Available online at www.sciencedirect.com
ScienceDirect
Journal of Multivariate Analysis 99 (2008) 1503–1517
Journal of Multivariate Analysis
www.elsevier.com/locate/jmva

Best linear unbiased prediction for linear combinations in general mixed linear models

Xu-Qing Liu[a,*], Jian-Ying Rong[b], Xiu-Ying Liu[c]

[a] *Department of Computing Science, Huaiyin Institute of Technology, Huai'an 223001, PR China*
[b] *Department of Foundation Courses, Huai'an College of Information Technology, Huai'an 223003, PR China*
[c] *Department of Mathematics, Heze University, Heze 274015, PR China*

Received 16 February 2006
Available online 30 January 2008

Abstract

The general mixed linear model can be written as $\boldsymbol{y} = \boldsymbol{X}\beta + \boldsymbol{Z}\boldsymbol{u} + \boldsymbol{e}$. In this paper, we mainly deal with two problems. Firstly, the problem of predicting a general linear combination of fixed effects and realized values of random effects in a general mixed linear model is considered and an explicit representation of the best linear unbiased predictor (BLUP) is derived. In addition, we apply the resulting conclusion to several special models and offer an alternative to characterization of BLUP. Secondly, we recall the notion of linear sufficiency and consider it as regards the BLUP problem and characterize it in several different ways. Further, we study the concepts of linear sufficiency, linear minimal sufficiency and linear completeness, and give relations among them. Finally, four concluding remarks are given.

AMS 2000 subject classifications: 62J05; 62F10; 15A04; 62B05

Keywords: Mixed linear model; Transformed model; Linear combination; Best linear unbiased predictor (BLUP); Linear sufficiency; Linear minimal sufficiency; Linear completeness

1. Introduction

The three classical small-area models, of Battese, Harter and Fuller [4], Dempster, Rubin and Tsutakawa [7], and Fay and Herriot [11], are all special cases of the general mixed linear model,

* Corresponding author.
E-mail addresses: liuxuqing688@gmail.com (X.-Q. Liu), rjy98@163.com (J.-Y. Rong), lxiuying666@yahoo.com.cn (X.-Y. Liu).

0047-259X/$ - see front matter
doi:10.1016/j.jmva.2008.01.004

denoted by

$$y = X\beta + Zu + e, \tag{1.1}$$

where y is an n-dimensional vector of observations, X and Z are known $n \times p$ and $n \times q$ matrices, respectively, β is a vector of fixed effects, u is a vector of random effects, and e is a vector of random errors. Assume that $\mathscr{E}(u) = 0$, $\mathscr{E}(e) = 0$, $\mathscr{D}(u) = \sigma^2 G$, $\mathscr{D}(e) = \sigma^2 R$, and $\mathscr{E}(ue') = \sigma^2 K$, in which $\mathscr{E}(\bullet)$ and $\mathscr{D}(\bullet)$ refer to expectation and dispersion matrices, σ^2 (>0) is unknown. The assumptions are expressible jointly as

$$\mathscr{E}\begin{pmatrix} u \\ e \end{pmatrix} = 0, \qquad \mathscr{D}\begin{pmatrix} u \\ e \end{pmatrix} = \sigma^2 \begin{pmatrix} G & K \\ K' & R \end{pmatrix}.$$

Defining $\Sigma = ZGZ' + ZK + K'Z' + R$, it is not difficult to see that $\mathscr{D}(y) = \sigma^2 \Sigma$. As we know, for the three small-area models it is customary to consider the problem of predicting some particular functions which are special cases of the general linear combination of fixed effects and realized values of random effects, say

$$f(l, m) = l'\beta + m'u, \tag{1.2}$$

for given vectors, l and m, of constants. In the early literature, this problem has been of great interest to many authors. Among them, Henderson [16] pointed out that the practitioner is usually concerned with the above problem (that is, of predicting linear functions of β and u jointly) while the animal breeding research worker is usually concerned with the problem of estimating some estimable linear function of β. See also Harville [13], Harville and Jeske [14], Prasad and Rao [24], Robinson [25], and Das et al. [6].

We call $f(l, m)$ predictable if $l'\beta$ is linearly estimable, that is $l \in \mathscr{R}(X')$, where $\mathscr{R}(\bullet)$ denotes the range (column space) of the matrix. A linear predictor $\tilde{f}(l, m) = c'y + c_0$ is said to be (linearly) unbiased with respect to (w.r.t.) $f(l, m)$ in the sense that $\mathscr{E}(\tilde{f}(l, m) - f(l, m)) = 0$ holds for all β. Clearly, $c'y + c_0$ is unbiased w.r.t. $f(l, m)$ if and only if $c_0 = 0$ and $X'c = l$. Thus, we only need to consider the homogeneous linear unbiased prediction class

$$\mathscr{C} = \left\{c'y | X'c = l\right\} \tag{1.3}$$

in the context. For the case of $K = 0$, Henderson [16] offered the best linear unbiased predictor (BLUP) for predictable $f(l, m)$ as $\tilde{f}(l, m) = l'\tilde{\beta} + m'GZ'\Sigma^{-1}(y - X\tilde{\beta})$ as regards the mean squared error (MSE) criterion, provided that Σ is nonsingular, where $\tilde{\beta}$ is any solution to the generalized least squares (GLS) equations $X'\Sigma^{-1}X\tilde{\beta} = X'\Sigma^{-1}y$. For the case of $K = 0$ and $\Sigma \geqslant 0$, Harville [13] obtained the essentially unique BLUP as $f^{\#}(l, m) = l'\beta^{\#} + m'GZ'\Sigma^{-1}(y - X\beta^{\#})$, in which $\beta^{\#}$ is any solution to the equation $X'\Sigma^{\#}X\beta^{\#} = X'\Sigma^{\#}y$, $\Sigma^{\#}$ is any particular generalized inverse of Σ satisfying $\mathrm{rk}\left(X'\Sigma^{\#}X\right) = \mathrm{rk}(X)$ and $X'\Sigma^{\#}\Sigma N_X = 0$ with the notation $N_X = I - P_X$, where P_X refers to the orthogonal projection onto the range of X (the same below). A coming problem, which has not hitherto appeared in the literature of this area, is how to deal with a more general situation and this will be solved in this paper.

Another main aim of this article is to consider those linear statistics which preserve sufficient information for predicting $f(l, m)$ and to characterize them. Such linear statistics will be called linearly (combination) sufficient statistics, combining the notion of linear sufficiency introduced by Drygas [8] when $X\beta$ was estimated and the notion of linear error-sufficiency investigated by Groß [12] while the random error term was predicted in a general Gauss–Markov model. This is as argued by Isotalo and Puntanen [20] that since the uniformly minimum variance unbiased estimators (UMVUEs) are based on the concepts of sufficiency and completeness and the well known Rao–Blackwell and Lehmann–Scheffé Theorems, linear versions of above can be defined correspondingly and then used as an alternative to obtaining the BLUE or BLUP.

The remainder is as follows. We first deduce an explicit expression of the BLUP for $f(\boldsymbol{l}, \boldsymbol{m})$ and offer an alternative method to characterizing BLUP after applying the resulting BLUP to some special situations in Section 2. Notions of linear sufficiency and linear minimal sufficiency in conjunction with linear completeness are recalled w.r.t. the BLUP problem in Section 3. Relations among them are provided in different ways. Finally, we give four concluding remarks.

2. Best linear unbiased prediction

For a particular predictable function $f(\boldsymbol{l}, \boldsymbol{m})$, we will derive the representation of its BLUP by virtue of Rao's *Unified Theory of Least Squares* and apply the resulting conclusions to the nested-error regression model and a random regression coefficients model with equi-correlated errors and in addition $\boldsymbol{X}$ having one column as the unit vector.

2.1. BLUP: General case

Define $\boldsymbol{T} = \boldsymbol{\Sigma} + \boldsymbol{XUX}'$, where $\boldsymbol{U}$ refers to any fixed arbitrary symmetric matrix such that $\boldsymbol{T}$ is symmetric nonnegative definite (s.n.n.d.) and $\mathscr{R}(\boldsymbol{T}) = \mathscr{R}(\boldsymbol{X}, \boldsymbol{\Sigma})$, or equivalently, $\mathscr{R}(\boldsymbol{X}) \subseteq \mathscr{R}(\boldsymbol{T})$. Without loss of generality, we suppose that $\mathscr{R}(\boldsymbol{U}) \subseteq \mathscr{R}(\boldsymbol{X}')$. By virtue of the above notation we put $f^*(\boldsymbol{l}, \boldsymbol{m}) = \boldsymbol{l}'\beta^* + \boldsymbol{m}'\boldsymbol{u}^*$, with

$$\begin{cases} \beta^* = (\boldsymbol{X}'\boldsymbol{T}^-\boldsymbol{X})^-\boldsymbol{X}'\boldsymbol{T}^-\boldsymbol{y}, \\ \boldsymbol{u}^* = (\boldsymbol{ZG} + \boldsymbol{K}')'\boldsymbol{T}^-(\boldsymbol{y} - \boldsymbol{X}\beta^*). \end{cases}$$

Note that $\boldsymbol{l} \in \mathscr{R}(\boldsymbol{X}')$. This combined with consistency of the model (1.1), i.e., $\boldsymbol{y} \in \mathscr{R}(\boldsymbol{X}, \boldsymbol{\Sigma}) = \mathscr{R}(\boldsymbol{T})$ almost surely, yields that $f^*(\boldsymbol{l}, \boldsymbol{m})$ is invariant w.r.t. the choices of generalized inverses involved, and thereby we can replace $(\bullet)^-$ with the corresponding $(\bullet)^+$ whenever necessary and vice versa. The following theorem concerns the BLUP of $f(\boldsymbol{l}, \boldsymbol{m})$.

Theorem 2.1. *For the general mixed linear model of form (1.1), assume that $f(\boldsymbol{l}, \boldsymbol{m})$ is predictable. Then $f^*(\boldsymbol{l}, \boldsymbol{m})$ is the essentially unique BLUP for $f(\boldsymbol{l}, \boldsymbol{m})$ w.r.t. the MSE criterion.*

Proof. Assume that $\boldsymbol{c}'\boldsymbol{y}$ is any fixed arbitrary linear unbiased predictor for $f(\boldsymbol{l}, \boldsymbol{m})$, that is $\boldsymbol{c}'\boldsymbol{y} \in \mathscr{C}$. We apply the Lagrange multipliers method below. Let $\mathrm{L}(\boldsymbol{c}, \lambda) = \sigma^{-2}\mathrm{MSE}(\boldsymbol{c}'\boldsymbol{y}, f) - 2\lambda'(\boldsymbol{l} - \boldsymbol{X}'\boldsymbol{c})$ be the Lagrange function, where λ is a p-dimensional vector of Lagrange multipliers, and

$$\begin{aligned} \mathrm{MSE}(\boldsymbol{c}'\boldsymbol{y}, f) &= \mathscr{E}(\boldsymbol{c}'\boldsymbol{y} - f)^2 = \mathscr{D}\begin{pmatrix} \boldsymbol{Z}'\boldsymbol{c} - \boldsymbol{m} \\ \boldsymbol{c} \end{pmatrix}'\begin{pmatrix} \boldsymbol{u} \\ \boldsymbol{e} \end{pmatrix} \\ &= \sigma^2 \begin{pmatrix} \boldsymbol{Z}'\boldsymbol{c} - \boldsymbol{m} \\ \boldsymbol{c} \end{pmatrix}' \begin{pmatrix} \boldsymbol{G} & \boldsymbol{K} \\ \boldsymbol{K}' & \boldsymbol{R} \end{pmatrix} \begin{pmatrix} \boldsymbol{Z}'\boldsymbol{c} - \boldsymbol{m} \\ \boldsymbol{c} \end{pmatrix} \end{aligned}$$

since $\boldsymbol{c}'\boldsymbol{y}$ is unbiased. It follows that $\mathrm{L}(\boldsymbol{c}, \lambda) = \boldsymbol{c}'\boldsymbol{\Sigma}\boldsymbol{c} + \boldsymbol{m}'\boldsymbol{G}\boldsymbol{m} - 2\boldsymbol{m}'(\boldsymbol{ZG} + \boldsymbol{K}')'\boldsymbol{c} - 2\lambda'(\boldsymbol{l} - \boldsymbol{X}'\boldsymbol{c})$. By means of standard formulas for partial derivatives of matrix functions, let the gradients of $\mathrm{L}(\boldsymbol{c}, \lambda)$ w.r.t. $\boldsymbol{c}$ and λ vanish. It follows that

$$\begin{cases} \dfrac{\partial \mathrm{L}(\boldsymbol{c}, \lambda)}{\partial \boldsymbol{c}} = \boldsymbol{0} \Leftrightarrow \boldsymbol{\Sigma}\boldsymbol{c} + \boldsymbol{X}\lambda = (\boldsymbol{ZG} + \boldsymbol{K}')\boldsymbol{m}, \\ \dfrac{\partial \mathrm{L}(\boldsymbol{c}, \lambda)}{\partial \lambda} = \boldsymbol{0} \Leftrightarrow \boldsymbol{X}'\boldsymbol{c} = \boldsymbol{l}, \end{cases}$$

or equivalently,

$$\begin{pmatrix} \boldsymbol{\Sigma} & \boldsymbol{X} \\ \boldsymbol{X}' & \boldsymbol{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{c} \\ \lambda \end{pmatrix} = \begin{pmatrix} (\boldsymbol{ZG}+\boldsymbol{K}')\boldsymbol{m} \\ \boldsymbol{l} \end{pmatrix}. \tag{2.1}$$

We first need to show the consistency of Eq. (2.1), that is,

$$\begin{pmatrix} (\boldsymbol{ZG}+\boldsymbol{K}')\boldsymbol{m} \\ \boldsymbol{l} \end{pmatrix} \in \mathscr{R}\begin{pmatrix} \boldsymbol{\Sigma} & \boldsymbol{X} \\ \boldsymbol{X}' & \boldsymbol{0} \end{pmatrix}. \tag{2.2}$$

Actually, employing the well known fact that

$$\begin{pmatrix} \boldsymbol{\Sigma} & \boldsymbol{X} \\ \boldsymbol{X}' & \boldsymbol{0} \end{pmatrix}^{-} = \begin{pmatrix} \boldsymbol{T}^- - \boldsymbol{T}^-\boldsymbol{XS}^-\boldsymbol{X}'\boldsymbol{T}^- & \boldsymbol{T}^-\boldsymbol{XS}^- \\ \boldsymbol{S}^-\boldsymbol{X}'\boldsymbol{T}^- & \boldsymbol{US}^-\boldsymbol{S}-\boldsymbol{S}^- \end{pmatrix}, \tag{2.3}$$

in which $\boldsymbol{S}=\boldsymbol{X}'\boldsymbol{T}^-\boldsymbol{X}$, we obtain

$$\begin{pmatrix} \boldsymbol{\Sigma} & \boldsymbol{X} \\ \boldsymbol{X}' & \boldsymbol{0} \end{pmatrix}\begin{pmatrix} \boldsymbol{\Sigma} & \boldsymbol{X} \\ \boldsymbol{X}' & \boldsymbol{0} \end{pmatrix}^{-} = \begin{pmatrix} \boldsymbol{TT}^- & \boldsymbol{X}(\boldsymbol{U}-\boldsymbol{USS}^-) \\ \boldsymbol{0} & \boldsymbol{SS}^- \end{pmatrix}$$

by direct operations, since $\mathscr{R}(\boldsymbol{X}) \subseteq \mathscr{R}(\boldsymbol{T})$, $\mathscr{R}(\boldsymbol{\Sigma}) \subseteq \mathscr{R}(\boldsymbol{T})$, and $\mathscr{R}(\boldsymbol{U}) \subseteq \mathscr{R}(\boldsymbol{X}') = \mathscr{R}(\boldsymbol{S})$. In view of the two facts $\mathscr{R}(\boldsymbol{A}) = \mathscr{R}(\boldsymbol{AA}^-)$ and $\mathscr{R}[\boldsymbol{X}(\boldsymbol{U}-\boldsymbol{USS}^-)] \subseteq \mathscr{R}(\boldsymbol{X}) \subseteq \mathscr{R}(\boldsymbol{T}) = \mathscr{R}(\boldsymbol{TT}^-)$, we get

$$\begin{aligned}\mathscr{R}\begin{pmatrix} \boldsymbol{\Sigma} & \boldsymbol{X} \\ \boldsymbol{X}' & \boldsymbol{0} \end{pmatrix} &= \mathscr{R}\left[\begin{pmatrix} \boldsymbol{\Sigma} & \boldsymbol{X} \\ \boldsymbol{X}' & \boldsymbol{0} \end{pmatrix}\begin{pmatrix} \boldsymbol{\Sigma} & \boldsymbol{X} \\ \boldsymbol{X}' & \boldsymbol{0} \end{pmatrix}^{-}\right] = \mathscr{R}\begin{pmatrix} \boldsymbol{TT}^- & \boldsymbol{X}(\boldsymbol{U}-\boldsymbol{USS}^-) \\ \boldsymbol{0} & \boldsymbol{SS}^- \end{pmatrix} \\ &= \mathscr{R}\begin{pmatrix} \boldsymbol{TT}^- & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{SS}^- \end{pmatrix} = \mathscr{R}\begin{pmatrix} \boldsymbol{T} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{S} \end{pmatrix} = \mathscr{R}\begin{pmatrix} \boldsymbol{\Sigma} & \boldsymbol{X} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{X}' \end{pmatrix}.\end{aligned}$$

Since $f(\boldsymbol{l},\boldsymbol{m})$ is estimable, we assume $\boldsymbol{l}=\boldsymbol{X}'\boldsymbol{d}$ for some vector $\boldsymbol{d}$, and therefore

$$\begin{pmatrix} (\boldsymbol{ZG}+\boldsymbol{K}')\boldsymbol{m} \\ \boldsymbol{l} \end{pmatrix} = \begin{pmatrix} \boldsymbol{ZG}+\boldsymbol{K}' & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{X}' \end{pmatrix}\begin{pmatrix} \boldsymbol{m} \\ \boldsymbol{d} \end{pmatrix}.$$

Thus to show (2.2), it suffices to justify $\mathscr{R}(\boldsymbol{ZG}+\boldsymbol{K}') \subseteq \mathscr{R}(\boldsymbol{T})$. Actually, writing

$$\begin{pmatrix} \boldsymbol{G} & \boldsymbol{K} \\ \boldsymbol{K}' & \boldsymbol{R} \end{pmatrix} = \begin{pmatrix} \boldsymbol{A} \\ \boldsymbol{B} \end{pmatrix}\begin{pmatrix} \boldsymbol{A} \\ \boldsymbol{B} \end{pmatrix}' = \begin{pmatrix} \boldsymbol{AA}' & \boldsymbol{AB}' \\ \boldsymbol{BA}' & \boldsymbol{BB}' \end{pmatrix},$$

and recalling that $\boldsymbol{\Sigma} = \boldsymbol{ZGZ}' + \boldsymbol{R} + \boldsymbol{ZK} + \boldsymbol{K}'\boldsymbol{Z}'$, we get

$$\begin{aligned}\mathscr{R}(\boldsymbol{ZG}+\boldsymbol{K}') &= \mathscr{R}\left[\begin{pmatrix} \boldsymbol{Z}' \\ \boldsymbol{I} \end{pmatrix}'\begin{pmatrix} \boldsymbol{G} \\ \boldsymbol{K}' \end{pmatrix}\right] = \mathscr{R}\left[\begin{pmatrix} \boldsymbol{Z}' \\ \boldsymbol{I} \end{pmatrix}'\begin{pmatrix} \boldsymbol{A} \\ \boldsymbol{B} \end{pmatrix}\boldsymbol{A}'\right] \subseteq \mathscr{R}\left[\begin{pmatrix} \boldsymbol{Z}' \\ \boldsymbol{I} \end{pmatrix}'\begin{pmatrix} \boldsymbol{A} \\ \boldsymbol{B} \end{pmatrix}\right] \\ &= \mathscr{R}\left[\begin{pmatrix} \boldsymbol{Z}' \\ \boldsymbol{I} \end{pmatrix}'\begin{pmatrix} \boldsymbol{A} \\ \boldsymbol{B} \end{pmatrix}\begin{pmatrix} \boldsymbol{A} \\ \boldsymbol{B} \end{pmatrix}'\begin{pmatrix} \boldsymbol{Z}' \\ \boldsymbol{I} \end{pmatrix}\right] = \mathscr{R}(\boldsymbol{\Sigma}) \subseteq \mathscr{R}(\boldsymbol{T}),\end{aligned}$$

and therefore the consistency of (2.1) is proved. Put

$$\begin{aligned}\boldsymbol{c}^* &= \left(\boldsymbol{T}^- - \boldsymbol{T}^-\boldsymbol{XS}^-\boldsymbol{X}'\boldsymbol{T}^- \quad \boldsymbol{T}^-\boldsymbol{XS}^-\right)\begin{pmatrix} (\boldsymbol{ZG}+\boldsymbol{K}')\boldsymbol{m} \\ \boldsymbol{l} \end{pmatrix} \\ &= (\boldsymbol{T}^- - \boldsymbol{T}^-\boldsymbol{XS}^-\boldsymbol{X}'\boldsymbol{T}^-)(\boldsymbol{ZG}+\boldsymbol{K}')\boldsymbol{m} + \boldsymbol{T}^-\boldsymbol{XS}^-\boldsymbol{l}\end{aligned}$$

(the same below). Note that $\boldsymbol{y}'\boldsymbol{c}^*$ is invariant w.r.t. the choices of generalized inverses involved with probability 1. Let $\boldsymbol{c}$ now be a fixed arbitrary vector satisfying $\boldsymbol{X}'\boldsymbol{c}=\boldsymbol{l}$. Then the MSE of $\boldsymbol{c}'\boldsymbol{y}$

w.r.t. $f(\boldsymbol{l}, \boldsymbol{m})$ is given as

$$\begin{aligned}\mathrm{MSE}(\boldsymbol{c}'\boldsymbol{y}, f(\boldsymbol{l},\boldsymbol{m})) &= \mathscr{E}(\boldsymbol{c}'\boldsymbol{y} - f(\boldsymbol{l},\boldsymbol{m}))^2 = \mathscr{E}\left(\boldsymbol{c}^{*\prime}\boldsymbol{y} - f(\boldsymbol{l},\boldsymbol{m}) + \boldsymbol{c}'\boldsymbol{y} - \boldsymbol{c}^{*\prime}\boldsymbol{y}\right)^2 \\ &= \mathrm{MSE}(\boldsymbol{c}^{*\prime}\boldsymbol{y}, f(\boldsymbol{l},\boldsymbol{m})) + \mathscr{E}(\boldsymbol{c}'\boldsymbol{y} - \boldsymbol{c}^{*\prime}\boldsymbol{y})^2 + 2\varrho,\end{aligned}$$

with

$$\begin{aligned}\varrho &= \mathscr{E}\left[(\boldsymbol{c}^{*\prime}\boldsymbol{y} - f(\boldsymbol{l},\boldsymbol{m}))(\boldsymbol{c}'\boldsymbol{y} - \boldsymbol{c}^{*\prime}\boldsymbol{y})\right] = \mathscr{E}\left\{\begin{pmatrix}\boldsymbol{Z}'\boldsymbol{c}^* - \boldsymbol{m}\\ \boldsymbol{c}^*\end{pmatrix}'\begin{pmatrix}\boldsymbol{u}\\ \boldsymbol{e}\end{pmatrix}\begin{pmatrix}\boldsymbol{u}\\ \boldsymbol{e}\end{pmatrix}'\begin{pmatrix}\boldsymbol{Z}'\\ \boldsymbol{I}\end{pmatrix}(\boldsymbol{c}^* - \boldsymbol{c})\right\} \\ &= \left[\boldsymbol{c}^{*\prime}\begin{pmatrix}\boldsymbol{Z}'\\ \boldsymbol{I}\end{pmatrix}'\begin{pmatrix}\boldsymbol{G} & \boldsymbol{K}\\ \boldsymbol{K}' & \boldsymbol{R}\end{pmatrix}\begin{pmatrix}\boldsymbol{Z}'\\ \boldsymbol{I}\end{pmatrix} - \begin{pmatrix}\boldsymbol{m}\\ \boldsymbol{0}\end{pmatrix}'\begin{pmatrix}\boldsymbol{G} & \boldsymbol{K}\\ \boldsymbol{K}' & \boldsymbol{R}\end{pmatrix}\begin{pmatrix}\boldsymbol{Z}'\\ \boldsymbol{I}\end{pmatrix}\right](\boldsymbol{c}^* - \boldsymbol{c}) \\ &= \left[\boldsymbol{c}^{*\prime}\boldsymbol{\Sigma} - \boldsymbol{m}'(\boldsymbol{ZG} + \boldsymbol{K}')'\right](\boldsymbol{c}^* - \boldsymbol{c}) = \left[\boldsymbol{c}^{*\prime}\boldsymbol{T} - \boldsymbol{m}'(\boldsymbol{ZG} + \boldsymbol{K}')'\right](\boldsymbol{c}^* - \boldsymbol{c}) = \boldsymbol{0},\end{aligned}$$

considering $\boldsymbol{X}'(\boldsymbol{c}^* - \boldsymbol{c}) = \boldsymbol{0}$. Consequently, $\mathrm{MSE}(\boldsymbol{c}^{*\prime}\boldsymbol{y}, f(\boldsymbol{l},\boldsymbol{m})) \leqslant \mathrm{MSE}(\boldsymbol{c}'\boldsymbol{y}, f(\boldsymbol{l},\boldsymbol{m}))$, with equality holding if and only if $\mathscr{E}(\boldsymbol{c}'\boldsymbol{y} - \boldsymbol{c}^{*\prime}\boldsymbol{y})^2 = 0$, or equivalently, $\boldsymbol{c}'\boldsymbol{y} = \boldsymbol{c}^{*\prime}\boldsymbol{y}$ almost surely. It follows that $f^*(\boldsymbol{l},\boldsymbol{m}) = \boldsymbol{y}'\boldsymbol{c}^* = \boldsymbol{l}'\beta^* + \boldsymbol{m}'\boldsymbol{u}^*$ is the essentially unique BLUP for $f(\boldsymbol{l},\boldsymbol{m})$ w.r.t. the MSE criterion. The proof is completed. ■

Note that the fact that $\mathscr{R}\begin{pmatrix}\boldsymbol{\Sigma} & \boldsymbol{X}\\ \boldsymbol{X}' & \boldsymbol{0}\end{pmatrix} = \mathscr{R}\begin{pmatrix}\boldsymbol{\Sigma} & \boldsymbol{X} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{X}'\end{pmatrix}$ can be justified by post-multiplying $\begin{pmatrix}\boldsymbol{\Sigma} & \boldsymbol{X}\\ \boldsymbol{X}' & \boldsymbol{0}\end{pmatrix}$ with a series of block elementary matrices. According to the proof of Theorem 2.1, by direct operations, one can justify that:

Theorem 2.2. *For the general mixed linear model (1.1), assume that $f(\boldsymbol{l},\boldsymbol{m})$ is predictable. Then*

$$\begin{aligned}\mathrm{MSE}\left(f^*(\boldsymbol{l},\boldsymbol{m}), f(\boldsymbol{l},\boldsymbol{m})\right) &= \begin{pmatrix}\boldsymbol{Z}'\boldsymbol{c}^* - \boldsymbol{m}\\ \boldsymbol{c}^*\end{pmatrix}'\begin{pmatrix}\boldsymbol{G} & \boldsymbol{K}\\ \boldsymbol{K}' & \boldsymbol{R}\end{pmatrix}\begin{pmatrix}\boldsymbol{Z}'\boldsymbol{c}^* - \boldsymbol{m}\\ \boldsymbol{c}^*\end{pmatrix} \\ &= \boldsymbol{l}'(\boldsymbol{S}^- - \boldsymbol{U})\boldsymbol{l} + \boldsymbol{m}'\boldsymbol{G}\boldsymbol{m} - 2\boldsymbol{m}'(\boldsymbol{ZG} + \boldsymbol{K}')'\boldsymbol{T}^-\boldsymbol{X}\boldsymbol{S}^-\boldsymbol{l} \\ &\quad - \boldsymbol{m}'(\boldsymbol{ZG} + \boldsymbol{K}')'\boldsymbol{Q}(\boldsymbol{ZG} + \boldsymbol{K}')\boldsymbol{m} \\ &= \alpha'\boldsymbol{S}^-\alpha - \boldsymbol{l}'\boldsymbol{U}\boldsymbol{l} + \boldsymbol{m}'\boldsymbol{G}\boldsymbol{m} - \boldsymbol{m}'(\boldsymbol{ZG} + \boldsymbol{K}')'\boldsymbol{T}^-(\boldsymbol{ZG} + \boldsymbol{K}')\boldsymbol{m},\end{aligned} \tag{2.4}$$

where $\alpha = \boldsymbol{l} - \boldsymbol{X}'\boldsymbol{T}^-(\boldsymbol{ZG} + \boldsymbol{K}')\boldsymbol{m}$, and $\boldsymbol{Q} = \boldsymbol{T}^+ - \boldsymbol{T}^+\boldsymbol{X}\boldsymbol{S}^-\boldsymbol{X}'\boldsymbol{T}^+$. ■

As we can see, the case of $\boldsymbol{Z} = \boldsymbol{0}$ and $\boldsymbol{l} = \boldsymbol{0}$ reduces to a special prediction problem in a special prediction model (cf. [20, p. 1012, Eq. (4)]), denoted by

$$\begin{pmatrix}\boldsymbol{y}\\ \boldsymbol{y}_f\end{pmatrix} = \begin{pmatrix}\boldsymbol{X}\\ \boldsymbol{0}\end{pmatrix}\beta + \begin{pmatrix}\boldsymbol{e}\\ \boldsymbol{u}\end{pmatrix}$$

with the same assumption as for (3.1). The BLUP of $\boldsymbol{m}'\boldsymbol{y}_f = f(\boldsymbol{0},\boldsymbol{m})|_{\boldsymbol{Z}=\boldsymbol{0}}$ is expressible as $\boldsymbol{K}(\boldsymbol{R}^- - \boldsymbol{R}^-\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{R}^-\boldsymbol{X})^-\boldsymbol{X}'\boldsymbol{R}^-)\boldsymbol{y}$. As to the case of $\boldsymbol{l} \neq \boldsymbol{0}$, it is a trivial situation and can be dealt with in a similar fashion.

2.2. BLUP: Applications

In the following, we apply our conclusions to the first small-area model, nested-error regression model, and a special random regression coefficients model (written $\mathscr{L}_\rho$) which is denoted by (1.1) with equi-correlated errors and in addition $\boldsymbol{X}$ having one column as the unit vector (we assume $\boldsymbol{X} = (\boldsymbol{1}_n, \boldsymbol{X}_0)$, without loss of generality).

2.2.1. Application to the nested-error regression model

The nested-error regression model, proposed by Battese, Harter, and Fuller [4] when they wanted to estimate mean acreage under a crop for counties in Iowa using Landsat satellite data in conjunction with survey data, is given by

$$y_{ij} = \boldsymbol{x}_{ij}'\beta + u_i + e_{ij}, \quad i = 1, \ldots, q, j = 1, \ldots, n_i,$$

where y_{ij} is the character of interest for the jth sampled unit in the ith small sample area, $\boldsymbol{x}_{ij} = (x_{ij1}, \ldots, x_{ijp})'$ is a p-dimensional vector of corresponding auxiliary values, $x_{ij1} = 1$, $\beta = (\beta_1, \ldots, \beta_p)'$ is a p-dimensional vector of unknown parameters, n_i is the number of sampled units observed in the ith small area and $\sum_{i=1}^{q} n_i = n$. The random errors u_i are assumed to be independent of $\mathscr{N}(0, \sigma_u^2)$, independent of the e_{ij}, which are assumed to be independent of $\mathscr{N}(0, \sigma_e^2)$. However, the normality assumption is not necessary in deriving the following results. The mean for the ith area may be written as

$$\mu_i = \overline{\boldsymbol{X}}_i'\beta + u_i (\triangleq \boldsymbol{l}'\beta + \boldsymbol{m}'\boldsymbol{u}),$$

which can be interpreted as the conditional mean of y_{ij} for the ith area given u_i. Note that μ_i is a linear combination of the fixed effects β and the realized value of the random effects $\boldsymbol{u}$, in which $\boldsymbol{l} = \overline{\boldsymbol{X}}_i$ is the sample mean of $\boldsymbol{X}_i$ for the ith area and $\boldsymbol{m} = (0, \ldots, 0, 1, 0, \ldots, 0)'$ with 1 in the ith position and 0's elsewhere. It is seen that the nested-error regression model can be written as (1.1), with

$$\begin{aligned}
&\boldsymbol{y} = (\boldsymbol{y}_1', \ldots, \boldsymbol{y}_q')', \qquad \boldsymbol{y}_i = (y_{i1}, \ldots, y_{in_i})', \qquad \boldsymbol{X} = (\boldsymbol{X}_1', \ldots, \boldsymbol{X}_q')',\\
&\boldsymbol{X}_i = (\boldsymbol{x}_{i1}, \ldots, \boldsymbol{x}_{in_i})', \quad x_{ij1} = 1,\\
&\boldsymbol{Z} = \operatorname{diag}(\boldsymbol{1}_{n_1}, \ldots, \boldsymbol{1}_{n_q}), \qquad \boldsymbol{u} = (u_1, \ldots, u_q)' \sim (\boldsymbol{0}, \sigma_u^2 \boldsymbol{I}_q), \quad \text{i.e., } \boldsymbol{G} = \tau \boldsymbol{I}_q,\\
&\qquad \text{with } \tau = \sigma_u^2/\sigma^2, \ \sigma^2 \triangleq \sigma_e^2,\\
&\boldsymbol{e} = (\boldsymbol{e}_1', \ldots, \boldsymbol{e}_q')' \sim (\boldsymbol{0}, \sigma_e^2 \boldsymbol{I}_n), \qquad \boldsymbol{e}_i = (e_{i1}, \ldots, e_{in_i})', \quad \text{i.e., } \boldsymbol{R} = \boldsymbol{I}_n, \ \boldsymbol{K} = \boldsymbol{0}.
\end{aligned}$$

It follows that $\boldsymbol{\Sigma} = \boldsymbol{ZGZ}' + \boldsymbol{R} + \boldsymbol{ZK} + \boldsymbol{K}'\boldsymbol{Z}' = \operatorname{diag}(\boldsymbol{\Sigma}_1, \ldots, \boldsymbol{\Sigma}_q)$ with $\boldsymbol{\Sigma}_i = \boldsymbol{I}_{n_i} + \tau \boldsymbol{1}_{n_i}\boldsymbol{1}_{n_i}'$. Notice that, on the one hand, $\boldsymbol{\Sigma}_i$ is not necessarily s.n.n.d. (nor is $\boldsymbol{\Sigma}$) in view of the fact that if $\boldsymbol{a} \in \mathscr{R}(\boldsymbol{A})$ and $\boldsymbol{b} \in \mathscr{R}(\boldsymbol{A}')$, then

$$1 + \operatorname{rk}(\boldsymbol{A} + \boldsymbol{ab}') = \operatorname{rk}\begin{pmatrix} \boldsymbol{A} & \boldsymbol{a} \\ -\boldsymbol{b}' & 1 \end{pmatrix} = \operatorname{rk}(\boldsymbol{A}) + \operatorname{rk}(1 + \boldsymbol{b}'\boldsymbol{A}^{-}\boldsymbol{a}),$$

and further $\boldsymbol{\Sigma} > \boldsymbol{0}$ iff $1 + n_i\tau \neq 0$ for $i = 1, 2, \ldots, q$; on the other hand, $\boldsymbol{U} = \operatorname{diag}(\boldsymbol{U}_1, \ldots, \boldsymbol{U}_q)$ is a suitable choice of $\boldsymbol{U}$ in view of $\boldsymbol{1}_{n_i}'\boldsymbol{X}_i^{\perp} = \boldsymbol{0}$, with $\boldsymbol{U}_i = \boldsymbol{X}_i^{+}(-\tau \boldsymbol{1}_{n_i}\boldsymbol{1}_{n_i}')\boldsymbol{X}_i'^{+}$, and thereby $\boldsymbol{T} = \operatorname{diag}(\boldsymbol{I}_{n_1}, \ldots, \boldsymbol{I}_{n_q}) = \boldsymbol{I}_n$. Rewrite $\boldsymbol{l} = \overline{\boldsymbol{X}}_i = \boldsymbol{X}'\boldsymbol{k}$, where $\boldsymbol{k} = \frac{1}{n_i}(\boldsymbol{0}_{n_1}', \ldots, \boldsymbol{0}_{n_{i-1}}', \boldsymbol{1}_{n_i}', \boldsymbol{0}_{n_{i+1}}', \ldots, \boldsymbol{0}_{n_q}')'$. Then, by Theorem 2.1, the BLUP of $\mu_i = \overline{\boldsymbol{X}}_i'\beta + u_i = \boldsymbol{l}'\beta + \boldsymbol{m}'\boldsymbol{u}$ is expressible as

$$\begin{aligned}
\mu_i^* &= f^*(\boldsymbol{l}, \boldsymbol{m})|_{\boldsymbol{l}=\boldsymbol{X}'\boldsymbol{k}, \boldsymbol{m}=(0,\ldots,0,1,0,\ldots,0)', \boldsymbol{G}=\tau\boldsymbol{I}_q, \boldsymbol{K}=\boldsymbol{0}, \boldsymbol{T}=\boldsymbol{I}_n} = \boldsymbol{k}' P_X \boldsymbol{y} + n_i\tau \boldsymbol{k}'\boldsymbol{Z}'(\boldsymbol{I} - P_X)\boldsymbol{y}\\
&= \boldsymbol{k}'[P_X + n_i\tau(\boldsymbol{I} - P_X)]\boldsymbol{y} = \overline{\boldsymbol{X}}_i'\hat{\beta} + n_i\tau\left(\bar{\boldsymbol{y}}_i - \overline{\boldsymbol{X}}_i'\beta\right),
\end{aligned}$$

in view of $\boldsymbol{m}'\boldsymbol{Z}' = n_i\boldsymbol{k}'$, where the symbol $\hat{\beta} = (\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{y}$ refers to the least squares (LS) solution. The choice of $\boldsymbol{U}$ seems to be complicated but the resulting representation of the BLUP is more concise than the one derived by Prasad and Rao [24].

2.2.2. Application to model $\mathscr{L}_\rho{}^{(1)}$

Dempster, Rubin and Tsutakawa [7] proposed a model with random regression coefficients. Their model of a special case, with single concomitant variable x and regression through the origin, can be written as

$$y_{ij} = \beta_i x_{ij} + e_{ij} = \beta x_{ij} + u_i x_{ij} + e_{ij}, \quad i = 1, \ldots, q,\ j = 1, \ldots, n_i,$$

where $\beta_i = \beta + u_i$, u_i and e_{ij} are as in the nested-error regression model (see Section 2.2.1). The mean for the ith area is given by $\mu_i = \overline{X}_i\beta + \overline{X}_i u_i$, which is a linear combination of the fixed effect β and the realized value of the random effect u_i. Thus, we may consider a more general case with a random regression coefficients model, denoted by

$$\boldsymbol{y} = \boldsymbol{X}\gamma + \boldsymbol{e}, \tag{2.5}$$

where $\boldsymbol{X}$ has one column as the unit vector, γ is supposed to be a vector of random regression coefficients with $\mathscr{E}(\gamma) = \beta$ and $\mathscr{D}\left[(\gamma', \boldsymbol{e}')'\right] = \sigma^2[(1-\rho)\boldsymbol{I}_{n+p} + \rho\boldsymbol{1}_{n+p}\boldsymbol{1}'_{n+p}]$. Writing $\gamma = \beta + \boldsymbol{e}$, the model reduces to $\mathscr{L}_\rho$ with $\boldsymbol{Z} = \boldsymbol{X}$ (and thereby $q = p$). That is, $\boldsymbol{1}_n \in \mathscr{R}(\boldsymbol{X})$, $\boldsymbol{G} = (1-\rho)\boldsymbol{I}_p + \rho\boldsymbol{1}_p\boldsymbol{1}'_p$, $\boldsymbol{K} = \rho\boldsymbol{1}_p\boldsymbol{1}'_n$, and $\boldsymbol{R} = (1-\rho)\boldsymbol{I}_n + \rho\boldsymbol{1}_n\boldsymbol{1}'_n$. We shall write this model as $\mathscr{L}_\rho{}^{(1)}$, for convenience. Clearly,

$$\boldsymbol{\Sigma} = \boldsymbol{Z}\boldsymbol{G}\boldsymbol{Z}' + \boldsymbol{Z}\boldsymbol{K} + \boldsymbol{K}'\boldsymbol{Z}' + \boldsymbol{R} = (1-\rho)\boldsymbol{I}_n + \left[(1-\rho)\boldsymbol{X}\boldsymbol{X}' + \rho(\boldsymbol{X}\boldsymbol{1}_p + \boldsymbol{1}_n)(\boldsymbol{X}\boldsymbol{1}_p + \boldsymbol{1}_n)'\right]$$

is nonsingular. In this case one may choose $\boldsymbol{U} = \boldsymbol{0}$ as a simple substitution. However, we will consider the choice as $\boldsymbol{U} = -\left[(1-\rho)P_{\boldsymbol{X}'} + \rho\boldsymbol{X}^+(\boldsymbol{X}\boldsymbol{1}_p + \boldsymbol{1}_n)(\boldsymbol{X}\boldsymbol{1}_p + \boldsymbol{1}_n)'\boldsymbol{X}'^+\right]$ considering $\boldsymbol{1}_n \in \mathscr{R}(\boldsymbol{X})$. Thereby, $\boldsymbol{T} = (1-\rho)\boldsymbol{I}_n$. By Theorem 2.1, the BLUP for $f(\boldsymbol{l}, \boldsymbol{m})$ is expressible as

$$f^*(\boldsymbol{l}, \boldsymbol{m}) = \boldsymbol{l}'(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}'\boldsymbol{y} \triangleq \boldsymbol{l}'\hat{\beta},$$

where $\hat{\beta}$ is the LS solution. Clearly, the above representation is independent of the scalar ρ. This fact would imply that the covariance factor ρ can be known or unknown in practice, and in addition, can be replaced with an arbitrary fixed scalar $\varrho \in (-1, 1)$. This process may be viewed as a problem of misspecification of the dispersion matrix.

2.3. BLUP: An alternative method

In this subsection, we offer an alternative method to the characterization of BLUP in a different way by means of so-called *linear zero functions* (i.e., unbiased linear estimators of zero) which are used widely in the literature; cf. Bhimasankaram and Sengupta [5]. Note that $\boldsymbol{c}'\boldsymbol{y}$ is unbiased w.r.t. $f(\boldsymbol{l}, \boldsymbol{m})$, in the sense $\mathscr{E}(\boldsymbol{c}'\boldsymbol{y} - f(\boldsymbol{l}, \boldsymbol{m})) = 0$ for any β, if and only if $\boldsymbol{c}'\boldsymbol{X} = \boldsymbol{l}'$. In this case, we have

$$\boldsymbol{c}'\boldsymbol{y} - f(\boldsymbol{l}, \boldsymbol{m}) = \boldsymbol{c}'(X\beta + \boldsymbol{Z}\boldsymbol{u} + \boldsymbol{e}) - (\boldsymbol{l}'\beta + \boldsymbol{m}'\boldsymbol{u}) = \begin{pmatrix} \boldsymbol{Z}'\boldsymbol{c} - \boldsymbol{m} \\ \boldsymbol{c} \end{pmatrix}' \begin{pmatrix} \boldsymbol{u} \\ \boldsymbol{e} \end{pmatrix}.$$

On the other hand, any linear zero function is of form

$$\boldsymbol{d}'(\boldsymbol{I} - P_{\boldsymbol{X}})\boldsymbol{y} = \boldsymbol{d}'(\boldsymbol{I} - P_{\boldsymbol{X}})(X\beta + \boldsymbol{Z}\boldsymbol{u} + \boldsymbol{e}) = \boldsymbol{d}'(\boldsymbol{I} - P_{\boldsymbol{X}})(\boldsymbol{Z}, \boldsymbol{I})\begin{pmatrix} \boldsymbol{u} \\ \boldsymbol{e} \end{pmatrix}.$$

As this is similar to [20, Theorem 2.1], it is not difficult to conclude that $\boldsymbol{c}'\boldsymbol{y} \in \mathscr{C}$ is the BLUP for $f(\boldsymbol{l}, \boldsymbol{m})$ if and only if $\mathrm{Cov}(\boldsymbol{c}'\boldsymbol{y} - f(\boldsymbol{l}, \boldsymbol{m}), \boldsymbol{d}'(\boldsymbol{I} - P_{\boldsymbol{X}})\boldsymbol{y}) = 0$ for any $\boldsymbol{d}$ (since

$\min\{\mathrm{MSE}[\boldsymbol{c}'\boldsymbol{y}, f(\boldsymbol{l},\boldsymbol{m})]\} \leftrightarrow \min\{\mathrm{MSE}[\boldsymbol{c}'\boldsymbol{y} - f(\boldsymbol{l},\boldsymbol{m}), 0]\}$). Note that

$$\begin{aligned}\mathrm{Cov}(\boldsymbol{c}'\boldsymbol{y} - f(\boldsymbol{l},\boldsymbol{m}), \boldsymbol{d}'(\boldsymbol{I} - P_X)\boldsymbol{y}) &= \mathscr{E}\left\{\begin{pmatrix}\boldsymbol{Z}'\boldsymbol{c}-\boldsymbol{m}\\ \boldsymbol{c}\end{pmatrix}'\begin{pmatrix}\boldsymbol{u}\\ \boldsymbol{e}\end{pmatrix}\begin{pmatrix}\boldsymbol{u}\\ \boldsymbol{e}\end{pmatrix}'\begin{pmatrix}\boldsymbol{Z}'\\ \boldsymbol{I}\end{pmatrix}(\boldsymbol{I}-P_X)\boldsymbol{d}\right\}\\ &= \left[\boldsymbol{c}'(\boldsymbol{Z},\ \boldsymbol{I}) - \boldsymbol{m}'(\boldsymbol{I},\ \boldsymbol{0})\right]\begin{pmatrix}\boldsymbol{G} & \boldsymbol{K}\\ \boldsymbol{K}' & \boldsymbol{R}\end{pmatrix}\begin{pmatrix}\boldsymbol{Z}'\\ \boldsymbol{I}\end{pmatrix}(\boldsymbol{I}-P_X)\boldsymbol{d}.\end{aligned}$$

Then, by direct operations, we derive an alternative characterization of BLUP in the following:

Theorem 2.3. *For the general mixed linear model (1.1), assume that $f(\boldsymbol{l},\boldsymbol{m})$ is predictable. Then the statements below are mutually equivalent:*

- *$\boldsymbol{c}'\boldsymbol{y}$ is the essentially unique BLUP for $f(\boldsymbol{l},\boldsymbol{m})$;*
- *$\boldsymbol{c}'(\boldsymbol{X}, \boldsymbol{\Sigma}\boldsymbol{X}^{\perp}) = (\boldsymbol{l}', \boldsymbol{m}'(\boldsymbol{Z}\boldsymbol{G}+\boldsymbol{K}')'\boldsymbol{X}^{\perp})$;*
- *$\boldsymbol{X}'\boldsymbol{c} = \boldsymbol{l}$ and $\boldsymbol{\Sigma}\boldsymbol{c} - (\boldsymbol{Z}\boldsymbol{G}+\boldsymbol{K}')\boldsymbol{m} \in \mathscr{R}(\boldsymbol{X})$.* ■

By Theorem 2.3, it is not difficult to see that for the model $\mathscr{L}_\rho{}^{(1)}$, $\boldsymbol{c}'\boldsymbol{y}$ is optimal for $\boldsymbol{l}'\gamma = \boldsymbol{l}'(\beta+\boldsymbol{e}) = f(\boldsymbol{l},\boldsymbol{l})$ if and only if $\boldsymbol{X}'\boldsymbol{c} = \boldsymbol{l}$ and $\boldsymbol{c} \in \mathscr{R}(\boldsymbol{X})$.

Clearly, $\boldsymbol{c}^*$ is such a $\boldsymbol{c}$ satisfying the second or the third condition that appeared in Theorem 2.3. Other explicit representations not necessarily equal to $\boldsymbol{c}^*$ can be obtained and thus BLUP has (potentially) different forms. Theorem 2.1 tells us that the BLUPs of different forms coincide with each other with probability 1, however.

3. Linear sufficiency

Write $\epsilon = \boldsymbol{Z}\boldsymbol{u}+\boldsymbol{e}$ in the model (1.1) if the interest is in estimating a linear function of β. Then ϵ is a random vector with null means and covariance matrix $\sigma^2\boldsymbol{\Sigma}$, and further (1.1) reduces to a linear model of the form

$$\boldsymbol{y} = \boldsymbol{X}\beta + \epsilon. \tag{3.1}$$

In some situations, we can get not all outputs of $\boldsymbol{y}$ but a particular linear transformed function $\boldsymbol{F}\boldsymbol{y}$ for some matrix $\boldsymbol{F}$ of suitable order. In this case, we have the transformed model given as

$$\boldsymbol{F}\boldsymbol{y} = \boldsymbol{F}\boldsymbol{X}\beta + \boldsymbol{F}\epsilon. \tag{3.2}$$

Thus it is reasonable to consider so-called linearly sufficient estimation as defined in [8]. The notion of *linear sufficiency* introduced by Drygas [8] is that $\boldsymbol{F}\boldsymbol{y}$ is said to be linearly sufficient if there is a linear function of $\boldsymbol{F}\boldsymbol{y}$ which is the best linear unbiased estimate (BLUE) of $\boldsymbol{X}\beta$. This classical notion was considered early on by many authors when they investigated those linear statistics which preserve enough information for obtaining BLUE of $\boldsymbol{X}\beta$. Among them Drygas [8–10], Baksalary and Mathew [3], Müeller [23], Heiligers and Markiewicz [15], Markiewicz [22] are mentioned. For a more general concept of linear sufficiency one can see Ip et al. [19]. Another three closely related notions are *quadratic sufficiency* (cf. [10,21]) and *linear error-sufficiency* (cf. [1] and Groß [12]), and *linear prediction sufficiency* (cf. Isotalo and Puntanen [20]). Furthermore, Ibarrola and Pérez-Palomares [17,18] applied linear sufficiency and linear completeness to a continuous time linear model and corresponding characterizations were offered.

Baksalary and Kala [2] proved that $\boldsymbol{F}\boldsymbol{y}$ is linearly sufficient if and only if

$$\mathscr{R}(\boldsymbol{X}'\boldsymbol{F}') = \mathscr{R}(\boldsymbol{X}') \tag{3.3}$$

and

$$\mathscr{R}(\boldsymbol{X}) \subseteq \mathscr{R}(\boldsymbol{TF}').$$

Note that the condition $\mathscr{R}(\boldsymbol{X}') = \mathscr{R}(\boldsymbol{X}'\boldsymbol{F}')$ is necessary for the natural requirement of unbiasedness. In the paper, we will always suppose that this (i.e., $\mathscr{R}(\boldsymbol{X}') = \mathscr{R}(\boldsymbol{X}'\boldsymbol{F}')$) is a precondition, and that

$$\mathscr{R}(\boldsymbol{F}') \subseteq \mathscr{R}(\boldsymbol{T}) \tag{3.4}$$

as argued by Müeller [23] and Ip et al. [19] that one can neglect all elements outside $\mathscr{R}(\boldsymbol{T})$ since they form a null set of the linear model (3.1).

As we can see, the above notion was proposed w.r.t. the problem of BLUE for $\boldsymbol{X}\beta$ (or all linearly estimable functions of β) when the interest is in estimating $\boldsymbol{X}\beta$. So, we may consider the notion w.r.t. the problem of BLUP for all predictable combinations of form $f(\boldsymbol{l}, \boldsymbol{m})$, with $\boldsymbol{l} \in \mathscr{R}(\boldsymbol{X}') = \mathscr{R}(\boldsymbol{X}'\boldsymbol{F}')$; see the natural requirement (3.3). Actually, the obtainable $\boldsymbol{Fy}$ combined with (1.1) gives a transformed model, denoted by

$$\boldsymbol{Fy} = \boldsymbol{FX}\beta + \boldsymbol{FZu} + \boldsymbol{Fe}, \tag{3.5}$$

or written as $\tilde{\boldsymbol{y}} = \tilde{\boldsymbol{X}}\beta + \tilde{\boldsymbol{Z}}\boldsymbol{u} + \tilde{\boldsymbol{e}}$, where $\tilde{\boldsymbol{y}} = \boldsymbol{Fy}$, $\tilde{\boldsymbol{X}} = \boldsymbol{FX}$, $\tilde{\boldsymbol{Z}} = \boldsymbol{FZ}$, and $\tilde{\boldsymbol{e}} = \boldsymbol{Fe}$, with

$$\mathscr{E}\begin{pmatrix}\boldsymbol{u}\\ \tilde{\boldsymbol{e}}\end{pmatrix} = \boldsymbol{0}, \qquad \mathscr{D}\begin{pmatrix}\boldsymbol{u}\\ \tilde{\boldsymbol{e}}\end{pmatrix} = \begin{pmatrix}\boldsymbol{G} & \boldsymbol{KF}'\\ \boldsymbol{FK}' & \boldsymbol{FRF}'\end{pmatrix}.$$

Put

$$\begin{aligned}
&\tilde{\boldsymbol{\Sigma}} = \tilde{\boldsymbol{Z}}\boldsymbol{G}\tilde{\boldsymbol{Z}}' + \boldsymbol{FRF}' + \tilde{\boldsymbol{Z}}\boldsymbol{KF}' + \boldsymbol{FK}'\tilde{\boldsymbol{Z}}' = \boldsymbol{F\Sigma F}', \qquad \tilde{\boldsymbol{T}} = \tilde{\boldsymbol{\Sigma}} + \tilde{\boldsymbol{X}}\boldsymbol{U}\tilde{\boldsymbol{X}}' = \boldsymbol{FTF}',\\
&\tilde{\boldsymbol{S}} = \tilde{\boldsymbol{X}}'\tilde{\boldsymbol{T}}^-\tilde{\boldsymbol{X}} = \boldsymbol{X}'\boldsymbol{F}'(\boldsymbol{FTF}')^-\boldsymbol{FX}, \qquad \tilde{\boldsymbol{Q}} = \tilde{\boldsymbol{T}}^- - \tilde{\boldsymbol{T}}^-\tilde{\boldsymbol{X}}\tilde{\boldsymbol{S}}^-\tilde{\boldsymbol{X}}'\tilde{\boldsymbol{T}}^-,\\
&\beta_* = \left(\tilde{\boldsymbol{X}}'\tilde{\boldsymbol{T}}^-\tilde{\boldsymbol{X}}\right)^-\tilde{\boldsymbol{X}}'\tilde{\boldsymbol{T}}^-\tilde{\boldsymbol{y}} = \tilde{\boldsymbol{S}}^-\boldsymbol{X}'\boldsymbol{F}'\tilde{\boldsymbol{T}}^-\boldsymbol{Fy}, \qquad \boldsymbol{u}_* = \left(\tilde{\boldsymbol{Z}}\boldsymbol{G} + \boldsymbol{FK}'\right)'\tilde{\boldsymbol{T}}^-\left(\tilde{\boldsymbol{y}} - \tilde{\boldsymbol{X}}\beta_*\right),\\
&\boldsymbol{c}_* = \tilde{\boldsymbol{Q}}\boldsymbol{F}(\boldsymbol{ZG} + \boldsymbol{K}')\boldsymbol{m} + (\boldsymbol{FTF}')^-\boldsymbol{FX}\left(\boldsymbol{X}'\boldsymbol{F}'(\boldsymbol{FTF}')^-\boldsymbol{FX}\right)^-\boldsymbol{l}.
\end{aligned}$$

Following from Theorems 2.1 and 2.2, we obtain the essentially unique BLUP for $f(\boldsymbol{l}, \boldsymbol{m})$ w.r.t. the model (3.5) as $f_*(\boldsymbol{l}, \boldsymbol{m}) = \boldsymbol{l}'\beta_* + \boldsymbol{m}'\boldsymbol{u}_* = \boldsymbol{c}_*'\boldsymbol{Fy}$ with

$$\begin{aligned}
\mathrm{MSE}\,(f_*(\boldsymbol{l}, \boldsymbol{m}), f(\boldsymbol{l}, \boldsymbol{m})) = {}& \boldsymbol{l}'(\tilde{\boldsymbol{S}}^- - \boldsymbol{U})\boldsymbol{l} + \boldsymbol{m}'\boldsymbol{Gm} - 2\boldsymbol{m}'(\boldsymbol{ZG} + \boldsymbol{K}')'\\
&\times \boldsymbol{F}'\tilde{\boldsymbol{T}}^-\boldsymbol{FX}\tilde{\boldsymbol{S}}^-\boldsymbol{l} - \boldsymbol{m}'(\boldsymbol{ZG} + \boldsymbol{K}')'\boldsymbol{F}'\tilde{\boldsymbol{Q}}\boldsymbol{F}(\boldsymbol{ZG} + \boldsymbol{K}')\boldsymbol{m}.
\end{aligned} \tag{3.6}$$

In the following, the notion of linear sufficiency w.r.t. the BLUP problem will be defined below. We consider this notion under the natural requirement (3.3) at all times in the context. Its characterization will be offered in a concise theorem version. Then, we apply the resulting conclusion to the model $\mathscr{L}_\rho$ and derive some attractive remarks.

3.1. Linear sufficiency w.r.t. BLUP

Definition 3.1. $\boldsymbol{Fy}$ is said to be linearly sufficient w.r.t. the BLUP problem if there is a linear function of $\boldsymbol{Fy}$ which is the BLUP of $f(\boldsymbol{l}, \boldsymbol{m})$ in the original model (1.1) for any given predictable $f(\boldsymbol{l}, \boldsymbol{m})$, and we define $\boldsymbol{F} \in \mathscr{F}_{\mathrm{MLM}}$. ■

It is clear that this notion reduces to the ordinary one defined by Drygas [8] if random effects vanish. Therefore, we may view ordinary linear sufficiency as a special situation of Definition 3.1. By the essential uniqueness of BLUP, we see that $\boldsymbol{F} \in \mathscr{F}_{\text{MLM}}$ if and only if $f_*(\boldsymbol{l}, \boldsymbol{m}) = f^*(\boldsymbol{l}, \boldsymbol{m})$ holds almost surely for any $\boldsymbol{l} \in \mathscr{R}(\boldsymbol{X}')$ $(= \mathscr{R}(\boldsymbol{X}'\boldsymbol{F}')$; cf. (3.3), the natural requirement) and for all $\boldsymbol{m}$, since both $f_*(\boldsymbol{l}, \boldsymbol{m}) = \boldsymbol{c}_*'\boldsymbol{F}\boldsymbol{y}$ and $f^*(\boldsymbol{l}, \boldsymbol{m}) = \boldsymbol{c}^{*\prime}\boldsymbol{y}$ are unbiased w.r.t. $f(\boldsymbol{l}, \boldsymbol{m})$. Note that if $\text{MSE}\,(f_*(\boldsymbol{l}, \boldsymbol{m}), f(\boldsymbol{l}, \boldsymbol{m})) = \text{MSE}\,(f^*(\boldsymbol{l}, \boldsymbol{m}), f(\boldsymbol{l}, \boldsymbol{m}))$, $f_*(\boldsymbol{l}, \boldsymbol{m})$ solves the problem of minimizing $\text{MSE}\left(\boldsymbol{c}'\boldsymbol{y}, f(\boldsymbol{l}, \boldsymbol{m})\right)$, and therefore

$$\boldsymbol{F} \in \mathscr{F}_{\text{MLM}} \Leftrightarrow f_* \overset{\text{a.s.}}{=\!=} f^* \Leftrightarrow f_* - f \overset{\text{a.s.}}{=\!=} f^* - f \Leftrightarrow \text{MSE}\,(f_*, f) = \text{MSE}\left(f^*, f\right)$$
$$\Leftrightarrow \left\{ \begin{array}{c} \boldsymbol{X}(\tilde{\boldsymbol{S}}^- - \boldsymbol{U})\boldsymbol{X}' = \boldsymbol{X}(\boldsymbol{S}^- - \boldsymbol{U})\boldsymbol{X}' \\ (\boldsymbol{ZG} + \boldsymbol{K}')'\boldsymbol{F}'\tilde{\boldsymbol{Q}}\boldsymbol{F}(\boldsymbol{ZG} + \boldsymbol{K}') = (\boldsymbol{ZG} + \boldsymbol{K}')'\boldsymbol{Q}(\boldsymbol{ZG} + \boldsymbol{K}') \\ \boldsymbol{X}\tilde{\boldsymbol{S}}^-\boldsymbol{X}'\boldsymbol{F}'\tilde{\boldsymbol{T}}^-\boldsymbol{F}(\boldsymbol{ZG} + \boldsymbol{K}') = \boldsymbol{X}\boldsymbol{S}^-\boldsymbol{X}'\boldsymbol{T}^-(\boldsymbol{ZG} + \boldsymbol{K}') \end{array} \right\} \tag{3.7}$$

in view of Eqs. (2.4) and (3.6), and $\mathscr{R}(\boldsymbol{S}) = \mathscr{R}(\boldsymbol{X}') = \mathscr{R}(\boldsymbol{X}'\boldsymbol{F}') = \mathscr{R}(\tilde{\boldsymbol{S}})$. The three conditions together given by (3.7) may be viewed as necessary and sufficient conditions (NSC) for $\boldsymbol{F} \in \mathscr{F}_{\text{MLM}}$. However, we have a concise version which will be stated in the following theorem.

Theorem 3.1. *$\boldsymbol{F} \in \mathscr{F}_{\text{MLM}}$ if and only if $\mathscr{R}(\boldsymbol{X}'\boldsymbol{F}') = \mathscr{R}(\boldsymbol{X}')$ and $\mathscr{R}(\boldsymbol{X}, \boldsymbol{ZG} + \boldsymbol{K}') \subseteq \mathscr{R}(\boldsymbol{TF}')$.*

Proof. Since we have viewed $\mathscr{R}(\boldsymbol{X}'\boldsymbol{F}') = \mathscr{R}(\boldsymbol{X}')$ as the precondition, we have $\boldsymbol{F} \in \mathscr{F}_{\text{MLM}} \Leftrightarrow$ (3.7). It is easily seen that $\boldsymbol{X}\tilde{\boldsymbol{S}}^-\boldsymbol{X}'$ and $\boldsymbol{X}\boldsymbol{S}^-\boldsymbol{X}'$ are invariant w.r.t. the choice of the generalized inverses involved considering the natural requirement and thereby we can replace them by corresponding Moore–Penrose inverses, i.e. $\boldsymbol{X}\tilde{\boldsymbol{S}}^-\boldsymbol{X}' = \boldsymbol{X}\tilde{\boldsymbol{S}}^+\boldsymbol{X}'$ and $\boldsymbol{X}\boldsymbol{S}^-\boldsymbol{X}' = \boldsymbol{X}\boldsymbol{S}^+\boldsymbol{X}'$. Let us now prove

$$\boldsymbol{X}\tilde{\boldsymbol{S}}^+\boldsymbol{X}' = \boldsymbol{X}\boldsymbol{S}^+\boldsymbol{X}' \Leftrightarrow \tilde{\boldsymbol{S}} = \boldsymbol{S} \Leftrightarrow \mathscr{R}(\boldsymbol{X}) \subseteq \mathscr{R}(\boldsymbol{TF}'). \tag{3.8}$$

Actually, pre-multiplying $\boldsymbol{X}\tilde{\boldsymbol{S}}^+\boldsymbol{X}' = \boldsymbol{X}\boldsymbol{S}^+\boldsymbol{X}'$ by $\boldsymbol{S}^+\boldsymbol{X}'\boldsymbol{T}^-$ and post-multiplying by $\boldsymbol{T}^-\boldsymbol{X}\boldsymbol{S}^+$, it follows that

$$\boldsymbol{X}\tilde{\boldsymbol{S}}^+\boldsymbol{X}' = \boldsymbol{X}\boldsymbol{S}^+\boldsymbol{X}' \Rightarrow \tilde{\boldsymbol{S}}^+ = \boldsymbol{S}^+\boldsymbol{S}\tilde{\boldsymbol{S}}^+\boldsymbol{S}\boldsymbol{S}^+ = \boldsymbol{S}^+\boldsymbol{S}\boldsymbol{S}^+\boldsymbol{S}\boldsymbol{S}^+ = \boldsymbol{S}^+ \Rightarrow \tilde{\boldsymbol{S}} = \boldsymbol{S},$$

and vice versa. The first NSC for $\boldsymbol{X}\tilde{\boldsymbol{S}}^+\boldsymbol{X}' = \boldsymbol{X}\boldsymbol{S}^+\boldsymbol{X}'$ under (3.3) is verified. Further,

$$\begin{aligned} \tilde{\boldsymbol{S}} = \boldsymbol{S} &\Leftrightarrow \boldsymbol{S} - \tilde{\boldsymbol{S}} = \boldsymbol{X}'\boldsymbol{T}^{+1/2}\left(\boldsymbol{I}_n - P_{\boldsymbol{T}^{1/2}\boldsymbol{F}'}\right)\boldsymbol{T}^{+1/2}\boldsymbol{X} = \boldsymbol{0} \\ &\Leftrightarrow \left(\boldsymbol{I}_n - P_{\boldsymbol{T}^{1/2}\boldsymbol{F}'}\right)\boldsymbol{T}^{+1/2}\boldsymbol{X} = \boldsymbol{0} \Leftrightarrow P_{\boldsymbol{T}^{1/2}\boldsymbol{F}'}\boldsymbol{T}^{+1/2}\boldsymbol{X} = \boldsymbol{T}^{+1/2}\boldsymbol{X} \\ &\Leftrightarrow \mathscr{R}(\boldsymbol{T}^{+1/2}\boldsymbol{X}) \subseteq \mathscr{R}(\boldsymbol{T}^{1/2}\boldsymbol{F}') \Leftrightarrow \mathscr{R}(\boldsymbol{X}) \subseteq \mathscr{R}(\boldsymbol{TF}'). \end{aligned}$$

(3.8) is thus proved. Under (3.3) and $\tilde{\boldsymbol{S}} = \boldsymbol{S}$ $(\Rightarrow \boldsymbol{F}'\tilde{\boldsymbol{T}}^-\boldsymbol{F}\boldsymbol{X} = \boldsymbol{T}^+\boldsymbol{X})$, the following would hold true:

$$\begin{aligned} &\boldsymbol{X}\tilde{\boldsymbol{S}}^-\boldsymbol{X}'\boldsymbol{F}'\tilde{\boldsymbol{T}}^-\boldsymbol{F}(\boldsymbol{ZG} + \boldsymbol{K}') = \boldsymbol{X}\boldsymbol{S}^-\boldsymbol{X}'\boldsymbol{T}^-(\boldsymbol{ZG} + \boldsymbol{K}') \\ &\quad \Leftrightarrow \boldsymbol{X}'\boldsymbol{F}'\tilde{\boldsymbol{T}}^-\boldsymbol{F}(\boldsymbol{ZG} + \boldsymbol{K}') = \boldsymbol{X}'\boldsymbol{T}^-(\boldsymbol{ZG} + \boldsymbol{K}') \\ &\quad \Rightarrow (\boldsymbol{ZG} + \boldsymbol{K}')'\boldsymbol{F}'\tilde{\boldsymbol{Q}}\boldsymbol{F}(\boldsymbol{ZG} + \boldsymbol{K}') = (\boldsymbol{ZG} + \boldsymbol{K}')'\boldsymbol{F}'\tilde{\boldsymbol{T}}^+\boldsymbol{F}(\boldsymbol{ZG} + \boldsymbol{K}') \\ &\qquad - (\boldsymbol{ZG} + \boldsymbol{K}')'\boldsymbol{F}'\tilde{\boldsymbol{T}}^+\boldsymbol{F}\boldsymbol{X}\tilde{\boldsymbol{S}}^-\boldsymbol{X}'\boldsymbol{F}'\tilde{\boldsymbol{T}}^+\boldsymbol{F}(\boldsymbol{ZG} + \boldsymbol{K}') \\ &\quad = (\boldsymbol{ZG} + \boldsymbol{K}')'\boldsymbol{F}'\tilde{\boldsymbol{T}}^-\boldsymbol{F}(\boldsymbol{ZG} + \boldsymbol{K}') - (\boldsymbol{ZG} + \boldsymbol{K}')'\boldsymbol{T}^-\boldsymbol{X}\boldsymbol{S}^-\boldsymbol{X}'\boldsymbol{T}^-(\boldsymbol{ZG} + \boldsymbol{K}'), \end{aligned}$$

in view of $\mathscr{R}(\boldsymbol{ZG}+\boldsymbol{K}') \subseteq \mathscr{R}(\boldsymbol{\Sigma}) \subseteq \mathscr{R}(\boldsymbol{T})$. Consequently,

$$\text{(3.7)} \Leftrightarrow \begin{cases} \mathscr{R}(\boldsymbol{X}) \subseteq \mathscr{R}(\boldsymbol{TF}') \\ (\boldsymbol{ZG}+\boldsymbol{K}')'\boldsymbol{F}'\tilde{\boldsymbol{T}}^{-}\boldsymbol{F}(\boldsymbol{ZG}+\boldsymbol{K}') = (\boldsymbol{ZG}+\boldsymbol{K}')'\boldsymbol{T}^{-}(\boldsymbol{ZG}+\boldsymbol{K}') \\ \boldsymbol{X}'\boldsymbol{F}'\tilde{\boldsymbol{T}}^{-}\boldsymbol{F}(\boldsymbol{ZG}+\boldsymbol{K}') = \boldsymbol{X}'\boldsymbol{T}^{-}(\boldsymbol{ZG}+\boldsymbol{K}') \end{cases}$$

$$\Leftrightarrow \begin{cases} \mathscr{R}(\boldsymbol{T}^{+1/2}\boldsymbol{X}) \subseteq \mathscr{R}(\boldsymbol{T}^{1/2}\boldsymbol{F}') \\ \left(\boldsymbol{T}^{+1/2}(\boldsymbol{ZG}+\boldsymbol{K}')\right)'\left(\boldsymbol{I}_n - P_{\boldsymbol{T}^{1/2}\boldsymbol{F}'}\right)\left(\boldsymbol{T}^{+1/2}(\boldsymbol{ZG}+\boldsymbol{K}')\right) = \boldsymbol{0} \\ \boldsymbol{X}'\boldsymbol{T}^{+1/2}\left(\boldsymbol{I}_n - P_{\boldsymbol{T}^{1/2}\boldsymbol{F}'}\right)\boldsymbol{T}^{+1/2}(\boldsymbol{ZG}+\boldsymbol{K}') = \boldsymbol{0}. \end{cases}$$

Observing that the third is a direct consequence of the first, and $\boldsymbol{I}_n - P_{\boldsymbol{T}^{1/2}\boldsymbol{F}'}$ is s.n.n.d., it follows that

$$\text{(3.7)} \Leftrightarrow \begin{cases} \mathscr{R}(\boldsymbol{X}) \subseteq \mathscr{R}(\boldsymbol{TF}') \\ \left(\boldsymbol{I}_n - P_{\boldsymbol{T}^{1/2}\boldsymbol{F}'}\right)\left(\boldsymbol{T}^{+1/2}(\boldsymbol{ZG}+\boldsymbol{K}')\right) = \boldsymbol{0} \\ \quad \Leftrightarrow \mathscr{R}\left(\boldsymbol{T}^{+1/2}(\boldsymbol{ZG}+\boldsymbol{K}')\right) \subseteq \mathscr{R}\left(\boldsymbol{T}^{1/2}\boldsymbol{F}'\right) \end{cases}$$

$$\Leftrightarrow \left\{\begin{array}{c} \mathscr{R}(\boldsymbol{X}) \subseteq \mathscr{R}(\boldsymbol{TF}') \\ \mathscr{R}\left(\boldsymbol{ZG}+\boldsymbol{K}'\right) \subseteq \mathscr{R}\left(\boldsymbol{TF}'\right) \end{array}\right\} \Leftrightarrow \mathscr{R}(\boldsymbol{X}, \boldsymbol{ZG}+\boldsymbol{K}') \subseteq \mathscr{R}(\boldsymbol{TF}').$$

This fact combined with $\boldsymbol{F} \in \mathscr{F}_{\text{MLM}} \Leftrightarrow$ (3.7) would mean that $\boldsymbol{F} \in \mathscr{F}_{\text{MLM}} \Leftrightarrow \mathscr{R}(\boldsymbol{X}, \boldsymbol{ZG}+\boldsymbol{K}') \subseteq \mathscr{R}(\boldsymbol{TF}')$ under the precondition (3.3). The proof is thus completed. ■

It should be noted that, under the condition $\mathscr{R}(\boldsymbol{X}, \boldsymbol{ZG}+\boldsymbol{K}') \subseteq \mathscr{R}(\boldsymbol{TF}')$, $\mathscr{R}(\boldsymbol{X}'\boldsymbol{F}') = \mathscr{R}(\boldsymbol{X}')$ will be satisfied inherently in view of the following implications:

$$\begin{aligned} &\mathscr{R}(\boldsymbol{X}) \subseteq \mathscr{R}(\boldsymbol{X}, \boldsymbol{ZG}+\boldsymbol{K}') \subseteq \mathscr{R}(\boldsymbol{TF}') \\ &\quad \Rightarrow \mathscr{R}(\boldsymbol{X}') = \mathscr{R}(\boldsymbol{X}'\boldsymbol{T}^{+}\boldsymbol{X}) \subseteq \mathscr{R}(\boldsymbol{X}'\boldsymbol{T}^{+}\boldsymbol{TF}') = \mathscr{R}(\boldsymbol{X}'\boldsymbol{F}') \subseteq \mathscr{R}(\boldsymbol{X}'). \end{aligned} \tag{3.9}$$

Based on this, the NSC for $\boldsymbol{F} \in \mathscr{F}_{\text{MLM}}$ obtained in Theorem 3.1 is reduced to $\mathscr{R}(\boldsymbol{X}, \boldsymbol{ZG}+\boldsymbol{K}') \subseteq \mathscr{R}(\boldsymbol{TF}')$. Let us now give an alternative characterization of $\mathscr{F}_{\text{MLM}}$ below, by using the method of Groß [12]. We will find that the NSC is independent of $\boldsymbol{U}$.

Theorem 3.2. $\boldsymbol{F} \in \mathscr{F}_{\text{MLM}}$ *iff* $\mathscr{N}(\boldsymbol{F}) \cap \left\{\mathscr{R}(\boldsymbol{X}, \boldsymbol{ZG}+\boldsymbol{K}') \oplus \mathscr{R}\left[\boldsymbol{\Sigma}(\boldsymbol{X}, \boldsymbol{ZG}+\boldsymbol{K}')^{\perp}\right]\right\} \subseteq \mathscr{R}\left[\boldsymbol{\Sigma}\,(\boldsymbol{X}, \boldsymbol{ZG}+\boldsymbol{K}')^{\perp}\right]$.

Proof. Observe that $\mathscr{R}(\boldsymbol{T})$ has direct sum decomposition $\mathscr{R}(\boldsymbol{T}) = \mathscr{R}(\boldsymbol{X}, \boldsymbol{ZG}+\boldsymbol{K}') \oplus \mathscr{R}\left[\boldsymbol{\Sigma}(\boldsymbol{X}, \boldsymbol{ZG}+\boldsymbol{K}')^{\perp}\right]$ and that $\boldsymbol{F} \in \mathscr{F}_{\text{MLM}} \Leftrightarrow \mathscr{R}(\boldsymbol{X}, \boldsymbol{ZG}+\boldsymbol{K}') \subseteq \mathscr{R}(\boldsymbol{TF}')$. It follows that

$$\begin{aligned} \boldsymbol{F} \in \mathscr{F}_{\text{MLM}} &\Leftrightarrow \mathscr{R}(\boldsymbol{TF}')^{\perp} \subseteq \mathscr{R}(\boldsymbol{X}, \boldsymbol{ZG}+\boldsymbol{K}')^{\perp} \\ &\Rightarrow \mathscr{R}\left(\boldsymbol{T}(\boldsymbol{TF}')^{\perp}\right) \subseteq \mathscr{R}\left(\boldsymbol{T}(\boldsymbol{X}, \boldsymbol{ZG}+\boldsymbol{K}')^{\perp}\right) \\ &\Leftrightarrow \mathscr{R}(\boldsymbol{F}')^{\perp} \cap \mathscr{R}(\boldsymbol{T}) \subseteq \mathscr{R}\left[\boldsymbol{\Sigma}\,(\boldsymbol{X}, \boldsymbol{ZG}+\boldsymbol{K}')^{\perp}\right] \end{aligned}$$

considering $\mathscr{R}(\boldsymbol{F}')^{\perp} = \mathscr{N}(\boldsymbol{F})$ and employing the well known fact that $\mathscr{R}\left(\boldsymbol{A}'(\boldsymbol{AB})^{\perp}\right) = \mathscr{R}(\boldsymbol{A}') \cap \mathscr{R}(\boldsymbol{B})^{\perp}$, and thus the necessity is proved. Conversely, provided $\mathscr{N}(\boldsymbol{F}) \cap \mathscr{R}(\boldsymbol{T}) \subseteq \mathscr{R}\left[\boldsymbol{\Sigma}\,(\boldsymbol{X}, \boldsymbol{ZG}+\boldsymbol{K}')^{\perp}\right]$, then

$$\mathscr{R}\left[(\boldsymbol{X}, \boldsymbol{ZG}+\boldsymbol{K}')'\boldsymbol{T}^{+}\boldsymbol{T}(\boldsymbol{TF}')^{\perp}\right] \subseteq \mathscr{R}\left[(\boldsymbol{X}, \boldsymbol{ZG}+\boldsymbol{K}')'\boldsymbol{T}^{+}\boldsymbol{T}(\boldsymbol{X}, \boldsymbol{ZG}+\boldsymbol{K}')^{\perp}\right] = \{\boldsymbol{0}\},$$

which means $(\boldsymbol{X}, \boldsymbol{ZG}+\boldsymbol{K}')'(\boldsymbol{TF}')^{\perp} = \boldsymbol{0}$ and therefore $\mathscr{R}(\boldsymbol{X}, \boldsymbol{ZG}+\boldsymbol{K}') \subseteq \mathscr{R}(\boldsymbol{TF}')$. The sufficiency is also proved. ■

Another alternative is by the Theorem 2.3. Actually, $\boldsymbol{F} \in \mathscr{F}_{\text{MLM}}$ iff for any predictable $f(\boldsymbol{l}, \boldsymbol{m})$ there is some $\boldsymbol{b}$ such that $\boldsymbol{b}'\boldsymbol{F}\boldsymbol{y}$ is the BLUP, i.e., $\boldsymbol{b}'\boldsymbol{F}(\boldsymbol{X}, \boldsymbol{\Sigma}\boldsymbol{X}^{\perp}) = (\boldsymbol{l}', \boldsymbol{m}'(\boldsymbol{ZG} + \boldsymbol{K}')'\boldsymbol{X}^{\perp}) = (\boldsymbol{d}', \boldsymbol{m}')\,\text{diag}(\boldsymbol{X}, (\boldsymbol{ZG} + \boldsymbol{K}')'\boldsymbol{X}^{\perp})$ with $\boldsymbol{l} = \boldsymbol{X}'\boldsymbol{d}$ for some $\boldsymbol{d}$. It follows that

$$\boldsymbol{F} \in \mathscr{F}_{\text{MLM}} \Leftrightarrow \mathscr{R}\begin{pmatrix} \boldsymbol{X}' & \boldsymbol{0} \\ \boldsymbol{0} & (\boldsymbol{X}^{\perp})'(\boldsymbol{ZG} + \boldsymbol{K}') \end{pmatrix} \subseteq \mathscr{R}\begin{pmatrix} \boldsymbol{X}'\boldsymbol{F}' \\ (\boldsymbol{X}^{\perp})'\boldsymbol{\Sigma}\boldsymbol{F}' \end{pmatrix}. \tag{3.10}$$

Theorem 3.3. $\boldsymbol{F} \in \mathscr{F}_{\text{MLM}}$ *if and only if* (3.10) *is satisfied.* ■

3.2. Linear sufficiency: Applications

Let us now consider the random regression coefficients model $\mathscr{L}_{\rho}^{(1)}$, which is a special situation of $\mathscr{L}_{\rho}$; we have $\mathscr{R}(\boldsymbol{TF}') = \mathscr{R}[(1-\rho)\boldsymbol{F}] = \mathscr{R}(\boldsymbol{F})$, and $\mathscr{R}(\boldsymbol{X}, \boldsymbol{ZG} + \boldsymbol{K}') = \mathscr{R}[\boldsymbol{X}, (1-\rho)\boldsymbol{X} + \rho\boldsymbol{X}\boldsymbol{1}_p\boldsymbol{1}_p' + \rho\boldsymbol{1}_n\boldsymbol{1}_p']$, and therefore we obtain the following result stated in a theorem version.

Theorem 3.4. *For the random regression coefficients model* $\mathscr{L}_{\rho}{}^{(1)}$, $\boldsymbol{F} \in \mathscr{F}_{\text{MLM}}$ *if and only if* $\mathscr{R}(\boldsymbol{X}) \subseteq \mathscr{R}(\boldsymbol{F}')$. ■

Consider another special case of $\mathscr{L}_{\rho}$ with the same assumptions as $\mathscr{L}_{\rho}{}^{(1)}$ but $\boldsymbol{Z} = \boldsymbol{I}_n$ (and thereafter $p = n$). The form of the model is a generalization of that of Fay and Herriot [11] in some sense. We write this model as $\mathscr{L}_{\rho}{}^{(2)}$. In a similar fashion, we have $\mathscr{R}(\boldsymbol{X}, \boldsymbol{ZG} + \boldsymbol{K}') = \mathbb{R}^n$, $\mathscr{R}(\boldsymbol{TF}') = \mathscr{R}(\boldsymbol{F})$, by choosing $\boldsymbol{U}$ suitably. Consequently, $\boldsymbol{F} \in \mathscr{F}_{\text{MLM}}$ if and only if $\boldsymbol{F}$ is of full column rank for the model $\mathscr{L}_{\rho}{}^{(2)}$. This is a trivial situation.

3.3. Linear sufficiency, linear minimal sufficiency, and linear completeness

We offer the definition to the linear minimal sufficiency w.r.t. BLUP as follows.

Definition 3.2. $\boldsymbol{Fy}$ is said to be linearly minimal sufficient w.r.t. the BLUP problem if $\boldsymbol{F} \in \mathscr{F}_{\text{MLM}}$ and for any other linearly sufficient statistic $\boldsymbol{F}_0\boldsymbol{y}$, there exists a matrix $\boldsymbol{B}$ such that $\boldsymbol{Fy} = \boldsymbol{BF}_0\boldsymbol{y}$ holds almost surely. We denote by $\mathscr{F}_{\text{MLM}-\min}$ the set of such $\boldsymbol{F}$'s. ■

We characterize $\mathscr{F}_{\text{MLM}-\min}$ in different ways in the following and, in addition, investigate relations among $\mathscr{F}_{\text{MLM}}$, $\mathscr{F}_{\text{MLM}-\min}$, and the notion of so-called linear completeness; cf. [8,12,20].

Theorem 3.5. *The following statements are mutually equivalent:*

(1) $\boldsymbol{F} \in \mathscr{F}_{\text{MLM}-\min}$*;*
(2) $\mathscr{R}(\boldsymbol{X}, \boldsymbol{ZG} + \boldsymbol{K}') = \mathscr{R}(\boldsymbol{TF}')$*;*
(3) $\mathscr{N}(\boldsymbol{F}) \cap \left\{\mathscr{R}(\boldsymbol{X}, \boldsymbol{ZG} + \boldsymbol{K}') \oplus \mathscr{R}\left[\Sigma\,(\boldsymbol{X}, \boldsymbol{ZG} + \boldsymbol{K}')^{\perp}\right]\right\} = \mathscr{R}\left[\boldsymbol{\Sigma}\,(\boldsymbol{X}, \boldsymbol{ZG} + \boldsymbol{K}')^{\perp}\right]$*;*
(4) $\mathscr{R}\begin{pmatrix} \boldsymbol{X}' & \boldsymbol{0} \\ \boldsymbol{0} & (\boldsymbol{X}^{\perp})'(\boldsymbol{ZG} + \boldsymbol{K}') \end{pmatrix} = \mathscr{R}\begin{pmatrix} \boldsymbol{X}'\boldsymbol{F}' \\ (\boldsymbol{X}^{\perp})'\boldsymbol{\Sigma}\boldsymbol{F}' \end{pmatrix}$.

Proof. The proof will be done by the following process, **(1)** $\Leftrightarrow$ **(2)** $\Leftrightarrow$ **(3)/(4)**:

(1)⇒(2) Let $\boldsymbol{Fy}$ be linearly minimally sufficient for all predictable functions of form $f(\boldsymbol{l}, \boldsymbol{m})$, and $\boldsymbol{F}_0\boldsymbol{y}$ be any other linearly sufficient statistic, i.e., $\boldsymbol{F} \in \mathscr{F}_{\text{MLM}-\min}$ and $\boldsymbol{F}_0 \in \mathscr{F}_{\text{MLM}}$. Then there exists a matrix $\boldsymbol{B}$ such that $\boldsymbol{Fy} = \boldsymbol{BF}_0\boldsymbol{y}$ holds almost surely, or equivalently,

$\boldsymbol{FT} = \boldsymbol{BF}_0\boldsymbol{T}$ $\big(\Rightarrow \mathscr{R}(\boldsymbol{TF}') \subseteq \mathscr{R}(\boldsymbol{TF}_0')\big)$, in which $\mathscr{R}(\boldsymbol{X}, \boldsymbol{ZG} + \boldsymbol{K}') \subseteq \mathscr{R}(\boldsymbol{TF}_0')$ and $\mathscr{R}(\boldsymbol{X}, \boldsymbol{ZG} + \boldsymbol{K}') \subseteq \mathscr{R}(\boldsymbol{TF}')$, or written jointly,

$$\mathscr{R}(\boldsymbol{X}, \boldsymbol{ZG} + \boldsymbol{K}') \subseteq \mathscr{R}(\boldsymbol{TF}') \subseteq \mathscr{R}(\boldsymbol{TF}_0'). \tag{3.11}$$

Take a special $\boldsymbol{F}_0$ such that $\text{rk}(\boldsymbol{X}, \boldsymbol{ZG} + \boldsymbol{K}') = \text{rk}(\boldsymbol{TF}_0')$, taking $\boldsymbol{F}_0 = (\boldsymbol{X}, \boldsymbol{ZG} + \boldsymbol{K}')'\boldsymbol{T}^+$ for instance. This combined with (3.11) would yield $\mathscr{R}(\boldsymbol{X}, \boldsymbol{ZG} + \boldsymbol{K}') = \mathscr{R}(\boldsymbol{TF}')$.

(2)⇒(1) It is clear.

(2)⇒(3) It is similar to the proof of Theorem 3.2.

(3)⇒(2) Provided $\mathscr{N}(\boldsymbol{F}) \cap \big\{\mathscr{R}(\boldsymbol{X}, \boldsymbol{ZG} + \boldsymbol{K}') \oplus \mathscr{R}\big[\boldsymbol{\Sigma}(\boldsymbol{X}, \boldsymbol{ZG} + \boldsymbol{K}')^\perp\big]\big\} = \mathscr{R}\big[\boldsymbol{\Sigma}(\boldsymbol{X}, \boldsymbol{ZG} + \boldsymbol{K}')^\perp\big]$. That is to say $\mathscr{R}\big(\boldsymbol{T}(\boldsymbol{TF}')^\perp\big) = \mathscr{R}\big(\boldsymbol{T}(\boldsymbol{X}, \boldsymbol{ZG} + \boldsymbol{K}')^\perp\big)$. Pre-multiplying by $\boldsymbol{F}$ gives $(\boldsymbol{TF}')'(\boldsymbol{X}, \boldsymbol{ZG} + \boldsymbol{K}')^\perp = \boldsymbol{0}$ and further one concludes that $\mathscr{R}(\boldsymbol{TF}') \subseteq \mathscr{R}(\boldsymbol{X}, \boldsymbol{ZG} + \boldsymbol{K}')$, while pre-multiplying with $(\boldsymbol{X}, \boldsymbol{ZG} + \boldsymbol{K}')'\boldsymbol{T}^+$ would mean that $(\boldsymbol{X}, \boldsymbol{ZG} + \boldsymbol{K}')'(\boldsymbol{TF}')^\perp = \boldsymbol{0}$ and further $\mathscr{R}(\boldsymbol{X}, \boldsymbol{ZG} + \boldsymbol{K}') \subseteq \mathscr{R}(\boldsymbol{TF}')$. Part **(2)** follows.

(2)⇒(4) Since $\boldsymbol{F}$ is supposed to belong to $\mathscr{F}_{\text{MLM}-\text{min}}$ (in view of **(2) ⇔ (1)**), $\boldsymbol{F} \in \mathscr{F}_{\text{MLM}}$ and thereby (3.10) is satisfied. Now, it suffices to show

$$\big(\text{rk}(\boldsymbol{X}) + \text{rk}[(\boldsymbol{I} - P_{\boldsymbol{X}})(\boldsymbol{ZG} + \boldsymbol{K}')] =\big)\, \text{rk}\begin{pmatrix} \boldsymbol{X}' & \boldsymbol{0} \\ \boldsymbol{0} & (\boldsymbol{X}^\perp)'(\boldsymbol{ZG} + \boldsymbol{K}') \end{pmatrix}$$
$$= \text{rk}\begin{pmatrix} \boldsymbol{X}'\boldsymbol{F}' \\ (\boldsymbol{X}^\perp)'\boldsymbol{\Sigma}\boldsymbol{F}' \end{pmatrix}$$

under the condition $\mathscr{R}(\boldsymbol{X}, \boldsymbol{ZG} + \boldsymbol{K}') = \mathscr{R}(\boldsymbol{TF}')$. This equality can be viewed as a direct consequence of the fact that $\text{rk}(\boldsymbol{X}, \boldsymbol{ZG} + \boldsymbol{K}') = \text{rk}(\boldsymbol{TF}') = \text{rk}[\boldsymbol{F}(\boldsymbol{X}, \boldsymbol{\Sigma}\boldsymbol{X}^\perp)]$ via the formula $\text{rk}(\boldsymbol{A}, \boldsymbol{B}) = \text{rk}(\boldsymbol{A}) + \text{rk}[(\boldsymbol{I} - P_{\boldsymbol{A}})\boldsymbol{B}]$.

(4)⇒(2) It is similar to **(3) ⇒ (2)** and is in view of the fact mentioned in the proof of **(2) ⇒ (4)**. ■

A statistic $\boldsymbol{Fy}$ is said to be linearly complete if each linear zero function of the form $\boldsymbol{AFy}$ vanishes almost surely. We denote by $\mathscr{F}_{\text{comp}}$ the set of such $\boldsymbol{F}$'s. By Drygas [8], $\boldsymbol{F} \in \mathscr{F}_{\text{comp}}$ if and only if $\mathscr{R}(\boldsymbol{F\Sigma}) \subseteq \mathscr{R}(\boldsymbol{FX})$. It is known that, when estimating $\boldsymbol{X}\beta$, linear completeness together with (ordinary) linear sufficiency is equivalent to (ordinary) linear minimal sufficiency. Isotalo and Puntanen [20] argued that the corresponding relation (among their linear prediction sufficiency, linear minimal prediction sufficiency, and linear completeness) does not seem to hold. However, this is not necessarily the case as regards the BLUP problem, when predicting $f(\boldsymbol{l}, \boldsymbol{m})$.

Theorem 3.6. *The following statements are mutually equivalent:*

(1) $\boldsymbol{F} \in \mathscr{F}_{\text{MLM}} \cap \mathscr{F}_{\text{comp}}$;
(2) $\boldsymbol{F} \in \mathscr{F}_{\text{MLM}-\text{min}}$ *and, in addition,* $\mathscr{R}\big[\boldsymbol{F}(\boldsymbol{ZG} + \boldsymbol{K}')\big] \subseteq \mathscr{R}(\boldsymbol{FX})$;
(3) $\mathscr{R}(\boldsymbol{X}) = \mathscr{R}(\boldsymbol{TF}')$.

Proof. It suffices to justify **(1)** ⇒ **(2)** ⇒ **(3)** ⇒ **(1)**.

(1)⇒(2) Provided $\boldsymbol{F} \in \mathscr{F}_{\text{MLM}} \cap \mathscr{F}_{\text{comp}}$. Observe that

$$\boldsymbol{F} \in \mathscr{F}_{\text{MLM}} \Leftrightarrow \mathscr{R}(\boldsymbol{X}, \boldsymbol{ZG} + \boldsymbol{K}') \subseteq \mathscr{R}(\boldsymbol{TF}'), \tag{3.12}$$
$$\boldsymbol{F} \in \mathscr{F}_{\text{comp}} \Leftrightarrow \mathscr{R}(\boldsymbol{F\Sigma}) \subseteq \mathscr{R}(\boldsymbol{FX}) \Rightarrow \text{rk}(\boldsymbol{FT}) = \text{rk}(\boldsymbol{FX}). \tag{3.13}$$

Combining (3.12) with (3.13), we obtain $\text{rk}(\boldsymbol{TF}') = \text{rk}(\boldsymbol{FX}) \leqslant \text{rk}(\boldsymbol{X}) \leqslant \text{rk}(\boldsymbol{X}, \boldsymbol{ZG} + \boldsymbol{K}') \leqslant \text{rk}(\boldsymbol{TF}')$, and thereby $\text{rk}(\boldsymbol{X}) = \text{rk}(\boldsymbol{X}, \boldsymbol{ZG} + \boldsymbol{K}') = \text{rk}(\boldsymbol{TF}')$. Consequently, the first equality combined with $\mathscr{R}(\boldsymbol{X}) \subseteq \mathscr{R}(\boldsymbol{X}, \boldsymbol{ZG}+\boldsymbol{K}')$ gives $\mathscr{R}(\boldsymbol{X}) = \mathscr{R}(\boldsymbol{X}, \boldsymbol{ZG}+\boldsymbol{K}')$ and therefore $\mathscr{R}(\boldsymbol{FX}) = \mathscr{R}\left[\boldsymbol{F}(\boldsymbol{X}, \boldsymbol{ZG}+\boldsymbol{K}')\right]$, which is further equivalent to $\mathscr{R}\left[\boldsymbol{F}(\boldsymbol{ZG}+\boldsymbol{K}')\right] \subseteq \mathscr{R}(\boldsymbol{FX})$; on the other hand, the second equality combined with $\mathscr{R}(\boldsymbol{X}, \boldsymbol{ZG}+\boldsymbol{K}') \subseteq \mathscr{R}(\boldsymbol{TF}')$ yields $\mathscr{R}(\boldsymbol{X}, \boldsymbol{ZG}+\boldsymbol{K}') = \mathscr{R}(\boldsymbol{TF}')$, which combined with Theorem 3.5 completes the proof of **(1)** $\Rightarrow$ **(2)**.

(2)⇒(3) In view of the following:

$$\begin{aligned}\mathscr{R}(\boldsymbol{X}, \boldsymbol{ZG}+\boldsymbol{K}') &= \mathscr{R}(\boldsymbol{TF}') \Rightarrow \operatorname{rk}\left[\boldsymbol{F}(\boldsymbol{X}, \boldsymbol{ZG}+\boldsymbol{K}')\right] = \operatorname{rk}\left(\boldsymbol{FTF}'\right)\\ &= \operatorname{rk}(\boldsymbol{TF}') = \operatorname{rk}(\boldsymbol{X}, \boldsymbol{ZG}+\boldsymbol{K}'),\\ \mathscr{R}\left[\boldsymbol{F}(\boldsymbol{ZG}+\boldsymbol{K}')\right] &\subseteq \mathscr{R}(\boldsymbol{FX}) \Leftrightarrow \mathscr{R}\left[\boldsymbol{F}(\boldsymbol{X}, \boldsymbol{ZG}+\boldsymbol{K}')\right]\\ &= \mathscr{R}(\boldsymbol{FX}) \Rightarrow \operatorname{rk}\left[\boldsymbol{F}(\boldsymbol{X}, \boldsymbol{ZG}+\boldsymbol{K}')\right] \leqslant \operatorname{rk}(\boldsymbol{X}),\end{aligned}$$

we obtain

$$\operatorname{rk}\left[\boldsymbol{F}(\boldsymbol{X}, \boldsymbol{ZG}+\boldsymbol{K}')\right] \leqslant \operatorname{rk}(\boldsymbol{X}) \leqslant \operatorname{rk}(\boldsymbol{X}, \boldsymbol{ZG}+\boldsymbol{K}') = \operatorname{rk}\left[\boldsymbol{F}(\boldsymbol{X}, \boldsymbol{ZG}+\boldsymbol{K}')\right].$$

Thus $\operatorname{rk}(\boldsymbol{X}) = \operatorname{rk}(\boldsymbol{TF}')$, which combined with $\mathscr{R}(\boldsymbol{X}) \subseteq \mathscr{R}(\boldsymbol{X}, \boldsymbol{ZG}+\boldsymbol{K}') = \mathscr{R}(\boldsymbol{TF}')$ gives $\mathscr{R}(\boldsymbol{X}) = \mathscr{R}(\boldsymbol{TF}')$.

(3)⇒(1) It is clear.

The proof is completed. ■

4. Concluding remarks

In this paper, we consider the problem of predicting a general linear combination of fixed effects and realized values of random effects in a general mixed linear model and we offered an explicit representation of the BLUP w.r.t. the MSE criterion and, in addition, applied the resulting conclusion to several special models. An alternative to characterization of BLUP was offered. Secondly, the notion of linear sufficiency w.r.t. the BLUP problem was considered and its characterization was investigated in three different ways. Finally, we studied concepts of linear minimal sufficiency w.r.t. BLUP and gave relations among $\mathscr{F}_{\text{MLM}}$, $\mathscr{F}_{\text{MLM}-\min}$ and $\mathscr{F}_{\text{comp}}$. Some significant remarks may be as follows:

1. *Relative efficiency*. It is seen that $\boldsymbol{Fy}$ is not always linearly sufficient. In this case, one can consider the problem of relative efficiency of f_* with respect to f^*, which varies in (0, 1]. If the value of the relative efficiency is near to 1^-, the loss will be trivial substituting f^* with f_*. Conversely, the loss is unacceptable if the value tends to 0^+. As we know, for this topic it is of interest to seek the upper or lower bound in practice.
2. *Objective function*. We mainly considered the scalar objective function in this article. This can be expanded readily to predicting $\boldsymbol{L}'\beta + \boldsymbol{M}'\boldsymbol{u}$ with $\boldsymbol{L}$ and $\boldsymbol{M}$ being matrices of suitable order w.r.t. the MSE criterion or MSE matrix criterion. It is not difficult to see that $\boldsymbol{L}'\beta^* + \boldsymbol{M}'\boldsymbol{u}^*$ has the minimal prediction MSE in the set of all linear unbiased predictors and further it possesses MSE matrix optimality in the Löwner sense.
3. The link of $\mathscr{F}_{\text{MLM}}$, $\mathscr{F}_{\text{MLM}-\min}$, $\mathscr{F}_{\text{comp}}$ w.r.t. BLUP and ordinary ones can be derived and viewed as an idea for how to deal with corresponding problems.
4. As pointed out by a referee, we can see that $\boldsymbol{Z} = \boldsymbol{0}$ in conjunction with $\boldsymbol{K} \neq \boldsymbol{0}$ results in an interesting topic. It is actually a combination of ordinary linear sufficiency and the so-called linear prediction sufficiency proposed by Isotalo and Puntanen [20] w.r.t. a special prediction model in some sense.

Acknowledgments

The authors are grateful to the two anonymous referees for many valuable comments and constructive suggestions resulting in the present version. Thanks also to an anonymous associate editor and the journal manager for helpful aspects. This research was supported by Grants HGQ0637 and HGQN0725 from the Huaiyin Institute of Technology.

References

[1] J.K. Baksalary, H.G. Drygas, A note on the concepts of sufficiency in the general Gauss–Markov model—a coordinatefree approach, Mathematische Schriften Kassel, Vordruck-Reihe des Fachbereichs 17 der Gesamthochschule Kassel, Preprint no. 92/2, 1992.
[2] J.K. Baksalary, R. Kala, Linear transformations preserving the best linear unbiased estimator in general Gauss–Markoff model, Ann. Statist. 9 (1981) 913–916.
[3] J.K. Baksalary, T. Mathew, Linear sufficiency and completeness in an incorrectly specified general Gauss–Markov model, Sankhya A 48 (1986) 169–180.
[4] G.E. Battese, R.M. Harter, W.A. Fuller, An error-components model for prediction of county crop area using survey and satellite data, J. Amer. Statist. Assoc. 83 (1988) 28–36.
[5] P. Bhimasankaram, D. Sengupta, The linear zero functions approach to linear models, Sankhyā, Ser. B 58 (1996) 338–351.
[6] K. Das, J. Jiang, J.N.K. Rao, Mean squared error of empirical predictor, Ann. Statist. 32 (2004) 818–840.
[7] A.P. Dempster, D.B. Rubin, R.K. Tsutakawa, Estimation in covariance component models, J. Amer. Statist. Assoc. 76 (1981) 341–353.
[8] H. Drygas, Sufficiency and completeness in general Gauss–Markoff model, Sankhya A 45 (1983) 88–98.
[9] H. Drygas, Linear sufficiency and some applications in multilinear estimation, J. Multivariate Anal. 16 (1985) 71–84.
[10] H. Drygas, Linear and quadratic sufficiency—the linear Rao–Blackwell theorem, in: S.K. Basu, B.K. Sinha (Eds.), Proceedings of the Calcutta Statistical Meeting, Norosa Publishing House, New Delhi, 1993, pp. 121–125. 27-12-1991–1-1-1992.
[11] R.E. Fay, R.A. Herriot, Estimates of income for small places: An application of James–Stein procedures to census data, J. Amer. Statist. Assoc. 74 (1979) 269–277.
[12] J. Groß, A note on the concepts of linear and quadratic sufficiency, J. Statist. Plann. Inference 70 (1998) 69–76.
[13] D.A. Harville, Extension of the Gauss–Markov Theorem to include the estimation of random effects, Ann. Statist. 4 (1976) 384–395.
[14] D.A. Harville, D.R. Jeske, Mean squared error of estimation or prediction under a general linear model, J. Amer. Statist. Assoc. 87 (1992) 724–731.
[15] B. Heiligers, A. Markiewicz, Linear sufficiency and admissibility in restricted linear models, Statist. Probab. Lett. 30 (1996) 105–111.
[16] C.R. Henderson, Best linear unbiased estimation and prediction under a selection model, Biometrics 31 (1975) 423–447.
[17] P. Ibarrola, A. Pérez-Palomares, Linear completeness in a continuous time Gauss–Markov model, Statist. Probab. Lett. 69 (2004) 143–149.
[18] P. Ibarrola, A. Pérez-Palomares, Linear sufficiency and linear admissibility in a continuous time Gauss–Markov model, J. Multivariate Anal. 87 (2003) 315–327.
[19] W.C. Ip, H. Wong, J.S. Liu, Sufficient and admissible estimators in general multivariate linear model, J. Statist. Plann. Inference 135 (2005) 371–383.
[20] J. Isotalo, S. Puntanen, Linear prediction sufficiency for new observations in the general Gauss–Markov model, Commun. Statist. A 35 (2006) 1011–1023.
[21] X. Liu, J. Rong, Nonnegative quadratic estimation and quadratic sufficiency in general linear models, J. Multivariate Anal. 98 (2007) 1180–1194.
[22] A. Markiewicz, Comparison of linear restricted models with respect to the validity of admissible and linearly sufficient estimators, Statist. Probab. Lett. 38 (1998) 347–354.
[23] J. Müeller, Sufficiency and completeness in the linear model, Multivariate Anal. 21 (1987) 312–323.
[24] N.G.N. Prasad, J.N.K. Rao, The estimation of the mean square error of small-area estimators, J. Amer. Statist. Assoc. 85 (1990) 163–171.
[25] G.K. Robinson, That BLUP is a good thing: The estimation of random effects, Statist. Sci. 6 (1991) 15–51.