Statistics and Applications {ISSN 2454-7395 (online)}
Special Issue in Memory of Prof. C R Rao
Volume 22, No. 3, 2024 (New Series), pp 95–118
http://www.ssca.org.in/journal

The Fundamental BLUE Equation in Linear Models Revisited

Stephen J. Haslett[1], Jarkko Isotalo[2], Augustyn Markiewicz[3] and Simo Puntanen[2]

[1] *School of Mathematical and Computational Sciences and Environmental Health Intelligence NZ, Massey University, Palmerston North, New Zealand; Research School of Finance, Actuarial Studies and Statistics, The Australian National University, Canberra, Australia; Faculty of Engineering and Information Sciences, University of Wollongong, Australia*

[2] *Faculty of Information Technology and Communication Sciences, Tampere University, FI-33014 Tampere, Finland*

[3] *Department of Mathematical and Statistical Methods, Poznań University of Life Sciences, Wojska Polskiego 28, PL-60637 Poznań, Poland*

Received: 22 April 2024; Revised: 07 June 2024; Accepted: 09 June 2024

Abstract

In the world of linear statistical models there is a particular matrix equation, $\mathbf{G}(\mathbf{X} : \mathbf{V}\mathbf{X}^{\perp}) = (\mathbf{X} : \mathbf{0})$, which is sufficiently important that it is sometimes called the fundamental BLUE equation. In this equation, $\mathbf{X}$ is a model matrix, $\mathbf{V}$ is the covariance matrix of $\mathbf{y}$ in the linear model $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, and we are interested in finding the best linear estimator, BLUE, of $\mathbf{X}\boldsymbol{\beta}$. Any solution $\mathbf{G}$ for this equation has the property that $\mathbf{G}\mathbf{y}$ provides a representation for the BLUE of $\mathbf{X}\boldsymbol{\beta}$: this is the message of the the fundamental BLUE equation, whose main developer was the late Professor C. R. Rao in early 1970s. In this article we revisit some interesting features and consequences of this equation. We do not provide essentially new results – the aim is to offer a compact easy-to-follow review including also some recent related results by the authors.

Key words: BLUE; BLUP; Covariance matrix; Equality of the BLUEs; Linear sufficiency; Misspecified model.

AMS Subject Classifications: 62J05, 62J10

1. Introduction

Our main focus in this paper is the linear model $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, denoted as triplet

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

Here $\mathbf{y}$ is an n-dimensional observable random vector, and $\boldsymbol{\varepsilon}$ is an unobservable random error vector with a known (possibly singular) covariance matrix $\mathrm{cov}(\boldsymbol{\varepsilon}) = \mathbf{V} = \mathrm{cov}(\mathbf{y})$ and

Corresponding Author: Simo Puntanen
Email: simo.puntanen@tuni.fi

expectation $\mathrm{E}(\boldsymbol{\varepsilon}) = \mathbf{0}$. The matrix $\mathbf{X}$ is a known $n \times p$ matrix, *i.e.*, $\mathbf{X} \in \mathbb{R}^{n\times p}$. Vector $\boldsymbol{\beta}$ is a vector of fixed (but unknown) parameters; here symbol $'$ stands for the transpose. We will also denote $\boldsymbol{\mu} = \mathbf{X}\boldsymbol{\beta}$. If we want to emphasize what is the covariance matrix, we may use notation $\mathscr{M}(\mathbf{V})$.

As for notations, the symbols $\mathrm{r}(\mathbf{A})$, $\mathbf{A}^-$, $\mathbf{A}^+$, $\mathscr{C}(\mathbf{A})$, and $\mathscr{C}(\mathbf{A})^\perp$, denote, respectively, the rank, 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 columnwise partitioned matrix with $\mathbf{A}_{a\times b}$ and $\mathbf{B}_{a\times c}$ as submatrices. By $\mathbf{A}^\perp$ we denote any matrix satisfying $\mathscr{C}(\mathbf{A}^\perp) = \mathscr{C}(\mathbf{A})^\perp$. We will write $\mathbf{P}_\mathbf{A} = \mathbf{A}\mathbf{A}^+ = \mathbf{A}(\mathbf{A}'\mathbf{A})^-\mathbf{A}'$ to denote the orthogonal projector onto $\mathscr{C}(\mathbf{A})$ and $\mathbf{Q}_\mathbf{A} = \mathbf{I}_a - \mathbf{P}_\mathbf{A}$, where $\mathbf{I}_a$ is the identity matrix of order a with a being the number of rows in $\mathbf{A}$. In particular, we denote

$$\mathbf{H} = \mathbf{P}_\mathbf{X}\,, \quad \mathbf{M} = \mathbf{I}_n - \mathbf{P}_\mathbf{X}\,, \quad \mathbf{M}_i = \mathbf{I}_n - \mathbf{P}_{\mathbf{X}_i}\,,\ i = 1, 2\,.$$

The following special cases or extensions of $\mathscr{M}$ will be considered in this paper:

(a) The partitioned linear model is denoted as

$$\mathscr{M}_{12} = \{\mathbf{y},\, \mathbf{X}\boldsymbol{\beta},\, \mathbf{V}\} = \{\mathbf{y},\, \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2,\, \mathbf{V}\} = \{\mathbf{y},\, \boldsymbol{\mu}_1 + \boldsymbol{\mu}_2,\, \mathbf{V}\}\,.$$

(b) In addition to the *full* model $\mathscr{M}_{12}$, we will consider the *small* models $\mathscr{M}_i = \{\mathbf{y},\, \mathbf{X}_i\boldsymbol{\beta}_i,\, \mathbf{V}\}$, $i = 1, 2$, and the *reduced* model

$$\mathscr{M}_{12\cdot 2} = \{\mathbf{M}_2\mathbf{y},\, \mathbf{M}_2\mathbf{X}_1\boldsymbol{\beta}_1,\, \mathbf{M}_2\mathbf{V}\mathbf{M}_2\},$$

which is obtained by premultiplying the model $\mathscr{M}_{12}$ by $\mathbf{M}_2 = \mathbf{I}_n - \mathbf{P}_{\mathbf{X}_2}$.

(c) Let $\mathbf{y}_*$ denote a $q \times 1$ unobservable random vector containing new observations. The new observations are assumed to be generated from

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

where $\mathbf{X}_*$ is a known $q\times p$ matrix, $\boldsymbol{\beta}$ is the same vector of fixed but unknown parameters as in $\mathscr{M}$, and $\boldsymbol{\varepsilon}_*$ is a q-dimensional random error vector. We further assume that

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

where $\boldsymbol{\Psi}$ is known. We denote this setup shortly as

$$\mathscr{M}_* = \left\{\begin{pmatrix}\mathbf{y}\\ \mathbf{y}_*\end{pmatrix},\, \begin{pmatrix}\mathbf{X}\\ \mathbf{X}_*\end{pmatrix}\boldsymbol{\beta},\, \begin{pmatrix}\mathbf{V} & \mathbf{V}_{12}\\ \mathbf{V}_{21} & \mathbf{V}_{22}\end{pmatrix}\right\}. \tag{1}$$

We call $\mathscr{M}_*$ "the linear model with new observations". Our main interest in $\mathscr{M}_*$ lies in predicting $\mathbf{y}_*$ on the basis of observable $\mathbf{y}$. Notice the crucial role of the cross-covariance matrix $\mathrm{cov}(\mathbf{y}, \mathbf{y}_*) = \mathbf{V}_{12} \in \mathbb{R}^{n\times q}$. The mixed linear model can be interpreted as a special case of $\mathscr{M}_*$; see Sec. 4.

Under the model $\mathscr{M} = \{\mathbf{y}, \mathbf{X}\boldsymbol{\beta}, \mathbf{V}\}$, the statistic $\mathbf{Gy}$, where $\mathbf{G}$ is an $n \times n$ matrix, is the best linear unbiased estimator, BLUE, of $\boldsymbol{\mu} = \mathbf{X}\boldsymbol{\beta}$ if $\mathbf{Gy}$ is unbiased, *i.e.*, $\mathbf{GX} = \mathbf{X}$, and it has the smallest covariance matrix in the Löwner sense among all unbiased linear estimators of $\mathbf{X}\boldsymbol{\beta}$; shortly denoted

$$\operatorname{cov}(\mathbf{Gy}) \leq_{\mathrm{L}} \operatorname{cov}(\mathbf{Cy}) \quad \text{for all } \mathbf{C} \in \mathbb{R}^{n\times n}\colon\ \mathbf{CX} = \mathbf{X}\,.$$

The BLUE of an estimable parametric function $\boldsymbol{\mu}_* = \mathbf{X}_*\boldsymbol{\beta}$, where $\mathbf{X}_* \in \mathbb{R}^{q\times p}$, is defined in the corresponding way. Estimability of $\mathbf{X}_*\boldsymbol{\beta}$ means that it has a linear unbiased estimator which happens if and only if $\mathscr{C}(\mathbf{X}_*') \subseteq \mathscr{C}(\mathbf{X}')$. In particular, $\boldsymbol{\mu}_1 = \mathbf{X}_1\boldsymbol{\beta}_1$ is estimable in the partitioned model if and only if

$$\mathscr{C}(\mathbf{X}_1) \cap \mathscr{C}(\mathbf{X}_2) = \{\mathbf{0}\}\,.$$

The random vector $\mathbf{Ay}$ is a linear unbiased predictor (LUP) of $\mathbf{y}_*$ if $\mathrm{E}(\mathbf{y}_* - \mathbf{Ay}) = \mathbf{0}$ for all $\boldsymbol{\beta} \in \mathbb{R}^p$. Such a matrix $\mathbf{A} \in \mathbb{R}^{q\times n}$ exists if and only if $\mathscr{C}(\mathbf{X}_*') \subseteq \mathscr{C}(\mathbf{X}')$, *i.e.*, $\mathbf{X}_*\boldsymbol{\beta}$ is estimable under $\mathscr{M}$ and then we say that $\mathbf{y}_*$ is predictable under $\mathscr{M}_*$. Now a LUP $\mathbf{Ay}$ is the best linear unbiased predictor, BLUP, for $\mathbf{y}_*$, if the covariance matrix of the prediction error, subject to the unbiasedness of the prediction, is minimized:

$$\operatorname{cov}(\mathbf{y}_* - \mathbf{Ay}) \leq_{\mathrm{L}} \operatorname{cov}(\mathbf{y}_* - \mathbf{A}_{\#}\mathbf{y}) \quad \text{for all } \mathbf{A}_{\#} : \mathbf{A}_{\#}\mathbf{X} = \mathbf{X}_*\,.$$

Our matrix expressions will use generalized inverses heavily and in this context it is essential to know whether the expressions are independent of the choice of the generalized inverses involved. Lemma 2.2.4 of Rao and Mitra (1971) gives the condition under which the matrix product $\mathbf{AB}^-\mathbf{C}$ is invariant with respect to the choice of $\mathbf{B}^-$.

Proposition 1: For nonnull matrices $\mathbf{A}$ and $\mathbf{C}$ the following holds:

(a) $\mathbf{AB}^-\mathbf{C} = \mathbf{AB}^+\mathbf{C}$ for all $\mathbf{B}^- \iff \mathscr{C}(\mathbf{C}) \subseteq \mathscr{C}(\mathbf{B})$ & $\mathscr{C}(\mathbf{A}') \subseteq \mathscr{C}(\mathbf{B}')$.

(b) $\mathbf{AA}^-\mathbf{C} = \mathbf{C}$ for some (and hence for all) $\mathbf{A}^- \iff \mathscr{C}(\mathbf{C}) \subseteq \mathscr{C}(\mathbf{A})$.

(c) $\mathbf{C}'\mathbf{A}^-\mathbf{A} = \mathbf{C}'$ for some (and hence for all) $\mathbf{A}^- \iff \mathscr{C}(\mathbf{C}) \subseteq \mathscr{C}(\mathbf{A}')$.

Suppose that the matrix equation

$$\mathbf{YB} = \mathbf{A} \tag{2}$$

is solvable for $\mathbf{Y}$, *i.e.*, $\mathscr{C}(\mathbf{A}') \subseteq \mathscr{C}(\mathbf{B}')$. Then it is well known, see, *e.g.*, Rao and Mitra (1971, p. 24), that the general solution $\mathbf{Y}_0$ to (2) can be written, for example, as

$$\mathbf{Y}_0 = \mathbf{AB}^+ + \mathbf{E}(\mathbf{I} - \mathbf{P}_{\mathbf{B}}) = \mathbf{AB}^+ + \mathbf{EQ}_{\mathbf{B}}\,, \quad \text{where } \mathbf{E} \text{ is free to vary,} \tag{3a}$$

$$\mathbf{Y}_0 = \{\text{one solution to } \mathbf{YB} = \mathbf{A}\} + \{\text{general solution to } \mathbf{YB} = \mathbf{0}\}. \tag{3b}$$

For later considerations, we collect some useful results into the following proposition.

Proposition 2: Consider the partitioned model $\mathscr{M}_{12}(\mathbf{V})$, and let "$\oplus$" refer to the direct sum and "$\boxplus$" to the direct sum of orthogonal subspaces. Then

(a) $\mathscr{C}(\mathbf{X}_1 : \mathbf{X}_2) = \mathscr{C}(\mathbf{X}_1 : \mathbf{M}_1\mathbf{X}_2)$, *i.e.*, $\mathscr{C}(\mathbf{X}) = \mathscr{C}(\mathbf{X}_1) \boxplus \mathscr{C}(\mathbf{M}_1\mathbf{X}_2)$.

(b) $\mathscr{C}(\mathbf{X} : \mathbf{V}) = \mathscr{C}(\mathbf{X} : \mathbf{VM}) = \mathscr{C}(\mathbf{X}) \oplus \mathscr{C}(\mathbf{VM}) = \mathscr{C}(\mathbf{X}) \boxplus \mathscr{C}(\mathbf{MV})$.

(c) $\mathbf{M} = \mathbf{I}_n - \mathbf{P}_{\mathbf{X}} = \mathbf{I}_n - (\mathbf{P}_{\mathbf{X}_1} + \mathbf{P}_{\mathbf{M}_1\mathbf{X}_2}) = \mathbf{M}_1\mathbf{Q}_{\mathbf{M}_1\mathbf{X}_2} = \mathbf{M}_1\mathbf{M}$.

(d) $\mathbf{Q}_{(\mathbf{X}:\mathbf{V})} = \mathbf{I}_n - (\mathbf{P}_{\mathbf{X}} + \mathbf{P}_{\mathbf{MV}}) = \mathbf{M} - \mathbf{P}_{\mathbf{MV}} = \mathbf{M}\mathbf{Q}_{\mathbf{MV}} = \mathbf{M}\mathbf{Q}_{(\mathbf{X}:\mathbf{V})}$.

(e) $\mathrm{r}(\mathbf{AB}) = \mathrm{r}(\mathbf{A}) - \dim \mathscr{C}(\mathbf{A}') \cap \mathscr{C}(\mathbf{B}^{\perp})$ for conformable $\mathbf{A}$ and $\mathbf{B}$.

We assume the model $\mathscr{M}(\mathbf{V})$ to be consistent in the sense that $\mathbf{y}$ lies in $\mathscr{C}(\mathbf{X} : \mathbf{V})$ with probability 1, *i.e.*, the observed numerical value of $\mathbf{y}$ satisfies

$$\mathbf{y} \in \mathscr{C}(\mathbf{X} : \mathbf{V}) = \mathscr{C}(\mathbf{X}) \oplus \mathscr{C}(\mathbf{VM}) = \mathscr{C}(\mathbf{X}) \boxplus \mathscr{C}(\mathbf{MV}),$$

so that

$$\mathbf{y} = \mathbf{Xa} + \mathbf{VMb} \quad \text{for some vectors } \mathbf{a} \in \mathbb{R}^p \text{ and } \mathbf{b} \in \mathbb{R}^n. \tag{4}$$

There is one special class of matrices worth particular attention and that is the set $\mathscr{W}_{\geq}$ of nonnegative definite matrices defined as

$$\mathscr{W}_{\geq} = \left\{ \mathbf{W} \in \mathbb{R}^{n\times n} : \mathbf{W} = \mathbf{V} + \mathbf{XUU'X'},\ \mathscr{C}(\mathbf{W}) = \mathscr{C}(\mathbf{X} : \mathbf{V}) \right\}. \tag{5}$$

In (5) $\mathbf{U}$ can be any matrix comprising p rows as long as $\mathscr{C}(\mathbf{W}) = \mathscr{C}(\mathbf{X} : \mathbf{V})$ is satisfied. One obvious choice is $\mathbf{U} = \mathbf{I}_p$. In particular, if $\mathscr{C}(\mathbf{X}) \subseteq \mathscr{C}(\mathbf{V})$, we can choose $\mathbf{U} = \mathbf{0}$. The extended version of $\mathscr{W}_{\geq}$ is

$$\mathscr{W} = \left\{ \mathbf{W} \in \mathbb{R}^{n\times n} : \mathbf{W} = \mathbf{V} + \mathbf{XTX'},\ \mathscr{C}(\mathbf{W}) = \mathscr{C}(\mathbf{X} : \mathbf{V}) \right\}. \tag{6}$$

Above $\mathbf{T} \in \mathbb{R}^{p\times p}$ is free to vary subject to condition $\mathscr{C}(\mathbf{W}) = \mathscr{C}(\mathbf{X} : \mathbf{V})$. Notice that $\mathbf{W}$ belonging to $\mathscr{W}$ is not necessarily nonnegative definite and it can be nonsymmetric. We may use the notations $\mathscr{W}(\mathscr{A})$ or $\mathscr{W}(\mathbf{V})$ to indicate that the model $\mathscr{A}$ or the covariance matrix $\mathbf{V}$ is under consideration. Proposition 3 collects together some properties of the class $\mathscr{W}$.

Proposition 3: Let $\mathbf{V} \in \mathbb{R}^{n\times n}$ be nonnegative definite and let $\mathbf{X} \in \mathbb{R}^{n\times p}$ and define $\mathbf{W}$ as $\mathbf{W} = \mathbf{V} + \mathbf{XTX'}$, where $\mathbf{T} \in \mathbb{R}^{p\times p}$. Then the following statements are equivalent:

(a) $\mathscr{C}(\mathbf{X} : \mathbf{V}) = \mathscr{C}(\mathbf{W})$,

(b) $\mathscr{C}(\mathbf{X}) \subseteq \mathscr{C}(\mathbf{W})$,

(c) $\mathbf{X}'\mathbf{W}^{-}\mathbf{X}$ is invariant for any choice of $\mathbf{W}^{-}$,

(d) $\mathscr{C}(\mathbf{X}'\mathbf{W}^{-}\mathbf{X}) = \mathscr{C}(\mathbf{X}')$ for any choice of $\mathbf{W}^{-}$,

(e) $\mathbf{X}(\mathbf{X}'\mathbf{W}^{-}\mathbf{X})^{-}\mathbf{X}'\mathbf{W}^{-}\mathbf{X} = \mathbf{X}$ for any choices of $\mathbf{W}^{-}$ and $(\mathbf{X}'\mathbf{W}^{-}\mathbf{X})^{-}$.

Observe that obviously $\mathscr{C}(\mathbf{W}) = \mathscr{C}(\mathbf{W}')$ since

$$\mathscr{C}(\mathbf{W}') = \mathscr{C}(\mathbf{V} + \mathbf{X}\mathbf{T}'\mathbf{X}') \subseteq \mathscr{C}(\mathbf{W}), \quad \mathrm{r}(\mathbf{W}') = \mathrm{r}(\mathbf{W}),$$

and hence in statements (a)–(e) $\mathbf{W}$ can be replaced with with $\mathbf{W}'$. For further properties of $\mathscr{W}$, see, *e.g.*, Baksalary and Mathew (1990, Th. 2), and Puntanen *et al.* (2011, Sec. 12.3). Haslett *et al.* (2022a) provide an extensive review of the class $\mathscr{W}$.

Let us cite Puntanen *et al.* (2011, Sec. 5.13):

Proposition 4: Consider the model $\mathscr{M} = \{\mathbf{y}, \mathbf{X}\boldsymbol{\beta}, \mathbf{V}\}$ and let $\mathbf{W} \in \mathscr{W}(\mathscr{M})$. Then

(a) $\mathscr{C}(\mathbf{VM})^{\perp} = \mathscr{C}(\mathbf{W}^{+}\mathbf{X} : \mathbf{Q}_{\mathbf{W}}) = \mathscr{C}[(\mathbf{W}^{+})'\mathbf{X} : \mathbf{Q}_{\mathbf{W}}]$,

(b) $\mathscr{C}(\mathbf{W}^{+}\mathbf{X})^{\perp} = \mathscr{C}(\mathbf{WM} : \mathbf{Q}_{\mathbf{W}}) = \mathscr{C}(\mathbf{VM} : \mathbf{Q}_{\mathbf{W}})$.

It appears to be useful to denote

$$\dot{\mathbf{M}} = \mathbf{M}(\mathbf{MVM})^{-}\mathbf{M}.$$

The matrix $\dot{\mathbf{M}}$ is unique with respect to the choice of the generalized inverse $(\mathbf{MVM})^{-}$ if and only if $\mathbb{R}^n = \mathscr{C}(\mathbf{X} : \mathbf{V})$. However, for example, $\mathbf{V}\dot{\mathbf{M}}\mathbf{P}_{\mathbf{W}}$ is always unique. It is noteworthy that using the Moore–Penrose inverse the following holds:

$$\mathbf{M}(\mathbf{MVM})^{+}\mathbf{M} = (\mathbf{MVM})^{+}\mathbf{M} = \mathbf{M}(\mathbf{MVM})^{+} = (\mathbf{MVM})^{+}. \tag{7}$$

In particular, for a positive definite $\mathbf{V}$ we have, for any $(\mathbf{MVM})^{-}$,

$$\begin{aligned}\mathbf{M}(\mathbf{MVM})^{-}\mathbf{M} &= \mathbf{V}^{-1/2}\mathbf{P}_{\mathbf{V}^{1/2}\mathbf{M}}\mathbf{V}^{-1/2} = \mathbf{V}^{-1/2}(\mathbf{I}_n - \mathbf{P}_{(\mathbf{V}^{1/2}\mathbf{M})^{\perp}})\mathbf{V}^{-1/2} \\ &= \mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{X}'\mathbf{V}^{-1} =: \mathbf{V}^{-1}(\mathbf{I}_n - \mathbf{P}_{\mathbf{X};\mathbf{V}^{-1}}),\end{aligned}$$

where we have used the obvious fact $\mathscr{C}(\mathbf{V}^{1/2}\mathbf{M})^{\perp} = \mathscr{C}(\mathbf{V}^{-1/2}\mathbf{X})$.

We will use the following notation:

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

Notice that $\mathbf{P}_{\mathbf{X};\mathbf{W}^{+}}$ and $\mathbf{P}_{\mathbf{X}_*;\mathbf{W}^{+}}$ are invariant for any choice of the generalized inverses $\mathbf{W}^{-}$ and $(\mathbf{X}'\mathbf{W}^{-}\mathbf{X})^{-}$ but this invariance does not concern the matrix

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

Proposition 5: Consider the linear model $\mathscr{M} = \{\mathbf{y}, \mathbf{X}\boldsymbol{\beta}, \mathbf{V}\}$. Let $\mathbf{T}$ be any $p \times p$ matrix such that the matrix $\mathbf{W} = \mathbf{V} + \mathbf{X}\mathbf{T}\mathbf{X}'$ satisfies the condition $\mathscr{C}(\mathbf{W}) = \mathscr{C}(\mathbf{X} : \mathbf{V})$, *i.e.*, $\mathbf{W} \in \mathscr{W}(\mathscr{M})$, and denote $\dot{\mathbf{M}} = \mathbf{M}(\mathbf{MVM})^{-}\mathbf{M}$. Then

(a) $\mathbf{P}_{\mathbf{W}}\mathbf{M}(\mathbf{MVM})^{-}\mathbf{M}\mathbf{P}_{\mathbf{W}} = \mathbf{W}^{+} - \mathbf{W}^{+}\mathbf{X}(\mathbf{X}'\mathbf{W}^{-}\mathbf{X})^{-}\mathbf{X}'\mathbf{W}^{+}$.

(b) $\mathbf{P_W M (MVM)^- M P_W} = \mathbf{(MVM)^+} = \mathbf{P_W \dot{M} P_W}$.

(c) $\mathbf{P}_{\mathbf{X};\mathbf{W}^+} = \mathbf{X(X'W^-X)^-X'W^+} = \mathbf{P_W} - \mathbf{VM(MVM)^-MP_W}$.

(d) $\mathbf{P}_{\mathbf{X};\mathbf{W}^+} = \mathbf{H} - \mathbf{HVM(MVM)^-MP_W} = \mathbf{H} - \mathbf{HVM(MVM)^+M}$, where $\mathbf{H} = \mathbf{P_X}$.

The result (a) is the the most crucial one in Proposition 5. For the proof of (a), see Puntanen *et al.* (2011, Prop. 15.2) and Isotalo *et al.* (2008, Cor. 2.2). Notice that in light of Proposition 2, we have

$$\mathbf{P_W} = \mathbf{P_X} + \mathbf{P_{MV}} = \mathbf{H} + \mathbf{P_{MVM}}\,, \quad \mathbf{MP_W} = \mathbf{MP_{MV}} = \mathbf{P_{MV}} = \mathbf{P_{MVM}}\,,$$

which implies (b) of Proposition 5. Premultiplying (a) by $\mathbf{W}$ and using $\mathscr{C}(\mathbf{X}) \subseteq \mathscr{C}(\mathbf{W})$ gives (c), *i.e.*,

$$\begin{aligned}\mathbf{X(X'W^-X)^-X'W^+} &= \mathbf{P_W} - \mathbf{VM(MVM)^-MP_W} = \mathbf{P_W} - \mathbf{VM(MVM)^+MP_W}\\ &= \mathbf{P_W} - \mathbf{V(MVM)^+P_W} = \mathbf{P_W} - \mathbf{V(MVM)^+}\\ &= \mathbf{P_W} - \mathbf{VM(MVM)^+M}\,,\end{aligned} \tag{8}$$

where we have used (7) and the fact that $\mathscr{C}[\mathbf{(MVM)^+}] = \mathscr{C}(\mathbf{MVM}) \subseteq \mathscr{C}(\mathbf{W})$. From (8) we immediately confirm that $\mathbf{X(X'W^-X)^-X'W^+}$ is invariant with respect to the choice of $\mathbf{W} \in \mathscr{W}$ supposing that $\mathscr{C}(\mathbf{X} : \mathbf{V}) = \mathscr{C}(\mathbf{W})$ is holding. Premultiplying (8) by $\mathbf{H} = \mathbf{P_X}$ gives

$$\begin{aligned}\mathbf{X(X'W^-X)^-X'W^+} &= \mathbf{H} - \mathbf{HVM(MVM)^-MP_W} = \mathbf{H} - \mathbf{HVM(MVM)^+MP_W}\\ &= \mathbf{H} - \mathbf{HV(MVM)^+P_W} = \mathbf{H} - \mathbf{HV(MVM)^+}\\ &= \mathbf{H} - \mathbf{HVM(MVM)^+M}\,.\end{aligned} \tag{9}$$

Remark 1: The equality (9) follows from (a) of Proposition 5. However, it is interesting to prove (9) directly. This is done by first observing that

$$\mathbf{X(X'W^-X)^-X'W^+(X : VM)} = [\mathbf{H} - \mathbf{HVM(MVM)^+MP_W}]\mathbf{(X : VM)}\,, \tag{10}$$

and then confirming that

$$\mathbf{X(X'W^-X)^-X'W^+Q_W} = [\mathbf{H} - \mathbf{HVM(MVM)^+MP_W}]\mathbf{Q_W}\,. \tag{11}$$

Now (10) and (11) together imply (9). □

2. The fundamental BLUE equation

Theorem 1 below provides so-called fundamental BLUE equations.

Theorem 1: [BLUE] Consider the linear model $\mathscr{M} = \{\mathbf{y}, \mathbf{X}\boldsymbol{\beta}, \mathbf{V}\}$.

(a) Then the linear estimator $\mathbf{Gy}$ is the BLUE for $\boldsymbol{\mu} = \mathbf{X}\boldsymbol{\beta}$ if and only if $\mathbf{G} \in \mathbb{R}^{n\times n}$ satisfies the equation

$$\mathbf{G}(\mathbf{X} : \mathbf{VX}^{\perp}) = (\mathbf{X} : \mathbf{0})\,. \tag{12}$$

(b) Let $\boldsymbol{\mu}_* = \mathbf{X}_*\boldsymbol{\beta}$, where $\mathbf{X}_* \in \mathbb{R}^{q\times p}$, be estimable so that $\mathscr{C}(\mathbf{X}_*') \subseteq \mathscr{C}(\mathbf{X}')$. Then $\mathbf{By}$ is the BLUE of $\boldsymbol{\mu}_*$ if and only if $\mathbf{B} \in \mathbb{R}^{q\times n}$ satisfies the equation

$$\mathbf{B}(\mathbf{X} : \mathbf{VX}^{\perp}) = (\mathbf{X}_* : \mathbf{0})\,. \tag{13}$$

(c) Let $\boldsymbol{\mu}_1 = \mathbf{X}_1\boldsymbol{\beta}_1$ be estimable in the partitioned model $\mathscr{M}_{12}$. Then $\mathbf{Cy}$ is the BLUE of $\boldsymbol{\mu}_1$ if and only if

$$\mathbf{C}(\mathbf{X}_1 : \mathbf{X}_2 : \mathbf{VX}^{\perp}) = (\mathbf{X}_1 : \mathbf{0} : \mathbf{0})\,. \tag{14}$$

For the proofs, see, *e.g.*, Rao (1973, p. 282) and for coordinate-free approach Drygas (1970, p. 55) and Zmyślony (1980). For further proofs see, for example, Groß (2004), Kala (1981, Th. 3.1), Puntanen *et al.* (2000), Puntanen *et al.* (2011, Th. 10), and Baksalary (2004).

For Theorem 2, characterizing the BLUP, see, *e.g.*, Christensen (2020, Th. 6.6.3), and Isotalo and Puntanen (2006, p. 1015).

Theorem 2: [BLUP] Consider the linear model with new observations defined as $\mathscr{M}_*$ in (1), where $\mathscr{C}(\mathbf{X}_*') \subseteq \mathscr{C}(\mathbf{X}')$, *i.e.*, $\mathbf{y}_*$ is predictable. Then:

(a) $\mathbf{Ay}$ is the BLUP for $\mathbf{y}_*$ if and only if $\mathbf{A}(\mathbf{X} : \mathbf{VX}^{\perp}) = (\mathbf{X}_* : \mathbf{V}_{21}\mathbf{X}^{\perp})\,.$

(b) $\mathbf{By}$ is the BLUE of $\boldsymbol{\mu}_* = \mathbf{X}_*\boldsymbol{\beta}$ if and only if $\mathbf{B}(\mathbf{X} : \mathbf{VX}^{\perp}) = (\mathbf{X}_* : \mathbf{0})\,.$

(c) $\mathbf{Dy}$ is the BLUP for $\boldsymbol{\varepsilon}_*$ if and only if $\mathbf{D}(\mathbf{X} : \mathbf{VX}^{\perp}) = (\mathbf{0} : \mathbf{V}_{21}\mathbf{X}^{\perp})\,.$

Theorems 1 and 2 offer extremely handy tools to check whether a given estimator/predictor is a BLUE/BLUP. Moreover, they provide convenient ways to introduce various representations for the BLUE/BLUP. The solutions are unique if and only if $\mathscr{C}(\mathbf{X} : \mathbf{VX}^{\perp}) = \mathbb{R}^n$. Trivially, one choice for $\mathbf{X}^{\perp}$ is $\mathbf{M} = \mathbf{I}_n - \mathbf{P_X}$. In view of (3b), the *general* solution for $\mathbf{G}$ in (12) can be expressed as

$$\{\text{one solution to } \mathbf{G}(\mathbf{X} : \mathbf{VM}) = (\mathbf{X} : \mathbf{0})\} + \{\text{general sol. to } \mathbf{G}(\mathbf{X} : \mathbf{VM}) = (\mathbf{0} : \mathbf{0})\}. \tag{15}$$

Suppose that $\mathbf{W} \in \mathscr{W}(\mathscr{M})$ where $\mathscr{M} = \{\mathbf{y}, \mathbf{X}\boldsymbol{\beta}, \mathbf{V}\}$. We observe immediately that

$$\mathbf{X}(\mathbf{X}'\mathbf{W}^-\mathbf{X})^-\mathbf{X}'\mathbf{W}^+(\mathbf{X} : \mathbf{VM}) = \mathbf{X}(\mathbf{X}'\mathbf{W}^-\mathbf{X})^-\mathbf{X}'\mathbf{W}^+(\mathbf{X} : \mathbf{WM}) = (\mathbf{X} : \mathbf{0})\,,$$

and so

$$\begin{aligned}\mathbf{P}_{\mathbf{X};\mathbf{W}^+}\mathbf{y} &= \mathbf{X}(\mathbf{X}'\mathbf{W}^-\mathbf{X})^-\mathbf{X}'\mathbf{W}^+\mathbf{y} \\ &= \mathbf{X}(\mathbf{X}'\mathbf{W}^-\mathbf{X})^-\mathbf{X}'\mathbf{W}^-\mathbf{y} = \mathbf{P}_{\mathbf{X}_*;\mathbf{W}^-}\mathbf{y} = \text{BLUE}(\mathbf{X}\boldsymbol{\beta}) = \tilde{\boldsymbol{\mu}}\,,\end{aligned}$$

where we have used the consistency condition (4) to replace $\mathbf{W}^+$ with $\mathbf{W}^-$. Correspondingly, for an estimable $\boldsymbol{\mu}_* = \mathbf{X}_*\boldsymbol{\beta}$ we have

$$\mathbf{P}_{\mathbf{X}_*;\mathbf{W}^+}\mathbf{y} = \mathbf{X}_*(\mathbf{X}'\mathbf{W}^-\mathbf{X})^-\mathbf{X}'\mathbf{W}^+\mathbf{y} = \text{BLUE}(\mathbf{X}_*\boldsymbol{\beta}) = \tilde{\boldsymbol{\mu}}_*\,. \tag{16}$$

Moreover, in view of (8) and (9) and the consistency of the model $\mathscr{M}$, we have

$$\tilde{\boldsymbol{\mu}} = \mathbf{P_W y} - \mathbf{VM(MVM)^- MP_W y} = \mathbf{y} - \mathbf{VM(MVM)^- My}, \tag{17a}$$

$$\begin{aligned}\tilde{\boldsymbol{\mu}} &= \mathbf{HP_W y} - \mathbf{HVM(MVM)^- MP_W y} = \mathbf{Hy} - \mathbf{HVM(MVM)^- My} \\ &= \mathrm{OLSE}(\boldsymbol{\mu}) - \mathbf{HVM(MVM)^- My},\end{aligned} \tag{17b}$$

which hold for all $\mathbf{y} \in \mathscr{C}(\mathbf{X} : \mathbf{V})$ and OLSE($\boldsymbol{\mu}$) refers to the ordinary least squares estimator of $\boldsymbol{\mu}$. One of the first references to (17b) is Albert (1973). Notice that in light of (17a) the BLUE's residual can be expressed as

$$\tilde{\boldsymbol{\varepsilon}} := \mathbf{y} - \tilde{\boldsymbol{\mu}} = \mathbf{VM(MVM)^- My}\,.$$

If $\mathscr{C}(\mathbf{X}) \subseteq \mathscr{C}(\mathbf{V})$, then $\mathscr{M}$ is said to be a weakly singular linear model. In this situation we can choose $\mathbf{T} = \mathbf{0}$ in (6) and thereby replace $\mathbf{W}$ with $\mathbf{V}$ so that

$$\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta}) = \mathbf{X}(\mathbf{X}'\mathbf{V}^-\mathbf{X})^-\mathbf{X}'\mathbf{V}^-\mathbf{y}\,. \tag{18}$$

3. How to solve the fundamental BLUE equation?

In the previous section we have shown that certain expressions satisfy the fundamental BLUE equation. It is another question how to end up into these solutions, *i.e.*, how to *introduce* these solutions. And this is just what we aim to do in this section. We believe that our approaches – some not much used in literature as straightforward as they are – may increase the insight of the meaning of the fundamental BLUE equations.

To begin, notice that part (b) of Theorem 1 can be expressed so that $\mathbf{By} = \mathrm{BLUE}(\boldsymbol{\mu}_*)$ if and only if the following two conditions are satisfied:

$$\text{(i) } \mathbf{By} \text{ is unbiased for } \boldsymbol{\mu}_*\,, \quad \text{(ii) } \mathbf{By} \text{ is uncorrelated with } \mathbf{My}. \tag{19}$$

How to solve (19)? As said, by simple substitution we can check that $\mathbf{P}_{\mathbf{X}_*;\mathbf{W}^+}\mathbf{y}$ is indeed the BLUE for $\boldsymbol{\mu}_* = \mathbf{X}_*\boldsymbol{\beta}$ under $\mathscr{M}$. We may now raise the question how to *introduce* a solution $\mathbf{B}$ for fundamental BLUE equation

$$\mathbf{B}(\mathbf{X} : \mathbf{VM}) = (\mathbf{X}_* : \mathbf{0})\,, \tag{20}$$

where $\mathbf{X}_* = \mathbf{LX}$ for some $\mathbf{L}$ so that $\boldsymbol{\mu}_* = \mathbf{X}_*\boldsymbol{\beta}$ is estimable. Notice that then

$$\mathbf{X}_*\mathbf{X}^+ = \mathbf{LXX}^+ = \mathbf{LP_X} = \mathbf{LH}\,.$$

■ **Solution 1:** The general solution to the unbiasedness condition (i) in (19), *i.e.*, to $\mathbf{BX} = \mathbf{X}_*$, can be expressed, *e.g.*, as

$$\mathbf{B}_0 = \mathbf{X}_*\mathbf{X}^+ + \mathbf{E}(\mathbf{I}_n - \mathbf{P_X}) = \mathbf{LH} + \mathbf{EM}\,, \quad \text{where } \mathbf{E} \text{ is free to vary.}$$

Hence the requirement (ii) in (19), *i.e.*, $\mathbf{B}_0\mathbf{VM} = \mathbf{0}$, is satisfied if and only if

$$\mathbf{LHVM} + \mathbf{EMVM} = \mathbf{0}\,, \quad \text{i.e.,} \quad \mathbf{EMVM} = -\mathbf{LHVM}\,,$$

from which we get the general expression for $\mathbf{E}$:

$$\mathbf{E}_0 = -\mathbf{LHVM}(\mathbf{MVM})^{+} + \mathbf{E}_1\mathbf{Q}_{\mathbf{MV}}\,, \quad \text{where } \mathbf{E}_1 \text{ is free to vary.} \tag{21}$$

In view of the decomposition

$$\mathbf{Q}_{(\mathbf{X}:\mathbf{V})} = \mathbf{I}_n - (\mathbf{P}_{\mathbf{X}} + \mathbf{P}_{\mathbf{MV}}) = -\mathbf{H} + \mathbf{Q}_{\mathbf{MV}}\,,$$

we have $\mathbf{Q}_{\mathbf{MV}} = \mathbf{H} + \mathbf{Q}_{(\mathbf{X}:\mathbf{V})}$, and thereby by (21) we have

$$\mathbf{E}_0 = -\mathbf{LHVM}(\mathbf{MVM})^{+} + \mathbf{E}_1(\mathbf{H} + \mathbf{Q}_{(\mathbf{X}:\mathbf{V})})\,,$$

and hence the expression for the general solution to $\mathbf{B}$ in (20) can be written as

$$\begin{aligned} \mathbf{B}_0 &= \mathbf{LH} + \mathbf{E}_0\mathbf{M} = \mathbf{LH} - \mathbf{LHVM}(\mathbf{MVM})^{+}\mathbf{M} + \mathbf{E}_1\mathbf{Q}_{(\mathbf{X}:\mathbf{V})}\mathbf{M} \\ &= \mathbf{L}[\mathbf{H} - \mathbf{HVM}(\mathbf{MVM})^{+}\mathbf{M}] + \mathbf{E}_1\mathbf{Q}_{(\mathbf{X}:\mathbf{V})} \\ &= \mathbf{LX}(\mathbf{X}'\mathbf{W}^{-}\mathbf{X})^{-}\mathbf{X}'\mathbf{W}^{+} + \mathbf{E}_1\mathbf{Q}_{(\mathbf{X}:\mathbf{V})} \\ &= \mathbf{X}_*(\mathbf{X}'\mathbf{W}^{-}\mathbf{X})^{-}\mathbf{X}'\mathbf{W}^{+} + \mathbf{E}_1\mathbf{Q}_{(\mathbf{X}:\mathbf{V})}\,, \end{aligned} \tag{22}$$

where $\mathbf{E}_1$ is free to vary. In (22) we have used (9).

■ **Solution 2:** An alternative way to introduce a representation for $\mathbf{B}$ is to start from $\mathbf{BVM} = \mathbf{0}$, which by Proposition 4 is equivalent to

$$\mathscr{C}(\mathbf{B}') \subseteq \mathscr{C}(\mathbf{VM})^{\perp} = \mathscr{C}\Big[(\mathbf{W}^{+})'\mathbf{X} : \mathbf{Q}_{\mathbf{W}}\Big]\,,$$

where $\mathbf{W} \in \mathscr{W}(\mathscr{M})$, so that

$$\mathbf{B}' = (\mathbf{W}^{+})'\mathbf{XR} + \mathbf{Q}_{\mathbf{W}}\mathbf{S}\,, \quad \text{for some } \mathbf{S} \text{ and } \mathbf{R}.$$

Now the unbiasedness condition $\mathbf{X}'\mathbf{B}' = \mathbf{X}_*'$ holds if and only if

$$\mathbf{X}'\mathbf{W}^{+}\mathbf{XR} + \mathbf{X}'\mathbf{Q}_{\mathbf{W}}\mathbf{S} = \mathbf{X}'\mathbf{W}^{+}\mathbf{XR} = \mathbf{X}_*'\,,$$

from which it follows that the general expression for $\mathbf{R}$ can be expressed as

$$\mathbf{R} = (\mathbf{X}'\mathbf{W}^{+}\mathbf{X})^{-}\mathbf{X}_*' + \mathbf{Q}_{\mathbf{X}'\mathbf{W}^{+}}\mathbf{E}_3 = (\mathbf{X}'\mathbf{W}^{+}\mathbf{X})^{-}\mathbf{X}_*' + \mathbf{Q}_{\mathbf{X}'}\mathbf{E}_3\,,$$

where $\mathbf{E}_3$ is free to vary. Hence the general expression for $\mathbf{B}'$ satisfying $\mathbf{B}(\mathbf{X} : \mathbf{VM}) = (\mathbf{X}_* : \mathbf{0})$ can be written as as

$$\begin{aligned} \mathbf{B}_0' &= (\mathbf{W}^{+})'\mathbf{X}(\mathbf{X}'\mathbf{W}^{+}\mathbf{X})^{-}\mathbf{X}_*' + (\mathbf{W}^{+})'\mathbf{X}\mathbf{Q}_{\mathbf{X}'}\mathbf{E}_3 + \mathbf{Q}_{\mathbf{W}}\mathbf{S} \\ &= (\mathbf{W}^{+})'\mathbf{X}(\mathbf{X}'\mathbf{W}^{+}\mathbf{X})^{-}\mathbf{X}_*' + \mathbf{Q}_{\mathbf{W}}\mathbf{S}\,, \end{aligned}$$

so that

$$\mathbf{B}_0 = \mathbf{X}_*(\mathbf{X}'\mathbf{W}^{+}\mathbf{X})^{-}\mathbf{X}'\mathbf{W}^{+} + \mathbf{S}'\mathbf{Q}_{\mathbf{W}} = \mathbf{P}_{\mathbf{X}_*;\mathbf{W}^{+}} + \mathbf{S}'\mathbf{Q}_{\mathbf{W}}\,,$$

where $\mathbf{S}$ is free to vary. Thus we have obtained the same presentation as in (22).

■ **Solution 3:** It is clear that there exists a matrix $\mathbf{X}^{\sim}$ such that $\mathbf{X}\mathbf{X}^{\sim}\mathbf{y}$ is the BLUE for $\mathbf{X}\boldsymbol{\beta}$, *i.e.*,

$$\mathbf{X}\mathbf{X}^{\sim}(\mathbf{X} : \mathbf{V}\mathbf{M}) = (\mathbf{X} : \mathbf{0})\,, \tag{23}$$

so that $\mathbf{X}^{\sim} \in \{\mathbf{X}^{-}\}$. According to Rao and Mitra (1971, Th. 2.4.1), the general expression for a generalized inverse $\mathbf{X}^{\sim} \in \{\mathbf{X}^{-}\}$ can be written as

$$\mathbf{X}^{\sim} = \mathbf{X}^{+} + \mathbf{E}_3(\mathbf{I}_n - \mathbf{P}_{\mathbf{X}}) + (\mathbf{I}_p - \mathbf{P}_{\mathbf{X}'})\mathbf{E}_4\,,$$

where $\mathbf{E}_3$ and $\mathbf{E}_4$ are free to vary. Now

$$\mathbf{X}\mathbf{X}^{\sim} = \mathbf{H} + \mathbf{X}\mathbf{E}_3\mathbf{M}\,, \tag{24}$$

and hence (23) holds if and only if

$$\mathbf{X}\mathbf{X}^{\sim}\mathbf{V}\mathbf{M} = \mathbf{H}\mathbf{V}\mathbf{M} + \mathbf{X}\mathbf{E}_3\mathbf{M}\mathbf{V}\mathbf{M} = \mathbf{0}\,,$$

i.e.,

$$\mathbf{X}\mathbf{E}_3\mathbf{M}\mathbf{V}\mathbf{M} = -\mathbf{H}\mathbf{V}\mathbf{M}\,. \tag{25}$$

One solution for $\mathbf{X}\mathbf{E}_3$ in (25) is $\mathbf{X}\mathbf{E}_3 = -\mathbf{H}\mathbf{V}\mathbf{M}(\mathbf{M}\mathbf{V}\mathbf{M})^{+}$, and thus $\mathbf{X}\mathbf{X}^{\sim}$ in (24) can be written as

$$\mathbf{X}\mathbf{X}^{\sim} = \mathbf{H} - \mathbf{H}\mathbf{V}\mathbf{M}(\mathbf{M}\mathbf{V}\mathbf{M})^{+}\mathbf{M}\,. \tag{26}$$

Notice that $\mathbf{X}^{\sim}$ satisfying (23) can be written as

$$\mathbf{X}^{=} = \mathbf{X}^{+} - \mathbf{X}^{+}\mathbf{V}\mathbf{M}(\mathbf{M}\mathbf{V}\mathbf{M})^{+}\mathbf{M}\,.$$

Another choice for $\mathbf{X}^{\sim}$ satisfying (23) is obviously

$$\mathbf{X}^{\#} = (\mathbf{X}'\mathbf{W}^{+}\mathbf{X})^{+}\mathbf{X}'\mathbf{W}^{+}.$$

It is easy to confirm that actually $\mathbf{X}^{=} = \mathbf{X}^{\#}$.

■ **Solution 4:** A very straightforward way to find a general solution to $\mathbf{B}(\mathbf{X} : \mathbf{V}\mathbf{M}) = (\mathbf{X}_{*} : \mathbf{0})$ is to write

$$\mathbf{B}_0 = (\mathbf{X}_{*} : \mathbf{0})(\mathbf{X} : \mathbf{V}\mathbf{M})^{+} + \mathbf{E}\mathbf{Q}_{(\mathbf{X}:\mathbf{V})} =: \mathbf{B}_1 + \mathbf{E}\mathbf{Q}_{(\mathbf{X}:\mathbf{V})}\,,$$

where the matrix $\mathbf{E}$ is free to vary. It is easy to confirm that $(\mathbf{X} : \mathbf{V}\mathbf{M})^{+}$ can be written as

$$(\mathbf{X} : \mathbf{V}\mathbf{M})^{+} = \begin{pmatrix} \mathbf{X}^{+}[\mathbf{I}_n - \mathbf{V}(\mathbf{M}\mathbf{V}\mathbf{M})^{+}] \\ (\mathbf{M}\mathbf{V}\mathbf{M})^{+} \end{pmatrix}. \tag{27}$$

Therefore, when $\mathbf{X}_{*} = \mathbf{L}\mathbf{X}$ for some $\mathbf{L} \in \mathbb{R}^{q\times n}$ so that $\mathbf{X}_{*}\mathbf{X}^{+} = \mathbf{L}\mathbf{H}$,

$$\begin{aligned}\mathbf{B}_1 &= (\mathbf{X}_{*} : \mathbf{0})(\mathbf{X} : \mathbf{V}\mathbf{M})^{+} = \mathbf{X}_{*}\mathbf{X}^{+}[\mathbf{I}_n - \mathbf{V}(\mathbf{M}\mathbf{V}\mathbf{M})^{+}] \\ &= \mathbf{L}[\mathbf{H} - \mathbf{H}\mathbf{V}(\mathbf{M}\mathbf{V}\mathbf{M})^{+}] = \mathbf{L}\mathbf{X}(\mathbf{X}'\mathbf{W}^{+}\mathbf{X})^{+}\mathbf{X}'\mathbf{W}^{+}.\end{aligned} \tag{28}$$

■ **Solution 5:** In the partitioned model $\mathscr{M}_{12} = \{\mathbf{y}, (\mathbf{X}_1 : \mathbf{X}_2)\boldsymbol{\beta}, \mathbf{V}\}$, one expression for the BLUE of $\boldsymbol{\mu}_1$ can be obtained from (16) yielding

$$\mathrm{BLUE}(\boldsymbol{\mu}_1 \mid \mathscr{M}_{12}) = \tilde{\boldsymbol{\mu}}_1(\mathscr{M}_{12}) = (\mathbf{X}_1 : \mathbf{0})(\mathbf{X}'\mathbf{W}^{-}\mathbf{X})^{-}\mathbf{X}'\mathbf{W}^{+}\mathbf{y} =: \mathbf{P}_{\#1}\mathbf{y}\,.$$

Premultiplying the model $\mathscr{M}_{12}$ by $\mathbf{M}_2$ yields the reduced model

$$\mathscr{M}_{12\cdot 2} = \{\mathbf{M}_2\mathbf{y},\, \mathbf{M}_2\mathbf{X}_1\boldsymbol{\beta}_1,\, \mathbf{M}_2\mathbf{V}\mathbf{M}_2\}\,.$$

The fundamental BLUE equation for estimating $\boldsymbol{\theta}_1 := \mathbf{M}_2\mathbf{X}_1\boldsymbol{\beta}_1$ under $\mathscr{M}_{12\cdot 2}$ is now

$$\mathbf{L}(\mathbf{M}_2\mathbf{X}_1 : \mathbf{M}_2\mathbf{V}\mathbf{M}_2 \cdot \mathbf{Q}_{\mathbf{M}_2\mathbf{X}_1}) = (\mathbf{M}_2\mathbf{X}_1 : \mathbf{0})\,. \tag{29}$$

To find a solution for $\mathbf{L}$ in (29), we observe that choosing $\mathbf{W} = \mathbf{V} + \mathbf{X}\mathbf{X}' \in \mathscr{W}(\mathscr{M}_{12})$ we have $\mathbf{M}_2\mathbf{W}\mathbf{M}_2 \in \mathscr{W}(\mathscr{M}_{12\cdot 2})$. Hence one solution for $\mathbf{L}$ in (29) is

$$\mathbf{L} = \mathbf{M}_2\mathbf{X}_1(\mathbf{X}_1'\dot{\mathbf{M}}_2\mathbf{X}_1)^-\mathbf{X}_1'\dot{\mathbf{M}}_2\mathbf{y} =: \mathbf{M}_2 \cdot \mathbf{P}_{\mathbf{X}_1;\dot{\mathbf{M}}_2}\,,$$

where

$$\dot{\mathbf{M}}_2 = \mathbf{M}_2(\mathbf{M}_2\mathbf{W}\mathbf{M}_2)^+\mathbf{M}_2\,,$$

and so $\mathbf{L}\mathbf{y} = \mathrm{BLUE}(\boldsymbol{\theta}_1 \mid \mathscr{M}_{12\cdot 2})$ and $\mathbf{P}_{\mathbf{X}_1;\dot{\mathbf{M}}_2}\mathbf{y} = \mathrm{BLUE}(\boldsymbol{\mu}_1 \mid \mathscr{M}_{12\cdot 2})$, *i.e.*,

$$\mathbf{X}_1(\mathbf{X}_1'\dot{\mathbf{M}}_2\mathbf{X}_1)^-\mathbf{X}_1'\dot{\mathbf{M}}_2\mathbf{y} = \mathbf{P}_{\mathbf{X}_1;\dot{\mathbf{M}}_2}\mathbf{y} = \tilde{\boldsymbol{\mu}}_1(\mathscr{M}_{12\cdot 2})\,.$$

It is easy to confirm that

$$\mathbf{P}_{\mathbf{X}_1;\dot{\mathbf{M}}_2}(\mathbf{X}_1 : \mathbf{X}_2 : \mathbf{V}\mathbf{M}) = (\mathbf{X}_1 : \mathbf{0} : \mathbf{0})\,,$$

so that the BLUEs of $\boldsymbol{\mu}_1$ under $\mathscr{M}_{12}$ and $\mathscr{M}_{12\cdot 2}$ coincide, which is the message of the Frisch–Waugh–Lovell theorem, see, *e.g.*, Groß and Puntanen (2000, Sec. 6).

Actually the following holds, see Haslett *et al.* (2023, Prop. 3.1),

$$\mathbf{P}_{1\#} = (\mathbf{X}_1 : \mathbf{0})(\mathbf{X}'\mathbf{W}^-\mathbf{X})^-\mathbf{X}'\mathbf{W}^+ = \mathbf{X}_1(\mathbf{X}_1'\dot{\mathbf{M}}_2\mathbf{X}_1)^-\mathbf{X}_1'\dot{\mathbf{M}}_2 = \mathbf{P}_{\mathbf{X}_1;\dot{\mathbf{M}}_2}\,,$$

and hence

$$\mathbf{P}_{1\#} = (\mathbf{X}_1 : \mathbf{0})\mathbf{X}^{\sim}, \text{ where } \mathbf{X}^{\sim} = (\mathbf{X}'\mathbf{W}^-\mathbf{X})^-\mathbf{X}'\mathbf{W}^+ \in \{\mathbf{X}^-\}\,,$$
$$\mathbf{P}_{\mathbf{X}_1;\dot{\mathbf{M}}_2} = \mathbf{X}_1\,\mathbf{X}_1^{\sim}, \qquad \text{where } \mathbf{X}_1^{\sim} = (\mathbf{X}_1'\dot{\mathbf{M}}_2\mathbf{X}_1)^-\mathbf{X}_1'\dot{\mathbf{M}}_2 \in \{\mathbf{X}_1^-\}\,.$$

■ **Solution 6:** Let $\mathbf{W} \in \mathscr{W}_{\geq}(\mathscr{M})$. Then it is clear that

$$\mathbf{G}(\mathbf{X} : \mathbf{V}\mathbf{M}) = (\mathbf{X} : \mathbf{0}) \iff \mathbf{G}(\mathbf{X} : \mathbf{W}\mathbf{M}) = (\mathbf{X} : \mathbf{0})\,.$$

Observing that $\mathscr{M}_{\mathbf{W}} = \{\mathbf{y}, \mathbf{X}\boldsymbol{\beta}, \mathbf{W}\}$ is a weakly singular linear model we can conclude, parallel to (18), that

$$\mathbf{X}(\mathbf{X}'\mathbf{W}^-\mathbf{X})^-\mathbf{X}'\mathbf{W}^-\mathbf{y} = \mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta} \mid \mathscr{M}_{\mathbf{W}}) = \mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta} \mid \mathscr{M})\,. \tag{30}$$

For (30) see also Christensen (2020, Th. 10.1.3).

■ **Solution 7:** (Pandora's Box.) Rao (1971, Th. 3.1) proved that the matrix $\mathbf{G}$ is a solution to the fundamental equation $\mathbf{G}(\mathbf{X} : \mathbf{V}\mathbf{M}) = (\mathbf{X} : \mathbf{0})$ if and only if there exists a matrix $\mathbf{L}$ such that $\mathbf{G}$ is a solution to

$$\begin{pmatrix} \mathbf{V} & \mathbf{X} \\ \mathbf{X}' & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{G}' \\ \mathbf{L} \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{X}' \end{pmatrix}.$$

Let us denote

$$\mathbf{\Gamma} = \begin{pmatrix} \mathbf{V} & \mathbf{X} \\ \mathbf{X}' & \mathbf{0} \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{21} & -\mathbf{C}_{22} \end{pmatrix} = \begin{pmatrix} \mathbf{V} & \mathbf{X} \\ \mathbf{X}' & \mathbf{0} \end{pmatrix}^{-} \in \{\mathbf{\Gamma}^{-}\},$$

so that $\mathbf{C}$ is a generalized inverse of $\mathbf{\Gamma}$. Rao (1971) showed that the matrix $\mathbf{C}$ is like a Pandora's Box, providing surprisingly many useful results concerning the model $\mathscr{M} = \{\mathbf{y}, \mathbf{X}\boldsymbol{\beta}, \mathbf{V}\}$. For example, denoting $\tilde{\boldsymbol{\mu}} = \mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta} \mid \mathscr{M})$, the following holds:

$$\tilde{\boldsymbol{\mu}} = \mathbf{X}\mathbf{C}_{12}'\mathbf{y}, \quad \mathrm{cov}(\tilde{\boldsymbol{\mu}}) = \mathbf{X}\mathbf{C}_{22}\mathbf{X}', \quad \tilde{\boldsymbol{\varepsilon}} = \mathbf{y} - \tilde{\boldsymbol{\mu}} = \mathbf{V}\mathbf{C}_{11}\mathbf{y}.$$

4. Solutions for BLUPs

Let us define the sets $\{\mathbf{P}_{\mathbf{y}_*|\mathscr{M}_*}\}$, $\{\mathbf{P}_{\mathbf{X}_*|\mathscr{M}_*}\}$, and $\{\mathbf{P}_{\boldsymbol{\varepsilon}_*|\mathscr{M}_*}\}$ as follows:

$$\mathbf{A} \in \{\mathbf{P}_{\mathbf{y}_*|\mathscr{M}_*}\} \iff \mathbf{A}(\mathbf{X} : \mathbf{V}\mathbf{M}) = (\mathbf{X}_* : \mathbf{V}_{21}\mathbf{M}), \tag{31a}$$

$$\mathbf{B} \in \{\mathbf{P}_{\mathbf{X}_*|\mathscr{M}_*}\} \iff \mathbf{B}(\mathbf{X} : \mathbf{V}\mathbf{M}) = (\mathbf{X}_* : \mathbf{0}), \tag{31b}$$

$$\mathbf{D} \in \{\mathbf{P}_{\boldsymbol{\varepsilon}_*|\mathscr{M}_*}\} \iff \mathbf{D}(\mathbf{X} : \mathbf{V}\mathbf{M}) = (\mathbf{0} : \mathbf{V}_{21}\mathbf{M}). \tag{31c}$$

Using (27), one solution to $\mathbf{A}(\mathbf{X} : \mathbf{V}\mathbf{M}) = (\mathbf{X}_* : \mathbf{V}_{21}\mathbf{M})$ can be written as

$$\mathbf{A}_1 = (\mathbf{X}_* : \mathbf{V}_{21}\mathbf{M})(\mathbf{X} : \mathbf{V}\mathbf{M})^{+} = \mathbf{B}_1 + \mathbf{V}_{21}(\mathbf{M}\mathbf{V}\mathbf{M})^{+},$$

where by (28), $\mathbf{B}_1 = \mathbf{X}_*(\mathbf{X}'\mathbf{W}^{+}\mathbf{X})^{+}\mathbf{X}'\mathbf{W}^{+}$. Putting (31b) and (31c) together yields

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

which implies that

$$(\mathbf{B} + \mathbf{D})(\mathbf{X} : \mathbf{V}\mathbf{M}) = (\mathbf{X}_* : \mathbf{V}_{21}\mathbf{M}),$$

and thereby $(\mathbf{B} + \mathbf{D})\mathbf{y}$ is a BLUP for $\mathbf{y}_*$ and we have the following result:

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

From part (c) of Theorem 2 we observe that $\mathbf{D}\mathbf{y}$ is the BLUP for $\boldsymbol{\varepsilon}_*$ if $\mathbf{D} = \mathbf{K}\mathbf{M}$ for some matrix $\mathbf{K} \in \mathbb{R}^{q\times n}$ such that $\mathbf{K}\mathbf{M}\mathbf{V}\mathbf{M} = \mathbf{V}_{21}\mathbf{M}$, from which one solution to $\mathbf{K}$ is $\mathbf{K} = \mathbf{V}_{21}\mathbf{M}(\mathbf{M}\mathbf{V}\mathbf{M})^{-}$ yielding the following expression:

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

Further representations, see Haslett *et al.* (2014, Th. 2), are

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

where $\mathbf{G} = \mathbf{X}(\mathbf{X}'\mathbf{W}^{-}\mathbf{X})^{-}\mathbf{X}'\mathbf{W}^{-}$. If $\mathbf{V}$ is positive definite and $\mathrm{r}(\mathbf{X}) = p$, we obtain

$$\mathrm{BLUP}(\mathbf{y}_*) = \mathrm{BLUE}(\mathbf{X}_*\boldsymbol{\beta}) + \mathrm{BLUP}(\boldsymbol{\varepsilon}_*) = \mathbf{X}_*\tilde{\boldsymbol{\beta}} + \mathbf{V}_{21}\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}) ,$$

where $\tilde{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y}$.

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

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

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 $\mathrm{E}(\mathbf{u}) = \mathbf{0}$, $\mathrm{cov}(\mathbf{u}) = \mathbf{D}$, $\mathrm{cov}(\mathbf{e}, \mathbf{u}) = \mathbf{0}$, and $\mathrm{E}(\mathbf{e}) = \mathbf{0}$, $\mathrm{cov}(\mathbf{e}) = \mathbf{R}$. In this situation we have

$$\mathrm{cov}\begin{pmatrix}\mathbf{e}\\ \mathbf{u}\end{pmatrix} = \begin{pmatrix}\mathbf{R} & \mathbf{0}\\ \mathbf{0} & \mathbf{D}\end{pmatrix} =: \boldsymbol{\Lambda} , \quad \mathrm{cov}\begin{pmatrix}\mathbf{y}\\ \mathbf{u}\end{pmatrix} = \begin{pmatrix}\boldsymbol{\Sigma} & \mathbf{ZD}\\ (\mathbf{ZD})' & \mathbf{D}\end{pmatrix} =: \boldsymbol{\Omega} .$$

The mixed model can be expressed as a version of the model with "new future observations", the new (unobservable) observations being, for example, in $\mathbf{u} = \mathbf{0}\boldsymbol{\beta} + \boldsymbol{\varepsilon}_*$:

$$\mathscr{L}_* := \left\{ \begin{pmatrix}\mathbf{y}\\ \mathbf{u}\end{pmatrix}, \begin{pmatrix}\mathbf{X}\\ \mathbf{0}\end{pmatrix}\boldsymbol{\beta}, \begin{pmatrix}\boldsymbol{\Sigma} & \mathbf{ZD}\\ (\mathbf{ZD})' & \mathbf{D}\end{pmatrix} \right\} . \tag{32}$$

Corresponding to (1) we have

$$\begin{aligned}
&\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} , \quad \boldsymbol{\varepsilon} = \mathbf{Z}\mathbf{u} + \mathbf{e} , \quad \mathrm{cov}(\boldsymbol{\varepsilon}) = \mathrm{cov}(\mathbf{y}) = \mathbf{ZDZ}' + \mathbf{R} =: \boldsymbol{\Sigma} ,\\
&\mathbf{y}_* = \mathbf{u} , \quad \mathbf{X}_* = \mathbf{0} ,\\
&\boldsymbol{\varepsilon}_* = \mathbf{u} , \quad \mathrm{cov}(\boldsymbol{\varepsilon}_*) = \mathbf{D}, \quad \mathrm{cov}(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}_*) = \mathbf{ZD}.
\end{aligned}$$

Now under the mixed model $\mathscr{L}$, $\mathbf{B}_1\mathbf{y}$ is the BLUE for $\boldsymbol{\mu} = \mathbf{X}\boldsymbol{\beta}$ and $\mathbf{B}_2\mathbf{y}$ is the BLUP for $\mathbf{u}$ if and only if

$$\begin{pmatrix}\mathbf{B}_1\\ \mathbf{B}_2\end{pmatrix}(\mathbf{X} : \boldsymbol{\Sigma}\mathbf{M}) = \begin{pmatrix}\mathbf{X} & \mathbf{0}\\ \mathbf{0} & (\mathbf{ZD})'\mathbf{M}\end{pmatrix} . \tag{34}$$

Thus the $\mathrm{BLUP}(\mathbf{u})$ can be written as

$$\mathrm{BLUP}(\mathbf{u}) = \mathbf{DZ}'\mathbf{W}^{-}(\mathbf{y} - \tilde{\boldsymbol{\mu}}) = \mathbf{DZ}'\mathbf{M}(\mathbf{M}\boldsymbol{\Sigma}\mathbf{M})^{-}\mathbf{M}\mathbf{y},$$

where $\mathbf{W} = \boldsymbol{\Sigma} + \mathbf{XX}'$. For example, in the simple situation when $\mathbf{X}$ has full column rank and $\boldsymbol{\Sigma} = \mathbf{ZDZ}' + \mathbf{R}$ is positive definite, we have

$$\mathrm{BLUP}(\mathbf{u}) = \mathbf{DZ}'\boldsymbol{\Sigma}^{-1}(\mathbf{I}_n - \mathbf{X}\tilde{\boldsymbol{\beta}}) , \quad \tilde{\boldsymbol{\beta}} = (\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{y}.$$

Remark 2: We can write up the mixed model (32) as

$$\begin{pmatrix}\mathbf{y}\\ \mathbf{u}\end{pmatrix} = \begin{pmatrix}\mathbf{X}\\ \mathbf{0}\end{pmatrix}\boldsymbol{\beta} + \begin{pmatrix}\boldsymbol{\varepsilon}\\ \boldsymbol{\varepsilon}_*\end{pmatrix}, \quad \text{where } \mathrm{cov}\begin{pmatrix}\boldsymbol{\varepsilon}\\ \boldsymbol{\varepsilon}_*\end{pmatrix} = \boldsymbol{\Omega} . \tag{35}$$

It noteworthy that even as (35) looks like a standard linear model it is not quite correct: the random vector $\mathbf{u}$ is *unobservable*. On the other hand, keeping $\mathbf{u}$ fixed (but unknown) and denoting $\mathbf{y}_0 = \mathbf{u} + \boldsymbol{\varepsilon}_*$ we get a fixed partitioned model with supplemented stochastic restrictions on $\mathbf{u}$:

$$\mathscr{F} := \left\{ \begin{pmatrix} \mathbf{y} \\ \mathbf{y}_0 \end{pmatrix}, \begin{pmatrix} \mathbf{X} & \mathbf{Z} \\ \mathbf{0} & \mathbf{I}_q \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta} \\ \mathbf{u} \end{pmatrix}, \begin{pmatrix} \mathbf{R} & \mathbf{0} \\ \mathbf{0} & \mathbf{D} \end{pmatrix} \right\}.$$

We get an interesting version of $\mathscr{F}$ by putting $\mathbf{y}_0 = \mathbf{0}$:

$$\mathscr{F}_\# := \left\{ \begin{pmatrix} \mathbf{y} \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \mathbf{X} & \mathbf{Z} \\ \mathbf{0} & \mathbf{I}_q \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta} \\ \mathbf{u} \end{pmatrix}, \begin{pmatrix} \mathbf{R} & \mathbf{0} \\ \mathbf{0} & \mathbf{D} \end{pmatrix} \right\} = \left\{ \mathbf{y}_\#,\, \mathbf{X}_\# \boldsymbol{\pi},\, \boldsymbol{\Lambda} \right\}.$$

Of course $\mathscr{F}_\#$ is not a proper model since $\mathbf{y}_0 = \mathbf{0}$. In the full rank case fitting the model $\mathscr{F}_\#$ yields to so-called Henderson equations and the BLUE of $\mathbf{X}\boldsymbol{\beta}$ and BLUP of $\mathbf{u}$ are obtained by minimizing the following quadratic form $f(\boldsymbol{\beta}, \mathbf{u})$ (keeping $\mathbf{u}$ as a non-random vector):

$$f(\boldsymbol{\beta}, \mathbf{u}) = (\mathbf{y}_\# - \mathbf{X}_\# \boldsymbol{\pi})' \boldsymbol{\Lambda}^{-1} (\mathbf{y}_\# - \mathbf{X}_\# \boldsymbol{\pi}).$$

For further references, see, *e.g.*, Henderson (1950, 1963) and Haslett *et al.* (2015). □

5. Two models with different covariance matrices

Suppose that we have two models $\mathscr{M}(\mathbf{V}_0) = \{\mathbf{y}, \mathbf{X}\boldsymbol{\beta}, \mathbf{V}_0\}$ and $\mathscr{M}(\mathbf{V}) = \{\mathbf{y}, \mathbf{X}\boldsymbol{\beta}, \mathbf{V}\}$, which have different covariance matrices. Then we can ask, for example, what is needed that every representation of the BLUE of $\boldsymbol{\mu} = \mathbf{X}\boldsymbol{\beta}$ under $\mathscr{M}(\mathbf{V}_0)$ remains BLUE under $\mathscr{M}(\mathbf{V})$. Mitra and Moore (1973, p. 139) give a very clear description of the different problems occurring:

(a) Problem MM-1: When is specific linear representation of the BLUE of $\mathbf{X}\boldsymbol{\beta}$ under $\mathscr{M}(\mathbf{V}_0)$ also BLUE under $\mathscr{M}(\mathbf{V})$?

(b) Problem MM-2: When does $\mathbf{X}\boldsymbol{\beta}$ have a common representation for the BLUE under $\mathscr{M}(\mathbf{V}_0)$ and $\mathscr{M}(\mathbf{V})$?

(c) Problem MM-3: When does every linear representation of the BLUE of $\mathbf{X}\boldsymbol{\beta}$ under $\mathscr{M}(\mathbf{V}_0)$ remain BLUE also under $\mathscr{M}(\mathbf{V})$?

As for MM-1, we may mention that Hauke *et al.* (2013) consider conditions under which

$$\mathbf{P}_{\mathbf{X};\mathbf{W}_0^+} \mathbf{y} = \mathbf{X}(\mathbf{X}'\mathbf{W}_0^- \mathbf{X})^- \mathbf{X}' \mathbf{W}_0^+ \mathbf{y} = \mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta} \mid \mathscr{M}(\mathbf{V})). \tag{36}$$

This happens if and only if $\mathbf{P}_{\mathbf{X};\mathbf{W}_0^+} \mathbf{V}\mathbf{M} = \mathbf{0}$, which further is equivalent to

$$\mathbf{X}'\mathbf{W}_0^+ \mathbf{V}\mathbf{M} = \mathbf{0}, \quad \text{i.e.,} \quad \mathscr{C}(\mathbf{V}\mathbf{M}) \subseteq \mathscr{C}(\mathbf{W}_0^+ \mathbf{X})^\perp = \mathscr{C}(\mathbf{V}_0 \mathbf{M} : \mathbf{Q}_{\mathbf{W}_0}),$$

where we have used Proposition 4. Denoting $\mathbf{Z} = (\mathbf{V}_0 \mathbf{M} : \mathbf{Q}_{\mathbf{W}_0})$, Hauke *et al.* (2013) showed that (36) holds if and only if $\mathbf{V}$ belongs to the class $\mathscr{V}_{mm1}$, say, defined as

$$\mathbf{V} \in \mathscr{V}_{mm1} \iff \mathbf{V} = \mathbf{X}\mathbf{A}\mathbf{A}'\mathbf{X}' + \mathbf{Z}\mathbf{B}\mathbf{B}'\mathbf{Z}' \quad \text{for some matrices } \mathbf{A} \text{ and } \mathbf{B}. \tag{37}$$

Let us take a closer look at MMM-3 in the spirit of Puntanen *et al.* (2011, Sec. 11.1). First, let us denote

$$\mathbf{G} \in \{\mathbf{P}_{\boldsymbol{\mu}\,|\,\mathbf{V}_0}\} \iff \mathbf{G}(\mathbf{X} : \mathbf{V}_0\mathbf{M}) = (\mathbf{X} : \mathbf{0})\,.$$

Let $\mathbf{G}$ be such a matrix that $\mathbf{Gy}$ is the BLUE for $\mathbf{X}\boldsymbol{\beta}$ under $\mathscr{M}(\mathbf{V}_0)$, Then we say that $\mathbf{Gy}$ remains BLUE under $\mathscr{M}(\mathbf{V})$ if the following implication holds:

$$\mathbf{G}(\mathbf{X} : \mathbf{V}_0\mathbf{M}) = (\mathbf{X} : \mathbf{0}) \implies \mathbf{G}(\mathbf{X} : \mathbf{VM}) = (\mathbf{X} : \mathbf{0})\,.$$

Moreover, let the set of all representations of BLUE of $\boldsymbol{\mu}$ under $\mathscr{M}(\mathbf{V}_0)$ be denoted as

$$\begin{aligned}\mathscr{B}(\boldsymbol{\mu}\,|\,\mathbf{V}_0) &= \{\mathrm{BLUE}(\boldsymbol{\mu}\,|\,\mathbf{V}_0)\} = \{\mathbf{Gy} : \mathbf{G}(\mathbf{X} : \mathbf{V}_0\mathbf{M}) = (\mathbf{X} : \mathbf{0})\} \\ &= \{\mathbf{Gy} : \mathbf{G} \in \{\mathbf{P}_{\boldsymbol{\mu}\,|\,\mathbf{V}_0}\}\}\,. \end{aligned} \tag{38}$$

It is important to understand that the notation of the above type (38) is merely symbolic. Our main interest lies in the *multipliers*, like the members of $\{\mathbf{P}_{\boldsymbol{\mu}\,|\,\mathbf{V}_0}\}$, of the response vector $\mathbf{y}$ which have specific properties. For the property that *every* representation of the BLUE of $\boldsymbol{\mu}$ under $\mathscr{M}(\mathbf{V}_0)$ remains BLUE of $\boldsymbol{\mu}$ under $\mathscr{M}(\mathbf{V})$ we will use the notation

$$\mathscr{B}(\boldsymbol{\mu}\,|\,\mathbf{V}_0) \subseteq \mathscr{B}(\boldsymbol{\mu}\,|\,\mathbf{V})\,, \quad \text{i.e.,} \quad \{\mathbf{P}_{\boldsymbol{\mu}\,|\,\mathbf{V}_0}\} \subseteq \{\mathbf{P}_{\boldsymbol{\mu}\,|\,\mathbf{V}}\}\,. \tag{39}$$

We may consider $\mathscr{M}(\mathbf{V}_0)$ as the original model and $\mathscr{M}(\mathbf{V})$ as the misspecified model; misspecification concerning only the covariance matrix.

Let us next show that (39) is equivalent to

$$\mathscr{C}(\mathbf{VM}) \subseteq \mathscr{C}(\mathbf{V}_0\mathbf{M})\,, \tag{40}$$

which is essentially Rao's result in Theorem 5.3 of his paper in 1971. This is a well-known old but yet a fundamental result whose proof is worth going through. Proceeding as Puntanen *et al.* (2011, p. 270), we observe that a general representation of a member in $\{\mathbf{P}_{\boldsymbol{\mu}\,|\,\mathbf{V}_0}\}$ can be expressed as

$$\mathbf{G}_0 = \mathbf{P}_{\mathbf{X};\mathbf{W}_0^+} + \mathbf{EQ}_{\mathbf{W}_0} = \mathbf{X}(\mathbf{X}'\mathbf{W}_0^-\mathbf{X})^-\mathbf{X}'\mathbf{W}_0^+ + \mathbf{E}(\mathbf{I}_n - \mathbf{P}_{\mathbf{W}_0})\,,$$

where $\mathbf{E}$ is free to vary and $\mathbf{W}_0 \in \mathscr{W}(\mathbf{V}_0)$. Now (39) holds if and only if

$$\mathbf{G}_0(\mathbf{X} : \mathbf{VM}) = (\mathbf{P}_{\mathbf{X};\mathbf{W}_0^+} + \mathbf{EQ}_{\mathbf{W}_0})(\mathbf{X} : \mathbf{VM}) = (\mathbf{X} : \mathbf{0})\,,$$

i.e.,

$$\mathbf{P}_{\mathbf{X};\mathbf{W}_0^+}\mathbf{VM} + \mathbf{EQ}_{\mathbf{W}_0}\mathbf{VM} = \mathbf{0} \quad \text{for all } \mathbf{E}, \tag{41}$$

which implies that $\mathbf{Q}_{\mathbf{W}_0}\mathbf{VM} = \mathbf{0}$, *i.e.*, $\mathscr{C}(\mathbf{VM}) \subseteq \mathscr{C}(\mathbf{W}_0) = \mathscr{C}(\mathbf{X} : \mathbf{V}_0\mathbf{M})$, which further means that

$$\mathbf{VM} = \mathbf{XR} + \mathbf{V}_0\mathbf{MS} \quad \text{for some } \mathbf{R} \text{ and } \mathbf{S}. \tag{42}$$

Substituting (42) into (41) shows that $\mathbf{XR} = \mathbf{0}$ and thereby (39) implies (40). The reverse relation is easy to check. It is worth noting that

$$\mathscr{C}(\mathbf{VM}) \subseteq \mathscr{C}(\mathbf{V}_0\mathbf{M}) \implies \mathscr{C}(\mathbf{X} : \mathbf{VM}) \subseteq \mathscr{C}(\mathbf{X} : \mathbf{V}_0\mathbf{M})$$

but the reverse implication does not hold.

Remark 3: Let us consider conditions under which

$$\mathbf{P}_{\mathbf{X};\mathbf{W}_0^-}\mathbf{y} = \mathbf{X}(\mathbf{X}'\mathbf{W}_0^-\mathbf{X})^-\mathbf{X}'\mathbf{W}_0^-\mathbf{y} = \mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta} \mid \mathscr{M}(\mathbf{V})) \ \text{ for all } \mathbf{W}_0^- \,, \tag{43}$$

i.e.,

$$\mathbf{X}'\mathbf{W}_0^-\mathbf{V}\mathbf{M} = \mathbf{0} \ \text{ for all } \mathbf{W}_0^- \,. \tag{44}$$

Now in view of Proposition 1, (44) holds if and only if

$$\mathbf{X}'\mathbf{W}_0^+\mathbf{V}\mathbf{M} = \mathbf{0} \ \text{ and } \ \mathscr{C}(\mathbf{V}\mathbf{M}) \subseteq \mathscr{C}(\mathbf{W}_0)\,, \tag{45}$$

i.e.,

$$\mathscr{C}(\mathbf{V}\mathbf{M}) \subseteq \mathscr{C}(\mathbf{W}_0^+\mathbf{X})^\perp = \mathscr{C}(\mathbf{V}_0\mathbf{M} : \mathbf{Q}_{\mathbf{W}_0}) \ \text{ and } \ \mathscr{C}(\mathbf{V}\mathbf{M}) \subseteq \mathscr{C}(\mathbf{W}_0)\,, \tag{46}$$

which together imply (40). □

It is clear that $\boldsymbol{\mu} = \mathbf{X}\boldsymbol{\beta}$ has a common representation for the BLUE under $\mathscr{M}(\mathbf{V}_0)$ and $\mathscr{M}(\mathbf{V})$, *i.e.*, $\{\mathbf{P}_{\boldsymbol{\mu}\,|\,\mathbf{V}_0}\} \cap \{\mathbf{P}_{\boldsymbol{\mu}\,|\,\mathbf{V}}\} \neq \{\emptyset\}$, if and only if the equation

$$\mathbf{G}(\mathbf{X} : \mathbf{V}_0\mathbf{M} : \mathbf{V}\mathbf{M}) = (\mathbf{X} : \mathbf{0} : \mathbf{0})$$

has a solution for $\mathbf{G}$, *i.e.*,

$$\mathscr{C}[(\mathbf{X} : \mathbf{0} : \mathbf{0})'] \subseteq \mathscr{C}[(\mathbf{X} : \mathbf{V}_0\mathbf{M} : \mathbf{V}\mathbf{M})']\,, \tag{47}$$

for which, according to Mitra and Moore (1973, Sec. 3), it is necessary and sufficient that

$$\mathscr{C}(\mathbf{V}_0\mathbf{M} : \mathbf{V}\mathbf{M}) \cap \mathscr{C}(\mathbf{X}) = \{\mathbf{0}\}\,.$$

Suppose that (47) holds. Given $\mathbf{V}_0$, how can we then characterize the class $\mathscr{V}_{mm2}$, say, of matrices $\mathbf{V}$ such that $\{\mathbf{P}_{\boldsymbol{\mu}\,|\,\mathbf{V}_0}\}$ and $\{\mathbf{P}_{\boldsymbol{\mu}\,|\,\mathbf{V}}\}$ are not disjoint? Mitra and Moore (1973, Sec. 3) showed that $\mathscr{V}_{mm2} = \mathscr{V}_{mm1}$ so that

$$\mathbf{V} \in \mathscr{V}_{mm2} \iff \mathbf{V} = \mathbf{X}\mathbf{A}\mathbf{A}'\mathbf{X}' + (\mathbf{V}_0\mathbf{M} : \mathbf{Q}_{\mathbf{W}_0}) \begin{pmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \end{pmatrix} (\mathbf{B}_1' : \mathbf{B}_2') \begin{pmatrix} \mathbf{M}\mathbf{V}_0 \\ \mathbf{Q}_{\mathbf{W}_0} \end{pmatrix}, \tag{48}$$

for some matrices $\mathbf{A}, \mathbf{B}_1$ and $\mathbf{B}_2$.

Let us next consider the following task: Given a covariance matrix $\mathbf{V}_0$, characterize the set $\mathscr{V}$ of covariance matrices such that every representation of the BLUE of $\mathbf{X}\boldsymbol{\beta}$ under $\mathscr{M}(\mathbf{V}_0)$ remains BLUE under $\mathscr{M}(\mathbf{V})$, *i.e.*,

$$\mathbf{V} \in \mathscr{V} \iff \mathscr{B}(\boldsymbol{\mu}\,|\,\mathbf{V}_0) \subseteq \mathscr{B}(\boldsymbol{\mu}\,|\,\mathbf{V})\,.$$

We will next show that a necessary condition for $\mathbf{V} \in \mathscr{V}$ is the following:

$$\mathbf{V} = \mathbf{X}\mathbf{A}\mathbf{A}'\mathbf{X}' + \mathbf{V}_0\mathbf{M}\mathbf{B}\mathbf{B}'\mathbf{M}\mathbf{V}_0 \ \text{ for some matrices } \mathbf{A} \text{ and } \mathbf{B}. \tag{49}$$

This is also given by Rao (1971, Th. 5.3) but we will give a slightly different proof. Notice that class $\mathscr{V}_{mm2}$ in (48) is wider than class $\mathscr{V}$ defined in (49).

Since $\mathscr{C}(\mathbf{X} : \mathbf{V}_0\mathbf{M} : \mathbf{Q}_{\mathbf{W}_0}) = \mathbb{R}^n$, where $\mathbf{W}_0 \in \mathscr{W}(\mathbf{V}_0)$, an arbitrary nonnegative definite matrix $\mathbf{V}$ can be expressed as $\mathbf{V} = \mathbf{U}\mathbf{U}'$ where

$$\mathbf{U} = \mathbf{X}\mathbf{L}_1 + \mathbf{V}_0\mathbf{M}\mathbf{L}_2 + \mathbf{Q}_{\mathbf{W}_0}\mathbf{L}_3\,,$$

for some matrices $\mathbf{L}_1, \mathbf{L}_2, \mathbf{L}_3$, so that

$$\mathbf{V} = \mathbf{X}\mathbf{L}_{11}\mathbf{X}' + \mathbf{V}_0\mathbf{M}\mathbf{L}_{22}\mathbf{M}\mathbf{V}_0 + \mathbf{Q}_{\mathbf{W}_0}\mathbf{L}_{33}\mathbf{Q}_{\mathbf{W}_0} + \mathbf{N} + \mathbf{N}', \tag{50}$$

where $\mathbf{L}_{ij} = \mathbf{L}_i\mathbf{L}_j'$, $j = 1, 2, 3$, and

$$\mathbf{N} = \mathbf{X}\mathbf{L}_{12}\mathbf{M}\mathbf{V}_0 + \mathbf{X}\mathbf{L}_{13}\mathbf{Q}_{\mathbf{W}_0} + \mathbf{V}_0\mathbf{M}\mathbf{L}_{23}\mathbf{Q}_{\mathbf{W}_0}\,.$$

Now

$$\mathbf{U}'\mathbf{M} = \mathbf{L}_2'\mathbf{M}\mathbf{V}_0\mathbf{M} + \mathbf{L}_3'\mathbf{Q}_{\mathbf{W}_0}\mathbf{M} = \mathbf{L}_2'\mathbf{M}\mathbf{V}_0\mathbf{M} + \mathbf{L}_3'\mathbf{Q}_{\mathbf{W}_0} =: \mathbf{S}\,,$$

where $\mathbf{Q}_{\mathbf{W}_0}\mathbf{M} = \mathbf{Q}_{\mathbf{W}_0}$ follows from part (d) of Proposition 2. Moreover,

$$\mathscr{C}(\mathbf{U}\mathbf{U}'\mathbf{M}) = \mathscr{C}(\mathbf{X}\mathbf{L}_1\mathbf{S} + \mathbf{V}_0\mathbf{M}\mathbf{L}_2\mathbf{S} + \mathbf{Q}_{\mathbf{W}_0}\mathbf{L}_3\mathbf{S}) \subseteq \mathscr{C}(\mathbf{V}_0\mathbf{M})$$

holds if and only if

$$\mathscr{C}(\mathbf{X}\mathbf{L}_1\mathbf{S} + \mathbf{Q}_{\mathbf{W}_0}\mathbf{L}_3\mathbf{S}) \subseteq \mathscr{C}(\mathbf{V}_0\mathbf{M})\,. \tag{51}$$

Premultiplying (51) by $\mathbf{Q}_{\mathbf{W}_0}$ shows that $\mathbf{Q}_{\mathbf{W}_0}\mathbf{L}_3\mathbf{S} = \mathbf{0}$, *i.e.*,

$$\mathbf{Q}_{\mathbf{W}_0}\mathbf{L}_{32}\mathbf{M}\mathbf{V}_0\mathbf{M} + \mathbf{Q}_{\mathbf{W}_0}\mathbf{L}_{33}\mathbf{Q}_{\mathbf{W}_0} = \mathbf{0}\,. \tag{52}$$

Postmultiplying (52) by $\mathbf{Q}_{\mathbf{W}_0}$ implies that $\mathbf{Q}_{\mathbf{W}_0}\mathbf{L}_{33}\mathbf{Q}_{\mathbf{W}_0} = \mathbf{0}$, *i.e.*,

$$\mathbf{L}_3\mathbf{Q}_{\mathbf{W}_0} = \mathbf{0}\,. \tag{53}$$

Substituting (53) into (51) yields

$$\mathscr{C}(\mathbf{X}\mathbf{L}_1\mathbf{S}) \subseteq \mathscr{C}(\mathbf{V}_0\mathbf{M})\,. \tag{54}$$

The disjointness of $\mathscr{C}(\mathbf{X})$ and $\mathscr{C}(\mathbf{V}_0\mathbf{M})$ implies that (54) holds if and only if

$$\mathbf{X}\mathbf{L}_1\mathbf{S} = \mathbf{X}\mathbf{L}_1(\mathbf{L}_2'\mathbf{M}\mathbf{V}_0\mathbf{M} + \mathbf{L}_3'\mathbf{Q}_{\mathbf{W}_0}) = \mathbf{0}\,,$$

which further is equivalent to

$$\mathbf{X}\mathbf{L}_{12}\mathbf{M}\mathbf{V}_0 = \mathbf{0}\,. \tag{55}$$

Substituting (53) and (55) into (50) proves that (49) is a necessary condition for $\mathbf{V} \in \mathscr{V}$. Its sufficiency is obvious.

Some equivalent statements to (39) are given as follows.

Proposition 6: Consider the linear models $\mathscr{M}(\mathbf{V}_0) = \{\mathbf{y}, \mathbf{X}\boldsymbol{\beta}, \mathbf{V}_0\}$ and $\mathscr{M}(\mathbf{V}) = \{\mathbf{y}, \mathbf{X}\boldsymbol{\beta}, \mathbf{V}\}$. Then the following statements are equivalent:

(a) $\mathscr{B}(\boldsymbol{\mu} \,|\, \mathbf{V}_0) \subseteq \mathscr{B}(\boldsymbol{\mu} \,|\, \mathbf{V})$, *i.e.*, $\mathbf{V} \in \mathscr{V}$.

(b) $\mathscr{C}(\mathbf{V}\mathbf{M}) \subseteq \mathscr{C}(\mathbf{V}_0\mathbf{M})$.

(c) $\mathbf{V} = \mathbf{X}\mathbf{A}\mathbf{A}'\mathbf{X}' + \mathbf{V}_0\mathbf{M}\mathbf{B}\mathbf{B}'\mathbf{M}\mathbf{V}_0$, for some matrices $\mathbf{A}$ and $\mathbf{B}$.

(d) $\mathbf{V} = \mathbf{V}_0 + \mathbf{X}\mathbf{C}\mathbf{C}'\mathbf{X}' + \mathbf{V}_0\mathbf{M}\mathbf{D}\mathbf{D}'\mathbf{M}\mathbf{V}_0$, for some matrices $\mathbf{C}$ and $\mathbf{D}$.

For the proof of Proposition 6 and related discussion, see, *e.g.*, Mitra and Moore (1973, Th. 4.1–4.2), Rao (1968, Lemma 5), Rao (1971, Th. 5.2, Th. 5.5), Rao (1973, p. 289), and Baksalary and Mathew (1986, Th. 3).

Consider then the special case when we have models $\mathscr{M}(\mathbf{I}) = \{\mathbf{y}, \mathbf{X}\boldsymbol{\beta}, \mathbf{I}\}$ and $\mathscr{M}(\mathbf{V}) = \{\mathbf{y}, \mathbf{X}\boldsymbol{\beta}, \mathbf{V}\}$. Then the BLUE of $\mathbf{X}\boldsymbol{\beta}$ under $\mathscr{M}(\mathbf{I})$ is $\mathbf{P}_{\mathbf{X}}\mathbf{y} = \mathbf{H}\mathbf{y}$ since the unique solution for $\mathbf{G}$ in $\mathbf{G}(\mathbf{X} : \mathbf{M}) = (\mathbf{X} : \mathbf{0})$ is $\mathbf{H}$. When is $\mathbf{H}\mathbf{y}$, *i.e.*, the ordinary least squares estimator (OLSE) BLUE for $\mathbf{X}\boldsymbol{\beta}$ under $\mathscr{M}(\mathbf{V})$? The answer is by part (b) of Proposition (6) the inclusion $\mathscr{C}(\mathbf{V}\mathbf{M}) \subseteq \mathscr{C}(\mathbf{M})$, which can be equivalently expressed as any of the following conditions:

$$\mathscr{C}(\mathbf{V}\mathbf{H}) \subseteq \mathscr{C}(\mathbf{H}), \quad \mathbf{H}\mathbf{V} = \mathbf{V}\mathbf{H}, \quad \mathbf{H}\mathbf{V}\mathbf{M} = \mathbf{0}\,.$$

For further references regarding the equality of OLSE and BLUE, see, *e.g.*, Rao (1967), Zyskind (1967), and Markiewicz *et al.* (2010, 2021).

Let $\mathscr{V}_{1/12}$ denote the set of nonnegative definite matrices $\mathbf{V}$ such that every representation of the BLUE of $\boldsymbol{\mu}_1$ under $\mathscr{M}(\mathbf{V}_0)$ remains BLUE under $\mathscr{M}(\mathbf{V})$, *i.e.*,

$$\mathbf{V} \in \mathscr{V}_{1/12} \iff \mathscr{B}(\boldsymbol{\mu}_1 \,|\, \mathbf{V}_0) \subseteq \mathscr{B}(\boldsymbol{\mu}_1 \,|\, \mathbf{V})\,.$$

In view of Haslett and Puntanen (2010a, Th. 2.1, 2023b, Th. 11.4), see also Mathew and Bhimasankaram (1983, Th. 2.1, Th. 2.4), the following holds:

Proposition 7: Consider the partitioned linear models $\mathscr{M}(\mathbf{V}_0)$ and $\mathscr{M}(\mathbf{V})$, where $\boldsymbol{\mu}_1 = \mathbf{X}_1\boldsymbol{\beta}_1$ is estimable. Then the following statements are equivalent:

(a) $\mathscr{B}(\boldsymbol{\mu}_1 \,|\, \mathbf{V}_0) \subseteq \mathscr{B}(\boldsymbol{\mu}_1 \,|\, \mathbf{V})$, *i.e.*, $\mathbf{V} \in \mathscr{V}_{1/12}$.

(b) $\mathscr{C}(\mathbf{M}_2\mathbf{V}\mathbf{M}) \subseteq \mathscr{C}(\mathbf{M}_2\mathbf{V}_0\mathbf{M})$.

(c) $\mathscr{C}(\mathbf{V}\mathbf{M}) \subseteq \mathscr{C}(\mathbf{X}_2 : \mathbf{V}_0\mathbf{M})$.

(d) The matrix $\mathbf{V}$ can be expressed, for some $\mathbf{L}_i$, $\mathbf{L}_{ij} = \mathbf{L}_i\mathbf{L}_j'$, as

$$\mathbf{V} = \mathbf{X}_1\mathbf{L}_{11}\mathbf{X}_1' + \mathbf{X}_2\mathbf{L}_{22}\mathbf{X}_2' + \mathbf{V}_0\mathbf{M}\mathbf{L}_{33}\mathbf{M}\mathbf{V}_0 + \mathbf{Z} + \mathbf{Z}',$$

where $\mathbf{Z} = \mathbf{X}_1\mathbf{L}_{12}\mathbf{X}_2' + \mathbf{X}_2\mathbf{L}_{23}\mathbf{M}\mathbf{V}_0$.

So far in this section we have been dealing with linear models $\mathscr{M}(\mathbf{V}_0) = \{\mathbf{y}, \mathbf{X}\boldsymbol{\beta}, \mathbf{V}_0\}$ and $\mathscr{M}(\mathbf{V}) = \{\mathbf{y}, \mathbf{X}\boldsymbol{\beta}, \mathbf{V}\}$. The corresponding considerations can be done for the two models with new future observations. For this purpose, denote

$$\mathscr{A}_1 = \left\{ \begin{pmatrix} \mathbf{y} \\ \mathbf{y}_* \end{pmatrix}, \begin{pmatrix} \mathbf{X} \\ \mathbf{X}_* \end{pmatrix} \boldsymbol{\beta}, \begin{pmatrix} \mathbf{V}_{11} & \mathbf{V}_{12} \\ \mathbf{V}_{21} & \mathbf{V}_{22} \end{pmatrix} \right\},$$

where $\mathscr{C}(\mathbf{X}_*') \subseteq \mathscr{C}(\mathbf{X}')$. Consider now another model $\mathscr{A}_2$, which may differ from $\mathscr{A}_1$ through its covariance matrix, *i.e.*,

$$\mathscr{A}_2 = \left\{ \begin{pmatrix} \mathbf{y} \\ \mathbf{y}_* \end{pmatrix}, \begin{pmatrix} \mathbf{X} \\ \mathbf{X}_* \end{pmatrix} \boldsymbol{\beta}, \begin{pmatrix} \underline{\mathbf{V}}_{11} & \underline{\mathbf{V}}_{12} \\ \underline{\mathbf{V}}_{21} & \underline{\mathbf{V}}_{22} \end{pmatrix} \right\}.$$

For the proof of the following result see Haslett and Puntanen (2010b).

Proposition 8: Consider the models $\mathscr{A}_1$ and $\mathscr{A}_2$ (with new unobserved future observations), where $\mathscr{C}(\mathbf{X}_*') \subseteq \mathscr{C}(\mathbf{X}')$. Then every representation of the BLUP for $\mathbf{y}_*$ under the model $\mathscr{A}_1$ is also a BLUP for $\mathbf{y}_*$ under the model $\mathscr{A}_2$ if and only if

$$\mathscr{C} \begin{pmatrix} \underline{\mathbf{V}}_{11}\mathbf{M} \\ \underline{\mathbf{V}}_{21}\mathbf{M} \end{pmatrix} \subseteq \mathscr{C} \begin{pmatrix} \mathbf{X} & \mathbf{V}_{11}\mathbf{M} \\ \mathbf{X}_* & \mathbf{V}_{21}\mathbf{M} \end{pmatrix}.$$

Consider then two mixed models:

$$\mathscr{B}_1 = \{\mathbf{y},\, \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u},\, \mathbf{D}_1\,, \mathbf{R}_1\}\,, \quad \mathscr{B}_2 = \{\mathbf{y},\, \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u},\, \mathbf{D}_2,\, \mathbf{R}_2\}\,.$$

The only difference above concerns the covariance matrices. We may denote $\boldsymbol{\Sigma}_i = \mathbf{Z}\mathbf{D}_i\mathbf{Z}' + \mathbf{R}_i\,,\ i = 1, 2\,$. For the next proposition, see Haslett and Puntanen (2011).

Proposition 9: Consider the mixed models $\mathscr{B}_1$ and $\mathscr{B}_2$. Then every representation of the BLUP for $\mathbf{u}$ under the model $\mathscr{B}_1$ is also the BLUP for $\mathbf{u}$ under the model $\mathscr{B}_2$ if and only if

$$\mathscr{C} \begin{pmatrix} \boldsymbol{\Sigma}_2\mathbf{M} \\ \mathbf{D}_2\mathbf{Z}'\mathbf{M} \end{pmatrix} \subseteq \mathscr{C} \begin{pmatrix} \mathbf{X} & \boldsymbol{\Sigma}_1\mathbf{M} \\ \mathbf{0} & \mathbf{D}_1\mathbf{Z}'\mathbf{M} \end{pmatrix}.$$

In particular, both the BLUE($\mathbf{X}\boldsymbol{\beta}$) under $\mathscr{B}_1$ continues to be BLUE($\mathbf{X}\boldsymbol{\beta}$) under $\mathscr{B}_2$ and BLUP($\mathbf{u}$) under $\mathscr{B}_1$ continues to be BLUP($\mathbf{u}$) under $\mathscr{B}_2$ if and only if

$$\mathscr{C} \begin{pmatrix} \boldsymbol{\Sigma}_2\mathbf{M} \\ \mathbf{D}_2\mathbf{Z}'\mathbf{M} \end{pmatrix} \subseteq \mathscr{C} \begin{pmatrix} \boldsymbol{\Sigma}_1\mathbf{M} \\ \mathbf{D}_1\mathbf{Z}'\mathbf{M} \end{pmatrix}.$$

6. Further remarks

In this section we very briefly review some recent articles by the authors. Fundamental BLUE/BLUP equations have instrumental role in these papers.

[**A**] Haslett *et al.* (2023), [**B**] Haslett *et al.* (2020).

In these articles we consider the partitioned linear model $\mathscr{M}_{12}$, and the corresponding small model $\mathscr{M}_1$. We focus on comparing the BLUEs of $\boldsymbol{\mu}_1$ under $\mathscr{M}_{12}$ and $\mathscr{M}_1$. Particular attention is paid on the consistency of the model, *i.e.*, whether the realized value of the response vector $\mathbf{y}$ belongs to the column space of $(\mathbf{X}_1 : \mathbf{V})$ or $(\mathbf{X}_1 : \mathbf{X}_2 : \mathbf{V})$. In [A] these models are supplemented with the new unobservable random vector $\mathbf{y}_*$, coming from $\mathbf{y}_* = \mathbf{X}_*\boldsymbol{\beta}_1 + \boldsymbol{\varepsilon}_*$. We will concentrate on comparing the BLUEs of $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}$, and BLUPs of $\mathbf{y}_*$ and $\boldsymbol{\varepsilon}_*$ under $\mathscr{M}_{12}$ and $\mathscr{M}_1$.

Let us shortly consider paper [A] to get an idea what kinds of problems we are dealing with here. Denote

$$\mathbf{G}_1 = \mathbf{X}_1(\mathbf{X}_1'\mathbf{W}_1^{+}\mathbf{X}_1)^{-}\mathbf{X}_1'\mathbf{W}_1^{+}, \quad \mathbf{P}_{\mathbf{X}_1;\dot{\mathbf{M}}_2} = \mathbf{X}_1(\mathbf{X}_1'\dot{\mathbf{M}}_2\mathbf{X}_1)^{-}\mathbf{X}_1'\dot{\mathbf{M}}_2\,,$$

where $\mathbf{W}_1 = \mathbf{V} + \mathbf{X}_1\mathbf{X}_1'$ so that $\mathbf{G}_1\mathbf{y} = \tilde{\boldsymbol{\mu}}_1(\mathscr{M}_1)$ and $\mathbf{P}_{\mathbf{X}_1;\dot{\mathbf{M}}_2}\mathbf{y} = \tilde{\boldsymbol{\mu}}_1(\mathscr{M}_{12})$. We might now be tempted to express the equality $\mathbf{G}_1\mathbf{y} = \mathbf{P}_{\mathbf{X}_1;\dot{\mathbf{M}}_2}\mathbf{y}$ as

$$\tilde{\boldsymbol{\mu}}_1(\mathscr{M}_1) = \tilde{\boldsymbol{\mu}}_1(\mathscr{M}_{12})\,, \quad \text{i.e.,} \quad \mathrm{BLUE}(\boldsymbol{\mu}_1 \mid \mathscr{M}_1) = \mathrm{BLUE}(\boldsymbol{\mu}_1 \mid \mathscr{M}_{12})\,. \tag{56}$$

However, the notation used in (56) can be problematic when the possible values of the response vector $\mathbf{y}$ are taken into account. Doing that, we can consider for example statements like

$$\mathbf{G}_1\mathbf{y} = \mathbf{P}_{\mathbf{X}_1;\dot{\mathbf{M}}_2}\mathbf{y} \ \text{ for all } \mathbf{y} \in \mathscr{C}(\mathbf{X}_1 : \mathbf{V})\,, \tag{57a}$$

$$\mathbf{G}_1\mathbf{y} = \mathbf{P}_{\mathbf{X}_1;\dot{\mathbf{M}}_2}\mathbf{y} \ \text{ for all } \mathbf{y} \in \mathscr{C}(\mathbf{X}_1 : \mathbf{X}_2 : \mathbf{V})\,. \tag{57b}$$

The claim (57a) appears to be equivalent to $\{\mathbf{P}_{\boldsymbol{\mu}_1 \mid \mathscr{M}_{12}}\} \subseteq \{\mathbf{P}_{\boldsymbol{\mu}_1 \mid \mathscr{M}_1}\}$.

[**C**] Haslett *et al.* (2021), [**D**] Haslett *et al.* (2023a).

In these articles we consider the partitioned fixed linear model $\mathscr{F}\colon \mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}$ and the corresponding mixed model $\mathscr{M}\colon \mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\mathbf{u} + \boldsymbol{\varepsilon}$, where $\boldsymbol{\varepsilon}$ is random error vector and $\mathbf{u}$ is a random effect vector. Isotalo *et al.* (2006) found conditions under which an arbitrary representation of the BLUE of an estimable parametric function of $\boldsymbol{\beta}_1$ in the fixed model $\mathscr{F}$ remains BLUE in the mixed model $\mathscr{M}$. In paper [C] we extend the results concerning further equalities arising from models $\mathscr{F}$ and $\mathscr{M}$. In paper [D] we establish upper bounds for the Euclidean norm of the difference between the BLUEs of an estimable parametric function of $\boldsymbol{\beta}_1$ under models $\mathscr{F}$ and $\mathscr{M}$.

[**E**] Haslett *et al.* (2023c), [**F**] Haslett *et al.* (2023b), [**G**] Haslett and Puntanen (2023).

We consider the partitioned linear model $\mathscr{M}_{12}(\mathbf{V}_0)$ and the corresponding small model $\mathscr{M}_1(\mathbf{V}_0)$. We define the set $\mathscr{V}_{1/12}$ of nonnegative definite matrices $\mathbf{V}$ such that every representation of the BLUE of $\boldsymbol{\mu}_1$ under $\mathscr{M}_{12}(\mathbf{V}_0)$ remains BLUE under $\mathscr{M}_{12}(\mathbf{V})$. Correspondingly, we can characterize the set $\mathscr{V}_1$ of matrices $\mathbf{V}$ such that every BLUE of $\boldsymbol{\mu}_1$ under $\mathscr{M}_1(\mathbf{V}_0)$ remains BLUE under $\mathscr{M}_1(\mathbf{V})$. In paper E we focus on the mutual relations between the sets $\mathscr{V}_1$ and $\mathscr{V}_{1/12}$.

In article [F] we focus on the mutual relations between the sets $\mathscr{V}_1$ and $\mathscr{V}_{12}$, where $\mathscr{V}_1$ is defined as in [E] and $\mathscr{V}_{12}$ is the set of nonnegative definite matrices $\mathbf{V}$ such that every representation of the BLUE of $\boldsymbol{\mu} = \mathbf{X}\boldsymbol{\beta}$ under $\mathscr{M}_{12}(\mathbf{V}_0)$ remains BLUE under $\mathscr{M}_{12}(\mathbf{V})$.

Structural insight into Rao's condition of 1971 can be gained by writing the quadratic form that is permitted to be added to the original covariance in block diagonal form. When the original full linear model is made smaller by reducing the number of regressors, block diagonal or diagonal matrices also provide insight into conditions for the entire set of full, small, and intermediate models each to retain their own BLUEs. The paper [G] outlines the role that such changes in error covariance structure can play in data confidentiality and data encryption, especially when the covariance of the BLUEs is also retained.

[H] Haslett *et al.* (2021)

A linear statistic $\mathbf{Fy}$ is called linearly sufficient for $\mathbf{X}_*\boldsymbol{\beta}$ under $\mathscr{M}(\mathbf{V})$ if there exists a matrix $\mathbf{A}$ such that $\mathbf{AFy}$ is the BLUE for $\mathbf{X}_*\boldsymbol{\beta}$, *i.e.*, there exists a matrix $\mathbf{A}$ such that

$$\mathbf{AF}(\mathbf{X} : \mathbf{VM}) = (\mathbf{X}_* : \mathbf{0}) .$$

Thus we can immediately recognize the crucial role of the fundamental BLUE equation in definition of the linear sufficiency. Originally the concept of linear sufficiency as done by Baksalary and Kala (1981, 1986). The article [H] provides an extensive review of this concept.

Acknowledgements

The authors are grateful to Professor Bikas Kumar Sinha for the invitation to prepare this article for possible publication in the SSCA journal *Statistics and Applications*. Thanks go also to Professor Vinod Kumar Gupta for editorial cooperation.

Part of this research has been done during the meetings of an International Research Group on Multivariate and Mixed Linear Models in the Mathematical Research and Conference Center, Będlewo, Poland, in October 2023 and April 2024, supported by the Stefan Banach International Mathematical Center. Thanks for helpful discussions go to Katarzyna Filipiak.

ORCID

Stephen J. Haslett `https://orcid.org/0000-0002-2775-5468`
Jarkko Isotalo `https://orcid.org/0000-0002-8068-6262`
Augustyn Markiewicz `https://orcid.org/0000-0001-5473-3419`
Simo Puntanen `https://orcid.org/0000-0002-6776-0173`

References

Albert, A. (1973). The Gauss-Markov Theorem for regression models with possibly singular covariances. *SIAM Journal on Applied Mathematics*, **24**, 182–187.

Baksalary, J. K. (2004). An elementary development of the equation characterizing best linear unbiased estimators. *Linear Algebra and its Applications*, **388**, 3–6.

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

Baksalary, J. K. and Kala, R. (1986). Linear sufficiency with respect to a given vector of parametric functions. *Journal of Statistical Planning and Inference*, **14**, 331–338.

Baksalary, J. K. and Mathew, T. (1986). Linear sufficiency and completeness in an incorrectly specified general Gauss–Markov model. *Sankhyā Series A*, **48**, 169–180.

Baksalary, J. K. and Mathew, T. (1990). Rank invariance criterion and its application to the unified theory of least squares. *Linear Algebra and its Applications*, **127**, 393–401.

Christensen, R. (2020). *Plane Answers to Complex Questions: The Theory of Linear Models.* Fifth Edition, Springer, New York.

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

Groß, J. (2004). The general Gauss–Markov model with possibly singular dispersion matrix. *Statistical Papers*, **45**, 311–336.

Groß, J. and Puntanen, S. (2000). Estimation under a general partitioned linear model. *Linear Algebra and its Applications*, **321**, 131–144.

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

Haslett, S. J., Isotalo, J., Kala, R., Markiewicz, A., and Puntanen, S. (2021). A review of the linear sufficiency and linear prediction sufficiency in the linear model with new observations. *Multivariate, Multilinear and Mixed Linear Models.* (K. Filipiak, A. Markiewicz, D. von Rosen, eds.) Springer, Cham, pp. 265–318.

Haslett, S. J., Isotalo, J., Markiewicz, A., and Puntanen, S. (2023a). Upper bounds for the Euclidean distances between the BLUEs under the partitioned linear fixed model and the corresponding mixed model. *Applied Linear Algebra, Probability and Statistics: A Volume in Honour of C.R. Rao and Arbind K. Lal.* (R.B. Bapat, K.M. Prasad, S.J. Kirkland, S.K. Neogy, S. Pati, S. Puntanen, eds.) Springer Singapore, Indian Statistical Institute Series. Chapter 3, pp. 27–43.

Haslett, S. J., Isotalo, J., Markiewicz, A., and Puntanen, S. (2023b). Permissible covariance structures for simultaneous retention of BLUEs in small and big linear models. *Applied Linear Algebra, Probability and Statistics: A Volume in Honour of C.R. Rao and Arbind K. Lal.* (R.B. Bapat, K.M. Prasad, S.J. Kirkland, S.K. Neogy, S. Pati, S. Puntanen, eds.) Springer Singapore, Indian Statistical Institute Series. Chapter 11, pp. 197–213.

Haslett, S. J., Isotalo, J., Markiewicz, A., and Puntanen, S. (2023c). Further remarks on permissible covariance structures for simultaneous retention of BLUEs in linear models. *Acta et Commentationes Universitatis Tartuensis de Mathematica*, **27**, 101–112.

Haslett, S. J., Isotalo, J., and Puntanen, S. (2021). Equalities between the BLUEs and BLUPs under the partitioned linear fixed model and the corresponding mixed model. *Acta et Commentationes Universitatis Tartuensis de Mathematica*, **25**, 239–257.

Haslett, S. J., Markiewicz, A., and Puntanen, S. (2020). Properties of BLUEs and BLUPs in full vs. small linear models with new observations. *Recent Developments in Multivariate and Random Matrix Analysis: Festschrift in Honour of Dietrich von Rosen.* (T. Holgersson, M. Singull, eds.) Springer, Cham, 123–146.

Haslett, S. J., Markiewicz, A., and Puntanen, S. (2022a). Properties of the matrix V + XTX' in linear statistical models. *Gujarat Journal of Statistical and Data Science* (formerly *Gujarat Statistical Review*), **38**, 107–131.

Haslett, S. J., Markiewicz, A., and Puntanen, S. (2023). Properties of BLUEs in full versus small linear models. *Communications in Statistics: Theory and Methods*, **52**, 7684–7698.

Haslett, S. J. and Puntanen, S. (2010a). Effect of adding regressors on the equality of the Best Linear Unbiased Estimators (BLUEs) under two linear models. *Journal of Statistical Planning and Inference*, **140**, 104–110.

Haslett, S. J. and Puntanen, S. (2010b). A note on the equality of the BLUPs for new observations under two linear models. *Acta et Commentationes Universitatis Tartuensis de Mathematica*, **14**, 27–33.

Haslett, S. J. and Puntanen, S. (2011). On the equality of the BLUPs under two linear mixed models. *Metrika*, **74**, 381–395.

Haslett, S. J. and Puntanen, S. (2023). Equality of BLUEs for full, small, and intermediate linear models under covariance change, with links to data confidentiality and encryption. *Applied Linear Algebra, Probability and Statistics: A Volume in Honour of C.R. Rao and Arbind K. Lal.* (R.B. Bapat, K.M. Prasad, S.J. Kirkland, S.K. Neogy, S. Pati, S. Puntanen, eds.) Springer Singapore, Indian Statistical Institute Series. Chapter 14, pp. 237–291.

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

Hauke, J., Markiewicz, A., and Puntanen, S. (2013). Revisiting the BLUE in a linear model via proper eigenvectors. *Combinatorial Matrix Theory and Generalized Inverses of Matrices.* (R.B. Bapat, S.J. Kirkland, K.M. Prasad, S. Puntanen, eds.) Springer, India, 73–83.

Henderson, C. R. (1950). Estimation of genetic parameters. *The Annals of Mathematical Statistics*, **21**, 309–310.

Henderson, C. R. (1963). Selection index and expected genetic advance. *Statistical Genetics and Plant Breeding*, National Academy of Sciences – National Research Council Publication No. 982, 141–163.

Isotalo, J., Möls, M., and Puntanen, S. (2006). Invariance of the BLUE under the linear fixed and mixed effects models. *Acta et Commentationes Universitatis Tartuensis de Mathematica*, **10**, 69–76.

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

Isotalo, J., Puntanen, S., and Styan, G. P. H. (2008). A useful matrix decomposition and its statistical applications in linear regression. *Communications in Statistics: Theory and Methods*, **37**, 1436–1457.

Kala, R. (1981). Projectors and linear estimation in general linear models. *Communications in Statistics: Theory and Methods*, **10**, 849–873.

Markiewicz, A., Puntanen, S., and Styan, G. P. H. (2010). A note on the interpretation of the equality of OLSE and BLUE. *Pakistan Journal of Statistics*, **26**, 127–134.

Markiewicz, A., Puntanen, S., and Styan, G. P. H. (2021). The legend of the equality of OLSE and BLUE: highlighted by C.R. Rao in 1967. *Methodology and Applications of Statistics: A Volume in Honor of C.R. Rao on the Occasion of his 100th Birthday.* (B.C. Arnold, N. Balakrishnan, C.A. Coelho, eds.) Springer, Cham. Pages 51–76.

Mathew, T. and Bhimasankaram, P. (1983). On the robustness of LRT in singular linear models. *Sankhyā Series A*, **45**, 301–312.

Mitra, S .K. and Moore, B. J. (1973). Gauss–Markov estimation with an incorrect dispersion matrix. *Sankhyā Series A*, **35**, 139–152.

Puntanen, S., Styan, G. P. H., and Werner, H. J. (2000). Two matrix-based proofs that the linear estimator Gy is the best linear unbiased estimator. *Journal of Statistical Planning and Inference*, **88**, 173–179.

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

Rao, C. R. (1967). Least squares theory using an estimated dispersion matrix and its application to measurement of signals. *Proc. Fifth Berkeley Symp. on Math. Statist. and Prob., Vol. 1.* (L.M. Le Cam, J. Neyman, eds.) Univ. of Calif. Press, Berkeley, pp. 355–372.

Rao, C. R. (1968). A note on a previous lemma in the theory of least squares and some further results. *Sankhyā Series A*, **30**, 259–266.

Rao, C. R. (1971). Unified theory of linear estimation. *Sankhyā Series A*, **33**, 371–394. [Corrigenda (1972): **34**, p. 194 and p. 477]

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

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

Zmyślony, R. (1980). A characterization of best linear unbiased estimators in the general linear model. *Mathematical Statistics and Probability Theory: In Proceedings Sixth International Conference (Wisła, Poland, 1978).* (W. Klonecki, A. Kozek, J. Rosiński, eds.) Springer, New York, pp. 365–373.

Zyskind, G. (1967). On canonical forms, non-negative covariance matrices and best and simple least squares linear estimators in linear models. *The Annals of Mathematical Statistics*, **38**, 1092–1109.