0 - w(X'X)^{-1}, C = (1-w)I_p,\ P_{XN_R} = XN_R\left(N_R'X'XN_R\right)^{-}N'X'.$$

Proof: For the model (4.49.), let $L_0X'y \in L_0^h$, risk function of under the EBLF is

$$\begin{aligned} R\left(L_0X'y;\ \beta, \sigma^2\right) &= \sigma^2 trX\tilde{L}'B\tilde{L}X' + \sigma^2 tr\left(t_1I_n - w^2X\left(X'X\right)^{-1}X'\right) \\ &\quad + \beta'\left[\tilde{L}X'X - C\right]' B\left[\tilde{L}X'X - C\right]\beta \\ &= \sigma^2 trX\tilde{L}'B\tilde{L}X' + \sigma^2 tr\left(t_1I_n - w^2X\left(X'X\right)^{-1}X'\right) \\ &\quad + \gamma'\left[\tilde{L}X'XN_R - CN_R\right]' B\left[\tilde{L}X'XN_R - CN_R\right]\gamma. \end{aligned}$$

$$B = X'X,\ \tilde{L} = L_0 - w(X'X)^{-1},\ C = (1-w)I_p,\ \beta = N_R\gamma,\ \gamma \in R^p.$$

Using Theorem 4.17, we know that, if we take β instant of $N_R\gamma$, $\gamma \in R^p$ in the model (4.49). Under the model (4.30)

$$L_0X'y \overset{L_0^h}{\in} \beta \quad \Leftrightarrow \quad \tilde{L}z \overset{\tilde{L}}{\in} CN_R\gamma.$$

Then, we get the necessary and sufficient condition $\tilde{L}z \overset{\tilde{L}}{\in} CN_R\gamma$ under the weighted QLF

$$(\beta_* - C\beta)' X'X(\beta_* - C\beta),$$

are as follows:

a) $\tilde{L} = \tilde{L}X'XN_R\left(N_R'X'XN_R\right)N_R'$,

b) $\tilde{L}X'P_{XN_R}X\tilde{L} \le \tilde{L}X'XN_R\left(N_R'X'XN_R\right)^- N_R'C$.

where,

$$B = X'X\ ,\ \tilde{L} = L_0 - w\left(X'X\right)^{-1}, C = (1-w)I_p,$$

$$P_{XN_R} = XN_R\left(N_R'X'XN_R\right)^- N_R'X'.$$

To get desired result we need to prove $XHX' = P_{XN_R}$. Since $HR' = 0$, we get

$$XHX'P_{XN_R} = XHN_R'X'P_{XN_R} = XHN_R'X' = XHX',$$

On the other hand

$$\begin{aligned} &XHX'P_{XN_R} \\ &= X\left[\left(X'X\right)^- - \left(X'X\right)^- R'\left(R\left(X'X\right)^- R'\right)^- R\left(X'X\right)^-\right] \\ &\times X'XN_R\left(N_R'X'XN_R\right)^- N_R'X'P_{XN_R} \\ &= P_{XN_R}. \end{aligned}$$

Then we have $XHX' = P_{XN_R}$. Then Theorem 4.20 is completed.

Corollary 4.3. In linear model (4.49.), if A_1 is the set of ALE of β in L under the QLF, and let A_1^* is the set of ALE of β under the EBLF. Then, A^* can be demonstrated as follows

$$A^* = \left\{ w\hat{\beta} + (1-w)\tilde{\beta}_{mse} : \tilde{\beta}_{mse} \in A \right\}$$
$$= \left\{ \hat{\beta} - (1-w)\left(X'X\right)^{-1/2} D\left(X'X\right)^{-1/2} X'y : \tilde{\beta}_{mse} \in A \right\}.$$

Proof: The proof is easily obtained according to the Theorem 4.3 and Corollary 4.1.

The Corollary 4.3. shows that the linear estimator under balanced loss function is a modification of the least squares' estimator. The correction term depends on a ALE of the RC in the equality restrictions model under the QLF. The second expression shows that the ALE under EBLF is a CC of least squares estimation and ALE under QLF, and the combination coefficient is a matrix.

Theorem 4.21. Under the model (4.49.) and EBLF, a linear estimator $Ly + a$ is β holds if and only if

a) $\tilde{L} = \tilde{L}X'XN_R\left(N_R'X'XN_R\right)N_R'$,

b) $\tilde{L}X'P_{XN_R}X\tilde{L} \le \tilde{L}X'XN_R\left(N_R'X'XN_R\right)^{-}N_R'C$,

c) $a \in \mu\left(\tilde{L}X'XN_R - CN_R\right)$.

hold simultaneously.

Proof: Details proof is similar to Theorem 4.4.

Example 4.5. Consider the LRM with assumptions $R\beta = r$. Show that a linear estimator $\tilde{\beta}$ is β if and only if

$$\tilde{\beta} = \left(wI_p + (1-w)S^{-1/2}DS^{1/2}\right)\hat{\beta} + (1-w)S^{-1/2}\left(I_p - D\right)S^{1/2}d,$$

where $S = X'X$, and D and d satisfy the conditions $D = D'$, $\sigma(D) \subset [0,1]$, $RS^{-1/2}D = 0$, and $wR\hat{\beta} + (1-w)Rd = r$.

Proof: Necessity

If a linear estimator $\tilde{\beta}$ is linearly admissible for β, using Theorem 4.3 and Theorem 4.4, $\tilde{\beta} = Ly + a$, L and a can be written as

$$\begin{aligned} L &= \left[w\left(X'X\right)^{-1} + (1-w)\left(X'X\right)^{-1/2} D\left(X'X\right)^{-1/2} \right] X' \\ &= wS^{-1}X' + (1-w)S^{-1}DS^{-1}X' \end{aligned},$$

$$a = \left(I_p - LX\right)d = (1-w)S^{-1/2}(I_p - D)S^{1/2}d\ .$$

Then

$$\begin{aligned} \tilde{\beta} &= Ly + a \\ &= \left(wI_p + (1-w)S^{-1/2}DS^{1/2}\right)\hat{\beta} + (1-w)S^{-1/2}\left(I_p - D\right)S^{1/2}d\ . \end{aligned}$$

From Corollary 4.1 and Theorem 4.5, we have $D = D'$, $\sigma(D) \subset [0,1]$.

Since, linear estimator satisfied the assumption $R\beta = r$.

$$\begin{aligned} R\tilde{\beta} &= R\left(wI_p + (1-w)S^{-1/2}DS^{1/2}\right)\hat{\beta} + (1-w)RS^{-1/2}\left(I_p - D\right)S^{1/2}d \\ &= wR\hat{\beta} + (1-w)RS^{-1/2}DS^{1/2}\hat{\beta} + (1-w)Rd - (1-w)RS^{-1/2}DS^{1/2}d \\ &= r. \end{aligned}$$

To be $R\tilde{\beta} = r$, satisfied the condition $RS^{-1/2}D = 0$, $wR\hat{\beta} + (1-w)Rd = r$.

Sufficiency.

If $\tilde{\beta} = \left(wI_p + (1-w)S^{-1/2}DS^{1/2}\right)\hat{\beta} + (1-w)S^{-1/2}\left(I_p - D\right)S^{1/2}d$, $D = D'$, $\sigma(D) \subset [0,1]$, $RS^{-1/2}D = 0$, $wR\hat{\beta} + (1-w)Rd = r$, where $S = X'X$.

Now we will proof $\tilde{\beta}$ linearly admissible for β .

$$\begin{aligned}\tilde{\beta} &= \left(wI_p + (1-w)S^{-1/2}DS^{1/2}\right)\hat{\beta} + (1-w)S^{-1/2}\left(I_p - D\right)S^{1/2}d \\ &= \left[wS^{-1} + (1-w)S^{-1/2}DS^{-1/2}\right]X'y + (1-w)S^{-1/2}\left(I_p - D\right)S^{1/2}d \\ &= Ly + a.\end{aligned}$$

where $L = \left[wS^{-1} + (1-w)S^{-1/2}DS^{-1/2}\right]X'$, $a = (1-w)S^{-1/2}\left(I_p - D\right)S^{1/2}d$.

Adding the condition $D = D'$, $\sigma(D) \subset [0,1]$ and from the $RS^{-1/2}D = 0$, $wR\hat{\beta} + (1-w)Rd = r$. Satisfied the $R\tilde{\beta} = r$. So using Theorem 4.5, we get $\tilde{\beta}$ is linearly admissible for β .

Case 2. When $r = R\beta + \phi$;

Durbin (1953) firstly to use sample and auxiliary information simultaneously, by developing a stepwise estimator for parameters. Then, Theil and Goldberger (1961) and Theil (1963) introduced the mixed estimation technique by combining the sample and prior knowledge in a common model

$$\begin{pmatrix} y \\ r \end{pmatrix} = \begin{pmatrix} X \\ R \end{pmatrix}\beta + \begin{pmatrix} \varepsilon \\ \phi \end{pmatrix}, \tag{4.50.}$$

where $r \in \Re^m$, $R \in \Re^{m\times p}$ with $rk(R) = m$, random error $\phi \sim (0, \sigma^2 W)$ and essential assumption is $E(\varepsilon\phi') = 0$. Then matrix of variance-covariance becomes

$$E\begin{pmatrix} \varepsilon \\ \phi \end{pmatrix}\begin{pmatrix} \varepsilon & \phi \end{pmatrix}' = \sigma^2\begin{pmatrix} V_1 & 0 \\ 0 & V_2 \end{pmatrix}.$$

If we let $\tilde{y}=\begin{pmatrix} y \\ r \end{pmatrix}$, $\tilde{X}=\begin{pmatrix} X \\ R \end{pmatrix}$, $\tilde{\varepsilon}=\begin{pmatrix} \varepsilon \\ \phi \end{pmatrix}$, Equation (4.50.) can be rewriting as

$$\tilde{y}=\tilde{X}\beta+\tilde{\varepsilon}, \qquad \tilde{\varepsilon} \sim (0, \sigma^2\tilde{V}). \tag{4.51.}$$

where $\tilde{V}>0$.

Theorem 4.22. An estimator $L\tilde{y}+a \overset{\text{L}}{\sim} G\beta$ holds under the stocastic restricted model (4.51.) if and only if

(*a*) $a \in C\left(L_1X+L_2R-G\right)=C\left(L\tilde{X}-G\right)$

(*b*) $L\tilde{y} \overset{\text{L}^h}{\sim} G\beta$;

Proof: Necessity.

(a) If suppose that $a \notin C\left(L\tilde{X}-G\right)$, write $a=a_1+a_2$, where $a_1 \in C\left(L\tilde{X}-G\right)$, $0 \neq a_2 \in C^{\perp}\left(L\tilde{X}-G\right)$, hence

$$a'a=\begin{pmatrix} a_1' & a_2' \end{pmatrix}\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}=a_1'a_1+a_2'a_2 \geq a_1'a_1. \tag{4.52.}$$

If we consider the decomposition of $L=\begin{pmatrix} L_1 & L_2 \end{pmatrix}$ such as $L\tilde{y}=L_1y+L_2r$, we have

$$\begin{aligned} R\left(L\tilde{y}+a;\beta,\sigma^2\right)=&E\left[t_1\left(\tilde{y}-\tilde{X}(L\tilde{y}+a)\right)'\left(\tilde{y}-\tilde{X}(L\tilde{y}+a)\right)\right] \\ &+E\left[t_2\left((L\tilde{y}+a)-\beta\right)'\tilde{X}'\tilde{X}\left((L\tilde{y}+a)-\beta\right)\right] \end{aligned}$$

$$
\begin{aligned}
&+\left[(1-t_1-t_2)\mathrm{E}\left(\tilde{X}(L\tilde{y}+a)-\tilde{y}\right)'\tilde{X}\left((L\tilde{y}+a)-\beta\right)\right]\\
&=tr\sigma_1^2L_1V_1L_1'B+tr\sigma_2^2L_2V_2L_2'B\\
&-2wtr\left(\sigma_1^2L_1V_1T^-X-\sigma_2^2L_2V_2T^-R\right)\left(X'T^-X+R'T^-R\right)^-G'\\
&+\left[\left(L_1X+L_2R-G\right)\beta+a\right]'B\left[\left(L_1X+L_2R-G\right)\beta+a\right]\\
&+t_1tr\left(\sigma_1^2G\left(X'T^-X+R'T^-R\right)^-X'T^-V_1T^-X\left(X'T^-X\right.\right.\\
&\left.\left.+R'T^-R\right)^-G'\right)+t_1tr\left(\sigma_2^2G\left(X'T^-X+R'T^-R\right)^-\right.\\
&\left.\times R'T^-V_2T^-R\left(X'T^-X+R'T^-R\right)^-G'\right)
\end{aligned}
\tag{4.53.}
$$

for any (β,σ^2), from the above we can get

$$
\begin{aligned}
&R\left(L\tilde{y}+a;\,\beta,\,\sigma^2\right)-R\left(L\tilde{y}+a_1;\,\beta,\,\sigma^2\right)\\
&\qquad=2\beta'\left(L_1X+L_2R-G\right)'B(a-a_1)+a'Ba-a_1'Ba_1\\
&\qquad>0.
\end{aligned}
$$

this means $L\tilde{y}+a_1$ is better than $L\tilde{y}+a$, which contradicts the assumption.

(b) We prove that $L\tilde{y}\in G\beta$, when $L\tilde{y}+a\overset{\mathrm{L}}{\in}G\beta$ holds.

Suppose, by contradiction, If $L\tilde{y}\notin G\beta$, there exists an estimator $K\tilde{y}\in\mathrm{L}^h$ such that

$$
R\left(K\tilde{y};\,\beta,\,\sigma^2\right)\le R\left(L\tilde{y};\,\beta,\,\sigma^2\right), \tag{4.54.}
$$

is right for all (β,σ^2) with the strict inequality holding at some point (β_1,σ_1^2). As $\alpha\in C\left(L\tilde{X}-G\right)$, there exists $a_0\in\Re^p$ such that $a=\left(L\tilde{X}-G\right)a_0$.

So, by (4.54.),

$$R\left(L\tilde{y}+a;\,\beta,\,\sigma^2\right)=R\left(L\tilde{y};\,\beta+a_0,\,\sigma^2\right)$$
$$\geq R\left(K\tilde{y};\,\beta+a_0,\,\sigma^2\right)=R\left(K\tilde{y}+(K\tilde{X}-I_p)a_0;\,\beta,\,\sigma^2\right),$$

holds for all (β,σ^2) and the strict inequality is true at point $\left(\beta_1-a_0,\sigma_1^2\right)$, which means that $K\tilde{y}+(K\tilde{X}-G)a_0$ is better than $L\tilde{y}+a$, this is contradictory to the admissibility of $L\tilde{y}+a$.

Sufficiency.

Suppose that there exists an estimator $K\tilde{y}+k$ is better than $L\tilde{y}+a$, means

$$R\left(K\tilde{y}+k;\,\beta,\,\sigma^2\right)\leq R\left(L\tilde{y}+a;\,\beta,\,\sigma^2\right).$$

Since proof of necessity, it can be assumed that $a=\left(L\tilde{X}-G\right)a_0$, $k=(K\tilde{X}-I_p)a_0$ substituting into the above formula, we get

$$R\left(K\tilde{y};\,\beta+a_0,\,\sigma^2\right)=R\left(K\tilde{y}+a_0;\,\beta,\,\sigma^2\right)$$
$$\leq R\left(L\tilde{y}+a_0;\,\beta,\,\sigma^2\right)=R\left(L\tilde{y};\,\beta+a_0,\,\sigma^2\right),$$

which means

$$R\left(K\tilde{y};\,\beta_0,\,\sigma^2\right)\leq R\left(L\tilde{y};\,\beta_0,\,\sigma^2\right),$$

for all $\beta_0\in\Re^p$ and $\sigma^2>0$, where $\beta_0=\beta+a_0$. Which means that $K\tilde{y}$ is better than $L\tilde{y}$, this is contradicting $L\tilde{y}\square\ \beta$. Hence, the sufficiency is proved.

4.5. Linear Admissibility Under Elliptical Restrictions

As we knew from previous information, parameter space is one of the important factors determining the admissibility of estimators. In practical

applications, we often know only that the regression parameter lies most probably in a certain bounded subset Θ such as equality and inequality restriction.

We define the ellipsoid

$$\Theta = \Theta_\beta = \left\{ \beta \in \Re^p : \beta' T \beta \le \sigma^2 \right\},$$

as parameter set, unknown parameters β and σ^2 satisfy the relationship $\beta' T \beta \le \sigma^2$ for some matrix $T \in \Re^{>}$.

Note that

$$\left(\beta - \beta_0\right)' T \left(\beta - \beta_0\right) \le \sigma^2,$$

is called incomplete non-central ellipsoidal restriction, where β_0 center of ellipse.

Theorem 4.23. Under the model (4.30) and EBLF, $Ly \overset{L^h}{\sqsubset} \beta[\Theta]$ holds if and only if L can be written as

$$L = \left[wS^{-1} + (1-w) S^{-1/2} D S^{-1/2} \right] X',$$

for some symmetric matrix $D \in \Re^{p \times p}$ satisfying the following conditions

$$tr\left\{ \left(I + D\right)^{-1} - I \right\} S^{-1/2} T S^{-1/2} \le I + trS^{-1/2} T S^{-1/2},$$

and

$$\lambda(D) \subset [0,1],$$

where $S = X'X$ and $0 \le w \le 1$.

Proof: From Theorem 4.2 and Theorem 2.19, we have

$$C^{-1}\tilde{L}Z \overset{mse}{\in} \beta[\Theta] \Leftrightarrow C^{-1}\tilde{L}X'y \overset{mse}{\in} \beta[\Theta] \Leftrightarrow Ly \overset{EBLF}{\in} \beta[\Theta],$$

then using Theorem 4.5 (Groß , 2003) and Theorem 2.18, we get

$$\begin{aligned} & B = C^{-1}\tilde{L}X' = S^{-1/2}DS^{-1/2}X' \\ \Leftrightarrow\ & \left(L_0 - wS^{-1}\right)X' = (1-w)S^{-1/2}DS^{-1/2}X' , \end{aligned}$$

Thus, we can easily get the

$$L = L_0X' = \left[wS^{-1} + (1-w)S^{-1/2}DS^{-1/2}\right]X' .$$

Since $P = C^{-1}\tilde{L} = L_0 - wS^{-1}$ in Theorem 2.18, one can easily get

$$\begin{aligned} & trST\left\{\left[I - (1-w)^{-1}\left(L_0 - wS^{-1}\right)S\right]^{-1} - I\right\} \le 1 \\ \Leftrightarrow\ & trS^{-1}T\left\{\left[I + S^{-1/2}DS^{1/2}\right]^{-1} - I\right\} \le 1 \\ \Leftrightarrow\ & tr\left[\left(I + D\right)^{-1} S^{-1/2}TS^{-1/2}\right] \le 1 + trS^{-1/2}TS^{-1/2} . \end{aligned}$$

Next step, we try to answer of how was implict characterizations.

Theorem 4.24. Under the linear model and EBLF, $Ly \overset{L^h}{\sqsubset} \beta[\Theta]$ holds if and only if the conditions

$$XL = L'X' \text{ , } \lambda(XL) \subset [0,1] \text{ ,}$$

and

$$tr\left[(1-w)^{-1} \left(LX + (1-2w)I \right)^{-1} S^{-1}T \right] \leq 1 + tr\left[S^{-1}T \right] ,$$

where $S = X'X$ and $0 \leq w \leq 1$.

Proof: We can obtain $XL = L'X'$, $\lambda(XL) \subset [0,1]$ using the same method as Theorem 4.5, where $L = \left[wS^{-1} + (1-w)S^{-1/2}DS^{-1/2} \right] X'$ for some matrix $D' = D, \lambda(D) \subset [0,1]$. For this matrix, the identity

$$LX + (1-2w)I_p = (1-w)S^{-1/2}\left[I + D \right] S^{1/2}$$

holds true, where $S = X'X$. Then, the desired results is obtained from the above Theorem 4.23.

5. AOLE IN THE GGM MODEL UNDER GEBLF

In this chapter, we proposes a new loss function GEBLF. AOLE is characterized in the GGM model with respect to GEBLF.

5.1. A New Loss Function

Consider the following GGM model

$$y = X\beta + \varepsilon, \;\; E(\varepsilon) = 0, \;\; E(\varepsilon\varepsilon') = \sigma^2 V \tag{5.1.}$$

where y is an n- dimensional vector of observations on the dependent variables such that $E(y) = X\beta$, $\text{var}(y) = \sigma^2 V$, V is a known n.n.d matrix ($V \geq 0$), $X, \beta, \sigma^2, \varepsilon$ are same as given in model (4.1.).

According to Rao's unified theory of least squares, the GLS estimator of β parameter in model (5.1.) is given by

$$\hat{\beta}_{GLSE} = S_{T_*}^{-} X' T_*^{+} y \,, \tag{5.2.}$$

where $S_{T_*} = X' T_*^{+} X$, $T_* = V + XX'$.

Combining Shalabh's EBLF and Rao's unified theory and following Cao (2014), Cao and He (2019) and Liu and Yin (2020), we propose the GEBLF as follows for GGM model

$$\begin{aligned} GEBLF(\beta_*, \beta, \lambda_1, \lambda_2, \lambda_3) = {} & \lambda_1 \left(y - X\beta_*\right)' T^{+} \left(y - X\beta_*\right) \\ & + \lambda_2 (\beta_* - \beta)' \tilde{S} (\beta_* - \beta) \\ & + \lambda_3 (X\beta_* - y)' T^{+} X (\beta_* - \beta), \end{aligned} \tag{5.3.}$$

where $0 \le \lambda_i \le 1$, $\sum_i \lambda_i = 1, i = 1,2,3$. The matrix $\tilde{S} > 0$ is known, β_* is an estimator of β and $T = V + XUX'$ with $rk(T) = rk(V : X)$ for arbitrary symmetric matrix $U \ge 0$. The corresponding risk function is defined by

$$R(\beta_*, \beta, \lambda_1, \lambda_2, \lambda_3) = E(GEBLF(\beta_*, \beta, \lambda_1, \lambda_2, \lambda_3)).$$

The main advantage of the GEBLF is that it is more general and more flexible in comparison to other criterion. GEBLF encompass the following particular cases:

(i) When $\lambda_1 = 1$, $\lambda_2 = 0$ and $\lambda_3 = 0$, the GEBLF in Equation (5.3.) reduces to the weighted square error loss

$$\left(y - X\beta_*\right)' T^+ \left(y - X\beta_*\right),$$

which is the criterion of goodness of fit of model . If $T^+ = I$, it turns to square error loss (SEL).

(ii) If we put $\lambda_1 = 0$, $\lambda_2 = 1$ and $\lambda_3 = 0$, the GEBLF in Equation (5.3.) reduces to the weighted QLF

$$(\beta_* - \beta)' \tilde{S} (\beta_* - \beta),$$

which is criterion of precision of estimation. When $\tilde{S} = I$, it turns to the unweighted QLF.

(iii) If $\lambda_1 = 0, \lambda_2 = 0$ and $\lambda_3 = 1$, the GEBLF in Equation (5.3.) reduces to the component $(X\beta_* - y)' T^+ X(\beta_* - \beta)$ which is the criterion of

interaction or covariation between the precision of estimation and goodness of fit.

(iv) When we take $\lambda_1 = \lambda$, $\lambda_2 = 1-\lambda$ and $\lambda_3 = 0$, we obtain the GBLF. If $\tilde{S} = X'X$, GBLF turns to ZBLF.

(v) Let $T^+ = I, \tilde{S} = X'X$, then GEBLF turns to EBLF.

(vi) If we take $\tilde{S} = C\text{ov}^{-1}(\beta)$ in (ii), GEBLF reduces to Fisherian Loss function (FLF).

(vii) When $\tilde{S} = Var^{-1}(\beta_*)$ in (ii), GEBLF turns to Mahalanobis Loss function (MLF).

5.2. Admissibility under the GEBLF

In this section, we will investigate the SNC for linear estimators to be admissible are obtained respectively, when $V \geq 0$ and $V > 0$ among the F and F ^{h} classes under the GGM model.

Since EBLF is the special case of GEBLF, there exists some relationship between the GEBLF and the quadratic loss. It is presented in Theorem 5.1 by creating the matrix $T = V + XUX'$ with $rk(T) = rk(V \vdots X)$, where $U \geq 0$ is any symmetric matrix.

Theorem 5.1. In GGM model, if $C\beta$ is estimable, then under the loss function $LF(F^*y, C\beta)$, $F^*y \sqsubset C\beta$ holds if and only if $Fy \sqsubset \beta$ holds under the GEBLF, where $Fy \in \text{F}^h$, $\lambda_1, \lambda_2 \in [0,1]$, $w = (1+\lambda_1-\lambda_2)/2$, $0 \leq w \leq 1$, $F^* = B^{1/2}\left(F - wB^{-1}X'T^+\right)$, $S_T = X'T^+X$, $C = B^{1/2}(I_p - wB^{-1}S_T)$, $T = V + XUX'$, $B = \lambda_2\tilde{S} + (1-\lambda_2)S_T$.

Proof: If $\beta_* = Fy \in \mathrm{F}^h$, using some properties of trace and Theorem 2.3, then risk function of Fy with respect to the GEBLF

$$\begin{aligned} R\left(Fy;\beta,\sigma^2\right) &= E\left[\lambda_1\left(y-XFy\right)'T^+\left(y-XFy\right)+\lambda_2\left(Fy-\beta\right)'\tilde{S}\left(Fy-\beta\right)\right] \\ &+\left[(1-\lambda_1-\lambda_2)\mathrm{E}\left(XFy-y\right)'T^+X\left(Fy-\beta\right)\right] \\ &= \sigma^2\,\mathrm{tr}(F^*VF^{*\prime})+\sigma^2 tr(\lambda_1 T^+V-w^2T^+XB^{-1}X'T^+) \\ &+\beta'\left(F^*X-C\right)'\left(F^*X-C\right)\beta \end{aligned} \tag{5.4.}$$

where $w=\left(1+\lambda_1-\lambda_2\right)/2$, $0\le w\le 1$, $T=V+XUX'$, $C=B^{1/2}(I_p-wB^{-1}S_T)$ and $F^*=B^{1/2}\tilde{F}=B^{1/2}\left(F-wB^{-1}X'T^+\right)$.

It is note that risk function of estimator F^*y under the quadratic loss $LF(F^*y,C\beta)$

$$E\left[LF(F^*y,C\beta)\right]=\sigma^2\,\mathrm{tr}(F^*VF^{*\prime})+\beta'\left(F^*X-C\right)'\left(F^*X-C\right)\beta, \tag{5.5.}$$

Let $Fy \sqsubset \beta$ under the GEBLF. In this case, for an arbitrary $My \in \mathrm{F}^h$, we have

$$R\left(Fy;\beta,\sigma^2\right)\le R\left(My;\beta,\sigma^2\right) \tag{5.6.}$$

for all $\beta\in\Re^p$ and $\sigma^2>0$.

From (5.4.), (5.5.), we get that (5.6.) holds if and only if

$$E\Big[LF(F^*y, C\beta)\Big] \le E\Big[LF(M^*y, C\beta)\Big], \tag{5.7.}$$

holds for all $\beta \in \Re^p$ and $\sigma^2 > 0$, where $M^* = B^{1/2}\tilde{M} = B^{1/2}\left(M - wB^{-1}XT^+\right)$, C and F^* is same as above mentioned. Thus, based on the (5.6.), (5.7.) and definition of admissibility, (5.6.) holds if and only if $F^*y \sqsubset C\beta$. The proof of Theorem 5.1 is completed.

Corollary 5.1. In GGM model and the GEBLF. Then $Fy \sqsubset \beta$ holds if and only if the following statements are valid:

(i) $FV = FP_{XT}V$,

(ii) $\tilde{F}XWX'\tilde{F}' \le \tilde{F}XWC'B^{-1/2}$, $\tilde{F}XWC'B^{-1/2}$ *is symmetric,*

(iii) $rk\Big[B^{1/2}(FX - I)WX'\Big] = rk\Big[B^{1/2}(FX - I)\Big]$,

where $P_{XT} = XS_{T_*}^- X'T_*^+$, $B = \lambda_2\tilde{S} + (1-\lambda_2)S_{T_*}$ and $\tilde{F} = F - wB^{-1}XT_*^+$, $W = S_{T_*}^- - I$, $T_* = V + XX'$, $S_{T_*} = X'T_*^+X$, $C = B^{1/2}(I_p - wB^{-1}S_{T_*})$, $w = \left(1 + \lambda_1 - \lambda_2\right)/2$, $0 \le w \le 1$.

Proof From Theorem 5.1, we have

$$F^*y \sqsubset C\beta \Leftrightarrow Fy \sqsubset \beta,$$

where $F^* = B^{1/2}\left(F - wB^{-1}XT^+\right)$ and $C = B^{1/2}(I_p - wB^{-1}S_T)$. Using Theorem 2.16(a), by simple computation, we get

$$F^*V = F^* P_{XT}V \Leftrightarrow FV = FP_{XT}V .$$

Since $C = B^{1/2}(I_p - wB^{-1}S_T)$, using Theorem 5.1 and Theorem 2.16 (b), we obtain

$$F^* XWC' - F^* XWX'F^{*\prime} \geq 0 \Leftrightarrow \tilde{F}XWX'\tilde{F}' \leq \tilde{F}XWC'B^{-1/2}. \qquad (5.8.)$$

Equation (5.7.) is equivalent to (ii). From Theorem 2.16 (c), we get

$$rk\{(F^* X - C)WX'\} = rk(F^* X - C)$$

$$\Leftrightarrow$$

$$rk\left[B^{1/2}(FX - I)WX'\right] = rk\left[B^{1/2}(FX - I)\right],$$

which shows (iii). Thus, the Corollary 5.1 is proved.

Corollary 5.2. If $V > 0$, then the following statements hold for (i) and (ii) in Corollary 5.1.

$(i)'$ $FV = FP_{XV}V,$

$(ii)'$ $\tilde{F}XS_V^- X'\tilde{F}' \leq \tilde{F}XS_V^-\left(I_p - wB^{-1}(I_p + S_V)^{-1}S_V\right)',$

$\tilde{F}XS_V^-\left(I_p - wB^{-1}(I_p + S_V)^{-1}S_V\right)'$ is symmetric,

where $\tilde{F} = F - wB^{-1}(I_p + S_V)^{-1}X'V^{-1}$, $B = \lambda_2\tilde{S} + (1-\lambda_2)(I_p + S_V)^{-1}S_V$,

$S_V = X'V^{-1}X$, $P_{XV} = XS_V^- X'V^{-1}$.

Proof: Since $T_* = V + XX'$, we have

$$\left(V + XX'\right)^{-1} = V^{-1} - V^{-1}X\left(I + S_V\right)^{-1}X'V^{-1}$$
$$S_V\left(I + S_V\right)^{-1} = I - \left(I + S_V\right)^{-1},$$

then

$$P_{XT} = P_{XV}. \tag{5.9.}$$

Using Corollary 5.1 (i) and Equation (5.9.),

$$FV = FP_{XT}V \Leftrightarrow FV = FP_{XV}V,$$

which shows $(i)'$. From Equation (5.9.), we get

$$XWX' = X\left[S_T^- - I\right]X' = XS_V^-X'V^{-1}\left(V + XX'\right) - XX'$$
$$= XS_V^-X' + XX' - XX' = XS_V^-X'.$$

and

$$C = B^{1/2}(I_p - wB^{-1}S_T) = B^{1/2}\left(I_p - wB^{-1}(I_p + S_V)^{-1}S_V\right)'.$$

So,

$$\tilde{F}XWX'\tilde{F}' \le \tilde{F}XWC'B^{-1/2},$$
$$\Leftrightarrow$$

$$\tilde{F}XS_V^- X'\tilde{F}' \le \tilde{F}XS_V^- \left(I_p - wB^{-1}(I_p + S_V)^{-1}S_V\right)',$$

which shows $(ii)'$.

Corollary 5.3. For the model $y = X\beta + \varepsilon,\ E(\varepsilon) = 0,\ E(\varepsilon\varepsilon') = \sigma^2 I_n$, if $C\beta$ is estimable, then $F^* y \underline{\in} C\beta$ holds under the loss function $LF(F^* y, C\beta)$ if and only if $Fy \underline{\in} \beta$ holds under the EBLF, where $Fy \in \mathrm{F}^h$, $F^* = S^{1/2}\left(F - wS^{-1}X'\right)$, $C = (1-w)S^{1/2}$ and $S = X'X$.

Proof: If $U = 0$, $\tilde{S} = S = X'X$, then $B = S$. From Theorem 5.1, since $F^* = B^{1/2}\left(F - wB^{-1}X'T^+\right)$, we can get $F^* = S^{1/2}\left(F - wS^{-1}X'\right)$ and $C = (1-w)S^{1/2}$.

Theorem 5.2. Consider the model $y = X\beta + \varepsilon,\ E(\varepsilon) = 0,\ E(\varepsilon\varepsilon') = \sigma^2 I_n$ and the EBLF. Then $Fy \underline{\in} \beta$ holds if and only if

(i) $F = FP_X$,

(ii) $(1-w)S^{1/2}\tilde{F}XS^{-1/2}$ is symmetric and $S^{1/2}\tilde{F}P_X\tilde{F}'S^{1/2} \le (1-w)S^{1/2}\tilde{F}XS^{-1/2}$.

hold simultaneously, where $P_X = XS^{-1}X'$, $\tilde{F} = F - wS^{-1}X'$, $S = X'X$.

Proof: If $Fy \underline{\in} \beta$, using Corollary 5.3, we know

$$Fy \underline{\in} \beta \Leftrightarrow F^* y \underline{\in} C\beta$$

where $Fy \in \mathrm{F}^h$, $F^* y \in \mathrm{F}^h$, $C = (1-w)S^{1/2}$, $F^* = S^{1/2}\left(F - wS^{-1}X'\right)$, $S = X'X$.

Further, if we consider the special case of Theorem 2.3 when $k = p$ and $C = (1-w)S^{1/2}$, then F^* turns to F^h and we get the sufficient and necessary condition for $F^* y \sqsubset C\beta$ under the loss function $LF(\beta_*, C\beta)$ are as follows:

(a) $F^* = F^* P_X$

(b) $F^* P_X F^{*\prime} \le (1-w)F^* XS^{-1/2}$, where $P_X = XS^{-1}X'$ and $(1-w)F^* XS^{-1/2}$ is symmetric.

From (a),(b) and $F^* = S^{1/2}\left(F - wS^{-1}X'\right)$, we can easily get

$$F^* = F^* P_X \Leftrightarrow F = FP_X$$

and

$$F^* P_X F^{*\prime} \le (1-w)F^* XS^{-1/2}$$

$$\Leftrightarrow$$

$$S^{1/2}\tilde{F}P_X\tilde{F}'S^{1/2} \le (1-w)S^{1/2}\tilde{F}XS^{-1/2}.$$

Since $(1-w)F^* XS^{-1/2}$ is symmetric, then $(1-w)S^{1/2}\tilde{F}XS^{-1/2}$ is also is symmetric. Which means (i) and (ii) are satisfied. Thus, the proof is completed.

Corollary 5.4 Consider the model (4.1.) and the EBLF, $Fy \sqsubset \beta$ holds if and only if F can be written as

$$F = \left[wS^{-1} + (1-w)S^{-1/2}DS^{-1/2} \right] X',$$

where $D = S^{1/2}FXS^{-1/2}$ is symmetric and all eigenvalues in [0,1] and $w = \left(1+\lambda_1-\lambda_2\right)/2, \lambda_1, \lambda_2 \in [0,1], 0 \le w \le 1$.

Proof*:* If we take $D = (1-w)^{-1}S^{1/2}\tilde{F}XS^{-1/2}$, using Theorem 5.2 (ii), we get

$$S^{1/2}\tilde{F}P_X\tilde{F}'S^{1/2} \le (1-w)S^{1/2}\tilde{F}XS^{-1/2} \;\Leftrightarrow\; DD' \le D$$

where D is symmetric and all eigenvalues in [0,1]. Since Theorem 5.2 (i), $F = FP_X$ and $\tilde{F} = F - wS^{-1}X'$, we have $\tilde{F} = \tilde{F}P_X$, then

$$\tilde{F} = \tilde{F}X\left(X'X\right)^{-1}X' = S^{-1/2}DS^{-1/2}X'.$$

From which we can get the desired results

$$F = wS^{-1}X' + (1-w)S^{-1/2}DS^{-1/2}X'.$$

This is same as the results in Mirezi and Kaçıranlar (2021).

Theorem 5.3. Under the GGM model (5.1) and GEBLF, $Fy + f \sqsubset \beta$ holds if and only if

(a) $Fy \sqsubset \beta$;

(b) $f \in C\left(FX - I_p\right)$ hold simultaneously.

Proof: Sufficiency.

If a linear estimator $\beta_* = Fy + f \in \mathrm{F}$, then, using Equation (5.3), we get the risk of its with respect to the GEBLF as follows:

$$R\left(Fy+f\,;\,\beta,\sigma^2\right)=\sigma^2 tr\left[\lambda_1\left(I_n-XF\right)'T^{+}\left(I_n-XF\right)+\lambda_2\left(F'SF\right)\right]V$$
$$+\left(1-\lambda_1-\lambda_2\right)\sigma^2 tr\left(\left(XF-I_n\right)'T^{+}XF\right)V$$
$$+\left[\left(FX-I_p\right)\beta+f\right]'B\left[\left(FX-I_p\right)\beta+f\right], \qquad (5.10.)$$

where $B=\lambda_2\tilde{S}+(1-\lambda_2)S_T$.

As $f\in C\left(FX-I_p\right)$, there exists $f_0\in\Re^p$ such that $f=\left(FX-I_p\right)f_0$. Hence, using (5.10.) we get

$$R\left(Fy+f\,;\,\beta,\sigma^2\right)=R\left(Fy;\,\beta+f_0,\sigma^2\right). \qquad (5.11.)$$

Suppose $Fy+f$ ▯ β, then there exists an estimator $Qy+q$ such that

$$R\left(Qy+q;\,\beta,\sigma^2\right)\le R\left(Fy+f\,;\,\beta,\sigma^2\right), \qquad (5.12.)$$

holds for all $\beta\in\Re^p$ and $\sigma^2>0$.

Let $\beta=-f_0$ in (5.12.), then by (5.11.) and (5.10.)

$$\begin{aligned}\left[q-\left(QX-I_p\right)f_0\right]' B\left[q-\left(QX-I_p\right)f_0\right] &= \lim_{\sigma^2\to 0^+} R\left(Qy+q;-f_0,\sigma^2\right)\\ &\le \lim_{\sigma^2\to 0^+} R\left(Fy+f;\beta,\sigma^2\right) \quad (5.13.)\\ &= \lim_{\sigma^2\to 0^+} R\left(Fy;0,\sigma^2\right).\end{aligned}$$

Since $B>0$ and the right side of inequality (5.13.) is approaching zero. Therefore,

$$q=\left(QX-I_p\right)f_0. \quad (5.14.)$$

From (5.11.), (5.12.)and (5.14.), we get

$$\begin{aligned}R\left(Qy;\beta+f_0,\sigma^2\right) &= R\left(Qy+f_0;\beta,\sigma^2\right)\\ &\le R\left(Fy+f_0;\beta,\sigma^2\right)=R\left(Fy;\beta+f_0,\sigma^2\right),\end{aligned}$$

which means

$$R\left(Qy;\beta_0,\sigma^2\right)\le R\left(Fy;\beta_0,\sigma^2\right),$$

for all $\beta_0\in\Re^p$ and $\sigma^2>0$, where $\beta_0=\beta+f_0$. This means Qy is better than Fy, it is contradicting $Fy\ \square\ \beta$. Thus, the proof of sufficiency is completed.

Necessity. Firstly, it is proved that $f\in C\left(FX-I_p\right)$ when $Fy+f\ \square\ \beta$.

Denote Υ as the orthogonal matrix of projection onto $C\left\{B^{1/2}\left(FX-I_p\right)\right\}$ and $\gamma=B^{-1/2}\Upsilon B^{1/2}f$. Let $\eta=B^{1/2}(FX-I_p)$, using the knowledge of projection matrix, we get

$$\begin{aligned}\Upsilon &= \eta(\eta'\eta)^{+}\eta' \\ &= B^{1/2}(FX - I_p)\left[(FX - I_p)'B(FX - I_p)\right]^{+}(FX - I_p)'B^{1/2}\end{aligned} \tag{5.15.}$$

and

$$\gamma = (FX - I_p)\left[(FX - I_p)'B(FX - I_p)\right]^{+}(FX - I_p)'Bf\ . \tag{5.16.}$$

So, $\gamma \in C\left(FX - I_p\right)$ holds. Using Moore - Penrose properties such as $\eta'\eta\left(\eta'\eta\right)^{+}\eta' = \eta'$, we have

$$\begin{aligned}&R\left(Fy + f\,;\,\beta,\,\sigma^2\right) - R\left(Fy + \gamma\,;\,\beta,\,\sigma^2\right) \\ &\qquad = f'Bf - \gamma'B\gamma = f'B^{1/2}\left(I_p - \Upsilon\right)B^{1/2}f \geq 0\end{aligned}$$

From which we get

$$R\left(Fy + f\,;\,\beta,\,\sigma^2\right) \geq R\left(Fy + \gamma\,;\,\beta,\,\sigma^2\right), \tag{5.17.}$$

for all $\beta \in \Re^p$ and $\sigma^2 > 0$. The inequality holds if and only if $B^{1/2}f = \Upsilon B^{1/2}f$ holds, which is $f = B^{-1/2}\Upsilon B^{1/2}f = \gamma$. Inequality (5.17.) and the admissibility of $Fy + f$ together result in $f = \gamma$. Thus $f \in C\left(FX - I_p\right)$ is right.

Next, we prove that $Fy \in \beta$, when $Fy + f \in \beta$ holds.

If $Fy \notin \beta$, there exists an estimator Ky such that

$$R\left(Ky\,;\,\beta,\,\sigma^2\right) \leq R\left(Fy\,;\,\beta,\,\sigma^2\right), \tag{5.18.}$$

is right for all (β,σ^2) with the strict inequality holding at some point (β_1,σ_1^2). As $f \in C\left(FX - I_p\right)$, there exists $f_0 \in \Re^p$ such that $f = \left(FX - I_p\right) f_0$.

So, by (5.18.)

$$R\left(Fy + f\,;\,\beta,\,\sigma^2\right) = R\left(Fy;\,\beta + f_0,\,\sigma^2\right)$$
$$\geq R\left(Ky;\,\beta + f_0,\,\sigma^2\right) = R\left(Ky + (KX - I_p) f_0\,;\,\beta,\,\sigma^2\right),$$

holds for all (β,σ^2) and the strict inequality is true at point $\left(\beta_1 - f_0,\sigma_1^2\right)$, which means that $Ky + (KX - I_p) f_0$ is better than $Ly + a$, this is contradictory to the admissibility of $Fy + f$. Hence necessity is proved. Therefore, the proof of Theorem 5.3 is completed.

Corollary 5.5. For the model (4.1), if $C\beta$ is estimable, then $L^* y \in C\beta$ holds under the loss function $LF(L^* y, C\beta)$ if and only if $Ly \in \beta$ holds under the EBLF, where $Ly \in \mathrm{L}^h$, $L^* = S^{1/2}\left(L - wS^{-1}X'\right)$, $C = (1-w)S^{1/2}$ and $S = X'X$.

Proof: If $U = 0$, $\tilde{S} = S = X'X$, then $B = S$. From Theorem 5.1, since $L^* = B^{1/2}\left(L - wB^{-1}X'T^{+}\right)$, we can get $L^* = S^{1/2}\left(L - wS^{-1}X'\right)$ and $C = (1-w)S^{1/2}$.

Theorem 5.4. Consider the model (4.1) and the EBLF. Then $Ly \in \beta$ holds if and only if

(i) $L = LP_X$,

(ii) $(1-w)S^{1/2}\tilde{L}XS^{-1/2}$ is symmetric and $S^{1/2}\,\tilde{L}P_X\tilde{L}'S^{1/2} \leq (1-w)S^{1/2}\tilde{L}XS^{-1/2}$,

hold simultaneously, where $P_X = XS^{-1}X'$, $\tilde{L} = L - wS^{-1}X'$, $S = X'X$, $w = (1+\lambda_1 - \lambda_2)/2$ and $0 \le w \le 1$, $\lambda_1, \lambda_2 \in [0,1]$.

Proof: If $Ly \in \beta$, using Corollary 5.5, we know

$$Ly \in \beta \Leftrightarrow L^* y \in C\beta,$$

where $Ly \in \mathrm{L}^h$, $L^* y \in \mathrm{L}^h$, $C = (1-w)S^{1/2}$, $L^* = S^{1/2}\left(L - wS^{-1}X'\right)$, $S = X'X$.
Further, if we consider the special case of Theorem 2.13 when $k = p$ and $C = (1-w)S^{1/2}$, then L^* turns into L^h and we get the sufficient and necessary condition for $L^* y \in C\beta$ under the loss function $LF(\beta_*, C\beta)$ are as follows:

(a) $L^* = L^* P_X$

(b) $L^* P_X L^{*\prime} \le (1-w)L^* XS^{-1/2}$, where $P_X = XS^{-1}X'$ and $(1-w)L^* XS^{-1/2}$ is symmetric.

From (a),(b) and $L^* = S^{1/2}\left(L - wS^{-1}X'\right)$, we can easily get

$$L^* = L^* P_X \Leftrightarrow L = LP_X$$

and

$$L^* P_X L^{*\prime} \le (1-w)L^* XS^{-1/2}$$
$$\Leftrightarrow$$
$$S^{1/2}\tilde{L}P_X\tilde{L}'S^{1/2} \le (1-w)S^{1/2}\tilde{L}XS^{-1/2}.$$

Since $(1-w)L^{*}XS^{-1/2}$ is symmetric, then $(1-w)S^{1/2}\tilde{L}XS^{-1/2}$ is also symmetric. Which means (i) and (ii) are satisfied. Thus, the proof is completed.

Corollary 5.6. Consider the model (4.1) and the EBLF, $Ly \square \beta$ holds if and only if L can be written as

$$L = \left[wS^{-1} + (1-w)S^{-1/2}DS^{-1/2} \right] X',$$

where $D = S^{1/2}FXS^{-1/2}$ is symmetric and all eigenvalues in [0,1] and $w = \left(1+\lambda_1-\lambda_2\right)/2$, $\lambda_1, \lambda_2 \in [0,1]$, $0 \le w \le 1$.

Proof: If we take $D = (1-w)^{-1}S^{1/2}\tilde{L}XS^{-1/2}$, using Theorem 5.4 (ii), we get

$$S^{1/2}\tilde{L}P_X\tilde{L}'S^{1/2} \le (1-w)S^{1/2}\tilde{L}XS^{-1/2} \;\Leftrightarrow\; DD' \le D$$

where D is symmetric and all eigenvalues in [0,1]. Since Theorem 5.4 (i), $L = LP_X$ and $\tilde{L} = L - wS^{-1}X'$, we have $\tilde{L} = \tilde{L}P_X$, then

$$\tilde{L} = \tilde{L}X\left(X'X\right)^{-1}X' = S^{-1/2}DS^{-1/2}X'.$$

From which we can get the desired results

$$L = wS^{-1}X' + (1-w)S^{-1/2}DS^{-1/2}X'.$$

This is same as the (4.9) results in Theorem 4.3.

5.2.1. Example.

Example 5.1. Let us consider the admissibility of GLS estimator $\hat{\beta}_{GLSE}$.

Since $\hat{\beta}_{GLSE} = S_{T_*}^{-} X' T_*^{+} y$, where $S_{T_*} = X' T_*^{+} X$, $T_* = V + XX'$. Then $F = S_{T_*}^{-} X' T_*^{+}$.

Using following notations $P_{XT} = X(X' T_*^{+} X)^{-} X' T_*^{+}$, $B = \lambda_2 \tilde{S} + (1-\lambda_2)(X' T_*^{+} X)$ and $\tilde{F} = (X' T_*^{+} X)^{-} X' T_*^{+} - wB^{-1} X' T_*^{+}$, $W = (X' T_*^{+} X)^{-} - I$, $C = B^{1/2}(I_p - wB^{-1} X' T_*^{+} X)$, $w = (1 + \lambda_1 - \lambda_2)/2$, $0 \le w \le 1$. It is easy to verify that $\hat{\beta}_{GLSE}$ satisfies the all condition of Corollary 5.1. This shows that the GLS estimator $\hat{\beta}_{GLSE}$ of $\hat{\beta}$ is admissible with respect to the GEBLF.

Example 5.2. In the same way, it is also easy to verify that $\hat{\beta}$ satisfies the all condition of Theorem 5.2. This shows that the best linear unbiased estimator $\hat{\beta}$ of $\hat{\beta}$ is admissible with respect to the EBLF.

Example 5.3. Let us consider the admissibility of generalized ridge estimator $\hat{\beta}_K = (S + K)^{-1} X' y$, where $S = X'X$ and $K \in \Re_s^{\geq}$.

We can write $\hat{\beta}_K = (S + K)^{-1} X' y = Fy$ with $F = (S + K)^{-1} X'$. Let $P_X = XS^{-1} X'$, it can easily be checked that the condition $F = FP_X$ is fulfilled, which means the condition (i) of Theorem 5.2 is satisfied.

Since $K \in \Re_s^{\geq}$,

$$S \leq_L S + K,$$

using Theorem 2.7 implying that

$$(S+K)^{-1} \leq_L S^{-1}.$$

Then,

$$\begin{aligned} S^{1/2}\tilde{F}P_X\tilde{F}'S^{1/2} &= S^{1/2}\left[(S+K)^{-1}-wS^{-1}\right]S\left[(S+K)^{-1}-wS^{-1}\right]' S^{1/2} \\ &\leq S^{1/2}\left[S^{-1}-wS^{-1}\right]S\left[(S+K)^{-1}-wS^{-1}\right]' S^{1/2} \\ &= (1-w)S^{1/2}\left[(S+K)^{-1}S-wI\right]' S^{-1/2} \\ &= (1-w)S^{1/2}\,\tilde{F}XS^{-1/2}, \end{aligned}$$

where $\tilde{F} = F - wS^{-1}X'$ and $(1-w)S^{1/2}\tilde{F}XS^{-1/2}$ is symmetric. So, the condition (ii) of Theorem 5.2 is satisfied. Thus, the generalized ridge estimator $\hat{\beta}_K$ of $\hat{\beta}$ is admissible with respect to the EBLF.

Remark 5.1. It can also easily be prove the admissibility of $\hat{\beta}_K$ by using Theorem 4.5.

5.3. AOLE for the stocastic RC under the GEBLF

In this section, we focused on investigating the linear admissibility of SRC with random effects in the GMM model and SNC are considered for it to be admissible in random effects models under the GEBLF and its special conditions.

Consider the following GGM random effects model

$$y = X\beta + \varepsilon, \tag{5.19.}$$

where y is an observable random vector, where $\beta \in \Re^p$ is a vector of SRC with mean $R\delta$ and covariance $\sigma^2\Phi$. Moreover $\Phi,\ V \geq 0\ (V \neq 0)$ and $R \in \Re^{p\times r}$ with

$rk(R) = r$. Nevertheless $(\delta, \sigma^2) \in \left(\Re^r, (0,\infty)\right)$ are unknown parameters, $\varepsilon \in \Re^n$ is an error vector, where $E(\varepsilon) = 0$, $\text{cov}(\varepsilon) = \sigma^2 V$ and $E(\beta\varepsilon') = 0$. From which, we easily get $\text{cov}(y) = \sigma^2(V + X\Phi X')$.

Model (5.19.) encompass the following cases such as:

a) when $\Phi = 0$, β turns into non-SRC.

b) if (j, j) element of Φ is zero, the j th element of β is also zero, then some elements of β are stochastic and others are non-stochastic.

c) when $V > 0$ and $\Phi \neq 0$, random effects model (5.19.) is turns in to nonsingular model. Singular random effects model often appears in theoretical studies. But this nonsingular random effects model is used more in practical problems than the singular random effects model (5.19.).

For the nonsingular model with $V > 0$ and $\Phi \neq 0$, we can select $U \neq 0$. Then the GEBLF turns into the following:

$$GEBLF(\beta_*, \beta, t_1, t_2, t_3) = t_1(y - X\beta_*)'V^{-1}(y - X\beta_*) + t_2(\beta_* - \beta)'S(\beta_* - \beta) + t_3(X\beta_* - y)'V^{-1}X(\beta_* - \beta). \tag{5.20.}$$

Theorem 5.5. Let $\mathrm{L}_0^h = \left\{FX'T^+ y : F \in \mathrm{R}^{p\times p}\right\}$. Then L_0^h is a complete class of L^h.

Proof. Let us notice to advance that

$$\text{cov}\begin{pmatrix}\beta \\ y\end{pmatrix} = \text{cov}\left(\begin{pmatrix}I & 0 \\ X & I\end{pmatrix}\begin{pmatrix}\beta \\ \varepsilon\end{pmatrix}\right) = \sigma^2\begin{pmatrix}\Phi & \Phi X' \\ X\Phi & V + X\Phi X'\end{pmatrix}.$$

If $\beta_* = Ly \in \mathrm{L}^h$, using some knowledge of matrix and GEBLF, we get

$$R\left(Ly;\delta,\sigma^2\right)=E\left[t_1\left(y-X\beta_*\right)'T^+\left(y-X\beta_*\right)+t_2\left(\beta_*-\beta\right)'S\left(\beta_*-\beta\right)\right]$$
$$+(1-t_1-t_2)\mathrm{E}\left(\left(X\beta_*-y\right)'T^+X\left(\beta_*-\beta\right)\right)$$
$$=\sigma^2 tr\left[\left(LX-I_p\right)\Phi\left(LX-I_p\right)'G+LVL'G-2wLVT^+X\right]$$
$$+t_1\sigma^2 tr(VT^+)+\delta'R'\left(LX-I_p\right)'G\left(LX-I_p\right)R\delta,$$

(5.21.)

where $G=t_2S+(1-t_2)S_T$, $S_T=X'T^+X$, $w=(1+t_1-t_2)/2$.

One easily gets, $LH_{XT}y\in \mathrm{L}_0^h$ from $Ly\in \mathrm{L}^h$, where $H_{XT}=XS_T^-X'T^+$. And using Theorem 2.4,

$$R\left(Ly;\beta,\sigma^2\right)-R\left(LH_{XT}y;\beta,\sigma^2\right)$$
$$=\sigma^2 tr\left[L\left(I_n-H_{XT}\right)V\left(I_n-H_{XT}\right)'L'G\right]\geq 0,$$

Moreover, the equality holds if and only if $LV=LH_{XT}V$. It is showed that L_0^h is a complete class of L^h.

Corollary 5.7. Let $\mathrm{L}_1^h=\left\{FX'y:F\in \mathrm{R}^{p\times p}\right\}$. Then L_1^h is a complete class of L^h.

Proof: If $U=0, V=I, R=I,$ then $T=I$. Since this is the special case of Theorem 5.5, then we can easily get the desired results.

Theorem 5.6. Consider the linear model

$$\begin{cases} z = S_T\beta + u; \\ E(u) = 0, E(uu') = \sigma^2 X'T^+VT^+X, \end{cases} \tag{5.22}$$

where z is a p dimensional observation vector. $\beta \in \Re^p$ is similar to the one in model (5.19). Let $C = (I - wG^{-1}S_T)$ and $\tilde{L} = \{Fz : F \in \Re^{p\times p}\}$, then $\hat{F}z \overset{\tilde{L}}{\sqsupset} C\beta$ holds under loss function

$$LF_I(\beta_*, C\beta) = (\beta_* - C\beta)'G(\beta_* - C\beta),$$

if and only if $FX'T^+y \overset{L_0}{\sqsupset} \beta$ holds under the model (5.19) and GEBLF, where $\hat{F} = F - wG^{-1}$, $G = t_2S + (1-t_2)S_T$, $S_T = X'T^+X$, $w = (1+t_1-t_2)/2$, β_* be an estimator of β.

Proof. Firstly, we get from (5.21)

$$\begin{aligned} \mathrm{R}(FX'T^+y;\delta,\sigma^2) &= EL(FX'T^+y,\beta) \\ &+ \sigma^2 tr\left[\left(FS_T - I_p\right)\Phi\left(FS_T - I_p\right)'G\right] \\ &+ \sigma^2 tr\left(FX'T^+VT^+XF'G - 2wFX'T^+VT^+X\right) \\ &+ t_1\sigma^2 tr(VT^+) + \delta'R'\left(FS_T - I_p\right)'G\left(FS_T - I_p\right)R\delta. \end{aligned} \tag{5.23}$$

It is noted that

$$
\begin{aligned}
ELF_G(\hat{F}z, C\beta) &= E\left[(\hat{F}z - C\beta)'G(\hat{F}z - C\beta)\right] \\
&= \sigma^2 tr\left[\left(FS_T - I_p\right)\Phi\left(FS_T - I_p\right)' G\right] \\
&+ \sigma^2 tr\left(FX'T^{+}VF'T^{+}XG - 2wFX'T^{+}VT^{+}X\right) \\
&+ w^2\sigma^2 tr(X'T^{+}VT^{+}XG^{-1}) + \delta' R'\left(FS_T - I_p\right)' G\left(FS_T - I_p\right)R\delta.
\end{aligned}
\tag{5.24}
$$

If we suppose that $FX'T^{+}y \overset{L_0}{\sqsubset} \beta$, for an arbitrary $PX'T^{+}y \in \mathrm{L}_0$, we have

$$
\mathrm{R}\,(FX'T^{+}y; \delta, \sigma^2) \le \mathrm{R}\,(PX'T^{+}y; \delta, \sigma^2)\,, \tag{5.25}
$$

for all $(\delta, \sigma^2) \in \Re^r \times (0, \infty)$ and the strict inequality holds at one point, say (δ_0, σ_0^2). From (5.23) and (5.24), we get that (5.25) holds if and only if

$$
ELF_G(\hat{F}z, C\beta) \le ELF_G(\hat{P}z, C\beta)\,, \tag{5.26}
$$

holds for all $(\delta, \sigma^2) \in \Re^r \times (0, \infty)$ with strict inequality holding at (δ_0, σ_0^2), where $\hat{P} = P - wG^{-1}$.

Thus from the definition of admissibility, (5.25) holds if and only if $\hat{F}z \overset{\tilde{\mathrm{L}}}{\sqsubset} C\beta$ in model (5.22). It is easy to know that the conclusion of Theorem 5.6 is holds.

Corollary 5.8. Given linear model

$$\begin{cases} z = X'X\beta + u; \\ E(u) = 0, E(uu') = \sigma^2 X'X, \end{cases}$$

where z is a p dimensional observation vector. $\beta \in \Re^p$ is a vector of SRC with mean δ and covariance $\sigma^2\Phi_1$. Let $C = (I - wG^{-1}X'X)$ and $\tilde{L} = \{Fz : F \in \Re^{p\times p}\}$, then $\hat{F}z \overset{\tilde{L}}{\sqcup} C\beta$ holds under loss function

$$LF_I(\beta_*, C\beta) = (d - C\beta)'G(d - C\beta),$$

if and only if $FX'y \overset{L_0}{\sqcup} \beta$ under the model (5.19) and GEBLF, where $\hat{F} = F - wG^{-1}$, $G = t_2 S + (1-t_2)S_T, S_T = X'T^+X$, $w = (1+t_1-t_2)/2$.

Proof. If we take $U = 0, V = I, R = I,$ in model (5.19) and the GEBLF, then $T = I$. From Theorem 5.6, we can easily get the desired results.

Theorem 5.7. Under GGM random effects model (5.19) and the GEBLF, $Ay \overset{L^h}{\sqcup} \beta$ holds if and only if following four propositions hold simultaneously:

(a) $AV = AH_{XT}V$,

(b) $\left\{\tilde{A}X - (I - wG^{-1}S_T)\Phi\left(RR' + \Upsilon_T\right)^+\right\}(I - \Psi)\Upsilon_T = 0$

(c) $\tilde{A}X\Psi\Upsilon\Psi'X'\tilde{A}' \le \tilde{A}X\Psi\Upsilon\Psi'(I - wG^{-1}S_T) + (I - wG^{-1}S_T)\Phi\Psi'(AX - I)',$

(d) $C\left[(AX - I)R\right] = C\left((AX - I)\Psi(S_T - I)\right),$

where $H_{XT} = XS_T^{-}X'T^{+}$, $\Psi = R\left[R'\left(RR'+\Upsilon_T\right)^{+}R\right]^{-}R'\left(RR'+\Upsilon_T\right)^{+}$, $S_T = X'T^{+}X$ $\Upsilon_T = S_T^{-1}+\Phi-I$, $G = t_2S+(1-t_2)S_T$, $\tilde{A} = A-wG^{-1}X'T^{+}$, $w=(1+t_1-t_2)/2$.

Proof. It is easy to get (a) from Theorem 5.5. According to the Theorem 5.6, we have

$$FX'T^{+}y \overset{L_0}{\sqsupset} \beta \;\Leftrightarrow\; \hat{F}z \overset{L}{\sqsupset} C\beta .$$

By Theorem 2.20 and Theorem 2.14, under the GEBLF, $FX'T^{+}y \overset{L_0}{\sqsupset} \beta$ holds if and only if

$$\begin{aligned}\hat{F}\Sigma &= \hat{F}X_0R(R'X_0'W^{+}X_0R)^{-}RX_0'W^{+}\Sigma \\ &\quad + C\Phi X_0'W^{+}\left[I-X_0R(R'X_0'W^{+}X_0R)^{-}R'X_0'W^{+}\right]\Sigma\end{aligned}, \tag{5.27.}$$

$$\begin{aligned}\hat{F}X_0RMR'X_0'\hat{F}' \le (\hat{F}X_0-C)R(R'X_0'W^{+}X_0R)^{-}R'X_0'W^{+}X_0\Phi C' \\ + CRMR'X_0'\hat{F}'\end{aligned}, \tag{5.28.}$$

and

$$C\left[(\hat{F}X_0-C)R\right] = C(N) \tag{5.29.}$$

holds simultaneously, where

$M = (R'X_0'W^+X_0R)^- R'X_0'W^+\Sigma W^+X_0R(R'X_0'W^+X_0R)^-$, $\Sigma = X_0\Phi X_0' + \Lambda$, $X_0 = S_T$

$N = (\hat{F}X_0 - C)R\left[(R'X_0'W^+X_0R)^- RX_0'W^+X_0\Phi C' - MR'X_0'\hat{F}'\right]$, $W = \Sigma + X_0RR'X_0'$,
$+ (\hat{F}X_0 - C)RMR'(\hat{F}X_0 - C)'$

$\hat{F} = F - wG^{-1}$, $C = (I - wG^{-1}S_T)$, $G = t_2S + (1-t_2)S_T$, $\Lambda = X'T^+VT^+X$ and $w = (1+t_1-t_2)/2$.

If we let

$$\Psi = R(R'X_0'W^+X_0R)^- R'X_0'W^+X_0 = R\left[R'\left(RR'+\Upsilon\right)^+R\right]^- R'\left(RR'+\Upsilon\right)^+,$$

where $\Upsilon = S_T^{-1} + \Phi - I$, using matrix knowledge it can be shown that (5.27) is equivalent to

$$\left\{\hat{F}S_T - (I - wG^{-1}S_T)\Phi\left(RR'+\Upsilon\right)^+\right\}(I-\Psi)\Upsilon = 0. \qquad (5.30.)$$

(5.28) is equivalent to

$$\begin{aligned}\hat{F}XT^+X\Psi\Upsilon\Psi'X'T^+X\hat{F}' \le{}& \hat{F}S_T\Psi\Upsilon\Psi'(I - wG^{-1}S_T) \\ &+ (I - wG^{-1}S_T)\Phi\Psi'(FS_T - I)',\end{aligned} \qquad (5.31.)$$

and (5.29.) is equivalent to

$$C\left[(FS_T - I)R\right] = C((FS_T - I)\Psi(S_T - I)), \qquad (5.32.)$$

where Ψ is the same as that of Theorem 5.7 So, from (5.30)- (5.32.).we can get the desired (a) –(d) results, then Theorem 5.7 is proved.

When $\Phi = 0$, GGM random model (5.19) is turns to GGM model with fixed effect. As $rk(X) = p$, $rk(R) = r$ and $\beta = R\delta$, one easily gets that δ is estimable. Without loss of generality, let $R = I$, and then further results can be obtained.

Corollary 5.9. Under model (5.19) with the above assumption and the GEBLF, $Ay \in \beta$ holds if and only if

(i) $AV = AH_{XT}V$;

(ii) $\tilde{A}X\left[S_T^{-1} - I\right]X'\tilde{A}' \le \tilde{A}X\left[S_T^{-1} - I\right](I - wG^{-1}S_T)$;

$(iii)\, rk\left[(AX - I)\right] = rk\left[(AX - I)(S_T - I)\right]$,

where $S_T = XT^{+}X$ hold simultaneously.

Further, since $rk\left[(AX - I)\right] = rk\left[(AX - I)(I - S_T)\right] = rk\left[(AX - I)(S_T^{-} - I)S_T\right]$, if we let $N = (AX - I)(S_T^{-} - I)$, $rk(NS_T) = rk(NX'T^{-}X) \le rk(NX')$ holds evidently. On the other hand, by Frobenius inequality (see Theorem 2.12)

$$rk(NX'T^{-}X) \ge rk(NX') + rk(X'T^{-}X) - rk(X'),$$

and

$$rk(X'T^{-}X) = rk(X'),$$

we have

$$rk(NX'T^{-}X) \ge rk(NX').$$

Then

$$rk(NX'T^{-}X) = rk(NX').$$

And since G is nonsingular

$$\begin{aligned} rk\left[G^{-1}(AX-I)\right] = rk\left[(AX-I)\right] &= rk(NX'T^{-}X) \\ &= rk(NX') = rk(G^{-1}(AX-I)WX') \end{aligned}.$$

Which means that Corollary 5.8 is same as the results of Mirezi and Kaçıranlar (2022).

Further, under the Corollary 5.8, if we assume that $V > 0$, then (5.19) is the classically nonsingular linear model. And the corresponding result can be obtained, which is main result of Mirezi and Kaçıranlar (2022). Here we omit it.

Corollary 5.9. Under the nonsingular model and loss function (5.20), $Ly \square \beta$ holds if and only if

(a') $AV = AH_V V$,

(b') $\left[\tilde{A}X - (I - wG_V^{-1}S_V)\Phi\left(RR' + \Upsilon_V\right)^{+}\right](I-\Psi)\Upsilon_V = 0$,

(c') $\tilde{A}X\Psi\Upsilon_V\Psi'X'\tilde{A}' \le \tilde{A}X\Psi\Upsilon_V\Psi'(I - wG_V^{-1}S_V) + (I - wG_V^{-1}S_V)\Phi\Psi'(AX-I)'$,

(d') $C\left[(AX-I)R\right] = C\left((AX-I)\Psi(S_V - I)\right)$,

hold simultaneously, where

$H_V = XS_V^{-1}X'V^{-1}$, $S_V = X'V^{-1}X$, $\Psi = R\left[R'\left(RR'+\Upsilon_V\right)^{+}R\right]^{-}R'\left(RR'+\Upsilon_V\right)^{+}$,

$\Upsilon_V = S_V^{-1} + \Phi - I$, $G_V = t_2 S + (1-t_2)S_V$.

Theorem 5.8. Under the GGM random effects model (5.19) and the GEBLF, $Ay + a \overset{\mathrm{L}}{\sqsubset} \beta$ holds if and only if $Ly \overset{\mathrm{L}^h}{\sqsubset} \beta$ is admissible and $a \in C\left[\left(AX - I\right)R\right]$.

Proof. Sufficiency.

If a linear estimator $\tilde{\beta} = Ay + a \in \mathrm{L}$, then, using GEBLF, we get the risk under the GEBLF as follows:

$$R\left(Ay;\delta,\sigma^2\right) = \sigma^2 tr\left[\left(AX - I_p\right)\Phi\left(AX - I_p\right)' G + AVA'G - 2wAVT^{+}X\right]$$
$$+ t_1\sigma^2 tr(VT^{+}) + \left[\left(AX - I_p\right)R\delta + a\right]' G\left[\left(AX - I_p\right)R\delta + a\right], \tag{5.33.}$$

where $G = t_2 S + (1-t_2)S_T$. As $a \in C\left[\left(AX - I\right)R\right]$, there exists $a_0 \in \Re^p$ such that $a = \left(AX - I\right)Ra_0$. Hence, one gets from (5.33)

$$R\left(Ay + a;\, \delta,\, \sigma^2\right) = R\left(Ay;\, \delta + a_0,\, \sigma^2\right). \tag{5.34.}$$

Suppose $Ay + a \sqsubset \beta$, then there exists an estimator $Dy + d$ such that

$$R\left(Dy + d;\, \delta,\, \sigma^2\right) \le R\left(Ay + a;\, \delta,\, \sigma^2\right), \tag{5.35.}$$

holds for all $(\beta,\sigma^2)\in\Re^p\times(0,\infty)$. Let $\delta=-a_0$ in (5.35.), then by (5.33.) and (5.34.)

$$\begin{aligned}\left[d-\left(DX-I_p\right)Ra_0\right]'G\left[d-\left(DX-I_p\right)Ra_0\right]&=\lim_{\sigma^2\to0^+}R\left(Dy+d;-a_0,\sigma^2\right)\\&\le\lim_{\sigma^2\to0^+}R\left(Ay+a;\beta,\sigma^2\right)\\&=\lim_{\sigma^2\to0^+}R\left(Ay;0,\sigma^2\right),\end{aligned}\tag{5.36.}$$

So

$$d=\left(DX-I_p\right)a_0.\tag{5.37.}$$

From (5.33.) - (5.36.), we get

$$R\left(Dy;\delta_0,\sigma^2\right)\le R\left(Ay;\delta_0,\sigma^2\right),$$

for all $(\delta_0,\sigma^2)\in\Re^p\times(0,\infty)$, where $\delta_0=\delta+a_0$. Which means that Dy is better than Ay, this is contradicting $Ay\in\beta$. Thus, the sufficiency is proved.

Necessity. Firstly, it is proved that $a\in C\left[\left(AX-I\right)R\right]$ if $Ay+a\overset{L}{\in}\beta$.

Denote B as the orthogonal matrix onto $C\left\{G^{1/2}\left(AX-I\right)R\right\}$ and $\eta=G^{-1/2}BG^{1/2}a$.

Let $\eta=C\left\{G^{1/2}\left(AX-I\right)R\right\}$, using the knowledge of projection matrix, we get

$$B=\eta(\eta'\eta)^+\eta=G^{1/2}(AX-I_p)\left[(AX-I_p)'G(AX-I_p)\right]^+(AX-I_p)'G^{1/2},\tag{5.38.}$$

and

$$\eta = (AX - I_p)\left[(AX - I_p)'G(AX - I_p)\right]^{+}(AX - I_p)'Ga\,. \tag{5.39.}$$

So, $\eta \in C\left[(AX - I)R\right]$ holds.

Notice that

$$R\left(Ay + a;\, \delta, \sigma^2\right) = R\left(Ay;\, \delta, \sigma^2\right) + \left[(AX - I)R\delta + a\right]' B\left[(AX - I)R\delta + a\right].$$

So, using Moore - Penrose properties such as $\eta'\eta\left(\eta'\eta\right)^{+}\eta' = \eta'$, we have

$$R\left(Ay + a;\, \delta, \sigma^2\right) - R\left(Ay + \eta;\, \delta, \sigma^2\right) = a'G^{1/2}\left(I_p - B\right)G^{1/2}a \geq 0.$$

From which we get

$$R\left(Ay + a;\, \delta, \sigma^2\right) \geq R\left(Ay + \eta;\, \delta, \sigma^2\right), \tag{5.40.}$$

for all $(\delta, \sigma^2) \in \Re^p \times (0, \infty)$. The inequality holds if and only if $G^{1/2}a = BG^{1/2}a$ holds, which is $a = G^{-1/2}BG^{1/2}a = \eta$. Inequality (5.35) and the admissibility of $Ay + a$ together result in $a = \eta$. Thus $a \in C\left[(AX - I)R\right]$ is right.

Next, we prove $Ay \overset{L^h}{\sqsubset} \beta$ if $Ay + a \overset{L}{\sqsubset} \beta$ holds.

If $Ay \not\sqsubset \beta$, there exists an estimator Zy such that

$$R\left(Zy;\, \delta, \sigma^2\right) \leq R\left(Ay;\, \delta, \sigma^2\right), \tag{5.41.}$$

is right for all (δ,σ^2) with the strict inequality holding at some point (δ_1,σ_1^2). As $a\in C\left[\left(AX-I\right)R\right]$, there exists $l_0\in\Re^p$ such that $a=\left(AX-I\right)Ra_0$.

So, by (5.36),

$$R\left(Ay+a;\,\delta,\,\sigma^2\right)=R\left(Ay;\,\delta+a_0,\,\sigma^2\right)$$
$$\geq R\left(Zy;\,\delta+a_0,\,\sigma^2\right)=R\left(Zy+(ZX-I_p)Ra_0;\,\delta,\,\sigma^2\right),$$

holds for all (δ,σ^2) and the strict inequality is true at point $\left(\delta_1-a_0,\sigma_1^2\right)$, which means that $Zy+(ZX-I)Ra_0$ is better than $Ay+a$, this is contradictory to the admissibility of $Ay+a$. Hence necessity is proved. Thus, the proof of Theorem 5.8 is completed.

Now we can easily get some special results corresponding to those obtained above under the special condition in model and loss function.

Theorem 5.9. If $T=I$, under model (5.19) with $E(\varepsilon)=0$, $\mathrm{cov}(\varepsilon)=\sigma^2 I$ and the EBLF, $Ay\overset{L^h}{\sqcup}\beta$ holds if and only if following propositions

(a) $A=AH$;

$$(b)\ \hat{A}(I+X\Phi_1X')\hat{A}'\leq\hat{L}(I+X\Phi_1X')X(X'X)^{-1}(I-wG^{-1}X'X)$$
$$+(I-wG^{-1}X'X)\Phi_1(AX-I)';$$

hold simultaneously, where $\hat{A}=A-wG^{-1}X'$, $H=X(X'X)^{-}X'$.

Proof: By the Theorem 5.6, we know

$$FX'y \overset{L_0}{\sqsubset} \beta \Leftrightarrow \hat{F}z \overset{\tilde{L}}{\sqsubset} C\beta .$$

and using Theorem 2.20, Theorem 2.14 and Theorem 5.7 with $U = 0, V = I, R = I,$ under the GEBLF, $FX'y \overset{L_0}{\sqsubset} \beta$ holds if and only if

$$\hat{F}\Sigma = \hat{F}\tilde{X}_*(\tilde{X}_*'W^{-1}\tilde{X}_*)^{-}\tilde{X}_*'W^{-1}\Sigma + C\Phi_1\tilde{X}_*'W^{-1}\left[I - \tilde{X}_*(\tilde{X}_*'W^{-1}\tilde{X}_*)^{-1}\tilde{X}_*'W^{-1}\right]\Sigma \tag{5.42.}$$

$$\hat{F}\tilde{X}_*M\tilde{X}_*'\hat{F}' \le (\hat{F}\tilde{X}_* - C)(\tilde{X}_*'W^{-1}\tilde{X}_*)^{-1}\tilde{X}_*'W^{-1}\tilde{X}_*\Phi_1C' + CM\tilde{X}_*'\hat{F}' , \tag{5.43.}$$

and

$$C\left[(\hat{F}\tilde{X}_* - C)R\right] = C(\hat{N}), \tag{5.44.}$$

holds simultaneously, where $M = (\tilde{X}_*'W^{-1}\tilde{X}_*)^{-1}\tilde{X}_*'W^{-1}\Sigma W^{-1}\tilde{X}_*(\tilde{X}_*'W^{-1}\tilde{X}_*)^{-1}$, $\Sigma = X_0\Phi_1X_0' + V_0$ and $W = \tilde{X}_*\tilde{X}_*' + \Sigma$, $\tilde{X}_* = V_0 = X'X$

$\hat{N} = (\hat{F}\tilde{X}_* - Q)\left[(\tilde{X}_*'W^{-1}\tilde{X}_*)^{-1}\tilde{X}_*'W^{-1}\tilde{X}_*\Phi_1Q' - M\tilde{X}_*'\hat{F}'\right] + (\hat{F}\tilde{X}_* - Q)M(\hat{F}\tilde{X}_* - Q)',.$

By the matrix calculations, Equations (5.42) and (5.44) are identical, whereas Equations (5.43) is equivalent to

$$\hat{F}X'(I + X\Phi_1X')X\hat{F}' \le \hat{F}X'(I + X\Phi_1X')X(X'X)^{-1}C' + C\Phi'(\hat{F}X'X - C)'. \tag{5.45.}$$

Hence, from Theorem 2.20 and Equation (5.45.), it follows that $Ly \overset{L^h}{\sqcup} \beta$ under the GEBLF with $T = I, R = I$ if and only if (a) and (b) hold simultaneously. Therefore, this completes the proof of Theorem 5.9.

Theorem 5.10. Under the same condition of Theorem 5.9 for model (5.19) and GEBLF, $Ly + l \overset{L}{\sqsubset} \beta$ holds if and only if following propositions

(i) $A = AH$,

(ii) $\hat{A}(I + X\Phi_1 X')\hat{A}' \le \hat{A}(I + X\Phi_1 X')X(X'X)^{-1}(I - wG^{-1}X'X)$
$+ (I - wG^{-1}X'X)\Phi_1(AX - I)'$,

(iii) $a \in C(AX - I)$,

where $H = X(X'X)^{-}X'$, $\hat{A} = A - wG^{-1}X'$ hold simultaneously.

Proof: Let $V = T = R = I$, it can be easily get from Theorem 5.8 and Theorem 5.9.

Remark 1. All results of this paper are unrelated to the selection of the inverse of T. Hence, we have used the M-P inversion throughout the article.

5.3.1. Example.

In this section, we give some examples to illustrate applications of our Theorems 5.7 - 5.8.

Example 5.1. Let

$X=\begin{pmatrix}0 & 1\\ 1 & 1\end{pmatrix}, U=\begin{pmatrix}1 & -1\\ -1 & 1\end{pmatrix}, \Phi=\begin{pmatrix}2 & 2\\ 2 & 2\end{pmatrix}, V=\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}, S_T=\begin{pmatrix}1 & 1\\ 1 & 2\end{pmatrix}$,

$A=\begin{pmatrix}-0.45 & 0.8\\ 0.85 & 0.1\end{pmatrix}$, then it is easy to calculate $G=\begin{pmatrix}1 & 1\\ 1 & 2\end{pmatrix}$,

$\tilde{A}=A-wG^{-1}XT^{+}=\begin{pmatrix}0.05 & 0.30\\ 0.35 & 0.10\end{pmatrix}$, The left of condition (c) is

$\tilde{A}X\Psi\Upsilon_T\Psi'X'\tilde{A}'=\begin{pmatrix}0.725 & 0.575\\ 0.575 & 0.525\end{pmatrix}$ and the right is $\begin{pmatrix}0.775 & 0.55\\ 0.525 & 0.55\end{pmatrix}$. Hence, the results of the right subtracting the left is $\begin{pmatrix}0.05 & -0.025\\ -0.05 & 0.025\end{pmatrix}$ which is p.d. So, $Ay \square \beta$ in a class of HL estimators.

Example 5.2 Under the GGM random model (5.19) and the GBLF, let's consider the AOLE for the SRC.

Udner the some condition which given in tablo, criterion GEBLF reduces to the GBLF.

$$GBLF(\beta_*,\beta,w)=w\left(y-X\beta_*\right)'T^{+}\left(y-X\beta_*\right)+(1-w)(\beta_*-\beta)'\tilde{S}(\beta_*-\beta).$$

So, from Theorem 5.8, ones easily get $G=wS_T+(1-w)S$, $\left(I-wG^{-1}S_T\right)=(1-w)G^{-1}S$. Then for GBLF, we can get results corresponding to Theorem 5.8 and Theorem 5.9.

On the other hand its same as the results of Cao (2014).

If we take $U=0$ and $V=I$, then GBLF reduced to the ZBLF. Then for ZBLF, we can get results corresponding to Theorem 5.9 and Theorem 5.10. Its same as the results of Cao et al. (2014).

Example 4.3. Under the GGM random effects model (5.19) and the QLF, FLF and MLF loss functions, respectively, let's consider the AOLE for the SRC.

When $t_1 = 0, t_2 = 1$, GEBLF reduce to the QLF, then $w = 0$ and under the QLF, we can easily get desired results corresponding Theorem 5.7 and Corollary 5.9. It means that if $Ly \in \beta$ holds if and only if following $(i) \sim (iv)$ hold simultaneously:

(i) $AV = AH_{XT}V$,

(ii) $\left\{\tilde{A}X - \Phi\left(RR' + \Upsilon_T\right)^{+}\right\}(I - \Psi)\Upsilon_T = 0$,

(iii) $\tilde{A}X\Psi\Upsilon_T\Psi'X'\tilde{A}' \le \tilde{A}X\Psi\Upsilon_T\Psi' + \Phi\Psi'(AX - I)'$,

(iv) $rk\left[(AX - I)R\right] = rk\left((AX - I)\Psi(S_T - I)\right)$.

Since QLF, FLF, MLF are special case of the GEBLF when $t_1 = 0, t_2 = 1$ and above $(i) \sim (iv)$ propositions does not dependent on S. Therefore, under the GGM random effects model (5.19) and the FLF (or MLF), then SNC for AOLE for the SRC are same as $(i) \sim (iv)$.

6. CC OF SOME ESTIMATORS

In this chapter, we will consider the Farebrother estimator and its propoerties, compared it with ridge estimator under mse. Furthermore, we propose a minimum matrix valued risk estimator which combined the OLS and RLS estimators.

6.1. Comparision of Farebrother and Ridge Estimator

Farebrother (1984) introduced an estimator combining the sample and a special auxiliary information, which was also called a generalisation of the ridge regression estimator (RRE). In this section, we first theoretically compare Farebrother's proposed estimator with the ridge estimator according to the MSE criterion under the three different restrictions.

Farebrother (1984) theoretically compared the proposed estimator with OLS and RLS with respect to the MSE criterion as r fixed (see also Groß, 2003, p.137-138). However, he defined the estimator when r was random. Now, we consider the comparisons when r is random.

6.1.1. Case 1 (Under the unbiased restrictions)

r is random with $r = R\beta + \phi$, $E(\phi) = 0$ and $\mathrm{Cov}(\phi) = (\sigma^2 / k) I_m$.

Dispersion matrices for the estimators given in (3.11.) and (3.12.) are, respectively,

$$V_1 = Cov(\hat{\beta}_M) = \sigma^2 (S_V + R'W^{-1}R)^{-1}, \tag{6.1.}$$

$$V_2 = Cov(\hat{\beta}_V(k)) = \sigma^2 S_k^{-1} S_V S_k^{-1}, \tag{6.2.}$$

where, $S_V = X'V^{-1}X$ and $S_k = (S_V + kI_p)$.

From (6.1.) and (6.2.) we have:

$$V_1 - V_2 = \sigma^2\left[(\mathrm{S}_V + R'W^{-1}R)^{-1} - S_k^{-1}\mathrm{S}_V\, S_k^{-1}\right].$$

Lemma 1. $V_1 - V_2 > 0 \Leftrightarrow \lambda_i^G(\mathrm{L}) < 1 \; or \; \lambda_{\max}(\mathrm{G}^{-1}\mathrm{L}) < 1$**,** where $G = 2kI_p + k^2 S_V^{-1}$ and $L = R'W^{-1}R$ (Güler and Kaçıranlar (2009)).

Lemma 2. Let $\lambda_i^G(\mathrm{L}) < 1$. Then $M_1 - M_2 = MMSE(\hat{\beta}_M) - MMSE(\hat{\beta}_V(\mathrm{k})) > 0 \Leftrightarrow$ $\beta'\left[2kI_p + k^2 S_V^{-1} - S_k S_V^{-1}R'(W + RS_V^{-1}R')^{-1}RS_V^{-1}S_k\right]^{-1}\beta < \sigma^2/k^2$. (Güler and Kaçıranlar (2009)).

In the case $V = I_n$ and $W = 1/k\, I_m$, we get the following results from Lemma 1 and Lemma 2, respectively.

Corollary 6.1

a) The sampling variance of $\hat{\beta}(k)$ is less than that of the $\hat{\beta}_F(k)$ if and only if $\lambda_i^{G^*}(L^*) < 1$**,** $G^* = 2kI_p + k^2 S^{-1}$ and $L^* = kR'R$.

b) The MSE of $\hat{\beta}(k)$ is less than that of the $\hat{\beta}_F(k)$ if and only if

$$\beta'\left[2kI_p + k^2 S^{-1} - S_k^* S^{-1}R'(\frac{1}{k}I_m + RS^{-1}R')^{-1}RS^{-1}S_k^*\right]^{-1}\beta < \sigma^2/k^2,$$

where $S_k^* = S + kI_p$.

6.1.2. Case 2 (Under the biased restrictions)

r is random with $r = R\beta + \delta + \phi$, $E(\phi) = 0$ and $\text{Cov}(\phi) = \frac{\sigma^2}{k} I_m$,

$E(r) - R\beta = \delta$, $\delta \neq 0$.

The comparisons of $\hat{\beta}_M$ and GLS estimator have been given by Rao and Toutenburg (1995) with respect to the different mse criterions under the biased restrictions. Now, we compare the $\hat{\beta}_F(k)$ with $\hat{\beta}(k)$.

Using Equation (3.3.), we get the following equations:

$$E(\hat{\beta}(k)) = \beta - k(S_k^*)^{-1}\beta,$$

$$\text{Bias}(\hat{\beta}(k)) = -k(S_k^*)^{-1}\beta,$$

$$\text{Cov}(\hat{\beta}(k)) = V_3 = \sigma^2(S_k^*)^{-1}S(S_k^*)^{-1}, \text{ and}$$

$$\begin{aligned}\text{MSE}(\hat{\beta}(k)) = M_3 &= \sigma^2(S_k^*)^{-1}S(S_k^*)^{-1} + k^2(S_k^*)^{-1}\beta\beta'(S_k^*)^{-1} \\ &= (S_k^*)^{-1}(\sigma^2 S + k^2\beta\beta')(S_k^*)^{-1} .\end{aligned} \tag{6.3.}$$

Similarly, using Equation (3.13.), we get the following equations:

$$E(\hat{\beta}_F(k)) = \beta + kT_k^{-1}R'(E(r) - R\beta),$$

$$\text{Bias}(\hat{\beta}_F(k)) = kT_k^{-1}R'\delta, \quad \delta = E(r) - R\beta,$$

$$\text{Cov}(\hat{\beta}_F(k)) = V_4 = \sigma^2 T_k^{-1}, \text{ and}$$

$$MSE(\hat{\beta}_F(k)) = M_4 = T_k^{-1}(\sigma^2 T_k + k^2 R'\delta\delta' R)T_k^{-1} \tag{6.4.}$$

where $T_k = S + k\,R'R$.

With regard to performance by the sampling variance-covariance matrices of $\hat{\beta}_F(k)$ and $\hat{\beta}(k)$ it is found that

$$\begin{aligned} V_4 - V_3 &= \sigma^2 T_k^{-1} - \sigma^2 (S_k^*)^{-1} S (S_k^*)^{-1} \\ &= \sigma^2 (T_k^{-1} - (S_k^*)^{-1} S (S_k^*)^{-1}) = \sigma^2 \Delta . \end{aligned}$$

Since S is (p.d.) $(S > 0)$ and $kR'R$ is also p.d., $S + kR'R > 0$. Similarly $(S_k^*)^{-1} S (S_k^*)^{-1} > 0$. So,

$$\begin{aligned} \Delta > 0 &\Leftrightarrow T_k^{-1} - (S_k^*)^{-1} S (S_k^*)^{-1} > 0 \\ &\Leftrightarrow S_k^* S^{-1} S_k^* - T_k > 0 \\ &\Leftrightarrow 2kI_p + k^2 S^{-1} - kR'R > 0 \end{aligned} .$$

Then $\Delta > 0 \Leftrightarrow G^* - L^* > 0$. In this case we get the same expression in Corollary 6.1 (a).

Now, the mse matrices of $\hat{\beta}_F(k)$ and $\hat{\beta}(k)$ are compared when both estimators are computed using the same value of k, and restrictions are stochastic and biased. Using Equations (6.3.) and (6.4), it is found that

$$\begin{aligned} M_4 - M_3 &= \sigma^2 \Delta + k^2 T_k^{-1} R' \delta \delta' R T_k^{-1} - k^2 (S_k^*)^{-1} \beta \beta' (S_k^*)^{-1} \\ &= \sigma^2 \Delta + t_1 t_1' - t_2 t_2'. \end{aligned}$$

Thus, using the results of Trenkler and Toutenburg (1990), Theorem 2.3 (b) can be proposed. Now, we have the following theorem.

Theorem 6.1.

a) The sampling variance of $\hat{\beta}_F(k)$ is less than that of the $\hat{\beta}(k)$ if and only if $\lambda_i^{G^*}(L^*) < 1$, $G^* = 2kI_p + k^2 S^{-1}$ and $L^* = kR'R$.

b) Let $\lambda_i^{G^*}(L^*) < 1$. Then $MSE(\hat{\beta}_F(k)) - MSE(\hat{\beta}(k)) = \sigma^2\Delta + t_1t_1' - t_2t_2'$ is p.d. if and only if $t_2'(\sigma^2\Delta + t_1t_1')t_2 < 1$ holds.

6.1.3. Case 3: r is fixed and $r = R\beta + \delta_1$

If we compare the mse matrices of $\hat{\beta}_F(k)$ and $\hat{\beta}(k)$, then we obtain the following results.

Theorem 6.2.

a) The sampling variance of $\hat{\beta}_F(k)$ is less than or equal that of the $\hat{\beta}(k)$ if and only if $\lambda_{\max}(S^{-1}A_1'SA_1) \le 1$.

b) Let $\lambda_{\max}(S^{-1}A_1'SA_1) \le 1$, and assume that $t_3 \in \Re(\Delta_1)$. Then $MSE(\hat{\beta}(k)) - MSE(\hat{\beta}_F(k)) = D'\left[\sigma^2\Delta_1 + t_3t_3' - t_4t_4'\right]D$ is n.n.d. if and only if $(t_3'\Delta_1^-t_3 + 1)(t_4'\Delta_1^-t_4 - 1) \le (t_3'\Delta_1^-t_4)^2$ holds.

Observe that the requirement $t_3 \in \Re(\Delta_1)$ is equivalent to $t_3 = \Delta_1\gamma$ for some vector γ . Hence, $t_3'\Delta_1^-t_3 = \gamma'\Delta_1'\Delta_1^-\Delta_1\gamma = \gamma'\Delta_1\Delta_1^-\Delta_1\gamma = \gamma'\Delta_1\gamma$ and $t_3'\Delta_1^-t_3$ is therefore seen to be invariant to the choice of the generalized inverse Δ_1^- (see also Rao and Toutenburg (1995), Kaçıranlar et al.(2011)).

c) Let $\lambda_{\max}(S^{-1}A_1'SA_1) \le 1$ **,** and assume that $t_3 \notin \Re(\Delta_1)$ and $t_4 \in \Re(\Delta_1 + t_3t_3')$.

Then $MSE(\hat{\beta}(k)) - MSE(\hat{\beta}_F(k))$ is n.n.d. if and only if

$$t_4'\Delta_1^- t_4 - 2\varphi(t_4'v)(t_4'u) + \gamma\varphi^2(u't_4)^2 \le 1,$$

where $u = (I - \Delta_1\Delta_1^-)t_3,\ v = \Delta_1^- t_3,\ \gamma = 1 + t_3'\Delta_1^- t_3,\ and\ \varphi = (u'u)^{-1}$.

Proof: In this case there will be no change in $\mathrm{Cov}(\hat{\beta}(k)) = V_3$ and

$MSE(\hat{\beta}(k)) = M_3$.

Using $(S_k^*)^{-1} = S^{-1} - S^{-1}(k^{-1}I_p + S^{-1})^{-1}S^{-1} = S^{-1}(I_p - C') = S^{-1}D' = DS^{-1}$,

where $C = S^{-1}(k^{-1}I_p + S^{-1})^{-1},\ D = I_p - C,\ M_3$ can be rewritten as

$$MSE(\hat{\beta}(k)) = M_3 = S^{-1}D'(\sigma^2 S + k^2\beta\beta')DS^{-1}. \qquad (6.5.)$$

But, when r is fixed and $r = R\beta + \delta_1$, we get the following equations:

$$\mathrm{Cov}(\hat{\beta}_F(k)) = V_5 = \sigma^2 T_k^{-1}ST_k^{-1},$$
$$MSE(\hat{\beta}_F(k)) = M_5 = T_k^{-1}(\sigma^2 S + k^2 R'\delta_1\delta_1'R)T_k^{-1}. \qquad (6.6.)$$

Using

$$\begin{aligned} T_k^{-1} &= S^{-1} - S^{-1}R'(k^{-1}I_m + RS^{-1}R')^{-1}RS^{-1} \\ &= S^{-1}(I_p - R'(k^{-1}I_m + RS^{-1}R')^{-1}RS^{-1}) \\ &= S^{-1}(I_p - B') = S^{-1}A' = AS^{-1} \end{aligned}$$

where $B = S^{-1}R'(k^{-1}I_m + RS^{-1}R')^{-1}R$, $A = I_p - B$,

Equation (6.6.) can be rewritten as

$$MSE(\hat{\beta}_F(k)) = M_5 = S^{-1}A'(\sigma^2 S + k^2 R'\delta_1\delta_1' R)AS^{-1} \qquad (6.7.)$$

where $\delta_1 = r - R\beta$. With regard to performance by the sampling variance-covariance matrices of $\hat{\beta}_F(k)$ and $\hat{\beta}(k)$, it is found that

$$\begin{aligned} V_3 - V_5 &= \sigma^2 (S_k^*)^{-1} S (S_k^*)^{-1} - \sigma^2 T_k^{-1} S T_k^{-1} \\ &= \sigma^2 D'(S - (D')^{-1} A' S A D^{-1}) D \\ &= \sigma^2 D S^{1/2} (I_p - S^{-1/2} D^{-1} A' S A D^{-1} S^{-1/2}) S^{1/2} D \\ &= \sigma^2 D S^{1/2} (I_p - S^{-1/2} A_1' S A_1 S^{-1/2}) S^{1/2} D = \sigma^2 \Delta_1 , \end{aligned}$$

where $A_1 = AD^{-1}$, $D' = D$.

$$\begin{aligned} \Delta_1 \geq 0 &\Leftrightarrow \mathrm{I}_p - S^{-1/2} A_1' S A_1 S^{-1/2} \geq 0 \\ &\Leftrightarrow \lambda_{\max}(S^{-1/2} A_1' S A_1 S^{-1/2}) \leq 1 \\ &\Leftrightarrow \lambda_{\max}(S^{-1} A_1' S A_1) \leq 1 \end{aligned}$$

where $S^{-1}A_1'SA_1$ has the same eigenvalues as $S^{-1/2}A_1'SA_1S^{-1/2}$. Thus, the proof of (a) is completed.

Now, the mse matrices of $\hat{\beta}_F(k)$ and $\hat{\beta}(k)$ are compared when both estimators are computed using the same value of k. Using Equations (6.5.) and (6.7.), it is found that

$$\begin{aligned} M_3 - M_5 &= D'(\sigma^2\Delta_1 + k^2\beta\beta' - k^2 A_1'R'\delta_1\delta_1'RA_1)D \\ &= D'(\sigma^2\Delta_1 + t_3t_3' - t_4t_4')D \end{aligned}.$$

Using the results of Trenkler and Toutenburg (1995), the proof of (b) and (c) is completed.

6.1.2. The optimal choice of k

A non - homogeneous GR estimator is of the form

$$\begin{aligned} \hat{\beta}_{k,\beta_0} &= \beta_0 + \left(X'X + K\right)^{-1} X'\left(y - X\beta_0\right) \\ &= (X'X + K)^{-1}(X'y - K\beta_0) \end{aligned}$$

where the $p\times 1$ vector β_0 and the $p\times p$ symmetric n.n.d. matrix K are β_0 chosen non-stocastically. Then $\hat{\beta}_{k,\beta_0}$ is also called a linear estimator of ridge type (see also Groß, 2003).

If we choose $K = kR'R,\quad \beta_0 = R'(RR')^{-1}r$, then $\hat{\beta}_{k,\beta_0}$ will become Farebrother's estimator. If we consider this estimator under the Case 1 before mentioned (r is random with $r = R\beta + \phi,\ E(\phi) = 0$ and $\mathrm{Cov}(\phi) = (\sigma^2/k)I_m$, $E(r) = R\beta$), then

$$E(\beta_0) = E(R'(RR')^{-1}r) = R'(RR')^{-1}R\beta,$$

$$Cov(\beta_0) = R'(RR')^{-1}Cov(r)(RR')^{-1}R = (\sigma^2/k)R'(RR')^{-2}R = V, \quad (6.8.)$$

$$\beta_0 \sim (R'(RR')^{-1}R\beta\ ,\ V).$$

Now, we consider the problem of estimating the parameter k . In Farebrother (1984) and Rao and Toutenburg (1995) (see p. 139-147), there are no methods available to estimate of k . We propose a procedure to estimate $1/k$ rather than k as follows.

Since $\mathrm{Cov}(\hat{\beta}, \beta_0) = 0$, prior information β_0 independent of $\hat{\beta}$. Noting that $\hat{\beta} - \beta_0$ has a distribution with mean $B\beta$ and covariance $(V + \sigma^2(X'X)^{-1})$, where

$$B = (I - R'(RR')^{-1}R), \ (B\beta,\ V + \sigma^2(X'X)^{-1}) \sim (\hat{\beta} - \beta_0)$$

Then

$$E[(\hat{\beta} - \beta_0)'(\hat{\beta} - \beta_0)] = tr(V + \sigma^2(X'X)^{-1}) + \beta'B'B\beta\ .$$

Using (6.8.), an unbiased estimate of $1/k$ is

$$\frac{1}{k} = \frac{(\hat{\beta} - \beta_0)'(\hat{\beta} - \beta_0) - \sigma^2 tr(X'X)^{-1} + \beta'B'B\beta}{\sigma^2 tr(R'(RR')^{-2}R)}$$

If σ^2 is known (Crouse et al.(1995)). The more common situation is when σ^2 is unknown. Then, σ^2 can be estimated by $s^2 = (y - X\hat{\beta})'(y - X\hat{\beta})/(n-p)$ and estimate of $1/k$ is

$$\frac{1}{k} = \frac{(\hat{\beta} - \beta_0)'(\hat{\beta} - \beta_0) - s^2 tr(X'X)^{-1} + \beta' B' B\beta}{s^2 tr(R'(RR')^{-2} R)}. \qquad (6.9.)$$

There are other estimates of σ^2 which may be used as well.

6.2. A Minimum Matrix Valued Risk Estimator Combining RLS and OLS Estimators

Farebrother's estimator is seemed to be CC of RLS and OLS, however, Farebrother's estimator is not really a CC estimator of OLS and RLS. Therefore, we will try to obtain the CC estimator of OLS and RLS. In this section, our attention is focused on the CC estimator, $\bar{\beta} = A\hat{\beta} + (I - A)\hat{\beta}_R$. We have compared that the matrix valued risk of the new convex matrix estimator cannot exceed the individual matrix valued risk of $\hat{\beta}$ and $\hat{\beta}_R$.

6.2.1. Introduction

To find an estimator which minimizes the quadratic risk function over a class of appropriate functions has always been a subject of curiosity (see also, Vinod and Ullah (1981), Toutenburg (1982), Trenkler (1985), Singh and Chaubey (1987), Odell and et al. (1989), Rao and Toutenburg (1995), Groß (2003), and etc.).

When the property of unbiasedness is ignored, biased estimators are widely used by many researchers. Employing the matrix valued risk, some improved estimators are derived in the literature, some of those are the minimum MSE

estimator of Theil (1958), Farebrother (1975) and a CC of the OLS estimator and the GR estimator of Singh and Chaubey (1987).

If the restrictions are incorrect, the reduction in sampling variance is still maintained but the estimator becomes a biased one. The danger in using exact constraints is that because of this potential bias, the estimator has risk that is greater than the OLS estimator.

Thus, we have two competing estimators of β here, namely OLS estimator and RLS estimator. The choice among them is not so straightforward. This has led to the development of many estimation methods with many different approaches in the statistical literature.

The purpose of the present section is to get some improvements using the CC of OLS estimator and RLS estimator. So, we represent some estimation methods and risk functions, in Section 6.2.2. In Section 6.2.3, we obtain the CC of OLS estimator and RLS estimator. We also show that the matrix valued risk of the new convex matrix estimator cannot exceed the individual matrix valued risk of OLS estimator and RLS estimator. In Section 6.2.4, we present a Monte Carlo simulation experiment to compare new estimator with OLS estimator and RLS estimator. Finally, we give a numerical example to demonstrate the theoretical results.

6.2.2. Estimation Methods and Risk Functions

As we know, the OLS estimator in the multiple LRM (2.1) is given as:

$$\hat{\beta} = S^{-1}X'y\,, \tag{6.10}$$

where $S = X'X$.

Under the ordinary LRM, sometimes additional information about the unknown parameter vector β is available, which can be expressed in terms of a linear equation

$$r = R\beta \tag{6.11.}$$

where r is a $m\times 1$ vector of known elements, R is a $m\times p$ $(m<p)$ full row rank matrix with known elements. The RLS estimator is given by the vector

$$\hat{\beta}_R = \hat{\beta} - S^{-1}R'(RS^{-1}R')^{-1}(R\hat{\beta} - r)\,. \tag{6.12.}$$

6.2.3. A Minimum mse – Matrix Estimator

When the property of unbiasedness is ignored, biased estimators are widely used by many researchers. Employing the MDE, some improved estimators are derived in the literature, one of which is the minimum matrix mean square error estimator of Theil (1958) defined as

$$\begin{aligned}\beta^* &= \beta\beta' X'(X\beta\beta' X' + \sigma^2 I_n)^{-1} y \\ &= \frac{\beta' X' y}{\sigma^2 + \beta' X' X\beta}\beta \\ &= \rho\beta\end{aligned}$$

where $\rho = \dfrac{\beta' X' y}{\sigma^2 + \beta' X' X\beta}$. An existing problem with this estimator is that it cannot provide any practical benefit due to its unknown parameters β and σ^2. To avoid this problem, Rao (1973a) suggested to consider some prior values of these parameters, while, Farebrother (1975) replaced β and σ^2 with $\hat{\beta}$ and s^2, respectively, in β^* where

$$\hat{\beta} = (X'X)^{-1}X'y,$$

and

$$s^2 = \frac{1}{n-p}(y - X\hat{\beta})'(y - X\hat{\beta}) .$$

Thus, an operational form of β^* is appeared as

$$\hat{\beta}_F^* = \frac{\hat{\beta}'X'y}{s^2 + \hat{\beta}'X'X\hat{\beta}}\hat{\beta} = \left[1 - \frac{s^2}{s^2 + \hat{\beta}'X'X\hat{\beta}}\right]\hat{\beta} = \hat{\rho}\hat{\beta} ,$$

where $\hat{\rho} = \frac{\hat{\beta}'X'y}{s^2 + \hat{\beta}'X'X\hat{\beta}}$.

Odell and et al.(1989) searched through the set of all $p \times p$ matrices for a pair of matrices, A_1 and A_2, so that a criterion , M or m, is minimized. It is desirable to partition the space of all matrices M_p into two special subspaces D_p and C_p where D_p is the collection of all $p \times p$ diagonal matrices and C_p is the collection of all $p \times p$ diagonal matrices with all diagonal elements equal. There are a lot of studies on combining estimators found in the statistical estimation literature. List of combined estimators is given by Odell and et al. (1989, p.1629).

CC of two estimators can be useful when both estimators appear to be appropriate in a specific situation and a choice between them is not desired. The observed loss (or risk) of the CC is bounded by the maximum of the observed losses of the individual estimators. In some cases, there is a possibility for the observed loss of the CC to be even smaller than both individual observed losses.

The finding of a CC of non-homogeneous and HL estimators RLS estimators and OLS estimator, respectively is suggested by Groß (2003). He also pointed out that $\bar{\beta} = \alpha\hat{\beta} + (1-\alpha)\hat{\beta}_R$ was considered by Bock (1975) with stochastic α.

Now, we consider the problem of how to optimally weight two vector estimators for a best third, with respect to minimum matrix mean squared error. Consider the convex matrix estimator model

$$\bar{\beta} = A\hat{\beta}_1 + (I - A)\hat{\beta}_2 .$$

The following notations will be used.

$$M = M(\bar{\beta}, \beta),\ M_{ij} = E\left(\hat{\beta}_i - \beta\right)\left(\hat{\beta}_j - \beta\right)',\ i, j = 1,2 \qquad (6.13.)$$

M is minimized when

$$A = \left[M_{22} - M_{21}\right]\left[M_{11} + M_{22} - M_{12} - M_{21}\right]^{-1} \qquad (6.14.)$$

hence, the minimal value of M is given by

$$M_0 = M_{22} - \left[M_{22} - M_{21}\right]\left[M_{11} + M_{22} - M_{12} - M_{21}\right]^{-1}\left[M_{22} - M_{12}\right] \qquad (6.15.)$$

Or

$$M_0 = M_{11} - \left[M_{11} - M_{12}\right]\left[M_{11} + M_{22} - M_{12} - M_{21}\right]^{-1}\left[M_{11} - M_{21}\right]. \qquad (6.16.)$$

Clearly, if $\hat{\beta}_1$ and $\hat{\beta}_2$ are independent, then $M_{12} = 0$. Note also that

$M_0 < M_{11}$ and $M_0 < M_{22}$,

(See details, Odell and et al. (1989)).

Now we obtain the CC of OLS estimator and RLS estimator following Odell and et al. (1989). We also show that the MDE of the new convex matrix estimator cannot exceed the individual matrix valued risk of $\hat{\beta}$ and $\hat{\beta}_R$.

Theorem 6.3. Let $\bar{\beta} = A\hat{\beta} + (I - A)\hat{\beta}_R$. Then the MDE of the new convex matrix estimator cannot exceed the individual MDE of $\hat{\beta}$ and $\hat{\beta}_R$, where $A = P\delta\delta'\left[\delta\delta' + \sigma^2(RS^{-1}R')\right]^{-1}(P'P)^{-1}P' + H(I - P(P'P)^{-1}P')$, $\delta = r - R\beta$, $P = S^{-1}R'(RS^{-1}R')^{-1}$ and H is any $p \times p$ matrix.

***Proof*:** Let $\hat{\beta}_1 = \hat{\beta}, \hat{\beta}_2 = \hat{\beta}_R$, thus the convex model can be written as

$$\bar{\beta} = A\hat{\beta} + (I - A)\hat{\beta}_R$$

for all A in M_p, the set of all $p \times p$ matrices . Using (6.13), we get

$$M_{11} = \sigma^2 S^{-1},$$
$$M_{22} = \sigma^2 S^{-1} - \sigma^2 C + S^{-1}R'(RS^{-1}R')^{-1}\delta\delta'(RS^{-1}R')^{-1}RS^{-1},$$
$$M_{12} = \sigma^2 S^{-1} - \sigma^2 C = M_{21},$$
$$M_{22} - M_{21} = S^{-1}R'(RS^{-1}R')^{-1}\delta\delta'(RS^{-1}R')^{-1}RS^{-1} = P\delta\delta'P', \qquad (6.17)$$

$$M_{11} - M_{12} = \sigma^2 C = \sigma^2 S^{-1}R'(RS^{-1}R')^{-1}RS^{-1} = \sigma^2 P(RS^{-1}R')P', \qquad (6.18)$$

where

$$\delta = r - R\beta ,$$

$$\hat{\beta}_R = Qy + q = (S^{-1}\mathrm{X}' - C\mathrm{X}')y + S^{-1}R'(RS^{-1}R')^{-1}r ,$$

$$Q = S^{-1}\mathrm{X}^{'} - C\mathrm{X}^{'} \quad , \quad q = S^{-1}R'(RS^{-1}R')^{-1}r ,$$

$$C = C^{'} = S^{-1}R'(RS^{-1}R')^{-1}RS^{-1} ,$$

$$P = S^{-1}R'(RS^{-1}R')^{-1} .$$

Then , we can write the MDE of $\tilde{\beta}$

$$\begin{aligned} M &= E\left(\bar{\beta} - \beta\right)\left(\bar{\beta} - \beta\right)' \\ &= ATA' - AW_1' - W_1A' + M_{22} \end{aligned}$$

where

$$T = M_{11} + M_{22} - M_{12} - M_{21} ,$$

$$W_1 = M_{22} - M_{21} .$$

If we take the derivative of M with respect to A and set equal to zero, we get

$$2AT - 2W_1 = 0 .$$

M is minimized if A is the solution to

$$AT = W_1 ,$$

Or

$$A = W_1 T^+ + H(I - TT^+),\ H \text{ is any } p \times p \text{ matrix}. \qquad (6.19.)$$

Note that A is a solution to equation (6.19.) if the matrix T is singular (Odell and et al. (1989) only considered the case of non-singular fundamental matrix (F), p.1631, it corresponds to the matrix T in our case), where T^+ shows Moore-Penrose inverse (see Graybill, 1983, p. 114), and the minimum in that case is $M = M_0$, where

$$\begin{aligned} M_0 &= M_{12} - AW_1' \\ &= M_{22} - (W_1 T^+ + H - HTT^+)W_1' \\ &= M_{22} - W_1 T^+ W_1' \end{aligned} \qquad (6.20.)$$

$$= M_{22} - [M_{22} - M_{21}][M_{11} + M_{22} - M_{12} - M_{21}]^+ [M_{22} - M_{21}]. \qquad (6.21.)$$

Using $(W_1^+ + W_2^+)^+ = W_1 - W_1(W_1 + W_2)^+ W_1 = W_2 - W_2(W_1 + W_2)^+ W_2$, (see Rao and Mitra (1971)), Equation (6.20.) can be rewritten as follows

$$\begin{aligned} M_0 &= M_{22} - W_1(W_1 + W_2)^+ W_1 \\ &= M_{22} + \left[(W_2 - W_1) - W_2(W_1 + W_2)^+ W_2 \right] \\ &= M_{11} - W_2(W_1 + W_2)^+ W_2 \\ &= M_{11} - [M_{11} - M_{12}][M_{11} + M_{22} - M_{12} - M_{21}]^+ [M_{11} - M_{12}] \end{aligned} \qquad (6.22.)$$

where $W_1 = M_{22} - M_{21}$, $W_2 = M_{11} - M_{12}$, $T = W_1 + W_2$. Since

$$
\begin{aligned}
T &= M_{11} + M_{22} - M_{12} - M_{21} \\
&= S^{-1}R'(RS^{-1}R')^{-1}\delta\delta'(RS^{-1}R')^{-1}RS^{-1} + \sigma^2 S^{-1}R'(RS^{-1}R')^{-1}RS^{-1} \\
&= S^{-1}R'(RS^{-1}R')^{-1}\left[\delta\delta' + \sigma^2(RS^{-1}R')\right](RS^{-1}R')^{-1}RS^{-1} \\
&= P\left[\delta\delta' + \sigma^2(RS^{-1}R')\right]P', \qquad (6.23.)
\end{aligned}
$$

using some properties of Moore - Penrose inverse, T^+ can be found as follows

$$
\begin{aligned}
T^+ &= \left[M_{11} + M_{22} - M_{12} - M_{21}\right]^+ \\
&= \left(P'\right)^+\left[\delta\delta' + \sigma^2(RS^{-1}R')\right]^{-1}P^+ \\
&= P\left(P'P\right)^{-1}\left[\delta\delta' + \sigma^2(RS^{-1}R')\right]^{-1}\left(P'P\right)^{-1}P' . \qquad (6.24.)
\end{aligned}
$$

Using (6.18.), (6.22.) and (6.24.), we get

$$
M_0 = \sigma^2 S^{-1} - \sigma^4 S^{-1}R'\left[\delta\delta' + \sigma^2(RS^{-1}R')\right]^{-1}RS^{-1} . \qquad (6.25.)
$$

Since $RS^{-1}R' \neq 0$ and $\delta\delta' \geq 0$, then $M_0 - M_{11} = -\sigma^4 S^{-1}R'\left[\delta\delta' + \sigma^2(RS^{-1}R')\right]^{-1}RS^{-1} < 0$, where 0 denotes zero matrix. It can be deduced that $M_0 < M_{11}$.

Using again (6.17.), (6.21.), (6.24.) and (6.25.), we get

$$
M_0 - M_{22} = -P\delta\delta'\left[\delta\delta' + \sigma^2(RS^{-1}R')\right]^{-1}\delta\delta'P' < 0 .
$$

So, we have $M_0 < M_{22}$; thus the proof is completed.

Theorem 6.4. Let $\hat{\beta}$ and $\hat{\beta}_R$ be given estimator for β. Then a minimum MDE estimator of β in the class $\overline{\beta} = A\hat{\beta} + (I - A)\hat{\beta}_R$ is expressible in the form $\overline{\beta} = (I - DR)\hat{\beta} + Dr$, where $\delta = r - R\beta$, $P = S^{-1}R'(RS^{-1}R')^{-1}$ and $D = \sigma^2 S^{-1}R'\left[\delta\delta' + \sigma^2(RS^{-1}R')\right]^{-1} = P\left[I + \frac{1}{\sigma^2}\delta\delta'(RS^{-1}R')^{-1}\right]$.

Proof: From (6.23.) and (6.24.), TT^+ is equal to

$$\begin{aligned} TT^+ &= P\left[\delta\delta' + \sigma^2(RS^{-1}R')\right]P'P\left(P'P\right)^{-1}\left[\delta\delta' + \sigma^2(RS^{-1}R')\right]^{-1}\left(P'P\right)^{-1}P' \\ &= P(P'P)^{-1}P' \ . \end{aligned} \tag{6.26.}$$

Substituting of these matrices into (6.19.), we find

$$\begin{aligned} A &= W_1T^+ + H(I - TT^+) \\ &= \left[M_{22} - M_{12}\right]T^+ + H(I - P(P'P)^{-1}P') \\ &= P\delta\delta'\left[\delta\delta' + \sigma^2(RS^{-1}R')\right]^{-1}(P'P)^{-1}P' + H(I - P(P'P)^{-1}P') \ . \end{aligned} \tag{6.27.}$$

Then , we get

$$I - A = I - P\delta\delta'\left[\delta\delta' + \sigma^2(RS^{-1}R')\right]^{-1}(P'P)^{-1}P' - H(I - P(P'P)^{-1}P') \tag{6.28.}$$

Thus we get the minimum matrix mean square error estimator, $\overline{\beta}$ as follows:

$$\overline{\beta} = A\hat{\beta} + (I - A)\hat{\beta}_R$$

$$= A\hat{\beta} + (I - A)\left(\hat{\beta} - P(R\hat{\beta} - r)\right)$$

$$= \hat{\beta} - \left\{\left[I - P\delta\delta'\left[\delta\delta^{'} + \sigma^2(RS^{-1}R')\right]^{-1}(P'P)^{-1}P' \right.\right.$$

$$\left.\left. -H(I - P(P'P)^{-1}P')\right]P(R\hat{\beta} - r)\right\}$$

$$= \hat{\beta} - P\left\{\left(I - \delta\delta'\left[\delta\delta^{'} + \sigma^2(RS^{-1}R')\right]^{-1}\right)(R\hat{\beta} - r)\right\}$$

$$= \hat{\beta} - P\left\{\sigma^2(RS^{-1}R')\left[\delta\delta^{'} + \sigma^2(RS^{-1}R')\right]^{-1}\right\}(R\hat{\beta} - r)$$

$$= \left[I - \sigma^2 S^{-1}R'\left[\delta\delta^{'} + \sigma^2(RS^{-1}R')\right]^{-1}R\right]\hat{\beta}$$

$$+\sigma^2 S^{-1}R'\left[\delta\delta^{'} + \sigma^2(RS^{-1}R')\right]^{-1}r$$

$$= \left(I - DR\right)\hat{\beta} + Dr\,, \qquad (6.29.)$$

where $D = \sigma^2 S^{-1}R'\left[\delta\delta^{'} + \sigma^2(RS^{-1}R')\right]^{-1} = P\left[I + \frac{1}{\sigma^2}\delta\delta'(RS^{-1}R')^{-1}\right]$, so the proof is completed.

An existing problem about this estimator is that it cannot provide any practical benefit due to its unknown parameters β and σ^2. To get rid of this problem, Following Rao (1973a) and Farebrother (1975), we get adaptive $\overline{\beta}$ as follows:

$$\overline{\beta}_A = \left(I - \hat{D}R\right)\hat{\beta} + \hat{D}r \qquad (6.30.)$$

where $\hat{D} = s^2 S^{-1}R'\left[\hat{\delta}\hat{\delta}^{'} + s^2(RS^{-1}R')\right]^{-1}$, $\hat{\delta} = r - R\hat{\beta}$, and $s^2 = \frac{1}{n-p}(y - X\hat{\beta})'(y - X\hat{\beta})$. Note that $\overline{\beta}_A$ does not have any minimum MDE

estimator property, because of the substitution of the estimates by the OLS estimator of β and σ^2.

6.2.4. A Numerical Example And Monte Carlo Simulation

In this section, we first perform an illustration of a real word dataset to display the performance of the new minimum MDE estimator. While doing this application, we focus on the scalar valued risk criterion since it is more convenient for empirical applications. In addition to the minimum MDE estimator, we also take account of the performance of the OLS estimator and the RLS estimator in data analysis. Based on a recent article by Martinez et al. (2018), the olive oil dataset is selected to examine the theoretical assertions of this article. The olive oil dataset is collected to investigate the quality problem of the production of olive oil. Olive oil is often categorized as virgin or extra virgin by olive oil producers in order to check the product quality. The production of olive oil has an important economic impact and hence managers of many companies pay attention to such a classification. As detecting the quality of olive oil, some physical and chemical features one of which is color are considered. Acidity level, peroxide level and ultraviolet (UV) absorbance are the physicochemical covariates that are employed for explaining the color of the olive oil. The olive oil dataset includes 16 olive samples where there are scores on six attributes of a sensory panel and measurements of five physicochemical quality characteristics. In this dataset, first five oils are Greek, the next five are Italian and the last six are Spanish and we only focus on the scores for one coloration type as an indicator of the physical property. Therefore, the response is determined as Y: Brown for the numerical example and the independent variables X_1: Acidity, X_2: Peroxide, X_3: UV absorbance measured at 232 nm (K232), X_4: UV absorbance measured at 270 nm (K270) and X_5: absorbance sweep measured in a range of 266 to 274 nm (DK) are used to explain the dependent variable.

The correlation matrix of the covariates is found as $\begin{bmatrix} 1.0000 & 0.0467 & 0.1070 & 0.4796 & 0.5289 \\ 0.0467 & 1.0000 & 0.8740 & 0.5535 & 0.4865 \\ 0.1070 & 0.8740 & 1.0000 & 0.7104 & 0.5356 \\ 0.4796 & 0.5535 & 0.7104 & 1.0000 & 0.3112 \\ 0.5289 & 0.4865 & 0.5356 & 0.3112 & 1.0000 \end{bmatrix}$. The estimated coefficients of the OLS estimator is found as $\hat{\beta}:\begin{pmatrix} -0.2232 & 0.5889 & 0.2201 & 0.2611 & -0.2834 \end{pmatrix}'$ and the estimated variance of this dataset is computed as $\hat{\sigma}^2 = 0.0196$. As for the additional information in the equation (6.11.), $R = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \end{pmatrix}$ and $r = 1$ is arbitrarily chosen. Then, the RLS estimator and the proposed minimum MDE estimator are computed respectively as $\hat{\beta}_R:\begin{pmatrix} 0.0142 & 0.7456 & 0.3236 & 0.2183 & -0.3017 \end{pmatrix}'$ and $\overline{\beta}:\begin{pmatrix} -0.1839 & 0.6148 & 0.2372 & 0.2540 & -0.2864 \end{pmatrix}'$.

It is clear that the estimated coefficients obtained by the OLS estimator and the minimum MDE estimator are close to each other for the olive oil data. Under these circumstances, the estimated scalar risk value of the OLS estimator is computed as 0.5329 while the estimated scalar risk value of the RLS estimator is computed as 0.6081. Consequently, we calculate the estimated scalar risk value of the newly proposed minimum MDE estimator as 0.5298 which is less than not only the estimated scalar risk value of the OLS estimator but also the estimated scalar risk value of the RLS estimator. This valuable outcome demonstrates that the minimum MDE estimator is superior to its competitors by means of scalar valued risk criterion for the olive oil data.

We already verify the outperformance of the minimum MDE estimator via an up to date data analysis. To reach more extensive results, we now aim at

conducting a Monte Carlo simulation study. We use the model given by McDonald and Galarneau (1975) and Kibria (2003)

$$x_{ij} = (1-\rho^2)^{1/2} z_{ij} + \rho z_{ip+1}, \quad i = 1,...,n, \quad j = 1,...,p,$$

where ρ represents the correlation between explanatory variables and z_{ij} are pseudo-random numbers from the standard normal distribution. Five different sets of correlations are considered as 0.1, 0.3, 0.5, 0.8 and 0.99 to investigate the effects of different degrees of multicollinearity that can be encountered in applications on the estimators. The regressors are used so that $p = 5$ when $n = 30$ and $n = 60$. Observations on responses are generated by

$$y_i = \beta_0 + \beta_1 x_{i1} + ... + \beta_p x_{ip} + e_i, \quad i = 1,...,n,$$

where e_i are independent normal pseudo-random numbers with mean 0 and variance σ^2. The standard deviations in the simulation study are investigated as 1 and 5. The additional information is chosen as in the numerical example.

The results of the simulation experiment are given in Tables 6.1.-6.2.(see Appendix 6-7). These tables include the estimated scalar risk values of the OLS estimator, RLS estimator and the minimum MDE estimator under different specifications of the parameters. Based on the outcomes of the Tables 6.1.-6.2., we can reach some general comments. As the standard deviation increases, the estimated scalar risk values of all the estimators increase, as expected. An increment in the sample size results in a decrement in the estimated scalar risk values of the mentioned estimators. Increasing level of correlation has an increasing effect on the estimated scalar risk values of the estimators. Moving to the comparison of the performance of estimators, we can state that the performance of the minimum MDE

estimator becomes more visible as the sample size increases. Because of the definition of the minimum MDE estimator, the superiority of it to the OLS estimator and the RLS estimator differs based on the specification of the parameters. Herein, the results of the empirical applications are summarized for reference to the researchers, of course, they can use different specifications of the parameters while employing the minimum MDE estimator for their own research.

7. RESULTS AND SUMMARY

We consider admissibility, one of the more specific and applicable criteria that measure the performance of loss functions and different estimators. There are many loss functions such as QLF, PLF, ZBLF, proposed by many researchers, we chose an EBLF that has a different structure from other classical loss functions.

Since EBLF is more general than QLF, mse and PLF, we can see that EBLF is reduced to them when when λ_1 and λ_2 take the different values in $[0,1]$. Therefore, using the complete class concept, we know that we should only look for admissible estimators in the complete class. And we found some relationships between EBLF and QLF.

Explicit and implicit characterization of the ALE according to the EBLF was investigated. The relationship between EBLF and QLF is presented in Theorem 4.2. For the admissibility of an estimator of form Ly under the EBLF, the structure of L is given in Theorem 4.3. We also show that the ALE under the EBLF is a CC of the ALE under the sum of square residuals and QLF. The relationship between the admissibility of homogeneous and heterogeneous estimators is given in Theorem 4.4. Furthermore, the alternative theorem for the admissibility of heterogeneous estimators is given as Theorem 4.5. These topics are presented for the first time in this study.

Furthermore, under the EBLF, we investigated the admissibility of some known estimators such as the OLS estimator, RLS estimator, GR estimator, and Bayes estimator. The relationship between the Obenchian and admissible estimators and the shrinkage properties of admissible estimators are discussed.

If we combine the ALEs, it is natural to ask whether a CC is again linear admissible. we also tried to answer this question. we giving the corresponding theorem showed that the answer is affirmative both in HHC under the special

condition. Also, we illustrated the performance of the CC admissible estimator by given example.

In the light of above results, we also discussed the admissibility of the linear estimator of RC in the linear model under the EBLF and different types of LRMs, respectively, non-restricted model, restricted model, and elliptical restricted models. The SNC for linear estimators to be admissible are obtained respectively in HHC.

Under the singular linear model and EBLF, we investigated that ALE on linear functions of RC when the design matrix has not full column rank. Then, according to the relationship between the singular model and the restricted model, the SNC for linear estimators under the singular model and various restricted linear models to be admissible are obtained respectively in HHC.

In Chapter 5, we proposed a more comprehensive new loss function GEBLF. The AOLE is characterized in the GGM model with respect to the GEBLF. The SNC for linear estimators to be admissible with a dispersion matrix possibly singular among the set of linear estimators are obtained.

The Table 7.1.(see Appendix 8) show a summary of the results obtained and the relationship between admissible estimators corresponding to the loss functions.

It is stated that since GEBLF is the most general criterion, the corresponding admissible estimator is also more general are obtained. The results obtained under special conditions of the GEBLF lead to results known in the literature. So, if an estimator is admissible under GEBLF then under some conditions, it is also admissible with respect to the criteria which is a special case of GEBLF. The reverse is not true.

As we understand the definition (5.3) of GEBLF, there is a different loss function for each specification of λ_1 , λ_2 , λ_3 . It will be arise a quations that admissibility under one GEBLF doesn't necessarily imply admissibility under any other?

It may not be satisfied in all situations. For example, EBLF, ZBLF, and QLF are special cases of GEBLF as we have seen in the properties of GEBLF (iv) and (v). Since the EBLF is more sensitive than QLF, which means that if an estimator is admissible under the EBLF, then it is also admissible with respect to the QLF (see table and examples). But, if an estimator is admissible under the EBLF, it is not admissible under ZBLF. Or admissibility under squared - error loss is not implied admissibility under QLF.

In addition, we investigated the admissibility of the SRC in a random effects model with respect to the GEBLF where the covariance of the error vector ε is $\text{cov}(\varepsilon) = \sigma^2 V$ and $E(\beta\varepsilon') = 0$. From this, we also get some results of the special case of the GGM random effects model and GEBLF where $T = I$ and $\text{cov}(\varepsilon) = \sigma^2 I$.

In Chapter 6, in three different restrictions, we first compared the Farebrother estimator with the ridge estimator according to the MSE criterion. Here are the things we deduced from the results we got as follows.

- Under the unbiased restrictions: r is random with $r = R\beta + \phi$, $E(\phi) = 0$ and $\text{Cov}(\phi) = (\sigma^2/k) I_m$, ridge estimator UNW than Farebrother estimator;
- Under the biased restrictions: r is random with $r = R\beta + \delta + \phi$, $E(\phi) = 0$ and $\text{Cov}(\phi) = (\sigma^2/k) I_m$, ridge estimator UNW than Farebrother estimator;
- When r is fixed with $r = R\beta + \delta_1$, Farebrother estimator UNW than ridge estimator.

Furthermore, a minimum matrix-valued risk estimator is proposed that combines the RLS and OLS estimators. We have obtained that the matrix-valued

risk of the new convex matrix estimator cannot exceed the individual matrix-valued risk of OLS and RLS estimators. Then, the olive oil dataset is selected to examine the theoretical results above, and the results are supported by numerical examples and Monte Carlo simulation. Finally, and an estimate of the Farebrother estimator parameter k is proposed.

As a future study, what will happen when D in structure L in Theorem 4.3 takes different values? Any chance of turning into an estimator we know? And if the RC and random error have correlated that means, the admissibility of this study would be a matter of further research. In addition, it is planned to reveal the performances of these estimators for different models by using EBLF criteria.

REFERENCES

Aitken, A. C., (1935). On Least Square and Linear Combinations of Observations. Proceedings of the Royal Society of Edinburgh, 55:42-48.

Akdeniz, F., Wan, A.T.K., Akdeniz, E., (2005). Generalized liu type estimators under zellner's balanced loss function. Commun. Stat. - Theory Methods. 34, 1725–1736.

Belsley, D. A., Kuh, E., (2004). Regression Diagnostics, Identifying Infuluental Data and Sources of Collinearity. New Jersey Wiley-Interscience.

Bock, M.E., (1975). Minimax estimators of the mean of a multivariate normal distributions. Annals of Statistics, 3: 209-218.

Baksalary, J.K., Markiewicz, A., (1985). Admissible linear estimators in restricted linear models. Linear Algebra Its Appl.70, 9–19.

Baksalary, J.K., Markiewicz, A., (1986). Characterizations of admissible linear estimators in restricted linear models. J. Stat. Plan. Inference, 13, 395–398.

Baksalary, J.K., Markiewicz, A., (1988). Admissible linear estimators in the general Gauss- Markov model. J. Stat. Plan. Inference. 19, 349–359.

Banerjee, S., Roy, A., (2014) Linear algebra and matrix analysis for statistic. Boca Raton: CRC.

Berger, J.O. (1985). Statistical Decision Theory and Bayesian Analysis (2nd ed.). Springer, New York.

Baye, M.R., Parker, D.F., (1984). Combining ridge and principal components regression: a mony demand illustration. Communications in Statistics, Theory and Methods. 13,197-205.

Cohen, A., (1966). All Admissible Linear Estimates of the Mean Vector. Ann. Math. Stat., 37, 458–463.

Chen, X.R., Wu, Q., Zhao, L., (1985). Parameter estimation theory for linear models. Beijing Science Press.

Crouse, R.H., Jin, C., Hanumara, R .C., (1995). Unbiased ridge estimation with prior information and ridge trace. Commun. Statist. Theor. Meth. 24(9): 2341-2354.

Cao, M. (2009). Φ admissibility for linear estimators on RC in a general multivariate linear model under balanced loss function. J. Stat. Plan. Inference. 139, 3354–3360.

Cao.M., Kong, F. C., (2013). General admissibility for linear estimators in a general multivariate linear model under balanced loss function. ACTA Math. Sin.-Engl. Ser., 29, 1823–1832.

Cao.M., (2014). Admissibility of linear estimators for the stochastic regression coefficient in a general Gauss–Markoff model under a balanced loss function. J. Multivar. Anal.124,25-30

Chaturvedi, A., Shalabh, (2014). Bayesian Estimation of Regression Coefficients Under Extended Balanced Loss Function. Commun. Stat. - Theory Methods. 43, 4253–4264.

Cao.M.X., Shena, G.J., (2016). Linearly admissible estimators of mean vector with respect to balanced loss function in multivariate statistics. Commun. Stat. - Theory Methods, 45, 1063–1069.

Cao.M.X., He, D.J., (2017). Admissibility of linear estimators of the common mean parameter in general linear models under a balanced loss function. J. Multivar. Anal. 153, 246–254.

Cao.M., He, D., (2019). Linearly admissible estimators on linear functions of regression coefficient under balanced loss function. Commun. Stat. - Theory Methods, 48, 2700–2706.

Cao.M.X., Xu, X.Z.,He, D.J., (2014). Linearly admissible estimators of stochastic regression coefficient under balanced loss function. Statistics: A Journal of Theoretical and Applied Statistics 48 359-366.

Durbin, J., (1953). A note on regression when there is extraneous information about one of the coefficients. J. Amer. Statist. Assoc., 48799-808.

Dey, D.K., Ghosh, M., Strawderman, W.E., (1999). On estimation with balanced loss functions. Stat. Probab. Lett. 45, 97–101.

Ferguson, T.S., (1967). Mathematical Statistics. Academic Press, New York.

Farebrother, R. W., (1975). The minimum mean square error linear estimator and ridge regression. Technometrics, 17 (1): 127–128.

Farebrother, R. W., (1984). The restricted least squares estimator and ridge regression. Commun . Statist. Theor. Meth, 13:191-196.

Giles, J.A., Giles, D.E.A., Ohtani, K. (1996). The exact risks of some pre-test and stein-type regression estimators umder balanced loss. Communications in Statistics - Theory and Methods. 25(12), 2901-2919.

Groß, J.G., Markiewicz, A., (2004). Characterizations of admissible linear estimators in the linear model. Linear Algebra Its Appl. 388, 239–248.

Gruber D.M.H.J., (2004). The Efficiency of Shrinkage Estimators with Respect to Zellner's Balanced Loss Function. Commun. Stat. - Theory Methods. 33, 235–249.

Güler, H., and Kaçıranlar, S., (2009). A comparison of mixed and ridge estimators of linear models. Commun. Stat. Simul. Comput. 38:368-401.

Graybill, F. A., (1983). Matrices With Application in Statistics. Wadsworth, California .

Groß, J. (2003). Linear Regression. Springer Verlag, NewYork . doi:10.1007/978-3-642-55864-1.

Hoerl, E., Kennard, R.W., (1970). Ridge regression: Biased for non-orthogonal problems. Technometrics,12 (1):55-67.

Hoffmann, K., (1995). All admissible linear estimators of the regression parameter vector in the case of an arbitrary parameter subset. J. Stat. Plan. Inference. 48, 371–377.

Hoffmann, K., (1977). Admissibility of linear estimators with respect to restricted parameter sets Mathematische Operationsforschung Statistik , Ser.Statistics, 8, 425- 438.

Hu, G., Peng, P., (2010). Admissibility for linear estimators of regression coefficient in a general Gauss–Markov model under balanced loss function. J. Stat. Plan. Inference. 140, 3365–3375.

Jafari Jozani, M., Marchand, E., Parsian, A., (2006). On estimation with weighted balanced-type loss function. Stat. Probab. Lett., 76, 773–780.

Kuks, J., and Olman, V., (1972). A minimax linear estimator of regression coefficients. Izvestia Akadamii Nauk Estonskoy SSR, 21:66-72.

Klonecki, W., Zontek, S., (1988). On the structure of admissible linear estimators. J. Multivar. Anal. 24, 11–30.

Kaçıranlar, S., Sakallıoğlu, S., Özkale, M. R., Güler, H., (2011). More on the restricted ridge regression estimation. Journal of Statistical Computation and Simulation, 81(11): 1433-1448.

Kaçıranlar.S., Dawoud, I., (2019). The optimal extended balanced loss function estimators. J. Comput. Appl. Math. 345, 86–98.

Liu, G., Yin, H., (2022). Admissibility in general Gauss–Markov model with respect to an ellipsoidal constraint under weighted balanced loss. Commun. Stat. - Theory Methods, 51(4), 1054-1066.

Lamotte. L.R., (1982). Admissiblity in linear estimation. Ann Stat., 10, 235–255.

Lu, C.Y., Shi, N.Z., (2002). Admissible linear estimators in linear models with respect to inequality constraints. Linear Algebra Its Appl. 354, 187–194.

Lehmann, E.L., G. Casella, G., (2005). Theory of Point Estimation, 2nd ed., New York: Springer-Verlag. DOI:10.1007/b98854.

Markiewicz, A., (1996). Characterization of general ridge estimators. Stat. Probab. Lett. 27, 145–148.

Mathew, T., Rao, C.R., Sinha, B.K., (1984). Admissible linear estimation in singular linear models. Commun. Stat. - Theory Methods, 13, 3033–3045.

Mirezi, B., Kaçıranlar, S., (2021a). A minimum matrix valued risk estimator combining restricted and ordinary least squares estimators. Commun. Stat. - Theory Methods, DOI:10.1080/03610926.2021.1934032.

Mirezi, B., Kaçıranlar, S., (2021b). Characterization of linearly admissible estimators under extended balanced loss function. Kybernetika. 57(4): 613-627.

Mirezi, B., Kaçıranlar, S., (2022). Admissible Linear estimators in the general Gauss – Markov model under generalized extended balanced loss function. Statistical Papers, DOI: 10.1007/s00362-022-01298-9 (accepted for publication).

Odell, P. L., Dorsett, D., Young, O., Igwe, J., (1989). Estimator Models For Combining Vector Estimators. Math.Comput.Modelling 12: 1627-1642. DOI :10.1016/0895-7177(89)90338-5

Ohtani, K., Giles, D.E.A., Giles, J.A., (1997). The exact risk performance of a pre-test estimator in a heteroskedastic linear regression model under the balanced loss function. Econom. Rev., 16, 119–130.

Ohtani, K., (1998). The exact risk of a weighted average estimator of the OLS and stein-rule estimators in regression under balanced loss. Stat. Risk Model, 16, 35–46.

Ohtani, K., (1999). Inadmissibility of the Stein-rule estimator under the balanced loss function. J. Econom. 88, 193–201.

Özbay, N., Kaçıranlar, S., (2017). The performance of the adaptive optimal estimator under the extended balanced loss function. Commun. Stat. - Theory Methods. 46, 11315–11326.

Özbay, N., Kaçıranlar, S., (2019). Risk performance of some shrinkage estimators. Commun. Stat. - Simul. Comput. https://doi.org/10.1080/03610918.2018.1554116

Rao, C.R., (1976). Estimation of Parameters in a Linear Model. Ann. Stat., 4, 1023–1037.

Rao, C. R., (1973a). Linear statistical inference and its applications. 2nd ed. Wiley, New York. DOI : 10.1112/jlms/s1-42.1.382b.

Rao, C. R., and S. K Mitra., (1971). Generalized Inverse of Matrices and its Applications. Wiley, New York.

Rao, C. R., and H. Toutenburg. (1995). Linear models: least squares and alternatives. Springer-Verlag , New York.

Rao, C. R., Toutenburg, H., (1995). Linear models least squares and alternatives. New York, Springer-Verlag.

Rodrigues, J., Zellner. A., (1994). Weighted balanced loss function and estimation of the mean time to failure. Commun. Stat. - Theory Methods. 23, 3609-3616.

Shalabh., (1995). Performance of Stein-rule procedure for simultaneous prediction of actual and average values of study variable in linear regression model. Jan Tinbergen Award Paper. Bull. 50th Sess. Int. Stat. Inst. 1375–1390.

Shalabh., Toutenburg, H., Heumann, C., (2009). Stein-rule estimation under an extended balanced loss function. J. Stat. Comput. Simul. 79, 1259–1273.

Stepniak, C., (1989). Admissible linear estimators in mixed linear models. J. Multivar. Anal., 31, 90–106.

Swindel, B.F., (1976). Good estimators based on prior information. Communication in. Statistics. Theory and Methods, 5: 1065-1075.

Theil, H., Goldberger, A. S., (1961). On pure and mixed statistical estimation in economics. International Economics Reviews, 2(1): 65-78.

Theil, H., (1963). On the use of incomplete prior information in regression analysis. Journal of the American Statistical Society, 58:401-414.

Trenkler, G., (1984). On the performance of biased estimatiors in the linear regression model with correlated or heteroscedastic errors. Journal of Econometrics, 25:179-190.

Terasvirta, T. (1979a). Some results on improving the least squares estimation of linear models by mixed estimation. *Discussion paper 7914,* Louvain, CORE.

Terasvirta, T. (1979b). The polynomial distributed lag revisited. Discussion paper 7919, Louvain, CORE.

Toutenburg, H. (1982). Prior Information in Linear Models. New York: John Wiley & Sons.

Trenkler, G and Toutenburg. (1990). Mean squared error matrix comparisons between biased estimators - An overview of recent results, *Statist*. 31:165-179.

Theil, H. (1958). Economic Forecasts and Policy. Amsterdam: North-Holland.

Theil, H. and Goldberger, A.S., (1961) On pure and mixed estiamtion in econometrics. Intl. Economic Rev., 2 65-78.

Theil, H., (1963). On the use of incomplete prior information in regression analysis. J. Amer. Statist. Assoc., 58, 401-414.

Trenkler, G. (1985). Mean square error matrix comparisons of estimators in linear regression. Commun . Statist. A, 14: 2495-2509.

Toutenburg , H., (1982). Prior information in Linear Models. Wiley, New York.

Vinod, H. D., and Ullah, A., (1981). Recent advances in regression models. Marcel Dekker, New York.

Varian, H. R., (1975). A Bayesian Approach to Real Estate Assessment. In:Fienberg, S. E. and Zellner, A. (Eds.), Studies Bayesian Econometrics and Statistics in Honor of Leonard J. Savage. North Holland, Amsterdam, 195–208.

Wu, Q., (1986). Admissibility of linear estimators of regression coefficient in a general Gauss-Markov model. Acta Math. Sci. 9, 251– 256.

Wu, Q., (1988). Several results on admissibility of a linear estimate of stochastic regression coefficients and parameters, Acta Math. Appl. Sin. 1195–106.(Chinese section).

Wu, Q., Dong, L., (1988). Necessary and sufficient conditions for linear estimators of stochastic regression coefficients and parameters to be admissible under quadratic loss. Acta Math. Sinica 31, 145–157

Wan, A.T.K., (1994). Risk comparison of inequality constrained least squares and other related estimators under balanced loss. Economics Letters. 46, 203–210.

Xu, X., Wu, Q., (2000). Linear admissible estimators of regression coefficient under balanced loss. Acta Math. Sci. 20, 468–473.

Zellner, A., (1986). Bayesian estimation and prediction using asymmetric loss functions. Journal of the American Statistical Association, 81 , 446-451.

Zellner, A., (1994). Bayesian and non-Bayesian estimation using balanced loss functions. Statistical Decision Theory and Related Topics V. New York, Springer. 377–390.

Zhang, S., Gui, W., (2014). Admissibility in general linear model with respect to an inequality constraint under balanced loss. J. Inequalities Appl. DOI:10.1186/1029-242X-2014-70.

Zinodiny, S., Rezaei, S., Nadarajah. S., (2017). Minimax estimation of the mean of the multivariate normal distribution. Commun. Stat. - Theory Methods, 46, 1422–1432.

CURRICULUM VITAE

Buatikan MIREZI. She graduated from Kashgar No 1 High School in 2005. Her undergraduate education at the Faculty of Science at Xi'an Jiaotong University in China in 2007. She graduated from the Department of Mathematics in 2011. She began her Master's education in the same year and in 2014 got her Master's Degree in Applied Mathematics. She was awarded a Turkish scholarship and came to the Çukurova university in Adana and started her Ph.D. in the statistic department of the institute of natural and applied science at Çukurova university.

APPENDICES

Appendix 1: Proof of Lemma 4.3

Proof: Note that , $\beta_* = Ly \in \mathrm{L}^h$, using some properties of trace and Theorem 2.3, from Equation (2.32.) and Equation (2.34), we get

$$\begin{aligned} R\left(Ly;\beta,\sigma^2\right) = \sigma^2 tr\Big[& t_1\left(I_n - XL\right)'\left(I_n - XL\right) + t_2\left(L'X'XL\right) \\ & + \left(1-t_1-t_2\right)\Big(\left(XL - I_n\right)' XL\Big)\Big] \\ & + \beta' X'\left(I_n - XL\right)'\left(I_n - XL\right) X\beta. \end{aligned}$$

On the other hand, if $Ly \in \mathrm{L}^h$, then $LP_X y = L_0 X'y \in \mathrm{L}_0^h$, where $P_X = X\left(X'X\right)^{-1} X'$ and $L_0 = LX(X'X)^{-1}$. So ,

$$\begin{aligned} & R\left(Ly;\beta,\sigma^2\right) - R\left(LP_X y;\beta,\sigma^2\right) \\ & \quad = \sigma^2 tr\Big[t_1\left(XLL'X'\right) + t_2\left(XLL'X'\right) - t_1\left(XLP_X L'X'\right) - t_2\left(XLP_X L'X'\right) \\ & \quad + \left(1-t_1-t_2\right)\left(L'X'XL - XLP_X L'X'\right)\Big] \\ & \quad = \sigma^2 tr\Big[t_1 L'X'XL\left(I_n - P_X\right) + t_2 L'X'XL\left(I_n - P_X\right)\Big] \\ & \quad + \sigma^2 tr\Big[\left(1-t_1-t_2\right)L'X'XL\left(I_n - P_X\right)\Big] \\ & \quad = \sigma^2 tr\Big[L'X'XL\left(I_n - P_X\right)\Big] \geq 0. \end{aligned}$$

Moreover, the equality holds if and only if $L = LP_X$.

Appendix 2: The proofs of Equations in Theorem 4.2.

Proof: Firstly, let $L_0X'y \in \mathrm{L}_0^h$ we get Equation (4.4) using (4.2.):

$$\begin{aligned} R\left(L_0X'y;\,\beta,\,\sigma^2\right) &= \sigma^2 tr\left[t_1\left(I_n - XL_0X'\right)'\left(I_n - XL_0X'\right)\right] \\ &\quad + t_2\sigma^2 tr\left(XL_0'X'XL_0X'\right) \\ &\quad + \left(1 - t_1 - t_2\right)\sigma^2 tr\left(\left(XL_0X' - I_p\right)' XL_0X'\right) \\ &\quad + \beta' X'\left(I_n - XL_0X'\right)'\left(I_n - XL_0X'\right)X\beta \\ &= \sigma^2 tr\left[XL_0'X'XLX' - \left(1 + t_1 - t_2\right)XL_0X'\right] + \sigma^2 tr\left(t_1 I_n\right) \\ &\quad + \beta'\left(L_0X'X - I_p\right)' X'X\left(L_0X'X - I_p\right)\beta \end{aligned}.$$

Since $0 \le t_1, t_2 \le 1$, we can find such a w denoted by $w = \left(1 + t_1 - t_2\right)/2$, where $0 \le w \le 1$. Then above equation can be rewritten as follows:

$$\begin{aligned} R\left(L_0X'y;\,\beta,\,\sigma^2\right) &= \sigma^2 tr\left[XL_0'X'XL_0X' - 2wXL_0X'\right] + \sigma^2 tr\left(t_1 I_n\right) \\ &\quad + \beta'\left(L_0X'X - I_p\right)' X'X\left(L_0X'X - I_p\right)\beta \\ &= \sigma^2 tr\left[\left(L_0X' - w\left(X'X\right)^{-1}X'\right)' X'X\left(L_0X' - w\left(X'X\right)^{-1}X'\right)\right] \\ &\quad + \sigma^2 tr\left(t_1 I_n - w^2 X\left(X'X\right)^{-1}X'\right) \\ &\quad + \beta'\left(L_0X'X - I_p\right)' X'X\left(L_0X'X - I_p\right)\beta. \end{aligned}$$

Let

$$B = X'X\ ,\ C = \left(1 - w\right)I_p \text{ and } \tilde{L} = L_0 - w\left(X'X\right)^{-1}.$$

Then, above equation equal to

$$\begin{aligned} R\left(L_0X'y;\,\beta,\,\sigma^2\right) &= \sigma^2 trX\tilde{L}'B\tilde{L}X' + \sigma^2 tr\left(t_1 I_n - w^2 X\left(X'X\right)^{-1}X'\right) \\ &\quad + \beta'\left[\tilde{L}X'X - C\right]' B\left[\tilde{L}X'X - C\right]\beta. \end{aligned}$$

Appendix 3: The proof of Equation (4.16.).

$$
\begin{aligned}
R\left(Ly+a;\beta,\sigma^2\right) &= t_1\Big[\sigma^2 tr\left(I_n-XL\right)'\left(I_n-XL\right)y+\beta' X'\left(I_n-XL\right)'\left(I_n-XL\right)X\beta \\
&\quad -a'X'(I_n-XL)X\beta-\beta' X'\left(I_n-XL\right)' Xa+a'X'Xa\Big] \\
&\quad +t_2\Big[\sigma^2 tr(L'X'XL)+\beta' X'L'X'XLX\beta+\beta' X'L'X'X(a-\beta) \\
&\quad +(a-\beta)'X'XLX\beta+(a-\beta)'XX(a-\beta)\Big] \\
&\quad +\left(1-t_1-t_2\right)\Big[\sigma^2 tr\left(XL-I_n\right)' XL \\
&\quad +\beta' X'\left(I_n-XL\right)' XLX\beta+a'X'XLX\beta \\
&\quad +\beta' X'\left(XL-I_n\right)' X(a-\beta)+a'X'X(a-\beta)\Big] \\
&= \sigma^2 tr\Big[t_1\left(I_n-XL\right)'\left(I_n-XL\right)+t_2\left(L'X'XL\right)+\left(1-t_1-t_2\right)\Big(\left(XL-I_n\right)' XL\Big)\Big] \\
&\quad +t_1\left\{\Big[\left(I_p-LX\right)\beta-a\Big]' X'X\Big[\left(I_p-LX\right)\beta-a\Big]\right\} \\
&\quad +t_2\left\{\Big[\left(LX-I_p\right)\beta+a\Big]' X'X\Big[\left(LX-I_p\right)\beta+a\Big]\right\} \\
&\quad +\left(1-t_1-t_2\right)\left\{\Big[\left(LX-I_p\right)\beta+a\Big]' X'X\Big[\left(LX-I_p\right)\beta+a\Big]\right\} \\
&= \sigma^2 tr\Big[t_1\left(I_n-XL\right)'\left(I_n-XL\right)+t_2\left(L'X'XL\right)+\left(1-t_1-t_2\right)\Big(\left(XL-I_n\right)' XL\Big)\Big] \\
&\quad +\Big[\left(LX-I_p\right)\beta+a\Big]' X'X\Big[\left(LX-I_p\right)\beta+a\Big].
\end{aligned}
$$

Appendix 4

Table 3.1. Linear ridge type estimators.

Estimator	symbol	$\hat{\beta}_{K,\beta_0}$
Ordinary least squares	$\hat{\beta}$	$K=0,\ \beta_0=0$
Marquardt	$\hat{\beta}^{(r,c)},\ c>0$	$K=((1-c)/c)U_2\Lambda_2U_2',\ \beta_0=0$
Ridge	$\hat{\beta}_k$	$K=kI_p,\ \beta_0=0$
Ridge	$\hat{\beta}_{k_1,\ldots,k_p}$	$K=Udiag(k_1,\ldots,k_p)U',\ \beta_0=0$
Farebrother	$\hat{\beta}_R(k), k<\infty$	$K=kR'R, \beta_0=R'(RR')^{-1}r$
Almost unbiased ridge	$\hat{\beta}_k^J$	$K=k^2(X'X+2kI_p)^{-1},\ \beta_0=0$
Iteration	$\hat{\beta}_{\delta,m}$	$K=X'X(T^{-1}-I_p),\ \beta_0=0,$ $T=I_p-(I_p-\delta X'X)^{m+1}$
Shrinkage	$\hat{\beta}(\rho)$	$K=\rho X'X,\ \beta_0=0$
Direction modified shrinkage	$\hat{\beta}_{\beta_0}(\rho)$	$K=\rho X'X$

Appendix 5

Table 4.1. Displays some well-known estimators from the $\mathcal{L}^{Ob}(\beta)$ with respect to a given spectral decomposition $X'X = U\Lambda U' = U_1\Lambda_1 U_1' + U_2\Lambda_2 U_2'$.

Estimator	symbol	$\tilde{L}\hat{\beta} \in \mathcal{L}^{ob}(\beta)$	$w=0,\ L\hat{\beta} \in \mathcal{L}^{Ob}(\beta)$
OLS	$\hat{\beta}$	$\tilde{L} = I_p$	$L_{mse} = I_p$
Principal componen t	$\hat{\beta}^{(r)}$	$\tilde{L} = U_1 U_1'$	$L_{mse} = U_1 U_1'$
Marquardt	$\hat{\beta}^{(r,c)}$	$\tilde{L} = U_1 U_1' + cU_2 U_2'$	$L_{mse} = U_1 U_1' + cU_2 U_2'$
Ridge	$\hat{\beta}_k$	$\tilde{L} = U(\Lambda + kI_p)^{-1}\Lambda U'$	$L_{mse} = U(\Lambda + kI_p)^{-1}\Lambda U'$
Baye-Parker	$\hat{\beta}_k^{(r)}$	$\tilde{L} = U_1(\Lambda_1 + kI_r)^{-1}\Lambda_1 U_1'$	$L_{mse} = U_1(\Lambda_1 + kI_r)^{-1}\Lambda_1 U_1'$
Almost unbiased ridge	$\hat{\beta}_k^J$	$\tilde{L} = U(\Lambda^2 + 2k\Lambda)^{-1}(\Lambda + kI_p)^{-2}U'$	$L_{mse} = U(\Lambda^2 + 2k\Lambda)^{-1}(\Lambda + kI_p)^{-2}U'$
iteration	$\hat{\beta}_{\delta,m}$	$\tilde{L} = U\left(I_p - (I_p - \delta\Lambda)^{m+1}\right)U'$	$L_{mse} = U\left(I_p - (I_p - \delta\Lambda)^{m+1}\right)U'$
shrinkage	$\hat{\beta}(\rho)$	$\tilde{L} = (1+\rho)^{-1} I_p$	$L_{mse} = (1+\rho)^{-1} I_p$

Appendix 6

Table 6.1. The results of the Monte Carlo simulation for $n = 30$

		Estimated scalar risk values of the estimators		
ρ	σ	**OLS estimator**	**RLS estimator**	**Minimum MDE estimator**
0.1	1	0.226018	1.090972	0.236232
	5	5.650469	5.058985	5.178957
0.3	1	0.234321	1.821804	0.239401
	5	5.858030	6.143654	5.634517
0.5	1	0.266258	0.729981	0.277721
	5	6.656455	5.959887	6.194491
0.8	1	0.520336	2.278703	0.527688
	5	13.008420	13.163018	12.709079
0.99	1	9.148035	47.217421	9.214254
	5	228.700898	243.998541	227.418942

Appendix 7

Table 6.2. The results of the Monte Carlo simulation for $n = 60$

		Estimated scalar risk values of the estimators		
ρ	σ	**OLS estimator**	**RLS estimator**	**Minimum MDE estimator**
0.1	1	0.094515	0.090517	0.089205
	5	2.362881	1.886220	2.108416
0.3	1	0.097293	0.236325	0.100766
	5	2.432343	2.187228	2.278394
0.5	1	0.110273	0.415454	0.111399
	5	2.756846	2.782712	2.700963
0.8	1	0.212963	0.562719	0.213367
	5	5.324086	5.499377	5.319596
0.99	1	0.097293	0.236325	0.100766
	5	2.432343	2.187228	2.278394

Appendix 8

Table 7.1. Admissible estimators under the GEBLF and its special different types of criteria (The table is given on the next page).

GEBLF $GEBLF(\beta,\hat{\beta}_0,\hat{\lambda}_1,\hat{\lambda}_2,\hat{\lambda}_3)=\hat{\lambda}_1(y-X\beta_0)^T(y-X\beta_0)+\hat{\lambda}_2(\beta-\beta)'\hat{S}(\beta-\beta)+\hat{\lambda}_3(X\beta_0-y)T^*X(\beta_0-\beta)$				
Relationship between the GEBLF and other Loss functions				
condition			Loss function	Linear admissible estimator
$T=VV^*+XUX^*$ $B=\lambda_1\hat{S}+(1-\lambda_1)S_y$	$0\le\hat{\lambda}_1,\hat{\lambda}_2\le 1$ $w=1+\hat{\lambda}_1-\hat{\lambda}_2/2$		GEBLF	$\beta_{(GEBLF)}=wB^{-1}X^*T^*y+(I-wB^{-1}X^*T^*X)(I-UX^*T^*X)\left\{\left[X^*T^*X(I-UX^*T^*X)\right]^{+}\right\}^{1/2}D\left\{\left[X^*T^*X(I-UX^*T^*X)\right]^{+}\right\}^{1/2}\left\{\right.X^*T$
$T\neq I_n$	$\hat{\lambda}_1=w,\hat{\lambda}_2=1-w$		GBLF	$\beta_{GBLF}=wB^{-1}X^*T^*y+(I-wB^{-1}S(I-UX^*T^*X)\left\{\left[X^*T^*X(I-UX^*T^*X)\right]^{+}\right\}^{1/2}D\left\{\left[X^*T^*X(I-UX^*T^*X)\right]^{+}\right\}^{1/2}\left\{\right.X^*T^*y$
	$\hat{\lambda}_1=1,\hat{\lambda}_2=0$		GLS	$\hat{\beta}=(X^*T^*X)^{-1}X^*T^*y$
$T=I_n$ $V=I_n$, $U=0$	$\hat{\lambda}_1=w,\hat{\lambda}_2=1-w$		ZBLF	$\beta_{ZBLF}=wB^{-1}X^*y+(I-wB^{-1}S(XX)^{-1/2}D(XX)^{-1/2}X^*y.$
	$\hat{S}=X^*X$		EBLF	$\beta_{EBLF}=w\hat{\beta}+(1-w)\beta_{ols}$
	$\lambda_1=0,\hat{\lambda}_2=1,\hat{S}=I$		QLF	$\beta_{ols}=(XX)^{-1/2}D(XX)^{-1/2}X^*y$
	$\lambda_1=0,\hat{\lambda}_2=1,\hat{S}=\mathrm{var}^{-1}(\beta)$		FLF	$\beta_{FLF}=(XX)^{-1/2}D(XX)^{-1/2}X^*y$
	$\lambda_1=0,\hat{\lambda}_2=1,\hat{S}=\mathrm{cov}^{-1}(\beta)$		MLF	$\beta_{MLF}=(XX)^{-1/2}D(XX)^{-1/2}X^*y$
	$\hat{\lambda}_1=1,\hat{\lambda}_2=0$		QLF	$\hat{\beta}_{OLS}=(X^*X)^{-1}X^*y$