

A new algebraic analysis to linear mixed models

Yongge Tian

China Economics and Management Academy, Central University of Finance and Economics, Beijing 100081, China

Abstract. This article presents a new investigation to the linear mixed model $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\varepsilon}$ with fixed effect $\mathbf{X}\boldsymbol{\beta}$ and random effect $\mathbf{Z}\boldsymbol{\gamma}$ under a general assumption via some novel algebraic tools in matrix theory, and reveals a variety of deep and profound properties hidden behind the linear mixed model. We first derive exact formulas for calculating the best linear unbiased predictor (BLUP) of a general vector $\boldsymbol{\phi} = \mathbf{F}\boldsymbol{\beta} + \mathbf{G}\boldsymbol{\gamma} + \mathbf{H}\boldsymbol{\varepsilon}$ of all unknown parameters in the model by solving a constrained quadratic matrix-valued function optimization problem in the Löwner partial ordering. We then consider some special cases of the BLUP for different choices of \mathbf{F} , \mathbf{G} , and \mathbf{H} in $\boldsymbol{\phi}$, and establish some fundamental decomposition equalities for the observed random vector \mathbf{y} and its covariance matrix.

Mathematics Subject Classifications: 62H12; 62J05

Keywords: linear mixed model, fixed effect, random effect, BLUP, BLUE, covariance matrix, decomposition

1 Introduction

Consider a linear model that includes fixed and random effects defined by

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\varepsilon}, \quad E \begin{bmatrix} \boldsymbol{\gamma} \\ \boldsymbol{\varepsilon} \end{bmatrix} = \mathbf{0}, \quad Cov \begin{bmatrix} \boldsymbol{\gamma} \\ \boldsymbol{\varepsilon} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix} = \boldsymbol{\Sigma}, \quad (1.1)$$

where

- $\mathbf{y} \in \mathbb{R}^{n \times 1}$ is an observable random vector,
- $\mathbf{X} \in \mathbb{R}^{n \times p}$ and $\mathbf{Z} \in \mathbb{R}^{n \times q}$ are known matrices of arbitrary ranks,
- $\boldsymbol{\beta} \in \mathbb{R}^{p \times 1}$ is a vector of fixed but unknown parameters (fixed effects),
- $\boldsymbol{\gamma} \in \mathbb{R}^{q \times 1}$ is a vector of unobservable random coefficients (random effects),
- $\boldsymbol{\varepsilon} \in \mathbb{R}^{n \times 1}$ is a random error vector,

$\boldsymbol{\Sigma} \in \mathbb{R}^{(q+n) \times (q+n)}$ is a known non-negative definite matrix of arbitrary rank.

Eq. (1.1) is usually called a Linear Mixed Model (LMM); see, e.g., Jiang (2007); McLean et al (1991); Muller & Stewart (2006); Verbeke & Molenberghs (2000).

Throughout this article, $\mathbb{R}^{m \times n}$ stands for the collection of all $m \times n$ real matrices. The symbols \mathbf{A}' , $r(\mathbf{A})$ and $\mathcal{R}(\mathbf{A})$ stand for the transpose, the rank and the range (column space) of a matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$, respectively. \mathbf{I}_m denotes the identity matrix of order m . The Moore–Penrose generalized inverse of \mathbf{A} , denoted by \mathbf{A}^+ , is defined to be the unique solution \mathbf{X} satisfying the four matrix equations $\mathbf{A}\mathbf{G}\mathbf{A} = \mathbf{A}$, $\mathbf{G}\mathbf{A}\mathbf{G} = \mathbf{G}$, $(\mathbf{A}\mathbf{G})' = \mathbf{A}\mathbf{G}$, and $(\mathbf{G}\mathbf{A})' = \mathbf{G}\mathbf{A}$. The symbols $\mathbf{P}_\mathbf{A}$, $\mathbf{E}_\mathbf{A}$, and $\mathbf{F}_\mathbf{A}$ stand for the three orthogonal projectors (symmetric idempotent matrices) $\mathbf{P}_\mathbf{A} = \mathbf{A}\mathbf{A}^+$, $\mathbf{A}^\perp = \mathbf{E}_\mathbf{A} = \mathbf{I}_m - \mathbf{A}\mathbf{A}^+$, and $\mathbf{F}_\mathbf{A} = \mathbf{I}_n - \mathbf{A}^+\mathbf{A}$, where both $\mathbf{E}_\mathbf{A}$ and $\mathbf{F}_\mathbf{A}$ satisfy $\mathbf{E}_\mathbf{A} = \mathbf{F}_\mathbf{A}'$ and $\mathbf{F}_\mathbf{A} = \mathbf{E}_\mathbf{A}'$, and the ranks of $\mathbf{E}_\mathbf{A}$ and $\mathbf{F}_\mathbf{A}$ are $r(\mathbf{E}_\mathbf{A}) = m - r(\mathbf{A})$ and $r(\mathbf{F}_\mathbf{A}) = n - r(\mathbf{A})$. Two symmetric matrices \mathbf{A} and \mathbf{B} of the same size are said to satisfy the inequality $\mathbf{A} \succcurlyeq \mathbf{B}$ in the Löwner partial ordering if $\mathbf{A} - \mathbf{B}$ is nonnegative definite. All about the orthogonal projectors $\mathbf{P}_\mathbf{A}$, $\mathbf{E}_\mathbf{A}$ and $\mathbf{F}_\mathbf{A}$ with their applications in the linear statistical models can be found, e.g., in Markiewicz & Puntanen (2015); Puntanen et al (2011); Rao & Mitra (1971). It is also well known that the Löwner partial ordering is a surprisingly strong and useful property between two symmetric matrices. For more issues about the Löwner partial ordering of symmetric matrices and applications in statistical analysis, see, e.g., Puntanen et al (2011).

We don't attach any further restrictions to the patterns of the submatrices $\boldsymbol{\Sigma}_{ij}$ in (1.1), although they are usually taken as certain prescribed forms for a specified in the statistical literature. In particular, the covariance matrices among the residual or error vector with other random factors in the model are usually assumed to be zero, say $\boldsymbol{\Sigma}_{12} = (\boldsymbol{\Sigma}_{21})' = \mathbf{0}$ in (1.1). Under the general assumptions in (1.1), the expectation and covariance matrix of \mathbf{y} , as well as the covariance matrices among $\boldsymbol{\gamma}$, $\boldsymbol{\varepsilon}$, and $\mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\varepsilon}$ are given by

$$E(\mathbf{y}) = \mathbf{X}\boldsymbol{\beta}, \quad (1.2)$$

$$Cov(\mathbf{y}) = Cov(\mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\varepsilon}) = \mathbf{Z}\boldsymbol{\Sigma}_{11}\mathbf{Z}' + \mathbf{Z}\boldsymbol{\Sigma}_{12} + \boldsymbol{\Sigma}_{21}\mathbf{Z}' + \boldsymbol{\Sigma}_{22} \triangleq \mathbf{V}, \quad (1.3)$$

$$Cov(\boldsymbol{\gamma}, \mathbf{y}) = Cov(\boldsymbol{\gamma}, \mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\varepsilon}) = \boldsymbol{\Sigma}_{11}\mathbf{Z}' + \boldsymbol{\Sigma}_{12} \triangleq \mathbf{V}_1, \quad (1.4)$$

$$Cov(\mathbf{Z}\boldsymbol{\gamma}, \mathbf{y}) = Cov(\mathbf{Z}\boldsymbol{\gamma}, \mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\varepsilon}) = \mathbf{Z}\boldsymbol{\Sigma}_{11}\mathbf{Z}' + \mathbf{Z}\boldsymbol{\Sigma}_{12} = \mathbf{Z}\mathbf{V}_1, \quad (1.5)$$

$$Cov(\boldsymbol{\varepsilon}, \mathbf{y}) = Cov(\boldsymbol{\varepsilon}, \mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\varepsilon}) = \boldsymbol{\Sigma}_{21}\mathbf{Z}' + \boldsymbol{\Sigma}_{22} \triangleq \mathbf{V}_2, \quad (1.6)$$

where \mathbf{V} , \mathbf{V}_1 , and \mathbf{V}_2 satisfy

$$\mathcal{R}(\mathbf{V}_1') \subseteq \mathcal{R}(\mathbf{V}), \quad \mathcal{R}(\mathbf{V}_2') \subseteq \mathcal{R}(\mathbf{V}), \quad \mathbf{V} = \mathbf{Z}\mathbf{V}_1 + \mathbf{V}_2 = \mathbf{V}_1'\mathbf{Z}' + \mathbf{V}_2'. \quad (1.7)$$

E-mail Address: yongge.tian@gmail.com

Eq. (1.1) is an important class of statistical models that have been widely used in many research areas, especially in the area of biomedical research, to analyze longitudinal and clustered data and multiple outcome data. Such data are encountered in a variety of fields including biostatistics, public health, psychometrics, educational measurement, and sociology. Like many other statistical models, (1.1) describes a relationship between a response variables and some regressors that have been measured or observed along with the response. In an LMM, both fixed and random effects contribute linearly to the response function and are very flexible and capable of fitting a large variety of data sets. Hence, LMMs are available to account for the variability of model parameters due to different factors that influence a response variable. LMMs are most widely-used models in statistical analysis, which are ubiquitous in many fields of research and are also the foundation for more advanced methods like logistic regression, survival analysis, multilevel modeling, and structural equation modeling, etc. LMMs have been a research topic of increasing interest in statistics for the past several decades, and a huge amount of literatures on BLUEs/BLUPs under LMMs can be found in the literature since Henderson (1975); see also Gumedze & Dunne (2011); Liu et al (2008); McLean et al (1991); Rao (1989); Robinson (1991); Searle (1988, 1997) to mention a few.

In order to establish some general results on predictions/estimations of all unknown parameters under the assumptions in (1.1), we construct a joint vector of parametric functions containing the fixed effect, random effect, and error term in (1.1) as follows

$$\boldsymbol{\phi} = [\mathbf{F}, \mathbf{G}, \mathbf{H}] \begin{bmatrix} \boldsymbol{\beta} \\ \boldsymbol{\gamma} \\ \boldsymbol{\varepsilon} \end{bmatrix} = \mathbf{F}\boldsymbol{\beta} + \mathbf{G}\boldsymbol{\gamma} + \mathbf{H}\boldsymbol{\varepsilon}, \quad (1.8)$$

where $\mathbf{F} \in \mathbb{R}^{s \times p}$, $\mathbf{G} \in \mathbb{R}^{s \times q}$, and $\mathbf{H} \in \mathbb{R}^{s \times n}$ are known matrices. In this case,

$$E(\boldsymbol{\phi}) = \mathbf{F}\boldsymbol{\beta}, \quad (1.9)$$

$$Cov(\boldsymbol{\phi}) = [\mathbf{G}, \mathbf{H}]\boldsymbol{\Sigma}[\mathbf{G}, \mathbf{H}]' = \mathbf{G}\boldsymbol{\Sigma}_{11}\mathbf{G}' + \mathbf{G}\boldsymbol{\Sigma}_{12}\mathbf{H}' + \mathbf{H}\boldsymbol{\Sigma}_{21}\mathbf{G}' + \mathbf{H}\boldsymbol{\Sigma}_{22}\mathbf{H}', \quad (1.10)$$

$$Cov\{\boldsymbol{\phi}, \mathbf{y}\} = [\mathbf{G}, \mathbf{H}]\boldsymbol{\Sigma}[\mathbf{Z}, \mathbf{I}_n]' = \mathbf{G}\mathbf{V}_1 + \mathbf{H}\mathbf{V}_2 = \mathbf{G}\boldsymbol{\Sigma}_{11}\mathbf{Z}' + \mathbf{G}\boldsymbol{\Sigma}_{12} + \mathbf{H}\boldsymbol{\Sigma}_{21}\mathbf{Z}' + \mathbf{H}\boldsymbol{\Sigma}_{22}. \quad (1.11)$$

Eq. (1.8) includes all vector operations in (1.1) as its special cases for different choices of \mathbf{F} , \mathbf{G} , and \mathbf{H} . For convenience of representation, we denote

$$\widehat{\mathbf{Z}} = [\mathbf{Z}, \mathbf{I}_n], \quad \mathbf{J} = [\mathbf{G}, \mathbf{H}], \quad \mathbf{C} = Cov\{\boldsymbol{\phi}, \mathbf{y}\} = \mathbf{J}\boldsymbol{\Sigma}\widehat{\mathbf{Z}}' = \mathbf{G}\mathbf{V}_1 + \mathbf{H}\mathbf{V}_2, \quad (1.12)$$

where \mathbf{V}_1 and \mathbf{V}_2 be as given in (1.4) and (1.6).

In statistical inference of a given linear regression model, a key issue is to choose an optimality criterion for predicting/estimating unknown parameters in the model. In comparison, the minimization of mean squared error (MSE) is a popular criterion for deriving best predictors in statistical inference. In case of linear predictor/estimators, it has the advantage that no further distributional assumptions of error terms need to be made, other than about the first- and second-order moments, while analytical formulas for predictors/estimators derived from the minimization of MSE and their statistical properties can easily be established. In order to establish a standard theory of the Best Linear Unbiased Predictor (BLUP) of all unknown parameters in (1.1), we need the following classic statistical concepts and definitions originated from Goldberger (1962).

Definition 1.1. The $\boldsymbol{\phi}$ in (1.8) is said to be *predictable* by \mathbf{y} in (1.1) if there exists a linear statistic $\mathbf{L}\mathbf{y}$ with $\mathbf{L} \in \mathbb{R}^{s \times n}$ such that

$$E(\mathbf{L}\mathbf{y} - \boldsymbol{\phi}) = \mathbf{0} \quad (1.13)$$

holds.

Definition 1.2. Let $\boldsymbol{\phi}$ be as given in (1.8). If there exists an $\mathbf{L}\mathbf{y}$ such that

$$Cov(\mathbf{L}\mathbf{y} - \boldsymbol{\phi}) = \min \quad \text{s.t.} \quad E(\mathbf{L}\mathbf{y} - \boldsymbol{\phi}) = \mathbf{0} \quad (1.14)$$

holds in the Löwner partial ordering, the linear statistic $\mathbf{L}\mathbf{y}$ is defined to be the BLUP of $\boldsymbol{\phi}$ in (1.8), and is denoted by

$$\mathbf{L}\mathbf{y} = \text{BLUP}(\boldsymbol{\phi}) = \text{BLUP}(\mathbf{F}\boldsymbol{\beta} + \mathbf{G}\boldsymbol{\gamma} + \mathbf{H}\boldsymbol{\varepsilon}). \quad (1.15)$$

If $\mathbf{G} = \mathbf{0}$ and $\mathbf{H} = \mathbf{0}$, the linear statistic $\mathbf{L}\mathbf{y}$ in (1.15) is the well-known Best Linear Unbiased Estimator (BLUE) of $\mathbf{F}\boldsymbol{\beta}$ under (1.1), and is denoted by

$$\mathbf{L}\mathbf{y} = \text{BLUE}(\mathbf{F}\boldsymbol{\beta}). \quad (1.16)$$

It is well known that optimal criteria for defining predictors/estimators under linear regression models have different choices. In comparison, BLUPs/BLEs have many simple and optimal algebraic and statistical properties, and therefore are the most welcome in both statistical theory and applications. To account for general prediction/estimation problems of unknown parameters in a linear model, it is common practice to first derive analytical expressions of predictors/estimators under the model. Along with some analytical methods of solving quadratic matrix-valued function problems were developed in recent years, it is now easy to handle various complicated matrix operations occurring in the statistical inferences under (1.1)–(1.11). The aim of the article is to solve the following problems:

(I) Derive exact expressions of the BLUEs/BLUPs of ϕ in (1.8) as well as $\mathbf{X}\beta$, $\mathbf{Z}\gamma$, and ε .

(II) Establish equalities among the BLUEs/BLUPs of $\mathbf{X}\beta$, $\mathbf{Z}\gamma$, and ε .

(III) Establish formulas for calculating the covariance matrices of the BLUEs/BLUPs of $\mathbf{X}\beta$, $\mathbf{Z}\gamma$, and ε .

Our approaches to these problems depend only on the general assumptions in (1.1) and pure matrix algebraic operations of the given matrices and vectors in (1.1) and (1.8). Thus, this work is more like rebuilding a new mathematical foundation in the statistical analysis of LMMs.

2 Preliminaries

BLUP was derived by Henderson in 1950 but the term "best linear unbiased predictor" (or "predictor") seems not to have been used until 1962. Note that BLUPs of random-effects are similar to BLUEs of fixed-effects from their definitions. The distinction arises because it is conventional to talk about estimating fixed effects, but predicting random effects. In the estimation theory of a regression model, the core part is to the choice and adoption of optimality criterion for estimating the parameters in the model. Based on the requirement of unbiasedness and minimum covariance matrix of an estimator, BLUPs under (1.1) were naturally proposed and have been main objects of study in regression analysis.

In order to establish various possible equalities and simplify matrix expressions composed by the Moore–Penrose inverses of matrices, we need the following well-known rank formulas for matrices and their Moore–Penrose inverses.

Lemma 2.1 (Marsaglia & Styan (1974)). *Let $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{B} \in \mathbb{R}^{m \times k}$, and $\mathbf{C} \in \mathbb{R}^{l \times n}$. Then*

$$r[\mathbf{A}, \mathbf{B}] = r(\mathbf{A}) + r(\mathbf{E}_A \mathbf{B}) = r(\mathbf{B}) + r(\mathbf{E}_B \mathbf{A}), \quad (2.1)$$

$$r \begin{bmatrix} \mathbf{A} \\ \mathbf{C} \end{bmatrix} = r(\mathbf{A}) + r(\mathbf{C} \mathbf{F}_A) = r(\mathbf{C}) + r(\mathbf{A} \mathbf{F}_C), \quad (2.2)$$

$$r \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{0} \end{bmatrix} = r(\mathbf{B}) + r(\mathbf{C}) + r(\mathbf{E}_B \mathbf{A} \mathbf{F}_C), \quad (2.3)$$

$$r \begin{bmatrix} \mathbf{A} \mathbf{A}' & \mathbf{B} \\ \mathbf{B}' & \mathbf{0} \end{bmatrix} = r[\mathbf{A}, \mathbf{B}] + r(\mathbf{B}). \quad (2.4)$$

Lemma 2.2 (Tian (2004)). *Under the conditions $\mathcal{R}(\mathbf{A}'_1) \subseteq \mathcal{R}(\mathbf{B}'_1)$, $\mathcal{R}(\mathbf{A}_2) \subseteq \mathcal{R}(\mathbf{B}_1)$, $\mathcal{R}(\mathbf{A}'_2) \subseteq \mathcal{R}(\mathbf{B}'_2)$, and $\mathcal{R}(\mathbf{A}_3) \subseteq \mathcal{R}(\mathbf{B}_2)$, the following two rank formulas hold*

$$r(\mathbf{A}_1 \mathbf{B}_1^+ \mathbf{A}_2) = r \begin{bmatrix} \mathbf{B}_1 & \mathbf{A}_2 \\ \mathbf{A}_1 & \mathbf{0} \end{bmatrix} - r(\mathbf{B}_1), \quad (2.5)$$

$$r(\mathbf{A}_1 \mathbf{B}_1^+ \mathbf{A}_2 \mathbf{B}_2^+ \mathbf{A}_3) = r \begin{bmatrix} \mathbf{0} & \mathbf{B}_2 & \mathbf{A}_3 \\ \mathbf{B}_1 & \mathbf{A}_2 & \mathbf{0} \\ \mathbf{A}_1 & \mathbf{0} & \mathbf{0} \end{bmatrix} - r(\mathbf{B}_1) - r(\mathbf{B}_2). \quad (2.6)$$

A well-known result in Penrose (1955) on the solvability and general solution of the linear matrix equation $\mathbf{Z}\mathbf{A} = \mathbf{B}$ is given below.

Lemma 2.3 (Penrose (1955)). *The linear matrix equation $\mathbf{Z}\mathbf{A} = \mathbf{B}$ is consistent if and only if $\mathcal{R}(\mathbf{A}') \supseteq \mathcal{R}(\mathbf{B}')$, or equivalently, $\mathbf{B}\mathbf{A}^+ \mathbf{A} = \mathbf{B}$. In this case, the general solution of $\mathbf{Z}\mathbf{A} = \mathbf{B}$ can be written in the following parametric form*

$$\mathbf{Z} = \mathbf{B}\mathbf{A}^+ + \mathbf{U}\mathbf{A}^\perp, \quad (2.7)$$

where \mathbf{U} is an arbitrary matrix.

In order to directly solve the matrix minimization problem in (1.14), we need the following result on analytical solution of a constrained quadratic matrix-valued function minimization problem, which was proved and utilized in Tian (2015a,b); Tian & Jiang (2016).

Lemma 2.4. *Let*

$$f(\mathbf{L}) = (\mathbf{L}\mathbf{C} + \mathbf{D})\mathbf{M}(\mathbf{L}\mathbf{C} + \mathbf{D})' \quad \text{s.t.} \quad \mathbf{L}\mathbf{A} = \mathbf{B}, \quad (2.8)$$

where $\mathbf{A} \in \mathbb{R}^{p \times q}$, $\mathbf{B} \in \mathbb{R}^{n \times q}$, $\mathbf{C} \in \mathbb{R}^{p \times m}$, and $\mathbf{D} \in \mathbb{R}^{n \times m}$ are given, $\mathbf{M} \in \mathbb{R}^{m \times m}$ is nnd, and the matrix equation $\mathbf{L}\mathbf{A} = \mathbf{B}$ is consistent. Then, there always exists a solution \mathbf{L}_0 of $\mathbf{L}_0 \mathbf{A} = \mathbf{B}$ such that

$$f(\mathbf{L}) \succcurlyeq f(\mathbf{L}_0) \quad (2.9)$$

holds for all solutions of $\mathbf{L}\mathbf{A} = \mathbf{B}$. In this case, the matrix \mathbf{L}_0 satisfying the above inequality is determined by the following consistent matrix equation

$$\mathbf{L}_0 \begin{bmatrix} \mathbf{A} & \mathbf{C} \mathbf{M} \mathbf{C}' \mathbf{A}^\perp \end{bmatrix} = \begin{bmatrix} \mathbf{B} & -\mathbf{D} \mathbf{M} \mathbf{C}' \mathbf{A}^\perp \end{bmatrix}. \quad (2.10)$$

In this case, the general expression of \mathbf{L}_0 and the corresponding $f(\mathbf{L}_0)$ and $f(\mathbf{L})$ are given by

$$\mathbf{L}_0 = \arg \min_{\mathbf{L} \mathbf{A} = \mathbf{B}} f(\mathbf{L}) = [\mathbf{B}, -\mathbf{D}\mathbf{M}\mathbf{C}'\mathbf{A}^\perp] [\mathbf{A}, \mathbf{C}\mathbf{M}\mathbf{C}'\mathbf{A}^\perp]^\perp + \mathbf{U} [\mathbf{A}, \mathbf{C}\mathbf{M}\mathbf{C}'\mathbf{A}^\perp]^\perp, \quad (2.11)$$

$$f(\mathbf{L}_0) = \min_{\mathbf{L} \mathbf{A} = \mathbf{B}} f(\mathbf{L}) = \mathbf{K}\mathbf{M}\mathbf{K}' - \mathbf{K}\mathbf{M}\mathbf{C}'\mathbf{T}\mathbf{C}\mathbf{M}\mathbf{K}', \quad (2.12)$$

$$f(\mathbf{L}) = f(\mathbf{L}_0) + (\mathbf{L}\mathbf{C} + \mathbf{D})\mathbf{M}\mathbf{C}'\mathbf{T}\mathbf{C}\mathbf{M}(\mathbf{L}\mathbf{C} + \mathbf{D})' \quad (2.13)$$

$$= f(\mathbf{L}_0) + (\mathbf{L}\mathbf{C}\mathbf{M}\mathbf{C}'\mathbf{A}^\perp + \mathbf{D}\mathbf{M}\mathbf{C}'\mathbf{A}^\perp) \mathbf{T} (\mathbf{L}\mathbf{C}\mathbf{M}\mathbf{C}'\mathbf{A}^\perp + \mathbf{D}\mathbf{M}\mathbf{C}'\mathbf{A}^\perp)', \quad (2.14)$$

where $\mathbf{K} = \mathbf{B}\mathbf{A}^+\mathbf{C} + \mathbf{D}$, $\mathbf{T} = (\mathbf{A}^\perp\mathbf{C}\mathbf{M}\mathbf{C}'\mathbf{A}^\perp)^\perp$, and $\mathbf{U} \in \mathbb{R}^{n \times p}$ is arbitrary.

3 Main results

In this section, we first introduce the concept of consistency of LMM in (1.1), and characterize the predictability of ϕ in (1.8). We then derive a fundamental linear matrix equation associated with the BLUP of ϕ , and present a variety of fundamental formulas for the BLUP and its covariance matrix operations.

Under the assumptions in (1.1), it is easy to see that

$$\begin{aligned} E[(\mathbf{I}_n - [\mathbf{X}, \mathbf{V}][\mathbf{X}, \mathbf{V}]^+) \mathbf{y}] &= (\mathbf{I}_n - [\mathbf{X}, \mathbf{V}][\mathbf{X}, \mathbf{V}]^+) \mathbf{X}\boldsymbol{\beta} = \mathbf{0}, \\ \text{Cov}[(\mathbf{I}_n - [\mathbf{X}, \mathbf{V}][\mathbf{X}, \mathbf{V}]^+) \mathbf{y}] &= (\mathbf{I}_n - [\mathbf{X}, \mathbf{V}][\mathbf{X}, \mathbf{V}]^+) \mathbf{V} (\mathbf{I}_n - [\mathbf{X}, \mathbf{V}][\mathbf{X}, \mathbf{V}]^+)' = \mathbf{0}. \end{aligned}$$

These two equalities imply that

$$\mathbf{y} \in \mathcal{R}[\mathbf{X}, \mathbf{V}] \text{ holds with probability 1.} \quad (3.1)$$

In this case, (1.1) is said to be consistent; see Rao (1971, 1973).

Note from (1.1) and (1.8) that the difference of $\mathbf{L}\mathbf{y} - \phi$ can be written as

$$\mathbf{L}\mathbf{y} - \phi = \mathbf{L}\mathbf{X}\boldsymbol{\beta} + \mathbf{L}\mathbf{Z}\boldsymbol{\gamma} + \mathbf{L}\boldsymbol{\varepsilon} - \mathbf{F}\boldsymbol{\beta} - \mathbf{G}\boldsymbol{\gamma} - \mathbf{H}\boldsymbol{\varepsilon} = (\mathbf{L}\mathbf{X} - \mathbf{F})\boldsymbol{\beta} + (\mathbf{L}\mathbf{Z} - \mathbf{G})\boldsymbol{\gamma} + (\mathbf{L} - \mathbf{H})\boldsymbol{\varepsilon}.$$

Then, the expectation vector and the covariance matrix of $\mathbf{L}\mathbf{y} - \phi$ can be written as

$$E(\mathbf{L}\mathbf{y} - \phi) = (\mathbf{L}\mathbf{X} - \mathbf{F})\boldsymbol{\beta}, \quad (3.2)$$

$$\text{Cov}(\mathbf{L}\mathbf{y} - \phi) = [\mathbf{L}\mathbf{Z} - \mathbf{G}, \mathbf{L} - \mathbf{H}]\boldsymbol{\Sigma}[\mathbf{L}\mathbf{Z} - \mathbf{G}, \mathbf{L} - \mathbf{H}]' = (\mathbf{L}\widehat{\mathbf{Z}} - \mathbf{J})\boldsymbol{\Sigma}(\mathbf{L}\widehat{\mathbf{Z}} - \mathbf{J})' \triangleq f(\mathbf{L}). \quad (3.3)$$

Hence, the constrained covariance matrix minimization problem in (1.14) converts to a mathematical problem of minimizing the quadratic matrix-valued function $f(\mathbf{L})$ subject to $(\mathbf{L}\mathbf{X} - \mathbf{F})\boldsymbol{\beta} = \mathbf{0}$. The equivalence of minimization problems of covariance matrices in statistical inference and matrix-valued function optimization problems in mathematics were first formulated by Rao (1989); see also Searle (1997).

Theorem 3.1. *The parameter vector ϕ in (1.8) is predictable by \mathbf{y} in (1.1) if and only if*

$$\mathcal{R}(\mathbf{X}') \supseteq \mathcal{R}(\mathbf{F}'). \quad (3.4)$$

Proof. It is obvious from (3.2) that

$$E(\mathbf{L}\mathbf{y} - \phi) = \mathbf{0} \Leftrightarrow \mathbf{L}\mathbf{X}\boldsymbol{\beta} = \mathbf{F}\boldsymbol{\beta} \text{ for all } \boldsymbol{\beta} \Leftrightarrow \mathbf{L}\mathbf{X} = \mathbf{F}.$$

From Lemma 2.3, the matrix equation $\mathbf{L}\mathbf{X} = \mathbf{F}$ is solvable if and only (3.4) holds. \square

Theorem 3.2. *Assume that ϕ in (1.8) is predictable by \mathbf{y} in (1.1). Then,*

$$E(\mathbf{L}\mathbf{y} - \phi) = \mathbf{0} \text{ and } \text{Cov}(\mathbf{L}\mathbf{y} - \phi) = \min \Leftrightarrow \mathbf{L}[\mathbf{X}, \text{Cov}(\mathbf{y})\mathbf{X}^\perp] = [\mathbf{F}, \text{Cov}\{\phi, \mathbf{y}\}\mathbf{X}^\perp]. \quad (3.5)$$

The matrix equation in (3.5), called the fundamental BLUP equation, is consistent, i.e.,

$$[\mathbf{F}, \text{Cov}\{\phi, \mathbf{y}\}\mathbf{X}^\perp][\mathbf{X}, \text{Cov}(\mathbf{y})\mathbf{X}^\perp]^\perp[\mathbf{X}, \text{Cov}(\mathbf{y})\mathbf{X}^\perp] = [\mathbf{F}, \text{Cov}\{\phi, \mathbf{y}\}\mathbf{X}^\perp] \quad (3.6)$$

holds under (3.4), while the general expression of \mathbf{L} , denoted by $\mathbf{L} = \mathbf{P}_{\mathbf{F}, \mathbf{G}, \mathbf{H}, \mathbf{X}, \mathbf{Z}, \boldsymbol{\Sigma}}$ and the corresponding BLUP(ϕ) can be written as

$$\begin{aligned} \text{BLUP}(\phi) &= \mathbf{P}_{\mathbf{F}, \mathbf{G}, \mathbf{H}, \mathbf{X}, \mathbf{Z}, \boldsymbol{\Sigma}} \mathbf{y} = \left([\mathbf{F}, \text{Cov}\{\phi, \mathbf{y}\}\mathbf{X}^\perp][\mathbf{X}, \text{Cov}(\mathbf{y})\mathbf{X}^\perp]^\perp + \mathbf{U}[\mathbf{X}, \text{Cov}(\mathbf{y})\mathbf{X}^\perp]^\perp \right) \mathbf{y} \\ &= \left([\mathbf{F}, \mathbf{C}\mathbf{X}^\perp][\mathbf{X}, \mathbf{V}\mathbf{X}^\perp]^\perp + \mathbf{U}[\mathbf{X}, \mathbf{V}\mathbf{X}^\perp]^\perp \right) \mathbf{y}, \end{aligned} \quad (3.7)$$

where $\mathbf{U} \in \mathbb{R}^{s \times n}$ is arbitrary. Further, the following results hold.

- (a) $r[\mathbf{X}, \mathbf{V}\mathbf{X}^\perp] = r[\mathbf{X}, \mathbf{V}]$ and $\mathcal{R}[\mathbf{X}, \mathbf{V}\mathbf{X}^\perp] = \mathcal{R}[\mathbf{X}, \mathbf{V}]$.
- (b) $\mathbf{P}_{\mathbf{F}, \mathbf{G}, \mathbf{H}, \mathbf{X}, \mathbf{Z}, \boldsymbol{\Sigma}}$ is unique if and only if $r[\mathbf{X}, \mathbf{V}] = n$.
- (c) BLUP(ϕ) is unique if and only if $\mathbf{y} \in [\mathbf{X}, \mathbf{V}]$ with probability 1, namely, (1.1) is consistent.

(d) The covariance matrix of $\text{BLUP}(\boldsymbol{\phi})$ is

$$\begin{aligned} \text{Cov}[\text{BLUP}(\boldsymbol{\phi})] &= \mathbf{P}_{\mathbf{F}, \mathbf{G}, \mathbf{H}, \mathbf{X}, \mathbf{Z}, \boldsymbol{\Sigma}} \text{Cov}(\mathbf{y}) \mathbf{P}'_{\mathbf{F}, \mathbf{G}, \mathbf{H}, \mathbf{X}, \mathbf{Z}, \boldsymbol{\Sigma}} \\ &= [\mathbf{F}, \mathbf{C}\mathbf{X}^\perp][\mathbf{X}, \mathbf{V}\mathbf{X}^\perp]^+ \mathbf{V}([\mathbf{F}, \mathbf{C}\mathbf{X}^\perp][\mathbf{X}, \mathbf{V}\mathbf{X}^\perp]^+)' ; \end{aligned} \quad (3.8)$$

the covariance matrix between $\text{BLUP}(\boldsymbol{\phi})$ and $\boldsymbol{\phi}$ is

$$\text{Cov}\{\text{BLUP}(\boldsymbol{\phi}), \boldsymbol{\phi}\} = [\mathbf{F}, \mathbf{C}\mathbf{X}^\perp][\mathbf{X}, \mathbf{V}\mathbf{X}^\perp]^+ (\mathbf{G}\mathbf{V}_1 + \mathbf{H}\mathbf{V}_2)' ; \quad (3.9)$$

the difference of $\text{Cov}(\boldsymbol{\phi})$ and $\text{Cov}[\text{BLUP}(\boldsymbol{\phi})]$ is

$$\text{Cov}(\boldsymbol{\phi}) - \text{Cov}[\text{BLUP}(\boldsymbol{\phi})] = \mathbf{J}\boldsymbol{\Sigma}\mathbf{J}' - [\mathbf{F}, \mathbf{C}\mathbf{X}^\perp][\mathbf{X}, \mathbf{V}\mathbf{X}^\perp]^+ \mathbf{V}([\mathbf{F}, \mathbf{C}\mathbf{X}^\perp][\mathbf{X}, \mathbf{V}\mathbf{X}^\perp]^+)' ; \quad (3.10)$$

the covariance matrix of $\boldsymbol{\phi} - \text{BLUP}(\boldsymbol{\phi})$ is

$$\text{Cov}[\boldsymbol{\phi} - \text{BLUP}(\boldsymbol{\phi})] = \left([\mathbf{F}, \mathbf{C}\mathbf{X}^\perp][\mathbf{X}, \mathbf{V}\mathbf{X}^\perp]^+ \widehat{\mathbf{Z}} - \mathbf{J}\right) \boldsymbol{\Sigma} \left([\mathbf{F}, \mathbf{C}\mathbf{X}^\perp][\mathbf{X}, \mathbf{V}\mathbf{X}^\perp]^+ \widehat{\mathbf{Z}} - \mathbf{J}\right)' . \quad (3.11)$$

(e) The ranks of the covariance matrices of $\text{BLUP}(\boldsymbol{\phi})$ and $\text{BLUP}(\boldsymbol{\phi}) - \boldsymbol{\phi}$ are

$$r\{\text{Cov}[\text{BLUP}(\boldsymbol{\phi})]\} = r \begin{bmatrix} \mathbf{V} & \mathbf{X} & \mathbf{0} \\ \mathbf{0} & \mathbf{F} & \mathbf{C} \\ \mathbf{0} & \mathbf{0} & \mathbf{X}' \end{bmatrix} - r[\mathbf{V}, \mathbf{X}] - r(\mathbf{X}), \quad (3.12)$$

$$r\{\text{Cov}[\boldsymbol{\phi} - \text{BLUP}(\boldsymbol{\phi})]\} = r \begin{bmatrix} \widehat{\mathbf{Z}}\boldsymbol{\Sigma} & \mathbf{X} \\ \mathbf{J}\boldsymbol{\Sigma} & \mathbf{F} \end{bmatrix} - r[\widehat{\mathbf{Z}}\boldsymbol{\Sigma}, \mathbf{X}]. \quad (3.13)$$

Proof. Eq. (1.14) is equivalent to finding a solution \mathbf{L}_0 of $\mathbf{L}_0\mathbf{X} = \mathbf{F}$ such that

$$f(\mathbf{L}) \succcurlyeq f(\mathbf{L}_0) \quad \text{for all solutions of } \mathbf{L}\mathbf{X} = \mathbf{F} \quad (3.14)$$

in the Löwner partial ordering. Since $\boldsymbol{\Sigma} \succcurlyeq \mathbf{0}$, (3.14) is a special case of (2.8). From Lemma 2.4, there always exists a solution \mathbf{L}_0 of $\mathbf{L}_0\mathbf{X} = \mathbf{F}$ such that $f(\mathbf{L}) \succcurlyeq f(\mathbf{L}_0)$ holds for all solutions of $\mathbf{L}\mathbf{X} = \mathbf{F}$, and the \mathbf{L}_0 is determined by the matrix equation

$$\mathbf{L}_0[\mathbf{X}, \widehat{\mathbf{Z}}\boldsymbol{\Sigma}\widehat{\mathbf{Z}}'\mathbf{X}^\perp] = [\mathbf{F}, \mathbf{J}\boldsymbol{\Sigma}\widehat{\mathbf{Z}}'\mathbf{X}^\perp], \quad (3.15)$$

where $\widehat{\mathbf{Z}}\boldsymbol{\Sigma}\widehat{\mathbf{Z}}' = \mathbf{V}$ and $\mathbf{J}\boldsymbol{\Sigma}\widehat{\mathbf{Z}}' = \mathbf{C}$, thus establishing the matrix equation in (3.5). This equation is consistent as well under (3.4), and the general solution of the equation and the corresponding BLUP are given in (3.7).

Result (a) is well known; see Markiewicz & Puntanen (2015); Puntanen et al (2011).

Results (b) and (c) follow from $[\mathbf{X}, \mathbf{V}\mathbf{X}^\perp]^\perp = \mathbf{0}$ and $[\mathbf{X}, \mathbf{V}\mathbf{X}^\perp]^\perp \mathbf{y} = \mathbf{0}$.

Taking covariance operation of (3.7) yields (3.8). Also from (1.8) and (3.7), the covariance matrix between $\text{BLUP}(\boldsymbol{\phi})$ and $\boldsymbol{\phi}$ is

$$\begin{aligned} \text{Cov}\{\text{BLUP}(\boldsymbol{\phi}), \boldsymbol{\phi}\} &= \text{Cov}\left\{([\mathbf{F}, \mathbf{C}\mathbf{X}^\perp][\mathbf{X}, \mathbf{V}\mathbf{X}^\perp]^+ + \mathbf{U}[\mathbf{X}, \mathbf{V}\mathbf{X}^\perp]^\perp)\mathbf{y}, \mathbf{G}\boldsymbol{\gamma} + \mathbf{H}\boldsymbol{\varepsilon}\right\} \\ &= [\mathbf{F}, \mathbf{C}\mathbf{X}^\perp][\mathbf{X}, \mathbf{V}\mathbf{X}^\perp]^+ \text{Cov}\left(\widehat{\mathbf{Z}} \begin{bmatrix} \boldsymbol{\gamma} \\ \boldsymbol{\varepsilon} \end{bmatrix}, \begin{bmatrix} \boldsymbol{\gamma} \\ \boldsymbol{\varepsilon} \end{bmatrix}\right) [\mathbf{G}, \mathbf{H}]' \\ &= [\mathbf{F}, \mathbf{C}\mathbf{X}^\perp][\mathbf{X}, \mathbf{V}\mathbf{X}^\perp]^+ \widehat{\mathbf{Z}}\boldsymbol{\Sigma}\mathbf{J}' \\ &= [\mathbf{F}, \mathbf{C}\mathbf{X}^\perp][\mathbf{X}, \mathbf{V}\mathbf{X}^\perp]^+ \mathbf{C}', \end{aligned}$$

thus establishing (3.9).

Combining (1.10) and (3.8) yields (3.10).

Substituting (3.7) into (3.3) and simplifying yields (3.11).

Applying (2.5) to (3.8) and simplifying yields

$$\begin{aligned} r\{\text{Cov}[\text{BLUP}(\boldsymbol{\phi})]\} &= r\left([\mathbf{F}, \mathbf{C}\mathbf{X}^\perp][\mathbf{X}, \mathbf{V}\mathbf{X}^\perp]^+ \mathbf{V}\right) \\ &= r \begin{bmatrix} [\mathbf{X}, \mathbf{V}\mathbf{X}^\perp] & \mathbf{V} \\ [\mathbf{F}, \mathbf{C}\mathbf{X}^\perp] & \mathbf{0} \end{bmatrix} - r[\mathbf{X}, \mathbf{V}\mathbf{X}^\perp] \\ &= r \begin{bmatrix} \mathbf{X} & \mathbf{0} & \mathbf{V} \\ \mathbf{F} & \mathbf{C} & \mathbf{0} \\ \mathbf{0} & \mathbf{X}' & \mathbf{0} \end{bmatrix} - r[\mathbf{X}, \mathbf{V}] - r(\mathbf{X}), \end{aligned}$$

as required for (3.12). Applying (2.5) to (3.11) and simplifying yields

$$\begin{aligned}
r\{Cov[\boldsymbol{\phi} - BLUP(\boldsymbol{\phi})]\} &= r\left([\mathbf{F}, \mathbf{CX}^\perp][\mathbf{X}, \mathbf{VX}^\perp]^+ \widehat{\mathbf{Z}}\boldsymbol{\Sigma} - \mathbf{J}\boldsymbol{\Sigma}\right) \\
&= r\begin{bmatrix} [\mathbf{X}, \mathbf{VX}^\perp] & \widehat{\mathbf{Z}}\boldsymbol{\Sigma} \\ [\mathbf{F}, \mathbf{CX}^\perp] & \mathbf{J}\boldsymbol{\Sigma} \end{bmatrix} - r[\mathbf{X}, \mathbf{VX}^\perp] \\
&= r\begin{bmatrix} \mathbf{X} & \widehat{\mathbf{Z}}\boldsymbol{\Sigma}\widehat{\mathbf{Z}}' & \widehat{\mathbf{Z}}\boldsymbol{\Sigma} \\ \mathbf{F} & \mathbf{J}\boldsymbol{\Sigma}\widehat{\mathbf{Z}}' & \mathbf{J}\boldsymbol{\Sigma} \\ \mathbf{0} & \mathbf{X}' & \mathbf{0} \end{bmatrix} - r[\mathbf{X}, \mathbf{V}] - r(\mathbf{X}) \\
&= r\begin{bmatrix} \mathbf{X} & \mathbf{0} & \widehat{\mathbf{Z}}\boldsymbol{\Sigma} \\ \mathbf{F} & \mathbf{0} & \mathbf{J}\boldsymbol{\Sigma} \\ \mathbf{0} & \mathbf{X}' & \mathbf{0} \end{bmatrix} - r[\mathbf{X}, \mathbf{V}] - r(\mathbf{X}) \\
&= r\begin{bmatrix} \mathbf{X} & \widehat{\mathbf{Z}}\boldsymbol{\Sigma} \\ \mathbf{F} & \mathbf{J}\boldsymbol{\Sigma} \end{bmatrix} - r[\mathbf{X}, \widehat{\mathbf{Z}}\boldsymbol{\Sigma}],
\end{aligned}$$

thus establishing (3.13). \square

Corollary 3.3. *Assume that $\boldsymbol{\phi}$ in (1.8) is predictable by \mathbf{y} in (1.1). Then, $\mathbf{F}\boldsymbol{\beta}$ is estimable by \mathbf{y} in (1.1) as well. In this case, the following results hold.*

(a) *The BLUP of $\boldsymbol{\phi}$ can be decomposed as the sum*

$$BLUP(\boldsymbol{\phi}) = BLUE(\mathbf{F}\boldsymbol{\beta}) + BLUP(\mathbf{G}\boldsymbol{\gamma}) + BLUP(\mathbf{H}\boldsymbol{\varepsilon}), \quad (3.16)$$

and they satisfy

$$Cov\{BLUE(\mathbf{F}\boldsymbol{\beta}), BLUP(\mathbf{G}\boldsymbol{\gamma} + \mathbf{H}\boldsymbol{\varepsilon})\} = \mathbf{0}, \quad (3.17)$$

$$Cov[BLUP(\boldsymbol{\phi})] = Cov[BLUE(\mathbf{F}\boldsymbol{\beta})] + Cov[BLUP(\mathbf{G}\boldsymbol{\gamma} + \mathbf{H}\boldsymbol{\varepsilon})]. \quad (3.18)$$

(b) *$\mathbf{T}\boldsymbol{\phi}$ is predictable for any matrix $\mathbf{T} \in \mathbb{R}^{t \times s}$, and*

$$BLUP(\mathbf{T}\boldsymbol{\phi}) = \mathbf{T}BLUP(\boldsymbol{\phi}). \quad (3.19)$$

Proof. Note that the matrix $[\mathbf{F}, \mathbf{CX}^\perp]$ in (3.7) can be rewritten as

$$[\mathbf{F}, \mathbf{CX}^\perp] = [\mathbf{F}, \mathbf{0}] + [\mathbf{0}, \mathbf{GV}_1\mathbf{X}^\perp] + [\mathbf{0}, \mathbf{HV}_2\mathbf{X}^\perp].$$

Correspondingly, $BLUP(\boldsymbol{\phi})$ in (3.7) can be rewritten as the sum

$$\begin{aligned}
BLUP(\boldsymbol{\phi}) &= \left([\mathbf{F}, \mathbf{CX}^\perp][\mathbf{X}, \mathbf{VX}^\perp]^+ + \mathbf{U}[\mathbf{X}, \mathbf{VX}^\perp]^\perp\right)\mathbf{y} \\
&= \left([\mathbf{F}, \mathbf{0}][\mathbf{X}, \mathbf{VX}^\perp]^+ + \mathbf{U}_1[\mathbf{X}, \mathbf{VX}^\perp]^\perp\right)\mathbf{y} \\
&\quad + \left([\mathbf{0}, \mathbf{GV}_1\mathbf{X}^\perp][\mathbf{X}, \mathbf{VX}^\perp]^+ + \mathbf{U}_2[\mathbf{X}, \mathbf{VX}^\perp]^\perp\right)\mathbf{y} + \\
&\quad + \left([\mathbf{0}, \mathbf{HV}_2\mathbf{X}^\perp][\mathbf{X}, \mathbf{VX}^\perp]^+ + \mathbf{U}_3[\mathbf{X}, \mathbf{VX}^\perp]^\perp\right)\mathbf{y} \\
&= BLUE(\mathbf{F}\boldsymbol{\beta}) + BLUP(\mathbf{G}\boldsymbol{\gamma}) + BLUP(\mathbf{H}\boldsymbol{\varepsilon}),
\end{aligned}$$

thus establishing (3.16). From (3.7), the covariance matrix between $BLUE(\mathbf{F}\boldsymbol{\beta})$ and $BLUP(\mathbf{G}\boldsymbol{\gamma} + \mathbf{H}\boldsymbol{\varepsilon})$ is

$$Cov\{BLUE(\mathbf{F}\boldsymbol{\beta}), BLUP(\mathbf{G}\boldsymbol{\gamma} + \mathbf{H}\boldsymbol{\varepsilon})\} = [\mathbf{F}, \mathbf{0}][\mathbf{X}, \mathbf{VX}^\perp]^+ \mathbf{V}([\mathbf{0}, (\mathbf{GV}_1 + \mathbf{HV}_2)\mathbf{X}^\perp][\mathbf{X}, \mathbf{VX}^\perp]^\perp)'$$

Applying (2.6) to the right-hand side of the above equality and simplifying, we obtain

$$\begin{aligned}
& r(\text{Cov}\{\text{BLUE}(\mathbf{F}\boldsymbol{\beta}), \text{BLUP}(\mathbf{G}\boldsymbol{\gamma} + \mathbf{H}\boldsymbol{\varepsilon})\}) \\
&= r\left([\mathbf{F}, \mathbf{0}][\mathbf{X}, \mathbf{V}\mathbf{X}^\perp]^+ \mathbf{V}([\mathbf{0}, (\mathbf{G}\mathbf{V}_1 + \mathbf{H}\mathbf{V}_2)\mathbf{X}^\perp][\mathbf{X}, \mathbf{V}\mathbf{X}^\perp]^+)' \right) \\
&= r \left[\begin{array}{cc} \mathbf{0} & \begin{bmatrix} \mathbf{X}' \\ \mathbf{X}^\perp \mathbf{V} \end{bmatrix} \\ \begin{bmatrix} \mathbf{X}, \mathbf{V}\mathbf{X}^\perp \\ \mathbf{F}, \mathbf{0} \end{bmatrix} & \begin{bmatrix} \mathbf{V} \\ \mathbf{0} \end{bmatrix} \end{array} \begin{bmatrix} \mathbf{X}^\perp (\mathbf{G}\mathbf{V}_1 + \mathbf{H}\mathbf{V}_2)' \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} \right] - 2r[\mathbf{X}, \mathbf{V}\mathbf{X}^\perp] \\
&= r \left[\begin{array}{cc} \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\mathbf{X}^\perp \mathbf{V}\mathbf{X}^\perp \end{bmatrix} & \begin{bmatrix} \mathbf{X}' \\ \mathbf{0} \end{bmatrix} \\ \begin{bmatrix} \mathbf{X}, \mathbf{0} \\ \mathbf{F}, \mathbf{0} \end{bmatrix} & \begin{bmatrix} \mathbf{V} \\ \mathbf{0} \end{bmatrix} \end{array} \begin{bmatrix} \mathbf{X}^\perp (\mathbf{G}\mathbf{V}_1 + \mathbf{H}\mathbf{V}_2)' \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} \right] - 2r[\mathbf{X}, \mathbf{V}] \\
&= r \begin{bmatrix} \mathbf{0} & \mathbf{X}' \\ \mathbf{X} & \mathbf{V} \\ \mathbf{F} & \mathbf{0} \end{bmatrix} + r[\mathbf{X}^\perp \mathbf{V}\mathbf{X}^\perp, \mathbf{X}^\perp (\mathbf{G}\mathbf{V}_1 + \mathbf{H}\mathbf{V}_2)'] - 2r[\mathbf{X}, \mathbf{V}] \\
&= r \begin{bmatrix} \mathbf{X} \\ \mathbf{F} \end{bmatrix} + r \begin{bmatrix} \mathbf{X}' \\ \mathbf{V} \end{bmatrix} + r[\mathbf{X}, \mathbf{V}\mathbf{X}^\perp, (\mathbf{G}\mathbf{V}_1 + \mathbf{H}\mathbf{V}_2)'] - r(\mathbf{X}) - 2r[\mathbf{X}, \mathbf{V}] \quad (\text{by (2.4)}) \\
&= r[\mathbf{X}, \mathbf{V}, \mathbf{V}_1' \mathbf{G}' + \mathbf{V}_2' \mathbf{H}'] - r[\mathbf{X}, \mathbf{V}] \quad (\text{by Theorem 3.2(a)}) \\
&= r[\mathbf{X}, \mathbf{V}, \mathbf{0}] - r[\mathbf{X}, \mathbf{V}] \quad (\text{by (1.7)}) \\
&= 0,
\end{aligned}$$

which implies that $\text{Cov}\{\text{BLUE}(\mathbf{F}\boldsymbol{\alpha}), \text{BLUP}(\mathbf{G}\boldsymbol{\gamma} + \mathbf{H}\boldsymbol{\varepsilon})\}$ is a zero matrix, thus establishing (3.17). Equation (3.18) follows from (3.16) and (3.17). \square

We next present some consequences of Theorem 3.2 for the different choices of three given matrices \mathbf{F} , \mathbf{G} , and \mathbf{H} in (1.8).

Corollary 3.4. *The following results hold.*

- (a) $\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\gamma}$, $\mathbf{X}\boldsymbol{\beta}$, $\mathbf{Z}\boldsymbol{\gamma}$, $\boldsymbol{\gamma}$, and $\boldsymbol{\varepsilon}$ in (1.1) are all predictable by \mathbf{y} in (1.1) and their BLUPs and BLUE are given by

$$\text{BLUP}(\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\gamma}) = \left([\mathbf{X}, \mathbf{Z}\mathbf{V}_1\mathbf{X}^\perp][\mathbf{X}, \mathbf{V}\mathbf{X}^\perp]^+ + \mathbf{U}_1[\mathbf{X}, \mathbf{V}]^\perp\right)\mathbf{y}, \quad (3.20)$$

$$\text{BLUE}(\mathbf{X}\boldsymbol{\beta}) = \left([\mathbf{X}, \mathbf{0}][\mathbf{X}, \mathbf{V}\mathbf{X}^\perp]^+ + \mathbf{U}_2[\mathbf{X}, \mathbf{V}]^\perp\right)\mathbf{y}, \quad (3.21)$$

$$\text{BLUP}(\mathbf{Z}\boldsymbol{\gamma}) = \left([\mathbf{0}, \mathbf{Z}\mathbf{V}_1\mathbf{X}^\perp][\mathbf{X}, \mathbf{V}\mathbf{X}^\perp]^+ + \mathbf{U}_3[\mathbf{X}, \mathbf{V}]^\perp\right)\mathbf{y}, \quad (3.22)$$

$$\text{BLUP}(\boldsymbol{\gamma}) = \left([\mathbf{0}, \mathbf{V}_1\mathbf{X}^\perp][\mathbf{X}, \mathbf{V}\mathbf{X}^\perp]^+ + \mathbf{U}_4[\mathbf{X}, \mathbf{V}]^\perp\right)\mathbf{y}, \quad (3.23)$$

$$\text{BLUP}(\boldsymbol{\varepsilon}) = \left([\mathbf{0}, \mathbf{V}_2\mathbf{X}^\perp][\mathbf{X}, \mathbf{V}\mathbf{X}^\perp]^+ + \mathbf{U}_5[\mathbf{X}, \mathbf{V}]^\perp\right)\mathbf{y}, \quad (3.24)$$

where $\mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3, \mathbf{U}_5 \in \mathbb{R}^{n \times n}$ and $\mathbf{U}_4 \in \mathbb{R}^{q \times n}$ are arbitrary.

- (b) \mathbf{y} satisfies the following decomposition equalities

$$\mathbf{y} = \text{BLUP}(\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\gamma}) + \text{BLUP}(\boldsymbol{\varepsilon}), \quad (3.25)$$

$$\mathbf{y} = \text{BLUE}(\mathbf{X}\boldsymbol{\beta}) + \text{BLUP}(\mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\varepsilon}), \quad (3.26)$$

$$\mathbf{y} = \text{BLUE}(\mathbf{X}\boldsymbol{\beta}) + \text{BLUP}(\mathbf{Z}\boldsymbol{\gamma}) + \text{BLUP}(\boldsymbol{\varepsilon}). \quad (3.27)$$

- (c) $\mathbf{X}\boldsymbol{\beta}$, $\mathbf{Z}\boldsymbol{\gamma}$, and $\boldsymbol{\varepsilon}$ satisfy the following covariance matrix equalities

$$\text{Cov}\{\text{BLUE}(\mathbf{X}\boldsymbol{\beta}), \text{BLUP}(\mathbf{Z}\boldsymbol{\gamma})\} = \mathbf{0}, \quad (3.28)$$

$$\text{Cov}\{\text{BLUE}(\mathbf{X}\boldsymbol{\beta}), \text{BLUP}(\boldsymbol{\varepsilon})\} = \mathbf{0}, \quad (3.29)$$

$$\text{Cov}\{\text{BLUE}(\mathbf{X}\boldsymbol{\beta}), \text{BLUP}(\mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\varepsilon})\} = \mathbf{0}, \quad (3.30)$$

$$\text{Cov}(\mathbf{y}) = \text{Cov}[\text{BLUE}(\mathbf{X}\boldsymbol{\beta})] + \text{Cov}[\text{BLUP}(\mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\varepsilon})]. \quad (3.31)$$

More precise consequences can be derived from the previous theorems and corollaries for special forms of (1.1). For instance, one of the well-known cases of (1.1) is the following linear random-effects model

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \quad \boldsymbol{\beta} = \mathbf{A}\boldsymbol{\alpha} + \boldsymbol{\gamma}, \quad (3.32)$$

where $\boldsymbol{\beta} \in \mathbb{R}^{p \times 1}$ is a vector of unobservable random variables, $\mathbf{A} \in \mathbb{R}^{p \times k}$ is known matrix of arbitrary rank, $\boldsymbol{\alpha} \in \mathbb{R}^{k \times 1}$ is a vector of fixed but unknown parameters (fixed effects), and $\boldsymbol{\gamma} \in \mathbb{R}^{p \times 1}$ is a vector of unobservable random variables (random effects). Substituting the second equation into the first equation in (3.32) yields an LMM as follows

$$\mathbf{y} = \mathbf{X}\mathbf{A}\boldsymbol{\alpha} + \mathbf{X}\boldsymbol{\gamma} + \boldsymbol{\varepsilon}. \quad (3.33)$$

The previous conclusions are also valid for (3.33), while this kind of work was presented in Tian (2015a).

In this article, we give a new investigation to the prediction/estimation problems of all unknown parameters in LMMs, revealed many intrinsic algebraic properties hidden behind LMMs, and obtained a variety of analytical formulas in the statistical analysis of LMMs. When the covariance matrix in (1.1) is given in certain block forms, such as, $\Sigma = \sigma^2 \text{diag}(\Sigma_1, \Sigma_2)$ or $\Sigma = \text{diag}(\tau^2 \mathbf{I}_q, \sigma^2 \mathbf{I}_n)$, where Σ_1 and Σ_2 are two known matrices, and σ^2 and τ^2 are unknown positive parameters, the formulas in the previous theorems and corollaries can be simplified further. Because this work is done under a most general assumption, it is believed that the article can serve as a standard reference in the statistical inference of LMMs for various specified model matrices \mathbf{X} and \mathbf{Z} , and specified covariance matrix Σ in (1.1).

References

- Goldberger, A.S., 1962. Best linear unbiased prediction in the generalized linear regression models. *J. Amer. Stat. Assoc.* 57, 369–375.
- Gumedze, F.N., Dunne, T.T., 2011. Parameter estimation and inference in the linear mixed model. *Linear Algebra Appl.* 435, 1920–1944.
- Henderson, C.R., 1975. Best linear unbiased estimation and prediction under a selection model. *Biometrics* 31, 423–447.
- Jiang, J., 2007. *Linear and Generalized Linear Mixed Models and Their Applications*. Springer, New York.
- Kubolawa, T., 1995. Estimation of variance componenets in mixed linear models. *J. multivariate Anal.* 53, 21–236.
- Liu, X.-Q., Rong, J.-Y., Liu, X.-Y., 2008. Best linear unbiased prediction for linear combinations in general mixed linear models. *J. Multivariate Anal.* 99, 1503–1517.
- Markiewicz, A., Puntanen, S., 2015. All about the \perp with its applications in the linear statistical models. *Open Math.* 13, 33–50.
- Marsaglia, G., Styan, G.P.H., 1974. Equalities and inequalities for ranks of matrices. *Linear Multilinear Algebra* 2, 269–292.
- McLean, R.A., Sanders, W.L., Stroup, W.W., 1991. A unified approach to mixed linear models. *Amer. Statist.* 45, 54–64.
- Muller, K.E., Stewart, P. W., 2006. *Linear Model Theory Univariate, Multivariate, and Mixed Models*. Wiley, New York.
- Penrose, R., 1955. A generalized inverse for matrices. *Proc. Cambridge Phil. Soc.* 51, 406–413.
- Puntanen, S., Styan, G.P.H., Isotalo, J., 2011. *Matrix Tricks for Linear Statistical Models: Our Personal Top Twenty*. Springer, Berlin.
- Rao, C.R., 1971. Unified theory of linear estimation. *Sankhyā, Ser A* 33, 371–394.
- Rao, C.R., 1973. Representations of best linear unbiased estimators in the Gauss–Markoff model with a singular dispersion matrix. *J. Multivariate Anal.* 3, 276–292.
- Rao, C.R., 1989. A lemma on optimization of matrix function and a review of the unified theory of linear estimation. In: *Statistical Data Analysis and Inference*, Y. Dodge (ed.), North Holland, pp. 397–417.
- Rao, C.R., Mitra, S.K., 1971. *Generalized Inverse of Matrices and Its Applications*. Wiley, New York.
- Robinson, G.K., 1991. That BLUP is a good thing: the estimation of random effects. *Statist. Sci.* 6, 15–32.
- Searle, S.R., 1988. Best linear unbiased estimation in mixed models of the analysis of variance. *J. Srivastava (Ed.), Probability and Statistics: Essays in Honor of Franklin A. Graybill*, North-Holland, Amsterdam, pp. 233–241.
- Searle, S.R., 1997. The matrix handling of BLUE and BLUP in the mixed linear model. *Linear Algebra Appl.* 264, 291–311.
- Tian, Y., 2004. Using rank formulas to characterize equalities for Moore–Penrose inverses of matrix products. *Appl. Math. Comput.* 147, 581–600.
- Verbeke, G., Molenberghs, G., 2000. *Linear Mixed Models for Longitudinal Data*. Springer, New York.
- Tian, Y., 2015a. A new derivation of BLUPs under random-effects model. *Metrika* 78, 905–918.
- Tian, Y., 2015b. A matrix handling of predictions under a general linear random-effects model with new observations. *Electron. J. Linear Algebra* 29, 30–45.
- Tian, Y., Jiang, B., 2016. An algebraic study of BLUPs under two linear random-effects models with correlated covariance matrices. *Linear Multilinear Algebra*, DOI:10.1080/03081087.2016.1155533.