AN ABSTRACT OF THE THESIS OF

Shaun S. Wulff for the degree of Doctor of Philosophy in Statistics, presented on June 2, 1999.

Title: Explicit Linear Maximum Likelihood Estimation in Mixed Models

Abstract approved: Redacted for Privacy

David S. Birkes

Mixed models have been widely used to model data from experiments which have fixed and random factors. Often there is interest in the estimation of fixed effects and variance components. The likelihood procedure is a general technique that has been applied to such problems. This procedure can be computationally difficult, as iterative algorithms are needed to solve for estimators that satisfy the likelihood equations. Previous research has been done to identify conditions under which there exists an explicit linear estimator for the full fixed effect vector or for the full variance component vector.

This thesis will examine explicit linear estimation in mixed models. The previous results will be extended to explicit linear estimation of a linear combination of the fixed effects or of a linear combination of the variance components. Specific results for the existence of an explicit linear estimator for a subvector of the full fixed effect vector or a subvector of the full variance component vector will also be presented.

The results of the thesis will be demonstrated using various models encountered in the experimental design setting. Applications will also be presented which include interpreting iterative procedures to solve for the estimators, saving computer time in profile likelihood calculations for fixed effects, and uniformly minimum variance unbiased estimation.

Explicit Linear Maximum Likelihood Estimation in Mixed Models

by

Shaun S. Wulff

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Presented June 2, 1999
Commencement June 2000

Doctor of Philosophy thesis of Shaun S. Wulff presented on June 2, 1999

APPROVED:

Redacted for Privacy

Major Professor, representing Statistics

Redacted for Privacy

Chair of Department of Statistics

Redacted for Privacy

Dean of Graduate School

I understand that my thesis will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my thesis to any reader upon request.

Redacted for Privacy

Shaun S. Wulff, Author

Acknowledgments

First, I thank my Lord and Saviour Jesus Christ whose compassion and strength I relied on daily. My faith in Christ brought me through difficult and challenging times. I give all glory, praise, and honor to Him. As the Holy Bible says, "Be careful for nothing; but in every thing by prayer and supplication with thanksgiving let your requests be made known unto God. And the peace of God, which passeth all understanding, shall keep your hearts and minds through Christ Jesus" (Philippians 4:6-7, KJV).

Thank you Heidi, my lovely and devoted wife, for supporting me throughout the entire process. I appreciate your willingness to move to western Oregon and the multitude of other sacrifices you made. You are directly responsible for this degree due to your assistance, compassion, understanding, and endurance. I love you and I am indebted to you.

Thanks to my mother, Carolyn, who encouraged me at a young age to treat my education as a job in which to work hard and give my best. Her encouragement and financial support made the pursuit of an advanced degree possible.

Thanks to Dr. David Birkes who patiently and tirelessly worked with me. I appreciate his willingness to answer questions and provide the details necessary to develop my understanding. His efforts made this research work academically rewarding and it was an honor to work with him.

Thanks to the department at Oregon State for its financial support. I enjoyed learning and working with the faculty and staff. In particular, thanks to my boss, Dr. Clifford Pereira, who worked hard to help me grow in the area of statistics. In addition, thanks to Dr. Scott Urquhart who also helped me apply statistical knowledge to consulting problems. Their efforts to clear my thinking and improve my statistical abilities are greatly appreciated.

I would also like to thank faculty members at Montana State University. Thank you Dr. William F. Quimby for encouraging me to pursue a statistics degree. Thank you Dr. John Borkowski for working with me in research and supporting me throughout my academic career. I also give thanks to Dr. Robert J. Boik for his advice, support, and efforts to bring me to Oregon State University.

I want to thank all current and former students, in which there are too many to mention, who have helped me with my studies and encouraged me throughout my schooling. I needed your help and the friendships made my college experience enjoyable.

Table of Contents

Table of Contents (Continued)

Table of Contents (Continued)

Table of Contents (Continued)

List of Tables

To my wife, Heidi, and my mother, Carolyn,
who have combined to provide love and support from
1st grade to the 24th.

Explicit Linear Maximum Likelihood Estimation in Mixed Models

1. Introduction

1.1. Motivation

Mixed linear models have been widely used to model data from experiments which have fixed and random factors. There is often interest in estimating the fixed effects and variance components in these models. Likelihood procedures have been used to solve this estimation problem. However, likelihood procedures can be computationally difficult, as iterative algorithms are needed to solve for the estimators that satisfy the likelihood equations. The estimators from the iterative procedure can also be hard to interpret and their performance can be difficult to assess.

Previous research has found conditions under which likelihood estimators of the vector of the fixed effects and the vector of the variance components are explicit and linear. These results characterize the full case where the complete parameter vector is under consideration. However, there are many cases under which such explicit linear estimators do not exist. The purpose of this study is to generalize these results by obtaining conditions under which the likelihood estimator of a linear combination of fixed effects or variance components is explicit and linear when explicit and linear estimators do not exist in the full case. Knowledge of the existence of explicit linear likelihood estimators for linear combinations of interest involving the fixed effects and variance components can be helpful for calculation, interpretation, and assessing performance.

1.2. Previous Results

The estimation of fixed effects and variance components has been a important statistical problem. Fixed effects can be estimated using least squares techniques. Ordinary least squares solutions are often inadequate in models with random effects since they do not account for the covariance. Generalized least squares can be used to account for the covariance when it is known. However, it is usually the case that the covariance depends on some unknown parameters. Estimated generalized least squares estimators can be used to estimate the fixed effects where the generalized least squares estimator is calculated using the estimated covariance matrix. However, the issue then is how to estimate the variance components (Searle et al.,1992). "For balanced data, it has been common practice to estimate these parameters by equating the means squares in the ANOVA table to their expectations" (Harville,1977). This method of estimation was generalized to the unbalanced case using techniques by Henderson (1953). Likelihood techniques have become a more popular alternative and these methods are reviewed in the next section.

1.2.1. Likelihood Estimation

The likelihood procedure is a general technique that can be applied to estimating variance components in balanced and unbalanced mixed models. This technique requires an assumption of a probability distribution for the data. It is typically assumed that the data are from a multivariate normal distribution (Searle et al.,1992). Thus, the multivariate normal distribution will be assumed in this thesis. Harville (1977) gives some of the advantages of using likelihood procedures in this setting:

> A maximum likelihood approach to the estimation of variance components has some attractive features. The maximum likelihood estimators are functions of every sufficient statistic and are consistent and asymptotically normal and efficient. Certain deficiencies of various other methods are not shared by maximum likelihood. In particular, the maximum likelihood approach is 'always' well-defined, even for the many useful generalizations of the ordinary ANOVA models, and, with maximum likelihood, nonnegativity constraints on the variance components or other constraints on the parameter space cause no conceptual difficulties.

On the other hand, complicated computational issues can arise when calculating likelihood estimators, since the solutions require solving nonlinear equations (Harville,1977). Iterative algorithms are necessary for finding such solutions and have been implemented with the use of modern computing software. Such computational algorithms and other issues related to likelihood estimation can be found in Harville (1977), Callanan and Harville (1991), and Searle et al. (1992).

1.2.2. Explicit Linear Likelihood Estimation

Due to the difficult computations necessary to solve the likelihood equations, there is an advantage to knowing when these estimators can be solved linearly and explicitly. In these cases, an iterative procedure is not necessary and the resulting estimates are easier to interpret and assess. This issue has been investigated by Rogers and Young (1977), Szatrowski (1980), and ElBassiouni (1983). All of these results pertain to the full case which involves the entire vector of fixed effects or the entire vector of variance components.

Rogers and Young (1977) identify conditions involving explicit linear maximum likelihood equation estimators for the entire vector of variance components. They examine when the inverse of the covariance matrix has linear structure. This allows the maximum likelihood equations to be solved linearly and explicitly.

Szatrowski (1980) finds conditions for the existence of explicit linear maximum likelihood equation estimators for the full case involving fixed effects and variance components. This approach involves obtaining models that have estimated generalized least squares estimators that correspond to solutions of the maximum likelihood equations. Under certain sufficient conditions, Szatrowski shows the estimated generalized least squares estimators for these models equal the least squares estimator. The least squares estimator satisfies the definition of an explicit linear maximum likelihood estimator.

ElBassiouni (1983) applies the results of Szatrowski (1980) to the restricted maximum likelihood procedure. Conditions are obtained under which the variance component vector has an explicit linear restricted maximum likelihood equation estimator.

The method of Szatrowski is of particular interest since it will be used in this study to extend the previous results.

1.2.3. Best Linear Unbiased Estimation

Best linear unbiased estimation is a concept which will be useful for obtaining conditions for the existence of explicit linear likelihood estimators. This type of estimation is defined by Puntanen and Styan (1989) and Seely(1996).

The relationship between best linear unbiased estimation and explicit maximum likelihood estimation can be explained for the fixed effects in a mixed effects linear model. Let this linear model be called the Y-Model. The best linear unbiased estimator for a given covariance matrix is the generalized least squares estimator (Searle et al.,1992). When the covariance matrix depends on an unknown variance component parameter that varies in some set, the generalized least squares estimator will not necessarily be the best over all possible parameter values. Under Zyskind's condition for the Y-Model (Zyskind,1967), for each value of the parameter, the associated generalized least squares estimators are equivalent and equal to the least squares estimator. In this case, the least squares estimator is the best linear unbiased estimator over all possible parameter values. Suppose the unknown variance component parameter is estimated using a solution to the maximum likelihood equations where the resulting estimate lies in the parameter set. Then Zyskind's condition can be used to show that the estimated generalized least squares estimator using the maximum likelihood equation estimator is equal to the least squares estimator. The least squares estimator is explicit, linear, and equivalent to the maximum likelihood equation estimator.

In order to apply similar results to variance components, it is necessary to obtain models to conduct quadratic estimation. Such models are formulated by Seely (1969,1971) and will be called linearized quadratic estimation models. A linearized quadratic estimation model can be defined for the maximum likelihood procedure and for the restricted maximum likelihood procedure. The response in such models involves quadratic forms of the original response. Generalizations of best linear unbiased estimation have been examined for these models by Seely and Zyskind (1969). In addition, Seely (1969) shows that Zyskind's condition in the linearized quadratic estimation model is equivalent to a quadratic subspace condition. Further discussion of least squares, generalized least squares, and best linear unbiased estimation is given by Puntanen and Styan (1989), Rao (1968), Seely (1996), and Birkes (1996).

1.3. Summary of Results

The approach of Szatrowski (1980) for the full case, which involves the entire vector of fixed effects or the entire vector of variance components, requires the use of the Y-Model and two linearized quadratic estimation models. These three models can be combined into a single underlying model in which Zyskind's condition can be investigated. The results derived for this underlying model will then be applied to the specific models of interest.

The results for the full case can be generalized to linear combinations involving the parameters. In this case, a full explicit linear estimator may not exist, but there may exist an explicit linear estimator for a linear combination of interest. This generalization is done using the underlying model by equating the linear combination involving the least squares estimator with the linear combination involving the generalized least squares estimator to obtain a generalized Zyskind's condition. This condition is applied to the particular models of interest to give results for the general case involving linear combinations of the fixed effects or linear combinations of the variance components in the maximum likelihood and restricted maximum likelihood procedures.

Another perspective relates to examining a subvector of the parameter vector. This perspective is useful for understanding and for checking the existence of an explicit linear estimator in specific examples. This is done using the underlying model by equating the subvector involving the least squares estimator with the subvector involving the generalized least squares estimator to obtain a generalized Zyskind's condition. The condition is again applied to the particular models of interest to obtain the associated results for estimating a subvector of the fixed effects vector or a subvector of the variance components vector in the maximum likelihood and restricted maximum likelihood procedures.

The existence of explicit linear estimators will be demonstrated for the full and general cases in mixed linear classification models. Specific examples will be examined as well as classes of examples that meet certain design and model conditions. A search of 3-way models under various designs is also presented.

The conditions for the existence of explicit linear estimators in the full and general cases will be applied to uniformly minimum variance unbiased estimation. The full case will be presented with respect to results from Seely(1971,1977) that prove the existence of a complete sufficient statistic for a family of normal distributions under the conditions. The general case will be presented to show that, under the conditions, the explicit linear estimator has uniformly minimum variance for all unbiased estimators in certain cases in the maximum likelihood and restricted maximum likelihood procedures. In addition, it is shown that an exact form can be obtained for the covariance of an explicit linear estimator.

The conditions for the existence of explicit linear estimators in the full and general cases will also be applied to data. An iterative procedure is defined and demonstrated through a data example using PROC MIXED in SAS. It is also demonstrated how an iterative procedure can be used to check the conditions. Data examples also demonstrate that computing time can be saved when accounting for explicit linear estimators. This savings is explained for the iterative procedure and for profile likelihood confidence intervals. A data example will be given to illustrate the savings in computing time.

This thesis will present notation and definitions pertaining to linear transformations in chapter 2. Chapter 3 will be used to define the models of interest, as well as results for these models. Chapter 4 gives the previous results of Szatrowski (1980) and ElBassiouni (1983), along with clarifications. Extensions involving the general cases will be given in chapters 5 and 6. Applications of the results pertaining to uniformly minimum variance unbiased estimation will be given in chapter 7, while chapter 8 discusses the application of the results to data. Chapter 9 provides a conclusion to the thesis. Appendices C and D have been included to help the reader. Appendix C gives a summary of the models and related theorems while Appendix D gives section numbers for common symbols and abbreviations.

2. Linear Transformations

The results of this study require knowledge of linear transformations. This chapter could be a review for a reader with a background in linear algebra. However, particular notation, definitions, and properties will be presented that provide an essential framework for later chapters. This chapter starts by presenting basic terminology and then gives particular results that will be useful.

2.1. Basic Terminology

The following definitions provide a foundation to build on. This study will focus on finite dimensional inner product spaces defined over the reals ($\mathcal{R}$). An inner product space is a vector space with an inner product. The following definitions can be found in Halmos (1958) and Seely (1996).

Definitions: Vector Space - $\mathcal{V}$ is a vector space provided that $\forall\ \alpha, \beta \in \mathcal{R},\ a, b, c \in \mathcal{V}$

i) $a + b = b + a$ ii) $a + (b + c) = (a + b) + c$

iii) $\exists$ unique $0 \ni a + 0 = a$ iv) $\exists$ unique $-a \ni a + (-a) = 0$

v) $\alpha(\beta a) = (\alpha\beta)a$ vi) $\exists$ unique $1 \ni 1a = a$

vii) $\alpha(a + b) = \alpha a + \alpha b$ viii) $(\alpha + \beta)a = \alpha a + \beta a$.

Subspace - A non-empty subset $\mathcal{U}$ of $\mathcal{V}$ is a subspace provided that $\forall\ \alpha, \beta \in \mathcal{R},\ a, b \in \mathcal{U},\ \alpha a + \beta b \in \mathcal{U}$.

Linear Transformation - If $\mathcal{W}$ and $\mathcal{V}$ are vector spaces, then the function $A : \mathcal{V} \rightarrow \mathcal{W}$ is a linear transformation from $\mathcal{V}$ into $\mathcal{W}$ provided that $A(\alpha a + \beta b) = \alpha A(a) + \beta A(b)\ \ \forall\ \alpha, \beta \in \mathcal{R},\ a, b \in \mathcal{V}$.

Linear Functional - l is a linear functional on $\mathcal{V}$ provided that $l : \mathcal{V} \rightarrow \mathcal{R}$ is a linear transformation.

Dual Space - The dual space of a vector space $\mathcal{V}$ is the vector space $\mathcal{V}^* = \{l : l$ a linear functional on $\mathcal{V}\}$.

Adjoint - The adjoint of the linear transformation A is given by $A^* : \mathcal{V}^* \rightarrow \mathcal{W}^*$ defined by $l(A(v)) = A^*(l(v))\ \ \forall\ v \in \mathcal{V},\ l \in \mathcal{W}^*$.

Self-Adjoint - A linear transformation $B : \mathcal{V} \rightarrow \mathcal{V}\ \ni\ B^* = B$.

Inner Product - For $\alpha, \beta \in \mathcal{R}$ and $a, b, c \in \mathcal{W}$, denote the inner product of a and b by $< a, b >$ where

i) $< a, b > \ = \ < b, a >$ ii) $< \alpha a + \beta b, c > \ = \alpha < a, c > + \beta < b, c >$

iii) $< a, a > \ \geq 0$ iv) $< a, a > \ = 0 \Leftrightarrow a = 0$.

Range Space - For the linear transformation $A : \mathcal{V} \to \mathcal{W}$, it is the subspace of $\mathcal{W}$ denoted by $\underline{R}(A) = A(\mathcal{V})$.

Rank - The dimension of the range space where $\dim(\underline{R}(A)) = \underline{r}(A)$.

Null Space - For the linear transformation $A : \mathcal{V} \to \mathcal{W}$, it is the subspace of $\mathcal{V}$ given by $\underline{N}(A) = \{z \in \mathcal{V} \mid A(z) = 0\}$.

Nullity - The dimension of the null space where $\dim(\underline{N}(A)) = \underline{n}(A)$.

Orthogonal Complement - For a subspace $\mathcal{U}$ of $\mathcal{W}$ in an inner product space, it is the subspace of $\mathcal{W}$ given by $\mathcal{U}^{\perp} = \{a \in \mathcal{W},\ b \in \mathcal{M} \mid < a, b > \ = 0\}$.

Non-Negative Definite (NND) - The linear transformation B is NND or $B \geq 0$ if B is self-adjoint and $< B(v), v > \ \geq 0 \quad \forall\, v \in \mathcal{V}$.

Positive Definite (PD) - The linear transformation B is PD or $B > 0$ if B is self-adjoint and $< B(v), v > \ > 0 \quad \forall\, v \in \mathcal{V}$.

Consider the inner product space $(\mathcal{W}, < , >)$. A norm and a metric can be defined by $||w|| = \ < w, w >^{\frac{1}{2}}$ and $d(w_1, w_2) = ||w_1 - w_2||$, respectively. For notational simplicity, a linear transformation A operating on an element v of a vector space will be denoted by Av instead of $A(v)$ as above. This should not be confused with matrix multiplication and should be clear from the context.

Any linear transformation can be expressed as a matrix. Consider a linear transformation $A : \mathcal{U} \to \mathcal{W}$ where $\{u_1, ..., u_p\}$ and $\{w_1, ..., w_n\}$ are bases for $\mathcal{U}$ and $\mathcal{W}$ respectively. Then for $j = 1, ..., p$ $Au_j = \sum_{i=1}^{n} m_{ij} w_i$. The matrix is a rectangular array of the np numbers m_{ij} given by $M_{n \times p} = \{m_{ij}\}$ which has column $\underline{c}_j = [m_{1j}, ..., m_{nj}]'$ and row $\underline{r}_i' = [m_{i1}, ..., m_{ip}]$ (Marcus and Minc,1965). The matrix M is a linear transformation from $\mathcal{R}^p$ into $\mathcal{R}^n$. The following definitions are given to summarize the notation which will be used for matrices.

Definitions: Matrix - Let $M : \mathcal{R}^p \to \mathcal{R}^n$ be denoted by $M_{n \times p} = \{m_{ij}\}_{n \times p} = M'_{p \times n}$.

Vector - Let $\underline{c} : \mathcal{R}^1 \to \mathcal{R}^n$ be denoted by $\underline{c}_{n \times 1} = \{c_i\}_{n \times 1}$.

Diagonal Matrix - The matrix $D_{n \times n} = \mathrm{diag}(\underline{d}_{n \times 1}) = \mathrm{diag}(\{d_i\}_{n \times 1}) = \begin{cases} d_i & i = j \\ 0 & i \neq j \end{cases}$.

Indexing and sizes will be suppressed when these values can easily be determined from the context or when they are not important. Notationally, A could represent a linear transformation or a matrix, but its representation should be clear from the context. In addition, let $\mathcal{S}_n$ = set of symmetric $n \times n$ matrices and $\mathcal{M}_{n \times m}$ = set of $n \times m$ matrices. Specific inner product spaces that will be considered include $(\mathcal{R}^n, < \underline{a}_{n\times 1}, \underline{b}_{n \times 1} > = \underline{a}'\underline{b})$ and $(\mathcal{S}_n, < A_{n\times n}, B_{n \times n} > = \text{tr}(AB))$.

2.2. Dual Spaces

The following propositions give some properties of the dual space. These propositions indicate that dual space $\mathcal{W}^*$ is isomorphic to $\mathcal{W}$ when $\mathcal{W}$ has an inner product. Thus, the dual space will be of little concern, since the main interest is in the real inner product space $(\mathcal{W}, < , >)$.

(Halmos,1958)

Proposition: $A^* : \mathcal{W}^* \rightarrow \mathcal{V}^*$ is a linear transformation.

proof: i) Let $l_{w^\circ} \in \mathcal{W}^*$ and consider $l_{v^\circ}(v) = l_{w^\circ}(A(v)) \in \mathcal{V}^*$. Then $\forall\ v \in \mathcal{V}$

$l_{v^\circ}(v) = l_{w^\circ}(A(v)) = A^*\, l_{w^\circ}(v)$ by definition of adjoint

$\Rightarrow l_{v^\circ} = A^*(\,l_{w^\circ})$ since above holds $\forall\ v \in \mathcal{V}$.

Thus, $\forall\, w^\circ \in \mathcal{W}\;\; A^*(\,l_{w^\circ}) = l_{v^\circ}$ for some $v^\circ \in \mathcal{V} \;\; \Rightarrow A^* : \mathcal{W}^* \rightarrow \mathcal{V}^*$.

ii) Let $w^* = \alpha_1 w_1^* + \alpha_2 w_2^*$ where $w_1^*, w_2^* \in \mathcal{W}^*$. Then $\forall\ v \in \mathcal{V}$

$$A^*(w^*(v)) \overset{(1)}{=} w^*(A(v)) = (\alpha_1 w_1^* + \alpha_2 w_2^*)(A(v)) \overset{(2)}{=} \alpha_1 w_1^*(A(v)) + \alpha_2 w_2^*(A(v))$$
$$\overset{(1)}{=} \alpha_1 A^*(w_1^*(v)) + \alpha_2 A^*(w_2^*(v)) \overset{(2)}{=} (\alpha_1 A^*(w_1^*) + \alpha_2 A^*(w_2^*))(v)$$

$\Rightarrow A^*(w^*) = \alpha_1 A^*(w_1^*) + \alpha_2 A^*(w_2^*)$ as the above holds $\forall\ v \in \mathcal{V}$

where (1) follows from definition of adjoint and (2) follows from linearity properties. ∎

(Halmos,1958)

Proposition: $\dim \mathcal{W}^* = \dim \mathcal{W}$.

proof: Suppose $\dim \mathcal{W} = n$ and $\{w_1, ..., w_n\}$ is a basis for $\mathcal{W}$.

i) Define $l : \mathcal{W} \rightarrow \mathcal{R}$ by $l(w) = \sum_{i=1}^{n} \alpha_i \beta_i$ where $w = \sum_{i=1}^{n} \alpha_i w_i \Rightarrow l(w_i) = \beta_i \Rightarrow l(w) = \sum_{i=1}^{n} \alpha_i l(w_i)$ $(*)$.

$\forall\, w, v \in \mathcal{W} \Rightarrow \gamma_1 w + \gamma_2 v \in \mathcal{W}$, so $\gamma_1 w + \gamma_2 v = (\sum_{j=1}^{n} \gamma_1 \alpha_{1j} w_j) + (\sum_{j=1}^{n} \gamma_2 \alpha_{2j} w_j) = \sum_{j=1}^{n} (\gamma_1 \alpha_{1j} + \gamma_2 \alpha_{2j}) w_j$

$$\Rightarrow l(\gamma_1 w + \gamma_2 v) = l(\sum_{j=1}^{n} (\gamma_1 \alpha_{1j} + \gamma_2 \alpha_{2j}) w_j) \overset{(*)}{=} \sum_{j=1}^{n} (\gamma_1 \alpha_{1j} + \gamma_2 \alpha_{2j}) l(w_j)$$
$$= \gamma_1 \sum_{j=1}^{n} \alpha_{1j} l(w_j) + \gamma_2 \sum_{j=1}^{n} \alpha_{2j} l(w_j) \overset{(*)}{=} \gamma_1\, l(\sum_{j=1}^{n} \alpha_{1j} w_j) + \gamma_2 l(\sum_{j=1}^{n} \alpha_{2j} w_j) = \gamma_1 l(w) + \gamma_2 l(v)$$

$\Rightarrow l \in \mathcal{W}^*$.

ii) Suppose $l_i\ i = 1, ..., n \in \mathcal{W}^*$ where $l_i(w_j) = \delta_{ij}$. Then $\forall\ w \in \mathcal{W}\ ,\ l \in \mathcal{W}^*$

$$l_i(w) = l_i(\sum_{j=1}^{n} \alpha_j w_j) = \sum_{j=1}^{n} \alpha_j l_i(w_j) = \sum_{j=1}^{n} \alpha_j \delta_{ij} = \alpha_i \Rightarrow l(w) \overset{(*)}{=} \sum_{i=1}^{n} \alpha_i \beta_i = \sum_{i=1}^{n} \beta_i\, l_i(w)$$

$\Rightarrow l = \sum_{i=1}^{n} \beta_i\, l_i$ since the above holds $\forall\ w \in \mathcal{W} \Rightarrow$ sp$\{ l_1, \dots , l_n\} = \mathcal{W}^*$ as the above holds $\forall\ l \in \mathcal{W}^*$.

iii) $\sum_{i=1}^{n} \beta_i\, l_i = 0 \Rightarrow \sum_{i=1}^{n} \beta_i\, l_i(w) = 0\ \forall\ w \in \mathcal{W}\ \Rightarrow \sum_{i=1}^{n} \beta_i\, l_i(w_j) = 0\ \ \forall\ j = 1, ..., n$

$\Rightarrow \sum_{i=1}^{n} \beta_i \delta_{ij} = 0\ \ \forall\ j = 1, ..., n\ \Rightarrow \beta_j = 0\ \forall\ j = 1, ..., n\ \Rightarrow \{l_1, \dots , l_n\}$ are linearly independent.

$\therefore \{l_1, \dots , l_n\}$ is a basis for $\mathcal{W}^* \Rightarrow \dim \mathcal{W}^* = n = \dim \mathcal{W}$. ∎

Proposition: Suppose $(\mathcal{W},\ <\ ,\ >\ )$ is a real inner product space. Then $\mathcal{W}^* = \mathcal{W}$.

proof: Then $\forall\ w \in \mathcal{W}$, define $l \in \mathcal{W}^*$ by $l(w) =\ < w, v >\ \ \forall\ v \in \mathcal{W}$.

Consider the linear operator $\Phi : \mathcal{W} \rightarrow \mathcal{W}^*$ defined by $\Phi(w) = l$. Note

$\Phi(w) = 0 \Rightarrow\ < w, v >\ = 0\ \forall\ v \in \mathcal{W}\ \Rightarrow\ < w, w >\ = 0 \Rightarrow w\ = 0 \Rightarrow \Phi$ is 1-1 (isomorphism)

$\Rightarrow \mathcal{W}^*$ is isomorphic to $\mathcal{W}$ since $\exists$ an isomorphism Φ and $\dim \mathcal{W}^* = \dim \mathcal{W}$ by above proposition. ∎

2.3. Subspaces

A number of relationships will be presented concerning subspaces. Many of these results will be used in later sections. Consider a finite dimensional inner product space given by $(\mathcal{W},\ <\ ,\ >\ )$.

(Halmos,1958) (Seely,1996)

Proposition: Let $\mathcal{T}$ and $\mathcal{U}$ be subspaces of $\mathcal{W}$. Then

i) $\mathcal{T} \subset \mathcal{U}$ and $\dim \mathcal{T} = \dim \mathcal{U}\ \Leftrightarrow \mathcal{T} = \mathcal{U}$

ii) $\mathcal{U} + \mathcal{U}^\perp = \mathcal{W}$

iii) $(\mathcal{T} + \mathcal{U})^\perp = \mathcal{T}^\perp \cap \mathcal{U}^\perp$

iv) $\dim(\mathcal{T} + \mathcal{U}) = \dim \mathcal{T} + \dim \mathcal{U} - \dim(\mathcal{T} \cap \mathcal{U})$

v) $\mathcal{U}^{\perp\perp} = \mathcal{U}$.

proof: i) Note $\mathcal{T} = \mathcal{U} \Leftrightarrow \mathcal{T} \subset \mathcal{U}$ and $\nexists\, a \in \mathcal{U} \ni a \notin \mathcal{T} \Leftrightarrow \mathcal{T} \subset \mathcal{U}$ and $\dim \mathcal{T} = \dim \mathcal{U}$.

ii) Let $S = \{u_1,\ ..., u_n\}$ be an orthonormal basis for $\mathcal{U}$ and $\forall\ w \in \mathcal{W}$ define $x = \sum_{i=1}^{m} < w, u_i > u_i \in \mathcal{U}$.

By the properties of the inner product and the orthonormal basis S,

$$< w - x, x >\ =\ < w, x >\ -\ < x, x >\ = (\sum_{i=1}^{m} < w, u_i >)(< w, u_i >\ -\ < x, u_i >)$$

$$= (\sum_{i=1}^{m} < w, u_i >)(< w, u_i >\ -\ < w, u_i >\ < u_i, u_i >) = 0 \quad \text{as } < u_i, u_i >\ = 1.$$

Thus, $w - x \in \mathcal{U}^\perp \Rightarrow w = x + (w - x) \in \mathcal{U} + \mathcal{U}^\perp \Rightarrow \mathcal{U} + \mathcal{U}^\perp = \mathcal{W}$ as the above holds $w \in \mathcal{W}$.

iii) $w \in (\mathcal{T}+\mathcal{U})^{\perp} \Leftrightarrow \; < w, v > \; = 0 \quad \forall \, v \in \mathcal{T}+\mathcal{U} \Leftrightarrow \; < w, x+y > \; = 0 \quad \forall \, x \in \mathcal{T}, y \in \mathcal{U}$

$\Leftrightarrow \; < w, x > \; = 0$ and $< w, y > \; = 0 \quad \forall \, x \in \mathcal{T}, y \in \mathcal{U} \Leftrightarrow w \in \mathcal{T}^{\perp} \cap \mathcal{U}^{\perp}$.

iv) Let $S_{\mathcal{T} \cap \mathcal{U}} = \{u_1, ..., u_k\}$ be a basis for $\mathcal{T} \cap \mathcal{U}$. Choose $v_1, ..., v_m \ni$

$S_{\mathcal{T}} = \{u_1, ..., u_k, v_1, ..., v_m\}$ is a basis for $\mathcal{T}$ and choose $w_1, ..., w_n \ni$

$S_{\mathcal{U}} = \{u_1, ..., u_k, w_1, ..., w_n\}$ is a basis for $\mathcal{U}$.

For purposes of contradiction $(*)$, suppose the elements of $S_{\mathcal{T}} \cup S_{\mathcal{U}}$ are linearly dependent

$\Rightarrow \exists$ linear combinations $\ni \sum_{i=1}^{m} \alpha_i v_i = \sum_{j=1}^{n} \beta_j w_j \in \mathcal{T} \cap \mathcal{U}$

$\Rightarrow \sum_{i=1}^{m} \alpha_i v_i = \sum_{j=1}^{n} \beta_j w_j = \sum_{l=1}^{k} \delta_l u_l$ for some linear combination of elements in $S_{\mathcal{T} \cap \mathcal{U}}$

$\Rightarrow S_{\mathcal{T}}, S_{\mathcal{U}}$ cannot be bases by definition since their elements are not independent $(*)$.

Thus, $S_{\mathcal{T}} \cup S_{\mathcal{U}}$ is a basis for $\mathcal{T}+\mathcal{U}$ as it is an independent spanning set for $\mathcal{T}+\mathcal{U}$. Also,

(1) $\dim(\mathcal{T}+\mathcal{U}) = k+m+n$ (2) $\dim \mathcal{T} + \dim \mathcal{U} - \dim(\mathcal{T} \cap \mathcal{U}) = k+m+k+n-k = k+m+n$

$\Rightarrow$ the result holds as (1) and (2) are equal.

v) By definition, $\mathcal{U}^{\perp\perp} = \{a \in \mathcal{W}, \; b \in \mathcal{U}^{\perp} \mid \; < a, b > \; = 0\} \supset \mathcal{U} \quad (*)$.

(1) Note $\dim \mathcal{W} = \dim(\mathcal{U}+\mathcal{U}^{\perp}) = \dim \mathcal{U} + \dim(\mathcal{U}^{\perp}) - \dim(\mathcal{U} \cap \mathcal{U}^{\perp})$ by ii) and iv)

$= \dim \mathcal{U} + \dim(\mathcal{U}^{\perp}) - \dim(\mathcal{W}^{\perp}) = \dim \mathcal{U} + \dim(\mathcal{U}^{\perp})$ by iii).

(2) Also, $\dim \mathcal{W} = \dim(\mathcal{U}^{\perp} + \mathcal{U}^{\perp\perp}) = \dim \mathcal{U}^{\perp} + \dim(\mathcal{U}^{\perp\perp}) - \dim(\mathcal{U}^{\perp} \cap \mathcal{U}^{\perp\perp})$ by ii) and iv)

$= \dim \mathcal{U}^{\perp} + \dim(\mathcal{U}^{\perp\perp}) - \dim(\mathcal{W}^{\perp}) = \dim \mathcal{U}^{\perp} + \dim(\mathcal{U}^{\perp\perp})$ by iii).

Thus, $\mathcal{U}^{\perp\perp} = \mathcal{U}$ by $(*)$, (1), (2). ■

By definition, range and null spaces are subspaces. The next proposition gives results for range spaces, null spaces, and ranks.

(Seely, 1996)

Proposition: Consider conformable linear transformations A and B. Then

i) $\underline{R}(A, B) = \underline{R}(A) + \underline{R}(B)$

ii) $\underline{r}(A, B) = \underline{r}(A) + \underline{r}(B) - \dim(\underline{R}(A) \cap \underline{R}(B))$

iii) $\underline{R}(A)^{\perp} = \underline{N}(A^*)$ $\qquad$ $\underline{N}(A)^{\perp} = \underline{R}(A^*)$

iv) $\underline{R}(AB) \subset \underline{R}(A)$ $\qquad$ $\underline{N}(B) \subset \underline{N}(AB)$

v) $\underline{r}(AB) = \underline{r}(B) - \dim(\underline{R}(B) \cap \underline{N}(A))$

vi) $\underline{R}(B^*B) = \underline{R}(B^*)$ $\qquad$ $\underline{N}(B^*B) = \underline{N}(B)$.

vii) $\underline{R}(A) \subset \underline{R}(B) \Leftrightarrow \underline{r}(A, B) = \underline{r}(B)$.

proof: i) Let $\mathcal{T}, \mathcal{U}_1, \mathcal{U}_2$ be subspaces and $A : \mathcal{U}_1 \to \mathcal{T}$ and $B : \mathcal{U}_2 \to \mathcal{T}$. Then

$\underline{R}(A,B) = \{[A \; B][u_1 u_2] \ni u_1 \in \mathcal{U}_1, \; u_2 \in \mathcal{U}_2\} = \{Au_1 + Bu_2 \mid u_1 \in \mathcal{U}_1, \; u_2 \in \mathcal{U}_2\}$

$= \{Au_1 \mid u_1 \in \mathcal{U}_1\} + \{Bu_2 \mid u_2 \in \mathcal{U}_2\} = \underline{R}(A) + \underline{R}(B)$.

ii) Follows from iv) in proposition above.

iii) Let A be defined as in the proof of i). Then

(1) $\underline{R}(A)^{\perp} = \{Au_1 | u_1 \in \mathcal{U}_1\}^{\perp} = \{w \in \mathcal{T} \mid < Au_1, w > \ = 0 \ \forall \ u_1 \in \mathcal{U}_1\}$

$= \{w \in \mathcal{T} \mid < u_1, A^*w > \ = 0 \ \forall \ u_1 \in \mathcal{U}_1 \}$

$= \{w \in \mathcal{T} \mid A^*w = 0 \} = \underline{N}(A^*)$ as above holds $\forall \ u_1 \in \mathcal{U}_1$.

(2) By (1) and v) in above prop, $\underline{N}(A) = \underline{R}(A^*)^{\perp} \Rightarrow \underline{N}(A)^{\perp} = \underline{R}(A^*)^{\perp\perp} = \underline{R}(A^*)$.

iv) Let $\mathcal{T}$, $\mathcal{U}$, $\mathcal{U}_2$ be subspaces and $A : \mathcal{U}_1 \to \mathcal{T}$ and $B : \mathcal{U}_2 \to \mathcal{U}_1$. Then

(1) $\underline{R}(AB) = \{ABw \mid w \in \mathcal{U}_2\} = \{A(Bw) \mid Bw \in \mathcal{U}_1\} \subset \{Av \mid v \in \mathcal{U}_1\} = \underline{R}(A)$.

(2) Let $t \in \underline{N}(B) \Rightarrow Bt = 0 \Rightarrow ABt = 0 \Rightarrow t \in \underline{N}(AB) \Rightarrow \underline{N}(B) \subset \underline{N}(AB)$.

v) Let T be a linear transformation $\ni T : \underline{R}(B) \to \mathcal{W}$ defined by $Tv = Av \quad \forall \ v \in \underline{R}(B)$. Then

$\underline{r}(B) = \underline{r}(T) + \underline{n}(T) = \underline{r}(AB) + \dim(\underline{R}(B) \cap \underline{N}(A))$ and the result follows.

vi) (1) $\underline{R}(B^*B) \subset \underline{R}(B^*)$ by iv) and by v). Also,

$\underline{r}(B^*B) = \underline{r}(B) - \dim(\underline{R}(B) \cap \underline{N}(B^*)) = \underline{r}(B) - \dim(\underline{R}(B) \cap \underline{R}(B)^{\perp}) = \underline{r}(B)$ by iii).

(2) By (1), $\underline{R}(B^*B) = \underline{R}(B^*) \Leftrightarrow \underline{R}(B^*B)^{\perp} = \underline{R}(B^*)^{\perp} \Leftrightarrow \underline{N}(B^*B) = \underline{N}(B)$ by iii).

vii) $\underline{r}(A, B) = \underline{r}(B) \Leftrightarrow \underline{r}(A) + \underline{r}(B) - \dim(\underline{R}(A) \cap \underline{R}(B)) = \underline{r}(B)$

$\underline{r}(A) = \dim(\underline{R}(A) \cap \underline{R}(B)) \Leftrightarrow \underline{R}(A) = \underline{R}(A) \cap \underline{R}(B)$ as $\underline{R}(A) \cap \underline{R}(B) \subset \underline{R}(A) \Leftrightarrow \underline{R}(A) \subset \underline{R}(B)$. ▪

Proposition: Let $\mathcal{T} = \{T_1, ..., T_t\}$ and U, W be conformable linear transformations. Then $\underline{R}(TU) \subset \underline{R}(W) \ \forall \ T \in \mathcal{T} \Leftrightarrow \underline{R}(TU) \subset \underline{R}(W) \ \forall \ T \in \text{sp}\mathcal{T}$.

proof: i) $\underline{R}(TU) \subset \underline{R}(W) \ \forall \ T \in \text{sp}\mathcal{T} \Rightarrow \underline{R}(TU) \subset \underline{R}(W) \ \forall \ T \in \mathcal{T}$ since $\mathcal{T} \subset \text{sp}\mathcal{T}$.

ii) $\underline{R}(TU) \subset \underline{R}(W) \ \forall \ T \in \mathcal{T} \Rightarrow \underline{R}(T_iU) \subset \underline{R}(W) \quad i = 1, ..., t$

$\Rightarrow T_i[\underline{R}(U)] \subset \underline{R}(W) \quad i = 1, ..., t$

$\Rightarrow \sum_{i=1}^{t} a_i T_i[\underline{R}(U)] \subset \underline{R}(W) \quad \forall \ a_i \in \mathcal{R}$ since $\underline{R}(W)$ is a subspace

$\Rightarrow T[\underline{R}(U)] \subset \underline{R}(W) \quad \forall \ a_i \in \mathcal{R}, T = \sum_{i=1}^{t} a_i T_i \in \text{sp}\mathcal{T}$

$\Rightarrow \underline{R}(TU) \subset \underline{R}(W) \quad \forall \ T \in \text{sp}\mathcal{T}$. ▪

2.4. Inverses

Under certain properties, a linear transformation A has an inverse (A^{-1}) or is invertible. These conditions and a useful proposition are presented below.

Definition: Invertible - A linear transformation $A : \mathcal{V} \to \mathcal{W}$ is invertible providing

(1) $Av_1 = Av_2 \Rightarrow v_1 = v_2$ (1-1) and (2) $\underline{R}(A) = \mathcal{W}$ (onto).

(Halmos,1958) (Marcus and Minc,1965)

Proposition: The linear transformation $A : \mathcal{V} \to \mathcal{V}$ is invertible if and only if $Av = 0 \Rightarrow v = 0$.

proof: i) Suppose A is invertible. Then $Av = 0 = A0 \Rightarrow v = 0$.

ii) Suppose $Av = 0 \Rightarrow v = 0$. Then $Av_1 = Av_2 \Rightarrow A(v_1 - v_2) = 0 \Rightarrow v_1 - v_2 = 0 \Rightarrow v_1 = v_2$.

In addition, suppose $\{b_1, ..., b_n\}$ is a basis for $\mathcal{V}$. Then

$$\sum_{i=1}^{n}\alpha_i A b_i = 0 \;\Rightarrow A(\sum_{i=1}^{n}\alpha_i b_i) = 0 \Rightarrow \sum_{i=1}^{n}\alpha_i b_i = 0 \Rightarrow \alpha_i = 0 \;\; i = 1, ..., n$$

$\Rightarrow \{Ab_1, ..., Ab_n\}$ is also basis for $\mathcal{V} \Rightarrow \underline{R}(A) = \mathcal{V}$. ∎

There will be interest in calculating the inverse for partitioned matrices containing linear transformations. This special setting will be described in a later chapter, but the result is given here. The inverse formulas can be verified by left and right multiplying the transformation and its inverse to obtain the identity transformation.

(Christensen,1996)

Inverse Formulas: Assuming all linear transformations are conformable, then

i) $[A + BCD]^{-1} = A^{-1} - A^{-1}B[C^{-1} + DA^{-1}B]^{-1}DA^{-1}$

ii) $\begin{bmatrix} A & B \\ B^* & D \end{bmatrix}^{-1} = \begin{bmatrix} A^{-1} + GE^{-1}G^* & -G^*E^{-1} \\ -E^{-1}G^* & E^{-1} \end{bmatrix} \qquad E = D - B^*A^{-1}B \qquad G = A^{-1}B.$

2.5. Projection Operators and Generalized Inverses

Projection operators play a fundamental role in later results. These operators have special properties, as indicated by their definitions. Another special transformation is the generalized inverse or g-inverse. These inverses are useful for characterizing projections and have special properties (Seely,1996).

Definitions: Projection Operator (PO) - P is a PO on $\underline{R}(P)$ along $\underline{N}(P) \;\Leftrightarrow\; P^2 = P$.

Orthogonal Projection Operatior (OPO) - P_A is an OPO on $\underline{R}(A) \Leftrightarrow \underline{R}(P_A) = \underline{R}(A),\; P_A = P_A^* = P_A^2$.

G-Inverse (A^-) - A^- is defined by the relation $AA^-A = A$.

Moore-Penrose Inverse (A^+) - A^+ is defined by the properties

1) $AA^+A = A$ 2) $A^+AA^+ = A^+$ 3) $(AA^+)^* = AA^+$ 4) $(A^+A)^* = A^+A$.

Note that P_A is used to represent an OPO on $\underline{R}(A)$ while the range and null space need to be specified for a PO. The next proposition establishes an alternative definition of projection operators. The second proposition uses the alternative definition to show that projections are unique linear operators. For subspaces $\mathcal{U}$ and $\mathcal{V}$, the direct sum ($\oplus$) is defined by $\mathcal{U} \oplus \mathcal{V} = \mathcal{U} + \mathcal{V} \ni \mathcal{U} \cap \mathcal{V} = \{0\}$ (Seely,1996).

(Halmos,1958)
Proposition: Let $P : \mathcal{V} \to \mathcal{V}$. Then $P^2 = P \Leftrightarrow \mathcal{V} = \underline{R}(P) \oplus \underline{N}(P)$ and $v \in \mathcal{V}$ can uniquely be expressed as $v = u + w$ where $Pw = 0$ and $Pv = u$.

proof: i) Suppose $P^2 = P$. (1) Let $v \in \mathcal{V}$. Then $v = Pv + (v - Pv) \in \underline{R}(P) + \underline{N}(P)$.
(2) Suppose $w \in \underline{R}(P) \cap \underline{N}(P)$. Then $w = Pv$ for some $v \in \mathcal{V}$ and $Pw = 0$. Thus,
$0 = Pw = PPv = Pv = w \Rightarrow \underline{R}(P) \cap \underline{N}(P) = \{0\}$. $\therefore$ by (1) and (2), $\mathcal{V} = \underline{R}(P) \oplus \underline{N}(P)$.
(3) By i), $v = u + w \in \underline{R}(P) + \underline{N}(P)$ where $u \in \underline{R}(P)$ and $w \in \underline{N}(P)$.
Let $u_i \in \underline{R}(P)$, $w_i \in \underline{N}(P)$ $i = 1, 2$ and assume $v = u_1 + w_1 = u_2 + w_2 \Rightarrow u_1 - u_2 = w_2 - w_1$
$\Rightarrow u_1 = u_2$, $w_1 = w_2$ since $u_1 - u_2 \in \underline{R}(P)$, $w_1 - w_2 \in \underline{N}(P)$, and $\underline{R}(P) \cap \underline{N}(P) = \{0\}$.
$\Rightarrow v \in \mathcal{V}$ can uniquely be expressed as $v = u + w$.
(4) Consider the unique expression in (3) given by $v = u + w$. Let $u = Pz$ for some z and note $Pw = 0$.
Then $Pv = Pu + Pw = PPz = Pz = u$ by (3).
ii) To show $P^2 = P$. By hypothesis, $v \in \mathcal{V}$ can uniquely be expressed as $v = u + w$ where $u \in \underline{R}(P)$, $w \in \underline{N}(P)$, and $Pv = u$. Thus, $P^2 v = PPv = Pu = u = Pv \Rightarrow P^2 = P$ as the above holds $\forall\, v \in \mathcal{V}$. ■

Proposition: i) P is a linear transformation.
ii) If $P_1^2 = P_1$, $P_2^2 = P_2$, $\underline{R}(P_1) = \underline{R}(P_2)$, and $\underline{N}(P_1) = \underline{N}(P_2)$, then $P_1 = P_2$.

proof: i) Let $v_i \in \mathcal{V}$ $i = 1, 2$. By the above proposition, $v_i = u_i + w_i$ where $u_i \in \underline{R}(P)$, $w_i \in \underline{N}(P)$, and $Pv_i = u_i$ $i = 1, 2$. Then $\alpha_1 v_1 + \alpha_2 v_2 = \alpha_1 u_1 + \alpha_2 u_2 + \alpha_1 w_1 + \alpha_2 w_2$ where
$\alpha_1 u_1 + \alpha_2 u_2 \in \underline{R}(P)$ and $\alpha_1 w_1 + \alpha_2 w_2 \in \underline{N}(P)$. By the above proposition,
$P(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 u_1 + \alpha_2 u_2 = \alpha_1 P v_1 + \alpha_2 P v_2$.
ii) By the above proposition, $\forall\, v \in \mathcal{V}$, $v = u + w$ where $u \in \underline{R}(P_1) = \underline{R}(P_2)$ and $w \in \underline{N}(P_1) = \underline{N}(P_2)$
$\Rightarrow P_1 v = u = P_2 v \Rightarrow P_1 = P_2$ as above holds $\forall\, v \in \mathcal{V}$. ■

A g-inverse for A may not be unique while the Moore-Penrose inverse for A is unique. The choice of a g-inverse in later applications will depend upon the context. Usually, a g-inverse will be used unless the specific properties of a Moore-Penrose inverse are needed. The following propositions demonstrate properties of A^- and A^+. The next six results are given in Seely (1996).

Proposition: For any linear transformation A, there exists A^{-}.

proof: Let P be a PO on $\underline{R}(A) \Rightarrow \exists\, G \ni AG = P \Rightarrow AGA = PA = A \Rightarrow G = A^{-}$. ■

Proposition: A g-inverse A^{-} satisfies

i) $(A^{*})^{-} = (A^{-})^{*}$ ii) If A is invertible, then $A^{-} = A^{-1}$

iii) AA^{-} is a PO on $\underline{R}(A)$ iv) $A^{-}A$ is a PO along $\underline{N}(A)$.

proof: i) By definition of A^{-}, $AA^{-}A = A \Rightarrow A^{*}(A^{-})^{*}A^{*} = A^{*} \Rightarrow (A^{-})^{*} = (A^{*})^{-}$.

ii) By definition of A^{-}, $AA^{-}A = A \Rightarrow A^{-1}AA^{-}AA^{-1} = A^{-1}AA^{-1} \Rightarrow A^{-} = A^{-1}$.

iii) $(AA^{-})(AA^{-}) = AA^{-}$ by definition of A^{-} and

$\underline{R}(A) = \underline{R}(AA^{-}A) \subset \underline{R}(AA^{-}) \subset \underline{R}(A) \Rightarrow \underline{R}(A) = \underline{R}(AA^{-})$.

iv) $(A^{-}A)(A^{-}A) = A^{-}A$ by definition of A^{-} and

$\underline{N}(A) \subset \underline{N}(A^{-}A) \subset \underline{N}(AA^{-}A) = \underline{N}(A) \Rightarrow \underline{N}(A) = \underline{R}(A^{-}A)$. ■

Proposition: $P_A = AGA^{*}$ where $G = (A^{*}A)^{-}$ is any g-inverse of A.

proof: Let $P = A(A^{*}A)^{-}A^{*}$ and note $A^{*}A(A^{*}A)^{-}$ is a PO on $\underline{R}(A^{*}A) = \underline{R}(A^{*})$

$\Rightarrow A^{*}A(A^{*}A)^{-}A^{*} = A^{*}P = A^{*}$. Then

i) $P^{*} = (A(A^{*}A)^{*}A^{*})^{*} = P$ since $G^{*} = G$

ii) $P^{2} = (A(A^{*}A)^{-}A^{*})(A(A^{*}A)^{-}A^{*}) = A(A^{*}A)^{-}A^{*} = P$

iii) $\underline{N}(A^{*}) \subset \underline{N}(A(A^{*}A)^{-}A^{*}) = \underline{N}(P) \subset \underline{N}(A^{*}P) = \underline{N}(A^{*})$

$\Rightarrow \underline{N}(A^{*}) = \underline{N}(P) \Rightarrow \underline{R}(A) = \underline{R}(P^{*}) = \underline{R}(P)$ by i).

$\therefore P_A = P \Rightarrow P_A = AGA^{*}$ where $G = (A^{*}A)^{-}$ is any g-inverse of A. ■

Proposition: If $A^{*} = A$, then $(A^{+})^{*} = A$.

proof: Let $C = (A^{+})^{*}$. Using the definition of A^{+} it can be shown that $C = (A^{*})^{+}$ since

1) $A^{*}CA^{*} = (AA^{+}A)^{*} = A^{*}$ 2) $CA^{*}C = (A^{+}AA^{+})^{*} = A^{+*} = C$

3) $(A^{*}C)^{*} = C^{*}A = A^{+}A = (A^{+}A)^{*} = A^{+}C$ 4) $(CA^{*})^{*} = AC^{*} = AA^{+} = (AA^{+})^{*} = CA^{*}$.

Hence, $C = (A^{+})^{*} = (A^{*})^{+} = A^{+}$ as $A^{*} = A$. ■

(Seely,1996) (Schott,1997)

Proposition: For a linear transformation A, there exists a unique A^{+}.

proof: i) Let $B = (A^*A)^- A^*$ and $C = (AA^*)^- A$ where B and C exist since the g-inverses exist. Define $G = C^*AB = A^*(AA^*)^- A(A^*A)^- A^*$.

i) To show A^+ exists and $A^+ = G$. Using the above expressions for P_A gives

1) $AGA = AC^*AB = AA^*(AA^*)^- A(A^*A)^- A^*A = AA^*(AA^*)^- P_A A$

$= AA^*(AA^*)^- A = A$ since $AA^*(AA^*)^-$ is a PO on $\underline{R}(AA^*) = \underline{R}(A)$.

2) $GAG = C^*ABAC^*AB = (A^*(AA^*)^- A(A^*A)^- A^*)A(A^*(AA^*)^- A(A^*A)^- A^*)$

$= A^*(AA^*)^- P_A A P_{A^*}(A^*A)^- A^* = A^*(AA^*)^- A(A^*A)^- A^* = C^*AB = G.$

3) $AG = AC^*AB = AA^*(AA^*)^- A(A^*A)^- A^* = AP_{A^*}(A^*A)^- A^* = A(A^*A)^- A^* = P_A$

$\Rightarrow (AG)^* = P_A^* = P_A = A.$

4) $GA = C^*ABA = A^*(AA^*)^- A(A^*A)^- A^*A = A^*(AA^*)^- P_A A = A^*(AA^*)^- A = P_{A^*}$

$\Rightarrow (GA)^* = P_{A^*}^* = P_{A^*} = GA$.

$\therefore A^+ = G \Rightarrow$ for an arbitrary matrix A there exists A^+ by definition.

ii) To show A^+ is unique. Suppose $\exists$ 2 Moore-Penrose inverses G_1 and G_2. By definition,

1) $AG_1 = (AG_1)^* = G_1^*A^* = G_1^*(AG_2A)^* = (AG_1)^*(AG_2)^* = AG_1AG_2 = AG_2$

2) $G_1A = (G_1A)^* = A^*G_1^* = (AG_2A)^*G_1^* = (G_2A)^*(G_1A)^* = G_2AG_1A = G_2A.$

$\therefore G_1 = G_1AG_1 = G_1AG_2 = G_2AG_2 = G_2.$ ▪

Corollary: For a linear transformation A, $AA^+ = P_A$ and $A^+A = P_{A^*}$.

The proof of the corollary follows from the proof of the above theorem. The next results will be useful in characterizing projection operators. The theorem was given its name in order to identify it easily.

(Seely, 1996)

General Projection Theorem: Suppose D and A are conformable linear transformations $\ni$ $\underline{r}(A^*DA) = \underline{r}(A^*)$ and G is a g-inverse of A^*DA. Then

i) $DAGA^*$ is the PO on $\underline{R}(DA)$ along $\underline{N}(A^*)$ ii) AGA^*D is the PO on $\underline{R}(A)$ along $\underline{N}(A^*D)$.

proof: i) Note (1) $(DAGA^*)(DAGA^*) = DAGA^*$ since A^*DAG is a PO on $\underline{R}(A^*DA) = \underline{R}(A^*)$

(2) $\underline{N}(A^*) \subset \underline{N}(DAGA^*) \subset \underline{N}(A^*DAGA^*) = \underline{N}(A^*)$ since $A^* = A^*DAGA^*$

(3) $\underline{R}(DA) = \underline{R}(DAGA^*DA) \subset \underline{R}(DAGA^*) \subset \underline{R}(DA)$ since $A = AGA^*DA$.

ii) Note (1) $(AGA^*D)(AGA^*D) = AGA^*D$ since A^*DAG is a PO on $\underline{R}(A^*DA) = \underline{R}(A^*)$

(2) $\underline{N}(A^*D) \subset \underline{N}(AGA^*D) \subset \underline{N}(A^*DAGA^*D) = \underline{N}(A^*D)$ since $A^* = A^*DAGA^*$

(3) $\underline{R}(A) = \underline{R}(AGA^*DA) \subset \underline{R}(AGA^*D) \subset \underline{R}(A)$ since $A = AGA^*DA$. ▪

Proposition: If $P : \mathcal{W} \to \mathcal{W}$ is an OPO, then $\exists\, T : \mathcal{T} \to \mathcal{W} \ni P = TT^*$ and $T^*T = I$.

proof: Let $\mathcal{T} = \underline{R}(P)$ and define $Tt = t \ \forall\, t \in \mathcal{T} \Rightarrow T : \mathcal{T} \to \mathcal{W}$. Then $\forall\, \omega \in \mathcal{W},\ t \in \mathcal{T}$

$< T^*\omega, t >_{\mathcal{T}} = \ < \omega, Tt >_{\mathcal{W}} = \ < \omega, t >_{\mathcal{W}} = \ < \omega, Pt >_{\mathcal{W}} = \ < P\omega, t >_{\mathcal{T}} \Rightarrow T^*\omega = P\omega.$

Now, $TT^* : \mathcal{W} \to \mathcal{W}$ and $\forall\, \omega \in \mathcal{W},\ TT^*\omega = TP\omega = P\omega \Rightarrow P = TT^*$.

Note $T^*T : \mathcal{T} \to \mathcal{T}$ where $\forall\, t \in \mathcal{T},\ T^*Tt = T^*t = Pt = t \Rightarrow T^*T = I.$ ∎

The next propositions are useful for describing combinations of projection operators.

(Halmos,1958) (Christensen,1996)

Proposition: If P_1 and P_2 are OPOs, then the following are equivalent for $P = P_1 + P_2$

i) $P^2 = P$ ii) $P_1P_2 = P_2P_1 = 0$ iii) P is the OPO on $\underline{R}(P_1) + \underline{R}(P_2)$.

proof: (1) i) $\Leftrightarrow P_1^2 + P_2^2 + P_1P_2 + P_2P_1 = P_1 + P_2 \Leftrightarrow P_1P_2 + P_2P_1 = 0 \ (*)$

$\Rightarrow P_1P_2 + P_1P_2P_1 = 0$ and $P_1P_2P_1 + P_2P_1 = 0$ by left and right multiplying by P_1

$\Rightarrow P_1P_2 = P_2P_1 \ (\circ) \Rightarrow$ ii) from $(*)$ and $(\circ)$.

(2) Note ii) $\Rightarrow P_1P_2 + P_2P_1 = 0 \Rightarrow$ i) from $(*)$ in (1).

(3) Note iii) $\Rightarrow$ i) by definition of OPO.

(4) i) $\Rightarrow P$ is an OPO since $P^2 = P$ and $P' = P$. In addition, $\underline{R}(P) \subset \underline{R}(P_1) + \underline{R}(P_2)$.

To show equality in the range spaces, let $v \in \underline{R}(P_1) + \underline{R}(P_2) \Rightarrow v = P_1u_1 + P_2u_2$ for some u_1, u_2.

Then $Pv = P(P_1u_1 + P_2u_2) = (P_1 + P_2)(P_1u_1 + P_2u_2) = P_1u_1 + P_1P_2u_2 + P_2P_1u_1 + P_2u_2$

$= P_1u_1 + P_2u_2 = v$ as i) $\Rightarrow$ ii) by (1). Thus, $v \in \underline{R}(P) \Rightarrow \underline{R}(P_1) + \underline{R}(P_2) \subset \underline{R}(P) \Rightarrow$ iii). ∎

(Halmos,1958)

Proposition: If P_1 and P_2 are OPOs, then the following are equivalent for $P = P_1P_2$

i) $P' = P$ ii) $\underline{R}(P) \subset \underline{R}(P_2)$ iii) P is the OPO on $\underline{R}(P_1) \cap \underline{R}(P_2)$.

proof: (1) i) $\Rightarrow P' = P \Rightarrow (P_1P_2)' = P_1P_2 \Rightarrow P_2P_1 = P_1P_2 \Rightarrow P = P_2P_1 \Rightarrow \underline{R}(P) \subset \underline{R}(P_2) \Rightarrow$ ii).

(2) Suppose ii). Then $\forall\, u \ \ Pu - P_2Pu \in \underline{R}(P_2) + \underline{R}(P) \subset \underline{R}(P_2)$ by ii) and

$0 = P_2(Pu - P_2Pu) \Rightarrow Pu - P_2Pu \in \underline{R}(P_2) \cap \underline{N}(P_2) = \{0\} \Rightarrow P_2P = P \ (*)$.

Using $(*)$, $P = P_2P = P_2'P_1'P_2' = (P_1P_2)'P_2' = P'P_2' = P' \Rightarrow$ i).

(3) Note iii) $\Rightarrow P' = P \Rightarrow$ i).

(4) i) and hypothesis $\Rightarrow P' = P$ and $P^2 = P \Rightarrow P$ is an OPO. By (1),

$P = P_1P_2 = P_2P_1 \Rightarrow \underline{R}(P) \subset \underline{R}(P_1) \cap \underline{R}(P_2)$. In order to show equality of the range spaces,

suppose $w \in \underline{R}(P_1) \cap \underline{R}(P_2) \Rightarrow w = P_1u_1 = P_2u_2$ for some u_1, u_2 and

$Pw = P_1P_2P_1u_1 = P_1P_2u_1 = P_2P_1u_1 = P_2w = P_2P_2u_2 = P_2u_2 = w \ \Rightarrow w \in \underline{R}(P)$

$\Rightarrow \underline{R}(P_1) \cap \underline{R}(P_2) \subset \underline{R}(P) \Rightarrow$ iii). ∎

2.6. Trace Operator

A few linear algebra results are needed involving eigenvalues, the spectral theorem, and the trace operator. Let A be a linear transformation on an n-dimensional vector space $\mathcal{W}$ and consider the following definitions and propositions (Halmos,1958).

Definitions: Eigenvalue - A scalar $\lambda \ni Ax = \lambda x$ for some non-zero x.

Multiplicity - If $C_\lambda =$ collection of all $x \ni Ax = \lambda x$, then the multiplicity of λ is $\dim C_\lambda$.

Trace - The trace operator is given by $\text{tr}(A) = \sum_{i=1}^{r} m_i \lambda_i = \sum_{i=1}^{n} \lambda_i$
where m_i is the multiplicity for the eigenvalue λ_i and $m_1 + \ldots + m_r = n$.

(Halmos,1958)
Spectral Theorem: For every self-adjoint linear transformation A on a finite-dimensional inner product space, $\exists\ \lambda_1, \ldots, \lambda_r \in \mathcal{R}$ and OPOs $E_1, \ldots, E_r \ni$

i) $\lambda_1, \ldots, \lambda_r$ are distinct ii) $E_i \neq 0 \quad E_i E_j = 0 \quad i \neq j = 1, \ldots, r$

iii) $\sum_{i=1}^{r} E_i = I$ iv) $A = \sum_{i=1}^{r} \lambda_i E_i$.

The value λ_i in the spectral theorem is an eigenvalue of A, because for $u \in \underline{R}(E_i)$ $(u \neq 0)$ $Au = (\sum_{i=1}^{r} \lambda_i E_i) u = \lambda_i u$. The multiplicity associated with λ_i is given by $\underline{r}(E_i)$. Also, $A^s = (\sum_{i=1}^{r} \lambda_i E_i)^s = \sum_{i=1}^{r} \lambda_i^s E_i$ due to the properties of the E_i's. The next proposition gives a corresponding spectral theorem for matrices.

(Christensen,1996)
Proposition: For a symmetric matrix $M_{n\times n}$, $\exists$ a symmetric matrix $R \ni R'MR = D = \text{diag}(\{\lambda_i\})$.

proof: Let $\underline{v}_1, \ldots, \underline{v}_n$ be an orthonormal set of eigenvectors of M corresponding to the eigenvalues $\lambda_1, \ldots, \lambda_n$. Then $R = [\underline{v}_1, \ldots, \underline{v}_n]$ and noting $\underline{v}_i' \underline{v}_j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$ gives
$R'MR = R'[M\underline{v}_1, \ldots, M\underline{v}_n] = R'[\lambda_1 \underline{v}_1, \ldots, \lambda_n \underline{v}_n] = \{\lambda_1 \underline{v}_i' \underline{v}_1 \ldots \lambda_n \underline{v}_i' \underline{v}_n\}_{n\times n} = \text{diag}(\{\lambda_i\}) = D.$ ∎

Since the trace operator plays a crucial role in the development of later results, it will be developed in this section. The next proposition gives an expression for the trace of a matrix. Then the properties of the trace are explored using both formulations. The matrix formulation will be most useful in later chapters.

The following results illustrate linear concepts concerning the relation between a linear transformation and its associated matrix representation.

Proposition: If the linear transformation A has matrix representation $M_{n\times n} = \{m_{ij}\}$, then $\text{tr}(A) = \sum_{i=1}^{n} m_{ii}$.

proof: i) Define the function $\tau(M) = \sum_{i=1}^{n} m_{ii}$.

To show if M and R are two matrix representations of A, then $\tau(M) = \tau(R)$.

Let $\text{sp}\{e_1, ..., e_n\} = \text{sp}\{f_1, ..., f_n\} = \mathcal{W}$ where $f_i = \sum_{j=1}^{n} b_{ij}e_j \ \forall\, i = 1, ..., n$ (1).

Since A is a mapping on $\mathcal{W}$, define $Ae_j = \sum_{l=1}^{n} m_{jl}e_l$ (2) and $Af_i = \sum_{j=1}^{n} r_{ij}f_j$ (3).

$$(4)\ \sum_{l=1}^{n}\Big(\sum_{j=1}^{n} b_{ij}m_{jl}\Big)e_l = \sum_{j=1}^{n} b_{ij}\Big(\sum_{l=1}^{n} m_{jl}e_l\Big) \overset{(2)}{=} \sum_{j=1}^{n} b_{ij}Ae_j \overset{(1)}{=} Af_i \overset{(3)}{=} \sum_{j=1}^{n} r_{ij}f_j$$

$$\overset{(1)}{=} \sum_{j=1}^{n} r_{ij} \sum_{l=1}^{n} b_{jl}e_l = \sum_{l=1}^{n}\Big(\sum_{j=1}^{n} r_{ij}b_{jl}\Big)e_l.$$

Hence, $\sum_{l=1}^{n}\Big(\sum_{j=1}^{n} b_{ij}m_{jl}\Big)e_l = \sum_{l=1}^{n}\Big(\sum_{j=1}^{n} r_{ij}b_{jl}\Big)e_l \Leftrightarrow \sum_{l=1}^{n}\Big(\sum_{j=1}^{n}(b_{ij}m_{jl} - r_{ij}b_{jl}\Big)e_l = 0$

$\Leftrightarrow \sum_{j=1}^{n}(b_{ij}m_{jl} - r_{ij}b_{jl}) = 0$ as e_l are linearly independent

$\Leftrightarrow \sum_{j=1}^{n} b_{ij}m_{jl} = \sum_{j=1}^{n} r_{ij}b_{jl} \Leftrightarrow BM = RB \Leftrightarrow R = BMB^{-1}$ where $B_{n\times n} = \{b_{ij}\}$ is 1-1 and onto by (1).

Let $D = B^{-1} = \{d_{li}\}$. Thus,

$$\tau(R) = \sum_{i=1}^{n} r_{ii} = \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{l=1}^{n} b_{ij}m_{jl}d_{li} = \sum_{j=1}^{n}\sum_{l=1}^{n} m_{jl}\Big(\sum_{i=1}^{n} d_{li}b_{ij}\Big) = \sum_{j=1}^{n}\sum_{l=1}^{n} m_{jl}\delta_{lj} = \sum_{i=1}^{n} m_{ii} = \tau(M).$$

ii) For the matrix M, $\exists$ a non-singular matrix $B \ni R = BMB^{-1}$ is triangular (Halmos,1958,p107).

To show $r_i\ i = 1, ..., n$ are the eigenvalues of A where r_i is the i^{th} diagonal element of R.

a) Define R^d to be a diagonal matrix with entries $r_i\ i = 1, ..., n$. Note $|R| = |R^d|$ (4)

and $R^d\underline{\delta}_i = r_i\underline{\delta}_i$ (5) where $\underline{\delta}_i = \{\delta_{ij}\}$ where $\delta_{ij} = 1$ if $i = j$ and 0 otherwise. By definition of eigenvalue, $|R^d - r_iI| = 0 \Rightarrow |(R - r_iI)^d| = 0 \overset{(4)}{\Rightarrow} |R - r_iI| = 0 \Rightarrow r_i$ is an eigenvalue of R.

b) $M\underline{x} = \lambda\underline{x} \Leftrightarrow B^{-1}BMB^{-1}B\underline{x} = \lambda\underline{x} \Leftrightarrow BMB^{-1}B\underline{x} = \lambda B\underline{x}$

$\Leftrightarrow R(B\underline{x}) = \lambda(B\underline{x})$. Thus, λ is an eigenvalue of $M \Leftrightarrow \lambda$ is an eigenvalue of R (Halmos,1958).

c) Let $\text{sp}\{e_1, ..., e_n\} = \mathcal{W}$. Then λ is an eigenvalue of A with eigenvector $x = \sum_{i=1}^{n} v_ie_i$

$$\Leftrightarrow Ax = \lambda x \Leftrightarrow \sum_{i=1}^{n} v_iAe_i = \sum_{i=1}^{n}\lambda v_ie_i \overset{(2)}{\Leftrightarrow} \sum_{i=1}^{n}\sum_{j=1}^{n} v_im_{ij}e_j = \sum_{i=1}^{n}\lambda v_ie_i \Leftrightarrow \sum_{j=1}^{n}\Big(\sum_{i=1}^{n} v_im_{ij}\Big)e_j = \sum_{j=1}^{n}\lambda v_je_j$$

$\Leftrightarrow \sum_{j=1}^{n}\Big(\Big(\sum_{i=1}^{n} v_im_{ij}\Big) - \lambda v_j\Big)e_j = 0 \Leftrightarrow \sum_{i=1}^{n} v_im_{ij} = \lambda v_j \quad j = 1, ..., n$ by linear independence

$\Leftrightarrow M\underline{v} = \lambda\underline{v}$ for $\underline{v} = [v_1, ..., v_n]' \Leftrightarrow \lambda$ is an eigenvalue of M (Marcus and Minc,1965).

Thus, $\sum_{i=1}^{n} m_{ii} = \tau(M) \overset{(i)}{=} \tau(R) = \sum_{i=1}^{n} r_i \overset{(ii)}{=} \text{tr}(A)$. ▪

Proposition: Consider linear transformations A and B on an n-dimensional vector space $\mathcal{W}$. Then

i) $\text{tr}(\alpha A + \beta B) = \alpha\text{tr}(A) + \beta\text{tr}(B)$ for $\alpha, \beta \in \mathcal{R}$

ii) $\text{tr}(AB) = \text{tr}(BA)$

iii) If $A^2 = A$, then $\text{tr}(A) = \underline{r}(A)$.

proof: The above proposition can be used to obtain the above equalities. Suppose $M_{n \times n} = \{m_{ij}\}$ and $R_{n \times n} = \{r_{ij}\}$ are matrix representations of A and B respectively.

i) The matrix representation of $\alpha A + \beta B$ is given by $\{\alpha m_{ij} + \beta r_{ij}\}$ by definition of matrix addition and scalar multiplication (Marcus and Minc,1965). Then

$\alpha\text{tr}(A) + \beta\text{tr}(B) = \alpha\sum_{i=1}^{n} m_{ii} + \beta\sum_{i=1}^{n} r_{ii} = \sum_{i=1}^{n}(\alpha m_{ii} + \beta r_{ii}) = \text{tr}(\alpha A + \beta B)$.

ii) The matrix representation of AB is given by $\{\sum_{s=1}^{n} m_{is} r_{sj}\}$ by definition of matrix multiplication (Marcus and Minc,1965). Then $\text{tr}(AB) = \sum_{i=1}^{n}(\sum_{j=1}^{n} m_{ij} r_{ji}) = \sum_{j=1}^{n}(\sum_{i=1}^{n} r_{ji} m_{ij}) = \text{tr}(BA)$.

iii) (1) Let $\text{sp}\{e_1, ..., e_n\} = \mathcal{W}$, $Ae_i = \sum_{j=1}^{n} m_{ij} e_j$, and $M = \{m_{ij}\}$. Consider the mapping $\Phi : \mathcal{W} \rightarrow \mathcal{R}^n$ given by $\Phi(u) = \{c_i\}_{n \times 1} = \underline{c}$ where $u = \sum_{i=1}^{n} c_i e_i$. In addition, suppose $v = \sum_{i=1}^{n} d_i e_i$ and $\alpha, \beta \in \mathcal{R}$. Then

a) $\Phi(\alpha u + \beta v) = \Phi(\sum_{i=1}^{n}(\alpha c_i + \beta d_i)e_i) = \{\alpha c_i + \beta d_i\}_{n \times 1} = \alpha\underline{c} + \beta\underline{d} = \alpha\Phi(u) + \beta\Phi(v)$

b) $\Phi(u) = \Phi(v) \Rightarrow \sum_{i=1}^{n} c_i e_i = \sum_{i=1}^{n} d_i e_i \Rightarrow \sum_{i=1}^{n}(c_i - d_i)e_i = 0 \Rightarrow c_i = d_i \; i = 1, ..., n \Rightarrow u = v$

c) $\underline{R}(\Phi) = \Phi(\mathcal{W}) = \Phi(\text{sp}\{e_1, ..., e_n\}) = \text{sp}\{\Phi(e_1), ..., \Phi(e_n)\} = \text{sp}\{\underline{\delta}_1, ..., \underline{\delta}_n\} = \mathcal{R}^n$

by i) with $\underline{\delta}_j = \{\delta_{ij}\}$ where $\delta_{ij} = 1$ if $i = j$ and 0 otherwise.

d) $\Phi(Au) = \Phi(\sum_{i=1}^{n} c_i A e_i) = \Phi(\sum_{i=1}^{n} c_i \sum_{j=1}^{n} m_{ij} e_j)$ where $Ae_i = \sum_{j=1}^{n} m_{ij} e_j$

$\overset{(a)}{=} \sum_{i=1}^{n}\sum_{j=1}^{n} c_i m_{ij} \Phi(e_j) = \sum_{i=1}^{n}\sum_{j=1}^{n} c_i m_{ij} \underline{\delta}_j = \sum_{i=1}^{n} c_i \underline{m}_i = M\underline{c} = M\Phi(u)$.

Hence, $\underline{r}(A) = \dim \underline{R}(A) \overset{(b,c)}{=} \dim \underline{R}(\Phi A) \overset{(d)}{=} \dim \underline{R}(M\Phi) \overset{(b,c)}{=} \dim \underline{R}(M) = \underline{r}(M)$.

(2) Consider the matrix representation of A given by M. Then $\exists$ a non-singular matrix $B \ni R = BMB^{-1}$ is triangular (Halmos,1958). Note $\underline{r}(M) = \underline{r}(R) =$ number of non-zero diagonal entries.

(3) Suppose $Ax = \lambda x$. Then $\lambda x = Ax = A^2 x = A(Ax) = \lambda Ax = \lambda^2 x \Rightarrow \lambda^2 = \lambda \Rightarrow \lambda = 0$ or 1 $\Rightarrow r_i = 0$ or 1 $i = 1, ..., n$ as diagonal elements of R (r_i) equal eigenvalues of A from ii) in above proof.

$\therefore \underline{r}(A) \overset{(1)}{=} \underline{r}(M) \overset{(2)}{=} \underline{r}(R) \overset{(2,3)}{=} \sum_{i=1}^{n} r_i \overset{(3)}{=} \text{tr}(A)$. ∎

2.7. Non-Negative Definite Linear Transformations

A few results will be presented in this section concerning non-negative definite (NND) and positive definite (PD) linear transformations which are defined in section 2.1. These linear transformations will be defined on the n-dimensional vector space $\mathcal{W}$ with inner product $<\, ,\, >$.

(Christensen,1996)

Proposition: i) The eigenvalues for a NND linear transformation are greater than or equal to zero.

ii) A self-adjoint linear transformation A is NND $\Leftrightarrow A = BB^*$ for some B.

proof: i) Let A be an NND transformation defined on $\mathcal{W}$. By the spectral theorem, $\exists\ \lambda_1, ..., \lambda_r \in \mathcal{R}$ and OPOs $E_1, ..., E_r \ni A = \sum_{i=1}^{r}\lambda_i E_i$. Then $< Av, v > \ \geq 0\ \ \forall\ v$ by definition of NND $\Rightarrow\ < Au, u > \ \geq 0$ for $u \in \underline{R}(E_i) \ni u \neq 0 \Rightarrow\ < (\sum_{i=1}^{r}\lambda_i E_i)u, u > \ \geq 0$ by the spectral theorem $\Rightarrow \lambda_i < u, u > \ \geq 0 \Rightarrow \lambda_i \geq 0$ as $< u, u > \ > 0$ since $u \neq 0$.

ii) (1) Suppose A is NND. By spectral theorem, $\exists\ \lambda_1, ..., \lambda_r \in \mathcal{R}$ and OPOs $E_1, ..., E_r \ni A = \sum_{i=1}^{r}\lambda_i E_i$ $\Rightarrow A^{\frac{1}{2}} = \sum_{i=1}^{r}\lambda_i^{\frac{1}{2}} E_i \Rightarrow A = A^{\frac{1}{2}}A^{\frac{1}{2}} = BB^*$.

(2) Suppose $A = BB^*$. Then A is self-adjoint and $< Av, v > \ = \ < BB^*v, v > \ = \ < B^*v, B^*v > \ \geq 0$ $\Rightarrow A$ is NND by definition. ■

Proposition: Consider conformable linear transformations D and V.

i) If V is NND, then $\underline{R}(D^*VD) = \underline{R}(D^*V)$. ii) If V is PD, then $\underline{R}(D^*VD) = \underline{R}(D^*)$.

proof: i) (1) Note $\underline{N}(D) \subset \underline{N}(VD) \subset \underline{N}(D^*VD)$.

(2) Because V is NND $\Rightarrow \exists\, B \ni V = BB^*$ by above proposition. Suppose $t \in \underline{N}(D^*VD) \Rightarrow D^*VDt = 0 \Rightarrow B^*Dt = 0 \Rightarrow BB^*Dt = 0 \Rightarrow VDt = 0\ (*) \Rightarrow t \in \underline{N}(VD)$.

Thus, by (1) and (2), $\underline{N}(D^*VD) = \underline{N}(VD) \Rightarrow \underline{R}(D^*VD) = \underline{R}(D^*V)$.

(3) From $(*)$ in (2) and since V is PD, $Dt = 0 \Rightarrow t \in \underline{N}(D)$. Thus, by (1) and (2) $\underline{N}(D^*VD) = \underline{N}(D) \Rightarrow \underline{R}(D^*VD) = \underline{R}(D^*)$. ■

Proposition: If V, W are NND linear transformations, then

i) $\text{tr}(W) = 0 \Leftrightarrow W = 0$ ii) $\text{tr}(VW) = 0 \Leftrightarrow VW = 0$.

proof: By the spectral theorem, $\exists\ \lambda_1, ..., \lambda_r \in \mathcal{R}$ and OPOs $E_1, ..., E_r \ni$ $W = \sum_{i=1}^{r}\lambda_i E_i = (\sum_{i=1}^{r}\lambda_i^{\frac{1}{2}} E_i)^2 = B^2$. Then (1) $W = 0 \Rightarrow \text{tr}(W) = \text{tr}(0) = 0$

(2) $0 = \text{tr}(W) = \text{tr}(\sum_{i=1}^{r}\lambda_i E_i) = \sum_{i=1}^{r}\lambda_i \text{tr}(E_i) \Rightarrow \text{tr}(E_i)\ \forall\ i$ by i) of above proposition as since $W \geq 0$ $\Rightarrow \underline{r}(E_i) = 0\ \forall i$ by iii) of proposition in section 2.6 since $E_i^2 = E_i \forall i \Rightarrow W = 0$.

ii) (1) $VW = 0 \Rightarrow \text{tr}(VW) = \text{tr}(0) = 0$.

(2) $0 = \text{tr}(VW) = \text{tr}(VB^2) = \text{tr}(BVB) \Rightarrow BVB = 0$ by i)(2) as BVB is NND $\Rightarrow VB = 0$ by above proposition $\Rightarrow VW = 0$. ■

2.8. Quadratic Subspaces

Quadratic spaces, developed by Seely (1969), will be useful in later results. This section pertains to symmetric $n \times n$ matrices $(\mathcal{S}_n)$. Definitions are given below where $\mathcal{C}$ is a subspace of $\mathcal{S}_n$.

Definitions: Quadratic Subspace (QS) - $\mathcal{C}$ is a QS provided $A^2 \in \mathcal{C} \ \ \forall \ A \in \mathcal{C}$.

Commutative Quadratic Subspace (CQS) - $\mathcal{C}$ is a CQS provided $\mathcal{C}$ is a QS and $\forall \ A, B \in \mathcal{C} \ \ AB = BA$.

The next proposition can be used to check whether or not a subspace is a QS or a CQS.

(Seely, 1969)

Proposition: i) $\mathcal{C}$ is a QS $\Leftrightarrow AB + BA \in \mathcal{C} \quad \forall \ A, B \in \mathcal{C}$.

ii) $\mathcal{C}$ is a CQS $\Leftrightarrow AB \in \mathcal{C} \quad \forall \ A, B \in \mathcal{C}$.

iii) If $\mathcal{C}$ is a QS, then $ABA \in \mathcal{C} \ \ \forall \ A, B \in \mathcal{C}$.

iv) Suppose $\exists \ D \in \mathcal{C} \ni AD = DA = A$. If $ABA \in \mathcal{C} \ \ \forall \ A, B \in \mathcal{C}$, then $\mathcal{C}$ is a QS.

proof: i) (1) $\mathcal{C}$ is a QS $\Rightarrow (A + B)^2 = A^2 + (AB + BA) + B^2 \in \mathcal{C} \Rightarrow AB + BA \in \mathcal{C}$ as $A^2, B^2 \in \mathcal{C}$

(2) $AB + BA \in \mathcal{C} \Rightarrow 2A^2 \in \mathcal{C}$ letting $A = B \Rightarrow A^2 \in \mathcal{C}$.

ii) (1) $\mathcal{C}$ is a CQS $\Rightarrow AB + BA = 2AB \in \mathcal{C}$ by i)

(2) $AB \in \mathcal{C} \Rightarrow AB = (AB)' \Rightarrow AB = BA$ since $A, B \in \mathcal{C}$. Hence, $AB + BA \in \mathcal{C} \Rightarrow \mathcal{C}$ is a CQS.

iii) $\mathcal{C}$ is a QS $\Rightarrow AD + DA \in \mathcal{C}$ with $D = AB + BA$ by (1)

$\Rightarrow A^2B + ABA + ABA + BA^2 \in \mathcal{C} \Rightarrow ABA \in \mathcal{C}$ as $A^2 \in \mathcal{C}$ so $A^2B + BA^2 \in \mathcal{C}$ by (1).

iv) Note $[A, B \in \mathcal{C} \Rightarrow ABA \in \mathcal{C}] \Rightarrow ADA \in \mathcal{C} \Rightarrow A^2 \in \mathcal{C}$ is a QS by definition of D. ∎

The definitions and proposition given in this section are sufficient to develop quadratic subspaces in later results.

2.9. Vec Operator and Horizontal Direct Product

Special matrix operators will also be of interest in later applications. These include the vec operator and the horizontal direct product. The vec operator allows matrices to be represented as vectors while the horizontal direct product combines matrices in a particular manner. The operators are defined below along with some of their properties:

vec operator - For $A = \{a_{ij}\} \in \mathcal{M}_n$, define vec$(A)$ by

$$\text{vec}(A) = [\, a_{11}, \ldots, a_{n1}, a_{12}, \ldots, a_{n2}, \ldots, a_{1n}, \ldots, a_{nn}]'_{n^2\times 1}.$$

horizontal direct product ($\odot$) - Let $A_{n\times r} = [\underline{a}_1, \ldots, \underline{a}_r]$ and $B_{n\times s} = [\underline{b}_1, \ldots, \underline{b}_s]$. Then

$[A \odot B]_{n\times rs} = \{a_{pu} b_{pv}\}_{(p,u,v)=(1,1,1)}^{(n,r,s)}$ where $a_{pu} = p^{th}$ element of $\underline{a}_u$.

The definition of the horizontal direct product does not specify a particular order for combining the column vectors of A and B. However, a consistent ordering should be used. The following propositions provide some elementary results involving vec and the horizontal direct product.

Proposition: i) For $\alpha, \beta \in \mathcal{R}$, $A, B \in \mathcal{M}_{n\times m}$, $\text{vec}(\alpha A + \beta B) = \alpha\text{vec}(A) + \beta\text{vec}(B)$.

ii) For $A, B \in \mathcal{M}_{n\times m}$, $\text{tr}(A'B) = \text{vec}(A)'\text{vec}(B)$.

proof: i) Result follows as scalar multiplication and addition operates same for matrices and vectors.
ii) Let $A = \{\underline{a}_i\}_{i=1}^m$ and $B = \{\underline{b}_i\}_{i=1}^m$. Then

$$\text{vec}(A)'\text{vec}(B) = [\underline{a}_1' \ldots \underline{a}_m'] \begin{bmatrix} \underline{b}_1 \\ \vdots \\ \underline{b}_m \end{bmatrix} = \sum_{i=1}^m \underline{a}_i' \underline{b}_i = \sum_{i=1}^m (A'B)_{ii} = \text{tr}(A'B). \quad \blacksquare$$

Proposition: i) $\underline{R}(A \odot B) = \underline{R}(B \odot A)$.

ii) $\underline{R}((A \odot B) \odot C) = \underline{R}(A \odot (B \odot C))$.

iii) $(A + B) \odot C = A \odot C + B \odot C$.

proof: i) Follows from definition where columns of $A \odot B$ and $B \odot A$ are identical, but interchanged.
ii) From definition where columns of $(A \odot B) \odot C$ and $A \odot (B \odot C)$ are identical, but interchanged.
iii) Then $(A + B) \odot C = \{(a_{ij} + b_{ij})c_{iu}\} = \{a_{ij}c_{iu}\} + \{b_{ij}c_{iu}\} = A \odot C + B \odot C$
which does not depend on the ordering associated with $\odot$. $\blacksquare$

This chapter provided notation, terminology, and results pertaining to linear transformations. These concepts will be used repeatedly in the later chapters. Linear transformation concepts will be particularly important for the models presented in the next chapter.

3. Linear Models

Models are tools which can be used to represent responses from random processes. For a random response in a linear space, a linear model assumes the expectation and variance exist. It is convenient to parameterize the model by expressing the expectation and variance in terms of unknown parameters that can be estimated from the random response.

(Seely,1996)
Linear Model: When the set of possible expectations of a random response is a linear subspace.

(Searle et al.,1992)
Linearly Mean-Parameterized Model : When the expectation of the random response is parameterized so that the expectation is a linear function of the parameter.

The latter model is usually called a linear model as well, but a distinction is made in this study. While these definitions are not exactly the same, the distinction is not critically important. Linearly mean-parameterized models will be presented which are not linear models, but they are essentially equivalent to linear models for purposes of this study, as will be demonstrated in section 3.3.

In order to use the approach of Szatrowski (1980), specific models need to be defined. These models include the Y-Model, linearized quadratic estimation models (LQEMs), and the Underlying Model (U-Model). The models, as well as their associated properties, are discussed in the following sections so they can be easily referenced for later chapters.

3.1. The Y-Model and Assumptions

3.1.1. Definitions and Assumptions

This study is particularly concerned with the linear model given in this section. The Y-Model is defined below for a random vector $\underline{Y} \in \mathcal{R}^n$.

Y-Model: $\mathrm{E}[\underline{Y}] \in \mathcal{E}^y = \{X\underline{\beta} \mid \underline{\beta} \in \mathcal{R}^p\} = \underline{R}(X) \qquad \mathrm{Cov}(\underline{Y}) \in \mathcal{V}^y = \{V_{\underline{\psi}} \mid \underline{\psi} \in \Xi\}$

The variance component vector $\underline{\psi}$ lies in a parameter set Ξ in which $V_{\underline{\psi}}$ is PD for all $\underline{\psi} \in \Xi$. This model does not assume any constraints on the fixed effect vector $\underline{\beta}$ and the matrix X may not have full column rank. The following assumptions will be required for some of the results, and will always be stated either in the result or at the start of the section.

Assumptions: Normality [N] - $\underline{Y} \sim \mathrm{N}_n(X\underline{\beta}, V_{\underline{\psi}})$

Linear Structure [L] - $\mathrm{Cov}(\underline{Y}) = V_{\underline{\psi}} = \sum_{i=1}^{k+1} \psi_i V_i,\ \underline{\psi} \in \Xi \subset \mathcal{R}^{k+1},\ V_{k+1} = I.$

Open Set [O] - Ξ contains a non-empty open set in $\mathcal{R}^{k+1}$.

Classification [C] - $\mathrm{E}[\underline{Y}] = \underline{1}\mu + \sum_{j=1}^{f} X_j \underline{\beta}_j$ and $\mathrm{Cov}(\underline{Y}) = \sum_{i=1}^{k} \sigma_i^2 Z_i Z_i' + \sigma_{k+1}^2 I$

where $X_1, \ldots, X_f, Z_1, \ldots, Z_k$ are classification matrices which are defined in section 3.1.3. Classification models are assumed to be proper as defined in section 3.1.3.

The Y-Model is a mixed effects model, as it contains both random and fixed effect parameters. The Y-Model under [L] has been referred to as a variance component model by Harville (1977) and Seely (1996), as having a patterned covariance matrix by Rogers (1977), and as having a covariance matrix with linear structure by Anderson (1969). A random effects linear model is a Y-Model under [L] with $X = \underline{1}$ and a fixed effects linear model is a Y-Model under [L] with $k = 0$.

The Y-Model under [C] has been referred to as a mixed classification model by Birkes (1996) and an ANOVA model by Harville (1977). This model is often expressed as $\underline{Y} = \underline{1}\mu + \sum_{j=1}^{f} X_j \underline{\beta}_j + \sum_{i=1}^{k} Z_i \underline{d}_i + \underline{e}$ where $\underline{d}_1, \ldots, \underline{d}_k, \underline{e}$ are uncorrelated random vectors with mean $\underline{0}$, $\mathrm{Cov}(\underline{d}_i) = \sigma_i^2 I$, and $\mathrm{Cov}(\underline{e}) = \sigma_{k+1}^2 I$. Sometimes σ_e^2 is used to denote σ_{k+1}^2.

The next sections develop linear model results that are needed for the Y-Model. These sections discuss the open set condition, balance, and likelihood estimation.

3.1.2. Open Set Condition

The section examines properties associated with the open set condition [O] which accompanies the linearity assumption [L]. This condition is also presented in a more general setting in sections 3.3.1 and 3.3.2. The following propositions illustrate some basic properties.

Proposition: Suppose Ξ contains a non-empty open set of dimension $k+1$. If $\mathcal{V} = \{V_{\underline{\psi}} = \sum_{i=1}^{k+1} \psi_i V_i \,|\, \underline{\psi} \in \Xi\}$, then $\mathrm{sp}\mathcal{V} = \mathrm{sp}\{V_1, \ldots, V_{k+1}\}$.

proof: Define the linear operator $V : \mathcal{R}^{k+1} \rightarrow \mathcal{S}_n$ by $V\underline{\psi} = \sum_{i=1}^{k+1} \psi_i V_i$. Since Ξ contains a non-empty open set of dimension $k+1$ and V is linear, $\mathrm{sp}\mathcal{V} = \mathrm{sp}V(\Xi) = V(\mathrm{sp}\,\Xi) = V(\mathcal{R}^{k+1}) = \mathrm{sp}\{V_1, \ldots, V_{k+1}\}$. ∎

Proposition: If $\Xi^* = \{\underline{\psi} \mid V_{\underline{\psi}} = \sum_{i=1}^{k+1} \psi_i V_i \text{ is PD}\}$, then Ξ^* is a non-empty open set in $\mathcal{R}^{k+1}$.

proof: i) The linear operator $V : \mathcal{R}^{k+1} \to \mathcal{S}_n$ defined in the preceding proof is continuous. Define $\mathcal{D}_n = \{M \in \mathcal{S}_n \mid M \text{ is PD}\}$ and $\underline{\delta}: \mathcal{S}_n \to \mathcal{R}^n$ by $\underline{\delta}(M) = [\delta_1(M), ..., \delta_n(M)]'$ where $\delta_j(M) = j^{th}$ principal determinant of M. Note δ_j, $j = 1, ..., n$, is continuous since the determinant is a sum of products of the entries of the matrix (Halmos,1958). Hence, $\underline{\delta}$ is continuous (Rudin,1976,4.10).
Note $M \in \mathcal{D}_n \Leftrightarrow \delta_j(M) > 0 \;\; j = 1, ..., n$ (Harville,1997,sec. 15.6)
$\Leftrightarrow \underline{\delta}(M) \in (0, \infty)^n$, which is an open subset of $\mathcal{R}^n$. Hence,
$\mathcal{D}_n = \underline{\delta}^{-1}((0, \infty)^n)$ is an open subset of $\mathcal{S}_n$ since $\underline{\delta}$ is continuous (Rudin,1976,4.8)
so $\Xi^* = V^{-1}(\mathcal{D}_n)$ is an open set in $\mathcal{R}^{k+1}$as V is continuous (Rudin,1976,4.8). ■

Another common form of Ξ which contains an open set in $\mathcal{R}^{k+1}$ is given by $\Xi = \{\underline{\psi} = [\sigma_1^2, \ldots, \sigma_k^2, \sigma_{k+1}^2]' \mid \sigma_1^2 \geq 0, \ldots, \sigma_k^2 \geq 0, \sigma_{k+1}^2 > 0\}$. The open set condition will be important to consider in later results.

3.1.3. Balance

Under the classification assumption [C], the number of observations in a class can be examined. Later results will consider patterns in the number of observations in a class or some sort of balance. This section establishes notation and definitions for balance in the Y-Model under [C]. The following notation and definitions are from VanLeeuwen et al. (1997) for p factors labelled $1, ..., p$:

Definitions:
The design for an p-way classification model is given by an p-dimensional incidence matrix $N = \{n_{i_1, \ldots, i_p}\}_{t_1 \times \ldots \times t_p}$, where $n_{i_1, \ldots, i_p}$ is the number of experimental units at level i_f of factor f with $i_f = 1, ..., t_f$ and $f = 1, ..., p$.

factor subsets - A subset $\mathcal{G} = \{f_1, ..., f_g\}$ with $g \leq p$ represents an effect corresponding to the interaction of the main effects of factors $f_1, ..., f_g$ or a nested effect such as when the effect of factor f_g is nested within factors $f_1, ..., f_{g-1}$.

containment - An effect associated with factor subset $\mathcal{G}$ is contained in an effect associated with factor subset $\mathcal{H}$ if $\mathcal{G} \subset \mathcal{H}$.

marginal incidence matrix ($N^{(\underline{g})}$) - For the factor subset $\mathcal{G} = \{f_1, ..., f_g\} \subset \{1, .., p\}$, $N^{(\underline{g})}$ is a g-dimensional matrix obtained from N by summing over the indices for the other $p - g$ factors. Let $\underline{g} = (f_1, ..., f_g)$ denote the vector form of $\mathcal{G}$.

classification matrix - The classification matrix G for the effect associated with factor subset $\mathcal{G}$ has a row for each observation in the data set and a column for each l combination of levels of factors in $\mathcal{G}$. In the row corresponding to a particular observation, all entries are 0 except for a 1 in the single column corresponding to the levels of the factors $f_1, ..., f_g$ that were applied to that observation. Columns with all zero entries are deleted. The sum of the j^{th} column of G corresponds to number of observations at level j.

completely balanced design - When $n_{i_1,...,i_p} = m \quad \forall\ i_1, ..., i_p$.

pseudo balance - When $n_{i_1,...,i_p} = m$ or $0 \quad \forall\ i_1, ..., i_p$.

balanced incidence matrix (Bal($\mathcal{G}$)) - the design is balanced with respect to a particular subset of factors $\mathcal{G}$ if all of the entries in $N^{(\underline{g})}$ are equal.

conditionally balanced (Bal($\mathcal{H}|\mathcal{G}$)) - N is balanced with respect to a particular subset of factors $\mathcal{H}$ given $\mathcal{G}$ if $\forall$ combination of levels of $\mathcal{G}$ the number of observations is the same for all combinations of levels of the factors in $\mathcal{H}$ that are not in $\mathcal{G}$.

balanced classification matrix - The classification matrix G is balanced if and only if each column of G has the same number of observations.

maximal rank - A classification matrix G has maximal rank provided that it has the same rank as when $N^{(\underline{g})}$ has all non-zero entries.

included effect - When the effect associated with some combination of factors is in the model.

proper classification model - Whenever $\mathcal{H}$ and $\mathcal{G}$ are random effect subsets then either $\mathcal{H} \cap \mathcal{G}$ is a random effect subset or it is contained in a fixed effect subset.

In almost all classification models that occur in practice, if the intersection of two included interaction effects is in the model and, if an included lower order effect is random, then all included higher order effects containing it must be random. Such models are proper. All mixed classification models that will be considered will be proper.

Complete balance is equivalent to a balanced incidence matrix with respect to the set of all factors. The notation Bal($\mathbb{G}$) or Bal($\mathbb{H}|\mathbb{G}$) will be used to denote Bal($\mathcal{G}$) or Bal($\mathcal{H}|\mathcal{G}$) for all factor subsets $\mathcal{G}$ and $\mathcal{H}$ in a collection of factor subsets defined by $\mathbb{G}$ and $\mathbb{H}$.

The following proposition characterizes properties of a classification matrix. These properties follow directly from the definition of a classification matrix.

(Birkes,1996) (Seely,1996)

Proposition: Let $H_{n\times s}$ be a classification matrix and n_j = # of 1's in the j^{th} column $j = 1, ..., s$. Then

i) $H'H = \text{diag}(n_1, \dots, n_s)$ ii) $H'\underline{1} = (n_1, \dots, n_s)'$ iii) $H\underline{1} = \underline{1}$ iv) $\underline{r}(H) = s$.

These properties of a classification matrix are helpful for examining balance. The following propositions demonstrate the relation between the classification matrix and balance.

(VanLeeuwen et al.,1997)

Proposition: If the incidence matrix is balanced with respect to $\mathcal{G}$, then the associated classification matrix G is balanced.

proof: Bal($\mathcal{G}$) $\Rightarrow$ all combinations of levels of the factors in $\mathcal{G}$ have the same number of observations $\Rightarrow$ all columns of G have the same number of observations $\Rightarrow$ G is balanced. ∎

(VanLeeuwen et al.,1997)

Proposition: Suppose H, G are associated classification matrices for $\mathcal{H}, \mathcal{G}$, respectively.

i) Bal($\mathcal{H}$) $\Rightarrow H'H = qI_t$ where H is $n \times t$ and q = # of observations in each column of H.

ii) Bal($\mathcal{H}$) $\Rightarrow P_H = \frac{1}{q}HH'$.

iii) Bal($\mathcal{H} \cup \mathcal{G}$) $\Rightarrow P_H P_G = P_K$, where K is a classification matrix of an included effect with $\underline{R}(K) = \underline{R}(H) \cap \underline{R}(G)$.

proof: i) Let q_i = # of observations in column of H $i = 1, ..., t$. Note

Bal($\mathcal{H}$) $\Rightarrow q_i = q \;\forall\; i = 1, ..., t$ by definition

$\Rightarrow H'H = \text{diag}(q_1, \dots, q_t) = \text{diag}(q, \dots, q)$ from classification matrix results

$\Rightarrow H'H = q\text{diag}(1, \dots, 1) = qI_t$.

ii) Since H is a classification matrix it has full column rank. Then $P_H = H(H'H)^{-1}H' = \frac{1}{q}HH'$ by i).

iii) The proof of this result is given by VanLeeuwen et al. (1997). ∎

This subsection concludes by showing that complete balance gives Zyskind's condition for the Y-Model under [C]. Zyskind's condition will be discussed further in section 3.3.5.

(Birkes,1996)

Proposition: If the Y-Model under [C] is completely balanced, then $\underline{R}(V_i X_j) \subset \underline{R}(X)\ \forall\ i, j$.

proof: Note $\underline{R}(V_i X_j) = \underline{R}(Z_i Z_i' X_j) = \underline{R}(P_{Z_i} P_{X_j}) = \underline{R}(Z_i) \cap \underline{R}(X_j)$ from above propositon $\Rightarrow \underline{R}(V_i X_j) \subset \underline{R}(X_j) \subset \underline{R}(X)$. ■

3.1.4. Likelihood Estimation

Likelihood estimation provides a way to estimate $\underline{\beta} \in \mathcal{R}^p$ and $\underline{\psi} \in \Xi$ in the Y-Model under [L], [O], and [N]. This estimation method identifies the parameter estimate that maximizes the likelihood function. The maximum likelihood estimate is the parameter point under which the observed sample is most likely to occur (Casella and Berger,1990). Thus, this type of estimation requires a distribution. Under normality, $\underline{Y} \sim \mathrm{N}_n(X\underline{\beta}, V_{\underline{\psi}})$. The density of $\underline{Y}$ and the likelihood function are given below assuming $V_{\underline{\psi}}$ is PD:

$$f(\underline{y}|\underline{\beta},\underline{\psi}) = (2\pi)^{-\frac{n}{2}}|V_{\underline{\psi}}|^{-\frac{1}{2}}\exp(\tfrac{-1}{2}(\underline{y} - X\underline{\beta})' V_{\underline{\psi}}^{-1}(\underline{y} - X\underline{\beta}))$$
$$l(\underline{\beta},\underline{\psi}) = \ln f(\underline{Y}\,|\underline{\beta},\underline{\psi}) = -\tfrac{n}{2}\ln 2\pi - \tfrac{1}{2}\ln|V_{\underline{\psi}}| - \tfrac{1}{2}(\underline{Y} - X\underline{\beta})' V_{\underline{\psi}}^{-1}(\underline{Y} - X\underline{\beta}).$$

In order to find the parameter points under which the samples are most likely to occur, the likelihood function can be maximized by differentiating with respect to $\underline{\beta}$ and $\underline{\psi}$, setting these derivatives equal to zero, and verifying these estimators generate a global maximum. A local maximum would exist when the matrix of second derivatives is negative definite. However, it can be difficult to determine the existence of a global maximum. Due to this difficulty, this study will focus on those estimators that are roots of the equations involving the first derivative. For differentiation, it is necessary to take the derivative of a matrix A which depends on a scalar t. The derivative of $A(t) = \{a_{ij}(t)\}$ is defined to be $\frac{d}{dt}A(t) = \{\frac{d}{dt}a_{ij}(t)\}$. The following matrix derivatives will be used where the first two require that A is invertible (Searle et al.,1992) (Harville,1997):

$$\tfrac{d}{dt}A^{-1}(t) = -A^{-1}(\tfrac{d}{dt}A(t))A^{-1} \qquad \tfrac{d}{dt}\ln|A(t)| = \mathrm{tr}(A^{-1}\tfrac{d}{dt}A(t)) \qquad \tfrac{d}{dt}\mathrm{tr}(A(t)) = \mathrm{tr}(\tfrac{d}{dt}A(t)).$$

The derivatives will now be taken assuming [L] and [O]. The maximum likelihood equations, maximum likelihood equation estimators, and the information matrix are (Searle et al.,1992,ch6):

(1) Maximum Likelihood (ML) Equations

$$0 = \frac{\partial l(\underline{\beta},\underline{\psi})}{\partial \underline{\beta}} = X'V_{\underline{\psi}}^{-1}\underline{Y} - X'V_{\underline{\psi}}^{-1}X\underline{\beta} = X'V_{\underline{\psi}}^{-1}(\underline{Y} - X\underline{\beta})$$
$$0 = \frac{\partial l(\underline{\beta},\underline{\psi})}{\partial \psi_i} = -\tfrac{1}{2}\mathrm{tr}(V_{\underline{\psi}}^{-1}V_i) + \tfrac{1}{2}(\underline{Y} - X\underline{\beta})'V_{\underline{\psi}}^{-1}V_iV_{\underline{\psi}}^{-1}(\underline{Y} - X\underline{\beta})$$

(2) Maximum Likelihood Equation Estimators (MLQE) - the solutions $(\widehat{\underline{\beta}}_{\text{MLQ}}, \widehat{\underline{\psi}}_{\text{MLQ}}) = (\widehat{\underline{\beta}}, \widehat{\underline{\psi}})$ given by

$$(X'V_{\widehat{\underline{\psi}}}^{-1}X)\widehat{\underline{\beta}} = X'V_{\widehat{\underline{\psi}}}^{-1}\underline{Y} \qquad \{\mathrm{tr}(V_{\widehat{\underline{\psi}}}^{-1}V_iV_{\widehat{\underline{\psi}}}^{-1}V_j)\}_{(k+1)\times(k+1)}\widehat{\underline{\psi}} = \{\underline{Y}'F_{\widehat{\underline{\psi}}}V_iF_{\widehat{\underline{\psi}}}\underline{Y}\}_{(k+1)\times 1}$$

where $F_{\underline{\psi}} = V_{\underline{\psi}}^{-1} - V_{\underline{\psi}}^{-1}X(X'V_{\underline{\psi}}^{-1}X)^{-}X'V_{\underline{\psi}}^{-1}$

(3) Information Matrix $(i(\underline{\beta},\underline{\psi}))$

$$u = \frac{\partial^2 l(\underline{\beta},\underline{\psi})}{\partial\underline{\beta}\,\partial\underline{\beta}'} = -X'V_{\underline{\psi}}^{-1}X$$
$$\underline{v}_i = \frac{\partial^2 l(\underline{\beta},\underline{\psi})}{\partial\underline{\beta}\,\partial\psi_i} = -X'V_{\underline{\psi}}^{-1}V_iV_{\underline{\psi}}^{-1}(\underline{Y} - X\underline{\beta})$$
$$w_{ij} = \frac{\partial^2 l(\underline{\beta},\underline{\psi})}{\partial\psi_i\partial\psi_j} = \tfrac{1}{2}\mathrm{tr}(V_{\underline{\psi}}^{-1}V_iV_{\underline{\psi}}^{-1}V_j) - (\underline{Y} - X\underline{\beta})'V_{\underline{\psi}}^{-1}V_iV_{\underline{\psi}}^{-1}V_jV_{\underline{\psi}}^{-1}(\underline{Y} - X\underline{\beta})$$

$$i(\underline{\beta},\underline{\psi}) = \begin{bmatrix} -\mathrm{E}[u] & -\{\mathrm{E}[\underline{v}_i]\}' \\ -\{\mathrm{E}[\underline{v}_i]\} & -\{\mathrm{E}[w_{ij}]\} \end{bmatrix} = \begin{bmatrix} X'V_{\underline{\psi}}^{-1}X & 0 \\ 0 & \{\tfrac{1}{2}\mathrm{tr}(V_{\underline{\psi}}^{-1}V_iV_{\underline{\psi}}^{-1}V_j)\} \end{bmatrix}_{(p+k+1)\times(p+k+1)}.$$

The ML equations in (1) and the information matrix in (3) can be obtained using the derivative rules. The matrix $F_{\underline{\psi}}$ defined in (2) will often be of use and is further discussed below. The MLQEs solve the equations in (2) where the equation for the variance components has been re-expressed using the following proposition.

Proposition: $\{\mathrm{tr}(V_{\underline{\psi}}^{-1}V_i)\}_{(k+1)\times 1} = \{\mathrm{tr}(V_{\underline{\psi}}^{-1}V_iV_{\underline{\psi}}^{-1}V_j)\}_{(k+1)\times(k+1)}\underline{\psi}$.

proof: $\{\mathrm{tr}(V_{\underline{\psi}}^{-1}V_i)\} = \{\mathrm{tr}(V_{\underline{\psi}}^{-1}V_iV_{\underline{\psi}}^{-1}V_{\underline{\psi}})\} = \{\mathrm{tr}(V_{\underline{\psi}}^{-1}V_iV_{\underline{\psi}}^{-1}\sum_{i=1}^{k+1}\psi_jV_j)\}$

$= \{\sum_{i=1}^{k+1}\mathrm{tr}(V_{\underline{\psi}}^{-1}V_iV_{\underline{\psi}}^{-1}V_j)\psi_j\} = \{\mathrm{tr}(V_{\underline{\psi}}^{-1}V_iV_{\underline{\psi}}^{-1}V_j)\}\underline{\psi}$. ∎

The MLQEs will be maximum likelihood estimators when $\widehat{\underline{\beta}}_{\text{MLQ}} \in \mathcal{R}^p$, $\widehat{\underline{\psi}}_{\text{MLQ}} \in \Xi$, and $(\widehat{\underline{\beta}}_{\text{MLQ}}, \widehat{\underline{\psi}}_{\text{MLQ}})$ maximizes the likelihood equation. This thesis will focus on the MLQE, which does not have to be in the parameter space and does not have to be a maximum.

Restricted maximum likelihood estimation is another likelihood method for estimating $\underline{\psi} \in \mathcal{R}^{k+1}$. Define the matrix $Q_{n\times q}$ for $q = n - \underline{r}(X)$ which has columns that form an orthonormal basis for $\underline{R}(X)^{\perp}$. Then $Q'Q = I$ and $QQ' = I - P_X = N_X$. The following proposition gives properties of $F_{\underline{\psi}}$, which will be useful for this estimation method.

$F_{\underline{\psi}}$-Lemma: i) $F_{\widehat{\underline{\psi}}}\underline{Y} = V_{\widehat{\underline{\psi}}}^{-1}(\underline{Y} - X\widehat{\underline{\beta}})$.

ii) $\underline{R}(V_{\underline{\psi}}N_X) = \underline{R}(V_{\underline{\psi}}Q) = \underline{N}(X'V_{\underline{\psi}}^{-1})$.

iii) $F_{\underline{\psi}} = V_{\underline{\psi}}^{-1} - V_{\underline{\psi}}^{-1}X(X'V_{\underline{\psi}}^{-1}X)^{-}X'V_{\underline{\psi}}^{-1} = Q(Q'V_{\underline{\psi}}Q)^{-1}Q' = N_X(N_XV_{\underline{\psi}}N_X)^{-}N_X$.

proof: i) $F_{\widehat{\underline{\psi}}}\underline{Y} = V_{\widehat{\underline{\psi}}}^{-1}\underline{Y} - V_{\widehat{\underline{\psi}}}^{-1}X(X'V_{\widehat{\underline{\psi}}}^{-1}X)^{-}X'V_{\widehat{\underline{\psi}}}^{-1}\underline{Y} = V_{\widehat{\underline{\psi}}}^{-1}(\underline{Y} - X\widehat{\underline{\beta}})$ where $\widehat{\underline{\beta}}$ is given in (2).

ii) (1) $\underline{R}(V_{\underline{\psi}}Q) = V_{\underline{\psi}}[\underline{R}(Q)] = V_{\underline{\psi}}[\underline{R}(QQ')] = V_{\underline{\psi}}[\underline{R}(N_X)] = \underline{R}(V_{\underline{\psi}}N_X)$.

(2) Note $\underline{t} \in \underline{R}(V_{\underline{\psi}}Q) \Leftrightarrow V_{\underline{\psi}}^{-1}\underline{t} \in \underline{R}(Q) = \underline{R}(QQ') = \underline{N}(X') \Leftrightarrow \underline{t} \in \underline{N}(X'V_{\underline{\psi}}^{-1})$.

iii) By the general projection theorem in section 2.5 and ii),

(1) $V_{\underline{\psi}}Q(Q'V_{\underline{\psi}}Q)^{-1}Q'$ is a PO on $\underline{R}(V_{\underline{\psi}}Q) = \underline{N}(X'V_{\underline{\psi}}^{-1})$ along $\underline{N}(Q') = \underline{R}(X)$

(2) $V_{\underline{\psi}}N_X(N_XV_{\underline{\psi}}N_X)^{-}N_X$ is a PO on $\underline{R}(V_{\underline{\psi}}N_X) = \underline{N}(X'V_{\underline{\psi}}^{-1})$ along $\underline{N}(N_X) = \underline{R}(X)$

(3) $I - X(X'V_{\underline{\psi}}^{-1}X)^{-}X'V_{\underline{\psi}}^{-1}$ is a PO on $\underline{N}(X'V_{\underline{\psi}}^{-1})$ along $\underline{R}(X)$

$\Rightarrow$ (1) = (2) = (3). Left multiplying by $V_{\underline{\psi}}^{-1}$ gives the result. ∎

Under normality, $\underline{Y} \sim \mathrm{N}_n(X\underline{\beta}, V_{\underline{\psi}})$ and so $Q'\underline{Y} \sim \mathrm{N}_q(\underline{0}, Q'V_{\underline{\psi}}Q)$. The density of $\underline{Z} = Q'\underline{Y}$ and the likelihood function are given below assuming $V_{\underline{\psi}}$ is PD:

$$f(\underline{z}|\underline{\psi}) = (2\pi)^{-\frac{n}{2}}|Q'V_{\underline{\psi}}Q|\exp(\tfrac{-1}{2}\underline{z}'(Q'V_{\underline{\psi}}Q)^{-1}\underline{z})$$

$$l_R(\underline{\psi}) = \ln f(\underline{Z}|\underline{\psi}) = -\tfrac{n-m}{2}\ln 2\pi - \tfrac{1}{2}\ln|Q'V_{\underline{\psi}}Q| - \tfrac{1}{2}\underline{Z}'(Q'V_{\underline{\psi}}Q)^{-1}\underline{Z}.$$

The derivatives will be taken assuming [L] and [O]. The restricted maximum likelihood equations, restricted maximum likelihood equation estimators, and the information matrix can now be given (Searle et al.,1992,ch 6):

(4) <u>Restricted Maximum Likelihood (REML) Equations</u>

$$0 = \tfrac{\partial l_R(\underline{\psi})}{\partial \psi_i} = -\tfrac{1}{2}\mathrm{tr}\left((Q'V_{\underline{\psi}}Q)^{-1}Q'V_iQ\right) + \tfrac{1}{2}\underline{Y}'Q(Q'V_{\underline{\psi}}Q)^{-1}Q'V_iQ(Q'V_{\underline{\psi}}Q)^{-1}Q'\underline{Y}.$$

(5) <u>Restricted Maximum Likelihood Equation Estimators (REMLQE)</u> - the solution $\widehat{\underline{\psi}}_{\mathrm{REMLQ}} = \widehat{\underline{\psi}}$ given by

$$\{\mathrm{tr}(F_{\widehat{\underline{\psi}}}V_iF_{\widehat{\underline{\psi}}}V_j)\}_{(k+1)\times(k+1)}\widehat{\underline{\psi}} = \{\underline{Y}'F_{\widehat{\underline{\psi}}}V_iF_{\widehat{\underline{\psi}}}\underline{Y}\}_{(k+1)\times 1}.$$

(6) Information Matrix ($i_R(\underline{\psi})$)

$$u_{ij} = \frac{\partial^2 l_R(\underline{\psi})}{\partial\psi_i\partial\psi_j} = \tfrac{1}{2}\mathrm{tr}((Q'V_{\underline{\psi}}Q)^{-1}Q'V_iQ\,(Q'V_{\underline{\psi}}Q)^{-1}Q'V_jQ)$$
$$-\underline{Y}'Q(Q'V_{\underline{\psi}}Q)^{-1}Q'V_iQ(Q'V_{\underline{\psi}}Q)^{-1}Q'V_jQ(Q'V_{\underline{\psi}}Q)^{-1}Q'\underline{Y}\,.$$

$$i_R(\underline{\psi}) = -\{\mathrm{E}[u_{ij}]\} = \{\tfrac{1}{2}\mathrm{tr}((Q'V_{\underline{\psi}}Q)^{-1}Q'V_iQ(Q'V_{\underline{\psi}}Q)^{-1}Q'V_jQ))\} = \{\tfrac{1}{2}\mathrm{tr}(F_{\widehat{\underline{\psi}}}V_iF_{\widehat{\underline{\psi}}}V_j)\}_{(k+1)\times(k+1)}.$$

The REML equations in (4) and the information matrix in (6) can be obtained using the derivative rules and the $F_{\underline{\psi}}$-lemma. The REML equations are identical to the MINVAR and the iterated MINQUE equations (Searle et al.,1992,section 11.3). The REMLQE for the variance components solves the equation in (5), which has been re-expressed using the following proposition.

Proposition: $\{\mathrm{tr}((Q'V_{\underline{\psi}}Q)^{-1}Q'V_iQ)\}_{(k+1)\times 1} = \{\mathrm{tr}(F_{\underline{\psi}}V_i)\}_{(k+1)\times 1} = \{\mathrm{tr}(F_{\underline{\psi}}V_iF_{\underline{\psi}}V_j)\}_{(k+1)\times(k+1)}\underline{\psi}\,.$

proof: By the general projection theorem in section 2.5, $F_{\underline{\psi}}V_{\underline{\psi}}$ is a PO, so
$(F_{\underline{\psi}}V_{\underline{\psi}})^2 = F_{\underline{\psi}}V_{\underline{\psi}} \Rightarrow F_{\underline{\psi}}V_{\underline{\psi}}F_{\underline{\psi}} = F_{\underline{\psi}}$ $(*)$. Then
$\{\mathrm{tr}((Q'V_{\underline{\psi}}Q)^{-1}Q'V_iQ\,)\} = \{\mathrm{tr}(F_{\underline{\psi}}V_i)\}$ since $F_{\underline{\psi}} = Q(Q'V_{\underline{\psi}}Q)^{-1}Q'$ by the $F_{\underline{\psi}}$-lemma
$= \{\mathrm{tr}(F_{\underline{\psi}}V_{\underline{\psi}}F_{\underline{\psi}}V_i)\} = \{\mathrm{tr}(F_{\underline{\psi}}V_iF_{\underline{\psi}}V_{\underline{\psi}})\}$ by $(*)$ and symmetry of trace operator
$= \{\mathrm{tr}(F_{\underline{\psi}}V_iF_{\underline{\psi}}\sum_{j=1}^{k+1}\psi_jV_j)\} = \{\sum_{i=1}^{k+1}\mathrm{tr}(F_{\underline{\psi}}V_iF_{\underline{\psi}}V_j)\psi_j\} = \{\mathrm{tr}(F_{\underline{\psi}}V_iF_{\underline{\psi}}V_j)\}\underline{\psi}.$ ∎

The REMLQE will be a restricted maximum likelihood estimator when $\widehat{\underline{\psi}}_{\text{REMLQ}} \in \Xi$ and $\widehat{\underline{\psi}}_{\text{REMLQ}}$ maximizes the restricted likelihood equation. This thesis will focus on the REMLQE, which may not be in the parameter space and does not have to be a maximum.

Maximum likelihood and restricted maximum likelihood estimation for the Y-Model under [L], [O], and [N] are of main interest for this study. Linearized quadratic estimation models will be defined in the next section to represent the likelihood equations in a convenient form.

3.2. Linearized Quadratic Estimation Models

3.2.1. Definitions

Linearized quadratic estimation models (LQEMs) will provide a modelling framework in which to conduct quadratic estimation. These models have been called dispersion-mean models by Searle et al. (1992) and were introduced by Seely (1971). They are useful for estimating variance components in the

Y-Model under [L], [O], and [N]. Some preliminary results are needed before presenting the models in this section. The next lemma gives the expectation and covariance of particular quadratic forms.

(Schott,1997)

Lemma: Let $\underline{Y} \sim N_n(\underline{\mu}, V)$ and A, B be symmetric $n \times n$ matrices. Then

i) $\mathrm{E}[\underline{Y}'A\underline{Y}] = \mathrm{tr}(AV) + \underline{\mu}'A\underline{\mu}$ ii) $\mathrm{Cov}(\underline{Y}'A\underline{Y}, \underline{Y}'B\underline{Y}) = 4\underline{\mu}'AVB\underline{\mu} + 2\mathrm{tr}(AVBV)$.

A special linear transformation will also be used. Consider $\Psi_D : \mathcal{S}_n \rightarrow \mathcal{S}_n$ given by $\Psi_D(A) = DA\,D$. The next propositon gives some properties of this mapping.

Proposition: i) Ψ_D is linear ii) $\Psi_D^* = \Psi_D$ iii) $\Psi_D^{-1} = \Psi_{D^{-1}}$ where D is invertible.

proof: i) For $A, B \in \mathcal{S}_n$ and $\alpha, \beta \in \mathcal{R}$,

$\Psi_D(\alpha A + \beta B) = D(\alpha A + \beta B)D = \alpha DAD + \beta DB\,D = \alpha\Psi_D(A) + \beta\Psi_D(B)$.

ii) $< \Psi_D(A), B > \; = \mathrm{tr}(\Psi_D(A)B) = \mathrm{tr}(DA\,DB) = \mathrm{tr}(ADBD)$

$= \mathrm{tr}(A\Psi_D(B)) = \; < A, \Psi_D(B) > \; \Rightarrow \Psi$ is self-adjoint.

iii) $\forall\, A \in \mathcal{S}_n$ $\Psi_{D^{-1}}\Psi_D(A) = D^{-1}(DA\,D)D^{-1} = A \Rightarrow \Psi_{D^{-1}}\Psi_D = I$ as true $\forall\ A \in \mathcal{S}_n$.

Thus, $\Psi_D^{-1} = \Psi_{D^{-1}}$ by definition of inverse. ■

Linearized quadratic estimation models are defined using a quadratic form $Y^\dagger = \underline{Z}\underline{Z}' \in \mathcal{S}_n$ where $\underline{Z} \sim N_n(\underline{0}, R_{\underline{\psi}})$ and $R_{\underline{\psi}} = \sum_{i=1}^{r} \psi_i R_i$ is a matrix having linear structure with $\underline{\psi} \in \Xi$. In addition, define the linear transformation $X^\dagger : \mathcal{R}^r \rightarrow \mathcal{S}_n$ by $X^\dagger \underline{u} = \sum_{i=1}^{r} u_i R_i$ and the mapping $V_{\underline{\psi}}^\dagger : \mathcal{S}_n \rightarrow \mathcal{S}_n$ by $V_{\underline{\psi}}^\dagger = 2\Psi_{R_{\underline{\psi}}}$. The next lemma indicates how these models are constructed.

(Seely,1971)

Lemma: $\mathrm{E}[Y^\dagger] \in \mathcal{U}^\dagger = \{X^\dagger \underline{\psi} \,|\, \underline{\psi} \in \Xi\}$ and $\mathrm{Cov}(Y^\dagger) \in \mathcal{V}^\dagger = \{V_{\underline{\psi}}^\dagger \,|\, \underline{\psi} \in \Xi\}$.

proof: i) $\mathrm{E}[Y^\dagger] = \mathrm{Cov}(\underline{Z}) + \mathrm{E}[\underline{Z}]\mathrm{E}[\underline{Z}]' = \sum_{i=1}^{r} \psi_i R_i = R_{\underline{\psi}} = X^\dagger \underline{\psi}$.

ii) Consider symmetric matrices A and B. Then using the trace inner product gives:

$\mathrm{Cov}(< A, Y^\dagger > , < B, Y^\dagger >) = \mathrm{Cov}(\mathrm{tr}(AY^\dagger), \mathrm{tr}(BY^\dagger))$

$= \mathrm{Cov}(\underline{Z}'A\underline{Z}, \underline{Z}'B\underline{Z}) = 2\mathrm{tr}(AR_{\underline{\psi}}BR_{\underline{\psi}})$ from above lemma

$= \; < A, 2R_{\underline{\psi}}BR_{\underline{\psi}} > \; = \; < A, \mathrm{Cov}(Y^\dagger)B > \; \Rightarrow \mathrm{Cov}(Y^\dagger)B = 2\Psi_{R_{\underline{\psi}}}(B) \Rightarrow \mathrm{Cov}(Y^\dagger) = 2\Psi_{R_{\underline{\psi}}} = V_{\underline{\psi}}^\dagger$.

iii) From i) and ii), $\mathrm{E}[Y^\dagger] \in \mathcal{U}^\dagger = \{X^\dagger \underline{\psi} \,|\, \underline{\psi} \in \Xi\}$ and $\mathrm{Cov}(Y^\dagger) \in \mathcal{V}^\dagger = \{V_{\underline{\psi}}^\dagger \,|\, \underline{\psi} \in \Xi\}$. ■

This lemma provides the definition for the linearized quadratic estimation model for $\underline{Z}$. This general model is summarized below.

LQEM for $\underline{Z}$: $\mathrm{E}[Y^{\dagger}] \in \mathcal{U}^{\dagger} = \{X^{\dagger}\underline{\psi} \mid \underline{\psi} \in \Xi\}$ $\quad$ $\mathrm{Cov}(Y^{\dagger}) \in \mathcal{V}^{\dagger} = \{V_{\underline{\psi}}^{\dagger} \mid \underline{\psi} \in \Xi\}$.

This model is not a linear model as defined by Seely unless $\mathcal{U}^{\dagger} = \mathrm{sp}\,\mathcal{U}^{\dagger}$. It is a linearly mean-parameterized model. The linearized part of the LQEM refers to Seely's notion of linearizing the expectation with respect to the parameter using a quadratic transformation of the original response vector.

Suppose Ξ contains a non-empty open set of dimension r. By the linearity of $X^{\dagger}$, $\mathrm{sp}\mathcal{U}^{\dagger} = \mathrm{sp}X^{\dagger}(\Xi) = X^{\dagger}(\mathrm{sp}\,\Xi) = X^{\dagger}(\mathcal{R}^{r}) = \underline{R}(X^{\dagger})$. Then the LQEM is a linear model when $\mathcal{U}^{\dagger} = \underline{R}(X^{\dagger})$. In addition, there is a functional relationship between the mean and the variance in the LQEM. It will be shown in section 3.3 that this is not a problem in this study. Also, the parametric vector $\underline{\psi}$ is estimable if and only if the R_i's are linearly independent.

Specific LQEMs are of interest which can be used to generate equations that correspond to the likelihood equations. These models are based on the Y-Model assuming [L], [O], and [N]. Four such models are stated below:

LQEM for $(\underline{Y} - X\underline{\beta})$: Let $\underline{Z} = \underline{Y} - X\underline{\beta}$ and $Y_1^{\circ} = \underline{Z}\underline{Z}'$. In addition, define $X^{\circ}\colon \mathcal{R}^{k+1} \to \mathcal{S}_n$ by $X^{\circ}\underline{u} = \sum_{i=1}^{k+1} u_i V_i$ and $V_{\underline{\psi}}^{\circ} : \mathcal{S}_n \to \mathcal{S}_n$ by $V_{\underline{\psi}}^{\circ} = 2\Psi_{V_{\underline{\psi}}}$. Then

$\mathrm{E}[Y_1^{\circ}] \in \mathcal{U}^{\circ} = \{X^{\circ}\underline{\psi} \mid \underline{\psi} \in \Xi\}$, $\mathrm{Cov}(Y_1^{\circ}) \in \mathcal{V}^{\circ} = \{V_{\underline{\psi}}^{\circ} \mid \underline{\psi} \in \Xi\}$.

ALQEM for $(\underline{Y} - X\widehat{\underline{\beta}})$: Let $\underline{Z} = \underline{Y} - X\widehat{\underline{\beta}}$ and $Y_2^{\circ} = \underline{Z}\underline{Z}'$ where $X\widehat{\underline{\beta}} = X(X'V_{\widehat{\underline{\psi}}}^{-1}X)^{-}X'V_{\widehat{\underline{\psi}}}^{-1}\underline{Y}$. Instead of using the true distribution of Y_2°, artificially assume the same model as above,

$\mathrm{E}[Y_2^{\circ}] \in \mathcal{U}^{\circ} = \{X^{\circ}\underline{\psi} \mid \underline{\psi} \in \Xi\}$, $\mathrm{Cov}(Y_2^{\circ}) \in \mathcal{V}^{\circ} = \{V_{\underline{\psi}}^{\circ} \mid \underline{\psi} \in \Xi\}$.

LQEM for $N_X\underline{Y}$: Let $\underline{Z} = N_X\underline{Y}$ and $Y^{\circ} = \underline{Z}\underline{Z}'$. In addition, define $X^{\circ} : \mathcal{R}^{k+1} \to \mathcal{S}_n$ by $X^{\circ}\underline{u} = \sum_{i=1}^{k+1} u_i N_X V_i N_X$ and $V_{\underline{\psi}}^{\circ} : \mathcal{S}_n \to \mathcal{S}_n$ by $V_{\underline{\psi}}^{\circ} = 2\Psi_{N_X V_{\underline{\psi}} N_X}$. Then

$\mathrm{E}[Y^{\circ}] \in \mathcal{U}^{\circ} = \{X^{\circ}\underline{\psi} \mid \underline{\psi} \in \Xi\}$, $\mathrm{Cov}(Y^{\circ}) \in \mathcal{V}^{\circ} = \{V_{\underline{\psi}}^{\circ} \mid \underline{\psi} \in \Xi\}$.

LQEM for $Q'\underline{Y}$: Let $\underline{Z} = Q'\underline{Y}$ and $Y^{\triangleright} = \underline{Z}\underline{Z}'$ for $Q_{n\times q} \ni QQ' = N_X$, $Q'Q = I$, $\underline{r}(Q) = q$. Define $X^{\triangleright} : \mathcal{R}^{k+1} \to \mathcal{S}_q$ by $X^{\triangleright}\underline{u} = \sum_{i=1}^{k+1} u_i Q' V_i Q$ and $V_{\underline{\psi}}^{\triangleright} : \mathcal{S}_q \to \mathcal{S}_q$ by $V_{\underline{\psi}}^{\triangleright} = 2\Psi_{Q' V_{\underline{\psi}} Q}$. Then

$\mathrm{E}[Y^{\triangleright}] \in \mathcal{U}^{\triangleright} = \{X^{\triangleright}\underline{\psi} \mid \underline{\psi} \in \Xi\}$, $\mathrm{Cov}(Y^{\triangleright}) \in \mathcal{V}^{\triangleright} = \{V_{\underline{\psi}}^{\triangleright} \mid \underline{\psi} \in \Xi\}$.

The LQEM for $(\underline{Y} - X\underline{\beta})$ can be used theoretically whether or not $\underline{\beta}$ is unknown. However, it cannot be directly applied when $\underline{\beta}$ is unknown. The ALQEM for $(\underline{Y} - X\widehat{\underline{\beta}})$ can be applied when $\underline{\beta}$ is unknown, given an estimate $\widehat{\underline{\psi}}$. This model is artificial (A), since it assumes the expectation and covariance of $\underline{Z}\underline{Z}'$ corresponds to the LQEM for $(\underline{Y} - X\underline{\beta})$ rather than the true expectation and covariance of $(\underline{Y} - X\widehat{\underline{\beta}})(\underline{Y} - X\widehat{\underline{\beta}})'$. This is important to remember when examining the unbiasedness of estimators with respect to this model.

A special case of the ALQEM for $(\underline{Y} - X\widehat{\underline{\beta}})$ that will be of interest is when $\underline{Y} - X\widehat{\underline{\beta}} = N_X\underline{Y}$. In this case, the model will be identified as the ALQEM for $N_X\underline{Y}$. This model is still artificial, which differentiates it from the LQEM for $N_X\underline{Y}$.

The LQEM for $N_X\underline{Y}$ and the LQEM for $Q'\underline{Y}$ are essentially the same for estimation purposes. The differences between these models are explored in the next section. The models of primary interest are the LQEM for $\underline{Z}$, the ALQEM for $(\underline{Y} - X\widehat{\underline{\beta}})$, the ALQEM for $N_X\underline{Y}$, and the LQEM for $N_X\underline{Y}$. The LQEM for $\underline{Z}$ is useful, as it incorporates the other LQEMs. The ALQEM for $(\underline{Y} - X\widehat{\underline{\beta}})$ will be used for the maximum likelihood method where $\underline{\beta}$ is unknown and the LQEM for $N_X\underline{Y}$ will be used for the restricted maximum likelihood method. These models are further examined in the following sections.

3.2.2. Covariance Properties

This study requires that the covariance be positive definite. The next proposition illustrates that the covariance matrices for the LQEMs do satisfy this property. For a linear space $\mathcal{W}$, let $\mathcal{L}_{\text{PD}}(\mathcal{W}, \mathcal{W})$ denote the set of positive definite transformations from $\mathcal{W} \to \mathcal{W}$ and consider the following two propositions.

Proposition: Let V be a PD matrix and $\mathcal{K} = \{N_XAN_X \mid A \in \mathcal{S}_n\}$. Then

i) $\Psi_V \in \mathcal{L}_{\text{PD}}(\mathcal{S}_n, \mathcal{S}_n)$ ii) $\Psi_{Q'VQ} \in \mathcal{L}_{\text{PD}}(\mathcal{S}_q, \mathcal{S}_q)$ iii) $\Psi_{N_XVN_X} \in \mathcal{L}_{\text{PD}}(\mathcal{K}, \mathcal{K})$.

proof: Since V is PD matrix $\Rightarrow \exists\, B \ni V = BB'$. Also, by definition of adjoint and inner product,
$< \Psi_{GG'}(A), A > \; = \; < GG'AGG', A > \; = \; < G'AG, G'AG > \; \geq 0 \Rightarrow \Psi_{GG'}$ is NND (1).
Also, $\Psi_{GG'}(A) = 0 \Rightarrow GG'AGG' = 0 \Rightarrow A = 0$ providing GG' is invertible (2).
i) Set $G = B \Rightarrow \Psi_V$ is NND by (1). Also, $GG' = V$ is invertible $\Rightarrow \Psi_V$ is invertible.
ii) Set $G = Q'B \Rightarrow \Psi_{Q'VQ}$ is NND. Also, $GG' = Q'VQ$ is invertible since Q has full rank $\Rightarrow \Psi_{Q'VQ}$ is invertible.
iii) Set $G = N_XB \Rightarrow \Psi_{N_XVN_X}$ is NND. Also, for $N_XAN_X \in \mathcal{K}$
$\Psi_{N_XVN_X}(N_XAN_X) = 0 \Rightarrow N_XVN_XAN_XVN_X = 0 \Rightarrow (N_XVN_XAN_XVN_X)AN_XVN_X = 0$
$\Rightarrow N_XVN_XAN_X = 0$ since $\underline{R}(N_XVN_X) = \underline{R}(N_X)$ when V is PD
$\Rightarrow N_XAN_XVN_XAN_X = 0 \Rightarrow N_XAN_X = 0$ since $\underline{R}(N_XVN_X) = \underline{R}(N_X)$ when V is PD.
Thus, $\Psi_{N_XVN_X}$ is invertible $\forall\; N_XAN_X \in \mathcal{K}$. ∎

Proposition: For the PD matrix V, $\Psi^{-1}_{N_XVN_X}(D) = \Psi_{(N_XVN_X)^+}(D) \quad \forall\ D \in \mathcal{K}$.

proof: i) Since V is PD matrix $\Rightarrow \exists\ B \ni V = BB'$. Note $\underline{N}(N_X) \subset \underline{N}(N_XVN_X)$. Also, suppose $\underline{t} \in \underline{N}(N_XVN_X) \Rightarrow N_XVN_X\underline{t} = 0 \Rightarrow N_XBB'N_X\underline{t} = 0 \Rightarrow B'N_X\underline{t} = 0 \Rightarrow VN_X\underline{t} = 0$
$\Rightarrow N_X\underline{t} = V^{-1}0 = 0 \Rightarrow \underline{t} \in \underline{N}(N_X) \Rightarrow \underline{N}(N_XVN_X) \subset \underline{N}(N_X)$. Hence,
$\underline{N}(N_XVN_X) = \underline{N}(N_X) \Rightarrow \underline{R}(N_XVN_X) = \underline{R}(N_X)$.
ii) Note $\forall\ D \in \mathcal{K} \Rightarrow D = N_XAN_X$ for some $A \in \mathcal{S}_n$. Then

$$\Psi_{(N_XVN_X)}\Psi_{(N_XVN_X)^+}(D) = (N_XVN_X)(N_XVN_X)^+(N_XAN_X)(N_XVN_X)^+(N_XVN_X)$$
$$= P_{(N_XVN_X)}(N_XAN_X)P_{(N_XVN_X)} = P_{N_X}(N_XAN_X)P_{N_X} = N_XAN_X = D$$

as $\underline{R}(N_XVN_X) = \underline{R}(N_X)$ when V is PD. Thus, $\Psi^{-1}_{N_XVN_X}(D) = \Psi_{(N_XVN_X)^+}(D) \quad \forall\ D \in \mathcal{K}$. ■

Since the LQEMs have a covariance which can be treated as positive definite, it will be assumed that the LQEM for $\underline{Z}$ has a positive definite covariance. The above proposition also illustrates the issue between using the LQEM for $N_X\underline{Y}$ and the LQEM for $Q'\underline{Y}$. Since the matrix N_X does not have full rank, the linear transformation $\Psi_{N_XVN_X}$ needs to be restricted to $\mathcal{K}$ in order to be invertible as $\Psi_{N_XVN_X}(P_X) = 0$ where $P_X \in \mathcal{S}_n$. This should not be a problem, since matrices of the form N_XVN_X are of primary interest. In addition, the identity matrix is not a possible covariance matrix for this model. This also is not a problem for discussing least square estimators and uniformly best linear unbiased estimators. The results in this study could be applied to the LQEM for $Q'\underline{Y}$. The decision of which to use is a matter of preference. Even though Q has full rank and the identity matrix is a possible covariance matrix, the LQEM for $N_X\underline{Y}$ seems easier to work with in applications.

3.2.3. Relation to Likelihood Estimation

This section will demonstrate the usefulness of the LQEMs. These models were defined in order to easily represent the likelihood equations. A preliminary result is needed to represent particular linear transformations for the LQEMs.

Lemma 1: Consider the mappings $X^\dagger$ and $V^\dagger_{\underline{\psi}}$ defined for the LQEM for $\underline{Z}$ in section 3.2.1. Then

i) $X^{\dagger *} : \mathcal{S}_n \to \mathcal{R}^r$ where $\{X^{\dagger *}B\}_i = \text{tr}(R_iB)$ for $i = 1, ..., r$ and $B \in \mathcal{S}_n$
ii) $X^{\dagger *}V^{\dagger -1}_{\underline{\psi}}Y^\dagger = \frac{1}{2}\{\underline{Z}'R^{-1}_{\underline{\psi}}R_iR^{-1}_{\underline{\psi}}\underline{Z}\}_{r\times 1}$
iii) $X^{\dagger *}V^{\dagger -1}_{\underline{\psi}}X^\dagger = \frac{1}{2}\{\text{tr}(R^{-1}_{\underline{\psi}}R_iR^{-1}_{\underline{\psi}}R_j\}_{r\times r}$
iv) $X^{\dagger *}X^\dagger = \frac{1}{2}\{\text{tr}(R_iR_j\}_{r\times r}$.

proof: i) Let $c_i = \{X^{\dagger *}B\}_i$ be the i^{th} element of $X^{\dagger *}B$. Hence,

$$< X^{\dagger}\underline{\psi}, B > \; = \; < \underline{\psi}, X^{\dagger *}B > \; \Rightarrow \mathrm{tr}((X^{\dagger}\underline{\psi})B) = \underline{\psi}' X^{\dagger *}B = \underline{\psi}'\{c_i\}$$

$$\Rightarrow \sum_{i=1}^{r} \psi_i \mathrm{tr}(R_i B) = \sum_{i=1}^{r} \psi_i c_i \Rightarrow c_i = \{X^{\dagger *}B\}_i = \mathrm{tr}(R_i B) \quad i = 1, ..., r.$$

ii) iii) For $A \in \mathcal{S}_n$ the results follow from i) where $X^{\dagger *}V_{\underline{\psi}}^{\dagger -1}A = \frac{1}{2}X^{\dagger *}R_{\underline{\psi}}^{-1}AR_{\underline{\psi}}^{-1} = \frac{1}{2}\{\mathrm{tr}(R_i R_{\underline{\psi}}^{-1}AR_{\underline{\psi}}^{-1})\}$.

iv) Follows from iii) where $V_{\underline{\psi}}^{\dagger} = I$. ∎

Corollary: If $R_i \; i = 1, ..., r$ are linearly independent, then $X^{\dagger *}X^{\dagger}$ and $X^{\dagger *}V_{\underline{\psi}}^{\dagger -1}X^{\dagger}$ are invertible.

proof: Define $R_{\underline{a}} = \sum_j a_j R_j$.

i) Let $M = \{\mathrm{tr}(R_i R_j)\}$. Then $M\underline{a} = \underline{0} \Rightarrow \sum_j \mathrm{tr}(R_i R_j)\, a_j = 0 \;\; \forall \; i \;\; \Rightarrow \mathrm{tr}(R_i R_{\underline{a}}) = 0 \;\; \forall \; i$

$\Rightarrow \sum_i a_i \mathrm{tr}(R_i R_{\underline{a}}) = 0 \Rightarrow \mathrm{tr}(R_{\underline{a}} R_{\underline{a}}) = 0 \Rightarrow R_{\underline{a}} R_{\underline{a}} = 0$ as $R_{\underline{a}}$ NND since sum of NND matrices is NND

$\Rightarrow R_{\underline{a}} = 0 \Rightarrow \underline{a} = 0$ since R_i's are linearly independent.

ii) Let $M = \{\mathrm{tr}(R_{\underline{\psi}}^{-1}R_i R_{\underline{\psi}}^{-1}R_j)\}$. Then $M\underline{a} = \underline{0}$

$\Rightarrow \mathrm{tr}(R_{\underline{\psi}}^{-1}R_{\underline{a}}R_{\underline{\psi}}^{-1}R_{\underline{a}}) = 0$ using same techniques as in i)

$\Rightarrow R_{\underline{\psi}}^{-1}R_{\underline{a}}R_{\underline{\psi}}^{-1}R_{\underline{a}} = 0$ as $R_{\underline{a}}R_{\underline{\psi}}^{-1}R_{\underline{a}}$ and $R_{\underline{\psi}}^{-1}$ NND

$\Rightarrow R_{\underline{\psi}}^{-1}R_{\underline{a}} \Rightarrow R_{\underline{a}} = 0 \Rightarrow \underline{a} = 0$ since R_i's are linearly independent. ∎

The ALQEM for $(\underline{Y} - X\widehat{\underline{\beta}})$ defined in section 3.2.1 can be used to obtain the maximum likelihood equations for estimating $\underline{\psi}$ when $\underline{\beta}$ is unknown. This is demonstrated in the next theorem.

ML Theorem: The ML equations for $\widehat{\underline{\psi}}_{\mathrm{MLQ}} = \widehat{\underline{\psi}}$ are given by $X^{\diamond *}V_{\widehat{\underline{\psi}}}^{\diamond -1}X^{\diamond}\widehat{\underline{\psi}} = X^{\diamond *}V_{\widehat{\underline{\psi}}}^{\diamond -1}Y_2^{\diamond}$.

proof: $X^{\diamond *}V_{\widehat{\underline{\psi}}}^{\diamond -1}X^{\diamond}\widehat{\underline{\psi}} = X^{\diamond *}V_{\widehat{\underline{\psi}}}^{\diamond -1}(\underline{Y} - X\widehat{\underline{\beta}})(\underline{Y} - X\widehat{\underline{\beta}})'$ where $\widehat{\underline{\beta}} = \widehat{\underline{\beta}}_{\mathrm{MLQ}}$

$\Leftrightarrow \frac{1}{2}\{\mathrm{tr}(V_i V_{\widehat{\underline{\psi}}}^{-1}V_{\widehat{\underline{\psi}}}V_{\widehat{\underline{\psi}}}^{-1}) = \frac{1}{2}\{\mathrm{tr}(V_i V_{\widehat{\underline{\psi}}}^{-1}(\underline{Y} - X\widehat{\underline{\beta}})(\underline{Y} - X\widehat{\underline{\beta}})'V_{\widehat{\underline{\psi}}}^{-1})\}$ by lemma 1

$\Leftrightarrow \{\mathrm{tr}(V_{\widehat{\underline{\psi}}}^{-1}V_i)\} = \{(\underline{Y} - X\widehat{\underline{\beta}})'V_{\widehat{\underline{\psi}}}^{-1}V_i V_{\widehat{\underline{\psi}}}^{-1}(\underline{Y} - X\widehat{\underline{\beta}})\}$

$\Leftrightarrow \{\mathrm{tr}(V_{\widehat{\underline{\psi}}}^{-1}V_i V_{\widehat{\underline{\psi}}}^{-1}V_j)\}\widehat{\underline{\psi}} = \{\underline{Y}'F_{\widehat{\underline{\psi}}}V_i F_{\widehat{\underline{\psi}}}\underline{Y}\}$ by proposition after ML equations and $F_{\underline{\psi}}$-lemma. ∎

Proposition: $X^{\circ *}X^{\circ}\widehat{\underline{\psi}} = X^{\circ *}Y_2^{\circ} \Leftrightarrow \{\mathrm{tr}(V_i V_j)\}_{(k+1)\times(k+1)}\widehat{\underline{\psi}} = \{\underline{Y}'N_X V_i N_X \underline{Y}\}_{(k+1)\times 1}$ when $\underline{Y} - X\widehat{\underline{\beta}} = N_X\underline{Y}$.

proof: $X^{\circ *}X^{\circ}\widehat{\underline{\psi}} = X^{\circ *}N_X\underline{Y}\underline{Y}'N_X \Leftrightarrow \{\mathrm{tr}(V_i V_{\widehat{\underline{\psi}}})\} = \{\mathrm{tr}(V_i N_X \underline{Y}\underline{Y}'N_X)\}$ by lemma 1

$\Leftrightarrow \{\mathrm{tr}(V_i V_j)\}\widehat{\underline{\psi}} = \{\underline{Y}'N_X V_i N_X \underline{Y}\}$ by proposition after ML equations. ∎

The equations in the above proposition will be of interest since they do not depend on the covariance. Thus, these equations are linear and explicit so that they can be solved without the use of an iterative procedure and they do not depend on any other unknown parameters. A goal in this study will be to characterize when the ML equations are equivalent to the equations given in the above proposition.

For the case where $\underline{\beta}$ is known, the ML equations would be obtained using the LQEM for $(\underline{Y} - X\underline{\beta})$. The equations would be given by $X^{\diamond *}V_{\widehat{\underline{\psi}}}^{\diamond -1}X^{\diamond}\widehat{\underline{\psi}} = X^{\diamond *}V_{\widehat{\underline{\psi}}}^{\diamond -1}Y_1^{\diamond}$.

The LQEM for $N_X\underline{Y}$ defined in section 3.2.1 can be used to obtain the restricted maximum likelihood equations for estimating $\underline{\psi}$. This is demonstrated in the next theorem. Additional propositions will be given which are related to this theorem.

REML Theorem: The REML equations for $\widehat{\underline{\psi}}_{\text{REMLQ}} = \widehat{\underline{\psi}}$ are given by $X^{\circ *}V_{\widehat{\underline{\psi}}}^{\circ -1}X^{\circ}\widehat{\underline{\psi}} = X^{\circ *}V_{\widehat{\underline{\psi}}}^{\circ -1}Y^{\circ}$.

proof: $X^{\circ *}V_{\widehat{\underline{\psi}}}^{\circ -1}X^{\circ}\widehat{\underline{\psi}} = X^{\circ *}V_{\widehat{\underline{\psi}}}^{\circ -1}N_X\underline{Y}\underline{Y}'N_X \Leftrightarrow \frac{1}{2}\{\text{tr}(N_XV_iN_X(N_XV_{\widehat{\underline{\psi}}}N_X)^+N_XV_{\widehat{\underline{\psi}}}N_X(N_XV_{\widehat{\underline{\psi}}}N_X)^+)\}$

$= \frac{1}{2}\{\text{tr}(N_XV_iN_X(N_XV_{\widehat{\underline{\psi}}}N_X)^+N_X\underline{Y}\underline{Y}'N_X(N_XV_{\widehat{\underline{\psi}}}N_X)^+)\}$ by lemma 1

$\Leftrightarrow \{\text{tr}((N_XV_{\widehat{\underline{\psi}}}N_X)^+N_XV_iN_X)\} = \{\underline{Y}'N_X(N_XV_{\widehat{\underline{\psi}}}N_X)^+N_XV_iN_X(N_XV_{\widehat{\underline{\psi}}}N_X)^+N_X\underline{Y}\}$

$\Leftrightarrow \{\text{tr}(F_{\widehat{\underline{\psi}}}V_i)\} = \{\underline{Y}'F_{\widehat{\underline{\psi}}}V_iF_{\widehat{\underline{\psi}}}\underline{Y}\}$ by the $F_{\underline{\psi}}$-lemma

$\Leftrightarrow \{\text{tr}(F_{\widehat{\underline{\psi}}}V_iF_{\widehat{\underline{\psi}}}V_j)\}\widehat{\underline{\psi}} = \{\underline{Y}'F_{\widehat{\underline{\psi}}}V_iF_{\widehat{\underline{\psi}}}\underline{Y}\}$ by the proposition after REML equations. ∎

Proposition: $X^{\circ *}X^{\circ}\widehat{\underline{\psi}} = X^{\circ *}Y^{\circ} \Leftrightarrow \{\text{tr}(V_iN_XV_jN_X)\}_{(k+1)\times(k+1)}\widehat{\underline{\psi}} = \{\underline{Y}'N_XV_iN_X\underline{Y}\}_{(k+1)\times 1}$.

proof: $X^{\circ *}X^{\circ}\widehat{\underline{\psi}} = X^{\circ *}N_X\underline{Y}\underline{Y}'N_X$

$\Leftrightarrow \frac{1}{2}\{\text{tr}(N_XV_iN_XN_XV_{\widehat{\underline{\psi}}}N_X)\} = \frac{1}{2}\{\text{tr}(N_XV_iN_X\underline{Y}\underline{Y}'N_X)\}$ by lemma 1

$\Leftrightarrow \{\text{tr}(V_iN_XV_jN_X)\}\widehat{\underline{\psi}} = \{\underline{Y}'N_XV_iN_X\underline{Y}\}$ by proposition after REML equations. ∎

The equations in the above proposition will be of interest since they do not depend on the covariance. Thus, these equations are linear and explicit, so they can be solved without the use of an iterative procedure and do not depend on any other unknown parameters. These equations are identical to the MINQUE0 or MIVQUE0 equations which can be obtained from the REML equations by plugging in $\underline{\psi}_0 = \{\psi_{i0}\}$ where $\psi_{i0} = \begin{cases} 0 & i = k \\ 1 & i = k+1 \end{cases}$ (Searle et al.,1992,section 11.3). A goal in this study will be to characterize when the REML equations given in the REML theorem are equivalent to the equations given in the above proposition.

The REML equations could also have been obtained using the LQEM for $Q'\underline{Y}$ given in section 3.2.1. The equations under this model would have the form $X^{\triangleright *}V_{\widehat{\underline{\psi}}}^{\triangleright -1}X^{\triangleright}\widehat{\underline{\psi}} = X^{\triangleright *}V_{\widehat{\underline{\psi}}}^{\triangleright -1}Y^{\triangleright}$.

3.3. The Underlying Model

3.3.1. Definitions

This section will establish an underlying model which incorporates those models presented in the previous sections. This will provide a convenient tool, since the results can be presented with respect to this model and applied to the other models as special cases. Thus, the results of this section can be applied to any of the previous sections in this chapter. A set of useful inner product spaces and linear transformations are listed below:

Spaces

- $(\mathcal{W}, <\,,\,>_{\mathcal{W}})$ = n-dimensional real inner product space = observation space
- $(\mathcal{P}, <\,,\,>_{\mathcal{P}})$ = p-dimensional real inner product space = mean parameter space
- $(\mathcal{H}, <\,,\,>_{\mathcal{H}})$ = h-dimensional real inner product space = estimation space
- $\mathcal{E}$ = linear subspace of $\mathcal{W}$ = expectation space

Linear Transformations

- $U : \mathcal{P} \to \mathcal{W}$ $\quad U^* : \mathcal{W} \to \mathcal{P}$ $\quad \underline{R}(U) = \mathcal{E}$
- $\Pi : \mathcal{H} \to \mathcal{P}$ $\quad \Pi^* : \mathcal{P} \to \mathcal{H}$
- $H : \mathcal{H} \to \mathcal{W}$ $\quad H^* : \mathcal{W} \to \mathcal{H}.$

Suppose $\omega \in \mathcal{W}$ is a random response. The expectation, $\mathrm{E}[\omega]$, and the covariance, $\mathrm{Cov}(\omega)$, are uniquely defined by:

$$\mathrm{E}[\omega] \text{ satisfies } \mathrm{E}[<a,\omega>_{\mathcal{W}}] = \;<a,\mathrm{E}[\omega]>_{\mathcal{W}} \qquad \forall\; a \in \mathcal{W}$$
$$\mathrm{Cov}(\omega) \text{ satisfies } \mathrm{Cov}(<a,\omega>_{\mathcal{W}},\, <b,\omega>_{\mathcal{W}}) = \;<a,\mathrm{Cov}(\omega)b>_{\mathcal{W}} \qquad \forall\; a,b \in \mathcal{W}.$$

It will be necessary to assume $\mathrm{Cov}(\omega) \in \mathcal{L}_{\mathrm{PD}}(\mathcal{W},\mathcal{W})$ or the set of positive definite linear transformations from $\mathcal{W} \to \mathcal{W}$. With these definitions, models can be used to represent $\mathrm{E}[\omega]$ and $\mathrm{Cov}(\omega)$. Two general models of interest are given below. Such models have been considered by Seely (1996).

<u>U-Model</u>: $(\mathrm{E}[\omega], \mathrm{Cov}(\omega)) \in \mathcal{T} \subset \mathcal{W} \times \mathcal{L}_{\mathrm{PD}}(\mathcal{W},\mathcal{W})$.

The underlying model (U-Model) allows the expectation and covariance of ω to be related. This is the most general representation of the expectation and covariance that is needed. Also, define the sets:

$\mathcal{U} = \{u \in \mathcal{W} \mid (u,v) \in \mathcal{T} \text{ for some } v\} \qquad \text{sp}\,\mathcal{U} = \mathcal{E}$

$\mathcal{V} = \{v \in \mathcal{L}_{\text{PD}}(\mathcal{W},\mathcal{W}) \mid (u,v) \in \mathcal{T} \text{ for some } u\}.$

Parameterizations will be used to provide a setting in which estimation can be defined with respect to the U-Model. Parameterization for the whole model, expectation, and the variance are given below.

Whole Model $\quad \tau : \Upsilon \to \mathcal{W} \times \mathcal{L}_{\text{PD}}(\mathcal{W},\mathcal{W})$ where $\tau(\Upsilon) = \mathcal{T}$

Expectation $\quad \tau_{\mathcal{U}} : \Upsilon_{\mathcal{U}} \to \mathcal{W}$ where $\tau_{\mathcal{U}}(\Upsilon_{\mathcal{U}}) = \mathcal{U}$ and $\text{sp}\Upsilon_{\mathcal{U}} = \mathcal{P}$

Variance $\quad \tau_{\mathcal{V}} : \Upsilon_{\mathcal{V}} \to \mathcal{L}_{\text{PD}}(\mathcal{W},\mathcal{W})$ where $\tau_{\mathcal{V}}(\Upsilon_{\mathcal{V}}) = \mathcal{V}$.

Despite the relationship between the mean and the variance, only the parameterization of the expectation is of concern. It is assumed that $\Upsilon_{\mathcal{U}} \subset \mathcal{P}$, $\text{sp}\Upsilon_{\mathcal{U}} = \mathcal{P}$, and the mapping $\tau_{\mathcal{U}}$ can be extended to a linear transformation $U : \mathcal{P} \to \mathcal{W}$. For $\theta \in \Upsilon_{\mathcal{U}}$, $\tau_{\mathcal{U}}(\theta) = U\theta \in \mathcal{U}$ and $\text{sp}\mathcal{U} = \text{sp}\{U\theta | \theta \in \Upsilon_{\mathcal{U}}\} = \text{sp}U(\Upsilon_{\mathcal{U}})$ $= U(\text{sp}\Upsilon_{\mathcal{U}}) = U(\mathcal{P}) = \underline{R}(U)$. The parameterization for the expectation is often expressed as $\text{E}_\theta[\omega] = U\theta,\ \theta \in \Upsilon_{\mathcal{U}}$. Such parameterizations can always be defined using

$$\Upsilon = \mathcal{T},\ \tau(u,v) = (u,v) \qquad \Upsilon_{\mathcal{U}} = \mathcal{U},\ \tau_{\mathcal{U}}(u) = u, Uu = u \qquad \Upsilon_{\mathcal{V}} = \mathcal{V},\ \tau_{\mathcal{V}}(v) = v.$$

Certain assumptions may be required for the U-Model. These assumptions are listed below, and will always be stated either in the result or at the start of the section. The reason behind the assumptions is demonstrated in the next section. A lemma is presented to demonstrate these assumptions do fit into the above framework as $\text{sp}\Upsilon_{\mathcal{U}} = \mathcal{P}$ under both [O] and [S], and that [O] is a stronger assumption than [S].

Assumptions: Open Set [O] - $\Upsilon_{\mathcal{U}}$ contains a non-empty open set in $\mathcal{P}$.

Spanning Condition [S] - $\text{sp}(\Upsilon_{\mathcal{U}} - \Upsilon_{\mathcal{U}}) = \mathcal{P}$.

O-S Lemma: i) Under [O], $\text{sp}\Upsilon_{\mathcal{U}} = \mathcal{P}$. ii) Under [S], $\text{sp}\Upsilon_{\mathcal{U}} = \mathcal{P}$. iii) [O] $\Rightarrow$ [S].

proof: i) Let $\mathcal{C} \subset \Upsilon_{\mathcal{U}}$ be a non-empty open set in $\mathcal{P}$, $\theta_0 \in \mathcal{C}$, and $u \in (\text{sp}\,\mathcal{C})^\perp = \mathcal{C}^\perp$. Because $\mathcal{C}$ is open, $\exists\, \epsilon > 0 \ni \theta_0 + \delta u \in \mathcal{C}\ \ \forall\ |\delta| < \epsilon$. Since $u \in \mathcal{C}^\perp$, $< u, \theta_0 + \delta u >_{\mathcal{P}} = 0\ \forall\ |\delta| < \epsilon$

$\Rightarrow\ < u, \theta_0 >_{\mathcal{P}} + \delta < u, u >_{\mathcal{P}} = 0\ \ \forall\ |\delta| < \epsilon \Rightarrow\ < u, u >_{\mathcal{P}} = 0 \Rightarrow u = 0.$

Thus, $\mathcal{C}^\perp = \{0\} \Rightarrow \text{sp}\,\mathcal{C} = \mathcal{P} \Rightarrow \text{sp}\Upsilon_{\mathcal{U}} = \mathcal{P}$ as $\mathcal{C} \subset \Upsilon_{\mathcal{U}}$ or $\text{sp}\mathcal{C} \subset \text{sp}\Upsilon_{\mathcal{U}}$.

ii) Note $\text{sp}(\Upsilon_{\mathcal{U}} - \Upsilon_{\mathcal{U}}) = \{\sum_{i=1}^{m} a_i(\gamma_{i1} - \gamma_{i2}) |\, m \geq 1, \gamma_{i1}, \gamma_{i2} \in \Upsilon_{\mathcal{U}}, a_i \in \mathcal{R}\}$

$\subset \{\sum_{i=1}^{m} a_i\gamma_i - \sum_{j=1}^{m} b_j\gamma_j) |\, m \geq 1, \gamma_i, \gamma_j \in \Upsilon_{\mathcal{U}}, a_i, b_j \in \mathcal{R}\} \subset \text{sp}(\Upsilon_{\mathcal{U}}).$

Thus, $\text{sp}(\Upsilon_{\mathcal{U}} - \Upsilon_{\mathcal{U}}) = \mathcal{P} \Rightarrow \text{sp}(\Upsilon_{\mathcal{U}}) = \mathcal{P}$.

iii) Using [O], let $\mathcal{C} \subset \Upsilon_{\mathcal{U}}$ be a non-empty open set in $\mathcal{P}$ and let $\gamma_0 \in \Upsilon_{\mathcal{U}}$. Then $\mathcal{C} - \gamma_0 \subset \Upsilon_{\mathcal{U}} - \{\gamma_0\}$ is a non-empty open set in $\mathcal{P}$. Hence, $\text{sp}(\Upsilon_{\mathcal{U}} - \{\gamma_0\}) = \mathcal{P}$ by i) $\Rightarrow \text{sp}(\Upsilon_{\mathcal{U}} - \Upsilon_{\mathcal{U}}) = \mathcal{P}$ [S] since $\Upsilon_{\mathcal{U}} - \{\gamma_0\} \subset \Upsilon_{\mathcal{U}} - \Upsilon_{\mathcal{U}}$ or $\text{sp}(\Upsilon_{\mathcal{U}} - \{\gamma_0\}) \subset \text{sp}(\Upsilon_{\mathcal{U}} - \Upsilon_{\mathcal{U}})$. ■

The U-Model is more general than a linear model. It is a linear model if and only if $\mathcal{U} = \text{sp}\,\mathcal{U} = \mathcal{E}$. Under the parameterization $\tau_{\mathcal{U}}$, the U-Model is a linearly mean-parameterized model since the expectation and the parameter are linearly related. For purposes of this thesis, the U-Model is equivalent to a linear model, namely the M-Model defined below. This model separates the mean and covariance. Seely (1996) refers to this model as the artificial model.

<u>M-Model</u> : $\text{E}[\omega] \in \mathcal{E}$ $\qquad\qquad$ $\text{Cov}(\omega) \in \mathcal{V}$.

Note that the M-Model is a special case of the U-Model when $\mathcal{T} = \mathcal{E} \times \mathcal{V}$. A parameterization could be defined for the M-Model, but it is not necessary for this study.

In the U-Model setting, the goal will be to estimate $\Pi^*\theta$ using estimators of the form $H^*\omega$. The next proposition gives the mean and variance of such an estimator.

Proposition: Suppose $\tau \in \mathcal{T}$. Then i) $\text{E}_\tau[H^*\omega] = H^*\text{E}_\tau[\omega]$ $\qquad$ ii) $\text{Cov}_\tau(H^*\omega) = H^*\text{Cov}_\tau(\omega)H$.

proof: i) $\text{E}_\tau[H^*\omega] \in \mathcal{H}$. Then $\forall\ h \in \mathcal{H}$ using the definition of expectation and adjoint

$< h, \text{E}_\tau[H^*\omega] >_{\mathcal{H}} = \text{E}_\tau[\, < h, H^*\omega >_{\mathcal{H}}] = \text{E}_\tau[\, < Hh, \omega >_{\mathcal{W}}]$

$= \ < Hh, \text{E}_\tau[\omega] >_{\mathcal{W}} = \ < h, H^*\text{E}_\tau[\omega] >_{\mathcal{H}}$.

ii) Note $H^*\omega \in \mathcal{H}$, so $\text{Cov}_\tau(H^*\omega) : \mathcal{H} \rightarrow \mathcal{H}$. Then $\forall\ h_1, h_2 \in \mathcal{H}$ using definition of covariance and adjoint

$< h_1, \text{Cov}_\tau(H^*\omega)h_2 >_{\mathcal{H}} = \text{Cov}_\tau(\, < h_1, H^*\omega >_{\mathcal{H}},\ < h_2, H^*\omega >_{\mathcal{H}})$

$= \text{Cov}_\tau(\, < Hh_1, \omega >_{\mathcal{W}},\ < Hh_2, \omega >_{\mathcal{W}}) = \ < Hh_1, \text{Cov}_\tau(\omega)Hh_2 >_{\mathcal{W}}$

$= \ < h_1, H^*\text{Cov}_\tau(\omega)Hh_2 >_{\mathcal{H}}$. ■

The purpose behind the U-Model is to have a linearly mean-parameterized model which is general enough to incorporate the particular models of interest. The results can then be derived for the general model and applied to the others as special cases. The U-Model fulfills this purpose as demonstrated through the following relations:

Y-Model :	$\Upsilon = \mathcal{R}^p \times \Xi,\ \Upsilon_{\mathcal{U}} = \mathcal{R}^p,\ \Upsilon_{\mathcal{V}} = \Xi$	$\mathcal{P} = \mathcal{R}^p$	$U\theta = X\underline{\beta}$
LQEM for $\underline{Z}$:	$\Upsilon = \Upsilon_{\mathcal{U}} = \Upsilon_{\mathcal{V}} = \Xi$	$\mathcal{P} = \mathcal{R}^r$	$U\theta = X^\dagger\underline{\psi}$.

The mean parameter set $\mathcal{R}^p$ always contains a non-empty open set for the Y-Model and Ξ will contain a non-empty open set in $\mathcal{R}^{k+1}$ under assumption [O] for the Y-Model which holds for the LQEM for $(\underline{Y} - X\widehat{\underline{\beta}})$ and for the LQEM for $N_X\underline{Y}$ in order for these models to generate the ML and REML equations, respectively.

3.3.2. Mean Estimability

Estimation is naturally only concerned with parameters that can be estimated. Thus, concepts related to estimability should be examined for the underlying model. Mean estimability will be examined for estimating $\Pi^*\theta$ in the U-Model. Consider the following definition:

Definition: Mean Estimable - The linear transformation $\Pi^*\theta$ is mean estimable provided
$U\theta_1 = U\theta_2 \Rightarrow \Pi^*\theta_1 = \Pi^*\theta_2 \;\; \forall \;\; \theta_1, \theta_2 \in \Upsilon_{\mathcal{U}}$.

The mean part of the definition indicates that the definition only applies to the behavior of the mean and not to the behavior of the variance. Some results concerning mean estimability are given below. The last two results demonstrate the necessity of assumption [S] in the U-Model.

Theorem: $U\theta$ is mean estimable.

proof: Follows directly from the definition. ∎

(Seely,1996)
Mean Estimability Theorem: Under [S], the following are equivalent:

i) $\Pi^*\theta$ is mean estimable
ii) $\underline{R}(\Pi) \subset \underline{R}(U^*)$
iii) $\exists\, H : \mathcal{H} \rightarrow \mathcal{W} \ni \mathrm{E}_\theta[H^*w] = \Pi^*\theta \;\; \forall \;\theta \in \mathcal{P}$.

proof: i) $\Leftrightarrow$ $[U\theta_1 = U\theta_2 \Rightarrow \Pi^*\theta_1 = \Pi^*\theta_2 \;\; \forall \;\; \theta_1, \theta_2 \in \Upsilon_{\mathcal{U}}\,] \Leftrightarrow [U\delta = 0 \Rightarrow \Pi^*\delta = 0 \;\; \forall\; \delta \in \Upsilon_{\mathcal{U}} - \Upsilon_{\mathcal{U}}\,]$
$\Leftrightarrow [U\delta = 0 \Rightarrow \Pi^*\delta = 0 \;\; \forall \;\delta \in \mathrm{sp}(\Upsilon_{\mathcal{U}} - \Upsilon_{\mathcal{U}}) = \mathcal{P}\;]$ using [S]
$\Leftrightarrow \underline{N}(U) \subset \underline{N}(\Pi^*) \Leftrightarrow \underline{R}(\Pi) \subset \underline{R}(U^*) \Leftrightarrow$ ii).
Also, ii) $\Leftrightarrow U^*H = \Pi$ for some $H : \mathcal{H} \rightarrow \mathcal{W} \;\Leftrightarrow H^*U = \Pi^*$
$\Leftrightarrow H^*U\theta = \Pi^*\theta \;\; \forall \;\theta \in \mathcal{P} \Leftrightarrow \mathrm{E}_\theta[H^*w] = \Pi^*\theta \;\; \forall \;\theta \in \mathcal{P} \Leftrightarrow$ iii). ∎

(Seely,1996)
Full Rank Theorem: Under [S], θ is mean estimable $\Leftrightarrow \underline{r}(U) = p$.

proof: Using the Mean Estimability theorem shows $\theta = I\theta$ is mean estimable

$\Leftrightarrow \underline{R}(I) \subset \underline{R}(U^*) \Leftrightarrow \underline{R}(U^*) = \mathcal{P} \Leftrightarrow \underline{r}(U^*) = \dim \mathcal{P} = p \Leftrightarrow \underline{r}(U) = p.$ ■

Estimation results can be given with respect to $U\theta$ since this linear transformation is mean estimable. The three cases in the Mean Estimability theorem are equivalent under [S]. The O-S lemma indicates that this condition is also satisfied under [O].

Some later results will require mean estimability of θ. If θ is not mean estimable, it is possible to obtain a parameterization which is full rank or one in which the parameter vector under the new parameterization is estimable. This is demonstrated in the next proposition.

Proposition: Suppose [S] and $\mathrm{E}_\theta[w] = U\theta$, $\theta \in \Upsilon_{\mathcal{U}}$, where θ is not mean estimable. The expectation can be reparameterized as $\mathrm{E}_{\underline{\alpha}}[w] = T\underline{\alpha}$ where $\underline{\alpha} \in A_{\mathcal{U}} = \{\underline{\alpha} \in \mathcal{R}^m | T\underline{\alpha} \in \mathcal{U}\}$, T is a linear transformation, and $\underline{\alpha}$ is mean estimable.

proof: i) Suppose $\underline{R}(U) = \mathrm{sp}\{u_1, ..., u_m\}$ where $\underline{r}(U) = m$ and define $T : \mathcal{R}^m \rightarrow \mathcal{W}$ by $T(\underline{\alpha}) = \sum_{i=1}^{m} \alpha_i u_i$.
Note $\underline{R}(U) = \underline{R}(T)$ and $\underline{N}(T) = \{\underline{0}\}$.
ii) $T(\mathrm{sp}(A_{\mathcal{U}} - A_{\mathcal{U}})) = \mathrm{sp}(T(A_{\mathcal{U}}) - T(A_{\mathcal{U}})) = \mathrm{sp}(\mathcal{U} - \mathcal{U}) = \mathrm{sp}(U(\Upsilon_{\mathcal{U}}) - U(\Upsilon_{\mathcal{U}}))$
$= U(\mathrm{sp}(\Upsilon_{\mathcal{U}} - \Upsilon_{\mathcal{U}}) = U(\mathcal{P}) = \underline{R}(U) = \underline{R}(T) = T(\mathcal{R}^m)$ by i)
$\Rightarrow \mathrm{sp}(A_{\mathcal{U}} - A_{\mathcal{U}}) = \mathcal{R}^m$ $([\mathrm{S}]_{A_{\mathcal{U}}})$ as $\underline{N}(T) = \{\underline{0}\}$ by i) and $\mathrm{sp}A_{\mathcal{U}} = \mathcal{R}^m$ by the O-S lemma.
iii) Then $\mathrm{E}_{\underline{\alpha}}[w] = T\underline{\alpha}$ is a parameterization of the expectation of w, because T is a linear transformation with $T(A_{\mathcal{U}}) = \mathcal{U}$ and $\mathrm{sp}A_{\mathcal{U}} = \mathcal{R}^m$ by ii). Now, $\underline{r}(T) = m - \underline{n}(T) = m - 0 = \dim \mathcal{R}^m$
$\Rightarrow \underline{\alpha}$ is mean estimable by the Full Rank theorem under $[\mathrm{S}]_{A_{\mathcal{U}}}$. ■

Methods of estimation can now be presented for parameters of interest that are mean estimable. The estimation methods include least squares and uniformly best linear unbiased estimation.

3.3.3. Least Squares Estimation

Least squares estimation will be presented in terms of the U-Model. This method of estimation may be more interpretable under the M-Model. However, in section 3.3.4, it will be shown that the estimators are equivalent under both models.

Consider the U-Model where $U\theta \in \mathcal{U}$ and $V \in \mathcal{V}$. The least squares and generalized least squares estimators are defined for θ which may or may not be mean estimable. When θ is not mean estimable, the least squares solution is not unique.

Definitions: Least Squares Estimator (LSE) - $\widehat{\theta}_I(\omega) = \widehat{\theta}_I$ is an LSE for θ provided it minimizes $< \omega - U\theta, \omega - U\theta >_W \ \ \forall\ \theta \in \mathcal{P}$.

Generalized Least Squares Estimator (GLSE) - Let V be the true $V \in \mathcal{V}$. Then $\widehat{\theta}_V(\omega) = \widehat{\theta}_V$ is a GLSE for θ provided it minimizes $< \omega - U\theta, V^{-1}(\omega - U\theta) >_W \ \ \forall\ \theta \in \mathcal{P}$.

The LSE does not depend upon the covariance, while the GLSE does depend upon the covariance. For the GLSE, the given variance V is fixed at the true $V \in \mathcal{V}$, whereas the θ in $U\theta$ varies over all $\theta \in \mathcal{P}$. This can be understood to mean that one covariance is selected from the set $\mathcal{V}$ and it is desired to estimate θ based on this V. The LSE theorem and the GLSE theorem provide representations for these quantities.

LSE Theorem: The following are equivalent: i) $\widehat{\theta}$ is an LSE ii) $U^*U\widehat{\theta} = U^*\omega$ iii) $U\widehat{\theta} = P_U\omega$.

proof: (1) $< \omega - U\theta, \omega - U\theta >_W = < \omega - P_U\omega + P_U\omega - U\theta, \omega - P_U\omega + P_U\omega - U\theta >_W$

$= < (I-P_U)\omega, (I-P_U)\omega >_W + < (I-P_U)\omega, P_U\omega - U\theta >_W$

$+ < P_U\omega - U\theta, (I-P_U)\omega >_W + < P_U\omega - U\theta, P_U\omega - U\theta >_W$ by linearity of inner product

$= < (I-P_U)\omega, (I-P_U)\omega >_W + < P_U\omega - U\theta, P_U\omega - U\theta >_W$ as $(I-P_U)\omega \perp P_U\omega - U\theta$

which is minimized when $U\widehat{\theta} = P_U\omega$. Hence, i) $\Leftrightarrow$ iii).

(2) Note $U\widehat{\theta}_I = P_U\omega \Rightarrow U^*U\widehat{\theta}_I = U^*\omega$ and $U^*U\widehat{\theta}_I = U^*\omega \Rightarrow U^*(U\widehat{\theta}_I - P_U\omega) = 0 \Rightarrow U\widehat{\theta}_I = P_U\omega$ as $U\widehat{\theta}_I - P_U\omega \in \underline{R}(U) \cap \underline{N}(U^*) = \{0\}$. Hence, ii) $\Leftrightarrow$ iii). ∎

GLSE Theorem: The following are equivalent for a given $V \in \mathcal{V}$

i) $\widehat{\theta}$ is a GLSE ii) $U^*V^{-1}U\widehat{\theta} = U^*V^{-1}\omega$ iii) $U\widehat{\theta} = U(U^*V^{-1}U)^{-}U^*V^{-1}\omega$.

proof: (1) By the spectral theorem, $\exists$ real numbers $\lambda_1, \ldots, \lambda_n$ and OPOs $E_1, \ldots, E_n \ni$ $V^{-1} = \sum_{i=1}^{n}\lambda_i E_i \Rightarrow V^{-\frac{1}{2}}$ exists where $V^{-\frac{1}{2}} = \sum_{i=1}^{n}\lambda_i^{\frac{1}{2}}E_i$, so

$< \omega - U\theta, V^{-1}(\omega - U\theta) >_W = < V^{-\frac{1}{2}}\omega - V^{-\frac{1}{2}}U\theta, V^{-\frac{1}{2}}\omega - V^{-\frac{1}{2}}U\theta >_W$

which is minimized when $(V^{-\frac{1}{2}}U)^*(V^{-\frac{1}{2}}U)\widehat{\theta} = (V^{-\frac{1}{2}}U)^*V^{-\frac{1}{2}}\omega$ by proof of LSE theorem

or equivalently when $U^*V^{-1}U\widehat{\theta} = U^*V^{-1}\omega$. Hence, i) $\Leftrightarrow$ ii).

(2) ii) $\Rightarrow U^*V^{-1}U\widehat{\theta} = U^*V^{-1}\omega \Rightarrow U(U^*V^{-1}U)^{-}U^*V^{-1}U\widehat{\theta} = U(U^*V^{-1}U)^{-}U^*V^{-1}\omega$

$\Rightarrow MU\widehat{\theta} = M\omega$ where M is the PO on $\underline{R}(U)$ along $\underline{N}(U^*V^{-1})$ by general projection theorem in 2.5.

$\Rightarrow U\widehat{\theta} = M\omega = U(U^*V^{-1}U)^{-}U^*V^{-1}\omega \Rightarrow$ iii).

(3) iii) $\Rightarrow U\widehat{\theta} = U(U^*V^{-1}U)^{-}U^*V^{-1}\omega \Rightarrow U^*V^{-1}U\widehat{\theta} = U^*V^{-1}U(U^*V^{-1}U)^{-}U^*V^{-1}\omega$

$\Rightarrow U^*V^{-1}U\widehat{\theta} = KU^*V^{-1}\omega$ where K is a PO on $\underline{R}(U^*V^{-1}U) = \underline{R}(U^*)$ using proposition in 2.3, 2.5

$\Rightarrow U^*V^{-1}U\widehat{\theta} = U^*V^{-1}\omega \Rightarrow$ ii). ∎

The LSE and GLSE are unbiased estimators of $U\theta$. In order to be a valid estimator, the GLSE requires the covariance to be known. If the covariance is unknown, then it has to be estimated. In this case, the GLSE would actually be an estimated GLSE or EGLSE where $\widehat{V} = \widehat{V}(\omega) \in \mathcal{L}_{\text{PD}}(\mathcal{W},\mathcal{W})$ $\forall\ \omega \in \mathcal{W}$. Using the above corollary, the EGLSE for $U\theta$ would be the value of $\widehat{\theta}_{\widehat{V}}$ which solves $U\widehat{\theta}_{\widehat{V}}(\omega) = U(U^*\widehat{V}^{-1}U)^{-}U^*\widehat{V}^{-1}\omega$. Because $\widehat{V} = \widehat{V}(\omega)$, the EGLSE is not necessarily linear or unbiased.

3.3.4. Uniformly Best Linear Unbiased Estimation

Uniformly best linear unbiased estimation provides a method to assess the performance of estimators. Consider estimators of the form $H^*\omega$ and let $\mathcal{L}_{\text{NND}}(\mathcal{H},\mathcal{H})$ be the set of NND linear transformations from $\mathcal{H}\to\mathcal{H}$ and denote $\text{Cov}(H^*\omega) \in \mathcal{L}_{\text{NND}}(\mathcal{W},\mathcal{W})$ by $\text{Cov}(H^*\omega) \geq 0$. In addition, let $\tau = (\text{E}[\omega], \text{Cov}(\omega)) \in \mathcal{T}$, $u = \text{E}[\omega] \in \mathcal{U}$, and $v = \text{Cov}(\omega) \in \mathcal{V}$.

Definition: Uniformly Best Linear Unbiased Estimator (UBLUE) - $H^*\omega$ is UBLUE for its expectation $\Leftrightarrow \text{Cov}_\tau(K^*\omega) \geq \text{Cov}_\tau(H^*\omega)\ \forall\ \tau \in \mathcal{T}\ \forall\ K^* : \mathcal{W}\to\mathcal{H} \ni \text{E}_\tau[K^*\omega] = \text{E}_\tau[H^*\omega]\ \forall\ \tau \in \mathcal{T}$.

The next result indicates that the existence of a UBLUE in the U-Model is equivalent to the existence of a UBLUE in the M-Model. This is done by showing that UBLUEs are equivalent under both models.

(Seely,1996)
U-M UBLUE Theorem: $H^*\omega$ is UBLUE in the U-Model if and only if it is UBLUE in the M-Model.

proof: i) $\mathcal{C}_{\text{U}} = \{K^*\omega |\ \text{E}_\tau[K^*\omega] = \text{E}_\tau[H^*\omega]\ \forall\ \tau \in \mathcal{T}\} = \{K^*\omega |\ K^*\text{E}_\tau[\omega] = H^*\text{E}_\tau[\omega]\ \forall\ \tau \in \mathcal{T}\}$
$= \{K^*\omega |\ K^*u = H^*u\ \ \forall\ u \in \mathcal{U}\}$
$= \{K^*\omega |\ K^*u = H^*u\ \ \forall\ u \in \mathcal{E}\}$ since the condition is linear it is the same under $\mathcal{U}$ and $\mathcal{E}$
$= \{K^*\omega |\ K^*\text{E}_u[\omega] = H^*\text{E}_u[\omega]\ \ \forall\ u \in \mathcal{E}\} = \{K^*\omega |\ \text{E}_u[K^*\omega] = \text{E}_u[H^*\omega]\ \ \forall\ u \in \mathcal{E}\} = \mathcal{C}_{\text{M}}$.
ii) By i) $\mathcal{C}_{\text{U}} = \mathcal{C}_{\text{M}} = \mathcal{C}$, so $\mathcal{D}_{\text{U}} = \{K^*\omega \in \mathcal{C} |\ \text{Cov}_\tau(K^*\omega) \geq \text{Cov}_\tau(H^*\omega)\ \forall\ \tau \in \mathcal{T}\}$
$= \{K^*\omega \in \mathcal{C} |\ K^*\text{Cov}_\tau(\omega)K \geq H^*\text{Cov}_\tau(\omega)H\ \ \forall\ \tau \in \mathcal{T}\}$
$= \{K^*\omega \in \mathcal{C} |\ K^*VK \geq H^*VH\ \ \forall\ V \in \mathcal{V}\}$
$= \{K^*\omega \in \mathcal{C} |\ K^*\text{Cov}_V(\omega)K \geq H^*\text{Cov}_V(\omega)H\ \ \forall\ V \in \mathcal{V}\}$
$= \{K^*\omega \in \mathcal{C} |\ \text{Cov}_V(K^*\omega) \geq \text{Cov}_V(H^*\omega)\ \ \forall\ V \in \mathcal{V}\} = \mathcal{D}_{\text{M}}$.
$\therefore\ \mathcal{C}_{\text{U}} = \mathcal{C}_{\text{M}}$ and $\mathcal{D}_{\text{U}} = \mathcal{D}_{\text{M}}$, then H^*W is UBLUE in the U-Model $\Leftrightarrow$ H^*W is UBLUE in M-Model. ▪

Additional definitions with respect to the U-Model are given below. These definitions will be used to develop properties of UBLUEs.

Definitions: IBLUE: A UBLUE with respect to $\mathcal{V} = \{I\}$.

VBLUE: A UBLUE with respect to $\mathcal{V} = \{V\}$ where V is given.

Full UBLUE (FUBLUE): A UBLUE for $\mathrm{E}[\omega]$.

The definition of a FUBLUE is for convenience, since it will be desirable to differentiate UBLUE properties in full and non-full cases. The next theorems will be used to identify UBLUEs and their uniqueness.

(Seely and Zyskind,1969)

Zyskind's Theorem: Assume $\mathrm{E}[\omega] \in \mathcal{W}$ and $\mathrm{Cov}(\omega) = V \geq 0$. Then $H^*\omega$ is VBLUE $\Leftrightarrow \underline{R}(VH) \subset \mathcal{E}$.

proof: i) Suppose $\underline{R}(VH) \subset \mathcal{E}$.

(1) Consider $K^*\omega \ni \mathrm{E}_\tau[K^*\omega] = \mathrm{E}_\tau[H^*\omega]\ \forall\ \tau \in \mathcal{T} \Leftrightarrow K^*u = H^*u\ \forall\, u \in \mathcal{E}$

$\Leftrightarrow (K^* - H^*)u = 0\ \forall\, u \in \mathcal{E}\ \Leftrightarrow F^*u = 0\ \forall\, u \in \mathcal{E}$ where $F = K - H \Leftrightarrow \underline{R}(F) \subset \mathcal{E}^\perp$.

(2) Note $F^*VH = 0$ by (1) since $\underline{R}(F) \subset \mathcal{E}^\perp$and $\underline{R}(VH) \subset \mathcal{E}$.

(3) Then $\mathrm{Cov}(K^*\omega) = \mathrm{Cov}((H + F)^*\omega) = H^*VH + H^*VF + F^*VH + F^*VF$

$= H^*VH + F^*VF$ by (2) $\Rightarrow \mathrm{Cov}(K^*\omega) = H^*VH + F^*VF \geq H^*VH = \mathrm{Cov}(H^*\omega)$.

ii) Suppose $H^*\omega$ is VBLUE.

(4) Let $\mathcal{B} = \{t|Vt \in \mathcal{E}\}$ and show $\mathcal{B} + \mathcal{E}^\perp = \mathcal{W}$.

Suppose $u \in \mathcal{B}^\perp \cap \mathcal{E}$ and note $\underline{N}(V) \subset \mathcal{B} \Rightarrow \mathcal{B}^\perp \subset \underline{N}(V)^\perp = \underline{R}(V)$.

Then $u \in \mathcal{B}^\perp \subset \underline{R}(V) \Rightarrow u = Vw$ for some w. Also, $u = Vw \in \mathcal{E}\ \Rightarrow w \in \mathcal{B}$.

Thus, $u'w = 0 \Rightarrow w'Vw = 0 \Rightarrow Vw = 0$ as V is NND $\Rightarrow u = 0$. Hence,

$\mathcal{B}^\perp \cap \mathcal{E} = \{0\} \Rightarrow \mathcal{B} + \mathcal{E}^\perp = \mathcal{W}$.

(5) From (4), can write $\mathcal{W} = \mathcal{B} + \mathcal{E}^\perp = \mathcal{C} \oplus \mathcal{E}^\perp$ where $\mathcal{C} \subset \mathcal{B}$. Define P to be a PO on $\mathcal{C}$ along $\mathcal{E}^\perp$

$\Rightarrow P^*$ is a PO on $\mathcal{E}$ along $\mathcal{C}^\perp$.

(6) Set $K = PH$, $N = I - P$, and $F = NH$. Note $H = K + F$. Then

a) $\mathrm{E}[K^*\omega] = \mathrm{E}[H^*P^*\omega] = H^*P^*\mathrm{E}[\omega] = H^*\mathrm{E}[\omega] = \mathrm{E}[H^*\omega]$ since P^* is a PO on $\mathcal{E}$ by (5)

b) $\underline{R}(K) \subset \underline{R}(P) \subset \mathcal{B}$ by (5) $\Rightarrow \underline{R}(VK) \subset \mathcal{E} \Rightarrow K^*\omega$ is a VBLUE by i).

c) $\underline{R}(F) \subset \underline{R}(N) = \underline{N}(P) = \mathcal{E}^\perp$ by (5).

Thus, by (6) and the hypothesis, $K^*\omega$ and $H^*\omega$ are VBLUE for $\mathrm{E}[H^*\omega]$. By definition of VBLUE,

$\mathrm{Cov}(H^*\omega) = \mathrm{Cov}(K^*\omega) \Leftrightarrow H^*VH = K^*VK \Leftrightarrow (K + F)^*V(K + F) = K^*VK$ by (6)

$\Leftrightarrow K^*VK + F^*VK + K^*VF + F^*VF = K^*VK\ \Leftrightarrow F^*VF = 0$ since $F^*VK = 0$ by (6)

$\Leftrightarrow VF = 0$ by proposition in 2.3 $\Leftrightarrow V(H - K) = 0$ by definition of F in (6)

$\Leftrightarrow VH = VK \Rightarrow \underline{R}(VH) = \underline{R}(VK) \subset \mathcal{E}$ by (6). ■

(Seely,1996)

Uniqueness Theorem: Assume V is PD and $H^*\omega$ is VBLUE. Then $K^*\omega$ is VBLUE for $\mathrm{E}[H^*\omega] \Leftrightarrow K = H$.

proof: i) If $K = H$, then the conclusion follows directly.
ii) Suppose $K^*\omega$ is VBLUE for $\mathrm{E}[H^*\omega]$. By (1) in the above proof, $F = K - H$ where $\underline{R}(F) \subset \mathcal{E}^{\perp}$.
In addition, $\underline{R}(VK) \subset \mathcal{E}$ by Zyskind's theorem. Hence,
$\mathrm{Cov}(K^*\omega) = \mathrm{Cov}(H^*\omega) \Leftrightarrow K^*VK = H^*VH \Leftrightarrow (F^* + H^*)V(F + H) = H^*VH$
$\Leftrightarrow F^*VF + H^*VH = H^*VH$ by (1) $\Leftrightarrow F^*VF = 0 \Rightarrow F = 0$ using proposition in 2.3 where V is PD
$\Rightarrow K = H$ by definition of F. ■

The above results can be applied to UBLUEs by noting that a UBLUE is a VBLUE $\forall\ V \in \mathcal{V}$ or equivalently $\forall\ V \in \mathrm{sp}\mathcal{V}$ by the linearity of the condition in Zyskind's theorem. Because this study is concerned with UBLUEs in the U-Model under a mean parameterization, the UBLUE definition is restated for the mean parameterized case.

Lemma: $H^*\omega$ is UBLUE in the U-Model for $\Pi^*\theta$ if and only if

i) $\mathrm{E}_\theta[H^*\omega] = \Pi^*\theta \ \ \forall\ \theta \in \mathcal{P}$

ii) $\mathrm{Cov}_V(H^*\omega) \leq \mathrm{Cov}_V(K^*\omega) \ \ \forall\ V \in \mathcal{V}$ and $\forall\ K^* : \mathcal{W} \to \mathcal{H} \ni \mathrm{E}_\theta[K^*\omega] = \Pi^*\theta \ \ \forall\ \theta \in \mathcal{P}$.

The first condition in the lemma defines unbiasedness for estimating $\Pi^*\theta$. The second condition indicates that the UBLUE is the best linear estimator for all possible covariances among all unbiased estimators. Zyskind's theorem can be applied to least squares and generalized least squares estimators to show these estimators are IBLUE and VBLUE, respectively.

Corollary: i) $U\widehat{\theta}_I$ is IBLUE for $U\theta$. ii) $U\widehat{\theta}_V$ is VBLUE for $U\theta$.

proof: i) (1) Note $\mathrm{E}_\theta[U\widehat{\theta}_I] = \mathrm{E}_\theta[P_U\omega] = P_UU\theta = U\theta$ by LSE theorem.
(2) $\underline{R}(P_U) = \underline{R}(U)$.
$\therefore U\widehat{\theta}_I$ is IBLUE for $U\theta$ by Zyskind's theorem.
ii) (1) Note $\mathrm{E}_\theta[U\widehat{\theta}_V] = \mathrm{E}_\theta[M\omega] = MU\theta = U\theta$ by the GLSE theorem.
(2) $\underline{R}(VM^*) = \underline{R}(U(U^*V^{-1}U)^-U^*) \subset \underline{R}(U)$ by the GLSE theorem.
$\therefore U\widehat{\theta}_V$ is VBLUE for $U\theta$ by Zyskind's theorem. ■

The LSE and GLSE are special cases of UBLUEs, so they have a correspondence to the LSE and GLSE in the M-Model. By the U-M UBLUE theorem and the uniqueness theorem, $U\widehat{\theta}_I$ is the unique IBLUE in the U-Model and the M-Model, and $U\widehat{\theta}_V$ is the unique VBLUE in the U-Model and the M-Model. Thus, least squares estimation in the U-Model is equivalent to least squares estimation in the M-Model, and generalized least squares estimation in the U-Model is equivalent to generalized least squares estimation in the M-Model.

The next theorem shows that linear combinations of UBLUEs are UBLUE for their expectation. An example of the importance of this result is given in the corollary concerning FUBLUEs.

(Seely,1996)
Linear Closure Property: If $H^*\omega$ and $K^*\omega$ are UBLUE, then $\begin{bmatrix} H^* \\ K^* \end{bmatrix}\omega$ and $L^*H^*\omega$ are UBLUE.

proof: Using Zyskind's theorem gives the following $\forall\, V \in \mathcal{V}$,
i) $\underline{R}(V[H\ \ K]) = \underline{R}(VH) + \underline{R}(VK) \subset \mathcal{E} + \mathcal{E} = \mathcal{E}$ as $H^*\omega$, $K^*\omega$ are UBLUE $\Rightarrow \begin{bmatrix} H^* \\ K^* \end{bmatrix}\omega$ is UBLUE.
ii) $\underline{R}(VHL) \subset \underline{R}(VH) \subset \mathcal{E}$ as $H^*\omega$ is UBLUE $\Rightarrow L^*H^*\omega$ is UBLUE. ▪

Corollary: If $U\theta$ has a FUBLUE, $\Pi^*\theta$ is estimable, and [S] holds, then $\Pi^*\theta$ has a UBLUE.

proof: Since $\Pi^*\theta$ is estimable $\Rightarrow \underline{R}(\Pi) \subset \underline{R}(U^*)$ by the Mean Estimability theorem under [S]
$\Rightarrow \Pi = U^*M$ for some linear transformation M
$\Rightarrow \Pi^* = M^*U$. Suppose $H^*\omega$ is the FUBLUE for $U\theta$ and consider $M^*H^*\omega$. Note
$\mathrm{E}[M^*H^*\omega] = M^*U\theta = \Pi^*\theta$ and $M^*H^*\omega$ is UBLUE by the linear closure property. ▪

The above corollary indicates that the UBLUE for $\Pi^*\theta$ can be derived from the FUBLUE. However, this may not always be the case, as a UBLUE may exist for $\Pi^*\theta$, but not for $U\theta$. For this study, it is convenient to distinguish between these two settings. If a FUBLUE exists, then this will be referred to as the full case and it is reasonable to think of the UBLUE for $\Pi^*\theta$ as a FUBLUE as it is derivable from the FUBLUE. If a FUBLUE does not exist, but a UBLUE exists for $\Pi^*\theta$, then this will be referred to as the general case. The full case is presented in chapter 4 while the general case is presented in chapters 5 and 6. The UBLUE conditions will provide some of the basic tools that will be examined in this study. These methods have been defined for the U-Model, and can be applied to the other models in this chapter.

3.3.5. Zyskind's Condition

Zyskind's condition leads to many nice properties. Some of these properties will be shown in this section. This condition has briefly been mentioned for the previous models, but it will be further developed in this section for the U-Model. The U-Model assumes that V is PD $\forall\ V \in \mathcal{V}$. This condition is defined as:

Definition: Zyskind's Condition (ZC) - The condition $\underline{R}(VU) \subset \underline{R}(U)\ \ \forall\ V \in \mathcal{V}$.

Zyskind's condition is linear and can be extended from $\mathcal{V}$ to sp$\mathcal{V}$. Also, LSE = GLSE $\forall\ V \in \mathcal{V}$ under ZC. This is demonstrated in the next theorem. The theorem also provides a condition under which V_1-GLSE = V_0-GLSE.

ZC Relation Theorem: i) For invertible V, $\underline{R}(VU) \subset \underline{R}(U) \Leftrightarrow U(U^*V^{-1}U)^-U^*V^{-1} = U(U^*U)^-U^*$.

ii) For invertible V where $\underline{N}(U) = \{0\}$, $\underline{R}(VU) \subset \underline{R}(U) \Leftrightarrow (U^*V^{-1}U)^{-1}U^*V^{-1} = (U^*U)^{-1}U^*$.

iii) If V_0, V_1 are invertible, then $\underline{R}(V_1^{-1}U) \subset \underline{R}(V_0^{-1}U) \Leftrightarrow \underline{R}(V_1V_0^{-1}U) \subset \underline{R}(U)$

$\Leftrightarrow U(U^*V_1^{-1}U)^-U^*V_1^{-1} = U(U^*V_0^{-1}U)^-U^*V_0^{-1}$.

proof: i) By general projection theorem, $A = U(U^*V^{-1}U)^-U^*V^{-1}$ is a PO on $\underline{R}(U)$ along $\underline{N}(U^*V^{-1})$.
By uniqueness of POs, $A = P_U \Leftrightarrow \underline{N}(U^*V_1^{-1}) = \underline{N}(U^*) \Leftrightarrow \underline{R}(V^{-1}U) = \underline{R}(U)$
$\Leftrightarrow \underline{R}(VU) = \underline{R}(U) \Leftrightarrow \underline{R}(VU) \subset \underline{R}(U)$ since $\underline{r}(VU) = \underline{r}(U)$ as V is invertible.
ii) Note that $\{0\} = \underline{N}(U) = \underline{N}(U^*U) \Rightarrow (U^*U)^{-1}$ exists.
(1) Suppose $(U^*V^{-1}U)^{-1}U^*V^{-1} = (U^*U)^{-1}U^*$.
Then $\underline{R}(VU) \subset \underline{R}(U)$ follows immediatedly from (1) by left multiplying by U and using i).
(2) Suppose $\underline{R}(VU) \subset \underline{R}(U)$. Then $A = P_U$ from i) $\Rightarrow U(U^*V^{-1}U)^{-1}U^*V^{-1} = U(U^*U)^{-1}U^*$
$\Rightarrow (U^*U)^{-1}U^*U(U^*V^{-1}U)^{-1}U^*V^{-1} = (U^*U)^{-1}U^*U(U^*U)^{-1}U^*$.
iii) Let $A_1 = U(U^*V_1^{-1}U)^-U^*V_1^{-1}$ and $A_0 = U(U^*V_0^{-1}U)^-U^*V_0^{-1}$. Then $A_1 = A_0$
$\Leftrightarrow \underline{N}(U^*V_1^{-1}) = \underline{N}(U^*V_0^{-1})$ by the general projection theorem in section 2.5
$\Leftrightarrow \underline{R}(V_1^{-1}U) = \underline{R}(V_0^{-1}U) \Leftrightarrow \underline{R}(U) = \underline{R}(V_1V_0^{-1}U)$
$\Leftrightarrow \underline{R}(V_1V_0^{-1}U) \subset \underline{R}(U)$ since have equality of ranks. ■

The next theorem uses the results of the preceding section to show LSE = UBLUE if and only if ZC. The theorem also provides conditions which are equivalent to V_0-GLSE = UBLUE. This theorem is the main result of this section.

Theorem: i) $U\widehat{\theta}_I$ is UBLUE for $U\theta \Leftrightarrow \underline{R}(VU) \subset \underline{R}(U) \quad \forall\, V \in \mathcal{V}$

ii) $U\widehat{\theta}_{V_0}$ is UBLUE for $U\theta \Leftrightarrow \underline{R}(V_0^{-1}U) \subset \underline{R}(V^{-1}U) \Leftrightarrow \underline{R}(VV_0^{-1}U) \subset \underline{R}(U) \quad \forall\, V \in \mathcal{V}$.

proof: i) (1) Note $\mathrm{E}_\theta[U\widehat{\theta}_I] = \mathrm{E}_\theta[P_U\omega] = P_U U\theta = U\theta$ by LSE theorem.

(2) $U\widehat{\theta}_I$ is UBLUE for $U\theta \Leftrightarrow U\widehat{\theta}_I$ is VBLUE for $U\theta \;\forall\, V \in \mathcal{V}$

$\Leftrightarrow \underline{R}(VP_U) = \underline{R}(VU) \subset \underline{R}(U) \;\forall\, V \in \mathcal{V}$ by Zyskind's theorem.

ii) (1) Note $\mathrm{E}_\theta[U\widehat{\theta}_0] = U(U^*V_0^{-1}U)^-U^*V_0^{-1}U\theta = U\theta$ as AA^- is a PO on $\underline{R}(A)$.

(2) By Zyskind's theorem, $U\widehat{\theta}_0$ is UBLUE for $U\theta$

$\Leftrightarrow \underline{R}(VV_0^{-1}U(U^*V_0^{-1}U)^-U^*) \subset \underline{R}(U) \quad \forall\, V \in \mathcal{V}$

$\Leftrightarrow \underline{R}(V_0^{-1}U(U^*V_0^{-1}U)^-U^*) \subset \underline{R}(V^{-1}U) \quad \forall\, V \in \mathcal{V}$

$\Leftrightarrow \underline{R}(V_0^{-1}U) \subset \underline{R}(V^{-1}U) \quad \forall\, V \in \mathcal{V}$ as $M_0 = V_0^{-1}U(U^*V_0^{-1}U)^-U^*$ is a PO on $\underline{R}(V_0^{-1}U)$

$\Leftrightarrow \underline{R}(VV_0^{-1}U) \subset \underline{R}(U) \quad \forall\, V \in \mathcal{V}$. ∎

The following proposition demonstrates a useful commutativity property involving projection operators and Zyskind's condition.

Proposition: $\underline{R}(VU) \subset \underline{R}(U) \Rightarrow P_UVP_U = VP_U = P_UV$ and $N_UVN_U = VN_U = N_UV$.

proof: $\underline{R}(VU) \subset \underline{R}(U) \Rightarrow \underline{R}(VP_U) \subset \underline{R}(U) \Rightarrow P_UVP_U = VP_U$ and $P_UVP_U = P_UV$ by symmetry.
Also, $N_UVN_U = (I - P_U)V(I - P_U) = (I - P_U)(V - VP_U) = (I - P_U)(V - P_UV) = N_UV$
and $N_UVN_U = VN_U$ by symmetry. ∎

Even though the U-Model assumes the covariance is PD, it is informative to consider the case where V is singular in order to establish the limitations of applying Zyskind's condition. Consider the case where the true V is singular. Puntanen and Styan (1989) define the GLSE by $U(U^*V^+U)^+U^*V^+\omega$. Even if V is singular, the result [LSE = VBLUE $\Leftrightarrow$ ZC] still holds. However, the following propositions indicate ZC is not sufficient to show GLSE = LSE and GLSE = VBLUE in this case.

Proposition: If V is singular, then $U(U^*V^+U)^+U^*V^+ = U(U^*U)^+U^* \Leftrightarrow \underline{R}(V^+U) = \underline{R}(U)$.

proof: (1) Note V is self-adjoint $\Rightarrow V^+$ is self-adjoint by section 2.5. Since V is NND, $V = BB^*$ for some B from section 2.7 $\Rightarrow V^+ = V^+VV^+ = V^+BB^*V^+ = CC^* \Rightarrow V^+$ is NND by section 2.7.

(2) $P = U(U^*V^+U)^+U^*V^+$ is a PO as $P^2 = U(U^*V^+U)^+U^*V^+U(U^*V^+U)^+U^*V^+ = P$.

Using (1), NND properties in section 2.7, and Moore-Penrose inverse properties in section 2.5

$\underline{R}(UU^*V^+) = \underline{R}(UU^*V^+UU^*) = \underline{R}(UU^*V^+U) = \underline{R}(U(U^*V^+U)^+)$

$\supset \underline{R}(P) \supset \underline{R}(U(U^*V^+U)^+U^*V^+U) = \underline{R}(UU^*V^+U) = \underline{R}(UU^*V^+UU^*) = \underline{R}(UU^*V^+)$.

Thus, $\underline{R}(P) = \underline{R}(UU^*V^+)$. Also,

$\underline{R}(V^+U) \supset \underline{R}(P^*) \supset \underline{R}(V^+U(U^*V^+U)^+U^*V^+) = \underline{R}(V^+P) = \underline{R}(V^+UU^*V^+) = \underline{R}(V^+U)$.

$\Rightarrow \underline{R}(P^*) = \underline{R}(V^+U) \Rightarrow \underline{N}(P) = \underline{N}(U^*V^+)$.

(3) From (2), $P = P_U \Leftrightarrow \underline{R}(UU^*V^+) = \underline{R}(U)$ and $\underline{R}(V^+U) = \underline{R}(U)$. However, $\underline{R}(V^+U) = \underline{R}(U)$

$\Rightarrow \underline{r}(U) = \underline{r}(V^+U) = \underline{r}(V^+UU^*) = \underline{r}(UU^*V^+) \Rightarrow \underline{R}(UU^*V^+) = \underline{R}(U)$. ∎

Proposition: i) If V is singular, then the GLSE = VBLUE $\Leftrightarrow \underline{R}(U) \subset \underline{R}(V)$.

ii) If V is singular and the GLSE = LSE, then the GLSE is VBLUE.

proof: i) By Zyskind's theorem, the GLSE = VBLUE $\Leftrightarrow \underline{R}(VV^+U(U^*V^+U)^+U^*) \subset \underline{R}(U)$

$\Leftrightarrow \underline{R}(VP^*) \subset \underline{R}(U)$ where P is the PO on $\underline{R}(UU^*V^+)$ along $\underline{N}(U^*V^+)$ by above corollary

$\Leftrightarrow \underline{R}(VV^+U) \subset \underline{R}(U)$ as $\underline{R}(P^*) = \underline{R}(V^+U)$

$\Leftrightarrow P_V\underline{R}(U) \subset \underline{R}(U)$ as $VV^+ = P_V$

$\Leftrightarrow \underline{R}(U) \subset \underline{R}(V)$.

ii) By the above proposition, GLSE = LSE $\Leftrightarrow \underline{R}(V^+U) = \underline{R}(U) \Rightarrow \underline{R}(U) \subset \underline{R}(V^+) = \underline{R}(V)$

$\Rightarrow$ GLSE = VBLUE by i). ∎

This section defined Zyskind's condition and demonstrated some of the results that can be derived from it. This condition will be referred to in the next chapters and is presented for the U-Model so that it can be applied directly to the other models of interest.

4. UBLUE for the Expectation

This chapter examines the existence of full uniformly best linear unbiased estimators (FUBLUEs) which are UBLUEs for the expectation of the model. These estimators are defined in section 3.3.4 FUBLUEs will first be identified in the underlying model and these results will be applied to the other models for estimating fixed effects and variance components. This chapter will also discuss explicit linear likelihood estimators under the maximum likelihood and restricted maximum likelihood procedures which have been presented by Rogers and Young (1977), Szatrowski (1980), and ElBassiouni (1983). The chapter concludes with an example.

4.1. FUBLUE for the Underlying Model

This section examines conditions under which a FUBLUE exists in the U-Model for the mean estimable quantity $U\theta$. The least squares estimator, the generalized least squares estimator, and the estimated generalized least squares estimator for $U\theta$ will be of interest in this section. These were given in the previous chapter using the LSE theorem, but they are listed here for reference:

LSE: $U\widehat{\theta}_I = U(U^*U)^- U^*\omega$

GLSE: $U\widehat{\theta}_V = U(U^*V^{-1}U)^- U^*V^{-1}\omega$ for the true $V \in \mathcal{L}_{\mathrm{PD}}(\mathcal{W}, \mathcal{W})$

EGLSE: $U\widehat{\theta}_{\widehat{V}} = U(U^*\widehat{V}^{-1}U)^- U^*\widehat{V}^{-1}\omega$ with $\widehat{V} = \widehat{V}(\omega) \in \mathcal{L}_{\mathrm{PD}}(\mathcal{W}, \mathcal{W})\ \forall\ \omega \in \mathcal{W}$.

The LSE has a simple linear form which does not depend on the covariance, the GLSE depends on the unknown true V, and the EGLSE is not linear and requires an estimate for the unknown covariance. The objective is to determine when the LSE is equal to the EGLSE. The UBLUE results of the previous chapter can be used to address this objective.

U-FUBLUE Theorem: The following are equivalent:

i) $U\widehat{\theta}_I$ is UBLUE for $U\theta$ in the U-Model

ii) $\underline{R}(VU) \subset \underline{R}(U)\ \ \forall\ V \in \mathcal{V}$ (ZC)

iii) $U(U^*V^{-1}U)^- U^*V^{-1} = U(U^*U)^- U^*\ \ \ \forall\ V \in \mathcal{V}$.

proof: By Zyskind's theorem i) $\Leftrightarrow$ ii) and by the ZC Relation theorem ii) $\Leftrightarrow$ iii). ∎

Proposition 1: If $V \in \mathrm{sp}\mathcal{V}$ and V is PD, then the conditions in the U-FUBLUE theorem imply $U\widehat{\theta}_V = U\widehat{\theta}_I$.

proof: $\underline{R}(VU) \subset \underline{R}(U)\ \forall\ V \in \mathcal{V} \Rightarrow \underline{R}(VU) \subset \underline{R}(U)\ \forall\ V \in \mathrm{sp}\mathcal{V}$ by section 2.3

$\Rightarrow U(U^*V^{-1}U)^-U^*V^{-1} = U(U^*U)^-U^*\ \forall\ V \in \mathrm{sp}\mathcal{V}$ by the ZC relation theorem where V is PD

$\Rightarrow U\widehat{\theta}_V = U\widehat{\theta}_I$. ∎

Proposition 2: If $I \in \mathrm{sp}\mathcal{V}$, then $U\theta$ has a UBLUE $\Leftrightarrow U\widehat{\theta}_I$ is UBLUE for $U\theta$.

proof: i) $U\widehat{\theta}_I$ is UBLUE for $U\theta \Rightarrow U\theta$ has a UBLUE.

ii) Suppose $U\theta$ has a UBLUE given by $U\widehat{\theta}_{V_0}$where $V_0 \in \mathcal{V}$

$\Rightarrow \underline{R}(V_0^{-1}U) \subset \underline{R}(V^{-1}U)\ \forall\ V \in \mathcal{V}$ by theorem in section 3.3.5 $\Rightarrow \underline{R}(VV_0^{-1}U) \subset \underline{R}(U)\ \ \forall V \in \mathcal{V}$

$\Rightarrow \underline{R}(VV_0^{-1}U) \subset \underline{R}(U)\ \forall\ V \in \mathrm{sp}\mathcal{V}$ by proposition in section 2.3 $\Rightarrow \underline{R}(V_0^{-1}U) \subset \underline{R}(U)$ as $I \in \mathrm{sp}\mathcal{V}$

$\Rightarrow U\widehat{\theta}_{V_0} = U\widehat{\theta}_I$ by ZC Relation theorem $\Rightarrow U\widehat{\theta}_I$ is UBLUE for $U\theta$ by Uniqueness theorem. ∎

The U-FUBLUE theorem indicates when the estimator $U\widehat{\theta}_I$ is best among unbiased linear estimators for $U\theta$. In this case, the best estimator has the same variance as the GLSE, since these estimators are equal. It is not necessary for $I \in \mathrm{sp}\mathcal{V}$ in order for $U\widehat{\theta}_I$ to be best. However, if $I \in \mathrm{sp}\mathcal{V}$ and $U\widehat{\theta}_I$ is not UBLUE, then there does not exist a UBLUE for $U\theta$.

Consider the set of transformations $\mathcal{L}_U(\mathcal{W},\mathcal{W}) = \{V \in \mathcal{L}_{\mathrm{PD}}(\mathcal{W},\mathcal{W}) | \underline{R}(VU) \subset \underline{R}(U)\}$. This set can be used to re-express the conditions in the U-FUBLUE theorem and apply them to the EGLSE. This is demonstrated by the following corollaries which follow directly from the U-FUBLUE theorem.

Corollary: ZC $\Leftrightarrow \mathcal{V} \subset \mathcal{L}_U(\mathcal{W},\mathcal{W})$.

Corollary: If $\mathcal{V} \subset \mathcal{L}_U(\mathcal{W},\mathcal{W})$, then GLSE = LSE = FUBLUE.

Corollary: If $\mathcal{V} \subset \mathcal{L}_U(\mathcal{W},\mathcal{W})$, then $\mathrm{sp}\mathcal{V} \cap \mathcal{L}_{\mathrm{PD}}(\mathcal{W},\mathcal{W}) \subset \mathcal{L}_U(\mathcal{W},\mathcal{W})$.

Corollary: If $U\theta$ has a FUBLUE and $\widehat{V} \in \mathcal{L}_U(\mathcal{W},\mathcal{W})$, then EGLSE = GLSE = LSE = FUBLUE.

A sufficient condition for the existence of a FUBLUE is Zyskinds condition (ZC). The FUBLUE is the LSE, which does not depend on the covariance matrix V. The results of this section will be applied to the other models of interest for purposes of estimating fixed effects and variance components.

4.2. FUBLUE for Fixed Effects

FUBLUEs for fixed effects can be obtained using the Y-Model and applying the results from the U-Model. The FUBLUE results can be used to determine the existence of an explicit linear maximum likelihood estimator for the full fixed effect vector $X\underline{\beta}$.

4.2.1. FUBLUE Results

This section examines conditions under which a FUBLUE exists in the Y-Model. The least squares estimator, the generalized least squares estimator, and the estimated generalized least squares estimator for $X\underline{\beta}$ will be of interest in this section. They are listed here for reference:

LSE: $X\widehat{\underline{\beta}}_I = X(X'X)^- X'\underline{Y}$

GLSE: $X\widehat{\underline{\beta}}_{\underline{\psi}} = X(X'V_{\underline{\psi}}^{-1}X)^- X'V_{\underline{\psi}}^{-1}\underline{Y}$ for a given $\underline{\psi}$ where $V_{\underline{\psi}}$ is PD

EGLSE: $X\widehat{\underline{\beta}}_{\widehat{\underline{\psi}}} = X(X'V_{\widehat{\underline{\psi}}}^{-1}X)^- X'V_{\widehat{\underline{\psi}}}^{-1}\underline{Y}$ where $V_{\widehat{\underline{\psi}}} = V_{\widehat{\underline{\psi}}(\underline{Y})}$ is PD $\forall\ \underline{Y} \in \mathcal{R}^n$.

The UBLUE result for the U-Model can be used to indicate when the LSE is a good estimator for estimating $X\underline{\beta}$. This is stated in the next theorem, which follows directly from the U-FUBLUE theorem.

Y-FUBLUE Theorem: The following are equivalent:

i) $X\widehat{\underline{\beta}}_I$ is UBLUE for $X\underline{\beta}$ in the Y-Model

ii) $\underline{R}(V_{\underline{\psi}}X) \subset \underline{R}(X) \quad \forall\ \underline{\psi} \in \Xi$

iii) $X(X'V_{\underline{\psi}}^{-1}X)^- X'V_{\underline{\psi}}^{-1} = X(X'X)^- X' \quad \forall\ \underline{\psi} \in \Xi$.

Proposition: Assume [L] and [O]. Then $\forall\ \underline{\psi} \in \mathcal{R}^{k+1} \ni V_{\underline{\psi}}$ is PD, the conditions in the Y-FUBLUE theorem imply that $X\widehat{\underline{\beta}}_{\underline{\psi}} = X\widehat{\underline{\beta}}_I$.

proof: Under [L] and [O], sp$\mathcal{V} = \{V_{\underline{\psi}} |\ \underline{\psi} \in \text{sp}\ \Xi\} = \{V_{\underline{\psi}} |\ \underline{\psi} \in \mathcal{R}^{k+1}\}$. Use proposition 1 in section 4.1. ∎

The Y-FUBLUE theorem does not make any assumptions about open sets, linear covariance structure, normality, or classification matrices. Only the appropriate form of ZC given in ii) is necessary to apply this theorem to the special cases of the Y-Model under [L], [O], and [C].

4.2.2. FELMLQE Results

For doing maximum likelihood estimation in the Y-Model, the assumptions of [L], [O], and [N] are used. Following the approach of Szatrowski (1980), the full explicit linear maximum likelihood equation estimator (FELMLQE) for $X\underline{\beta}$ satisfies $X\widehat{\underline{\beta}}_{\text{MLQ}} = A\underline{Y}$ for a constant fixed matrix A, which is not random and does not depend on any parameters. In this case, the estimator $X\widehat{\underline{\beta}}_{\text{MLQ}}$ is linear and the explicit part indicates that A is constant. The following lemma shows that the MLQE $=$ EGLSE in this setting.

Lemma: $X\widehat{\underline{\beta}}_{\text{MLQ}} = X\widehat{\underline{\beta}}_{\widehat{\underline{\psi}}_{\text{MLQ}}}$.

proof: Note from the likelihood equations and the definition of the EGLSE both are a solution to $X\widehat{\underline{\beta}}_{\widehat{\underline{\psi}}} = X(X'V_{\widehat{\underline{\psi}}}^{-1}X)^{-}X'V_{\widehat{\underline{\psi}}}^{-1}\underline{Y}$ where $\widehat{\underline{\psi}} = \widehat{\underline{\psi}}_{\text{MLQ}}$. ■

The next theorem presents a sufficient condition for the existence of an FELMLQE for $X\underline{\beta}$. This condition is ZC for the Y-Model.

(Szatrowski,1980)

Y-FELMLQE Theorem: Consider the Y-Model under [L], [O], and [N]. If $\widehat{\underline{\psi}}_{\text{MLQ}}$ exists $\ni V_{\widehat{\underline{\psi}}_{\text{MLQ}}}$ is PD and $\underline{R}(V_{\underline{\psi}}X) \subset \underline{R}(X)\ \ \forall\, \underline{\psi} \in \Xi$, then $X\widehat{\underline{\beta}}_I$ is an FELMLQE for $X\underline{\beta}$.

proof: Since $V_{\widehat{\underline{\psi}}_{\text{MLQ}}}$ is PD by hypothesis, $X\widehat{\underline{\beta}}_{\widehat{\underline{\psi}}_{\text{MLQ}}} = X\widehat{\underline{\beta}}_I$ by proposition in section 4.2.1.
By the above lemma, $X\widehat{\underline{\beta}}_{\text{MLQ}} = X\widehat{\underline{\beta}}_{\widehat{\underline{\psi}}_{\text{MLQ}}}$. Thus, $X\widehat{\underline{\beta}}_{\text{MLQ}} = X\widehat{\underline{\beta}}_I = A\underline{Y}$ where $A = X(X'X)^{-}X'$
$\Rightarrow X\widehat{\underline{\beta}}_I$ is an FELMLQE for $X\underline{\beta}$ by definition. ■

The proof of the theorem shows the MLQE is equivalent to the LSE under the assumptions. Because the LSE is linear and explicit, there exists a FELMLQE for $X\underline{\beta}$. The condition for the existence of an FELMLQE is sufficient, but not necessary. This is due to the fact that the relation $X\widehat{\underline{\beta}}_{\widehat{\underline{\psi}}} = A\underline{Y}$ for some A and $\widehat{\underline{\psi}} \in \Xi$ does not necessarily imply $\underline{R}(V_{\underline{\psi}}\, X) \subset \underline{R}(X)\ \ \forall\, \underline{\psi} \in \Xi$. The result in the following proposition does hold.

Proposition: If $X\widehat{\underline{\beta}}_{\underline{\psi}} = A\underline{Y}\ \ \forall\, \underline{\psi} \in \Xi$ for some constant A and $\exists\, \underline{\psi}_0 \in \text{sp}\,\Xi \ni V_{\underline{\psi}_0} = I$, then $\underline{R}(V_{\underline{\psi}}\, X) \subset \underline{R}(X)\ \ \forall\, \underline{\psi} \in \Xi$.

proof: $X\widehat{\underline{\beta}}_{\underline{\psi}} = X(X'V_{\underline{\psi}}^{-1}X)^{-}X'V_{\underline{\psi}}^{-1}\underline{Y} = A\underline{Y}\ \ \forall\, \underline{\psi} \in \text{sp}\,\Xi \ni V_{\underline{\psi}}$ is PD by Proposition 1 in section 4.1
$\Rightarrow X(X'V_{\underline{\psi}_0}^{-1}X)^{-}X'V_{\underline{\psi}_0}^{-1}\underline{Y} = X(X'X)^{-}X'\underline{Y} = A\underline{Y}$ for $\underline{\psi}_0 \in \text{sp}\,\Xi$. Thus, $\forall\, \underline{\psi} \in \Xi$,
$X(X'V_{\underline{\psi}}^{-1}X)^{-}X'V^{-1} = X(X'X)^{-}X' \Rightarrow \underline{R}(V_{\underline{\psi}}\, X) \subset \underline{R}(X)\ \ \forall\, \underline{\psi} \in \Xi$ by the ZC Relation theorem. ■

For completely balanced mixed classification models, ZC is satisfied so there exists an FELMLQE for $X\underline{\beta}$. However, a more general result involving balance can be stated.

(VanLeeuwen et al.,1997)

Corollary: For the Y-Model under [C] and [N], let $\mathbb{H}$ be the set of factors corresponding to all random effects and $\mathbb{G}$ be the set of factors corresponding to all fixed effects. If the design is Bal($\mathbb{H}|\mathbb{G}$), then $\exists$ an FELMLQE for $X\underline{\beta}$.

proof: VanLeeuwen et al. (1997) show this balance condition implies ZC. This gives the desired result from the Y-FELMQE theorem. ∎

4.3. FUBLUE for Variance Components

FUBLUEs for the variance components can be obtained using linearized quadratic estimation models and applying the results from the U-Model. The FUBLUE results can be used to determine the existence of a linear estimator for the full variance component vector. The results will be applied to maximum likelihood estimators and restricted maximum likelihood estimators. The FUBLUE results will be presented first for the LQEM for $\underline{Z}$ and applied to the ALQEM for $(\underline{Y} - X\widehat{\underline{\beta}})$ and the LQEM for $N_X\underline{Y}$.

4.3.1. FUBLUE Results

The U-FUBLUE theorem indicates ZC is a sufficient condition for the least squares estimator to be FUBLUE. The first task will be to characterize ZC for the LQEMs. Two results will be presented to show that ZC for the LQEM for $\underline{Z}$ is equivalent to a QS condition.

Lemma 1: Let $B, R_1, ..., R_r \in \mathcal{S}_n$, $R_{\underline{\psi}} = \sum_{i=1}^{r} \psi_i R_i$ for $\underline{\psi} \in \mathcal{R}^r$, and Ξ contain a non-empty open set in $\mathcal{R}^r$. Then $\text{sp}\{R_{\underline{\psi}} B R_{\underline{\psi}} | \underline{\psi} \in \Xi\} = \text{sp}\{R_{\underline{\psi}} B R_{\underline{\psi}} | \underline{\psi} \in \mathcal{R}^r\} = \text{sp}\{R_i B R_j + R_j B R_i | 1 \le i \le j \le r\}$.

proof: (1) Define $T(\underline{\psi}) = \underline{\psi}\underline{\psi}'$ and $D(C) = \sum_i \sum_j c_{ij} R_i B R_j$ for $C = \{c_{ij}\} \in \mathcal{S}_r$. Then $R_{\underline{\psi}} B R_{\underline{\psi}} = \sum_i \sum_j \psi_i \psi_j R_i B R_j = D(T(\underline{\psi}))$.

(2) To show sp$T(\Xi) = \mathcal{S}_r$. Suppose $F \in (\text{sp}\,T(\Xi))^{\perp}$. Then $\text{tr}(F\underline{\psi}\underline{\psi}') = \underline{\psi}' F \underline{\psi} = 0 \ \forall \ \underline{\psi} \in \Xi$ (*).
Let $\underline{u} \in \mathcal{R}^r$ and G be a non-empty open set contained in Ξ. Choose $\underline{\psi}_0 \in G$. Since G is open $\exists \ \epsilon > 0$ $\ni \underline{\psi}_0 + \delta\underline{u} \in G \ \forall \ |\delta| < \epsilon$. By (*), $0 = (\underline{\psi}_0 + \delta\underline{u})' F (\underline{\psi}_0 + \delta\underline{u}) = \underline{\psi}_0' F \underline{\psi}_0 + 2\delta\underline{\psi}_0' F \underline{u} + \delta^2 \underline{u}' F \underline{u}$. Because the quadratic polynomial in δ is $0 \ \forall \ |\delta| < \epsilon$, its coefficients must be 0. Hence, $\underline{u}' F \underline{u} = 0 \ \forall \ \underline{u} \in \mathcal{R}^r$ $\Rightarrow F = 0$. Thus, $(\text{sp}T(\Xi))^{\perp} = \{0\} \Rightarrow \text{sp}\,T(\Xi) = \mathcal{S}_r$.
(3) Since D is linear, $D(T(\mathcal{R}^r)) \subset D(\mathcal{S}_r) \overset{(2)}{=} D(\text{sp}T(\Xi)) = \text{sp}D(T(\Xi))$. Thus, $\text{sp}D(T(\mathcal{R}^r)) = \text{sp}D(T(\Xi))$ as $T(\Xi) \subset T(\mathcal{R}^r)$. This establishes the first equality.

(4) Note $\forall\ \underline{\psi} \in \mathcal{R}^r\ R_{\underline{\psi}}BR_{\underline{\psi}} = \sum_{i=1}^{r}\psi_i^2 R_iBR_i + \sum\sum_{1\le i\le j\le r}\psi_i\psi_j(R_iBR_j + R_jBR_i)$

$\Rightarrow$ sp$\{R_{\underline{\psi}}BR_{\underline{\psi}}|\underline{\psi} \in \mathcal{R}^r\} \subset$ sp$\{R_iBR_j + R_jBR_i|1 \le i \le j \le r\}$.

(5) Let $\underline{\psi}^{(l,m)} = \{\psi_k\}$ where $\psi_k = \begin{cases} 1 & k = l \text{ or } m \\ 0 & \text{otherwise} \end{cases}$. Note $\underline{\psi}^{(l,m)} \in \mathcal{R}^r$. Then $\forall\, i, j \ni 1 \le i < j \le r$,

$R_iBR_j + R_jBR_i = (R_iBR_i + R_jBR_j + R_iBR_j + R_jBR_i) - R_iBR_i + R_jBR_j$

$= R_{\underline{\psi}^{(i,j)}}BR_{\underline{\psi}^{(i,j)}} - R_{\underline{\psi}^{(i,i)}}BR_{\underline{\psi}^{(i,i)}} - R_{\underline{\psi}^{(j,j)}}BR_{\underline{\psi}^{(j,j)}}$

$\Rightarrow$ sp$\{R_iBR_j + R_jBR_i|1 \le i \le j \le r\} \subset$ sp$\{R_{\underline{\psi}}BR_{\underline{\psi}}|\underline{\psi} \in \mathcal{R}^r\}$.

Hence, the last equality follows from (4) and (5). ∎

Lemma 2: Let $R_1, ..., R_r \in \mathcal{S}_n$ and for $\underline{\psi} \in \mathcal{R}^r$ define $R_{\underline{\psi}} = \sum_{i=1}^{r}\psi_i R_i$. Consider a set of symmetric matrices $\mathcal{A} = \{R_{\underline{\psi}}|\underline{\psi} \in \Xi\}$ where Ξ contains a non-empty open set in $\mathcal{R}^r$. Assume $\exists\ M \in$ sp$\mathcal{A}$ $\ni MA = A\ \forall\ A \in \mathcal{A}$. Then the following are equivalent:

i) $ABA \in$ sp$\mathcal{A}\ \ \forall\ A, B \in \mathcal{A}$ ii) $ABA \in$ sp$\mathcal{A}\ \ \forall\ A \in \mathcal{A}, B \in$ sp$\mathcal{A}$

iii) $ABA \in$ sp$\mathcal{A}\ \ \forall\ A, B \in$ sp$\mathcal{A}$ iv) sp$\mathcal{A}$ is a QS.

proof: (1) Note iii) $\Leftrightarrow$ iv) by QS results in section 2.7 and since $\exists\ M \in$ sp$\mathcal{A}$ $\ni MA = A\ \ \forall\ A \in \mathcal{A}$.

(2) Also, iii) $\Rightarrow$ ii) $\Rightarrow$ i) as $\mathcal{A} \subset$ sp$\mathcal{A}$.

(3) Suppose i) holds and fix $A \in \mathcal{A}$. Define $\alpha : \mathcal{S}_n \to \mathcal{S}_n$ by $\alpha(B) = ABA$. By i), $\alpha(\mathcal{A}) \subset$ sp$\mathcal{A}$. Since α is a linear transformation, $\alpha($sp$\mathcal{A}) =$ sp$\,\alpha(\mathcal{A}) \subset$ sp$\mathcal{A}$. Thus, $ABA \in$ sp$\mathcal{A}\ \ \forall\ A \in \mathcal{A}, B \in$ sp$\mathcal{A} \Rightarrow$ ii).

(4) Assume ii) and fix $B \in$ sp$\mathcal{A}$. Let $R_{\underline{\psi}} \in$ sp$\mathcal{A} = \{R_{\underline{\psi}}|\underline{\psi} \in \mathcal{R}^r\}$ as sp$\,\Xi = \mathcal{R}^r$ by O-S lemma in 3.3.1. By lemma 1 and using ii), sp$\{R_{\underline{\psi}}BR_{\underline{\psi}}|\underline{\psi} \in \mathcal{R}^r\} =$ sp$\{R_{\underline{\psi}}BR_{\underline{\psi}}|\underline{\psi} \in \Xi\} \subset$ sp$\mathcal{A} \Rightarrow$ iii). ∎

The matrix M is necessary in order to have a matrix in the set which acts like the identity. The next theorem uses the above result to represent ZC for the LQEM for $\underline{Z}$, which is given by $\underline{R}(V_{\underline{\psi}}^{\dagger}X^{\dagger}) \subset \underline{R}(X^{\dagger})$ $\forall\ \underline{\psi} \in \Xi$. Recall, $\underline{R}(X^{\dagger}) =$ sp$\,\mathcal{U}^{\dagger}$.

QS Theorem: If Ξ contains a non-empty open set in $\mathcal{R}^r$ and $\exists\, M \in$ sp$\mathcal{U}^{\dagger} \ni MR = R\ \forall\, R \in \mathcal{U}^{\dagger}$, then $\underline{R}(V_{\underline{\psi}}^{\dagger}X^{\dagger}) \subset \underline{R}(X^{\dagger})\ \forall\ \underline{\psi} \in \Xi \Leftrightarrow$ sp$\mathcal{U}^{\dagger}$ is a QS.

proof: $\underline{R}(V_{\underline{\psi}}^{\dagger}X^{\dagger}) \subset \underline{R}(X^{\dagger})\ \ \forall\ \underline{\psi} \in \Xi \Leftrightarrow V_{\underline{\psi}}^{\dagger}(S) \in$ sp$\mathcal{U}^{\dagger} =$ sp$\{R_1, ..., R_r\}\ \ \forall\, S \in$ sp$\mathcal{U}^{\dagger},\ \forall\ \underline{\psi} \in \Xi$

$\Leftrightarrow\ 2\Psi_R(S) \in$ sp$\mathcal{U}^{\dagger}\ \ \ \forall\ S \in$ sp$\mathcal{U}^{\dagger}, R \in \mathcal{U}^{\dagger} \Leftrightarrow RSR \in$ sp$\mathcal{U}^{\dagger}\ \ \forall\, S \in$ sp$\mathcal{U}^{\dagger}, R \in \mathcal{U}^{\dagger}$

$\Leftrightarrow$ sp$\mathcal{U}^{\dagger}$ is a QS from lemma 2 as Ξ contains a non-empty open set in $\mathcal{R}^r$ and $\exists\, M \in$ sp$\mathcal{U}^{\dagger}$. ∎

The QS theorem and the U-FUBLUE theorem can be used to obtain a corresponding FUBLUE theorem for the LQEM for $\underline{Z}$. The associated least squares, generalized least squares, and estimated generalized least squares estimator for $X^{\dagger}\underline{\psi}$ will be of interest. These are listed for reference:

LSE: $X^{\dagger}\widehat{\underline{\psi}}_I = X^{\dagger}(X^{\dagger *}X^{\dagger})^{-}X^{\dagger *}Y^{\dagger}$

GLSE: $X^{\dagger}\widehat{\underline{\psi}}_{\underline{\psi}} = X^{\dagger}(X^{\dagger *}V_{\underline{\psi}}^{\dagger -1}X^{\dagger})^{-}X^{\dagger *}V_{\underline{\psi}}^{\dagger -1}Y^{\dagger}$ for a given $\underline{\psi}$ where $V_{\underline{\psi}}^{\dagger}$ is PD

EGLSE: $X^{\dagger}\widehat{\underline{\psi}} = X^{\dagger}(X^{\dagger *}V_{\widehat{\underline{\psi}}}^{\dagger -1}X^{\dagger})^{-}X^{\dagger *}V_{\widehat{\underline{\psi}}}^{\dagger -1}Y^{\dagger}$ where $V_{\widehat{\underline{\psi}}}^{\dagger} = V_{\widehat{\underline{\psi}}(\underline{Y})}^{\dagger}$ is PD $\forall\ \underline{Y} \in \mathcal{R}^n$.

The UBLUE result for the U-Model can be used to indicate when the LSE is the same as the EGLSE when estimating $X^{\dagger}\underline{\psi}$. This is stated in the next theorem. The proof of the theorem follows directly from the proof of the U-FUBLUE theorem and the QS theorem.

LQZ-FUBLUE Theorem: The following are equivalent for the LQEM for $\underline{Z}$ when Ξ contains a non-empty open set in $\mathcal{R}^r$ and $\exists\, M \in \text{sp}\,\mathcal{U}^{\dagger} \ni MR = R\ \forall\ R \in \mathcal{U}^{\dagger}$:

i) $X^{\dagger}\widehat{\underline{\psi}}_I$ is UBLUE for $X^{\dagger}\underline{\psi}$

ii) $\underline{R}(V_{\underline{\psi}}^{\dagger}X^{\dagger}) \subset \underline{R}(X^{\dagger})\quad \forall\ \underline{\psi} \in \Xi$

iii) $X^{\dagger}(X^{\dagger *}V_{\underline{\psi}}^{\dagger -1}X^{\dagger})^{-}X^{\dagger *}V_{\underline{\psi}}^{\dagger -1} = X^{\dagger}(X^{\dagger *}X^{\dagger})^{-}X^{\dagger *}\quad \forall\ \underline{\psi} \in \Xi$

iv) $\text{sp}\,\mathcal{U}^{\dagger} = \text{sp}\{R_1, ..., R_r\}$ is a QS .

Proposition: For any $\underline{\psi} \in \mathcal{R}^r \ni R_{\underline{\psi}}$ is PD, the conditions in the LQZ-FUBLUE theorem imply $X^{\dagger}\widehat{\underline{\psi}}_{\underline{\psi}} = X^{\dagger}\widehat{\underline{\psi}}_I$.

proof: Condition ii) in the LQZ-FUBLUE theorem $\Rightarrow \underline{R}(V_{\underline{\psi}}^{\dagger}X^{\dagger}) \subset \underline{R}(X^{\dagger})\ \forall\ \underline{\psi} \in \Xi$

$\Rightarrow V_{\underline{\psi}}^{\dagger}X^{\dagger}\underline{u} \in \underline{R}(X^{\dagger})\ \forall\ \underline{\psi} \in \Xi, \underline{u} \in \mathcal{R}^r$

$\Rightarrow 2R_{\underline{\psi}}B_{\underline{u}}R_{\underline{\psi}} \in \underline{R}(X^{\dagger})\quad \forall\ \underline{\psi} \in \Xi, \underline{u} \in \mathcal{R}^r$ where $B_{\underline{u}} = X^{\dagger}\underline{u} \in \mathcal{S}_n$

$\Rightarrow \{R_{\underline{\psi}}B_{\underline{u}}R_{\underline{\psi}} | \underline{\psi} \in \Xi\} \subset \underline{R}(X^{\dagger})\ \forall \underline{u} \in \mathcal{R}^r \Rightarrow \text{sp}\{R_{\underline{\psi}}B_{\underline{u}}R_{\underline{\psi}} | \underline{\psi} \in \Xi\} \subset \underline{R}(X^{\dagger})\ \forall\ \underline{u} \in \mathcal{R}^r$

$\Rightarrow \text{sp}\{R_{\underline{\psi}}B_{\underline{u}}R_{\underline{\psi}} | \underline{\psi} \in \mathcal{R}^r\} \subset \underline{R}(X^{\dagger})\ \forall\ \underline{u} \in \mathcal{R}^r$ by lemma 1 $\Rightarrow R_{\underline{\psi}}B_{\underline{u}}R_{\underline{\psi}} \in \underline{R}(X^{\dagger})\ \forall\ \underline{\psi} \in \mathcal{R}^r, \underline{u} \in \mathcal{R}^r$

$\Rightarrow V_{\underline{\psi}}^{\dagger}X^{\dagger}\underline{u} \in \underline{R}(X^{\dagger})\ \forall\ \underline{\psi} \in \mathcal{R}^r, \underline{u} \in \mathcal{R}^r \Rightarrow \underline{R}(V_{\underline{\psi}}^{\dagger}X^{\dagger}) \subset \underline{R}(X^{\dagger})\ \forall\ \underline{\psi} \in \mathcal{R}^r$

$\Rightarrow$ condition iii) in the LQZ-FUBLUE theorem $\forall\ \underline{\psi} \in \mathcal{R}^r \ni R_{\underline{\psi}}$ is PD

$\Rightarrow X^{\dagger}\widehat{\underline{\psi}}_{\underline{\psi}} = X^{\dagger}\widehat{\underline{\psi}}_I\ \forall\ \underline{\psi} \in \mathcal{R}^r \ni R_{\underline{\psi}}$ is PD. ∎

This theorem was stated for the LQEM for $\underline{Z}$, so it can be applied to both the ALQEM for $(\underline{Y} - X\widehat{\underline{\beta}})$ and the LQEM for $N_X\underline{Y}$. The following sections use these models to examine the maximum likelihood and restricted maximum likelihood procedures.

4.3.2. FELMLQE Results

Maximum likelihood estimation was defined for the Y-Model under [L], [O], and [N]. This estimation procedure for the variance components generally requires an iterative procedure. Following a similar approach to Szatrowski (1980), the full explicit linear maximum likelihood estimator (FELMLQE) for $X^\circ \underline{\psi}$ satisfies $X^\circ \widehat{\underline{\psi}}_{\text{MLQ}} = AY_2^\circ$ where A is a linear transformation which is not random and does not depend on any parameters and Y_2° does not depend on any estimators or unknown parameters.

The associated least squares, generalized least squares, and the estimated generalized least squares estimators for the ALQEM for $(\underline{Y} - X\widehat{\underline{\beta}})$ will be of importance in this section. They are listed below:

LSE: $X^\circ \widehat{\underline{\psi}}_I = X^\circ (X^{\circ *} X^\circ)^- X^{\circ *} Y_2^\circ$

GLSE: $X^\circ \widehat{\underline{\psi}}_{\underline{\psi}} = X^\circ (X^{\circ *} V_{\underline{\psi}}^{\circ -1} X^\circ)^- X^{\circ *} V_{\underline{\psi}}^{\circ -1} Y_2^\circ$ for a given $\underline{\psi}$ where $V_{\underline{\psi}}^\circ$ is PD

EGLSE: $X^\circ \widehat{\underline{\psi}} = X^\circ (X^{\circ *} V_{\widehat{\underline{\psi}}}^{\circ -1} X^\circ)^- X^{\circ *} V_{\widehat{\underline{\psi}}}^{\circ -1} Y_2^\circ$ where $V_{\widehat{\underline{\psi}}}^\circ = V_{\widehat{\underline{\psi}}(\underline{Y})}^\circ$ is PD $\forall\ \underline{Y} \in \mathcal{R}^n$.

For clarity, the EGLSE will often be labelled as $\widehat{\underline{\psi}}_{\text{EGLS}}$. Recall, the ALQEM for $(\underline{Y} - X\widehat{\underline{\beta}})$ was defined so that the equations for the EGLSE correspond to the ML equations.

Lemma: $X^\circ \widehat{\underline{\psi}}_{\text{MLQ}} = X^\circ \widehat{\underline{\psi}}_{\text{EGLS}}$.

proof: Note from the ML theorem and the definition of the EGLSE for the ALQEM for $(\underline{Y} - X\widehat{\underline{\beta}})$ both are given by $X^\circ \widehat{\underline{\psi}} = X^{\circ *}(X^{\circ *} V_{\widehat{\underline{\psi}}}^{\circ -1} X^\circ)^{-1} X^{\circ *} V_{\widehat{\underline{\psi}}}^{\circ -1} Y_2^\circ$. ∎

For the ALQEM for $(\underline{Y} - X\widehat{\underline{\beta}})$, the response $Y_2^\circ = (\underline{Y} - X\widehat{\underline{\beta}})(\underline{Y} - X\widehat{\underline{\beta}})'$ will in general depend on $\widehat{\underline{\psi}}$ as $X\widehat{\underline{\beta}} = X(X'V_{\widehat{\underline{\psi}}}^{-1}X)^- X'V_{\widehat{\underline{\psi}}}^{-1}\underline{Y}$. In this setting, it will not be possible to obtain a FELMLQE for $X^\circ \underline{\psi}$ by definition. The additional assumption ZC for the Y-Model will be applied so that the ALQEM for $(\underline{Y} - X\widehat{\underline{\beta}})$ can be used to obtain conditions for the existence of a FELMLQE for $X^\circ \underline{\psi}$. As shown in section 4.2.2, this condition is sufficient for $\underline{Y} - X\widehat{\underline{\beta}}_{\underline{\psi}} = N_X \underline{Y}\ \ \forall\ \underline{\psi} \in \mathcal{R}^{k+1}$ such that $V_{\underline{\psi}}$ is PD.

(Szatrowski,1980)

ALQNY-FELMLQE Theorem: Consider the Y-Model under [L], [O], [N], and ZC. If $\widehat{\underline{\psi}}_{\text{MLQ}}$ exists $\ni$ $V_{\widehat{\underline{\psi}}_{\text{MLQ}}}$ is PD and sp$\mathcal{V} = \text{sp}\{V_1, ..., V_k, I\}$ is a QS, then $X^\circ \widehat{\underline{\psi}}_I$ is a FELMLQE for $X^\circ \underline{\psi}$.

proof: Since $V_{\widehat{\underline{\psi}}_{\text{MLQ}}}$ is PD by hypothesis and $X^\circ \widehat{\underline{\psi}}_{\text{MLQ}} = X^\circ \widehat{\underline{\psi}}_{\text{EGLS}}$ by the above lemma, $X^\circ \widehat{\underline{\psi}}_{\text{MLQ}} = X^\circ \widehat{\underline{\psi}}_I = AY_2^\circ$ by the proposition in section 4.3.1 where $A = X^{\circ *}(X^{\circ *} X^\circ)^- X^{\circ *}$ and Y_2° does not depend on $\underline{\psi}$ or $\widehat{\underline{\psi}}$ by ZC $\Rightarrow$ $X^\circ \widehat{\underline{\psi}}_I$ is an FELMLQE for $X^\circ \underline{\psi}$ by definition. ∎

Note $X^\circ\widehat{\underline{\psi}}_I$ may not be an unbiased estimator for $X^\circ\underline{\psi}$, but it does satisfy the requirement of a FELMLQE. The explicit expression for $X^\circ\widehat{\underline{\psi}}_I$ is given in section 3.2.3. An example of a case in which there exists an FELMLQE is given in the following theorem.

Theorem: A completely balanced nested mixed classification linear model has FELMLQEs for $X\underline{\beta}$, $X^\circ\underline{\psi}$.

proof: Note ZC holds for the Y-Model in balanced classification models. In addition, balance gives $\mathrm{sp}\mathcal{V} = \mathrm{sp}\{P_1, \ldots, P_k\}$. Hence, for $i \leq j = 1, \ldots, k$ (assuming ordered by nesting)
$P_iP_j + P_jP_i = 2P_i \in \mathrm{sp}\,\mathcal{V}$ using nesting $\Rightarrow$ $\mathrm{sp}\mathcal{V}$ is a QS
$\Rightarrow \exists$ ELMLQEs for $X\underline{\beta}$ and $X^\circ\underline{\psi}$ by the Y-ELMLQE and ALQNY-FELMLQE theorems. ∎

4.3.3. FELREMLQE Results

Restricted maximum likelihood estimation was defined for the Y-Model under [L], [O], and [N]. This estimation procedure for the variance components generally requires an iterative procedure. Following a similar approach to Szatrowski (1980), the FELREMLQE for $X^\circ\underline{\psi}$ satisfies $X^\circ\widehat{\underline{\psi}}_{\text{REMLQ}} = AY^\circ$, where A is a linear transformation that is not random and does not depend on any parameters and Y° does not depend on any estimators or unknown parameters.

The associated least squares, generalized least squares, and estimated generalized least squares estimators for $X^\circ\underline{\psi}$ in the LQEM for $N_X\underline{Y}$ will be of importance in this section. They are listed below:

LSE: $X^\circ\widehat{\underline{\psi}}_I = X^\circ(X^{\circ *}X^\circ)^{-1}X^{\circ *}Y^\circ$
GLSE: $X^\circ\widehat{\underline{\psi}}_{\underline{\psi}} = X^\circ(X^{\circ *}V^{\circ -1}_{\underline{\psi}}X^\circ)^- X^{\circ *}V^{\circ -1}_{\underline{\psi}}Y^\circ$ for a given $\underline{\psi}$ where $V^\circ_{\underline{\psi}}$ is PD
EGLSE: $X^\circ\widehat{\underline{\psi}} = X^\circ(X^{\circ *}V^{\circ -1}_{\widehat{\underline{\psi}}}X^\circ)^- X^{\circ *}V^{\circ -1}_{\widehat{\underline{\psi}}}Y^\circ$ where $V^\circ_{\widehat{\underline{\psi}}} = V^\circ_{\widehat{\underline{\psi}}(\underline{Y})}$ is PD $\forall\ \underline{Y} \in \mathcal{R}^n$.

For clarity, the EGLSE will often be labelled as $\widehat{\underline{\psi}}_{\text{EGLS}}$. Recall, the LQEM for $N_X\underline{Y}$ was defined so that the equations for the EGLSE correspond to the REML equations.

Lemma: $X^\circ\widehat{\underline{\psi}}_{\text{REMLQ}} = X^\circ\widehat{\underline{\psi}}_{\text{EGLS}}$.

proof: Note from REML theorem and the definition of the EGLSE for LQEM for $N_X\underline{Y}$ both are given by $X^\circ\widehat{\underline{\psi}} = X^{\circ *}(X^{\circ *}V^{\circ -1}_{\widehat{\underline{\psi}}}X^\circ)^- X^{\circ *}V^{\circ -1}_{\widehat{\underline{\psi}}}Y^\circ$. ∎

The main theorem can now be stated concerning the existence of a FELREMLQE for $X^\circ\underline{\psi}$.

(ElBassiouni,1983).

LQNY-FELREMLQE Theorem: Consider the Y-Model under [L], [O], and [N]. If $\widehat{\underset{\sim}{\psi}}_{\text{REMLQ}}$ exists $\ni V_{\widehat{\underset{\sim}{\psi}}_{\text{REMLQ}}}$ is PD and $\text{sp}\mathcal{V}_{N_X} = \text{sp}\{N_X V_1 N_X, ..., N_X V_k N_X, N_X\}$ is a QS, then $X^\circ \widehat{\underset{\sim}{\psi}}_I$ is a FELREMLQE for $X^\circ \underset{\sim}{\psi}$.

proof: Since $V_{\widehat{\underset{\sim}{\psi}}_{\text{REMLQ}}}$ is PD by hypothesis and $X^\circ \widehat{\underset{\sim}{\psi}}_{\text{REMLQ}} = X^\circ \widehat{\underset{\sim}{\psi}}_{\text{EGLS}}$ by the above lemma, $X^\circ \widehat{\underset{\sim}{\psi}}_{\text{REMLQ}} = X^\circ \widehat{\underset{\sim}{\psi}}_I = AY^\circ$ by the proposition in section 4.3.1 where $A = X^{\circ *}(X^{\circ *}X^\circ)^- X^{\circ *}$ and Y° does not depend on $\underset{\sim}{\psi}$ or $\widehat{\underset{\sim}{\psi}} \Rightarrow X^\circ \widehat{\underset{\sim}{\psi}}_I$ is an FELREMLQE for $X^\circ \underset{\sim}{\psi}$ by definition. ∎

The explicit expression for $X^\circ \widehat{\underset{\sim}{\psi}}_I$ is given in section 3.2.3. The following corollaries establish cases in which there exists a FELREMLQE for $X^\circ \underset{\sim}{\psi}$.

Corollary: If the sufficient conditions hold for the existence of an FELMLQE for $X\underset{\sim}{\beta}$ and $X^\circ \underset{\sim}{\psi}$, then there is a FELREMLQE for $X^\circ \underset{\sim}{\psi}$.

proof: Hence, ZC holds for the Y-Model and QS holds for $\text{sp}\mathcal{V}$. Using ZC, $\forall \; V, W \in \mathcal{V}$

$N_X W N_X V N_X + N_X V N_X W N_X = N_X(WV + VW)N_X = N_X U N_X \in \text{sp}\mathcal{V}_{N_X}$

where $U \in \text{sp}\mathcal{V}$ since $\text{sp}\mathcal{V}$ is a QS $\Rightarrow \text{sp}\mathcal{V}_{N_X}$ is a QS using QS results in section 2.8. ∎

Corollary: In a completely balanced mixed model, $\exists$ an FELREMLQE for $X^\circ \underset{\sim}{\psi}$.

proof: Recall ZC holds in balanced mixed models. For any two matrices F and G, $Q_{FG} = N_X P_F N_X N_X P_G N_X + N_X P_G N_X N_X P_F N_X = 2N_X P_H N_X$ using ZC and balance results from section 3.1.3. Since H is a matrix in the model, then it corresponds to a fixed or random effect. If H corresponds to a fixed effect, then $Q_{FG} = 0 \in \text{sp}\mathcal{V}_{N_X}$ and if H corresponds to a random effect, then $Q_{FG} \in \text{sp}\mathcal{V}_{N_X}$
$\Rightarrow \text{sp}\mathcal{V}_{N_X}$ is a QS $\Rightarrow \exists$ an FELREMLQE for $X^\circ \underset{\sim}{\psi}$ by the LQNY-FELREMLQE theorem. ∎

Corollary: If there are 2 variance components including σ_d^2 with associated matrix BB' and $\underset{\sim}{r}(N_X BB') = 1$, then $\text{sp}\mathcal{V}_{N_X}$ is a QS.

proof: Note $\underset{\sim}{R}(N_X BB' N_X) \subset \underset{\sim}{R}(N_X BB') \subset \underset{\sim}{R}(N_X B) = \underset{\sim}{R}(N_X BB' N_X) \Rightarrow \underset{\sim}{r}(N_X BB' N_X) = 1$
$\Rightarrow N_X BB' N_X = c\underline{h}\underline{h}' \in \mathcal{V}_{N_X}$ for some $\underline{h}$. Hence, $(N_X BB' N_X)^2 = c^2 \underline{h}\underline{h}'\underline{h}\underline{h}' = d\underline{h}\underline{h}' \in \text{sp}\mathcal{V}_{N_X}$
$\Rightarrow \text{sp}\mathcal{V}_{N_X}$ is a QS as all other combinations are in $\text{sp}\mathcal{V}_{N_X}$. ∎

(VanLeeuwen et al.,1997)

Corollary: For the Y-Model under [C] and [N], let $\mathbb{H}$ be the set of factors corresponding to all random effects and $\mathbb{G}$ be the set of factors corresponding to all fixed effects. If Bal$(\mathcal{H}_1 \cup \mathcal{H}_2)$ $\forall\, \mathcal{H}_1, \mathcal{H}_2 \in \mathbb{H}$ and Bal$(\mathbb{H}|\mathbb{G})$, then $\exists$ an FELREMLQE for $X^{\circ}\underline{\psi}$.

proof: VanLeeuwen et al. (1997) show these balance conditions produce a QS for sp$\mathcal{V}_{N_X}$. The desired result follows from the LQNY-FELREMLQE theorem. ∎

The general result of the last corollary may be clarified with an example. Consider the mixed model $y_{ijkl} = \mu + \alpha_i + b_j + c_k + e_{ijkl}$ where α_i is fixed and b_j, c_k are random. The balance conditions in the above corollary would be equivalent to :

$$\text{i) Bal}(\mathcal{H}_1 \cup \mathcal{H}_2) \Leftrightarrow \text{Bal}(\{2\} \cup \{3\}) = \text{Bal}(\{2,3\})$$
$$\text{ii) Bal}(\mathbb{H}|\mathbb{G}) \Leftrightarrow \text{Bal}(\{2\}|\{1\}) \text{ and } \text{Bal}(\{3\}|\{1\}) \text{ .}$$

If each factor had two levels, then these conditions would be satisfied under the incidence matrix:

$$\{n_{ij1}\} = \begin{bmatrix} 5 & 4 \\ 2 & 3 \end{bmatrix} \qquad \{n_{ij2}\} = \begin{bmatrix} 4 & 5 \\ 3 & 2 \end{bmatrix}$$
$$\text{i) } \{n_{\cdot jk}\} = \begin{bmatrix} 7 & 7 \\ 7 & 7 \end{bmatrix} \Rightarrow \text{Bal}(\{2,3\})$$
$$\text{ii) } \{n_{ij\cdot}\} = \{n_{i\cdot k}\} = \begin{bmatrix} 9 & 9 \\ 5 & 5 \end{bmatrix} \Rightarrow \text{Bal}(\{2\}|\{1\}),\text{Bal}(\{3\}|\{1\})\,.$$

4.4. Example: Balanced Random 1-Way Model

The balanced random 1-way model will be used to illustrate the ease of computation of ML and REML estimators when these estimators are linear and explicit. Consider the Y-model given by

$$y_{ij} = \mu + b_i + e_{ij} \qquad i = 1, \ldots, b \quad j = 1, \ldots, r \text{ or } \underline{Y}_{br\times 1} = \underline{1}\,\mu + B_{br\times b}\,\underline{b} + \underline{e} \ .$$

The usual case would require a maximization of the following density with respect to $\mu, \sigma_b^2, \sigma_e^2$ where $\mu \in \mathcal{R}$ and $\Xi = \{[\sigma_b^2 \;\; \sigma_e^2]' | \sigma_b^2 \geq 0,\, \sigma_e^2 > 0\}$:

$$f(\underline{y}\,|\mu, \sigma_b^2, \sigma_e^2) = (2\pi)^{-\frac{n}{2}} |\,\sigma_b^2 BB' + \sigma_e^2 I\,|^{-\frac{1}{2}} \exp[-\tfrac{1}{2}(\underline{Y} - \underline{1}\mu)'(\sigma_b^2 BB' + \sigma_e^2 I)^{-1}(\underline{Y} - \underline{1}\mu)].$$

However, explicit forms for the estimators can be found, since the sufficient conditions are satisfied under complete balance for this model. These estimators are derived below:

i) MLQE for μ.

Note $\underline{R}(V_{\underline{\psi}}X) = \underline{R}((\sigma_b^2 BB' + \sigma_e^2 I)\underline{1}) = \underline{R}((r\sigma_b^2 P_B + \sigma_e^2 I)\underline{1}) = \underline{R}(r\sigma_b^2\underline{1} + \sigma_e^2\underline{1}) = \underline{R}(\underline{1})$

$\Rightarrow$ FELMLQE exists for μ by Y-FELMLQE theorem and is given by

$$\widehat{\mu} = (X'V_{\widehat{\underline{\psi}}}^{-1}X)^{-1}X'V_{\widehat{\underline{\psi}}}^{-1}\underline{Y} = (\underline{1}'\underline{1})^{-1}\underline{1}'\underline{Y} = \tfrac{1}{br}\underline{1}'\underline{Y} = \bar{Y}_{..}\ .$$

ii) MLQE for $\underline{\psi} = [\sigma_b^2\ \sigma_e^2]'$.

Note sp$\mathcal{V}$ = sp$\{I, BB'\}$ = sp$\{I, P_B\}$ forms a QS

$\Rightarrow \exists$ FELMLQE for $\underline{\psi}$ by the ALQNY-FELMLQE theorem given by

$$\{\text{tr}(V_iV_j)\}_{2\times 2}\widehat{\underline{\psi}}_{\text{MLQ}} = \{\underline{Y}'NV_iN\underline{Y}\}_{2\times 1}$$

$$\Rightarrow \widehat{\underline{\psi}}_{\text{MLQ}} = \begin{bmatrix} r^2\text{tr}(P_B) & r\text{tr}(P_B) \\ r\text{tr}(P_B) & \text{tr}(I) \end{bmatrix}^{-1} \begin{bmatrix} r\underline{Y}'N_XP_BN_X\underline{Y} \\ \underline{Y}'N_X\underline{Y} \end{bmatrix}$$

$$= \frac{1}{r^2brb-(rb)^2}\begin{bmatrix} rb & -rb \\ -rb & r^2b \end{bmatrix}\begin{bmatrix} r\underline{Y}'(P_B - P_{\underline{1}})\underline{Y} \\ \underline{Y}'(I - P_{\underline{1}})\underline{Y} \end{bmatrix}$$

$$= \frac{1}{(rb)^2(r-1)}\begin{bmatrix} r^2b\,\underline{Y}'(P_B - P_{\underline{1}})\underline{Y} - rb\underline{Y}'(I - P_B + P_B - P_{\underline{1}})\underline{Y} \\ -r^2b\,\underline{Y}'(P_B - P_{\underline{1}})\underline{Y} + r^2b\,\underline{Y}'(I - P_{\underline{1}})\underline{Y} \end{bmatrix}$$

$$= \frac{1}{(rb)^2(r-1)}\begin{bmatrix} rb(r-1)\underline{Y}'(P_B - P_{\underline{1}})\underline{Y} - rb\underline{Y}'(I - P_B)\underline{Y} \\ r^2b\,\underline{Y}'(I - P_B)\underline{Y} \end{bmatrix} = \begin{bmatrix} \frac{b-1}{rb}\text{MSA} - \frac{1}{r}\text{MSE} \\ \text{MSE} \end{bmatrix}.$$

iii) REMLQE for $\underline{\psi} = [\sigma_b^2, \sigma_e^2]'$.

Let $N_X = I - P_{\underline{1}}$ and note sp$\mathcal{V}_{N_X}$ = sp$\{N_X, N_XBB'N_X\}$ = sp$\{I - P_{\underline{1}}, P_B - P_{\underline{1}}\}$ is a QS

$\Rightarrow \exists$ FELREMLQE for $\underline{\psi}$ by the LQNY-FELREMLQE theorem given by

$$\{\text{tr}(V_iN_XV_jN_X)\}_{2\times 2}\,\widehat{\underline{\psi}}_{\text{REMLQ}} = \{\underline{Y}'N_XV_iN_X\underline{Y}\}_{2\times 1}$$

$$\Rightarrow \widehat{\underline{\psi}}_{\text{REMLQ}} = \begin{bmatrix} r^2\text{tr}(P_BN_XP_BN_X) & r\text{tr}(P_BN_XIN_X) \\ r\text{tr}(IN_XP_BN_X) & \text{tr}(IN_XIN_X) \end{bmatrix}^{-1} \begin{bmatrix} r\underline{Y}'N_XP_BN_X\underline{Y} \\ \underline{Y}'N_X\underline{Y} \end{bmatrix}$$

$$= \begin{bmatrix} r^2\text{tr}(P_B - P_{\underline{1}}) & r\text{tr}(P_B - P_{\underline{1}}) \\ r\text{tr}(P_B - P_{\underline{1}}) & \text{tr}(I - P_{\underline{1}}) \end{bmatrix}^{-1} \begin{bmatrix} r\underline{Y}'(P_B - P_{\underline{1}})\underline{Y} \\ \underline{Y}'(I - P_{\underline{1}})\underline{Y} \end{bmatrix}$$

$$= \frac{1}{r^2(b-1)(br-1)-(r(b-1))^2}\begin{bmatrix} br-1 & -r(b-1) \\ -r(b-1) & r^2(b-1) \end{bmatrix}\begin{bmatrix} r\underline{Y}'(P_B - P_{\underline{1}})\underline{Y} \\ \underline{Y}'(I - P_{\underline{1}})\underline{Y} \end{bmatrix}$$

$$= \frac{1}{br^2(b-1)(r-1)}\begin{bmatrix} (br-1)r\,\underline{Y}'(P_B - P_{\underline{1}})\underline{Y} - r(b-1)\underline{Y}'(I - P_B + P_B - P_{\underline{1}})\underline{Y} \\ -r^2(b-1)\,\underline{Y}'(P_B - P_{\underline{1}})\underline{Y} + r^2(b-1)\,\underline{Y}'(I - P_{\underline{1}})\underline{Y} \end{bmatrix}$$

$$= \frac{1}{br^2(b-1)(r-1)}\begin{bmatrix} br(r-1)\underline{Y}'(P_B - P_{\underline{1}})\underline{Y} - r(b-1)\underline{Y}'(I - P_B)\underline{Y} \\ r^2(b-1)\,\underline{Y}'(I - P_B)\underline{Y} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{r}(\text{MSA} - \text{MSE}) \\ \text{MSE} \end{bmatrix} = \text{ANOVA} \quad \text{(Searle et al.,1992,p59).}$$

The estimates under both the ML and REML procedures are explicit and linear. Note that the FELMLQE for μ has an interpretative expression. The FELMLQE and FELRELMQE for σ_e^2 are identical while the FELMLQE and the FELRELMQE for σ_b^2 are not the same.

5. UBLUE for Mean Estimable Functions

The results of this chapter extend the work of Szatrowski (1980) and ElBassiouni (1983). This chapter examines explicit linear representations involving mean estimable linear combinations of the parameter vector. The assumption of a full rank model is not necessary in this chapter and will be discussed further in chapter 6. The results will first be presented for the underlying model and then applied to the particular models of interest for examining linear combinations of the fixed effects and linear combinations of the variance components.

5.1. UBLUE for the Underlying Model

This section examines conditions under which a UBLUE exists for a linear transformation of the mean parameters in the U-Model. Consider the linear transformation defined in section 3.3.1 given by $\Pi^* : \mathcal{P} \to \mathcal{H}$, also denoted $\Pi^*\theta$. Assumptions for this section are that $\Pi^*\theta$ is mean estimable in the U-Model under [S]. Assumption [S] is needed to apply the Mean Estimability theorem in section 3.3.2 which indicates that $\Pi^*\theta$ is mean estimable if and only if $\underline{R}(\Pi) \subset \underline{R}(U^*)$. Two useful lemmas are given below.

Lemma 1: If $\underline{R}(\Pi) \subset \underline{R}(U^*)$ and V is PD, then $\Pi = U^*V^{-1}U(U^*V^{-1}U)^-\Pi$.

proof: Let $P = U^*V^{-1}U(U^*V^{-1}U)^- \Rightarrow P$ is a PO on $\underline{R}(U^*V^{-1}U)$ from section 2.5 and $\underline{R}(U^*V^{-1}U) = \underline{R}(U^*)$ from section 2.7. Thus,
$U^*V^{-1}U(U^*V^{-1}U)^-\Pi = P\Pi = \Pi$ as $\underline{R}(\Pi) \subset \underline{R}(U^*)$ by hypothesis. ∎

Lemma 2: The following are equivalent under [S]:

i) $\Pi^*\widehat{\theta}_{V_0} = \Pi^*(U^*V_0^{-1}U)^-U^*V_0^{-1}\omega$ is UBLUE for $\Pi^*\theta$ in the U-Model

ii) $\underline{R}(VV_0^{-1}U(U^*V_0^{-1}U)^-\Pi) \subset \underline{R}(U) \quad \forall\, V \in \mathcal{V}$.

iii) $\Pi^*(U^*V^{-1}U)^-U^*V^{-1} = \Pi^*(U^*V_0^{-1}U)^-U^*V_0^{-1} \quad \forall\, V \in \mathcal{V}$.

proof: (1) Note $E[\Pi^*\widehat{\theta}_{V_0}] = \Pi^*(U^*V_0^{-1}U)^-U^*V_0^{-1}U\theta = \Pi^*\theta$ by lemma 1.

(2) i) $\Leftrightarrow \Pi^*\widehat{\theta}_{V_0}$ is UBLUE for $\Pi^*\theta \Leftrightarrow \underline{R}(VV_0^{-1}U(U^*V_0^{-1}U)^-\Pi) \subset \underline{R}(U) \;\forall\, V \in \mathcal{V}$

$\Leftrightarrow$ ii) by Zyskind's theorem and (1).

(3) Then ii) $\Leftrightarrow \underline{R}(VV_0^{-1}U(U^*V_0^{-1}U)^-\Pi) \subset \underline{R}(U) \quad \forall\, V \in \mathcal{V}$

$\Leftrightarrow \underline{R}(V_0^{-1}U(U^*V_0^{-1}U)^-\Pi) \subset \underline{R}(V^{-1}U) \quad \forall\, V \in \mathcal{V}$

$\Leftrightarrow V_0^{-1}U(U^*V_0^{-1}U)^-\Pi = V^{-1}U(U^*V^{-1}U)^-U^*V_0^{-1}U(U^*V_0^{-1}U)^-\Pi \;\; \forall\, V \in \mathcal{V}$ by projection theorem

$\Leftrightarrow V_0^{-1}U(U^*V_0^{-1}U)^-\Pi = V^{-1}U(U^*V^{-1}U)^-\Pi \quad \forall\, V \in \mathcal{V}$ by lemma 1

$\Leftrightarrow \Pi^*(U^*V^{-1}U)^-U^*V^{-1} = \Pi^*(U^*V_0^{-1}U)^-U^*V_0^{-1} \;\; \forall\, V \in \mathcal{V} \Leftrightarrow$ iii). ∎

For purposes of discussing the least squares estimator, interest is in $V_0 = I$. The least squares estimator, the generalized least squares estimator, and the estimated generalized least squares estimator for $\Pi^*\theta$ will be of particular interest. Since $\Pi^*\theta$ is mean estimable and $\Pi = U^*M$ for some M, the least squares estimators for $\Pi^*\theta$ are linear combinations of those derived in section 3.3.3. They are given by:

LSE: $\Pi^*\underline{\widehat{\theta}}_I = \Pi^*(U^*U)^-U^*\omega$

GLSE: $\Pi^*\widehat{\theta}_V = \Pi^*(U^*V^{-1}U)^-U^*V^{-1}\omega$ for a given $V \in \mathcal{L}_{\mathrm{PD}}(\mathcal{W},\mathcal{W})$

EGLSE: $\Pi^*\widehat{\theta}_{\widehat{V}} = \Pi^*(U^*\widehat{V}^{-1}U)^-U^*\widehat{V}^{-1}\omega$ with $\widehat{V} = \widehat{V}(\omega) \in \mathcal{L}_{\mathrm{PD}}(\mathcal{W},\mathcal{W})\ \forall\ \omega \in \mathcal{W}$.

The LSE has a simple linear form which does not depend on the covariance, the GLSE depends on the unknown true V, and the EGLSE is not linear and requires an estimate of the unknown variance. The objective is to determine when the LSE is equal to the EGLSE. The above lemma can be used to generate the main UBLUE results of this section. These results are now presented.

U-UBLUE Theorem: The following are equivalent under [S]:

i) $\Pi^*\widehat{\theta}_I$ is UBLUE for $\Pi^*\theta$ for the U-Model

ii) $\underline{R}(VU(U^*U)^-\Pi) \subset \underline{R}(U)\ \ \forall\ V \in \mathcal{V}$

iii) $\Pi^*(U^*V^{-1}U)^-U^*V^{-1} = \Pi^*(U^*U)^-U^*\ \ \forall\ V \in \mathcal{V}$.

proof: Apply the above lemma where $V_0 = I$. ■

Proposition 1: If $V \in \mathrm{sp}\mathcal{V}$ and V is PD, then the conditions in the above theorem imply $\Pi^*\underline{\widehat{\theta}}_I = \Pi^*\widehat{\theta}_V$.

proof: $\underline{R}(VU(U^*U)^-\Pi) \subset \underline{R}(U)\ \ \forall\ V \in \mathcal{V} \Rightarrow \underline{R}(VU(U^*U)^-\Pi) \subset \underline{R}(U)\ \ \forall\ V \in \mathrm{sp}\mathcal{V}$ by section 2.3
$\Rightarrow \Pi^*(U^*V^{-1}U)^-U^*V^{-1} = \Pi^*(U^*U)^-U^*\ \forall\ V \in \mathrm{sp}\mathcal{V} \ni V$ is PD by the U-UBLUE theorem. ■

Proposition 2: If $I \in \mathrm{sp}\mathcal{V}$, then $\Pi^*\theta$ has a UBLUE $\Leftrightarrow \Pi^*\widehat{\theta}_I$ is UBLUE for $\Pi^*\theta$.

proof: i) $\Pi^*\widehat{\theta}_I$ is UBLUE for $\Pi^*\theta \Rightarrow \Pi^*\theta$ has a UBLUE.

ii) Suppose $\Pi^*\theta$ has a UBLUE given by $\Pi^*\widehat{\theta}_{V_0}$ where $V_0 \in \mathcal{V}$

$\Rightarrow \underline{R}(VV_0^{-1}U(U^*V_0^{-1}U)^-\Pi) \subset \underline{R}(U)\qquad \forall\ V \in \mathcal{V}$ by lemma

$\Rightarrow \underline{R}(VV_0^{-1}U(U^*V_0^{-1}U)^-\Pi) \subset \underline{R}(U)\qquad \forall\ V \in \mathrm{sp}\mathcal{V}$ from section 2.3.

$\Rightarrow \underline{R}(V_0^{-1}U(U^*V_0^{-1}U)^-\Pi) \subset \underline{R}(U) = \underline{R}(P_U)\qquad$ since $I \in \mathrm{sp}\ \mathcal{V}$

$\Rightarrow V_0^{-1}U(U^*V_0^{-1}U)^-\Pi = U(U^*U)^+U^*V_0^{-1}U(U^*V_0^{-1}U)^-\Pi = U(U^*U)^-\Pi$ by lemma 1

$\Rightarrow\ \Pi^*(U^*V_0^{-1}U)^-U^*V_0^{-1} = \Pi^*(U^*U)^-U^* \Rightarrow \Pi^*\widehat{\theta}_{V_0} = \Pi^*\widehat{\theta}_I$

$\Rightarrow \Pi^*\widehat{\theta}_I$ is UBLUE for $\Pi^*\theta$ by the Uniqueness theorem in section 3.3.4. ■

Condition ii) of the U-UBLUE theorem is a general form of Zyskind's condition (ZC) and will be called the Generalized Zyskind's condition (GZC). This condition is equivalent to ZC when $\Lambda = U^*$. Condition i) of the U-UBLUE theorem indicates the UBLUE has an expression which does not depend on V. The conditions of the U-UBLUE theorem can be re-expressed in a convenient form. Define the set $\mathcal{L}_{\Pi}(\mathcal{W},\mathcal{W}) = \{V \in \mathcal{L}_{\mathrm{PD}}(\mathcal{W},\mathcal{W}) | \underline{R}(VU(U^*U)^-\Pi) \subset \underline{R}(U)\}$. The first three corollaries restate the previous results while the last one applies the results to the EGLSE.

Corollary: GZC $\Leftrightarrow \mathcal{V} \subset \mathcal{L}_{\Pi}(\mathcal{W},\mathcal{W})$.

Corollary: If $\mathcal{V} \subset \mathcal{L}_{\Pi}(\mathcal{W},\mathcal{W})$, then GLSE = LSE = UBLUE.

Corollary: If $\mathcal{V} \subset \mathcal{L}_{\Pi}(\mathcal{W},\mathcal{W})$, then $\mathrm{sp}\mathcal{V} \cap \mathcal{L}_{\mathrm{PD}}(\mathcal{W},\mathcal{W}) \subset \mathcal{L}_{\Pi}(\mathcal{W},\mathcal{W})$.

Corollary: If $\Pi^*\theta$ has a UBLUE and $\widehat{V} \in \mathcal{L}_{\Pi}(\mathcal{W},\mathcal{W})$, then EGLSE = GLSE = LSE = UBLUE.

It is important to know that the GZC does not depend upon the parameterization of the expectation. This property will be examined by defining a reparameterization for the mean of the U-Model. Consider the mean parameterization defined in section 3.3.1 and the reparameterization $T : \mathcal{Q} \to \mathcal{W}$ where T is a linear transformation, $T(A_{\mathcal{U}}) = \mathcal{U}$, and $\mathrm{sp}A_{\mathcal{U}} = \mathcal{Q}$. Thus, $\underline{R}(T) = T(\mathcal{Q}) = T(\mathrm{sp}A_{\mathcal{U}}) = \mathrm{sp}T(A_{\mathcal{U}})$ $= \mathrm{sp}\mathcal{U} = \underline{R}(U)$. The following definition is useful for relating two linear transformations $\Pi^*\theta$ and $\Gamma^*\alpha$ into the same space where $\Pi^* : \mathcal{P} \to \mathcal{H}$ and $\Gamma^* : \mathcal{Q} \to \mathcal{H}$.

Definition: <u>Mean Correspondence</u> - $\Pi^*\theta$ and $\Gamma^*\alpha$ have mean correspondence ($\Pi^*\theta \doteq \Gamma^*\alpha$) provided that $\forall\, \theta \in \Upsilon_{\mathcal{U}}, \alpha \in A_{\mathcal{U}}$, $[U\theta = T\alpha \Rightarrow \Pi^*\theta = \Gamma^*\alpha]$.

The next lemma is useful for defining Γ in order to have mean correspondence. This lemma and the reparameterization defined above will be used to generate the theorem.

Lemma: Consider the U-Model under [S] where $\Pi^*\theta$ is estimable and $\Pi = U^*M$ for some M. Then $\Pi^*\theta \doteq \Gamma^*\alpha \Leftrightarrow \Gamma = T^*M$.

proof: i) By the Mean Estimability theorem assuming [S], $\Pi^*\theta$ is estimable $\Leftrightarrow \underline{R}(\Pi) \subset \underline{R}(U^*)$ $\Leftrightarrow \Pi = U^*M$ for some M.

ii) Suppose $\Pi^*\theta \doteq \Gamma^*\alpha$. Then $T(A_{\mathcal{U}}) = \mathcal{U} = U(\Upsilon_{\mathcal{U}}) \Rightarrow \forall \alpha \in A_{\mathcal{U}}\ \exists \theta \in \Upsilon_{\mathcal{U}} \ni U\theta = T\alpha$ $\Rightarrow \forall \alpha \in A_{\mathcal{U}}\ \exists \theta \in \Upsilon_{\mathcal{U}} \ni M^*U\theta = M^*T\alpha \Rightarrow \forall \alpha \in A_{\mathcal{U}}\ \exists \theta \in \Upsilon_{\mathcal{U}} \ni \Pi^*\theta = M^*T\alpha$ by assumption

$\Rightarrow \forall\, \alpha \in A_{\mathcal{U}} \quad \Gamma^*\alpha = M^*T\alpha \quad$ since $\Pi^*\theta \doteq \Gamma^*\alpha$

$\Rightarrow \forall\, \alpha \in \text{sp}A_{\mathcal{U}} = \mathcal{Q} \quad \Gamma^*\alpha = M^*T\alpha \Rightarrow \Gamma^* = M^*T \Rightarrow \Gamma = T^*M.$

iii) Suppose $\Gamma = T^*M$. Then $\forall\, \theta \in \Upsilon_{\mathcal{U}}, \alpha \in A_{\mathcal{U}}\ [U\theta = T\alpha \Rightarrow M^*U\theta = M^*T\alpha \Rightarrow \Pi^*\theta = \Gamma^*\alpha]$

$\Rightarrow \Pi^*\theta \doteq \Gamma^*\alpha$ by definition. ∎

Theorem: If $\Pi^*\theta \doteq \Gamma^*\alpha$ under the two parameterizations described above, where $\Pi^*\theta$ is estimable in the U-Model under [S], then the GZC for $\Pi^*\theta$ is equivalent to the GZC for $\Gamma^*\alpha$.

proof: Note $\underline{R}(U) = \underline{R}(T)$. By the Mean Estimability theorem assuming [S], $\Pi = U^*M$ for some M. Then $\underline{R}(VU(U^*U)^-\Pi) = \underline{R}(VU(U^*U)^-U^*M) = \underline{R}(VP_UM)$

$= \underline{R}(VP_TM) = \underline{R}(VT(T^*T)^-T^*M) = \underline{R}(VT(T^*T)^-\Gamma)$ by the above lemma.

Hence, $\underline{R}(VU(U^*U)^-\Pi) \subset \underline{R}(U) \Leftrightarrow \underline{R}(VT(T^*T)^-\Gamma) \subset \underline{R}(T)$. ∎

The next theorem shows the relationship between ZC and GZC. The results assuming GZC are more general than the results assuming ZC as ZC implies GZC.

Theorem: i) If $\underline{r}(\Pi) = \underline{r}(U)$, then GZC $\Rightarrow$ ZC $\qquad$ ii) ZC $\Rightarrow$ GZC.

proof: i) By mean estimability $\underline{R}(\Pi) \subset \underline{R}(U^*) \Rightarrow \underline{R}(\Pi) = \underline{R}(U^*)$ as have equality of ranks.

Thus, GZC $\Rightarrow \underline{R}(VU(U^*U)^-\Pi) \subset \underline{R}(U) \quad \forall\, V \in \mathcal{V}$

$\Rightarrow \underline{R}(VU(U^*U)^-U^*) \subset \underline{R}(U) \qquad \forall\, V \in \mathcal{V}$ as $\underline{R}(\Pi) = \underline{R}(U^*)$

$\Rightarrow \underline{R}(VP_U) \subset \underline{R}(U)\ \forall\, V \in \mathcal{V} \Rightarrow \underline{R}(VU) \subset \underline{R}(U) \quad \forall\, V \in \mathcal{V}.$

ii) By mean estimability $\underline{R}(\Pi) \subset \underline{R}(U^*) \Rightarrow \Pi = U^*M$. Then ZC $\Rightarrow \underline{R}(VU) \subset \underline{R}(U)\ \forall\, V \in \mathcal{V}$

$\Rightarrow U(U^*V^{-1}U)^-U^*V^{-1} = U(U^*U)^-U^* \quad \forall\, V \in \mathcal{V}$ by the ZC Relation theorem

$\Rightarrow M^*U(U^*V^{-1}U)^-U^*V^{-1} = M^*U(U^*U)^-U^* \quad \forall\, V \in \mathcal{V}$

$\Rightarrow \Pi^*(U^*V^{-1}U)^-U^*V^{-1} = \Pi^*(U^*U)^-U^* \quad \forall\, V \in \mathcal{V}$

$\Rightarrow \underline{R}(VU(U^*U)^-\Pi) \subset \underline{R}(U) \quad \forall\, V \in \mathcal{V} \Rightarrow$ GZC by the U-UBLUE theorem. ∎

A sufficient condition for the existence of a UBLUE is the Generalized Zyskind's condition. Under this condition, the UBLUE is the LSE, which does not depend on the covariance matrix V. The results of this section will be applied to the other models of interest for purposes of estimating fixed effects and variance components.

5.2. UBLUE for Fixed Effects

UBLUEs for fixed effects can be obtained using the Y-Model and applying the results from the U-Model. Consider the estimable linear combination of the fixed effect vector $\Lambda'\underline{\beta}$. Recall $\underline{\beta} \in \mathcal{R}^p$ and $\mathcal{R}^p$ contains a non-empty open set, so the Mean Estimability theorem can be applied. First, results will be presented for UBLUEs which will then be applied to maximum likelihood estimation.

5.2.1. UBLUE Results

This section examines conditions under which a UBLUE exists in the Y-Model for $\Lambda'\underline{\beta}$. The least squares estimator, the generalized least squares estimator, and the estimated generalized least squares estimator for $\Lambda'\underline{\beta}$ will be of interest in this section. They are listed here for reference:

LSE: $\Lambda'\widehat{\underline{\beta}}_I = \Lambda'(X'X)^-X'\underline{Y}$

GLSE: $\Lambda'\widehat{\underline{\beta}}_{\underline{\psi}} = \Lambda'(X'V_{\underline{\psi}}^{-1}X)^-X'V_{\underline{\psi}}^{-1}\underline{Y}$ for a given $\underline{\psi}$ where $V_{\underline{\psi}}$ is PD

EGLSE: $\Lambda'\widehat{\underline{\beta}}_{\widehat{\underline{\psi}}} = \Lambda'(X'V_{\widehat{\underline{\psi}}}^{-1}X)^-X'V_{\widehat{\underline{\psi}}}^{-1}\underline{Y}$ where $V_{\widehat{\underline{\psi}}} = V_{\widehat{\underline{\psi}}(\underline{Y})}$ is PD $\forall\ \underline{Y} \in \mathcal{R}^n$.

The U-UBLUE theorem for the U-Model can be directly applied in this setting to indicate when the LSE is equal to the GLSE for estimating $\Lambda'\underline{\beta}$. This is stated in the next theorem and restated in the following proposition under assumptions [L] and [O].

Y-UBLUE Theorem: The following are equivalent:

i) $\Lambda'\widehat{\underline{\beta}}_I$ is UBLUE for $\Lambda'\underline{\beta}$ in the Y-Model

ii) $\underline{R}(V_{\underline{\psi}}X(X'X)^-\Lambda) \subset \underline{R}(X)\ \ \forall\ \underline{\psi} \in \Xi$

iii) $\Lambda'(X'V_{\underline{\psi}}^{-1}X)^-X'V_{\underline{\psi}}^{-1} = \Lambda'(X'X)^-X'\ \ \forall\ \underline{\psi} \in \Xi$.

Proposition: Assume [L] and [O]. Then $\forall\ \underline{\psi} \in \mathcal{R}^{k+1} \ni V_{\underline{\psi}}$ is PD, the conditions in the Y-FUBLUE theorem imply that $\Lambda'\widehat{\underline{\beta}}_{\underline{\psi}} = \Lambda'\widehat{\underline{\beta}}_I$.

proof: Under [L], [O], sp$\mathcal{V} = \{V_{\underline{\psi}} |\ \underline{\psi} \in \text{sp}\ \Xi\} = \{V_{\underline{\psi}} | \underline{\psi} \in \mathcal{R}^{k+1}\}$. Apply proposition 1 in section 5.1. ∎

The Y-UBLUE theorem does not make any assumptions about normality, linear covariance structure, or classification matrices. Only the form of GZC given in ii) is necessary to apply this theorem to the special cases of the Y-Model under [L], [O], and [C].

5.2.2. ELMLQE Results

For doing maximum likelihood estimation in the Y-Model, the assumptions [L], [O], and [N] were used. Extending the definition of Szatrowski (1980), the explicit linear maximum likelihood equation estimator (ELMLQE) for $\Lambda'\underline{\beta}$ satisfies $\Lambda'\widehat{\underline{\beta}}_{\text{MLQ}} = A\underline{Y}$ for a constant matrix A which is not random and does not depend on any parameters. In this case, the estimator $\Lambda'\widehat{\underline{\beta}}_{\text{MLQ}}$ is linear and the explicit part indicates that A is constant. The following lemma shows that the MLQE = EGLSE in this setting.

Lemma: $\Lambda'\widehat{\underline{\beta}}_{\text{MLQ}} = \Lambda'\widehat{\underline{\beta}}_{\widehat{\underline{\psi}}_{\text{MLQ}}}$.

proof: Since $\Lambda'\underline{\beta}$ is estimable $\Rightarrow \underline{R}(\Lambda) \subset \underline{R}(X') \Rightarrow \Lambda = X'M$ for some M. Recall, it has been shown $X\widehat{\underline{\beta}}_{\text{MLQ}} = X\widehat{\underline{\beta}}_{\widehat{\underline{\psi}}_{\text{MLQ}}} \Rightarrow M'X\widehat{\underline{\beta}}_{\text{MLQ}} = M'X\widehat{\underline{\beta}}_{\widehat{\underline{\psi}}_{\text{MLQ}}} \Rightarrow \Lambda'\widehat{\underline{\beta}}_{\text{MLQ}} = \Lambda'\widehat{\underline{\beta}}_{\widehat{\underline{\psi}}_{\text{MLQ}}}$. ∎

The next theorem presents a sufficient condition for the existence of an ELMLQE for $\Lambda'\underline{\beta}$. This condition is GZC for the Y-Model.

Y-ELMLQE Theorem: Consider the Y-Model under [L], [O], and [N]. If $\widehat{\underline{\psi}}_{\text{MLQ}}$ exists $\ni V_{\widehat{\underline{\psi}}_{\text{MLQ}}}$ is PD and $\underline{R}(V_{\underline{\psi}}X(X'X)^{-}\Lambda) \subset \underline{R}(X)\ \forall\ \underline{\psi} \in \Xi$, then $\Lambda'\widehat{\underline{\beta}}_I$ is an ELMLQE for $\Lambda'\underline{\beta}$.

proof: Since $V_{\widehat{\underline{\psi}}_{\text{MLQ}}}$ is PD by hypothesis, $\Lambda'\widehat{\underline{\beta}}_{\widehat{\underline{\psi}}_{\text{MLQ}}} = \Lambda'\widehat{\underline{\beta}}_I$ by the proposition in section 5.2.1.
By the above lemma, $\Lambda'\widehat{\underline{\beta}}_{\text{MLQ}} = \Lambda'\widehat{\underline{\beta}}_{\widehat{\underline{\psi}}_{\text{MLQ}}}$. Thus, $\Lambda'\widehat{\underline{\beta}}_{\text{MLQ}} = \Lambda'\widehat{\underline{\beta}}_I = A\underline{Y}$ where $A = \Lambda'(X'X)^{-}X'$
$\Rightarrow \Lambda'\widehat{\underline{\beta}}_I$ is an ELMLQE for $\Lambda'\underline{\beta}$ by definition. ∎

The proof of the theorem shows the MLQE is equivalent to the LSE under the assumptions. Because the LSE is linear and explicit, there exists an ELMLQE for $\Lambda'\underline{\beta}$. An application of the Y-ELMLQE theorem is given in section 5.4.

5.3. UBLUE for Variance Components

UBLUEs for the variance components can be obtained using linearized quadratic estimation models and applying the results from the U-Model. Consider a vector of mean estimable linear combinations of the variance components $\Gamma'\underline{\psi}$ where $\Gamma : \mathcal{R}^s \rightarrow \mathcal{R}^r$ and $\underline{\psi} \in \Xi$. In order to apply the results of the Mean Estimability theorem, it is sufficient to assume Ξ contains a non-empty open set in $\mathcal{R}^r$. This assumption does hold for the ALQEM for $(\underline{Y} - X\widehat{\underline{\beta}})$ and the LQEM for $N_X\underline{Y}$. The UBLUE results will be presented first for the LQEM for $\underline{Z}$, and applied to the ALQEM for $(\underline{Y} - X\widehat{\underline{\beta}})$ for maximum likelihood estimation and to the LQEM for $N_X\underline{Y}$ for restricted maximum likelihood estimation.

5.3.1. UBLUE Results

The UBLUE results will be given for the LQEM for $\underline{Z}$ under the open set assumption for Ξ. The associated least squares, generalized least squares, and estimated generalized least squares estimator for $\Gamma'\underline{\psi}$ will be of interest. These are listed here for reference:

LSE: $\Gamma'\widehat{\underline{\psi}}_I = \Gamma'(X^{\dagger *}X^{\dagger})^{-}X^{\dagger *}Y^{\dagger}$

GLSE: $\Gamma'\widehat{\underline{\psi}}_{\underline{\psi}} = \Gamma'(X^{\dagger *}V_{\underline{\psi}}^{\dagger -1}X^{\dagger})^{-}X^{\dagger *}V_{\underline{\psi}}^{\dagger -1}Y^{\dagger}$ for a given $\underline{\psi}$ where $V_{\underline{\psi}}^{\dagger}$ is PD

EGLSE: $\Gamma'\widehat{\underline{\psi}} = \Gamma'(X^{\dagger *}V_{\widehat{\underline{\psi}}}^{\dagger -1}X^{\dagger})^{-}X^{\dagger *}V_{\widehat{\underline{\psi}}}^{\dagger -1}Y^{\dagger}$ where $V_{\widehat{\underline{\psi}}}^{\dagger} = V_{\widehat{\underline{\psi}}(Y)}^{\dagger}$ is PD $\forall\ \underline{Y} \in \mathcal{R}^n$.

The UBLUE result for the U-Model can be used to indicate when the LSE is the same as the EGLSE when estimating $\Gamma'\underline{\psi}$. This is stated in the next theorem. The proof of the theorem follows directly from the proof of the U-UBLUE theorem and is restated in the following proposition.

LQZ-UBLUE Theorem: The following are equivalent when Ξ contains a non-empty open set in $\mathcal{R}^r$:

i) $\Gamma'\widehat{\underline{\psi}}_I$ is UBLUE for $\Gamma'\underline{\psi}$ in the LQEM for $\underline{Z}$

ii) $\underline{R}(V_{\underline{\psi}}^{\dagger}X^{\dagger}(X^{\dagger *}X^{\dagger})^{-}\Gamma) \subset \underline{R}(X^{\dagger}) \forall\ \underline{\psi} \in \Xi$

iii) $\Gamma'(X^{\dagger *}V_{\underline{\psi}}^{\dagger -1}X^{\dagger})^{-}X^{\dagger *}V_{\underline{\psi}}^{\dagger -1} = \Gamma'(X^{\dagger *}X^{\dagger})^{-}X^{\dagger *}\quad \forall\ \underline{\psi} \in \Xi$.

Proposition: For any $\underline{\psi} \in \mathcal{R}^r \ni R_{\underline{\psi}}$ is PD, the conditions in the LQZ-UBLUE theorem imply $\Gamma'\widehat{\underline{\psi}}_{\underline{\psi}} = \Gamma'\widehat{\underline{\psi}}_I$.

proof: Condition ii) in the LQZ-UBLUE theorem $\Rightarrow \underline{R}(V_{\underline{\psi}}^{\dagger}X^{\dagger}(X^{\dagger *}X^{\dagger})^{-}\Gamma) \subset \underline{R}(X^{\dagger})\ \forall\ \underline{\psi} \in \Xi$

$\Rightarrow V_{\underline{\psi}}^{\dagger}X^{\dagger}(X^{\dagger *}X^{\dagger})^{-}\Gamma\underline{u} \in \underline{R}(X^{\dagger})\ \forall\ \underline{\psi} \in \Xi, \underline{u} \in \mathcal{R}^s$

$\Rightarrow 2R_{\underline{\psi}}B_{\underline{u}}R_{\underline{\psi}} \in \underline{R}(X^{\dagger})\quad \forall\ \underline{\psi} \in \Xi, \underline{u} \in \mathcal{R}^s$ where $B_{\underline{u}} = X^{\dagger}(X^{\dagger *}X^{\dagger})^{-}\Gamma\underline{u} \in \mathcal{S}_n$

$\Rightarrow \{R_{\underline{\psi}}B_{\underline{u}}R_{\underline{\psi}} | \underline{\psi} \in \Xi\} \subset \underline{R}(X^{\dagger}) \forall \underline{u} \in \mathcal{R}^s \Rightarrow \mathrm{sp}\{R_{\underline{\psi}}B_{\underline{u}}R_{\underline{\psi}} | \underline{\psi} \in \Xi\} \subset \underline{R}(X^{\dagger}) \forall\ \underline{u} \in \mathcal{R}^s$

$\Rightarrow \mathrm{sp}\{R_{\underline{\psi}}B_{\underline{u}}R_{\underline{\psi}} | \underline{\psi} \in \mathcal{R}^r\} \subset \underline{R}(X^{\dagger}) \forall\ \underline{u} \in \mathcal{R}^s$ by lemma 1 in section 4.3.1

$\Rightarrow R_{\underline{\psi}}B_{\underline{u}}R_{\underline{\psi}} \in \underline{R}(X^{\dagger})\ \forall\ \underline{\psi} \in \mathcal{R}^r, \underline{u} \in \mathcal{R}^s \Rightarrow V_{\underline{\psi}}^{\dagger}X^{\dagger}(X^{\dagger *}X^{\dagger})^{-}\Gamma\underline{u} \in \underline{R}(X^{\dagger})\ \forall\ \underline{\psi} \in \mathcal{R}^r, \underline{u} \in \mathcal{R}^s$

$\Rightarrow \underline{R}(V_{\underline{\psi}}^{\dagger}X^{\dagger}(X^{\dagger *}X^{\dagger})^{-}\Gamma) \subset \underline{R}(X^{\dagger})\ \forall\ \underline{\psi} \in \mathcal{R}^r$

$\Rightarrow$ condition iii) in the LQZ-UBLUE theorem $\forall\ \underline{\psi} \in \mathcal{R}^r \ni R_{\underline{\psi}}$ is PD

$\Rightarrow \Gamma'\widehat{\underline{\psi}}_{\underline{\psi}} = \Gamma'\widehat{\underline{\psi}}_I\ \forall\ \underline{\psi} \in \mathcal{R}^r \ni R_{\underline{\psi}}$ is PD. ■

This theorem was stated for the LQEM for $\underline{Z}$, so it applies to both the ALQEM for $N_X\underline{Y}$ and the LQEM for $N_X\underline{Y}$. The following sections use these models to examine the maximum likelihood and restricted maximum likelihood procedures.

5.3.2. ELMLQE Results

For doing maximum likelihood estimation in the Y-Model, the assumptions [L], [O], and [N] were used. Extending the definition of Szatrowski (1980), the explicit linear maximum likelihood estimator (ELMLQE) for $\Gamma'\underline{\psi}$ satisfies $\Gamma'\widehat{\underline{\psi}}_{\text{MLQ}} = AY_2^\diamond$, where A is a linear transformation that is not random and does not depend on any parameters and $Y_2^\diamond$ does not depend on any estimators or unknown parameters.

The associated least squares, generalized least squares, and the estimated generalized least squares estimators for the ALQEM for $(\underline{Y} - X\widehat{\underline{\beta}})$ will be of importance in this section. They are listed below:

LSE: $\Gamma'\widehat{\underline{\psi}}_I = \Gamma'(X^{\diamond *}X^\diamond)^- X^{\diamond *}Y_2^\diamond$

GLSE: $\Gamma'\widehat{\underline{\psi}}_{\underline{\psi}} = \Gamma'(X^{\diamond *}V_{\underline{\psi}}^{\diamond -1}X^\diamond)^- X^{\diamond *}V_{\underline{\psi}}^{\diamond -1}Y_2^\diamond$ for a given $\underline{\psi}$ where $V_{\underline{\psi}}^\diamond$ is PD

EGLSE: $\Gamma'\widehat{\underline{\psi}} = \Gamma'(X^{\diamond *}V_{\widehat{\underline{\psi}}}^{\diamond -1}X^\diamond)^- X^{\diamond *}V_{\widehat{\underline{\psi}}}^{\diamond -1}Y_2^\diamond$ where $V_{\widehat{\underline{\psi}}}^\diamond = V_{\widehat{\underline{\psi}}(\underline{Y})}^\diamond$ is PD $\forall\ \underline{Y} \in \mathcal{R}^n$.

For clarity, the EGLSE will often be labelled as $\widehat{\underline{\psi}}_{\text{EGLS}}$. The ALQEM for $(\underline{Y} - X\widehat{\underline{\beta}})$ was defined so that the equations for the EGLSE correspond to the ML equations for $\underline{\psi}$. The following lemma establishes the equivalence between the EGLSE and the MLQE for $\Gamma'\underline{\psi}$.

Lemma: $\Gamma'\widehat{\underline{\psi}}_{\text{MLQ}} = \Gamma'\widehat{\underline{\psi}}_{\text{EGLS}}$.

proof: Since $\Gamma'\underline{\psi}$ is estimable $\Rightarrow \underline{R}(\Gamma) \subset \underline{R}(X^{\diamond *}) \Rightarrow \Gamma = X^{\diamond *}M$ for some M. From section 4.3.2, $X^\diamond\widehat{\underline{\psi}}_{\text{MLQ}} = X^\diamond\widehat{\underline{\psi}}_{\text{EGLS}} \Rightarrow M^*X^\diamond\widehat{\underline{\psi}}_{\text{MLQ}} = M^*X^\diamond\widehat{\underline{\psi}}_{\text{EGLS}} \Rightarrow \Gamma'\widehat{\underline{\psi}}_{\text{MLQ}} = \Gamma'\widehat{\underline{\psi}}_{\text{EGLS}}$. ∎

Under the ALQEM for $(\underline{Y} - X\widehat{\underline{\beta}})$, the response $Y_2^\diamond = (\underline{Y} - X\widehat{\underline{\beta}})(\underline{Y} - X\widehat{\underline{\beta}})'$ will generally depend on $\widehat{\underline{\psi}}$ as $X\widehat{\underline{\beta}} = X(X'V_{\widehat{\underline{\psi}}}^{-1}X)^- X'V_{\widehat{\underline{\psi}}}^{-1}\underline{Y}$. This is likely to present a difficulty in satisfying the definition of an ELMLQE for $\Gamma'\underline{\psi}$. In order to deal with this difficulty, it will be assumed that ZC holds for the Y-Model where $\underline{R}(V_{\underline{\psi}}X) \subset \underline{R}(X)\ \forall\ \underline{\psi} \in \Xi$. As shown in section 4.2.2, this condition is sufficient for $\underline{Y} - X\widehat{\underline{\beta}}_{\underline{\psi}} = N_X\underline{Y}\ \ \forall\, \underline{\psi} \in \mathcal{R}^{k+1}$ such that $V_{\underline{\psi}}$ is PD. Under ZC for the Y-Model, the ALQEM for $(\underline{Y} - X\widehat{\underline{\beta}})$ can be used to obtain conditions for the existence of an ELMLQE for $\Gamma'\underline{\psi}$. However, it is possible that a weaker condition could suffice for some examples as only the linear combination $\Gamma'\underline{\psi}$ is of interest. Still, the following theorem assumes ZC.

ALQNY-ELMLQE Theorem: Consider the Y-Model under [L], [O], [N], and ZC. If $\widehat{\underline{\psi}}_{\text{MLQ}}$ exists $\ni$ $V_{\widehat{\underline{\psi}}_{\text{MLQ}}}$ is PD and $\underline{R}(V_{\underline{\psi}}^{\diamond} X^{\diamond}(X^{\diamond *}X^{\diamond})^{-}\Gamma) \subset \underline{R}(X^{\diamond})$ $\forall$ $\underline{\psi} \in \Xi$, then $\Gamma'\widehat{\underline{\psi}}_I$ is an ELMLQE for $\Gamma'\underline{\psi}$.

proof: Since $V_{\widehat{\underline{\psi}}_{\text{MLQ}}}$ is PD by hypothesis and $\Gamma'\widehat{\underline{\psi}}_{\text{MLQ}} = \Gamma'\widehat{\underline{\psi}}_{\text{EGLS}}$ by the above lemma, $\Gamma'\widehat{\underline{\psi}}_{\text{MLQ}} = \Gamma'\widehat{\underline{\psi}}_I = AY_2^{\diamond}$ by the proposition in section 5.3.1 where $A = \Gamma'(X^{\diamond *}X^{\diamond})^{-}X^{\diamond *}$ and $Y_2^{\diamond} = N_X\underline{Y}\underline{Y}'N_X$ does not depend $\underline{\psi}$ or $\widehat{\underline{\psi}}$ by ZC $\Rightarrow$ $\Gamma'\widehat{\underline{\psi}}_I$ is an ELMLQE for $\Gamma'\underline{\psi}$ by definition. ∎

Note $\Gamma'\widehat{\underline{\psi}}_I$ may not be an unbiased estimator for $\Gamma'\underline{\psi}$, but it does satisfy the requirement of an ELMLQE. The explicit expression $\Gamma'\widehat{\underline{\psi}}_I$ is a linear combination of the equations given in section 3.2.3.

5.3.3. ELREMLQE Results

For doing restricted maximum likelihood estimation in the Y-Model, the assumptions [L], [O], and [N] were used. Extending the definition of Szatrowski (1980), an explicit linear restricted maximum likelihood estimator (ELREMLQE) for $\Gamma'\underline{\psi}$ satisfies $\Gamma'\widehat{\underline{\psi}}_{\text{REMLQ}} = AY^{\circ}$, where A is a linear transformation that is not random and does not depend on any parameters and Y° does not depend on any estimators or unknown parameters.

The associated least squares, generalized least squares, and estimated generalized least squares estimators for $\Gamma'\underline{\psi}$ in the LQEM for $N_X\underline{Y}$ will be of importance in this section. They are listed below:

LSE: $\Gamma'\widehat{\underline{\psi}}_I = \Gamma'(X^{\circ *}X^{\circ})^{-}X^{\circ *}Y^{\circ}$

GLSE: $\Gamma'\widehat{\underline{\psi}}_{\underline{\psi}} = \Gamma'(X^{\circ *}V_{\underline{\psi}}^{\circ -1}X^{\circ})^{-}X^{\circ *}V_{\underline{\psi}}^{\circ -1}Y^{\circ}$ for a given $\underline{\psi}$ where $V_{\underline{\psi}}^{\circ}$ is PD

EGLSE: $\Gamma'\widehat{\underline{\psi}} = \Gamma'(X^{\circ *}V_{\widehat{\underline{\psi}}}^{\circ -1}X^{\circ})^{-}X^{\circ *}V_{\widehat{\underline{\psi}}}^{\circ -1}Y^{\circ}$ where $V_{\widehat{\underline{\psi}}}^{\circ} = V_{\widehat{\underline{\psi}}(\underline{Y})}$ is PD $\forall$ $\underline{Y} \in \mathcal{R}^n$.

For clarity, the EGLSE will often be labelled as $\widehat{\underline{\psi}}_{\text{EGLS}}$. The LQEM for $N_X\underline{Y}$ was defined so that the equations for the EGLSE correspond to the REML equations for $\underline{\psi}$. The following lemma establishes the equivalence between the EGLSE and the REMLQE for $\Gamma'\underline{\psi}$.

Lemma: $\Gamma'\widehat{\underline{\psi}}_{\text{REMLQ}} = \Gamma'\widehat{\underline{\psi}}_{\text{EGLS}}$.

proof: Since $\Gamma'\underline{\psi}$ is estimable $\Rightarrow$ $\underline{R}(\Gamma) \subset \underline{R}(X^{\circ *}) \Rightarrow \Gamma = X^{\circ *}M$ for some M. From section 4.3.3, $X^{\circ}\widehat{\underline{\psi}}_{\text{REMLQ}} = X^{\circ}\widehat{\underline{\psi}}_{\text{EGLS}} \Rightarrow M^{*}X^{\circ}\widehat{\underline{\psi}}_{\text{MLQ}} = M^{*}X^{\circ}\widehat{\underline{\psi}}_{\text{EGLS}} \Rightarrow \Gamma'\widehat{\underline{\psi}}_{\text{MLQ}} = \Gamma'\widehat{\underline{\psi}}_{\text{EGLS}}$. ∎

The main theorem can now be stated concerning the existence of an ELREMLQE for $\Gamma'\underline{\psi}$.

LQNY-ELREMLQE Theorem: Consider the Y-Model under [L], [O], and [N]. If $\widehat{\underline{\psi}}_{\text{REMLQ}}$ exists $\ni$ $V_{\widehat{\underline{\psi}}_{\text{REMLQ}}}$ is PD and $\underline{R}(V^{\circ}_{\underline{\psi}}X^{\circ}(X^{\circ *}X^{\circ})^{-}\Gamma) \subset \underline{R}(X^{\circ})\ \forall\ \underline{\psi} \in \Xi$, then $\Gamma'\widehat{\underline{\psi}}_I$ is an ELREMLQE for $\Gamma'\underline{\psi}$.

proof: Since $V_{\widehat{\underline{\psi}}_{\text{REMLQ}}}$ is PD by hypothesis and $\Gamma'\widehat{\underline{\psi}}_{\text{REMLQ}} = \Gamma'\widehat{\underline{\psi}}_{\text{EGLS}}$ by the above lemma, $\Gamma'\widehat{\underline{\psi}}_{\text{REMLQ}} = \Gamma'\widehat{\underline{\psi}}_I = AY^{\circ}$ by the proposition in section 5.3.1 where $A = \Gamma'(X^{\circ *}X^{\circ})^{-}X^{\circ *}$ and $Y^{\circ} = N_X\underline{Y}\underline{Y}'N_X$ does not depend $\underline{\psi}$ or $\widehat{\underline{\psi}}$ $\Rightarrow \Gamma'\widehat{\underline{\psi}}_I$ is an ELREMLQE for $\Gamma'\underline{\psi}$ by definition. ∎

The ELREMLQE for $\Gamma'\underline{\psi}$ is given by $\Gamma'\widehat{\underline{\psi}}_I$, which is linear and UBLUE for $\Gamma'\underline{\psi}$. An application of this result is discussed in section 5.5.

5.4. ML Example: 2-Way Mixed Model with No Interaction

This example will be used to demonstrate the results for an ELMLQE for the fixed effects in a 2-way mixed model with no interaction. Consider the following notation:

$$y_{ijk} = \mu + \alpha_i + b_j + e_{ij} \quad i = 1, ..., t \quad j = 1, ..., r \quad k = 1, ..., n_{ij}$$
$$\underline{Y} = \underline{1}\mu + A\underline{\alpha} + B\underline{b} + \underline{e},\ \text{E}[\underline{Y}] = \underline{1}\mu + A\underline{\alpha} = X\underline{\beta},$$
$$\text{Cov}(\underline{Y}) = V_{\underline{\psi}} = \sigma_b^2 BB' + \sigma_e^2 I\,,\ \Xi = \{\underline{\psi} = [\sigma_b^2\ \sigma_e^2]' \,|\, \sigma_b^2 \geq 0,\ \sigma_e^2 > 0\},$$
$$N = A'B = \{n_{ij}\}_{t\times r}\,,\ \ n_{\cdot j} = \sum_{i=1}^{t} n_{ij}\cdot\ .$$

Assuming $\underline{R}(A) \cap \underline{R}(B) = \underline{R}(\underline{1}_n)$ which is equivalent to $\underline{r}(A, B) = t + r - 1$, the problem is to determine the conditions under which an ELMLQE exists for $\Lambda'\underline{\beta} = \begin{bmatrix} \alpha_1 - \alpha_t \\ \vdots \\ \alpha_{t-1} - \alpha_t \end{bmatrix}$.

Lemma: $\underline{r}(N) = 1 \Rightarrow \underline{R}(V_{\underline{\psi}}X(X'X)^{-1}\Lambda) \subset \underline{R}(X)\ \forall\ \underline{\psi} \in \Xi$.

proof: i) Note $\underline{R}(X) = \underline{R}([\,\underline{1}\ A\,]) = \underline{R}(A)$. Since the GZC is invariant under a reparameterization by proposition in section 5.1, it is helpful to consider the reparameterization given by

$$\underline{1}\mu + A\underline{\alpha} \doteq A\underline{\tau} \Rightarrow \mu + \alpha_i \doteq \tau_i \Rightarrow \alpha_i - \alpha_t = (\mu + \alpha_i) - (\mu + \alpha_t) \doteq \tau_i - \tau_t$$
$$\Rightarrow \Lambda'\underline{\beta} = [\,\underline{0}_{t-1}\ I_{t-1}\ \ -\underline{1}_{t-1}]\underline{\beta} = \begin{bmatrix} \alpha_1 - \alpha_t \\ \vdots \\ \alpha_{t-1} - \alpha_t \end{bmatrix} \doteq \begin{bmatrix} \tau_1 - \tau_t \\ \vdots \\ \tau_{t-1} - \tau_t \end{bmatrix} = [I_{t-1}\ \ -\underline{1}_{t-1}]\underline{\tau} = \Pi'\underline{\tau}$$

where $\underline{1}_t'\Pi = 0$, $\underline{r}(\Pi) = t - 1$, $\underline{n}(\underline{1}_t') = t - 1$. Thus, $\underline{R}(\Pi) = \underline{R}(\underline{1}_t)^{\perp} = \underline{N}(\underline{1}_t')$.

ii) Let $\underline{w} \in \underline{R}(\Pi) = \underline{N}(\underline{1}_t')$ by i). Note

$\underline{r}(N) = 1 \Rightarrow A'B = \underline{g}\underline{h}' \Rightarrow n_{ij} = g_i h_j \Rightarrow n_{i\cdot} = g_i h. \Rightarrow \frac{1}{n_{i\cdot}} g_i = \frac{1}{\underline{1}_r'\underline{h}}$.

Thus, $[(A'A)^{-1}\underline{g}\,]'\underline{w} \;=\; \begin{bmatrix} \frac{1}{n_{1\cdot}}g_1 \\ \vdots \\ \frac{1}{n_{t\cdot}}g_t \end{bmatrix}' \underline{w} \;=\; (\,\frac{1}{\underline{1}_r'\underline{h}}\underline{1}_t\,)'\underline{w} = 0.$

iii) For $\underline{\psi} \in \Xi$, $\; V_{\underline{\psi}}A(A'A)^{-1}\underline{w} = (\sigma_b^2 BB' + \sigma_e^2 I\,)A(A'A)^{-1}\underline{w}$

$= \sigma_b^2 B\underline{h}\underline{g}'(A'A)^{-1}\underline{w} + \sigma_e^2 A(A'A)^{-1}\underline{w}$ by ii) $A'B = \underline{g}\underline{h}'$

$= \sigma_e^2 A(A'A)^{-1}\underline{w} \in \underline{R}(A)$ since $\underline{g}'(A'A)^{-1}\underline{w} = 0$ by ii).

Hence, $\underline{R}(V_{\underline{\psi}}A(A'A)^{-1}\Pi) \subset \underline{R}(A)$ as the above is true $\forall\, \underline{w} \in R(\Pi\,)$ and $\forall\, \underline{\psi} \in \Xi$

$\Rightarrow \underline{R}(V_{\underline{\psi}}X(X'X)^{-1}\Lambda) \subset \underline{R}(X) \;\; \forall\; \underline{\psi} \in \Xi$ as the condition is invariant under a reparameterization. ■

Lemma: $\underline{R}(V_{\underline{\psi}}X(X'X)^{-1}\Lambda) \subset \underline{R}(X)\; \forall\; \underline{\psi} \in \Xi \Rightarrow \underline{r}(N) = 1.$

proof: $\underline{R}(V_{\underline{\psi}}X(X'X)^{-1}\Lambda) \subset \underline{R}(X) \;\; \forall\; \underline{\psi} \in \Xi$

$\Rightarrow \underline{R}(V_{\underline{\psi}}A(A'A)^{-1}\Pi) \subset \underline{R}(A) \;\; \forall\; \underline{\psi} \in \Xi$ under the reparameterization defined in above lemma i)

$\Rightarrow \underline{R}((\sigma_b^2 BB' + \sigma_e^2 I\,)A(A'A)^{-1}\Pi) \subset \underline{R}(A) \;\; \forall\; \underline{\psi} = [\sigma_b^2 \;\; \sigma_e^2]' \in \Xi$

$\Rightarrow \underline{R}(BB'A(A'A)^{-1}\Pi\,) \subset \underline{R}(A)$

$\Rightarrow \underline{R}(BB'A(A'A)^{-1}\Pi\,) \subset \underline{R}(A) \cap \underline{R}(B) = \underline{R}(\underline{1}_n) = \underline{R}(B\underline{1}_r)$ by assumption

$\Rightarrow \underline{R}(B'A(A'A)^{-1}\Pi\,) \subset \underline{R}(\underline{1}_r)$ since $\underline{N}(B) = \{\underline{0}\}$

$\Rightarrow \underline{R}(\{\frac{n_{ij}}{n_{i\cdot}}\}_{r\times t}\Pi\,) \subset \underline{R}(\underline{1}_r) \Rightarrow \underline{R}(\{\frac{n_{ij}}{n_{i\cdot}} - \frac{n_{tj}}{n_{t\cdot}}\}_{r\times t-1}) \subset \underline{R}(\underline{1}_r)$ by definition of Π in i) of above lemma

$\Rightarrow \frac{n_{ij}}{n_{i\cdot}} - \frac{n_{tj}}{n_{t\cdot}} = c_i \;\forall\, i,j \Rightarrow n_{ij} = n_{i\cdot}(c_i + \frac{n_{tj}}{n_{t\cdot}}) \;\; \forall\, i,j$

$\Rightarrow n_{i\cdot} = n_{i\cdot}(rc_i + \frac{n_{t\cdot}}{n_{t\cdot}}) = n_{i\cdot}(rc_i + 1) \;\; \forall\, i$ summing over $j \Rightarrow c_i = 0 \;\forall\, i\,.$

Thus, $n_{ij} = n_{i\cdot}\frac{n_{tj}}{n_{t\cdot}} = g_i h_j \;\; \forall\, i,j \;\Rightarrow N = \underline{g}\underline{h}' \Rightarrow \underline{r}(N) = 1.$ ■

Theorem: For the 2-Way Mixed Model with No Interaction, GZC holds $\Leftrightarrow \underline{r}(N) = 1.$

The proof of the theorem follows from the above two lemmas. Note that $\underline{r}(N) = 1$ if and only if there are proportional frequencies. By the Y-ELMLQE theorem, an ELMLQE exists for $\Lambda'\underline{\beta}$ and is given by $\Lambda'\widehat{\underline{\beta}}_I = \Lambda'(X'X)^{+}X'\underline{Y} = \Pi'(A'A)^{-1}A'\underline{Y} = \Pi'[\text{diag}(n_{1\cdot}, \ldots, n_{t\cdot})\,]^{-1}\{\text{diag}\,(\underline{1}_{n_i}')\}_{t\times n}\underline{Y}$

$= [I_{t-1} \;\; -\underline{1}_{t-1}] \begin{bmatrix} \overline{Y}_{1\cdot} \\ \vdots \\ \overline{Y}_{t\cdot} \end{bmatrix} = \begin{bmatrix} \overline{Y}_{1\cdot} - \overline{Y}_{t\cdot} \\ \vdots \\ \overline{Y}_{t-1\cdot} - \overline{Y}_{t\cdot} \end{bmatrix}$. The solution to the ML equations has a simple interpretation in this explicit case. However, if the MLQE was not explicit, such a simple interpretable formula would not be obtained.

5.5. REML Example: 3-Way Mixed Model under Pseudo Balance

The best design has complete balance. For the ML procedure, it is possible not to have an FELMLQE for $\underline{\psi}$ in a completely balanced design. For the REML procedure, there is an FELREMLQE for $\underline{\psi}$ in balanced designs as shown in section 4.3.3. The next type of balance to examine for the REML procedure is pseudo balance, which is defined in section 3.1.3. Under pseudo balance, there may not be an FELREMLQE for $\underline{\psi}$, but there might be an ELREMLQE for a linear combination of interest. This section will describe such examples.

For some 3-way mixed models with two-level factors, there exists an ELREMLQE for $\underline{1}'\underline{\psi} = \sum_{i=1}^{k+1} \psi_i$, or the variance of a single observation, under pseudo balance. An example of an incidence matrix with pseudo balance in which $\underline{\psi}$ is estimable under the REML procedure is given by:
$\{n_{ij1}\} = \begin{bmatrix} r & 0 \\ r & 0 \end{bmatrix}$ $\{n_{ij2}\} = \begin{bmatrix} r & r \\ 0 & r \end{bmatrix}$ where $r > 1$. A search was conducted over all proper two-level models with an incidence matrix of the above form in which there were at least 2 variance components. Models that had an ELREMLQE for $\underline{1}'\underline{\psi}$ are listed in Table 5.1 using the notation:
-1 = fixed effect, 0 = omitted, and 1 = random effect. These models also have an ELREMLQE for σ_e^2, but not for the other components individually.

Table 5.1. 3-Way Models with ELREMLQE for Sum of Variance Components and Residual

a	b	c	ab	ac	bc	abc	e
-1	0	0	1	1	0	1	1
-1	1	1	0	1	1	1	1
-1	1	1	1	0	1	1	1
-1	1	1	1	1	0	1	1
-1	1	1	1	1	1	0	1
0	1	1	1	0	1	1	1
0	1	1	0	1	1	1	1
1	1	1	1	1	1	1	1

6. UBLUE for the Full Rank Case

This chapter discusses a generalization of the results of Szatrowski (1980) and ElBassiouni (1983) under the assumption of a full rank model or one whose parameter vector is estimable. The results of the chapter are a special case of those in the previous chapter. The previous chapter examined the existence of a UBLUE for a linear combination of the parameter, while this chapter examines the existence of a UBLUE for a subvector of the parameter vector. This formulation provides another way to think about the problem which is convenient to work with. The results will first be presented for the underlying model and then applied to the particular models of interest for examining a subvector of the fixed effects vector and a subvector of the variance components vector.

6.1 UBLUE for the Full Rank Underlying Model

The conditions for a UBLUE in this model could be derived using the results in chapter 5. However, it is informative to construct the conditions in a different manner. Consider the U-Model under [S], where $\Upsilon_{\mathcal{U}}$ is a subset of $\mathcal{R}^p$ such that $\mathrm{sp}\Upsilon_{\mathcal{U}} = \mathcal{R}^p$, $U \in \mathcal{L}(\mathcal{R}^p, \mathcal{W})$, and $\underline{\theta} \in \Upsilon_{\mathcal{U}}$. It is assumed that $\underline{\theta}$ is estimable. By the Full Rank theorem in section 3.2.2 assuming [S], $\underline{\theta}$ is estimable if and only if $\underline{r}(U) = p$. In this case, $\underline{N}(U) = \{0\}$. Let $\underline{e}_i = \{e_{ij}\} \in \mathcal{R}^p$ where $e_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$ and suppose $b_i = U\underline{e}_i$. Then $U\underline{\theta} = U\sum_{i=1}^{p}\theta_i\underline{e}_i = \sum_{i=1}^{p}\theta_i b_i$.

Consider a partition of the parameter vector where $\underline{\theta}_{p\times 1} = [\underline{\theta}'_{1p_1\times 1}\ \underline{\theta}'_{2p_2\times 1}]'$. Suppose interest is in estimating $\underline{\theta}_2$. Consider the following notation, which partitions the U-Model accordingly:

i) $\mathrm{E}[\omega] = U\underline{\theta} = \sum_{i=1}^{p}\theta_i b_i = \sum_{i=1}^{p_1}\theta_i b_i + \sum_{i=p_1+1}^{p}\theta_i b_i = U_1\underline{\theta}_1 + U_2\underline{\theta}_2$ for $U_j \in \mathcal{L}(\mathcal{R}^{p_j}, \mathcal{W})$

ii) $V \in \mathcal{V} \subset \mathcal{L}_{\mathrm{PD}}(\mathcal{W}, \mathcal{W})$

iii) $F_{1V} : \mathcal{W} \to \mathcal{W}$ $\qquad F_{1V} = V^{-1} - V^{-1}U_1(U_1^*V^{-1}U_1)^{-1}U_1^*V^{-1}$

iv) $N_{U_1} : \mathcal{W} \to \mathcal{W}$ $\qquad N_{U_1} = I - U_1(U_1^*U_1)^{-1}U_1^*$

v) Define $Q_1 \ni N_{U_1} = Q_1Q_1^*$ and $Q_1^*Q_1 = I_{p_1}$.

The notation in i) provides a partition of the expectation, while the covariance in ii) remains unchanged. The notation in iii), iv), and v) will be used in the next four lemmas to obtain a UBLUE condition for $\underline{\theta}_2$.

Lemma 1: i) $F_{1V}U_1 = 0$ $\qquad$ ii) $F_{1V} = Q_1(Q_1^*VQ_1)^{-1}Q_1^*$

iii) VF_{1V} is a PO on $\underline{N}(U_1^*V^{-1})$ along $\underline{R}(U_1)$

iv) $F_{1V}V$ is a PO on $\underline{N}(U_1^*)$ along $\underline{R}(V^{-1}U_1)$.

proof: i) From the definition of F_{1V}. ii) Follows from the $F_{\underline{V}}$-lemma in section 3.1.4. iii) iv) From the general projection theorem in section 2.5. ∎

Lemma 2: If $M_V = (U^*V^{-1}U)^{-1}U^*V^{-1} \in \mathcal{L}(\mathcal{W},\mathcal{R}^p)$ is partitioned as $M_V = \begin{bmatrix} M_{1V} \\ M_{2V} \end{bmatrix}$ where $M_{1V} \in \mathcal{L}(\mathcal{W},\mathcal{R}^{p_1})$ and $M_{2V} \in \mathcal{L}(\mathcal{W},\mathcal{R}^{p_2})$, then $M_{2V} = (U_2^*F_{1V}U_2)^{-1}U_2^*F_{1V}$ and $M_{1V} = (U_1^*V^{-1}U_1)^{-1}U_1^*V^{-1}(I - U_2M_{2V})$.

proof: (1) Note $< U\underline{\theta},\omega >_{\mathcal{W}} = \; < U_1\underline{\theta}_1 + U_2\underline{\theta}_2,\omega >_{\mathcal{W}} = \; < U_1\underline{\theta}_1,\omega >_{\mathcal{W}} + < U_2\underline{\theta}_2,\omega >_{\mathcal{W}}$

$$= \; < \underline{\theta}_1, U_1^*\omega >_{\mathcal{R}^{p_1}} + < \underline{\theta}_2, U_2^*\omega >_{\mathcal{R}^{p_2}} = \underline{\theta}_1'U_1^*\omega + \underline{\theta}_2'U_2^*\omega$$

$$= \begin{bmatrix} \underline{\theta}_1 \\ \underline{\theta}_2 \end{bmatrix}' \begin{bmatrix} U_1^*\omega \\ U_2^*\omega \end{bmatrix} = \; < \begin{bmatrix} \underline{\theta}_1 \\ \underline{\theta}_2 \end{bmatrix}, \begin{bmatrix} U_1^*\omega \\ U_2^*\omega \end{bmatrix} >_{\mathcal{R}^p} \Rightarrow U^* = \begin{bmatrix} U_1^* \\ U_2^* \end{bmatrix}.$$

(2) $(U^*V^{-1}U)^{-1}$ is invertible, so the inverse formula in section 2.4 gives

$$M_V = (U^*V^{-1}U)^{-1}U^*V^{-1} = (\begin{bmatrix} U_1^* \\ U_2^* \end{bmatrix} V^{-1}[U_1 \; U_2])^{-1}\begin{bmatrix} U_1^* \\ U_2^* \end{bmatrix} V^{-1} \quad \text{by (1)}$$

$$= \begin{bmatrix} U_1^*V^{-1}U_1 & U_1^*V^{-1}U_2 \\ U_2^*V^{-1}U_1 & U_2^*V^{-1}U_2 \end{bmatrix}^{-1} \begin{bmatrix} U_1^*V^{-1} \\ U_2^*V^{-1} \end{bmatrix} = \begin{bmatrix} A & B \\ B' & D \end{bmatrix}^{-1} \begin{bmatrix} U_1^*V^{-1} \\ U_2^*V^{-1} \end{bmatrix}$$

$$= \begin{bmatrix} A^{-1} + GE^{-1}G^* & -GE^{-1} \\ -E^{-1}G^* & E^{-1} \end{bmatrix} \begin{bmatrix} U_1^*V^{-1} \\ U_2^*V^{-1} \end{bmatrix}$$

where $E = D - B'A^{-1}B = U_2^*F_{1V}U_2$ and $G = A^{-1}B = (U_1^*V^{-1}U_1)^{-1}U_1^*V^{-1}U_2$. Thus,

$M_{2V} = -E^{-1}G^*U_1^*V^{-1} + E^{-1}U_2^*V^{-1} = (U_2^*F_{1V}U_2)^{-1}U_2^*F_{1V}$ and

$$M_{1V} = (A^{-1} + GE^{-1}G^*)U_1^*V^{-1} - GE^{-1}U_2^*V^{-1} = A^{-1}U_1^*V^{-1} + GE^{-1}(G^*U_1^*V^{-1} - U_2^*V^{-1})$$

$$= (U_1^*V^{-1}U_1)^{-1}U_1^*V^{-1} - GM_{2V} = (U_1^*V^{-1}U_1)^{-1}U_1^*V^{-1}(I - U_2M_{2V}). \quad ∎$$

Lemma 3: The following are equivalent:

i) $M_{2V} = (U_2^*F_{1V}U_2)^{-1}U_2^*F_{1V} = (U_2^*F_{1V_0}U_2)^{-1}U_2^*F_{1V_0} = M_{2V_0}$

ii) $\underline{R}(F_{1V_0}U_2) \subset \underline{R}(F_{1V}U_2)$

iii) $\underline{R}(N_{U_1}VF_{1V_0}U_2) \subset \underline{R}(N_{U_1}U_2)$

iv) $\underline{R}(VF_{1V_0}U_2) \subset \underline{R}(U)$.

proof: (1) By the general projection theorem in section 2.5, U_2M_{2V} is a PO on $\underline{R}(U_2)$ along $\underline{N}(U_2^*F_{1V})$. Since $\underline{N}(U_2) = \{\underline{0}\}$, i) $\Leftrightarrow U_2M_{2V} = U_2M_{2V_0} \Leftrightarrow \underline{N}(U_2^*F_{1V}) = \underline{N}(U_2^*F_{1V_0}) \Leftrightarrow \underline{R}(F_{1V}U_2) = \underline{R}(F_{1V_0}U_2)$ $\Leftrightarrow \underline{R}(F_{1V_0}U_2) \subset \underline{R}(F_{1V}U_2) \Leftrightarrow$ ii). The second to last equivalence is true $\forall$ PD V because

$\underline{r}(F_{1V}U_2) = \underline{r}(U_2) - \dim[\underline{R}(U_2) \cap \underline{N}(F_{1V})] = \underline{r}(U_2) - \dim[\underline{R}(U_2) \cap \underline{N}(VF_{1V})]$

$= \underline{r}(U_2) - \dim[\underline{R}(U_2) \cap \underline{R}(U_1)]$ by previous lemma 1 iii)

$= \underline{r}(U_2)$ as U has full column rank.

(2) ii) $\Leftrightarrow \underline{R}(F_{1V_0}U_2) \subset \underline{R}(F_{1V}U_2) \Leftrightarrow F_{1V_0}U_2 = F_{1V}U_2B$ for some B

$\Leftrightarrow Q_1(Q_1^*V_0Q_1)^{-1}Q_1^*U_2 = Q_1(Q_1^*VQ_1)^{-1}Q_1^*U_2B$ by lemma 1 ii)

$\Leftrightarrow (Q_1Q_1^*V)Q_1(Q_1^*V_0Q_1)^{-1}Q_1^*U_2 = Q_1Q_1^*U_2B$ left multiplying by $Q_1Q_1^*V$ or $Q_1(Q_1^*V_0Q_1)^{-1}Q_1^*$

$\Leftrightarrow N_{U_1}VF_{1V_0}U_2 = N_{U_1}U_2B$ by lemma 1 $N_{U_1} = Q_1Q_1^*$

$\Leftrightarrow \underline{R}(N_{U_1}VF_{1V_0}U_2) \subset \underline{R}(N_{U_1}U_2) \Leftrightarrow$ iii).

(3) iv) $\Rightarrow \underline{R}(VF_{1V_0}U_2) \subset \underline{R}(U) \Rightarrow N_{U_1}[\underline{R}(VF_{1V_0}U_2)] \subset N_{U_1}[\underline{R}(U)]$

$\Rightarrow \underline{R}(N_{U_1}VF_{1V_0}U_2) \subset \underline{R}(N_{U_1}U_2) \Rightarrow$ iii) as $\underline{R}(N_{U_1}U) = \underline{R}(N_{U_1}[U_1\ U_2]) = \underline{R}([0\ N_{U_1}U_2])$.

(4) iii) $\Rightarrow \underline{R}(N_{U_1}VF_{1V_0}U_2) \subset \underline{R}(N_{U_1}U_2) \Rightarrow N_{U_1}VF_{1V_0}U_2 = N_{U_1}U_2B$ for some B

$\Rightarrow N_{U_1}(VF_{1V_0}U_2 - U_2B) = 0 \Rightarrow \underline{R}(VF_{1V_0}U_2 - U_2B) \subset \underline{N}(N_{U_1}) = \underline{R}(U_1) \subset \underline{R}(U)$

$\Rightarrow \underline{R}(VF_{1V_0}U_2) \subset \underline{R}(U) \Rightarrow$ iv). ▪

Lemma 4: $\widehat{\underline{\theta}}_{2V_0} = (U_2^*F_{1V_0}U_2)^{-1}U_2^*F_{1V_0}\omega$ is UBLUE for $\underline{\theta}_2 \Leftrightarrow \underline{R}(VF_{1V_0}U_2) \subset \underline{R}(U)\ \forall\ V \in \mathcal{V}$.

proof: By lemma 1, $\mathrm{E}[\widehat{\underline{\theta}}_{2V_0}] = (U_2^*F_{1V_0}U_2)^{-1}U_2^*F_{1V_0}[U_1\underline{\theta}_1 + U_2\underline{\theta}_2] = (U_2^*F_{1V_0}U_2)^{-1}U_2^*F_{1V_0}U_2\underline{\theta}_2 = \underline{\theta}_2$. By Zyskind's theorem, $\widehat{\underline{\theta}}_{2V_0}$ is UBLUE for $\underline{\theta}_2 \Leftrightarrow \underline{R}(VF_{1V_0}U_2(U_2^*F_{1V_0}U_2)^{-1}) \subset \underline{R}(U)\ \ \forall\ V \in \mathcal{V}$

$\Leftrightarrow \underline{R}(VF_{1V_0}U) \subset \underline{R}(U)\ \ \forall\ V \in \mathcal{V}$. ▪

For purposes of discussing the least squares estimator, interest is in $V_0 = I$. Lemma 2 partitions the GLSE in the U-Model to obtain an expression for the GLSE for $\underline{\theta}_1$ and the GLSE for $\underline{\theta}_2$. Condition iv) of lemma 3, can be used to indicate when the GLSE for $\underline{\theta}_2$ and the LSE for $\underline{\theta}_2$ are equal. The least squares, generalized least squares, and estimated generalized least squares estimators for $\underline{\theta}_2$ are listed below:

LSE: $\widehat{\underline{\theta}}_{2I} = M_{2I}\omega = (U_2^*N_{U_1}U_2)^{-1}U_2^*N_{U_1}\omega$

GLSE: $\widehat{\underline{\theta}}_{2V} = M_{2V}\omega = (U_2^*F_{1V}U_2)^{-1}U_2^*F_{1V}\omega$ for a given $V \in \mathcal{L}_{\mathrm{PD}}(\mathcal{W},\mathcal{W})$

EGLSE: $\widehat{\underline{\theta}}_{2\widehat{V}} = M_{2\widehat{V}}\omega = (U_2^*F_{1\widehat{V}}U_2)^{-1}U_2^*F_{1\widehat{V}}\omega$ with $\widehat{V} = \widehat{V}(\omega) \in \mathcal{L}_{\mathrm{PD}}(\mathcal{W},\mathcal{W})\ \forall\ \omega \in \mathcal{W}$.

The LSE has a simple linear form which does not depend on the covariance, the GLSE depends on the unknown true V, and the EGLSE is not linear and requires an estimate for the unknown variance. Lemma 3 can be used to indicate when the LSE is equal to the GLSE. The following UBLUE theorem is a special case of the U-UBLUE theorem in the full rank (FR) setting.

U-UBLUE$_{FR}$ Theorem: The following are equivalent under [S]:

i) $\widehat{\underline{\theta}}_{2I}$ is UBLUE for $\underline{\theta}_2$ in the U-Model

ii) $\underline{R}(VN_{U_1}U_2) \subset \underline{R}(U)\ \ \forall\ V \in \mathcal{V}$

iii) $(U_2^*F_{1V}U_2)^{-1}U_2^*F_{1V} = (U_2^*N_{U_1}U_2)^{-1}U_2^*N_{U_1}\ \ \forall\ V \in \mathcal{V}$.

proof: ii) $\Leftrightarrow$ iii) from lemma 3 and i) $\Leftrightarrow$ ii) by lemma 4. ▪

Proposition 1: If $V \in \mathrm{sp}\mathcal{V}$ and V is PD, then the conditions in the U-UBLUE$_{\mathrm{FR}}$ theorem imply $\widehat{\underline{\theta}}_{2V} = \widehat{\underline{\theta}}_{2I}$.

proof: $\underline{R}(VN_{U_1}U_2) \subset \underline{R}(U)\ \forall\ V \in \mathcal{V} \Rightarrow \underline{R}(VN_{U_1}U_2) \subset \underline{R}(U)\ \forall\ V \in \mathrm{sp}\mathcal{V}$ by proposition in section 2.3 $\Rightarrow M_{2V} = M_{2I}\ \forall\ V \in \mathrm{sp}\mathcal{V} \ni V$ is PD by the U-UBLUE$_{\mathrm{FR}}$ theorem. ∎

Proposition 2: If $I \in \mathrm{sp}\mathcal{V}$, then $\underline{\theta}_2$ has a UBLUE $\Leftrightarrow \widehat{\underline{\theta}}_{2I}$ is UBLUE for $\underline{\theta}_2$.

proof: i) $\widehat{\underline{\theta}}_{2I}$ is UBLUE for $\underline{\theta}_2 \Rightarrow \underline{\theta}_2$ has a UBLUE.
ii) Suppose $\underline{\theta}_2$ has a UBLUE given by $\widehat{\underline{\theta}}_{2V_0}$ where $V_0 \in \mathcal{V}$
$\Rightarrow \underline{R}(VF_{1V_0}U_2) \subset \underline{R}(U)\ \forall\ V \in \mathcal{V}$ by lemma 4 $\Rightarrow \underline{R}(VF_{1V_0}U_2) \subset \underline{R}(U)\ \forall\ V \in \mathrm{sp}\mathcal{V}$
$\Rightarrow \underline{R}(F_{1V_0}U_2) \subset \underline{R}(F_{1V}U_2)\ \forall\ V \in \mathrm{sp}\mathcal{V}$ by lemma 3 $\Rightarrow \underline{R}(F_{1V_0}U_2) \subset \underline{R}(N_{U_1}U_2)$ as $I \in \mathrm{sp}\mathcal{V}$
$\Rightarrow M_{2V_0} = M_{2I}$ by lemma 2 $\Rightarrow \widehat{\underline{\theta}}_{2V_0} = \widehat{\underline{\theta}}_{2I} \Rightarrow \widehat{\underline{\theta}}_{2I}$ is UBLUE for $\underline{\theta}_2$ by the Uniqueness theorem. ∎

Condition i) of the U-UBLUE$_{\mathrm{FR}}$ theorem indicates the UBLUE has an expression which does not depend on V. Condition ii) gives the GZC for $\underline{\theta}_2$ in this full rank setting, which will be denoted GZC$_{\mathrm{FR}}$. The conditions of the U-UBLUE$_{\mathrm{FR}}$ theorem can be re-expressed in a convenient form. Define the set $\mathcal{L}_{U_2}(\mathcal{W},\mathcal{W}) = \{V \in \mathcal{L}_{\mathrm{PD}}(\mathcal{W},\mathcal{W}) |\ \underline{R}(VN_{U_1}U_2) \subset \underline{R}(U)\}$. The first three corollaries restate the previous results, while the last one applies the results to the EGLSE.

Corollary: GZC$_{\mathrm{FR}} \Leftrightarrow \mathcal{V} \subset \mathcal{L}_{U_2}(\mathcal{W},\mathcal{W})$.

Corollary: If $\mathcal{V} \subset \mathcal{L}_{U_2}(\mathcal{W},\mathcal{W})$, then GLSE = LSE = UBLUE.

Corollary: If $\mathcal{V} \subset \mathcal{L}_{U_2}(\mathcal{W},\mathcal{W})$, then $\mathrm{sp}\mathcal{V} \cap \mathcal{L}_{\mathrm{PD}}(\mathcal{W},\mathcal{W}) \subset \mathcal{L}_{U_2}(\mathcal{W},\mathcal{W})$.

Corollary: If $\underline{\theta}_2$ has a UBLUE and $\widehat{V} \in \mathcal{L}_{U_2}(\mathcal{W},\mathcal{W})$, then EGLSE = GLSE = LSE = UBLUE.

Assume $V = V_{\underline{\theta}}$ and let $\widehat{V} = V_{\widehat{\underline{\theta}}}$. Suppose the GZC$_{\mathrm{FR}}$ holds and $\widehat{V} \in \mathcal{L}_{U_2}(\mathcal{W},\mathcal{W})$. An iterative procedure would still be required to solve for $\widehat{\underline{\theta}}_{1\widehat{V}}$. The $(i+1)^{th}$ solution in the iterative procedure is given by $\widehat{\underline{\theta}}_{1\widehat{V}}^{(i+1)} = (U_1^* V^{-1}_{\widehat{\underline{\theta}}^{(i)}_{1\widehat{V}}, \widehat{\underline{\theta}}_{2I}} U_1)^{-1} U_1^* V^{-1}_{\widehat{\underline{\theta}}^{(i)}_{1\widehat{V}}, \widehat{\underline{\theta}}_{2I}} (\omega - U_2 \widehat{\underline{\theta}}_{2I})$.

The next theorem shows the relationship between ZC and GZC$_{\mathrm{FR}}$. The results assuming GZC$_{\mathrm{FR}}$ are more general than the results assuming ZC, as ZC implies GZC$_{\mathrm{FR}}$.

Theorem: i) If $U_1 = 0$, then GZC$_{\text{FR}} \Rightarrow$ ZC

ii) If $\underline{R}(VU_1) \subset \underline{R}(U_1)\ \forall\ V \in \mathcal{V}$, then GZC$_{\text{FR}} \Rightarrow$ ZC

iii) ZC $\Rightarrow$ GZC$_{\text{FR}}$.

proof: Note $\underline{R}(N_{U_1}U) = \underline{R}(N_{U_1}[U_1\ U_2]) = \underline{R}([\,0\ N_{U_1}U_2\,])$.

i) Since $U_1 = 0$ by hypothesis, $N_{U_1} = I$. Thus, $\forall\ V \in \mathcal{V}$,

$\underline{R}(VN_{U_1}U_2) \subset \underline{R}(U) \Rightarrow \underline{R}(VN_{U_1}U) \subset \underline{R}(U) \Rightarrow \underline{R}(VU) \subset \underline{R}(U)$.

ii) $\forall\ V \in \mathcal{V}$, $\underline{R}(VN_{U_1}U_2) \subset \underline{R}(U) \Rightarrow \underline{R}(VN_{U_1}U) \subset \underline{R}(U) \Rightarrow VN_{U_1}U = UB$ for some B

$\Rightarrow VU = VP_{U_1}U + UB$

$\Rightarrow \underline{R}(VU) \subset \underline{R}(VP_{U_1}U + UB) \subset \underline{R}(VU_1) + \underline{R}(U) \subset \underline{R}(U_1) + \underline{R}(U) = \underline{R}(U)$ as $\underline{R}(VU_1) \subset \underline{R}(U_1)$.

iii) Note $N_{U_1}U_2 = U_2 - P_{U_1}U_2 \Rightarrow \underline{R}(N_{U_1}U_2) \subset \underline{R}(U)$. Then $\forall\ V \in \mathcal{V}$,

$\underline{R}(VN_{U_1}U_2) \subset \underline{R}(VU) \subset \underline{R}(U)$ under the assumption of ZC. ∎

6.2. The Generalized Zyskind's Conditions

The section explores the relationship between GZC and GZC$_{\text{FR}}$. First, consider the full rank setting to examine how the GZC$_{\text{FR}}$ can be used to obtain the GZC. Note $\underline{\theta}_2 = \Gamma'\underline{\theta}$ where $\Gamma' = [0_{p_2\times p_1}\ I_{p_2\times p_2}]$. The following theorem gives the equivalence in this case.

Theorem 1: Let the U-Model under [S] have full rank. Then

$\underline{R}(VN_{U_1}U_2) \subset \underline{R}(U) \Leftrightarrow \underline{R}(VU(U^*U)^{-1}\Gamma) \subset \underline{R}(U)$.

proof: i) $\underline{R}(VU(U^*U)^{-1}\Gamma) \subset \underline{R}(U)$

$\Leftrightarrow \Gamma'(U^*V^{-1}U)^{-1}U^*V^{-1} = \Gamma'(U^*U)^{-1}U^*$ by U-UBLUE theorem

$\Leftrightarrow [0\ I\,]M_V = [0\ I\,]M_I \Leftrightarrow M_{2V} = M_{2I}$ from lemma 2 and definition of Γ

$\Leftrightarrow \underline{R}(VN_{U_1}U_2) \subset \underline{R}(U)$ by U-UBLUE$_{\text{FR}}$ theorem. ∎

Assume the GZC involving an estimable function $\Pi'\underline{\theta}$ where $\Pi : \mathcal{R}^q \to \mathcal{R}^p$ and $\underline{r}(\Pi) = q$ and consider translating it to the GZC$_{\text{FR}}$ for a corresponding parameter $\xi_{2(q\times 1)}$. This direction is useful for checking the GZC condition in the full rank setting.

Since $\underline{r}(\Pi) = q$ there exists a matrix $\Delta_{p\times p-q} \ni K_{p\times p} = [\Delta\ \Pi]$ is invertible. Thus, $K'\underline{\theta} = \xi$ $\Rightarrow \underline{\theta} = (K')^{-1}\xi$ where $(K')^{-1} = \{h_{ij}\}$ and so $\theta_i = \sum_{j=1}^{p} h_{ij}\xi_j$. These definitions give $U\underline{\theta} = \sum_{i=1}^{p}\theta_i b_i = \sum_{i=1}^{p}\sum_{j=1}^{p} h_{ij}\xi_j b_i = \sum_{j=1}^{p}\xi_j(\sum_{i=1}^{p} h_{ij}b_i) = \sum_{j=1}^{p}\xi_j c_j = T\xi$ where $T \in \mathcal{L}(\mathcal{R}^p, \mathcal{W})$. Consider the following notation for estimating $\Pi'\underline{\theta} = \xi_2 = \Gamma'\xi$ where $\Gamma' = [0_{q\times(p-q)}\ I_{q\times q}]$:

i) $\mathrm{E}[\omega] = U\underline{\theta} = T\underline{\xi} = T_1\underline{\xi}_1 + T_2\underline{\xi}_2 \qquad T_j \in \mathcal{L}(\mathcal{R}^{v_j}, \mathcal{W})$ with $v_1 = p - q,\ v_2 = q$

ii) $N_{T_1} : \mathcal{W} \to \mathcal{W} \qquad N_{T_1} = I - T_1(T_1^* T_1)^{-1} T_1^*$.

Theorem 2: For the estimable function $\Pi'\underline{\theta} = \underline{\xi}_2 = \Gamma'\underline{\xi}$ and using the transformations defined above, $\underline{R}(VU(U^*U)^{-1}\Pi) \subset \underline{R}(U) \Leftrightarrow \underline{R}(VN_{T_1}T_2) \subset \underline{R}(T)$.

proof: $\underline{R}(VN_{T_1}T_2) \subset \underline{R}(T) \Leftrightarrow \underline{R}(VT(T^*T)^{-1}\Gamma) \subset \underline{R}(T) \Leftrightarrow \underline{R}(VU(U^*U)^{-1}\Pi) \subset \underline{R}(U)$
by theorem 1 as the GZC is invariant under a reparameterizaton from the theorem in section 5.1. ∎

6.3. UBLUE for Estimable Fixed Effects

This section examines UBLUE results for the subvector of the fixed effects given by $\underline{\beta}_2$ in the Y-model that has full rank or where $\underline{r}(X_{n\times p}) = p$. Recall $\underline{\beta} \in \mathcal{R}^p$ and $\mathcal{R}^p$ contains a non-empty open set, so the Full Rank theorem can be applied. First, results will be presented for UBLUEs, which will then be applied to maximum likelihood estimation for $\underline{\beta}_2$.

6.3.1. UBLUE Results

This section examines conditions under which a UBLUE exists in the Y-Model for $\underline{\beta}_2$. Consider the following definitions:

i) $\mathrm{E}[\underline{Y}] = X\underline{\beta} = \sum_{i=1}^{p} \underline{x}_i\beta_i = \sum_{i=1}^{p_1} \underline{x}_i\beta_i + \sum_{i=p_1+1}^{p} \underline{x}_i\beta_i = X_1\underline{\beta}_1 + X_2\underline{\beta}_2$

ii) $F_{1V_{\underline{\psi}}} = V_{\underline{\psi}}^{-1} - V_{\underline{\psi}}^{-1}X_1(X_1^*V_{\underline{\psi}}^{-1}X_1)^{-1}X_1^*V_{\underline{\psi}}^{-1} \qquad$ iii) $N_{X_1} = I - P_{X_1}$.

The definition in i) shows the partition for the expectation, while ii) and iii) define matrices that will be of interest. The least squares estimator, the generalized least squares estimator, and the estimated generalized least squares estimator for $\underline{\beta}_2$ will also be of interest in this section. They are listed below:

LSE: $\quad \widehat{\underline{\beta}}_{2I} = (X_2'N_{X_1}X_2)^{-1}X_2'N_{X_1}\underline{Y}$

GLSE: $\widehat{\underline{\beta}}_{2\underline{\psi}} = (X_2^*F_{1V_{\underline{\psi}}}X_2)^{-1}X_2^*F_{1V_{\underline{\psi}}}\underline{Y} \qquad$ for a given $\underline{\psi}$ where $V_{\underline{\psi}}$ is PD

EGLSE: $\widehat{\underline{\beta}}_{2\widehat{\underline{\psi}}} = (X_2^*F_{1V_{\widehat{\underline{\psi}}}}X_2)^{-1}X_2^*F_{1V_{\widehat{\underline{\psi}}}}\underline{Y} \qquad$ where $V_{\widehat{\underline{\psi}}} = V_{\widehat{\underline{\psi}}(\underline{Y})}$ is PD $\forall\ \underline{Y} \in \mathcal{R}^n$.

The UBLUE result for the U-Model in the full rank setting can be directly applied in this setting to indicate when the LSE is equal to the GLSE for estimating $\underline{\beta}_2$. This is stated in the next theorem and restated in the following proposition under assumptions [L] and [O].

Y-UBLUE$_{FR}$ Theorem: The following are equivalent:

i) $\widehat{\underline{\beta}}_{2I}$ is UBLUE for $\underline{\beta}_2$ in the Y-Model

ii) $\underline{R}(V_{\underline{\psi}} N_{X_1} X_2) \subset \underline{R}(X)\ \ \forall\ \underline{\psi} \in \Xi$

iii) $(X_2^* F_{1V_{\underline{\psi}}} X_2)^{-1} X_2^* F_{1V_{\underline{\psi}}} = (X_2^* N_{X_1} X_2)^{-1} X_2^* N_{X_1}\ \ \forall\ \underline{\psi} \in \Xi.$

Proposition: Assume [L] and [O]. Then $\forall\ \underline{\psi} \in \mathcal{R}^{k+1} \ni V_{\underline{\psi}}$ is PD, the conditions in the Y-FUBLUE$_{\text{FR}}$ theorem imply that $\widehat{\underline{\beta}}_{2\underline{\psi}} = \widehat{\underline{\beta}}_{2I}$.

proof: Under [L], [O], sp$\mathcal{V} = \{V_{\underline{\psi}} |\ \underline{\psi} \in \text{sp}\ \Xi\} = \{V_{\underline{\psi}} | \underline{\psi} \in \mathcal{R}^{k+1}\}$. Apply proposition 1 in section 6.1. ■

The Y-UBLUE theorem does not make any assumptions about normality, linear covariance structure, or classfication matrices. Only the form of GZC$_{\text{FR}}$ given in ii) is necessary to apply this theorem to the special cases of the Y-Model under [L], [O], and [C].

6.3.2. ELMLQE Results

For doing maximum likelihood estimation in the Y-Model, the assumptions [L], [O], and [N] were used. Extending the definition of Szatrowski (1980), an explicit linear maximum likelihood equation estimator (ELMLQE) for $\underline{\beta}_2$ satisfies $\widehat{\underline{\beta}}_{2\text{MLQ}} = A\underline{Y}$ for a constant matrix A which is not random and does not depend on any parameters. In this case, the estimator $\widehat{\underline{\beta}}_{2\text{MLQ}}$ is linear and the explicit part indicates that A is constant. The following lemma shows that the MLQE = EGLSE in this setting.

Lemma: $\widehat{\underline{\beta}}_{2\text{MLQ}} = \widehat{\underline{\beta}}_{2\widehat{\underline{\psi}}_{\text{MLQ}}}$.

proof: It has been shown that $X\widehat{\underline{\beta}}_{\text{MLQ}} = X\widehat{\underline{\beta}}_{\widehat{\underline{\psi}}_{\text{MLQ}}}$

$\Rightarrow \widehat{\underline{\beta}}_{\text{MLQ}} = \widehat{\underline{\beta}}_{\widehat{\underline{\psi}}_{\text{MLQ}}}$ multiplying both sides by $(X'X)^{-1}X$ as X has full column rank where $\widehat{\underline{\beta}}_{\widehat{\underline{\psi}}_{\text{MLQ}}}$ can be partitioned as in lemma 2. ■

The next theorem presents a sufficient condition for the existence of an ELMLQE for $\underline{\beta}_2$. This condition is the GZC$_{\text{FR}}$ for the Y-Model.

Y-ELMLQE$_{FR}$ Theorem: Consider the Y-Model under [L], [O], and [N]. If $\widehat{\underline{\psi}}_{\text{MLQ}}$ exists $\ni V_{\widehat{\underline{\psi}}_{\text{MLQ}}}$ is PD and $\underline{R}(V_{\underline{\psi}} N_{X_1} X_2) \subset \underline{R}(X)\ \forall\ \underline{\psi} \in \Xi$, then $\widehat{\underline{\beta}}_{2I}$ is an ELMLQE for $\underline{\beta}_2$.

proof: Since $V_{\hat{\psi}_{\text{MLQ}}}$ is PD by hypothesis, $\hat{\beta}_{2\hat{\psi}_{\text{MLQ}}} = \hat{\beta}_{2I}$ by the proposition in section 6.3.1.
By the above lemma, $\hat{\beta}_{2\text{MLQ}} = \hat{\beta}_{2\hat{\psi}_{\text{MLQ}}}$. Thus, $\hat{\beta}_{2\text{MLQ}} = \hat{\beta}_{2I} = A\underline{Y}$ where $A = (X_2'N_{X_1}X_2)^{-1}X_2'N_{X_1}$
$\Rightarrow \hat{\beta}_{2I}$ is an ELMLQE for β_2 by definition. ∎

The proof of the theorem shows the MLQE is equivalent to the LSE under the assumptions. Because the LSE is linear and explicit, there exists an ELMLQE for β_2.

6.4. UBLUE for Estimable Variance Components

UBLUEs for the estimable variance component vector can be obtained using linearized quadratic estimation models and applying the results from the U-Model. The UBLUE results can be used to determine the existence of a linear estimator for a subvector $\underline{\psi}_2$ of the variance component vector $\underline{\psi} \in \Xi$. In order to apply the results of the Full Rank theorem, it is sufficient to assume Ξ contains a non-empty open set in $\mathcal{R}^r$. This assumption does hold for the ALQEM for $(\underline{Y} - X\hat{\beta})$ and the LQEM for $N_X\underline{Y}$. The UBLUE results will be presented for the LQEM for $\underline{Z}$, and then applied to the ALQEM for $(\underline{Y} - X\hat{\beta})$ for maximum likelihood estimation and the LQEM for $N_X\underline{Y}$ for restricted maximum likelihood estimation.

6.4.1. UBLUE Results

The UBLUE results will be given for the LQEM for $\underline{Z}$ under the open set assumption for Ξ. Consider the following notation where $r_2 = r - r_1$:

i) $\text{E}[Y^\dagger] = X^\dagger\underline{\psi} = R_{\underline{\psi}} = \sum_{i=1}^{r}\psi_i R_i = \sum_{i=1}^{r_1}\psi_i R_i + \sum_{i=r_1+1}^{r}\psi_i R_i = X_1^\dagger\underline{\psi}_1 + X_2^\dagger\underline{\psi}_2$ for $X_j^\dagger \in \mathcal{L}(\mathcal{R}^{r_j}, \mathcal{S}_n)$

ii) $V_{\underline{\psi}}^\dagger = 2\Psi_{R_{\underline{\psi}}} \in \mathcal{V}^\dagger \subset \mathcal{L}_{\text{PD}}(\mathcal{S}_n, \mathcal{S}_n)$

ii) $F_{1V_{\underline{\psi}}^\dagger} : \mathcal{S}_n \rightarrow \mathcal{S}_n \qquad F_{1V_{\underline{\psi}}^\dagger} = V_{\underline{\psi}}^{\dagger-1} - V_{\underline{\psi}}^{\dagger-1}X_1^\dagger(X_1^{\dagger *}V_{\underline{\psi}}^{\dagger-1}X_1^\dagger)^{-1}X_1^{\dagger *}V_{\underline{\psi}}^{\dagger-1}$

iii) $N_{X_1^\dagger} : \mathcal{S}_n \rightarrow \mathcal{S}_n \qquad N_{X_1^\dagger} = I - X_1^\dagger(X_1^{\dagger *}X_1^\dagger)^{-1}X_1^{\dagger *} = I - P_{X_1^\dagger}$.

The definition in i) gives a partition of the expectation and an expression for the covariance. The notation in ii) and iii) will be useful for the following results. The objective is to estimate $\underline{\psi}_2$ in the full rank model where $\underline{r}(X^\dagger) = r$, or equivalently when the R_i's $i = 1, ..., r$ are linearly independent. The associated least squares, generalized least squares, and estimated generalized least squares estimator for $\underline{\psi}_2$ will be of interest. These are listed here for reference:

LSE:	$\widehat{\underline{\psi}}_{2I} = (X_2^{\dagger *} N_{X_1^\dagger} X_2^\dagger)^{-1} X_2^{\dagger *} N_{X_1^\dagger} Y^\dagger$	
GLSE:	$\widehat{\underline{\psi}}_{2\underline{\psi}} = (X_2^{\dagger *} F_{1V_{\underline{\psi}}^\dagger} X_2^\dagger)^{-1} X_2^{\dagger *} F_{1V_{\underline{\psi}}^\dagger} Y^\dagger$	for a given $\underline{\psi}$ where $V_{\underline{\psi}}^\dagger$ is PD
EGLSE:	$\widehat{\underline{\psi}}_2 = (X_2^{\dagger *} F_{1V_{\widehat{\underline{\psi}}}^\dagger} X_2^\dagger)^{-1} X_2^{\dagger *} F_{1V_{\widehat{\underline{\psi}}}^\dagger} Y^\dagger$	where $V_{\widehat{\underline{\psi}}}^\dagger = V_{\widehat{\underline{\psi}}(\underline{Y})}^\dagger$ is PD $\forall\ \underline{Y} \in \mathcal{R}^n$.

The UBLUE result for the U-Model can be used to indicate when the LSE is the same as the GLSE when estimating $\underline{\psi}_2$. This is stated in the next theorem. The proof of the theorem follows directly from the proof of the U-UBLUE$_{\text{FR}}$ theorem. The result is restated in the following proposition. The proof of this proposition follows from the proposition in section 5.3.1, as it is a special case.

LQZ-UBLUE$_{FR}$ Theorem: The following are equivalent when Ξ contains a non-empty open set in $\mathcal{R}^r$:

i) $\widehat{\underline{\psi}}_{2I}$ is UBLUE for $\underline{\psi}_2$ in the LQEM for $\underline{Z}$

ii) $\underline{R}(V_{\underline{\psi}}^\dagger N_{X_1^\dagger} X_2^\dagger) \subset \underline{R}(X^\dagger) \quad \forall\ \underline{\psi} \in \Xi$

iii) $(X_2^{\dagger *} F_{1V_{\underline{\psi}}^\dagger} X_2^\dagger)^{-1} X_2^{\dagger *} F_{1V_{\underline{\psi}}^\dagger} = (X_2^{\dagger *} N_{X_1^\dagger} X_2^\dagger)^{-1} X_2^{\dagger *} N_{X_1^\dagger} \quad \forall\ \underline{\psi} \in \Xi.$

Proposition: For any $\underline{\psi} \in \mathcal{R}^r \ni R_{\underline{\psi}}$ is PD, the conditions in the LQZ-UBLUE theorem imply $\widehat{\underline{\psi}}_{2\underline{\psi}} = \widehat{\underline{\psi}}_{2I}$.

This theorem was stated for the LQEM for $\underline{Z}$, so it applies to both the ALQEM for $N_X\underline{Y}$ and the LQEM for $N_X\underline{Y}$. The following sections use these models to examine the maximum likelihood and restricted maximum likelihood procedures.

The next corollary gives another condition which has interpretative value. There are a few cases in which it can be applied. The condition can also be applied to the other LQEMs of interest.

Jordan Ideal Condition: If $A \in \underline{R}(X^\dagger)$, $B \in \underline{R}(N_{X_1^\dagger} X_2^\dagger) \Rightarrow ABA \in \underline{R}(N_{X_1^\dagger} X_2^\dagger)$ $(\circ)$, then $\underline{R}(V_{\underline{\psi}}^\dagger N_{X_1^\dagger} X_2^\dagger) \subset \underline{R}(X^\dagger)$.

proof: $(\circ) \Rightarrow V_{\underline{\psi}}^\dagger[\underline{R}(N_{X_1^\dagger} X_2^\dagger)] \subset \underline{R}(N_{X_1^\dagger} X_2^\dagger) \Rightarrow \underline{R}(V_{\underline{\psi}}^\dagger N_{X_1^\dagger} X_2^\dagger)] \subset \underline{R}(N_{X_1^\dagger} X_2^\dagger)$

$\Rightarrow \underline{R}(N_{X_1^\dagger} V_{\underline{\psi}}^\dagger N_{X_1^\dagger} X_2^\dagger) \subset \underline{R}(N_{X_1^\dagger} X_2^\dagger) \Rightarrow \underline{R}(V_{\underline{\psi}}^\dagger N_{X_1^\dagger} X_2^\dagger) \subset \underline{R}(X^\dagger)$ by lemma 3 in section 6.1. ∎

The next two propositions examine the linear transformation $N_{X_1^\dagger}$. The first proposition indicates $N_{X_1^\dagger}$ is the same whether $\underline{R}(X_1^\dagger)^\perp$ is taken with respect to $\underline{R}(X^\dagger)$ or $\mathcal{S}_n$, while the second proposition gives expressions for $P_{X_1^\dagger}$ and linear tranformations involving $N_{X_1^\dagger}$.

Proposition: $N_{X_1^\dagger}$ defined on $\underline{R}(X_1^\dagger)^\perp$ is the restriction of $N_{X_1^\dagger}$ on $\mathcal{S}_n$.

proof: i) Let $\mathcal{A} = \underline{R}(X_1^\dagger)$, $\mathcal{B} = \underline{R}(X^\dagger)$, and $\mathcal{S} = \mathcal{S}_n$. Then
$\mathcal{A}^{\perp\mathcal{B}} = \{B \in \mathcal{B} \mid < B, A > = 0 \ \forall \ A \in \mathcal{A}\}$ and $\mathcal{A}^{\perp\mathcal{S}} = \{ S \in \mathcal{S} \mid < S, A > = 0 \ \forall \ A \in \mathcal{A}\}$
$\Rightarrow \mathcal{A}^{\perp\mathcal{B}} = \mathcal{A}^{\perp\mathcal{S}} \cap \mathcal{B}$.
ii) Let $\overset{\perp}{\oplus}$ = perpendicular sum (Seely,1996). Then $\mathcal{B} = \mathcal{A} \overset{\perp}{\oplus} \mathcal{A}^{\perp\mathcal{B}}$ and $\mathcal{S} = \mathcal{B} \overset{\perp}{\oplus} \mathcal{B}^{\perp\mathcal{S}}$
$\Rightarrow \mathcal{S} = \mathcal{A} \overset{\perp}{\oplus} \mathcal{A}^{\perp\mathcal{B}} \overset{\perp}{\oplus} \mathcal{B}^{\perp\mathcal{S}} \Rightarrow \mathcal{A}^{\perp\mathcal{S}} = \mathcal{A}^{\perp\mathcal{B}} \overset{\perp}{\oplus} \mathcal{B}^{\perp\mathcal{S}}$.
iii) For $B \in \mathcal{B}$, $B = A \overset{\perp}{\oplus} C$ where $A \in \mathcal{A}$, $C \in \mathcal{A}^{\perp\mathcal{B}}$ and $B = A \overset{\perp}{\oplus} C$ where $A \in \mathcal{A}$, $C \in \mathcal{A}^{\perp\mathcal{S}}$ by ii).
Thus, $P_{\mathcal{A}}^{\mathcal{B}}(B) = A = P_{\mathcal{A}}^{\mathcal{S}}(B)$.
iv) By iii), $N_{\mathcal{A}}^{\mathcal{B}} = P_{\mathcal{A}^{\perp\mathcal{B}}}^{\mathcal{B}}(B) = (I^{\mathcal{B}} - P_{\mathcal{A}}^{\mathcal{B}})(B) = B - P_{\mathcal{A}}^{\mathcal{B}}(B) = B - P_{\mathcal{A}}^{\mathcal{S}}(B)$
$= (I^{\mathcal{S}} - P_{\mathcal{A}}^{\mathcal{S}})(B) = P_{\mathcal{A}^{\perp\mathcal{S}}}^{\mathcal{S}}(B) = N_{\mathcal{A}}^{\mathcal{S}}$. ∎

Proposition: Let $T = \{\mathrm{tr}(R_j R_l)\}_{r\times r} = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix}$ where T_{fg} is $r_f \times r_g$. For $A \in \mathcal{S}_n$, let
$\underline{u}_A = \{\mathrm{tr}(R_h A)\}_{r\times 1} = \begin{bmatrix} \underline{u}_{A1} \\ \underline{u}_{A2} \end{bmatrix}$ where $\underline{u}_{Af}$ is $r_f \times 1$. Then

i) $P_{X_1^\dagger} A = \sum_{h=1}^{r_1} a_h R_h$ where $\underline{a} = T_{11}^{-1} \underline{u}_{A1}$

ii) $X_2^{\dagger *} N_{X_1^\dagger} A = \underline{u}_{A2} - T_{21} T_{11}^{-1} \underline{u}_{A1}$

iii) $X_2^{\dagger *} N_{X_1^\dagger} X_2^\dagger = T_{22} - T_{21} T_{11}^{-1} T_{12}$.

proof: i) $P_{X_1^\dagger} A = W$ where $A = W + Z$ with $W \in \underline{R}(X_1^\dagger)$ and $Z \in \underline{R}(X_1^\dagger)^\perp$
$\Rightarrow W = \sum_{h=1}^{r_1} a_h R_h$ for some $\underline{a}$. In addition, $\forall \ m = 1, \dots, r_1 \ < Z, R_m > = 0$
$\Rightarrow \mathrm{tr}(Z R_m) = 0 \Rightarrow \mathrm{tr}(A R_m) - \mathrm{tr}(W R_m) = 0 \Rightarrow \mathrm{tr}(A R_m) = \sum_{h=1}^{r_1} a_h \mathrm{tr}(R_h R_m)$
$\Rightarrow \{\mathrm{tr}(A R_m)\}_{r_1\times 1} = \{\mathrm{tr}(R_h R_m)\}_{r_1\times r_1}\underline{a} \Rightarrow \underline{a} = \{\mathrm{tr}(R_h R_m)\}^{-1}\{\mathrm{tr}(A R_m)\} = T_{11}^{-1}\underline{u}_{A1}$.
ii) $X_2^{\dagger *} N_{X_1^\dagger} A = X_2^{\dagger *} A - X_2^{\dagger *} P_{X_1^\dagger} A = X_2^{\dagger *} A - \sum_{h=1}^{r_1} a_h X_2^{\dagger *} R_h$ by i)
$= \{\mathrm{tr}(R_i A)\}_{i=r_1+1}^{r} - \sum_{h=1}^{r_1} a_h \{\mathrm{tr}(R_i R_h)\}_{i=r_1+1}^{r}$ by lemma 1 i) in section 3.2.3
$= \underline{u}_{A2} - T_{21} T_{11}^{-1} \underline{u}_{A1}$ by the definition of $\underline{a}$.
iii) For $\underline{v} = \{v_i\}_{i=r_1+1}^{r} \in \mathcal{R}^{r_2}$, $X_2^{\dagger *} N_{X_1^\dagger} X_2^\dagger \underline{v} = \sum_{i=r_1+1}^{r} v_i X_2^{\dagger *} N_{X_1^\dagger} R_i$
$= \sum_{i=r_1+1}^{r} v_i (\underline{u}_{R_i 2} - T_{21} T_{11}^{-1} \underline{u}_{R_i 1})$ by ii)
$= \sum_{i=r_1+1}^{r} \underline{u}_{R_i 2} v_i - T_{21} T_{11}^{-1} \sum_{i=r_1+1}^{r} \underline{u}_{R_i 1} v_i = T_{22}\underline{v} - T_{21} T_{11}^{-1} T_{12} \underline{v} = (T_{22} - T_{21} T_{11}^{-1} T_{12})\underline{v}$. ∎

6.4.2. ELMLQE Results

For doing maximum likelihood estimation in the Y-Model, assumptions [L], [O], and [N] were used. Extending the definition of Szatrowski (1980), the explicit linear maximum likelihood estimator (ELMLQE) for $\underline{\psi}_{2(k_2 \times 1)}$, with $k_2 = k + 1 - k_1$, satisfies $\widehat{\underline{\psi}}_{2\text{MLQ}} = AY_2^\circ$, where A is a linear transformation that is not random and does not depend on any parameters and Y_2° does not depend on any estimators or unknown parameters.

Consider the following notation for the ALQEM for $(\underline{Y} - X\widehat{\underline{\beta}})$, where it is artificially assumed that $\mathrm{E}[Y_2^\circ] \in \{X^\circ \underline{\psi} \mid \underline{\psi} \in \Xi\}$ as indicated in section 3.2.1 :

i) $\mathrm{E}[Y_2^\circ] = X^\circ \underline{\psi} = V_{\underline{\psi}} = \sum_{i=1}^{k+1} \psi_i V_i = \sum_{i=1}^{k_1} \psi_i V_i + \sum_{i=k_1+1}^{k+1} \psi_i V_i = X_1^\circ \underline{\psi}_1 + X_2^\circ \underline{\psi}_2$ for $X_j^\circ \in \mathcal{L}(\mathcal{R}^{k_j}, \mathcal{S}_n)$

ii) $V_{\underline{\psi}}^\circ = 2\Psi_{V_{\underline{\psi}}} \in \mathcal{V}^\circ \subset \mathcal{L}_{PD}(\mathcal{S}_n, \mathcal{S}_n)$

iii) $F_{1V_{\underline{\psi}}^\circ} : \mathcal{S}_n \rightarrow \mathcal{S}_n \qquad F_{1V_{\underline{\psi}}^\circ} = V_{\underline{\psi}}^{\circ -1} - V_{\underline{\psi}}^{\circ -1} X_1^\circ (X_1^{\circ *} V_{\underline{\psi}}^{\circ -1} X_1^\circ)^{-1} X_1^{\circ *} V_{\underline{\psi}}^{\circ -1}$

iv) $N_{X_1^\circ} : \mathcal{S}_n \rightarrow \mathcal{S}_n \qquad N_{X_1^\circ} = I - X_1^\circ (X_1^{\circ *} X_1^\circ)^{-1} X_1^{\circ *} = I - P_{X_1^\circ}$.

The definition in i) gives a partition of the expectation while ii) gives the form of the covariance. The notation in iii) and iv) define linear transformations that will be of interest in later results. Assuming the V_i's are linearly independent, then $\underline{r}(X^\circ) = k + 1$ and this model fits into the full rank setting. The EGLSE for $\underline{\psi}$ can be partitioned to give an EGLSE for $\underline{\psi}_2$ which can be compared to the LSE for $\underline{\psi}_2$. The least squares, generalized least squares, and estimated generalized least squares estimators for the ALQEM for $N_X \underline{Y}$ are:

LSE: $\widehat{\underline{\psi}}_{2I} = (X_2^{\circ *} N_{X_1^\circ} X_2^\circ)^{-1} X_2^{\circ *} N_{X_1^\circ} Y_2^\circ$

GLSE: $\widehat{\underline{\psi}}_{2\underline{\psi}} = (X_2^{\circ *} F_{1V_{\underline{\psi}}^\circ} X_2^\circ)^{-1} X_2^{\circ *} F_{1V_{\underline{\psi}}^\circ} Y_2^\circ$ for a given $\underline{\psi}$ where $V_{\underline{\psi}}^\circ$ is PD

EGLSE: $\widehat{\underline{\psi}}_2 = (X_2^{\circ *} F_{1V_{\widehat{\underline{\psi}}}^\circ} X_2^\circ)^{-1} X_2^{\circ *} F_{1V_{\widehat{\underline{\psi}}}^\circ} Y_2^\circ$ where $V_{\widehat{\underline{\psi}}}^\circ = V_{\widehat{\underline{\psi}}(\underline{Y})}^\circ$ is PD $\forall\ \underline{Y} \in \mathcal{R}^n$.

For clarity, the EGLSE will often be labelled as $\widehat{\underline{\psi}}_{\text{EGLS}}$. The ALQEM for $(\underline{Y} - X\widehat{\underline{\beta}})$ was defined so that the equations for the EGLSE correspond to the ML equations for $\underline{\psi}$. The following lemma establishes the equivalence between the EGLSE and the MLQE for $\underline{\psi}_2$.

Lemma: $\widehat{\underline{\psi}}_{2\text{MLQ}} = \widehat{\underline{\psi}}_{2\text{EGLS}}$.

proof: From section 4.3.2, $X^\circ \widehat{\underline{\psi}}_{\text{MLQ}} = X^\circ \widehat{\underline{\psi}}_{\text{EGLS}} \Rightarrow \widehat{\underline{\psi}}_{\text{MLQ}} = \widehat{\underline{\psi}}_{\text{EGLS}}$ multiplying both sides by $(X^{\circ *} X^\circ)^{-1} X^\circ$ as X° has full column rank. Then $\widehat{\underline{\psi}}_{\text{EGLS}}$ can be partitioned as in lemma 2 in section 6.1. ∎

Under the ALQEM for $(\underline{Y} - X\widehat{\underline{\beta}})$, the response $Y_2^\diamond = (\underline{Y} - X\widehat{\underline{\beta}})\,(\underline{Y} - X\widehat{\underline{\beta}})'$ will generally depend on $\widehat{\underline{\psi}}$ as $X\widehat{\underline{\beta}} = X(X'V_{\widehat{\underline{\psi}}}^{-1}X)^{-}X'V_{\widehat{\underline{\psi}}}^{-1}\underline{Y}$. This is likely to present a difficulty in satisfying the definition of an ELMLQE for $\underline{\psi}_2$. In order to deal with this difficulty, it will be assumed that ZC holds for the Y-Model where $\underline{R}(V_{\underline{\psi}}X) \subset \underline{R}(X)\ \forall\ \underline{\psi} \in \Xi$. As shown in section 4.2.2, this condition is sufficient for $\underline{Y} - X\widehat{\underline{\beta}}_{\underline{\psi}} = N_X\underline{Y}\ \ \forall\ \underline{\psi} \in \mathcal{R}^{k+1}$ such that $V_{\underline{\psi}}$ is PD. Under ZC for the Y-Model, the ALQEM for $(\underline{Y} - X\widehat{\underline{\beta}})$ can be used to obtain conditions for the existence of an ELMQE for $\Gamma'\underline{\psi}$. However, it is possible that a weaker condition could suffice for some examples as only the subvector $\underline{\psi}_2$ is of interest. Still, the following theorem will assume ZC.

ALQNY-ELMLQE$_{FR}$ Theorem: Consider the Y-Model under [L], [O], [N], and ZC. If $\widehat{\underline{\psi}}_{\text{MLQ}}$ exists $\ni$ $V_{\widehat{\underline{\psi}}_{\text{MLQ}}}$ is PD and $\underline{R}(V_{\underline{\psi}}^\diamond N_{X_1^\diamond}X_2^\diamond) \subset \underline{R}(X^\diamond)\ \forall\ \underline{\psi} \in \Xi$, then $\widehat{\underline{\psi}}_{2I}$ is an ELMLQE for $\underline{\psi}_2$.

proof: Since $V_{\widehat{\underline{\psi}}_{\text{MLQ}}}$ is PD by hypothesis and $\widehat{\underline{\psi}}_{2\text{MLQ}} = \widehat{\underline{\psi}}_{2\text{EGLS}}$ by the above lemma, $\widehat{\underline{\psi}}_{2\text{MLQ}} = \widehat{\underline{\psi}}_{2I} = AY_2^\diamond$ by the proposition in section 6.4.1 where $A = (X_2^{\diamond *}N_{X_1^\diamond}X_2^\diamond)^{-1}X_2^{\diamond *}N_{X_1^\diamond}$ and $Y_2^\diamond = N_X\underline{Y}\underline{Y}'N_X$ does not depend $\underline{\psi}$ or $\widehat{\underline{\psi}}$ by ZC $\Rightarrow \widehat{\underline{\psi}}_{2I}$ is an ELMLQE for $\underline{\psi}_2$ by definition. ▪

Note $\widehat{\underline{\psi}}_{2I}$ may not be an unbiased estimator for $\underline{\psi}_2$, but it does satisfy the requirement of an ELMLQE. An example will be presented in section 6.6. The explicit expression for $\widehat{\underline{\psi}}_{2I}$ is given below.

Lemma: Let $T = \{\text{tr}(V_iV_j)\}_{(k+1)\times(k+1)} = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix}$ and $\underline{u} = \{\underline{Y}'N_XV_iN_X\underline{Y}\}_{(k+1)\times 1} = \begin{bmatrix} \underline{u}_1 \\ \underline{u}_2 \end{bmatrix}$. For the ALQEM for $(\underline{Y} - X\widehat{\underline{\beta}})$ under ZC for the Y-Model, $\widehat{\underline{\psi}}_{2I} = (T_{22} - T_{21}T_{11}^{-1}T_{12})^{-1}(\underline{u}_2 - T_{21}T_{11}^{-1}\underline{u}_1)$.

proof: Since ZC holds for the Y-Model, then $\underline{Y} - X\widehat{\underline{\beta}} = N_X\underline{Y}$ by the Y-FUBLUE theorem $\Rightarrow \{\text{tr}(V_iY_2^\diamond)\} = \{\text{tr}(V_i\underline{Y}'N_XN_X\underline{Y})\} = \{\underline{Y}'N_XV_iN_X\underline{Y}\} = \underline{u}$. By the last proposition in section 6.4.1, $\widehat{\underline{\psi}}_{2I} = (X_2^{\diamond *}N_{X_1^\diamond}X_2^\diamond)^{-1}X_2^{\diamond *}N_{X_1^\diamond}Y_2^\diamond = (T_{22} - T_{21}T_{11}^{-1}T_{12})^{-1}(\underline{u}_2 - T_{21}T_{11}^{-1}\underline{u}_1)$. ▪

6.4.3. ELREMLQE Results

For doing restricted maximum likelihood estimation in the Y-Model, assumptions [L], [O], and [N] were used. Extending the definition of Szatrowski (1980), the explicit linear restricted maximum likelihood estimator (ELREMLQE) for $\underline{\psi}_{2(k_2\times 1)}$, with $k_2 = k + 1 - k_1$, satisfies $\widehat{\underline{\psi}}_{2\text{REMLQ}} = AY^\circ$, where A is a linear transformation which is not random and does not depend on any parameters and Y°does not depend on any estimators or unknown parameters. Consider the notation listed below for the LQEM for $N_X\underline{Y}$ where $\tilde{V} = N_XVN_X$:

i) $\mathrm{E}[Y^\circ] = X^\circ\underline{\psi} = \tilde{V}_{\underline{\psi}} = \sum_{i=1}^{k+1}\psi_i\tilde{V}_i = \sum_{i=1}^{k_1}\psi_i\tilde{V}_i + \sum_{i=k_1+1}^{k}\psi_i\tilde{V}_i = X_1^\circ\underline{\psi}_1 + X_2^\circ\underline{\psi}_2$ for $X_j^\circ \in \mathcal{L}(\mathcal{R}^{k_j}, \mathcal{S}_n)$

ii) $V_{\underline{\psi}}^\circ = 2\Psi_{\tilde{V}_{\underline{\psi}}} \in \mathcal{V}^\circ \subset \mathcal{L}_{\mathrm{PD}}(\mathcal{S}_n, \mathcal{S}_n)$

iii) $F_{1V_{\underline{\psi}}^\circ} : \mathcal{S}_n \rightarrow \mathcal{S}_n \qquad F_{1V_{\underline{\psi}}^\circ} = V_{\underline{\psi}}^{\circ-1} - V_{\underline{\psi}}^{\circ-1}X_1^\circ(X_1^{\circ*}V_{\underline{\psi}}^{\circ-1}X_1^\circ)^{-1}X_1^{\circ*}V_{\underline{\psi}}^{\circ-1}$

iv) $N_{X_1^\circ} : \mathcal{S}_n \rightarrow \mathcal{S}_n \qquad N_{X_1^\circ} = I - X_1^\circ(X_1^{\circ*}X_1^\circ)^{-1}X_1^{\circ*} = I - P_{X_1^\circ}$.

The definition in i) gives a partition of the expectation while ii) gives the form of the covariance. The notation in iii) and iv) give linear transformations that will be of interest in later results. Assuming the $N_X V_i N_X$'s are linearly independent, then $\underline{r}(X^\circ) = k + 1$ and this model fits into the full rank setting. The EGLSE for $\underline{\psi}$ can be partitioned to give an EGLSE for $\underline{\psi}_2$ which can be compared to the LSE for $\underline{\psi}_2$. The associated least squares, generalized least squares, and estimated generalized least squares estimators for the LQEM for $N_X\underline{Y}$ are given by:

LSE: $\widehat{\underline{\psi}}_{2I} = (X_2^{\circ*}N_{X_1^\circ}X_2^\circ)^{-1}X_2^{\circ*}N_{X_1^\circ}Y^\circ$

GLSE: $\widehat{\underline{\psi}}_{2\underline{\psi}} = (X_2^{\circ*}F_{1V_{\underline{\psi}}^\circ}X_2^\circ)^{-1}X_2^{\circ*}F_{1V_{\underline{\psi}}^\circ}Y^\circ$ for a given $\underline{\psi}$ where $V_{\underline{\psi}}^\circ$ is PD

EGLSE: $\widehat{\underline{\psi}}_2 = (X_2^{\circ*}F_{1V_{\hat{\underline{\psi}}}^\circ}X_2^\circ)^{-1}X_2^{\circ*}F_{1V_{\hat{\underline{\psi}}}^\circ}Y^\circ$ where $V_{\hat{\underline{\psi}}}^\circ = V_{\hat{\underline{\psi}}(\underline{Y})}^\circ$ is PD $\forall\ \underline{Y} \in \mathcal{R}^n$.

For clarity, the EGLSE will often be labelled as $\widehat{\underline{\psi}}_{\mathrm{EGLS}}$. The LQEM for $N_X\underline{Y}$ was defined so that the equations for the EGLSE correspond to the REML equations for $\underline{\psi}$. The following lemma establishes the equivalence between the EGLSE and the REMLQE for $\underline{\psi}_2$.

Lemma: $\widehat{\underline{\psi}}_{2\mathrm{REMLQ}} = \widehat{\underline{\psi}}_{2\mathrm{EGLS}}$.

proof: From section 4.3.3, $X^\circ\widehat{\underline{\psi}}_{\mathrm{REMLQ}} = X^\circ\widehat{\underline{\psi}}_{\mathrm{EGLS}} \Rightarrow \widehat{\underline{\psi}}_{\mathrm{REMLQ}} = \widehat{\underline{\psi}}_{\mathrm{EGLS}}$ multiplying both sides by $(X^{\circ*}X^\circ)^{-1}X^\circ$ as X° has full column rank. Then $\widehat{\underline{\psi}}_{\mathrm{EGLS}}$ can be partitioned as in lemma 2 in section 6.1. ∎

The main theorem can now be stated concerning the existence of an ELREMLQE for $\underline{\psi}_2$.

LQNY-ELREMLQE$_{FR}$ Theorem: Consider the Y-Model under [L], [O], and [N]. If $\widehat{\underline{\psi}}_{\mathrm{REMLQ}}$ exists $\ni$ $V_{\widehat{\underline{\psi}}_{\mathrm{REMLQ}}}$ is PD and $\underline{R}(V_{\underline{\psi}}^\circ N_{X_1^\circ}X_2^\circ) \subset \underline{R}(X^\circ)\ \forall\ \underline{\psi} \in \Xi$, then $\widehat{\underline{\psi}}_{2I}$ is an ELREMLQE for $\underline{\psi}_2$.

proof: Since $V_{\widehat{\underline{\psi}}_{\mathrm{REMLQ}}}$ is PD by hypothesis and $\widehat{\underline{\psi}}_{2\mathrm{REMLQ}} = \widehat{\underline{\psi}}_{2\mathrm{EGLS}}$ by the above lemma, $\widehat{\underline{\psi}}_{2\mathrm{REMLQ}} = \widehat{\underline{\psi}}_{2I} = AY^\circ$ by the proposition in section 6.4.1 where $A = (X_2^{\circ*}N_{X_1^\circ}X_2^\circ)^{-1}X_2^{\circ*}N_{X_1^\circ}$ and $Y^\circ = N_X\underline{Y}\underline{Y}'N_X$ does not depend $\underline{\psi}$ or $\widehat{\underline{\psi}} \Rightarrow \widehat{\underline{\psi}}_{2I}$ is an ELREMLQE for $\underline{\psi}_2$ by definition. ∎

The ELREMLQE for $\underline{\psi}_2$ is given by $\widehat{\underline{\psi}}_{2I}$ which is linear and UBLUE for $\underline{\psi}_2$. Examples will be presented in section 6.7 and 6.8. A lemma is now given which shows how to calculate $\widehat{\underline{\psi}}_{2I}$. Note that $\underline{\tilde{u}} = \underline{u}$ as given in the last lemma in section 6.4.2.

Lemma: Let $\tilde{T} = \{\mathrm{tr}(\tilde{V}_i\tilde{V}_j)\}_{(k+1)\times(k+1)} = \begin{bmatrix} \tilde{T}_{11} & \tilde{T}_{12} \\ \tilde{T}_{21} & \tilde{T}_{22} \end{bmatrix}$ and $\underline{\tilde{u}} = \{\underline{Y}'N_X\tilde{V}_iN_X\underline{Y}\}_{(k+1)\times 1} = \begin{bmatrix} \underline{\tilde{u}}_1 \\ \underline{\tilde{u}}_2 \end{bmatrix}$.
For the LQEM for $N_X\underline{Y}$, $\widehat{\underline{\psi}}_{2I} = (\tilde{T}_{22} - \tilde{T}_{21}\tilde{T}_{11}^{-1}\tilde{T}_{12})^{-1}(\underline{\tilde{u}}_2 - \tilde{T}_{21}\tilde{T}_{11}^{-1}\underline{\tilde{u}}_1)$.

proof: Note $\{\mathrm{tr}(\tilde{V}_iY^\circ)\} = \{\mathrm{tr}(\tilde{V}_iN_X\underline{Y}\underline{Y}'N_X)\} = \{\underline{Y}'N_X\tilde{V}_iN_X\underline{Y}\} = \underline{\tilde{u}}$. By the last proposition in section 6.4.1, $\widehat{\underline{\psi}}_{2I} = (X_2^{\circ *}N_{X_1^\circ}X_2^\circ)^{-1}X_2^{\circ *}N_{X_1^\circ}Y^\circ = (\tilde{T}_{22} - \tilde{T}_{21}\tilde{T}_{11}^{-1}\tilde{T}_{12})^{-1}(\underline{\tilde{u}}_2 - \tilde{T}_{21}\tilde{T}_{11}^{-1}\underline{\tilde{u}}_1)$. ■

6.5. Checking the Conditions

This section discusses methods and issues involved in checking the UBLUE conditions under assumptions [L] and [O] for the Y-Model. These checks are designed for a programming language that can handle matrix computations. The following result is useful for performing the checks.

Lemma: i) $\underline{R}(A) \subset \underline{R}(B) \not\Rightarrow \underline{R}(\mathrm{vec}A) \subset \underline{R}(\mathrm{vec}B)$.

ii) $A_1, \ldots, A_n \in \mathrm{sp}\,\{B_1, \ldots, B_m\} \Leftrightarrow \underline{R}([\,\mathrm{vec}A_1, \ldots, \mathrm{vec}A_n)] \subset \underline{R}([\,\mathrm{vec}B_1, \ldots, \mathrm{vec}B_m)\,]$.

proof: i) Counterexample: $A = \begin{bmatrix} 1 & 2 \\ 2 & 2 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 4 \\ 2 & 4 \end{bmatrix}$, $\begin{bmatrix} (\mathrm{vec}A)' \\ (\mathrm{vec}B)' \end{bmatrix} = \begin{bmatrix} 1 & 2 & 2 & 2 \\ 1 & 2 & 4 & 4 \end{bmatrix}$.

ii) $A_1, \ldots, A_n \in \mathrm{sp}\{B_1, \ldots, B_m\} \Leftrightarrow A_i = \sum_{j=1}^{m}\alpha_j^{(i)}B_j$ for some $\underline{\alpha}^{(i)} \in \mathcal{R}^m \;\; \forall\; i$

$\Leftrightarrow \mathrm{vec}\,A_i = \sum_{j=1}^{m}\alpha_j^{(i)}\mathrm{vec}\,B_j$ by the linearity of vec

$\Leftrightarrow \underline{R}([\,\mathrm{vec}A_1, \ldots, \mathrm{vec}A_n)\,] \subset \underline{R}([\,\mathrm{vec}B_1, \ldots, \mathrm{vec}B_m)\,]$ by the definition of containment. ■

Assuming [L] and [O], consider the following checks that can be made for Zyskind's Condition in the Y-Model (1), the full rank assumption in the LQEM for $\underline{Z}$ (2), the Quadratic Subspace Condition in the $\underline{Z}$-Model (3), and the Generalized Zyskind's Condition for the full rank setting in the LQEM for $\underline{Z}$ (4).

(1) *ZC Check*: $\underline{R}(V_{\underline{\psi}}X) \subset \underline{R}(X) \;\; \forall\, \underline{\psi} \in \Xi$.

$\underline{R}(V_{\underline{\psi}}X) \subset \underline{R}(X) \;\; \forall\; \underline{\psi} \in \Xi \Leftrightarrow \underline{R}(V_{\underline{\psi}}X) \subset \underline{R}(X) \;\; \forall\, \underline{\psi} \in \mathcal{R}^{k+1}$ as Ξ contains a non-empty open set

$\Leftrightarrow \underline{R}(V_lX) \subset \underline{R}(X) \quad l = 1, \ldots, k$

$\Leftrightarrow \underline{r}([V_lX\;\; X]) - \underline{r}(X) = 0 \quad l = 1, \ldots, k$ from the proposition in section 2.3

$\Leftrightarrow \underline{r}([V_1X\;\; V_2X\; \ldots\; V_kX\;\; X]) - \underline{r}(X) = 0.$

(2) *Full Rank Check*: $\underline{r}(X^{\dagger}) = \dim \text{sp}\mathcal{V}^{\dagger} = r$.

Consider $S = \{\text{vec}(R_j)\}_{n^2 \times r}$. The model has full rank provided $\underline{r}(S) = r \Leftrightarrow \underline{r}(S) - r = 0$.

(3) *QS Check*: $\underline{R}(X^{\dagger}) = \text{sp}\{R_1, \ldots, R_r\}$ is a QS.

Let $d = \binom{r}{1} + \binom{r}{2} = r + \frac{1}{2}r(r-1) = \frac{1}{2}r(r+1)$. For $j = 1, \ldots, r$, $l = 1, \ldots, r$, $j \leq l$, let $M_{jl} = R_jR_l + R_lR_j$. Then sp$\mathcal{V}$ is a QS $\Leftrightarrow M_{jl} \in \text{sp}\mathcal{V}\ \forall\, j, l$ from the proposition in section 2.8

$\Leftrightarrow \underline{R}(M) \subset \underline{R}(S)$ where $M = \{\text{vec}(M_{jl})\}_{n^2 \times d}$ and $S = \{\text{vec}(R_j)\}_{n^2 \times r}$ by the above lemma

$\Leftrightarrow \underline{r}(M, S) - \underline{r}(S) = 0$ from the proposition in section 2.3.

(4) *GZC*$_{FR}$ *Check*: $\underline{R}(V_{\underline{\psi}}^{\dagger} N_{X_1^{\dagger}} X_2^{\dagger}) \subset \underline{R}(X^{\dagger})\ \forall\, \underline{\psi} \in \Xi$.

i) Characterize $P_{X_1^{\dagger}} R_i$.

Let $i = r_1 + 1, \ldots, r$, $h = 1, \ldots, r_1$, $m = 1, \ldots, r_1$, and $r_2 = r - r_1$. By last proposition in section 6.4.1,

$\text{vec}(P_{X_1^{\dagger}} R_i) = \text{vec}(\sum_{h=1}^{r_1} a_h^{(i)} R_h) = \sum_{h=1}^{r_1} a_h^{(i)} \text{vec}(R_h) = [\text{vec}(R_1) \ldots \text{vec}(R_{r_1})] T_{11}^{-1} \underline{u}_{R_i1}$

where $\underline{a}^{(i)} = T_{11}^{-1} \underline{u}_{R_i1} = \{\text{tr}(R_m R_h)\}_{r_1 \times r_1}^{-1} \{\text{tr}(R_i R_h)\}_{r_1 \times 1}$. Then unvec $\text{vec}(P_{X_1^{\dagger}} R_i)$ to obtain $P_{X_1^{\dagger}} R_i$.

ii) Find a spanning set for $\underline{R}(N_{X_1^{\dagger}} X_2^{\dagger})$.

Note $\underline{R}(N_{X_1^{\dagger}} X_2^{\dagger}) = N_{X_1^{\dagger}}[\underline{R}(X_2^{\dagger})] = \text{sp}\{E_{r_1+1}, \ldots, E_r\}$ where $E_i = N_{X_1^{\dagger}} R_i = R_i - P_{X_1^{\dagger}} R_i$.

iii) Find a spanning set for $\text{sp}\{\underline{R}(V_{\underline{\psi}}^{\dagger} N_{X_1^{\dagger}} X_2^{\dagger}) | \underline{\psi} \in \Xi\}$.

$\underline{R}(V_{\underline{\psi}}^{\dagger} N_{X_1^{\dagger}} X_2^{\dagger}) = V_{\underline{\psi}}^{\dagger}[\underline{R}(N_{X_1^{\dagger}} X_2^{\dagger})] = V_{\underline{\psi}}^{\dagger}[\text{sp}\{E_{r_1+1}, \ldots, E_r\}] = \text{sp}\{V_{\underline{\psi}}^{\dagger} E_{r_1+1}, \ldots, V_{\underline{\psi}}^{\dagger} E_r\}$ by ii)

$= \text{sp}\{R_{\underline{\psi}} E_{r_1+1} R_{\underline{\psi}}, \ldots, R_{\underline{\psi}} E_r R_{\underline{\psi}}\}$. Thus,

$\text{sp}\{\underline{R}(V_{\underline{\psi}}^{\dagger} N_{X_1^{\dagger}} X_2^{\dagger}) | \underline{\psi} \in \Xi\} = \text{sp}\{R_{\underline{\psi}} E_{r_1+1} R_{\underline{\psi}}, \ldots, R_{\underline{\psi}} E_r R_{\underline{\psi}} | \underline{\psi} \in \Xi\}$

$= \text{sp}\{G_{ijl} | r_1 + 1 \leq i \leq r, 1 \leq j \leq l \leq r\}$ where $G_{ijl} = R_j E_i R_l + R_l E_i R_j$ by lemma 1 in section 4.3.1.

iv) Thus, GZC$_{\text{FR}}$ holds $\Leftrightarrow \underline{R}(V_{\underline{\psi}}^{\dagger} N_{X_1^{\dagger}} X_2^{\dagger}) \subset \underline{R}(X^{\dagger})\ \forall\, \underline{\psi} \in \Xi$

$\Leftrightarrow G_{ijl} \in \text{sp}\{R_j | 1 \leq j \leq r\}\ \forall i, j, l \ni r_1 + 1 \leq i \leq r, 1 \leq j \leq l \leq r$ by iii)

$\Leftrightarrow \underline{R}(G) \subset \underline{R}(S)$ where $G = \{\text{vec}\, G_{ijl}\}_{n^2 \times r_2 d}$ and $S = \{\text{vec}\, R_j\}_{n^2 \times r}$ by the above lemma

$\Leftrightarrow \underline{r}(G, S) - \underline{r}(S) = 0$ from the proposition in section 2.3.

The checks (2)-(4) have been demonstrated using the notation from the LQEM for $\underline{Z}$. This was done to demonstrate their applicability to both the ML and REML estimation procedures using the ALQEM for $(\underline{Y} - X\widehat{\underline{\beta}})$ (a) and the LQEM for $N_X\underline{Y}$ where $\tilde{V}_j = N_X V_j N_X$ (b). The respective changes for checks (2)-(4) would be as follows for these two models:

	S	M	G	E_i
(a)	$\{\text{vec}(V_j)\}$	$\{\text{vec}(V_jV_l + V_lV_j)\}$	$\{\text{vec}(V_jE_iV_l + V_lE_iV_j)\}$	$N_{X_1^{\circ}} V_i$
(b)	$\{\text{vec}(\tilde{V}_j)\}$	$\{\text{vec}(\tilde{V}_j\tilde{V}_l + \tilde{V}_l\tilde{V}_j)\}$	$\{\text{vec}\,(\tilde{V}_j\tilde{E}_i\tilde{V}_l + \tilde{V}_l\tilde{E}_i\tilde{V}_j)\}$	$N_{X_1^{\circ}} \tilde{V}_i$

For purposes of characterizing examples which meet the UBLUE conditions, certain tools are helpful in addition to the checks described above. The first item is an RTABLE which examines pairwise products of the matrices in the set $\mathrm{sp}\mathcal{V}^\dagger$. The entries of the table assess the balance in the design. The r_{jl} entry of the RTABLE is given by:

$$r_{jl} = \begin{cases} -1 \text{ if } R_jR_l \neq cR_p \text{ for some } c, p \\ 0 \text{ if } R_jR_l = 0 \\ 1 \text{ if } R_jR_l = cR_p \text{ for some } c \neq 0, p \end{cases} .$$

The second tool involves the elements of $\underline{a}^{(i)}$ defined by the relation $\sum_{h=1}^{r_1} a_h^{(i)} \operatorname{tr}(R_mR_h) = \operatorname{tr}(R_mR_i)$ for $1 \leq m \leq r_1$ and $r_1 + 1 \leq i \leq r$ which is given in part i) of the last proposition in section 6.4.1. If $a_h^{(i)} \neq 0$, then it is likely that $r_{mh} = 0$ or 1 would be needed in order for the relation to hold. However, if $a_h^{(i)} = 0$, then factor h would not need to have balance properties with respect to factor m. This tool provides model-based conditions, as it indicates whether or not the effect associated with factor h should be in the model given the design. The significance of $\underline{a}^{(i)}$ will be evident in later examples.

It is helpful to know that when checking for the existence of an ELMLQE or an ELREMLQE for a subvector of the variance component vector it is only necessary to consider models with more than two variance components. This is due to the fact that a model with two variance components which satisfies the $\mathrm{GZC_{FR}}$ will automatically satisfy the QS condition.

Theorem: Assume $\mathrm{sp}\mathcal{V}^\dagger = \{R_1, R_2\}$ where $R_2^2 = R_2, R_1R_2 = R_1$, and $\operatorname{tr}(R_1) \neq 0$. If $\mathrm{GZC_{FR}}$ holds for $\underline{\psi}_1$ or $\underline{\psi}_2$, then $\mathrm{sp}\mathcal{V}^\dagger$ is a QS.

proof: i) $\mathrm{GZC_{FR}}$ for $\underline{\psi}_1 \Rightarrow G_{112} \in \mathrm{sp}\mathcal{V}^\dagger$ where G_{112} is defined in the $\mathrm{GZC_{FR}}$ check. Thus,
$R_1(R_1 - aR_2)R_2 + R_2(R_1 - aR_2)R_1 = R_1^2 - aR_1 + R_1^2 - aR_1 = 2(R_1^2 - aR_1) \in \mathrm{sp}\mathcal{V}^\dagger$
$\Rightarrow R_1^2 \in \mathrm{sp}\mathcal{V}^\dagger$ since $R_1 \in \mathrm{sp}\mathcal{V}^\dagger$. Thus, $\mathrm{sp}\mathcal{V}^\dagger$ is a QS given the properties of R_2.
ii) $\mathrm{GZC_{FR}}$ for $\underline{\psi}_2 \Rightarrow G_{212} \in \mathrm{sp}\mathcal{V}^\dagger$
$\Rightarrow R_1(R_2 - aR_1)R_2 + R_2(R_2 - aR_1)R_1 = R_1 - aR_1^2 + R_1 - aR_1^2 = 2(R_1 - aR_1^2) \in \mathrm{sp}\mathcal{V}^\dagger$.
By the last proposition in section 6.4.1, $a = \frac{\operatorname{tr}(R_1R_2)}{\operatorname{tr}(R_1^2)} = \frac{\operatorname{tr}(R_1)}{\operatorname{tr}(R_1^2)} \neq 0$ as $\operatorname{tr}(R_1) \neq 0$ by hypothesis.
Hence, $R_1^2 \in \mathrm{sp}\mathcal{V}^\dagger$ since $R_1 \in \mathrm{sp}\mathcal{V}^\dagger$. Thus, $\mathrm{sp}\mathcal{V}^\dagger$ is a QS given the properties of R_2. ∎

Consider the variance component vector $\underline{\psi}_{r\times 1} = [\underline{\psi}'_{1(r_1\times 1)}\, \underline{\psi}'_{2(r_2\times 1)}]'$ where there is interest in $\underline{\psi}_2 = \{\psi_i\}_{r_2\times 1}$. For the purposes of checking $\mathrm{GZC_{FR}}$ for $\underline{\psi}_2$, it is possible to check simultaneously for the vector $\underline{\psi}_2$ or individually for the components ψ_i $i = r_1 + 1, ..., r$. The check described above was presented for the simultaneous case. However, it may be computationally easier to check for the $\mathrm{GZC_{FR}}$

individually. In this case, the check would be applied to ψ_i for $i = r_1 + 1, ..., r$ where ψ_i is a subcomponent of $\underline{\psi}^{(i)}_{r\times 1} = [\underline{\psi}'_{1(r-1\times 1)}\ \psi_i]'$.

Theorem: GZC$_{\text{FR}}$ for $\underline{\psi}_{2(r_2\times 1)} \Leftrightarrow$ GZC$_{\text{FR}}$ for $\psi_i\ \forall\, i = r_1 + 1, ..., r$.

proof: i) Suppose GZC$_{\text{FR}}$ for $\underline{\psi}_2$. Then $\widehat{\underline{\psi}}_{2I}$ is UBLUE for $\underline{\psi}_2$ by the LQZ-UBLUE$_{\text{FR}}$ theorem

$\Rightarrow \underline{\delta}'_i\widehat{\underline{\psi}}_{2I}$ is UBLUE for $\psi_i\ \forall\, i = r_1 + 1, ..., r$ by the Linear Closure Property where

$\underline{\delta}_i = \{\delta_{ij}\} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \Rightarrow$ GZC$_{\text{FR}}$ for $\psi_i\ \forall\ i = r_1 + 1, ..., r$ by the LQZ-UBLUE$_{\text{FR}}$ theorem.

ii) Suppose GZC$_{\text{FR}}$ for $\psi_i\,\forall\, i = r_1 + 1, ..., r$. Then for $i = r_1 + 1, ..., r$

$\widehat{\psi}_{iI}$ is UBLUE for ψ_i by the LQZ-UBLUE$_{\text{FR}}$ theorem

$\Rightarrow\ \{\widehat{\psi}_{iI}\}_{k_2\times 1}$ is UBLUE for $\underline{\psi}_2$ by the Linear Closure Property

$\Rightarrow$ GZC$_{\text{FR}}$ for $\underline{\psi}_{2\,(k_2\times 1)}$ by the LQZ-UBLUE$_{\text{FR}}$ theorem. ∎

The methods presented here for checking the conditions are used to verify the existence of examples that satisfy the GZC$_{\text{FR}}$. Such examples are given in sections 6.6, 6.7, and 6.8. The methods were also used to search for examples among 3-way models which are presented in section 6.9.

6.6. ML Examples: Balanced Models with Random Highest Possible Order Effect

This example assumes complete balance and that the highest possible order effect is random and included in the model. The highest order effect may correspond to an interaction or nested effect. It is necessary to assume that $\underline{\psi}$ is estimable in this setting, or equivalently that $\dim \text{sp}\mathcal{V}^\diamond = k + 1$, in order to use the results of this chapter.

Theorem: Consider the Y-Model under [L], [O], [C], and [N]. Suppose the design is completely balanced and the model contains the highest possible order interaction as a random effect. If $\widehat{\underline{\psi}}_{\text{MLQ}}$ exists and $V_{\widehat{\underline{\psi}}_{\text{MLQ}}}$ is PD, then σ_e^2 has an ELMLQE.

proof: Let $r =$ number of replicates and T be the design matrix associated with the highest order interaction. Let $k + 1$ identify the residual term and k identify the random effect associated with the highest order effect. Note $V_k = rP_T$ due to complete balance.

By last proposition in section 6.4.1 with $m = 1, ..., k$,

$$\sum_{h=1}^{k} a_h \text{tr}(V_hV_m) = \text{tr}(IV_m) \Rightarrow \sum_{h=1}^{k-1} a_h\text{tr}(V_hV_m) + a_k\text{tr}(rP_TV_m) = \text{tr}(V_m)$$

$$\Rightarrow\ \sum_{h=1}^{k-1} a_h\text{tr}(V_hV_m) + ra_k\text{tr}(V_m) = \text{tr}(V_m) \text{ as } T \text{ is associated with the highest order effect}$$

$\Rightarrow\ a_h = \begin{cases} \frac{1}{r} & h = k \\ 0 & h \neq k \end{cases}$ as $\underline{a}$ is uniquely determined since $\underline{\psi}$ is estimable. Thus, $P_{X_1^\circ} I = \frac{1}{r} r P_T = P_T$

$\Rightarrow\ E_e = (I - P_{X_1^\circ})I = I - P_T$. Then

(a) if j or $l \neq e \quad G_{ejl} = V_j(I - P_T)V_l + V_l(I - P_T)V_j = 0 \in \underline{R}(X^\circ)$ as $\underline{R}(V_j) \subset \underline{R}(P_T)$

(b) if $j = l = e \quad G_{eee} = I(I - P_T)I + I(I - P_T)I = 2(I - P_T) = 2V_e - \frac{2}{r}V_k \in \underline{R}(X^\circ)$.

$\therefore$ By the GZC$_{\text{FR}}$ check and the ALQNY-ELMLQE$_{\text{FR}}$, $\exists$ an ELMLQE for σ_e^2 as $G_{jel} \in \underline{R}(X^\circ)\ \forall j,\ l$. ∎

6.7. REML Examples: Random Pseudo Balanced Models

Particular models have ELREMLQEs for the residual variance component under pseudo balance. Pseudo balance is defined in section 3.1.3. Consider a random model that has the highest possible order effect as random and at least one other random effect. Let H denote the matrix associated with the highest possible order effect and let G denote the matrix for any other random effect in the model not including the residual error term. Let V_g, V_h, and V_e denote the covariance matrices associated with the arbitrary random effect, the highest order effect, and the residual term. It is also necessary to assume $\underline{\psi}$ is estimable, or equivalently that $\dim \mathrm{sp}\mathcal{V}^\circ = k + 1$, in order to apply the results of this chapter.

Lemma: For the random model described above under pseudo balance:

i) $\tilde{V}_g = (I - P_{\underline{1}})GG'(I - P_{\underline{1}}) \qquad \tilde{V}_h = r(P_H - P_{\underline{1}}) \qquad \tilde{V}_e = I - P_{\underline{1}}$

ii) $\tilde{V}_g\tilde{V}_h = r\tilde{V}_g \quad \tilde{V}_g\tilde{V}_e = \tilde{V}_g$

iii) $E_e = (I - P_{X_1^\circ})\tilde{V}_e = \tilde{V}_e - \frac{1}{r}\tilde{V}_h$.

proof: i) Note $H'H = rI$ by pseudo balance. Thus,

$\tilde{V}_h = (I - P_{\underline{1}})HH'(I - P_{\underline{1}}) = r(I - P_{\underline{1}})P_H(I - P_{\underline{1}}) = r(P_H - P_{\underline{1}})$

and the other covariance matrices follow from definition under the REML procedure.

ii) $\tilde{V}_g\tilde{V}_h = r(I - P_{\underline{1}})GG'(P_H - P_{\underline{1}}) = r(I - P_{\underline{1}})G(G' - G'P_{\underline{1}}) = r\tilde{V}_g$ by i) and $\underline{R}(G) \subset \underline{R}(H)$.

In addition, $\tilde{V}_g\tilde{V}_e = (I - P_{\underline{1}})GG'(I - P_{\underline{1}}) = \tilde{V}_g$.

iii) From the last proposition in section 6.4.1, $P_{X_1^\circ}\tilde{V}_e = \sum_{m=1}^{k} a_m\tilde{V}_m$ where for $g = 1, .., k$

$\sum_{m=1}^{k} a_m \mathrm{tr}(\tilde{V}_g\tilde{V}_m) = \mathrm{tr}(\tilde{V}_g\tilde{V}_e) = \mathrm{tr}(\tilde{V}_g)$ by ii). Note $\mathrm{tr}(\tilde{V}_g\tilde{V}_h) = r\mathrm{tr}(\tilde{V}_g)$ by ii) $\Rightarrow a_h = \frac{1}{r}$ and $a_g = 0$

for $g = 1, .., k$ as $\underline{a}$ is uniquely determined since $\underline{\psi}$ is estimable. Thus, $E_e = (I - P_{X_1^\circ})\tilde{V}_e = \tilde{V}_e - \frac{1}{r}\tilde{V}_h$. ∎

Theorem: Consider the Y-Model under [L], [O], [C], and [N]. Suppose the design is pseudo balanced and that the model contains the highest possible order interaction as a random effect. If $\hat{\underline{\psi}}_{\text{REMLQ}}$ exists and $V_{\hat{\underline{\psi}}_{\text{REMLQ}}}$ is PD, then σ_e^2 has an ELREMLQE.

proof: Note $G_{jel} = \tilde{V}_j E_e \tilde{V}_l + \tilde{V}_l E_e \tilde{V}_j$. By the lemma,

$\tilde{V}_g E_e = \tilde{V}_g - \frac{1}{r} r \tilde{V}_g = 0 \Rightarrow G_{gel} = 0 \ \ \forall g = 1, .., k, \ \ l = 1, ..., k+1$

$\tilde{V}_e E_e = \tilde{V}_e - \frac{1}{r}\tilde{V}_h = E_e \Rightarrow G_{eee} = 2E_e \in \underline{R}(X^{\circ})$. Thus, $G_{jel} \in \underline{R}(X^{\circ}) \ \forall j, \ l$.

$\therefore$ By the GZC$_{\text{FR}}$ check and the LQNY-ELREMLQE$_{\text{FR}}$ theorem, an ELREMLQE exists for σ_e^2. ∎

6.8. REML Examples: Random Models

This section provides conditions for the existence of ELREMLQEs for variance components in a class of random models. In order to achieve a result that includes all random models in this class, the notation is quite cumbersome. Consider p factors in a classification model with the random vector $\underline{Y}$ indexed as $Y_{\underline{f}^* r}$, where $\underline{f}^*$ is a vector of indices indicating the levels of the factors associated with the observation $Y_{\underline{f}^* r}$. Additonal notation is given below using definitions from section 3.1.3:

Notation: $\mathcal{F}^* = \{1, ..., p\}$ = the complete set of all factors

$\mathbb{F} = \{\mathcal{F} \mid \mathcal{F} \subset \mathcal{F}^*, \mathcal{F}\text{-effects included in model}\}$ = collection of factor subsets of all included effects

$\underline{F} = [f_1, ..., f_u]'$ = vector form of $\mathcal{F} = \{f_1, ..., f_u\}$

$\underline{f} = [i_1, ..., i_u]'$ = vector of indices of levels of factors in $\underline{F}$ where $i_j \in \{1, ..., t_{f_j}\}$ for $j = 1, ..., u$

$t_{\underline{F}} = \prod_{j=1}^{u} t_{f_j}$ = number of levels associated with $\underline{f}$

$\#\mathbb{F}$ = number of elements in $\mathbb{F}$

$\underline{F}' \sqcap \underline{F}$ = vector listing the factors associated with $\mathcal{F}' \cap \mathcal{F}$

$\mathcal{R} = \{1, ..., p, p+1\}$ = factor subset associated with the residual error term

$\underline{R}$ = vector listing residual factor associated with $\mathcal{R}$.

In order to include the residual, $\mathcal{R}$, let $\mathbb{F}^{\dagger} = \mathbb{F} \cup \mathcal{R}$. Now examine a partition of $\mathcal{F}^*$ where $\mathcal{F}^* = \mathcal{I}^* \uplus \mathcal{J}^* \uplus \mathcal{K}^*$, using the symbol $\uplus$ to denote disjoint union. Consider the additional notation and assumptions for the above partition:

Notation: $\mathbb{I} = \{\mathcal{I} \mid \mathcal{I} \subset \mathcal{I}^*, \mathcal{I} \text{ non-empty}, \mathcal{I}\text{-effects included}\}$ $\qquad \mathbb{H} = \mathbb{F}^{\dagger} \backslash \mathbb{I} = \mathbb{I}^c$

$\mathbb{J} = \{\mathcal{J} \mid \mathcal{J} \subset \mathcal{J}^*, \mathcal{J} \text{ non-empty}, \mathcal{J}\text{-effects included}\}$

$\mathbb{K} = \{\mathcal{K} \mid \mathcal{K} \subset \mathcal{K}^*, \mathcal{K} \text{ non-empty}, \mathcal{K}\text{-effects included}\}$

$\mathbb{I} \overset{*}{\sqcup} \mathbb{J} = \{\mathcal{I} \cup \mathcal{J} \mid \mathcal{I} \in \mathbb{I}, \mathcal{J} \in \mathbb{J}\}$ $\qquad \mathbb{J} \overset{*}{\sqcup} \mathbb{K} = \{\mathcal{J} \cup \mathcal{K} \mid \mathcal{J} \in \mathbb{J}, \mathcal{K} \in \mathbb{K}\}$

$\mathbb{I} \sqcup \mathbb{J} = \{\mathcal{I} \cup \mathcal{J} \mid \mathcal{I} \in \mathbb{I}, \mathcal{J} \in \mathbb{J}, (\mathcal{I} \cup \mathcal{J})\text{-effects included}\} = \mathbb{I} \overset{*}{\sqcup} \mathbb{J} \cap \mathbb{F}$

$\mathbb{J} \sqcup \mathbb{K} = \{\mathcal{J} \cup \mathcal{K} \mid \mathcal{J} \in \mathbb{J}, \mathcal{K} \in \mathbb{K}, (\mathcal{J} \cup \mathcal{K})\text{-effects included}\} = \mathbb{J} \overset{*}{\sqcup} \mathbb{K} \cap \mathbb{F}$

$\mathbb{I} \sqcup \mathbb{K} = \{\mathcal{I} \cup \mathcal{K} \mid \mathcal{I} \in \mathbb{I}, \mathcal{K} \in \mathbb{K}, (\mathcal{I} \cup \mathcal{K})\text{-effects included}\}$

$\mathbb{I} \sqcup \mathbb{J} \sqcup \mathbb{K} = \{\mathcal{I} \cup \mathcal{J} \cup \mathcal{K} \mid \mathcal{I} \in \mathbb{I}, \mathcal{J} \in \mathbb{J}, \mathcal{K} \in \mathbb{K}, (\mathcal{I} \cup \mathcal{J} \cup \mathcal{K})\text{-effects included}\}$

Assumptions: i) $\mathbb{I} \sqcup \mathbb{J} = \mathbb{I} \overset{*}{\sqcup} \mathbb{J}$ ii) $\mathbb{I} \sqcup \mathbb{K} = \emptyset$ iii) $\mathbb{I} \sqcup \mathbb{J} \sqcup \mathbb{K} = \emptyset$

iv) It is possible for $\mathbb{K} = \emptyset$ or $\mathbb{J} \sqcup \mathbb{K} = \emptyset$, but not both.

Under the partition, notation, and assumptions, the set of all factors can be represented as $\mathbb{F} = \mathbb{I} \uplus \mathbb{J} \uplus \mathbb{K} \uplus (\mathbb{I} \sqcup \mathbb{J}) \uplus (\mathbb{J} \sqcup \mathbb{K})$. The associated model can now be written as $Y_{\underline{i}^* \underline{j}^* \underline{k}^* r}$ for $r = 1, ..., n_{\underline{i}^* \underline{j}^* \underline{k}^*}$ where:

$$Y_{\underline{i}^* \underline{j}^* \underline{k}^* r} = \mu + \sum_{\mathcal{I} \in \mathbb{I}} a_{\underline{i}}^{\underline{\mathbf{I}}} + \sum_{\mathcal{J} \in \mathbb{J}} b_{\underline{j}}^{\underline{\mathbf{J}}} + \sum_{\mathcal{K} \in \mathbb{K}} c_{\underline{k}}^{\underline{\mathbf{K}}} + \sum_{\mathcal{I} \cup \mathcal{J} \in \mathbb{I} \sqcup \mathbb{J}} (ab)_{\underline{ij}}^{\underline{\mathbf{IJ}}} + \sum_{\mathcal{J} \cup \mathcal{K} \in \mathbb{J} \sqcup \mathbb{K}} (bc)_{\underline{jk}}^{\underline{\mathbf{JK}}} + e_{\underline{i}^* \underline{j}^* \underline{k}^* r}$$

$$\underline{Y} = \underline{1}\mu + \sum_{\mathcal{I} \in \mathbb{I}} Z_{\underline{\mathbf{I}}} \underline{a}^{\mathbf{I}} + \sum_{\mathcal{J} \in \mathbb{J}} Z_{\underline{\mathbf{J}}} \underline{b}^{\mathbf{J}} + \sum_{\mathcal{K} \in \mathbb{K}} Z_{\underline{\mathbf{K}}} \underline{c}^{\mathbf{K}} + \sum_{\mathcal{I} \cup \mathcal{J} \in \mathbb{I} \sqcup \mathbb{J}} Z_{\underline{\mathbf{IJ}}} (\underline{ab})^{\mathbf{IJ}} + \sum_{\mathcal{J} \cup \mathcal{K} \in \mathbb{J} \sqcup \mathbb{K}} Z_{\underline{\mathbf{JK}}} (\underline{ab})^{\mathbf{JK}} + \underline{e} \ .$$

In addition, define $N_X = I - P_{\underline{1}}$, $\tilde{V}_{\underline{\mathbf{F}}} = N_X Z_{\underline{\mathbf{F}}} Z_{\underline{\mathbf{F}}}' N_X = N_X V_{\underline{\mathbf{F}}} N_X$, and denote the incidence matrix by $N = \{n[\underline{i}^* \underline{j}^* \underline{k}^*]\} = \{n[\underline{ii^c}\, \underline{jj^c}\, \underline{kk^c}]\}$. If the incidence matrix is summed over $\underline{f}^*$, then its position will be replaced by '$*$'. If the incidence matrix is summed over a subvector of $\underline{f}^c$, then its position will be replaced by '$\cdot$'. The following lemmas will be used to show that if the set $\mathbb{J}$ has a dominating factor, then $\mathrm{Bal}(\mathbb{I} \sqcup \mathbb{J})$ and $\mathrm{Bal}(\mathbb{I} \,|\, \mathbb{J} \overset{*}{\sqcup} \mathbb{K})$ implies there exists an ELREMLQE for $\underline{\psi}_{\mathbb{I}}$ or for all variance components in the $\mathbb{I}$-set. It is necessary to assume that $\underline{\psi}$ is estimable, or equivalently, that the $\tilde{V}_{\underline{\mathbf{F}}}$'s are linearly independent $\forall\, \mathcal{F} \in \mathbb{F}$.

Trace Formula for Quadratic Expression: Let $C = \{c_{ij}\}, D = \mathrm{diag}(\{d_{ii}\})$, then $\mathrm{tr}(C'DC) = \sum_i \sum_j d_{ii} c_{ij}^2$.

proof: Note $(C'DC)_{jj'} = \sum_i c_{ij}\, d_{ii} c_{ij'} = \sum_i d_{ii} c_{ij}\, c_{ij'}$. ■

Lemma 1: i) $\mathrm{Bal}(\mathbb{I} \sqcup \mathbb{J}) \Leftrightarrow n[\underline{i} \cdot \underline{j} \cdot *] = n[\underline{1} \cdot \underline{1} \cdot *] \quad \forall\, \mathcal{I} \in \mathbb{I}, \mathcal{J} \in \mathbb{J}$.

ii) If $\mathrm{Bal}(\mathbb{I} \sqcup \mathbb{J})$, then $\forall\, \mathcal{I} \in \mathbb{I}, \mathcal{J} \in \mathbb{J}$

$$\begin{aligned} \tilde{V}_{\underline{\mathbf{I}}} &= m_{\underline{\mathbf{I}}}(P_{\underline{\mathbf{I}}} - P_{\underline{1}}) & m_{\underline{\mathbf{I}}} &= n[\underline{1} \cdot * *] \\ \tilde{V}_{\underline{\mathbf{J}}} &= m_{\underline{\mathbf{J}}}(P_{\underline{\mathbf{J}}} - P_{\underline{1}}) & m_{\underline{\mathbf{J}}} &= n[* \underline{1} \cdot *] \\ \tilde{V}_{\underline{\mathbf{IJ}}} &= m_{\underline{\mathbf{IJ}}}(P_{\underline{\mathbf{IJ}}} - P_{\underline{1}}) & m_{\underline{\mathbf{IJ}}} &= n[\underline{1} \cdot \underline{1} \cdot *] \quad \text{where} \quad m_{\underline{\mathbf{I}}} = t_{\underline{\mathbf{J}}} m_{\underline{\mathbf{IJ}}} m_{\underline{\mathbf{J}}} = t_{\underline{\mathbf{I}}} m_{\underline{\mathbf{IJ}}} . \end{aligned}$$

proof: i) $\mathrm{Bal}(\mathbb{I} \sqcup \mathbb{J}) \Leftrightarrow \mathrm{Bal}(\mathcal{I} \cup \mathcal{J})\ \forall\, \mathcal{I} \in \mathbb{I}, \mathcal{J} \in \mathbb{J}$

$\Leftrightarrow \forall\, \mathcal{I} \in \mathbb{I}, \mathcal{J} \in \mathbb{J} \quad n[\underline{i} \cdot \underline{j} \cdot *] = \sum_{\mathcal{I}' \in \mathbb{I} \setminus \mathcal{I}} \sum_{\mathcal{J}' \in \mathbb{J} \setminus \mathcal{J}} \sum_{\mathcal{K} \in \mathbb{K}} n[\underline{i}^* \underline{j}^* \underline{k}^*] = n[\underline{1} \cdot \underline{1} \cdot *]$.

ii) $\mathrm{Bal}(\mathbb{I} \sqcup \mathbb{J}) \Leftrightarrow \mathrm{Bal}(\mathcal{I} \cup \mathcal{J}) \quad \forall\, \mathcal{I} \in \mathbb{I},\ \mathcal{J} \in \mathbb{J}$

$\Rightarrow\ V_{\underline{\mathbf{I}}} = m_{\underline{\mathbf{I}}} P_{\underline{\mathbf{I}}}, V_{\underline{\mathbf{J}}} = m_{\underline{\mathbf{J}}} P_{\underline{\mathbf{J}}}\,, V_{\underline{\mathbf{IJ}}} = m_{\underline{\mathbf{IJ}}} P_{\underline{\mathbf{IJ}}}$ where $m_{\underline{\mathbf{IJ}}} = n[\underline{i} \cdot \underline{j} \cdot *] = n[\underline{1} \cdot \underline{1} \cdot *]$

$m_{\underline{\mathbf{I}}} = n[\underline{i} \cdot * *] = n[\underline{1} \cdot * *] = t_{\underline{\mathbf{J}}}\, n[\underline{1} \cdot \underline{1} \cdot *] = t_{\underline{\mathbf{J}}}\, m_{\underline{\mathbf{IJ}}}$

$m_{\underline{\mathbf{J}}} = n[* \underline{j} \cdot *] = n[* \underline{1} \cdot *] = t_{\underline{\mathbf{I}}}\, n[\underline{1} \cdot \underline{1} \cdot *] = t_{\underline{\mathbf{I}}}\, m_{\underline{\mathbf{IJ}}}$. ■

Lemma 2: i) $\mathrm{Bal}(\mathbb{I}\,|\mathbb{J}\overset{*}{\sqcup}\mathbb{K}) \Leftrightarrow n[\underline{\imath}\cdot\underline{\jmath}\cdot\underline{k}\cdot\,] = n[\underline{1}\cdot\underline{\jmath}\cdot\underline{k}\cdot\,] \;\; \forall\, \mathcal{I}\in\mathbb{I}, \mathcal{J}\in\mathbb{J}, \mathcal{K}\in\mathbb{K}$.

ii) $\mathrm{Bal}(\mathbb{I}\,|\mathbb{J}\overset{*}{\sqcup}\mathbb{J}\overset{*}{\sqcup}\mathbb{K}) \Leftrightarrow n[\underline{\imath}\cdot\underline{\jmath}^a\underline{\jmath}^\circ\underline{\jmath}^b\cdot\underline{k}\cdot\,] = n[\underline{1}\cdot\underline{\jmath}^a\underline{\jmath}^\circ\underline{\jmath}^b\cdot\underline{k}\cdot\,]$

$\forall\, \mathcal{I}\in\mathbb{I}, \mathcal{J}\in\mathbb{J}$ with $\underline{\jmath}=\underline{\jmath}^a\underline{\jmath}^\circ$, $\mathcal{J}'\in\mathbb{J}$ with $\underline{\jmath}'=\underline{\jmath}^\circ\underline{\jmath}^b$, $\mathcal{K}\in\mathbb{K}$.

iii) $\mathrm{Bal}(\mathbb{I}\,|\mathbb{J}\overset{*}{\sqcup}\mathbb{J}\overset{*}{\sqcup}\mathbb{K}) \Rightarrow \mathrm{Bal}(\mathbb{I}\,|\mathbb{J}\overset{*}{\sqcup}\mathbb{K})$.

iv) If $\exists\, \mathcal{J}^d\in\mathbb{J} \ni \mathcal{J}\subset\mathcal{J}^d \;\; \forall\, \mathcal{J}\in\mathbb{J}$ (D), then $\mathrm{Bal}(\mathbb{I}\,|\mathbb{J}\overset{*}{\sqcup}\mathbb{J}\overset{*}{\sqcup}\mathbb{K}) \Leftrightarrow \mathrm{Bal}(\mathbb{I}\,|\mathbb{J}\overset{*}{\sqcup}\mathbb{K})$.

proof: i) $\mathrm{Bal}(\mathbb{I}\,|\mathbb{J}\overset{*}{\sqcup}\mathbb{K}) \Leftrightarrow \mathrm{Bal}(\mathcal{I}|\mathcal{J}\cup\mathcal{K}) \;\; \forall\, \mathcal{I}\in\mathbb{I}, \mathcal{J}\in\mathbb{J}, \mathcal{K}\in\mathbb{K}$

$\Leftrightarrow \forall\, \mathcal{I}\in\mathbb{I}, \mathcal{J}\in\mathbb{J}, \mathcal{K}\in\mathbb{K} \;\; n[\underline{\imath}\cdot\underline{\jmath}\cdot\underline{k}\cdot\,] = \sum\limits_{\mathcal{I}'\in\mathbb{I}\setminus\mathcal{I}}\sum\limits_{\mathcal{J}'\in\mathbb{J}\setminus\mathcal{J}}\sum\limits_{\mathcal{K}'\in\mathbb{K}\setminus\mathcal{K}} n[\underline{\imath}^*\underline{\jmath}^*\underline{k}^*] = n[\underline{1}\cdot\underline{\jmath}\cdot\underline{k}\cdot\,]$.

ii) $\mathrm{Bal}(\mathbb{I}\,|\mathbb{J}\overset{*}{\sqcup}\mathbb{J}\overset{*}{\sqcup}\mathbb{K}) \Leftrightarrow \mathrm{Bal}(\mathcal{I}|\mathcal{J}\cup\mathcal{J}'\cup\mathcal{K}) \;\; \forall\, \mathcal{I}\in\mathbb{I}, \mathcal{J}\in\mathbb{J}, \mathcal{J}'\in\mathbb{J}, \mathcal{K}\in\mathbb{K}$

$\Leftrightarrow \forall\, \mathcal{I}\in\mathbb{I}, \mathcal{J}\in\mathbb{J}, \mathcal{J}'\in\mathbb{J}, \mathcal{K}\in\mathbb{K}$

$n[\underline{\imath}\cdot\underline{\jmath}^a\underline{\jmath}^\circ\underline{\jmath}^b\cdot\underline{k}\cdot\,] = \sum\limits_{\mathcal{I}'\in\mathbb{I}\setminus\mathcal{I}}\sum\limits_{\mathcal{J}^\dagger\in\mathbb{J}\setminus\mathcal{J}\cup\mathcal{J}'}\sum\limits_{\mathcal{K}'\in\mathbb{K}\setminus\mathcal{K}} n[\underline{\imath}^*\underline{\jmath}^*\underline{k}^*] = n[\underline{1}\cdot\underline{\jmath}^a\underline{\jmath}^\circ\underline{\jmath}^b\cdot\underline{k}\cdot\,]$.

iii) By ii), $\forall\, \mathcal{I}\in\mathbb{I}, \mathcal{J}\in\mathbb{J}$ with $\underline{\jmath}=\underline{\jmath}^a\underline{\jmath}^\circ$, $\mathcal{J}'\in\mathbb{J}$ with $\underline{\jmath}'=\underline{\jmath}^\circ\underline{\jmath}^b$, $\mathcal{K}\in\mathbb{K}$

$\mathrm{Bal}(\mathbb{I}\,|\mathbb{J}\overset{*}{\sqcup}\mathbb{J}\overset{*}{\sqcup}\mathbb{K}) \Leftrightarrow n[\underline{\imath}\cdot\underline{\jmath}^a\underline{\jmath}^\circ\underline{\jmath}^b\cdot\underline{k}\cdot\,] = n[\underline{1}\cdot\underline{\jmath}^a\underline{\jmath}^\circ\underline{\jmath}^b\cdot\underline{k}\cdot\,] \;\; \forall\, \mathcal{I}\in\mathbb{I}, \mathcal{J}\in\mathbb{J}, \mathcal{K}\in\mathbb{K}$

$\Rightarrow n[\underline{\imath}\cdot\underline{\jmath}\cdot\underline{k}\cdot\,] = \sum\limits_{\mathcal{J}'\in\mathbb{J}\setminus\mathcal{J}} n[\underline{\imath}\cdot\underline{\jmath}^a\underline{\jmath}^\circ\underline{\jmath}^b\cdot\underline{k}\cdot\,] = n[\underline{\imath}\cdot\underline{\jmath}^a\underline{\jmath}^\circ\cdot\underline{k}\cdot\,] = n[\underline{1}\cdot\underline{\jmath}\cdot\underline{k}\cdot\,] \;\; \forall\, \mathcal{I}\in\mathbb{I}, \mathcal{J}\in\mathbb{J}, \mathcal{K}\in\mathbb{K}$

$\Leftrightarrow \mathrm{Bal}(\mathbb{I}\,|\mathbb{J}\overset{*}{\sqcup}\mathbb{K})$ by i).

iv) $\mathrm{Bal}(\mathbb{I}\,|\mathbb{J}\overset{*}{\sqcup}\mathbb{J}\overset{*}{\sqcup}\mathbb{K}) \Leftrightarrow \mathrm{Bal}(\mathcal{I}|\mathcal{J}\cup\mathcal{J}'\cup\mathcal{K}) \;\; \forall\, \mathcal{I}\in\mathbb{I}, \mathcal{J}\in\mathbb{J}, \mathcal{J}'\in\mathbb{J}, \mathcal{K}\in\mathbb{K}$

$\overset{(D)}{\Leftrightarrow} \mathrm{Bal}(\mathcal{I}|\mathcal{J}\cup\mathcal{J}^d\cup\mathcal{K}) \;\; \forall\, \mathcal{I}\in\mathbb{I}, \mathcal{J}\in\mathbb{J}, \mathcal{K}\in\mathbb{K} \overset{(D)}{\Leftrightarrow} \mathrm{Bal}(\mathcal{I}|\mathcal{J}^d\cup\mathcal{K}) \;\; \forall\, \mathcal{I}\in\mathbb{I}, \mathcal{K}\in\mathbb{K}$

$\overset{(D)}{\Leftrightarrow} \mathrm{Bal}(\mathcal{I}|\mathcal{J}\cup\mathcal{K}) \;\; \forall\, \mathcal{I}\in\mathbb{I}, \mathcal{J}\in\mathbb{J}, \mathcal{K}\in\mathbb{K} \Leftrightarrow \mathrm{Bal}(\mathbb{I}\,|\mathbb{J}\overset{*}{\sqcup}\mathbb{K})$. ■

Lemma 3: i) $\mathrm{Bal}(\mathbb{I}\sqcup\mathbb{J}), \mathrm{Bal}(\mathbb{I}\,|\mathbb{J}\overset{*}{\sqcup}\mathbb{K}) \Rightarrow \tilde{V}_{\underline{\mathrm{K}}}\tilde{V}_{\underline{\mathrm{I}}} = 0 \;\; \forall\, \mathcal{I}\in\mathbb{I}, \mathcal{K}\in\mathbb{K}$.

ii) $\mathrm{Bal}(\mathbb{I}\sqcup\mathbb{J}), \mathrm{Bal}(\mathbb{I}\,|\mathbb{J}\overset{*}{\sqcup}\mathbb{K}) \Rightarrow \tilde{V}_{\underline{\mathrm{JK}}}\tilde{V}_{\underline{\mathrm{I}}} = 0 \;\; \forall\, \mathcal{I}\in\mathbb{I}, \mathcal{J}\cup\mathcal{K}\in\mathbb{J}\sqcup\mathbb{K}$.

iii) $\mathrm{Bal}(\mathbb{I}\sqcup\mathbb{J}), \mathrm{Bal}(\mathbb{I}\,|\mathbb{J}\overset{*}{\sqcup}\mathbb{K}) \Rightarrow \tilde{V}_{\underline{\mathrm{K}}}(\tilde{V}_{\underline{\mathrm{J}}} - t_{\mathrm{I}}\tilde{V}_{\underline{\mathrm{IJ}}}) = 0 \;\; \forall\, \mathcal{I}\in\mathbb{I}, \mathcal{J}\in\mathbb{J}, \mathcal{K}\in\mathbb{K}$.

iv) $\mathrm{Bal}(\mathbb{I}\sqcup\mathbb{J}), \mathrm{Bal}(\mathbb{I}\,|\mathbb{J}\overset{*}{\sqcup}\mathbb{J}\overset{*}{\sqcup}\mathbb{K}) \Rightarrow \tilde{V}_{\underline{\mathrm{J'K}}}\tilde{V}_{\underline{\mathrm{J}}} - t_{\mathrm{I}}\tilde{V}_{\underline{\mathrm{J'K}}}\tilde{V}_{\underline{\mathrm{IJ}}} = 0 \;\; \forall\, \mathcal{I}\in\mathbb{I}, \mathcal{J}\in\mathbb{J}, \mathcal{J}'\cup\mathcal{K}\in\mathbb{J}\sqcup\mathbb{K}$.

proof: i) (1) Note $\mathrm{tr}(Z'_{\underline{\mathrm{K}}}(P_{\underline{\mathrm{I}}} - P_{\underline{1}})Z_{\underline{\mathrm{K}}}) = \mathrm{tr}(Z'_{\underline{\mathrm{K}}}Z_{\underline{\mathrm{I}}}(Z'_{\underline{\mathrm{I}}}Z_{\underline{\mathrm{I}}})^{-1}Z'_{\underline{\mathrm{I}}}Z_{\underline{\mathrm{K}}} - Z'_{\underline{\mathrm{K}}}\underline{1}(\underline{1}'\underline{1})^{-1}\underline{1}'Z_{\underline{\mathrm{K}}})$

$= \mathrm{tr}(\{n[\underline{\imath}\cdot *\,\underline{k}\cdot\,]\}'^{(\underline{k},\underline{\imath})}_{t_{\mathrm{K}}\times t_{\mathrm{I}}}\mathrm{diag}(\{\tfrac{1}{n[\underline{\imath}\cdot **]}\})\{n[\underline{\imath}\cdot *\,\underline{k}\cdot\,]\}^{(\underline{\imath},\underline{k})}_{t_{\mathrm{I}}\times t_{\mathrm{K}}} - \tfrac{1}{n[***]}\{n[**\,\underline{k}\cdot\,]\}'^{(\underline{k})}_{t_{\mathrm{K}}\times 1}\{n[**\,\underline{k}\cdot\,]\}^{(\underline{k})}_{1\times t_{\mathrm{K}}})$

$= \sum\limits_{\underline{k}=1}^{t_{\mathrm{K}}}(\sum\limits_{\underline{\imath}=1}^{t_{\mathrm{I}}}\frac{n^2[\underline{\imath}\cdot *\,\underline{k}\cdot]}{n[\underline{\imath}\cdot **]} - \frac{n^2[**\,\underline{k}\cdot]}{n[***]})] = \sum\limits_{\underline{k}=1}^{t_{\mathrm{K}}}(\sum\limits_{\underline{\imath}=1}^{t_{\mathrm{I}}}\frac{n^2[\underline{1}\cdot *\,\underline{k}\cdot]}{n[\underline{1}\cdot **]} - \frac{t_{\mathrm{I}}^2 n^2[\underline{1}\cdot *\,\underline{k}\cdot]}{t_{\mathrm{I}}n[\underline{1}\cdot **]}) = 0$ by lemmas 1 and 2

$\Rightarrow Z'_{\underline{\mathrm{K}}}(P_{\underline{\mathrm{I}}} - P_{\underline{1}})Z_{\underline{\mathrm{K}}} = 0$ since NND $\Rightarrow Z'_{\underline{\mathrm{K}}}(P_{\underline{\mathrm{I}}} - P_{\underline{1}}) = 0$.

(2) By lemma 1 iii) $\forall\, \mathcal{I}\in\mathbb{I}, \mathcal{K}\in\mathbb{K}$,

$\tilde{V}_{\underline{\mathrm{K}}}\tilde{V}_{\underline{\mathrm{I}}} = (I - P_{\underline{1}})Z_{\underline{\mathrm{K}}}Z'_{\underline{\mathrm{K}}}(I - P_{\underline{1}})m_{\mathrm{I}}(P_{\underline{\mathrm{I}}} - P_{\underline{1}}) = m_{\mathrm{I}}(I - P_{\underline{1}})Z_{\underline{\mathrm{K}}}Z'_{\underline{\mathrm{K}}}(P_{\underline{\mathrm{I}}} - P_{\underline{1}}) = 0$ by (1).

ii) (1) Note $\mathrm{tr}(Z'_{\underline{\mathrm{JK}}}(P_{\underline{\mathrm{I}}} - P_{\underline{1}})Z_{\underline{\mathrm{JK}}}) = \mathrm{tr}(Z'_{\underline{\mathrm{JK}}}Z_{\underline{\mathrm{I}}}(Z'_{\underline{\mathrm{I}}}Z_{\underline{\mathrm{I}}})^{-1}Z'_{\underline{\mathrm{I}}}Z_{\underline{\mathrm{JK}}} - Z'_{\underline{\mathrm{JK}}}\underline{1}(\underline{1}'\underline{1})^{-1}\underline{1}'Z_{\underline{\mathrm{JK}}})$

$= \mathrm{tr}(\{n[\underline{\imath}\cdot\underline{\jmath}\cdot\underline{k}\cdot\,]\}'^{(\underline{\jmath}\underline{k},\underline{\imath})}_{t_{\mathrm{J}}t_{\mathrm{K}}\times t_{\mathrm{I}}}\mathrm{diag}(\{\tfrac{1}{n[\underline{\imath}\cdot **]}\})\{n[\underline{\imath}\cdot\underline{\jmath}\cdot\underline{k}\cdot\,]\}^{(\underline{\imath},\underline{\jmath}\underline{k})}_{t_{\mathrm{I}}\times t_{\mathrm{J}}t_{\mathrm{K}}}$

$- \tfrac{1}{n[****]}\{n[*\,\underline{\jmath}\cdot\underline{k}\cdot\,]\}'^{(\underline{\jmath}\underline{k})}_{t_{\mathrm{J}}t_{\mathrm{K}}\times 1}\{n[*\,\underline{\jmath}\cdot\underline{k}\cdot\,]\}^{(\underline{\jmath}\underline{k})}_{1\times t_{\mathrm{J}}t_{\mathrm{K}}})$

$$= \sum_{\jmath=1}^{t_{\mathrm{J}}} \sum_{\underline{k}=1}^{t_{\underline{\mathrm{K}}}} \Big(\sum_{\underline{i}=1}^{t_{\mathrm{I}}} \tfrac{n^2[\underline{i}\cdot\jmath\cdot\underline{k}\cdot]}{n[\underline{i}\cdot**]} - \tfrac{n^2[*\,\jmath\cdot\underline{k}\cdot]}{n[****]} \Big) = \sum_{\jmath=1}^{t_{\mathrm{J}}} \sum_{\underline{k}=1}^{t_{\underline{\mathrm{K}}}} \Big(\sum_{\underline{i}=1}^{t_{\mathrm{I}}} \tfrac{n^2[\underline{1}\cdot\jmath\cdot\underline{k}\cdot]}{n[\underline{1}\cdot**]} - \tfrac{t_{\mathrm{I}}^2 n^2[\underline{1}\cdot\jmath\cdot\underline{k}\cdot]}{t_{\mathrm{I}}\, n[****]} \Big) = 0 \text{ by lemmas 1 and 2}$$

$\Rightarrow Z'_{\underline{\mathrm{JK}}}(P_{\underline{\mathrm{I}}} - P_{\underline{1}})Z_{\underline{\mathrm{JK}}} = 0$ since NND $\Rightarrow Z'_{\underline{\mathrm{JK}}}(P_{\underline{\mathrm{I}}} - P_{\underline{1}}) = 0.$

(2) By lemma 1 iii) $\forall\, \mathcal{I} \in \mathbb{I},\ \mathcal{J} \cup \mathcal{K} \in \mathbb{J} \sqcup \mathbb{K}$,

$\tilde{V}_{\underline{\mathrm{JK}}}\tilde{V}_{\underline{\mathrm{I}}} = m_{\underline{\mathrm{I}}}(I - P_{\underline{1}})Z_{\underline{\mathrm{JK}}}Z'_{\underline{\mathrm{JK}}}(I - P_{\underline{1}})(P_{\underline{\mathrm{I}}} - P_{\underline{1}}) = m_{\underline{\mathrm{I}}}(I - P_{\underline{1}})Z_{\underline{\mathrm{JK}}}Z'_{\underline{\mathrm{JK}}}(P_{\underline{\mathrm{I}}} - P_{\underline{1}}) = 0$ by (1) .

iii) (1) Note $\mathrm{tr}(Z'_{\underline{\mathrm{K}}}(P_{\underline{\mathrm{IJ}}} - P_{\underline{\mathrm{J}}})Z_{\underline{\mathrm{K}}}) = \mathrm{tr}(Z'_{\underline{\mathrm{K}}}Z_{\underline{\mathrm{IJ}}}(Z'_{\underline{\mathrm{IJ}}}Z_{\underline{\mathrm{IJ}}})^{-1}Z'_{\underline{\mathrm{IJ}}}Z_{\underline{\mathrm{K}}} - Z'_{\underline{\mathrm{K}}}Z_{\underline{\mathrm{J}}}(Z'_{\underline{\mathrm{J}}}Z_{\underline{\mathrm{J}}})^{-1}Z'_{\underline{\mathrm{J}}}Z_{\underline{\mathrm{K}}})$

$$= \mathrm{tr}(\{n[\underline{i}\cdot\jmath\cdot\underline{k}\cdot]\}'^{(\underline{k},\underline{i}\jmath)}_{t_{\underline{\mathrm{K}}}\times t_{\mathrm{I}}t_{\mathrm{J}}}\mathrm{diag}(\{\tfrac{1}{n[\underline{i}\cdot\jmath\cdot*]}\})\{n[\underline{i}\cdot\jmath\cdot\underline{k}\cdot]\}^{(\underline{i}\jmath,\underline{k})}_{t_{\mathrm{I}}t_{\mathrm{J}}\times t_{\underline{\mathrm{K}}}}$$

$$- \{n[*\,\jmath\cdot\underline{k}\cdot]\}'^{(\underline{k},\jmath)}_{t_{\underline{\mathrm{K}}}\times t_{\mathrm{J}}}\mathrm{diag}\{\tfrac{1}{n[*\,\jmath\cdot*]}\}\{n[*\,\jmath\cdot\underline{k}\cdot]\}^{(\jmath,\underline{k})}_{t_{\mathrm{J}}\times t_{\underline{\mathrm{K}}}})$$

$$= \sum_{\jmath=1}^{t_{\mathrm{J}}} \sum_{\underline{k}=1}^{t_{\underline{\mathrm{K}}}} \Big(\sum_{\underline{i}=1}^{t_{\mathrm{I}}} \tfrac{n^2[\underline{i}\cdot\jmath\cdot\underline{k}\cdot]}{n[\underline{i}\cdot\jmath\cdot*]} - \tfrac{n^2[*\,\jmath\cdot\underline{k}\cdot]}{n[*\,\jmath\cdot*]} \Big) = \sum_{\jmath=1}^{t_{\mathrm{J}}} \sum_{\underline{k}=1}^{t_{\underline{\mathrm{K}}}} \Big(\sum_{\underline{i}=1}^{t_{\mathrm{I}}} \tfrac{n^2[\underline{1}\cdot\jmath\cdot\underline{k}\cdot]}{n[\underline{1}\cdot\underline{1}\cdot*]} - \tfrac{t_{\mathrm{I}}^2 n^2[\underline{1}\cdot\jmath\cdot\underline{k}\cdot]}{t_{\mathrm{I}}\, n[\underline{1}\cdot\underline{1}\cdot*]} \Big) = 0 \quad \text{by lemmas 1 and 2}$$

$\Rightarrow Z'_{\underline{\mathrm{K}}}(P_{\underline{\mathrm{IJ}}} - P_{\underline{\mathrm{J}}})Z_{\underline{\mathrm{K}}} = 0$ since NND $\Rightarrow Z'_{\underline{\mathrm{K}}}(P_{\underline{\mathrm{IJ}}} - P_{\underline{\mathrm{J}}}) = 0 \Rightarrow Z'_{\underline{\mathrm{K}}}(P_{\underline{\mathrm{J}}} - P_{\underline{\mathrm{IJ}}}) = 0$.

(2) By lemma 1 iii) $\forall\, \mathcal{I} \in \mathbb{I},\ \mathcal{J} \in \mathbb{J},\ \mathcal{K} \in \mathbb{K}$,

$\tilde{V}_{\underline{\mathrm{K}}}(\tilde{V}_{\underline{\mathrm{J}}} - t_{\mathrm{I}}\tilde{V}_{\underline{\mathrm{IJ}}}) = \tilde{V}_{\underline{\mathrm{K}}}(m_{\underline{\mathrm{J}}}(P_{\underline{\mathrm{J}}} - P_{\underline{1}}) - t_{\mathrm{I}}m_{\underline{\mathrm{IJ}}}(P_{\underline{\mathrm{IJ}}} - P_{\underline{1}})) = m_{\underline{\mathrm{J}}}\tilde{V}_{\underline{\mathrm{K}}}(P_{\underline{\mathrm{J}}} - P_{\underline{\mathrm{IJ}}})$

$= m_{\underline{\mathrm{J}}}(I - P_{\underline{1}})Z_{\underline{\mathrm{K}}}Z'_{\underline{\mathrm{K}}}(I - P_{\underline{1}})(P_{\underline{\mathrm{J}}} - P_{\underline{\mathrm{IJ}}}) = m_{\underline{\mathrm{J}}}(I - P_{\underline{1}})Z_{\underline{\mathrm{K}}}Z'_{\underline{\mathrm{K}}}(P_{\underline{\mathrm{J}}} - P_{\underline{\mathrm{IJ}}}) = 0$ by (1).

iv) (1) For $\mathcal{J} \in \mathbb{J}$, let $\jmath = \jmath^a\jmath^\circ$ and for $\mathcal{J}' \in \mathbb{J}$, let $\jmath' = \jmath^\circ\jmath^b$. Note

$\mathrm{tr}(Z'_{\underline{\mathrm{J'K}}}(P_{\underline{\mathrm{IJ}}} - P_{\underline{\mathrm{J}}})Z_{\underline{\mathrm{J'K}}}) = \mathrm{tr}(Z'_{\underline{\mathrm{J'K}}}Z_{\underline{\mathrm{IJ}}}(Z'_{\underline{\mathrm{IJ}}}Z_{\underline{\mathrm{IJ}}})^{-1}Z'_{\underline{\mathrm{IJ}}}Z_{\underline{\mathrm{J'K}}} - Z'_{\underline{\mathrm{J'K}}}Z_{\underline{\mathrm{J}}}(Z'_{\underline{\mathrm{J}}}Z_{\underline{\mathrm{J}}})^{-1}Z'_{\underline{\mathrm{J}}}Z_{\underline{\mathrm{J'K}}})$

$$= \mathrm{tr}(\{\delta_{\jmath^{\circ\prime}\jmath^\circ}n[\underline{i}\cdot\jmath^a\jmath^\circ\jmath^b\cdot\underline{k}\cdot]\}'^{(\jmath'\underline{k},\underline{i}\jmath)}_{t_{\mathrm{J}'}t_{\underline{\mathrm{K}}}\times t_{\mathrm{I}}t_{\mathrm{J}}}\mathrm{diag}(\{\tfrac{1}{n[\underline{i}\cdot\jmath\cdot*]}\})\{\delta_{\jmath^{\circ\prime}\jmath^\circ}n[\underline{i}\cdot\jmath^a\jmath^\circ\jmath^b\cdot\underline{k}\cdot]\}^{(\underline{i}\jmath,\jmath'\underline{k})}_{t_{\mathrm{I}}t_{\mathrm{J}}\times t_{\mathrm{J}'}t_{\underline{\mathrm{K}}}}$$

$$- \{\delta_{\jmath^{\circ\prime}\jmath^\circ}n[*\,\jmath^a\jmath^\circ\jmath^b\cdot\underline{k}\cdot]\}'^{(\jmath'\underline{k},\jmath)}_{t_{\mathrm{J}'}t_{\underline{\mathrm{K}}}\times t_{\mathrm{J}}}\mathrm{diag}(\{\tfrac{1}{n[*\,\jmath\cdot*]}\})\{\delta_{\jmath^{\circ\prime}\jmath^\circ}n[*\,\jmath^a\jmath^\circ\jmath^b\cdot\underline{k}\cdot]\}^{(\jmath,\jmath'\underline{k})}_{t_{\mathrm{J}}\times t_{\mathrm{J}'}t_{\underline{\mathrm{K}}}})$$

$$= \sum_{\jmath=1}^{t_{\mathrm{J}}} \sum_{\jmath'=1}^{t_{\mathrm{J}'}} \sum_{\underline{k}=1}^{t_{\underline{\mathrm{K}}}} \Big(\sum_{\underline{i}=1}^{t_{\mathrm{I}}} \tfrac{\delta_{\jmath^{\circ\prime}\jmath^\circ}n^2[\underline{i}\cdot\jmath^a\jmath^\circ\jmath^b\cdot\underline{k}\cdot]}{n[\underline{i}\cdot\jmath\cdot*]} - \tfrac{\delta_{\jmath^{\circ\prime}\jmath^\circ}n[*\,\jmath^a\jmath^\circ\jmath^b\cdot\underline{k}\cdot]}{n[*\,\jmath\cdot*]} \Big) = \sum_{\jmath=1}^{t_{\mathrm{J}}} \sum_{\underline{k}=1}^{t_{\underline{\mathrm{K}}}} \Big(\sum_{\underline{i}=1}^{t_{\mathrm{I}}} \tfrac{n^2[\underline{i}\cdot\cdot\jmath^a\jmath^\circ\jmath^b\cdot\underline{k}\cdot]}{n[\underline{i}\cdot\jmath\cdot*]} - \tfrac{n[*\cdot\jmath^a\jmath^\circ\jmath^b\cdot\underline{k}\cdot]}{n[*\,\jmath\cdot*]} \Big)$$

$$= \sum_{\jmath=1}^{t_{\mathrm{J}}} \sum_{\underline{k}=1}^{t_{\underline{\mathrm{K}}}} \Big(\sum_{\underline{i}=1}^{t_{\mathrm{I}}} \tfrac{n^2[\underline{1}\cdot\cdot\jmath^a\jmath^\circ\jmath^b\cdot\underline{k}\cdot]}{n[\underline{1}\cdot\underline{1}\cdot*]} - \tfrac{t_{\mathrm{I}}^2 n[\underline{1}\cdot\jmath^a\jmath^\circ\jmath^b\cdot\underline{k}\cdot]}{t_{\mathrm{I}}\, n[\underline{1}\cdot\underline{1}\cdot*]} \Big) = 0 \text{ by lemma 1, lemma 2, and}$$

using the trace formula for a quadratic expression where the first part of the difference has

$$C' = Z'_{\underline{\mathrm{JK}}}Z_{\underline{\mathrm{IJ}}} = \{c_{(\jmath'\underline{k})(\underline{i}\jmath)}\} = \{\delta_{\jmath^{\circ\prime}\jmath^\circ}n[\underline{i}\cdot\cdot\jmath^a\jmath^\circ\jmath^b\cdot\underline{k}\cdot]\}$$
$$D = (Z'_{\underline{\mathrm{IJ}}}Z_{\underline{\mathrm{IJ}}})^{-1} = \{d_{(\underline{i}\jmath)(\underline{i}\jmath)}\} = \{\tfrac{1}{n[\underline{i}\cdot\jmath\cdot*]}\} \Rightarrow \mathrm{tr}(C'DC) = \sum_i\sum_j d_{ii}c_{ij}^2 \ .$$

Then $Z'_{\underline{\mathrm{J'K}}}(P_{\underline{\mathrm{IJ}}} - P_{\underline{\mathrm{J}}})Z_{\underline{\mathrm{J'K}}} = 0$ since NND $\Rightarrow Z'_{\underline{\mathrm{J'K}}}(P_{\underline{\mathrm{IJ}}} - P_{\underline{\mathrm{J}}}) = 0 \Rightarrow Z'_{\underline{\mathrm{J'K}}}(P_{\underline{\mathrm{J}}} - P_{\underline{\mathrm{IJ}}}) = 0.$

(2) By lemma 1 iii) $\forall\, \mathcal{I} \in \mathbb{I},\ \mathcal{J} \in \mathbb{J},\ \mathcal{J}' \cup \mathcal{K} \in \mathbb{J} \sqcup \mathbb{K}$,

$\tilde{V}_{\underline{\mathrm{J'K}}}(\tilde{V}_{\underline{\mathrm{J}}} - t_{\mathrm{I}}\tilde{V}_{\underline{\mathrm{IJ}}}) = \tilde{V}_{\underline{\mathrm{J'K}}}(m_{\underline{\mathrm{J}}}(P_{\underline{\mathrm{J}}} - P_{\underline{1}}) - t_{\mathrm{I}}m_{\underline{\mathrm{IJ}}}(P_{\underline{\mathrm{IJ}}} - P_{\underline{1}}))$

$= m_{\underline{\mathrm{J}}}\tilde{V}_{\underline{\mathrm{J'K}}}(P_{\underline{\mathrm{J}}} - P_{\underline{\mathrm{IJ}}}) = m_{\underline{\mathrm{J}}}(I - P_{\underline{1}})Z_{\underline{\mathrm{J'K}}}Z'_{\underline{\mathrm{J'K}}}(I - P_{\underline{1}})(P_{\underline{\mathrm{J}}} - P_{\underline{\mathrm{IJ}}})$

$= m_{\underline{\mathrm{J}}}(I - P_{\underline{1}})Z_{\underline{\mathrm{J'K}}}Z'_{\underline{\mathrm{J'K}}}(P_{\underline{\mathrm{J}}} - P_{\underline{\mathrm{IJ}}}) = 0$ by (1). ■

Lemma 4: Assume Bal($\mathbb{I} \sqcup \mathbb{J}$), Bal($\mathbb{I}\,|\,\mathbb{J} \overset{*}{\sqcup} \mathbb{J} \overset{*}{\sqcup} \mathbb{K}$) and consider $T_{\mathrm{I}}\underline{p}^{(\mathrm{I})} = \underline{c}^{(\mathrm{I})}\ \forall\, \mathcal{I} \in \mathbb{I}$ where $\{\mathrm{tr}(\tilde{V}_{\underline{\mathrm{H}}}\tilde{V}_{\underline{\mathrm{H}'}})\}_{(\#\mathbb{H})\times(\#\mathbb{H})}\{p^{(\mathrm{I})}_{\underline{\mathrm{H}'}}\} = \{\mathrm{tr}(\tilde{V}_{\underline{\mathrm{H}}}\tilde{V}_{\underline{\mathrm{I}}})\}_{(\#\mathbb{H})\times 1}$. Then $\forall\, \mathcal{J} \in \mathbb{J}\ \exists\ \underline{p}^{(\mathrm{I})} \ni$ $T_{\mathrm{I}}\underline{p}^{(\mathrm{I})} = \underline{c}^{(\mathrm{I})}$, $p^{(\mathrm{I})}_{\underline{\mathrm{IJ}'}} = -\,t_{\mathrm{I}}p^{(\mathrm{I})}_{\underline{\mathrm{J}'}}$, and $p^{(\mathrm{I})}_{\underline{\mathrm{H}'}} = 0\ \ \forall\ \underline{\mathrm{H}}' \neq \underline{\mathrm{J}}', \underline{\mathrm{IJ}}'$ provided

(c1) $\{\frac{t_{\mathrm{J}\sqcap\mathrm{J}'}}{t_{\mathrm{J}'}}\}_{(\#\mathbb{J})\times(\#\mathbb{J})}\{p^{(\mathrm{I})}_{\underline{\mathrm{J}'}}\} = \frac{-1}{t_{\mathrm{I}}}\underline{1}_{(\#\mathbb{J})\times 1}$ and (c2) $\sum_{\mathcal{J}'\in\mathbb{J}} p^{(\mathrm{I})}_{\underline{\mathrm{J}'}} = \frac{-1}{t_{\mathrm{I}}}$.

proof: By the definition of $\underline{p}^{(\mathrm{I})}$, $\forall\ \mathcal{I} \in \mathbb{I}, \forall\ \mathcal{H} \in \mathbb{H}\ \sum\limits_{\mathcal{H}' \in \mathbb{H}} \mathrm{tr}(\tilde{V}_{\underline{\mathrm{H}}}\tilde{V}_{\underline{\mathrm{H}}'})p^{(\mathrm{I})}_{\underline{\mathrm{H}}'} = \mathrm{tr}(\tilde{V}_{\underline{\mathrm{H}}}\tilde{V}_{\mathrm{I}})$

$$\Leftrightarrow \sum_{\mathcal{J}' \in \mathbb{J}} \mathrm{tr}(\tilde{V}_{\underline{\mathrm{H}}}\tilde{V}_{\underline{\mathrm{J}}'})p^{(\mathrm{I})}_{\underline{\mathrm{J}}'} + \sum_{\mathcal{I} \cup \mathcal{J}' \in \mathbb{I} \sqcup \mathbb{J}'} \mathrm{tr}(\tilde{V}_{\underline{\mathrm{H}}}\tilde{V}_{\underline{\mathrm{IJ}}'})p^{(\mathrm{I})}_{\underline{\mathrm{IJ}}'} = \mathrm{tr}(\tilde{V}_{\underline{\mathrm{H}}}\tilde{V}_{\mathrm{I}})$$

$$\Leftrightarrow \sum_{\mathcal{J}' \in \mathbb{J}} p^{(\mathrm{I})}_{\underline{\mathrm{J}}'}[\mathrm{tr}(\tilde{V}_{\underline{\mathrm{H}}}\tilde{V}_{\underline{\mathrm{J}}'}) - t_{\mathrm{I}}\mathrm{tr}(\tilde{V}_{\underline{\mathrm{H}}}\tilde{V}_{\underline{\mathrm{IJ}}'})] = \mathrm{tr}(\tilde{V}_{\underline{\mathrm{H}}}\tilde{V}_{\mathrm{I}}) \qquad (*) \quad \text{as } \underline{\mathrm{I}} \text{ stays the same.}$$

The goal is to verify that $\underline{p}^{(\mathrm{I})}$ satisfies $(*)$ $\forall\ \mathcal{H} \in \mathbb{H}$ under (c1) and (c2). The conditions (c1) and (c2) are shown to hold for cases i), ii), iii), but they are obtained using cases iv), v). Thus, consider the following:

i) $\forall\ \mathcal{J} \in \mathbb{J}\ \ \sum\limits_{\mathcal{J}' \in \mathbb{J}} p^{(\mathrm{I})}_{\underline{\mathrm{J}}'}[\mathrm{tr}(\tilde{V}_{\underline{\mathrm{J}}}\tilde{V}_{\underline{\mathrm{J}}'}) - t_{\mathrm{I}}\mathrm{tr}(\tilde{V}_{\underline{\mathrm{J}}}\tilde{V}_{\underline{\mathrm{IJ}}'})]$

$$= \sum_{\mathcal{J}' \in \mathbb{J}} p^{(\mathrm{I})}_{\underline{\mathrm{J}}'}[m_{\underline{\mathrm{J}}}m_{\underline{\mathrm{J}}'}\mathrm{tr}(P_{\underline{\mathrm{J}} \sqcap \underline{\mathrm{J}}'} - P_{\underline{1}}) - t_{\mathrm{I}}m_{\underline{\mathrm{J}}}m_{\underline{\mathrm{IJ}}'}\mathrm{tr}(P_{\underline{\mathrm{J}} \sqcap \underline{\mathrm{J}}'} - P_{\underline{1}})] \quad \text{by lemma 1}$$

$$= \sum_{\mathcal{J}' \in \mathbb{J}} p^{(\mathrm{I})}_{\underline{\mathrm{J}}'}[m_{\underline{\mathrm{J}}}m_{\underline{\mathrm{J}}'}\mathrm{tr}(P_{\underline{\mathrm{J}} \sqcap \underline{\mathrm{J}}'} - P_{\underline{1}}) - m_{\underline{\mathrm{J}}}m_{\underline{\mathrm{J}}'}\mathrm{tr}(P_{\underline{\mathrm{J}} \sqcap \underline{\mathrm{J}}'} - P_{\underline{1}})] \quad = 0 \quad \text{by lemma 1}$$

$$= \mathrm{tr}(\tilde{V}_{\underline{\mathrm{J}}}\tilde{V}_{\mathrm{I}}) \quad \text{by lemma 1.}$$

ii) $\forall\ \mathcal{K} \in \mathbb{K}\ \ \sum\limits_{\mathcal{J}' \in \mathbb{J}} p^{(\mathrm{I})}_{\underline{\mathrm{J}}'}[\mathrm{tr}(\tilde{V}_{\underline{\mathrm{K}}}\tilde{V}_{\underline{\mathrm{J}}'}) - t_{\mathrm{I}}\mathrm{tr}(\tilde{V}_{\underline{\mathrm{K}}}\tilde{V}_{\underline{\mathrm{IJ}}'})] = \sum\limits_{\mathcal{J}' \in \mathbb{J}} p^{(\mathrm{I})}_{\underline{\mathrm{J}}'}[\mathrm{tr}(\tilde{V}_{\underline{\mathrm{K}}}(\tilde{V}_{\underline{\mathrm{J}}} - t_{\mathrm{I}}\tilde{V}_{\underline{\mathrm{IJ}}})] = 0$ by lemma 3 iii)

$$= \mathrm{tr}(\tilde{V}_{\underline{\mathrm{K}}}\tilde{V}_{\mathrm{I}}) \quad \text{by lemma 3 i).}$$

iii) $\forall\ \mathcal{J} \cup \mathcal{K} \in \mathbb{J} \sqcup \mathbb{K}\ \ \ \sum\limits_{\mathcal{J}' \in \mathbb{J}} p^{(\mathrm{I})}_{\underline{\mathrm{J}}'}[\mathrm{tr}(\tilde{V}_{\underline{\mathrm{JK}}}\tilde{V}_{\underline{\mathrm{J}}'}) - t_{\mathrm{I}}\mathrm{tr}(\tilde{V}_{\underline{\mathrm{JK}}}\tilde{V}_{\underline{\mathrm{IJ}}'})$

$$= \sum_{\mathcal{J}' \in \mathbb{J}} p^{(\mathrm{I})}_{\underline{\mathrm{J}}'}[\mathrm{tr}(\tilde{V}_{\underline{\mathrm{JK}}}(\tilde{V}_{\underline{\mathrm{J}}'} - t_{\mathrm{I}}\tilde{V}_{\underline{\mathrm{IJ}}'})] = 0 \quad \text{by lemma 3 iv)}$$

$$= \mathrm{tr}(\tilde{V}_{\underline{\mathrm{JK}}}\tilde{V}_{\mathrm{I}}) \quad \text{by lemma 3 ii).}$$

The next two cases are where the effort is needed to obtain the conditions (c1) and (c2).

iv) $\forall\ \mathcal{I}^+ \cup \mathcal{J} \in \mathbb{I} \sqcup \mathbb{J}\ \ \ \sum\limits_{\mathcal{J}' \in \mathbb{J}} p^{(\mathrm{I})}_{\underline{\mathrm{J}}'}[\mathrm{tr}(\tilde{V}_{\underline{\mathrm{I}^+\mathrm{J}}}\tilde{V}_{\underline{\mathrm{J}}'}) - t_{\mathrm{I}}\mathrm{tr}(\tilde{V}_{\underline{\mathrm{I}^+\mathrm{J}}}\tilde{V}_{\underline{\mathrm{IJ}}'})] = \mathrm{tr}(\tilde{V}_{\underline{\mathrm{I}^+\mathrm{J}}}\tilde{V}_{\mathrm{I}})$

$$\Leftrightarrow \sum_{\mathcal{J}' \in \mathbb{J}} p^{(\mathrm{I})}_{\underline{\mathrm{J}}'}[m_{\underline{\mathrm{I}^+\mathrm{J}}}m_{\underline{\mathrm{J}}'}\mathrm{tr}(P_{\underline{\mathrm{J}} \sqcap \underline{\mathrm{J}}'} - P_{\underline{1}}) - t_{\mathrm{I}}m_{\underline{\mathrm{I}^+\mathrm{J}}}m_{\underline{\mathrm{IJ}}'}\mathrm{tr}(P_{(\mathrm{I}^+ \sqcap \mathrm{I})(\underline{\mathrm{J}} \sqcap \underline{\mathrm{J}}')} - P_{\underline{1}})]$$

$$= m_{\underline{\mathrm{I}^+\mathrm{J}}}m_{\mathrm{I}}\mathrm{tr}(P_{\mathrm{I}^+ \sqcap \underline{\mathrm{I}}} - P_{\underline{1}}) \quad \text{by lemma 1}$$

$$\Leftrightarrow \sum_{\mathcal{J}' \in \mathbb{J}} -p^{(\mathrm{I})}_{\underline{\mathrm{J}}'}m_{\underline{\mathrm{J}}'}[\mathrm{tr}(P_{(\mathrm{I}^+ \sqcap \mathrm{I})(\underline{\mathrm{J}} \sqcap \underline{\mathrm{J}}')} - P_{\underline{\mathrm{J}} \sqcap \underline{\mathrm{J}}'})] = m_{\mathrm{I}}\mathrm{tr}(P_{\mathrm{I}^+ \sqcap \underline{\mathrm{I}}} - P_{\underline{1}})$$

$$\Leftrightarrow \sum_{\mathcal{J}' \in \mathbb{J}} -p^{(\mathrm{I})}_{\underline{\mathrm{J}}'}m_{\underline{\mathrm{J}}'}t_{\underline{\mathrm{J}} \sqcap \underline{\mathrm{J}}'}(t_{\mathrm{I}^+ \sqcap \underline{\mathrm{I}}} - 1) = m_{\mathrm{I}}(t_{\mathrm{I}^+ \sqcap \underline{\mathrm{I}}} - 1) \qquad (\circ)\,.$$

Note $(\circ)$ holds trivially when $t_{\mathrm{I}^+ \sqcap \underline{\mathrm{I}}} = 1$, so consider the case where $t_{\mathrm{I}^+ \sqcap \underline{\mathrm{I}}} > 1$

$$\Rightarrow \sum_{\mathcal{J}' \in \mathbb{J}} -p^{(\mathrm{I})}_{\underline{\mathrm{J}}'}m_{\underline{\mathrm{J}}'}t_{\underline{\mathrm{J}} \sqcap \underline{\mathrm{J}}'} = m_{\mathrm{I}} \Leftrightarrow \sum_{\mathcal{J}' \in \mathbb{J}} -p^{(\mathrm{I})}_{\underline{\mathrm{J}}'}\tfrac{n}{t_{\underline{\mathrm{J}}'}}t_{\underline{\mathrm{J}} \sqcap \underline{\mathrm{J}}'} = \tfrac{n}{t_{\mathrm{I}}} \quad \text{as } n = t_{\mathrm{I}}m_{\mathrm{I}} = t_{\underline{\mathrm{J}}'}m_{\underline{\mathrm{J}}'} \text{ by lemma 1}$$

$$\Leftrightarrow \{\tfrac{t_{\underline{\mathrm{J}} \sqcap \underline{\mathrm{J}}'}}{t_{\underline{\mathrm{J}}'}}\}_{(\#\mathbb{J}) \times (\#\mathbb{J})}\{p^{(\mathrm{I})}_{\underline{\mathrm{J}}'}\} = \tfrac{-1}{t_{\mathrm{I}}}\underline{1}_{(\#\mathbb{J}) \times 1} \qquad \text{(c1)}\,.$$

v) for $\mathcal{R}$, $\sum\limits_{\mathcal{J}' \in \mathbb{J}} p^{(\mathrm{I})}_{\underline{\mathrm{J}}'}[\mathrm{tr}(\tilde{V}_{\underline{\mathrm{R}}}\tilde{V}_{\underline{\mathrm{J}}'}) - t_{\mathrm{I}}\mathrm{tr}(\tilde{V}_{\underline{\mathrm{R}}}\tilde{V}_{\underline{\mathrm{IJ}}'})] = \mathrm{tr}(\tilde{V}_{\underline{\mathrm{R}}}\tilde{V}_{\mathrm{I}})$

$$\Leftrightarrow \sum_{\mathcal{J}' \in \mathbb{J}} p^{(\mathrm{I})}_{\underline{\mathrm{J}}'}[m_{\underline{\mathrm{J}}'}\mathrm{tr}(P_{\underline{\mathrm{J}}'} - P_{\underline{1}}) - t_{\mathrm{I}}m_{\underline{\mathrm{IJ}}'}\mathrm{tr}(P_{\underline{\mathrm{IJ}}'} - P_{\underline{1}})] = m_{\mathrm{I}}\mathrm{tr}(P_{\underline{\mathrm{I}}} - P_{\underline{1}}) \quad \text{by lemma 1}$$

$$\Leftrightarrow \sum_{\mathcal{J}' \in \mathbb{J}} -p^{(\mathrm{I})}_{\underline{\mathrm{J}}'}m_{\underline{\mathrm{J}}'}[\mathrm{tr}(P_{\underline{\mathrm{IJ}}'} - P_{\underline{\mathrm{J}}'})] = m_{\mathrm{I}}\mathrm{tr}(P_{\underline{\mathrm{I}}} - P_{\underline{1}}) \Leftrightarrow \sum_{\mathcal{J}' \in \mathbb{J}} -p^{(\mathrm{I})}_{\underline{\mathrm{J}}'}m_{\underline{\mathrm{J}}'}t_{\underline{\mathrm{J}}'}(t_{\underline{\mathrm{I}}} - 1) = m_{\mathrm{I}}(t_{\underline{\mathrm{I}}} - 1) \quad (\circ\circ).$$

Note $(\circ\circ)$ holds trivially when $t_{\mathrm{I}^+ \sqcap \underline{\mathrm{I}}} = 1$, so consider the case where $t_{\mathrm{I}^+ \sqcap \underline{\mathrm{I}}} > 1 \Rightarrow \sum\limits_{\mathcal{J}' \in \mathbb{J}} -p^{(\mathrm{I})}_{\underline{\mathrm{J}}'}m_{\underline{\mathrm{J}}'}t_{\underline{\mathrm{J}}'} = m_{\mathrm{I}}$

$$\Leftrightarrow \sum_{\mathcal{J}' \in \mathbb{J}} -p^{(\mathrm{I})}_{\underline{\mathrm{J}}'}n = \tfrac{n}{t_{\mathrm{I}}} \Leftrightarrow \sum_{\mathcal{J}' \in \mathbb{J}} p^{(\mathrm{I})}_{\underline{\mathrm{J}}'} = \tfrac{-1}{t_{\mathrm{I}}} \ \text{(c2)} \ \text{as } n = t_{\mathrm{I}}m_{\mathrm{I}} = t_{\underline{\mathrm{J}}'}m_{\underline{\mathrm{J}}'} \text{ by lemma 1.} \ \blacksquare$$

The next lemma is given here for convenience to show that covariance matrices which correspond to the effects in $\mathbb{I}$, $\mathbb{J}$, $\mathbb{I} \sqcup \mathbb{J}$ have a closure property.

Lemma 5: Let $\gamma = \mathrm{sp}\{\mathcal{I} \in \mathbb{I}, \mathcal{J} \in \mathbb{J} \mid \tilde{V}_{\underline{I}}, \tilde{V}_{\underline{J}}, \tilde{V}_{\underline{IJ}}\}$. Under $\mathrm{Bal}(\mathbb{I} \sqcup \mathbb{J})$,
$\tilde{V}_{\underline{I'J'}}\tilde{V}_{\underline{S}} \in \gamma \;\; \forall\, \mathcal{I}' \cup \mathcal{J}' \in \mathbb{I} \sqcup \mathbb{J}\,, \forall\; \tilde{V}_{\underline{S}} \in \gamma\,.$

proof: Since $\tilde{V}_{\underline{S}} \in \gamma \Rightarrow \tilde{V}_{\underline{S}} = k_{\underline{I}}\tilde{V}_{\underline{I}} + k_{\underline{J}}\tilde{V}_{\underline{J}} + k_{\underline{IJ}}\tilde{V}_{\underline{IJ}}$ for some $k_{\underline{I}}, k_{\underline{J}}\,, k_{\underline{IJ}}$. By $\mathrm{Bal}(\mathbb{I} \sqcup \mathbb{J})$,

$$\begin{aligned}\tilde{V}_{\underline{I'J'}}\tilde{V}_{\underline{S}} &= m_{\underline{I'J'}}(k_{\underline{I}}m_{\underline{I}}(P_{\underline{I'}\sqcap\underline{I}} - P_{\underline{1}}) + k_{\underline{J}}m_{\underline{J}}(P_{\underline{J'}\sqcap\underline{J}} - P_{\underline{1}}) + k_{\underline{IJ}}m_{\underline{IJ}}(P_{(\underline{I'}\sqcap\underline{I})(\underline{J'}\sqcap\underline{J})} - P_{\underline{1}})\\ &= l_{\underline{I'}\sqcap\underline{I}}\tilde{V}_{\underline{I'}\sqcap\underline{I}} + l_{\underline{J'}\sqcap\underline{J}}\tilde{V}_{\underline{J'}\sqcap\underline{J}} + l_{(\underline{I'}\sqcap\underline{I})(\underline{J'}\sqcap\underline{J})}\tilde{V}_{(\underline{I'}\sqcap\underline{I})(\underline{J'}\sqcap\underline{J})} \in \gamma\end{aligned}$$

as $\mathcal{I}' \sqcap \mathcal{I} \in \mathbb{I}, \mathcal{J}' \sqcap \mathcal{J} \in \mathbb{J}$ and $l_{\underline{F'}\sqcap\underline{F}} = \begin{cases} m_{\underline{I'J'}}k_{\underline{F}}\frac{m_{\underline{F}}}{m_{\underline{F'}\sqcap\underline{F}}} & \text{if } \mathcal{F}' \cap \mathcal{F} \neq \emptyset \\ 0 & \text{if } \mathcal{F}' \cap \mathcal{F} = \emptyset \end{cases}$. ∎

The following theorem gives a main result which states that an ELREMLQE exists for all variance components simultaneously for the set of factors $\mathbb{I}$ under model based conditions (c1) and (c2), as well as the design based conditions $\mathrm{Bal}(\mathbb{I} \sqcup \mathbb{J})$ and $\mathrm{Bal}(\mathbb{I} \,|\, \mathbb{J} \overset{*}{\sqcup} \mathbb{J} \overset{*}{\sqcup} \mathbb{K})$. The factors in $\mathbb{I}$ are examined simultaneously so that $\mathbb{H} = \mathbb{F}^{\dagger} \backslash \mathbb{I} = \mathbb{I}^{c}$.

Theorem: Consider the Y-Model under [L], [O], [C], and [N] under the notation and assumptions given in this section. In addition, assume conditions (c1) and (c2) in lemma 4 hold and that the design has $\mathrm{Bal}(\mathbb{I} \sqcup \mathbb{J})$ and $\mathrm{Bal}(\mathbb{I} \,|\, \mathbb{J} \overset{*}{\sqcup} \mathbb{J} \overset{*}{\sqcup} \mathbb{K})$. If $\hat{\psi}_{\mathrm{REMLQ}}$ exists and $V_{\hat{\psi}_{\mathrm{REMLQ}}}$ is PD, then $\exists$ an ELREMLQE for $\underline{\psi}_{\mathbb{I}}$.

proof: By lemma 4 under (c1) and (c2), $\forall\; \mathcal{I} \in \mathbb{I} \quad E_{\underline{I}} = \tilde{V}_{\underline{I}} - \sum_{\mathcal{H} \in \mathbb{H}} p_{\underline{H}}^{(\mathrm{I})}\tilde{V}_{\underline{H}} = \tilde{V}_{\underline{I}} + \sum_{\mathcal{J} \in \mathbb{J}} p_{\underline{J}}^{(\mathrm{I})}[\tilde{V}_{\underline{J}} - t_{\underline{I}}\tilde{V}_{\underline{IJ}}]$

$\Rightarrow G_{\underline{F'}\underline{IF}} = \tilde{V}_{\underline{F'}}E_{\underline{I}}\tilde{V}_{\underline{F}} + \tilde{V}_{\underline{F}}E_{\underline{I}}\tilde{V}_{\underline{F'}}$. To show $G_{\underline{F'}\underline{IF}} \in \underline{R}(X^{\circ}) \;\; \forall\, \mathcal{F}, \mathcal{F}' \in \mathbb{F}^{\dagger}$. Then

i) $\forall\; \mathcal{I}' \in \mathbb{I} \;\; \tilde{V}_{\underline{I'}}E_{\underline{I}} = m_{\underline{I'}}m_{\underline{I}}(P_{\underline{I'}\sqcap\underline{I}} - P_{\underline{1}}) + \sum_{\mathcal{J} \in \mathbb{J}} p_{\underline{J}}^{(\mathrm{I})}[0 - t_{\underline{I}}m_{\underline{I'}}m_{\underline{IJ}}(P_{\underline{I'}\sqcap\underline{I}} - P_{\underline{1}})]$

$= (m_{\underline{I'}}m_{\underline{I}} - \sum_{\mathcal{J} \in \mathbb{J}} p_{\underline{J}}^{(\mathrm{I})}t_{\underline{I}}m_{\underline{I'}}m_{\underline{IJ}})(P_{\underline{I'}\sqcap\underline{I}} - P_{\underline{1}}) \propto \tilde{V}_{\underline{I'}\sqcap\underline{I}}$ or $0 \in \underline{R}(X^{\circ})$

$\Rightarrow \forall\; \mathcal{I}^{+} \in \mathbb{I} G_{\underline{I'}\underline{II^{+}}} \propto \tilde{V}_{\underline{I'}\sqcap\underline{I}}\tilde{V}_{\underline{I^{+}}} \propto \tilde{V}_{\underline{I'}\sqcap\underline{I}\sqcap\underline{I^{+}}}$ or $0 \in \underline{R}(X^{\circ})$ by lemma 1

$\Rightarrow \forall\; \mathcal{J}' \in \mathbb{J} \quad G_{\underline{I'}\underline{IJ'}} \propto \tilde{V}_{\underline{I'}\sqcap\underline{I}}\tilde{V}_{\underline{J'}} = 0 \in \underline{R}(X^{\circ})$ by lemma 1

$\Rightarrow \forall\; \mathcal{K}' \in \mathbb{K} \quad G_{\underline{I'}\underline{IK'}} \propto \tilde{V}_{\underline{I'}\sqcap\underline{I}}\tilde{V}_{\underline{K'}} = 0 \in \underline{R}(X^{\circ})$ by lemma 3 i)

$\Rightarrow \forall\, \mathcal{I}^{+} \cup \mathcal{J}^{+} \in \mathbb{I} \sqcup \mathbb{J} \quad G_{\underline{I'}\underline{II^{+}J^{+}}} \propto \tilde{V}_{\underline{I'}\sqcap\underline{I}}\tilde{V}_{\underline{I^{+}J^{+}}} \propto \tilde{V}_{\underline{I'}\sqcap\underline{I}\sqcap\underline{I^{+}}}$ or $0 \in \underline{R}(X^{\circ})$ by assumption i)

$\Rightarrow \forall\, \mathcal{J}' \cup \mathcal{K}' \in \mathbb{J} \sqcup \mathbb{K} \quad G_{\underline{I'}\underline{IJ'K'}} \propto \tilde{V}_{\underline{I'}\sqcap\underline{I}}\tilde{V}_{\underline{J'K'}} = 0 \in \underline{R}(X^{\circ})$ by lemma 3 ii)

$\Rightarrow G_{\underline{I'}\underline{IR}} \propto \tilde{V}_{\underline{I'}\sqcap\underline{I}}\tilde{V}_{\underline{R}} = \tilde{V}_{\underline{I'}\sqcap\underline{I}}$ or $0 \in \underline{R}(X^{\circ})$.

ii) $\forall\; \mathcal{J}' \in \mathbb{J} \quad \tilde{V}_{\underline{J'}}E_{\underline{I}} = 0 + \sum_{\mathcal{J} \in \mathbb{J}} p_{\underline{J}}^{(\mathrm{I})}[m_{\underline{J'}}m_{\underline{J}}(P_{\underline{J'}\sqcap\underline{J}} - P_{\underline{1}}) - t_{\underline{I}}m_{\underline{J'}}m_{\underline{IJ}}(P_{\underline{J'}\sqcap\underline{J}} - P_{\underline{1}})]$

$= \sum_{\mathcal{J} \in \mathbb{J}} p_{\underline{J}}^{(\mathrm{I})}[m_{\underline{J'}}m_{\underline{J}}(P_{\underline{J'}\sqcap\underline{J}} - P_{\underline{1}}) - m_{\underline{J'}}m_{\underline{J}}(P_{\underline{J'}\sqcap\underline{J}} - P_{\underline{1}})] = 0$ by lemma 1

$\Rightarrow \forall\; \mathcal{F} \in \mathbb{F}^{\dagger} \quad G_{\underline{J'}\underline{IF}} = 0 \in \underline{R}(X^{\circ})$.

iii) $\forall\ \mathcal{K} \in \mathbb{K} \qquad \tilde{V}_{\underline{K}} E_{\underline{I}} = \tilde{V}_{\underline{K}} \tilde{V}_{\underline{I}} + \sum_{\mathcal{J} \in \mathbb{J}} p_{\underline{I}}^{(1)} \tilde{V}_{\underline{K}} (\tilde{V}_{\underline{J}} - t_{\underline{I}} \tilde{V}_{\underline{IJ}}) = 0$ by lemma 3 i) iii)

$\Rightarrow \forall\ \mathcal{F} \in \mathbb{F}^{\dagger} \quad G_{\underline{KIF}} = 0 \in \underline{R}(X^{\circ})$.

iv) $\forall\ \mathcal{J}' \cup \mathcal{K}' \in \mathbb{J} \sqcup \mathbb{K} \quad \tilde{V}_{\underline{J'K'}} E_{\underline{I}} = \tilde{V}_{\underline{J'K'}} \tilde{V}_{\underline{I}} + \sum_{\mathcal{J} \in \mathbb{J}} p_{\underline{I}}^{(1)} \tilde{V}_{\underline{J'K'}} [\tilde{V}_{\underline{J}} - t_{\underline{I}} \tilde{V}_{\underline{IJ}}] = 0$ by lemma 3 ii) iv)

$\Rightarrow \forall\ \mathcal{F} \in \mathbb{F}^{\dagger} \quad G_{\underline{J'K'IF}} = 0 \in \underline{R}(X^{\circ})$.

v) $\forall\ \mathcal{I}' \cup \mathcal{J}' \in \mathbb{I} \sqcup \mathbb{J}$

$\tilde{V}_{\underline{I'J'}} E_{\underline{I}} = m_{\underline{I'J'}} m_{\underline{I}} (P_{\underline{I'} \cap \underline{I}} - P_{\underline{1}}) + \sum_{\mathcal{J} \in \mathbb{J}} p_{\underline{I}}^{(1)} [m_{\underline{I'J'}} m_{\underline{J}} (P_{\underline{J'} \cap \underline{J}} - P_{\underline{1}}) - t_{\underline{I}} m_{\underline{I'J'}} m_{\underline{IJ}} (P_{(\underline{I'} \cap \underline{I})(\underline{J'} \cap \underline{J})} - P_{\underline{1}})]$

$= \tilde{V}_{\underline{S}}$ or $0 \in \mathrm{sp}\{\mathcal{I} \in \mathbb{I}, \mathcal{J} \in \mathbb{J} \mid \tilde{V}_{\underline{I}}, \tilde{V}_{\underline{J}}, \tilde{V}_{\underline{IJ}}\} = \gamma$.

Then $\forall\ \mathcal{I}^{+} \cup \mathcal{J}^{+} \in \mathbb{I} \sqcup \mathbb{J} \quad G_{\underline{I'J'II^+J^+}} \propto \tilde{V}_{\underline{S}} \tilde{V}_{\underline{I^+J^+}} = \tilde{V}_{\underline{S'}} \in \gamma \in \underline{R}(X^{\circ})$ by lemma 5

and $G_{\underline{I'J'IR}} \propto \tilde{V}_{\underline{S}} \tilde{V}_{\underline{R}} = \tilde{V}_{\underline{S}}$ or $0 \in \underline{R}(X^{\circ})$.

$\therefore$ By the GZC$_{\text{FR}}$ check and the LQNY-ELREMLQE$_{\text{FR}}$ theorem, an ELREMLQE exists for $\underline{\psi}_{\mathbb{I}}$ as $G_{\underline{F'IF}} \in \underline{R}(X^{\circ}) \ \forall\ \mathcal{F}, \mathcal{F}' \in \mathbb{F}^{\dagger}$ as the above holds $\forall \mathcal{I} \in \mathbb{I}$. ∎

The model based conditons (c1) and (c2) are helpful for identifying random models that have ELREMLQEs for variance components corresponding to main effects. However, the conditions are somewhat abstract, so possible structures in $\mathbb{J}$ will be examined which satisfy these two conditions. Additional notation will be needed to describe such structures in $\mathbb{J}$.

Let $\mathbb{J} = \{\mathcal{J}_1, ..., \mathcal{J}_r\}$ and $\mathcal{J}^* = \{1, ..., s\}$. The set $\mathbb{J}$ has a dominating factor providing $\exists\ \mathcal{J}^d \in \mathbb{J} \ni \mathcal{J} \subset \mathcal{J}^d \ \forall\ \mathcal{J} \in \mathbb{J}$. Note that a dominating factor exists when $\mathbb{J}$ has a nested or complete structure. In addition, let $M_1 = (\{\frac{t_{\underline{J} \cap \underline{J'}}}{t_{\underline{J'}}}\}_{\mathcal{J},\mathcal{J}'})_{r \times r}$ and $M_2 = (\{t_{\underline{J} \cap \underline{J'}}\}_{\mathcal{J},\mathcal{J}'})_{r \times r}$ where $\underline{R}(M_1) = \underline{R}(M_2)$.

A complete structure in $\mathbb{J}$ will be useful for later results. Some special notation will be defined for this case. For purposes of convenience in this setting, let $\mathcal{J}_0 = \emptyset \in \mathbb{J}$ so $r = 2^s$. Then $M_2^* = \{t_{\underline{J} \cap \underline{J'}}\}_{2^s \times 2^s}$ has a row and column of 1's. Note M_2 is a principal submatrix of M_2^*. In addition, define $T = \mathrm{diag}(\{\bigodot_{i=1}^{s}(1, t_i - 1)\})$ and $G = \bigodot_{i=1}^{s}(\underline{1}, \underline{g}_{j_i})$ where $\underline{g}_{j_i} = \{g_{j_k j_i}\} = \begin{cases} 1 & \mathcal{J}_i \subset \mathcal{J}_k \\ 0 & \mathcal{J}_i \not\subset \mathcal{J}_k \end{cases}$. The horizontal direct product $\odot$ is defined in section 2.9.

This notation can be demonstrated by an example where $\mathcal{J}^* = \{1, 2\}$ and $t_{j_0} = 1$. Then

$$M_2^* = \{t_{\underline{J} \cap \underline{J'}}\} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & t_1 & 1 & t_1 \\ 1 & 1 & t_2 & t_2 \\ 1 & t_1 & t_2 & t_1 t_2 \end{bmatrix} = \begin{bmatrix} 1 & \underline{1}_{1 \times 3} \\ \underline{1}_{3 \times 1} & M_2 \end{bmatrix},$$

$$G = (\underline{1}, \underline{g}_1) \odot (\underline{1}, \underline{g}_2) = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix} \odot \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}, \text{ and}$$

$$T = \mathrm{diag}(\{(\underline{1}, t_1 - 1) \odot (\underline{1}, t_2 - 1)\}) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & t_2 - 1 & 0 & 0 \\ 0 & 0 & t_1 - 1 & 0 \\ 0 & 0 & 0 & (t_1 - 1)(t_2 - 1) \end{bmatrix}.$$

The following three lemmas establish properties that will be necessary for the final result.

Lemma 6: If $A_{n\times bc} = B_{n\times b} \odot C_{n\times c}$ and $D_{bc\times bc} = \text{diag}(\underline{e}'_{1\times b} \odot \underline{f}'_{1\times c})$, then $AD = (B' \odot \underline{e})' \odot (C' \odot \underline{f})'$.

proof: Let $j = 1, ..., b$, $k = 1, ..., c$, $i = 1, ..., n$ and $B = \{b_{ij}\}$, $C = \{c_{ik}\}$, $\underline{e} = \{e_j\}$, $\underline{f} = \{f_k\}$. Then

$$AD = \{\underline{a}_{jk}\}\text{diag}(\{e_j f_k\}) = \{\underline{b}_j \odot \underline{c}_k\}\text{diag}(\{e_j f_k\}) = \{b_{ij}c_{ik}\}\text{diag}(\{e_j f_k\})$$
$$= \{b_{ij}c_{ik}e_j f_k\}_{n\times bc} = \{b_{ij}e_j\}_{n\times b} \odot \{c_{ik}f_k\}_{n\times c}$$
$$= (\underline{b}_i \odot \underline{e}) \odot \{\underline{c}_i \odot \underline{f}\} = (B'_{b\times n} \odot \underline{e}_{b\times 1})' \odot (C'_{c\times n} \odot \underline{f}_{c\times 1})'. \quad \blacksquare$$

Lemma 7: Let $m = \#(\mathcal{J} \cap \mathcal{J}')$. Then $\prod\limits_{j\in\mathcal{J}\cap\mathcal{J}'} t_j = 1 + \sum\limits_{\{j_1,...,j_u\} \subset \mathcal{J}\cap\mathcal{J}'} (t_{j_1} - 1)...(t_{j_u} - 1)$

where $\{j_1, ..., j_u\}$ is non-empty and $1 \leq u \leq m$.

proof: i) For $m = 1$, $\prod\limits_{j\in\mathcal{J}\cap\mathcal{J}'} t_j = t_{j_1} = 1 + (t_{j_1} - 1) = 1 + \sum\limits_{(j_1)\subset\mathcal{J}\cap\mathcal{J}'} (t_{j_1} - 1)$

$$= 1 + \sum_{\{j_1,...,j_u\} \subset \mathcal{J}\cap\mathcal{J}'} (t_{j_1} - 1)...(t_{j_u} - 1)\,.$$

ii) Assume the relation holds for $m = k$ where $\mathcal{J} \cap \mathcal{J}' = (j_1, ..., j_k)$. Then the relation can be expressed as $\prod\limits_{i=1}^{k} t_{j_i} = 1 + \sum\limits_{(j_1,...,j_u)\subset(j_1,...,j_k)} (t_{j_1} - 1)...(t_{j_u} - 1) = A \quad (1)$.

Let $\mathcal{U}_k = \{(j_1, ..., j_u) \subset (j_1, ..., j_k) |\, u \in [1, k]\}$

$$\Rightarrow\ \mathcal{U}_{k+1} = \{(j_1, ..., j_u) \subset (j_1, ..., j_k) |\, u \in [0, k+1]\} = \{(j_1, ..., j_u)\backslash(k+1) \subset (j_1, ..., j_k) |\, u \in [0, k]\}$$
$$\uplus\ \{(j_1, ..., j_u) \sqcup (k+1) \subset (j_1, ..., j_k) |\, u \in [0, k]\} = \mathcal{U}_k \uplus \mathcal{W} \qquad (2).$$

For $m = k+1$, $\prod\limits_{i=1}^{k+1} t_{j_i} = t_{k+1}A = (1 + (t_{k+1} - 1))A = A + (t_{k+1} - 1)A$

$$= [1 + \sum_{\mathcal{U}_k}(t_{j_1} - 1)...(t_{j_u} - 1)] + [1 + \sum_{\mathcal{U}_k}(t_{j_1} - 1)...(t_{j_u} - 1)](t_{k+1} - 1) \text{ by (1)}$$
$$= [1 + \sum_{\mathcal{U}_k}(t_{j_1} - 1)...(t_{j_u} - 1)] + \sum_{\mathcal{W}}(t_{j_1} - 1)...(t_{j_u} - 1)(t_{k+1} - 1) \text{ by definition of } \mathcal{W}$$
$$= 1 + \sum_{\mathcal{U}_{k+1}} (t_{j_1} - 1)...(t_{j_u} - 1) \quad \text{by (2).}$$

$$\therefore\ \prod_{j\in\mathcal{J}\cap\mathcal{J}'} t_j = 1 + \sum_{\{j_1,...,j_u\} \subset \mathcal{J}\cap\mathcal{J}'} (t_{j_1} - 1)...(t_{j_u} - 1) \text{ by induction.} \quad \blacksquare$$

Lemma 8: i) The $(\mathcal{J}, \mathcal{J}')$ entry of G is $1 \Leftrightarrow \mathcal{J}' \subset \mathcal{J}$, and 0 otherwise.

ii) The $(\mathcal{J}, \mathcal{J}')$ entry of $(\underline{g}_{j_1} \odot ... \odot \underline{g}_{j_u})(\underline{g}_{j_1} \odot ... \odot \underline{g}_{j_u})'$ is $1 \Leftrightarrow \{j_1, ..., j_u\} \subset \mathcal{J} \cap \mathcal{J}'$, and 0 otherwise.

proof: i) Follows from the definition of G.

ii) The column vector $(\underline{g}_{j_1} \odot \dots \odot \underline{g}_{j_u})$ has entry 1 in row $\mathcal{J} \Leftrightarrow \{j_1, \dots, j_u\} \subset \mathcal{J}$ (1) by i).

The row vector $(\underline{g}_{j_1} \odot \dots \odot \underline{g}_{j_u})'$ has entry 1 in column $\mathcal{J}' \Leftrightarrow \{j_1, \dots, j_u\} \subset \mathcal{J}'$ (2) by i).

Thus, $(\underline{g}_{j_1} \odot \dots \odot \underline{g}_{j_u})(\underline{g}_{j_1} \odot \dots \odot \underline{g}_{j_u})'$ has entry 1 at $(\mathcal{J}, \mathcal{J}') \Leftrightarrow \{j_1, \dots, j_u\} \subset \mathcal{J} \cap \mathcal{J}'$

as both (1) and (2) must hold for the entry to be 1. ∎

Corollary: Consider the Y-Model under [L], [O], [C], and [N] under the notation and assumptions given in this section. In addition, assume $\mathrm{Bal}(\mathbb{I} \sqcup \mathbb{J})$, $\mathrm{Bal}(\mathbb{I} \,|\, \mathbb{J} \overset{*}{\sqcup} \mathbb{K})$, and that the set $\mathbb{J}$ has a dominating factor. If $\widehat{\underline{\psi}}_{\mathrm{REMLQ}}$ exists and $V_{\widehat{\underline{\psi}}_{\mathrm{REMLQ}}}$ is PD, then $\exists$ an ELREMLQE for $\underline{\psi}_{\mathbb{I}}$.

proof: i) Suppose $\mathbb{J} = \{\mathcal{J}_1, \dots, \mathcal{J}_r\}$ has a complete structure. Then $\mathcal{J}_1, \dots, \mathcal{J}_s$, which contain only a single element, are also in $\mathbb{J}$. Hence,

$$
\begin{aligned}
(GTG')_{2^s \times 2^s} &= [(\underline{1}, \underline{g}_{j_1}) \odot \dots \odot (\underline{1}, \underline{g}_{j_s})][\mathrm{diag}(\{(\underline{1}, t_{j_1} - 1) \odot \dots \odot (\underline{1}, t_{j_s} - 1)\})]G' \\
&= [\underset{i=1}{\overset{s}{\odot}} (\begin{bmatrix} \underline{1}' \\ \underline{g}'_{j_i} \end{bmatrix} \odot \begin{bmatrix} 1 \\ t_{j_i} - 1 \end{bmatrix})'\,]\, G' \text{ repeatedly applying Lemma 6} \\
&= [\underset{i=1}{\overset{s}{\odot}} (\underline{1}, (t_{j_i} - 1)\underline{g}_{j_i})]\, G' \\
&= [(\underline{1}, (t_{j_1} - 1)\underline{g}_{j_1}) \odot \dots \odot (\underline{1}, (t_{j_s} - 1)\underline{g}_{j_s})][(\underline{1}, \underline{g}_{j_1}) \odot \dots \odot (\underline{1}, \underline{g}_{j_s})]' \\
&= [(\underline{1}, \dots, (t_{j_1} - 1)\dots(t_{j_u} - 1)(\underline{g}_{j_1} \odot \dots \odot \underline{g}_{j_u}), \dots][\underline{1}, \dots, (\underline{g}_{j_1} \odot \dots \odot \underline{g}_{j_u}), \dots]' \\
&= \underline{1}\underline{1}' + \dots + (t_{j_1} - 1)\dots(t_{j_u} - 1)(\underline{g}_{j_1} \odot \dots \odot \underline{g}_{j_u})(\underline{g}_{j_1} \odot \dots \odot \underline{g}_{j_u})' \\
&= \{1 + \sum_{\{j_1, \dots, j_u\} \subset \mathcal{J} \cap \mathcal{J}'} (t_{j_1} - 1)\dots(t_{j_u} - 1)\}_{\mathcal{J}, \mathcal{J}'} \text{ by Lemma 8 where } \{j_1, \dots, j_u\} \text{ non-empty} \\
&= \{\prod_{j \in \mathcal{J} \cap \mathcal{J}'} t_j\} \text{ by Lemma 7} \\
&= \{t_{\underline{I} \cap \underline{I}'}\} = M_2^*.
\end{aligned}
$$

ii) Order the matrix G to obtain G° so that the $\mathcal{J}_0$ column comes first ($\mathcal{J}_0 = \emptyset$), the $\mathcal{J}_i$ columns which contain exactly 1 member come next, and so on. This results in G° having ordered columns $\ni$ if $\mathcal{J} \subset \mathcal{J}'$, then the column corresponding to $\mathcal{J}$ precedes the column corresponding to $\mathcal{J}'$. Order the rows of G° in the same manner. On the diagonal, the $(\mathcal{J}, \mathcal{J})$ entry of G° is 1 since $\mathcal{J} \subset \mathcal{J}$ by Lemma 8. Above the diagonal, the $(\mathcal{J}, \mathcal{J}')$ entry of G° is 0 because if it were 1, then $\mathcal{J}' \subset \mathcal{J}$ and $\mathcal{J}'$ would precede $\mathcal{J}$. This cannot be the case as $\mathcal{J}$ must precede $\mathcal{J}'$ in order for the $\mathcal{J}, \mathcal{J}'$ entry to be above the diagonal.

$\therefore$ G° is lower triangular and nonsingular.

iii) Note $\underline{r}(G) = 2^s$ by ii) $\Rightarrow M_2^*$ is PD by i) as it is NND and has full rank

$\Rightarrow$ every principal submatrix of M_2^* is PD $\Rightarrow M_2$ is PD as it is a principal submatrix of M_2^*

$\Rightarrow M_1$ is non-singular since $\underline{R}(M_2) = \underline{R}(M_1)$. Thus, for any structure within $\mathbb{J}$ and corresponding matrix $(M_1)_{(\#\mathbb{J}) \times (\#\mathbb{J})}$, $\exists$ a unique solution $\underline{p}^{(\mathrm{I})} \ni M_1\{p_{\underline{I}'}^{(\mathrm{I})}\} = \frac{-1}{t_1}\underline{1}$ (c1).

iv) When $\mathbb{J}$ has a dominating factor, then for the last row of M_1 corresponding to $\mathcal{J}^d$,

$\mathcal{J}' \subset \mathcal{J}^d$ $\forall$ columns correponding to $\mathcal{J}'$

$\Rightarrow \frac{t_{\mathcal{J}^d \sqcap \mathcal{J}'}}{t_{\mathcal{J}'}} = \frac{t_{\mathcal{J}'}}{t_{\mathcal{J}'}} = 1$ and so (c1) $\Rightarrow M_1\{p_{\mathcal{J}'}^{(\mathrm{I})}\} = \frac{-1}{t_{\mathrm{I}}}\underline{1} \Rightarrow \sum_{\mathcal{J}' \in \mathbb{J}} p_{\mathcal{J}'}^{(\mathrm{I})} = \frac{-1}{t_{\mathrm{I}}}$ (c2) .

v) By lemma 2 iv), $\mathbb{J}$ has a dominating factor $\Rightarrow$ Bal($\mathbb{I}\,|\mathbb{J} \overset{*}{\sqcup} \mathbb{J} \overset{*}{\sqcup} \mathbb{K}$) $\Leftrightarrow$ Bal($\mathbb{I}\,|\mathbb{J} \overset{*}{\sqcup} \mathbb{K}$).

$\therefore$ $\exists$ a unique solution $\underline{p}^{(\mathrm{I})}$ which satisfies (c1) and (c2) by i)-iv). By v) and the above theorem, under Bal($\mathbb{I} \sqcup \mathbb{J}$) and Bal($\mathbb{I}\,|\mathbb{J} \overset{*}{\sqcup} \mathbb{K}$), if the set $\mathbb{J}$ has a dominating factor, then $\exists$ an ELREMLQE for $\underline{\psi}_{\mathbb{I}}$. ∎

The conditions for the existence of an ELREMLQE for $\underline{\psi}_{\mathbb{I}}$ in random models are dependent on a partition corresponding to $\mathbb{I}$, $\mathbb{J}$, $\mathbb{K}$, $\mathbb{I} \sqcup \mathbb{J}$, $\mathbb{J} \sqcup \mathbb{K}$ under (c1) and (c2). The conditions (c1) and (c2) are satisfied when $\mathbb{J}$ has a dominating factor such as when there is a nested or complete structure. The design needs to have Bal($\mathbb{I} \sqcup \mathbb{J}$) and Bal($\mathbb{I}\,|\mathbb{J} \overset{*}{\sqcup} \mathbb{K}$) in order for this model to have the ELREMLQE for $\underline{\psi}_{\mathbb{I}}$. The conditions for the random models may provide insight into conditions for other classification models.

6.9. Searching for Examples Involving 3-Way Models

In order to identify examples that satisfy the GZC$_{\mathrm{FR}}$ for the variance component vector, a search was conducted for 3-way classification models with 2 levels of each factor. Patterns were examined to identify classes of examples such as those proven in sections 6.6, 6.7, and 6.8. These examples are tabled in this section for reference.

Table 6.1 identifies the incidence matrices that were used for the search. These were chosen to reflect types of balance (bal). Some of the designs are permutations of one another where these permutations were used to identify the behavior of particular factors.

Table 6.2 lists models and the associated designs that had ELMLQEs or ELREMLQEs, but not FELMLQEs or FELREMLQEs for the variance components. The results were obtained from a search of all possible proper 2 level 3-way classification models under the designs listed in Table 6.1. However, duplicate cases involving permutations of the factors were removed. In addition, REML cases do not include those involving pseudo balance for random models that contain the highest possible order interaction term and have an ELREMLQE for the residual component. These cases were proven in section 6.7. Also, the ML cases do not include completely balanced models that have the highest possible order interaction and have an ELMLQE for the residual component. These cases were proven in section 6.6. Such cases were removed to keep the table succinct.

For example, consider the first line of Table 6.2 denoted by ($*$). This line shows that for the ML method with ZC for the Y-Model under incidence matrix 1 in Table 6.1, $\exists$ an ELMLQE for σ_e^2 in a mixed model with effect A fixed, effects B C BC random, and effects AB AC ABC omitted. The other lines of the table follow in the same manner.

Table 6.1. Particular Incidence Matrices for 3-Way Models with 2 Levels

<table>
<tr><th>#</th><th>a × bc</th><th>a × b</th><th>a × c</th><th>b × c</th><th>a</th><th>b</th><th>c</th><th>balance</th></tr>
<tr><td>1</td><td>$\begin{bmatrix} 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 \end{bmatrix}$</td><td>bal</td><td>bal</td><td>bal</td><td>bal</td><td>bal</td><td>bal</td><td>bal(abc)</td></tr>
<tr><td>2</td><td>$\begin{bmatrix} 1 & 1 & 2 & 2 \\ 1 & 1 & 2 & 2 \end{bmatrix}$</td><td>bal</td><td>$\begin{bmatrix} 2 & 4 \\ 2 & 4 \end{bmatrix}$</td><td>$\begin{bmatrix} 2 & 4 \\ 2 & 4 \end{bmatrix}$</td><td>bal</td><td>bal</td><td>$\begin{bmatrix} 4 \\ 8 \end{bmatrix}$</td><td>bal(ab|c)</td></tr>
<tr><td>3</td><td>$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 \end{bmatrix}$</td><td>$\begin{bmatrix} 2 & 2 \\ 4 & 4 \end{bmatrix}$</td><td>$\begin{bmatrix} 2 & 2 \\ 4 & 4 \end{bmatrix}$</td><td>bal</td><td>$\begin{bmatrix} 4 \\ 8 \end{bmatrix}$</td><td>bal</td><td>bal</td><td>bal(bc|a)</td></tr>
<tr><td>4</td><td>$\begin{bmatrix} 1 & 2 & 1 & 2 \\ 1 & 2 & 1 & 2 \end{bmatrix}$</td><td>$\begin{bmatrix} 2 & 4 \\ 2 & 4 \end{bmatrix}$</td><td>bal</td><td>$\begin{bmatrix} 2 & 2 \\ 4 & 4 \end{bmatrix}$</td><td>bal</td><td>$\begin{bmatrix} 4 \\ 8 \end{bmatrix}$</td><td>bal</td><td>bal(ac|b)</td></tr>
<tr><td>5</td><td>$\begin{bmatrix} 1 & 1 & 2 & 2 \\ 2 & 2 & 1 & 1 \end{bmatrix}$</td><td>bal</td><td>$\begin{bmatrix} 2 & 4 \\ 4 & 2 \end{bmatrix}$</td><td>bal</td><td>bal</td><td>bal</td><td>bal</td><td>bal(b|ac)</td></tr>
<tr><td>6</td><td>$\begin{bmatrix} 1 & 2 & 1 & 2 \\ 2 & 1 & 2 & 1 \end{bmatrix}$</td><td>$\begin{bmatrix} 2 & 4 \\ 4 & 2 \end{bmatrix}$</td><td>bal</td><td>bal</td><td>bal</td><td>bal</td><td>bal</td><td>bal(c|ab)</td></tr>
<tr><td>7</td><td>$\begin{bmatrix} 1 & 2 & 2 & 1 \\ 1 & 2 & 2 & 1 \end{bmatrix}$</td><td>bal</td><td>bal</td><td>$\begin{bmatrix} 2 & 4 \\ 4 & 2 \end{bmatrix}$</td><td>bal</td><td>bal</td><td>bal</td><td>bal(a|bc)</td></tr>
<tr><td>8</td><td>$\begin{bmatrix} 1 & 2 & 2 & 1 \\ 2 & 1 & 1 & 2 \end{bmatrix}$</td><td>bal</td><td>bal</td><td>bal</td><td>bal</td><td>bal</td><td>bal</td><td></td></tr>
<tr><td>9</td><td>$\begin{bmatrix} 1 & 1 & 2 & 2 \\ 2 & 2 & 3 & 3 \end{bmatrix}$</td><td>$\begin{bmatrix} 3 & 3 \\ 5 & 5 \end{bmatrix}$</td><td>$\begin{bmatrix} 2 & 4 \\ 4 & 6 \end{bmatrix}$</td><td>$\begin{bmatrix} 3 & 5 \\ 3 & 5 \end{bmatrix}$</td><td>$\begin{bmatrix} 6 \\ 10 \end{bmatrix}$</td><td>bal</td><td>$\begin{bmatrix} 6 \\ 10 \end{bmatrix}$</td><td>bal(b|ac)</td></tr>
<tr><td>10</td><td>$\begin{bmatrix} 1 & 1 & 3 & 3 \\ 2 & 2 & 4 & 4 \end{bmatrix}$</td><td>$\begin{bmatrix} 4 & 4 \\ 6 & 6 \end{bmatrix}$</td><td>$\begin{bmatrix} 2 & 6 \\ 4 & 8 \end{bmatrix}$</td><td>$\begin{bmatrix} 3 & 7 \\ 3 & 7 \end{bmatrix}$</td><td>$\begin{bmatrix} 8 \\ 12 \end{bmatrix}$</td><td>bal</td><td>$\begin{bmatrix} 6 \\ 14 \end{bmatrix}$</td><td>bal(b|ac)</td></tr>
<tr><td>11</td><td>$\begin{bmatrix} 1 & 2 & 1 & 2 \\ 2 & 3 & 2 & 3 \end{bmatrix}$</td><td>$\begin{bmatrix} 2 & 4 \\ 4 & 6 \end{bmatrix}$</td><td>$\begin{bmatrix} 3 & 3 \\ 5 & 5 \end{bmatrix}$</td><td>$\begin{bmatrix} 3 & 3 \\ 5 & 5 \end{bmatrix}$</td><td>$\begin{bmatrix} 6 \\ 10 \end{bmatrix}$</td><td>$\begin{bmatrix} 6 \\ 10 \end{bmatrix}$</td><td>bal</td><td>bal(c|ab)</td></tr>
<tr><td>12</td><td>$\begin{bmatrix} 1 & 2 & 3 & 2 \\ 2 & 3 & 2 & 1 \end{bmatrix}$</td><td>bal</td><td>$\begin{bmatrix} 3 & 5 \\ 5 & 3 \end{bmatrix}$</td><td>$\begin{bmatrix} 3 & 5 \\ 5 & 3 \end{bmatrix}$</td><td>bal</td><td>bal</td><td>bal</td><td></td></tr>
<tr><td>13</td><td>$\begin{bmatrix} 1 & 2 & 2 & 3 \\ 2 & 3 & 3 & 4 \end{bmatrix}$</td><td>$\begin{bmatrix} 3 & 5 \\ 5 & 7 \end{bmatrix}$</td><td>$\begin{bmatrix} 3 & 5 \\ 5 & 7 \end{bmatrix}$</td><td>$\begin{bmatrix} 3 & 5 \\ 5 & 7 \end{bmatrix}$</td><td>$\begin{bmatrix} 8 \\ 12 \end{bmatrix}$</td><td>$\begin{bmatrix} 8 \\ 12 \end{bmatrix}$</td><td>$\begin{bmatrix} 8 \\ 12 \end{bmatrix}$</td><td></td></tr>
<tr><td>14</td><td>$\begin{bmatrix} 1 & 2 & 2 & 4 \\ 2 & 3 & 4 & 6 \end{bmatrix}$</td><td>$\begin{bmatrix} 3 & 6 \\ 6 & 9 \end{bmatrix}$</td><td>$\begin{bmatrix} 3 & 6 \\ 5 & 10 \end{bmatrix}$</td><td>$\begin{bmatrix} 3 & 6 \\ 5 & 10 \end{bmatrix}$</td><td>$\begin{bmatrix} 9 \\ 15 \end{bmatrix}$</td><td>$\begin{bmatrix} 9 \\ 15 \end{bmatrix}$</td><td>$\begin{bmatrix} 8 \\ 16 \end{bmatrix}$</td><td></td></tr>
</table>

Table 6.2. Particular 3-Way Models with ELMLQE and ELREMLQE for Variance Components

f = FIXED . = NOT IN MODEL 1 = RANDOM EXPLICIT 0 = RANDOM NOT EXPLICIT

	DESIGN	TYPE	A	B	C	AB	AC	BC	ABC	e	
ML+ZC	1	mixed	f	0	0	.	.	0	.	1	(*)
	3	mixed	f	0	0	.	.	0	.	1	
	1	mixed	f	0	0	.	.	0	1	1	
	1	mixed	f	0	0	.	1	0	.	1	
	1	mixed	f	0	0	.	1	0	1	1	
	1	mixed	f	0	0	1	.	0	.	1	
	1	mixed	f	0	0	1	.	0	1	1	
	1	random	.	0	0	.	.	0	.	1	
	3	random	.	0	0	.	.	0	.	1	
	5	random	.	0	0	.	.	0	.	1	
	6	random	.	0	0	.	.	0	.	1	
	8	random	.	0	0	.	.	0	.	1	
	15	random	.	0	0	.	.	0	.	1	
	1	random	.	0	0	.	.	0	1	1	
	1	random	.	0	0	.	1	0	.	1	
	1	random	.	0	0	.	1	0	1	1	
	1	random	.	0	0	1	.	0	.	1	
	1	random	.	0	0	1	.	0	1	1	
	1	random	0	0	0	0	.	0	.	1	
	1	random	0	0	0	0	.	0	1	1	
	1	random	0	0	0	0	0	.	.	1	
	1	random	0	0	0	0	0	.	1	1	
REML	3	mixed	1	.	f	.	0	0	0	0	
	5	mixed	1	.	f	.	0	0	0	0	
	3	mixed	1	0	f	.	0	.	0	0	
	5	mixed	1	0	f	.	0	.	0	0	
	3	mixed	1	0	f	.	0	0	0	0	
	5	mixed	1	0	f	.	0	0	0	0	
	2	mixed	1	0	f	0	.	0	.	0	
	6	mixed	0	0	f	1	0	0	0	0	
	8	mixed	0	0	f	1	0	0	0	0	
	3	mixed	1	f	f	0	.	0	0	0	
	6	mixed	1	f	f	0	.	0	0	0	
	8	mixed	0	f	f	1	.	0	0	0	
	6	mixed	0	f	f	1	0	0	0	0	
	8	mixed	0	f	f	1	1	0	0	0	
	6	mixed	f	f	f	1	0	0	0	0	
	3	mixed	1	f	f	0	.	.	0	0	
	6	mixed	1	f	f	0	.	.	0	0	
	3	mixed	1	f	.	0	.	.	0	0	
	6	mixed	1	f	.	0	.	.	0	0	
	2	random	1	0	.	0	.	0	.	0	
	7	random	1	0	.	0	.	0	.	0	
	7	random	1	0	0	0	.	.	.	0	
	2	random	1	0	0	0	.	0	.	0	
	7	random	1	0	0	0	.	0	.	0	

7. UMVUE in the Full and General Case

This chapter applies the results in chapters 4 and 5 to uniformly minimum variance unbiased estimation. It has been shown that the conditions for the existence of an ELMLQE or an ELREMLQE are equivalent to the existence of a UBLUE for the associated model. This section examines the relationship of the ELMLQE and the ELREMLQE to the uniformly minimum variance unbiased estimator (UMVUE). The UMVUE is defined below for euclidean vectors $\underline{T}$, $\underline{S}$, $\underline{\theta}$, and $\underline{d}$:

(Casella and Berger,1990)
UMVUE: An estimator $\underline{T}$ is UMVUE for a parameter $\underline{\theta}$ if it satisfies $\mathrm{E}_{\underline{\theta}}[\underline{T}] = \underline{\theta} \ \forall \ \underline{\theta}$, and for any other estimator $\underline{S} \ni \mathrm{E}_{\underline{\theta}}[\underline{S}] = \underline{\theta} \ \forall \ \underline{\theta}, \ \mathrm{Var}_{\underline{\theta}}(\underline{d}'\underline{T}) \leq \mathrm{Var}_{\underline{\theta}}(\underline{d}'\underline{S}) \ \forall \ \underline{\theta}, \ \underline{d}$.

The definition indicates that the UMVUE has minimum variance over all unbiased estimators. Note that the UMVUE is model dependent through the expectation and variance. The results will first be presented for the full case using the results of chapter 4 and then for the general case using the results of chapter 5. Section 7.2 gives exact forms of the covariance of the ELMLQE and the ELREMLQE. For the ELMLQE involving the fixed effects and the ELREMLQE involving the variance components, the covariance can shown to be a function of the information matrix.

7.1. UMVUE in the Full Case

This section demonstrates that $\Lambda'\widehat{\underline{\beta}}_{\mathrm{MLQ}}$, $\Gamma'\widehat{\underline{\psi}}_{\mathrm{MLQ}}$, and $\Gamma'\widehat{\underline{\psi}}_{\mathrm{REMLQ}}$ are UMVUE for their expectation in the appropriate model under the full UBLUE conditions. This purpose of this section is to show how the results of this thesis are related to previous results concerning UMVUEs from Seely (1969,1971,1977). The previous results prove that, under the full UBLUE conditions, a complete sufficient statistic (CSS) exists for the normal family of distributions under both the ML and REML methods. This is established in the following two theorems for the given family of normal distributions. This section assumes that $\underline{\beta}$ is mean estimable in the Y-Model, $\underline{\psi}$ is mean estimable in the ALQEM for $(\underline{Y} - X\widehat{\underline{\beta}})$, and $\underline{\psi}$ is mean estimable in the LQEM for $Q'\underline{Y}$.

(Seely,1971)
Theorem 1: Consider the Y-Model under [L], [O], [N], ZC, and QS. Then $(\underline{Y}'V_1\underline{Y}, \ldots, \underline{Y}'V_k\underline{Y}, \underline{Y}'\underline{Y})'$ and $X'\underline{Y}$ are jointly a CSS.

proof: (1) Note $\underline{R}(V_{\underline{\psi}}X) \subset \underline{R}(X) \Rightarrow \underline{R}(V_{\underline{\psi}}X) = \underline{R}(X)$
$\Rightarrow V_{\underline{\psi}}P_X = P_XV_{\underline{\psi}}$ and $P_XV_{\underline{\psi}}^{-1} = V_{\underline{\psi}}^{-1}P_X$ by the proposition in section 3.3.5.

(2) From Seely(1971), sp$\mathcal{V}$ is a QS $\Rightarrow V_{\underline{\psi}}^{-1} = \sum_{i=1}^{k+1}\theta_i(\underline{\psi})V_i$ where

$\underline{\theta}(\underline{\psi}) = [\theta_1(\underline{\psi}), ..., \theta_k(\underline{\psi})]'$ is an open mapping.

(3) $f(\underline{y}|\underline{\beta},\underline{\psi}) = (2\pi)^{-\frac{n}{2}}|V_{\underline{\psi}}|^{-\frac{1}{2}}\exp(\frac{-1}{2}(\underline{y} - X\underline{\beta})'V_{\underline{\psi}}^{-1}(\underline{y} - X\underline{\beta}))$ as $\underline{Y} \sim N(X\underline{\beta}, V_{\underline{\psi}})$

$= h(\underline{\beta},\underline{\psi})\exp(\frac{-1}{2}\underline{y}'V_{\underline{\psi}}^{-1}\underline{y} + \underline{\beta}'X'V_{\underline{\psi}}^{-1}\underline{y}) = h(\underline{\beta},\underline{\psi})\exp(\frac{-1}{2}\underline{y}'V_{\underline{\psi}}^{-1}\underline{y} + \underline{\beta}'X'P_XV_{\underline{\psi}}^{-1}\underline{y})$

$= h(\underline{\beta},\underline{\psi})\exp(\frac{-1}{2}\underline{y}'V_{\underline{\psi}}^{-1}\underline{y} + \underline{\beta}'X'V_{\underline{\psi}}^{-1}P_X\underline{y})$ by (1)

$= h(\underline{\beta},\underline{\psi})\exp(\frac{-1}{2}\sum_{i=1}^{k+1}\theta_i(\underline{\psi})\underline{y}'V_i\underline{y} + \underline{\beta}'X'(\sum_{i=1}^{k+1}\theta_i(\underline{\psi})V_i)P_X\underline{y})$ using (2)

$= h(\underline{\beta},\underline{\psi})\exp(\sum_{i=1}^{k+1}\phi_i(\underline{\psi})\underline{y}'V_i\underline{y} + [\underline{\delta}(\underline{\beta},\underline{\psi})]'X'\underline{y})$.

(4) Let $\Omega = \{[\delta_1(\underline{\beta},\underline{\psi}), ..., \delta_p(\underline{\beta},\underline{\psi}), \phi_1(\underline{\psi}), ..., \phi_{k+1}(\underline{\psi})]' | \underline{\beta} \in \mathcal{R}^p, \underline{\psi} \in \Xi\}$.

Note $\mathcal{R}^p \times (\phi_1(\underline{\psi}), ..., \phi_{k+1}(\underline{\psi})) \in \Omega$ since for fixed $\underline{\psi}$, $\underline{\beta}$ ranges over $\mathcal{R}^p$ and so $\underline{\delta}(\underline{\beta},\underline{\psi})$ ranges over $\mathcal{R}^p$. Thus, $\Omega = \mathcal{R}^p \times \underline{\phi}(\Xi)$ contains a non-empty open set as $\mathcal{R}^p$ contains a non-empty open set, Ξ contains a non-empty open set by [O], and $\underline{\phi}$ is an open mapping by (2).

$\therefore$ $(\underline{Y}'V_1\underline{Y}, ..., \underline{Y}'V_k\underline{Y}, \underline{Y}'\underline{Y})'$, $X'\underline{Y}$ are jointly CSS from Lehmann (1986,Theorem 4.3.1). ∎

The restricted maximum likelihood estimation method was presented in section 3.1.4. Consider the matrix $Q_{n\times q}$ for $q = n - \underline{r}(X)$ which has columns that form an orthonormal basis for $\underline{R}(X)^{\perp}$. Then $Q'Q = I$ and $QQ' = I - P_X = N_X$. For the Y-Model under [N], $\underline{Y} \sim \mathrm{N}_n(X\underline{\beta}, V_{\underline{\psi}})$ which implies $Q'\underline{Y} \sim \mathrm{N}_q(\underline{0}, Q'V_{\underline{\psi}}Q)$. The latter model will be denoted the QY-Model. In addition, let sp$\mathcal{V}^{\triangleright}$ = sp$\{Q'V_1Q, ..., Q'V_kQ, I\}$.

(Seely,1971)

Theorem 2: Consider the Y-Model under [L], [O], and [N] where sp$\mathcal{V}^{\triangleright}$ is a QS. Then $(\underline{Y}'N_XV_1N_X\underline{Y}, ..., \underline{Y}'N_XV_kN_X\underline{Y}, \underline{Y}'N_X\underline{Y})'$ is a CSS in the QY-Model.

proof: (1) Let $V^{\triangleright} = Q'VQ$. From Seely(1971), sp$\mathcal{V}^{\triangleright}$ is a QS $\Rightarrow V_{\underline{\psi}}^{\triangleright -1} = \sum_{i=1}^{k+1}\theta_i(\underline{\psi})V_i^{\triangleright}$ where

$\underline{\theta}(\underline{\psi}) = [\theta_1(\underline{\psi}), ..., \theta_k(\underline{\psi})]'$ is an open mapping.

(2) $f(Q'\underline{y}|\underline{\psi}) = (2\pi)^{-\frac{n}{2}}|Q'V_{\underline{\psi}}Q|^{-\frac{1}{2}}\exp(\frac{-1}{2}(\underline{y}'Q(Q'V_{\underline{\psi}}Q)^{-1}Q'\underline{y})$ as $Q'\underline{Y} \sim N(\underline{0}, Q'V_{\underline{\psi}}Q)$

$= h(\underline{\psi})\exp(\frac{-1}{2}\sum_{i=1}^{k+1}\theta_i(\underline{\psi})\underline{y}'QQ'V_iQQ'\underline{y})$ using (1).

(3) Under [O], Ξ contains a non-empty open set

$\Rightarrow (\underline{Y}'N_XV_1N_X\underline{Y}, ..., \underline{Y}'N_XV_kN_X\underline{Y}, \underline{Y}'N_X\underline{Y})'$ is a CSS from Lehmann (1986,Theorem 4.3.1). ∎

Seely (1977) also shows that the conditions in theorems 1 and 2 are necessary and sufficient for the existence of a CSS. However, sufficiency is adequate for this section.

It will be shown that the ELMLQE and ELREMLQE are functions of the CSS. This indicates that the these quantities are UMVUE by the Lehmann-Scheffe theorem (Casella and Berger,1990,p320).

Lemma 1: If $T(\underline{W})$ is sufficient for the family $\mathcal{P} = \{f_{\underline{\theta}}(\underline{w}) \mid \underline{\theta} \in \Theta\}$ and if $\widehat{\underline{\theta}}$ is a solution to the ML equations, then $\widehat{\underline{\theta}}$ is a function of $T(\underline{W})$.

proof: By the Factorization theorem (Lehmann,1983,Theorem 5.2), $T(\underline{W})$ is sufficient for the family $\mathcal{P}$ $\Leftrightarrow f_{\underline{\theta}}(\underline{w}) = g_{\underline{\theta}}(T(\underline{w}))h(\underline{w})$ for some $g_{\underline{\theta}}, h$. Thus, $\frac{\partial}{\partial \underline{\theta}} \ln f_{\underline{\theta}}(\underline{w}) = 0 \Leftrightarrow \frac{\partial}{\partial \underline{\theta}} \ln(g_{\underline{\theta}}(T(\underline{w}))h(\underline{w})) = 0$ $\Leftrightarrow \frac{\partial}{\partial \underline{\theta}} \ln(g_{\underline{\theta}}(T(\underline{w})) + \frac{\partial}{\partial \underline{\theta}} \ln h(\underline{w}) = 0 \Leftrightarrow \frac{\partial}{\partial \underline{\theta}} \ln(g_{\underline{\theta}}(T(\underline{w})) = 0$ $(*)$. Then $\widehat{\underline{\theta}}$ is a solution to $(*)$ which depends only on $T(\underline{w}) \Rightarrow \widehat{\underline{\theta}}$ is a function of $T(\underline{W})$. ∎

Theorem 3: Consider the Y-model under [L], [O], [N], ZC, and QS where $\widehat{\underline{\psi}}_{\mathrm{MLQ}}$ exists and $V_{\widehat{\underline{\psi}}_{\mathrm{MLQ}}}$ is PD. Then $\Lambda'\widehat{\underline{\beta}}_I$ and $\Gamma'\widehat{\underline{\psi}}_I$ are FELMLQE and UMVUE for $\Lambda'\underline{\beta}$ and $\mathrm{E}[\Gamma'\widehat{\underline{\psi}}_I]$.

proof: Since ZC holds and sp$\mathcal{V}$ is a QS where $V_{\widehat{\underline{\psi}}_{\mathrm{MLQ}}}$ is PD by hypothesis, $\Lambda'\widehat{\underline{\beta}}_{\mathrm{MLQ}} = \Lambda'\widehat{\underline{\beta}}_I$ and $\Gamma'\widehat{\underline{\psi}}_{\mathrm{MLQ}} = \Gamma'\widehat{\underline{\psi}}_I$ are FELMLQEs by the Y-FELMLQE and ALQNY-FELMLQE theorems. In addition, $\exists$ a complete sufficient statistic for the family of distributions by theorem 1 where $\Lambda'\widehat{\underline{\beta}}_{\mathrm{MLQ}}$ and $\Gamma'\widehat{\underline{\psi}}_{\mathrm{MLQ}}$ are functions of the sufficient statistic by lemma 1. Thus, $\Lambda'\widehat{\underline{\beta}}_I$ and $\Gamma'\widehat{\underline{\psi}}_I$ are UMVUE for their expectation by the Lehmann-Scheffe theorem (Casella and Berger,1990,p320). ∎

Theorem 4: Consider the Y-Model under [L], [O], and [N] where sp$\mathcal{V}^{\triangleright}$ is a QS, $\widehat{\underline{\psi}}_{\mathrm{REMLQ}}$ exists, and $V_{\widehat{\underline{\psi}}_{\mathrm{REMLQ}}}$ is PD. Then $\Gamma'\widehat{\underline{\psi}}_I$ is FELREMLQE and UMVUE in the QY-Model for $\Gamma'\underline{\psi}$.

proof: Since sp$\mathcal{V}^{\triangleright}$ is a QS by hypothesis, $\Gamma'\widehat{\underline{\psi}}_{\mathrm{REMLQ}} = \Gamma'\widehat{\underline{\psi}}_I$ are FELREMLQE by the LQNY-FELREMLQE theorem. In addition, $\exists$ a complete sufficient statistic for the family of distributions by theorem 2 where $\Gamma'\widehat{\underline{\psi}}_{\mathrm{REMLQ}}$ is a function of the sufficient statistic by lemma 1. Thus, $\Gamma'\widehat{\underline{\psi}}_I$ is UMVUE in the QY-Model for $\Gamma'\underline{\psi}$ by the Lehmann-Scheffe theorem (Casella and Berger,1990, p320). ∎

These results only apply to the full case and cannot be extended to the general case, since the family of distributions do not necessarily admit a complete sufficient statistic under the general UBLUE conditions. Results for the general case are given in the next section.

7.2. UMVUE in the General Case

The previous section examined UMVUE properties under the UBLUE conditions in the full case presented in chapter 4. This section will examine UMVUE properties under the UBLUE conditions in the general case presented in chapter 5. These conditions will be used to provide expressions for the covariance of $\Lambda'\widehat{\underline{\beta}}_{\text{MLQ}}$, $\Gamma'\widehat{\underline{\psi}}_{\text{MLQ}}$, and $\Gamma'\widehat{\underline{\psi}}_{\text{REMLQ}}$. Without the UBLUE conditions, such exact expressions cannot be obtained. In addition, it will be proven that $\Lambda'\widehat{\underline{\beta}}_{\text{MLQ}}$, $\Gamma'\widehat{\underline{\psi}}_{\text{MLQ}}$ where $\underline{\beta}$ is known, and $\Gamma'\widehat{\underline{\psi}}_{\text{REMLQ}}$ are UMVUE in the appropriate model for $\Lambda'\underline{\beta}$ and $\Gamma'\underline{\psi}$ under the UBLUE conditions. An estimator can be shown to be UMVUE by showing that the covariance of the estimator attains a lower bound over all unbiased estimators.

7.2.1. The Covariance Inequality

The next result gives the lower bound for the variance of an estimator. It will be used to identify the existence of a UMVUE.

(Lehmann,1983,Theorem 2.7.1)
Covariance Inequality: For an estimator δ of $g(\underline{\theta})$ and any function $\underline{\alpha}(\underline{\theta})$, which depends on the data, and has finite second moments, $\text{Var}(\delta) \geq \text{Cov}(\delta, \underline{\alpha})[\text{Cov}(\underline{\alpha})]^{-1}\text{Cov}(\underline{\alpha}, \delta)$ where equality holds if and only if $\delta = \text{Cov}(\delta, \underline{\alpha})[\text{Cov}(\underline{\alpha})]^{-1}\underline{\alpha} + c$ for some constant c.

proof: i) $\text{Var}(\delta - \text{Cov}(\delta, \underline{\alpha})[\text{Cov}(\underline{\alpha})]^{-1}\underline{\alpha}) \geq 0$
$\Leftrightarrow \text{Var}(\delta) + \text{Cov}(\delta, \underline{\alpha})[\text{Cov}(\underline{\alpha})]^{-1}[\text{Cov}(\underline{\alpha})][\text{Cov}(\underline{\alpha})]^{-1}\text{Cov}(\underline{\alpha}, \delta) - 2\text{Cov}(\delta, \underline{\alpha})[\text{Cov}(\underline{\alpha})]^{-1}\text{Cov}(\underline{\alpha}, \delta) \geq 0$
$\Leftrightarrow \text{Var}(\delta) - \text{Cov}(\delta, \underline{\alpha})[\text{Cov}(\underline{\alpha})]^{-1}\text{Cov}(\underline{\alpha}, \delta) \geq 0.$
ii) The inequality in i) is an equality $\Leftrightarrow \delta = \text{Cov}(\delta, \underline{\alpha})[\text{Cov}(\underline{\alpha})]^{-1}\underline{\alpha} + c$ for some constant c. ∎

The Covariance Inequality cannot be used directly since the right hand side depends on δ through $\text{Cov}(\delta, \underline{\alpha})$. For particular choices of δ and $\underline{\alpha}$, it will be the case that $\text{Cov}(\delta, \underline{\alpha})$ only depends on the parameter $\underline{\theta}$ and not on δ. For this purpose, some definitions and notation will be used from likelihood theory (Lehmann,1983,Ch.2). Consider a family of distributions $P_{\underline{\theta}}$ for $\underline{\theta} \in \Upsilon \subset \mathcal{R}^p$ where Υ contains a non-empty open set in $\mathcal{R}^p$. Suppose the distribution $P_{\underline{\theta}}$ has density $p_{\underline{\theta}}$. The following definitions and notation are useful:

Definitions: Likelihood Function: $L(\underline{\theta}|\underline{w}) = p_{\underline{\theta}}(\underline{w})$

Log-Likelihood Function: $l(\underline{\theta}|\underline{w}) = \ln L(\underline{\theta}|\underline{w})$

Score Statistic: $\underline{u}(\underline{\theta}|\underline{w}) = \{\frac{\partial}{\partial \theta_i} l(\underline{\theta}|\underline{w})\}_{p\times 1}$

Information Matrix: $i(\underline{\theta}) = \mathrm{Cov}_{\underline{\theta}}(\underline{u}(\underline{\theta}|\underline{w})) = \{\mathrm{Cov}_{\underline{\theta}}(\frac{\partial}{\partial \theta_i} l(\underline{\theta}|\underline{w}), \frac{\partial}{\partial \theta_j} l(\underline{\theta}|\underline{w}))\}_{p\times p}$.

These definitions as well as the following results require certain regularity conditions pertaining to the existence of derivatives, the existence of expectations, and the ability to interchange differentiation and expectation (Lehmann,1983,p125-6). These conditions are met for the normal family of distributions where the parameter space contains a non-empty open set (Lehmann,1986,Theorem 2.9). Only this family of distributions is of particular interest, so it will be assumed that the regularity conditions are satisfied.

(Lehmann,1983,Lemma 2.6.1)

Lemma 2: i) $\mathrm{E}_{\underline{\theta}}[\underline{u}(\underline{\theta}|\underline{w})] = 0$

ii) $\mathrm{Cov}_{\underline{\theta}}(\delta, \underline{u}(\underline{\theta}|\underline{w})) = \frac{\partial}{\partial \underline{\theta}} \mathrm{E}_{\underline{\theta}}[\delta] \;\; \forall\, \delta$ with finite second moments.

iii) $i(\underline{\theta}) = \{ -\mathrm{E}[\frac{\partial^2}{\partial\theta_i \partial\theta_j} p_{\underline{\theta}}]\}$.

proof: i) $\mathrm{E}_{\underline{\theta}}[\underline{u}(\underline{\theta}|\underline{w})] = \mathrm{E}_{\underline{\theta}}[\frac{\partial}{\partial \underline{\theta}} \ln p_{\underline{\theta}}(\underline{w})\,] = \mathrm{E}_{\underline{\theta}}[\frac{\frac{\partial}{\partial \underline{\theta}} p_{\underline{\theta}}(\underline{w})}{p_{\underline{\theta}}(\underline{w})}] = \int \frac{\frac{\partial}{\partial \underline{\theta}} p_{\underline{\theta}}(\underline{w})}{p_{\underline{\theta}}(\underline{w})} p_{\underline{\theta}}(\underline{w}) d\underline{w} = \frac{\partial}{\partial \underline{\theta}} \int p_{\underline{\theta}}(\underline{w}) d\underline{w} = \frac{\partial}{\partial \underline{\theta}} 1 = 0.$

ii) $\mathrm{Cov}_{\underline{\theta}}(\delta, \underline{u}(\underline{\theta}|\underline{w})) = \mathrm{E}_{\underline{\theta}}[\delta, \underline{u}(\underline{\theta}|\underline{w})] - \mathrm{E}_{\underline{\theta}}[\delta]\mathrm{E}_{\underline{\theta}}[\underline{u}(\underline{\theta}|\underline{w})] = \mathrm{E}_{\underline{\theta}}[\delta, \underline{u}(\underline{\theta}|\underline{w})]$ by i)

$= \int \delta(\underline{w}) \underline{u}(\underline{\theta}|\underline{w}) p_{\underline{\theta}}(\underline{w}) d\underline{w} = \int \delta(\underline{w}) (\frac{\partial}{\partial \underline{\theta}} \ln p_{\underline{\theta}}(\underline{w})) p_{\underline{\theta}}(\underline{w}) d\underline{w}$ by definition of $\underline{u}(\underline{\theta}|\underline{w})$

$= \int \delta(\underline{w}) \frac{\partial}{\partial \underline{\theta}} p_{\underline{\theta}}(\underline{w}) d\underline{w} = \frac{\partial}{\partial \underline{\theta}} \int \delta(\underline{w}) p_{\underline{\theta}}(\underline{w}) d\underline{w} = \frac{\partial}{\partial \underline{\theta}} \mathrm{E}_{\underline{\theta}}[\delta\,]$.

iii) Note $\frac{\partial^2}{\partial\theta_i \partial\theta_j} \ln p_{\underline{\theta}}(\underline{w}) = \frac{1}{p_{\underline{\theta}}} (\frac{\partial^2}{\partial\theta_i \partial\theta_j} p_{\underline{\theta}}(\underline{w})) - \frac{1}{p_{\underline{\theta}}^2} \frac{\partial}{\partial\theta_i} p_{\underline{\theta}}(\underline{w}) \frac{\partial}{\partial\theta_j} p_{\underline{\theta}}(\underline{w})$. By definition of $i(\underline{\theta})$,

$i_{ij}(\underline{\theta}) = \mathrm{Cov}_{\underline{\theta}}(\frac{\partial}{\partial\theta_i} l(\underline{\theta}|\underline{w}), \frac{\partial}{\partial\theta_j} l(\underline{\theta}|\underline{w})) = \mathrm{E}_{\underline{\theta}}[\frac{\partial}{\partial\theta_i} l(\underline{\theta}|\underline{w}) \frac{\partial}{\partial\theta_j} l(\underline{\theta}|\underline{w})]$ by i)

$= \mathrm{E}_{\underline{\theta}}[\frac{1}{p_{\underline{\theta}}^2} \frac{\partial}{\partial\theta_i} p_{\underline{\theta}}(\underline{w}) \frac{\partial}{\partial\theta_j} p_{\underline{\theta}}(\underline{w})] = \mathrm{E}[\frac{-\partial^2}{\partial\theta_i\partial\theta_j} p_{\underline{\theta}}(\underline{w}) + \frac{1}{p_{\underline{\theta}}} (\frac{\partial^2}{\partial\theta_i\partial\theta_j} p_{\underline{\theta}}(\underline{w}))]$

$= \mathrm{E}[\frac{-\partial^2}{\partial\theta_i\partial\theta_j} p_{\underline{\theta}}(\underline{w})\,] + \frac{\partial}{\partial\theta_i} \mathrm{E}_{\underline{\theta}}[\frac{\partial}{\partial\theta_j} \ln p_{\underline{\theta}}(\underline{w})\,] = -\mathrm{E}[\frac{-\partial^2}{\partial\theta_i\partial\theta_j} p_{\underline{\theta}}\,]$ by i). ∎

(Lehmann,1983,Theorem 2.7.3)

Lemma 3: For an unbiased estimator $\widehat{\underline{\theta}}$ of $g(\underline{\theta})$ and $\forall\, \underline{d}$, $\mathrm{Var}_{\underline{\theta}}(\underline{d}'\widehat{\underline{\theta}}) \geq \underline{d}'[i(\underline{\theta})]^{-1}\underline{d}$ with equality if and only if $\underline{d}'\widehat{\underline{\theta}} = \underline{d}'[i(\underline{\theta})]^{-1}\underline{u}(\underline{\theta}) + c$ for some constant $c = c(\underline{\theta}, \underline{d})$.

proof: Apply the Covariance Inequality with $\delta = \underline{d}'\widehat{\underline{\theta}}$ and $\underline{\alpha}(\underline{\theta}|\underline{w}) = \underline{u}(\underline{\theta}|\underline{w})$ noting that $\mathrm{Cov}(\underline{u}(\underline{\theta}|\underline{w})) = i(\underline{\theta})$ by definition and $\mathrm{Cov}(\underline{d}'\widehat{\underline{\theta}}, \underline{u}(\underline{\theta}|\underline{w})) = \frac{\partial}{\partial \underline{\theta}} \mathrm{E}_{\underline{\theta}}[\underline{d}'\widehat{\underline{\theta}}] = \frac{\partial}{\partial \underline{\theta}} \underline{d}'\underline{\theta} = \underline{d}$ by lemma 2. ∎

7.2.2. UMVUE Results

The next step is to examine whether the quantities $\Lambda'\widehat{\underline{\beta}}_{\widehat{\underline{\psi}}_{MLQ}}$, $\Gamma'\widehat{\underline{\psi}}_{MLQ}$, and $\Gamma'\widehat{\underline{\psi}}_{REMLQ}$ can be written in the linear form given in lemma 3 for attaining the lower bound. For the ML and REML methods, assumptions [L], [O], and [N] are being used for the Y-Model. Lemma 1 in section 3.2.3 can provide a convenient representation of the score function and the information matrix, given in section 3.1.4, for the ML and REML estimation methods. These representations are given below:

$$\text{ML:} \quad \underline{u}(\underline{\beta},\underline{\psi}) = \begin{bmatrix} \underline{u}_1(\underline{\beta},\underline{\psi}) \\ \underline{u}_2(\underline{\beta},\underline{\psi}) \end{bmatrix} = \begin{bmatrix} X'V_{\underline{\psi}}^{-1}\underline{Y} - X'V_{\underline{\psi}}^{-1}X\underline{\beta} \\ X^{\diamond *}V_{\underline{\psi}}^{\diamond -1}Y_1^{\diamond} - X^{\diamond *}V_{\underline{\psi}}^{\diamond -1}X^{\diamond}\underline{\psi} \end{bmatrix}$$

$$i(\underline{\beta},\underline{\psi}) = \text{diag}(i_{11}(\underline{\beta},\underline{\psi}), i_{22}(\underline{\beta},\underline{\psi})) = \begin{bmatrix} X'V_{\underline{\psi}}^{-1}X & 0 \\ 0 & X^{\diamond *}V_{\underline{\psi}}^{\diamond -1}X^{\diamond} \end{bmatrix}$$

$$\text{REML:} \quad \underline{u}_R(\underline{\psi}) = X^{\circ *}V_{\underline{\psi}}^{\circ -1}Y^{\circ} - X^{\circ *}V_{\underline{\psi}}^{\circ -1}X^{\circ}\underline{\psi} \qquad i_R(\underline{\psi}) = X^{\circ *}V_{\underline{\psi}}^{\circ -1}X^{\circ}.$$

Note that $\underline{u}_2(\underline{\beta},\underline{\psi})$ depends on $Y_1^{\diamond}$, not $Y_2^{\diamond}$, as $\underline{\beta}$ and $\underline{\psi}$ are not being estimated. In order for the information matrix to be PD, it is necessary to assume $\underline{\beta}$ is mean estimable in the Y-Model, $\underline{\psi}$ is mean estimable in the ALQEM for $(\underline{Y} - X\widehat{\underline{\beta}})$, and $\underline{\psi}$ is mean estimable in the LQEM for $N_X\underline{Y}$. The following lemma demonstrates that the GLSE for these models are linearly related to the score statistic. The expression for the GLSE's can be found in chapter 4. Due to the issue of the response in the ALQEM for the ML method, let $\widehat{\underline{\psi}}_{\underline{\psi}}(Y_2^{\diamond})$ denote the GLSE given in section 4.3.2 and let $\widehat{\underline{\psi}}_{\underline{\psi}}(Y_1^{\diamond})$ denote the same expression using $Y_1^{\diamond}$ instead of $Y_2^{\diamond}$. In addition, let $\widehat{\underline{\psi}}(Y_2^{\diamond})$ denote the EGLSE given in section 4.3.2.

Lemma 4: i) For the ML method, $\widehat{\underline{\beta}}_{\underline{\psi}} = \underline{\beta} + [i_{11}(\underline{\beta},\underline{\psi})]^{-1}\underline{u}_1(\underline{\beta},\underline{\psi})$.

ii) For the ML method, $\widehat{\underline{\psi}}_{\underline{\psi}}(Y_1^{\diamond}) = \underline{\psi} + [i_{22}(\underline{\beta},\underline{\psi})]^{-1}\underline{u}_2(\underline{\beta},\underline{\psi})$.

iii) For the REML method, $\widehat{\underline{\psi}}_{\underline{\psi}} = \underline{\psi} + [i_R(\underline{\psi})]^{-1}\underline{u}_R(\underline{\psi})$.

proof: i) $\widehat{\underline{\beta}}_{\underline{\psi}} = (X'V_{\underline{\psi}}^{-1}X)^{-1}X'V_{\underline{\psi}}^{-1}\underline{Y} = (X'V_{\underline{\psi}}^{-1}X)^{-1}(X'V_{\underline{\psi}}^{-1}\underline{Y} - X'V_{\underline{\psi}}^{-1}X\underline{\beta}) + \underline{\beta}$

$= \underline{\beta} + [i_{11}(\underline{\beta},\underline{\psi})]^{-1}\underline{u}_1(\underline{\beta},\underline{\psi})$.

ii) $\widehat{\underline{\psi}}_{\underline{\psi}}(Y_1^{\diamond}) = (X^{\diamond *}V_{\underline{\psi}}^{\diamond}X^{\diamond})^{-1}X^{\diamond *}V_{\underline{\psi}}^{\diamond -1}Y_1^{\diamond} = (X^{\diamond *}V_{\underline{\psi}}^{\diamond}X^{\diamond})^{-1}[X^{\diamond *}V_{\underline{\psi}}^{\diamond -1}Y_1^{\diamond} - X^{\diamond *}V_{\underline{\psi}}^{\diamond}X^{\diamond}\underline{\psi}] + \underline{\psi}$

$= \underline{\psi} + [i_{22}(\underline{\beta},\underline{\psi})]^{-1}\underline{u}_2(\underline{\beta},\underline{\psi})$.

iii) $\widehat{\underline{\psi}}_{\underline{\psi}} = (X^{\circ *}V_{\underline{\psi}}^{\circ}X^{\circ})^{-1}X^{\circ *}V_{\underline{\psi}}^{\circ -1}Y^{\circ} = (X^{\circ *}V_{\underline{\psi}}^{\circ -1}X^{\circ})^{-1}[X^{\circ *}V_{\underline{\psi}}^{\circ -1}Y^{\circ} - X^{\circ *}V_{\underline{\psi}}^{\circ -1}X^{\circ}\underline{\psi}] + \underline{\psi}$

$= \underline{\psi} + [i_R(\underline{\psi})]^{-1}\underline{u}_R(\underline{\psi})$. ∎

Corollary: For the ML method, $\widehat{\underline{\psi}}(Y_2^\circ) = \widehat{\underline{\psi}} + [i_{22}(\widehat{\underline{\beta}},\widehat{\underline{\psi}})]^{-1}\underline{u}_2(\widehat{\underline{\beta}},\widehat{\underline{\psi}})$.

proof: $\widehat{\underline{\psi}}(Y_2^\circ) = (X^{\circ *}V_{\widehat{\underline{\psi}}}^\circ X^\circ)^{-1}X^{\circ *}V_{\widehat{\underline{\psi}}}^{\circ -1}Y_2^\circ = (X^{\circ *}V_{\widehat{\underline{\psi}}}^\circ X^\circ)^{-1}[X^{\circ *}V_{\widehat{\underline{\psi}}}^{\circ -1}Y_2^\circ - X^{\circ *}V_{\widehat{\underline{\psi}}}^\circ X^\circ\widehat{\underline{\psi}}] + \widehat{\underline{\psi}}$

$= \widehat{\underline{\psi}} + [i_{22}(\widehat{\underline{\beta}},\widehat{\underline{\psi}})]^{-1}\underline{u}_2(\widehat{\underline{\beta}},\widehat{\underline{\psi}})$. ■

The following two results indicate that under the appropriate GZC conditions, $\Lambda'\widehat{\underline{\beta}}_{\widehat{\underline{\psi}}_{\text{MLQ}}}$, $\Gamma'\widehat{\underline{\psi}}_{\text{MLQ}}(Y_1^\circ)$ where $\underline{\beta}$ is known, and $\Gamma'\widehat{\underline{\psi}}_{\text{REMLQ}}$ are UMVUE. These expressions are the same as those for the GLSE except they use $\widehat{\underline{\psi}}_{\text{MLQ}}$ and $\widehat{\underline{\psi}}_{\text{REMLQ}}$ as estimates in place of $\underline{\psi}$. Such expressions, as well as expressions for the LSEs, are given in chapter 5. The following results assume $\widehat{\underline{\psi}}_{\text{MLQ}}$ and $\widehat{\underline{\psi}}_{\text{REMLQ}}$ exist. The GZC conditions are conveniently referenced below. Note that the conditions do not depend upon the response.

GZC-1: $\underline{R}(V_{\underline{\psi}}X(X'X)^-\Lambda) \subset \underline{R}(X) \quad \forall\ \underline{\psi} \in \Xi$

GZC-2: $\underline{R}(V_{\underline{\psi}}^\circ X^\circ(X^{\circ *}X^\circ)^-\Gamma) \subset \underline{R}(X^\circ) \quad \forall\ \underline{\psi} \in \Xi$

GZC-3: $\underline{R}(V_{\underline{\psi}}^\circ X^\circ(X^{\circ *}X^\circ)^-\Gamma) \subset \underline{R}(X^\circ) \quad \forall\ \underline{\psi} \in \Xi$.

Theorem 5: i) If $V_{\widehat{\underline{\psi}}_{\text{MLQ}}}$ is PD, then GZC-1 $\Leftrightarrow \Lambda'\widehat{\underline{\beta}}_{\widehat{\underline{\psi}}_{\text{MLQ}}} = \Lambda'\underline{\beta} + \Lambda'[i_{11}(\underline{\beta},\underline{\psi})]^{-1}\underline{u}_1(\underline{\beta},\underline{\psi})\ \forall\ \underline{\psi} \in \Xi$.
ii) If $V_{\widehat{\underline{\psi}}_{\text{MLQ}}}$ is PD and $\underline{\beta}$ is known, then GZC-2 $\Leftrightarrow \Gamma'\widehat{\underline{\psi}}_{\widehat{\underline{\psi}}_{\text{MLQ}}}(Y_1^\circ) = \Gamma'\underline{\psi} + \Gamma'[i_{22}(\underline{\beta},\underline{\psi})]^{-1}\underline{u}_2(\underline{\beta},\underline{\psi})\ \forall \underline{\psi} \in \Xi$.
iii) If $V_{\widehat{\underline{\psi}}_{\text{REMLQ}}}$ is PD, then GZC-3 $\Leftrightarrow \Gamma'\widehat{\underline{\psi}}_{\widehat{\underline{\psi}}_{\text{REMLQ}}} = \Gamma'\underline{\psi} + \Gamma'[i_R(\underline{\psi})]^{-1}\underline{u}_R(\underline{\psi})\ \forall\ \underline{\psi} \in \Xi$.

proof: i) (1) Suppose GZC-1. By the proposition in section 5.2.1, $\forall\ \underline{\psi} \in \mathcal{R}^{k+1} \ni V_{\underline{\psi}}$ is PD,
$\Lambda'\widehat{\underline{\beta}}_{\widehat{\underline{\psi}}_{\text{MLQ}}} = \Lambda'\widehat{\underline{\beta}}_I = \Lambda'\widehat{\underline{\beta}}_{\underline{\psi}} = \Lambda'(X'V_{\underline{\psi}}^{-1}X)^{-1}X'V_{\underline{\psi}}^{-1}\underline{Y} = \Lambda'\underline{\beta} + \Lambda'[i_{11}(\underline{\beta},\underline{\psi})]^{-1}\underline{u}_1(\underline{\beta},\underline{\psi})$ by lemma 4.
(2) Let $\underline{\psi}_0 = [0, ..., 0, 1]' \in \Xi \Rightarrow V_{\underline{\psi}_0} = I$. Suppose $\Lambda'\widehat{\underline{\beta}}_{\widehat{\underline{\psi}}_{\text{MLQ}}} = \Lambda'\underline{\beta} + \Lambda'[i_{11}(\underline{\beta},\underline{\psi})]^{-1}\underline{u}_1(\underline{\beta},\underline{\psi})\ \forall\ \underline{\psi} \in \Xi$
$\Rightarrow \Lambda'\widehat{\underline{\beta}}_{\widehat{\underline{\psi}}_{\text{MLQ}}} = \Lambda'(X'V_{\underline{\psi}}^{-1}X)^{-1}X'V_{\underline{\psi}}^{-1}\underline{Y}\ \ \forall \underline{\psi} \in \Xi$ by lemma 4. In addition, the above holds for $\underline{\psi}_0 \in \Xi$, so $\Lambda'\widehat{\underline{\beta}}_{\widehat{\underline{\psi}}_{\text{MLQ}}} = \Lambda'(X'X)^{-1}X'\underline{Y}$ from above. Thus, $\Lambda'(X'V_{\underline{\psi}}^{-1}X)^{-1}X'V_{\underline{\psi}}^{-1}\underline{Y} = \Lambda'(X'X)^{-1}X'\underline{Y}\ \forall \underline{\psi} \in \Xi$
$\Rightarrow$ GZC-1 by the Y-UBLUE theorem.
ii) Since $\underline{\beta}$ is known, $\widehat{\underline{\psi}}_{\widehat{\underline{\psi}}_{\text{MLQ}}}$ is a function of Y_1° which can be seen in section 3.2.3.
(1) Suppose GZC-2. By the proposition in section 5.3.1, $\forall\ \underline{\psi} \in \mathcal{R}^{k+1} \ni V_{\underline{\psi}}$ is PD,
$\Gamma'\widehat{\underline{\psi}}_{\widehat{\underline{\psi}}_{\text{MLQ}}}(Y_1^\circ) = \Gamma'\widehat{\underline{\psi}}_I(Y_1^\circ) = \Gamma'\widehat{\underline{\psi}}_{\underline{\psi}}(Y_1^\circ) = \Gamma'(X^{\circ *}V_{\underline{\psi}}^{\circ -1}X^\circ)^{-1}X^{\circ *}V_{\underline{\psi}}^{\circ -1}Y_1^\circ$
$= \Gamma'\underline{\psi} + \Gamma'[i_{22}(\underline{\beta},\underline{\psi})]^{-1}\underline{u}_2(\underline{\beta},\underline{\psi})$ by lemma 4.
(2) Let $\underline{\psi}_0 = [0, ..., 0, 1]' \in \Xi \Rightarrow V_{\underline{\psi}_0} = I \Rightarrow V_{\underline{\psi}_0}^\circ = I$.
Suppose $\Gamma'\widehat{\underline{\psi}}_{\widehat{\underline{\psi}}_{\text{MLQ}}}(Y_1^\circ) = \Gamma'\underline{\psi} + \Gamma'[i_{22}(\underline{\beta},\underline{\psi})]^{-1}\underline{u}_2(\underline{\beta},\underline{\psi})\ \ \forall\ \underline{\psi} \in \Xi$
$\Rightarrow \Gamma'\widehat{\underline{\psi}}_{\widehat{\underline{\psi}}_{\text{MLQ}}}(Y_1^\circ) = \Gamma'(X^{\circ *}V_{\underline{\psi}}^{\circ -1}X^\circ)^{-1}X^{\circ *}V_{\underline{\psi}}^{\circ -1}Y_1^\circ\ \forall\ \underline{\psi} \in \Xi$ by lemma 4.
In addition the above holds for $\underline{\psi}_0 \in \Xi$, so $\Gamma'\widehat{\underline{\psi}}_{\widehat{\underline{\psi}}_{\text{MLQ}}}(Y_1^\circ) = \Gamma'(X^{\circ *}X^\circ)^{-1}X^{\circ *}Y_1^\circ$ from above.

Thus, $\Gamma'(X^{\circ *}V_{\underline{\psi}}^{\circ -1}X^{\circ})^{-1}X^{\circ *}V_{\underline{\psi}}^{\circ -1}Y_1^{\circ} = \Gamma'(X^{\circ *}X^{\circ})^{-1}X^{\circ *}Y_1^{\circ} \ \ \forall\, \underline{\psi} \in \Xi$

$\Rightarrow$ GZC-2 by the LQZ-UBLUE theorem.

iii) (1) Suppose GZC-3. By the proposition in section 5.3.1, $\forall\, \underline{\psi} \in \mathcal{R}^{k+1} \ni V_{\underline{\psi}}$ is PD,

$\Gamma'\widehat{\underline{\psi}}_{\widehat{\underline{\psi}}_{\text{REMLQ}}} = \Gamma'\widehat{\underline{\psi}}_I = \Gamma'\widehat{\underline{\psi}}_{\underline{\psi}} = \Gamma'(X^{\circ *}V_{\underline{\psi}}^{\circ -1}X^{\circ})^{-1}X^{\circ *}V_{\underline{\psi}}^{\circ -1}Y^{\circ} = \Gamma'\underline{\psi} + \Gamma'[i_R(\underline{\psi})]^{-1}\underline{u}_R(\underline{\psi})$ by lemma 4.

(2) Let $\underline{\psi}_0 = [0, \ldots, 0, 1]' \in \Xi \Rightarrow V_{\underline{\psi}_0} = I \Rightarrow V_{\underline{\psi}_0}^{\circ} = N_X$. By lemma 1 in section 3.2.3,

$X^{\circ *}V_{\underline{\psi}_0}^{\circ -1}A = \frac{1}{2}\{\text{tr}(N_X V_i N_X (N_X)^+ A (N_X)^+)\} = \frac{1}{2}\{\text{tr}((N_X)^+ N_X N_X V_i N_X N_X (N_X)^+ A)\}$

$= \frac{1}{2}\{\text{tr}(P_{N_X} N_X V_i N_X P_{N_X} A)\}$ as $P_{N_X} = (N_X)^+ N_X = N_X (N_X)^+$.

$= \frac{1}{2}\{\text{tr}(N_X V_i N_X A)\} = X^{\circ *}A$ by lemma 1 in section 3.2.3.

Suppose $\Gamma'\widehat{\underline{\psi}}_{\widehat{\underline{\psi}}_{\text{REMLQ}}} = \Gamma'\underline{\psi} + \Gamma'[i_R(\underline{\psi})]^{-1}\underline{u}_R(\underline{\psi}) \ \ \forall\, \underline{\psi} \in \Xi$

$\Rightarrow \Gamma'\widehat{\underline{\psi}}_{\widehat{\underline{\psi}}_{\text{REMLQ}}} = \Gamma'(X^{\circ *}V_{\underline{\psi}}^{\circ -1}X^{\circ})^{-1}X^{\circ *}V_{\underline{\psi}}^{\circ -1}Y^{\circ} \ \ \forall\, \underline{\psi} \in \Xi$ by lemma 4.

In addition the above holds for $\underline{\psi}_0 \in \Xi$, so $\Gamma'\widehat{\underline{\psi}}_{\widehat{\underline{\psi}}_{\text{REMLQ}}} = \Gamma'(X^{\circ *}X^{\circ})^{-1}X^{\circ *}Y^{\circ}$ from above.

Thus, $\Gamma'(X^{\circ *}V_{\underline{\psi}}^{\circ -1}X^{\circ})^{-1}X^{\circ *}V_{\underline{\psi}}^{\circ -1}Y^{\circ} = \Gamma'(X^{\circ *}X^{\circ})^{-1}X^{\circ *}Y^{\circ} \ \ \forall\, \underline{\psi} \in \Xi$

$\Rightarrow$ GZC-3 by the LQZ-UBLUE theorem. ■

Theorem 6: i) $V_{\widehat{\underline{\psi}}_{\text{MLQ}}}$ PD and GZC-1 $\Rightarrow \Lambda'\widehat{\underline{\beta}}_I$ is ELMLQE, UBLUE, and UMVUE in the Y-Model.

ii) $V_{\widehat{\underline{\psi}}_{\text{MLQ}}}$ PD, $\underline{\beta}$ known, and GZC-2 $\Rightarrow \Gamma'\widehat{\underline{\psi}}_I(Y_1^{\circ})$ is ELMLQE, UBLUE, and UMVUE in the Y-Model.

iii) $V_{\widehat{\underline{\psi}}_{\text{REMLQ}}}$ PD and GZC-3 $\Rightarrow \Gamma'\widehat{\underline{\psi}}_I$ is ELREMLQE, UBLUE, and UMVUE in the QY-Model.

proof: i) From proof of theorem 5, $\Lambda'\widehat{\underline{\beta}}_{\widehat{\underline{\psi}}_{\text{MLQ}}} = \Lambda'\widehat{\underline{\beta}}_I = \Lambda'\widehat{\underline{\beta}}_{\underline{\psi}} = \Lambda'\underline{\beta} + \Lambda'[i_{11}(\underline{\beta},\underline{\psi})]^{-1}\underline{u}_1(\underline{\beta},\underline{\psi}) \ \ \forall\, \underline{\psi} \in \Xi$

$\Rightarrow \text{Var}_{\underline{\psi}}(\underline{d}'\Lambda'\widehat{\underline{\beta}}_I)$ is a minimum for unbiased estimators of $\underline{d}'\Lambda'\underline{\beta} \ \ \forall\, \underline{\psi} \in \Xi, \underline{d}$ by lemma 3

$\Rightarrow \Lambda'\widehat{\underline{\beta}}_I$ is a UMVUE in the Y-Model by definition.

Also, by the Y-UBLUE and Y-ELMLQE theorems, $\Lambda'\widehat{\underline{\beta}}_I$ is UBLUE and ELMLQE.

ii) From theorem 5, $\Gamma'\widehat{\underline{\psi}}_{\widehat{\underline{\psi}}_{\text{MLQ}}}(Y_1^{\circ}) = \Gamma'\widehat{\underline{\psi}}_I(Y_1^{\circ}) = \Gamma'\widehat{\underline{\psi}}_{\underline{\psi}}(Y_1^{\circ}) = \Gamma'\underline{\psi} + \Gamma'[i_{22}(\underline{\beta},\underline{\psi})]^{-1}\underline{u}_2(\underline{\beta},\underline{\psi}) \ \ \forall\, \underline{\psi} \in \Xi$

$\Rightarrow \text{Var}_{\underline{\psi}}(\underline{d}'\Gamma'\widehat{\underline{\psi}}_I(Y_1^{\circ}))$ is a minimum for unbiased estimators of $\underline{d}'\Gamma'\underline{\psi} \ \ \forall\, \underline{\psi} \in \Xi, \underline{d}$ by lemma 3

$\Rightarrow \Gamma'\widehat{\underline{\psi}}_I(Y_1^{\circ})$ is a UMVUE in the Y-Model by definition.

Also, by the LQZ-UBLUE and ALQNY-ELMLQE theorems, $\Gamma'\widehat{\underline{\psi}}_I(Y_1^{\circ})$ is UBLUE and ELMLQE as Y_1° does not depend on any unknown parameters.

iii) As shown in proof of theorem 5, $\Gamma'\widehat{\underline{\psi}}_{\widehat{\underline{\psi}}_{\text{REMLQ}}} = \Gamma'\widehat{\underline{\psi}}_I = \Gamma'\widehat{\underline{\psi}}_{\underline{\psi}} = \Gamma'\underline{\psi} + \Gamma'[i_R(\underline{\psi})]^{-1}\underline{u}_R(\underline{\psi}) \ \ \forall\, \underline{\psi} \in \Xi$

$\Rightarrow \text{Var}_{\underline{\psi}}(\underline{d}'\Gamma'\widehat{\underline{\psi}}_I)$ is a minimum for unbiased estimators of $\underline{d}'\Gamma'\underline{\psi} \ \ \forall\, \underline{\psi} \in \Xi, \underline{d}$ by lemma 3

$\Rightarrow \Gamma'\widehat{\underline{\psi}}_I$ is a UMVUE in the QY-Model by definition.

Also, by the LQZ-UBLUE and LQNY-ELREMLQE theorems, $\Gamma'\widehat{\underline{\psi}}_I$ is UBLUE and ELREMLQE. ■

Theorem 6 indicates that, given a PD covariance, the GZC is a sufficient condition for the existence of a UMVUE in the appropriate family of distributions. The GZC and UMVUE results stated in theorem 6 are not equivalent since the UMVUE does not necessarily imply that the CRLB is attained (Casella and Berger,1990,p314).

The results of theorem 5 can be used to obtain expresssions for $\mathrm{Cov}_{\underset{\sim}{\psi}}(\Lambda'\widehat{\underset{\sim}{\beta}}_{\mathrm{MLQ}})$, $\mathrm{Cov}_{\underset{\sim}{\psi}}(\Gamma'\widehat{\underset{\sim}{\psi}}_{\mathrm{MLQ}}(Y_1^\diamond))$, and $\mathrm{Cov}_{\underset{\sim}{\psi}}(\Gamma'\widehat{\underset{\sim}{\psi}}_{\mathrm{REMLQ}})$. Generally, such covariance expressions cannot be obtained, since these quantities are not linear in terms of $\underset{\sim}{Y}$, $Y_1^\diamond$, and Y°, respectively.

Corollary: Suppose the conditions in theorem 5 hold for the three cases. Then

i) $\mathrm{Cov}_{\underset{\sim}{\psi}}(\Lambda'\widehat{\underset{\sim}{\beta}}_{\mathrm{MLQ}}) = \Lambda'(X'V_{\underset{\sim}{\psi}}^{-1}X)^{-1}\Lambda = \Lambda'[i_{11}(\underset{\sim}{\beta},\underset{\sim}{\psi})]^{-1}\Lambda$

ii) $\mathrm{Cov}_{\underset{\sim}{\psi}}(\Gamma'\widehat{\underset{\sim}{\psi}}_{\mathrm{MLQ}}(Y_1^\diamond)) = \Gamma'(X^{\diamond *}V_{\underset{\sim}{\psi}}^{\diamond -1}X^\diamond)^{-1}\Gamma = \Gamma'[i_{22}(\underset{\sim}{\beta},\underset{\sim}{\psi})]^{-1}\Gamma$

iii) $\mathrm{Cov}_{\underset{\sim}{\psi}}(\Gamma'\widehat{\underset{\sim}{\psi}}_{\mathrm{REMLQ}}) = \Gamma'(X^{\circ *}V_{\underset{\sim}{\psi}}^{\circ -1}X^\circ)^{-1}\Gamma = \Gamma'[i_R(\underset{\sim}{\beta},\underset{\sim}{\psi})]^{-1}\Gamma.$

proof: These above covariance expressions can be obtained using the expressions in theorem 5, where the information matrix is constant and the score statistic is a random quantity. ∎

In general, such expressions for the covariance cannot be obtained. Searle et al. (1992) and Miller (1977), recommend using the expressions in i), ii), and iii) as approximations for the covariance of $\Lambda'\widehat{\underset{\sim}{\beta}}_{\mathrm{MLQ}}$, $\Gamma'\widehat{\underset{\sim}{\psi}}_{\mathrm{MLQ}}$, and $\Gamma'\widehat{\underset{\sim}{\psi}}_{\mathrm{REMLQ}}$, respectively. Their recommendation is due to the fact that these are the asymptotic expressions for the covariance. However, under the UBLUE conditions, the above corollary shows these covariance expressions are exact.

SAS (1996) uses the recommended approximations in its covariance calculations for the ML and REML methods in the PROC MIXED procedure. Since the unknown parameter $\underset{\sim}{\psi}$ is involved in the expression, it must be estimated. This is typically done using the ML or REML estimate of $\underset{\sim}{\psi}$ to calculate the estimated covariance. It should be noted that the estimated covariance is not equivalent to the exact expression, even under the UBLUE conditions. It is not clear how well these estimates perform (Searle et al.,1992,p320).

An exact expression still has not been examined for $\mathrm{Cov}_{\underset{\sim}{\psi}}(\Gamma'\widehat{\underset{\sim}{\psi}}_{\mathrm{MLQ}})$ when $\underset{\sim}{\beta}$ is unknown. This situation requires the use of the response $Y_2^\diamond$ which did not fit into the above formulation. However, an exact form can be given under GZC-2 and ZC for the Y-Model. The exact form is not the same as the expression in ii) of the above corollary since $\mathrm{Cov}_{\underset{\sim}{\psi}}(Y_2^\diamond) = V_{\underset{\sim}{\psi}}^\circ$ instead of $V_{\underset{\sim}{\psi}}^\diamond$ under ZC for the Y-Model.

Theorem: For the ALQEM for $(\underset{\sim}{Y} - X\widehat{\underset{\sim}{\beta}})$ under GZC-2 where ZC holds for the Y-Model and $V_{\widehat{\underset{\sim}{\psi}}_{\mathrm{MLQ}}}$ is PD,

$$\mathrm{Cov}_{\underset{\sim}{\psi}}(\Gamma'\widehat{\underset{\sim}{\psi}}_{\mathrm{MLQ}}(Y_2^\diamond)) = \Gamma'(X^{\diamond *}X^\diamond)^{-1}X^{\diamond *}V_{\underset{\sim}{\psi}}^\circ X^\diamond(X^{\diamond *}X^\diamond)^{-1}\Gamma.$$

proof: By the proposition in section 4.2.1, $(\underline{Y} - X\widehat{\underline{\beta}}) = N_X\underline{Y}$ under ZC $\Rightarrow Y_2^\diamond = N_X\underline{Y}\underline{Y}'N_X$.
By the proposition in section 5.3.1, $\forall\, \underline{\psi} \in \mathcal{R}^{k+1} \ni V_{\underline{\psi}}$ is PD, $\Gamma'\widehat{\underline{\psi}}_{\widehat{\underline{\psi}}_{\mathrm{MLQ}}}(Y_2^\diamond) = (X^{\diamond *}X^\diamond)^{-1}X^{\diamond *}Y_2^\diamond$. Thus,
$\mathrm{Cov}_{\underline{\psi}}(\Gamma'(X^{\diamond *}X^\diamond)^{-1}X^{\diamond *}Y_2^\diamond) = \Gamma'(X^{\diamond *}X^\diamond)^{-1}X^{\diamond *}\mathrm{Cov}(N_X\underline{Y}\underline{Y}'N_X)X^\diamond(X^{\diamond *}X^\diamond)^{-1}\Gamma$ from above
$= \Gamma'(X^{\diamond *}X^\diamond)^{-1}X^{\diamond *}V_{\underline{\psi}}^\diamond X^\diamond(X^{\diamond *}X^\diamond)^{-1}\Gamma$ from section 3.2.1. ■

This theorem gives the exact form for the covariance of the MLQ when GZC-2 and ZC hold. It would be interesting to compare the covariance estimates using the exact form in the above theorem with the asymptotic form in corollary ii).

8. Data Applications

This chapter applies the general UBLUE results obtained in chapters 4, 5, and 6 to issues that arise in the analysis of data. The applications include an iterative procedure for obtaining MLQEs and REMLQEs as well as profile likelihood calculations and computing time. The Battery Life Example (Montgomery,1991,p207) is used throughout this chapter to demonstrate the applications. The PROC MIXED procedure in SAS (SAS,1996) will also be discussed, since it is a standard statistical tool for analyzing data from mixed models.

8.1. An Iterative Procedure for Obtaining MLQEs and REMLQEs

This section presents a general procedure for calculating MLQEs and REMLQEs in the general case for linear combinations of the fixed effects $\Lambda'\underline{\beta}$ and linear combinations of the variance components $\Gamma'\underline{\psi}$. An iterative procedure will be given with respect to the existence of an explicit linear solution. Data examples will be used to illustrate the procedure.

8.1.1. The Procedure

Consider estimating the variance component vector ($\underline{\psi}$) under the ML and REML estimation methods. The likelihood equation for the ML method can be written as function of $\underline{\psi}$ only by substituting $(X'V_{\underline{\psi}}^{-1}X)^{-1}X'V_{\underline{\psi}}^{-1}\underline{Y}$ in place of $\underline{\beta}$ where $\underline{\beta}$ is estimable. This is a convenient way to have the ML and REML equations depend on $\underline{\psi}$ only (Harville,1977).

Linear quadratic estimation models were defined so that the EGLSE would correspond to either the MLQE or the REMLQE. Section 3.2.3 shows how the ALQEM for $(\underline{Y} - X\widehat{\underline{\beta}})$ can be used to obtain the ML equations for $\underline{\psi}$ and how the LQEM for $N_X\underline{Y}$ can be used to obtain the REML equations for $\underline{\psi}$. These equations are given below assuming that $\underline{\psi}$ is mean estimable in both models which is equivalent to the V_i's and the $N_XV_iN_X$'s being linearly independent. Recall $F_{\widehat{\underline{\psi}}} = V_{\widehat{\underline{\psi}}}^{-1} - V_{\widehat{\underline{\psi}}}^{-1}X(X'V_{\widehat{\underline{\psi}}}^{-1}X)^{-}X'V_{\widehat{\underline{\psi}}}^{-1}$:

$$\text{MLQE: } \widehat{\underline{\psi}} = (X^{\diamond *}V_{\widehat{\underline{\psi}}}^{\diamond -1}X^{\diamond})^{-1}X^{\diamond *}V_{\widehat{\underline{\psi}}}^{\diamond -1}Y_2^{\diamond} = \{\mathrm{tr}(V_{\widehat{\underline{\psi}}}^{-1}V_iV_{\widehat{\underline{\psi}}}^{-1}V_j)\}^{-1}\{\underline{Y}'F_{\widehat{\underline{\psi}}}V_iF_{\widehat{\underline{\psi}}}\underline{Y}\}$$
$$\text{REMLQE: } \widehat{\underline{\psi}} = (X^{\circ *}V_{\widehat{\underline{\psi}}}^{\circ -1}X^{\circ})^{-1}X^{\circ *}V_{\widehat{\underline{\psi}}}^{\circ -1}Y^{\circ} = \{\mathrm{tr}(F_{\widehat{\underline{\psi}}}V_iF_{\widehat{\underline{\psi}}}V_j)\}^{-1}\{\underline{Y}'F_{\widehat{\underline{\psi}}}V_iF_{\widehat{\underline{\psi}}}\underline{Y}\}\,.$$

These equations demonstrate that an iterative procedure is needed to identify the solution given by the MLQE and the REMLQE as both sides of the equations involve $\widehat{\underline{\psi}}$. Such an iterative procedure based on the above equations is called Anderson's Iterative Algorithm (Harville,1977). The following steps define the iterative procedure, assuming there are no parameter constraints:

i) Choose an initial starting value $\widehat{\underline{\psi}}^{(0)}$.

For $i = 0, 1, 2, \ldots$, repeat the following steps given $\widehat{\underline{\psi}}^{(i)}$.

ii) Find the covariance matrix $V_{\widehat{\underline{\psi}}^{(i)}}$ and use it to calculate the right side of the equation.

iii) Let the result in ii) be $\widehat{\underline{\psi}}^{(i+1)}$.

iv) Check if $\| \widehat{\underline{\psi}}^{(i+1)} - \widehat{\underline{\psi}}^{(i)} \| < \epsilon$. If yes, then stop, else continue.

v) Replace i by $i+1$ and go to ii).

Conditions were obtained in sections 5.3.2 and 5.3.3 under which there exists an ELMLQE or an ELREMLQE for a linear combination of the variance components given by $\Gamma'\underline{\psi}$. The ELMLQE corresponds to the LSE in the ALQEM for $(\underline{Y} - X\widehat{\underline{\beta}})$ under Zyskind's condition and the ELREMLQE corresponds to the LSE in the LQEM for $N_X\underline{Y}$. The equations for the ELMLQE and ELREMLQE are:

$$\text{ELMLQE: } \Gamma'\widehat{\underline{\psi}} = \Gamma'(X^{\circ *}X^{\circ})^{-1}X^{\circ *}Y_2^{\circ} = \Gamma'\{\mathrm{tr}(V_iV_j)\}^{-1}\{\underline{Y}'N_XV_iN_X\underline{Y}\}$$
$$\text{ELREMLQE: } \Gamma'\widehat{\underline{\psi}} = \Gamma'(X^{\circ *}X^{\circ})^{-1}X^{\circ *}Y^{\circ} = \Gamma'\mathrm{tr}(V_iN_XV_jN_X)\}^{-1}\{\underline{Y}'N_XV_iN_X\underline{Y}\}.$$

When the sufficient conditions for an ELMLQE or an ELREMLQE are satisfied, then the iterative procedure will converge in a single iteration for the linear combination $\Gamma'\underline{\psi}$ as the right side of the equation does not involve $\widehat{\underline{\psi}}$.

The scoring method is another iterative procedure that can be used to estimate $\underline{\psi}$ (Searle, et al., 1992). Consider the notation for the information matrix and the score statistic given in section 7.2.2. Note that the ML equations, when substituting $(X'V_{\underline{\psi}}^{-1}X)^{-1}X'V_{\underline{\psi}}^{-1}\underline{Y}$ in place of $\underline{\beta}$, no longer depend on $\underline{\beta}$. Thus, the score statistic can be represented as $\underline{u}(\underline{\psi}) = \underline{u}_2(\underline{\beta},\underline{\psi})$ and the information matrix as $i(\underline{\psi}) = i_{22}(\underline{\beta},\underline{\psi})$. The iterative scoring equations, for the ML and REML methods are then given by:

$$\text{MLQE: } \quad \widehat{\underline{\psi}}^{(i+1)} = \widehat{\underline{\psi}}^{(i)} + [i(\widehat{\underline{\psi}}^{(i)})]^{-1}\underline{u}(\widehat{\underline{\psi}}^{(i)})$$
$$\text{REMLQE: } \quad \widehat{\underline{\psi}}^{(i+1)} = \widehat{\underline{\psi}}^{(i)} + [i_R(\widehat{\underline{\psi}}^{(i)})]^{-1}\underline{u}_R(\widehat{\underline{\psi}}^{(i)}).$$

The same iterative steps i)-v) can be used to solve these equations. By the corollary to lemma 4 and lemma 4 iii) in section 7.2.2, the equations from the scoring method are the same as the equations in Anderson's Iterative Algorithm. From the ML theorem in section 3.2.3, $Y_2^{\diamond}$ is the appropriate response in the equations for the scoring method, since $\underline{\psi}$ is being estimated. Thus, if an ELMLQE or an ELREMLQE exist, then the equations in the scoring method will converge in a single iteration. Other iterative procedures for $\underline{\psi}$ are presented and compared in Harville (1977) and (Searle, et al.,1992).

Now consider estimating the fixed effect vector $\underline{\beta}$, where $\underline{\beta}$ is mean estimable or equivalently that the matrix X has full rank. From section 4.2.1, the EGLSE is given by:

EGLSE: $\widehat{\underline{\beta}}_{\widehat{\underline{\psi}}} = (X'V_{\widehat{\underline{\psi}}}^{-1}X)^{-1}X'V_{\widehat{\underline{\psi}}}^{-1}\underline{Y}$.

For the ML method, the solution for the fixed effect vector $\underline{\beta}$ can now be obtained using the variance component estimate from the above ML equations. In section 4.2.2, it is shown that the MLQE for $\underline{\beta}$ is the same as the EGLSE for $\underline{\beta}$ where $\widehat{\underline{\psi}} = \widehat{\underline{\psi}}_{\text{MLQ}}$. Thus, the MLQE for $\widehat{\underline{\beta}}$ depends on $\widehat{\underline{\psi}}$. However, if the sufficient conditions for an ELMLQE are satisfied for a linear combination $\Lambda'\underline{\beta}$, then $\Lambda'\widehat{\underline{\beta}}$ does not depend on $\widehat{\underline{\psi}}$ as shown in the following equation from section 5.2:

ELMLQE: $\Lambda'\widehat{\underline{\beta}} = \Lambda'(X'X)^{-1}X'\underline{Y}$.

This suggests that when an ELMLQE exists, the value of $\Lambda'\widehat{\underline{\beta}}$ will not change with the value of $\widehat{\underline{\psi}}$. One method to evaluate this is to calculate $\Lambda'\widehat{\underline{\beta}}$ for each iterative value $\widehat{\underline{\psi}}^{(i)}$ to see whether the quantity $\Lambda'\widehat{\underline{\beta}}_{\widehat{\underline{\psi}}^{(i)}}$ changes. Another method is to estimate $\underline{\psi}$ under a different procedure to obtain an estimate $\widehat{\underline{\psi}}_{\text{P}}$ where $V_{\widehat{\underline{\psi}}_{\text{P}}}$ is PD, calculate the EGLSE with $\widehat{\underline{\psi}} = \widehat{\underline{\psi}}_{\text{P}}$, and determine whether $\Lambda'\widehat{\underline{\beta}}_{\widehat{\underline{\psi}}_{\text{P}}} = \Lambda'\widehat{\underline{\beta}}_{\widehat{\underline{\psi}}_{\text{MLQ}}}$. If $\Lambda'\widehat{\underline{\beta}}_{\widehat{\underline{\psi}}_{\text{P}}} \neq \Lambda'\widehat{\underline{\beta}}_{\widehat{\underline{\psi}}_{\text{MLQ}}}$, then $\Lambda'\widehat{\underline{\beta}}_{\widehat{\underline{\psi}}_{\text{P}}}$ is not the MLQE for $\Lambda'\underline{\beta}$ and an ELMLQE does not exist for $\Lambda'\underline{\beta}$. For example, it is common practice in data analysis to calculate the EGLSE with $\widehat{\underline{\psi}} = \widehat{\underline{\psi}}_{\text{REMLQ}}$. In general, the resulting estimate $\widehat{\underline{\beta}}_{\widehat{\underline{\psi}}_{\text{REMLQ}}}$ is not the MLQE, nor the REMLQE. It is the EGLSE using the variance component estimate from the REML procedure as REML is preferred by many over ML for estimating variance components (Searle et al.,1992,sections 6.6-6.8).

The equations described in this section can be calculated using any computing language that has matrix computation ability. In particular, the PROC MIXED procedure in SAS will provide calculations of the above quantities (SAS,1996). The ML procedure in SAS does represent the maximum likelihood equations in terms of $\underline{\psi}$ only by substituting $(X'V_{\underline{\psi}}^{-1}X)^{-1}X'V_{\underline{\psi}}^{-1}\underline{Y}$ in place of $\underline{\beta}$ (SAS,1996). However, it is necessary to adjust the defaults of the MIXED procedure in SAS in order to implement the scoring method. The following options should be specified:

a) method - specifies the estimation method ML, REML, or MIVQUE0

b) nobound - no boundary constraints on the variance components

c) noprofile - includes residual component in iterations

d) scoring - uses expected hessian in estimation method (need to specify for all iterations).

The MIXED procedure allows the user to choose among the ML, REML, or MIVQUE0 methods for estimating variance components. The MIVQUE0 method is described in Searle et al. (1992,Section 11.3) and the resulting equations are identical to those for the FELREMLQE presented in section 3.2.3.

The remaining three options are necessary since the scoring method does not assume any constraints, includes the residual component as part of the overall calculations, and uses the expected hessian matrix for all iterative calculations. The expected hessian corresponds to the information matrix which is given in the above equations for the scoring method. By default, the MIXED procedure uses the observed hessian which is the matrix of second derivatives. In addition, SAS uses the information matrix with the estimated variance component vector as its estimate of the asymptotic covariance matrix of the variance components (SAS,1996) (Searle, et al.,1992,Chapter 6).

There are also options available in SAS that are helpful for interpreting the output. The itdetails option outputs the variance component parameter values at each iteration. This output indicates whether the iterative procedure for the variance components converges in a single iteration or whether some linear combination converges in a single iteration. The solution option outputs the estimates of the fixed effects. This output can indicate whether the estimates of the fixed effects are the same over different estimation methods. The asycov option outputs the asymptotic covariance matrix of the the variance components. This can be useful for purposes of interpretation. SAS offers a variety of choices for stopping rules. The default, under the absolute option, iterates until $\underline{u}(\widehat{\underline{\psi}}^{(i)})'\,[i(\widehat{\underline{\psi}}^{(i)})]^{-1}\underline{u}(\widehat{\underline{\psi}}^{(i)}) < 1 \times 10^{-8}$ where $\underline{u}(\widehat{\underline{\psi}}^{(i)})$ is the score function at $\widehat{\underline{\psi}}^{(i)}$ and $i(\widehat{\underline{\psi}}^{(i)})$ is the information matrix at $\widehat{\underline{\psi}}^{(i)}$. The absolute option prevents the criterion from being scaled by a multiple of the log likelihood function evaluated at $\widehat{\underline{\psi}}^{(i)}$ (SAS,1996).

The PROC MIXED procedure in SAS will be applied to the examples described in the following three sections. These examples illustrate the applicability of the UBLUE results to data examples.

8.1.2. Battery Life Example I

The data for this example is from Montgomery (1991,p207) and is shown in Table 8.1. The responses represent battery life (in hours) for batteries with certain material types at given temperatures. It should be noted that the design is balanced as there are four observations per treatment combination.

Table 8.1. Data for Battery Life Example

	Temperature (°F)		
Material	15	70	125
1	130 74 155 180	34 80 40 75	20 82 70 58
2	150 159 188 126	136 106 122 115	25 58 70 45
3	$\underline{138}$ $\underline{168}$ 110 160	$\underline{174}$ $\underline{150}$ 120 139	$\underline{96}$ $\underline{82}$ 104^x 60^x

Assume material [M] and temperature [T] represent random effects in a 2-way random model with interaction M*T. The variance components and the overall mean will be estimated using the MIXED procedure in SAS. The output from this procedure is summarized in Table 7.2. The SAS code used to generate this output for the ML method is given in Appendix A.

Table 8.2. SAS Output for Battery Life Example I

REML

Estimation Iteration History

M	T	M*T	RESIDUAL	ITERATION
0	0	0	1	0
244.8681	1429.6597	432.0579	675.2130	1

Asymptotic Covariance Matrix of Estimates

Cov Parm	Row	T	M	M*T	Residual
T	1	1555452.2	-71357.3	-34965.5	0
M	2	-71357.3	210768.1	-58117.4	0
T*M	3	-34965.5	-58117.4	182061.1	-8442.8
Residual	4	0	0	-8442.8	33771.3

Solution for Fixed Effects

Effect	Estimate	Std Error
INTERCEPT	105.5278	24.9988

ML

Estimation Iteration History

M	T	M*T	RESIDUAL	ITERATION
0	0	0	1	0
191.2087	439.3520	591.4815	675.2130	1
55.2815	1007.2364	511.5329	675.2130	2
205.0174	843.4741	461.2602	675.2130	3
163.4037	955.1122	459.7966	675.2130	4
185.0232	917.4251	456.5319	675.2130	5
177.0576	934.4312	457.0915	675.2130	6
180.4469	927.8164	456.7242	675.2130	7
179.0962	930.5594	456.8477	675.2130	8
179.6498	929.4543	456.7929	675.2130	9
179.4256	929.9052	456.8143	675.2130	10
179.5169	929.7222	456.8055	675.2130	11
179.4798	929.7966	456.8090	675.2130	12

Asymptotic Covariance Matrix of Estimates

Cov Parm	Row	T	M	M*T	Residual
T	1	966242.4	-24197.8	-47838.6	0
M	2	-24197.8	166293.3	-62957.2	0
T*M	3	-47838.6	-62957.2	196944.9	-8442.8
Residual	4	0	0	-8442.8	33771.3

Solution for Fixed Effects

Effect	Estimate	Std Error
INTERCEPT	105.5278	20.9587

Since the design is balanced, Zyskind's condition holds for the Y-Model. By the Y-FELMLQE theorem in section 4.2.2, there exists an FELMLQE for μ which is the overall mean $\bar{Y}_{...}$. This can be seen in the output since the two methods of estimation yield the same fixed effect estimate for the intercept. The estimate does not depend on the variance component estimates, which are not the same for the ML and REML methods. An exact expression for the standard error of the FELMLQE $\bar{Y}_{...}$ is $\frac{1}{36}\sqrt{\underline{1}'V_{\underline{\psi}}\underline{1}}$. The standard error estimate in the output is obtained by plugging in $\widehat{\underline{\psi}}_{\text{MLQ}}$ or $\widehat{\underline{\psi}}_{\text{REMLQ}}$ into the exact expression.

By the corollary in section 4.3.3, there is an FELREMLQE for $\underline{\psi}$ in balanced designs. This can be seen in the REML estimation iteration history as only a single iteration is needed to obtain the solution. As shown section 7.2.2, there exists an exact expression for the covariance of the REML estimate. The values in the asymptotic covariance matrix are obtained by plugging in $\widehat{\underline{\psi}}_{\text{REMLQ}}$ into the exact expression. Thus, these values are estimates from an exact expression, rather than estimates of the asymptotic expression.

The ML estimation iteration history indicates there is not an explicit linear solution for the variance components under the ML procedure since it takes 12 iterations to converge. There does not exist an FELMLQE for $\underline{\psi}$ since $\text{sp}\mathcal{V} = \{P_M, P_T, P_{M\times T}, I\}$ is not a QS as $P_M P_T + P_T P_M = P_{\underline{1}} \notin \text{sp}\mathcal{V}$. However, by example 6.6, there exists an ELMLQE for σ_e^2 as this model includes the highest possible order term. This is evident from the ML estimation iteration history where the estimate of the residual component does not change over the iterations. It is interesting to note that the ELMLQE and ELREMLQE for σ_e^2 are the same in this example. Since the estimate of σ_e^2 is the same for both methods, the exact estimate of $\text{Cov}_{\underline{\psi}}(\sigma_e^2)$ will also be the same.

8.1.3. Battery Life Example II

For illustrative purposes, consider a modification of Battery Life Example I. For this example, the first two observations are removed for material 3 at each temperature level. The deleted observations are underlined in Table 7.1. This results in an unbalanced design which has the incidence matrix $\begin{bmatrix} 4 & 4 & 4 \\ 4 & 4 & 4 \\ 2 & 2 & 2 \end{bmatrix}$ where each row denotes a material and each column denotes a temperature. Also, suppose temperature is a fixed factor in a model which does not include the interaction term M*T. Let Y_{ijk} be the response for temperature i, material j, and observation k. Also, let α_i be the treatment effect associated with temperature i. Then the expectation for this model is given by $\text{E}[Y_{ijk}] = \mu + \alpha_i$ for all j, k.

The variance components and the fixed effects will be estimated using the MIXED procedure in SAS under the ML, REML, and MIVQUE0 methods. The MIVQUE0 method is presented since it is not the same as the REML method in this case. The output from this procedure is given in Table 8.3. The SAS code used to generate this output for the REML method is given in Appendix A.

Table 8.3. SAS Output for Battery Life Example II

REML

Covariance Parameter Estimates

M	194.4	Residual	929.8

Solution for Fixed Effects

Parameter	Effect	T	Estimate	Std Error
$\mu + \alpha_3$	INT		60.78227290	12.64288343
$\alpha_1 - \alpha_3$	T	1	84.00000000	13.63633746
$\alpha_2 - \alpha_3$	T	2	37.50000000	13.63633746
0	T	3	0.00000000	.

ML

Covariance Parameter Estimates

M	104.5	Residual	860.6

Solution for Fixed Effects

Parameter	Effect	T	Estimate	Std Error
$\mu + \alpha_3$	INT		60.42803353	11.06201142
$\alpha_1 - \alpha_3$	T	1	84.00000000	13.11971264
$\alpha_2 - \alpha_3$	T	2	37.50000000	13.11971264
0	T	3	0.00000000	.

MIVQUE0

Covariance Parameter Estimates

M	211.3	Residual	921.8

Solution for Fixed Effects

Parameter	Effect	T	Estimate	Std Error
$\mu + \alpha_3$	INT		60.83984741	12.83522304
$\alpha_1 - \alpha_3$	T	1	84.00000000	13.57776055
$\alpha_2 - \alpha_3$	T	2	37.50000000	13.57776055
0	T	3	0.00000000	.

The variance component estimates for the ML and REML methods required iterations. The values of these estimates over the iterations are not of interest in this example. However, note how the final estimates differ across the estimation methods.

The solutions for the fixed effects are from SAS's default parameterization, which provides estimates of $\mu + \alpha_3$, $\alpha_1 - \alpha_3$, and $\alpha_2 - \alpha_3$. Despite the different variance component estimates, the output shows that the estimates for the treatment effect differences $\alpha_1 - \alpha_3$ and $\alpha_2 - \alpha_3$ remain the same across the estimation methods. As shown in section 5.4, there indeed exists an ELMLQE for the treatment effect differences when the rank of the incidence matrix is 1, or equivalently when the incidence matrix has proportional frequencies.

This model did not include the interaction term M*T. If the model did include this term, then an ELMLQE would not exist for the treatment effect differences when the rank of the incidence matrix is 1.

From section 7.2.2, the standard error of the treatment differences has an exact expression. The associated standard error estimates for an estimation method are obtained by plugging in either $\widehat{\psi}_{\text{MLQ}}$, $\widehat{\psi}_{\text{REMLQ}}$, or $\widehat{\psi}_{\text{MIVQUE0}}$ into the exact expression.

8.1.4. Battery Life Example III

For illustrative purposes, consider a modification of Battery Life Example I. For this example, the last two observations are removed from the combination material 3, temperature 3. The two deleted observations are marked by 'x' in Table 7.1. This results in an unbalanced design which has the incidence matrix $\begin{bmatrix} 4 & 4 & 4 \\ 4 & 4 & 4 \\ 4 & 4 & 2 \end{bmatrix}$ where each row denotes a material and each column denotes a temperature. Also, suppose temperature is a fixed factor. The interaction term M*T will be included in the model as a random effect. The expectation for this model is also given by $\mathrm{E}[Y_{ijk}] = \mu + \alpha_i$ for all j, k as described in section 8.1.3.

The variance components and the fixed effects will be estimated using the MIXED procedure in SAS under the ML, REML, and MIVQUE0 methods. The MIVQUE0 method is presented since it is not the same as the REML method in this case. The output from this procedure is given in Table 8.4. The SAS code used to generate the output for the MIVQUE0 method is in Appendix A.

Table 8.4. SAS Output for Battery Life Example III

REML

Covariance Parameter Estimates

M	247.8	M*T	446.3	Residual	686.4

Solution for Fixed Effects

Parameter	Effect	T	Estimate	Std Error
$\mu + \alpha_1$	T	1	144.83333333	16.98671077
$\mu + \alpha_2$	T	2	107.58333333	16.98671077
$\mu + \alpha_3$	T	3	64.32324833	17.46732374

ML

Covariance Parameter Estimates

M	159.9	M*T	243.2	Residual	684.6

Solution for Fixed Effects

Parameter	Effect	T	Estimate	Std Error
$\mu + \alpha_1$	T	1	144.83333333	13.83577091
$\mu + \alpha_2$	T	2	107.58333333	13.83577091
$\mu + \alpha_3$	T	3	63.90099970	14.38532825

MIVQUE0

Covariance Parameter Estimates

M	205.7	M*T	491.6	Residual	683.1

Solution for Fixed Effects

Parameter	Effect	T	Estimate	Std Error
$\mu + \alpha_1$	T	1	144.83333333	17.01108447
$\mu + \alpha_2$	T	2	107.58333333	17.01108447
$\mu + \alpha_3$	T	3	64.29243163	17.49171658

The variance component estimates for the ML and REML methods required iterations. The values of these estimates over the iterations are not of interest in this example. However, note how the final estimates differ across the estimation methods.

The solutions for the fixed effects are from SAS's parameterization under the NOINT option. The output from this parameterization gives the estimate for $\mu + \alpha_1$, $\mu + \alpha_2$, and $\mu + \alpha_3$. Despite the difference in the variance component estimates, the output shows that the estimates for the first two treatment means are the same across the estimation methods. A check of the UBLUE conditions reveals that there indeed exists an ELMLQE for $\mu + \alpha_1$ and $\mu + \alpha_2$.

From section 7.2.2, the standard error of the first two treatment mean has an exact expression. The associated standard error estimates for the estimation methods are obtained by plugging in either $\widehat{\underline{\psi}}_{\text{MLQ}}$, $\widehat{\underline{\psi}}_{\text{REMLQ}}$, or $\widehat{\underline{\psi}}_{\text{MIVQUE0}}$ into the exact expression.

The examples in these last two sections illustrate that the ability to identify an ELMLQE for a linear combination of the fixed effects may depend upon the parameterization. For example, it was easier to identify the ELMLQE in the previous section under the SAS default parameterization while it was easier to identify the ELMLQE in this section using the NOINT option.

In the previous examples, the existence of an ELMLQE and ELREMLQE were already proven in previous chapters. However, in this section, the existence of an ELMLQE had to be verified separately using a matrix computing language to perform the check described in section 6.5. This leads to the question of whether it is possible to verify the existence of an ELMLQE or ELREMLQE using the iterative procedure. This question is investigated in the next section.

8.1.5. Checking the Conditions Using the Iterative Procedure

This section describes a method to check the conditions using the iterative procedure presented in section 8.1.1. Consider checking the GZC presented in chapter 5 which is sufficient for the existence of an ELMLQE or ELREMLQE for a linear combination of the fixed effects or a linear combination of the variance components. The previous sections in this chapter show the outcomes of the iterative procedure when the GZC does and does not hold. However, misleading conclusions could be drawn by observing these outcomes without knowledge of whether or not GZC holds. This is due to the dependence of the iterative procedure on particular data values $\underline{y}$ and particular parameter values $\underline{\psi}$. However, it is possible to verify the conditions, with a degree of certainty, using the iterative procedure. This section only presents the method and its interpretation. A formal justification of the method is given in Appendix B.

In order to implement the method, it is necessary to have two items. The first item is a randomly generated observation $\underline{y}^r$ from a continuous distribution with support that contains a non-empty open set. In data analysis problems, it is likely that the observations can be assumed to be randomly observed from a normal distribution. The second item is a randomly generated value of $\underset{\sim}{\psi}^r$ from a continuous distribution with support $\Xi^* = \{\underset{\sim}{\psi} \in \mathcal{R}^{k+1} | V_{\underset{\sim}{\psi}} \text{ is PD}\}$. In addition, the value $\underset{\sim}{\psi}^r$ must be generated independent of $\underset{\sim}{y}^r$.

Consider checking the GZC for the ML procedure for a linear combination of the fixed effects $\Lambda'\underset{\sim}{\beta}$. The GZC, in this case, is the sufficient condition for the existence of an ELMLQE for $\Lambda'\underset{\sim}{\beta}$ (section 5.2.2). This method does not need to involve the iterative procedure to solve for $\underset{\sim}{\psi}$, but it is necessary to have a given value of $\underset{\sim}{\psi}$. For a fixed value of $\underset{\sim}{\psi}$ given by $\underset{\sim}{\psi}^f$, $\Lambda'\widehat{\underset{\sim}{\beta}}_{\underset{\sim}{\psi}^f}(\underset{\sim}{y}^r)$ is the GLSE based on $\underset{\sim}{\psi}^f$ and $\underset{\sim}{y}^r$ (section 5.2.1). Method A is as follows:

1. For any given value, $\underset{\sim}{\psi}^g \in \Xi^*$, calculate $\Lambda'\widehat{\underset{\sim}{\beta}}_{\underset{\sim}{\psi}^g}(\underset{\sim}{y}^r)$ (ex. $\underset{\sim}{\psi}^g = [0\ 0 \ldots 0\ 1]'$).
2. For the random value $\underset{\sim}{\psi}^r$, calculate $\Lambda'\widehat{\underset{\sim}{\beta}}_{\underset{\sim}{\psi}^r}(\underset{\sim}{y}^r)$.
3. Does $\Lambda'\widehat{\underset{\sim}{\beta}}_{\underset{\sim}{\psi}^g}(\underset{\sim}{y}^r) = \Lambda'\widehat{\underset{\sim}{\beta}}_{\underset{\sim}{\psi}^r}(\underset{\sim}{y}^r)$?

Consider checking the GZC for the ML or REML procedure for a linear combination of the variance components $\Gamma'\underset{\sim}{\psi}$. The GZC, in this case, is the sufficient condition for the existence of an ELMLQE or ELREMLQE for $\Gamma'\underset{\sim}{\psi}$ (section 5.3.2, 5.3.3). Let $\Gamma'\widehat{\underset{\sim}{\psi}}(\underset{\sim}{y}^r)$ denote the EGLSE based on $\underset{\sim}{y}^r$ (section 5.3.2, 5.3.3). In order to calculate the EGLSE, it is necessary to use the iterative procedure to solve for $\underset{\sim}{\psi}$ where the starting value is specified. Method B is as follows:

1. Use the random value $\underset{\sim}{\psi}^r$ as a starting value in the iterative procedure.
2. Calculate $\Gamma'\widehat{\underset{\sim}{\psi}}(\underset{\sim}{y}^r)$.
3. Does the iterative procedure converge in a single iteration for $\Gamma'\widehat{\underset{\sim}{\psi}}(\underset{\sim}{y}^r)$?

Methods A and B can be implemented in a computing language that can perform the iterative procedure described in section 8.1.1, generate $\underset{\sim}{y}^r$ and $\underset{\sim}{\psi}^r$, fix the variance component values at $\underset{\sim}{\psi}^g$ or $\underset{\sim}{\psi}^r$, and specify $\underset{\sim}{\psi}^r$ as a starting value in the iterative procedure. For instance, it may be the case that these methods can be used in SAS in the PROC MIXED procedure with the PARMS statement (SAS,1996).

Consider the random variables $\underset{\sim}{Y}$ and $\underset{\sim}{\psi}$. For either Method A or B, let $S(\underset{\sim}{\psi}, \underset{\sim}{Y}) = 1$ if the answer to step 3 is yes and let $S(\underset{\sim}{\psi}, \underset{\sim}{Y}) = 0$ if the answer is no. Also, let $\xi = \begin{cases} 1 & \text{if GZC holds} \\ 0 & \text{if GZC does not hold} \end{cases}$ and P_ξ be the joint probability distribution of the independent random variables $\underset{\sim}{Y}$ and $\underset{\sim}{\psi}$. From the results in Appendix B, $\xi = 1$ implies $P_\xi(S(\underset{\sim}{\psi}, \underset{\sim}{Y}) = 1) = 1$ and $\xi = 0$ implies $P_\xi(S(\underset{\sim}{\psi}, \underset{\sim}{Y}) = 0) = 1$.

However, the goal is to use $S(\underset{\sim}{\psi}, \underset{\sim}{Y})$ as a statistic to draw inference about the unknown parameter ξ. Informally, one could ignore the probability measure P_ξ and say $\xi = 1$ if and only if $S(\underset{\sim}{\psi}, \underset{\sim}{Y}) = 1$ and $\xi = 0$ if and only if $S(\underset{\sim}{\psi}, \underset{\sim}{Y}) = 0$. For either Method A or B, this means that the answer to step 3 is 'YES' if and only if the GZC holds and the answer to step 3 is 'NO' if and only if the GZC does not hold.

Formally, it is necessary to account for the probability measure when drawing inference to ξ. This can be done using a confidence region. A 100% confidence region for ξ is given by $C(S(\underline{\psi}, \underline{Y})) = \begin{cases} \{1\} & \text{if } S(\underline{\psi}, \underline{Y}) = 1 \\ \{0\} & \text{if } S(\underline{\psi}, \underline{Y}) = 0 \end{cases}$. This confidence region indicates that if $S(\underline{\psi}, \underline{Y}) = 1$, then $C(S(\underline{\psi}, \underline{Y})) = \{1\}$ contains the true value of ξ with 100% confidence. On the other hand, if $S(\underline{\psi}, \underline{Y}) = 0$, then $C(S(\underline{\psi}, \underline{Y})) = \{0\}$ contains the true value of ξ with 100% confidence. Thus, the coverage probability is $P_\xi(\xi \in C(\underline{\psi}, \underline{Y})) = 1$ as shown in Appendix B.

This section has demonstrated how the iterative procedure in section 8.1.1 can be used to check the GZC for fixed effects and variance components. Justification of the results in this section is given in Appendix B. One problem with these methods is that the numbers randomly generated from a computer are not truly random, rather they are 'pseudo-random'.

8.2. Profile Likelihood Calculations and Computing Time

Suppose there exists an ELMLQE for a subvector of an estimable parameter vector which could consist of fixed effects or variance components. This section demonstrates that computing time and resources could be saved by accounting for the ELMLQE in the iterative procedure given in section 8.1.1. In particular, profile likelihood calculations are examined. This section discusses adjusting the iterative procedure for the ELMLQE and computing profile likelihood confidence intervals. These results will be applied to Battery Life Example I.

8.2.1. Adjusting the Iterative Procedure

The iterative procedure presented in section 8.1 can be altered when there are explicit linear likelihood estimators. This alteration may be helpful for saving computing time and resources. For large data sets with numerous variables, the savings could be dramatic. The adjustments for the iterative procedure will be presented with respect to the maximum likelihood procedure as fixed effects are of interest.

Consider the variance component vector $\underline{\psi} = [\underline{\psi}'_{1\,k_1\times 1}\ \underline{\psi}'_{2\,k_2\times 1}]'$. If there is an ELMLQE for the subvector $\underline{\psi}_2$, then the iterative procedure can be adjusted to account for the simple explicit linear expression given by $\widehat{\underline{\psi}}_2$. From the formulas in given in section 6.4.2, the MLQE for $\underline{\psi}$ would consist of the subvectors given by:

ELMLQE: $\widehat{\underline{\psi}}_{2I} = (X_2^{\diamond *} N_{X_1^\diamond} X_2^\diamond)^{-1} X_2^{\diamond *} N_{X_1^\diamond} Y_2^\diamond$

MLQE: $\widehat{\underline{\psi}}_1 = (X_1^{\diamond *} V_{\widehat{\underline{\psi}}}^{\diamond -1} X_1^\diamond)^{-1} X_1^{\diamond *} V_{\widehat{\underline{\psi}}}^{\diamond -1} (Y_2^\diamond - X_2^\diamond \widehat{\underline{\psi}}_{2I})$.

The ELMLQE can be identified on a single iteration, while the MLQE will require an iterative procedure. However, the MLQE for $\underline{\psi}$ can now be calculated using the inverse of a $k_1 \times k_1$ matrix instead of a $(k_1 + k_2) \times (k_1 + k_2)$ matrix. It should be noted these expressions assume ZC holds, so there must exist a FELMLQE for $X\underline{\beta}$ as indicated by the Y-FELMLQE theorem.

Consider the case where there does not exist an FELMLQE for $\underline{\beta}$ or for $\underline{\psi}$. The iterative procedure discussed in section 8.1 shows how to find the MLQE for $\underline{\beta}$ after an MLQE for $\underline{\psi}$ has been obtained. However, it may be informative to calculate $\underline{\beta}$ for each iterative solution of $\underline{\psi}$ in this case. There is potential to save computing time and resources in this case as well. Suppose $\underline{\beta} = [\underline{\beta}'_{1\,p_1\times 1}\,\underline{\beta}'_{2\,p_2\times 1}]'$ where there exists an ELMLQE for $\underline{\beta}_2$. Using the formulas in section 6.2, the MLQE for $\underline{\beta}$ would be given by:

ELMLQE: $\widehat{\underline{\beta}}_{2I} = (X_2' N_{X_1} X_2)^{-1} X_2' N_{X_1} \underline{Y}$

MLQE: $\widehat{\underline{\beta}}_1 = (X_1' V_{\widehat{\underline{\psi}}}^{-1} X_1)^{-1} X_1' V_{\widehat{\underline{\psi}}}^{-1} (\underline{Y} - X_2 \widehat{\underline{\beta}}_{2I})$.

The ELMLQE can be identified on a single iteration while the MLQE will require an iterative procedure. However, the MLQE for $\underline{\beta}$ can now be calculated using the inverse of a inverse $p_1 \times p_1$ matrix instead $(p_1 + p_2) \times (p_1 + p_2)$ matrix.

8.2.2. Computing Profile Likelihood Confidence Intervals

The profile likelihood procedure is a technique which provides inference about a parameter in the presence of nuisance parameters. Particular interest in this section is obtaining a profile likelihood confidence interval. This procedure can be defined as in McCullagh and Nelder (1983). Suppose there is interest in $\underline{\theta}_1$ where $\underline{\theta} = [\underline{\theta}'_{1\,t_1\times 1}\;\; \underline{\theta}'_{2\,t_2\times 1}]'$ and consider the following definitions:

Definitions: Profile Log Likelihood Function (for $\underline{\theta}_1$): $l_p(\underline{\theta}_1) = \sup_{\underline{\theta}_2} \ln L(\underline{\theta}_1, \underline{\theta}_2)$

Likelihood Ratio Test (LRT): For testing $H_0 : \underline{\theta}_1 = \underline{\theta}_{10}$ vs $H_A : \underline{\theta}_1 \neq \underline{\theta}_{10}$, the LRT is given by

$$T_p(\underline{\theta}_{10}) = 2[\, l_p(\widehat{\underline{\theta}}_1) - l_p(\underline{\theta}_{10})\,] \stackrel{.}{\sim} \chi^2_{p_2} \text{ under } H_0$$

Approximate $100(1-\alpha)\%$ Confidence Region (for $\underline{\theta}_1$) : The region given by

$$\{\,\underline{\theta}_1 : T_p(\underline{\theta}_1) \le \chi^2_{t_2,1-\alpha}\} = \{\underline{\theta}_1 : 2[\, l_p(\widehat{\underline{\theta}}_1) - l_p(\underline{\theta}_1)\,] \le \chi^2_{t_2,1-\alpha}\}\,.$$

Using the profile likelihood function, a likelihood ratio test statistic can be obtained and inverted to produce an approximate confidence region for $\underline{\theta}_1$. In order to calculate this approximate confidence region, it is necessary to find $l_p(\underline{\theta}_1) = l(\underline{\theta}_1, \widehat{\underline{\theta}}_{2\underline{\theta}_1})$ or $\widehat{\underline{\theta}}_{2\underline{\theta}_1}$ for each value of $\underline{\theta}_1$. Thus, there is a computational advantage if there exists an explicit likelihood estimator for $\widehat{\underline{\theta}}_{2\underline{\theta}_1}$ or some of its components.

Saving computer time is important in this case since an iterative procedure must be repeated for each value of $\underline{\theta}_1$. The iterative procedure would be adjusted as indicated in section 8.2.1. The difficulty is how to incorporate the given value of $\underline{\theta}_1$ when calculating $l_p(\underline{\theta}_1)$.

The following changes can be made to account for a given value of $\underline{\theta}_1$ when $\underline{\theta}_1 = \underline{\beta}_1$. For a particular value $\underline{\beta}_{10}$, the response is $\underline{Z}_0 = \underline{Y} - X_1\underline{\beta}_{10}$ where $\mathrm{E}[\underline{Z}_0] = X_2\underline{\beta}_2$ and $\mathrm{Cov}(\underline{Z}_0) = \mathrm{Cov}(\underline{Y}) = V_{\underline{\psi}}$. Zyskind's condition for the $\underline{Z}_0$-Model would be $\underline{R}(V_{\underline{\psi}}X_2) \subset \underline{R}(X_2)\ \forall \underline{\psi} \in \Xi$. The corresponding LQEM has response $Y_2^{\diamond} = (\underline{Z}_0 - X_2\underline{\beta}_2)(\underline{Z}_0 - X_2\underline{\beta}_2)'$. The iterative procedure to estimate $\widehat{\underline{\psi}}$ and $\widehat{\underline{\beta}}_2$ would be applied to this LQEM as in section 8.1.1. The conditions for the existence of an ELMLQE for the subvector of $\underline{\psi}$ will not be changed since these conditions do not depend on $Y_2^{\diamond}$.

The following changes can be made to account for a given value of $\underline{\theta}_1$ when $\underline{\theta}_1 = \underline{\psi}_1$. It is more difficult in this case to account for an ELMLQE. Let $\underline{\psi}^* = [\underline{\psi}'_{10\,k_1\times 1}\ \ \underline{\psi}'_{2k_2\times 1}\ \ \underline{\psi}'_{3k_3\times 1}]'$ where $\underline{\psi}_{10}$ is a particular value and it is of interest whether there exists an ELMLQE for $\underline{\psi}_3$. Hence, the covariance matrix is $\mathrm{Cov}(\underline{Y}) = V_{\underline{\psi}^*} = V_{\underline{\psi}_{10}} + V_{\underline{\psi}_2} + V_{\underline{\psi}_3}$ and $\mathrm{E}[Y_1^{\diamond} - V_{\underline{\psi}_{10}}] = V_{\underline{\psi}^*} - K_0 = V_{\underline{\psi}_2} + V_{\underline{\psi}_3} = X_2^{\diamond}\underline{\psi}_2 + X_3^{\diamond}\underline{\psi}_3$. From the U-UBLUE$_{\mathrm{FR}}$ theorem, an ELMLQE would exist for $\underline{\psi}_3$ when $\underline{R}(V_{\underline{\psi}^*}^{\diamond}N_{X_2^{\diamond}}X_3^{\diamond}) \subset \underline{R}([X_2^{\diamond} + X_3^{\diamond}])$ for all $\underline{\psi}^*$ where $\mathrm{Cov}(Y_1^{\diamond}) = V_{\underline{\psi}^*}^{\diamond} = 2\Psi_{V_{\underline{\psi}^*}}$. Thus, the sufficient condition for the existence of an ELMLQE is affected by $\underline{\psi}_{10}$ through $\underline{\psi}^*$. The following proposition gives an example in which an ELMLQE exists in this setting.

Proposition: Consider the balanced random 1-way model where $\underline{\psi} = [\sigma_a^2\ \ \sigma_e^2]$. If $\widehat{\underline{\psi}}_{\mathrm{MLQ}}$ exists and $V_{\widehat{\underline{\psi}}_{\mathrm{MLQ}}}$ is PD, then $\exists$ an ELMLQE for $\sigma_a^2\ \ \forall\, \underline{\psi}^* = [\sigma_a^2\ \ \sigma_{e0}^2\,]$ where σ_{e0}^2 is a fixed value of σ_e^2.

proof: $V_{\underline{\psi}^*} = m\sigma_a^2 P_A + \sigma_{e0}^2 I = V_{\underline{\psi}_3} + V_{\underline{\psi}_{10}} \Rightarrow \mathrm{E}[Y_1^{\diamond} - \sigma_{e0}^2 I\,] = m\sigma_a^2 P_A$ and

$\mathrm{Cov}(Y_1^{\diamond} - \sigma_{e0}^2 I\,) = 2\Psi_{V_{\underline{\psi}^*}} = V_{\underline{\psi}^*}^{\diamond}$. Note $\underline{R}(V_{\underline{\psi}^*}^{\diamond}N_{X_2^{\diamond}}X_3^{\diamond}) = \underline{R}(V_{\underline{\psi}^*}^{\diamond}X_3^{\diamond})$ since $X_2^{\diamond} = 0$.

Thus, $\forall\, \underline{\psi}^* = [\sigma_a^2\ \ \sigma_{e0}^2\,]$ and $u \in \mathcal{R}$, $V_{\underline{\psi}^*}^{\diamond}X_3^{\diamond}u = 2V_{\underline{\psi}^*}V_{\underline{\psi}_3}uV_{\underline{\psi}^*}$

$$= 2u(\sigma_{e0}^2 I + m\sigma_a^2 P_A)m\sigma_a^2 P_A(\sigma_{e0}^2 I + m\sigma_a^2 P_A) = 2um\sigma_a^2(\sigma_{e0}^2 P_A + m\sigma_a^2 P_A)(\sigma_{e0}^2 I + m\sigma_a^2 P_A)$$

$$= 2um\sigma_a^2(\sigma_{e0}^2 + m\sigma_a^2)^2 P_A \in \underline{R}(X_3^{\diamond})\ \ \forall\, \sigma_{e0}^2, \sigma_a^2$$

$\Rightarrow \exists$ an ELMLQE for σ_a^2 by the LQNY-ELMLQE$_{\mathrm{FR}}$ theorem. ∎

8.2.3. Battery Life Example I

Battery Life Example I was presented in section 8.1.2. In this section, it will be used to indicate the saving of computing time for calculating a profile likelihood confidence interval for the fixed effect $\theta_1 = \mu$ where the complete parameter vector is given by $\underline{\theta} = [\mu \;\; \sigma_M^2 \;\; \sigma_T^2 \; \sigma_{M\times T}^2 \; \sigma_e^2]' = [\theta_1 \;\; \underline{\theta}_{2(4\times1)}']'$. For this example, it was observed that there exists an ELMLQE for the subcomponent $\underline{\theta}_2$ consisting of σ_e^2. Computing time could be saved by removing σ_e^2 from the iterative procedure as described in section 8.2.1. Also, note that ZC holds in this example, so there is an ELMLQE for μ as well.

The CPU time was measured for finding the profile likelihood when acounting for the ELMLQE for σ_e^2 and when ignoring the ELMLQE for σ_e^2. The MATLAB program was used on a Pentium II 200 MHz computer and generated an approximate 95% profile likelihood confidence interval for μ given by $(49.25, 161.75)$. The CPU time for a particular computer varies on a run and depends on the parameter range, step size, and convergence criteria. For purposes of illustration, the values were set to [0, 200], .5, and .1 respectively. The times are given in Table 8.5 for a single run. This table indicates that over 15 minutes were saved by accounting for the ELMLQE for σ_e^2. The time savings were substantial in this example. The savings would be much larger for examples with more observations and more variance components.

Table 8.5. CPU Time for Profile Likelihood for Mean in Battery Life Example I

CALCULATION	CPUTIME
Account for ELMLQE for σ_e^2	35 min 55 sec
Do not account for ELMLQE for σ_e^2	51 min 5 sec (>15 min)

9. Conclusion

9.1. Summary

Szatrowski (1980) and Elbassiouni (1983) establish conditions for the existence of a full ELMLQE and a full ELREMLQE for the fixed effect vector and the variance component vector. These results are presented in chapter 4. This thesis presents the previous results using models carefully defined in chapter 3. The sufficient conditions were related to the UBLUE conditions in chapter 4.

This thesis extends the results of Szatrowski (1980) and Elbassiouni (1983) to identify conditions for the existence of an ELMLQE and an ELREMLQE for linear combinations involving the fixed effects and variance components. The general case was formulated in chapter 5 and the most general version is given in section 5.1. A special case involving conditions for the existence an ELMLQE and an ELREMLQE for a subvector of the fixed effect vector or a subvector of the variance component vector is presented in chapter 6. The general procedure for obtaining these conditions involves deriving UBLUE results for the underlying model defined in section 3.3.1. These UBLUE results can be applied to the specific models to obtain conditions for existence of the ELMLQE and the ELREMLQE. Under the UBLUE conditions, the ELMLQE and ELREMLQE are given by the least squares estimators with respect to the models of interest.

This thesis also presents examples in which the ELMLQE and ELREMLQE conditions hold in the general case. The most comprehensive example is given in section 6.8 and defines a class of random models under specific design and model conditions that have an ELREMLQE for a subvector of the variance component vector. Other examples that have ELMLQEs or ELREMLQEs for a subvector of the variance component vector are discussed at the end of chapters 5 and 6. Tables are given in section 6.9, which illustrate 3-way models that have an ELMLQE or ELRELMQE for the variance components under various designs.

Chapter 7 applies the UBLUE conditions to UMVUE's in the full and general cases. In the full case in section 7.1, there exists a complete sufficient statistic statistic for the family of normal distributions under the ML and REML procedures (Seely,1971). The ELMLQE and ELREMLQE can be shown to be functions of the complete sufficient statistic. In section 7.2, it is shown that the ELMLQE for a linear combination of fixed effects and the ELREMLQE for a linear combination of variance components are UMVUE. This is done by showing that the covariance attains the lower bound for unbiased estimators. This section also gives exact expressions for the covariance of the ELMLQE and ELREMLQE.

This thesis also applies these results in chapter 8 to an iterative procedure for obtaining the MLQE and REMLQE. Section 8.1 discusses the procedure and how to implement it in SAS using PROC MIXED. Sections 8.1.2, 8.1.3, and 8.1.4 apply the procedure to data examples. Section 8.1.5 shows how to use the iterative procedure to check the conditions in the iterative algorithm. Methods to save computing time are given in section 8.2. The savings is shown to be dramatic for profile likelihood calculations. The methods are demonstrated using a data example in section 8.2.3.

The UBLUE conditions have been used to generalize the results of Szatrowski (1980) and Elbassiouni (1983) to linear combinations of the parameters. The underlying linear model establishes a framework in which to extend the results, so that they can be applied to the particular models of interest for the purposes of ML and REML estimation. Under the UBLUE conditions, the ELMLQE and ELREMLQE are given by the least squares estimator in the appropriate model. Such estimators are easy to compute, simple to interpret, and have optimal properties. The general idea behind these results can be applied to any situation where least squares and generalized least squares estimation is applicable.

9.2. Further Research

This study has identified interesting questions for future research. The UBLUE conditions, which are mentoned below, refer to the GZC or those conditions presented in chapter 5 for estimating the fixed effects and variance components. The questions for further research are listed below:

(1) Apply the UBLUE conditions to hypothesis testing in mixed models. For instance, these conditions may useful for identifying the existence of exact F-tests.

(2) Generalize the results to the case where the covariance is not PD. Christensen (1996,section 12.5) discusses maximum likelihood estimation for singular normal distributions. Sections 3.3.4 and 3.3.5 provide results where the covariance is NND.

(3) The UBLUE conditions indicate when the maximum likelihood and restricted maximum likelihood estimators are unbiased. Additional work could be done to determine whether these estimators are equal to analysis of variance estimators (ANOVA) (Searle et al.,1992).

(4) Determine whether other iterative procedures converge in a single iteration under the UBLUE conditions (Searle, et al.,1992). In particular, examine convergence subject to constraints on the variance components (Harville,1977).

(5) Examine whether the results pertaining to the class of random models presented in section 6.8 can be extended to a class of mixed models.

(6) Identify conditions under which the MLQE and REMLQE exist. Such conditions could be used to show when ML and REML procedures are applicable.

(7) Determine whether it is possible to use a less restrictive condition than ZC for the Y-Model when applying the ALQEM for $(\underline{Y} - X\hat{\underline{\beta}})$ to identify the existence of an ELMLQE in the partial case. This issue is discussed in sections 5.3.2 and 6.4.2. For these cases, it may be possible to use a weaker condition since the full variance component vector is not of interest.

(8) Examine whether the existence of an ELMLQE or ELREMLQE for a subcomponent of the variance component vector is equivalent to part of the inverse of the covariance matrix being explicit. Rogers and Young (1977) and Seely (1971) examine this relationship in the full case.

(9) Examine whether the UBLUE conditions can be applied to generalized linear models. In particular, consider generalized estimating equations.

(10) Extend the UBLUE conditions to general covariance structures. For example, this could include repeated measures designs. It would require a reformulation of the conditions to covariances that do not have the linear structure.

(11) Derive design based conditions in which $\underline{\psi}$ is estimable in the ALQEM for $(\underline{Y} - X\hat{\underline{\beta}})$ and the LQEM for $N_X\underline{Y}$ for particular classes of models.

(12) Find out whether the existence of a partial ELMLQE implies the existence of a partial ELREMLQE.

Bibliography

Anderson, T.W. (1969). "Statistical Inference for Covariance Matrices with Linear Structure". *Multivariate Analysis-II* , ed. P.R. Krishnaiah, New York: Academic Press, 55-66.

Birkes, D. (1996). Linear Model Notes, unpublished. Oregon State University.

Callanan, T.R. and Harville, D.A. (1991). "Some New Algorithms for Computing Restricted Maximum Likelihood Estimates of Variance Components". *Journal of Statistical Computing and Simulation*, v 38, 239-259.

Casella, G. and Berger, R.L. (1990). *Statistical Inference.* Pacific Grove, CA: Wadsworth&Brooks/Cole.

Christensen, R. (1996). *Plane Answers to Complex Questions*, 2nd Ed. New York: Springer.

ElBassiouni, M. Y. (1983). "On the Existence of Explicit Restricted Maximum Likelihood Estimators in Multivariate Normal Models". *The Indian Journal of Statistics*, v 45, ser B, pt 2, 303-306.

Halmos, P.R. (1958). *Finite-Dimensional Vector Spaces.* New York: Springer-Verlag.

Harville, D.A. (1977). "Maximum Likelihood Approaches to Variance Component Estimation and to Related Problems". *Journal of the American Statistical Association*, v 72, no 358, 320-338.

Harville, D.A. (1997). *Matrix Algebra from a Statistician's Perspective.* New York: Springer-Verlag.

Henderson, C.R. (1953). "Estimation of Variance and Covariance Components". *Biometrics*, v 9, 226-252.

Krantz, S.G. and Parks, H.R. (1992). *A Primer of Real Analytic Functions.* Boston: Birkhauser Verlag.

Lehmann, E.L. (1986). *Testing Statistical Hypotheses.* New York: Chapman and Hall.

Lehmann, E.L. (1983). *Theory of Point Estimation.* Belmont, CA: Wadsworth, Inc.

Marcus, Marvin and Minc, Henryk. (1965). *Introduction to Linear Algebra.* New York: Dover Publications, Inc.

McCullagh, P. and Nelder, J.A. (1983). *Generalized Linear Models.* London: Chapman and Hall.

Miller, J.J. (1977). "Asymptotic Properties of Maximum Likelihood Estimates in the Mixed Model of the Analysis of Variance". *The Annals of Statistics*, v5, no 4, 746-763.

Montgomery, D.C. (1991). *Design and Analysis of Experiments*, 3rd Ed. New York: John Wiley & Sons.

Puntanen, S. and Styan, P.H. (1989). "The Equality of the Ordinary Least Squares Estimator and the Best Linear Unbiased Estimator". *The American Statistician*, v 43, no 3, 153-161.

Rao, C.R. (1968). "A Note on a Previous Lemma in the Theory of Least Squares and Some Further Results". *Sankhyā*, ser A, v 30, 245-252.

Rogers, G.S. and Young, D.L. (1977). "Explicit Maximum Likelihood Estimators for Certain Patterned Covariance Matrices". *Communications in Statistics*, A6(2), 121-133.

Royden, H.L. (1988). *Real Analysis*, 3rd Ed. New York: Macmillan Publishing.

Rudin, Walter. (1976). *Principles of Mathematical Analysis*, 3rd Ed. New York: McGraw-Hill, Inc.

SAS Institute Inc. (1989). *SAS/STAT® User's Guide, Version 6*, 4th Ed. Cary, NC: SAS Institute Inc.

SAS Institute Inc. (1996). *SAS/STAT® Software: Changes and Enhancements through Release 6.11.* Cary, NC: SAS Institute Inc.

Schott, James R. (1997). *Matrix Analysis for Statistics*. New York: John Wiley & Sons, Inc.

Searle, S.R. and Casella, G. and McCulloch, C.E. (1992). *Variance Components*. New York: John Wiley and Sons.

Seely, J. and Zyskind, G. (1969). "Linear Spaces and Minimum Variance Unbiased Estimation". Technical Report No. 12, Department of Statistics, Oregon State University, Corvallis, Oregon.

Seely, J. (1969). "Quadratic Subspaces and Completeness for a Family of Normal Distributions". Technical Report No. 17, Department of Statistics, Oregon State University, Corvallis, Oregon.

Seely, J. (1971). "Quadratic Subspaces and Completeness". *The Annals of Mathematical Statistics*, v 42, no 2, 710-721.

Seely, J. (1977). "Minimal Sufficient Statistics and Completeness for Multivariate Normal Families". *Sankhyā*, v 39, ser A, pt 2, 176-185.

Seely, J. (1996). Linear Model Notes, unpublished. Oregon State University.

Smith, K.T. (1971). *Primer of Modern Analysis*. Belmont, CA: Bogden and Quigley, Inc.

Szatrowski, T.H. & Miller, John J. (1980). "Explicit Maximum Likelihood estimates from Balanced Data in the Mixed Model of the Analysis of Variance". *The Annals of Statistics*, v 8 , no 4, 811-819.

Szatrowski, T.H. (1980). "Necessary and Sufficient Conditions for Explicit Solutions in the Multivariate Normal Estimation Problem for Patterned Means and Covariances". *The Annals of Statistics*, v 8, no 4, 802-810.

VanLeeuwen, D.M. and Birkes, D.S. and Seely, J.F. (1997). "Balance and Orthogonality in Designs for Mixed Classification Models". Technical Report No. 168, Department of Statistics, Oregon State University, Corvallis, Oregon.

VanLeeuwen, D.M. and Seely, J.F. and Birkes, D.S. (1998). "Sufficient Conditions for Orthogonal Designs in Mixed Linear Models". *Journal of Statistical Planning and Inference*, v 73, 373-389.

Zyskind, G. (1967). "On Canonical Forms, Non-Negative Covariance Matrices and Best and Simple Least Squares Linear Estimators in Linear Models". *Annals of Mathematical Statistics*, v 38, 1092-1109.

Appendices

Appendix A - SAS Code Used for the Battery Life Examples

The SAS code is given below that was used to generate the output for the Battery Life Example I given in Table 8.2 for the ML procedure using the data set batt1:

```
proc mixed data=batt1 method=ml nobound noprofile scoring=30 itdetails asycov absolute;
   class T M;
   model Y = / solution;
   random T M T*M;
run;
```

The SAS code is given below that was used to generate the output for the Battery Life Example II given in Table 8.3 for the REML procedure using the data set batt2:

```
proc mixed data=batt2 method=reml nobound noprofile scoring=30 asycov absolute;
   class T M;
   model Y = T / solution;
   random M;
run;
```

The SAS code is given below that was used to generate the output for the Battery Life Example III given in Table 8.4 for the MIVQUE0 procedure using the data set batt3:

```
proc mixed data=batt3 method=mivque0 nobound noprofile scoring=30 asycov absolute;
   class T M;
   model Y = T / noint solution;
   random M T*M;
run;
```

Appendix B - Details For Checking the Conditions Using the Iterative Procedure

The purpose of this appendix is to provide details to accompany the discussion in section 8.1.5 pertaining to checking the conditions using the iterative procedure presented in section 8.1.1. The justification is complex and requires results and definitions concerning real analytic varieties and measure theory. Real analytic varieties are used to characterize the UBLUE conditions for the Y-Model in section 5.2.1 and the LQEM for $\underline{Z}$ in section 5.3.1. Measure theory results allow probabilistic conclusions to be made about the UBLUE conditions based on information which can be obtained from the iterative procedure. The following definitions will be useful in this section.

(Krantz and Parks,1992,p25)

real analytic function - A function $f : \Upsilon \to \mathcal{R}$ where Υ is a non-empty open set in $\mathcal{R}^m \ni \forall\ \underline{\theta} \in \Upsilon$ f can be represented by a convergent power series in some neighborhood of $\underline{\theta}$.

(Krantz and Parks,1992,p152)

real analytic variety - Set of common zeros in Υ of a finite set of real analytic functions.

(Smith,1971,p255)

regular function - $\underline{F} : \mathcal{R}^m \to \mathcal{R}^q$ is regular if $\frac{\partial \underline{F}(\underline{\theta})}{\partial \underline{\theta}}$ exists, is continous, and has maximal rank $\forall\ \underline{\theta} \in \mathcal{R}^m$.

(Smith,1971,p255)

smooth manifold - A smooth manifold of dimension k in $\mathcal{R}^m$ is a set $M \ni \forall\ \underline{a} \in M\ \exists$ a function $\underline{F} : \mathcal{R}^m \to \mathcal{R}^{m-k}$ which is regular on an open set Ω containing $\underline{a}$ and is such that $M \cap \Omega = \{\underline{\theta} \in \mathcal{R}^m | \underline{F}(\underline{\theta}) = \underline{0}\} \cap \Omega$.

(Lehmann,1983,p9)

lebesque measure - A probability measure λ_m defined on the smallest σ-algebra containing all open rectangles in $\mathcal{R}^m$. For $T = \{\underline{x} \in \mathcal{R}^m | a_i < x_i < b_i\ \ i = 1, ..., m\}$, $\lambda_m(T) = \prod_{i=1}^{m}(b_i - a_i)$.

These definitions are incorporated into the following three lemmas which will be used to derive the theorem.

(Krantz and Parks,1992,p25)

Lemma 1: Suppose f and g are real analytic functions with domain $\Upsilon \ni g(\underline{\theta}) \neq 0\ \forall\ \underline{\theta} \in \Upsilon$. Then $f + g$ and f/g are real analytic functions.

(Smith,1971,p299)

Lemma 2: A smooth submanifold in $\mathcal{R}^m$ of dimension $< m$ has m-dimensional lebesque measure 0.

Lemma 3: For $\underline{\theta} \in \Upsilon$ where Υ is a non-empty open set in $\mathcal{R}^m$, consider real polynomials $p(\underline{\theta})$ and $q(\underline{\theta})$ where $q(\underline{\theta}) \neq 0 \ \forall \ \underline{\theta} \in \Upsilon$, so that $r(\underline{\theta}) = \frac{p(\underline{\theta})}{q(\underline{\theta})}$ is a non-zero rational function. If $\mathcal{A} = \{\underline{\theta} \in \Upsilon | r(\underline{\theta}) = 0\}$, then $\lambda_m(\mathcal{A}) = 0$.

proof: i) By definition, p and q are real analytic functions

$\Rightarrow r = \frac{p}{q}$ is a real analytic function with domain Υ by lemma 1 as $q(\underline{\theta}) \neq 0 \ \forall \ \underline{\theta} \in \Upsilon$

$\Rightarrow \mathcal{A}$ is an real analytic variety in $\mathcal{R}^m$ by definition.

ii) Consider the notation and results in theorem 5.2.3 of Krantz and Parks (1992,p154) which establishes that a real analytic variety is the finite union of real analytic smooth submanifolds of dimensions $< m$.

For each $\underline{\theta}^0 \in \Upsilon$, define $r^*_{\underline{\theta}^0}(\underline{\theta}) = r(\underline{\theta} + \underline{\theta}^0)$, $Q^*_{\underline{\theta}^0} = \{\underline{\theta} \,|\, ||\underline{\theta}|| < \delta\}$, $Z^*_{\underline{\theta}^0} = \{\underline{\theta} \in Q^*_{\underline{\theta}^0} | r^*_{\underline{\theta}^0}(\underline{\theta}) = 0\} = \overset{f_{\underline{\theta}^0}}{\underset{j=1}{\cup}} \mathcal{B}^*_{\underline{\theta}^0,j}$

where $\mathcal{B}^*_{\underline{\theta}^0,j} = \{\underline{\theta} \in \mathcal{R}^m | [\theta_1, ..., \theta_k]' \in \Omega$ and $\underline{F}(\underline{\theta}) = 0\}$ with Ω open in $\mathcal{R}^k$ and

$\underline{F}(\underline{\theta}) = \begin{bmatrix} g_{k+1}(\theta_1, ..., \theta_k) - \theta_{k+1} \\ \vdots \\ g_m(\theta_1, ..., \theta_k) - \theta_m \end{bmatrix}$ for real analytic functions g_l defined on Ω. Note $\mathcal{B}^*_{\underline{\theta}^0,j}$ can be written as

$\mathcal{B}^*_{\underline{\theta}^0,j} = \{\underline{\theta} \in \mathcal{R}^m | \underline{F}(\underline{\theta}) = 0 \cap (\Omega \times \mathcal{R}^{m-k})\}$ where $M = \Omega \times \mathcal{R}^{m-k}$ is open in $\mathcal{R}^m$. In addition,

$\frac{\partial \underline{F}}{\partial \underline{\theta}} = [\frac{\partial \underline{g}}{\partial \underline{\theta}} \quad - I_{m-k}]$ has maximal rank $m - k \Rightarrow \underline{F}$ is regular on M by definition.

iii) Let $Q_{\underline{\theta}^0} = Q^*_{\underline{\theta}^0} + \underline{\theta}^0$. Note that $\underline{\theta}^1 \in Z^*_{\underline{\theta}^0} + \underline{\theta}^0 \Leftrightarrow \underline{\theta}^1 - \underline{\theta}^0 \in Z^*_{\underline{\theta}^0}$

$\Leftrightarrow \underline{\theta}^1 - \underline{\theta}^0 \in Q^*_{\underline{\theta}^0}$ and $r^*_{\underline{\theta}^0}(\underline{\theta}^1 - \underline{\theta}^0) = 0 \Leftrightarrow \underline{\theta}^1 \in Q^*_{\underline{\theta}^0} + \underline{\theta}^0 = Q_{\underline{\theta}^0}$ and $r(\underline{\theta}^1 - \underline{\theta}^0 + \underline{\theta}^0) = 0$

$\Leftrightarrow \underline{\theta}^1 \in \{\underline{\theta} \in Q_{\underline{\theta}^0} | r(\underline{\theta}) = 0\} = Z_{\underline{\theta}^0}$. Thus, $Z_{\underline{\theta}^0} = Z^*_{\underline{\theta}^0} + \underline{\theta}^0 = Q_{\underline{\theta}^0} \cap \mathcal{A}$ as the above holds $\forall \ \underline{\theta}^1$.

iv) From ii), $Q^*_{\underline{\theta}^0}$ is open $\Rightarrow Q_{\underline{\theta}^0}$ is open. Then $\underline{\theta}^0 \in Q_{\underline{\theta}^0} \ \forall \ \underline{\theta}^0 \Rightarrow \Upsilon \subset \underset{\underline{\theta} \in \Upsilon}{\cup} Q_{\underline{\theta}}$

$\Rightarrow \Upsilon \subset \overset{\infty}{\underset{i=1}{\cup}} Q_{\underline{\theta}^i}$ as $\mathcal{R}^m$ is separable (Royden,1988,p142)

$\Rightarrow \mathcal{A} = \Upsilon \cap \mathcal{A} \subset (\overset{\infty}{\underset{i=1}{\cup}} Q_{\underline{\theta}^i}) \cap \mathcal{A} = \overset{\infty}{\underset{i=1}{\cup}} (Q_{\underline{\theta}^i} \cap \mathcal{A}) = \overset{\infty}{\underset{i=1}{\cup}} Z_{\underline{\theta}^i}$ by iii).

v) From ii), $\mathcal{B}^*_{\underline{\theta}^0,j}$ is a smooth manifold of dimension $k < m$ by definition $\forall \ j, \underline{\theta}^0$

$\Rightarrow \lambda_m(\mathcal{B}^*_{\underline{\theta}^0,j}) = 0$ by lemma 2 $\forall \ j, \underline{\theta}^0$

$\Rightarrow \lambda_m(Z^*_{\underline{\theta}^0}) = \lambda_m(\overset{f_{\underline{\theta}^0}}{\underset{j=1}{\cup}} \mathcal{B}^*_{\underline{\theta}^0,j}) \leq \sum_{j=1}^{f_{\underline{\theta}^0}} \lambda_m(\mathcal{B}^*_{\underline{\theta}^0,j}) = 0 \ \ \forall \ \underline{\theta}^0$ by ii) and subaddivity Royden (1988, p57)

$\Rightarrow \lambda_m(Z_{\underline{\theta}^0}) = \lambda_m(Z^*_{\underline{\theta}^0} + \underline{\theta}^0) = \lambda_m(Z^*_{\underline{\theta}^0}) = 0 \ \ \forall \ \underline{\theta}^0$ by iii) and translation invariance Royden (1988, p58)

$\Rightarrow \lambda_m(\mathcal{A}) = \lambda_m(\overset{\infty}{\underset{i=1}{\cup}} Z_{\underline{\theta}^i}) \leq \sum_{i=1}^{\infty} \lambda_m(Z_{\underline{\theta}^i})$ by iv) and subadditivity Royden (1988). ∎

Theorem: Let $\Xi^* = \{\underline{\psi} \in \mathcal{R}^{k+1} | V_{\underline{\psi}} \text{ is PD}\}$.

i) If $\mathcal{Z} = \{\underline{\psi} \in \Xi^* | \Lambda'(X'V_{\underline{\psi}}^{-1}X)^{-1}X'V_{\underline{\psi}}^{-1} = \Lambda'(X'X)^{-1}X'\}$, then $\lambda_{k+1}(\mathcal{Z}) = 0$ or $\mathcal{Z} = \Xi^*$.

ii) If $\mathcal{Z} = \{\underline{\psi} \in \Xi^* | \Gamma'(X^{\dagger *}V_{\underline{\psi}}^{\dagger -1}X^{\dagger})^{-1}X^{\dagger *}V_{\underline{\psi}}^{\dagger -1} = \Gamma'(X^{\dagger *}X^{\dagger})^{-1}X^{\dagger *}\}$, then $\lambda_{k+1}(\mathcal{Z}) = 0$ or $\mathcal{Z} = \Xi^*$.

proof: From section 3.1.2, Ξ^* is open in $\mathcal{R}^{k+1}$.

i) The entries of the matrix $\Lambda'(X'V_{\underline{\psi}}^{-1}X)^{-1}X'V^{-1} - \Lambda'(X'X)^{-1}X'$ are ratios of polynomials in $\psi_1, ..., \psi_{k+1}$ given by $r_{ij}(\underline{\psi}) = \frac{p_{ij}(\underline{\psi})}{q_{ij}(\underline{\psi})}$. The denominators $q_{ij}(\underline{\psi}) \neq 0 \ \forall \ \underline{\psi} \in \Xi^*$ since $V_{\underline{\psi}}$ is PD $\forall \ \underline{\psi} \in \Xi^*$.

(1) If at least one entry $r_{ij}(\underline{\psi}) \neq 0$, then $\lambda_{k+1}(\mathcal{Z}) = 0$ by lemma 3.

(2) If $r_{ij}(\underline{\psi}) = 0 \ \forall \ i, j$, then $\mathcal{Z} = \Xi^*$ by definition of $\mathcal{Z}$.

ii) For $A \in \mathcal{S}_n$, define $\mathcal{Z}_A = \{\underline{\psi} \in \Xi^* | \Gamma'(X^{\dagger *}V_{\underline{\psi}}^{\dagger -1}X^{\dagger})^{-1}X^{\dagger *}V_{\underline{\psi}}^{\dagger -1}A = \Gamma'(X^{\dagger *}X^{\dagger})^{-1}X^{\dagger *}A\}$. Note $\Gamma'(X^{\dagger *}V_{\underline{\psi}}^{\dagger -1}X^{\dagger})^{-1}X^{\dagger *}V_{\underline{\psi}}^{\dagger -1}A = \Gamma'\{\frac{1}{2}\text{tr}(R_iR_{\underline{\psi}}^{-1}R_jR_{\underline{\psi}}^{-1})\}^{-1}\{\frac{1}{2}\text{tr}(R_iR_{\underline{\psi}}^{-1}AR_{\underline{\psi}}^{-1})\}$ by lemma 1 section 3.2.3. The entries of the matrix $\Gamma'(X^{\dagger *}V_{\underline{\psi}}^{\dagger -1}X^{\dagger})^{-1}X^{\dagger *}V_{\underline{\psi}}^{\dagger -1}A - \Gamma'(X^{\dagger *}X^{\dagger})^{-1}X^{\dagger *}A$ are ratios of polynomials in $\psi_1, ..., \psi_{k+1}$ given by $r_{ij}(\underline{\psi}) = \frac{p_{ij}(\underline{\psi})}{q_{ij}(\underline{\psi})}$. The denominators $q_{ij}(\underline{\psi}) \neq 0 \forall \ \underline{\psi} \in \Xi^*$ since $V_{\underline{\psi}}$ is PD $\forall \ \underline{\psi} \in \Xi^*$.

(1) If for some $A \in \mathcal{S}_n$, $\exists$ at least one entry $r_{ij}(\underline{\psi}) \neq 0$, then $\lambda_{k+1}(\mathcal{Z}_A) = 0$ by lemma 3. Then $\mathcal{Z} = \bigcap_{A \in \mathcal{S}_n} \mathcal{Z}_A \Rightarrow \lambda_{k+1}(\mathcal{Z}) \leq \lambda_{k+1}(\mathcal{Z}_A) = 0$ by monotonicity (Royden,1988,p55).

(2) If $\forall \ A \in \mathcal{S}_n$, $r_{ij}(\underline{\psi}) = 0 \ \forall \ i, j$, then $\mathcal{Z} = \bigcap_{A \in \mathcal{S}_n} \mathcal{Z}_A = \bigcap_{A \in \mathcal{S}_n} \Xi^* = \Xi^*$. ∎

By the Y-UBLUE and LQZ-UBLUE theorems in chapter 5, the UBLUE condition, or GZC, is equivalent to $\mathcal{Z} = \Xi^*$. Consider the unknown parameter $\xi = \begin{cases} 1 & \mathcal{Z} = \Xi^* \\ 0 & \mathcal{Z} \neq \Xi^* \end{cases}$. When $\xi = 1$, the GZC will hold $\forall \ \underline{\psi} \in \Xi^*$. When $\xi = 0$, the GZC will not hold $\forall \ \underline{\psi} \in \Xi^*$.

In order to devise a method for checking the conditions based on the results of the above theorem, it will be useful to consider probability measures P defined on $\mathcal{R}^m$ that are absolutely continuous with respect to λ_m. The probability measure on the set $\mathcal{A}$ can be written as $P(\mathcal{A}) = \int_{\mathcal{A}} p \, d\lambda_m$ where $0 \leq p \leq 1$ is the probability density of P (Lehmann,1983). Then $\lambda_m(\mathcal{A}) = 0 \Rightarrow P(\mathcal{A}) = 0$ since $0 \leq P(\mathcal{A}) = \int_{\mathcal{A}} p \, d\lambda_m \leq \int_{\mathcal{A}} (1) d\lambda_m = \lambda_m(\mathcal{A}) = 0$.

Suppose $\underline{\psi}$ is randomly distributed with respect to an absolutely continous probability distribution $P_\xi^{\underline{\psi}}$ with support Ξ^*. Also, suppose the random observation vector $\underline{Y}$ is distributed with respect to an absolutely continous probability distribution $P_\xi^{\underline{Y}}$ with support $\mathcal{Y}$ that contains a non-empty open set in $\mathcal{R}^n$. Also, assume that $\underline{\psi}$ and $\underline{Y}$ are independent. Let P_ξ denote the joint probability distribution defined on $\mathcal{R}^n \times \mathcal{R}^{k+1}$. For $\mathcal{A} = \mathcal{A}_1 \times \mathcal{A}_2$ where $\mathcal{A} \subset \mathcal{R}^n \times \mathcal{R}^{k+1}$, $\mathcal{A}_1 \subset \mathcal{R}^n$, $\mathcal{A}_2 \subset \mathcal{R}^{k+1}$, the joint probability distribution is defined by $P_\xi(\mathcal{A}) = \int_{\mathcal{A}} p_\xi d\lambda_m = \int_{\mathcal{A}_1}[\int_{\mathcal{A}_2} p_\xi^{\underline{\psi}} p_\xi^{\underline{Y}} d\lambda_m] d\lambda_m = P_\xi^{\underline{\psi}}(\mathcal{A}_1)P_\xi^{\underline{Y}}(\mathcal{A}_2)$ using the independence of $\underline{\psi}$ and $\underline{Y}$ (Lehmann,1986,p40).

Let $D(\underline{\psi})$ be the difference between the two quantities given in the set $\mathcal{Z}$ and let $B(\underline{Y})$ be the response for either the Y-Model, the ALQEM for $(\underline{Y} - X\widehat{\underline{\beta}})$, or the LQEM for $N_X\underline{Y}$. Then $T(\underline{\psi}, \underline{Y}) = D(\underline{\psi})B(\underline{Y})$ corresponds to the difference between the GLSE and the LSE in the appropriate model.

The statistic $T(\underline{\psi}, \underline{Y})$ and the probability distribution P_ξ will be used to draw inference about the unknown parameter ξ. This will be accomplished by generating a confidence region for ξ. Define the confidence region $C(T(\underline{\psi}, \underline{Y})) = \begin{cases} \{1\} & \text{if } T(\underline{\psi}, \underline{Y}) = \underline{0} \\ \{0\} & \text{if } T(\underline{\psi}, \underline{Y}) \neq \underline{0} \end{cases}$. The following theorem gives the coverage probability for this confidence region (Casella and Berger,p404).

Theorem: $C(T(\underline{\psi}, \underline{Y}))$ is a 100% confidence region for ξ as $P_\xi(\xi \in C(T(\underline{\psi}, \underline{Y}))) = 1$.

proof: i) $\xi = 1 \Rightarrow \mathcal{Z} = \Xi^* \Rightarrow D(\underline{\psi}) = 0 \ \forall \ \underline{\psi} \in \Xi^* \Rightarrow T(\underline{\psi}, \underline{Y}) = D(\underline{\psi})B(\underline{Y}) = 0 \ \ \forall \ \underline{\psi} \in \Xi^*, \ \underline{Y} \in \mathcal{R}^n$

$\Rightarrow P_\xi(T(\underline{\psi}, \underline{Y}) = 0) = 1$

$\Rightarrow C(T(\underline{\psi}, \underline{Y})) = \{1\} \ \ \forall \ \underline{\psi} \in \Xi^*, \ \underline{Y} \in \mathcal{R}^n$ w.p. 1 with respect to P_ξ by definition of $C(T(\underline{\psi}, \underline{Y}))$

$\Rightarrow P_\xi(1 \in C(T(\underline{\psi}, \underline{Y}))) = 1.$

ii) $\xi = 0 \Rightarrow \mathcal{Z} \neq \Xi^* \Rightarrow \lambda_{k+1}(\mathcal{Z}) = 0$ by the above theorem

$\Rightarrow P_\xi^{\underline{\psi}}(\underline{\psi} \in \mathcal{Z}) = 0$ by absolute continuity $\Rightarrow P_\xi^{\underline{\psi}}(\underline{\psi} \notin \mathcal{Z}) = 1$

$\Rightarrow \underline{\psi} \notin \mathcal{Z}$ w.p. 1 with respect to $P_\xi^{\underline{\psi}} \Rightarrow D(\underline{\psi}) \neq 0$ w.p. 1 with respect to $P_\xi^{\underline{\psi}}$

$\Rightarrow P_\xi^{\underline{Y}}(T(\underline{\psi}, \underline{Y}) = D(\underline{\psi})B(\underline{Y}) = 0 | \underline{\psi}) = 0$ w.p. 1 with respect to $P_\xi^{\underline{\psi}}$

$\Rightarrow \mathrm{E}[P_\xi^{\underline{Y}}(T(\underline{\psi}, \underline{Y}) = 0 | \underline{\psi})] = 0$

$\Rightarrow \int_F P_\xi^{\underline{Y}}(T(\underline{\psi}, \underline{Y}) = 0 | \underline{\psi}) p_\xi^{\underline{\psi}}(\underline{\psi}) d\lambda_m = 0$ where $F = \{\underline{\psi} | \, T(\underline{\psi}, \underline{Y}) = 0\}$

$\Rightarrow \int_F [\int_G p_\xi^{\underline{Y}}(\underline{Y} | \underline{\psi}) p_\xi^{\underline{\psi}}(\underline{\psi}) d\lambda_m] d\lambda_m = 0$ where $G = \{\underline{Y} | \, T(\underline{\psi}, \underline{Y}) = 0\}$

$\Rightarrow \int_F [\int_G p_\xi^{\underline{Y}}(\underline{Y}) p_\xi^{\underline{\psi}}(\underline{\psi}) d\lambda_m] d\lambda_m = 0$ as $\underline{\psi}$ and $\underline{Y}$ are independent

$\Rightarrow \int_{F \times G} p_\xi(\underline{\psi}, \underline{Y}) d\lambda_m = 0$ as $\underline{\psi}$ and $\underline{Y}$ are independent

$\Rightarrow P_\xi(T(\underline{\psi}, \underline{Y}) = 0) = 0 \Rightarrow P_\xi(T(\underline{\psi}, \underline{Y}) \neq 0) = 1$

$\Rightarrow C(T(\underline{\psi}, \underline{Y})) = \{0\} \ \ \forall \ \underline{\psi} \in \Xi^*, \ \underline{Y} \in \mathcal{R}^n$ w.p. 1 with respect to P_ξ by definition of $C(T(\underline{\psi}, \underline{Y}))$

$\Rightarrow P_\xi(0 \in C(\underline{\psi}, \underline{Y})) = 1.$

$\therefore$ By i) and ii), $C(T(\underline{\psi}, \underline{Y}))$ is a 100% confidence region for ξ as $P_\xi(\xi \in C(T(\underline{\psi}, \underline{Y}))) = 1$. ∎

Note the above proof also shows that $\xi = 1$ implies $P_\xi(T(\underline{\psi}, \underline{Y}) = 0) = 1$ and $\xi = 0$ implies $P_\xi(T(\underline{\psi}, \underline{Y}) \neq 0) = 1$. The above two theorems in this section prove the details given in section 8.1.5. Section 8.1.5 provides a method for checking the GZC using the iterative procedure presented in section 8.1.1. Methods A and B generate the statistic $T(\underline{\psi}, \underline{Y})$, or equivalently $S(\underline{\psi}, \underline{Y})$, and state the results using the 100% confidence region given above (section 8.1.5).

Appendix C - Summary of Models and Theorems

This appendix gives a summary of the models that are considered in this thesis as well as the associated theorems that are of main importance. It is hoped that this summary will provide an easy reference to help the reader.

The Underlying Model

Purpose: Examine the UBLUE in a general framework that can be applied to the special cases.

Notation: $\omega \in \mathcal{W}$, $\theta \in \Upsilon_{\mathcal{U}} \subset \mathcal{P}$, $U \in \mathcal{L}(\mathcal{P}, \mathcal{W})$, $\mathcal{V} \subset \mathcal{L}_{\mathrm{PD}}(\mathcal{W}, \mathcal{W})$.

Definition: $\mathrm{E}_{\theta}[\omega] = U\theta \qquad \mathrm{Cov}(\omega) = V \in \mathcal{V}$ (section 3.3.1).

UBLUE for Full Case: U-FUBLUE Theorem (section 4.1).

UBLUE for General Case: U-UBLUE Theorem (section 5.1).

UBLUE for General Case in Full Rank Setting: U-UBLUE$_{\mathrm{FR}}$ Theorem (section 6.1).

Special Cases of the Underlying Model

A. Y-Model

Purpose: Use to model the fixed effects for the ML method.

Notation: $\underline{Y} \in \mathcal{R}^n$, $\underline{\beta} \in \mathcal{R}^p$, $X_{n \times p}$, $V_{\underline{\psi}\, n \times n}$ is PD for all $\underline{\psi} \in \Xi$.

Definition: $\mathrm{E}_{\underline{\beta}}[\underline{Y}] = X\underline{\beta} \quad \mathrm{Cov}(\underline{Y}) = V_{\underline{\psi}}$ (section 3.1.1).

UBLUE for Full Case: U-FUBLUE Theorem (section 4.2.1).

UBLUE for General Case: Y-UBLUE Theorem (section 5.2.1).

UBLUE for General Case in Full Rank Setting: Y-UBLUE$_{\mathrm{FR}}$ Theorem (section 6.3.1).

ELMLQE for Full Case: Y-FELMLQE Theorem (section 4.2.2).

ELMLQE for General Case: Y-ELMLQE Theorem (section 5.2.2).

ELMLQE for General Case in Full Rank Setting: Y-ELMLQE$_{\mathrm{FR}}$ Theorem (section 6.3.2).

B. ALQEM for $(\underline{Y} - X\widehat{\underline{\beta}})$

Purpose: Use to model the variance components for the ML method.

Notation: $\underline{Z} = \underline{Y} - X\widehat{\underline{\beta}} = (I - X(X'V_{\widehat{\underline{\psi}}}^{-1}X)^{-}X'V_{\widehat{\underline{\psi}}}^{-1})\underline{Y}$, $Y_2^{\diamond} = \underline{Z}\underline{Z}'$, $\underline{\psi} \in \Xi$.

Definition: $\mathrm{E}_{\underline{\psi}}[Y_2^{\diamond}] = X^{\diamond}\underline{\psi}$ $\quad \mathrm{Cov}(Y_2^{\diamond}) = V_{\underline{\psi}}^{\diamond} = 2\Psi_{V_{\underline{\psi}}}$ (section 3.2.1).

UBLUE for Full Case: LQZ-FUBLUE Theorem (section 4.3.1).

UBLUE for General Case: LQZ-FUBLUE Theorem (section 5.3.1).

UBLUE for General Case in Full Rank Setting: LQZ-UBLUE$_{\mathrm{FR}}$ Theorem (section 6.4.1).

ELMLQE for Full Case: ALQNY-FELMLQE Theorem (section 4.3.2).

ELMLQE for General Case: ALQNY-ELMLQE Theorem (section 5.3.2).

ELMLQE for General Case in Full Rank Setting: ALQNY-ELMLQE$_{\mathrm{FR}}$ Theorem (section 6.4.2).

C. LQEM for $N_X\underline{Y}$

Purpose: Use to the model variance components for the REML method.

Notation: $\underline{Z} = N_X\underline{Y}$, $Y^{\circ} = \underline{Z}\underline{Z}'$, $\underline{\psi} \in \Xi$.

Definition: $\mathrm{E}_{\underline{\psi}}[Y^{\circ}] = X^{\circ}\underline{\psi}$ $\quad \mathrm{Cov}(Y^{\circ}) = V_{\underline{\psi}}^{\circ} = 2\Psi_{N_XV_{\underline{\psi}}N_X}$ (section 3.2.1).

UBLUE for Full Case: LQZ-FUBLUE Theorem (section 4.3.1).

UBLUE for General Case: LQZ-FUBLUE Theorem (section 5.3.1).

UBLUE for General Case in Full Rank Setting: LQZ-UBLUE$_{\mathrm{FR}}$ Theorem (section 6.4.1).

ELREMLQE for Full Case: LQNY-FELREMLQE Theorem (section 4.3.3).

ELREMLQE for General Case: LQNY-ELREMLQE Theorem (section 5.3.3).

ELREMLQE for General Case in Full Rank Setting: LQNY-ELREMLQE$_{\mathrm{FR}}$ Theorem (section 6.4.3).

Appendix D - Abbreviations and Symbols

CQS = Commutative Quadratic Subspace (2.8)

EGLSE = Estimated Generalized Least Squares Estimator (3.3.3)

ELMLQE = Explicit Linear Maximum Likelihood Equation Estimator (5.2, 5.3)

ELREMLQE = Explicit Linear Restricted Maximum Likelihood Equation Estimator (5.3)

FELMLQE = Full ELMLQE (4.2, 4.3)

FELREMLQE = Full ELREMLQE (4.3)

FUBLUE = Full UBLUE (3.3.4)

g-inverse = Generalized Inverse (2.5)

GLSE = Generalized Least Squares Estimator (3.3.3)

GZC = Generalized Zyskinds Condition (5.2)

IBLUE = Best Linear Unbiased Estimation with respect to multiple of identity transformation I (3.3.4)

LQEM = Linearized Quadratic Estimation Models (3.2)

LSE = Least Squares Estimator (3.3.3)

ML = Maximum Likelihood (2.4.4)

MLQE = Maximum Likelihood Equation Estimator (3.1.4)

NND = Non-Negative Definite (2.1.1)

OPO = Orthogonal Projection Operator (2.1)

QS = Quadratic Subspace (2.7)

PD = Positive Definite (2.1.1)

PO = Projection Operator (2.1)

REML = Restricted Maximum Likelihood (3.1.4)

REMLQE = Restricted Maximum Likelihood Equation Estimator (3.1.4)

UBLUE = Uniformly Best Linear Unbiased Estimator (3.3.4)

VBLUE = Best Linear Unbiased Estimator with respect to NND transformation V (3.3.4)

ZC = Zyskinds Condition (3.3.5)

$\ni$ = such that

$\exists$ = there exists

$\Rightarrow$ = implies

$\Leftrightarrow$ = if and only if

A = matrix or linear transformation (2.1)

A^* = adjoint of A when A is a linear transformation (2.1)

A' = transpose of A if A is a matrix (2.1)

$\underline{a}$ = vector (2.1)

$\underline{R}(A)$ = range of A (2.1)

$\underline{r}(A)$ = rank of A (2.1)

$\underline{N}(A)$ = null space of A (2.1)

$\underline{R}(A)$ = range space of A (2.1)

$\mathcal{U}$ = subspace (2.1)

$\mathcal{U}^{\perp}$ = orthogonal complement of $\mathcal{U}$ (2.1)

$\dim \mathcal{U}$ = dimension of $\mathcal{U}$ (2.1)

$\mathcal{S}_n$ = set of symmetric $n \times n$ matrices (2.1)

$\mathcal{M}_{n \times m}$ = set of $n \times m$ matrices (2.1)

P_A = OPO on $\underline{R}(A)$ (2.5)

A^- = g-inverse of A (2.5)

A^+ = Moore-Penrose Inverse for A (2.5)

$\mathrm{tr}(A)$ = trace of A (2.6)

$\mathrm{vec}(A)$ = vector form of the matrix A (2.9)

$A \odot B$ = horizontal direct product between the matrices A and B (2.9)

$\mathcal{U} \oplus \mathcal{V}$ = direct sum $= \mathcal{U} + \mathcal{V} \ni \mathcal{U} \cap \mathcal{V} = \{0\}$ (2.5)

Y-Model = the original model of interest (3.1)

[L] = linearity assumption for the covariance matrix in the Y-Model (3.1)

[O] = open set assumption for parameters of the covariance matrix in the Y-Model or in U-Model (3.1)

[N] = normality assumption for the Y-Model (3.1)

[C] = classification assumption for the Y-Model (3.1)

Bal($\mathcal{G}$) = balance with respect to a particular subset of factors $\mathcal{G}$ (3.1.3)

LQEM for Z = Linearized Quadratic Estimation Model for the random vector $\underline{Z}$ (3.2)

Ψ = A linear transformation from $\mathcal{S}_n \rightarrow \mathcal{S}_n$ given by $\Psi_\Sigma(A) = \Sigma A \Sigma$ (3.2.2)

U-Model = Underlying Model (3.3)

[S] = spanning assumption in the U-Model (3.3)

$\mathcal{L}_{\mathrm{PD}}(\mathcal{W}, \mathcal{W})$ = the set of PD linear transformations mapping $\mathcal{W} \rightarrow \mathcal{W}$ (3.2.2, 3.3.1)

$\mathcal{L}_{\mathrm{NND}}(\mathcal{W}, \mathcal{W})$ = the set of NND linear transformations mapping $\mathcal{W} \rightarrow \mathcal{W}$ (3.3.4)