This is the accepted manuscript of the article, which has been published in Tez, M. & von Rosen, D. (eds) Trends and Perspectives in Linear Statistical Inference: Proceedings of the LINSTAT2016 meeting held 22-25 August 2016 in Istanbul, Turkey. Cham: Springer. Contributions to Statistics. ISBN: 978-3-319-73240-4. ISSN: 1431-1968. https://doi.org/10.1007/978-3-319-73241-1_8

Some properties of linear prediction sufficiency in the linear model

Jarkko Isotalo

Department of Forest Sciences, P.O. Box 27
FI-00014 University of Helsinki, Finland
email: jarkko.isotalo@helsinki.fi

Augustyn Markiewicz

Department of Mathematical and Statistical Methods
Poznań University of Life Sciences
Wojska Polskiego 28, PL-60637 Poznań, Poland
email: amark@up.poznan.pl

Simo Puntanen[1]

Faculty of Natural Sciences
FI-33014 University of Tampere, Finland
email: simo.puntanen@uta.fi

Abstract. A linear statistic $\mathbf{Fy}$ is called linearly prediction sufficient, or shortly BLUP-sufficient, for the new observation $\mathbf{y}_*$, say, if there exists a matrix $\mathbf{A}$ such that $\mathbf{AFy}$ is the best linear unbiased predictor, BLUP, for $\mathbf{y}_*$. We review some properties of linear prediction sufficiency that have not been received much attention in the literature and provide some clarifying comments. In particular, we consider the best linear unbiased prediction of the error term related to $\mathbf{y}_*$. We also explore some interesting properties of mixed linear models including the connection between a particular extended linear model and its transformed version.

Key words and phrases: Best linear unbiased estimator, BLUE, Best linear unbiased predictor, BLUP, linear sufficiency, orthogonal projector, transformed linear model.

MSC: 62J05, 62J10.

[1]Author responsible for correspondence

1 Introduction

To make the article more self-readable we go through some basic concepts related to linear sufficiency. So, let us get started with the general linear model $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, shortly denoted as a triplet

$$\mathscr{M} = \{\mathbf{y}, \mathbf{X}\boldsymbol{\beta}, \mathbf{V}\},$$

where $\mathbf{X}_{n\times p}$ is a known model matrix, the vector $\mathbf{y}$ is an observable n-dimensional random vector, $\boldsymbol{\beta}$ is a $p \times 1$ vector of unknown parameters, and $\boldsymbol{\varepsilon}$ is an unobservable vector of random errors with expectation $\mathrm{E}(\boldsymbol{\varepsilon}) = \mathbf{0}$, and covariance matrix $\mathrm{cov}(\boldsymbol{\varepsilon}) = \mathbf{V}$, where the nonnegative definite matrix $\mathbf{V}$ is known and can be singular. Premultiplying the model $\mathscr{M}$ by $\mathbf{F}_{f\times n}$ yields the transformed model

$$\mathscr{M}_t = \{\mathbf{F}\mathbf{y}, \mathbf{F}\mathbf{X}\boldsymbol{\beta}, \mathbf{F}\mathbf{V}\mathbf{F}'\},$$

which will have a crucial role in our considerations.

Let $\mathbf{y}_*$ denote a $q \times 1$ unobservable random vector containing new future observations. The new observations are assumed to follow the linear model

$$\mathbf{y}_* = \mathbf{X}_*\boldsymbol{\beta} + \boldsymbol{\varepsilon}_*,$$

where $\mathbf{X}_*$ is a known $q \times p$ matrix, $\boldsymbol{\beta}$ is the same vector of unknown parameters as in $\mathscr{M}$, and $\boldsymbol{\varepsilon}_*$ is a q-dimensional random error vector. The expectation and the covariance matrix are

$$\mathrm{E}\begin{pmatrix}\mathbf{y}\\ \mathbf{y}_*\end{pmatrix} = \begin{pmatrix}\mathbf{X}\boldsymbol{\beta}\\ \mathbf{X}_*\boldsymbol{\beta}\end{pmatrix} = \begin{pmatrix}\mathbf{X}\\ \mathbf{X}_*\end{pmatrix}\boldsymbol{\beta}, \quad \mathrm{cov}\begin{pmatrix}\mathbf{y}\\ \mathbf{y}_*\end{pmatrix} = \begin{pmatrix}\mathbf{V} & \mathbf{V}_{12}\\ \mathbf{V}_{21} & \mathbf{V}_{22}\end{pmatrix} = \boldsymbol{\Gamma},$$

where the covariance matrix matrix $\boldsymbol{\Gamma}$ is assumed to be known. For brevity, we denote the linear model with new observations as

$$\mathscr{M}_* = \left\{\begin{pmatrix}\mathbf{y}\\ \mathbf{y}_*\end{pmatrix}, \begin{pmatrix}\mathbf{X}\\ \mathbf{X}_*\end{pmatrix}\boldsymbol{\beta}, \begin{pmatrix}\mathbf{V} & \mathbf{V}_{12}\\ \mathbf{V}_{21} & \mathbf{V}_{22}\end{pmatrix}\right\}.$$

Our main interest in $\mathscr{M}_*$ lies in predicting $\mathbf{y}_*$ on the basis of observable $\mathbf{y}$.

Suppose we transform $\mathscr{M}$ into $\mathscr{M}_t$ and do the prediction in this situation. Corresponding to $\mathscr{M}_*$ we have now the following setup:

$$\mathscr{M}_{t*} = \left\{\begin{pmatrix}\mathbf{F}\mathbf{y}\\ \mathbf{y}_*\end{pmatrix}, \begin{pmatrix}\mathbf{F}\mathbf{X}\\ \mathbf{X}_*\end{pmatrix}\boldsymbol{\beta}, \begin{pmatrix}\mathbf{F}\mathbf{V}\mathbf{F}' & \mathbf{F}\mathbf{V}_{12}\\ \mathbf{V}_{21}\mathbf{F}' & \mathbf{V}_{22}\end{pmatrix}\right\}.$$

As for notation, let $\mathbb{R}^{m\times n}$ denote the set of $m \times n$ real matrices. The symbols $\mathbf{A}'$, $\mathbf{A}^-$, $\mathbf{A}^+$, $\mathscr{C}(\mathbf{A})$, and $\mathscr{C}(\mathbf{A})^\perp$, denote, respectively, the transpose, a generalized inverse, the Moore–Penrose inverse, the column space, and the orthogonal complement of the column space of the matrix $\mathbf{A}$. By $(\mathbf{A} : \mathbf{B})$

we denote the partitioned matrix with $\mathbf{A}_{m\times n}$ and $\mathbf{B}_{m\times k}$ as submatrices. By $\mathbf{A}^{\perp}$ we denote any matrix satisfying $\mathscr{C}(\mathbf{A}^{\perp}) = \mathscr{C}(\mathbf{A})^{\perp}$. Furthermore, we will write $\mathbf{P}_{\mathbf{A}} = \mathbf{A}\mathbf{A}^{+} = \mathbf{A}(\mathbf{A}'\mathbf{A})^{-}\mathbf{A}'$ to denote the orthogonal projector (with respect to the standard inner product) onto $\mathscr{C}(\mathbf{A})$, and $\mathbf{Q}_{\mathbf{A}} = \mathbf{I} - \mathbf{P}_{\mathbf{A}}$. In particular, we denote $\mathbf{M} = \mathbf{I}_n - \mathbf{P}_{\mathbf{X}}$. One choice for $\mathbf{X}^{\perp}$ is of course $\mathbf{M}$.

The linear estimator $\mathbf{Gy}$ is the best linear unbiased estimator, BLUE, of $\mathbf{X}\boldsymbol{\beta}$ whenever $\mathbf{Gy}$ is unbiased and it has the smallest covariance matrix (in the Löwner sense) among all linear unbiased estimators of $\mathbf{X}\boldsymbol{\beta}$. The following lemma characterises the BLUE; see, e.g., Drygas (1970, p. 55), Rao (1973, p. 282), and more recently Baksalary & Trenkler (2009).

Lemma 1.1. *Consider the general linear model* $\mathscr{M} = \{\mathbf{y}, \mathbf{X}\boldsymbol{\beta}, \mathbf{V}\}$. *Then the estimator* $\mathbf{Gy}$ *is the* BLUE *for* $\mathbf{X}\boldsymbol{\beta}$ *if and only if* $\mathbf{G}$ *satisfies the equation*

$$\mathbf{G}(\mathbf{X} : \mathbf{V}\mathbf{X}^{\perp}) = (\mathbf{X} : \mathbf{0})\,. \tag{1.1}$$

The corresponding condition for $\mathbf{Ay}$ *to be the* BLUE *of an estimable parametric function* $\mathbf{K}\boldsymbol{\beta}$, *i.e.,* $\mathscr{C}(\mathbf{K}') \subset \mathscr{C}(\mathbf{X}')$, *is*

$$\mathbf{A}(\mathbf{X} : \mathbf{V}\mathbf{X}^{\perp}) = (\mathbf{K} : \mathbf{0})\,.$$

We assume the model $\mathscr{M}$ to be consistent in the sense that the observed value of $\mathbf{y}$ lies in $\mathscr{C}(\mathbf{X} : \mathbf{V})$ with probability 1. Hence we assume that under $\mathscr{M}$

$$\mathbf{y} \in \mathscr{C}(\mathbf{X} : \mathbf{V}) = \mathscr{C}(\mathbf{X} : \mathbf{V}\mathbf{X}^{\perp}) = \mathscr{C}(\mathbf{X} : \mathbf{V}\mathbf{M})\,.$$

The corresponding consistency is assumed in all models that we will consider. Moreover, in the consistent linear model $\mathscr{M}$, the estimators $\mathbf{G}_1\mathbf{y}$ and $\mathbf{G}_2\mathbf{y}$ are said to be equal with probability 1 if

$$\mathbf{G}_1\mathbf{y} = \mathbf{G}_2\mathbf{y} \quad \text{for all } \mathbf{y} \in \mathscr{C}(\mathbf{X} : \mathbf{V})\,.$$

The linear predictor $\mathbf{By}$ is said to be unbiased for $\mathbf{y}_*$ if $\mathrm{E}(\mathbf{y}_* - \mathbf{By}) = \mathbf{0}$ for all $\boldsymbol{\beta} \in \mathbb{R}^p$. This is equivalent to $\mathbf{X}_*' = \mathbf{X}'\mathbf{B}'$. The inclusion $\mathscr{C}(\mathbf{X}_*') \subset \mathscr{C}(\mathbf{X}')$ is the well-known condition for the estimability of $\mathbf{X}_*\boldsymbol{\beta}$ under $\mathscr{M}$. When $\mathscr{C}(\mathbf{X}_*') \subset \mathscr{C}(\mathbf{X}')$ holds, will say that $\mathbf{y}_*$ is predictable under $\mathscr{M}_*$. Now a linear unbiased predictor $\mathbf{By}$ is the best linear unbiased predictor, BLUP, for $\mathbf{y}_*$, if the Löwner ordering

$$\mathrm{cov}(\mathbf{y}_* - \mathbf{By}) \leq_{\mathrm{L}} \mathrm{cov}(\mathbf{y}_* - \mathbf{Cy})$$

holds for all $\mathbf{C}$ such that $\mathbf{Cy}$ is an unbiased linear predictor for $\mathbf{y}_*$.

The following lemma characterises the BLUP; for the proof, see, e.g., Christensen (2011, p. 294), and Isotalo & Puntanen (2006, p. 1015).

Lemma 1.2. *Consider the linear model* $\mathscr{M}_*$, *where* $\mathscr{C}(\mathbf{X}_*') \subset \mathscr{C}(\mathbf{X}')$, *i.e.,* $\mathbf{y}_*$ *is predictable. The linear predictor* $\mathbf{By}$ *is the best linear unbiased predictor* (BLUP) *for* $\mathbf{y}_*$ *if and only if* $\mathbf{B}$ *satisfies the equation*

$$\mathbf{B}(\mathbf{X} : \mathbf{V}\mathbf{X}^{\perp}) = (\mathbf{X}_* : \mathbf{V}_{21}\mathbf{X}^{\perp}) = (\mathbf{X}_* : \mathrm{cov}(\mathbf{y}_*, \mathbf{y})\mathbf{X}^{\perp})\,.$$

We will frequently utilise Lemma 2.2.4 of Rao & Mitra (1971), which says that for nonnull matrices $\mathbf{A}$ and $\mathbf{C}$ the following holds:

$$\mathbf{AB}^-\mathbf{C} = \mathbf{AB}^+\mathbf{C} \iff \mathscr{C}(\mathbf{C}) \subset \mathscr{C}(\mathbf{B}) \ \& \ \mathscr{C}(\mathbf{A}') \subset \mathscr{C}(\mathbf{B}') \,. \tag{1.2}$$

One well-known solution for $\mathbf{G}$ in (1.1) (which is always solvable) is

$$\mathbf{P}_{\mathbf{X};\mathbf{W}^-} := \mathbf{X}(\mathbf{X}'\mathbf{W}^-\mathbf{X})^-\mathbf{X}'\mathbf{W}^-,$$

where $\mathbf{W}$ is a matrix belonging to the set of nonnegative definite matrices defined as

$$\mathcal{W} = \left\{ \mathbf{W} \in \mathbb{R}^{n\times n} : \mathbf{W} = \mathbf{V} + \mathbf{XUU}'\mathbf{X}',\ \mathscr{C}(\mathbf{W}) = \mathscr{C}(\mathbf{X} : \mathbf{V}) \right\}.$$

Denoting

$$\mathbf{P}_{\mathbf{X};\mathbf{W}^+} := \mathbf{X}(\mathbf{X}'\mathbf{W}^-\mathbf{X})^-\mathbf{X}'\mathbf{W}^+,$$

we observe, in view of (1.2), that $\mathbf{P}_{\mathbf{X};\mathbf{W}^-}\mathbf{y} = \mathbf{P}_{\mathbf{X};\mathbf{W}^+}\mathbf{y}$ for all $\mathbf{y} \in \mathscr{C}(\mathbf{W})$.

The structure of the contribution is as follows. In Section 2 we recall some well-known conditions for the BLUE- and BLUP-sufficiency and in particular clarify and extend some concepts related to BLUP-sufficiency. In Section 3 we introduce some representations for the BLUPs and and explore the corresponding sufficiency relations. Section 4 provides some representations for the BLUPs and BLUEs and in Section 5 we apply our results to the linear mixed models. While writing this contribution, our attempt has been to call well-known (or pretty well-known) results Lemmas, while Theorems refer to our own contributions or clarifications.

2 Conditions for linear sufficiency and linear prediction sufficiency

A linear statistic $\mathbf{Fy}$, where $\mathbf{F} \in \mathbb{R}^{f\times n}$, is called linearly sufficient for $\mathbf{X}\boldsymbol{\beta}$ under the model $\mathscr{M} = \{\mathbf{y}, \mathbf{X}\boldsymbol{\beta}, \mathbf{V}\}$, if there exists a matrix $\mathbf{A} \in \mathbb{R}^{n\times f}$ such that $\mathbf{AFy}$ is the BLUE for $\mathbf{X}\boldsymbol{\beta}$. Correspondingly, $\mathbf{Fy}$ is linearly sufficient for estimable $\mathbf{K}\boldsymbol{\beta}$, where $\mathbf{K} \in \mathbb{R}^{k\times p}$, if there exists a matrix $\mathbf{A} \in \mathbb{R}^{k\times f}$ such that $\mathbf{AFy}$ is the BLUE for $\mathbf{K}\boldsymbol{\beta}$. To have a slightly shorter terminology, we often will use the phrase "BLUE-sufficient" and the notation $\mathbf{Fy} \in \mathcal{S}(\mathbf{K}\boldsymbol{\beta})$.

For the following Lemma 2.1 and Lemma 2.2, see, e.g., Baksalary & Kala (1981, 1986), Drygas (1983), Tian & Puntanen (2009, Th. 2.8), and Kala et al. (2017, Th. 2).

Lemma 2.1. *The statistic* $\mathbf{Fy}$ *is* BLUE-*sufficient for* $\mathbf{X}\boldsymbol{\beta}$ *under the model* $\mathscr{M} = \{\mathbf{y}, \mathbf{X}\boldsymbol{\beta}, \mathbf{V}\}$ *if and only if any of the following equivalent statements holds:*

(a) $\mathscr{C}\begin{pmatrix}\mathbf{X}'\\ \mathbf{0}\end{pmatrix} \subset \mathscr{C}\begin{pmatrix}\mathbf{X}'\mathbf{F}'\\ \mathbf{MVF}'\end{pmatrix}$,

(b) $\mathscr{C}(\mathbf{X}) \subset \mathscr{C}(\mathbf{WF}')$, *where* $\mathbf{W} \in \mathcal{W}$,

(c) $\mathscr{C}(\mathbf{X}'\mathbf{F}') = \mathscr{C}(\mathbf{X}')$ *and* $\mathscr{C}(\mathbf{FX}) \cap \mathscr{C}(\mathbf{FVX}^{\perp}) = \{\mathbf{0}\}$.

Let $\mathbf{K}\boldsymbol{\beta}$ *be estimable under* $\mathscr{M}$. *Then* $\mathbf{Fy}$ *is* BLUE-*sufficient for* $\mathbf{K}\boldsymbol{\beta}$ *if and only if*

(d) $\mathscr{C}\begin{pmatrix}\mathbf{K}' \\ \mathbf{0}\end{pmatrix} \subset \mathscr{C}\begin{pmatrix}\mathbf{X}'\mathbf{F}' \\ \mathbf{MVF}'\end{pmatrix}$.

Let $\mathbf{F}_0$ be a matrix with property $\mathscr{C}(\mathbf{F}_0') = \mathscr{C}(\mathbf{F}')$. Then Lemma 2.1 immediately implies the following:

$$\mathbf{Fy} \in \mathcal{S}(\mathbf{K}\boldsymbol{\beta}) \iff \mathbf{F}_0\mathbf{y} \in \mathcal{S}(\mathbf{K}\boldsymbol{\beta}). \tag{2.1}$$

If $\mathscr{C}(\mathbf{F}_0') \subset \mathscr{C}(\mathbf{F}')$, then the implication "$\Longleftarrow$" is holding in (2.1).

Lemma 2.2. *Consider the model* $\mathscr{M} = \{\mathbf{y}, \mathbf{X}\boldsymbol{\beta}, \mathbf{V}\}$ *and its transformed version* $\mathscr{M}_t = \{\mathbf{Fy}, \mathbf{FX}\boldsymbol{\beta}, \mathbf{FVF}'\}$, *and let* $\mathbf{K}\boldsymbol{\beta}$ *be estimable under* $\mathscr{M}$ *and* $\mathscr{M}_t$. *Then the following statements are equivalent:*

(a) $\mathbf{Fy}$ *is* BLUE-*sufficient for* $\mathbf{K}\boldsymbol{\beta}$.

(b) $\mathrm{BLUE}(\mathbf{K}\boldsymbol{\beta} \mid \mathscr{M}) = \mathrm{BLUE}(\mathbf{K}\boldsymbol{\beta} \mid \mathscr{M}_t)$ *with probability* 1.

(c) *There exists at least one representation of* BLUE *of* $\mathbf{K}\boldsymbol{\beta}$ *under* $\mathscr{M}$ *which is the* BLUE *also under the transformed model* $\mathscr{M}_t$.

Notice that the parametric function $\mathbf{K}\boldsymbol{\beta}$ is estimable under $\mathscr{M}$ as well as under $\mathscr{M}_t$ if and only if

$$\mathscr{C}(\mathbf{K}') \subset \mathscr{C}(\mathbf{X}') \cap \mathscr{C}(\mathbf{X}'\mathbf{F}') = \mathscr{C}(\mathbf{X}'\mathbf{F}'), \tag{2.2}$$

while $\mathbf{X}\boldsymbol{\beta}$ is estimable under $\mathscr{M}_t$ whenever

$$\mathscr{C}(\mathbf{X}') = \mathscr{C}(\mathbf{X}'\mathbf{F}'), \quad \text{i.e.,} \quad \mathrm{rank}(\mathbf{X}) = \mathrm{rank}(\mathbf{FX}).$$

The concept of linear prediction sufficiency is defined analogically as follows: Let $\mathbf{y}_*$ be predictable under the model $\mathscr{M}_*$, i.e., $\mathscr{C}(\mathbf{X}_*') \subset \mathscr{C}(\mathbf{X}')$. Then $\mathbf{Fy}$ is called linearly prediction sufficient for $\mathbf{y}_*$ if there exists a matrix $\mathbf{A}$ such that $\mathbf{AFy}$ is the BLUP for $\mathbf{y}_*$; that is, there exists a matrix $\mathbf{A}$ such that

$$\mathbf{AF}(\mathbf{X} : \mathbf{VM}) = (\mathbf{X}_* : \mathbf{V}_{21}\mathbf{M}). \tag{2.3}$$

Corresponding to the phrase "BLUE-sufficient", we may use the term "BLUP-sufficient" and the notation $\mathbf{Fy} \in \mathcal{S}(\mathbf{y}_*)$.

The following theorem collects together some important properties of the linear prediction sufficiency.

Theorem 2.1. *Suppose that* $\mathbf{y}_*$ *is predictable under* $\mathscr{M}_*$ *and* $\mathscr{M}_{t*}$. *Then:*

(a) *If* $\mathbf{Fy}$ *is* BLUP*-sufficient for* $\mathbf{y}_*$*, then every representation of the* BLUP *for* $\mathbf{y}_*$ *under the transformed model* $\mathscr{M}_{t*}$ *is* BLUP *also under the original model* $\mathscr{M}_*$.

Moreover, the following statements are equivalent:

(b) $\mathbf{Fy}$ *is* BLUP*-sufficient for* $\mathbf{y}_*$*, or shortly* $\mathbf{Fy} \in \mathcal{S}(\mathbf{y}_*)$.

(c) $\mathscr{C}\begin{pmatrix}\mathbf{X}'_*\\ \mathbf{MV}_{12}\end{pmatrix} \subset \mathscr{C}\begin{pmatrix}\mathbf{X}'\mathbf{F}'\\ \mathbf{MVF}'\end{pmatrix}$.

(d) $\mathrm{BLUP}(\mathbf{y}_* \mid \mathscr{M}_*) = \mathrm{BLUP}(\mathbf{y}_* \mid \mathscr{M}_{t*})$ *with probability* 1.

(e) *There exists at least one representation of* BLUP *of* $\mathbf{y}_*$ *under* $\mathscr{M}_*$ *which is* BLUP *also under the transformed model* $\mathscr{M}_{t*}$.

Proof. The claim (a) was proved by Isotalo & Puntanen (2006, Th. 3.2); see also Remark 2.1 below. The equivalence of (b) and (c) is obvious because (b) means that there exists a matrix $\mathbf{A}$ such that (2.3) holds. Suppose that (2.3) holds for some $\mathbf{A}$. Then the same multiplier $\mathbf{AF}$ gives the BLUP for $\mathbf{y}_*$ under the transformed model $\mathscr{M}_{t*}$ if and only if

$$\mathbf{A}(\mathbf{FX} : \mathbf{FVF}'\mathbf{Q}_{\mathbf{FX}}) = (\mathbf{X}_* : \mathbf{V}_{21}\mathbf{F}'\mathbf{Q}_{\mathbf{FX}}). \tag{2.4}$$

In view of Markiewicz & Puntanen (2017, Lemma 5) and Rao & Mitra (1971, Compl. 7, p. 118), the following holds:

$$\mathscr{C}(\mathbf{F}'\mathbf{Q}_{\mathbf{FX}}) = \mathscr{C}(\mathbf{F}') \cap \mathscr{C}(\mathbf{M}), \tag{2.5a}$$

$$\mathbf{F}'\mathbf{Q}_{\mathbf{FX}} = \mathbf{MF}'\mathbf{Q}_{\mathbf{FX}}. \tag{2.5b}$$

Substituting (2.5b) into (2.4) we immediately see that (2.3) implies (2.4) i.e., (b) implies (e). The statement (d) means that we have the equality

$$\mathbf{B}(\mathbf{X} : \mathbf{V}) = \mathbf{CF}(\mathbf{X} : \mathbf{V}) \tag{2.6}$$

for some $\mathbf{B}$ and $\mathbf{C}$ satisfying

$$\mathbf{B}(\mathbf{X} : \mathbf{VM}) = (\mathbf{X}_* : \mathbf{V}_{21}\mathbf{M}),$$
$$\mathbf{C}(\mathbf{FX} : \mathbf{FVF}'\mathbf{Q}_{\mathbf{FX}}) = (\mathbf{X}_* : \mathbf{V}_{21}\mathbf{F}'\mathbf{Q}_{\mathbf{FX}}).$$

Now, $\mathbf{By}$ is BLUP for $\mathbf{y}_*$ under $\mathscr{M}_*$ and and hence, in light of (2.6), $\mathbf{CFy}$ is also BLUP for $\mathbf{y}_*$ under $\mathscr{M}_*$, and thus by definition, $\mathbf{Fy}$ is BLUP-sufficient for $\mathbf{y}_*$. Hence we have shown that (d) implies (b). It is obvious that (e) implies (d) and thereby the proof is completed. □

The above proof is parallel to that of Kala et al. (2017, Th. 2) concerning Lemma 2.2.

Remark 2.1. Regarding the claim (a) in Theorem 2.1, Isotalo & Puntanen (2006, Th. 3.2) state the following: "Every representation of the BLUP for $\mathbf{y}_*$ under the transformed model $\mathscr{M}_{t*}$ is BLUP also under the original model $\mathscr{M}_*$ and vice versa." Stated in this way, the *vice versa* part is not quite correct and may result in wrong or confusing interpretations. Hence we will clarify the meaning of the *vice versa* part below. The corresponding considerations for the BLUE of estimable parametric function are done in Kala et al. (2017, Sec. 4) and here we proceed along their lines.

To do this, we take a look at the multipliers of the response vector $\mathbf{y}$ when obtaining the BLUPs. Let $\mathbf{y}_*$ be predictable under the models $\mathscr{M}_*$ and $\mathscr{M}_{t*}$ and denote

$$\begin{aligned}\mathcal{A} &= \{\mathbf{A} : \mathbf{AFy} = \mathrm{BLUP}(\mathbf{y}_* \mid \mathscr{M}_*)\} \\ &= \{\mathbf{A} : \mathbf{AF}(\mathbf{X} : \mathbf{VM}) = (\mathbf{X}_* : \mathbf{V}_{21}\mathbf{M})\}\,, \\ \mathcal{C} &= \{\mathbf{C} : \mathbf{CFy} = \mathrm{BLUP}(\mathbf{y}_* \mid \mathscr{M}_{t*})\} \\ &= \big\{\mathbf{C} : \mathbf{C}(\mathbf{FX} : \mathbf{FVF}'\mathbf{Q}_{\mathbf{FX}}) = (\mathbf{X}_* : \mathbf{V}_{21}\mathbf{F}'\mathbf{Q}_{\mathbf{FX}})\big\}\,.\end{aligned}$$

Proceeding along the same lines as Kala et al. (2017, Th. 3) in their BLUE-considerations, we can obtain the following result.

Theorem 2.2. *Suppose that* $\mathbf{Fy}$ *is* BLUP*-sufficient for the predictable* $\mathbf{y}_*$ *under the model* $\mathscr{M}_*$*, and let the sets of matrices* $\mathcal{A}$ *and* $\mathcal{C}$ *be defined as above. Then* $\mathcal{A} = \mathcal{C}$*.*

To describe more statistically the meaning of Theorem 2.2, let $\mathbf{Fy}$ be BLUP-sufficient for $\mathbf{y}_*$ under $\mathscr{M}_*$. Then, for each matrix $\mathbf{C}$ such that $\mathbf{CFy}$ is the BLUP of $\mathbf{y}_*$ in the transformed model $\mathscr{M}_{t*}$, the statistic $\mathbf{CFy}$ is also the BLUP of $\mathbf{y}_*$ in the original model $\mathscr{M}_*$, and vice versa. Notice that in this statement the "vice versa" means that we consider such $\mathbf{C}$ for which $\mathbf{CFy}$ is BLUP under $\mathscr{M}_*$, not the set of matrices $\mathbf{B}$ such that $\mathbf{By}$ is BLUP under $\mathscr{M}_*$. □

3 Some representations for the BLUPs

Let us start by considering the BLUP for $\boldsymbol{\varepsilon}_*$ which offers interesting views. Theorem 3.1 below could be proved directly using Lemma 1.2 by choosing $\boldsymbol{\varepsilon}_*$ as the "new future observations". However, we find it illustrative to give an alternative proof.

Theorem 3.1. *Under the model* $\mathscr{M}_*$*, the statistic* $\mathbf{Cy}$ *is the* BLUP *for* $\boldsymbol{\varepsilon}_*$ *if and only if*

$$\mathbf{C}(\mathbf{X} : \mathbf{VM}) = (\mathbf{0} : \mathbf{V}_{21}\mathbf{M})\,,$$

or, equivalently, $\mathbf{C} = \mathbf{AM}$ *for some matrix* $\mathbf{A}$ *such that*

$$\mathbf{AMVM} = \mathbf{V}_{21}\mathbf{M}\,. \tag{3.1}$$

Proof. The predictor $\mathbf{Cy}$ is unbiased for $\boldsymbol{\varepsilon}_*$ if and only if $\mathrm{E}(\boldsymbol{\varepsilon}_* - \mathbf{Cy}) = \mathbf{0}$ and so $\mathbf{CX} = \mathbf{0}$ and hence necessarily $\mathbf{C} = \mathbf{AM}$ for some matrix $\mathbf{A}$. Now $\mathbf{AMy}$ is the BLUP for $\boldsymbol{\varepsilon}_*$ if $\mathbf{A}$ is such that the covariance matrix of the prediction error $\boldsymbol{\varepsilon}_* - \mathbf{AMy}$ is minimal in the Löwner sense. We recall that for any matrix $\mathbf{A}$, we have the Löwner ordering

$$\mathrm{cov}(\boldsymbol{\varepsilon}_* - \mathbf{AMy}) \geq_{\mathrm{L}} \mathrm{cov}[\boldsymbol{\varepsilon}_* - \mathbf{V}_{21}\mathbf{M}(\mathbf{MVM})^-\mathbf{My}]\,, \tag{3.2}$$

where

$$\mathrm{cov}(\boldsymbol{\varepsilon}_*, \mathbf{My})[\mathrm{cov}(\mathbf{My})]^- = \mathbf{V}_{21}\mathbf{M}(\mathbf{MVM})^-.$$

For the Löwner inequality in (3.2), see Puntanen et al. (2011, Th. 9). We have thus found that $\mathrm{BLUP}(\boldsymbol{\varepsilon}_*)$ has a representation

$$\mathrm{BLUP}(\boldsymbol{\varepsilon}_*) = \mathbf{V}_{21}\mathbf{M}(\mathbf{MVM})^-\mathbf{My}.$$

On the other hand, according to Puntanen et al. (2011, Cor. 9.1), for any matrix $\mathbf{A}$,

$$\mathrm{cov}(\boldsymbol{\varepsilon}_* - \mathbf{AMy}) \geq_{\mathrm{L}} \mathrm{cov}(\boldsymbol{\varepsilon}_* - \mathbf{A}_1\mathbf{My})$$

if and only if $\mathbf{A}_1$ is a solution to (3.1). □

In view of the identity, see Haslett et al. (2014, Sec. 2),

$$\begin{aligned}\mathbf{P}_{\mathbf{X};\mathbf{W}^+} &= \mathbf{X}(\mathbf{X}'\mathbf{W}^-\mathbf{X})^-\mathbf{X}'\mathbf{W}^+ \\ &= \mathbf{P_W} - \mathbf{VM}(\mathbf{MVM})^-\mathbf{MP_W}\,,\end{aligned} \tag{3.3}$$

the $\mathrm{BLUP}(\boldsymbol{\varepsilon}_*)$ can be expressed, for example, as follows:

$$\begin{aligned}\mathrm{BLUP}(\boldsymbol{\varepsilon}_*) &= \mathbf{V}_{21}\mathbf{M}(\mathbf{MVM})^-\mathbf{My} \\ &= \mathbf{V}_{21}\mathbf{W}^-(\mathbf{I}_n - \mathbf{G})\mathbf{y} \\ &= \mathbf{V}_{21}\mathbf{V}^-(\mathbf{I}_n - \mathbf{G})\mathbf{y},\end{aligned}$$

where $\mathbf{W} \in \mathcal{W}$, $\mathbf{y} \in \mathscr{C}(\mathbf{W})$, and $\mathbf{G} = \mathbf{X}(\mathbf{X}'\mathbf{W}^-\mathbf{X})^-\mathbf{X}'\mathbf{W}^- = \mathbf{P}_{\mathbf{X};\mathbf{W}^-}$.

It is well known that the general solution to $\mathbf{A}(\mathbf{X} : \mathbf{VM}) = (\mathbf{X}_* : \mathbf{0})$ can be written, for example, as

$$\mathbf{A}_0 = (\mathbf{X}_* : \mathbf{0})(\mathbf{X} : \mathbf{VM})^+ + \mathbf{N}_1\mathbf{Q_W} := \mathbf{A}_1 + \mathbf{N}_1\mathbf{Q_W}\,,$$

where $\mathbf{N}_1 \in \mathbb{R}^{q\times n}$ is free to vary and $\mathbf{Q_W} = \mathbf{I}_n - \mathbf{P_W}$, $\mathbf{W} \in \mathcal{W}$. Similarly, the general solution to $\mathbf{B}(\mathbf{X} : \mathbf{VM}) = (\mathbf{X}_* : \mathbf{V}_{21}\mathbf{M})$ can be written as

$$\mathbf{B}_0 = (\mathbf{X}_* : \mathbf{V}_{21}\mathbf{M})(\mathbf{X} : \mathbf{VM})^+ + \mathbf{N}_2\mathbf{Q_W} := \mathbf{B}_1 + \mathbf{N}_2\mathbf{Q_W}\,,$$

where the matrix $\mathbf{N}_2 \in \mathbb{R}^{q\times n}$ is free to vary. Consider then the equation

$$\mathbf{C}(\mathbf{X} : \mathbf{VM}) = (\mathbf{0} : \mathbf{V}_{21}\mathbf{M})\,,$$

for which the general solution is

$$\mathbf{C}_0 = (\mathbf{0} : \mathbf{V}_{21}\mathbf{M})(\mathbf{X} : \mathbf{VM})^{+} + \mathbf{N}_3\mathbf{Q}_{\mathbf{W}} := \mathbf{C}_1 + \mathbf{N}_3\mathbf{Q}_{\mathbf{W}}\,,$$

where the matrix $\mathbf{N}_3 \in \mathbb{R}^{q\times n}$ is free to vary. Then $\mathbf{B}_1 = \mathbf{A}_1 + \mathbf{C}_1$ and

$$\mathbf{B}_0 = \mathbf{A}_0 + \mathbf{C}_0 + \mathbf{N}_0\mathbf{Q}_{\mathbf{W}}\,,$$

where $\mathbf{N}_0$ is free to vary. In other words, if

$$\begin{pmatrix}\mathbf{A}\\ \mathbf{C}\end{pmatrix}(\mathbf{X} : \mathbf{VM}) = \begin{pmatrix}\mathbf{X}_* & \mathbf{0}\\ \mathbf{0} & \mathbf{V}_{21}\mathbf{M}\end{pmatrix},$$

then

$$(\mathbf{A}+\mathbf{C})(\mathbf{X} : \mathbf{VM}) = (\mathbf{X}_* : \mathbf{V}_{21}\mathbf{M})\,,$$

and so

$$(\mathbf{A}+\mathbf{C})\mathbf{y} = \mathrm{BLUP}(\mathbf{y}_*)\,.$$

Of course,

$$\mathbf{Ay} = \mathrm{BLUE}(\mathbf{X}_*\boldsymbol{\beta})\,, \quad \mathbf{Cy} = \mathrm{BLUP}(\boldsymbol{\varepsilon}_*)\,,$$

so that we have obtained the following result:

Theorem 3.2. *Under the linear model $\mathscr{M}_*$, where $\mathbf{y}_*$ is predictable, the following decomposition holds (with probability* 1*):*

$$\mathrm{BLUP}(\mathbf{y}_*) = \mathrm{BLUE}(\mathbf{X}_*\boldsymbol{\beta}) + \mathrm{BLUP}(\boldsymbol{\varepsilon}_*)\,.$$

Next we consider the BLUP-sufficiency of $\mathbf{Fy}$ for $\boldsymbol{\varepsilon}_*$.

Theorem 3.3. *The statistic $\mathbf{Fy}$ is* BLUP*-sufficient for $\boldsymbol{\varepsilon}_*$ under $\mathscr{M}_*$ if and only if any of the following equivalent conditions holds:*

(a) $\mathscr{C}\begin{pmatrix}\mathbf{0}\\ \mathbf{MV}_{12}\end{pmatrix} \subset \mathscr{C}\begin{pmatrix}\mathbf{X}'\mathbf{F}'\\ \mathbf{MVF}'\end{pmatrix}$.

(b) $\mathscr{C}(\mathbf{MV}_{12}) \subset \mathscr{C}(\mathbf{MVF}'\mathbf{Q}_{\mathbf{FX}}) = \mathscr{C}(\mathbf{MVMF}'\mathbf{Q}_{\mathbf{FX}})$.

(c) $\mathrm{BLUP}(\boldsymbol{\varepsilon}_* \mid \mathscr{M}_*) = \mathrm{BLUP}(\boldsymbol{\varepsilon}_* \mid \mathscr{M}_{t*})$ *with probability* 1.

(d) *There exists at least one representation of* BLUP *of $\boldsymbol{\varepsilon}_*$ under $\mathscr{M}_*$ which is* BLUP *also under the transformed model $\mathscr{M}_{t*}$.*

In particular, if $\mathbf{Fy}$ is BLUE*-sufficient for $\mathbf{X}\boldsymbol{\beta}$, then* (b) *becomes*

(e) $\mathscr{C}(\mathbf{MV}_{12}) \subset \mathscr{C}(\mathbf{MVF}')$.

Proof. The statistic $\mathbf{Fy}$ is BLUP-sufficient for $\boldsymbol{\varepsilon}_*$ under $\mathscr{M}_*$ if the equation

$$\mathbf{AF}(\mathbf{X} : \mathbf{VM}) = (\mathbf{0} : \mathbf{V}_{21}\mathbf{M}) \tag{3.4}$$

has a solution for $\mathbf{A}$ which obviously happens if and only if (a) holds. The condition (a) means that there exists a matrix $\mathbf{N}$ such that

$$\mathbf{0} = \mathbf{X}'\mathbf{F}'\mathbf{N}, \quad \mathbf{MV}_{12} = \mathbf{MVF}'\mathbf{N}\,.$$

Hence $\mathbf{N} = \mathbf{Q}_{\mathbf{FX}}\mathbf{N}_1$ for some matrix $\mathbf{N}_1$ and

$$\mathbf{MV}_{12} = \mathbf{MVF}'\mathbf{Q}_{\mathbf{FX}}\mathbf{N}_1\,. \tag{3.5}$$

The equality (3.5) holds for some matrix $\mathbf{N}_1$ if and only if

$$\mathscr{C}(\mathbf{MV}_{12}) \subset \mathscr{C}(\mathbf{MVF}'\mathbf{Q}_{\mathbf{FX}}) = \mathscr{C}(\mathbf{MVMF}'\mathbf{Q}_{\mathbf{FX}})\,,$$

where we have used (2.5b).

Suppose that (a) holds so that there exists some matrix $\mathbf{A}$ such that (3.4) holds. Then the same multiplier $\mathbf{AF}$ gives the BLUP for $\boldsymbol{\varepsilon}_*$ under the transformed model $\mathscr{M}_{t*}$ if and only if $\mathbf{A}$ satisfies the equation

$$\mathbf{A}(\mathbf{FX} : \mathbf{FVF}'\mathbf{Q}_{\mathbf{FX}}) = (\mathbf{0} : \mathbf{V}_{21}\mathbf{F}'\mathbf{Q}_{\mathbf{FX}})\,.$$

Proceeding onwards as in the proof of Theorem 2.1, the equivalence between (a), (c) and (d) can be shown.

To prove (e), let us assume that $\mathbf{Fy}$ is BLUE-sufficient for $\mathbf{X}\boldsymbol{\beta}$. It is clear that

$$\mathscr{C}(\mathbf{MVF}'\mathbf{Q}_{\mathbf{FX}}) \subset \mathscr{C}(\mathbf{MVF}')\,. \tag{3.6}$$

Using the rank rule of the matrix product, see Marsaglia & Styan (1974, Cor. 6.2),

$$\begin{aligned} \operatorname{rank}(\mathbf{MVF}'\mathbf{Q}_{\mathbf{FX}}) &= \operatorname{rank}(\mathbf{MVF}') - \dim \mathscr{C}(\mathbf{FVM}) \cap \mathscr{C}(\mathbf{FX}) \\ &= \operatorname{rank}(\mathbf{MVF}')\,, \end{aligned} \tag{3.7}$$

because in view of part (c) of Lemma 2.1, we have $\dim \mathscr{C}(\mathbf{FVM}) \cap \mathscr{C}(\mathbf{FX}) = \{\mathbf{0}\}$. This means that we get equality in (3.6) and so the proof of (e) is completed. □

It is of course clear that the corresponding property as (a) in Theorem 2.1, holds as well for the BLUP($\boldsymbol{\varepsilon}_*$).

Theorem 3.4. *Consider the following three statements:*

(a) $\mathbf{Fy}$ *is* BLUE-*sufficient for* $\mathbf{X}_*\boldsymbol{\beta}$.

(b) $\mathbf{Fy}$ *is* BLUP-*sufficient for* $\boldsymbol{\varepsilon}_*$.

(c) $\mathbf{Fy}$ *is* BLUP-*sufficient for* $\mathbf{y}_*$.

Then above, any two conditions together imply the third one. Moreover, if

$$\mathscr{C}(\mathbf{X}_*) \cap \mathscr{C}(\mathbf{V}_{21}\mathbf{M}) = \{\mathbf{0}\}\,,$$

then

$$\text{(c)} \quad \Longrightarrow \quad \text{(a) } \textit{and } \text{(b)}\,.$$

Proof. Denote

$$\mathbf{A} = \begin{pmatrix} \mathbf{X}_* \\ \mathbf{0} \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} \mathbf{0} \\ \mathbf{MV}_{12} \end{pmatrix} \quad \mathbf{C} = \begin{pmatrix} \mathbf{X}'\mathbf{F}' \\ \mathbf{MVF}' \end{pmatrix}.$$

Now (c) holds if and only if

$$\mathbf{P_C}(\mathbf{A}+\mathbf{B}) = \mathbf{A}+\mathbf{B}\,, \tag{3.8}$$

which is equivalent to

$$\mathbf{P_C A} - \mathbf{A} = -(\mathbf{P_C B} - \mathbf{B})\,,$$

from which the first part of the theorem follows. To prove the second part, we have to show that if

$$\mathscr{C}(\mathbf{A}') \cap \mathscr{C}(\mathbf{B}') = \{\mathbf{0}\}\,, \tag{3.9}$$

then

$$\mathscr{C}(\mathbf{A}+\mathbf{B}) \subset \mathscr{C}(\mathbf{C}) \quad \Longrightarrow \quad \mathscr{C}(\mathbf{A}) \subset \mathscr{C}(\mathbf{C})\,. \tag{3.10}$$

Postmultiplying (3.8) by $\mathbf{Q_{B'}}$ yields

$$\mathbf{P_C A Q_{B'}} = \mathbf{A Q_{B'}}\,. \tag{3.11}$$

If $\operatorname{rank}(\mathbf{AQ_{B'}}) = \operatorname{rank}(\mathbf{A})$, which happens if and only if (3.9) holds, we can, in light of the rank cancellation rule of Marsaglia & Styan (1974, Th. 2), cancel the right-most $\mathbf{Q_{B'}}$ in each side of (3.11) and obtain $\mathbf{P_C A} = \mathbf{A}$ as claimed in (3.10). □

Remark 3.1. The notion of linear error-sufficiency was introduced by Groß (1998), while considering linear sufficient statistics for the prediction of the random error term $\boldsymbol{\varepsilon}$ in the general linear model. This is nothing but the BLUP-sufficiency of $\boldsymbol{\varepsilon}$. Proceeding along the lines of Theorem 3.1, we can conclude that under the model $\mathscr{M}$, the statistic $\mathbf{Cy}$ is the BLUP for $\boldsymbol{\varepsilon}$ if and only if

$$\mathbf{C}(\mathbf{X} : \mathbf{VM}) = (\mathbf{0} : \mathbf{VM})\,,$$

and one explicit solution is

$$\mathrm{BLUP}(\boldsymbol{\varepsilon} \mid \mathscr{M}) = \mathbf{VM}(\mathbf{MVM})^{-}\mathbf{My} = \mathbf{y} - \mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta} \mid \mathscr{M})\,.$$

Obviously $\mathbf{Fy}$ is BLUP-sufficient for $\boldsymbol{\varepsilon}$ if and only if

$$\mathscr{N}(\mathbf{FX} : \mathbf{FVM}) \subset \mathscr{N}(\mathbf{0} : \mathbf{VM})\,.$$

For the BLUP of $\boldsymbol{\varepsilon}$, see also Arendacká & Puntanen (2015, Lemma 1). □

4 Representations for the BLUP in the transformed model

When doing the "BLUP-hunting" in models $\mathscr{M}_*$ and $\mathscr{M}_{t*}$ we assume that the parametric function $\mathbf{X}_*\boldsymbol{\beta}$ is estimable under $\mathscr{M}$ as well as under $\mathscr{M}_t$, which, in light of (2.2), happens if and only if $\mathscr{C}(\mathbf{X}_*') \subset \mathscr{C}(\mathbf{X}'\mathbf{F}')$, so that

$$\mathbf{X}_* = \mathbf{LFX} \quad \text{for some matrix } \mathbf{L}\,. \tag{4.1}$$

Similarly, $\mathbf{X}\boldsymbol{\beta}$ is required to be estimable under $\mathscr{M}_t$ so that $\mathscr{C}(\mathbf{X}') = \mathscr{C}(\mathbf{X}'\mathbf{F}')$.

Denote

$$\begin{aligned}\mathbf{G} &= \mathbf{X}(\mathbf{X}'\mathbf{W}^-\mathbf{X})^-\mathbf{X}'\mathbf{W}^- = \mathbf{P}_{\mathbf{X};\mathbf{W}^-}\,,\\ \mathbf{P}_{\mathbf{FX};(\mathbf{FWF}')^-} &= \mathbf{FX}[\mathbf{X}'\mathbf{F}'(\mathbf{FWF}')^-\mathbf{FX}]^-\mathbf{X}'\mathbf{F}'(\mathbf{FWF}')^-,\\ \mathbf{G}_t &= \mathbf{X}[\mathbf{X}'\mathbf{F}'(\mathbf{FWF}')^-\mathbf{FX}]^-\mathbf{X}'\mathbf{F}'(\mathbf{FWF}')^-\mathbf{F}\,,\end{aligned}$$

so that $\mathbf{FG}_t = \mathbf{P}_{\mathbf{FX};(\mathbf{FWF}')^-}\mathbf{F}$.

Estimator $\mathbf{BFy}$ is the $\mathrm{BLUE}(\mathbf{FX}\boldsymbol{\beta} \mid \mathscr{M}_t)$ if and only if $\mathbf{B}$ satisfies

$$\mathbf{B}(\mathbf{FX} : \mathbf{FVF}'\mathbf{Q}_{\mathbf{FX}}) = (\mathbf{FX} : \mathbf{0})\,,$$

so that one expression for $\mathbf{B}$ is $\mathbf{B} = \mathbf{P}_{\mathbf{FX};(\mathbf{FWF}')^-} := \mathbf{FXA}$ and then

$$\mathbf{FXA}(\mathbf{FX} : \mathbf{FVF}'\mathbf{Q}_{\mathbf{FX}}) = (\mathbf{FX} : \mathbf{0})\,. \tag{4.2}$$

Because $\mathrm{rank}(\mathbf{FX}) = \mathrm{rank}(\mathbf{X})$, we can cancel the left-most $\mathbf{F}$ from both sides of (4.2) resulting

$$\mathbf{X}[\mathbf{X}'\mathbf{F}'(\mathbf{FWF}')^-\mathbf{FX}]^-\mathbf{X}'\mathbf{F}'(\mathbf{FWF}')^-(\mathbf{FX} : \mathbf{FVF}'\mathbf{Q}_{\mathbf{FX}}) = (\mathbf{X} : \mathbf{0})\,.$$

Thus $\mathbf{G}_t\mathbf{y}$ is the BLUE for $\mathbf{X}\boldsymbol{\beta}$ under $\mathscr{M}_t$ and

$$\mathbf{G}_t(\mathbf{X} : \mathbf{VF}'\mathbf{Q}_{\mathbf{FX}}) = (\mathbf{X} : \mathbf{0})\,. \tag{4.3}$$

An alternative expression for $\mathrm{BLUE}(\mathbf{FX}\boldsymbol{\beta} \mid \mathscr{M}_t)$ can be obtained using the corresponding identity as in (3.3):

$$\begin{aligned}\mathbf{P}_{\mathbf{FX};(\mathbf{FWF}')^+} &= \mathbf{FX}[\mathbf{X}'\mathbf{F}'(\mathbf{FWF}')^-\mathbf{FX}]^-\mathbf{X}'\mathbf{F}'(\mathbf{FWF}')^+\\ &= \mathbf{P}_{\mathbf{FW}} - \mathbf{FVF}'\mathbf{Q}_{\mathbf{FX}}(\mathbf{Q}_{\mathbf{FX}}\mathbf{FVF}'\mathbf{Q}_{\mathbf{FX}})^-\mathbf{Q}_{\mathbf{FX}}\mathbf{P}_{\mathbf{FW}}\,.\end{aligned}$$

Namely, for $\mathbf{y} \in \mathscr{C}(\mathbf{W})$ and, noting that $\mathbf{P}_{\mathbf{FW}}\mathbf{Fy} = \mathbf{Fy}$, we get

$$\begin{aligned}\mathrm{BLUE}(\mathbf{FX}\boldsymbol{\beta} \mid \mathscr{M}_t) &= \mathbf{FG}_t\mathbf{y}\\ &= \mathbf{P}_{\mathbf{FX};(\mathbf{FWF}')^+}\mathbf{Fy}\\ &= \mathbf{Fy} - \mathbf{FVF}'\mathbf{Q}_{\mathbf{FX}}(\mathbf{Q}_{\mathbf{FX}}\mathbf{FVF}'\mathbf{Q}_{\mathbf{FX}})^-\mathbf{Q}_{\mathbf{FX}}\mathbf{Fy}.\end{aligned}$$

It is interesting to observe that the matrix

$$\mathbf{G}_{\#} = \mathbf{I}_n - \mathbf{VF}'\mathbf{Q}_{\mathbf{FX}}(\mathbf{Q}_{\mathbf{FX}}\mathbf{FVF}'\mathbf{Q}_{\mathbf{FX}})^{-}\mathbf{Q}_{\mathbf{FX}}\mathbf{F}$$

satisfies (4.3), i.e.,

$$\mathbf{G}_{\#}(\mathbf{X} : \mathbf{VF}'\mathbf{Q}_{\mathbf{FX}}) = (\mathbf{X} : \mathbf{0})\,.$$

However, $\mathbf{G}_{\#}$ and $\mathbf{G}_t$ are not necessarily equal; their difference is

$$\mathbf{G}_t - \mathbf{G}_{\#} = \mathbf{NQ}_{(\mathbf{X}:\mathbf{VF}'\mathbf{Q}_{\mathbf{FX}})}$$

for some matrix $\mathbf{N}$.

Consider then the expressions for the BLUP of $\boldsymbol{\varepsilon}_*$ under the transformed model $\mathscr{M}_{t*}$. One way to do this is to use Theorem 3.1, which says that $\mathbf{DFy}$ is the $\mathrm{BLUP}(\boldsymbol{\varepsilon}_* \mid \mathscr{M}_{t*})$ if $\mathbf{D}$ is a solution to

$$\mathbf{D}(\mathbf{FX} : \mathbf{FVF}'\mathbf{Q}_{\mathbf{FX}}) = (\mathbf{0} : \mathbf{V}_{21}\mathbf{F}'\mathbf{Q}_{\mathbf{FX}})\,.$$

Thus the BLUP of $\boldsymbol{\varepsilon}_*$ under $\mathscr{M}_{t*}$ can be expressed as

$$\mathrm{BLUP}(\boldsymbol{\varepsilon}_* \mid \mathscr{M}_{t*}) = \mathbf{V}_{21}\mathbf{F}'\mathbf{Q}_{\mathbf{FX}}(\mathbf{Q}_{\mathbf{FX}}\mathbf{FVF}'\mathbf{Q}_{\mathbf{FX}})^{-}\mathbf{Q}_{\mathbf{FX}}\mathbf{Fy}. \tag{4.4}$$

Recall that in (4.4), $\mathbf{F}'\mathbf{Q}_{\mathbf{FX}}$ can be replaced with $\mathbf{MF}'\mathbf{Q}_{\mathbf{FX}}$. One alternative expression is

$$\mathrm{BLUP}(\boldsymbol{\varepsilon}_* \mid \mathscr{M}_{t*}) = \mathbf{V}_{21}\mathbf{F}'(\mathbf{FVF}')^{-}\mathbf{F}(\mathbf{I}_n - \mathbf{G}_t)\mathbf{y}.$$

We complete this section by giving some alternative expressions for the BLUP of $\mathbf{y}_*$. Using (4.1), let us denote

$$\boldsymbol{\mu}_* = \mathbf{X}_*\boldsymbol{\beta} = \mathbf{LFX}\boldsymbol{\beta}, \quad \boldsymbol{\mu} = \mathbf{X}\boldsymbol{\beta}\,.$$

The $\mathrm{BLUP}(\mathbf{y}_*)$ under $\mathscr{M}_*$ can be written as

$$\begin{aligned}
\mathrm{BLUP}(\mathbf{y}_* \mid \mathscr{M}_*) &= \mathrm{BLUE}(\boldsymbol{\mu}_* \mid \mathscr{M}) + \mathbf{V}_{21}\mathbf{V}^{-}[\mathbf{y} - \mathrm{BLUE}(\boldsymbol{\mu} \mid \mathscr{M})] \\
&= \mathbf{LFGy} + \mathbf{V}_{21}\mathbf{V}^{-}(\mathbf{I}_n - \mathbf{G})\mathbf{y} \\
&= \mathbf{LFGy} + \mathbf{V}_{21}\mathbf{M}(\mathbf{MVM})^{-}\mathbf{My} \\
&= \mathrm{BLUE}(\boldsymbol{\mu}_* \mid \mathscr{M}) + \mathrm{BLUP}(\boldsymbol{\varepsilon}_* \mid \mathscr{M}_*)\,,
\end{aligned} \tag{4.5}$$

or shortly,

$$\tilde{\mathbf{y}}_* = \tilde{\boldsymbol{\mu}}_* + \tilde{\boldsymbol{\varepsilon}}_*\,.$$

Under the transformed model we have

$$\begin{aligned}
\mathrm{BLUP}(\mathbf{y}_* \mid \mathscr{M}_{t*}) &= \mathrm{BLUE}(\boldsymbol{\mu}_* \mid \mathscr{M}_t) + \mathbf{V}_{21}\mathbf{F}'(\mathbf{FVF}')^{-}\mathbf{F}[\mathbf{y} - \mathrm{BLUE}(\boldsymbol{\mu} \mid \mathscr{M}_t)] \\
&= \mathbf{LFG}_t\mathbf{y} + \mathbf{V}_{21}\mathbf{F}'(\mathbf{FVF}')^{-}\mathbf{F}(\mathbf{I}_n - \mathbf{G}_t)\mathbf{y} \\
&= \mathbf{LFG}_t\mathbf{y} + \mathbf{V}_{21}\mathbf{F}'\mathbf{Q}_{\mathbf{FX}}(\mathbf{Q}_{\mathbf{FX}}\mathbf{FVF}'\mathbf{Q}_{\mathbf{FX}})^{-}\mathbf{Q}_{\mathbf{FX}}\mathbf{Fy} \\
&= \mathrm{BLUE}(\boldsymbol{\mu}_* \mid \mathscr{M}_t) + \mathrm{BLUP}(\boldsymbol{\varepsilon}_* \mid \mathscr{M}_{t*})\,,
\end{aligned} \tag{4.6}$$

or shortly,

$$\tilde{\mathbf{y}}_{t*} = \tilde{\boldsymbol{\mu}}_{t*} + \tilde{\varepsilon}_{t*}\,. \tag{4.7}$$

In (4.5) and (4.6) the matrix $\mathbf{V}$ can be replaced with $\mathbf{W} \in \mathcal{W}$. In passing we may notice that under $\mathscr{M}_*$, $\tilde{\boldsymbol{\mu}}_*$ and $\tilde{\boldsymbol{\varepsilon}}_*$ are uncorrelated and hence $\operatorname{cov}(\tilde{\mathbf{y}}_*) = \operatorname{cov}(\tilde{\boldsymbol{\mu}}_*) + \operatorname{cov}(\tilde{\boldsymbol{\varepsilon}}_*)$. The corresponding property holds also for the terms of (4.7). For further representations for the $\mathrm{BLUP}(\mathbf{y}_* \mid \mathscr{M}_*)$ we refer to Haslett et al. (2014).

5 Linear mixed model

One application of the model $\mathscr{M}_*$ is the linear mixed model

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \boldsymbol{\varepsilon}, \quad \text{or shortly,} \quad \mathscr{L} = \{\mathbf{y}, \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u}, \mathbf{D}, \mathbf{R}, \mathbf{S}\},$$

where $\mathbf{X}_{n\times p}$ and $\mathbf{Z}_{n\times q}$ are known matrices, $\boldsymbol{\beta} \in \mathbb{R}^p$ is a vector of unknown fixed effects, $\mathbf{u}$ is an unobservable vector (q elements) of random effects with $\mathrm{E}(\mathbf{u}) = \mathbf{0}$, $\operatorname{cov}(\mathbf{u}) = \mathbf{D}_{q\times q}$, $\operatorname{cov}(\boldsymbol{\varepsilon}, \mathbf{u}) = \mathbf{S}_{n\times q}$, and $\mathrm{E}(\boldsymbol{\varepsilon}) = \mathbf{0}$, $\operatorname{cov}(\boldsymbol{\varepsilon}) = \mathbf{R}_{n\times n}$. In this situation

$$\operatorname{cov}\begin{pmatrix} \boldsymbol{\varepsilon} \\ \mathbf{u} \end{pmatrix} = \begin{pmatrix} \mathbf{R} & \mathbf{S} \\ \mathbf{S}' & \mathbf{D} \end{pmatrix},$$

and

$$\operatorname{cov}(\mathbf{y}) = \mathbf{Z}\mathbf{D}\mathbf{Z}' + \mathbf{R} + \mathbf{Z}\mathbf{S}' + \mathbf{S}\mathbf{Z}' := \boldsymbol{\Sigma}.$$

The mixed model can be expressed as a version of the model with “new observations”, the new observations being now in $\mathbf{u}$:

$$\left\{ \begin{pmatrix} \mathbf{y} \\ \mathbf{u} \end{pmatrix}, \begin{pmatrix} \mathbf{X} \\ \mathbf{0} \end{pmatrix} \boldsymbol{\beta}, \begin{pmatrix} \boldsymbol{\Sigma} & \mathbf{Z}\mathbf{D} + \mathbf{S} \\ \mathbf{D}\mathbf{Z}' + \mathbf{S}' & \mathbf{D} \end{pmatrix} \right\}.$$

Moreover, choosing the “new observations” as $\mathbf{g} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u}$, we get

$$\left\{ \begin{pmatrix} \mathbf{y} \\ \mathbf{g} \end{pmatrix}, \begin{pmatrix} \mathbf{X} \\ \mathbf{X} \end{pmatrix} \boldsymbol{\beta}, \begin{pmatrix} \boldsymbol{\Sigma} & (\mathbf{Z}\mathbf{D} + \mathbf{S})\mathbf{Z}' \\ \mathbf{Z}(\mathbf{D}\mathbf{Z}' + \mathbf{S}') & \mathbf{Z}\mathbf{D}\mathbf{Z}' \end{pmatrix} \right\}.$$

Thus, see, e.g., Haslett et al. (2015), under the mixed model $\mathscr{L}$ the following statements hold:

(a) $\mathbf{A}\mathbf{y}$ is the BLUE for $\mathbf{X}\boldsymbol{\beta}$ if and only if

$$\mathbf{A}(\mathbf{X} : \boldsymbol{\Sigma}\mathbf{M}) = (\mathbf{X} : \mathbf{0})\,. \tag{5.1}$$

(b) $\mathbf{B}\mathbf{y}$ is the BLUP for $\mathbf{u}$ if and only if

$$\mathbf{B}(\mathbf{X} : \boldsymbol{\Sigma}\mathbf{M}) = \big[\mathbf{0} : (\mathbf{D}\mathbf{Z}' + \mathbf{S}')\mathbf{M}\big] = \big[\mathbf{0} : \operatorname{cov}(\mathbf{u}, \mathbf{y})\mathbf{M}\big]\,.$$

(c) $\mathbf{Cy}$ is the BLUP for $\mathbf{g} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Zu}$ if and only if

$$\mathbf{C}(\mathbf{X} : \boldsymbol{\Sigma}\mathbf{M}) = \big[\mathbf{X} : \mathbf{Z}(\mathbf{DZ}' + \mathbf{S}')\mathbf{M}\big] = \big[\mathbf{X} : \operatorname{cov}(\mathbf{g}, \mathbf{y})\mathbf{M}\big] . \tag{5.2}$$

Thus we have, corresponding to Theorem 3.2,

$$\mathrm{BLUP}(\mathbf{X}\boldsymbol{\beta} + \mathbf{Zu} \mid \mathscr{L}) = \mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta} \mid \mathscr{L}) + \mathrm{BLUP}(\mathbf{Zu} \mid \mathscr{L}) ,$$

so that one representation for the BLUP of $\mathbf{g}$ under $\mathscr{L}$ is

$$\begin{aligned}\mathrm{BLUP}(\mathbf{g}) &= \mathbf{Ty} + \mathbf{Z}(\mathbf{DZ}' + \mathbf{S}')\mathbf{W}_{\boldsymbol{\Sigma}}^{-}(\mathbf{y} - \mathbf{Ty}) \\ &= \mathbf{Ty} + \mathbf{Z}(\mathbf{DZ}' + \mathbf{S}')\mathbf{M}(\mathbf{M}\boldsymbol{\Sigma}\mathbf{M})^{-}\mathbf{My},\end{aligned}$$

where $\mathbf{T} = \mathbf{X}(\mathbf{X}'\mathbf{W}_{\boldsymbol{\Sigma}}^{-}\mathbf{X})^{-}\mathbf{X}'\mathbf{W}_{\boldsymbol{\Sigma}}^{-}$ and

$$\mathbf{W}_{\boldsymbol{\Sigma}} = \boldsymbol{\Sigma} + \mathbf{XUU}'\mathbf{X}', \quad \mathscr{C}(\mathbf{W}_{\boldsymbol{\Sigma}}) = \mathscr{C}(\mathbf{X} : \boldsymbol{\Sigma}) .$$

Conditions for $\mathbf{Fy}$ being linearly sufficient or linearly prediction sufficient for $\mathbf{X}\boldsymbol{\beta}$, $\mathbf{u}$, and $\mathbf{g} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Zu}$, respectively, can be straightforwardly derived from (5.1)–(5.2). For example, $\mathbf{Fy}$ is BLUP-sufficient for $\mathbf{g}$ if and only if

$$\mathscr{C}\begin{pmatrix} \mathbf{X}' \\ \mathbf{M}(\mathbf{ZD} + \mathbf{S})\mathbf{Z}' \end{pmatrix} \subset \mathscr{C}\begin{pmatrix} \mathbf{X}'\mathbf{F}' \\ \mathbf{M}\boldsymbol{\Sigma}\mathbf{F}' \end{pmatrix} . \tag{5.3}$$

Corresponding properties as under $\mathscr{M}_*$ in Theorem 3.4 for $\mathbf{X}_*\boldsymbol{\beta}$, $\boldsymbol{\varepsilon}_*$, and $\mathbf{y}_*$ hold also under $\mathscr{L}$ for $\mathbf{X}\boldsymbol{\beta}$, $\mathbf{Zu}$, and $\mathbf{g}$.

For the linear sufficiency in the mixed model, see also Liu et al. (2008, Sec. 3). They defined the BLUP-sufficiency in a slightly different manner which we will not handle here. Inspired by their Theorem 3.1, we will now show that $\mathbf{Fy}$ is BLUP-sufficient for $\mathbf{g} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Zu}$ if

$$\mathscr{C}(\mathbf{X} : \mathbf{ZD} + \mathbf{S}) \subset \mathscr{C}(\mathbf{W}_{\boldsymbol{\Sigma}}\mathbf{F}') . \tag{5.4}$$

In view of part (b) of Lemma 2.1, the "first part" of (5.4), $\mathscr{C}(\mathbf{X}) \subset \mathscr{C}(\mathbf{W}_{\boldsymbol{\Sigma}}\mathbf{F}')$, is equivalent to

$$\mathscr{C}\begin{pmatrix} \mathbf{X}' \\ \mathbf{0} \end{pmatrix} \subset \mathscr{C}\begin{pmatrix} \mathbf{X}'\mathbf{F}' \\ \mathbf{M}\boldsymbol{\Sigma}\mathbf{F}' \end{pmatrix} , \tag{5.5}$$

which means that $\mathbf{Fy} \in \mathcal{S}(\mathbf{X}\boldsymbol{\beta})$. If $\mathbf{Fy}$ would be also BLUP-sufficient for $\mathbf{Zu}$, that is,

$$\mathscr{C}\begin{pmatrix} \mathbf{0} \\ \mathbf{M}(\mathbf{ZD} + \mathbf{S})\mathbf{Z}' \end{pmatrix} \subset \mathscr{C}\begin{pmatrix} \mathbf{X}'\mathbf{F}' \\ \mathbf{M}\boldsymbol{\Sigma}\mathbf{F}' \end{pmatrix} , \tag{5.6}$$

then (5.3) would hold. Now (5.6) can be equivalently expressed as

$$\mathscr{C}[\mathbf{M}(\mathbf{ZD} + \mathbf{S})\mathbf{Z}'] \subset \mathscr{C}(\mathbf{M}\boldsymbol{\Sigma}\mathbf{F}'\mathbf{Q}_{\mathbf{FX}}) = \mathscr{C}(\mathbf{M}\boldsymbol{\Sigma}\mathbf{F}') = \mathscr{C}(\mathbf{M}\mathbf{W}_{\boldsymbol{\Sigma}}\mathbf{F}') , \tag{5.7}$$

where the equality follows from (3.7). Premultiplying (5.4) by $\mathbf{M}$ gives (5.7) at once. Thus we have proved that (5.4) implies (5.3). Notice that in the light of the second part of Theorem 3.4 the implication (5.3) $\implies$ (5.5) holds in the situation when

$$\mathscr{C}(\mathbf{X}) \cap \mathscr{C}[\mathbf{Z}(\mathbf{D}\mathbf{Z}' + \mathbf{S}')\mathbf{M}] = \{\mathbf{0}\}\,.$$

There is one further interesting link connecting the mixed model and the following extended partitioned model:

$$\begin{aligned}\mathscr{A} &= \{\dot{\mathbf{y}},\, \dot{\mathbf{X}}\boldsymbol{\pi},\, \dot{\mathbf{V}}\} \\ &= \left\{ \begin{pmatrix} \mathbf{y} \\ \mathbf{y}_0 \end{pmatrix}, \begin{pmatrix} \mathbf{X} & \mathbf{Z} \\ \mathbf{0} & -\mathbf{I}_q \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta} \\ \boldsymbol{\gamma} \end{pmatrix}, \begin{pmatrix} \mathbf{R} & \mathbf{S} \\ \mathbf{S}' & \mathbf{D} \end{pmatrix} \right\},\end{aligned}$$

where both $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are *fixed* effects parameters. Expressed in error terms we have

$$\begin{aligned}\mathbf{y} &= \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\varepsilon}\,, \\ \mathbf{y}_0 &= \qquad\quad -\boldsymbol{\gamma} + \boldsymbol{\varepsilon}_0\,,\end{aligned}$$

where $\operatorname{cov}\binom{\mathbf{y}}{\mathbf{y}_0} = \operatorname{cov}\binom{\boldsymbol{\varepsilon}}{\boldsymbol{\varepsilon}_0} = \dot{\mathbf{V}}$. Premultiplying the model $\mathscr{A}$ by the matrix

$$\mathbf{T} = (\mathbf{I}_n : \mathbf{Z})\,,$$

as in Arendacká & Puntanen (2015, Sec. 2), yields the equation

$$\mathbf{y} + \mathbf{Z}\mathbf{y}_0 = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\varepsilon}_0 + \boldsymbol{\varepsilon}\,. \tag{5.8}$$

We see that (5.8) defines a mixed model, say $\mathscr{B}$, where the observable response is $\mathbf{w} = \mathbf{y} + \mathbf{Z}\mathbf{y}_0$ and $\boldsymbol{\varepsilon}_0$ is the unobservable random effect, and

$$\operatorname{cov}(\mathbf{y} + \mathbf{Z}\mathbf{y}_0) = \operatorname{cov}(\mathbf{Z}\boldsymbol{\varepsilon}_0 + \boldsymbol{\varepsilon}) = \mathbf{Z}\mathbf{D}\mathbf{Z}' + \mathbf{R} + \mathbf{Z}\mathbf{S}' + \mathbf{S}\mathbf{Z}' = \boldsymbol{\Sigma}\,.$$

We can denote the resulting mixed model as

$$\mathscr{B} = \{\mathbf{y} + \mathbf{Z}\mathbf{y}_0,\, \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\varepsilon}_0,\, \mathbf{D},\, \mathbf{R},\, \mathbf{S}\}\,.$$

We can also interpret $\mathscr{B}$ as a fixed effect model and write it as $\mathscr{B} = \{\mathbf{w}, \mathbf{X}\boldsymbol{\beta}, \boldsymbol{\Sigma}\}$, where the random effect is not written up explicitly.

It is now interesting to know whether the BLUEs of $\mathbf{X}\boldsymbol{\beta}$ under $\mathscr{A}$ and $\mathscr{B}$ are equal. We answer to this question using the linear sufficiency concept while Haslett et al. (2015) and Arendacká & Puntanen (2015) solved this problem using different approach. To do this, we write $\mathscr{A}$ as

$$\mathscr{A} = \{\dot{\mathbf{y}},\, \dot{\mathbf{X}}\boldsymbol{\pi},\, \dot{\mathbf{V}}\} = \{\dot{\mathbf{y}},\, \dot{\mathbf{X}}_1\boldsymbol{\beta} + \dot{\mathbf{X}}_2\boldsymbol{\gamma},\, \dot{\mathbf{V}}\}\,.$$

First we notice that $\dot{\mathbf{X}}_1\boldsymbol{\beta}$ (and thereby $\mathbf{X}\boldsymbol{\beta}$) is estimable because $\mathscr{C}(\dot{\mathbf{X}}_1)$ and $\mathscr{C}(\dot{\mathbf{X}}_2)$ are disjoint. Then we observe that

$$\mathbf{T}' = \begin{pmatrix} \mathbf{I}_n \\ \mathbf{Z}' \end{pmatrix} \in \left\{ \begin{pmatrix} \mathbf{Z} \\ -\mathbf{I}_q \end{pmatrix}^{\perp} \right\} = \{\dot{\mathbf{X}}_2^{\perp}\},$$

i.e., $\mathbf{T}\dot{\mathbf{X}}_2 = \mathbf{0}$ and $\operatorname{rank}(\mathbf{T}) = \operatorname{rank}(\dot{\mathbf{X}}_2^{\perp})$. It is well known by Frisch–Waugh–Lowell Theorem that premultiplying $\mathscr{A}$ by orthogonal projector $\dot{\mathbf{M}}_2 = \mathbf{I}_{n+q} - \mathbf{P}_{\dot{\mathbf{X}}_2}$ yields the reduced model under which the BLUE of $\dot{\mathbf{X}}_1\boldsymbol{\beta}$ is the same as in $\mathscr{A}$, that is, $\dot{\mathbf{M}}_2\dot{\mathbf{y}}$ is linearly sufficient for $\dot{\mathbf{X}}_1\boldsymbol{\beta}$. Now $\mathscr{C}(\mathbf{T}') = \mathscr{C}(\dot{\mathbf{M}}_2)$ and hence, in view of (2.1), $\mathbf{T}\dot{\mathbf{y}}$ is also linearly sufficient for $\dot{\mathbf{X}}_1\boldsymbol{\beta}$ and thereby

$$\operatorname{BLUE}(\mathbf{X}\boldsymbol{\beta} \mid \mathscr{A}) = \operatorname{BLUE}(\mathbf{X}\boldsymbol{\beta} \mid \mathscr{B})\,.$$

For the linear sufficiency in the partitioned model, see also Kala et al. (2017, Sec. 5).

Haslett et al. (2015) and Arendacká & Puntanen (2015) also showed the following:

$$\operatorname{BLUP}(\boldsymbol{\varepsilon}_0 \mid \mathscr{A}) = \operatorname{BLUP}(\boldsymbol{\varepsilon}_0 \mid \mathscr{B}) = \operatorname{BLUE}(\boldsymbol{\gamma} \mid \mathscr{A}) + \mathbf{y}_0\,.$$

The connection between the models $\mathscr{A}$ and $\mathscr{B}$ can be used as a tool to calculate the BLUEs and BLUPs in mixed model and it is often referred to as a Henderson's method; see, e.g., Henderson et al. (1959) and McCulloch et al. (2008, Ch. 8).

Acknowledgements

Thanks go to Professor Xu-Qing Liu and the anonymous referees for helpful comments. Part of this research was done during the meeting of an International Research Group on Multivariate and Mixed Linear Models in the Mathematical Research and Conference Center, Będlewo, Poland, November 2016, supported by the Stefan Banach International Mathematical Center.

References

Arendacká, B., Puntanen, S. (2015). Further remarks on the connection between fixed linear model and mixed linear model. Stat. Pap. **56**, 1235–1247. DOI

Baksalary, J.K., Kala, R. (1981). Linear transformations preserving best linear unbiased estimators in a general Gauss–Markoff model. Ann. Stat. **9**, 913–916. DOI

Baksalary, J.K., Kala, R. (1986). Linear sufficiency with respect to a given vector of parametric functions. J. Stat. Plan. Inf. **14**, 331–338. DOI

Baksalary, O.M., Trenkler, G. (2009). A projector oriented approach to the best linear unbiased estimator. Stat. Pap. **50**, 721–733. DOI

Christensen, R. (2011). Plane Answers to Complex Questions: the Theory of Linear Models, 4th Edition. Springer, New York.

Drygas, H. (1970). The Coordinate-Free Approach to Gauss–Markov Estimation. Springer, Berlin.

Drygas, H. (1983). Sufficiency and completeness in the general Gauss–Markov model. Sankhyā Ser. A **45**, 88–98.

Groß, J. (1998). A note on the concepts of linear and quadratic sufficiency. J. Stat. Plan. Inf. **70**, 88–98.

Haslett, S.J., Isotalo, J., Liu, Y., Puntanen, S. (2014). Equalities between OLSE, BLUE and BLUP in the linear model. Stat. Pap. **55**, 543–561. DOI

Haslett, S.J., Puntanen, S., Arendacká, B. (2015). The link between the mixed and fixed linear models revisited. Stat. Pap. **56**, 849–861. DOI

Henderson, C.R., Kempthorne, O., Searle, S.R., von Krosigh, C.N. (1959). The estimation of environmental and genetic trends from records subject to culling. Biometrics **15**, 192–218.

Isotalo, J., Puntanen, S. (2006). Linear prediction sufficiency for new observations in the general Gauss–Markov model. Commun. Stat. Theory Methods **35**, 1011–1023. DOI

Kala, R., Markiewicz, A., Puntanen, S. (2017). Some further remarks on the linear sufficiency in the linear model. *Applied and Computational Matrix Analysis: MatTriad, Coimbra, Portugal, September 2015, Selected, Revised Contributions* (Natália Bebiano, editor), Springer Proceedings in Mathematics & Statistics, vol. 192, 275–294. DOI

Kala, R., Puntanen, S., Tian, Y. (2017). Some notes on linear sufficiency. Stat. Pap., **58**, 1–17. DOI

Liu, X.-Q., Rong, J.-Y., Liu, J.-Y. (2008). Best linear unbiased prediction for linear combinations in general mixed linear models. J. Multivar. Anal. **99**, 1503–1517. DOI

Markiewicz, A., Puntanen, S. (2017). Further properties of the linear sufficiency in the partitioned linear model. Matrices, Statistics and Big Data: Proceedings of the 25th International Workshop on Matrices and Statistics, IWMS-2016, held in Funchal, Madeira. Portugal, 6–9 June 2016. Springer, in press.

Marsaglia, G., Styan, G.P.H. (1974). Equalities and inequalities for ranks of matrices. Linear and Multilinear Algebra **2**, 269–292. DOI

McCulloch, C.E., Searle, S.R., Neuhaus, J.M. (2008). Generalized, Linear, and Mixed Models. 2nd Edition. Wiley, New York.

Puntanen, S., Styan, G.P.H., Isotalo, J. (2011). Matrix Tricks for Linear Statistical Models: Our Personal Top Twenty. Springer, Heidelberg. DOI

Rao, C.R. (1973). Representations of best linear estimators in the Gauss–Markoff model with a singular dispersion matrix. J. Multivariate Anal. **3**, 276–292. DOI

Rao, C.R., Mitra, S.K. (1971). Generalized Inverse of Matrices and Its Applications. Wiley, New York.

Tian, Y., Puntanen, S. (2009). On the equivalence of estimations under a general linear model and its transformed models. Linear Algebra Appl. **430**, 2622–2641. DOI