SIMULTANEOUS OPTIMAL PREDICTIONS UNDER TWO SEEMINGLY UNRELATED LINEAR RANDOM-EFFECTS MODELS

YONGGE TIAN* AND PENGYANG XIE

College of Business and Economics, Shanghai Business School, Shanghai, China

(Communicated by Aviv Gibali)

ABSTRACT. This paper considers simultaneous optimal prediction and estimation problems in the context of linear random-effects models. Assume a pair of seemingly unrelated linear random-effects models (SULREMs) with the random-effects and the error terms correlated. Our aim is to find analytical formulas for calculating best linear unbiased predictors (BLUPs) of all unknown parameters in the two models by means of solving a constrained quadratic matrix optimization problem in the Löwner sense. We also present a variety of theoretical and statistical properties of the BLUPs under the two models.

1. Introduction. Linear regression models are classic issues in statistical theory and are the common roots of many branches of current statistical theory. Although there has been a relatively systematical research results concerning linear regression models and their applications in the past centuries, one can still propose many theoretical and applied problems on linear models and approach these problems by way of various mathematical and statistical tools. In statistical data analysis and inference, people often encounter linear regression models that include random effects, or namely, linear random-effects models (LREMs). Such a kind of models are commonly used to analyze longitudinal and correlated data, which occur in a variety of fields including biostatistics, public health, psychometrics, educational measurement, and sociology. LREMs are available to account for the variability of model parameters due to different factors that influence a response variable. The problems of statistical inference on LREMs is now an important part in the data analysis, and a huge amount of literature on LREMs spreads in the fields of statistics and other disciplines. Seemingly unrelated linear random-effects models (SULREMs) are extensions of LREMs that allow correlated errors between the matrix regression equations in the models.

We now introduce the modeling framework concerning LREMs. In statistical analysis of data collected from different time periods, we may meet with the cases where two or more observable random vectors. For example, let $y_1$ and $y_2$ be two $n_1 \times 1$ and $n_2 \times 1$ vectors of observable random variables that have the following

2020 Mathematics Subject Classification. Primary: 62F10, 62J05; Secondary: 90C22, 90C90.

Key words and phrases. SULREM, predictability, estimability, BLUP, BLUE, generalized inverses, rank formula.

* Corresponding author: Yongge Tian.
two different model structures

\[ M_1 : y_1 = X_1 \beta_1 + \varepsilon_1, \]
\[ M_2 : y_2 = X_2 \beta_2 + \varepsilon_2, \]

where \( y_i \in \mathbb{R}^{n_i \times 1} \) are vectors of observable response variables, \( X_i \in \mathbb{R}^{n_i \times p_i} \) are known matrices of arbitrary ranks, \( \varepsilon_i \in \mathbb{R}^{n_i \times 1} \) are vectors of unobservable random errors, \( \beta_i \in \mathbb{R}^{p_i \times 1} \) are an unknown vectors of satisfying

\[ \beta_i = A_i \alpha_i + \gamma_i, \quad i = 1, 2, \]

where \( A_i \in \mathbb{R}^{p_i \times k_i} \) are two known matrices of arbitrary ranks, \( \alpha_i \in \mathbb{R}^{k_i \times 1} \) are two vectors of fixed but unknown parameters, and \( \gamma_i \in \mathbb{R}^{p_i \times 1} \) are two vectors of unobservable random variables. In this situation, (1.1) and (1.2) are also classified as two-level hierarchical linear models; see e.g., [2, 8, 9, 29] for more exposition of the background of hierarchical linear models in statistical data analysis and inference. On the other hand, LREMs can be viewed as special forms of linear mixed-effects models. To see this, we substitute (1.3) into (1.1) and (1.2) to obtain the following two seemingly unrelated linear mixed-effects models

\[ M_1 : y_1 = X_1 A_1 \alpha_1 + X_1 \gamma_1 + \varepsilon_1, \]
\[ M_2 : y_2 = X_2 A_2 \alpha_2 + X_2 \gamma_2 + \varepsilon_2 \]

with the fixed-effects vector \( \alpha_i \) and the random-effects vector \( \gamma_i, \quad i = 1, 2. \) The two models are said to be seemingly unrelated because there are no common unknown parameter vectors in them. However, there still exists a possibility to do statistical analysis simultaneously and obtain more accurate inference results concerning the two models under certain assumptions. One of such cases is to assume that the random vectors in the two models are correlated. In this paper, we assume that the expectation and the covariance matrix of the joint vector in (1.4) and (1.5) have the following general forms

\[
\begin{bmatrix}
\gamma_1 \\
\gamma_2 \\
\varepsilon_1 \\
\varepsilon_2
\end{bmatrix}
\begin{bmatrix}
0 \\
0 \\
0 \\
0
\end{bmatrix},
\begin{bmatrix}
\Sigma_{11} & \Sigma_{12} & \Sigma_{13} & \Sigma_{14} \\
\Sigma_{21} & \Sigma_{22} & \Sigma_{23} & \Sigma_{24} \\
\Sigma_{31} & \Sigma_{32} & \Sigma_{33} & \Sigma_{34} \\
\Sigma_{41} & \Sigma_{42} & \Sigma_{43} & \Sigma_{44}
\end{bmatrix}
\triangleq \Sigma
\]

(1.6)

\[
\begin{bmatrix}
\Sigma_{11} & \Sigma_{12} \\
\Sigma_{21} & \Sigma_{22}
\end{bmatrix}
\triangleq \Sigma_1,
\begin{bmatrix}
\Sigma_{31} & \Sigma_{32} \\
\Sigma_{41} & \Sigma_{42}
\end{bmatrix}
\triangleq \Sigma_2
\]

(1.7)

where \( \Sigma \in \mathbb{R}^{(n+p) \times (n+p)} \) is a known nonnegative definite matrix of arbitrary rank and the submatrices \( \Sigma_{ij} \) are nonzero for \( i \neq j \), and \( n = n_1 + n_2 \) and \( p = p_1 + p_2 \). Here we give no further restrictions to the patterns of the submatrices \( \Sigma_{ij} \) in (1.6) although they are usually taken as certain prescribed forms for a given linear random-effects model in the statistical literature. In other words, if \( \Sigma \) is assumed to be unknown or is given with some parametric forms, such as,

\[ \Sigma = \text{diag}\{\Sigma_{11}, \Sigma_{22}, \Sigma_{33}, \Sigma_{44}\}, \Sigma = \text{diag}\{\sigma^2_1 I_{p_1}, \sigma^2_2 I_{p_2}, \sigma^2_3 I_{n_1}, \sigma^2_4 I_{n_2}\}, \Sigma = \sigma^2_{p+n}, \]

etc, where \( \sigma^2 \) and \( \sigma^2 \) are arbitrary positive scaling factors. In practice, if the matrices or parameters are unknown, people may firstly estimate them using the observed data in (1.1) and (1.2) and then substitute them into the models to make further statistical inference to (1.1) and (1.2). In order to make inference under (1.6), simultaneously, we assemble the two regression equations in (1.4) and (1.5)
as follows

\[ \mathbb{M} : \ y = X\beta + \epsilon = X\Lambda\alpha + X\gamma + \epsilon, \]  

(1.8)

where

\[
y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}, \quad X = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix}, \quad A = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix},
\]

\[
\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}, \quad \alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}, \quad \gamma = \begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix}, \quad \epsilon = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \end{bmatrix}.
\]

In this situation, (1.1) and (1.2) can be obtained from the transformed model

\[ T_1y = T_1X\beta + T_1\epsilon = T_1X\Lambda\alpha + T_1X\gamma + T_1\epsilon \]  

(1.9)

for \( T_1 = [I_{n1}, 0] \) and \( T_2 = [0, I_{n2}] \). So that (1.4) and (1.5) can also be regarded as two sub-sample models of the specified linear mixed model in (1.8). The transmitted model (1.8) is simpler in form, which will be utilized when predicting and estimating the unknown parameter vectors in the two models (1.4) and (1.5). As the amount of data in the simultaneous model increases, we should make full use of the simultaneous model to carry out statistical inference instead of using a separate model. On the other hand, people are also interested in the relationship between the estimators under the two separated models, and therefore want to consider comparison problems of estimators under two models.

Under the assumptions and notation in (1.1)–(1.8), a group of subsequent formulas for the expectations and covariance matrices of \( y_i \) and \( y \) in (1.1), (1.2), and (1.8) are given as follows

\[
E(y) = X\Lambda\alpha, \quad E(y_1) = X_1\Lambda_1\alpha_1, \quad E(y_2) = X_2\Lambda_2\alpha_2, \quad \]  

(1.10)

\[
\text{Cov}(y) = \text{Cov}(X\gamma + \epsilon) = [X, I_{n}]\Sigma[X, I_{n}]', \quad \tilde{X}\tilde{\Sigma}\tilde{X}' \triangleq V, \]  

(1.11)

\[
\text{Cov}(y_1) = \text{Cov}(X_1\gamma_1 + \epsilon_1) = [X_1, I_{n_1}]\Sigma_1[X_1, I_{n_1}]', \quad \tilde{X}_1\tilde{\Sigma}_1\tilde{X}_1' \triangleq V_1, \]  

(1.12)

\[
\text{Cov}(y_2) = \text{Cov}(X_2\gamma_2 + \epsilon_2) = [X_2, I_{n_2}]\Sigma_2[X_2, I_{n_2}]', \quad \tilde{X}_2\tilde{\Sigma}_2\tilde{X}_2' \triangleq V_2. \]  

(1.13)

These covariance matrices are known under the assumptions in (1.1)–(1.8), and will occur in the problems of statistical inference under (1.4)–(1.8), as demonstrated in Section 4 below.

Prediction analysis is a general inference method for predicting the accuracy of quantitative experiments. The method is applicable to experiments in which the data is to be analyzed by means of various optimization methods, such as the least squares method, the weighted least squares method, the best linear unbiased prediction method, etc. A convenient way of simultaneously estimating/predicting all unknown parameters in (1.4), (1.5), and (1.8) is to construct three general vectors as follows

\[
\phi_i = F_i\alpha_i + G_i\gamma_i + H_i\epsilon_i, \quad i = 1, 2, \]

(1.14)

\[
\phi = F\alpha + G\gamma + H\epsilon, \]

(1.15)

which encompass all the unknown vectors in (1.4), (1.5), and (1.8) as their special cases, where \( F_i \in \mathbb{R}^{s \times k_i}, \ F \in \mathbb{R}^{s \times (k_1 + k_2)}, \ G_i \in \mathbb{R}^{s \times p_i}, \ G \in \mathbb{R}^{s \times p}, \ H_i \in \mathbb{R}^{s \times n_i}, \) and \( H \in \mathbb{R}^{s \times n} \) are known matrices of arbitrary ranks. So that (1.14) and (1.15) include all vector operations in (1.1)–(1.9) as their special cases. In these settings, we can
readily obtain the following results
\[
E(\phi_i) = F_i \alpha_i, \quad E(\phi) = FA, \quad \text{(1.16)}
\]
\[
\text{Cov}(\phi_i) = [G_i, H_i] \Sigma_i [G_i, H_i]', \quad \text{Cov}(\phi) = [G, H] \Sigma [G, H]', \quad \text{(1.17)}
\]
\[
\text{Cov}(\phi_i, y_i) = \text{Cov}(\phi_i, X_i, \gamma_i + \epsilon_i) = [G_i, H_i] \Sigma_i [X_i, I_n]', \quad \text{(1.18)}
\]
for \(i = 1, 2\). These preparations show that we can estimate/predict (1.14) and (1.15) from (1.1) and (1.2) separately or simultaneously. This idea for constructing general vectors as given in (1.14) and (1.15) was first given in [26] who showed a lemma on optimization of a matrix function in the L"owner partial ordering and established a unified theory of linear estimations/predictions of all unknown parameters in general linear models with fixed or mixed effects; see also [28, Lemma 4.7]. In addition, the work on separate and simultaneous estimations/predictions of unknown parameters in different models can be found in [1, 3–6, 10–14, 16, 25, 30, 38–40]; some new results concerning simultaneous linear estimations/predictions of all unknown parameters in LREMs with original and future observations were obtained in [33, 34] by solving certain constrained quadratic matrix-valued function optimization problems in the L"owner partial ordering.

To account for general prediction/estimation problems of unknown parameters in a given linear regression model, it is common practice to first adopt a feasible procedure for obtaining exact expressions of predictors/estimators of the unknown parameters in the model. Tian recently developed an analytical method in [33, 34] to solve certain types of constrained quadratic matrix-valued function optimization problem in the L"owner partial ordering, and used the method to examine some simultaneous linear estimations/predictions of all unknown parameters in LREMs with original and future observations; see also [7, 11, 15, 17, 18, 31, 32, 35–37] for a series of related approaches.

Similarly to the preceding work on LREMs, we are able to derive a group of theoretical inference conclusions under the assumptions in (1.1)–(1.18), including the analytical formulas for calculating the best linear unbiased predictors (BLUPs) of \(\phi_i\) and \(\phi\) under the general assumptions in (1.1)–(1.18), and various mathematical and statistical properties and performances of these BLUPs under the given assumptions.

2. Notation and preliminaries. Let \(\mathbb{R}^{m \times n}\) denote the collection of all \(m \times n\) real matrices, \(A', r(A),\) and \(\mathcal{R}(A)\) denote the transpose, the rank, and the range (column space) of a matrix \(A \in \mathbb{R}^{m \times n}\), respectively, \(I_m\) denote the identity matrix of order \(m\). The Moore–Penrose generalized inverse of \(A\), denoted by \(A^+\), is defined to be the unique solution \(X\) satisfying the four Penrose matrix equations \(AGA = A, GAG = G, (AG)' = AG,\) and \((GA)' = GA\). The Moore–Penrose inverse of a matrix \(A\) was specially studied and recognized because \(AA^+, A^+A, I_m - AA^+,\) and \(I_n - A^+A\) are orthogonal projectors onto the ranges and kernels of \(A\) and \(A^*\), respectively, so that it optimizes a number of interesting properties of many matrix computation problems. In this paper, we denote by \(A^\perp = E_A = I_m - AA^+\) and \(F_A = I_n - A^+A\) the two orthogonal projectors induced from \(A\) ("\(\perp\)" denotes the orthogonal projector onto the orthogonal complement of the range of a matrix), both of which 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 inequality \(A \succ B\) in the L"owner partial ordering if \(A - B\) is nonnegative definite. We refer the reader to the literature [19, 20, 22, 27] for expositions of generalized inverse of matrices and applications to linear models.
It is well known that the statistics provides mathematicians with various challenging and exciting problems at different levels, since most of the problems in statistics arise from the real-life activities. To solve these problems, statisticians utilize knowledge from all parts of mathematics, from those very abstract to numerical computation and interpretation of the results. Moreover, every statistician is expected to find an optimal solution to a real problem by appeal to various optimization techniques. There are plenty of classic and novel discussions in literature on derivations and representations of BLUPs under linear regression models, which motivate some deeper considerations and explorations of universal-algebraic methods dealing with the BLUP problems. Subsequently, we present two known fundamental and significant results and facts in mathematics concerning analytic solutions of a matrix equation and a constrained quadratic matrix optimization problem, which we shall use as active study tools in the establishment of BLUP theory under the preceding model assumptions.

**Lemma 2.1** ([21]). 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.

**Lemma 2.2** ([33]). Let

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

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, $M \in \mathbb{R}^{m \times m}$ is nnd, and the matrix equation $LA = B$ is consistent. Then, there always exists a solution $L_0$ of $L_0A = B$ such that

$$f(L) \succ f(L_0)$$

holds for all solutions of $LA = B$. In such case, the matrix $L_0$ satisfying the above inequality is determined by the following consistent matrix equation

$$L_0 [A, CMC'A^+] = [B, -DMCA^+].$$

In this case, the general expression of $L_0$ and the corresponding $f(L_0)$ and $f(L)$ are given by

$$L_0 = \arg \min_{LA = B} f(L) = [B, -DMCA^+] [A, CMC'A^+] + U [A, CMC']^{-},$$

$$f(L_0) = \min_{LA = B} f(L) = KMK' - KMC'TCMK',$$

$$f(L) = f(L_0) + (LC + D)MC'TCM(LC + D)'$$

$$= f(L_0) + (LCMC'A^+ + DMCA^+) T (LCMC'A^+ + DMCA^+)',$$

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

3. Consistency of SULREMs and predictability of all unknown parameters in SULREMs. Parametric statistical inference is concerned with predication/estimation of unknown parameters in a given parametric regression model, which is certain mathematical and computational process of drawing conclusions about scientific truths hidden behind the observed data. There are many modes of performing statistical inference, including statistical modeling, data oriented strategies, and explicit use of designs and randomization in analyses. In our approach to
the simultaneous optimal predictions under two seemingly unrelated linear random-effects models described above, we use the conventional concepts, definitions, and statistical techniques in the mainstream regression theory.

We first introduce a couple of notations and definitions on consistency and estimability under the general assumptions in (1.1)–(1.18). Throughout, we denote by

$$\tilde{X}_i = X_iA_i, \quad \tilde{X}_i = [X_i, I_{n_i}], \quad J_i = [G_i, H_i], \quad C_i = \text{Cov}\{\phi_i, y_i\} = J_i\Sigma_i'$$

for $i = 1, 2$. Under the assumptions in (1.1), (1.2), (1.6), and (1.7), we can deduce that

$$E\left(\frac{I_{n_i} - [\tilde{X}_i, V_i][\tilde{X}_i, V_i]^+}{\tilde{X}_i}y_i\right)$$

$$= (I_{n_i} - [\tilde{X}_i, V_i][\tilde{X}_i, V_i]^+)^+\tilde{X}_i\alpha_i = 0,$$

$$\text{Cov}\left(\frac{I_{n_i} - [\tilde{X}_i, V_i][\tilde{X}_i, V_i]^+}{\tilde{X}_i}y_i\right)$$

$$= (I_{n_i} - [\tilde{X}_i, V_i][\tilde{X}_i, V_i]^+)^+V_i(I_{n_i} - [\tilde{X}_i, V_i][\tilde{X}_i, V_i]^+)^T = 0$$

for $i = 1, 2$. These equalities imply that

$$y_i \in \mathcal{R}[\tilde{X}_i, V_i] \text{ hold with probability 1, } i = 1, 2.$$

In this case, (1.1) and (1.2) are said to be consistent, respectively (cf. [23, 24]).

It is well known in statistics that the establishment of the BLUP theory under linear regression models is quite straightforward, which only requires that expectations and covariance matrices related to unknown parameters in the models are given. We next introduce the definitions of the predictability and the BLUPs of $\phi_i$ and $\phi$ in (1.14) and (1.15) for $i = 1, 2$.

**Definition 3.1.** The vectors $\phi_i$ in (1.14) are said to be predictable under (1.1) and (1.2), respectively, if there exist linear statistics $L_iy_i$ with $L_i \in \mathbb{R}^{s \times n_i}$ such that $E(L_iy_i - \phi_i) = 0$ hold, $i = 1, 2$. The $\phi$ in (1.15) is said to be predictable by $y$ in (1.8) if there exists a linear statistic $Ly$ with $L \in \mathbb{R}^{s \times n}$ such that $E(Ly - \phi) = 0$ holds.

**Definition 3.2.** Let $\phi_1$, $\phi_2$, and $\phi$ be as given in (1.14) and (1.15), respectively. If there exist $L_iy_i$ such that

$$\text{Cov}(L_iy_i - \phi_i) = \min \text{ s.t. } E(L_iy_i - \phi_i) = 0, \quad i = 1, 2$$

hold in the Löwner partial ordering, the linear statistics $L_iy_i$ are defined to be the BLUPs of $\phi_i$ in (1.14), and are denoted by

$$L_iy_i = \text{BLUP}_{\mathcal{M}_i}(\phi_i) = \text{BLUP}_{\mathcal{M}_i}(F_i\alpha_i + G_i\gamma_i + H_i\varepsilon_i), \quad i = 1, 2.$$  

(3.3)

If $G_i = 0$ and $H_i = 0$, the linear statistics $L_iy_i$ in (3.3) are the well-known BLUEs of $F_i\alpha_i$ under (1.1) and (1.2), and are denoted by

$$L_iy_i = \text{BLUE}_{\mathcal{M}_i}(F_i\alpha_i), \quad i = 1, 2.$$  

(3.4)

If there exists an $Ly$ such that

$$\text{Cov}(Ly - \phi) = \min \text{ s.t. } E(Ly - \phi) = 0$$

holds in the Löwner partial ordering, the linear statistic $Ly$ is defined to be the BLUP of $\phi$ in (1.15), and is denoted by

$$Ly = \text{BLUP}_{\mathcal{M}}(\phi) = \text{BLUP}_{\mathcal{M}}(F\alpha + G\gamma + H\varepsilon).$$  

(3.6)
If $G = 0$ and $H = 0$, the linear statistic $L_y$ in (3.6) is the well-known BLUE of $F_\alpha$ under (1.8) and is denoted by

$$Ly = \text{BLUE}_{\mathcal{M}}(F_\alpha) .$$

(3.7)

Note from (1.4), (1.5) and (1.14) that

$$L_i y_i - \phi_i = L_i \tilde{X}_i \alpha_i + L_i X_i \gamma_i + L_i \varepsilon_i - F_i \alpha_i - G_i \gamma_i - H_i \varepsilon_i = (L_i \tilde{X}_i - F_i) \alpha_i + (L_i X_i - G_i) \gamma_i + (L_i - H_i) \varepsilon_i, \quad i = 1, 2 .$$

Then, the expectations and covariance matrices of $L_i y_i - \phi_i$ can be written as

$$E(L_i y_i - \phi_i) = (L_i \tilde{X}_i - F_i) \alpha_i, \quad i = 1, 2,$$

$$\text{Cov}(L_i y_i - \phi_i) = \text{Cov}[(L_i X_i - G_i) \gamma_i + (L_i - H_i) \varepsilon_i] = [L_i X_i - G_i, L_i - H_i] \Sigma_i [L_i X_i - G_i, L_i - H_i]^\prime$$

$$= (L_i \tilde{X}_i - J_i) \Sigma_i (L_i \tilde{X}_i - J_i)^\prime \overset{\Delta}{=} f_i(L_i), \quad i = 1, 2 .$$

With these definitions and statistical facts in mind, we can convert the constrained covariance matrix minimization problem in (3.2) to a underlying mathematical minimization problem on quadratic matrix-valued function minimization problem of $f_i(L_i)$ subject to $L_i \tilde{X}_i = F_i$, $i = 1, 2 .

4. Main results. From the definitions, the BLUPs under LREMs in (3.2) and (3.5) are equivalent to certain mathematical optimization problems that are characterized by given objective and constraint functions. After preparing the background materials, we are now ready to establish the BLUP theory under the assumptions in (1.1)–(1.18).

**Theorem 4.1.** Let $\mathcal{M}_1$ and $\mathcal{M}_2$ be as given in (1.1) and (1.2), respectively. Then, the $\phi_i$ in (1.14) are predictable by $y_i$ in (1.1) and (1.2) respectively if and only if

$$\mathcal{R}(\tilde{X}_i^\prime) \supseteq \mathcal{R}(F_i^\prime), \quad i = 1, 2 .$$

(4.1)

In these cases,

$$\text{Cov}(L_i y_i - \phi_i) = \min s.t. \ E(L_i y_i - \phi_i) = 0$$

$$\Leftrightarrow L_i [\tilde{X}_i, V_i \tilde{X}_i] = [F_i, C_i \tilde{X}_i], \quad i = 1, 2 .$$

(4.2)

The matrix equations in (4.2), called the fundamental BLUP equations of $\phi_i$ under (1.1) and (1.2), respectively, are consistent as well under (4.1), while the general solution $\tilde{L}_i$ of the equations and the corresponding BLUP$_{\mathcal{M}_i}(\phi_i)$ can be written as

$$\text{BLUP}_{\mathcal{M}_i}(\phi_i) = L_i y_i = \left[ F_i, C_i \tilde{X}_i + \tilde{Y}_i, V_i \tilde{X}_i \right]^\dagger + U_i [\tilde{X}_i, V_i]^\dagger \right] y_i ,$$

(4.3)

where $U_i \in \mathbb{R}^{r \times n_i}$ are arbitrary, $i = 1, 2$. In addition, the following results hold.

(a) $r[\tilde{X}_i, V_i \tilde{X}_i] = r[\tilde{X}_i, \tilde{X}_i V_i] = r[\tilde{X}_i, V_i], and \mathcal{R}(\tilde{X}_i, V_i \tilde{X}_i) = \mathcal{R}(\tilde{X}_i, \tilde{X}_i V_i)$

(b) $L_i$ is unique if and only if $r[\tilde{X}_i, V_i] = n_i, i = 1, 2 .

(c) BLUP$_{\mathcal{M}_i}(\phi_i)$ is unique if and only if $y_i \in \mathcal{R}(\tilde{X}_i, V_i)$ holds with probability 1, $i = 1, 2$, i.e., (1.1) and (1.2) are consistent, respectively.

(d) The covariance matrix of BLUP$_{\mathcal{M}_i}(\phi_i)$, as well as the covariance matrix between BLUP$_{\mathcal{M}_i}(\phi_i)$ and $\phi_i$ are unique, and satisfy the equalities.
Proof. It is obvious from (1.14) that equivalent to finding solutions $\hat{\mathbf{Y}}_i$. From Lemma 2.1, the matrix equations are consistent respectively if and only if

$$\text{Cov}(\mathbf{B}, \mathbf{A}) = \text{Cov}(\mathbf{A}, \mathbf{B})$$

(4.4)

Thus establishing (4.5). The two equalities in (4.6) follow from (1.17) and (4.4). From (3.1) and (4.3),

$$\text{Cov}\{\mathbf{B}, \mathbf{A}\} = \text{Cov}(\mathbf{A}, \mathbf{B})$$

(4.7)

for $i = 1, 2$.

(c) The following BLUP decomposition equalities hold

$$\text{BLUP}_{\mathcal{M}_i}(\mathbf{A}) = \text{BLUE}_{\mathcal{M}_i}(\mathbf{F}, \mathbf{A}) + \text{BLUP}_{\mathcal{M}_i}(\mathbf{G}, \mathbf{G}) + \text{BLUP}_{\mathcal{M}_i}(\mathbf{H}, \mathbf{E})$$

(4.8)

for i = 1, 2.

(f) If $\phi_1$ and $\phi_2$ are predictable under (1.1) and (1.2), respectively, then $T_1\phi_1$ and $T_2\phi_2$ are predictable as well under (1.1) and (1.2), respectively, and $\text{BLUP}_{\mathcal{M}_i}(T_1, \phi_1) = T_1\text{BLUP}_{\mathcal{M}_i}(\phi_1)$ hold for any matrices $T_i \in \mathbb{R}^{i \times i}, i = 1, 2$.

\textbf{Proof.} It is obvious from (1.14) that

$$\text{E}(\mathbf{L}_i \mathbf{Y}_i - \mathbf{F}) = 0 \Leftrightarrow \mathbf{L}_i \hat{\mathbf{X}}_i \mathbf{A} - \mathbf{F} \mathbf{A} = 0 \text{ for all } \mathbf{A}, \mathbf{L}_i \hat{\mathbf{X}}_i = \mathbf{F}, \text{ } i = 1, 2.$$  

From Lemma 2.1, the matrix equations are consistent respectively if and only if (4.1) hold. In these cases, we see from Lemma 2.1 that the first parts of (4.2) hold in the Löwner partial ordering. Further from Lemma 2.2, there always exist solutions $\mathbf{L}_i$ of $\mathbf{L}_i \hat{\mathbf{X}}_i = \mathbf{F}_i$ such that (4.8) hold, and the $\mathbf{L}_i$ are determined by the matrix equations

$$\mathbf{L}_i [\hat{\mathbf{X}}_i, \hat{\mathbf{V}}_i \hat{\mathbf{X}}_i^+] = [\mathbf{F}_i, \mathbf{C}_i \hat{\mathbf{X}}_i^+]$$

(4.9)

such that (4.8) hold, and the $\mathbf{L}_i$ are determined by the matrix equations

$$\mathbf{L}_i [\hat{\mathbf{X}}_i, \hat{\mathbf{V}}_i \hat{\mathbf{X}}_i^+] = [\mathbf{F}_i, \mathbf{C}_i \hat{\mathbf{X}}_i^+]$$

(4.10)

thus establishing the matrix equations in (4.2). Solving the matrix equations by Lemma 2.1 gives the $\mathbf{L}_i$ in (4.3).

Result (a) is well known. Results (b) and (c) follow directly from (4.3). Taking covariance matrices of (4.3) yields (4.4). From (3.1) and (4.3),

$$\text{Cov}\{\mathbf{B}, \mathbf{A}\} = \text{Cov}(\mathbf{A}, \mathbf{B})$$

(4.11)

for $i = 1, 2$.

Thus establishing (4.5). The two equalities in (4.6) follow from (1.17) and (4.4). Results (e) and (f) are direct consequences of (4.3).

One of most important features of BLUPs is concerned with the universal sum decompositions of BLUPs.

\textbf{Theorem 4.2.} Let $\mathcal{M}_1$ and $\mathcal{M}_2$ be as given in (1.1) and (1.2), respectively. Then, $\mathbf{X}_i \beta_i, \mathbf{X}_i \mathbf{A}, \mathbf{X}_i \mathbf{A}, \mathbf{X}_i \gamma_i$ and $\mathbf{E}_i$ are all predictable and estimable under (1.1) and (1.2), respectively, and the following decomposition equalities

$$\mathbf{Y}_i = \text{BLUP}_{\mathcal{M}_i}(\mathbf{X}_i \beta_i) + \text{BLUP}_{\mathcal{M}_i}(\mathbf{E}_i),$$

(4.12)

$$\mathbf{Y}_i = \text{BLUP}_{\mathcal{M}_i}(\mathbf{X}_i \mathbf{A}, \mathbf{A}) + \text{BLUP}_{\mathcal{M}_i}(\mathbf{X}_i \gamma_i) + \text{BLUP}_{\mathcal{M}_i}(\mathbf{E}_i)$$

(4.13)
always hold, \( i = 1, 2, \) or equivalently,

\[
\begin{align*}
y &= \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \text{BLUP}_\mathcal{M}(X_1 \beta_1) + \text{BLUP}_\mathcal{M}(\varepsilon_1) \\ \text{BLUP}_\mathcal{M}(X_2 \beta_2) + \text{BLUP}_\mathcal{M}(\varepsilon_2) \end{bmatrix}, \\
y &= \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \text{BLUP}_\mathcal{M}(X_1 A_1 \alpha_1) + \text{BLUP}_\mathcal{M}(X_1 \gamma_1) + \text{BLUP}_\mathcal{M}(\varepsilon_1) \\ \text{BLUP}_\mathcal{M}(X_2 A_2 \alpha_2) + \text{BLUP}_\mathcal{M}(X_2 \gamma_2) + \text{BLUP}_\mathcal{M}(\varepsilon_2) \end{bmatrix}.
\end{align*}
\]

(4.11) (4.12)

Proof. Setting \( \phi_i = y_i, \ i = 1, 2, \) in (4.7) yields (4.9)–(4.12).

In the remainder of this section, we derive the BLUPs of \( \phi_1, \phi_2, \phi \) in (1.14) and (1.15). We denote

\[
M_1 = [I, \ 0], \ M_2 = [0, I], \ N_1 = [I, \ 0], \ N_2 = [0, I],
\]

\[
T_1 = [I, \ 0], \ T_2 = [0, I], \ \tilde{X} = XA,
\]

\[
J = [G, H], \ \tilde{J}_i = [G_i N_i, H_i T_i], \ i = 1, 2,
\]

\[
C = \text{Cov}\{\phi, y\} = J\Sigma\tilde{X}, \ D_i = \text{Cov}\{\phi_i, y\} = \tilde{J}_i \Sigma \tilde{X}^T, \ i = 1, 2.
\]

It follows from (1.8), (1.14), and (1.15) that

\[
K_y - \phi = K \tilde{X} \alpha + K X \gamma + K \varepsilon - F \alpha - G \gamma - H \varepsilon
\]

\[
= (K \tilde{X} - F) \alpha + (KX - G) \gamma + (K - H) \varepsilon,
\]

(4.13)

\[
K_i y - \phi_i = K_i \tilde{X} \alpha + K_i X \gamma + K_i \varepsilon - F_i \alpha - G_i \gamma_i - H_i \varepsilon_i
\]

\[
= (K_i, \tilde{X} - F, M_i) \alpha + (K X - G N_i) \gamma + (K_i - H_i T_i) \varepsilon
\]

(4.14)

for \( i = 1, 2. \) Hence, the expectations and the covariance matrices of \( K y - \phi \) and \( K_i y - \phi_i \) can be written as

\[
E(Ky - \phi) = (K \tilde{X} - F) \alpha,
\]

(4.15)

\[
E(K_i y - \phi_i) = (K_i \tilde{X} - F_i M_i) \alpha, \ i = 1, 2,
\]

(4.16)

\[
\text{Cov}(Ky - \phi) = \text{Cov} [(KX - G) \gamma + (K - H) \varepsilon]
\]

\[
= [KX - G, K - H] \Sigma [KX - G, K - H]^T
\]

\[
= (K|X, I_n| - |G, H|) \Sigma (K|X, I_n| - |G, H|)^T
\]

\[
= (K \tilde{X} - J) \Sigma (K \tilde{X} - J)^T \overset{\triangle}{=} g(K)
\]

(4.17)

\[
\text{Cov}(K_i y - \phi_i) = \text{Cov} [(K_i X - G_i N_i) \gamma + (K_i - H_i T_i) \varepsilon]
\]

\[
= [K_i X - G_i N_i, K_i - H_i T_i] \Sigma [K_i X - G_i N_i, K_i - H_i T_i]^T
\]

\[
= (K_i|X, I_n| - |G_i N_i, H_i T_i|) \Sigma (K_i|X, I_n| - |G_i N_i, H_i T_i|)^T
\]

\[
= (K_i \tilde{X} - \tilde{J}_i) \Sigma (K_i \tilde{X} - \tilde{J}_i)^T \overset{\triangle}{=} g_i(K_i), \ i = 1, 2
\]

(4.18)

Applying Lemmas 2.1 and 2.2 to (4.15)–(4.18), we obtain the following results.

Theorem 4.3. Let \( \mathcal{M} \) be as given in (1.8). Then, the following results hold.

(a) The parameter vector \( \phi \) in (1.15) is predictable by \( y \) in (1.8) if and only if

\[
\mathcal{R}(\tilde{X}) \supseteq \mathcal{R}(F^T).
\]

(4.19)

In this case,

\[
\text{Cov}(Ky - \phi) = \min \text{ s.t. } E(Ky - \phi) = 0 \Leftrightarrow K|\tilde{X}. \ V \tilde{X}^T| = [F, C \tilde{X}^T].
\]

(4.20)
The parameter vectors \( \hat{\phi} \) are predictable by \( y \) under (1.8) if and only if
\[
\mathcal{R}(\hat{\mathbf{X}}') \supseteq \mathcal{R}(F_i'), \quad i = 1, 2.
\] (4.22)

In this case,
\[
\operatorname{Cov}(K_i y - \phi_i) = \operatorname{min} \text{ s.t. } E(K_i y - \phi_i) = 0 \\
\iff K_i [\hat{\mathbf{X}}, \mathbf{V}\hat{\mathbf{X}}^\perp] = [F_i M_i, D_i \hat{\mathbf{X}}^\perp], \quad i = 1, 2.
\] (4.23)

The matrix equations in (4.23) are consistent, respectively, under (4.22), and the general solutions \( K_i \) of the equations and the corresponding BLUPs of \( \phi_i \) under (1.8) are given by
\[
\text{BLUP}_M(\phi_i) = K_i y = \left( [F_i M_i, D_i \hat{\mathbf{X}}^\perp][\hat{\mathbf{X}}, \mathbf{V}\hat{\mathbf{X}}^\perp]^+ + U_i [\hat{\mathbf{X}}, \mathbf{V}\hat{\mathbf{X}}^\perp]^+ \right) y,
\] where \( U_i \in \mathbb{R}^{s \times n} \) are arbitrary, \( i = 1, 2 \).

(c) \( r[\hat{\mathbf{X}}, \mathbf{V}\hat{\mathbf{X}}^\perp] = r[\hat{\mathbf{X}}, \mathbf{X}^\perp] = r[\mathbf{X}, \mathbf{V}], \mathcal{R}[\hat{\mathbf{X}}, \mathbf{V}\hat{\mathbf{X}}^\perp] = \mathcal{R}[\hat{\mathbf{X}}, \hat{\mathbf{X}}^\perp] = \mathcal{R}[\hat{\mathbf{X}}, \mathbf{V}]. \)

(d) \( K \) is unique if and only if \( r[\mathbf{X}, \mathbf{V}] = n_1 + n_2 = n \).

(e) BLUP\(_M(\phi)\) is unique if and only if \( y \in \mathcal{R}[\hat{\mathbf{X}}, \mathbf{V}] \) holds with probability 1, namely, (1.8) is consistent.

(f) The following covariance matrix equalities
\[
\operatorname{Cov} [\text{BLUP}_M(\phi)] = \left( [F, \mathbf{C}\hat{\mathbf{X}}^\perp][\hat{\mathbf{X}}, \mathbf{V}\hat{\mathbf{X}}^\perp]^+ \right) \mathbf{V} \left( [F, \mathbf{C}\hat{\mathbf{X}}^\perp][\hat{\mathbf{X}}, \mathbf{V}\hat{\mathbf{X}}^\perp]^+ \right)'
\]
\[
\operatorname{Cov} [\text{BLUP}_M(\phi), \phi] = [F, \mathbf{C}\hat{\mathbf{X}}^\perp][\hat{\mathbf{X}}, \mathbf{V}\hat{\mathbf{X}}^\perp]^+ \mathbf{C}'
\]
\[
\operatorname{Cov}(\hat{\phi}) = \operatorname{Cov} [\text{BLUP}_M(\phi)]
\]
\[
= J\Sigma J' - \left( [F, \mathbf{C}\hat{\mathbf{X}}^\perp][\hat{\mathbf{X}}, \mathbf{V}\hat{\mathbf{X}}^\perp]^+ \right) \mathbf{V} \left( [F, \mathbf{C}\hat{\mathbf{X}}^\perp][\hat{\mathbf{X}}, \mathbf{V}\hat{\mathbf{X}}^\perp]^+ \right)'
\]

hold, \( i = 1, 2 \).

(g) The BLUPs of \( \phi \) and \( \phi_i \) under (1.8) can be decomposed as the sums
\[
\text{BLUP}_M(\phi) = \text{BLUE}_M(\mathbf{F}\alpha) + \text{BLUP}_M(\mathbf{G}\gamma) + \text{BLUP}_M(\mathbf{H}\epsilon),
\]
\[
\text{BLUP}_M(\phi_i) = \text{BLUE}_M(\mathbf{F}_i\alpha_i) + \text{BLUP}_M(\mathbf{G}_i\gamma_i) + \text{BLUP}_M(\mathbf{H}_i\epsilon_i), \quad i = 1, 2.
\]

(h) If \( \phi \) is predictable under (1.8), then \( T\phi \) is predictable as well under (1.8), and
\[
\text{BLUP}_M(T\phi) = T\text{BLUP}_M(\phi) \text{ holds for any matrix } T \in \mathbb{R}^{t \times s}.
\]
(i) $X\beta, X_i\beta_i, X\alpha, \hat{X}, X\gamma, \hat{X}\gamma_i, \varepsilon,$ and $\varepsilon_i$ are always predictable and estimable under (1.8), respectively, $i = 1, 2,$ and the following BLUP decompositions under (1.8) always hold

$$y = \text{BLUP}_{\theta}(X\beta) + \text{BLUP}_{\theta}(\varepsilon),$$
$$y = \text{BLUE}_{\theta}(X\alpha) + \text{BLUP}_{\theta}(X\gamma) + \text{BLUP}_{\theta}(\varepsilon),$$

or equivalently,

$$\begin{align*}
\mathbf{y} &= \left[ \text{BLUP}_{\theta}(X_1\beta_1) + \text{BLUP}_{\theta}(\varepsilon_1) \right] \\
\mathbf{y} &= \left[ \text{BLUP}_{\theta}(X_2\beta_2) + \text{BLUP}_{\theta}(\varepsilon_2) \right] \\
\mathbf{y}_1 &= \left[ \text{BLUE}_{\theta}(X_1\alpha_1) + \text{BLUP}_{\theta}(X_1\gamma_1) + \text{BLUP}_{\theta}(\varepsilon_1) \right] \\
\mathbf{y}_2 &= \left[ \text{BLUE}_{\theta}(X_2\alpha_2) + \text{BLUP}_{\theta}(X_2\gamma_2) + \text{BLUP}_{\theta}(\varepsilon_2) \right].
\end{align*}$$

**Proof.** It is obvious from (4.15), (4.16), and Lemma 2.1 that

$$E(K\mathbf{y} - \phi) = 0 \iff K\hat{X}\alpha = F\alpha \text{ for all } \alpha \iff K\hat{X} = F \iff \mathcal{R}(\hat{X}) \supseteq \mathcal{R}(F'),$$
$$E(K_i\mathbf{y} - \phi_i) = 0 \iff K_i\hat{X}\alpha = F_iM_i\alpha \text{ for all } \alpha \iff K_i\hat{X} = F_iM_i$$

$$\iff \mathcal{R}(\hat{X}) \supseteq \mathcal{R}(F_i) \iff \mathcal{R}(\hat{X}_i') \supseteq \mathcal{R}(F_i')$$

for $i = 1, 2,$ thus establishing (4.19) and (4.22), respectively. In these cases, applying Lemma 2.2 to the matrix equations in (4.20) and (4.23) leads to the conclusions in the theorem. The details of the proofs are omitted here due to space limitation. ☐

Because Theorems 4.1–4.3 are a group of theoretical inference results under the general covariance matrix assumptions in (1.6) and (1.7), they cannot directly be used to make statistical inference if the covariance matrix structure in (1.6) is totally unknown. Instead, we can obtain various concrete conclusions with respect to the specified covariance matrix structures in (1.6). Moreover, Theorems 4.1–4.3 provide a group of standard criteria for the comparison of the efficiency of other kinds of estimators, such as, the ordinary least-square estimators and weighted least-square estimators of $\hat{\beta}_i$ and $\hat{a}_i$ in (1.1) and (1.2), as well as, (1.4) and (1.5), which are obtained without using the assumptions in (1.6) and (1.7) and have statistical performances different from the BLUPs/BLUEs of the unknown parameters. We shall discuss these problems in forthcoming papers.

5. **Concluding remarks.** In summary, we remark that this study concentrates primarily to some current points of interest in simultaneous optimal predictions under two seemingly unrelated linear random-effects models, in which, we set up theoretical analysis to the optimization prediction problems through use of some powerful mathematical and statistical optimization methods. We carefully and sensitively extends some classic conception and knowledge on BLUPs to reflect the wealth of relevant novel results revealed in the past several years. Because all the formulas and facts in the preceding theorems are represented in certain analytical expressions or formulas, they can easily be reduced to various specified conclusions when the model matrices and covariance matrix in (1.1) and (1.2) are given in certain prescribed formulations. For example, both (1.1) and (1.2) encompass certain types of LREMs that have partially common parameter vectors as their special cases, such as,

$$y_i = X_i\beta + \varepsilon_i, \quad \beta = A\alpha + \gamma, \quad i = 1, 2,$$
$$y_i = X_i\beta_i + \varepsilon_i, \quad \beta_i = A_i\alpha + \gamma_i, \quad i = 1, 2.$$
under which more specified statistical inference results can be obtained; please refer to \[15, 36\] for the corresponding work. Besides, similar theoretical approaches can be conducted to some general types of seemingly unrelated regression models, for example, the following two seemingly unrelated linear mixed models

\[ y_i = X_i \alpha_i + Z_i \gamma_i + \varepsilon_i, \quad i = 1, 2, \]

where the two observed random vectors \( y_1 \) and \( y_2 \) are correlated statistically. We believe that more theoretical results and facts about BLUPs and BLUEs under different kinds of linear regression models can be established by a similar approach. It is no doubt that previous and recent studies show that the classic concepts like BLUPs and BLUEs have vital roles in the statistical inference of regression models, which have thrown up many difficult problems concerning optimal prediction and estimations under various parametric model assumptions, and thus have led to a broad and deep approaches in statistics and data analysis.

Acknowledgments. The authors are grateful to anonymous referees and the handling editor for their helpful comments and suggestions on an earlier version of this paper. This study is supported in part by the Shandong Provincial Natural Science Foundation \#ZR2019MA065.

REFERENCES

[1] N. K. Bansal and K. J. Miescke, Simultaneous selection and estimation in general linear models, \textit{J. Stat. Plann. Inference}, \textbf{104} (2002), 377–390.
[2] A. S. Bryk, S. W. Raudenbush and R. T. Congdon, \textit{Hierarchical Linear and Nonlinear Modeling with HLM/2L and HLM/3L Programs}, Scientific Software International, Chicago, IL, 1996.
[3] A. Chaturvedi, S. Kesarwani and R. Chandra, Simultaneous prediction based on shrinkage estimator, in: \textit{Recent Advances in Linear Models and Related Areas, Essays in Honour of Helge Toutenburg}, Springer, 2008, pp. 181–204.
[4] A. Chaturvedi, A. T. K. Wan and S. P. Singh, Improved multivariate prediction in a general linear model with an unknown error covariance matrix, \textit{J. Multivariate Anal.}, \textbf{83} (2002), 166–182.
[5] M. Dube and V. Manocha, Simultaneous prediction in restricted regression models, \textit{J. Appl. Statist. Sci.}, \textbf{11} (2002), 277–288.
[6] B. Effron and C. Morris, Combining possibly related estimation problems (with discussion), \textit{J. Roy. Stat. Soc. B}, \textbf{35} (1973), 379–421. \url{https://www.jstor.org/stable/2985106}
[7] S. Gan, C. Lu and Y. Tian, Computation and comparison of estimators under different linear random-effects models, \textit{Commun. Statist. Simul. Comput.}, \textbf{49} (2020), 1210–1222.
[8] A. Gelman and J. Hill, \textit{Data Analysis Using Regression and Multilevel/Hierarchical Models}, Cambridge University Press, 2007.
[9] H. Goldstein and J. D. Leeuw, \textit{Handbook of Multilevel Analysis}, Springer New York, 2008.
[10] C. A. Gotway and N. Cressie, Improved multivariate prediction under a general linear model, \textit{J. Multivariate Anal.}, \textbf{45} (1993), 56–72.
[11] N. Güler and M. E. Büyükkaya, Rank and inertia formulas for covariance matrices of BLUPs in general linear mixed models, \textit{Commun. Statist. Theor. Meth.}, 2020.
[12] S. J. Haslett and S. Puntanen, Equality of BLUEs or BLUPs under two linear models using stochastic restrictions, \textit{Stat. Papers}, \textbf{51} (2010), 465–475.
[13] S. J. Haslett and S. Puntanen, A note on the equality of the BLUPs for new observations under two linear models, \textit{Acta Comm. Univ. Tartu. Math.}, \textbf{14} (2010), 27–33.
[14] S. J. Haslett and S. Puntanen, On the equality of the BLUPs under two linear mixed models, \textit{Metrika}, \textbf{74} (2011), 381–395.
[15] J. Hou and B. Jiang, Predictions and estimations under a group of linear models with random coefficients, \textit{Comm. Statist. Simul. Comput.}, \textbf{47} (2018), 510–525.
[16] H. Jiang, J. Qian and Y. Sun, Best linear unbiased predictors and estimators under a pair of constrained seemingly unrelated regression models, \textit{Stat. Probab. Lett.}, \textbf{158} (2020), 108669.
[17] C. Lu, Y. Sun and Y. Tian, A comparison between two competing fixed parameter constrained general linear models with new regressors, Statistics, 52 (2018), 769–781.
[18] C. Lu, Y. Sun and Y. Tian, Two competing linear random-effects models and their connections, Stat. Papers, 59 (2018), 1101–1115.
[19] A. Markiewicz and S. Puntanen, All about the ⊥ with its applications in the linear statistical models, Open Math., 13 (2015), 33–50.
[20] S. K. Mitra, Generalized inverse of matrices and applications to linear models, in: Handbook of Statistics, P.K. Krishnaiah, ed., Vol. 1, North-Holland, pp. 471–512, 1980.
[21] R. Penrose, A generalized inverse for matrices, Proc. Cambridge Phil. Soc., 51 (1955), 406–413.
[22] S. Puntanen, G. P. H. Styan and J. Isotalo, Matrix Tricks for Linear Statistical Models, Our Personal Top Twenty, Springer, Berlin, 2011.
[23] C. R. Rao, Unified theory of linear estimation, Sankhyā, Ser. A, 33 (1971), 371–394.
[24] C. R. Rao, Representations of best linear unbiased estimators in the Gauss–Markoff model with a singular dispersion matrix, J. Multivariate Anal., 3 (1973), 276–292.
[25] C. R. Rao, Simultaneous estimation of parameters in different linear models and applications to biometric problems, Biometrics, 31 (1975), 545–554.
[26] C. R. Rao, 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, 1989, pp. 397–417.
[27] C. R. Rao and S. K. Mitra, Generalized Inverse of Matrices and Its Applications, Wiley, New York, 1971.
[28] C. R. Rao, H. Toutenburg, Shalabh and C. Heumann, Linear Models and Generalizations: Least Squares and Alternatives, 3rd edition, Springer, Berlin, 2008.
[29] S. W. Raudenbush and A. S. Bryk, Hierarchical Linear Models: Applications and Data Analysis Methods, 2nd edition, Sage, Thousand Oaks, 2002.
[30] Shalabh, Performance of Stein-rule procedure for simultaneous prediction of actual and average values of study variable in linear regression models, Bull. Internat. Stat. Instit., 56 (1995), 1375–1390.
[31] Y. Sun, B. Jiang and H. Jiang, Computations of predictors/estimators under a linear random-effects model with parameter restrictions, Comm. Statist. Theory Meth., 48 (2019), 3482–3497.
[32] Y. Sun, H. Jiang and Y. Tian, A prediction analysis in a constrained multivariate general linear model with future observations, Comm. Statist. Theory Meth., 2020.
[33] Y. Tian, A new derivation of BLUPs under random-effects model, Metrika, 78 (2015), 905–918.
[34] Y. Tian, A matrix handling of predictions under a general linear random-effects model with new observations, Electron. J. Linear Algebra, 29 (2015), 30–45.
[35] Y. Tian, Transformation approaches of linear random-effects models, Statist. Meth. Appl., 26 (2017), 583–608.
[36] Y. Tian and B. Jiang, An algebraic study of BLUPs under two linear random-effects models with correlated covariance matrices, Linear Multilinear Algebra, 64 (2016), 2351–2367.
[37] Y. Tian and J. Wang, Some remarks on fundamental formulas and facts in the statistical analysis of a constrained general linear model, Commun. Statist. Theory Meth., 49 (2020), 1201–1216.
[38] H. Toutenburg, Prior Information in Linear Models. Wiley, New York, 1982.
[39] H. Toutenburg and Shalabh, Predictive performance of the methods of restricted and mixed regression estimators, Biometr. J., 38 (1996), 951–959.
[40] H. Toutenburg and Shalabh, Improved prediction in linear regression model with stochastic linear constraints, Biometr. J., 42 (2000), 71–86.

Received May 2020; revised September 2020.

E-mail address: yongge.tian@gmail.com
E-mail address: xiepengyang@163.com