The Equality Between Linear Transforms of Ordinary Least Squares and Best Linear Unbiased Estimator

Jürgen Groß & Götz Trenkler
Department of Statistics
University of Dortmund
Vogelpothsweg 87
*D-44221 Dortmund, Germany**

Abstract. The best linear unbiased estimator $\mathrm{BLUE}(\mathbf{CX}\boldsymbol{\beta})$ of a linear transform $\mathbf{CX}\boldsymbol{\beta}$ in the general Gauss-Markov model $\{\mathbf{y}, \mathsf{E}(\mathbf{y}) = \mathbf{X}\boldsymbol{\beta}, \mathsf{Cov}(\mathbf{y}) = \sigma^2\mathbf{V}\}$ is the linear transform $\mathbf{C}\,\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta})$ of the best linear unbiased estimator $\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta})$ of $\mathbf{X}\boldsymbol{\beta}$. Similarly, for the ordinary least squares estimator, $\mathrm{OLSE}(\mathbf{CX}\boldsymbol{\beta}) = \mathbf{C}\,\mathrm{OLSE}(\mathbf{X}\boldsymbol{\beta})$. The problem of equality of $\mathrm{OLSE}(\mathbf{X}\boldsymbol{\beta})$ and $\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta})$ has been widely discussed in the literature. In this note, characterizations of the equality $\mathbf{C}\,\mathrm{OLSE}(\mathbf{X}\boldsymbol{\beta}) = \mathbf{C}\,\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta})$ are given in terms of projectors and subspaces.

Keywords: Ordinary least squares estimator, best linear unbiased estimator, prediction, linear transform, orthogonal projector

AMS 1991 Mathematics Subject Classification: 62J05

1. Introduction.

Consider the general Gauss-Markov model denoted by

$$M = \{\mathbf{y}, \mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{V}\}, \tag{1.1}$$

where $\mathbf{y}$ is an observable $n \times 1$ random vector with $\mathsf{E}(\mathbf{y}) = \mathbf{X}\boldsymbol{\beta}$ and $\mathsf{Cov}(\mathbf{y}) = \sigma^2\mathbf{V}$, $\mathbf{X}$ is a known $n \times p$ matrix of rank r, $0 < r < n$, $\mathbf{V}$ is a known $n \times n$ symmetric nonnegative definite matrix (possibly singular), $\boldsymbol{\beta}$ is a $p \times 1$ vector of unknown parameters, and $\sigma^2 > 0$ is either known or unknown.

Our interest focuses on estimation of a linear transform $\mathbf{CX}\boldsymbol{\beta}$ of $\mathbf{X}\boldsymbol{\beta}$, where $\mathbf{C}$ is a given $k \times n$ matrix. Recall that estimation of $\mathbf{X}_0\boldsymbol{\beta}$ with known $\mathbf{X}_0 = \mathbf{CX}$ can also be seen as (classical) prediction of an unobservable random vector $\mathbf{y}_0$ satisfying $\mathsf{E}(\mathbf{y}_0) = \mathbf{X}_0\boldsymbol{\beta}$.

*e-mail: gross@amadeus.statistik.uni-dortmund.de, trenkler@amadeus.statistik.uni-dortmund.de

It is well known [10, 11] that a representation of the best linear unbiased estimator (BLUE) of $\mathbf{CX}\boldsymbol{\beta}$ is given by

$$\mathrm{BLUE}(\mathbf{CX}\boldsymbol{\beta}) = \mathbf{CX}(\mathbf{X}'\mathbf{T}^{+}\mathbf{X})^{+}\mathbf{X}'\mathbf{T}^{+}\mathbf{y}, \tag{1.2}$$

where $\mathbf{T} = \mathbf{V}+\mathbf{XX}'$. Here, $\mathbf{A}'$ and $\mathbf{A}^{+}$ denote the transpose and the Moore-Penrose inverse of an arbitrary matrix $\mathbf{A}$, respectively. Since we have

$$\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta}) = \mathbf{X}(\mathbf{X}'\mathbf{T}^{+}\mathbf{X})^{+}\mathbf{X}'\mathbf{T}^{+}\mathbf{y}, \tag{1.3}$$

it is clear that $\mathrm{BLUE}(\mathbf{CX}\boldsymbol{\beta}) = \mathbf{C}\,\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta})$. Although there exist further representations of $\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta})$, all of them coincide almost surely, so that without loss of generality we may confine ourselves to (1.3). Consider now the competing estimator $\mathbf{C}\,\mathrm{OLSE}(\mathbf{X}\boldsymbol{\beta})$, where

$$\mathrm{OLSE}(\mathbf{X}\boldsymbol{\beta}) = \mathbf{XX}^{+}\mathbf{y} \tag{1.4}$$

is known as the ordinary least squares estimator (OLSE) of $\mathbf{X}\boldsymbol{\beta}$. The problem of equality of $\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta})$ and $\mathrm{OLSE}(\mathbf{X}\boldsymbol{\beta})$ has been widely discussed in the literature. See e.g. [2] where two different versions of this problem are investigated, and [9] for an excellent overview. As one (among many other) necessary and sufficient condition for $\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{OLSE}(\mathbf{X}\boldsymbol{\beta})$, Puntanen and Styan [9, cond. AS2] state

$$\mathbf{P_X V} = \mathbf{P_X V P_X}, \tag{1.5}$$

where the symbol $\mathbf{P_A} = \mathbf{AA}^{+}$ is used to represent the orthogonal projector onto the range (column space) $\mathcal{R}(\mathbf{A})$ of an arbitrary $n \times p$ matrix $\mathbf{A}$. The symbol $\mathbf{Q_A} = \mathbf{I}_n - \mathbf{P_A}$ will denote the orthogonal projector onto the orthogonal complement of $\mathcal{R}(\mathbf{A})$, compare also [6, Chap. 12].

In the following we derive a condition similar to (1.5) for the less restrictive equality $\mathbf{C}\,\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta}) = \mathbf{C}\,\mathrm{OLSE}(\mathbf{X}\boldsymbol{\beta})$.

2. Equality of estimators.

By confining ourselves to the representation (1.3) of $\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta})$ we observe that $\mathbf{C}\,\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta}) = \mathbf{C}\,\mathrm{OLSE}(\mathbf{X}\boldsymbol{\beta})$ almost surely if and only if

$$\mathbf{CXX}^{+}\mathbf{y} = \mathbf{CX}(\mathbf{X}'\mathbf{T}^{+}\mathbf{X})^{+}\mathbf{X}'\mathbf{T}^{+}\mathbf{y} \quad \text{for all } \mathbf{y} \in \mathcal{R}(\mathbf{X}) + \mathcal{R}(\mathbf{V}), \tag{2.1}$$

the latter being equivalent to the identities

$$\mathbf{C}\mathbf{X}\mathbf{X}^{+}\mathbf{X} = \mathbf{C}\mathbf{X}(\mathbf{X}'\mathbf{T}^{+}\mathbf{X})^{+}\mathbf{X}'\mathbf{T}^{+}\mathbf{X} \tag{2.2}$$

and

$$\mathbf{C}\mathbf{X}\mathbf{X}^{+}\mathbf{V} = \mathbf{C}\mathbf{X}(\mathbf{X}'\mathbf{T}^{+}\mathbf{X})^{+}\mathbf{X}'\mathbf{T}^{+}\mathbf{V}. \tag{2.3}$$

But since we have $\mathbf{X} = \mathbf{X}\mathbf{X}^{+}\mathbf{X}$ and

$$\mathbf{X} = \mathbf{X}(\mathbf{X}'\mathbf{T}^{+}\mathbf{X})^{+}\mathbf{X}'\mathbf{T}^{+}\mathbf{X}, \tag{2.4}$$

the latter being true in view of $\mathcal{R}(\mathbf{X}') = \mathcal{R}(\mathbf{X}'\mathbf{T}^{+}\mathbf{X})$, see e.g. [3, Theorem 2], condition (2.2) is always met. Thus, the equality under study holds if and only if (2.3) is satisfied.

PROPOSITION 1. *Under model $M = \{\mathbf{y}, \mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{V}\}$, equality of* $\mathbf{C}\,\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta})$ *and* $\mathbf{C}\,\mathrm{OLSE}(\mathbf{X}\boldsymbol{\beta})$ *holds almost surely if and only if*

$$\mathbf{C}\,\mathbf{P}_{\mathbf{X}}\mathbf{V} = \mathbf{C}\,\mathbf{P}_{\mathbf{X}}\mathbf{V}\mathbf{P}_{\mathbf{X}}. \tag{2.5}$$

Proof. It remains to show that (2.5) is equivalent to (2.3). Since by (2.4) and $\mathbf{X}'\mathbf{T}^{+}\mathbf{T} = \mathbf{X}'$ we have

$$\begin{aligned}
\mathbf{X}(\mathbf{X}'\mathbf{T}^{+}\mathbf{X})^{+}\mathbf{X}'\mathbf{T}^{+}\mathbf{V} &= \mathbf{X}(\mathbf{X}'\mathbf{T}^{+}\mathbf{X})^{+}\mathbf{X}'\mathbf{T}^{+}(\mathbf{T} - \mathbf{X}\mathbf{X}') \\
&= \mathbf{X}(\mathbf{X}'\mathbf{T}^{+}\mathbf{X})^{+}\mathbf{X}'\mathbf{T}^{+}\mathbf{T} - \mathbf{X}\mathbf{X}' \\
&= \mathbf{X}(\mathbf{X}'\mathbf{T}^{+}\mathbf{X})^{+}\mathbf{X}' - \mathbf{X}\mathbf{X}',
\end{aligned} \tag{2.6}$$

it is clear that (2.5) follows from (2.3) by right-multiplication of (2.3) with $\mathbf{X}\mathbf{X}^{+}$. Conversely assume that (2.5) is satisfied. Then,

$$\begin{aligned}
\mathbf{C}\mathbf{P}_{\mathbf{X}}\mathbf{V} &= \mathbf{C}\mathbf{P}_{\mathbf{X}}\mathbf{V}(\mathbf{X}^{+})'\mathbf{X}' \\
&= \mathbf{C}\mathbf{P}_{\mathbf{X}}\mathbf{V}(\mathbf{X}^{+})'\mathbf{X}'\mathbf{T}^{+}\mathbf{X}(\mathbf{X}'\mathbf{T}^{+}\mathbf{X})^{+}\mathbf{X}' \\
&= \mathbf{C}\mathbf{P}_{\mathbf{X}}\mathbf{V}\mathbf{P}_{\mathbf{X}}\mathbf{T}^{+}\mathbf{X}(\mathbf{X}'\mathbf{T}^{+}\mathbf{X})^{+}\mathbf{X}' \\
&= \mathbf{C}\mathbf{P}_{\mathbf{X}}\mathbf{V}\mathbf{T}^{+}\mathbf{X}(\mathbf{X}'\mathbf{T}^{+}\mathbf{X})^{+}\mathbf{X}'.
\end{aligned}$$

But since in view of (2.6) we have

$$\mathbf{V}\mathbf{T}^{+}\mathbf{X}(\mathbf{X}'\mathbf{T}^{+}\mathbf{X})^{+}\mathbf{X}' = \mathbf{X}(\mathbf{X}'\mathbf{T}^{+}\mathbf{X})^{+}\mathbf{X}'\mathbf{T}^{+}\mathbf{V},$$

it follows that

$$\mathbf{C}\mathbf{P}_{\mathbf{X}}\mathbf{V} = \mathbf{C}\mathbf{P}_{\mathbf{X}}\mathbf{X}(\mathbf{X}'\mathbf{T}^{+}\mathbf{X})^{+}\mathbf{X}'\mathbf{T}^{+}\mathbf{V} = \mathbf{C}\mathbf{X}(\mathbf{X}'\mathbf{T}^{+}\mathbf{X})^{+}\mathbf{X}'\mathbf{T}^{+}\mathbf{V},$$

showing (2.3). ∎

As an immediate corollary we obtain the following.

Corollary. *Under model* $M = \{\mathbf{y}, \mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{V}\}$, *the following statements are equivalent* :

(i) $\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{OLSE}(\mathbf{X}\boldsymbol{\beta})$,

(ii) $\mathbf{X}'\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta}) = \mathbf{X}'\mathrm{OLSE}(\mathbf{X}\boldsymbol{\beta})$,

(iii) $\mathbf{X}^{+}\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta}) = \mathbf{X}^{+}\mathrm{OLSE}(\mathbf{X}\boldsymbol{\beta})$,

Proof. Observe that for $\mathbf{C} = \mathbf{X}'$ and $\mathbf{C} = \mathbf{X}^{+}$, equations (2.5) and (1.5) are equivalent. ∎

Obviously each of the numerous equivalent conditions for $\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{OLSE}(\mathbf{X}\boldsymbol{\beta})$ is sufficient for $\mathbf{C}\,\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta}) = \mathbf{C}\,\mathrm{OLSE}(\mathbf{X}\boldsymbol{\beta})$, including

$$\mathbf{P}_{\mathbf{X}}\mathbf{V} = \mathbf{V}\mathbf{P}_{\mathbf{X}}, \tag{2.7}$$

which is called Zyskind's condition in [9].

It is clear that (2.5) may be reformulated as

$$\mathbf{C}\mathbf{P}_{\mathbf{X}}\mathbf{V}\mathbf{Q}_{\mathbf{X}} = \mathbf{0}, \tag{2.8}$$

where $\mathbf{Q}_{\mathbf{X}} = \mathbf{I}_n - \mathbf{P}_{\mathbf{X}}$. If we are interested in characterizing all matrices $\mathbf{C}$ satisfying $\mathbf{C}\,\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta}) = \mathbf{C}\,\mathrm{OLSE}(\mathbf{X}\boldsymbol{\beta})$ almost surely, then we simply have to inspect the general solution to (2.8) with respect to $\mathbf{C}$, being

$$\mathbf{C} = \mathbf{Z}(\mathbf{I}_n - \mathbf{P}_{\mathbf{X}}\mathbf{V}\mathbf{Q}_{\mathbf{X}}(\mathbf{P}_{\mathbf{X}}\mathbf{V}\mathbf{Q}_{\mathbf{X}})^{+}), \tag{2.9}$$

where $\mathbf{Z}$ is an arbitrary $k \times n$ matrix. In view of

$$\mathrm{rank}(\mathbf{P}_{\mathbf{X}}\mathbf{V}\mathbf{Q}_{\mathbf{X}}) \leq \mathrm{rank}(\mathbf{Q}_{\mathbf{X}}) = n - \mathrm{rank}(\mathbf{X}) < n$$

it is obvious that there always exists more than one (trivial) solution to (2.8).

If we are interested in characterizing all nonnegative definite matrices $\mathbf{V}$ satisfying $\mathbf{C}\,\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta}) = \mathbf{C}\,\mathrm{OLSE}(\mathbf{X}\boldsymbol{\beta})$ almost surely, then we have to inspect the general nonnegative definite solution to (2.8) with respect to $\mathbf{V}$. It is clear that (2.8) is equivalent to $\mathbf{M}\mathbf{V}\mathbf{Q}_{\mathbf{X}} = \mathbf{0}$, where $\mathbf{M} = \mathbf{P}_{\mathbf{X}}\mathbf{C}'\mathbf{C}\mathbf{P}_{\mathbf{X}}$, see [6, Lemma 11.6.2], which in turn is equivalent to

$$\mathbf{P}_{\mathbf{M}}\mathbf{V}\mathbf{Q}_{\mathbf{X}} = \mathbf{0}. \tag{2.10}$$

The general nonnegative definite solution to (2.10) can be derived from Theorem 2.5 in [7]. By letting $\mathbb{R}_{m\times n}$ denote the set of $m \times n$ real matrices and $\mathbb{R}^{\geq}_{n\times n}$ denote the set of $n \times n$ real (symmetric) nonnegative definite matrices, we may state the following.

Proposition 2. *Under model $M = \{\mathbf{y}, \mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{V}\}$, the following two statements hold:*

(i) *For given $\mathbf{X} \in \mathbb{R}_{n\times p}$ and $\mathbf{V} \in \mathbb{R}^{\geq}_{n\times n}$ the set of all matrices $\mathbf{C} \in \mathbb{R}^{k\times n}$ satisfying $\mathbf{C}\,\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta}) = \mathbf{C}\,\mathrm{OLSE}(\mathbf{X}\boldsymbol{\beta})$ almost surely is given by*

$$\{\mathbf{C} = \mathbf{Z}\mathbf{Q}_{\mathbf{L}} | \mathbf{Z} \in \mathbb{R}_{k\times n}\},$$

where $\mathbf{L} = \mathbf{P}_{\mathbf{X}}\mathbf{V}\mathbf{Q}_{\mathbf{X}}$.

(ii) *For given $\mathbf{X} \in \mathbb{R}_{n\times p}$ and $\mathbf{C} \in \mathbb{R}_{k\times n}$ the set of all matrices $\mathbf{V} \in \mathbb{R}^{\geq}_{n\times n}$ satisfying $\mathbf{C}\,\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta}) = \mathbf{C}\,\mathrm{OLSE}(\mathbf{X}\boldsymbol{\beta})$ almost surely is given by*

$$\{\mathbf{V} = \mathbf{P}_{\mathbf{M}}\mathbf{Z}_1\mathbf{P}_{\mathbf{M}} + \mathbf{Q}_{\mathbf{X}}\mathbf{Z}_2\mathbf{Q}_{\mathbf{X}} + (\mathbf{P}_{\mathbf{X}} - \mathbf{P}_{\mathbf{M}})\mathbf{Z}_3(\mathbf{P}_{\mathbf{X}} - \mathbf{P}_{\mathbf{M}}) | \mathbf{Z}_1, \mathbf{Z}_2, \mathbf{Z}_3 \in \mathbb{R}^{\geq}_{n\times n}\},$$

where $\mathbf{M} = \mathbf{P}_{\mathbf{X}}\mathbf{C}'\mathbf{C}\mathbf{P}_{\mathbf{X}}$.

Proof. The proof of (i) is clear from the above considerations. For the proof of (ii) observe beforehand that

$$\mathbf{P}_{\mathbf{X}}\mathbf{P}_{\mathbf{M}} = \mathbf{P}_{\mathbf{M}} = \mathbf{P}_{\mathbf{M}}\mathbf{P}_{\mathbf{X}}, \tag{2.11}$$

$$(\mathbf{P}_{\mathbf{M}} + \mathbf{Q}_{\mathbf{X}})^+ = \mathbf{P}_{\mathbf{M}} + \mathbf{Q}_{\mathbf{X}}, \tag{2.12}$$

see also [4, Theorem 3.1.1] for (2.12). From Theorem 2.5 in [7], the general nonnegative definite solution to (2.10) with respect to $\mathbf{V}$ is given by

$$\begin{aligned}\mathbf{V} \;=\; & (\mathbf{P}_{\mathbf{M}} + \mathbf{Q}_{\mathbf{X}})^+(\mathbf{A} + \mathbf{B})(\mathbf{P}_{\mathbf{M}} + \mathbf{Q}_{\mathbf{X}})^+ \\ & + [\mathbf{I}_n - (\mathbf{P}_{\mathbf{M}} + \mathbf{Q}_{\mathbf{X}})^+(\mathbf{P}_{\mathbf{M}} + \mathbf{Q}_{\mathbf{X}})]\mathbf{Z}_3[\mathbf{I}_n - (\mathbf{P}_{\mathbf{M}} + \mathbf{Q}_{\mathbf{X}})^+(\mathbf{P}_{\mathbf{M}} + \mathbf{Q}_{\mathbf{X}})],\end{aligned}$$

where $\mathbf{Z}_3 \in \mathbb{R}^{\geq}_{n\times n}$ is arbitrary, and $\mathbf{A}$ and $\mathbf{B}$ are arbitrary nonnegative definite solutions of

$$\mathbf{A}(\mathbf{P_M} + \mathbf{Q_X})^{+}\mathbf{Q_X} = \mathbf{0}, \tag{2.13}$$

$$\mathbf{P_M}(\mathbf{P_M} + \mathbf{Q_X})^{+}\mathbf{B} = \mathbf{0}, \tag{2.14}$$

such that $\mathbf{A} + \mathbf{B}$ is nonnegative definite. By using (2.11) and (2.12) it follows that (2.13) is equivalent to $\mathbf{A}\mathbf{Q_X} = \mathbf{0}$ with general nonnegative definite solution, see [7, Theorem 2.2],

$$\mathbf{A} = (\mathbf{I}_n - \mathbf{Q_X})\mathbf{Z}_1(\mathbf{I}_n - \mathbf{Q_X}) = \mathbf{P_X}\mathbf{Z}_1\mathbf{P_X}, \tag{2.15}$$

where $\mathbf{Z}_1 \in \mathbb{R}^{\geq}_{n\times n}$ is arbitrary. Moreover, it is seen that (2.14) is equivalent to $\mathbf{P_M}\mathbf{B} = \mathbf{0}$ with general nonnegative definite solution, see [7, Theorem 2.2],

$$\mathbf{B} = (\mathbf{I}_n - \mathbf{P_M})\mathbf{Z}_2(\mathbf{I}_n - \mathbf{P_M}) = \mathbf{Q_M}\mathbf{Z}_2\mathbf{Q_M}, \tag{2.16}$$

where $\mathbf{Z}_2 \in \mathbb{R}^{\geq}_{n\times n}$ is arbitrary. Clearly $\mathbf{A} + \mathbf{B}$ is nonnegative definite for all choices of $\mathbf{Z}_1, \mathbf{Z}_2 \in \mathbb{R}^{\geq}_{n\times n}$. By using again (2.11) and (2.12), we observe that

$$\mathbf{I}_n - (\mathbf{P_M} + \mathbf{Q_X})^{+}(\mathbf{P_M} + \mathbf{Q_X}) = \mathbf{P_X} - \mathbf{P_M}.$$

Therefore, the general nonnegative definite solution to (2.10) is given by

$$\mathbf{V} = (\mathbf{P_M} + \mathbf{Q_X})(\mathbf{A} + \mathbf{B})(\mathbf{P_M} + \mathbf{Q_X}) + (\mathbf{P_X} - \mathbf{P_M})\mathbf{Z}_3(\mathbf{P_X} - \mathbf{P_M}),$$

where $\mathbf{A}$ and $\mathbf{B}$ are as in (2.15) and (2.16), respectively. By writing

$$\begin{aligned}
&(\mathbf{P_M} + \mathbf{Q_X})(\mathbf{A} + \mathbf{B})(\mathbf{P_M} + \mathbf{Q_X}) \\
&\quad = (\mathbf{P_M} + \mathbf{Q_X})(\mathbf{P_X}\mathbf{Z}_1\mathbf{P_X} + \mathbf{Q_M}\mathbf{Z}_2\mathbf{Q_M})(\mathbf{P_M} + \mathbf{Q_X}) \\
&\quad = \mathbf{P_M}\mathbf{P_X}\mathbf{Z}_1\mathbf{P_X}\mathbf{P_M} + \mathbf{Q_X}\mathbf{Q_M}\mathbf{Z}_2\mathbf{Q_M}\mathbf{Q_X} \\
&\quad = \mathbf{P_M}\mathbf{Z}_1\mathbf{P_M} + \mathbf{Q_X}\mathbf{Z}_2\mathbf{Q_X},
\end{aligned}$$

we arrive at

$$\mathbf{V} = \mathbf{P_M}\mathbf{Z}_1\mathbf{P_M} + \mathbf{Q_X}\mathbf{Z}_2\mathbf{Q_X} + (\mathbf{P_X} - \mathbf{P_M})\mathbf{Z}_3(\mathbf{P_X} - \mathbf{P_M}),$$

where $\mathbf{Z}_1, \mathbf{Z}_2, \mathbf{Z}_3 \in \mathbb{R}^{\geq}_{n\times n}$ are arbitrary. ■

A related but somewhat different problem is to determine the subspace of possible observation vectors $\mathbf{y}$ for fixed $\mathbf{C}$, $\mathbf{X}$ and $\mathbf{V}$ satisfying $\mathbf{C}\,\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta}) = \mathbf{C}\,\mathrm{OLSE}(\mathbf{X}\boldsymbol{\beta})$. Under the assumptions $r = \mathrm{rank}(\mathbf{X}) = p$, $\mathrm{rank}(\mathbf{V}) = n$ and $\mathbf{C} = \mathbf{B}\mathbf{X}^{+}$, where $\mathbf{B}$ is an arbitrary (but fixed) $k \times p$ matrix, this subspace has been identified in [5] to be

$$\mathfrak{E} = \mathcal{R}(\mathbf{X}) \oplus \left[\mathcal{R}(\mathbf{X}^{\perp}) \cap \left[\mathcal{R}(\mathbf{V}\mathbf{X}^{\perp}) \oplus [\mathcal{R}(\mathbf{X}) \cap \mathcal{N}(\mathbf{C})]\right]\right], \tag{2.17}$$

where $\mathbf{X}^{\perp}$ denotes any matrix of maximal rank such that $\mathbf{X}'\mathbf{X}^{\perp} = \mathbf{0}$, and $\mathcal{N}(\mathbf{C})$ denotes the null space of $\mathbf{C}$. In case $\mathbf{B} = \mathbf{I}_p$, $k = p$, the subspace $\mathfrak{E}$ reduces to

$$\mathfrak{E} = \mathcal{R}(\mathbf{X}) \oplus \left[\mathcal{R}(\mathbf{X}^{\perp}) \cap \mathcal{R}(\mathbf{V}\mathbf{X}^{\perp})\right], \tag{2.18}$$

which has been observed earlier in [8].

We will now demonstrate that the subspace (2.17) remains the appropriate choice under the more general assumptions of model (1.1), when in addition $\mathbf{y}$ is restricted to be in $\mathcal{R}(\mathbf{X})+\mathcal{R}(\mathbf{V})$. The latter guarantees that different choices of representations of $\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta})$ cannot lead to different estimates of $\mathbf{X}\boldsymbol{\beta}$.

Proposition 3. *Under model $M = \{\mathbf{y}, \mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{V}\}$, for given $\mathbf{X} \in \mathbb{R}_{n\times p}$, $\mathbf{V} \in \mathbb{R}^{\geq}_{n\times n}$ and $\mathbf{C} \in \mathbb{R}_{k\times n}$, the set $\mathfrak{E}$ of all vectors $\mathbf{y} \in \mathcal{R}(\mathbf{X})+\mathcal{R}(\mathbf{V})$ satisfying $\mathbf{C}\,\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta}) = \mathbf{C}\,\mathrm{OLSE}(\mathbf{X}\boldsymbol{\beta})$ is given by*

$$\mathfrak{E} = \mathcal{R}(\mathbf{X}) \oplus [\mathcal{R}(\mathbf{X}^{\perp}) \cap \mathfrak{F}],$$

where $\mathfrak{F} = \mathcal{R}(\mathbf{V}\mathbf{X}^{\perp}) \oplus [\mathcal{R}(\mathbf{X}) \cap \mathcal{N}(\mathbf{C})]$.

Proof. The set of all vectors $\mathbf{y} \in \mathfrak{E}$ is

$$\mathfrak{E} = \mathcal{N}[\mathbf{C}(\mathbf{P}_{\mathbf{X}} - \mathbf{R})] \cap [\mathcal{R}(\mathbf{X}) + \mathcal{R}(\mathbf{V})], \tag{2.19}$$

where $\mathbf{R} = \mathbf{X}(\mathbf{X}'\mathbf{T}^{+}\mathbf{X})^{+}\mathbf{X}'\mathbf{T}^{+}$. Let $\mathfrak{F} = [\mathcal{R}(\mathbf{X}) + \mathcal{R}(\mathbf{V})] \cap \mathcal{N}(\mathbf{C}\mathbf{R})$. Then

$$\mathfrak{F} = \mathcal{R}(\mathbf{V}\mathbf{X}^{\perp}) \oplus [\mathcal{R}(\mathbf{X}) \cap \mathcal{N}(\mathbf{C})] \tag{2.20}$$

follows similarly as in the proof of Lemma 1 in [5], and it remains to show

$$\mathcal{N}[\mathbf{C}(\mathbf{P}_{\mathbf{X}} - \mathbf{R})] \cap [\mathcal{R}(\mathbf{X}) + \mathcal{R}(\mathbf{V})] = \mathcal{R}(\mathbf{X}) \oplus \left[\mathcal{R}(\mathbf{X}^{\perp}) \cap \mathfrak{F}\right]. \tag{2.21}$$

Let $\mathbf{y}$ be a vector belonging to the left-hand subspace of (2.21). Such a vector $\mathbf{y}$ can be written as $\mathbf{y} = \mathbf{a} + \mathbf{b}$ for some $\mathbf{a} \in \mathcal{R}(\mathbf{X})$ and some $\mathbf{b} \in \mathcal{R}(\mathbf{X}^{\perp})$. Obviously, $\mathbf{b} =$

$\mathbf{y}-\mathbf{a} \in \mathcal{R}(\mathbf{X})+\mathcal{R}(\mathbf{V})$. Moreover we have $\mathbf{CP_X y} = \mathbf{Ca}$ and $\mathbf{CRy} = \mathbf{CRa}+\mathbf{CRb} = \mathbf{Ca} + \mathbf{CRb}$, yielding $\mathbf{0} = \mathbf{C}(\mathbf{P_X} - \mathbf{R})\mathbf{y} = -\mathbf{CRb}$. Therefore, $\mathbf{b} \in \mathfrak{F} = \mathcal{N}(\mathbf{CR}) \cap [\mathcal{R}(\mathbf{X}) + \mathcal{R}(\mathbf{V})]$, showing that the left-hand subspace of (2.21) is contained in the right-hand subspace of (2.21). To demonstrate the reverse inclusion let $\mathbf{y} = \mathbf{Xa}+\mathbf{b}$, where $\mathbf{a} \in \mathbb{R}^p$ and $\mathbf{b} \in \mathcal{R}(\mathbf{X}^\perp) \cap \mathfrak{F}$, where clearly $\mathcal{R}(\mathbf{X}^\perp) \cap \mathfrak{F} \subseteq \mathcal{R}(\mathbf{X}) + \mathcal{R}(\mathbf{V})$, and therefore $\mathbf{y} \in \mathcal{R}(\mathbf{X}) + \mathcal{R}(\mathbf{V})$. Moreover, $\mathbf{X}'\mathbf{b} = \mathbf{0}$ and $\mathbf{CRb} = \mathbf{0}$, and in view of $(\mathbf{P_X} - \mathbf{R})\mathbf{X} = \mathbf{0}$ we obtain $\mathbf{C}(\mathbf{P_X} - \mathbf{R})\mathbf{y} = \mathbf{C}(\mathbf{P_X} - \mathbf{R})\mathbf{Xa} + \mathbf{C}(\mathbf{P_X} - \mathbf{R})\mathbf{b} = \mathbf{C}(\mathbf{P_X} - \mathbf{R})\mathbf{b} = \mathbf{C}(\mathbf{X}^+)'\mathbf{X}'\mathbf{b} - \mathbf{CRb} = \mathbf{0} - \mathbf{0} = \mathbf{0}$, showing $\mathbf{y} \in \mathcal{N}[\mathbf{C}(\mathbf{P_X} - \mathbf{R})]$. Hence, the right-hand subspace of (2.21) is contained in the left-hand subspace of (2.21). ∎

Note that for the special choice $\mathbf{C} = \mathbf{I}_n$, $k = n$, we have $\mathfrak{F} = \mathcal{R}(\mathbf{VX}^\perp)$, showing that the set $\mathfrak{E}$ of all vectors $\mathbf{y} \in \mathcal{R}(\mathbf{X})+\mathcal{R}(\mathbf{V})$ satisfying $\text{BLUE}(\mathbf{X}\boldsymbol{\beta}) = \text{OLSE}(\mathbf{X}\boldsymbol{\beta})$ is given by (2.18).

3. Example.

Consider the one-way classification model

$$y_{ij} = \mu + \alpha_i + e_{ij}, \quad i = 1, \dots, a, \quad j = 1, \dots n_i,$$

where the e_{ij}'s are uncorrelated random variables with means 0 and variances $d_{ij}\sigma^2$.

Assume for a numerical example $a = 3$, $n_1 = 3$, $n_2 = 2$, $n_1 = 1$, and $d_{ij} = 1$ if $(i, j) \neq (1, 3)$. Assume in addition $d_{13} \neq 1$ but otherwise unknown. Then the error variances are not homogenous within groups, and from Corollary 4 in [1] it follows that we do not have equality of OLSE and BLUE of any parametric function. However, if we consider the contrast $\alpha_2 - \alpha_3 = \mathbf{c}'\mathbf{X}\boldsymbol{\beta}$, where $\mathbf{c}' = (0, 0, 0, -\frac{1}{2}, -\frac{1}{2}, 1)$, it follows easily from our Proposition that $\text{OLSE}(\alpha_2 - \alpha_3) = \text{BLUE}(\alpha_2 - \alpha_3) = y_{31} - \frac{1}{2}y_{21} - \frac{1}{2}y_{22}$.

Acknowledgements. Support by Deutsche Forschungsgemeinschaft under grants Tr 253/2-3 and SFB 475 is gratefully acknowledged.

REFERENCES

1. J.K. Baksalary, Criteria for the equality between ordinary least squares and best linear unbiased estimators under certain linear models, *The Canadian Journal of Statistics*, 16, pp. 97–102 (1988).

2. J.K. Baksalary and A.C. van Eijnsbergen, A comparison of two criteria for ordinary-least-squares estimators to be best linear unbiased estimators, *The American Statistician*, 42, pp. 205–208 (1988).

3. J.K. Baksalary, S. Puntanen and G.P.H. Styan, A property of the dispersion matrix of the best linear unbiased estimator in the general Gauss-Markov model, *Sankhyā, Ser. A*, 52, pp. 279–296 (1990).

4. S.L. Campbell and C.D. Meyer, Jr., *Generalized Inverses of Linear Transformations*, Pitman, London (1979).

5. J. Groß and G. Trenkler, When do linear transforms of ordinary least squares and Gauss-Markov estimator coincide?, *Sankhyā, Ser. A*, 59, pp. 175–178 (1997).

6. D.A. Harville, *Matrix Algebra From a Statistician's Perspective*, Springer, New York (1997).

7. C.G. Khatri and S.K. Mitra, Hermitian and nonnegative definite solutions of linear matrix equations, *SIAM Journal of Applied Mathematics*, 31, pp. 579–585.

8. W. Krämer, A note on the equality of ordinary least squares and Gauss-Markov estimates in the general linear model, *Sankhyā, Ser. A*, 42, pp. 130–131 (1971).

9. S. Puntanen and G.P.H. Styan, The equality of the ordinary least squares estimator and the best linear unbiased estimator (with discussion), *The American Statistician*, 43, pp. 153–164 (1989).

10. C.R. Rao, Unified theory of linear estimation, *Sankhyā, Ser. A*, 33, pp. 370–396 (1971).

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