A matrix handling of predictions of new observations under a general random-effects model

Yongge Tian
China Economics and Management Academy, Central University of Finance and Economics, Beijing 100081, China, yongge.tian@gmail.com

Follow this and additional works at: http://repository.uwyo.edu/ela

Recommended Citation
DOI: https://doi.org/10.13001/1081-3810.2895
A MATRIX HANDLING OF PREDICTIONS UNDER A GENERAL LINEAR RANDOM-EFFECTS MODEL WITH NEW OBSERVATIONS*

YONGGE TIAN†

Abstract. Linear regression models that include random effects are commonly used to analyze longitudinal and correlated data. Assume that a general linear random-effects model \( y = X\beta + \varepsilon \) with \( \beta = A\alpha + \gamma \) is given, and new observations in the future follow the linear model \( y_f = X_f\beta + \varepsilon_f \). This paper shows how to establish a group of matrix equations and analytical formulas for calculating the best linear unbiased predictor (BLUP) of the vector \( \phi = F\alpha + G\gamma + H\varepsilon + H_f\varepsilon_f \) of all unknown parameters in the two models under a general assumption on the covariance matrix among the random vectors \( \gamma, \varepsilon \) and \( \varepsilon_f \) via solving a constrained quadratic matrix-valued function optimization problem. Many consequences on the BLUPs of \( \phi \) and their covariance matrices, as well as additive decomposition equalities of the BLUPs with respect to its components are established under various assumptions.

Key words. Linear random-effects model, Quadratic matrix-valued function, Löwner partial ordering, BLUP, BLUE, Covariance matrix.

AMS subject classifications. 15A09, 62H12, 62J05.

Dedicated to Professor R. B. Bapat on the occasion of his 60th birthday

1. Introduction. Throughout this paper, \( \mathbb{R}^{m \times n} \) stands for the collection of all \( m \times n \) real matrices. The symbols \( A' \), \( r(A) \) and \( \mathcal{R}(A) \) stand for the transpose, the rank and the range (column space) of a matrix \( A \in \mathbb{R}^{m \times n} \), respectively. \( I_m \) denotes the identity matrix of order \( m \). The Moore–Penrose inverse of \( A \), denoted by \( A^+ \), is defined to be the unique solution \( G \) satisfying the four matrix equations \( AGA = A, \ GAG = G, \ (AG)' = AG, \ \text{and} \ (GA)' = GA. \ P_A, E_A, \) and \( F_A \) stand for the three orthogonal projectors (symmetric idempotent matrices) \( P_A = A A^+ \), \( E_A = A^+ = I_m - A A^+ \), and \( F_A = I_n - A^+ A \), where \( E_A \) and \( F_A \) satisfy \( E_A = F_A \) and \( F_A = E_A \). Two symmetric matrices \( A \) and \( B \) of the same size are said to satisfy the Löwner partial ordering \( A \succeq B \) if \( A - B \) is nonnegative definite.

Consider a general Linear Random-effects Model (LRM) defined by
\[
y = X\beta + \varepsilon, \quad \beta = A\alpha + \gamma, \quad (1.1)
\]
or marginally,
\[
y = XA\alpha + X\gamma + \varepsilon, \quad (1.2)
\]

*Received by the editors on February 22, 2015. Accepted for publication on May 28, 2015. Handling Editor: Simo Puntanen.
†China Economics and Management Academy, Central University of Finance and Economics, Beijing 100081, China (yongge.tian@gmail.com). Supported by the National Natural Science Foundation of China (Grant No. 11271384).
Predictions under LRM

where in the first-stage model, \( y \in \mathbb{R}^{n \times 1} \) is a vector of observable response variables, \( X \in \mathbb{R}^{n \times p} \) is a known matrix of arbitrary rank, \( \varepsilon \in \mathbb{R}^{n \times 1} \) is a vector of unobservable random errors, while in the second-stage model, \( \beta \in \mathbb{R}^{p \times 1} \) is a vector of unobservable random variables, \( A \in \mathbb{R}^{p \times k} \) is known matrix of arbitrary rank, \( \alpha \in \mathbb{R}^{k \times 1} \) is a vector of fixed but unknown parameters (fixed effects), \( \gamma \in \mathbb{R}^{p \times 1} \) is a vector of unobservable random variables (random effects). Concerning the expectation and covariance matrix of random vectors \( \gamma \) and \( \varepsilon \) in (1.1), we adopt the following general assumption

\[
E \left[ \begin{array}{c} \gamma \\ \varepsilon \end{array} \right] = 0, \quad Cov \left[ \begin{array}{c} \gamma \\ \varepsilon \end{array} \right] = \left[ \begin{array}{cc} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{array} \right] := \Sigma, \tag{1.3} \]

where \( \Sigma_{11} \in \mathbb{R}^{p \times p} \), \( \Sigma_{12} = \Sigma_{21} \in \mathbb{R}^{p \times n} \), and \( \Sigma_{22} \in \mathbb{R}^{n \times n} \) are known, while \( \Sigma \in \mathbb{R}^{(p+n) \times (p+n)} \) is non-negative definite (nnd) matrix of arbitrary rank.

One of the ultimate goals of statistical modelling is to be able to predict future observations based on currently available information. Assume that new observations of response variables in the future follow the model

\[
y_f = X_f \beta + \varepsilon_f = X_f A \alpha + X_f \gamma + \varepsilon_f, \tag{1.4} \]

where \( X_f \in \mathbb{R}^{n_f \times p} \) is a known model matrix associated with the new observations, \( \beta \) is the same vector of unknown parameters as in (1.1), and \( \varepsilon_f \in \mathbb{R}^{n_f \times 1} \) is a vector of measurement errors associated with new observations. Combining (1.2) and (1.4) yields the following marginal model

\[
\tilde{y} = \tilde{X} \alpha + \tilde{X} \gamma + \tilde{\varepsilon}, \quad \tilde{y} = \begin{bmatrix} y \\ y_f \end{bmatrix}, \quad \tilde{X} = \begin{bmatrix} X \\ X_f \end{bmatrix}, \quad \tilde{\varepsilon} = \begin{bmatrix} \varepsilon \\ \varepsilon_f \end{bmatrix}. \tag{1.5} \]

In order to establish some general results on prediction analysis of (1.5), we assume that the expectation and covariance matrix of \( \gamma, \varepsilon, \) and \( \varepsilon_f \) are given by

\[
E \left[ \begin{array}{c} \gamma \\ \varepsilon \\ \varepsilon_f \end{array} \right] = 0, \quad Cov \left[ \begin{array}{c} \gamma \\ \varepsilon \\ \varepsilon_f \end{array} \right] = \left[ \begin{array}{ccc} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} \\ \Sigma_{21} & \Sigma_{22} & \Sigma_{23} \\ \Sigma_{31} & \Sigma_{32} & \Sigma_{33} \end{array} \right] := \tilde{\Sigma}, \tag{1.6} \]

where we don’t attach any further restrictions to the patterns of the submatrices \( \Sigma_{ij} \) in (1.6), although they are usually taken as certain prescribed forms for a specified LRM in the statistical literature. In particular, the covariances among the residual or error vector with other random factors in the model are usually assumed to be zero. This assumption is ordinarily applied to most practical applications in the biological sciences when the assumption is invalid.

Linear regression models that include random effects are commonly used to analyze longitudinal and correlated data, which are available to account for the variability of model parameters due to different factors that influence a response variable. The LRM in (1.1) is also called nested model or two-level model in the statistical literature, where the two equations are called the first-stage model and the second-stage model, respectively. Statistical inference on LMRs is now an important part in data analysis,
and a huge amount of literature spreads in statistics and other disciplines. As usual, a main task in the investigation of LRM is to establish predictors/estimators of all unknown parameters in the model. Recall that the Best Linear Unbiased Predictors (BLUPs) of unknown random parameters and the Best Linear Unbiased Estimators (BLUEs) of fixed but unknown parameters in LRM are fundamental concepts in current regression analysis, which are defined directly from the requirement of both unbiasedness and minimum covariance matrices of predictors/estimators of the unknown parameters. In fact, BLUPs/BLUEs are primary choices in all possible predictors/estimators due to their simple and optimality properties, and have wide applications in both pure and applied disciplines of statistical inferences. The theory of BLUPs/BLUEs under linear regression models belongs to the classical methods of mathematical statistics. Along with recent development of optimization methods in matrix theory, it is now easy to deal with various complicated matrix operations occurring in the statistical inference of (1.5). In [25], the present author established a group of fundamental matrix equations and analytical formulas for calculating the BLUPs/BLUEs of all unknown parameters in (1.1) via solving a constrained quadratic matrix-valued function optimization problem, and also formulated an open problem of establishing matrix equations and formulas for calculating the BLUPs/BLUEs of the future $y_f$, $X_f \beta$, $X_f \alpha_n$, $X_f \gamma$, and $\epsilon_f$ in (1.4) from the observed response vector $y$ in (1.1). For convenience of representation, let

$$R = \begin{bmatrix} R_1 \\ R_2 \end{bmatrix} = \begin{bmatrix} X & I_n & 0 \\ X_f & 0 & I_{n_f} \end{bmatrix} = [\bar{X}, I_{n+n_f}] , \quad (1.7)$$

Under the general assumptions in (1.3) and (1.6), the covariance matrix of the combined random vector $\tilde{y}$ in (1.5) is given by

$$\text{Cov}(\tilde{y}) = \begin{bmatrix} \text{Cov}(y) \\ \text{Cov}(y_f, y) \\ \text{Cov}(y_f) \end{bmatrix} = R\bar{\Sigma}R' := V , \quad (1.8)$$

where

$$\text{Cov}(y) = R_1\bar{\Sigma}R_1': := V_{11} , \quad \text{Cov}(y_f, y) = R_1\bar{\Sigma}R_2': := V_{12} , \quad \text{Cov}(y_f) = R_2\bar{\Sigma}R_2' := V_{22} . \quad (1.9)$$

They are all known matrices under the assumptions in (1.1)–(1.6), and will occur in the statistical inference of (1.5). Assumptions in (1.1)–(1.6) are so general that they include almost all LRM s with different structures of covariance matrices as their special cases. Note from (1.5) that under the general assumptions in (1.1)–(1.6), $y$ and $y_f$ are correlated. Hence, it is desirable to give predictions of the future observations $y_f$, as well as $X_f \beta$ and $\epsilon_f$ in (1.4) from the observed response vector $y$ in (1.1) under the assumptions in (1.1)–(1.6). It is of great practical interest to simultaneously identify the important predictors that correspond to both the fixed- and random-effects components in LRM. Some previous and recent work on simultaneous estimations and predictions of combined unknown parameters under regression models can be found in [3, 17, 21, 25, 26].
In order to predict/estimate all unknown parameters in (1.5) simultaneously, we construct a vector containing the fixed effects, random effects, and error terms in (1.5) as follows

$$\phi = F\alpha + G\gamma + H\varepsilon + H_f\varepsilon_f,$$

(1.11)

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

$$E(\phi) = F\alpha, \quad Cov(\phi) = J\tilde{\Sigma}J', \quad Cov\{\phi, y\} = J\tilde{\Sigma}R_{1}', \quad J = [G, H, H_f].$$

(1.12)

Eq. (1.11) contains all possible matrix and vector operations in (1.1)–(1.5) as its special cases. For instance, if $F = T\tilde{X}A$, $G = T\tilde{X}$, and $[H, H_f] = T$, then (1.11) becomes

$$\phi = T\tilde{X}A\alpha + T\tilde{X}\gamma + T\tilde{\varepsilon} = T\tilde{y},$$

(1.13)

which contains $y$, $y_f$, and $\tilde{y}$ as its special cases for different choices of $T$. Another well-known form of $\phi$ in (1.11) is the following target function discussed in [3, 4, 21, 26], which allows the prediction of both $y_f$ and $E(y_f)$,

$$\tau = \lambda y_f + (1 - \lambda)E(y_f) = X_fA\alpha + \lambda X_f\gamma + \lambda \varepsilon_f,$$

(1.14)

where $\lambda (0 \leq \lambda \leq 1)$ is a non-stochastic scalar assigning weights to actual and expected values of $y_f$. Clearly, the problem of predicting a linear combination of the fixed- and random-effects can be formulated as a special case of the general prediction problem on $\phi$ in (1.11). Thus, the simultaneous statistical inference of all unknown parameters in (1.11) is a comprehensive work, and will play prescriptive role for various special statistical inference problems under (1.1) from both theoretical and applied points of view. Note that there are 13 given matrices in (1.1)–(1.6) and (1.11). Hence, statistical inference of $\phi$ in (1.11) is not easy task, we will encounter many tedious matrix operations for the given 13 matrices, as demonstrated in Section 3 below.

The paper is organized as follows. Section 2 introduces the definitions of the BLUPs/BLUEs of all unknown parameters in (1.5). A linear statistic $Ly$ under (1.1), where $L \in \mathbb{R}^{s \times n}$, is
Y. Tian said to have the same expectation with \( \phi \) in (1.11) if and only if \( E(Ly - \phi) = 0 \) holds. If there exists an \( L_0y - \phi \) such that

\[
E(L_0y - \phi) = 0 \quad \text{and} \quad \text{Cov}(Ly - \phi) \succeq \text{Cov}(L_0y - \phi) \quad \text{s.t.} \quad E(Ly - \phi) = 0
\] (2.1)

hold, where \( \text{Cov}(Ly - \phi) = E[(Ly - \phi)(Ly - \phi)'] \) is the matrix mean squared error (MMSE) of \( \phi \) under \( E(Ly - \phi) = 0 \), then the linear statistic \( L_0y \) is defined to be the BLUP of \( \phi \) in (1.11), and is denoted by

\[
L_0y = \text{BLUP}(\phi) = \text{BLUP}(F\alpha + G\gamma + H\varepsilon + H_f\varepsilon_f).
\] (2.2)

If \( G = 0, H = 0, \) and \( H_f = 0 \), in (1.11), the \( L_0y \) satisfying (2.1) is called the BLUE of \( F\alpha \) under (1.1), and is denoted by

\[
L_0y = \text{BLUE}(F\alpha).
\] (2.3)

It should be pointed out that (2.1) can equivalently be converted to certain constrained matrix-valued function optimization problem in the Löwner partial ordering. This kind of equivalences between covariance matrix minimization problems and matrix-valued function minimization problems were firstly characterized in [17]; see also [19]. When \( A = I_p \) and \( \Sigma_{11} = 0, \) (1.1) is the well-known general linear fixed-effects model. In this instance, the work on predictions of new observations was widely considered since 1970s; see, e.g., [5, 7, 8, 9, 10, 11, 12]. On the other hand, (1.1) is a special case of general Linear Mixed-effects Models (LMMs), and some previous results on BLUPs/BLUEs under LMMs can be found in the literature; see, e.g., [1, 6, 16, 18, 19].

The following lemma is well known; see [15].

**Lemma 2.1.** The linear matrix equation \( AX = B \) is consistent if and only if \( r[A, B] = r(A), \) or equivalently, \( AA'B = B. \) In this case, the general solution of the equation can be written in the following parametric form \( X = A^+B + (I-A^+A)U, \) where \( U \) is an arbitrary matrix.

We also need the following known formulas on ranks of matrices; see [13, 22].

**Lemma 2.2.** Let \( A \in \mathbb{R}^{m \times n}, B \in \mathbb{R}^{m \times k}, \) and \( C \in \mathbb{R}^{l \times n}. \) Then

\[
r[A, B] = r(A) + r(E_AB) = r(B) + r(E_BA),
\] (2.4)

\[
r\begin{bmatrix} A \\ C \end{bmatrix} = r(A) + r(CF_A) = r(C) + r(AF_C),
\] (2.5)

\[
r\begin{bmatrix} AA' \\ B' \\ 0 \end{bmatrix} = r[A, B] + r(B).
\] (2.6)
Predictions under LRM

If \( \mathcal{R}(A_1) \subseteq \mathcal{R}(B_1) \), \( \mathcal{R}(A_2) \subseteq \mathcal{R}(B_1) \), \( \mathcal{R}(A_2^\perp) \subseteq \mathcal{R}(B_2^\perp) \) and \( \mathcal{R}(A_3) \subseteq \mathcal{R}(B_2) \), then

\[
\begin{align*}
    r(A_1B_1^+A_2) &= r \begin{bmatrix} B_1 & A_2 \\ A_1 & 0 \end{bmatrix} - r(B_1), \\
    r(A_1B_1^+A_2B_2^+A_3) &= r \begin{bmatrix} B_1 & A_2 & 0 \\ A_1 & 0 & 0 \end{bmatrix} - r(B_1) - r(B_2).
\end{align*}
\]

The following result on analytical solutions of a constrained matrix-valued function optimization problem was given in [25].

**Lemma 2.3.** Let

\[ f(L) = (LC + D)M(LC + D)' \quad \text{s.t.} \quad LA = B, \tag{2.9} \]

where \( A \in \mathbb{R}^{p \times q} \), \( B \in \mathbb{R}^{n \times q} \), \( C \in \mathbb{R}^{p \times m} \), and \( D \in \mathbb{R}^{n \times m} \) are given, \( M \in \mathbb{R}^{n \times m} \) is nd, and the matrix equation \( LA = B \) is consistent. Then, there always exists a solution \( L_0 \) of \( L_0A = B \) such that

\[ f(L) \succ f(L_0) \tag{2.10} \]

holds for all solutions of \( LA = B \). In this case, the matrix \( L_0 \) satisfying (2.10) is determined by the following consistent matrix equation

\[ L_0[A, CMC'A^\perp] = [B, -DMC'A^\perp], \tag{2.11} \]

while the general expression of \( L_0 \) and the corresponding \( f(L_0) \) are given by

\[ L_0 = [B, -DMC'A^\perp][A, CMC'A^\perp] + U[A, CMC'], \tag{2.12} \]

\[ f(L_0) = KMK' - KMC'(A^\perp CMC'A^\perp) + CMC', \tag{2.13} \]

where \( K = BA^\perp C + D \), and \( U \in \mathbb{R}^{n \times p} \) is arbitrary.

Many optimization problems in parametric statistical inferences, as demonstrated below, can be converted to the minimization of (2.9) in the Löwner partial ordering, while analytical solutions to these optimization problems in statistics can be derived from Lemma 2.3. More results on (constrained) quadratic matrix-valued function optimization problems in the Löwner partial ordering can be found in [23, 24].

### 3. Equations and formulas for BLUPs/BLUEs of all unknown parameters in LRM

In what follows, we assume that (1.1) is consistent, i.e., \( y \in \mathcal{F}[XA, V_{11}] \) holds with probability 1. In this section, we first show how to derive the BLUP of the vector \( \phi \) in (1.11), and then give some direct consequences under different assumptions.

**Lemma 3.1.** The vector \( \phi \) in (1.11) is predictable by \( y \) in (1.1) if and only if there exists \( L \in \mathbb{R}^{n \times n} \) such that \( LXA = F \), or equivalently,

\[ \mathcal{F}[(XA)'] \supseteq \mathcal{F}(F'). \tag{3.1} \]
**Proof.** It is obvious that $E(Ly - \phi) = 0 \iff LXA\alpha - F\alpha = 0$ for all $\alpha \iff LXA = F$. From Lemma 2.1, the matrix equation is consistent if and only if (3.1) holds. \(\square\)

**Theorem 3.2.** Assume that $\phi$ in (1.11) is predictable by $y$ in (1.1), namely, (3.1) holds, and let $\tilde{X}$, $R_1$, $V_{ij}$, and $J$ be as given in (1.5), (1.7), (1.9), (1.10), and (1.12). Also denote $\tilde{X} = XA$. Then

\[
E(Ly - \phi) = 0 \quad \text{and} \quad \text{Cov}(Ly - \phi) = \min \Rightarrow L[\tilde{X}, \text{Cov}(y)\tilde{X}^+] = [F, \text{Cov}(\phi, y)\tilde{X}^+].
\]

The matrix equation in (3.2), called the fundamental equation for BLUP, is consistent as well under (3.1). In this case, the general solution of $L$ and BLUP($\phi$) can be written as

\[
\text{BLUP}(\phi) = Ly = \left( [F, J\Sigma R_1\tilde{X}^+] [\tilde{X}, V_{11}\tilde{X}^+]^+ + U[\tilde{X}, V_{11}\tilde{X}^+] \right) y,
\]

where $U \in \mathbb{R}^{s \times n}$ is arbitrary. In particular,

\[
\begin{align*}
\text{BLUP}(X\beta) &= \left( [\tilde{X}, [X, 0, 0]\tilde{R}_1\tilde{X}^+] [\tilde{X}, V_{11}\tilde{X}^+]^+ + U_1[\tilde{X}, V_{11}\tilde{X}^+] \right) y, \\
\text{BLUP}(X_f\beta) &= \left( [X_fA, [X_f, 0, 0]\tilde{R}_1\tilde{X}^+] [\tilde{X}, V_{11}\tilde{X}^+]^+ + U_2[\tilde{X}, V_{11}\tilde{X}^+] \right) y, \\
\text{BLUE}(XA\alpha) &= \left( [\tilde{X}, 0][\tilde{X}, V_{11}\tilde{X}^+]^+ + U_3[\tilde{X}, V_{11}\tilde{X}^+] \right) y, \\
\text{BLUE}(X_fA\alpha) &= \left( [X_fA, 0][\tilde{X}, V_{11}\tilde{X}^+]^+ + U_4[\tilde{X}, V_{11}\tilde{X}^+] \right) y, \\
\text{BLUP}(Xg) &= \left( [0, [X, 0, 0]\tilde{R}_1\tilde{X}^+] [\tilde{X}, V_{11}\tilde{X}^+]^+ + U_5[\tilde{X}, V_{11}\tilde{X}^+] \right) y, \\
\text{BLUP}(X_fg) &= \left( [0, [X_f, 0, 0]\tilde{R}_1\tilde{X}^+] [\tilde{X}, V_{11}\tilde{X}^+]^+ + U_6[\tilde{X}, V_{11}\tilde{X}^+] \right) y, \\
\text{BLUP}(\varepsilon) &= \left( [0, [0, I_n, 0]\tilde{R}_1\tilde{X}^+] [\tilde{X}, V_{11}\tilde{X}^+]^+ + U_7[\tilde{X}, V_{11}\tilde{X}^+] \right) y, \\
\text{BLUP}(\varepsilon_f) &= \left( [0, [0, I_n, I_n]\tilde{R}_1\tilde{X}^+] [\tilde{X}, V_{11}\tilde{X}^+]^+ + U_8[\tilde{X}, V_{11}\tilde{X}^+] \right) y,
\end{align*}
\]

where $U_i$ are arbitrary matrices of appropriate sizes, $i = 1, 2, \ldots, 8$. Further, the following results hold.

(a) $r[\tilde{X}, V_{11}\tilde{X}^+] = r[\tilde{X}, R_1\tilde{\Sigma}], \mathcal{R}[\tilde{X}, V_{11}\tilde{X}^+] = \mathcal{R}[\tilde{X}, R_1\tilde{\Sigma}]$, and $\mathcal{R}(\tilde{X}) \cap \mathcal{R}(V_{11}\tilde{X}^+) = \{0\}$.  
(b) $L_0$ is unique if and only if $r[\tilde{X}, V_{11}] = n$.  
(c) BLUP($\phi$) is unique with probability 1 if and only if $y \in \mathcal{R}[\tilde{X}, V_{11}]$, i.e., (1.1) is consistent.
(d) BLUP(\(\phi\)) satisfies

\[
\begin{align*}
\text{Cov}[\text{BLUP}(\phi)] &= [\mathbf{F}, J\Sigma R'_I \hat{\mathbf{X}}^\perp] \mathbb{E}[\hat{\mathbf{X}}, V_{11} \hat{\mathbf{X}}^\perp] + V_{11}([\mathbf{F}, J\Sigma R'_I \hat{\mathbf{X}}^\perp] \mathbb{E}[\hat{\mathbf{X}}, V_{11} \hat{\mathbf{X}}^\perp])', \\
\text{Cov}(\phi - \text{BLUP}(\phi)) &= J\Sigma' - [\mathbf{F}, J\Sigma R'_I \hat{\mathbf{X}}^\perp] \mathbb{E}[\hat{\mathbf{X}}, V_{11} \hat{\mathbf{X}}^\perp] + V_{11}([\mathbf{F}, J\Sigma R'_I \hat{\mathbf{X}}^\perp] \mathbb{E}[\hat{\mathbf{X}}, V_{11} \hat{\mathbf{X}}^\perp])', \\
\text{Cov}(\phi - \text{BLUP}(\phi)) &= ([\mathbf{F}, J\Sigma R'_I \hat{\mathbf{X}}^\perp] \mathbb{E}[\hat{\mathbf{X}}, V_{11} \hat{\mathbf{X}}^\perp])^\perp R_1 - J\Sigma([\mathbf{F}, J\Sigma R'_I \hat{\mathbf{X}}^\perp] \mathbb{E}[\hat{\mathbf{X}}, V_{11} \hat{\mathbf{X}}^\perp])^\perp R_1 - J').
\end{align*}
\]

**Proof.** By noticing that

\[
\begin{align*}
L\mathbf{y} - \phi &= L\hat{\mathbf{X}}\alpha + LX\gamma + L\varepsilon - F\alpha - G\gamma - H\varepsilon - H_f\varepsilon_f \\
&= (L\hat{\mathbf{X}} - F)\alpha + (LX - G)\gamma + (L - H)\varepsilon - H_f\varepsilon_f,
\end{align*}
\]

we see that the covariance matrix of \(L\mathbf{y} - \phi\) can be written as

\[
\begin{align*}
\text{Cov}(L\mathbf{y} - \phi) &= \text{Cov}((LX - G)\gamma + (L - H)\varepsilon - H_f\varepsilon_f) \\
&= (L[X, I_n, 0] - [G, H, H_f])\Sigma(L[X, I_n, 0] - [G, H, H_f])' \\
&= (LR_1 - J)\Sigma(LR_1 - J)' := f(L).
\end{align*}
\]

In this setting, we see from Lemma 2.3 that the first part of (3.2) is equivalent to finding a solution \(L_0\) of the consistent matrix equation \(L_0\hat{\mathbf{X}} = \mathbf{F}\) such that

\[
f(L) \succeq f(L_0) \quad \text{s.t.} \quad L\hat{\mathbf{X}} = \mathbf{F}
\]

holds in the Löwner partial ordering. Further from Lemma 2.3, there always exists a solution \(L_0\) of \(L_0\hat{\mathbf{X}} = \mathbf{F}\) such that (3.17) holds, and the \(L_0\) is determined by the matrix equation

\[
L_0[\hat{\mathbf{X}}, V_{11} \hat{\mathbf{X}}^\perp] = [\mathbf{F}, J\Sigma R'_I \hat{\mathbf{X}}^\perp],
\]

establishing the matrix equation in (3.2). Solving the equation by Lemma 2.1 gives the \(L_0\) in (3.3). Also from (2.13),

\[
f(L_0) = \text{Cov}(L_0\mathbf{y} - \phi) = ([\mathbf{F}, J\Sigma R'_I \hat{\mathbf{X}}^\perp] \mathbb{E}[\hat{\mathbf{X}}, V_{11} \hat{\mathbf{X}}^\perp])^\perp R_1 - J\Sigma \\
\times ([\mathbf{F}, J\Sigma R'_I \hat{\mathbf{X}}^\perp] \mathbb{E}[\hat{\mathbf{X}}, V_{11} \hat{\mathbf{X}}^\perp])^\perp R_1 - J)',
\]

as required for (3.15). Result (a) is well known.

Results (b) and (c) follow directly from (3.3).
Taking the covariance matrix of (3.3), and simplifying by (1.9) and $R(V_{11}) \subseteq R(\hat{X}, V_{11}\hat{X}^\perp)$ yield (3.12). From (1.11) and (3.3),
\[
\begin{align*}
\text{Cov}\{\text{BLUP}(\phi), \phi\} &= \text{Cov}\{L_0y, \phi\} \\
&= \text{Cov}\left\{\left[\begin{array}{c}
F, J\Sigma R_1'\hat{X}^\perp
\end{array}\right][\hat{X}, V_{11}\hat{X}^\perp]^{+}R_1 \left[\begin{array}{c}
\beta \\
\bar{\varepsilon}
\end{array}\right], J \left[\begin{array}{c}
\beta \\
\bar{\varepsilon}
\end{array}\right]\right\} \\
&= [F, J\Sigma R_1'\hat{X}^\perp][\hat{X}, V_{11}\hat{X}^\perp]^{+}R_1\Sigma'J',
\end{align*}
\]
equation (3.13). Eq. (3.14) follows from (1.12) and (3.12).

The matrix equation in (3.2) shows that the BLUPs/BLUEs of all unknown parameters in (1.5) generally depend on the covariance matrix of the observed random vector $y$, and the covariance matrix between $\phi$ and $y$. Because the matrix equation in (3.2) and the formulas in (3.3)–(3.15) are presented by common operations of the given matrices and their generalized inverses, Theorem 3.2 and its proof in fact provide a standard procedure of handling matrix operations that occur in the theory of BLUPs/BLUEs under general LRMs. From the fundamental equations and formulas in Theorem 3.2, we are now able to derive many new and valuable consequences on properties of BLUPs/BLUEs under various conditions.

**Corollary 3.3.** Let $\phi$ be as given in (1.11). Then, the following results hold.

(a) If $\phi$ is predictable by $y$ in (1.1), then $T\phi$ is predictable by $y$ in (1.1) as well for any matrix $T \in \mathbb{R}^{t \times s}$, and
\[
\text{BLUP}(T\phi) = T\text{BLUP}(\phi)
\] (3.19)
holds.

(b) If $\phi$ is predictable by $y$ in (1.1), then $F\alpha$ is estimable by $y$ in (1.1) as well, and the BLUP of $\phi$ can be decomposed as the sum
\[
\text{BLUP}(\phi) = \text{BLUE}(F\alpha) + \text{BLUP}(G\gamma) + \text{BLUP}(H\varepsilon) + \text{BLUP}(Hf\varepsilon_f)
\] (3.20)
and the following formulas for covariance matrices hold
\[
\text{Cov}\{\text{BLUE}(F\alpha), \text{BLUP}(G\gamma + H\varepsilon + H_f\varepsilon_f)\} = 0,
\] (3.21)
\[
\text{Cov}[\text{BLUP}(\phi)] = \text{Cov}[\text{BLUE}(F\alpha)] + \text{Cov}[\text{BLUE}(G\gamma + H\varepsilon + H_f\varepsilon_f)].
\] (3.22)

(c) If $\alpha$ in (1.1) is estimable by $y$ in (1.1), i.e., $r(XA) = k$, then $\phi$ is predictable by $y$ in (1.1) as well. In this case, the following BLUP/BLUE decomposition equalities
\[
\begin{align*}
\text{BLUP} & \begin{bmatrix}
\alpha \\
\gamma \\
\varepsilon \\
\varepsilon_f
\end{bmatrix} = \begin{bmatrix}
\text{BLUE}(\alpha) \\
\text{BLUP}(\gamma) \\
\text{BLUP}(\varepsilon) \\
\text{BLUP}(\varepsilon_f)
\end{bmatrix}, \\
\text{BLUP}(\phi) &= F\text{BLUE}(\alpha) + G\text{BLUP}(\gamma) + H\text{BLUP}(\varepsilon) + H_f\text{BLUP}(\varepsilon_f)
\end{align*}
\] (3.24)
hold.
Proof. The predictability of $T\phi$ follows from $\mathcal{R}(X'A') \supseteq \mathcal{R}(F') \supseteq \mathcal{R}(F'T')$.

Also from (3.3),

$$\text{BLUP}(T\phi) = \left( [TF, TJ\Sigma R_1' \hat{X}^\perp] [\hat{X}, V_{11} \hat{X}^\perp]^+ + U_1[\hat{X}, V_{11} \hat{X}^\perp]^+ \right) y$$

$$= T \left( [F, J\Sigma R_1' \hat{X}^\perp] [\hat{X}, V_{11} \hat{X}^\perp]^+ + U_1[\hat{X}, V_{11} \hat{X}^\perp]^+ \right) y$$

$$= T \text{BLUP}(\phi),$$

where $U = TU_1$, establishing (3.19).

Note that $Ly$ in (3.3) can be decomposed as

$$Ly = (S_1 + S_2 + S_3 + S_4)y = S_1y + S_2y + S_3y + S_4y,$$

where

$$S_1 = [F, 0][\hat{X}, V_{11} \hat{X}^\perp]^+ + U_1[\hat{X}, V_{11} \hat{X}^\perp]^+,$$

$$S_2 = [0, G, 0, 0][\hat{X}, V_{11} \hat{X}^\perp]^+ + U_2[\hat{X}, V_{11} \hat{X}^\perp]^+,$$

$$S_3 = [0, H, 0][\hat{X}, V_{11} \hat{X}^\perp]^+ + U_3[\hat{X}, V_{11} \hat{X}^\perp]^+,$$

$$S_4 = [0, 0, H_f][\hat{X}, V_{11} \hat{X}^\perp]^+ + U_4[\hat{X}, V_{11} \hat{X}^\perp]^+,$$

and

$$\text{BLUE}(F\alpha) = S_1y, \quad \text{BLUP}(G\gamma) = S_2y, \quad \text{BLUP}(H\varepsilon) = S_3y, \quad \text{BLUP}(H_f\varepsilon_f) = S_4y,$$

establishing (3.20).

We also obtain from (3.3) that

$$\text{Cov} \{\text{BLUE}(F\alpha), \text{BLUP}(G\gamma + H\varepsilon + H_f\varepsilon_f)\}$$

$$= \text{Cov} \left\{ \left( [F, 0][\hat{X}, V_{11} \hat{X}^\perp]^+ + U_1[\hat{X}, V_{11} \hat{X}^\perp]^+ \right) y, \right. \left( [0, J\Sigma R_1' \hat{X}^\perp] [\hat{X}, V_{11} \hat{X}^\perp]^+ + U_2[\hat{X}, V_{11} \hat{X}^\perp]^+ \right) y \right\}$$

$$= [F, 0][\hat{X}, V_{11} \hat{X}^\perp]^+ V_{11}y, [0, J\Sigma R_1' \hat{X}^\perp] [\hat{X}, V_{11} \hat{X}^\perp]^+ \right)'.$$

Applying (2.8) to (3.25) and simplifying, we obtain

$$r(\text{Cov} \{\text{BLUE}(F\alpha), \text{BLUP}(G\gamma + H\varepsilon + H_f\varepsilon_f)\})$$

$$= r \left( [F, 0][\hat{X}, V_{11} \hat{X}^\perp]^+ V_{11}y, [0, J\Sigma R_1' \hat{X}^\perp] [\hat{X}, V_{11} \hat{X}^\perp]^+ \right)$$

$$= r \left[ \begin{array}{ccc} 0 & \hat{X}' & 0 \\ \hat{X}^\perp V_{11} & 0 & 0 \\ [\hat{X}, V_{11} \hat{X}^\perp] & V_{11} & 0 \\ [F, 0] & 0 & 0 \end{array} \right] - 2r[\hat{X}, V_{11} \hat{X}^\perp]$$

$$= r \left[ \begin{array}{ccc} 0 & 0 & \hat{X}' \\ 0 & 0 & 0 \\ [\hat{X}, 0] & V_{11} & 0 \\ [F, 0] & 0 & 0 \end{array} \right] - 2r[\hat{X}, V_{11}]$$

$$\text{Predictions under LRM}$$

9
the coefficient matrices of BLUP. The pair of matrix equations have a common solution if and only if

where by block elementary matrix operations,

\[ r \begin{bmatrix} 0 & \hat{X}' \\ \hat{X} & V_{11} \end{bmatrix} + r[\hat{X}^{-1} V_{11} \hat{X}^{-1}, \hat{X}^{-1} R_1 \hat{\Sigma}'J'] - 2r[\hat{X}, V_{11}] \]

\[ = r \begin{bmatrix} \hat{X} \\ F \end{bmatrix} + r \begin{bmatrix} \hat{X}' \\ V_{11} \end{bmatrix} + r[\hat{X}, V_{11} \hat{X}^{-1}, R_1 \hat{\Sigma}'J'] \]

\[-r(\hat{X}) - 2r[\hat{X}, V_{11}] \quad \text{(by (2.4) and (2.6))} \]

\[ = r[\hat{X}, V_{11}, R_1 \hat{\Sigma}'J'] - r[\hat{X}, V_{11}] \]

\[ = r[\hat{X}, V_{11}] - r[\hat{X}, V_{11}] \quad \text{(by Theorem 3.2(a))} \]

\[ = 0, \]

which implies that Cov{BLUE($F\alpha$), BLUP($G\gamma + H\epsilon + H_f\epsilon_f$)} is a null matrix, establishing (3.21). Eq. (3.22) follows from (3.20) and (3.21). Eqs. (3.23), and (3.24) follow from (1.11), (3.19) and (3.20). □

**Corollary 3.4.** Let

\[ \phi_1 = F_1 \alpha + G_1 \gamma + H_1 \epsilon + H_{f1} \epsilon_{f1}, \quad \phi_2 = F_2 \alpha + G_2 \gamma + H_2 \epsilon + H_{f2} \epsilon_{f2}, \]

where $F_1, F_2 \in \mathbb{R}^{s \times k}, G_1, G_2 \in \mathbb{R}^{s \times p}, H_1, H_2 \in \mathbb{R}^{s \times n}$, and $H_{f1}, H_{f2} \in \mathbb{R}^{s \times n_f}$ are known matrices, and assume that they are predictable by $y$ in (1.1). Then, the following results hold.

(a) The sum $\phi_1 + \phi_2$ is predictable by $y$ in (1.1), and the BLUP of $\phi_1 + \phi_2$ satisfies

\[ \text{BLUP}(\phi_1 + \phi_2) = \text{BLUP}(\phi_1) + \text{BLUP}(\phi_2). \quad (3.26) \]

(b) BLUP($\phi_1$) = BLUP($\phi_2$) $\iff$ $F_1 = F_2$ and $\mathcal{H}(R_1 \hat{\Sigma}'J' - R_1 \hat{\Sigma}'J_2) \subseteq \mathcal{H}(\hat{X})$, where $J_1 = [G_1, H_1, H_{f1}]$ and $J_2 = [G_2, H_2, H_{f2}]$.

**Proof.** Eq. (3.26) follows from (3.20). From Theorem 3.2, the two equations for the coefficient matrices of BLUP($\phi_1$) = $L_1y$ and BLUP($\phi_2$) = $L_2y$ are given by

\[ L_1[\hat{X}, V_{11} \hat{X}^{-1}] = [F_1, J_1 \hat{\Sigma}R_1' \hat{X}^{-1}], \quad L_2[\hat{X}, V_{11} \hat{X}^{-1}] = [F_2, J_2 \hat{\Sigma}R_2' \hat{X}^{-1}]. \]

The pair of matrix equations have a common solution if and only if

\[ r \begin{bmatrix} \hat{X} & V_{11} \hat{X}^{-1} & \hat{X} & V_{11} \hat{X}^{-1} \\ F_1 & J_1 \hat{\Sigma}R_1' \hat{X}^{-1} & F_2 & J_2 \hat{\Sigma}R_2' \hat{X}^{-1} \end{bmatrix} = r[\hat{X}, V_{11} \hat{X}^{-1}, \hat{X}, V_{11} \hat{X}^{-1}], \quad (3.27) \]

where by block elementary matrix operations

\[ r \begin{bmatrix} \hat{X} & V_{11} \hat{X}^{-1} & \hat{X} & V_{11} \hat{X}^{-1} \\ F_1 & J_1 \hat{\Sigma}R_1' \hat{X}^{-1} & F_2 & J_2 \hat{\Sigma}R_2' \hat{X}^{-1} \end{bmatrix} = r \begin{bmatrix} \hat{X} & V_{11} \hat{X}^{-1} & 0 & 0 \\ 0 & 0 & F_2 - F_1 & (J_2 \hat{\Sigma}R_2' - J_1 \hat{\Sigma}R_1') \hat{X}^{-1} \end{bmatrix} \]

\[ = r[\hat{X}, V_{11} \hat{X}^{-1}] + r[F_2 - F_1, (J_2 \hat{\Sigma}R_2' - J_1 \hat{\Sigma}R_1') \hat{X}^{-1}], \]

\[ r[\hat{X}, V_{11} \hat{X}^{-1}, \hat{X}, V_{11} \hat{X}^{-1}] = r[\hat{X}, V_{11} \hat{X}^{-1}]. \]
Hence, (3.27) is equivalent to \([F_2 - F_1, (J_2\hat{\Sigma}R'_1 - J_1\hat{\Sigma}_R'_1)\hat{X}] = 0\), which is further equivalent to (b) by Lemma 2.1.

As direct consequences of Theorem 3.2 and Corollary 3.3, we next give the BLUPs under (1.4).

**Corollary 3.5.** Let \(R, R_1, \) and \(R_2\) be as given in (1.7), and denote \(Z_1 = \text{Cov}\{X_f\beta, y\} = X_f\Sigma_{11}X' + X_f\Sigma_{12}\) and \(Z_2 = \text{Cov}\{\varepsilon_f, y\} = \Sigma_{31}X' + \Sigma_{32}\). Then, the following results hold.

(a) The future observation vector \(y_f\) in (1.4) is predictable by \(y\) in (1.1) if and only if \(\mathcal{R}(\hat{X}') \supseteq \mathcal{R}[(X_fA)']\). In this case,

\[
\text{BLUP}(y_f) = \left( [X_fA, V_{21}\hat{X}]^\perp [\hat{X}, V_{11}\hat{X}]^\perp + U[\hat{X}, V_{11}]^\perp \right) y, \quad (3.28)
\]

and

\[
\text{Cov}[\text{BLUP}(y_f)] = [X_fA, V_{21}\hat{X}]^\perp [\hat{X}, V_{11}\hat{X}]^\perp + V_{11}([X_fA, V_{21}\hat{X}]^\perp [\hat{X}, V_{11}\hat{X}]^\perp)', \quad (3.29)
\]

\[
\text{Cov}[\text{BLUP}(y_f'), y_f] = [X_fA, V_{21}\hat{X}]^\perp [\hat{X}, V_{11}\hat{X}]^\perp + V_{12}, \quad (3.30)
\]

\[
\text{Cov}(y_f - \text{BLUP}(y_f)) = ([X_fA, V_{21}\hat{X}]^\perp [\hat{X}, V_{11}\hat{X}]^\perp + R_1 - R_2)\hat{\Sigma} \times ([X_fA, V_{21}\hat{X}]^\perp [\hat{X}, V_{11}\hat{X}]^\perp + R_1 - R_2)', \quad (3.31)
\]

where \(U \in \mathbb{R}^{n_{y} \times n}\) is arbitrary.

(b) \(X_f\beta\) in (1.4) is predictable by \(y\) in (1.1) if and only if \(\mathcal{R}(\hat{X}') \supseteq \mathcal{R}[(X_fA)']\).

In this case,

\[
\text{BLUP}(X_f\beta) = \left( [X_fA, Z_{11}\hat{X}]^\perp [\hat{X}, V_{11}\hat{X}]^\perp + U_1[\hat{X}, V_{11}\hat{X}]^\perp \right) y, \quad (3.33)
\]

and

\[
\text{Cov}[\text{BLUP}(X_f\beta)] = [X_fA, Z_{11}\hat{X}]^\perp [\hat{X}, V_{11}\hat{X}]^\perp + V_{11}([X_fA, Z_{11}\hat{X}]^\perp [\hat{X}, V_{11}\hat{X}]^\perp)', \quad (3.34)
\]

\[
\text{Cov}[\text{BLUP}(X_f\beta), X_f\beta] = [X_fA, Z_{11}\hat{X}]^\perp [\hat{X}, V_{11}\hat{X}]^\perp + Z_1', \quad (3.35)
\]

\[
\text{Cov}(X_f\beta - \text{BLUP}(X_f\beta)) = X_f\Sigma_{11}X_f - [X_fA, Z_{11}\hat{X}]^\perp [\hat{X}, V_{11}\hat{X}]^\perp + V_{11}([X_fA, Z_{11}\hat{X}]^\perp [\hat{X}, V_{11}\hat{X}]^\perp)', \quad (3.36)
\]

\[
\text{Cov}[X_f\beta - \text{BLUP}(X_f\beta)] = ([X_fA, Z_{11}\hat{X}]^\perp [\hat{X}, V_{11}\hat{X}]^\perp + R_1 - [X_f, 0, 0])\hat{\Sigma} \times ([X_fA, Z_{11}\hat{X}]^\perp [\hat{X}, V_{11}\hat{X}]^\perp + R_1 - [X_f, 0, 0])', \quad (3.37)
\]

where \(U_1 \in \mathbb{R}^{n_{y} \times n}\) is arbitrary.
(c) $\varepsilon_f$ in (1.4) is always predictable by $y$ in (1.1), and

$$BLUP(\varepsilon_f) = \left( Z_2(\bar{X}^\perp V_{11} \bar{X}^\perp)^+ + U_2[\bar{X}, V_{11}]^\perp \right) y, \quad (3.38)$$

$$Cov[BLUP(\varepsilon_f)] = Cov\{BLUP(\varepsilon_f), \varepsilon_f\}$$

$$= Z_2(\bar{X}^\perp V_{11} \bar{X}^\perp)^+ Z'_2,$$  \hspace{1cm} (3.39)

$$Cov[\varepsilon_f - BLUP(\varepsilon_f)] = Cov(\varepsilon_f) - Cov[BLUP(\varepsilon_f)]$$

$$= \Sigma_{33} - Z_2(\bar{X}^\perp V_{11} \bar{X}^\perp)^+ Z'_2,$$  \hspace{1cm} (3.40)

where $U_2 \in \mathbb{R}^{n_f \times n}$ is arbitrary.

Finally, we give a group of fundamental decomposition equalities of the BLUPs of $y$, $y_f$, and $\tilde{y}$ in (1.5).

**Corollary 3.6.** The vector $\tilde{y}$ in (1.5) is predictable by $y$ in (1.1) if and only if $\mathcal{R}(Xa) \supseteq \mathcal{R}(X_f a)$. In this case, the following decomposition equalities

$$y = BLUP(y) = BLUE(\textbf{X}a) + BLUP(\textbf{X}\gamma) + BLUP(\varepsilon), \quad (3.41)$$

$$BLUP(y_f) = BLUE(X_f a) + BLUP(X_f \gamma) + BLUP(\varepsilon_f), \quad (3.42)$$

$$BLUP(\tilde{y}) = \begin{bmatrix} BLUP(y) \\ BLUP(y_f) \end{bmatrix} = \begin{bmatrix} y \\ BLUP(y_f) \end{bmatrix} \quad (3.43)$$

always hold.

The additive decomposition equalities of the BLUPs in (3.41) and (3.42) are in fact built-in restrictions to BLUPs/BLUEs, which sufficiently demonstrate the key roles of BLUPs/BLUEs in statistical inferences of LRM. Some previous discussions on built-in restrictions to BLUPs/BLUEs can be found; e.g., in [2, 14, 19, 20].

### 4. Remarks.

This paper established a general theory on BLUPs/BLUEs under LRM with original and future observations, and obtained many equations and formulas for calculating BLUPs/BLUEs of all unknown parameters in the LRM via solving a constrained quadratic matrix-valued function optimization problem in the Löwner partial ordering. Because the BLUPs/BLUEs in the previous sections are formulated by common operations of the given matrices and their generalized inverses in LRM, we can easily derive many new mathematical and statistical properties of the BLUPs/BLUEs under various assumptions. It is expected more interesting results on statistical inferences of LRM can be derived from the equations and formulas in the previous sections. Here we mention a few:

(a) Derive closed-form formulas for calculating the ranks and inertias of the difference $Cov[\phi - BLUP(\phi)] - A$, and use them to derive necessary and sufficient conditions for the following equality and inequalities

$$Cov[\phi - BLUP(\phi)] = A \quad (\succ A \quad \succ A, \quad \prec A, \quad \preceq A)$$

to hold under the assumptions in (1.1)–(1.12), where $A$ is symmetric matrix, say, $A = Cov(\phi) - Cov[BLUP(\phi)]$. 


Predictions under LRM

(b) Establish necessary and sufficient conditions for the following decomposition equalities

\[

cov[\text{BLUP}(\phi)] = cov[\text{BLUE}(F\alpha)] + cov[\text{BLUP}(G\gamma)]
\]

\[
+cov[\text{BLUP}(H\epsilon)] + cov[\text{BLUP}(H_f\epsilon_f)],
\]

\[
cov[\phi - \text{BLUP}(\phi)] = cov[\text{BLUE}(F\alpha)] + cov[\text{BLUP}(G\gamma - \text{BLUP}(G\gamma)]
\]

\[
+cov[H\epsilon - \text{BLUP}(H\epsilon)]
\]

\[
+cov[H_f\epsilon_f - \text{BLUP}(H_f\epsilon_f)]
\]

to hold respectively under the assumptions in (1.1)–(1.12).

(c) Establish necessary and sufficient condition for BLUP(\phi) = \phi (like fixed points of matrix map) to hold. A special case is that BLUP(Ty) = Ty always holds for any matrix T, and is this case unique?

It is expected that this type of work will bring deep understanding of statistical inferences of BLUPs/BLUEs from many new aspects.

Finally, we give a general matrix formulation and derivation of BLUPs under random vector equations. Note that (1.1) is a special case of the following equation of random vectors

\[
A_1y_1 + A_2y_2 + A_3y_3 = 0,
\]

where A_1, A_2, and A_3 are three random matrices of appropriate sizes, and y_1, y_2 and y_3 are three random vectors of appropriate sizes satisfying

\[
E\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix},
\]

\[
cov\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} \\ \Sigma_{21} & \Sigma_{22} & \Sigma_{23} \\ \Sigma_{31} & \Sigma_{32} & \Sigma_{33} \end{bmatrix},
\]

in which, b_1, b_2 and b_3 are constant vectors, or fixed but unknown parameter vectors. In this setting, taking expectation and covariance matrix of (4.1) yields

\[
A_1b_1 + A_2b_2 + A_3b_3 = 0
\]

and

\[
[A_1, A_2, A_3] \begin{bmatrix} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} \\ \Sigma_{21} & \Sigma_{22} & \Sigma_{23} \\ \Sigma_{31} & \Sigma_{32} & \Sigma_{33} \end{bmatrix} \begin{bmatrix} A_1' \\ A_2' \\ A_3' \end{bmatrix} = 0.
\]

Assume now that one of y_1, y_2 and y_3 is observed, say, y_1, and the other two are unobservable, and we want to predict the vector L_2y_2 + L_3y_3 from (4.1) and (4.2). In this case, we let

\[
S = \{L_1y_1 - L_2y_2 - L_3y_3 \mid E(L_1y_1 - L_2y_2 - L_3y_3) = 0\}.
\]

If there exists matrix \(\hat{L}_1\) such that

\[
\begin{bmatrix} \text{cov}(L_1y_1 - L_2y_2 - L_3y_3) \\ \text{s.t. } L_1y_1 - L_2y_2 - L_3y_3 \in S, \end{bmatrix} \]

\[
\Rightarrow \text{cov}(\hat{L}_1y_1 - L_2y_2 - L_3y_3)
\]

(4.5)
the linear statistic $\hat{L}_1 y_1$ is defined to the BLUP of $L_2 y_2 + L_3 y_3$ under (4.1) and is denoted by $\hat{L}_1 y_1 = \text{BLUP}(L_2 y_2 + L_3 y_3)$. Note that

$$\text{Cov}(L_1 y_1 - L_2 y_2 - L_3 y_3) = [L_1, -L_2, -L_3] \begin{bmatrix} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} \\ \Sigma_{21} & \Sigma_{22} & \Sigma_{23} \\ \Sigma_{31} & \Sigma_{32} & \Sigma_{33} \end{bmatrix} \begin{bmatrix} L_1' \\ -L_2' \\ -L_3' \end{bmatrix} := f(L_1).$$

Thus, (4.5) is equivalent to

$$f(L_1) \succ f(\hat{L}_1) \quad \text{s.t.} \quad L_1 b_1 = L_2 b_2 + L_3 b_3. \quad (4.6)$$

By a similar approach as given in the previous sections, we can establish a group of equations as follows

$$A_1 y_1 + \text{BLUP}(A_2 y_2) + \text{BLUP}(A_3 y_3) = 0 \quad \text{if } y_1 \text{ is observable},$$

$$\text{BLUP}(A_1 y_1) + A_2 y_2 + \text{BLUP}(A_3 y_3) = 0 \quad \text{if } y_2 \text{ is observable},$$

$$\text{BLUP}(A_2 y_1) + \text{BLUP}(A_2 y_2) + A_3 y_3 = 0 \quad \text{if } y_3 \text{ is observable}.$$ 

Eqs. (4.1)–(4.3) contain almost all linear structures occurred in regression analysis. It is expected that more valuable results can be obtained on (4.1)–(4.6), which can serve as a mathematical foundation under various regression model assumptions.

REFERENCES

Predictions under LRM


