UNIVERSITY OF CALIFORNIA
RIVERSIDE

Search for the Optimum Variance Components Estimates in Mixed Effects Models

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Applied Statistics

by

Luyao Peng

September 2019

Dissertation Committee:

Dr. Subir Ghosh, Chairperson
Dr. Weixin Yao
Dr. Zhenyu Jia

The Dissertation of Luyao Peng is approved:

Committee Chairperson

University of California, Riverside

Acknowledgments

I would like to gratefully and sincerely thank my advisor Dr. Subir Ghosh for his guidance, understanding, patience, and consistent encouragement during my graduate study. In Fall 2013, I took the class of Dr. Ghosh for the first time, which got me to know Dr. Ghosh and opened up my vision and knowledge in statistical world. After that, I started my study in Statistics starting from undergraduate to graduate courses, transitioned to the master program and then the PhD program in Statistcs. Dr. Ghosh provided me tremendous help and support, without which I could not have made it through this long process. After working with Dr. Ghosh, not only has he been a great mentor in my research, but he has also guided and advised me in my life (especially when I experienced two important personal changes in my life). I will keep in mind of the candid and pertinent advises from Dr. Ghosh, which encourage me to keep on fixing my shortcommings and making progress. In Chinese, people always say thanks to the help of the benefactors no matter how difficult the life is. I believe Dr. Ghosh is the benefactor in my life.

Besides my advisor, I would like to thank the rest of my thesis committee members Dr. Weixin Yao for giving incredible suggestions on my research and offering me the opportunity of transferring to the statistics program, and Dr. Zhenyu Jia, a super friendly professor, for his incredible suggestions, support and time.

I am grateful to all the faculty members for helping me establish solid academic background knownedge. Especially, I want to express my gratitude to Dr. Jun Li for her help, academic suggestions and time, she is more like a friend of mine than a professor, Dr. Shujie Ma for sharing her personal academic and growth experiences. I also want to extend my thanks to all my friends in UCR for their caring, love, and support. Thank you for always being there for me and encouraging me whenever I am frustrated or disturbed.

I owe my deepest thanks to my husband, who is my spiritual support, beloved person and soul mate. I also owe my deepest thanks to my families, even thousands miles away, you have always been my pillar, joy, and driving force.

ABSTRACT OF THE DISSERTATION

Search for the Optimum Variance Components Estimates in Mixed Effects Models

by

Luyao Peng

Doctor of Philosophy, Graduate Program in Applied Statistics
University of California, Riverside, September 2019
Dr. Subir Ghosh, Chairperson

This dissertation aims at searching for the optimum variance components estimates in the mixed-effects model. Traditional estimation methods of the variance components include the analysis of variance/method of moment (ANOVA/MoM) estimation, which is the optimum estimation (OPE) when the data are balanced, the maximum likelihood estimation (MLE) and the restricted maximum likelihood estimation (REMLE). However, when the data have small sample sizes and unbalanced structures, the optimum estimates do not exist, ML estimates are biased, MLE and REMLE cannot provide the closed-form expressions of the estimates to study their small-sample statistical properties. To solve those problems, we proposed the near optimum estimation (NOPE) method and the average optimum estimation (AOPE) method when the data are unbalanced in DOE. When estimating σ_2^2 and a linear function of variance components $\sigma_1^2 + p_2\sigma_2^2$ in SAE, we proposed methods of finding the unbiased quadratic estimators with smaller variances than the corresponding MoM estimators. We presented simulation studies to evaluate the estimation performance of our proposed methods and compare them with MoM, ML and REML. All of our pro-

posed estimators have closed-form expressions and do not require the functional form in the distributional assumptions.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Literature Review

Mixed-effects models are widely used in analyzing correlated data, especially the data with repeated measurements made on the same observational unit or the measurements are made within clusters of related observational units. When drawing statistical inferences with respect to the fixed effects on the dependent variable, the estimates of the fixed effects coefficients are dependent on the estimated variance components in the mixed-effects model, therefore, the estimation of the variance components is important in drawing inferences in mixed-effects models.

A large literature is available for the estimation methods of variance components in the mixed-effects model. Optimum estimation (OPE) method is the best quadratic unbiased estimator among the class of unbiased quadratic estimators for the variance components. When the data are balanced, the optimum estimators exist and include the analysis of variance (ANOVA) estimators / Method of Moment (MoM) estimators[12]. Under the normality assumptions, the optimum estimators include the Minimum Variance (MINVAR) quadratic estimators[20], the Minimum Mean Square (MIMS) quadratic estimators[16]. Other standard estimation methods are the Maximum Likelihood (ML) estimation methods (9 (), 10 ()), and the Restricted Maximum Likelihood (REML) method[18]. Both MLE and REMLE are based on distributional assumptions and they have some nice properties such as asymptotically consistency, efficiency and normality under some regularity conditions.

Despite the merits of those estimation methods, there are two major problems. First, when the data are unbalanced, the OP estimators do not exist, even though ML and REML can give numeric values of the estimated variance components under certain distributional assumptions, the explicit forms of those estimators remain unknown, so the small-sample statistical properties of the estimators cannot be studied, which is not desirable for the fields of study where small samples are common, such as the experimental design and the small area estimation. Second, both ML and REML estimators require normality assumptions and may not be robust in estimation when the normality assumption is not valid.

We aim to solve for those problems in two areas, the design of experiment (DOE) and the small area estimation (SAE). In DOE, we propose the method of near optimum estimation (NOPE), which do not require the functional form of the distribution assumptions and is near optimum with explicit form when the optimum estimators do not exist. When the experimental design is replicated, the average optimum estimation (AOPE) method is proposed. In SAE, we propose to use the MoM estimators as a benchmark to find the unbiased estimators with explicit forms and smaller variances than the corresponding MoM estimators. We also propose to find the approximated $\boldsymbol{A}$ matrix for ML and REML estiamtes.

1.2 Thesis Contribution

This thesis searches for the optimum quadratic estimator, $\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}$, to estimate the variance components in the mixed-effects model.

In SAE, our proposed estimators are unbiased and having smaller variances than corresponding MoM estimators when the optimum estimates do not exist, which is demonstrated by using an example. Our proposed estimators also have closed-form expressions and do not require functional form of the distribution assumption. In addition, we found approximated $\boldsymbol{A}$ matrix for ML and REML estiamtes.

In DOE, we proposed the NOPE and AOPE using replications. The simulations demonstrated the comparable estimation performance of our methods with ML and REML under normality assumptions. Simulations also demonstrated the robustness of our methods against the departure from normality by using the skew normal distrubution for both AOPE and NOPE compared with MoM, ML and REML.

1.3 Thesis Outline

In Chapter 2, the quadratic forms, the expecation, variance and covariance expressions are introduced. In Chapter 3, the mixed-effects models are introduced first, the unbiased quadratic estimators for the variance components in the mixed-effects model as well as the variance of the quadratic estimators will be reviewed and illustrated using an example. In Chapter 4, for the mixed-effects models in DOE, the NOPE and AOPE methods are introduced using the Unbalanced Incomplete Block Design (UIBD) and the UIBD with replications. In Chapter 5, two mixed-effects models in SAE, Fay-Herriot model and the nested-error mixed-effects model, are introduced. For each SAE model, the algorithmic way of finding the class of unbiased quadratic estimators for the variance components is introduced using the MoM estimators as a benchmark, then the procedures of finding a subclass of unbiased quadratic estimators with smaller variance than that of the corresponding MoM estimators are illustrated using an example.

Chapter 2

The Quadratic Forms

2.1 Introduction

Let $\boldsymbol{B}$ be a real and symmetric $k \times k$ matrix and $\boldsymbol{b}$ be a $k \times 1$ random vector. An expression of the form $\boldsymbol{b}'\boldsymbol{B}\boldsymbol{b}$ is called a quadratic form. We can write the quadratic form as

$$\boldsymbol{b}'\boldsymbol{B}\boldsymbol{b} = \sum_{i=1}^{k}\sum_{j=1}^{k} b_i B_{i,j} b_j. \tag{2.1.1}$$

2.2 Expectation

Theorem 2.1. For $\boldsymbol{b} = (b_1, b_2, \ldots, b_k)'$, we assume that b_i's are independent with mean 0 and variance σ_i^2. Then, we have

$$E(\boldsymbol{b}'\boldsymbol{B}\boldsymbol{b}) = tr(\boldsymbol{B}\Delta_1), \tag{2.1.2}$$

where

$$\Delta_1 = \begin{bmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_k^2 \end{bmatrix}.$$

Proof:

Since b_i's are independent with mean 0, variance σ_i^2, we have

$$Var(\boldsymbol{b}) = \Delta_1 = \begin{bmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_k^2 \end{bmatrix},$$

$$\begin{aligned} E(\boldsymbol{b}'\boldsymbol{B}\boldsymbol{b}) &= \sum_{i=1}^{k}\sum_{j=1}^{k} E(b_i B_{i,j} b_j) \\ &= \sum_{i=1}^{k}\sum_{j=1}^{k} B_{i,j} E(b_i b_j) \\ &= \sum_{i=1}^{k}\sum_{j=1}^{k} B_{i,j} [Var(\boldsymbol{b})]_{i,j}. \end{aligned}$$

Since $Var(\boldsymbol{b}) = \Delta_1$, which is a diagonal matrix, only the terms with index $i = j$ are non-zero, then we have

$$\begin{aligned} \sum_{i=1}^{k}\sum_{j=1}^{k} B_{i,j} [Var(\boldsymbol{b})]_{i,j} &= \sum_{i=1}^{k} [\boldsymbol{B}\Delta_1]_{i,i} \\ &= tr(\boldsymbol{B}\Delta_1). \end{aligned}$$

2.3 Variance

Theorem 2.2. For $\boldsymbol{b} = (b_1, b_2, \ldots, b_k)'$, we assume that b_i's are independent with the mean of 0, variance σ_i^2 and kurtosis γ_i. Then, we have

$$Var(\boldsymbol{b}'\boldsymbol{B}\boldsymbol{b}) = 2tr(\boldsymbol{B}\Delta_1\boldsymbol{B}\Delta_1) + tr(\tilde{\boldsymbol{B}}\Delta_2\tilde{\boldsymbol{B}}), \tag{2.3.1}$$

where $\tilde{\boldsymbol{B}}$ is the diagonal matrix with the same diagonal elements as the matrix $\boldsymbol{B}$ and

$$\Delta_1 = \begin{bmatrix} \sigma_1^2 & 0 & \ldots & 0 \\ 0 & \sigma_2^2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \sigma_k^2 \end{bmatrix}, \Delta_2 = \begin{bmatrix} \gamma_1\sigma_1^4 & 0 & \ldots & 0 \\ 0 & \gamma_2\sigma_2^4 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \gamma_k\sigma_k^4 \end{bmatrix}.$$

Proof:

For a scalar $\boldsymbol{b}'\boldsymbol{B}\boldsymbol{b}$, it is known that

$$Var(\boldsymbol{b}'\boldsymbol{B}\boldsymbol{b}) = E[(\boldsymbol{b}'\boldsymbol{B}\boldsymbol{b})^2] - [E(\boldsymbol{b}'\boldsymbol{B}\boldsymbol{b})]^2. \tag{2.3.2}$$

Since $\boldsymbol{b}'\boldsymbol{B}\boldsymbol{b} = \sum_{i=1}^{k}\sum_{j=1}^{k} b_i B_{i,j} b_j$ in (2.1.1), we can write

$$(\boldsymbol{b}'\boldsymbol{B}\boldsymbol{b})^2 = \sum\sum\sum_{1 \le i,j,m,n \le k}\sum B_{i,j}B_{m,n}b_ib_jb_mb_n. \tag{2.3.3}$$

Because b_i's are independent with mean 0 and variance σ_i^2, we have $E(b_i) = 0, E(b_i^2) = \sigma_i^2, E(b_i^4) = (\gamma_i + 3)\sigma_i^4$,

$$E(b_ib_jb_nb_m) = \begin{cases} (\gamma_i + 3)\sigma_i^4, \text{ for } i = j = k = l, \\ \sigma_i^2\sigma_j^2, \text{ for } i = j \neq m = n \text{ or } i = m \neq j = n \text{ or } i = n \neq m = j, \\ 0, \text{ otherwise,} \end{cases}$$

and $\boldsymbol{B}$ is symmetric matrix, i.e. $B_{i,j} = B_{j,i}$, the first term in (2.3.2) is equal to

$$\begin{aligned} E[(\boldsymbol{b}'\boldsymbol{B}\boldsymbol{b})^2] &= \sum_{i=1}^{k} B_{i,i}^2(\gamma_i + 3)\sigma_i^4 + \sum\sum_{1\le i\ne m\le k} B_{i,i}B_{m,m}\sigma_i^2\sigma_m^2 + \sum\sum_{1\le i\ne j\le k} B_{i,j}B_{i,j}\sigma_i^2\sigma_j^2 \\ &\quad + \sum\sum_{1\le i\ne m\le k} B_{i,m}B_{m,i}\sigma_i^2\sigma_m^2 \\ &= \sum_{i=1}^{k} B_{i,i}^2\gamma_i\sigma_i^4 + 3\sum_{i=1}^{k} B_{i,i}^2\sigma_i^4 + \sum\sum_{1\le i\ne j\le k} B_{i,i}B_{j,j}\sigma_i^2\sigma_j^2 + 2\sum\sum_{1\le i\ne j\le k} B_{i,j}B_{i,j}\sigma_i^2\sigma_j^2. \end{aligned} \tag{2.3.4}$$

It is previously shown that

$$\begin{aligned} E(\boldsymbol{b}'\boldsymbol{B}\boldsymbol{b}) &= \sum_{i=1}^{k} B_{i,i}\sigma_i^2, \\ [E(\boldsymbol{b}'\boldsymbol{B}\boldsymbol{b})]^2 &= \sum_{i=1}^{k}\sum_{j=1}^{k} B_{i,i}B_{j,j}\sigma_i^2\sigma_j^2. \end{aligned} \tag{2.3.5}$$

Then, we have

$$\begin{aligned} Var(\boldsymbol{b}'\boldsymbol{B}\boldsymbol{b}) &= E[(\boldsymbol{b}'\boldsymbol{B}\boldsymbol{b})^2] - [E(\boldsymbol{b}'\boldsymbol{B}\boldsymbol{b})]^2 \\ &= \sum_{i=1}^{k} B_{i,i}^2\gamma_i\sigma_i^4 + 3\sum_{i=1}^{k} B_{i,i}^2\sigma_i^4 + \sum\sum_{1\le i\ne j\le k} B_{i,i}B_{j,j}\sigma_i^2\sigma_j^2 \\ &\quad + 2\sum\sum_{1\le i\ne j\le k} B_{i,j}B_{i,j}\sigma_i^2\sigma_j^2 - \left[\sum_{i=1}^{k}\sum_{j=1}^{k} B_{i,i}B_{j,j}\sigma_i^2\sigma_j^2\right] \\ &= \sum_{i=1}^{k} B_{i,i}^2\gamma_i\sigma_i^4 + 2\sum_{i=1}^{k} B_{i,i}^2\sigma_i^4 + 2\sum\sum_{1\le i\ne j\le k} B_{i,j}B_{i,j}\sigma_i^2\sigma_j^2 \\ &= \sum_{i=1}^{k} B_{i,i}^2\gamma_i\sigma_i^4 + 2\sum\sum_{1\le i,j\le k} B_{i,j}B_{i,j}\sigma_i^2\sigma_j^2 \\ &= tr(\tilde{\boldsymbol{B}}\Delta_2\tilde{\boldsymbol{B}}) + 2tr(\Delta_1\boldsymbol{B}\Delta_1\boldsymbol{B}). \end{aligned} \tag{2.3.6}$$

2.4 Covariance

Theorem 2.3. For $\boldsymbol{b} = (b_1, b_2, \ldots, b_k)'$, we assume that b_i's are independent with the mean of 0, variance σ_i^2 and kurtosis γ_i. For two quadratic forms $\boldsymbol{b}'\boldsymbol{B}\boldsymbol{b}$ and $\boldsymbol{b}'\boldsymbol{F}\boldsymbol{b}$, where $\boldsymbol{B}$ and $\boldsymbol{F}$

are both $k \times k$ symmetric matrices, we have

$$Cov(\boldsymbol{b}'\boldsymbol{B}\boldsymbol{b}, \boldsymbol{b}'\boldsymbol{F}\boldsymbol{b}) = tr(\tilde{\boldsymbol{B}}\Delta_2\tilde{\boldsymbol{F}}) + 2tr(\Delta_1\boldsymbol{B}\Delta_1\boldsymbol{F}), \tag{2.4.1}$$

where $\tilde{\boldsymbol{B}}$ is the diagonal matrix with the same diagonal elements as the matrix $\boldsymbol{B}$, $\tilde{\boldsymbol{F}}$ is the diagonal matrix with the same diagonal elements as the matrix $\boldsymbol{F}$, Δ_1 and Δ_2 are given in (2.3.1).

Proof:

For scalar $\boldsymbol{b}'\boldsymbol{B}\boldsymbol{b}$ and $\boldsymbol{b}'\boldsymbol{F}\boldsymbol{b}$, we have

$$Cov(\boldsymbol{b}'\boldsymbol{B}\boldsymbol{b}, \boldsymbol{b}'\boldsymbol{F}\boldsymbol{b}) = E[(\boldsymbol{b}'\boldsymbol{B}\boldsymbol{b})(\boldsymbol{b}'\boldsymbol{F}\boldsymbol{b})] - E(\boldsymbol{b}'\boldsymbol{B}\boldsymbol{b})E(\boldsymbol{b}'\boldsymbol{F}\boldsymbol{b}). \tag{2.4.2}$$

Since $\boldsymbol{b}'\boldsymbol{B}\boldsymbol{b} = \sum_{i=1}^{k}\sum_{j=1}^{k} b_i B_{i,j} b_j$ and $\boldsymbol{b}'\boldsymbol{F}\boldsymbol{b} = \sum_{i=1}^{k}\sum_{j=1}^{k} b_i F_{i,j} b_j$, we can write

$$(\boldsymbol{b}'\boldsymbol{B}\boldsymbol{b})(\boldsymbol{b}'\boldsymbol{F}\boldsymbol{b}) = \sum\sum_{1\le i,j,m,n\le k}\sum\sum B_{i,j}F_{m,n}b_i b_j b_m b_n. \tag{2.4.3}$$

Since b_i's are independent with mean 0 and variance σ_i^2, $E(b_i) = 0, E(b_i^2) = \sigma_i^2, E(b_i^4) = (\gamma_i + 3)\sigma_i^4$, and

$$E(b_i b_j b_n b_m) = \begin{cases} (\gamma_i + 3)\sigma_i^4, \text{ for } i = j = k = l, \\ \sigma_i^2\sigma_j^2, \text{ for } i = j \neq m = n \text{ or } i = m \neq j = n \text{ or } i = n \neq m = j, \\ 0, \text{ otherwise.} \end{cases}$$

Because $\boldsymbol{B}$ and $\boldsymbol{F}$ are both symmetric matrix, therefore, the first term in (2.4.2) is equal to

$$
\begin{aligned}
E[(\boldsymbol{b}'\boldsymbol{B}\boldsymbol{b})(\boldsymbol{b}'\boldsymbol{F}\boldsymbol{b})] &= \sum_{i=1}^{k} B_{i,i}F_{i,i}(\gamma_i+3)\sigma_i^4 + \sum\sum_{1\le i\ne m\le k} B_{i,i}F_{m,m}\sigma_i^2\sigma_m^2 + \sum\sum_{1\le i\ne j\le k} B_{i,j}F_{i,j}\sigma_i^2\sigma_j^2 \\
&\quad + \sum\sum_{1\le i\ne m\le k} B_{i,m}F_{m,i}\sigma_i^2\sigma_m^2 \\
&= \sum_{i=1}^{k} B_{i,i}F_{i,i}\gamma_i\sigma_i^4 + 3\sum_{i=1}^{k} B_{i,i}F_{i,i}\sigma_i^4 + \sum\sum_{1\le i\ne j\le k} B_{i,i}F_{j,j}\sigma_i^2\sigma_j^2 \\
&\quad +2\sum\sum_{1\le i\ne j\le k} B_{i,j}F_{i,j}\sigma_i^2\sigma_j^2.
\end{aligned}
\tag{2.4.4}
$$

It is previously shown that

$$
\begin{aligned}
E(\boldsymbol{b}'\boldsymbol{B}\boldsymbol{b}) &= \sum_{i=1}^{k} B_{i,i}\sigma_i^2, \\
E(\boldsymbol{b}'\boldsymbol{F}\boldsymbol{b}) &= \sum_{i=1}^{k} F_{i,i}\sigma_i^2, \\
[E(\boldsymbol{b}'\boldsymbol{B}\boldsymbol{b})]\,[E(\boldsymbol{b}'\boldsymbol{F}\boldsymbol{b})] &= \sum_{i=1}^{k}\sum_{j=1}^{k} B_{i,i}F_{j,j}\sigma_i^2\sigma_j^2,
\end{aligned}
\tag{2.4.5}
$$

we have

$$
\begin{aligned}
Cov(\boldsymbol{b}'\boldsymbol{B}\boldsymbol{b}, \boldsymbol{b}'\boldsymbol{F}\boldsymbol{b}) &= E[(\boldsymbol{b}'\boldsymbol{B}\boldsymbol{b})(\boldsymbol{b}'\boldsymbol{F}\boldsymbol{b})] - E(\boldsymbol{b}'\boldsymbol{B}\boldsymbol{b})E(\boldsymbol{b}'\boldsymbol{F}\boldsymbol{b}) \\
&= \sum_{i=1}^{k} B_{i,i}F_{i,i}\gamma_i\sigma_i^4 + 3\sum_{i=1}^{k} B_{i,i}F_{i,i}\sigma_i^4 + \sum\sum_{1\le i\ne j\le k} B_{i,i}F_{j,j}\sigma_i^2\sigma_j^2 \\
&\quad +2\sum\sum_{1\le i\ne j\le k} B_{i,j}F_{i,j}\sigma_i^2\sigma_j^2 - \left[\sum_{i=1}^{k}\sum_{j=1}^{k} B_{i,i}F_{j,j}\sigma_i^2\sigma_j^2\right] \\
&= \sum_{i=1}^{k} B_{i,i}F_{i,i}\gamma_i\sigma_i^4 + 2\sum_{i=1}^{k} B_{i,i}F_{i,i}\sigma_i^4 + 2\sum\sum_{1\le i\ne j\le k} B_{i,j}F_{i,j}\sigma_i^2\sigma_j^2 \\
&= \sum_{i=1}^{k} B_{i,i}F_{i,i}\gamma_i\sigma_i^4 + 2\sum\sum_{1\le i,j\le k} B_{i,j}F_{i,j}\sigma_i^2\sigma_j^2 \\
&= tr(\tilde{\boldsymbol{B}}\Delta_2\tilde{\boldsymbol{F}}) + 2tr(\Delta_1\boldsymbol{B}\Delta_1\boldsymbol{F}).
\end{aligned}
\tag{2.4.6}
$$

2.5 Example

Consider an illustrative example to compute the expectation, variance and covariance of a quadratic form.

For a random vector $\boldsymbol{b} = (b_1, b_2, b_3)'$, we assumed that b_i's are indpendent with mean 0, common variance σ^2 and kurtosis γ, we have

$$E(\boldsymbol{b}) = \boldsymbol{0}, Var(\boldsymbol{b}) = \Delta_1 = \sigma^2 \boldsymbol{I}_3, \tag{2.5.1}$$

where $\boldsymbol{I}_3$ is an identity matrix of order 3.

Consider the quadratic form $\boldsymbol{b}'\boldsymbol{B}\boldsymbol{b}$ with $\boldsymbol{B}$ given as

$$\boldsymbol{B} = \begin{bmatrix} B_{1,1} & B_{1,2} & B_{1,3} \\ B_{1,2} & B_{2,2} & B_{2,3} \\ B_{1,3} & B_{2,3} & B_{3,3} \end{bmatrix}. \tag{2.5.2}$$

The expectation of $\boldsymbol{b}'\boldsymbol{B}\boldsymbol{b}$ is

$$\begin{aligned} E(\boldsymbol{b}'\boldsymbol{B}\boldsymbol{b}) &= tr(\boldsymbol{B}\Delta_1) \\ &= \sigma^2(B_{11} + B_{22} + B_{33}). \end{aligned} \tag{2.5.3}$$

The variance of $\boldsymbol{b}'\boldsymbol{B}\boldsymbol{b}$ is

$$\begin{aligned} Var(\boldsymbol{b}'\boldsymbol{B}\boldsymbol{b}) &= tr(\tilde{\boldsymbol{B}}\Delta_2\tilde{\boldsymbol{B}}) + 2tr(\Delta_1\boldsymbol{B}\Delta_1\boldsymbol{B}) \\ &= \sigma^4(B_{11}^2\gamma + 2B_{11}^2 + 4B_{12}^2 + 4B_{13}^2 + B_{22}^2\gamma + 2B_{22}^2 + 4B_{23}^2 + B_{33}^2\gamma \\ &\quad +2B_{33}^2). \end{aligned} \tag{2.5.4}$$

Consider another quadratic form $\boldsymbol{b}'\boldsymbol{F}\boldsymbol{b}$, where $\boldsymbol{F}$ is given as

$$\boldsymbol{F} = \begin{bmatrix} F_{1,1} & F_{1,2} & F_{1,3} \\ F_{1,2} & F_{2,2} & F_{2,3} \\ F_{1,3} & F_{2,3} & F_{3,3} \end{bmatrix}. \tag{2.5.5}$$

The covariance of $\boldsymbol{b}'\boldsymbol{B}\boldsymbol{b}$ and $\boldsymbol{b}'\boldsymbol{F}\boldsymbol{b}$ is

$$\begin{aligned} Cov(\boldsymbol{b}'\boldsymbol{B}\boldsymbol{b}, \boldsymbol{b}'\boldsymbol{F}\boldsymbol{b}) &= 2tr(\boldsymbol{B}\Delta_1\boldsymbol{F}\Delta_1) + tr(\tilde{\boldsymbol{B}}\Delta_2\tilde{\boldsymbol{F}}) \\ &= \sigma^4(B_{11}F_{11}\gamma + 2B_{11}F_{11} + B_{22}F_{22}\gamma + 2B_{22}F_{22} + B_{33}F_{33}\gamma \\ &\quad +2B_{33}F_{33}). \end{aligned} \tag{2.5.6}$$

Chapter 3

The Unbiased Estimation of the Variance Components in the Mixed Effects Model

3.1 Introduction

Statistics is concerned with the partitions and the estimations of the observed variation of data due to different sources, which is an important step in the procedure of drawing statistical inferences about the effects of certain factors on the response variables.

Variation among data can be studied through different classes of linear models, one of which is called the mixed-effects model. The mixed-effects model is widely used in analyzing the correlated data, especially the data with repeated measurements made on the same observational unit or the measurements are made on clusters of related observational units.

In the mixed-effects model, the effects of a factor have two kinds. The first are fixed effects, whose effects on the response variable are attributable to a finite set of levels of a factor. The second kind of effects are random effects, whose effects on the response variable are attributable to a infinite set of levels of a factor, the levels selected in the data are only a random sample of those infinite levels of the factor.

Here is an example of the fixed effects and the random effects. Consider an experiment with balanced incomplete block design to evaluate the effects of two hormones on the duration of the reepithelisation of cornea in rabbits. The effects of cortisone and desoxycorticosterone are compared with the control treatment (saline solution). It is assumed that the two eyes of a rabbit are two independent experimental units forming a block, it is also assumed that the block effects follow certain distribution. The resulting design is arrayed in Table 3.1:

Table 3.1: Example of the Fixed Effects and the Random Effects in Experiment

		Blocks					
		Rabbit1		Rabbit2		Rabbit3	
		1	2	1	2	1	2
Treatments	cortisone						
	desoxycorticosterone						
	control						

In this experiment, the effects for the treatment factor are the fixed effects, because the researcher is only interested in the effects of the two hormones by estimating the treatment effects. The effects due to the rabbit blocks are considered as the random effects if the rabbits in the experiments are a random sample from the pool of rabbits with assumed distributions. The effects of the random errors are also considered as the random effects with certain distributional assumptions in the experiment.

Since distributional assumptions are made to those random effects, the statistical interests in them lie in estimating the variances of those effects. Those variances are known as the variance components, their sum is the variance of the variable being observed[26].

A large literature are available for the estimation methods of variance components in the mixed-effects model, including the Method of Moment (MoM), the maximum likelihood (ML) method and the restricted maximum likelihood (REML) method. All of those methods have some nice properties under some regularity conditions. However, when the data are finite, such as the data in DOE, ML estimators are generally biased. Besides, ML and REML require distributional assumptions and do not have closed-form expressions for the estimators of the variance components, so the finite-sample properties of those estimators cannot be studied.

In this section, we will find the class of unbiased quadratic estimators for the variance components with explicit forms under the general mixed-effects model.

3.2 The Mixed Effects Model

Consider the linear mixed effect model

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{U}_1\boldsymbol{b}_1 + \cdots + \boldsymbol{U}_{s-1}\boldsymbol{b}_{s-1} + \boldsymbol{e}, \tag{3.2.1}$$

where $\boldsymbol{y}$ is a $n \times 1$ vector of observations, $\boldsymbol{X}$ is a $n \times p$ known matrix, $\boldsymbol{\beta}$ is a $p \times 1$ vector of fixed effect coefficients. For $i = 1, 2, \ldots, s-1$, $\boldsymbol{b}_i$ is a $q_i \times 1$ vector of random effects. $\boldsymbol{U}_i$ is a $n \times q_i$ known incidence matrices, $\boldsymbol{e}$ is a $n \times 1$ vector of random errors. It is assumed the q_i components in $\boldsymbol{b}_i$ are independent with mean 0, variance σ_i^2 and kurtosis γ_i, the n components in $\boldsymbol{e}$ are independent with mean 0, variance σ_e^2 and kurtosis γ_e, and the vectors

$\boldsymbol{b}_i$'s and $\boldsymbol{e}$ are independent, then we have $Var(\boldsymbol{y}) = \sum_{i=1}^{s-1} \sigma_i^2 \boldsymbol{U}_i \boldsymbol{U}_i' + \sigma_e^2 I_n$. Here, $\sigma_1^2, \sigma_2^2, \ldots, \sigma_e^2$ are the unknown variance components in the mixed-effects model (3.2.1).

Define

$$\begin{cases} q_s = n, \\ \boldsymbol{U}_s = \boldsymbol{I}_n, \\ \boldsymbol{b}_s = \boldsymbol{e}, \end{cases} \tag{3.2.2}$$

the mixed-effects model in (3.2.1) can also be expressed as

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{U}_1 \boldsymbol{b}_1 + \cdots + \boldsymbol{U}_{s-1} \boldsymbol{b}_{s-1} + \boldsymbol{U}_s \boldsymbol{b}_s. \tag{3.2.3}$$

Define

$$
\begin{cases}
\boldsymbol{U} = [\boldsymbol{U}_1, \boldsymbol{U}_2, \ldots, \boldsymbol{U}_s], \\
\boldsymbol{b}' = [\boldsymbol{b}_1', \boldsymbol{b}_2', \ldots, \boldsymbol{b}_s'], \\
Var(\boldsymbol{b}) = \begin{bmatrix} \sigma_1^2 \boldsymbol{I}_{q_1} & 0 & \ldots & 0 \\ 0 & \sigma_2^2 \boldsymbol{I}_{q_2} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \sigma_s^2 \boldsymbol{I}_n \end{bmatrix}, \\
\boldsymbol{V}_i = \boldsymbol{U}_i \boldsymbol{U}_i', i = 1, 2, \ldots, s-1, \\
\boldsymbol{V}_s = \boldsymbol{I}_n, \\
\boldsymbol{V} = Var(\boldsymbol{y}),
\end{cases}
\tag{3.2.4}
$$

the mixed-effects model in (3.2.3) can be expressed as

$$
\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{U}\boldsymbol{b}. \tag{3.2.5}
$$

The variance of $\boldsymbol{y}$ is

$$
\begin{aligned}
\boldsymbol{V} &= \boldsymbol{U} Var(\boldsymbol{b}) \boldsymbol{U}' \\
&= \sum_{i=1}^{s-1} \sigma_i^2 \boldsymbol{V}_i + \sigma_s^2 \boldsymbol{I}_n
\end{aligned}
\tag{3.2.6}
$$

3.3 The Unbiased Quadratic Estimators of the Variance Components

We have introduced the general mixed-effects model and the variance components in the general mixed-effects model. In this section, we will present the unbiased quadratic estimators for the variance components in the general mixed-effects model.

Recall that in the general mixed-effects model (3.2.1), $\sigma_1^2, \sigma_2^2, \ldots, \sigma_e^2$ are s unknown variance components. Consider using a quadratic function $\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}$ to estimate a linear function of the variance components, $\boldsymbol{p}'\boldsymbol{\sigma}^2 = \sum_{i=1}^{s} p_i\sigma_i^2$, where p_i's are known constants and $\sigma_s^2 = \sigma_e^2$, The matrix $\boldsymbol{A}$ in $\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}$ is chosen according to the following criteria:

- Symmetry: $\boldsymbol{A}$ should be a $n \times n$ symmetric matrix, i.e. $\boldsymbol{A} = \boldsymbol{A}'$.
- Unbiasedness: $\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}$ is an unbiased estimator for $\boldsymbol{p}'\boldsymbol{\sigma}^2$, i.e. $E(\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}) = \boldsymbol{p}'\boldsymbol{\sigma}^2$.

For the general mixed-effects model in (3.2.1), and based on the expectation in (2.1.2), we have

$$\begin{aligned} E(\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}) &= E[(\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{U}\boldsymbol{b})'\boldsymbol{A}(\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{U}\boldsymbol{b})] \\ &= E(\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{A}\boldsymbol{U}\boldsymbol{b} + \boldsymbol{b}'\boldsymbol{U}'\boldsymbol{A}\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{b}'\boldsymbol{U}'\boldsymbol{A}\boldsymbol{U}\boldsymbol{b}) \\ &= E(\boldsymbol{b}'\boldsymbol{U}'\boldsymbol{A}\boldsymbol{U}\boldsymbol{b}) + E(\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}\boldsymbol{\beta}) \qquad (3.3.1) \\ &= tr(\boldsymbol{A}\boldsymbol{V}) + (\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}\boldsymbol{\beta}) \\ &= \sum_{i=1}^{s} tr(\boldsymbol{A}\boldsymbol{U}_i\boldsymbol{U}_i')\sigma_i^2 + (\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}\boldsymbol{\beta}). \end{aligned}$$

If requiring $\boldsymbol{A}\boldsymbol{X} = \boldsymbol{0}$ and $tr(\boldsymbol{A}\boldsymbol{U}_i\boldsymbol{U}_i') = p_i$ for $i = 1, 2, \ldots, s$, then $E(\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}) = \boldsymbol{p}'\boldsymbol{\sigma}^2$, i.e. the quadratic estimator $\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}$ is an unbiased estimator for $\boldsymbol{p}'\boldsymbol{\sigma}^2$.

Consider a class of matrices $\mathscr{A}$ satisfying the conditions for unbiasedness. The class of matrices $\mathscr{A}$ is defined as

$$\mathscr{A} = \left\{ \boldsymbol{A} | \boldsymbol{A} = \boldsymbol{A}', \boldsymbol{AX} = 0, tr(\boldsymbol{AV}_i) = p_i, i = 1, 2, \ldots, s \right\}. \tag{3.3.2}$$

For any $\boldsymbol{A} \in \mathscr{A}$, $\boldsymbol{y}'\boldsymbol{Ay}$ is an unbiased estimator of $\boldsymbol{p}'\boldsymbol{\sigma}^2$, see 5 ().

3.4 The Variance of the Quadratic Estimators for the Variance Components

For the matrix $\boldsymbol{U}$ in (3.2.4), and any $\boldsymbol{A} \in \mathscr{A}$ in (3.3.2), define

$$\boldsymbol{B} = \boldsymbol{U}'\boldsymbol{AU}. \tag{3.4.1}$$

Let $\tilde{\boldsymbol{B}}$ be a diagonal matrix with the same diagonal elements of the matrix $\boldsymbol{B}$ in (3.4.1). Define

$$\Delta_1 = \begin{bmatrix} \sigma_1^2 \boldsymbol{I}_{q_1} & \boldsymbol{0} & \ldots & \boldsymbol{0} \\ \boldsymbol{0} & \sigma_2^2 \boldsymbol{I}_{q_2} & \ldots & \boldsymbol{0} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \ldots & \sigma_s^2 \boldsymbol{I}_{q_s} \end{bmatrix}, \Delta_2 = \begin{bmatrix} \gamma_1 \sigma_1^4 \boldsymbol{I}_{q_1} & \boldsymbol{0} & \ldots & \boldsymbol{0} \\ \boldsymbol{0} & \gamma_2 \sigma_2^4 \boldsymbol{I}_{q_2} & \ldots & \boldsymbol{0} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \ldots & \gamma_s \sigma_s^4 \boldsymbol{I}_{q_s} \end{bmatrix}. \tag{3.4.2}$$

For any $\boldsymbol{A} \in \mathscr{A}$ in (3.3.2), the variance of $\boldsymbol{y}'\boldsymbol{Ay}$ is

$$Var(\boldsymbol{y}'\boldsymbol{Ay}) = tr(\tilde{\boldsymbol{B}}\Delta_2\tilde{\boldsymbol{B}}) + 2tr(\Delta_1\boldsymbol{B}\Delta_1\boldsymbol{B}). \tag{3.4.3}$$

3.5 Example of the Unbiased Quadratic Estimators and Their Variance

Consider an illustrative example presented in Table 3.2. The data are obtained from an experiment with split-plot design, the design evaluates the effects of two treatment factors, which are curing temperature and coating material. Because it is expensive to replicate the experiment on each single bar for each temperature and coating material, the two treatment factors are applied to two different experimental units. Among 24 bars, six bars are randomly assigned to each of the four coating materials, the experimental unit for coating materials is the individual bar. Each curing temperature is applied to a group of four bars and is replicated twice, the experimental unit for curing temperature is a group of four bars (Heat). Because the design has two different experimental units for the two treatment factors, it also has two random errors corresponding to the two experimental units, which are the random effects of the group of four bars (Heat) and radom error of each individual bar.

Table 3.2: Split Plot Design

Temperature	Heat	Coating1	Coating2	Coating3	Coating4
360F	1	67	73	83	89
	6	33	8	46	54
370F	2	65	91	87	86
	5	140	142	121	150
380F	3	155	127	147	212
	4	108	100	90	153

Consider the following mixed-effects model,

$$y_{ijk} = \mu + T_i + e_{1ij} + C_k + (TC)_{ik} + e_{2ijk}, i = 1, 2, 3, j = 1, 2, k = 1, 2, 3, 4, \qquad (3.5.1)$$

where μ is the grand mean, T_i is the fixed effect of the ith temperature, C_k is the fixed effect of the kth coating material, $(TC)_{ik}$ is the interaction between the ith temperature and the kth coating material, e_{1ij} is the random effect of the jth replication (heat) in the ith temperature, e_{2ijk} is the random error of individual bar in the kth coating materials in the jth replication in the ith temperature. It is assumed that $e'_{1ij}s$ are independent with mean 0, common variance σ_1^2 and kurtosis γ_1, $e'_{2ijk}s$ are independent with mean 0, common variance σ_2^2 and kurtosis γ_2, $e'_{1ij}s$ and $e'_{2ijk}s$ are independent. Note that the functional form of the distribution is not needed for the random effects.

Define

$$\begin{cases} \boldsymbol{j}_b = \text{ an column vector of order b with all entries equal to 1 }, \\ \mathbf{0}_b = \text{ an column vector of order b with all entries equal to 0 }, \\ \boldsymbol{I}_b = \text{ an identity matrix of order b }, \\ \mathbf{0}_{b,b} = \text{ a } b \times b \text{ matrix with all entries equal to 0}, \\ \boldsymbol{J}_b = \text{ a } b \times b \text{ matrix with all entries equal to 1}. \end{cases} \qquad (3.5.2)$$

The model can also be expressed in the matrix notation as

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{U}_1\boldsymbol{b} + \boldsymbol{e}, \qquad (3.5.3)$$

where

$$\boldsymbol{y} = \begin{bmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \end{bmatrix} \text{ with } \boldsymbol{y}_1 = \begin{bmatrix} y_{111} \\ y_{112} \\ y_{113} \\ y_{114} \\ y_{211} \\ y_{212} \\ y_{213} \\ y_{214} \\ y_{311} \\ y_{312} \\ y_{313} \\ y_{314} \end{bmatrix} = \begin{bmatrix} 67 \\ 73 \\ 83 \\ 89 \\ 65 \\ 91 \\ 87 \\ 86 \\ 155 \\ 127 \\ 147 \\ 212 \end{bmatrix}, \boldsymbol{y}_2 = \begin{bmatrix} y_{121} \\ y_{122} \\ y_{123} \\ y_{124} \\ y_{221} \\ y_{222} \\ y_{223} \\ y_{224} \\ y_{321} \\ y_{322} \\ y_{323} \\ y_{324} \end{bmatrix} = \begin{bmatrix} 33 \\ 8 \\ 46 \\ 54 \\ 140 \\ 142 \\ 121 \\ 150 \\ 108 \\ 100 \\ 90 \\ 153 \end{bmatrix}, \boldsymbol{\beta} = \begin{bmatrix} \mu \\ T_1 \\ T_2 \\ C_1 \\ C_2 \\ C_3 \\ T_1C_1 \\ T_1C_2 \\ T_1C_3 \\ T_2C_1 \\ T_2C_2 \\ T_2C_3 \end{bmatrix},$$

$$\boldsymbol{X} = \begin{bmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_1 \end{bmatrix} \text{ with } \boldsymbol{X}_1 = \begin{bmatrix} \boldsymbol{j}_3 & \boldsymbol{j}_3 & \mathbf{0}_3 & \boldsymbol{I}_3 & \boldsymbol{I}_3 & \mathbf{0}_{3,3} \\ 1 & 1 & 0 & \mathbf{0}_3' & \mathbf{0}_3' & \mathbf{0}_3' \\ \boldsymbol{j}_3 & \mathbf{0}_3 & \boldsymbol{j}_3 & \boldsymbol{I}_3 & \mathbf{0}_{3,3} & \boldsymbol{I}_3 \\ 1 & 0 & 1 & \mathbf{0}_3' & \mathbf{0}_3' & \mathbf{0}_3' \\ \boldsymbol{j}_3 & \mathbf{0}_3 & \mathbf{0}_3 & \boldsymbol{I}_3 & \mathbf{0}_{3,3} & \mathbf{0}_{3,3} \\ 1 & 0 & 0 & \mathbf{0}_3' & \mathbf{0}_3' & \mathbf{0}_3' \end{bmatrix},$$

$$\boldsymbol{U}_1 = \begin{bmatrix} \boldsymbol{Z}_1 & \mathbf{0} \\ \mathbf{0} & \boldsymbol{Z}_1 \end{bmatrix} \text{ with } \boldsymbol{Z}_1 = \begin{bmatrix} \boldsymbol{j}_4 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{j}_4 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{j}_4 \end{bmatrix}, \boldsymbol{b} = \begin{bmatrix} b_{11} \\ b_{21} \\ b_{31} \\ b_{12} \\ b_{22} \\ b_{32} \end{bmatrix}.$$

Under the model (3.5.3), we have

$$\begin{aligned} E(\boldsymbol{y}) &= \boldsymbol{X\beta}, \\ Var(\boldsymbol{y}) &= \boldsymbol{V} = \sigma_1^2 \boldsymbol{U}_1 \boldsymbol{U}_1' + \sigma_2^2 \boldsymbol{I}_{24}, \end{aligned} \tag{3.5.4}$$

where $\boldsymbol{U}_1\boldsymbol{U}_1' = \begin{bmatrix} \boldsymbol{Z}_1\boldsymbol{Z}_1' & \mathbf{0} \\ \mathbf{0} & \boldsymbol{Z}_1\boldsymbol{Z}_1' \end{bmatrix} = \begin{bmatrix} \boldsymbol{J}_4 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{J}_4 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{J}_4 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{J}_4 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{J}_4 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{J}_4 \end{bmatrix}.$

For the model (3.5.3), σ_1^2 and σ_2^2 in (3.5.4) are the unknown variance components to be estimated.

3.5.1 The Unbiased Quadratic Estimator for σ_1^2

Consider $\boldsymbol{y'Ay}$ as an estimator of σ_1^2, where $\boldsymbol{A}$ is a 24×24 symmetric matrix defined as

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}'_{12} & \boldsymbol{A}_{22} \end{bmatrix}. \tag{3.5.1.1}$$

$\boldsymbol{A}_{11}$, $\boldsymbol{A}_{12}$ and $\boldsymbol{A}_{22}$ are matrices of order 12×12, $\boldsymbol{A}_{11}$ and $\boldsymbol{A}_{22}$ are symmetric matrices. Under the model in (3.5.3), if $\boldsymbol{A}$ satisfy the following conditions for unbiasedness,

$$\mathscr{A} = \left\{ \boldsymbol{A} : \boldsymbol{A} = \boldsymbol{A}', \boldsymbol{AX} = \boldsymbol{0}, tr(\boldsymbol{A}\boldsymbol{U}_1\boldsymbol{U}_1') = 1, tr(\boldsymbol{A}) = 0 \right\}, \tag{3.5.1.2}$$

for any $\boldsymbol{A} \in \mathscr{A}$ in (3.5.1.2), $\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}$ will be an unbiased estimator of σ_1^2.

For the conditions $\boldsymbol{AX} = \boldsymbol{0}$ in (3.5.1.2), it can be shown that

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_{11} & -\boldsymbol{A}_{11} \\ -\boldsymbol{A}_{11} & \boldsymbol{A}_{11} \end{bmatrix}, \tag{3.5.1.3}$$

where $\boldsymbol{A}_{11}$ is any symmetric matrix of order 12×12.

For the condition $tr(\boldsymbol{A}) = 0$ and $tr(\boldsymbol{A}\boldsymbol{U}_1\boldsymbol{U}_1') = 1$ in (3.5.1.2), we have

$$\begin{cases} tr(\boldsymbol{A}_{11}) = 0, \\ tr(\boldsymbol{A}_{11}\boldsymbol{U}_1\boldsymbol{U}_1') = \frac{1}{2}, \end{cases} \tag{3.5.1.4}$$

where $\boldsymbol{U}_1\boldsymbol{U}_1' = \begin{bmatrix} \boldsymbol{J}_4 & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{J}_4 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{J}_4 \end{bmatrix}$. In other word, as long as we find a matrix $\boldsymbol{A}_{11}$ that is a symmetric matrix with $tr(\boldsymbol{A}_{11}) = 0$ and $tr(\boldsymbol{A}_{11}\boldsymbol{U}_1\boldsymbol{U}_1') = \frac{1}{2}$, the matrix $\boldsymbol{A}$ will satisfy the unbiasedness conditions in (3.5.1.2) and give an unbiased quadratic estimator $\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}$ for σ_1^2.

3.5.2 The Unbiased Quadratic Estimator for σ_2^2

Let $\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}$ be an estimator of σ_2^2, where $\boldsymbol{A}$ is a 24×24 symmetric matrix defined as

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}'_{12} & \boldsymbol{A}_{22} \end{bmatrix}. \tag{3.5.2.1}$$

$\boldsymbol{A}_{11}$, $\boldsymbol{A}_{12}$ and $\boldsymbol{A}_{22}$ are matrices of order 12×12, $\boldsymbol{A}_{11}$ and $\boldsymbol{A}_{22}$ are symmetric matrices. Under the model in (3.5.3), if $\boldsymbol{A}$ satisfy the following conditions for unbiasedness

$$\mathscr{A} = \left\{\boldsymbol{A} : \boldsymbol{A} = \boldsymbol{A}', \boldsymbol{A}\boldsymbol{X} = \boldsymbol{0}, tr(\boldsymbol{A}\boldsymbol{U}_1\boldsymbol{U}_1') = 0, tr(\boldsymbol{A}) = 1\right\}, \tag{3.5.2.2}$$

for any $\boldsymbol{A} \in \mathscr{A}$ in (3.5.2.2), $\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}$ will be an unbiased estimator of σ_2^2.

To satisfy the condition $\boldsymbol{A}\boldsymbol{X} = \boldsymbol{0}$ in (3.5.2.2), it can be shown that

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_{11} & -\boldsymbol{A}_{11} \\ -\boldsymbol{A}_{11} & \boldsymbol{A}_{11} \end{bmatrix} \tag{3.5.2.3}$$

where $\boldsymbol{A}_{11}$ is any symmetric matrix of order 12×12.

For the condition $tr(\boldsymbol{A}) = 1$ and $tr(\boldsymbol{A}\boldsymbol{U}_1\boldsymbol{U}_1') = 0$ in (3.5.2.2), we have

$$\begin{cases} tr(\boldsymbol{A}_{11}) = \frac{1}{2}, \\ tr(\boldsymbol{A}_{11}\boldsymbol{U}_1\boldsymbol{U}_1') = 0. \end{cases} \tag{3.5.2.4}$$

As long as we find a matrix $\boldsymbol{A}_{11}$ that is symmetric with $tr(\boldsymbol{A}_{11}) = \frac{1}{2}$ and $tr(\boldsymbol{A}_{11}\boldsymbol{U}_1\boldsymbol{U}_1') = 0$, the matrix $\boldsymbol{A}$ will satisfy the conditions in (3.5.2.2) and give an unbiased quadratic estimator $\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}$ for σ_2^2.

Chapter 4

The Near Optimum Estimation of the Variance Components in the Mixed Effects Model in Experimental Design

4.1 Introduction

The mixed-effects models are frequently used in experimental design. When the design is balanced, the Optimum Estimator (OPE) exists and are often used to estimate the variance components in the mixed-effects model. However, when the design is unbalanced, OPE does not exist, therefore, the near optimum estimation (NOPE) method is proposed to estimate the variance components.

When the experiment involves replications, it is difficult to find out the NOPE for the full data, therefore, the average optimum estimation (AOPE) method is proposed based on the NOPE and the replications in the experiment.

In this section, the method of NOPE will be introduced first in the unbalanced experimental design with one replication, following which the method of AOPE will be introduced under the unbalanced experimental design with replications.

4.2 Unbalanced Incomplete Block Design (UIBD)

Consider an UIBD with 1 replication in Table 4.1

Table 4.1: Unbalanced Incomplete Block Design with One Replication

	Block1	Block2	Block3
TreatmentA	A	A	A
TreatmentB	B		B
TreatmentC		C	

The data are

Table 4.2: Data for Unbalanced Incomplete Block Design with One Replication

	Block1	Block2	Block3
TreatmentA	y_{11}	y_{21}	y_{31}
TreatmentB	y_{12}		y_{32}
TreatmentC		y_{23}	

We fit the following mixed-effects model to the data

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{U}_1\boldsymbol{b}_1 + \boldsymbol{e}, \tag{4.2.1}$$

where

$$\boldsymbol{y}=\begin{bmatrix} y_{11}\\ y_{12}\\ y_{21}\\ y_{23}\\ y_{31}\\ y_{32}\end{bmatrix},\boldsymbol{X}=\begin{bmatrix}1&0&0\\0&1&0\\1&0&0\\0&0&1\\1&0&0\\0&1&0\end{bmatrix},\boldsymbol{\beta}=\begin{bmatrix}\beta_1\\ \beta_2\\ \beta_3\end{bmatrix},\boldsymbol{U}_1=\begin{bmatrix}1&0&0\\1&0&0\\0&1&0\\0&1&0\\0&0&1\\0&0&1\end{bmatrix},\boldsymbol{b}=\begin{bmatrix}b_1\\ b_2\\ b_3\end{bmatrix}.$$

Let $\boldsymbol{b}'=(\boldsymbol{b}_1',\boldsymbol{e}'),\boldsymbol{U}=(\boldsymbol{U}_1,\boldsymbol{I}),\boldsymbol{V}_1=\boldsymbol{U}_1\boldsymbol{U}_1',\boldsymbol{V}_2=\boldsymbol{I}_6$, then it is assumed that

$$\begin{cases} E(\boldsymbol{b}_1)=\boldsymbol{0},E(\boldsymbol{e})=\boldsymbol{0},\\ Var(\boldsymbol{b}_1)=\sigma_1^2\boldsymbol{I}_3,Var(\boldsymbol{e})=\sigma_2^2\boldsymbol{I}_6,\\ Kurtosis(\boldsymbol{b}_1)=\gamma_1\boldsymbol{1}_3,Kurtosis(\boldsymbol{e})=\gamma_2\boldsymbol{1}_6,\\ Var(\boldsymbol{b})=Var\left(\begin{bmatrix}\boldsymbol{b}_1\\ \boldsymbol{e}\end{bmatrix}\right)=\Delta_1=\begin{bmatrix}\sigma_1^2\boldsymbol{I}_3 & \boldsymbol{0}\\ \boldsymbol{0} & \sigma_2^2\boldsymbol{I}_6\end{bmatrix},\\ E(\boldsymbol{y})=\boldsymbol{X\beta},\\ Var(\boldsymbol{y})=\sigma_1^2\boldsymbol{U}_1\boldsymbol{U}_1'+\sigma_2^2\boldsymbol{I}_6=\sigma_1^2\boldsymbol{V}_1+\sigma_2^2\boldsymbol{V}_2.\end{cases} \tag{4.2.2}$$

Note that the functional forms of the random effects distributions are not needed. σ_1^2 and σ_2^2 are the unknown variance components to be estimated.

4.3 Near Optimum Estimation

4.3.1 The Unbiased Estimators for σ_1^2 and σ_2^2

Consider using a quadratic estimator $\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}$ to estimate the variance components.

For the design matrices $\boldsymbol{X}$ and $\boldsymbol{U}_1$ in (4.2.1), a class of matrices $\mathscr{A}$ satisfying the conditions for unbiasedness is defined as

$$\mathscr{A} = \left\{ \boldsymbol{A} | \boldsymbol{A} = \boldsymbol{A}', \boldsymbol{A}\boldsymbol{X} = \boldsymbol{0}, tr(\boldsymbol{A}\boldsymbol{V}_i) = p_i, i = 1, 2 \right\}. \quad (4.3.1.1)$$

For any $\boldsymbol{A} \in \mathscr{A}$ with $p_1 = 1, p_2 = 0$, the quadratic estimator $\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}$ is an unbiased estimator of σ_1^2. For any $\boldsymbol{A} \in \mathscr{A}$ with $p_1 = 0, p_2 = 1$, the quadratic estimator $\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}$ is an unbiased estimator of σ_2^2.

Consider another class of matrices $\mathscr{C}$ defined as

$$\mathscr{C} = \left\{ \boldsymbol{C} | \boldsymbol{C} = \boldsymbol{C}', \boldsymbol{C}\boldsymbol{X} = \boldsymbol{0}, tr(\boldsymbol{C}\boldsymbol{V}_i) = 0, i = 1, 2 \right\}. \quad (4.3.1.2)$$

For any $\boldsymbol{C} \in \mathscr{C}$, the quadratic estimator $\boldsymbol{y}'\boldsymbol{C}\boldsymbol{y}$ is an unbiased estimator of 0.

For any $\boldsymbol{A} \in \mathscr{A}$, if $\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}$ satisfies

$$Cov(\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}, \boldsymbol{y}'\boldsymbol{C}\boldsymbol{y}) = 0, \text{ for all } \boldsymbol{C} \in \mathscr{C} \text{ and } \sigma_i^2 > 0, i = 1, 2, \quad (4.3.1.3)$$

$\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}$ will be the uniformly minimum variance quadratic unbiased estimator (UMVQUE/OPE) for $\sigma_i^2, i = 1, 2$, see 21 ().

However, the design in Table 4.1 is unbalanced, so the OPEs do not exist. Therefore, we propose NOPE method to find the estimates that are near optimum for the variance components being estimated.

4.3.2 The Near Optimum Estimation of σ_1^2 and σ_2^2

Under the mixed-effects model in (4.2.1), any $\boldsymbol{A} \in \mathscr{A}$ for $p_1 = 1, p_2 = 0$ in (4.3.1.1) satisfy

$$\mathscr{A} = \left\{\boldsymbol{A}|\boldsymbol{A} = \boldsymbol{A}', \boldsymbol{AX} = \boldsymbol{0}, tr(\boldsymbol{AV}_1) = 1, tr(\boldsymbol{AV}_2) = 0\right\},$$

and can be expressed as

$$\boldsymbol{A} = \begin{bmatrix} a_1 & a_2 & -a_1 - a_3 - a_4 & 0 & a_3 + a_4 & -a_2 \\ a_2 & a_3 & 1/2 - 2a_2 & 0 & a_2 - 1/2 & -a_3 \\ -a_1 - a_3 - a_4 & 1/2 - 2a_2 & a_4 & 0 & a_1 + a_3 & 2a_2 - 1/2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ a_3 + a_4 & a_2 - 1/2 & a_1 + a_3 & 0 & -a_1 - a_3 - a_4 & 1/2 - a_2 \\ -a_2 & a_3 & 2a_2 - 1/2 & 0 & 1/2 - a_2 & a_3 \end{bmatrix}, \tag{4.3.2.1}$$

where a_1, a_2, a_3, a_4 are any reals. $\boldsymbol{y}'\boldsymbol{Ay}$ will be an unbiased estimator for σ_1^2.

Any $\boldsymbol{C} \in \mathscr{C}$ defined in (4.3.1.2) can be written as

$$
\boldsymbol{C} = \begin{bmatrix} c_1 & c_2 & -c_1-c_3-c_4 & 0 & c_3+c_4 & -c_2 \\ c_2 & c_3 & -2c_2 & 0 & c_2 & -c_3 \\ -c_1-c_3-c_4 & -2c_2 & c_4 & 0 & c_1+c_3 & 2c_2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ c_3+c_4 & c_2 & c_1+c_3 & 0 & -c_1-c_3-c_4 & -c_2 \\ -c_2 & c_3 & 2c_2 & 0 & -c_2 & c_3 \end{bmatrix}, \tag{4.3.2.2}
$$

where c_1, c_2, c_3, c_4 are any reals. $\boldsymbol{y}'\boldsymbol{C}\boldsymbol{y}$ will be an unbiased estimator for 0.

Define

$$
\boldsymbol{B} = \boldsymbol{U}\boldsymbol{A}\boldsymbol{U}', \tilde{\boldsymbol{B}} = diag(\boldsymbol{B}), \boldsymbol{F} = \boldsymbol{U}\boldsymbol{F}\boldsymbol{U}', \tilde{\boldsymbol{F}} = diag(\boldsymbol{F}),
$$

$$
\Delta_1 = \begin{bmatrix} \sigma_1^2 \boldsymbol{I}_3 & \boldsymbol{0} \\ \boldsymbol{0} & \sigma_2^2 \boldsymbol{I}_6 \end{bmatrix}, \Delta_2 = \begin{bmatrix} \gamma_1 \sigma_1^4 \boldsymbol{I}_3 & \boldsymbol{0} \\ \boldsymbol{0} & \gamma_2 \sigma_2^4 \boldsymbol{I}_6 \end{bmatrix}.
$$

According to Theorem 2.4, the general form of $Cov(\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}, \boldsymbol{y}'\boldsymbol{C}\boldsymbol{y})$ under the mixed-effects model in (4.2.1) is

$$
Cov(\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}, \boldsymbol{y}'\boldsymbol{C}\boldsymbol{y}) = tr(\tilde{\boldsymbol{B}}\Delta_2\tilde{\boldsymbol{F}}) + 2tr(\Delta_1\boldsymbol{B}\Delta_1\boldsymbol{F}). \tag{4.3.2.3}
$$

For $\boldsymbol{A}$ in (4.3.2.1) and $\boldsymbol{C}$ in (4.3.2.2), the $Cov(\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}, \boldsymbol{y}'\boldsymbol{C}\boldsymbol{y})$ in (4.3.2.3) can be expressed as

$$
Cov(\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}, \boldsymbol{y}'\boldsymbol{C}\boldsymbol{y}) = (\boldsymbol{Q}\boldsymbol{\eta} - \boldsymbol{g})'\boldsymbol{\delta}, \tag{4.3.2.4}
$$

where

$$\boldsymbol{g} = \begin{bmatrix} g_1 \\ g_2 \\ g_3 \\ g_4 \end{bmatrix} = \begin{bmatrix} \gamma_1\sigma_1^4 + 6\sigma_1^2\left(\sigma_1^2+\sigma_2^2\right) \\ 2\gamma_1\sigma_1^4 + 12\left(\sigma_1^2+\sigma_2^2\right)^2 \\ \gamma_2\sigma_1^4 + 6\sigma_1^2\sigma_2^2 + 6\sigma_1^4 \\ \gamma_2\sigma_1^4 + 6\sigma_1^2\sigma_2^2 + 6\sigma_1^4 \end{bmatrix},$$

$$\boldsymbol{Q} = \begin{bmatrix} 2g_2+6g_3-8g_4 & 4g_1 & 2g_2+6g_3-8g_4 & g_2+3g_3-4g_4 \\ 4g_1 & 8g_2 & 4g_1 & 2g_1 \\ 2g_2+6g_3-8g_4 & 4g_1 & g_2+g_3 & g_2+g_4 \\ g_2+3g_3-4g_4 & 2g_1 & g_2+g_4 & 2g_2+6g_3-8g_4 \end{bmatrix},$$

$$\boldsymbol{\eta} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix}, \boldsymbol{\delta} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix}.$$

If there exists an $\boldsymbol{\eta}$ that makes $\boldsymbol{Q}\boldsymbol{\eta} - \boldsymbol{g}$ to be a vector close to a null vector, then regardless of $\boldsymbol{\delta}$, we declare that $Cov(\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}, \boldsymbol{y}'\boldsymbol{C}\boldsymbol{y})$ will be close to 0, and the $\boldsymbol{\eta}$ will yield an $\boldsymbol{A}$ matrix resulting in a near optimum estimator for σ_1^2.

Consider $\boldsymbol{\eta} = \begin{bmatrix} 0 \\ 1/4 \\ 0 \\ 0 \end{bmatrix}$, we have

$$\begin{aligned}
\boldsymbol{\phi} &= \boldsymbol{Q\eta} - \boldsymbol{g} \\
&= \begin{bmatrix} 0 \\ 0 \\ 0 \\ -g_1/2 \end{bmatrix}, \qquad (4.3.2.5) \\
Cov(\boldsymbol{y'Ay}, \boldsymbol{y'Cy}) &= -\sigma_1^2 \left(3\sigma_1^2 + 3\sigma_2^2 + \tfrac{1}{2}\gamma_1\sigma_1^2\right) c_4.
\end{aligned}$$

We declare that $\boldsymbol{\phi}$ will be close to a null vector if the value g_1 is small, then the value of the covariance will be close to 0.

The corresponding NOPE under the model in (4.2.2) for σ_1^2 is

$$\hat{\sigma}_1^2 = \tfrac{1}{2}(y_{11} - y_{31})(y_{12} - y_{32}) \qquad (4.3.2.6)$$

The procedures of finding the $\boldsymbol{A}$ in $\boldsymbol{y'Ay}$ for σ_2^2 are similar as the procedure of finding the $\boldsymbol{A}$ in $\boldsymbol{y'Ay}$ for σ_1^2.

Under the mixed-effects model in (4.2.1), any $\boldsymbol{A} \in \mathscr{A}$ for $p_1 = 0, p_2 = 1$ in (4.3.1.1) satisfy

$$\mathscr{A} = \left\{\boldsymbol{A} | \boldsymbol{A} = \boldsymbol{A}', \boldsymbol{AX} = \boldsymbol{0}, tr(\boldsymbol{AV}_1) = 0, tr(\boldsymbol{AV}_2) = 1\right\},$$

and can be written as

$$
\boldsymbol{A} = \begin{bmatrix} a_1 & a_3 & a_2 - a_1 & 0 & a_2 & -a_3 \\ -a_3 & 1/2 - a_1 + a_2 - a_4 & -1/2 - 2a_3 & 0 & a_3 + 1/2 & -1/2 + a_1 - a_2 + a_4 \\ a_2 - a_1 & -1/2 - 2a_3 & a_1 - 2a_2 + a_4 & 0 & a_2 - a_4 & 1/2 + 2a_3 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -a_2 & 1/2 + a_3 & a_2 - a_4 & 0 & a_4 & -1/2 - a_3 \\ -a_3 & -1/2 + a_1 - a_2 + a_4 & 1/2 + 2a_3 & 0 & -1/2 - a_3 & 1/2 - a_1 + a_2 - a_4 \end{bmatrix}, \tag{4.3.2.7}
$$

where a_1, a_2, a_3, a_4 are any reals. $\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}$ will be an unbiased estimator for σ_2^2.

$\boldsymbol{C} \in \mathscr{C}$ is defined in (4.3.2.2).

For $\boldsymbol{A}$ in (4.3.2.7) and $\boldsymbol{C}$ in (4.3.2.2), $Cov(\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}, \boldsymbol{y}'\boldsymbol{C}\boldsymbol{y})$ in (4.3.2.3) can be expressed as

$$
Cov(\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}, \boldsymbol{y}'\boldsymbol{C}\boldsymbol{y}) = (\boldsymbol{Q}\boldsymbol{\eta} - \boldsymbol{g})'\boldsymbol{\delta}. \tag{4.3.2.8}
$$

Consider $\boldsymbol{\eta} = \begin{bmatrix} 1/4 \\ 1/4 \\ -1/4 \\ 1/4 \end{bmatrix}$, we have

$$
\begin{aligned}
\boldsymbol{\phi} &= \boldsymbol{Q}\boldsymbol{\eta} - \boldsymbol{g} \\
&= \begin{bmatrix} -\sigma_2^4(\gamma_2 + 6)/4 \\ \sigma_2^4(\gamma_2 + 6)/2 \\ 0 \\ -\sigma_2^4(\gamma_2 + 6)/4 \end{bmatrix}, \qquad (4.3.2.9) \\
Cov(\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}, \boldsymbol{y}'\boldsymbol{C}\boldsymbol{y}) &= \sigma_2^4(\gamma_2 + 6)/4(-c_1 - 2c_2 - c_4).
\end{aligned}
$$

If $\sigma_2^4(\gamma_2+6)$ is small, then both $\sigma_2^4(\gamma_2+6)/4$ and $\sigma_2^4(\gamma_2+6)/2$ are close to 0, we declare that $\boldsymbol{\phi}$ is close to null vector, and the covariance will be close to 0.

The corresponding NOPE under the model in (4.2.2) for σ_2^2 is

$$\hat{\sigma}_2^2 = \frac{1}{4}(y_{11} - y_{12} - y_{31} + y_{32})^2. \tag{4.3.2.10}$$

4.3.3 Simulations of NOPE

In section 4.2, under the UIBD in Table 4.1 with the mixed-effects model in (4.2.1), the NOPE for σ_1^2 and σ_2^2 are

$$\begin{array}{rcl} \hat{\sigma}_{1NOPE}^2 & = & \frac{1}{2}(y_{11}-y_{31})(y_{12}-y_{32}), \\ \hat{\sigma}_{2NOPE}^2 & = & \frac{1}{4}(y_{11}-y_{12}-y_{31}+y_{32})^2 \end{array}.$$

Simulations of NOPEs for σ_1^2 and σ_2^2 are conducted by the following steps:

- Assuming the mixed-effects model in (4.2.1) with the variance components, σ_1^2 and σ_2^2, set to $(0.05, 0.005), (0.005, 0.05), (1, 0.5), (0.5, 1)$. The fixed effects are $\beta_1 = 5, \beta_2 = 7, \beta_3 = 10$.

- For each set of the values of the variance components, generate 3 block random effects in the vector $\boldsymbol{b}$ from $N(0, \sigma_1^2)$ and generate 6 random errors in the vector $\boldsymbol{e}$ from $N(0, \sigma_2^2)$ for each dataset.

- Creating the observation vectors by $y_{ij} = \beta_i + b_j + e_{ij}$.

- 100,000 datasets are simulated.

- For each simulated data, calculate NOPE for σ_1^2 and σ_2^2, respectively, using $\hat{\sigma}^2_{1NOPE}$ and $\hat{\sigma}^2_{2NOPE}$ above.

- $\hat{\sigma}^2_{1NOPE}$ can be negative, then replace the negative $\hat{\sigma}^2_{1NOPE}$ by 0. $\hat{\sigma}^2_{2NOPE}$ will always be non-negative.

- Using the same simulated data, the ML and REML estimators are also computed. All negative estimates on σ_1^2 are replaced by 0.

- The critera to evaluate the estimation performances of MLE, REMLE and NOPE are

$$\begin{cases} \text{Root Mean Squared Error (RMSE)} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\hat{\theta}_i - \theta\right)} \\ \text{Mean Absolute Deviation (MAD)} = \frac{1}{n}\sum_{i=1}^{n}\left|\hat{\theta}_i - \theta\right| \\ \text{Absolute Bias(AB)} = \left|\frac{1}{n}\sum_{i=1}^{n}\hat{\theta}_i - \theta\right| \end{cases}$$

The performance comparisons of NOPE, MLE, and REMLE for σ_1^2 and σ_2^2 are given in Table 4.3 using the three criterion functions RMSE, MAD and AB. Smaller value of criterion function means better the estimate performance.

Table 4.3: Simulations for $\sigma_1^2 = 1, \sigma_2^2 = 0.5$

Parameters	True Values	Criteria	MLE	REMLE	NOPE
σ_1^2	1	RMSE	0.94189	1.35247	1.7929
		MAD	0.75511	0.96801	1.22631
		AB	0.23329	0.09039	0.00197
σ_2^2	0.5	RMSE	0.40785	0.55735	0.7076
		MAD	0.37897	0.41879	1.22631
		AB	0.32882	0.08328	0.00111

Table 4.4: Simulations for $\sigma_1^2 = 0.5, \sigma_2^2 = 1$

Parameters	True Values	Criteria	MLE	REMLE	NOPE
σ_1^2	0.5	RMSE	0.82351	1.22267	1.57486
		MAD	0.57231	0.77899	1.04633
		AB	0.08431	0.31208	0.00169
σ_2^2	1	RMSE	0.79572	0.90738	1.40633
		MAD	0.74516	0.74201	0.96255
		AB	0.32882	0.08328	0.00111

Table 4.5: Simulations for $\sigma_1^2 = 0.05, \sigma_2^2 = 0.005$

Parameters	True Values	Criteria	MLE	REMLE	NOPE
σ_1^2	0.05	RMSE	0.03927	0.0539	0.07463
		MAD	0.03284	0.03953	0.05091
		AB	0.01578	0.00026	0.017
σ_2^2	0.005	RMSE	0.00702	0.007	0.00702
		MAD	0.00483	0.00474	0.00483
		AB	0.00002	0.00015	0.00002

Table 4.6: Simulations for $\sigma_1^2 = 0.005, \sigma_2^2 = 0.05$

Parameters	True Values	Criteria	MLE	REMLE	NOPE
σ_1^2	0.05	RMSE	0.02913	0.04224	0.00702
		MAD	0.01627	0.02308	0.00483
		AB	0.23329	0.09039	0.00197
σ_2^2	0.005	RMSE	0.0383	0.04167	0.07083
		MAD	0.03535	0.03511	0.04864
		AB	0.01134	0.01824	0.00028

The performance of NOPE for estimating σ_1^2 and σ_2^2 under the criterion function AB is better than both MLE and REMLE as demonstrated in Table 4.3 for σ_1^2 and $\sigma_2^2 = (1, 0.5), (0.5, 1)$. The performances of MLE and REMLE are better than NOPE under the criterion function RMSE and MAD. However, the differences between NOPE and REMLE or MLE with respect to RMSE and MAD become very small for the values of σ_1^2 and $\sigma_2^2 = (0.05, 0.005), (0.005, 0.05)$, respectively.

Comparisons are also made between the estimations under normality assumtions and non-normality assumptions for the random errors on criteria of MSE, MAD and AB. The simulation procedures are

- Assuming the mixed-effects model in (4.2.1) with the variance components of the random effects are $\sigma_1^2 = 0.5, \sigma_2^2 = 0.5$, the fixed effects are $\beta_1 = 5, \beta_2 = 7, \beta_3 = 10$.
- Generate 3 block random effects in the vector $\boldsymbol{b}$ from $N(0, \sigma_1^2 = 0.5)$. Generate 6 random errors in the vector $\boldsymbol{e}$ from $SkewNormal(0, \sigma_2^2 = 0.5, \alpha)$, where α is the skewness and equal to $0, 0.2, 0.5, 0.8$, respectively.
- Creating the observation vectors by having $y_{ij} = \beta_i + b_j + e_{ij}$.
- For each simulated data, calculate the NOPE for σ_1^2 and σ_2^2, respectively, using $\hat{\sigma}^2_{1NOPE}$ and $\hat{\sigma}^2_{2NOPE}$.
- $\hat{\sigma}^2_{1NOPE}$ can be negative, then replace the negative $\hat{\sigma}^2_{1NOPE}$ by 0. $\hat{\sigma}^2_{2NOPE}$ will be non-negative.
- Using the same simulated data, the estimates using MoM, ML and REML are also computed. All negative estimates on σ_1^2 are replaced by 0.

Constraints are that we select the rows with both estimates are between 0 and 1 for all methods.

Table 4.7: Simulations for Skew Normal Assumption of the Random Errors for Skewness=0, 0.2

		skewness=0		skewness=0.2	
		σ_1^2	σ_2^2	σ_1^2	σ_2^2
MSE	ML	0.1406987	0.1507905	0.1410539	0.1502407
	REML	0.1428146	0.1206559	0.1429001	0.1208729
	MoM	0.177789	0.1270317	0.1772525	0.1268744
	NOPE	0.1541742	0.1289118	0.1538197	0.1287292
MAD	ML	0.3375406	0.3617736	0.338001	0.3606924
	REML	0.3401189	0.3115702	0.3401595	0.311835
	MoM	0.3536078	0.3222834	0.3529085	0.321595
	NOPE	0.3594773	0.3253927	0.3589814	0.3247065
AB	ML	0.245857	0.3320308	0.2478642	0.3311886
	REML	0.2313799	0.2204798	0.2323231	0.2201566
	MoM	0.2314204	0.2304625	0.2309076	0.2306565
	NOPE	0.2783875	0.2382985	0.2765637	0.2381765

Table 4.8: Simulations for Skew Normal Assumption of the Random Errors for Skewness=0.5, 0.8

		skewness=0.5		skewness=0.8	
		σ_1^2	σ_2^2	σ_1^2	σ_2^2
MSE	ML	0.141179	0.150242	0.1411782	0.1503219
	REML	0.1428403	0.1208147	0.1430371	0.1209708
	MOM	0.1765556	0.1270954	0.1764622	0.1274011
	NOPE	0.153747	0.128984	0.1538815	0.1292965
MAD	ML	0.3383945	0.360726	0.3384362	0.3608316
	REML	0.3400015	0.3116567	0.3402305	0.3117297
	MoM	0.3526914	0.3218177	0.3524321	0.3224994
	NOPE	0.358954	0.325065	0.3591034	0.3256738
AB	ML	0.2469016	0.331376	0.247268	0.3313993
	REML	0.2334383	0.2193637	0.2334936	0.2191658
	MoM	0.2305244	0.2314463	0.2316282	0.2315001
	NOPE	0.2763293	0.2389698	0.2766649	0.23941

Under the departure from normality assumption using skew normal distribution, NOPE performs better in estimating σ_2^2 than that of MLEs under all three criterion functions for all skewness. NOPE also has smaller MSE than MoM estimators for all four skewness.

4.4 The Average Optimum Estimation

4.4.1 Introduction

In section 4.3, we introduced an UIBD with one replication. Consider a block design with more than one replication, the number of observations will become large, and the dimension of the $\boldsymbol{A}$ matrix in the NOPE will be large too, which is computationally difficult to find the $\boldsymbol{A}$ matrix for the full data. When data are in replicated unbalanced block design, an averaged $\boldsymbol{A}$ matrix can be obtained by averaging the $\boldsymbol{A}$ matrix in NOPE for each of the possible combinations of blocks with the same design, the resulted quadratic estimator is called the average optimum estimator (AOPE).

Two things are important in deriving AOPE, one is the NOPE introduced in section 4.3, another one is the replication in the block design, because replication is a convenient way of increasing sample size and precision of the estimation of variance components.

Consider a randomized block design D_0 with n observations. The NOPE for the ith variance component can be obtained under D_0, denoted by $\boldsymbol{y}_0'\boldsymbol{A}_{i0}\boldsymbol{y}_0$, for $i = 1, 2, \ldots, s$. Consider another block design D with r replications of D_0, let the total number of possible combinations of the blocks that form D_0 design be R.

Definition 4.1. Let $\boldsymbol{y}_{j,0}$ be the vector of observations for the jth replicated D_0 among R combinations, $j = 1, 2, \ldots R$, the NOPE for the ith variance component for the jth combination is $\boldsymbol{y}_{j,0}'\boldsymbol{A}_{i,0}\boldsymbol{y}_{j,0}$, then the AOPE for the ith variance component under D is the average of the NOPEs across R combinations of blocks, which can be expressed as

$$\hat{\sigma}^2_{iAOPE} = \frac{1}{R}\sum_{j=1}^{R} \boldsymbol{y}'_{j,0}\boldsymbol{A}_{i,0}\boldsymbol{y}_{j,0} \tag{4.4.1.1}$$

4.4.2 Example

Consider UIBD in Table 4.1 with replication of 3:

Table 4.9: Unbalanced Incomplete Block Design with Replication of 3

Block1	Block2	Block3	Block4	Block5	Block6	Block7	Block8	Block9
A	A	A	A	A	A	A	A	A
B	C	B	B	C	B	B	C	B

The data are

Table 4.10: Data under Unbalanced Incomplete Block Design with Replication of 3

Block1	Block2	Block3	Block4	Block5	Block6	Block7	Block8	Block9
y_{11}	y_{21}	y_{31}	y_{41}	y_{51}	y_{61}	y_{71}	y_{81}	y_{91}
y_{12}	y_{23}	y_{32}	y_{42}	y_{53}	y_{62}	y_{72}	y_{83}	y_{92}

In this design, Block1, 2 and 3 for a D_0. In section 4.3, we obtained the NOPEs for σ_1^2 and σ_2^2, respectively, for one replication. Under the data in Table 4.10, we will compute the AOPEs in the following steps.

- In Table 4.10, there are 6 blocks taking treatment A and B, 3 blocks taking treatment A and C. The number of combinations of blocks forming a D_0 is $R = \binom{6}{2}\binom{3}{1} = 45$. But notice that NOPEs for σ_1^2 and σ_2^2 in (4.3.2.6) and (4.3.2.10) only involve the observations within the blocks taking treatment A and B, then we can make combinations among the 6 blocks with treatment A and B by $\binom{6}{2}$, which results in $R = 15$ possible combinations of blocks.

- NOPEs are calculated for σ_1^2 and σ_2^2 for each of the 15 combinations.

- Among the 15 NOPEs for σ_1^2, the negative values among are deleted. For the 15 NOPEs of σ_2^2, all of them are positive.

- Take the average of the NOPEs for $\hat{\sigma}_1^2$ and $\hat{\sigma}_2^2$, respectively, then the averaged NOPEs are the AOPEs under UIBD with replications, denoted by $\hat{\sigma}^2_{1,AOPE}$ and $\hat{\sigma}^2_{2,AOPE}$, respectively.

4.4.3 Simulations of AOPE

Simulation is conducted under UIBD with 3 replications to compare the performance of AOPE, MLE and REMLE. Three performance measures are used, which are RMSE, MAD and AB.

Assuming the mixed-effects model in (4.2.1) for the UIBD with 3 replications in Table 4.9, the fixed effects are the effects of three treatments and are set to be $\beta_1 = 5, \beta_2 = 7, \beta_3 = 10$, the variance components of the random effects are $\sigma_1^2 = 1, \sigma_2^2 = 0.5$. 100000 datasets were simulated, each of which has 18 observations in 9 blocks. We estimated σ_1^2 and σ_2^2 using MLE, REMLE and AOPE, respectively. Different constraints for the estimation results are illustrated in Table 4.11:

Based on the results, if all positive estimates from three methods are compared, REMLE has the smallest AB for σ_1^2, while AOPE has smallest AB for σ_2^2; AOPE has relatively larger RMSE and MAD for both σ_1^2 and σ_2^2 compared to REMLE, but has smaller RMSE and MAD compared to MLE.

Table 4.11: Simulations for AOPEs under Different Constraints

	$\hat{\sigma}_1^2 > 0$, $\hat{\sigma}_2^2 > 0$	RMSE		MAD		AB	
	number	σ_1^2	σ_2^2	σ_1^2	σ_2^2	σ_1^2	σ_2^2
ML	86246	6.7702039	0.2653412	1.6850741	0.2075662	1.21615089	0.0921169127
RML	97510	0.6409854	0.258142	0.498303	0.2031989	0.02772392	0.0088170505
AOPE	99999	0.9981198	0.3167611	0.7107089	0.2447469	0.41937281	0.0009931301
	$\hat{\sigma}_1^2$<2, $\hat{\sigma}_2^2$<1	RMSE		MAD		AB	
	number	σ_1^2	σ_2^2	σ_1^2	σ_2^2	σ_1^2	σ_2^2
ML	72345	0.4854838	0.2215651	0.4092232	0.1884476	0.12349361	0.1020785
REML	85530	0.4871395	0.2158477	0.4099007	0.1802718	0.08827059	0.0373453
AOPE	72268	0.4888204	0.2399181	0.4111224	0.2037107	0.04022179	0.0634323
	$\hat{\sigma_1}^2$<2.5, $\hat{\sigma}_2^2$<1.5	RMSE		MAD		AB	
	number	σ_1^2	σ_2^2	σ_1^2	σ_2^2	σ_1^2	σ_2^2
ML	77531	0.5465144	0.2394244	0.4502371	0.1985373	0.06533076	0.08973339
REML	94490	0.5539084	0.2520327	0.4556334	0.2007752	0.02760647	0.01017646
AOPE	87331	0.6236932	0.2900552	0.5063284	0.2338359	0.17667736	0.01494005
	$\hat{\sigma_1}^2$<3, $\hat{\sigma}_2^2$<2	RMSE		MAD		AB	
	number	σ_1^2	σ_2^2	σ_1^2	σ_2^2	σ_1^2	σ_2^2
ML	79441	0.5963306	0.2481871	0.4774922	0.2018711	0.029131999	0.087485111
REML	96576	0.5981806	0.2576327	0.4800423	0.2030464	0.005110412	0.00843176
AOPE	93713	0.7369966	0.310142	0.5786544	0.2425551	0.268230017	0.003120061

When comparing the estimates within the intervals [0,2] and [0,1] for σ_1^2 and σ_2^2, respectively, AOPE has smallest AB for σ_1^2, REMLE has the smallest AB for σ_2^2; REMLE and MLE have smaller RMSE and MAD for both σ_1^2 and σ_2^2 than that of AOPE.

When we widen the right bound of the interval, the AB of AOPE gets larger for σ_1^2 and smaller for σ_2^2, while REMLE and MLE's AB become smaller for both σ_1^2 and σ_2^2; the RMSE and MAD of AOPE also get larger for both σ_1^2 and σ_2^2.

Chapter 5

Mixed Effects Model in Small Area Estimation

5.1 Introduction of Small Area Estimation

Surveys are used to provide reliable predictions for certain characteristics not only for the total population of interest but also for a variety of subpopulations, those subpopulations are called domains. Examples of domain include a geographical domain (state, county or school district within a geographic area), a socio-demographic domain (a specific age-sex-race group) or an industrial domain (a set of firms belonging to a spe1,cific industry).

In sample surveys, a 'direct' domain estimator is only based on domain-specific sample data. A domain is regarded as 'large' if the sample within the domain is large enough to yield 'direct estimates' of adequate precison. A domain is regarded as 'small' if the domain-specific sample is not large enough to provide direct estimates at adequate precision. In this study, we will refer to those small domains as small areas (SAs).

There are various reasons for the scarcity of domain-specific data available in SAs. Typically, when the sampling design aims to provide reliable estimation for large areas and pays little attention to SAs of interest, the samples for those SAs could be small or even unavailable. For example, the overall sample size of a statewide telephone survey in the state of Nebraska is 4300, which is large enough to produce reliable direct estimates of the prevalance of alcohol abuse for the state or some large counties, but there are only 14 observations available in Boone county, a small county in Nebraska. The problem is even more severe for direct survey estimation of the prevalence for white female in the age-group 25-44 in this county, since only one observation is available from the survey, see 17 ().

When the sample sizes in SAs are too small, the traditional direct survey estimation methods are likely to yield large sampling errors. In order to reduce the samping errors, it is necessary to borrow information on related characteristics of the SAs from administrative data records, and/or combine the survey outcomes from relavant domains or period. Those information are called auxilliary information. Small Area Estimation (SAE) is the method of incorporating auxiliary variables to produce more reliable estimates of characteristics such as means, counts, quantiles, etc. for SAs and assessing the precision of estimations.

5.2 Examples

Consider M mutually exclusive domains in a population, a sample is drawn from the total population based on a specific sampling design. Some domains out of those M domains could have large sample sizes while others have small sample sizes (SAs). Suppose there are m SAs and each has sample size $n_i, i = 1, \ldots, m$. The data are given below

Table 5.1: Survey Data in SA

SA	Sampled Observations			
1	y_{11}	y_{12}	$\cdots$	y_{1n_1}
2	y_{21}	y_{22}	$\cdots$	y_{2n_2}
$\vdots$	$\vdots\ \vdots$	$\vdots$	$\vdots\ \vdots$	
m	y_{m1}	y_{m2}	$\cdots$	y_{mn_m}

Since the sample sizes in SAs $(n_i, i = 1, \ldots, m)$ are small, it is necessary to incorporate auxiliary variables from other data sources or related domains. The SA data could be at either area level or unit level. Data at different level impiy the application of different models for variance component estimations and statistical inferences, which will be introduced in the following sections.

5.2.1 Area-level Data

Consider the auxiliary variable (X) is the average gross income of firms, the characteristics to predict (Y) is the average wages and salaries of firms, the goal is to predict the true average wages and salaries of firms for each of the m SAs. For $i = 1, \ldots, m$, the number of sampled firms in the ith SA is n_i and is small. Suppose only area-specific auxiliary information is known, that is the average gross income of firms for the ith SA, denoted as x_i, the direct estimators of the true average wages and salaries of firms for each SA are also known, denoted as $\hat{y}_i$, the data in Table 5.2 illustrate an example of area-level data

Table 5.2: Area-level Data for m SAs

SA	Sample Size	x_i	$\hat{y}_i$
1	n_1	x_1	$\hat{y}_1$
2	n_2	x_2	$\hat{y}_2$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
m	n_m	x_m	$\hat{y}_m$

5.2.2 Unit-level Data

Consider the survey data at unit level. Suppose the auxiliary variable (average gross income) for each sampled firm is available, the average wages and salaries of each firm (y_{ij}) is also available from the survey data, the data in Table 5.3 illustrates an example of unit-level data

Table 5.3: Unit-level Data for m SAs

Area	Sample Size	x_{ij}	y_{ij}
1	n_1	x_{11}	y_{11}
		$\vdots$	
		x_{1n_1}	y_{1n_1}
2	n_2	x_{21}	y_{21}
		$\vdots$	$\vdots$
		x_{2n_2}	y_{2n_2}
$\vdots$	$\vdots$	$\vdots$	$\vdots$
m	n_m	x_{m1}	y_{m1}
		$\vdots$	$\vdots$
		x_{mn_m}	y_{mn_m}

5.3 The Estimation Methods in SAE

5.3.1 The Indirect Estimation Methods

When making predictions for SAs, it is very common that the sample sizes in SAs are too small to provide reliable direct estimates, so indirect estimation methods are commonly used. The indirect estimation methods improve the prediction precision by borrowing the 'strength' from other related areas or time periods with more reliable predictions on the same characteristics. There are four common indirect SAE methods, they are synthetic estimators, composite estimators, James-Stein estimators and model-based indirect estimators.

An estimator is called synthetic estimator if a reliable direct estimator for a large area, covering several SAs, is used to derive an indirect estimator for a SA under the assumption that the SAs have the same estimated characteristics as the large area[6]. For example, a synthetic estimator for the ith SA total ($\hat{y}_{i,Synthetic}$) without auxiliary information is

$$\hat{y}_{i,Synthetic} = \hat{y}, \tag{5.3.1.1}$$

where $\hat{y}$ is the direct estimator of the total of the larger domain. However the estimator in (5.3.1.1) can be biased if the true SA total is not approximately equal the true total of the larger area.

A synthetic estimator can also be obtained by using auxiliary variables. For example, for $i = 1, \ldots, m$ out of M SAs, when the direct estimator ($\hat{y}_i$) for the area total and the related area-level auxiliary variables ($x_{i1}, \ldots, x_{ip}$) are available, a linear regression model can be applied to the data ($\hat{y}_i, x_{i1}, \ldots, x_{ip}$) and the resulted $\hat{\beta}_0, \ldots \hat{\beta}_p$ lead to regression-synthetic predictors for M areas

$$\hat{y}_{i,Synthetic} = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \cdots + \hat{\beta}_p x_{ip}, i = 1, \ldots, M. \tag{5.3.1.2}$$

However, the estimator in (5.3.1.2) can be heavily biased if the underlying model assumptions are not valid, see 23 ().

Synthetic estimation ($\hat{y}_{i,Synthetic}$) has potential bias if the underlying model assumptions are violated. Even though the direct estimator ($\hat{y}_{i,Direct}$) is unbiased estimator, it has large sampling errors for SAs due to small sample sizes. To balance the drawbacks of the two estimators, a combination of the two estimators lead to the composite estimator, which is expressed as

$$\hat{y}_{i,Composite} = \phi_i \hat{y}_{iS} + (1 - \phi_i)\hat{y}_{iD}, i = 1, \ldots, M, \tag{5.3.1.3}$$

for a suitably chosen weight $\phi_i, (0 < \phi_i < 1)$ for the ith SA, see 23 ().

Another approach of composite estimation is using a common weight for all areas, that is $\phi_i = \phi$, and then minimizing the total MSE with respect to ϕ. This estimation method is called James-Stein method. Using common weight ensures good overall efficiency in estimation but not necessarily for each of the SAs, see 23 ().

The above indirect estimation methods reduce the sampling errors due to small sample sizes by only incorporating the auxiliary variables into the model. The model-based estimation methods not only incorporate auxiliary information, but also take into account the between-area variations that could not be explained by the auxiliary variables, which also help reduce the sampling errors, see 14 (). In the next section, the model-based estimation methods will be introduced.

5.3.2 Models in SAE

The model-based estimation methods in SAE are based on the mixed-effects models. There are two commonly used mixed-effects models in SAE, they are the nested-error mixed-effects model and the Fay-Herriot model.

1 () were interested in estimating the area of corn and soybeans for each of the 12 counties in North-Central Iowa. Each county was then divided into area segments (sampling unit), and there were 37 area segments sampled from the entire North-Central Iowa to ascertain the area under corn and soybean by interviewing farm operator. The number of sampled segments in a county, n_i, ranged from 1 to 6. They also collected the auxiliary information, which are the number of pixels classified as corn and soybeens for each segment by satellite readings. Since the sampled survey data and the auxiliary data are available for each segment, and the area segments are nested within each county (SA), Battese and Fuller proposed the nested-error mixed-effects model. For the jth unit in the ith SA, the general form of the nested-error mixed-effects model is

$$y_{ij} = \beta_0 + \boldsymbol{x}_{ij}\boldsymbol{\beta} + b_i + e_{ij}; i = 1, \dots m; j = 1, \dots, n_i, \tag{5.3.2.1}$$

where m is the number of the SAs in the survey data, and n_i is the number of sampled units from the ith SA, y_{ij} is the reported value of the study characteristics for the jth observation in the ith SA, $\boldsymbol{x}_{ij}$ is the vector of corresponding values of p auxiliary variables for the jth observation in the ith SA. In this example, $m = 12$, n_i ranges from 1 to 6, y_{ij} is the reported area for the jth segment in the ith county, x_{ij1}, x_{ij2} are the number of pixels classified as corn or soybean, respectively. The effects of the auxiliary variables are assumed to be fixed, $\boldsymbol{\beta}$ is the 2×1 coefficients of these auxiliary variables, b_i is the random effects of the ith county, and e_{ij} is the random error of the jth observation in the ith SA. It is assumed that $b_i \overset{iid}{\sim} N(0, \sigma_1^2)$, $e_{ij} \overset{iid}{\sim} N(0, \sigma_2^2)$, b_i's and e_{ij}'s are independent.

Another commonly used model is the Fay-Herriort model (F-H model) proposed by 4 () in the context of estimating per capita income (PCI) for SAs with populations less than 1000. Specifically, the direct estimate of the PCI for the ith SA, $\hat{y}_{i,Direct}$, is available based on past studies, the auxiliary variables for each SA were obtainied from the associated county PCI, the tax return data for 1969, and the data on housing from the 1970 census. Since the direct estimates and the auxiliary variables are all area-specific, the F-H model is applicable. The general form of the F-H model is:

$$\hat{y}_{i,Direct} = \boldsymbol{x}_i'\boldsymbol{\beta} + b_i + e_i, i = 1, \ldots, m \tag{5.3.2.2}$$

$\hat{y}_{i,Direct}$ is the direct survey estimate for the mean PCI in the ith SA, $\boldsymbol{x}_i'$ is a $1 \times p$ vector of values of the auxiliary variables for the ith SA (places with population less than 1000), $\boldsymbol{\beta}$ is a $p \times 1$ vector of the coefficients of auxiliary variables, e_i is the sampling error of the survey estimate in the ith SA, b_i is the random effect of the ith SA. It is assumed that $b_i \overset{iid}{\sim} N(0, \sigma_1^2)$, $e_i \overset{iid}{\sim} N(0, \sigma_2^2/n_i)$, where n_i is the number of sampled observations in the ith SA, σ_2^2 is assumed to be known, b_i's and e_i's are independent.

5.3.3 The Best Linear Unbiased Prediction (BLUP) in SAE

The interest in SAE is to predict the true means $\boldsymbol{\mu}$ (or totals) of an interested variable in SAs. In the corn and soybean example, the total area of corn and soybeans for each county are estimated; in the CPI example, the mean CPI for each area with population less than 1000 are estimated. Under the mixed-effects model, the Best Linear Unbiased Prediction (BLUP) is commonly used to predict the true SA mean (or total).

Consider the general mixed-effects model in (3.2.1) under normality assumptions, it is assumed that $\boldsymbol{b}_i \sim MVN(\boldsymbol{0}, \sigma_i^2 \boldsymbol{I}_{q_i}), i = 1, 2, \ldots, s-1$, $\boldsymbol{e} \sim MVN(\boldsymbol{0}, \sigma_s^2 \boldsymbol{I}_n)$, $\boldsymbol{b}_i$'s and $\boldsymbol{e}$ are independent. Define

$$\begin{aligned}
\boldsymbol{b}' &= (\boldsymbol{b}_1', \boldsymbol{b}_2', \ldots, \boldsymbol{b}_{s-1}'), \\
\boldsymbol{U} &= (\boldsymbol{U}_1, \boldsymbol{U}_2, \ldots, \boldsymbol{U}_{s-1}), \\
Var(\boldsymbol{b}) &= diag(\sigma_i^2 \boldsymbol{I}_{q_i}) = \boldsymbol{G}, \\
Var(\boldsymbol{e}) &= \boldsymbol{R},
\end{aligned}$$

we have

$$\boldsymbol{V} = Var(\boldsymbol{y}) = \boldsymbol{U}\boldsymbol{G}\boldsymbol{U}' + \boldsymbol{R}. \tag{5.3.3.1}$$

Under the general mixed-effects model in (3.2.1), the true SA means we are interested in predicting can be represented as a linear combination of the fixed effects and the random effects, which is

$$\boldsymbol{\mu} = \boldsymbol{\lambda}'\boldsymbol{\beta} + \boldsymbol{m}'\boldsymbol{b}, \tag{5.3.3.2}$$

where $\boldsymbol{\lambda}$ and $\boldsymbol{m}$ are vectors of known constants.

Assuming the variance components are known under the general mixed-effects model with normality assumptions, the Best Linear Unbiased Predictor (BLUP) of the random effects of $\boldsymbol{b}$ by 13 () is

$$\hat{\boldsymbol{b}}_{BLUP} = \boldsymbol{G}\boldsymbol{U'}\boldsymbol{V}^{-1}(\boldsymbol{y} - \boldsymbol{X'}\hat{\beta}), \tag{5.3.3.3}$$

where $\hat{\boldsymbol{\beta}} = (\boldsymbol{X'}\boldsymbol{V^{-1}}\boldsymbol{X})^{-1}\boldsymbol{X'}\boldsymbol{V^{-1}}\boldsymbol{y}$.

The BLUP of $\boldsymbol{\mu}$ is

$$\begin{aligned} \hat{\boldsymbol{\mu}}_{BLUP} &= \boldsymbol{\lambda'}\hat{\boldsymbol{\beta}} + \boldsymbol{m'}\hat{\boldsymbol{b}} \\ &= \boldsymbol{\lambda'}\hat{\boldsymbol{\beta}} + \boldsymbol{m'}\boldsymbol{G}\boldsymbol{U'}\boldsymbol{V}^{-1}(\boldsymbol{y} - \boldsymbol{X'}\hat{\beta}). \end{aligned} \tag{5.3.3.4}$$

Under the F-H model, true mean for the ith SA is represented as $\mu_i = \boldsymbol{x'}\boldsymbol{\beta} + b_i$, its BLUP is

$$\hat{\mu}_i = \frac{\sigma_2^2/n_i}{\sigma_2^2/n_i + \sigma_1^2}\boldsymbol{x'}\hat{\boldsymbol{\beta}} + \frac{\sigma_1^2}{\sigma_2^2/n_i + \sigma_1^2}\hat{y}_{i,Direct},$$

where $\hat{y}_{i,Direct}$ is the direct estimate for the ith SA, see 19 ().

Under the nested-error mixed-effects model, assuming the population means of auxiliary variables for the ith SA, $\bar{\boldsymbol{X}}_i$, are known , the BLUP of μ_i is

$$\hat{\mu}_i = \bar{\boldsymbol{X}}_i'\hat{\boldsymbol{\beta}} + \gamma_i(\bar{y}_i - \bar{\boldsymbol{x}}_i'\hat{\boldsymbol{\beta}}),$$

where $\bar{y}_i, \bar{\boldsymbol{x}}_i$ are the sample averages of the dependent variable and auxiliary variables, respectively, $\gamma_i = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2/n_i}$, see 19 ().

5.4 A Class of Unbiased Estimators for the Variance Components in the F-H Model

The results of BLUPs under the mixed-effects models are based on the assumption that the variance components are known. When the variance components are not known, the BLUPs need to be estimated by replacing the estimated values of those variance components. When the variance components in BLUP are replaced by their estimated values, the predictor in (5.3.3.4) becomes the Empirical Best Linear Unbiased Predictor (EBLUP). The estimation of the variance components is an important step in deriving EBLUPs for SAs.

In this section, we will find a class of unbiased quadratic estimators for the variance components in the F-H model.

5.4.1 F-H Model

When the data $(\hat{y}_{i,Direct}, \boldsymbol{x}_i)$ are area-specific, the F-H model applied is

$$\hat{y}_{i,Direct} = \boldsymbol{x}_i'\boldsymbol{\beta} + b_i + e_i, i = 1, 2, \ldots, m. \tag{5.4.1.1}$$

It is assumed that $b_i's$ are independent with mean 0, variance σ_1^2 and kurtosis γ_1, $e_{ij}'s$ are independent with mean 0, variance $\frac{\sigma_2^2}{n_i}$ and kurtosis γ_2, b_i's and e_i's are independent. Note that the functional form of normality is not needed to derive the unbiased estimators for the variance components using our methods.

In matrix form, (5.4.1.1) can be expressed as

$$\hat{\boldsymbol{y}}_{Direct} = \boldsymbol{X\beta} + \boldsymbol{b} + \boldsymbol{e}, \tag{5.4.1.2}$$

where $\hat{\boldsymbol{y}}_{Direct} = \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \vdots \\ \hat{y}_m \end{bmatrix}$, $\boldsymbol{X} = \begin{bmatrix} 1 & x_{12} & \dots & x_{1p} \\ 1 & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{m2} & \dots & x_{mp} \end{bmatrix}$.

It is assumed that $\boldsymbol{X}$ has full column rank, $E(\boldsymbol{b}) = \boldsymbol{0}$, $Var(\boldsymbol{b}) = \sigma_1^2 \boldsymbol{I}_m$, Kurtosis$(\boldsymbol{b}) = \gamma_1 \boldsymbol{1}_m$, $E(\boldsymbol{e}) = \boldsymbol{0}$, $Var(\boldsymbol{e}) = \sigma_2^2 \boldsymbol{D}$, Kurtosis$(\boldsymbol{e}) = \gamma_2 \boldsymbol{1}_m$, where $\boldsymbol{I}_m$ is an identity matrix of order m, $\boldsymbol{D}$ is an diagonal matrix of order m with $\frac{1}{n_i}$ as the ith diagonal element, $\boldsymbol{1}$ is a vector of ones. It is also assumed that $\boldsymbol{b}$ and $\boldsymbol{e}$ are independent. $\boldsymbol{V} = Var(\boldsymbol{y}) = \sigma_1^2 \boldsymbol{I} + \sigma_2^2 \boldsymbol{D}$.

In this model, σ_1^2 and σ_2^2 are the unknown variance components to be estimated.

5.4.2 The Method of Finding the Class of Unbiased Quadratic Estimators

To estimate σ_1^2 and σ_2^2 in (5.4.1.2), define for $\boldsymbol{X}$ and $\boldsymbol{V}$ in (5.4.1.2) a class of $\mathscr{A}_u$ matrices of order m, which satisfy the following conditions for unbiasedness,

$$\text{For } u = 1, 2, \mathscr{A}_u = \left\{ \boldsymbol{A}_u : \boldsymbol{A}_u = \boldsymbol{A}_u', \boldsymbol{A}_u \boldsymbol{X} = \boldsymbol{0}, tr(\boldsymbol{A}_u) = 2 - u, tr(\boldsymbol{A}_u \boldsymbol{D}) = u - 1 \right\}, \tag{5.4.2.1}$$

for all $\boldsymbol{A}_u \in \mathscr{A}_u$, we have $E(\boldsymbol{y}' \boldsymbol{A}_u \boldsymbol{y}) = \sigma_u^2$, in other words, $\boldsymbol{y}' \boldsymbol{A}_u \boldsymbol{y}$ is an unbiased estimator of σ_u^2, $u = 1, 2$.

For the condition $\boldsymbol{A}_u = \boldsymbol{A}_u', \boldsymbol{A}_u \boldsymbol{X} = \boldsymbol{0}$ in (5.4.2.1), we can partition $\boldsymbol{A}_u$ and $\boldsymbol{X}$ as

$$\begin{bmatrix} \boldsymbol{A}_{u11} & \boldsymbol{A}_{u12} \\ \boldsymbol{A}_{u12}' & \boldsymbol{A}_{u22} \end{bmatrix} \begin{bmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{0} \end{bmatrix}, \tag{5.4.2.2}$$

where $\boldsymbol{A}_{u11}$ has full rank with order $(m-p)\times(m-p)$, $\boldsymbol{A}_{u12}$ has order $(m-p)\times p$ and $\boldsymbol{A}_{u22}$ has order $p\times p$, $\boldsymbol{X}_1$ contains $(m-p)$ rows in $\boldsymbol{X}$, $\boldsymbol{X}_2$ contains the remaining p rows in $\boldsymbol{X}$. Note that the order of the observations may be permuted such that $\boldsymbol{X}$ can be partitioned in a way that $\boldsymbol{X}_2$ must be a full rank $p\times p$ matrix.

From (5.4.2.2), we have

$$\begin{cases} \boldsymbol{A}_{u11}\boldsymbol{X}_1 + \boldsymbol{A}_{u12}\boldsymbol{X}_2 = \boldsymbol{0}, & \textcircled{1} \\ \boldsymbol{A}'_{u12}\boldsymbol{X}_1 + \boldsymbol{A}_{u22}\boldsymbol{X}_2 = \boldsymbol{0}. & \textcircled{2} \end{cases} \tag{5.4.2.3}$$

Rearrange (5.4.2.3), we have

$$\begin{cases} \boldsymbol{X}_1 = -\boldsymbol{A}_{u11}^{-1}\boldsymbol{A}_{u12}\boldsymbol{X}_2, & \textcircled{3} \\ \boldsymbol{A}_{u22}\boldsymbol{X}_2 = \boldsymbol{A}'_{u12}\boldsymbol{A}_{u11}^{-1}\boldsymbol{A}_{u12}\boldsymbol{X}_2. & \textcircled{4} \end{cases} \tag{5.4.2.4}$$

From ③ , $\boldsymbol{A}_{u11}$ and $\boldsymbol{X}_2$ have full rank, we can represent $\boldsymbol{A}_{u12}$ in terms of $\boldsymbol{A}_{11}, \boldsymbol{X}_1$ and $\boldsymbol{X}_2$ as

$$\boldsymbol{A}_{u12} = -\boldsymbol{A}_{u11}\boldsymbol{X}_1\boldsymbol{X}_2^{-1}. \tag{5.4.2.5}$$

Rearranging ④ in (5.4.2.4), we have

$$\left(\boldsymbol{A}_{u22} - \boldsymbol{A}'_{u12}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{u12}\right)\boldsymbol{X}_2 = \boldsymbol{0}. \tag{5.4.2.6}$$

Again, since $\boldsymbol{X}_2$ in (5.4.2.6) has full rank, then (5.4.2.6) becomes

$$\boldsymbol{A}_{u22} = \boldsymbol{A}'_{u12}\boldsymbol{A}_{u11}^{-1}\boldsymbol{A}_{u12}. \tag{5.4.2.7}$$

Plugging $\boldsymbol{A}_{u12}$ in (5.4.2.5) into (5.4.2.7), we can also express $\boldsymbol{A}_{u22}$ in terms of $\boldsymbol{A}_{u11}, \boldsymbol{X}_1, \boldsymbol{X}_2$

$$\boldsymbol{A}_{u22} = \boldsymbol{X}_2^{\prime -1}\boldsymbol{X}'_1\boldsymbol{A}_{u11}\boldsymbol{X}_1\boldsymbol{X}_2^{-1}. \tag{5.4.2.8}$$

Using (5.4.2.5) and (5.4.2.8), we have

$$\boldsymbol{A}_u = \begin{bmatrix} \boldsymbol{A}_{u11} & -\boldsymbol{A}_{u11}\boldsymbol{X}_1\boldsymbol{X}_2^{-1} \\ -\boldsymbol{X}_2^{'-1}\boldsymbol{X}_1'\boldsymbol{A}_{u11} & \boldsymbol{X}_2^{'-1}\boldsymbol{X}_1'\boldsymbol{A}_{u11}\boldsymbol{X}_1\boldsymbol{X}_2^{-1} \end{bmatrix}. \tag{5.4.2.9}$$

For the condition $tr(\boldsymbol{A}_u) = 2 - u$ in (5.4.2.1), using the expression of $\boldsymbol{A}_u$ in (5.4.2.9), we have

$$tr(\boldsymbol{A}_{u11}) + tr(\boldsymbol{X}_2^{'-1}\boldsymbol{X}_1'\boldsymbol{A}_{u11}\boldsymbol{X}_1\boldsymbol{X}_2^{-1}) = 2 - u, \tag{5.4.2.10}$$

implying

$$tr(\boldsymbol{A}_{u11}) + tr[\boldsymbol{A}_{u11}\boldsymbol{X}_1(\boldsymbol{X}_2'\boldsymbol{X}_2)^{-1}\boldsymbol{X}_1'] = 2 - u. \tag{5.4.2.11}$$

Let $\boldsymbol{U} = \boldsymbol{X}_1(\boldsymbol{X}_2'\boldsymbol{X}_2)^{-1}\boldsymbol{X}_1'$, (5.4.2.11) can be written as

$$\begin{aligned} tr(\boldsymbol{A}_{u11}) + tr(\boldsymbol{A}_{u11}\boldsymbol{U}) &= 2 - u, \\ tr[\boldsymbol{A}_{u11}(\boldsymbol{I} + \boldsymbol{U})] &= 2 - u. \end{aligned} \tag{5.4.2.12}$$

Let $\boldsymbol{W} = \boldsymbol{I} + \boldsymbol{U}$, (5.4.2.12) can be written as

$$tr(\boldsymbol{A}_{u11}\boldsymbol{W}) = 2 - u. \tag{5.4.2.13}$$

Define

$$\boldsymbol{A}_{u11} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1,(m-p)} \\ a_{12} & a_{22} & \cdots & a_{2,(m-p)} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1,(m-p)} & a_{2,(m-p)} & \cdots & a_{(m-p),(m-p)} \end{bmatrix},$$

$$
\boldsymbol{W} = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1,(m-p)} \\ w_{12} & w_{22} & \cdots & w_{2,(m-p)} \\ \vdots & \vdots & \ddots & \vdots \\ w_{1,(m-p)} & w_{2,(m-p)} & \cdots & w_{(m-p),(m-p)} \end{bmatrix},
$$

$\boldsymbol{A}_{u11}$ and $\boldsymbol{W}$ are both symmetric, $\boldsymbol{A}_{u11}$ has full rank. From (5.4.2.12), we have

$$
tr(\boldsymbol{A}_{u11}\boldsymbol{W}) = \sum_{i=j} a_{ij}w_{ij} + 2\sum_{i<j} a_{ij}w_{ij} = 2 - u. \tag{5.4.2.14}
$$

For the condition $tr(\boldsymbol{A}_u\boldsymbol{D}) = u - 1$ in (5.4.2.1), we partition $\boldsymbol{D}$ in (5.4.1.2) according to the partition of $\boldsymbol{X}$ as

$$
\boldsymbol{D} = \begin{bmatrix} \boldsymbol{D}_1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{D}_2 \end{bmatrix},
$$

where $\boldsymbol{D}_1$ has order $(m-p) \times (m-p)$, $\boldsymbol{D}_2$ has order $p \times p$.

Using the expression of $\boldsymbol{A}_u$ in (5.4.2.9), we have

$$
tr(\boldsymbol{A}_{u11}\boldsymbol{D}_1) + tr(\boldsymbol{X}_2^{\prime -1}\boldsymbol{X}_1^{\prime}\boldsymbol{A}_{u11}\boldsymbol{X}_1\boldsymbol{X}_2^{-1}\boldsymbol{D}_2) = u - 1, \tag{5.4.2.15}
$$

implying

$$
tr(\boldsymbol{A}_{u11}\boldsymbol{D}_1) + tr[\boldsymbol{A}_{u11}\boldsymbol{X}_1\boldsymbol{X}_2^{-1}\boldsymbol{D}_2\boldsymbol{X}_2^{\prime -1}\boldsymbol{X}_1^{\prime}] = u - 1. \tag{5.4.2.16}
$$

Define $\boldsymbol{E} = \boldsymbol{X}_1\boldsymbol{X}_2^{-1}\boldsymbol{D}_2\boldsymbol{X}_2^{\prime -1}\boldsymbol{X}_1^{\prime}$, (5.4.2.16) can be written as

$$
\begin{aligned} tr(\boldsymbol{A}_{u11}\boldsymbol{D}_1) + tr(\boldsymbol{A}_{u11}\boldsymbol{E}) &= 1 - u, \\ tr[\boldsymbol{A}_{u11}(\boldsymbol{D}_1 + \boldsymbol{E})] &= 1 - u. \end{aligned} \tag{5.4.2.17}
$$

Define $\boldsymbol{F} = \boldsymbol{D}_1 + \boldsymbol{E}$, (5.4.2.17) can be written as

$$tr(\boldsymbol{A}_{u11}\boldsymbol{F}) = 1 - u. \tag{5.4.2.18}$$

Define

$$\boldsymbol{F} = \begin{bmatrix} f_{11} & f_{12} & \cdots & f_{1,(m-p)} \\ f_{12} & f_{22} & \cdots & f_{2,(m-p)} \\ \vdots & \vdots & \ddots & \vdots \\ f_{1,(m-p)} & f_{2,(m-p)} & \cdots & f_{(m-p),(m-p)}, \end{bmatrix}$$

$$tr(\boldsymbol{A}_{u11}\boldsymbol{F}) = \sum_{i=j} a_{ij} f_{ij} + 2 \sum_{i<j} a_{ij} f_{ij} = u - 1. \tag{5.4.2.19}$$

We need to choose a symmetric $\boldsymbol{A}_{u11}$ satisfying equation (5.4.2.14) and equation (5.4.2.19), then the resulted quadratic estimator $\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}$ with $\boldsymbol{A}_{u11}$ will be unbiased for $\sigma_u^2, u = 1, 2$.

5.4.3 Example

In this section, we present a numerical example to illustrate the procedure of finding the unbiased class of matrices $\boldsymbol{A}$ in (5.4.2.1).

Consider the following data, which contain four SAs. For $i = 1, 2, 3, 4$, the auxiliary variable x_i is the sample average of the gross income of firms in the ith SA, $\hat{y}_{i,Direct}$ is the direct estimates of the true average wages and salaries of firms in the ith SA:

Table 5.4: Numeric Data for F-H Model

SA	n_i	x_i	$\hat{y}_{i,Direct}$
1	2	35	14.143
2	3	31	7.258
3	3	95	24.211
4	2	135	18.519

Since only area-level information of the auxiliary variable and area-specific direct estimates are available, we apply F-H model to this data, which is

$$\hat{\boldsymbol{y}}_{Direct} = \boldsymbol{X\beta} + \boldsymbol{b} + \boldsymbol{e}, \tag{5.4.3.1}$$

where $\hat{\boldsymbol{y}}_{Direct} = \begin{bmatrix} 14.143 \\ 7.258 \\ 24.211 \\ 18.519 \end{bmatrix}, \boldsymbol{X} = \begin{bmatrix} 1 & 35 \\ 1 & 31 \\ 1 & 95 \\ 1 & 135 \end{bmatrix}.$

The rank of $\boldsymbol{X}$ is 2. It is assumed that $E(\boldsymbol{b}) = \boldsymbol{0}, Var(\boldsymbol{b}) = \sigma_1^2\boldsymbol{I}_4$, $E(\boldsymbol{e}) = \boldsymbol{0}, Var(\boldsymbol{e}) = \sigma_2^2\boldsymbol{D}$, where $\boldsymbol{D} = diag(1/2, 1/3, 1/3, 1/2)$, $\boldsymbol{b}$ and $\boldsymbol{e}$ are independent. $\boldsymbol{V} = \mathrm{Var}(\boldsymbol{y}) = \sigma_1^2\boldsymbol{I}_4 + \sigma_2^2\boldsymbol{D}$. To estimate σ_1^2 and σ_2^2 in (5.4.3.1), define for $\boldsymbol{X}$ and $\boldsymbol{V}$ in (5.4.3.1) a class of $\boldsymbol{A}_u$ matrices with order of 4 as

$$\text{for } u = 1, 2, \boldsymbol{A}_u = \begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ a_2 & a_5 & a_6 & a_7 \\ a_3 & a_6 & a_8 & a_9 \\ a_4 & a_7 & a_9 & a_{10} \end{bmatrix}. \tag{5.4.3.2}$$

If the class of $\boldsymbol{A}_u$ matrices satisfy the following conditions,

$$u = 1, 2, \mathscr{A}_u = \left\{\boldsymbol{A}_u : \boldsymbol{A}_u = \boldsymbol{A}_u', \boldsymbol{A}_u\boldsymbol{X} = \boldsymbol{0}, tr(\boldsymbol{A}_u) = 2 - u, tr(\boldsymbol{A}_u\boldsymbol{D}) = u - 1\right\}. \tag{5.4.3.3}$$

For all $\boldsymbol{A}_u \in \mathscr{A}_u$, we have $E(\boldsymbol{y}'\boldsymbol{A}_u\boldsymbol{y}) = \sigma_u^2$, in other words, $\boldsymbol{y}'\boldsymbol{A}_u\boldsymbol{y}$ is an unbiased estimator of σ_u^2.

For the condition $\boldsymbol{A}_u\boldsymbol{X} = \boldsymbol{0}$ in (5.4.3.3), we can partition $\boldsymbol{A}_u$ and $\boldsymbol{X}$ such that $\boldsymbol{X}_2$ has full rank, that is

$$\begin{bmatrix} \boldsymbol{A}_{u11} & \boldsymbol{A}_{u12} \\ \boldsymbol{A}'_{u12} & \boldsymbol{A}_{u22} \end{bmatrix} \begin{bmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{0} \end{bmatrix}, \tag{5.4.3.4}$$

where $\boldsymbol{A}_{u11} = \begin{bmatrix} a_1 & a_2 \\ a_2 & a_3 \end{bmatrix}$, $\boldsymbol{X}_1 = \begin{bmatrix} 1 & 35 \\ 1 & 31 \end{bmatrix}$, $\boldsymbol{X}_2 = \begin{bmatrix} 1 & 95 \\ 1 & 135 \end{bmatrix}$.

From (5.4.2.9), we have

$$\boldsymbol{A}_u = \left[\begin{array}{cc|cc} a_1 & a_2 & -\frac{5a_1}{2} - \frac{13a_2}{5} & \frac{3a_1}{2} + \frac{8a_2}{5} \\ a_2 & a_3 & -\frac{5a_2}{2} - \frac{13a_3}{5} & \frac{3a_2}{2} + \frac{8a_3}{5} \\ \hline -\frac{5a_1}{2} - \frac{13a_2}{5} & -\frac{5a_2}{2} - \frac{13a_3}{5} & \frac{25a_1}{4} + 13a_2 + \frac{169a_3}{25} & -\frac{15a_1}{4} - \frac{79a_2}{10} - \frac{104a_3}{25} \\ \frac{3a_1}{2} + \frac{8a_2}{5} & \frac{3a_2}{2} + \frac{8a_3}{5} & -\frac{15a_1}{4} - \frac{79a_2}{10} - \frac{104a_3}{25} & \frac{9a_1}{4} + \frac{24a_2}{5} + \frac{64a_3}{25} \end{array}\right], \tag{5.4.3.5}$$

and

$$\boldsymbol{U} = \begin{bmatrix} \frac{17}{2} & \frac{89}{10} \\ \frac{89}{10} & \frac{233}{25} \end{bmatrix}, \boldsymbol{W} = \begin{bmatrix} \frac{19}{2} & \frac{89}{10} \\ \frac{89}{10} & \frac{258}{25} \end{bmatrix}, \boldsymbol{E} = \begin{bmatrix} 3.2083 & 3.3667 \\ 3.3667 & 3.5333 \end{bmatrix}, \boldsymbol{F} = \begin{bmatrix} 3.7083 & 3.3667 \\ 3.3667 & 3.8667 \end{bmatrix}. \tag{5.4.3.6}$$

From $\boldsymbol{W}$ and $\boldsymbol{F}$, we have

$$tr(\boldsymbol{A}_{u11}\boldsymbol{W}) = \frac{19a_1}{2} + \frac{89a_2}{5} + \frac{258a_3}{25} = 2 - u, \tag{5.4.3.7}$$

$$tr(\boldsymbol{A}_{u11}\boldsymbol{F}) = 3.7083a_1 + 6.7333a_2 + 3.8667a_3 = u - 1. \tag{5.4.3.8}$$

From (5.4.3.7) and (5.4.3.8), we can express a_1, a_3 in terms of a_2, that is

$$\begin{cases} a_1 = -0.430368763557484a_2 - 2.51626898047722, \\ a_3 = -1.32863340563991a_2 + 2.41323210412148. \end{cases} \tag{5.4.3.9}$$

Therefore, for $u = 1, 2$, $\boldsymbol{A}_u$ satisfying the conditions in (5.4.3.3) can be expressed as

$$\boldsymbol{A}_u = \begin{bmatrix} -0.4304a_2+9.232u-11.75 & a_2 & -1.524a_2-23.08u+29.37 & 0.9544a_2+13.85u-17.62 \\ a_2 & -1.329a_2-8.595u+11.01 & 0.9544a_2+22.35u-28.62 & -0.6258a_2-13.75u+17.61 \\ -1.524a_2-23.08u+29.37 & 0.9544a_2+22.35u-28.62 & 1.329a_2-0.4046u+0.9913 & -0.759a_2+1.137u-1.74 \\ 0.9544a_2+13.85u-17.62 & -0.6258a_2-13.75u+17.61 & -0.759a_2+1.137u-1.74 & 0.4304a_2-1.232u+1.748 \end{bmatrix}, \tag{5.4.3.10}$$

where a_2 is any reals.

For $u = 1$, (5.4.3.10) is

$$\boldsymbol{A}_1 = \begin{bmatrix} -0.4304a_2 - 2.516 & a_2 & -1.524a_2 + 6.291 & 0.9544a_2 - 3.774 \\ a_2 & -1.329a_2 + 2.413 & 0.9544a_2 - 6.274 & -0.6258a_2 + 3.861 \\ -1.524a_2 + 6.291 & 0.9544a_2 - 6.274 & 1.329a_2 + 0.5868 & -0.759a_2 - 0.603 \\ 0.9544a_2 - 3.774 & -0.6258a_2 + 3.861 & -0.759a_2 - 0.603 & 0.4304a_2 + 0.5163 \end{bmatrix}, \tag{5.4.3.11}$$

where a_2 is any reals.

For $u = 2$, (5.4.3.10) is

$$\boldsymbol{A}_2 = \left[\begin{array}{cc|cc} -0.4304a_2 + 6.716 & a_2 & -1.524a_2 - 16.79 & 0.9544a_2 + 10.07 \\ a_2 & -1.329a_2 - 6.182 & 0.9544a_2 + 16.07 & -0.6258a_2 - 9.891 \\ \hline -1.524a_2 - 16.79 & 0.9544a_2 + 16.07 & 1.329a_2 + 0.1822 & -0.759a_2 + 0.5336 \\ 0.9544a_2 + 10.07 & -0.6258a_2 - 9.891 & -0.759a_2 + 0.5336 & 0.4304a_2 - 0.7158 \end{array}\right], \tag{5.4.3.12}$$

where a_2 is any reals.

$\boldsymbol{A}_1$ and $\boldsymbol{A}_2$ in (5.4.3.11) and (5.4.3.12) are within the unbiased class of matrices $\mathscr{A}_1$ and $\mathscr{A}_2$, so the quadratic estimators, $\boldsymbol{y}'\boldsymbol{A}_1\boldsymbol{y}$ and $\boldsymbol{y}'\boldsymbol{A}_2\boldsymbol{y}$ are unbiased quadratic estimators for σ_1^2 and σ_2^2, respectively.

5.5 Search for the Optimum Estimator for the Variance Components in the Nested Error Mixed-effects Model

Consider another example for predicting the average wages and salaries of firms for 4 SAs. A sample of 10 firms were collected from the 4 SAs, 2 firms in the sample were from each of SAs 1 and 4, 3 firms in the sample were from each of SAs 2 and 3. The records of the average gross income of a firm (the auxiliary variable X) and the average wage and salary of a firm (the dependent variable Y) were collected in the sample. The data are presented in Table 5.5.

Table 5.5: Firm Data for 4 SAs

Area	Sample Size	X	Y
1	2	30	6
		40	12
2	3	35	2
		28	3
		30	10
3	3	100	40
		110	25
		75	10
4	2	150	23
		120	17

5.5.1 The Nested Error Mixed-effects Model

To obtain the predictions of the average wages and salaries of firms for the 4 SAs when firm-specific data are available, we consider the following linear mixed-effects model,

$$y_{ij} = \boldsymbol{x}'_{ij}\boldsymbol{\beta} + b_i + e_{ij}, i = 1, 2, 3, 4, j = 1, \ldots, n_i, \tag{5.5.1.1}$$

where y_{ij} is the dependent variable (the average wage and salary) for the jth firm from the ith SA, $\boldsymbol{x}'_{ij} = (x_{ij1}, x_{ij2})$ is a vector of the corresponding values for the intercept and the auxiliary variable for the jth firm from the ith SA, $\boldsymbol{\beta} = (\beta_0, \beta_1)'$ is a vector of fixed unknown parameters, b_i is the random effect of the ith SA on the dependent variable for the jth firm from the ith SA, e_{ij} is the random error for the jth firm from the ith SA. It is assumed that $E(b_i) = 0, Var(b_i) = \sigma_1^2$, kurtosis$(b_i) = \gamma_1$, $E(e_{ij}) = 0, Var(e_{ij}) = \sigma_2^2$, kurtosis$(e_{ij}) = \gamma_2$, b_i and e_{ij} are independent. Note the functional forms of the distribution are not needed here. Based on the variances and the independence assumptions of b_i's and e_{ij}'s, we have

$$Cov(y_{ij}, y_{i'j'}) = \begin{cases} \sigma_1^2 + \sigma_2^2, & i = i', j = j', \\ \sigma_1^2, & i = i', j \neq j', \\ 0, & i \neq i', j \neq j'. \end{cases} \tag{5.5.1.2}$$

To express the model in (5.5.1.2) in matrix form, define

$$\boldsymbol{y} = \begin{bmatrix} 6 \\ 12 \\ 2 \\ 3 \\ 10 \\ 40 \\ 25 \\ 10 \\ 23 \\ 17 \end{bmatrix}, \boldsymbol{X} = \begin{bmatrix} 1 & 30 \\ 1 & 40 \\ 1 & 35 \\ 1 & 28 \\ 1 & 30 \\ 1 & 100 \\ 1 & 110 \\ 1 & 75 \\ 1 & 150 \\ 1 & 120 \end{bmatrix}, \boldsymbol{U} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

the model (5.5.1.2) can be expresses as

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{U}_1\boldsymbol{b} + \boldsymbol{e}, \tag{5.5.1.3}$$

where $\boldsymbol{y}$ is a vector of 10 observations, $\boldsymbol{X}$ is the known matrix with values corresponding to the fixed effects parameters, $\boldsymbol{\beta}$ is the vector of the fixed effects parameters, $\boldsymbol{U}$ is the known incidence matrix, each row of $\boldsymbol{U}_1$ has one on the ith element and zeroes on the remaining elements, which represents the jth firm belongs to the ith SA, $\boldsymbol{b}$ and $\boldsymbol{e}$ are the vectors of independent random variables. It is assumed that $E(\boldsymbol{b}) = \boldsymbol{0}, Var(\boldsymbol{b}) = \sigma_1^2\boldsymbol{I}_4$, Kurtosis$(\boldsymbol{b}) = \gamma_1\mathbf{1}_4$, $E(\boldsymbol{e}) = \boldsymbol{0}, Var(\boldsymbol{e}) = \sigma_2^2\boldsymbol{I}_{10}$, Kurtosis$(\boldsymbol{e}) = \gamma_2\mathbf{1}_{10}$. $\boldsymbol{b}$ and $\boldsymbol{e}$ are independent. Define $\boldsymbol{V}_1 = \boldsymbol{U}_1\boldsymbol{U}_1'$, by the variances and the independence of $\boldsymbol{b}$ and $\boldsymbol{e}$, the variance of $\boldsymbol{y}$ can be expressed as

$$Var(\boldsymbol{y}) = \sigma_1^2\boldsymbol{U}_1\boldsymbol{U}_1' + \sigma_2^2\boldsymbol{I}_{10} = \sigma_1^2\boldsymbol{V}_1 + \sigma_2^2\boldsymbol{I}_{10}. \tag{5.5.1.4}$$

In this model, σ_1^2 and σ_2^2, are the unknown variance components to be estimated.

5.5.2 Search for the Optimum Unbiased Quadratic Estimators for σ_2^2

In this section, we propose three methods to construct a matrix $\boldsymbol{A}_2$ within the unbiased class of $\mathscr{A}_2$ by using the MoM estimator as a benchmark. Then we compare the variances of the proposed estimators with that of the MoM estimators.

Method 1

Consider $\boldsymbol{y}'\boldsymbol{A}_2\boldsymbol{y}$ as a quadratic estimator of σ_2^2, where $\boldsymbol{A}_2$ is a 10×10 matrix. To have an unbiased estimator for σ_2^2, the $\boldsymbol{A}_2$ matrix in the estimator $\boldsymbol{y}'\boldsymbol{A}_2\boldsymbol{y}$ for σ_2^2 will satisfy the conditions for unbiasedness

- $\boldsymbol{A}_2$ is a symmetric matrix, that is

$$\boldsymbol{A}_2 = \boldsymbol{A}_2'. \tag{5.5.2.1}$$

- For $\boldsymbol{X}$ and $\boldsymbol{U}_1$ in (5.5.1.3), define $[\boldsymbol{X}, \boldsymbol{U}_1]$ as the concatenation of $\boldsymbol{X}$ and $\boldsymbol{U}_1$ columnwise. Note the last column in $[\boldsymbol{X}, \boldsymbol{U}_1]$ is dependent on the column 1, 3, 4 and 5, so we drop the last column in $[\boldsymbol{X}, \boldsymbol{U}_1]$ so that the matrix $[\boldsymbol{X}, \boldsymbol{U}_1]$ has full column rank when estimating σ_2^2. Denote the $\boldsymbol{U}_1$ matrix without the last column as $\boldsymbol{U}_1^*$ and define $\boldsymbol{W} = [\boldsymbol{X}, \boldsymbol{U}_1^*]$, the second unbiased requirement for $\boldsymbol{A}_2$ is

$$\boldsymbol{A}_2\boldsymbol{W} = \boldsymbol{0}. \tag{5.5.2.2}$$

- The trace of $\boldsymbol{A}_2$ is equal to 1

$$tr(\boldsymbol{A}_2) = 1. \tag{5.5.2.3}$$

- For $\boldsymbol{V}_1$ in (5.5.1.4), the trace of the matrix $\boldsymbol{A}_2\boldsymbol{V}_1$ is equal to 0

$$tr(\boldsymbol{A}_2\boldsymbol{V}_1) = 0. \tag{5.5.2.4}$$

Consider a class $\mathscr{A}_2$ of matrices satisfying the conditions for unbiasedness as

$$\mathscr{A}_2 = \left\{\boldsymbol{A}_2 | \boldsymbol{A}_2 = \boldsymbol{A}_2', \boldsymbol{A}_2\boldsymbol{W} = \boldsymbol{0}, tr(\boldsymbol{A}_2) = 1, tr(\boldsymbol{A}_2\boldsymbol{V}_1) = 0\right\}, \tag{5.5.2.5}$$

for any $\boldsymbol{A}_2 \in \mathscr{A}_2$, $\boldsymbol{y}'\boldsymbol{A}_2\boldsymbol{y}$ is an unbiased estimator of σ_2^2, i.e. $E(\boldsymbol{y}'\boldsymbol{A}_2\boldsymbol{y}) = \sigma_2^2$.

Based on the variance of a quadratic form in (3.4.3), given the $\boldsymbol{X}$ and $\boldsymbol{U}_1$ in (5.5.1.3), we have

$$Var(\boldsymbol{y}'\boldsymbol{A}_2\boldsymbol{y}) = tr(\tilde{\boldsymbol{B}}\Delta_2\tilde{\boldsymbol{B}}) + 2tr(\Delta_1\boldsymbol{B}\Delta_1\boldsymbol{B}), \tag{5.5.2.6}$$

where

$$\begin{aligned} \boldsymbol{B} &= \begin{bmatrix} \boldsymbol{U}_1' \\ \boldsymbol{I}_{10} \end{bmatrix} \boldsymbol{A}_2[\boldsymbol{U}_1, \boldsymbol{I}_{10}], \\ \Delta_1 &= \begin{bmatrix} \sigma_1^2\boldsymbol{I}_4 & \boldsymbol{0} \\ \boldsymbol{0} & \sigma_2^2\boldsymbol{I}_{10} \end{bmatrix}, \Delta_2 = \begin{bmatrix} \gamma_1\sigma_1^4\boldsymbol{I}_4 & \boldsymbol{0} \\ \boldsymbol{0} & \gamma_2\sigma_2^4\boldsymbol{I}_{10} \end{bmatrix}. \end{aligned} \tag{5.5.2.7}$$

Since the $\boldsymbol{A}_2$ matrix in the MoM estimator for σ_2^2, denoted by $\boldsymbol{A}_{2MoM}$, is a member in the class $\mathscr{A}_2$ matrices, we will first present the $\boldsymbol{A}_{2MoM}$ matrix[26], and then we will present the method of constructing a general $\boldsymbol{A}_2$ matrices satisfying all the unbiasedness conditions in (5.5.2.5) by using the matrix structure of $\boldsymbol{A}_{2MoM}$.

Define $\boldsymbol{H} = \boldsymbol{W}(\boldsymbol{W}'\boldsymbol{W})^{-1}\boldsymbol{W}'$, we have

$$\boldsymbol{A}_{2MoM} = \frac{\boldsymbol{I}_{10} - \boldsymbol{H}}{tr(\boldsymbol{I}_{10} - \boldsymbol{H})}, \tag{5.5.2.8}$$

where

$$
\boldsymbol{I}_{10}-\boldsymbol{H} = \frac{1}{1176}\left[\begin{array}{cc|ccc|ccc|cc}
563 & -563 & 20 & -15 & -5 & 25 & 75 & -100 & 75 & -75 \\
-563 & 563 & -20 & 15 & 5 & -25 & -75 & 100 & -75 & 75 \\
\hline
20 & -20 & 768 & -380 & -388 & -20 & -60 & 80 & -60 & 60 \\
-15 & 15 & -380 & 775 & -395 & 15 & 45 & -60 & 45 & -45 \\
-5 & 5 & -388 & -395 & 783 & 5 & 15 & -20 & 15 & -15 \\
\hline
25 & -25 & -20 & 15 & 5 & 759 & -467 & -292 & -75 & 75 \\
75 & -75 & -60 & 45 & 15 & -467 & 559 & -92 & -225 & 225 \\
-100 & 100 & 80 & -60 & -20 & -292 & -92 & 384 & 300 & -300 \\
\hline
75 & -75 & -60 & 45 & 15 & -75 & -225 & 300 & 363 & -363 \\
-75 & 75 & 60 & -45 & -15 & 75 & 225 & -300 & -363 & 363
\end{array}\right],
$$

$$
tr(\boldsymbol{I}_{10}-\boldsymbol{H}) = 5 \tag{5.5.2.9}
$$

To express $\boldsymbol{A}_{2MoM}$ in (5.5.2.8) in the block matrix form, define

$$
\boldsymbol{A}_{11} = 563\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \boldsymbol{A}_{12} = 5\begin{bmatrix} 4 & -3 & -1 \\ -4 & 3 & 1 \end{bmatrix}, \boldsymbol{A}_{13} = 25\begin{bmatrix} 1 & 3 & -4 \\ -1 & -3 & 4 \end{bmatrix},
$$

$$
\boldsymbol{A}_{14} = 75\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \boldsymbol{A}_{22} = \begin{bmatrix} 768 & -380 & -388 \\ -380 & 775 & -395 \\ -388 & -395 & 783 \end{bmatrix}, \boldsymbol{A}_{23} = 5\begin{bmatrix} -4 & -12 & 16 \\ 3 & 9 & -12 \\ 1 & 3 & -4 \end{bmatrix},
$$

$$
\boldsymbol{A}_{24} = 15\begin{bmatrix} -4 & 4 \\ 3 & -3 \\ 1 & -1 \end{bmatrix}, \boldsymbol{A}_{33} = \begin{bmatrix} 759 & -467 & -292 \\ -467 & 559 & -92 \\ -292 & -92 & 384 \end{bmatrix}, \boldsymbol{A}_{34} = 75\begin{bmatrix} -1 & 1 \\ -3 & 3 \\ 4 & -4 \end{bmatrix},
$$

$$\boldsymbol{A}_{44} = 363 \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

The $\boldsymbol{A}_{2MoM}$ in (5.5.2.8) can be expressed as

$$\boldsymbol{A}_{2MoM} = \frac{1}{5880} \begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \boldsymbol{A}_{13} & \boldsymbol{A}_{14} \\ \boldsymbol{A}'_{12} & \boldsymbol{A}_{22} & \boldsymbol{A}_{23} & \boldsymbol{A}_{24} \\ \boldsymbol{A}'_{13} & \boldsymbol{A}'_{23} & \boldsymbol{A}_{33} & \boldsymbol{A}_{34} \\ \boldsymbol{A}'_{14} & \boldsymbol{A}'_{24} & \boldsymbol{A}'_{34} & \boldsymbol{A}_{44} \end{bmatrix}. \tag{5.5.2.10}$$

Using the formula $Var(\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y})$ in (5.5.2.6),

$$Var(\boldsymbol{y}'\boldsymbol{A}_{2MoM}\boldsymbol{y}) = 0.1080867\gamma_2\sigma_2^4 + 0.4\sigma_2^4. \tag{5.5.2.11}$$

In order to construct a general $\boldsymbol{A}_2$ matrices for σ_2^2 satisfying all the unbiasedness conditions in (5.5.2.5) by using the structure of $\boldsymbol{A}_{2MoM}$, we need to make some elements in $\boldsymbol{A}_{2MoM}$ free while maintain the structure of $\boldsymbol{A}_{2MoM}$ so that the constructed $\boldsymbol{A}_2$ still satisfy the unbiasedness conditions.

Suppose the common elements in block matrix $\boldsymbol{A}_{11}, \boldsymbol{A}_{12}, \boldsymbol{A}_{13}, \boldsymbol{A}_{14}, \boldsymbol{A}_{23}, \boldsymbol{A}_{24}, \boldsymbol{A}_{34}, \boldsymbol{A}_{44}$, and the elements in $\boldsymbol{A}_{22}$ and $\boldsymbol{A}_{33}$ are unknown, then the block matrices become

$$\boldsymbol{A}_{11} = a_{11} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \boldsymbol{A}_{12} = a_{12} \begin{bmatrix} 4 & -3 & -1 \\ -4 & 3 & 1 \end{bmatrix}, \boldsymbol{A}_{13} = a_{13} \begin{bmatrix} 1 & 3 & -4 \\ -1 & -3 & 4 \end{bmatrix},$$

$$\boldsymbol{A}_{14} = a_{14} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \boldsymbol{A}_{22} = \begin{bmatrix} -w_1 - w_2 & w_1 & w_2 \\ w_1 & -w_1 - w_3 & w_3 \\ w_2 & w_3 & -w_2 - w_3 \end{bmatrix},$$

$$A_{23} = a_{23}\begin{bmatrix} -4 & -12 & 16 \\ 3 & 9 & -12 \\ 1 & 3 & -4 \end{bmatrix}, A_{24} = a_{24}\begin{bmatrix} -4 & 4 \\ 3 & -3 \\ 1 & -1 \end{bmatrix},$$

$$A_{33} = \begin{bmatrix} -w_4 - w_5 & w_4 & w_5 \\ w_4 & -w_4 - w_6 & w_6 \\ w_5 & w_6 & -w_5 - w_6 \end{bmatrix}, A_{34} = a_{34}\begin{bmatrix} -1 & 1 \\ -3 & 3 \\ 4 & -4 \end{bmatrix}, A_{44} = a_{44}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

Replace the block matrices in $\boldsymbol{A}_{2MoM}$ in (5.5.2.10) by those corresponding block matrices $\boldsymbol{A}_{11}, \boldsymbol{A}_{12}, \boldsymbol{A}_{13}, \boldsymbol{A}_{14}, \boldsymbol{A}_{23}, \boldsymbol{A}_{24}, \boldsymbol{A}_{34}, \boldsymbol{A}_{44}$ with unknown elements, denote the new matrix as $\boldsymbol{A}_{2OE}$, where OE indicates our estimator, we have

$$\boldsymbol{A}_{2OE} = \frac{1}{C}\begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \boldsymbol{A}_{13} & \boldsymbol{A}_{14} \\ \boldsymbol{A}'_{12} & \boldsymbol{A}_{22} & \boldsymbol{A}_{23} & \boldsymbol{A}_{24} \\ \boldsymbol{A}'_{13} & \boldsymbol{A}'_{23} & \boldsymbol{A}_{33} & \boldsymbol{A}_{34} \\ \boldsymbol{A}'_{14} & \boldsymbol{A}'_{24} & \boldsymbol{A}'_{34} & \boldsymbol{A}_{44} \end{bmatrix}, \quad (5.5.2.12)$$

where $C = 2a_{11} + 2a_{44} - 2w_1 - 2w_2 - 2w_3 - 2w_4 - 2w_5 - 2w_6$.

The unknown elements in $\boldsymbol{A}_2$ are $a_{11}, a_{12}, a_{13}, a_{14}, a_{23}, a_{24}, a_{34}, a_{44}, w_1, w_2, w_3, w_4, w_5, w_6$. Note that when $a_{11} = 563, a_{12} = 5, a_{13} = 25, a_{14} = 75, a_{23} = 5, a_{24} = 15, a_{34} = 75, a_{44} = 363, w_1 = -380, w_2 = -388, w_3 = -395, w_4 = -467, w_5 = -292, w_6 = -92$, $\boldsymbol{A}_{2OE}$ is exactly equal to $\boldsymbol{A}_{2MoM}$.

It can be shown that $\boldsymbol{A}_2$ in (5.5.2.12) satisfies all the unbiasedness conditions in (5.5.2.5) except $\boldsymbol{A}_{2OE}\boldsymbol{W} = \boldsymbol{0}$.

Presenting $\boldsymbol{A}_{2OE}\boldsymbol{W}$ explicitly, we have

$$\boldsymbol{A}_2\boldsymbol{W} = \begin{bmatrix} 0 & -10a_{11}+26a_{12}+130a_{13}+30a_{14} & 0 & 0 & 0 \\ 0 & 10a_{11}-26a_{12}-130a_{13}-30a_{14} & 0 & 0 & 0 \\ 0 & -40a_{12}-520a_{23}-120a_{24}-7w_1-5w_2 & 0 & 0 & 0 \\ 0 & 30a_{12}+390a_{23}+90a_{24}+7w_1+2w_3 & 0 & 0 & 0 \\ 0 & 10a_{12}+130a_{23}+30a_{24}+5w_2-2w_3 & 0 & 0 & 0 \\ 0 & -10a_{13}-26a_{23}-30a_{34}+10w_4-25w_5 & 0 & 0 & 0 \\ 0 & -30a_{13}-78a_{23}-90a_{34}-10w_4-35w_6 & 0 & 0 & 0 \\ 0 & 40a_{13}+104a_{23}+120a_{34}+25w_5+35w_6 & 0 & 0 & 0 \\ 0 & -10a_{14}-26a_{24}-130a_{34}+30a_{44} & 0 & 0 & 0 \\ 0 & 10a_{14}+26a_{24}+130a_{34}-30a_{44} & 0 & 0 & 0 \end{bmatrix}. \quad (5.5.2.13)$$

To satisfy the condition that $\boldsymbol{A}_{2OE}\boldsymbol{W} = \mathbf{0}$, the system of equations in the second column of the matrix in (5.5.2.13) needs to be equal to a null vector. Now the problem becomes equating the expressions in the second column in (5.5.2.13) to $\mathbf{0}$ and solved for the 14 unknown elements $(a_{11}, a_{12}, a_{13}, a_{14}, a_{23}, a_{24}, a_{34}, a_{44}, w_1, w_2, w_3, w_4, w_5, w_6)$. Note that 6 equations out of these 10 equations are independent, so we can solve for 6 unknown elements and expressed them in terms of the remaining 8 unknown elements. The system of equations to be solved is

$$
\begin{cases}
-10a_{11} + 26a_{12} + 130a_{13} + 30a_{14} = 0, \\
10a_{11} - 26a_{12} - 130a_{13} - 30a_{14} = 0, \\
-40a_{12} - 520a_{23} - 120a_{24} - 7w_1 - 5w_2 = 0, \\
30a_{12} + 390a_{23} + 90a_{24} + 7w_1 + 2w_3 = 0, \\
10a_{12} + 130a_{23} + 30a_{24} + 5w_2 - 2w_3 = 0, \\
-10a_{13} - 26a_{23} - 30a_{34} + 10w_4 - 25w_5 = 0, \\
-30a_{13} - 78a_{23} - 90a_{34} - 10w_4 - 35w_6 = 0, \\
40a_{13} + 104a_{23} + 120a_{34} + 25w_5 + 35w_6 = 0, \\
-10a_{14} - 26a_{24} - 130a_{34} + 30a_{44} = 0, \\
10a_{14} + 26a_{24} + 130a_{34} - 30a_{44} = 0.
\end{cases}
\tag{5.5.2.14}
$$

To solve (5.5.2.14), define

$$
\boldsymbol{M}_1 = \begin{bmatrix}
-10 & 26 & 0 & 0 & 0 & 0 \\
0 & -40 & -520 & -120 & -7 & 0 \\
0 & 30 & 390 & 90 & 7 & 0 \\
0 & 0 & -26 & 0 & 0 & 10 \\
0 & 0 & -78 & 0 & 0 & -10 \\
0 & 0 & 0 & -26 & 0 & 0
\end{bmatrix},
$$

$$
\boldsymbol{M}_2 = \begin{bmatrix} 130 & 30 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 \\ -10 & 0 & -30 & 0 & 0 & 0 & -25 & 0 \\ -30 & 0 & -90 & 0 & 0 & 0 & 0 & -35 \\ 0 & -10 & -130 & 30 & 0 & 0 & 0 & 0 \end{bmatrix},
$$

$$
\boldsymbol{\alpha}_1 = \begin{bmatrix} a_{11} \\ a_{12} \\ a_{23} \\ a_{24} \\ w_2 \\ w_5 \end{bmatrix}, \boldsymbol{\alpha}_2 = \begin{bmatrix} a_{13} \\ a_{14} \\ a_{34} \\ a_{44} \\ w_1 \\ w_3 \\ w_4 \\ w_6 \end{bmatrix}.
$$

Since $\boldsymbol{M}_1$ is constructed to have full rank, then (5.5.2.14) can be rewritten as

$$
[\boldsymbol{M}_1, \boldsymbol{M}_2] \begin{bmatrix} \boldsymbol{\alpha}_1 \\ \boldsymbol{\alpha}_2 \end{bmatrix} = \mathbf{0}, \tag{5.5.2.15}
$$

which implies

$$
\begin{aligned}
\boldsymbol{\alpha}_1 &= -\boldsymbol{M}_1^{-1}\boldsymbol{M}_2\boldsymbol{\alpha}_2 \\
&= \begin{cases}
26a_{13} + 3a_{14} - \frac{39a_{24}}{5} + 39a_{34} - \frac{91w_1}{150} - \frac{13w_3}{75} + \frac{13w_4}{3} + \frac{91w_6}{6}, \\
5a_{13} - 3a_{24} + 15a_{34} - \frac{7w_1}{30} - \frac{w_3}{15} + \frac{5w_4}{3} + \frac{35w_6}{6}, \\
-\frac{5a_{13}}{13} - \frac{15a_{34}}{13} - \frac{5w_4}{39} - \frac{35w_6}{78}, \\
\frac{a_{14}}{3} + \frac{13a_{24}}{15} + \frac{13a_{34}}{3}, \\
\frac{7w_1}{15} + \frac{8w_3}{15}, \\
\frac{8w_4}{15} + \frac{7w_6}{15}.
\end{cases}
\end{aligned} \tag{5.5.2.16}
$$

Plugging (5.5.2.16) back into $\boldsymbol{A}_{2OE}$ in (5.5.2.12) and define

$$
\begin{aligned}
C &= 52a_{13} + \frac{20a_{14}}{3} - \frac{208a_{24}}{15} + \frac{260a_{34}}{3} - \frac{311w_1}{75} - \frac{256w_3}{75} + \frac{28w_4}{5} + \frac{137w_6}{5}, \\
\boldsymbol{A}_{11} &= \left(26a_{13} + 3a_{14} - \frac{39a_{24}}{5} + 39a_{34} - \frac{91w_1}{150} - \frac{13w_3}{75} + \frac{13w_4}{3} + \frac{91w_6}{6}\right) \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \\
\boldsymbol{A}_{12} &= \left(5a_{13} - 3a_{24} + 15a_{34} - \frac{7w_1}{30} - \frac{w_3}{15} + \frac{5w_4}{3} + \frac{35w_6}{6}\right) \begin{bmatrix} 4 & -3 & -1 \\ -4 & 3 & 1 \end{bmatrix}, \\
\boldsymbol{A}_{13} &= a_{13} \begin{bmatrix} 1 & 3 & -4 \\ -1 & -3 & 4 \end{bmatrix}, \\
\boldsymbol{A}_{14} &= a_{14} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix},
\end{aligned}
$$

$$\boldsymbol{A}_{22} = \begin{bmatrix} -w_1 - \left(\frac{7w_1}{15} + \frac{8w_3}{15}\right) & w_1 & \frac{7w_1}{15} + \frac{8w_3}{15} \\ w_1 & -w_1 - w_3 & w_3 \\ \frac{7w_1}{15} + \frac{8w_3}{15} & w_3 & -\left(\frac{7w_1}{15} + \frac{8w_3}{15}\right) - w_3 \end{bmatrix},$$

$$\boldsymbol{A}_{23} = \left(-\frac{5a_{13}}{13} - \frac{15a_{34}}{13} - \frac{5w_4}{39} - \frac{35w_6}{78}\right) \begin{bmatrix} -4 & -12 & 16 \\ 3 & 9 & -12 \\ 1 & 3 & -4 \end{bmatrix},$$

$$\boldsymbol{A}_{24} = a_{24} \begin{bmatrix} -4 & 4 \\ 3 & -3 \\ 1 & -1 \end{bmatrix},$$

$$\boldsymbol{A}_{33} = \begin{bmatrix} -w_4 - \left(\frac{8w_4}{15} + \frac{7w_6}{15}\right) & w_4 & \frac{8w_4}{15} + \frac{7w_6}{15} \\ w_4 & -w_4 - w_6 & w_6 \\ \frac{8w_4}{15} + \frac{7w_6}{15} & w_6 & -\left(\frac{8w_4}{15} + \frac{7w_6}{15}\right) - w_6 \end{bmatrix},$$

$$\boldsymbol{A}_{34} = a_{34} \begin{bmatrix} -1 & 1 \\ -3 & 3 \\ 4 & -4 \end{bmatrix},$$

$$\boldsymbol{A}_{44} = \left(\frac{a_{14}}{3} + \frac{13a_{24}}{15} + \frac{13a_{34}}{3}\right) \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

$\boldsymbol{A}_{2OE}$ in (5.5.2.12) becomes

$$\boldsymbol{A}_{2OE} = \frac{1}{C} \begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \boldsymbol{A}_{13} & \boldsymbol{A}_{14} \\ \boldsymbol{A}'_{12} & \boldsymbol{A}_{22} & \boldsymbol{A}_{23} & \boldsymbol{A}_{24} \\ \boldsymbol{A}'_{13} & \boldsymbol{A}'_{23} & \boldsymbol{A}_{33} & \boldsymbol{A}_{34} \\ \boldsymbol{A}'_{14} & \boldsymbol{A}'_{24} & \boldsymbol{A}'_{34} & \boldsymbol{A}_{44} \end{bmatrix}. \tag{5.5.2.17}$$

Now the $\boldsymbol{A}_{2OE}$ matrix in (5.5.2.17) satisfies all the unbiased consitions in (5.5.2.5), i.e. $E(\boldsymbol{y}'\boldsymbol{A}_{2OE}\boldsymbol{y}) = \sigma_2^2$.

Note that there are 8 unknown elements in the $\boldsymbol{A}_2$ matrix in (5.5.2.17), which are $a_{13}, a_{14}, a_{24}, a_{34}$, w_1, w_3, w_4, w_6. To further reduce the number of unknown elements, we impose the following relations for those 8 unknown elements,

$$\begin{cases} a_{13} = \frac{5}{3}a_{24}, \\ a_{14} = 5a_{24}, \\ a_{34} = 5a_{24}, \\ w_1 = -\frac{76}{3}a_{24}, \\ w_4 = -\frac{467}{15}a_{24}, \\ w_3 = -\frac{79}{3}a_{24}. \end{cases} \tag{5.5.2.18}$$

Plugging (5.5.2.18) into $\boldsymbol{A}_{2OE}$ in (5.5.2.17), define $p = \frac{w_6}{a_{24}}$ and

$$C = \frac{42004}{75} + \frac{137p}{5}, \boldsymbol{A}_{11} = \left(\frac{1175}{9} + \frac{91p}{6}\right)\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \boldsymbol{A}_{12} = \left(\frac{325}{9} + \frac{35p}{6}\right)\begin{bmatrix} 4 & -3 & -1 \\ -4 & 3 & 1 \end{bmatrix},$$

$$\boldsymbol{A}_{13} = \frac{5}{3}\begin{bmatrix} 1 & 3 & -4 \\ -1 & -3 & 4 \end{bmatrix}, \boldsymbol{A}_{14} = 5\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \boldsymbol{A}_{22} = \begin{bmatrix} \frac{256}{5} & -\frac{76}{3} & -\frac{388}{15} \\ -\frac{76}{3} & \frac{155}{3} & -\frac{79}{3} \\ -\frac{388}{15} & -\frac{79}{3} & \frac{261}{5} \end{bmatrix},$$

$$\boldsymbol{A}_{23} = \left(-\frac{283}{117} - \frac{35p}{78}\right)\begin{bmatrix} -4 & -12 & 16 \\ 3 & 9 & -12 \\ 1 & 3 & -4 \end{bmatrix}, \boldsymbol{A}_{24} = \begin{bmatrix} -4 & 4 \\ 3 & -3 \\ 1 & -1 \end{bmatrix},$$

$$\boldsymbol{A}_{33} = \begin{bmatrix} \frac{10741}{225} - \frac{7p}{15} & -\frac{467}{15} & -\frac{3736}{225} + \frac{7p}{15} \\ -\frac{467}{15} & \frac{467}{15} - p & p \\ -\frac{3736}{225} + \frac{7p}{15} & p & \frac{3736}{225} - \frac{22p}{15} \end{bmatrix}, \boldsymbol{A}_{34} = 5 \begin{bmatrix} -1 & 1 \\ -3 & 3 \\ 4 & -4 \end{bmatrix}, \boldsymbol{A}_{44} = \frac{121}{5} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

$\boldsymbol{A}_{2OE}$ in (5.5.2.17) becomes

$$\boldsymbol{A}_{2OE} = \frac{1}{C} \begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \boldsymbol{A}_{13} & \boldsymbol{A}_{14} \\ \boldsymbol{A}'_{12} & \boldsymbol{A}_{22} & \boldsymbol{A}_{23} & \boldsymbol{A}_{24} \\ \boldsymbol{A}'_{13} & \boldsymbol{A}'_{23} & \boldsymbol{A}_{33} & \boldsymbol{A}_{34} \\ \boldsymbol{A}'_{14} & \boldsymbol{A}'_{24} & \boldsymbol{A}'_{34} & \boldsymbol{A}_{44} \end{bmatrix}. \tag{5.5.2.19}$$

Now p is the only unknown elements in $\boldsymbol{A}_{2OE}$. The general $\boldsymbol{A}_{2OE}$ satisfy all the unbiased conditions in (5.5.2.5).

Given $\boldsymbol{A}_{2OE}$ in (5.5.2.19) and $\boldsymbol{y}$ in (5.5.1.3), the quadratic estimator of σ_2^2 can expressed as

$$\hat{\sigma}_2^2 = \boldsymbol{y}'\boldsymbol{A}_{2OE}\boldsymbol{y} = \tfrac{50(-1353p+488260)}{13(2055p+42004)}, \tag{5.5.2.20}$$

which is a function of p.

Theorem 5.1. For the general matrix $\boldsymbol{A}_2$ in (5.5.2.19), $\hat{\sigma}_2^2$ is positive if and only if

$$-\frac{42004}{2055} < p < \frac{488260}{1353}. \tag{5.5.2.21}$$

By using the formula $Var(\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y})$ in (5.5.2.6), we have

$$Var(\boldsymbol{y}'\boldsymbol{A}_{2OE}\boldsymbol{y}) = \frac{\sigma_2^4\,(c_1\gamma_2 + c_2)}{d}, \tag{5.5.2.22}$$

where

$$
\begin{aligned}
c_1 &= 46921725p^2 + 786190260p + 4738540304, \\
c_2 &= 180p\,(5329155p + 70204328) + 47644195776, \\
d &= 18\left(4223025p^2 + 172636440p + 1764336016\right).
\end{aligned}
$$

It is worth noting that for each set of observations $\boldsymbol{y}$, one can obtain $\hat{\sigma}_2^2$ using $\boldsymbol{A}_{2OE}$ in (5.5.2.20) and $\hat{\sigma}_{2MoM}^2$ using $\boldsymbol{A}_{2MoM}$ in (5.5.2.10). If equating $\boldsymbol{y}'\boldsymbol{A}_{2OE}\boldsymbol{y}$ to its corresponding MoM estimate, $\hat{\sigma}_{2MoM}^2$, to solve for p, we can always get the same $p = -\frac{92}{15}$ for different $\boldsymbol{y}$. Because $\boldsymbol{A}_{2MoM}$ is dependent on the fixed design matrices $\boldsymbol{X}$ and $\boldsymbol{U}_1$, $\boldsymbol{A}_{2MoM}$ will not change with different $\boldsymbol{y}$, and $\boldsymbol{A}_{2OE}$ is constructed based on the structure of $\boldsymbol{A}_{2MoM}$, so $\boldsymbol{A}$ will not change with $\boldsymbol{y}$ either.

When $p = -\frac{92}{15}$, $\boldsymbol{A}_{2OE} = \boldsymbol{A}_{2MoM}$. We will run simulations to compare the MSE, MAD and AB of $\boldsymbol{A}_{2OE}$ by varying the value of p to see if there exists a set of p values in $\boldsymbol{A}_{2OE}$ that give better estimation performances than that of the MoM estimators.

Under the mixed-effects model in (5.5.1.3), keep $\boldsymbol{x}_i$ the same as in (5.5.1.3), the coefficients of the intercept and the fixed effect are set to be $\beta_0 = 1.03, \beta_1 = 0.19$, the variance components of the random effects are $\sigma_1^2 = 14, \sigma_2^2 = 65$. For each dataset, simulate 4 independent random effects, b_i, from $N(0, \sigma_1^2)$, simulate 10 independent random errors, e_{ij}, from $N(0, \sigma_2^2)$, then the jth simulated observation in the ith SA is $y_{ij} = \boldsymbol{x}_i\boldsymbol{\beta} + b_i + e_{ij}$. 100000 datasets are simulated, each of which has 10 observations. We estimated σ_2^2 using our method and MoM method for each dataset.

Table 5.6 shows the comparisons of the estimation performances of the quadratic estimators for σ_2^2 by using $\boldsymbol{A}_{2OE}$ with p and $\boldsymbol{A}_{2MoM}$,

Table 5.6: Simulations in Estimating σ_2^2

p	MAD	MSE	AB
-6.3	31.714079	1692.397825	64.831730
-6.25	31.689239	1688.415453	64.831031
-6.2	31.674003	1685.751993	64.830336
-6.15	31.667689	1684.380098	64.829646
-6.1	31.669250	1684.272988	64.828961
-6.05	31.677989	1685.404439	64.828281
-6	31.694280	1687.748773	64.827606
-5.95	31.718212	1691.280841	64.826935
$-\frac{92}{15}$ (MoM)	**31.667327**	**1684.205169**	64.829417

From Table 5.6, since $\boldsymbol{A}_{2OE}$ and $\boldsymbol{A}_{2MoM}$ are both within the class of unbiased matrices $\mathscr{A}_2$, the resulted quadratic estimators $\boldsymbol{y}'\boldsymbol{A}_{2OE}\boldsymbol{y}$ and $\boldsymbol{y}'\boldsymbol{A}_{2MoM}\boldsymbol{y}$ are unbiased for σ_2^2. However, the MoM estimates for σ_2^2 has slightly better estimation performances in MSE and MAD compared to the estimates with other values of p in $\boldsymbol{A}_{2OE}$ matrix.

Method 2

Recall in (5.5.2.2),

$$\begin{aligned} \boldsymbol{W} &= [\boldsymbol{X}, \boldsymbol{U}_1^*], \\ \boldsymbol{H} &= \boldsymbol{W}(\boldsymbol{W}'\boldsymbol{W})^{-1}\boldsymbol{W}', \end{aligned} \tag{5.5.2.23}$$

the rank of $\boldsymbol{H}$ is 5.

Let $\boldsymbol{H}_1$ be a matrix consisting the 5 independent columns in $\boldsymbol{I} - \boldsymbol{H}$, then

$$\boldsymbol{H}_1 = \frac{1}{1176}\begin{bmatrix} 563 & 20 & -15 & 25 & 75 \\ -563 & -20 & 15 & -25 & -75 \\ 20 & 768 & -380 & -20 & -60 \\ -15 & -380 & 775 & 15 & 45 \\ -5 & -388 & -395 & 5 & 15 \\ 25 & -20 & 15 & 759 & -467 \\ 75 & -60 & 45 & -467 & 559 \\ -100 & 80 & -60 & -292 & -92 \\ 75 & -60 & 45 & -75 & -225 \\ -75 & 60 & -45 & 75 & 225 \end{bmatrix}, \tag{5.5.2.24}$$

the rank of $\boldsymbol{H}_1$ is 5 and $\boldsymbol{H}_1'\boldsymbol{W} = \boldsymbol{0}$.

Let $\mathscr{Q}$ be a class of 5×5 full rank symmetric matrices, for any matrix $\boldsymbol{Q} \in \mathscr{Q}$, define

$$\boldsymbol{P} = \boldsymbol{H}_1\boldsymbol{Q}, \tag{5.5.2.25}$$

since $\boldsymbol{H}_1'\boldsymbol{W} = \boldsymbol{0}$, then $\boldsymbol{P}'\boldsymbol{W} = \boldsymbol{0}$.

By using $\boldsymbol{P}$ in (5.5.2.25), our proposed $\boldsymbol{A}_{2OE}$ matrix that satisfies the unbiasedness conditions in (5.5.2.5) is

$$\boldsymbol{A}_{2OE} = \frac{\boldsymbol{P}\boldsymbol{P}'}{tr(\boldsymbol{P}\boldsymbol{P}')}. \tag{5.5.2.26}$$

Consider a $\boldsymbol{Q}$ matrix as

$$\boldsymbol{Q} = \begin{bmatrix} a & b & b & b & b \\ b & a & b & b & b \\ b & b & a & b & b \\ b & b & b & a & b \\ b & b & b & b & a \end{bmatrix}, \tag{5.5.2.27}$$

we have

$$\begin{aligned} \boldsymbol{A}_{2OE} &= \frac{\boldsymbol{P}\boldsymbol{P}'}{tr(\boldsymbol{P}\boldsymbol{P}')} \\ &= \frac{\boldsymbol{H}_1\boldsymbol{Q}\boldsymbol{Q}'\boldsymbol{H}_1'}{tr(\boldsymbol{H}_1\boldsymbol{Q}\boldsymbol{Q}'\boldsymbol{H}_1')}, \end{aligned} \tag{5.5.2.28}$$

the variance for $\boldsymbol{y}'\boldsymbol{A}_{2OE}\boldsymbol{y}$ in (5.5.2.28) is

$$Var(\boldsymbol{y}'\boldsymbol{A}_{2OE}\boldsymbol{y}) = \sigma_2^4\frac{(c_1\gamma_2 + c_2)}{c_3}, \tag{5.5.2.29}$$

where

$$\begin{aligned} c_1 &= 9205207903a^4 - 24064788374a^3b + 62262239567a^2b^2 - 22105876784ab^3 + 585095553738b^4, \\ c_2 &= 34004008836a^4 - 98509034736a^3b + 341287532712a^2b^2 + 257020903440ab^3 \\ &\quad +441302714748b^4, \\ c_3 &= 86436(732736a^4 - 1304544a^3b + 4485716a^2b^2 - 3476244ab^3 + 5202961b^4). \end{aligned}$$

Define $d = \frac{a}{b}$, dividing the numerator and denomenator of (5.5.2.29) by b^4, (5.5.2.29) becomes

$$Var(\boldsymbol{y}'\boldsymbol{A}_{2OE}\boldsymbol{y}) = \sigma_2^4 \frac{c_1\gamma_2 + c_2}{c_3}, \tag{5.5.2.30}$$

where

$$\begin{aligned}
c_1 &= 58509553738 - 22105876784d + 62262239567d^2 - 24064788374d^3 + 9205207903d^4, \\
c_2 &= 441302714748 + 257020903440d + 341287532712d^2 - 98509034736d^3 \\
&\quad +34004008836d^4, \\
c_3 &= 86436(5202961 - 3476244d + 4485716d^2 - 1304544d^3 + 732736d^4).
\end{aligned}$$

Notice the coefficients in $Var(\boldsymbol{y}'\boldsymbol{A}_{2OE}\boldsymbol{y})$ is a function of d, and recall the coefficients with and without γ_2 in the variance of the $\boldsymbol{y}'\boldsymbol{A}_{2MoM}\boldsymbol{y}$ is 0.1080867 and 0.4, respectively. By changing the value of d in (5.5.2.30), the comparisons of the coefficients with and without γ_2 are displayed in Figure 5.1 and Figure 5.2, respectively,

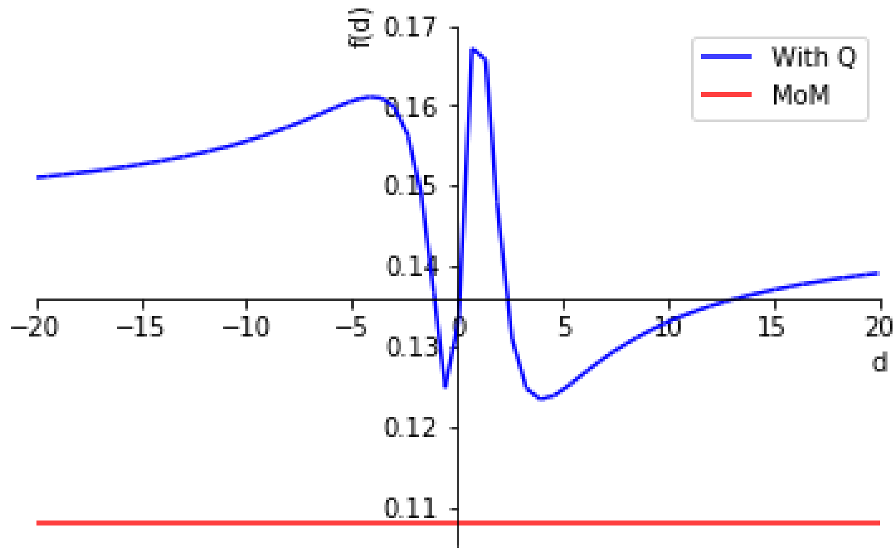

Figure 5.1: Comparison of the Coefficients in Variances with Kurtosis

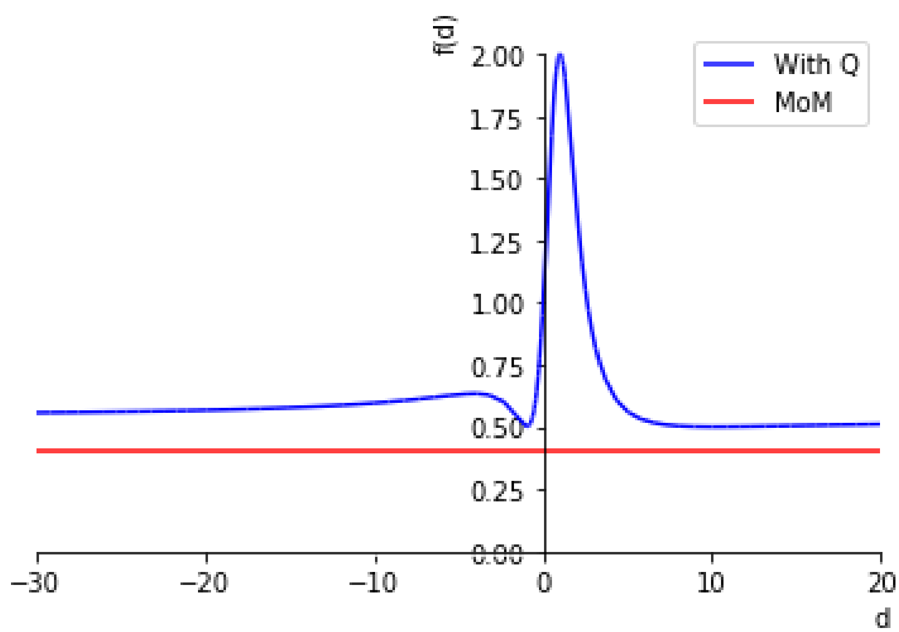

Figure 5.2: Comparison of the Coefficients in Variances without Kurtosis

Both coefficients, $\frac{c_1}{d}, \frac{c_2}{d}$ in (5.5.2.30), are greater than that of the corresponding coefficients in variances for MoM estimators for $d \in (-20, 20)$. (d is not equal to 4 and d not equal to 1, otherwise $\boldsymbol{Q}$ will not be full rank matrix). The $\boldsymbol{Q}$ in (5.5.2.27) matrix is special class in the class of full rank 5×5 symmetric matrix $\mathscr{Q}$, the coefficients in $Var(\boldsymbol{y}'\boldsymbol{A}_{2OE}\boldsymbol{y})$ based on those $\boldsymbol{Q}$ is not smaller than the corresponding coefficients in $Var(\boldsymbol{y}'\boldsymbol{A}_{2MoM}\boldsymbol{y})$ over the tested range of d. However, other classes of $\boldsymbol{Q}$ in $\mathscr{Q}$ may be considered to construct a matrix $\boldsymbol{A}_2$ with $Var(\boldsymbol{y}'\boldsymbol{A}_{2OE}\boldsymbol{y})$ smaller than that of the MoM estimators.

Method 3

Recall in (5.5.2.2),

$$\begin{aligned} \boldsymbol{W} &= [\boldsymbol{X}, \boldsymbol{U}_1^*], \\ \boldsymbol{H} &= \boldsymbol{W}(\boldsymbol{W}'\boldsymbol{W})^{-1}\boldsymbol{W}', \end{aligned} \tag{5.5.2.31}$$

the rank of $\boldsymbol{H}$ is 5.

Let $\boldsymbol{H}_1$ be a matrix consisting the 5 independent columns in $\boldsymbol{I} - \boldsymbol{H}$, then

$$\boldsymbol{H}_1 = \frac{1}{1176}\begin{bmatrix} 563 & 20 & -15 & 25 & 75 \\ -563 & -20 & 15 & -25 & -75 \\ 20 & 768 & -380 & -20 & -60 \\ -15 & -380 & 775 & 15 & 45 \\ -5 & -388 & -395 & 5 & 15 \\ 25 & -20 & 15 & 759 & -467 \\ 75 & -60 & 45 & -467 & 559 \\ -100 & 80 & -60 & -292 & -92 \\ 75 & -60 & 45 & -75 & -225 \\ -75 & 60 & -45 & 75 & 225 \end{bmatrix}. \qquad (5.5.2.32)$$

the rank of $\boldsymbol{H}_1$ is 5.

Let $\boldsymbol{H}_2$ be a matrix consisting the remaining 5 columns not included in $\boldsymbol{H}_1$,

$$
\boldsymbol{H}_2 = \frac{1}{1176}\begin{bmatrix}
-563 & -5 & -100 & 75 & -75 \\
563 & 5 & 100 & -75 & 75 \\
-20 & -388 & 80 & -60 & 60 \\
15 & -395 & -60 & 45 & -45 \\
5 & 783 & -20 & 15 & -15 \\
-25 & 5 & -292 & -75 & 75 \\
-75 & 15 & -92 & -225 & 225 \\
100 & -20 & 384 & 300 & -300 \\
-75 & 15 & 300 & 363 & -363 \\
75 & -15 & -300 & -363 & 363
\end{bmatrix}. \tag{5.5.2.33}
$$

Consider a $\boldsymbol{Q}$ matrix as

$$
\boldsymbol{Q} = \begin{bmatrix}
a & 0 & 0 & 0 & 0 \\
0 & a & 0 & 0 & 0 \\
0 & 0 & a & 0 & 0 \\
0 & 0 & 0 & a & 0 \\
0 & 0 & 0 & 0 & a
\end{bmatrix}. \tag{5.5.2.34}
$$

Define a $\boldsymbol{P}$ matrix as

$$
\boldsymbol{P} \quad = \quad [\boldsymbol{H}_1, \boldsymbol{H}_2\boldsymbol{Q}] \tag{5.5.2.35}
$$

It can be shown that $\boldsymbol{P}'\boldsymbol{W} = \boldsymbol{0}, tr(\boldsymbol{P}\boldsymbol{P}') = \frac{307a^2}{147} + \frac{428}{147}$.

By using $\boldsymbol{P}$ in (5.5.2.35) , our proposed $\boldsymbol{A}_{2OE}$ matrix in the unbiased quadratic estimatior

$\boldsymbol{y}'\boldsymbol{A}_{2OE}\boldsymbol{y}$ for estimating σ_2^2 is

$$\boldsymbol{A}_{2OE} = \frac{\boldsymbol{P}\boldsymbol{P}'}{tr(\boldsymbol{P}\boldsymbol{P}')}. \tag{5.5.2.36}$$

It can be shown that $\boldsymbol{A}_{2OE}$ in (5.5.2.36) satisfy all the unbiased conditions in (5.5.2.5), i.e. $\boldsymbol{y}'\boldsymbol{A}_{2OE}\boldsymbol{y}$ is an unbiased estimator for σ_2^2.

Using the formula $Var(\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y})$ in (5.5.2.6), we have

$$Var(\boldsymbol{y}'\boldsymbol{A}_{2OE}\boldsymbol{y}) = \sigma_2^4 \frac{c_1\gamma_2 + c_2}{c_3}, \tag{5.5.2.37}$$

where

$$\begin{aligned}
c_1 &= 4120975645a^4 + 6862197574a^2 + 9205207903, \\
c_2 &= 21704511780a^4 + 19003300344a^2 + 34004008836, \\
c_3 &= 345744(94249a^4 + 262792a^2 + 183184).
\end{aligned}$$

Notice that the coefficients in $Var(\boldsymbol{y}'\boldsymbol{A}_{2OE}\boldsymbol{y})$ are functions of a, and recall the coefficients with and without γ_2 in the variance of $\boldsymbol{y}'\boldsymbol{A}_{2MoM}\boldsymbol{y}$ is 0.1080867 and 0.4, respectively. By changing the value of a in (5.5.2.37), the comparisons of the coefficients with and without γ_2 are displayed in Figure 5.3 and Figure 5.4, respectively.

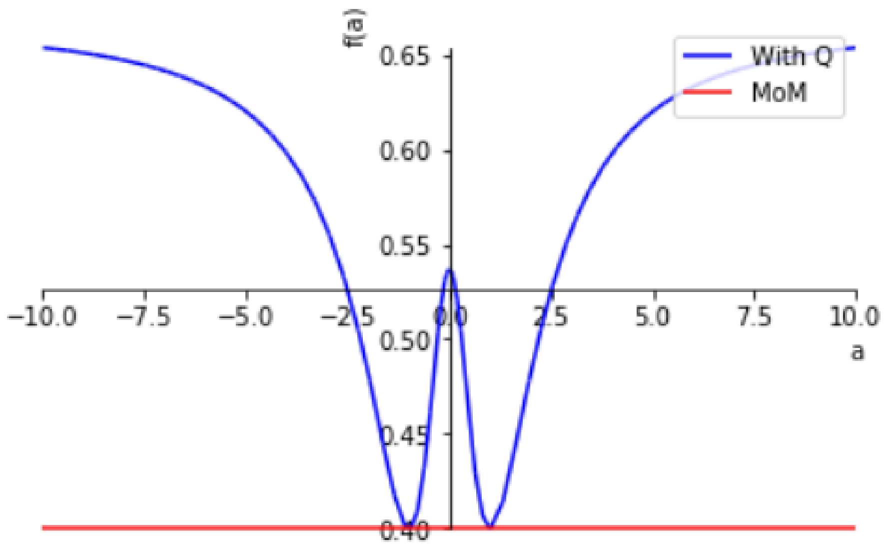

Figure 5.3: Comparison of the Coefficients in Variances without Kurtosis

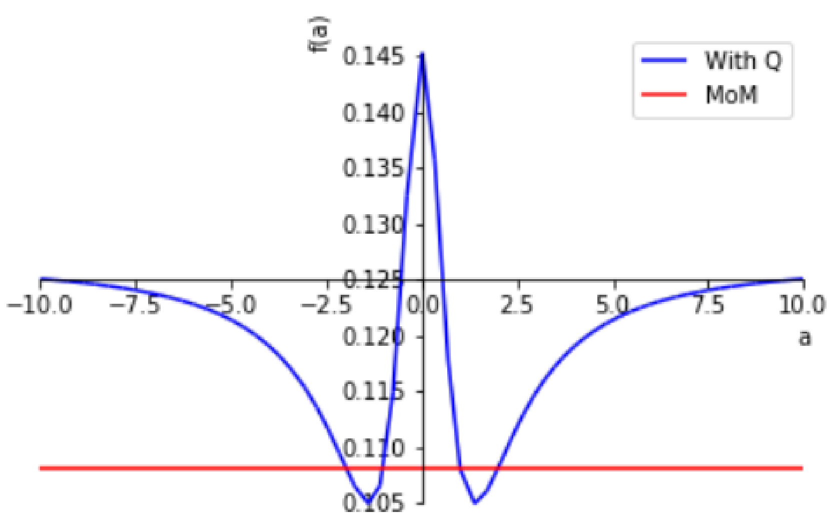

Figure 5.4: Comparison of the Coefficients in Variances with Kurtosis

We compare $\sigma_2^{-4}Var(\boldsymbol{y}'\boldsymbol{A}_{2OE}\boldsymbol{y})$ and $\sigma_2^{-4}Var(\boldsymbol{y}'\boldsymbol{A}_{2MoM}\boldsymbol{y})$ by varying the value of a for $\gamma_2 = 3$, which is shown in Figure 5.5,

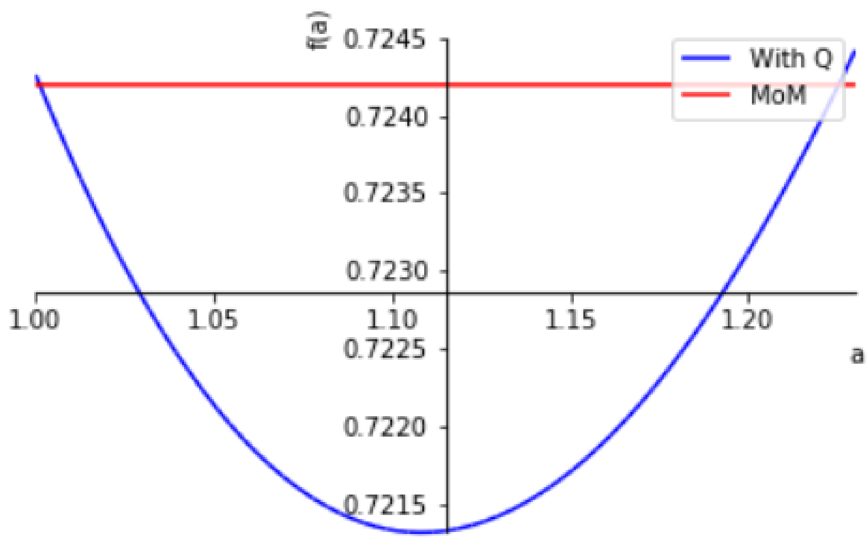

Figure 5.5: Comparison of the Variances between OE and MoM

From Figure 5.5, when $a \in (1.22539257845605, 1.00104933271506)$, the variance of our estimator is smaller than that of the MoM estimator. Since $Var(\boldsymbol{y}'\boldsymbol{A}_{2OE}\boldsymbol{y})$ in (5.5.2.37) is a symmetric function, when $a \in (-1.22539257845605, -1.00104933271506)$, the values of $Var(\boldsymbol{y}'\boldsymbol{A}_{2OE}\boldsymbol{y})$ will be the same as that of the corresponding positive value of a. When $a = 1.1082739093703766$, $\gamma_2 = 3$, $\sigma_2^{-4}Var(\boldsymbol{y}'\boldsymbol{A}_{2OE}\boldsymbol{y}) = 0.721313$, which reaches the minimum in Figure 5.5, (recall $\sigma_2^{-4}Var(\boldsymbol{y}'\boldsymbol{A}_{2MoM}\boldsymbol{y}) = 0.724201$ for $\gamma_2 = 3$).

Figure 5.6 shows the comparison of $\sigma_2^{-4}Var(\boldsymbol{y}'\boldsymbol{A}_{2OE}\boldsymbol{y})$ and $\sigma_2^{-4}Var(\boldsymbol{y}'\boldsymbol{A}_{2MoM}\boldsymbol{y})$ by changing the value of $\gamma_2 \in (2.96, 3.04)$ and the value of a,

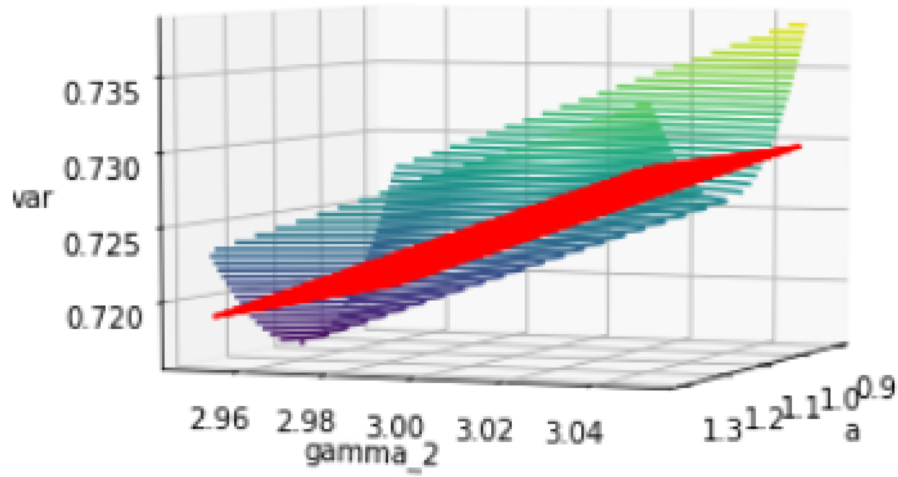

Figure 5.6: Comparison of the Coefficients in Variances of OE and MoM Estimators

For $\gamma_2 \in (2.96, 3.04)$, some part of $\sigma_2^{-4} Var(\boldsymbol{y}'\boldsymbol{A}_{2OE}\boldsymbol{y})$, which is the blue curved plain, are always smaller than the corresponding $\sigma_2^{-4} Var(\boldsymbol{y}'\boldsymbol{A}_{2MoM}\boldsymbol{y})$.

Figure 5.7 shows the comparison of the coefficients in variances of our estimators at the value of a giving the minimum coefficients in variances for our estimators ($\min_{a=a_{min}} \sigma_2^{-4} Var(\boldsymbol{y}'\boldsymbol{A}_{2OE}\boldsymbol{y})$) and MoM estimates ($\sigma_2^{-4} Var(\boldsymbol{y}'\boldsymbol{A}_{2MoM}\boldsymbol{y})$) for $2.9 < \gamma_2 < 3.1$,

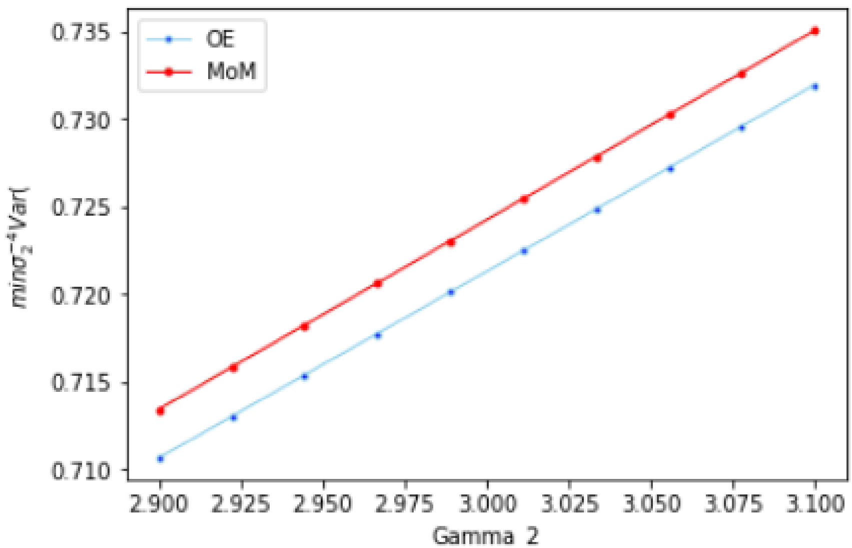

Figure 5.7: Comparison of the Coefficients in Variances of OE and MoM Estimators at a_{min} and different γ_2

For $2.9 < \gamma_2 < 3.1$, $\min_{a=a_{min}} \sigma_2^{-4} Var(\boldsymbol{y}'\boldsymbol{A}_{2OE}\boldsymbol{y})$ for our estimators are always smaller than $\sigma_2^{-4} Var(\boldsymbol{y}'\boldsymbol{A}_{2MoM}\boldsymbol{y})$ of the corresponding MoM estimators.

For $\gamma_2 > 0$, the following table shows the comparison of $\min_{a=a_{min}} \sigma_2^{-4} Var(\boldsymbol{y}'\boldsymbol{A}_{2OE}\boldsymbol{y})$ for our estimators and $\sigma_2^{-4} Var(\boldsymbol{y}'\boldsymbol{A}_{2MoM}\boldsymbol{y})$ for MoM estimators,

Table 5.7: Comparison of the Coefficients in Variances of OE and MoM Estimators at different γ_2

γ_2	Range of a	a_{min}	OE	MoM	Differences
0	(1, 1)	1	0.4	0.4	0
1	(1, 1.0909)	1.0454	0.5077	0.5081	0.0004
2	(1, 1.1649)	1.0824	0.6147	0.6162	0.0015
3	(1, 1.2267)	1.1083	0.7213	0.7243	0.0030
4	(1, 1.2792)	1.1396	0.8276	0.8323	0.0047
5	(1, 1.3247)	1.1623	0.9337	0.9404	0.0067

It can be seen that our estimators have smaller coefficients in variances than the corresponding MoM estimators for $\gamma_2 \in (0, 5]$, and as γ_2 increases, the differences in the coefficients in variances become larger.

Compared with ML and REML, which do not provide closed-form expression of the estimates, we can find the approximated $\boldsymbol{A}_2$ matrix for ML and REML estimates for σ_2^2 using our proposed estimators.

Table 5.8: Approximating the $\boldsymbol{A}_2$ Matrices for REML and ML Estimates for σ_2^2 at $\gamma_2 = 3$

(a, OE)	REML	ML
(± 1.0095, 64.8106)	64.8106	
(±1.4377, 59.1236)		59.1236

With a in $\boldsymbol{A}_{2OE}$ equal to ± 1.0095 and ± 1.4377, the estimates closest to REMLE and MLE, respectively, can be found.

Table 5.9: Comparison of the Estimated Variances with the Approximated $\boldsymbol{A}_2$ matrix for $\gamma_2 = 3$

	a	Estimate	σ_2^{-4}Var(Estimate)	$\hat{Var}$(Estimate)
OE	±1.1082	63.3127	0.7213	2891.3295
MoM		64.9605	0.7243	3056.4493
ML^*	±1.4377	59.12356	0.7400	2586.5770
$REML^*$	±1.0095	64.81062	0.7237	3040.0462

where ML^* and $REML^*$ are the approximated ML and REML estimates using $\boldsymbol{A}_{2OE}$ matrix. In Table 5.9, the estimated variance for ML^* are smaller than that of $REML^*$.

Consider a more general $\boldsymbol{Q}$ matrix, that is

$$\boldsymbol{Q} = \begin{bmatrix} a & b & b & b & b \\ b & a & b & b & b \\ b & b & a & b & b \\ b & b & b & a & b \\ b & b & b & b & a \end{bmatrix}. \tag{5.5.2.38}$$

Using $\boldsymbol{Q}$ in (5.5.2.38), we can obtain $\boldsymbol{P}$ from (5.5.2.35) and $\boldsymbol{A}_{2OE}$ from (5.5.2.36), the resulted variance of $\boldsymbol{y}'\boldsymbol{A}_{2OE}\boldsymbol{y}$ is

$$Var(\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}) = \sigma_2^4 \tfrac{c_1\gamma_2 + c_2}{c_3}, \tag{5.5.2.39}$$

where

$$\begin{aligned} c_1 &= 4120975645a^4 - 3363619902a^3b + 31574726299a^2b^2 + 6862197574a^2 \\ &\quad -2497426632ab^3 - 968351910ab + 53971680640b^4 + 25996262431b^2 \\ &\quad +9205207903, \\ c_2 &= 21704511780a^4 - 27758748528a^3b + 256714401384a^2b^2 + 19003300344a^2 \\ &\quad +263114295696ab^3 - 499600080ab + 4613316664668b^4 + 75263801256b^2 \\ &\quad +34004008836, \\ c_3 &= 86436(376996a^4 - 341384a^3b + 2581176a^2b^2 + 1051168a^2 - 1133684ab^3 \\ &\quad -475936ab + 4157521b^4 + 3490768b^2 + 732736. \end{aligned}$$

By fixing $a = 1.1082739093703766$ (the a value giving the minimum variance compared to that of MoM estimators for $\boldsymbol{Q} = diag(a)$), $\gamma_2 = 3$, b is the only unknown elements in $\sigma_2^{-4}Var(\boldsymbol{y}'\boldsymbol{A}_{2OE}\boldsymbol{y})$.

Figure 5.8 shows the comparison of $\sigma_2^{-4}Var(\boldsymbol{y}'\boldsymbol{A}_{2OE}\boldsymbol{y})$ and $\sigma_2^{-4}Var(\boldsymbol{y}'\boldsymbol{A}_{2MoM}\boldsymbol{y})$ by varying the value of b in $\sigma_2^{-4}Var(\boldsymbol{y}'\boldsymbol{A}_{2OE}\boldsymbol{y})$ at $a = 1.1082739093703766, \gamma_2 = 3$,

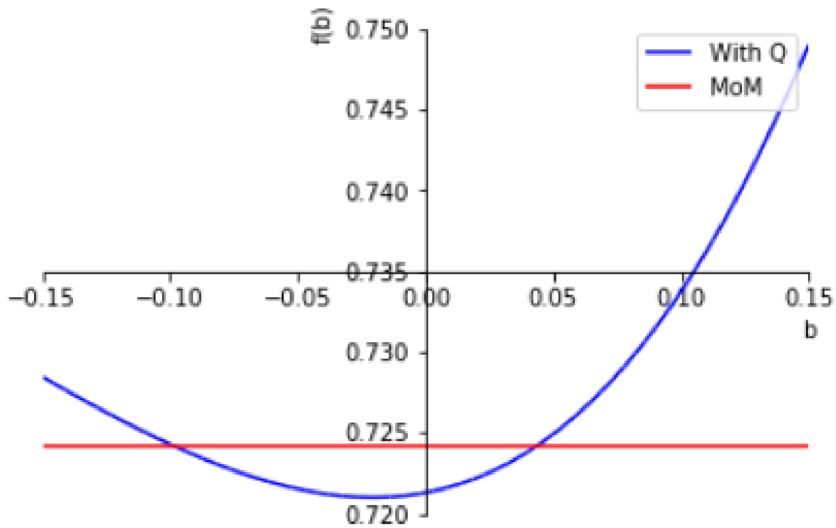

Figure 5.8: Comparison of the Coefficients in Variances between OE and MoM Estimator

For $a = 1.1082739093703766, \gamma_2 = 3$, when $b = -0.0205697543545269$, $\sigma_2^{-4}Var(\boldsymbol{y}'\boldsymbol{A}_{2OE}\boldsymbol{y}) = 0.721019908439413$, which reaches the minimum in Figure 5.8 (recall $\sigma_2^{-4}Var(\boldsymbol{y}'\boldsymbol{A}_{2MoM}\boldsymbol{y}) = 0.724201$ for $\gamma_2 = 3$).

5.5.3 Search for the Optimum Unbiased Quadratic Estimators for $\sigma_1^2 + p_2\sigma_2^2$

In order to estimate σ_1^2, we first obtain $\hat{\sigma}_2^2$, which is introduced in Section 5.5.2, then we need to estimate $\sigma_1^2 + p_2\sigma_2^2$, by subtracting $p_2\hat{\sigma}_2^2$ from $\widehat{\sigma_1^2 + p_2\sigma_2^2}$, the estimate for σ_1^2 can be obtained.

Let $\boldsymbol{y}'\boldsymbol{A}_{12}\boldsymbol{y}$ be an estimator of $\sigma_1^2 + p_2\sigma_2^2$, where $\boldsymbol{A}_{12}$ is a 10×10 matrix. To have an unbiased estimatr for σ_1^2, the $\boldsymbol{A}_{12}$ in the estimator $\boldsymbol{y}'\boldsymbol{A}_{12}\boldsymbol{y}$ will satisfy the conditions for unbiasedness

- $\boldsymbol{A}_{12}$ is a symmetric matrix, that is

$$\boldsymbol{A}_{12} = \boldsymbol{A}'_{12}. \tag{5.5.3.1}$$

- For $\boldsymbol{X}$ in (5.5.1.3),

$$\boldsymbol{A}_{12}\boldsymbol{X} = \boldsymbol{0}. \tag{5.5.3.2}$$

- The trace of $\boldsymbol{A}_1$ is equal to 0,

$$tr(\boldsymbol{A}_{12}) = p_2. \tag{5.5.3.3}$$

- For $\boldsymbol{V}_1$ in (5.5.1.4), the trace of $\boldsymbol{A}_1\boldsymbol{V}_1$ is equal to 1,

$$tr(\boldsymbol{A}_{12}\boldsymbol{V}_1) = 1. \tag{5.5.3.4}$$

Consider a class $\mathscr{A}_{12}$ of matrices satisfying the conditions for unbiasedness as

$$\mathscr{A}_{12} = \left\{\boldsymbol{A}_{12}|\boldsymbol{A}_{12} = \boldsymbol{A}'_{12}, \boldsymbol{A}_{12}\boldsymbol{X} = \boldsymbol{0}, tr(\boldsymbol{A}_{12}) = p_2, tr(\boldsymbol{A}_{12}\boldsymbol{V}_1) = 1\right\}. \tag{5.5.3.5}$$

For any $\boldsymbol{A}_{12} \in \mathscr{A}_{12}$, $\boldsymbol{y}'\boldsymbol{A}_{12}\boldsymbol{y}$ is an unbiased estimator of $\sigma_1^2 + p_2\sigma_2^2$, i.e. $E(\boldsymbol{y}'\boldsymbol{A}_{12}\boldsymbol{y}) = \sigma_1^2 + p_2\sigma_2^2$.

Since the $\boldsymbol{A}_{12}$ matrix in the MoM estimator for $\sigma_1^2+p_2\sigma_2^2$, denoted by $\boldsymbol{A}_{12MoM}$, is a member in the class $\mathscr{A}_{12}$ matrices, we will first present $\boldsymbol{A}_{12MoM}$ matrix[26].

For $\boldsymbol{X}$ in (5.5.1.3), define $\boldsymbol{H} = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'$, the rank of $\boldsymbol{H}$ is 2.

$$\boldsymbol{A}_{12MoM} = \frac{\boldsymbol{I} - \boldsymbol{H}}{tr((\boldsymbol{I} - \boldsymbol{H})\boldsymbol{U}_1\boldsymbol{U}'_1)}, \tag{5.5.3.6}$$

$\boldsymbol{A}_{12MoM}$ is a member of the unbiased class of $\mathscr{A}_{12}$.

Using the formula $Var(\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y})$ in (5.5.2.6), define $r = \frac{\sigma_1}{\sigma_2}$

$$Var(\boldsymbol{y}'\boldsymbol{A}_{12MoM}\boldsymbol{y}) = \left(1.7685r^4 + 0.7737r^2 + 1.3250\right)\sigma_2^4. \tag{5.5.3.7}$$

Let $\boldsymbol{H}_1$ be a matrix consisting the 8 independent columns in $\boldsymbol{I} - \boldsymbol{H}$, then

$$\boldsymbol{H}_1 = \frac{1}{92408}\begin{bmatrix} 74431 & -15887 & -16932 & -18395 & -17977 & -3347 & -1257 & -8572 \\ -15887 & 78111 & -15092 & -16205 & -15887 & -4757 & -3167 & -8732 \\ -16932 & -15092 & 76396 & -17300 & -16932 & -4052 & -2212 & -8652 \\ -18395 & -16205 & -17300 & 73575 & -18395 & -3065 & -875 & -8540 \\ -17977 & -15887 & -16932 & -18395 & 74431 & -3347 & -1257 & -8572 \\ -3347 & -4757 & -4052 & -3065 & -3347 & 79191 & -14627 & -9692 \\ -1257 & -3167 & -2212 & -875 & -1257 & -14627 & 75871 & -9852 \\ -8572 & -8732 & -8652 & -8540 & -8572 & -9692 & -9852 & 83116 \\ 7103 & 3193 & 5148 & 7885 & 7103 & -20267 & -24177 & -10492 \\ 833 & -1577 & -372 & 1315 & 833 & -16037 & -18447 & -10012 \end{bmatrix}, \tag{5.5.3.8}$$

the rank of $\boldsymbol{H}_1$ is 8.

Let $\boldsymbol{H}_2$ be a matrix consisting the remaining 2 columns not included in $\boldsymbol{H}_1$,

$$\boldsymbol{H}_2 = \frac{1}{92408}\begin{bmatrix} 7103 & 833 \\ 3193 & -1577 \\ 5148 & -372 \\ 7885 & 1315 \\ 7103 & 833 \\ -20267 & -16037 \\ -24177 & -18447 \\ -10492 & -10012 \\ 52591 & -28087 \\ -28087 & 71551 \end{bmatrix}. \tag{5.5.3.9}$$

Consider a $\boldsymbol{Q}$ matrix as

$$\boldsymbol{Q} = \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}. \tag{5.5.3.10}$$

Define a $\boldsymbol{P}$ matrix as

$$\boldsymbol{P} = [\boldsymbol{H}_1, \boldsymbol{H}_2\boldsymbol{Q}]. \tag{5.5.3.11}$$

By using $\boldsymbol{P}$ in (5.5.3.11) , our proposed quadratic estimatior for $\sigma_1^2 + p_2\sigma_2^2$ is

$$\boldsymbol{A}_{12OE} = \frac{\boldsymbol{P}\boldsymbol{P}'}{tr(\boldsymbol{P}\boldsymbol{P}'\boldsymbol{U}_1\boldsymbol{U}_1')}. \tag{5.5.3.12}$$

It can be shown that $\boldsymbol{A}_{12OE}$ satisfy all the unbiased conditions in (5.5.3.5), i.e. $E(\boldsymbol{y}'\boldsymbol{A}_{12OE}\boldsymbol{y}) = \sigma_1^2 + p_2\sigma_2^2$.

Using the formula $Var(\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y})$ in (5.5.2.6) for $\gamma_1 = 3, \gamma_2 = 3$,

$$Var(\boldsymbol{y}'\boldsymbol{A}_{12OE}\boldsymbol{y}) = \left(\frac{c_1 + c_2 + c_3}{c_4}\right)\sigma_2^4, \tag{5.5.3.13}$$

where

$$\begin{aligned}
c_1 &= 6.2762 \cdot 10^{14}a^4r^4 + 3.9297 \cdot 10^{14}a^4r^2 + 9.2930 \cdot 10^{14}a^4 \\
c_2 &= 2.6074 \cdot 10^{15}a^2r^4 + 9.8917 \cdot 10^{14}a^2r^2 + 6.0184 \cdot 10^{14}a^2 \\
c_3 &= 7.1527 \cdot 10^{15}r^4 + 3.2264 \cdot 10^{15}r^2 + 6.2516 \cdot 10^{15} \\
c_4 &= 1.9303 \cdot 10^{14}a^4 + 1.7436 \cdot 10^{15}a^2 + 3.9372 \cdot 10^{15}
\end{aligned}$$

Note that $\sigma_2^{-4}Var(\boldsymbol{y}'\boldsymbol{A}_{12MoM}\boldsymbol{y})$ is a function of r and $\sigma_2^{-4}Var(\boldsymbol{y}'\boldsymbol{A}_{12OE}\boldsymbol{y})$ is a function of a and r. Figure 5.9 shows the comparison of $\sigma_2^{-4}Var(\boldsymbol{y}'\boldsymbol{A}_{12OE}\boldsymbol{y})$ and $\sigma_2^{-4}Var(\boldsymbol{y}'\boldsymbol{A}_{12MoM}\boldsymbol{y})$ for $a \in (0.7, 1.3)$, $r \in (0.1, 0.3)$, $\gamma_1 = 3, \gamma_2 = 3$

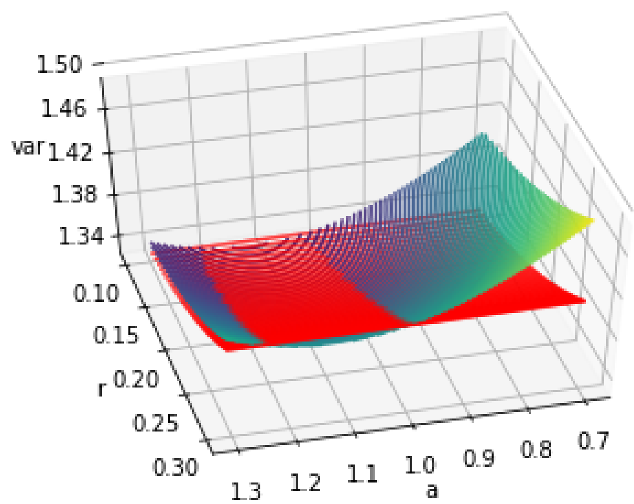

Figure 5.9: Comparison of the Coefficients in Variances of OE and MoM Estimators for $\sigma_1^2 + p_2\sigma_2^2$

At ratio $\in (0.1, 0.3)$, there always exists a set of a yielding $\sigma_2^{-4} Var(\boldsymbol{y}'\boldsymbol{A}_{12OE}\boldsymbol{y})$ smaller than $\sigma_2^{-4} Var(\boldsymbol{y}'\boldsymbol{A}_{12MoM}\boldsymbol{y})$.

Table 5.10 shows the comparison of $\min\limits_{a=a_{min}} \sigma_2^{-4} Var(\boldsymbol{y}'\boldsymbol{A}_{12OE}\boldsymbol{y})$ and $\sigma_2^{-4} Var(\boldsymbol{y}'\boldsymbol{A}_{12MoM}\boldsymbol{y})$ at different ratios (r) for $\gamma_1 = 3, \gamma_2 = 3$,

Table 5.10: Comparison of the Minimum Coefficients in Variances at Different Ratios for $\gamma_1 = 3, \gamma_2 = 3$

Ratio (r)	Range of a	a_{min}	OE	MoM
0.15	(1, 1.1180)	1.059	1.3358	1.3432
0.21	(1, 1.1167)	1.058	1.3551	1.3625
1.00	(1, 1.0406)	1.0203	3.8647	3.8672
5.00	(0.8911, 1)	0.9123	1125.0793	1125.9606

From Table 5.10, it can be seen that at ratio = (0.15, 0.21, 1, 5), the minimum coefficients of our estimators, $\min\limits_{a=a_{min}} \sigma_2^{-4} Var(\boldsymbol{y}'\boldsymbol{A}_{12OE}\boldsymbol{y})$, are always smaller than the coefficients of MoM estimators $\sigma_2^{-4} Var(\boldsymbol{y}'\boldsymbol{A}_{12MoM}\boldsymbol{y})$ at each ratio.

Chapter 6

Conclusions

This dissertation aims at searching for the optimum variance components estimates in the mixed-effects model. Methods are illustrated in DOE and SAE.

When the optimum estimates do not exist, we propose the near optimum estimators for the variance components in DOE. The simulation results show that under the normality assumptions for both variance components, NOPE for estimating σ_1^2 and σ_2^2 are better under the criterion function AB than both MLE and REMLE, while MLE and REMLE are better than NOPE under the criterion function RMSE and MAD across 4 sets of simulated variance components. Under the departure from normality by using the skew normal distrubution, NOPE performs better in estimating σ_2^2 than that of MLEs under all three criterion functions for all skewedness. NOPE also has smaller MSE than MoM estimators for all four skewness for both σ_1^2 and σ_2^2.

When the experimental design is replicated, we propose AOPE to estimate the variance components for the full data. Simulation results show that when considering all positive estimates, AOPE has smallest AB for σ_2^2, AOPE has relatively larger RMSE and MAD for both σ_1^2 and σ_2^2 compared to REMLE, but has smaller RMSE and MAD compared to MLE. To estimate the variance components in SAE, our proposed estimators for the variance components are unbiased and have smaller variances than the corresponding MoM estimators for both σ_2^2 and $\sigma_1^2 + p_2\sigma_2^2$ at tested range of γ_1, γ_2 and ratio r.

All of our proposed estimators have closed-form expressions and do not require the functional form in the distributional assumptions.

References

George E Battese and Wayne A Fuller. Prediction of county crop areas using survey and satellite data. In *Proceedings of the section on survey research methods, American Statistical Association*, volume 5, page 5. Am. Stat. Assoc Alexandria VA, 1981.

WG Cochran. The use of the analysis of variance in enumeration by sampling. *Journal of the American Statistical Association*, 34(207):492–510, 1939.

Gauri S Datta and Partha Lahiri. A unified measure of uncertainty of estimated best linear unbiased predictors in small area estimation problems. *Statistica Sinica*, pages 613–627, 2000.

Robert E Fay III and Roger A Herriot. Estimates of income for small places: an application of james-stein procedures to census data. *Journal of the American Statistical Association*, 74(366a):269–277, 1979.

Subir Ghosh, Li Guo, and Luyao Peng. Variance component estimators ope, nope and aope in linear mixed effects models. *Australian & New Zealand Journal of Statistics*, 60(4):481–505, 2018.

M.E. Gonzalez. Use and estimation of synthetic estimators. *In proceedings of the social statistics section*, pages 33–36, 1973.

Franklin A Graybill et al. On quadratic estimates of variance components. *The Annals of Mathematical Statistics*, 25(2):367–372, 1954.

Franklin A Graybill and AW Wortham. A note on uniformly best unbiased estimators for variance components. *Journal of the American Statistical Association*, 51(274):266–268, 1956.

Herman O Hartley and Jon NK Rao. Maximum-likelihood estimation for the mixed analysis of variance model. *Biometrika*, 54(1-2):93–108, 1967.

David A Harville. Maximum likelihood approaches to variance component estimation and to related problems. *Journal of the American Statistical Association*, 72(358):320–338, 1977.

David Heckerman, Deepti Gurdasani, Carl Kadie, Cristina Pomilla, Tommy Carstensen, Hilary Martin, Kenneth Ekoru, Rebecca N Nsubuga, Gerald Ssenyomo, Anatoli Kamali, et al. Linear mixed model for heritability estimation that explicitly addresses environmental variation. *Proceedings of the National Academy of Sciences*, 113(27):7377–7382, 2016.

Charles R Henderson. Estimation of variance and covariance components. *Biometrics*, 9(2):226–252, 1953.

Charles R Henderson. Best linear unbiased estimation and prediction under a selection model. *Biometrics*, pages 423–447, 1975.

Jiming Jiang and Partha Lahiri. Mixed model prediction and small area estimation. *Test*, 15(1):1, 2006.

Raghu N Kackar and David A Harville. Approximations for standard errors of estimators of fixed and random effects in mixed linear models. *Journal of the American Statistical Association*, 79(388):853–862, 1984.

Lynn Roy Lamotte. On non-negative quadratic unbiased estimation of variance components. *Journal of the American Statistical Association*, 68(343):728–730, 1973.

Jane. L. Meza and P Lahiri. Estimation of lifetime alcohol abuse for nebraska counties. 2003.

H Desmond Patterson and Robin Thompson. Recovery of inter-block information when block sizes are unequal. *Biometrika*, 58(3):545–554, 1971.

NG Narasimha Prasad and Jon NK Rao. The estimation of the mean squared error of small-area estimators. *Journal of the American statistical association*, 85(409):163–171, 1990.

C Radhakrishna Rao. Minimum variance quadratic unbiased estimation of variance components. *Journal of Multivariate Analysis*, 1(4):445–456, 1971.

C Radhakrishna Rao. Estimation of variance and covariance components in linear models. *Journal of the American Statistical Association*, 67(337):112–115, 1972.

Calyampudi Radhakrishna Rao, Calyampudi Radhakrishna Rao, Mathematischer Statistiker, Calyampudi Radhakrishna Rao, and Calyampudi Radhakrishna Rao. *Linear statistical inference and its applications*, volume 2. Wiley New York, 1973.

John NK Rao. Small-area estimation. *Wiley StatsRef: Statistics Reference Online*, pages 1–8, 2014.

Shayle R Searle. Large sample variances of maximum likelihood estimators of variance components using unbalanced data. *Biometrics*, pages 505–524, 1970.

Shayle R Searle. The matrix handling of blue and blup in the mixed linear model. *Linear* *algebra and its applications*, 264:291–311, 1997.

Shayle R Searle, G Casella, and CE McCulloch. Variance components. 1992. *Hoboken: John Wiley & Sons*.

Wang Songgui and Yin Suju. A new estimate of the parameters in linear mixed models. *Science in China Series A: Mathematics*, 45(10):1301–1311, 2002.