This is the accepted manuscript of the article, which has been published in Matrices, Statistics and Big Data : Selected Contributions from IWMS 2016. 2019, 1-22.
https://doi.org/10.1007/978-3-030-17519-1_1

Chapter 1
Further properties of the linear sufficiency in the partitioned linear model

Augustyn Markiewicz and Simo Puntanen

Abstract A linear statistic $\mathbf{Fy}$, where $\mathbf{F}$ is an $f \times n$ matrix, is called linearly sufficient for estimable parametric function $\mathbf{K}\beta$ under the model $\mathscr{M} = \{\mathbf{y}, \mathbf{X}\beta, \mathbf{V}\}$, if there exists a matrix $\mathbf{A}$ such that $\mathbf{AFy}$ is the BLUE for $\mathbf{K}\beta$. In this paper we consider some particular aspects of the linear sufficiency in the partitioned linear model where $\mathbf{X} = (\mathbf{X}_1 : \mathbf{X}_2)$ with β being partitioned accordingly. We provide new results and new insightful proofs for some known facts, using the properties of relevant covariance matrices and their expressions via certain orthogonal projectors. Particular attention will be paid to the situation under which adding new regressors (in $\mathbf{X}_2$) does not affect the linear sufficiency of $\mathbf{Fy}$.

Key words: Best linear unbiased estimator, generalized inverse, linear model, linear sufficiency, orthogonal projector, Löwner ordering, transformed linear model.

1.1 Introduction

In this paper we consider the partitioned linear model $\mathbf{y} = \mathbf{X}_1\beta_1 + \mathbf{X}_2\beta_2 + \varepsilon$, or shortly denoted

$$\mathscr{M}_{12} = \{\mathbf{y}, \mathbf{X}\beta, \mathbf{V}\} = \{\mathbf{y}, \mathbf{X}_1\beta_1 + \mathbf{X}_2\beta_2, \mathbf{V}\}, \tag{1.1}$$

where we may drop off the subscripts from $\mathscr{M}_{12}$ if the partitioning is not essential in the context. In (1.1), $\mathbf{y}$ is an n-dimensional observable response variable, and ε is an unobservable random error with a known covariance matrix $\mathrm{cov}(\varepsilon) = \mathbf{V} =$

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

Simo Puntanen
Faculty of Natural Sciences, FI-33014 University of Tampere, Finland, e-mail: simo.puntanen@uta.fi

$\operatorname{cov}(\mathbf{y})$ and expectation $\mathrm{E}(\varepsilon) = \mathbf{0}$. The matrix $\mathbf{X}$ is a known $n \times p$ matrix, i.e., $\mathbf{X} \in \mathbb{R}^{n \times p}$, partitioned columnwise as $\mathbf{X} = (\mathbf{X}_1 : \mathbf{X}_2)$, $\mathbf{X}_i \in \mathbb{R}^{n \times p_i}$, $i = 1, 2$. Vector $\beta = (\beta_1', \beta_2')' \in \mathbb{R}^p$ is a vector of fixed (but unknown) parameters; here symbol $'$ stands for the transpose. Sometimes we will denote $\mu = \mathbf{X}\beta$, $\mu_i = \mathbf{X}_i\beta_i$, $i = 1, 2$.

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 $\mathbf{A}^\perp$ we denote any matrix satisfying $\mathscr{C}(\mathbf{A}^\perp) = \mathscr{C}(\mathbf{A})^\perp$. Furthermore, we will write $\mathbf{P_A} = \mathbf{P}_{\mathscr{C}(\mathbf{A})} = \mathbf{A}\mathbf{A}^+ = \mathbf{A}(\mathbf{A}'\mathbf{A})^-\mathbf{A}'$ to denote the orthogonal projector (with respect to the standard inner product) onto $\mathscr{C}(\mathbf{A})$. In particular, we denote $\mathbf{M} = \mathbf{I}_n - \mathbf{P_X}$, $\mathbf{M}_i = \mathbf{I}_n - \mathbf{P}_{\mathbf{X}_i}$, $i = 1, 2$.

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

$$\mathscr{M}_{12\cdot 2} = \{\mathbf{M}_2\mathbf{y}, \mathbf{M}_2\mathbf{X}_1\beta_1, \mathbf{M}_2\mathbf{V}\mathbf{M}_2\}, \tag{1.2}$$

which is obtained by premultiplying the model $\mathscr{M}_{12}$ by $\mathbf{M}_2 = \mathbf{I}_n - \mathbf{P}_{\mathbf{X}_2}$. There is one further model that takes lot of our attention, it is the *transformed* model

$$\mathscr{M}_t = \{\mathbf{F}\mathbf{y}, \mathbf{F}\mathbf{X}\beta, \mathbf{F}\mathbf{V}\mathbf{F}'\} = \{\mathbf{F}\mathbf{y}, \mathbf{F}\mathbf{X}_1\beta_1 + \mathbf{F}\mathbf{X}_2\beta_2, \mathbf{F}\mathbf{V}\mathbf{F}'\}, \tag{1.3}$$

which is obtained be premultiplying $\mathscr{M}_{12}$ by matrix $\mathbf{F} \in \mathbb{R}^{f \times n}$.

We assume that the models under consideration are consistent which in the case of $\mathscr{M}$ means that the observed value of the response variable satisfies

$$\mathbf{y} \in \mathscr{C}(\mathbf{X} : \mathbf{V}) = \mathscr{C}(\mathbf{X} : \mathbf{V}\mathbf{X}^\perp) = \mathscr{C}(\mathbf{X}) \oplus \mathscr{C}(\mathbf{V}\mathbf{X}^\perp), \tag{1.4}$$

where "$\oplus$" refers to the direct sum of column spaces.

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

$$\operatorname{cov}(\mathbf{G}\mathbf{y}) \leq_{\mathrm{L}} \operatorname{cov}(\mathbf{C}\mathbf{y}) \quad \text{for all } \mathbf{C} \in \mathbb{R}^{n \times n}\text{: } \mathbf{C}\mathbf{X} = \mathbf{X}. \tag{1.5}$$

The BLUE of an estimable parametric function $\mathbf{K}\beta$, where $\mathbf{K} \in \mathbb{R}^{k \times p}$, is defined in the corresponding way. Recall that $\mathbf{K}\beta$ is said to be estimable if it has a linear unbiased estimator which happens if and only if $\mathscr{C}(\mathbf{K}') \subset \mathscr{C}(\mathbf{X}')$, i.e.,

$$\mathbf{K}\beta \text{ is estimable under } \mathscr{M} \iff \mathscr{C}(\mathbf{K}') \subset \mathscr{C}(\mathbf{X}'). \tag{1.6}$$

The structure of our paper is as follows. In Section 1.2 we provide some preliminary results that are not only needed later on but they have some matrix-algebraic interest in themselves. In Sections 1.3 and 1.4 we consider the estimation of $\mu = \mathbf{X}\beta$ and $\mu_1 = \mathbf{X}_1\beta_1$, respectively. In Section 1.5 we study the linear sufficiency under $\mathscr{M}_1$ vs. $\mathscr{M}_{12}$. We characterize the linearly sufficient statistic $\mathbf{F}\mathbf{y}$ by using the covariance matrices of the BLUEs under $\mathscr{M}_{12}$ and under its transformed version $\mathscr{M}_t$. In particular, certain orthogonal projectors appear useful in our considerations. From a different angle, the linear sufficiency in a partitioned linear model has been treated, e.g., in Isotalo & Puntanen (2006, 2009), Markiewicz & Puntanen (2009), and Kala & Pordzik (2009). Baksalary (1984, 1987, §3.3, §5) considered linear sufficiency under $\mathscr{M}_{12}$ and $\mathscr{M}_1$ assuming that $\mathbf{V} = \mathbf{I}_n$. Dong et al. (2014) study interesting connections between the BLUEs under two transformed models using so called matrix-rank method.

1.2 Some preliminary results

For the proof of the following fundamental lemma, see, e.g., Rao (1973, p. 282).

Lemma 1. *Consider the general linear model $\mathscr{M} = \{\mathbf{y}, \mathbf{X}\beta, \mathbf{V}\}$. Then the statistic $\mathbf{Gy}$ is the* BLUE *for $\mathbf{X}\beta$ if and only if $\mathbf{G}$ satisfies the equation*

$$\mathbf{G}(\mathbf{X} : \mathbf{V}\mathbf{X}^{\perp}) = (\mathbf{X} : \mathbf{0}), \tag{1.7}$$

in which case we denote $\mathbf{G} \in \{\mathbf{P}_{\mathbf{X}|\mathbf{V}\mathbf{X}^{\perp}}\}$. The corresponding condition for $\mathbf{By}$ to be the BLUE *of an estimable parametric function $\mathbf{K}\beta$ is*

$$\mathbf{B}(\mathbf{X} : \mathbf{V}\mathbf{X}^{\perp}) = (\mathbf{K} : \mathbf{0}). \tag{1.8}$$

Two estimators $\mathbf{G}_1\mathbf{y}$ and $\mathbf{G}_2\mathbf{y}$ are said to be equal (with probability 1) whenever $\mathbf{G}_1\mathbf{y} = \mathbf{G}_2\mathbf{y}$ for all $\mathbf{y} \in \mathscr{C}(\mathbf{X} : \mathbf{V}) = \mathscr{C}(\mathbf{X} : \mathbf{V}\mathbf{X}^{\perp})$. When talking about the equality of estimators we sometimes may drop the phrase “with probability 1”. Thus for any $\mathbf{G}_1, \mathbf{G}_2 \in \{\mathbf{P}_{\mathbf{X}|\mathbf{V}\mathbf{X}^{\perp}}\}$ we have $\mathbf{G}_1(\mathbf{X} : \mathbf{V}\mathbf{X}^{\perp}) = \mathbf{G}_2(\mathbf{X} : \mathbf{V}\mathbf{X}^{\perp})$, and thereby $\mathbf{G}_1\mathbf{y} = \mathbf{G}_2\mathbf{y}$ with probability 1.

One well-known solution for $\mathbf{G}$ in (1.7) (which is always solvable) is

$$\mathbf{P}_{\mathbf{X};\mathbf{W}^{-}} := \mathbf{X}(\mathbf{X}'\mathbf{W}^{-}\mathbf{X})^{-}\mathbf{X}'\mathbf{W}^{-}, \tag{1.9}$$

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

$$\mathscr{W} = \left\{\mathbf{W} \in \mathbb{R}^{n\times n} : \mathbf{W} = \mathbf{V} + \mathbf{X}\mathbf{U}\mathbf{U}'\mathbf{X}',\ \mathscr{C}(\mathbf{W}) = \mathscr{C}(\mathbf{X} : \mathbf{V})\right\}. \tag{1.10}$$

For clarity, we may use the notation $\mathscr{W}_{\mathscr{A}}$ to indicate which model is under consideration. Similarly, $\mathbf{W}_{\mathscr{A}}$ may denote a member of class $\mathscr{W}_{\mathscr{A}}$. We will also use the phrase “$\mathbf{W}_{\mathscr{A}}$ is a $\mathbf{W}$-matrix under the model $\mathscr{A}$”.

For the partitioned linear model $\mathscr{M}_{12}$ we will say that $\mathbf{W} \in \mathscr{W}$ if the following properties hold:

$$\begin{aligned}\mathbf{W} = \mathbf{V} + \mathbf{X}\mathbf{U}\mathbf{U}'\mathbf{X}' &= \mathbf{V} + (\mathbf{X}_1 : \mathbf{X}_2)\begin{pmatrix}\mathbf{U}_1\mathbf{U}_1' & \mathbf{0}\\ \mathbf{0} & \mathbf{U}_2\mathbf{U}_2'\end{pmatrix}\begin{pmatrix}\mathbf{X}_1'\\ \mathbf{X}_2'\end{pmatrix}\\ &= \mathbf{V} + \mathbf{X}_1\mathbf{U}_1\mathbf{U}_1'\mathbf{X}_1' + \mathbf{X}_2\mathbf{U}_2\mathbf{U}_2'\mathbf{X}_2'\,, \end{aligned} \tag{1.11a}$$

$$\mathbf{W}_i = \mathbf{V} + \mathbf{X}_i\mathbf{U}_i\mathbf{U}_i'\mathbf{X}_i', \quad i = 1,2, \tag{1.11b}$$

$$\mathscr{C}(\mathbf{W}) = \mathscr{C}(\mathbf{X} : \mathbf{V})\,, \quad \mathscr{C}(\mathbf{W}_i) = \mathscr{C}(\mathbf{X}_i : \mathbf{V})\,, \quad i = 1,2. \tag{1.11c}$$

For example, the following statements concerning $\mathbf{W} \in \mathscr{W}$ are equivalent:

$$\mathscr{C}(\mathbf{X} : \mathbf{V}) = \mathscr{C}(\mathbf{W}),\ \ \mathscr{C}(\mathbf{X}) \subset \mathscr{C}(\mathbf{W}),\ \ \mathscr{C}(\mathbf{X}'\mathbf{W}^-\mathbf{X}) = \mathscr{C}(\mathbf{X}')\,. \tag{1.12}$$

Instead of $\mathscr{W}$, several corresponding properties also hold in the extended set

$$\mathscr{W}_* = \big\{\mathbf{W} \in \mathbb{R}^{n\times n} : \mathbf{W} = \mathbf{V} + \mathbf{X}\mathbf{N}\mathbf{X}',\ \mathscr{C}(\mathbf{W}) = \mathscr{C}(\mathbf{X} : \mathbf{V})\big\}\,, \tag{1.13}$$

where $\mathbf{N} \in \mathbb{R}^{p\times p}$ can be any (not necessarily nonnegative definite) matrix satisfying $\mathscr{C}(\mathbf{W}) = \mathscr{C}(\mathbf{X} : \mathbf{V})$. However, in this paper we consider merely the set $\mathscr{W}$. For further properties of $\mathscr{W}_*$, see, e.g., Puntanen et al. (2011, §12.3), and Kala et al. (2017).

Using (1.9), the BLUEs of $\mu = \mathbf{X}\beta$ and of estimable $\mathbf{K}\beta$, respectively, can be expressed as

$$\mathrm{BLUE}(\mathbf{X}\beta \mid \mathscr{M}) = \tilde{\mu}(\mathscr{M}) = \mathbf{X}(\mathbf{X}'\mathbf{W}^-\mathbf{X})^-\mathbf{X}'\mathbf{W}^-\mathbf{y}, \tag{1.14a}$$

$$\mathrm{BLUE}(\mathbf{K}\beta \mid \mathscr{M}) = \mathbf{K}(\mathbf{X}'\mathbf{W}^-\mathbf{X})^-\mathbf{X}'\mathbf{W}^-\mathbf{y}, \tag{1.14b}$$

where $\mathbf{W}$ belongs to the class $\mathscr{W}$. The representations (1.14a)–(1.14b) are invariant with respect to the choice of generalized inverses involved; this can be shown using (1.12) and the fact that for any nonnull $\mathbf{A}$ and $\mathbf{C}$ the following holds [Rao & Mitra (1971, Lemma 2.2.4)]:

$$\mathbf{A}\mathbf{B}^-\mathbf{C} = \mathbf{A}\mathbf{B}^+\mathbf{C} \text{ for all } \mathbf{B}^- \iff \mathscr{C}(\mathbf{C}) \subset \mathscr{C}(\mathbf{B}) \text{ and } \mathscr{C}(\mathbf{A}') \subset \mathscr{C}(\mathbf{B}')\,. \tag{1.15}$$

Notice that part $\mathbf{X}(\mathbf{X}'\mathbf{W}^-\mathbf{X})^-\mathbf{X}'$ of $\mathbf{P}_{\mathbf{X};\mathbf{W}^-}$ in (1.9) is invariant with respect to the choice of generalized inverses involved but

$$\mathbf{P}_{\mathbf{X};\mathbf{W}^+} = \mathbf{X}(\mathbf{X}'\mathbf{W}^+\mathbf{X})^+\mathbf{X}'\mathbf{W}^+ = \mathbf{X}(\mathbf{X}'\mathbf{W}^-\mathbf{X})^-\mathbf{X}'\mathbf{W}^+ \tag{1.16}$$

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

The concept of linear sufficiency was introduced by Baksalary & Kala (1981) and Drygas (1983) who considered linear statistics, which are “sufficient” for $\mathbf{X}\beta$ under $\mathscr{M}$, or in other words, “linear transformations preserving best linear unbiased estimators”. A linear statistic $\mathbf{Fy}$, where $\mathbf{F} \in \mathbb{R}^{f\times n}$, is called linearly sufficient for $\mathbf{X}\beta$ under the model $\mathscr{M}$ if there exists a matrix $\mathbf{A} \in \mathbb{R}^{n\times f}$ such that $\mathbf{AFy}$ is the BLUE for $\mathbf{X}\beta$. Correspondingly, $\mathbf{Fy}$ is linearly sufficient for estimable $\mathbf{K}\beta$, where $\mathbf{K} \in \mathbb{R}^{k\times p}$, if there exists a matrix $\mathbf{A} \in \mathbb{R}^{k\times f}$ such that $\mathbf{AFy}$ is the BLUE for $\mathbf{K}\beta$.

Sometimes we will denote shortly $\mathbf{Fy} \in \mathrm{S}(\mathbf{X}\beta)$ or $\mathbf{Fy} \in \mathrm{S}(\mathbf{X}\beta \mid \mathscr{M})$, to indicate that $\mathbf{Fy}$ is linearly sufficient for $\mathbf{X}\beta$ under the model $\mathscr{M}$ (if the model is not obvious from the context).

Drygas (1983) introduced the concept of linear minimal sufficiency and defined it as follows: $\mathbf{Fy}$ is linearly minimal sufficient if for any other linearly sufficient statistics $\mathbf{Sy}$, there exists a matrix $\mathbf{A}$ such that $\mathbf{Fy} = \mathbf{ASy}$ almost surely.

In view of Lemma 1, $\mathbf{Fy}$ is linearly sufficient for $\mathbf{X}\beta$ if and only if the equation

$$\mathbf{AF}(\mathbf{X} : \mathbf{VM}) = (\mathbf{X} : \mathbf{0}) \tag{1.17}$$

has a solution for $\mathbf{A}$. Baksalary & Kala (1981) and Drygas (1983) proved part (a) and Baksalary & Kala (1986) part (b) of the following:

Lemma 2. *Consider the model* $\mathscr{M} = \{\mathbf{y}, \mathbf{X}\beta, \mathbf{V}\}$ *and let* $\mathbf{K}\beta$ *be estimable. Then:*

(a) *The statistic* $\mathbf{Fy}$ *is linearly sufficient for* $\mathbf{X}\beta$ *if and only if*

$$\mathscr{C}(\mathbf{X}) \subset \mathscr{C}(\mathbf{WF}'), \quad \text{where } \mathbf{W} \in \mathscr{W}. \tag{1.18}$$

Moreover, $\mathbf{Fy}$ *is linearly minimal sufficient for* $\mathbf{X}\beta$ *if and only if* $\mathscr{C}(\mathbf{X}) = \mathscr{C}(\mathbf{WF}')$.

(b) *The statistic* $\mathbf{Fy}$ *is linearly sufficient for* $\mathbf{K}\beta$ *if and only if*

$$\mathscr{C}[\mathbf{X}(\mathbf{X}'\mathbf{W}^{-}\mathbf{X})^{-}\mathbf{K}'] \subset \mathscr{C}(\mathbf{WF}'), \quad \text{where } \mathbf{W} \in \mathscr{W}. \tag{1.19}$$

Moreover, $\mathbf{Fy}$ *is linearly minimal sufficient for* $\mathbf{K}\beta$ *if and only if equality holds in* (1.19)*.*

Actually, Kala et al. (2017) showed that in Lemma 2 the class $\mathscr{W}$ can be replaced with the more general class $\mathscr{W}_*$ defined in (1.13). For further related references, see Baksalary & Mathew (1986) and Müller (1987).

Supposing that $\mathbf{Fy}$ is linearly sufficient for $\mathbf{X}\beta$, one could expect that both $\mathscr{M}$ and its transformed version $\mathscr{M}_t = \{\mathbf{Fy}, \mathbf{FX}\beta, \mathbf{FVF}'\}$ provide the same basis for obtaining the BLUE of $\mathbf{X}\beta$. This connection was proved by Baksalary & Kala (1981, 1986). Moreover, Tian & Puntanen (2009, Th. 2.8) and Kala et al. (2017, Th. 2) showed the following:

Lemma 3. *Consider the model* $\mathscr{M} = \{\mathbf{y}, \mathbf{X}\beta, \mathbf{V}\}$ *and* $\mathscr{M}_t = \{\mathbf{Fy}, \mathbf{FX}\beta, \mathbf{FVF}'\}$*, and let* $\mathbf{K}\beta$ *be estimable under* $\mathscr{M}_{12}$ *and* $\mathscr{M}_t$*. Then the following statements are equivalent:*

(a) $\mathbf{Fy}$ *is linearly sufficient for* $\mathbf{K}\beta$*.*
(b) $\mathrm{BLUE}(\mathbf{K}\beta \mid \mathscr{M}) = \mathrm{BLUE}(\mathbf{K}\beta \mid \mathscr{M}_t)$ *with probability* 1*.*
(c) *There exists at least one representation of* BLUE *of* $\mathbf{K}\beta$ *under* $\mathscr{M}$ *which is the* BLUE *also under the transformed model* $\mathscr{M}_t$*.*

Later we will need the following Lemma 4. The proofs are parallel to those in Puntanen et al. (2011, §5.13), and Markiewicz & Puntanen (2015, Th. 5.2). In this lemma the notation $\mathbf{A}^{1/2}$ stands for the nonnegative definite square root of a nonnegative definite matrix $\mathbf{A}$. Similarly $\mathbf{A}^{+1/2}$ denotes the Moore–Penrose inverse of $\mathbf{A}^{1/2}$. Notice that in particular $\mathbf{P}_{\mathbf{A}} = \mathbf{A}^{1/2}\mathbf{A}^{+1/2} = \mathbf{A}^{+1/2}\mathbf{A}^{1/2}$.

Lemma 4. *Let* $\mathbf{W}, \mathbf{W}_1$ *and* $\mathbf{W}_2$ *be defined as in* (1.11a)–(1.11c). *Then:*

(a) $\mathscr{C}(\mathbf{VM})^{\perp} = \mathscr{C}(\mathbf{WM})^{\perp} = \mathscr{C}(\mathbf{W}^{+}\mathbf{X} : \mathbf{Q}_{\mathbf{W}})$, *where* $\mathbf{Q}_{\mathbf{W}} = \mathbf{I}_n - \mathbf{P}_{\mathbf{W}}$,
(b) $\mathscr{C}(\mathbf{W}^{1/2}\mathbf{M})^{\perp} = \mathscr{C}(\mathbf{W}^{+1/2}\mathbf{X} : \mathbf{Q}_{\mathbf{W}})$,
(c) $\mathscr{C}(\mathbf{W}^{1/2}\mathbf{M}) = \mathscr{C}(\mathbf{W}^{+1/2}\mathbf{X} : \mathbf{Q}_{\mathbf{W}})^{\perp} = \mathscr{C}(\mathbf{W}^{+1/2}\mathbf{X})^{\perp} \cap \mathscr{C}(\mathbf{W})$,
(d) $\mathbf{P}_{\mathbf{W}^{1/2}\mathbf{M}} = \mathbf{P}_{\mathbf{W}} - \mathbf{P}_{\mathbf{W}^{+1/2}\mathbf{X}} = \mathbf{P}_{\mathscr{C}(\mathbf{W}) \cap \mathscr{C}(\mathbf{W}^{+1/2}\mathbf{X})^{\perp}}$.

Moreover, in (a)–(d) *the matrices* $\mathbf{X}$, $\mathbf{M}$ *and* $\mathbf{W}$ *can be replaced with* $\mathbf{X}_i$, $\mathbf{M}_i$ *and* $\mathbf{W}_i$, $i = 1, 2$, *respectively, so that, for example,* (a) *becomes*

(e) $\mathscr{C}(\mathbf{VM}_i)^{\perp} = \mathscr{C}(\mathbf{W}_i\mathbf{M}_i)^{\perp} = \mathscr{C}(\mathbf{W}_i^{+}\mathbf{X}_i : \mathbf{Q}_{\mathbf{W}_i}), \quad i = 1, 2.$

Similarly, reversing the roles of $\mathbf{X}$ *and* $\mathbf{M}$, *the following, for example, holds:*

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

Also the following lemma appears to be useful for our considerations.

Lemma 5. *Consider the partitioned linear model* $\mathscr{M}_{12}$ *and suppose that* $\mathbf{F}$ *is an* $f \times n$ *matrix and* $\mathbf{W} \in \mathscr{W}$. *Then*

(a) $\mathscr{C}(\mathbf{F}'\mathbf{Q}_{\mathbf{FX}_2}) = \mathscr{C}(\mathbf{F}') \cap \mathscr{C}(\mathbf{M}_2)$, *where* $\mathbf{Q}_{\mathbf{FX}_2} = \mathbf{I}_f - \mathbf{P}_{\mathbf{FX}_2}$,
(b) $\mathscr{C}(\mathbf{WF}'\mathbf{Q}_{\mathbf{FX}_2}) = \mathscr{C}(\mathbf{WF}') \cap \mathscr{C}(\mathbf{WM}_2)$,
(c) $\mathscr{C}(\mathbf{W}^{1/2}\mathbf{F}'\mathbf{Q}_{\mathbf{FX}_2}) = \mathscr{C}(\mathbf{W}^{1/2}\mathbf{F}') \cap \mathscr{C}(\mathbf{W}^{1/2}\mathbf{M}_2)$,
(d) $\mathbf{F}'\mathbf{Q}_{\mathbf{FX}_2} = \mathbf{M}_2\mathbf{F}'\mathbf{Q}_{\mathbf{FX}_2}$.

Proof. In light of Rao & Mitra (1971, Complement 7, p. 118), we get

$$\mathscr{C}(\mathbf{F}') \cap \mathscr{C}(\mathbf{M}_2) = \mathscr{C}[\mathbf{F}'(\mathbf{FM}_2^{\perp})^{\perp}] = \mathscr{C}(\mathbf{F}'\mathbf{Q}_{\mathbf{FX}_2}), \tag{1.20}$$

and so (a) is proved. In view of Lemma 4, we have $\mathscr{C}(\mathbf{W}^{1/2}\mathbf{M}_2)^{\perp} = \mathscr{C}(\mathbf{W}^{+1/2}\mathbf{X}_2 : \mathbf{Q}_{\mathbf{W}})$, and hence

$$\begin{aligned}\mathscr{C}(\mathbf{W}^{1/2}\mathbf{F}') \cap \mathscr{C}(\mathbf{W}^{1/2}\mathbf{M}_2) &= \mathscr{C}\{\mathbf{W}^{1/2}\mathbf{F}'[\mathbf{FW}^{1/2}(\mathbf{W}^{1/2}\mathbf{M}_2)^{\perp}]^{\perp}\} \\ &= \mathscr{C}\{\mathbf{W}^{1/2}\mathbf{F}'[\mathbf{FW}^{1/2}(\mathbf{W}^{+1/2}\mathbf{X}_2 : \mathbf{Q}_{\mathbf{W}})]^{\perp}\} \\ &= \mathscr{C}[\mathbf{W}^{1/2}\mathbf{F}'(\mathbf{FX}_2)^{\perp}] = \mathscr{C}(\mathbf{W}^{1/2}\mathbf{F}'\mathbf{Q}_{\mathbf{FX}_2}). \end{aligned} \tag{1.21}$$

Obviously in (1.21) $\mathbf{W}^{1/2}$ can be replaced with $\mathbf{W}$. The statement (d) follows immediately from the inclusion $\mathscr{C}(\mathbf{F}'\mathbf{Q}_{\mathbf{FX}_2}) \subset \mathscr{C}(\mathbf{M}_2)$. □

Next we present an important lemma characterizing the estimability under $\mathscr{M}_{12}$ and $\mathscr{M}_t$.

Lemma 6. *Consider the models* $\mathscr{M}_{12}$ *and its transformed version* $\mathscr{M}_t$ *and let* $\mathbf{F}$ *be an* $f \times n$ *matrix. Then the followings statements hold:*

(a) $\mathbf{X}\beta$ *is estimable under* $\mathscr{M}_t$ *if and only if*

$$\mathscr{C}(\mathbf{X}') = \mathscr{C}(\mathbf{X}'\mathbf{F}'), \quad \textit{i.e.,} \quad \mathscr{C}(\mathbf{X}) \cap \mathscr{C}(\mathbf{F}')^{\perp} = \{\mathbf{0}\}. \tag{1.22}$$

(b) $\mathbf{X}_1\beta_1$ *is estimable under* $\mathscr{M}_{12}$ *if and only if*

$$\mathscr{C}(\mathbf{X}_1') = \mathscr{C}(\mathbf{X}_1'\mathbf{M}_2), \quad \textit{i.e.,} \quad \mathscr{C}(\mathbf{X}_1) \cap \mathscr{C}(\mathbf{X}_2) = \{\mathbf{0}\}. \tag{1.23}$$

(c) $\mathbf{X}_1\beta_1$ *is estimable under* $\mathscr{M}_t$ *if and only if*

$$\mathscr{C}(\mathbf{X}_1') = \mathscr{C}(\mathbf{X}_1'\mathbf{F}'\mathbf{Q}_{\mathbf{F}\mathbf{X}_2}), \tag{1.24}$$

or, equivalently, if and only if

$$\mathscr{C}(\mathbf{X}_1') = \mathscr{C}(\mathbf{X}_1'\mathbf{F}') \quad \textit{and} \quad \mathscr{C}(\mathbf{F}\mathbf{X}_1) \cap \mathscr{C}(\mathbf{F}\mathbf{X}_2) = \{\mathbf{0}\}. \tag{1.25}$$

(d) β *is estimable under* $\mathscr{M}_{12}$ *if and only if* $\mathrm{r}(\mathbf{X}) = p$.
(e) β_1 *is estimable under* $\mathscr{M}_{12}$ *if and only if* $\mathrm{r}(\mathbf{X}_1'\mathbf{M}_2) = p_1$.
(f) β_1 *is estimable under* $\mathscr{M}_t$ *if and only if* $\mathrm{r}(\mathbf{X}_1'\mathbf{F}'\mathbf{Q}_{\mathbf{F}\mathbf{X}_2}) = \mathrm{r}(\mathbf{X}_1) = p_1$.

Proof. In view of (1.6), $\mathbf{X}\beta$ is estimable under $\mathscr{M}_t$ if and only if $\mathscr{C}(\mathbf{X}') \subset \mathscr{C}(\mathbf{X}'\mathbf{F}')$, i.e., $\mathscr{C}(\mathbf{X}') = \mathscr{C}(\mathbf{X}'\mathbf{F}')$. The alternative claim in (a) follows from

$$\mathrm{r}(\mathbf{F}\mathbf{X}) = \mathrm{r}(\mathbf{X}) - \mathscr{C}(\mathbf{X}) \cap \mathscr{C}(\mathbf{F}')^{\perp}, \tag{1.26}$$

where we have used the rank rule of Marsaglia & Styan (1974, Cor. 6.2) for the matrix product. For the claim (b), see, e.g., Puntanen et al. (2011, §16.1). To prove (c), we observe that $\mathbf{X}_1\beta_1 = (\mathbf{X}_1 : \mathbf{0})\beta$ is estimable under $\mathscr{M}_t$ if and only if

$$\mathscr{C}\begin{pmatrix}\mathbf{X}_1' \\ \mathbf{0}\end{pmatrix} \subset \mathscr{C}\begin{pmatrix}\mathbf{X}_1'\mathbf{F}' \\ \mathbf{X}_2'\mathbf{F}'\end{pmatrix}, \text{ i.e., } \mathbf{X}_1' = \mathbf{X}_1'\mathbf{F}'\mathbf{A} \text{ and } \mathbf{0} = \mathbf{X}_2'\mathbf{F}'\mathbf{A}, \tag{1.27}$$

for some $\mathbf{A}$. The equality $\mathbf{0} = \mathbf{X}_2'\mathbf{F}'\mathbf{A}$ means that $\mathbf{A} = \mathbf{Q}_{\mathbf{F}\mathbf{X}_2}\mathbf{B}$ for some $\mathbf{B}$, and thereby $\mathbf{X}_1' = \mathbf{X}_1'\mathbf{F}'\mathbf{Q}_{\mathbf{F}\mathbf{X}_2}\mathbf{B}$ which holds if and only if $\mathscr{C}(\mathbf{X}_1') = \mathscr{C}(\mathbf{X}_1'\mathbf{F}'\mathbf{Q}_{\mathbf{F}\mathbf{X}_2})$. Thus we have proved condition (1.24). Notice that (1.24) is equivalent to

$$\begin{aligned}\mathrm{r}(\mathbf{X}_1') = \mathrm{r}(\mathbf{X}_1'\mathbf{F}'\mathbf{Q}_{\mathbf{F}\mathbf{X}_2}) &= \mathrm{r}(\mathbf{X}_1'\mathbf{F}') - \dim\mathscr{C}(\mathbf{F}\mathbf{X}_1) \cap \mathscr{C}(\mathbf{F}\mathbf{X}_2) \\ &= \mathrm{r}(\mathbf{X}_1) - \dim\mathscr{C}(\mathbf{X}_1) \cap \mathscr{C}(\mathbf{F}')^{\perp} - \dim\mathscr{C}(\mathbf{F}\mathbf{X}_1) \cap \mathscr{C}(\mathbf{F}\mathbf{X}_2),\end{aligned} \tag{1.28}$$

which confirms (1.25). The proofs of (d)–(f) are obvious. □

For the proof Lemma 7, see, e.g., Puntanen et al. (2011, p. 152).

Lemma 7. *The following three statements are equivalent:*

$$\mathbf{P}_{\mathbf{A}} - \mathbf{P}_{\mathbf{B}} \textit{ is an orthogonal projector}, \;\; \mathbf{P}_{\mathbf{A}} - \mathbf{P}_{\mathbf{B}} \geq_{\mathrm{L}} \mathbf{0}, \;\; \mathscr{C}(\mathbf{B}) \subset \mathscr{C}(\mathbf{A}). \tag{1.29}$$

If any of the above conditions holds then $\mathbf{P}_{\mathbf{A}} - \mathbf{P}_{\mathbf{B}} = \mathbf{P}_{\mathscr{C}(\mathbf{A}) \cap \mathscr{C}(\mathbf{B})^{\perp}} = \mathbf{P}_{(\mathbf{I}-\mathbf{P}_{\mathbf{B}})\mathbf{A}}$.

1.3 Linearly sufficient statistic for $\mu = \mathbf{X}\beta$ in $\mathscr{M}_{12}$

Let us consider a partitioned linear model $\mathscr{M}_{12} = \{\mathbf{y}, \mathbf{X}_1\beta_1 + \mathbf{X}_2\beta_2, \mathbf{V}\}$, and its transformed version $\mathscr{M}_t = \{\mathbf{F}\mathbf{y}, \mathbf{F}\mathbf{X}_1\beta_1 + \mathbf{F}\mathbf{X}_2\beta_2, \mathbf{F}\mathbf{V}\mathbf{F}'\}$. Choosing $\mathbf{W} = \mathbf{V} + \mathbf{X}\mathbf{U}\mathbf{U}'\mathbf{X}' \in \mathscr{W}$, we have, for example, the following representations for the covariance matrix of the BLUE for $\mu = \mathbf{X}\beta$:

$$\begin{aligned}\operatorname{cov}(\tilde{\mu} \mid \mathscr{M}_{12}) &= \mathbf{V} - \mathbf{V}\mathbf{M}(\mathbf{M}\mathbf{V}\mathbf{M})^{-}\mathbf{M}\mathbf{V} = \mathbf{W} - \mathbf{W}\mathbf{M}(\mathbf{M}\mathbf{W}\mathbf{M})^{-}\mathbf{M}\mathbf{W} - \mathbf{T} \\ &= \mathbf{W}^{1/2}(\mathbf{I}_n - \mathbf{P}_{\mathbf{W}^{1/2}\mathbf{M}})\mathbf{W}^{1/2} - \mathbf{T} = \mathbf{W}^{1/2}\mathbf{P}_{\mathbf{W}^{+1/2}\mathbf{X}}\mathbf{W}^{1/2} - \mathbf{T} \\ &= \mathbf{X}(\mathbf{X}'\mathbf{W}^{+}\mathbf{X})^{-}\mathbf{X}' - \mathbf{T} = \mathbf{X}(\mathbf{X}'\mathbf{W}^{+1/2}\mathbf{W}^{+1/2}\mathbf{X})^{-}\mathbf{X}' - \mathbf{T}, \end{aligned} \tag{1.30}$$

where $\mathbf{T} = \mathbf{X}\mathbf{U}\mathbf{U}'\mathbf{X}'$. Above we have used Lemma 4d which gives

$$\mathbf{I}_n - \mathbf{P}_{\mathbf{W}^{1/2}\mathbf{M}} = \mathbf{Q}_{\mathbf{W}} + \mathbf{P}_{\mathbf{W}^{+1/2}\mathbf{X}}. \tag{1.31}$$

Consider then the transformed model $\mathscr{M}_t$ and assume that $\mathbf{X}\beta$ is estimable under $\mathscr{M}_t$, i.e., (1.22) holds. Under $\mathscr{M}_t$ we can choose the $\mathbf{W}$-matrix as

$$\mathbf{W}_{\mathscr{M}_t} = \mathbf{F}\mathbf{V}\mathbf{F}' + \mathbf{F}\mathbf{X}\mathbf{U}\mathbf{U}'\mathbf{X}'\mathbf{F}' = \mathbf{F}\mathbf{W}\mathbf{F}' \in \mathscr{W}_{\mathscr{M}_t}, \tag{1.32}$$

and so, denoting $\mathbf{T} = \mathbf{X}\mathbf{U}\mathbf{U}'\mathbf{X}$, we have

$$\begin{aligned}\tilde{\mu}(\mathscr{M}_t) &= \operatorname{BLUE}(\mathbf{X}\beta \mid \mathscr{M}_t) =: \mathbf{G}_t\mathbf{y} \\ &= \mathbf{X}[\mathbf{X}'\mathbf{F}'(\mathbf{F}\mathbf{W}\mathbf{F}')^{-}\mathbf{F}\mathbf{X}]^{-}\mathbf{X}'\mathbf{F}'(\mathbf{F}\mathbf{W}\mathbf{F}')^{-}\mathbf{F}\mathbf{y}, \end{aligned} \tag{1.33}$$

$$\begin{aligned}\operatorname{cov}(\tilde{\mu} \mid \mathscr{M}_t) &= \mathbf{X}[\mathbf{X}'\mathbf{F}'(\mathbf{F}\mathbf{W}\mathbf{F}')^{-}\mathbf{F}\mathbf{X}]^{-}\mathbf{X}' - \mathbf{T} \\ &= \mathbf{X}(\mathbf{X}'\mathbf{W}^{+1/2}\mathbf{P}_{\mathbf{W}^{1/2}\mathbf{F}'}\mathbf{W}^{+1/2}\mathbf{X})^{-}\mathbf{X}' - \mathbf{T}. \end{aligned} \tag{1.34}$$

Of course, by the definition of the BLUE, we always have the Löwner ordering

$$\operatorname{cov}(\tilde{\mu} \mid \mathscr{M}_{12}) \leq_{\mathrm{L}} \operatorname{cov}(\tilde{\mu} \mid \mathscr{M}_t). \tag{1.35}$$

However, it is of interest to confirm (1.35) algebraically. To do this we see at once that

$$\mathbf{X}'\mathbf{W}^{+1/2}\mathbf{W}^{+1/2}\mathbf{X} \geq_{\mathrm{L}} \mathbf{X}'\mathbf{W}^{+1/2}\mathbf{P}_{\mathbf{W}^{1/2}\mathbf{F}'}\mathbf{W}^{+1/2}\mathbf{X}. \tag{1.36}$$

Now (1.36) is equivalent to

$$(\mathbf{X}'\mathbf{W}^{+1/2}\mathbf{W}^{+1/2}\mathbf{X})^{+} \leq_{\mathrm{L}} (\mathbf{X}'\mathbf{W}^{+1/2}\mathbf{P}_{\mathbf{W}^{1/2}\mathbf{F}'}\mathbf{W}^{+1/2}\mathbf{X})^{+}. \tag{1.37}$$

Notice that the equivalence of (1.36) and (1.37) holds in view of the following result: Let $\mathbf{0} \leq_{\mathrm{L}} \mathbf{A} \leq_{\mathrm{L}} \mathbf{B}$. Then $\mathbf{A}^{+} \geq_{\mathrm{L}} \mathbf{B}^{+}$ if and only if $\mathrm{r}(\mathbf{A}) = \mathrm{r}(\mathbf{B})$; see Milliken & Akdeniz (1977). Now $\mathrm{r}(\mathbf{X}'\mathbf{W}^{+}\mathbf{X}) = \mathrm{r}(\mathbf{X}'\mathbf{W}) = \mathrm{r}(\mathbf{X})$, and

$$\begin{aligned}\mathrm{r}(\mathbf{X}'\mathbf{W}^{+1/2}\mathbf{P}_{\mathbf{W}^{1/2}\mathbf{F}'}\mathbf{W}^{+1/2}\mathbf{X}) &= \mathrm{r}(\mathbf{X}'\mathbf{W}^{+1/2}\mathbf{P}_{\mathbf{W}^{1/2}\mathbf{F}'}) \\ &= \mathrm{r}(\mathbf{X}'\mathbf{P}_{\mathbf{W}}\mathbf{F}') = \mathrm{r}(\mathbf{X}'\mathbf{F}') = \mathrm{r}(\mathbf{X})\,, \end{aligned} \tag{1.38}$$

where the last equality follows from the estimability condition (1.25). Now (1.37) implies

$$\mathbf{X}(\mathbf{X}'\mathbf{W}^{+1/2}\mathbf{W}^{+1/2}\mathbf{X})^{-}\mathbf{X}' \leq_{\mathrm{L}} \mathbf{X}(\mathbf{X}'\mathbf{W}^{+1/2}\mathbf{P}_{\mathbf{W}^{1/2}\mathbf{F}'}\mathbf{W}^{+1/2}\mathbf{X})^{-}\mathbf{X}', \tag{1.39}$$

which is just (1.35).

Now $\mathrm{E}(\mathbf{G}_t\mathbf{y}) = \mathbf{X}\beta$, and hence by Lemma 3, $\mathbf{Fy}$ is linearly sufficient for $\mathbf{X}\beta$ if and only if

$$\mathrm{cov}(\tilde{\mu} \mid \mathscr{M}_{12}) = \mathrm{cov}(\tilde{\mu} \mid \mathscr{M}_t)\,. \tag{1.40}$$

Next we show directly that (1.40) is equivalent to (1.18). First we observe that (1.40) holds if and only if

$$\mathbf{X}(\mathbf{X}'\mathbf{W}^{+1/2}\mathbf{W}^{+1/2}\mathbf{X})^{-}\mathbf{X}' = \mathbf{X}(\mathbf{X}'\mathbf{W}^{+1/2}\mathbf{P}_{\mathbf{W}^{1/2}\mathbf{F}'}\mathbf{W}^{+1/2}\mathbf{X})^{-}\mathbf{X}'. \tag{1.41}$$

Pre- and postmultiplying (1.41) by $\mathbf{X}^{+}$ and by $(\mathbf{X}')^{+}$, respectively, and using the fact that $\mathbf{P}_{\mathbf{X}'} = \mathbf{X}^{+}\mathbf{X}$, gives an equivalent form to (1.41):

$$(\mathbf{X}'\mathbf{W}^{+1/2}\mathbf{W}^{+1/2}\mathbf{X})^{+} = (\mathbf{X}'\mathbf{W}^{+1/2}\mathbf{P}_{\mathbf{W}^{1/2}\mathbf{F}'}\mathbf{W}^{+1/2}\mathbf{X})^{+}. \tag{1.42}$$

Obviously (1.42) holds if and only if $\mathscr{C}(\mathbf{W}^{+1/2}\mathbf{X}) \subset \mathscr{C}(\mathbf{W}^{1/2}\mathbf{F}')$, which further is equivalent to

$$\mathscr{C}(\mathbf{X}) \subset \mathscr{C}(\mathbf{WF}')\,, \tag{1.43}$$

which is precisely the condition (1.18) for $\mathbf{Fy}$ being linearly sufficient for $\mathbf{X}\beta$. As a summary we can write the following:

Theorem 1. *Let $\mu = \mathbf{X}\beta$ be estimable under $\mathscr{M}_t$ and let $\mathbf{W} \in \mathscr{W}$. Then*

$$\mathrm{cov}(\tilde{\mu} \mid \mathscr{M}_{12}) \leq_{\mathrm{L}} \mathrm{cov}(\tilde{\mu} \mid \mathscr{M}_t)\,. \tag{1.44}$$

Moreover, the following statements are equivalent:

(a) $\mathrm{cov}(\tilde{\mu} \mid \mathscr{M}_{12}) = \mathrm{cov}(\tilde{\mu} \mid \mathscr{M}_t)$,
(b) $\mathbf{X}(\mathbf{X}'\mathbf{W}^{+}\mathbf{X})^{-}\mathbf{X}' = \mathbf{X}(\mathbf{X}'\mathbf{W}^{+1/2}\mathbf{P}_{\mathbf{W}^{1/2}\mathbf{F}'}\mathbf{W}^{+1/2}\mathbf{X})^{-}\mathbf{X}$,
(c) $\mathbf{X}'\mathbf{W}^{+}\mathbf{X} = \mathbf{X}'\mathbf{W}^{+1/2}\mathbf{P}_{\mathbf{W}^{1/2}\mathbf{F}'}\mathbf{W}^{+1/2}\mathbf{X}$,
(d) $\mathscr{C}(\mathbf{W}^{+1/2}\mathbf{X}) \subset \mathscr{C}(\mathbf{W}^{1/2}\mathbf{F}')$,
(e) $\mathscr{C}(\mathbf{X}) \subset \mathscr{C}(\mathbf{WF}')$,
(f) $\mathbf{Fy}$ *is linearly sufficient for* $\mu = \mathbf{X}\beta$ *under* $\mathscr{M}_{12}$.

1.4 Linearly sufficient statistic for $\mu_1 = \mathbf{X}_1\beta_1$ in $\mathscr{M}_{12}$

Consider then the estimation of $\mu_1 = \mathbf{X}_1\beta_1$ under $\mathscr{M}_{12}$. We assume that (1.23) holds so that μ_1 is estimable under $\mathscr{M}_{12}$. Premultiplying the model $\mathscr{M}_{12}$ by $\mathbf{M}_2 = \mathbf{I}_n - \mathbf{P}_{\mathbf{X}_2}$ yields the reduced model

$$\mathscr{M}_{12\cdot 2} = \{\mathbf{M}_2\mathbf{y},\ \mathbf{M}_2\mathbf{X}_1\beta_1,\ \mathbf{M}_2\mathbf{V}\mathbf{M}_2\}\,. \tag{1.45}$$

Now the well-known Frisch–Waugh–Lovell theorem, see, e.g., Groß & Puntanen (2000), states that the BLUEs of μ_1 under $\mathscr{M}_{12}$ and $\mathscr{M}_{12\cdot 2}$ coincide (with probability 1):

$$\mathrm{BLUE}(\mu_1 \mid \mathscr{M}_{12}) = \mathrm{BLUE}(\mu_1 \mid \mathscr{M}_{12\cdot 2})\,. \tag{1.46}$$

Hence, we immediately see that $\mathbf{M}_2\mathbf{y}$ is linearly sufficient for μ_1.

Now any matrix of the form

$$\mathbf{M}_2\mathbf{V}\mathbf{M}_2 + \mathbf{M}_2\mathbf{X}_1\mathbf{U}_1\mathbf{U}_1'\mathbf{X}_1'\mathbf{M}_2 \tag{1.47}$$

satisfying $\mathscr{C}(\mathbf{M}_2\mathbf{V} : \mathbf{M}_2\mathbf{X}_1\mathbf{U}_1) = \mathscr{C}(\mathbf{M}_2\mathbf{V} : \mathbf{M}_2\mathbf{X}_1)$, is a $\mathbf{W}$-matrix in $\mathscr{M}_{12\cdot 2}$. We may denote this class as $\mathscr{W}_{\mathscr{M}_{12\cdot 2}}$, and

$$\mathbf{W}_{\mathscr{M}_{12\cdot 2}} = \mathbf{M}_2\mathbf{W}\mathbf{M}_2 = \mathbf{M}_2\mathbf{W}_1\mathbf{M}_2 \in \mathscr{W}_{\mathscr{M}_{12\cdot 2}}\,, \tag{1.48}$$

where $\mathbf{W}$ and $\mathbf{W}_1$ are defined as in (1.11a)–(1.11c).

It is interesting to observe that in (1.47) the matrix $\mathbf{U}_1$ can be chosen as a null matrix if and only if

$$\mathscr{C}(\mathbf{M}_2\mathbf{X}_1) \subset \mathscr{C}(\mathbf{M}_2\mathbf{V})\,, \tag{1.49}$$

which can be shown to be equivalent to

$$\mathscr{C}(\mathbf{X}_1) \subset \mathscr{C}(\mathbf{X}_2 : \mathbf{V})\,. \tag{1.50}$$

Namely, it is obvious that (1.50) implies (1.49) while the reverse implication follows from the following:

$$\mathscr{C}(\mathbf{X}_1) \subset \mathscr{C}(\mathbf{X}_1 : \mathbf{X}_2) = \mathscr{C}(\mathbf{X}_2 : \mathbf{M}_2\mathbf{X}_1) \subset \mathscr{C}(\mathbf{X}_2 : \mathbf{M}_2\mathbf{V}) = \mathscr{C}(\mathbf{X}_2 : \mathbf{V})\,. \tag{1.51}$$

This means that

$$\mathbf{M}_2\mathbf{V}\mathbf{M}_2 \in \mathscr{W}_{\mathscr{M}_{12\cdot 2}} \iff \mathscr{C}(\mathbf{X}_1) \subset \mathscr{C}(\mathbf{X}_2 : \mathbf{V})\,. \tag{1.52}$$

One expression for the BLUE of $\mu_1 = \mathbf{X}_1\beta_1$, obtainable from $\mathscr{M}_{12\cdot 2}$, is

$$\mathrm{BLUE}(\mu_1 \mid \mathscr{M}_{12}) = \tilde{\mu}_1(\mathscr{M}_{12}) = \mathbf{X}_1(\mathbf{X}_1'\dot{\mathbf{M}}_{2W}\mathbf{X}_1)^{-}\mathbf{X}_1'\dot{\mathbf{M}}_{2W}\mathbf{y}\,, \tag{1.53}$$

where

$$\dot{\mathbf{M}}_{2W} = \mathbf{M}_2 \mathbf{W}^-_{\mathscr{M}_{12\cdot 2}} \mathbf{M}_2 = \mathbf{M}_2(\mathbf{M}_2\mathbf{W}\mathbf{M}_2)^-\mathbf{M}_2 \,. \tag{1.54}$$

In particular, if (1.50) holds then we can choose $\mathbf{W}_{\mathscr{M}_{12\cdot 2}} = \mathbf{M}_2\mathbf{V}\mathbf{M}_2$, and

$$\dot{\mathbf{M}}_{2W} = \mathbf{M}_2(\mathbf{M}_2\mathbf{V}\mathbf{M}_2)^-\mathbf{M}_2 =: \dot{\mathbf{M}}_2 \,. \tag{1.55}$$

Notice that by Lemma 4d, we have

$$\begin{aligned} \mathbf{P_W}\dot{\mathbf{M}}_{2W}\mathbf{P_W} &= \mathbf{P_W}\mathbf{M}_2(\mathbf{M}_2\mathbf{W}\mathbf{M}_2)^-\mathbf{M}_2\mathbf{P_W} \\ &= \mathbf{W}^{+1/2}\mathbf{P}_{\mathbf{W}^{1/2}\mathbf{M}_2}\mathbf{W}^{+1/2} \\ &= \mathbf{W}^{+1/2}(\mathbf{P_W} - \mathbf{P}_{\mathbf{W}^{+1/2}\mathbf{X}_2})\mathbf{W}^{+1/2} \\ &= \mathbf{W}^+ - \mathbf{W}^+\mathbf{X}_2(\mathbf{X}_2'\mathbf{W}^+\mathbf{X}_2)^-\mathbf{X}_2'\mathbf{W}^+ , \end{aligned} \tag{1.56}$$

and hence, for example,

$$\begin{aligned} \mathbf{W}\dot{\mathbf{M}}_{2W}\mathbf{X}_1 &= \mathbf{W}[\mathbf{W}^+ - \mathbf{W}^+\mathbf{X}_2(\mathbf{X}_2'\mathbf{W}^+\mathbf{X}_2)^-\mathbf{X}_2'\mathbf{W}^+]\mathbf{X}_1 \\ &= [\mathbf{I}_n - \mathbf{X}_2(\mathbf{X}_2'\mathbf{W}^+\mathbf{X}_2)^-\mathbf{X}_2'\mathbf{W}^+]\mathbf{X}_1 \,. \end{aligned} \tag{1.57}$$

Observe that in (1.54), (1.56) and (1.57) the matrix $\mathbf{W}$ can be replaced with $\mathbf{W}_1$. For a thorough review of the properties of $\dot{\mathbf{M}}_{2W}$, see Isotalo et al. (2008).

In the next theorem we collect some interesting properties of linearly sufficient estimators of μ_1.

Theorem 2. *Let $\mu_1 = \mathbf{X}_1\beta_1$ be estimable under $\mathscr{M}_{12}$ and let $\mathbf{W} \in \mathscr{W}$. Then the statistic $\mathbf{Fy}$ is linearly sufficient for μ_1 under $\mathscr{M}_{12}$ if and only if*

$$\mathscr{C}(\mathbf{W}\dot{\mathbf{M}}_{2W}\mathbf{X}_1) \subset \mathscr{C}(\mathbf{W}\mathbf{F}') , \tag{1.58}$$

or, equivalently,

$$\mathscr{C}\{[\mathbf{I}_n - \mathbf{X}_2(\mathbf{X}_2'\mathbf{W}^+\mathbf{X}_2)^-\mathbf{X}_2'\mathbf{W}^+]\mathbf{X}_1\} \subset \mathscr{C}(\mathbf{W}\mathbf{F}') , \tag{1.59}$$

where $\dot{\mathbf{M}}_{2W} = \mathbf{M}_2(\mathbf{M}_2\mathbf{W}\mathbf{M}_2)^-\mathbf{M}_2$. Moreover,

(a) $\mathbf{M}_2\mathbf{y}$ *is linearly sufficient for* μ_1.
(b) $\dot{\mathbf{M}}_{2W}\mathbf{y} = \mathbf{M}_2(\mathbf{M}_2\mathbf{W}\mathbf{M}_2)^-\mathbf{M}_2\mathbf{y}$ *is linearly sufficient for* μ_1.
(c) $\mathbf{X}_1'\dot{\mathbf{M}}_{2W}\mathbf{y}$ *is linearly minimal sufficient for* μ_1.
(d) *If* $\mathscr{C}(\mathbf{X}_1) \subset \mathscr{C}(\mathbf{X}_2 : \mathbf{V})$, (1.58) *becomes*

$$\mathscr{C}(\mathbf{W}\dot{\mathbf{M}}_2\mathbf{X}_1) \subset \mathscr{C}(\mathbf{W}\mathbf{F}') , \quad \text{where } \dot{\mathbf{M}}_2 = \mathbf{M}_2(\mathbf{M}_2\mathbf{V}\mathbf{M}_2)^-\mathbf{M}_2 \,. \tag{1.60}$$

(e) *If* $\mathbf{V}$ *is positive definite,* (1.58) *becomes* $\mathscr{C}(\dot{\mathbf{M}}_2\mathbf{X}_1) \subset \mathscr{C}(\mathbf{F}')$.
(f) *If* β_1 *is estimable under* $\mathscr{M}_{12}$, *then*

$$\mathbf{Fy} \in \mathrm{S}(\mathbf{X}_1\beta_1 \mid \mathscr{M}_{12}) \iff \mathbf{Fy} \in \mathrm{S}(\beta_1 \mid \mathscr{M}_{12}) \,. \tag{1.61}$$

Proof. The sufficiency condition (1.58) was proved by Kala et al. (2017, §3), and, using a different approach, by Isotalo & Puntanen (2006, Th. 2). Claims (a), (b), (c) and (e) are straightforward to confirm and (d) was considered already before the Theorem. Let us confirm part (f). If $\mathbf{Fy} \in \mathrm{S}(\mathbf{X}_1\beta_1 \mid \mathscr{M}_{12})$, then there exists a matrix $\mathbf{A}$ such that

$$\mathbf{AF}(\mathbf{X}_1 : \mathbf{X}_2 : \mathbf{VM}) = (\mathbf{X}_1 : \mathbf{0} : \mathbf{0})\,. \tag{1.62}$$

Because of the estimability of β_1, the matrix $\mathbf{X}_1$ has a full column rank. Premultiplying (1.62) by $(\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'$ yields

$$\mathbf{BF}(\mathbf{X}_1 : \mathbf{X}_2 : \mathbf{VM}) = (\mathbf{I}_{p_1} : \mathbf{0} : \mathbf{0})\,, \tag{1.63}$$

where $\mathbf{B} = (\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{A}$, and thereby $\mathbf{Fy} \in \mathrm{S}(\mathbf{X}_1\beta_1 \mid \mathscr{M}_{12})$ implies $\mathbf{Fy} \in \mathrm{S}(\beta_1 \mid \mathscr{M}_{12})$. The reverse direction can be proved in the corresponding way. Thus we have confirmed that claim (e) indeed holds. □

The covariance matrix of the BLUE of $\mu_1 = \mathbf{X}_1\beta_1$ under $\mathscr{M}_{12}$ can be expressed as

$$\begin{aligned}\mathrm{cov}(\tilde{\mu}_1 \mid \mathscr{M}_{12}) &= \mathbf{X}_1(\mathbf{X}_1'\dot{\mathbf{M}}_{2W}\mathbf{X}_1)^-\mathbf{X}_1' - \mathbf{T}_1\\ &= \mathbf{X}_1[\mathbf{X}_1'\mathbf{M}_2(\mathbf{M}_2\mathbf{W}\mathbf{M}_2)^-\mathbf{M}_2\mathbf{X}_1]^-\mathbf{X}_1' - \mathbf{T}_1\\ &= \mathbf{X}_1[\mathbf{X}_1'\mathbf{W}^{+1/2}\mathbf{P}_{\mathbf{W}^{1/2}\mathbf{M}_2}\mathbf{W}^{+1/2}\mathbf{X}_1]^-\mathbf{X}_1' - \mathbf{T}_1\,,\end{aligned} \tag{1.64}$$

where $\mathbf{T}_1 = \mathbf{X}_1\mathbf{U}_1\mathbf{U}_1'\mathbf{X}_1'$ and $\mathbf{W}$ can be replaced with $\mathbf{W}_1$.

Remark 1. The rank of the covariance matrix of the BLUE(β), as well as that of BLUE($\mathbf{X}\beta$), under $\mathscr{M}_{12}$ is

$$\mathrm{r}[\mathrm{cov}(\tilde{\beta} \mid \mathscr{M}_{12})] = \dim \mathscr{C}(\mathbf{X}) \cap \mathscr{C}(\mathbf{V})\,; \tag{1.65}$$

see, e.g., Puntanen et al. (2011, p. 137). Hence for estimable β,

$$\mathscr{C}(\mathbf{X}) \subset \mathscr{C}(\mathbf{V}) \iff \mathrm{cov}(\tilde{\beta} \mid \mathscr{M}_{12}) \text{ is positive definite.} \tag{1.66}$$

Similarly, for estimable β_1,

$$\begin{aligned}\mathrm{r}[\mathrm{cov}(\tilde{\beta}_1 \mid \mathscr{M}_{12})] = \mathrm{r}[\mathrm{cov}(\tilde{\beta}_1 \mid \mathscr{M}_{12\cdot 2})] &= \dim \mathscr{C}(\mathbf{M}_2\mathbf{X}_1) \cap \mathscr{C}(\mathbf{M}_2\mathbf{V}\mathbf{M}_2)\\ &= \dim \mathscr{C}(\mathbf{M}_2\mathbf{X}_1) \cap \mathscr{C}(\mathbf{M}_2\mathbf{V}) \le \mathrm{r}(\mathbf{M}_2\mathbf{X}_1)\,.\end{aligned} \tag{1.67}$$

The estimability of β_1 means that $\mathrm{r}(\mathbf{M}_2\mathbf{X}_1) = p_1$ and thereby

$$\mathrm{r}[\mathrm{cov}(\tilde{\beta}_1 \mid \mathscr{M}_{12})] = p_1 \iff \mathscr{C}(\mathbf{M}_2\mathbf{X}_1) \subset \mathscr{C}(\mathbf{M}_2\mathbf{V})\,. \tag{1.68}$$

Thus, by the equivalence of (1.49) and (1.50), for estimable β_1 the following holds:

$$\mathscr{C}(\mathbf{X}_1) \subset \mathscr{C}(\mathbf{X}_2 : \mathbf{V}) \iff \mathrm{cov}(\tilde{\beta}_1 \mid \mathscr{M}_{12}) \text{ is positive definite.} \quad \square \tag{1.69}$$

What is the covariance matrix of the BLUE of $\mu_1 = \mathbf{X}_1\beta_1$ under $\mathscr{M}_t$? First we need to make sure that $\mathbf{X}_1\beta_1$ is estimable under $\mathscr{M}_t$, i.e., (1.24) holds.

Let us eliminate the $\mathbf{FX}_2\beta_2$-part by premultiplying $\mathscr{M}_t$ by $\mathbf{Q_{FX_2}} = \mathbf{I}_f - \mathbf{P_{FX_2}}$. Thus we obtain the reduced transformed model

$$\begin{aligned}\mathscr{M}_{t\cdot 2} &= \{\mathbf{Q_{FX_2}Fy}, \mathbf{Q_{FX_2}FX}_1\beta_1, \mathbf{Q_{FX_2}FVF'Q_{FX_2}}\} \\ &= \{\mathbf{N'y}, \mathbf{N'X}_1\beta_1, \mathbf{N'VN}\}\,, \end{aligned} \tag{1.70}$$

where $\mathbf{N} = \mathbf{F'Q_{FX_2}} \in \mathbb{R}^{n\times f}$, and, see Lemma 5, the matrix $\mathbf{N}$ has the property

$$\mathscr{C}(\mathbf{N}) = \mathscr{C}(\mathbf{F'}) \cap \mathscr{C}(\mathbf{M}_2)\,. \tag{1.71}$$

Notice also that in view of (1.71) and part (c) of Lemma 6,

$$\begin{aligned}\mathrm{r}(\mathbf{N'X}_1) &= \mathrm{r}(\mathbf{X}_1) - \dim\mathscr{C}(\mathbf{X}_1)\cap\mathscr{C}(\mathbf{N})^{\perp} \\ &= \mathrm{r}(\mathbf{X}_1) - \dim\mathscr{C}(\mathbf{X}_1)\cap\mathscr{C}[(\mathbf{F'})^{\perp} : \mathbf{X}_2] = \mathrm{r}(\mathbf{X}_1)\,,\end{aligned} \tag{1.72}$$

so that

$$\mathrm{r}(\mathbf{X}_1'\mathbf{W}^{+1/2}\mathbf{P}_{\mathbf{W}^{1/2}\mathbf{N}}\mathbf{W}^{+1/2}\mathbf{X}_1) = \mathrm{r}(\mathbf{X}_1'\mathbf{W}^{+1/2}\mathbf{W}^{1/2}\mathbf{N}) = \mathrm{r}(\mathbf{X}_1'\mathbf{N}) = \mathrm{r}(\mathbf{X}_1)\,. \tag{1.73}$$

Correspondingly, we have

$$\mathrm{r}(\mathbf{X}_1'\mathbf{W}^{+1/2}\mathbf{P}_{\mathbf{W}^{1/2}\mathbf{M}_2}\mathbf{W}^{+1/2}\mathbf{X}_1) = \mathrm{r}(\mathbf{X}_1)\,. \tag{1.74}$$

The $\mathbf{W}$-matrix under $\mathscr{M}_{t\cdot 2}$ can be chosen as

$$\mathbf{W}_{\mathscr{M}_{t\cdot 2}} = \mathbf{Q_{FX_2}FW}_1\mathbf{F'Q_{FX_2}} = \mathbf{N'W}_1\mathbf{N}\,, \tag{1.75}$$

where $\mathbf{W}_1$ can be replaced with $\mathbf{W}$. In view of the Frisch–Waugh–Lowell theorem, the BLUE of $\mu_1 = \mathbf{X}_1\beta_1$ is

$$\tilde{\mu}_1(\mathscr{M}_t) = \tilde{\mu}_1(\mathscr{M}_{t\cdot 2}) = \mathbf{X}_1[\mathbf{X}_1'\mathbf{N}(\mathbf{N'WN})^-\mathbf{N'X}_1]^-\mathbf{X}_1'\mathbf{N}(\mathbf{N'WN})^-\mathbf{N'y}\,, \tag{1.76}$$

while the corresponding covariance matrix is

$$\begin{aligned}\mathrm{cov}(\tilde{\mu}_1 \mid \mathscr{M}_t) &= \mathbf{X}_1[\mathbf{X}_1'\mathbf{N}(\mathbf{N'WN})^-\mathbf{N'X}_1]^-\mathbf{X}_1' - \mathbf{T}_1 \\ &= \mathbf{X}_1(\mathbf{X}_1'\mathbf{W}^{+1/2}\mathbf{P}_{\mathbf{W}^{1/2}\mathbf{N}}\mathbf{W}^{+1/2}\mathbf{X}_1)^-\mathbf{X}_1' - \mathbf{T}_1\,,\end{aligned} \tag{1.77}$$

where $\mathbf{T}_1 = \mathbf{X}_1\mathbf{U}_1\mathbf{U}_1'\mathbf{X}_1'$ (and $\mathbf{W}$ can be replaced with $\mathbf{W}_1$).

By definition we of course have

$$\mathrm{cov}(\tilde{\mu}_1 \mid \mathscr{M}_{12}) \leq_{\mathrm{L}} \mathrm{cov}(\tilde{\mu}_1 \mid \mathscr{M}_t)\,, \tag{1.78}$$

but it is illustrative to confirm this also algebraically. First we observe that in view of Lemma 5,

$$\mathscr{C}(\mathbf{W}^{1/2}\mathbf{F}'\mathbf{Q}_{\mathbf{FX}_2}) = \mathscr{C}(\mathbf{W}^{1/2}\mathbf{F}') \cap \mathscr{C}(\mathbf{W}^{1/2}\mathbf{M}_2)\,, \tag{1.79}$$

and thereby Lemma 7 implies that $\mathbf{P}_{\mathbf{W}^{1/2}\mathbf{M}_2} - \mathbf{P}_{\mathbf{W}^{1/2}\mathbf{F}'\mathbf{Q}_{\mathbf{FX}_2}} = \mathbf{P}_{\mathbf{Z}}$, where

$$\mathscr{C}(\mathbf{Z}) = \mathscr{C}(\mathbf{W}^{1/2}\mathbf{M}_2) \cap \mathscr{C}(\mathbf{W}^{+1/2}\mathbf{F}'\mathbf{Q}_{\mathbf{FX}_2})^{\perp}. \tag{1.80}$$

Hence we have the following equivalent inequalities:

$$\mathbf{X}_1'\mathbf{W}^{+1/2}(\mathbf{P}_{\mathbf{W}^{1/2}\mathbf{M}_2} - \mathbf{P}_{\mathbf{W}^{1/2}\mathbf{F}'\mathbf{Q}_{\mathbf{FX}_2}})\mathbf{W}^{+1/2}\mathbf{X}_1 \geq_{\mathrm{L}} \mathbf{0}\,, \tag{1.81}$$

$$\mathbf{X}_1'\mathbf{W}^{+1/2}\mathbf{P}_{\mathbf{W}^{1/2}\mathbf{M}_2}\mathbf{W}^{+1/2}\mathbf{X}_1 \geq_{\mathrm{L}} \mathbf{X}_1'\mathbf{W}^{+1/2}\mathbf{P}_{\mathbf{W}^{1/2}\mathbf{F}'\mathbf{Q}_{\mathbf{FX}_2}}\mathbf{W}^{+1/2}\mathbf{X}_1\,, \tag{1.82}$$

$$(\mathbf{X}_1'\mathbf{W}^{+1/2}\mathbf{P}_{\mathbf{W}^{1/2}\mathbf{M}_2}\mathbf{W}^{+1/2}\mathbf{X}_1)^{+} \leq_{\mathrm{L}} (\mathbf{X}_1'\mathbf{W}^{+1/2}\mathbf{P}_{\mathbf{W}^{1/2}\mathbf{F}'\mathbf{Q}_{\mathbf{FX}_2}}\mathbf{W}^{+1/2}\mathbf{X}_1)^{+}. \tag{1.83}$$

The equivalence between (1.82) and (1.83) is due to the fact that the matrices on each side of (1.82) have the same rank, which is $\mathrm{r}(\mathbf{X}_1)$; see (1.73) and (1.74). The equivalence between (1.83) and (1.78) follows by the same argument as that between (1.41) and (1.42).

The equality in (1.82) holds if and only if

$$\mathbf{P}_{\mathbf{Z}}\mathbf{W}^{+1/2}\mathbf{X}_1 = (\mathbf{P}_{\mathbf{W}^{1/2}\mathbf{M}_2} - \mathbf{P}_{\mathbf{W}^{1/2}\mathbf{F}'\mathbf{Q}_{\mathbf{FX}_2}})\mathbf{W}^{+1/2}\mathbf{X}_1 = \mathbf{0}\,, \tag{1.84}$$

which is equivalent to

$$\mathbf{W}^{1/2}(\mathbf{P}_{\mathbf{W}^{1/2}\mathbf{M}_2} - \mathbf{P}_{\mathbf{W}^{1/2}\mathbf{F}'\mathbf{Q}_{\mathbf{FX}_2}})\mathbf{W}^{+1/2}\mathbf{X}_1 = \mathbf{0}\,. \tag{1.85}$$

Writing up (1.85) yields

$$\mathbf{W}\dot{\mathbf{M}}_{2W}\mathbf{X}_1 = \mathbf{W}\mathbf{M}_2(\mathbf{M}_2\mathbf{W}\mathbf{M}_2)^{-}\mathbf{M}_2\mathbf{X}_1 = \mathbf{W}\mathbf{N}(\mathbf{N}'\mathbf{W}\mathbf{N})^{-}\mathbf{N}\mathbf{X}_1\,. \tag{1.86}$$

We observe that in view of (1.80) we have

$$\mathscr{C}(\mathbf{Z})^{\perp} = \mathscr{C}(\mathbf{W}^{+1/2}\mathbf{X}_2 : \mathbf{Q}_{\mathbf{W}} : \mathbf{W}^{1/2}\mathbf{F}'\mathbf{Q}_{\mathbf{FX}_2})\,, \tag{1.87}$$

where we have used the Lemma 4 giving us $\mathscr{C}(\mathbf{W}^{1/2}\mathbf{M}_2)^{\perp} = \mathscr{C}(\mathbf{W}^{+1/2}\mathbf{X}_2 : \mathbf{Q}_{\mathbf{W}})$. Therefore (1.84) holds if and only if

$$\mathscr{C}(\mathbf{W}^{+1/2}\mathbf{X}_1) \subset \mathscr{C}(\mathbf{Z})^{\perp} = \mathscr{C}(\mathbf{W}^{+1/2}\mathbf{X}_2 : \mathbf{Q}_{\mathbf{W}} : \mathbf{W}^{1/2}\mathbf{F}'\mathbf{Q}_{\mathbf{FX}_2})\,. \tag{1.88}$$

Premultiplying the above inclusion by $\mathbf{W}^{1/2}$ yields an equivalent condition:

$$\begin{aligned}\mathscr{C}(\mathbf{X}_1) \subset \mathscr{C}(\mathbf{X}_2 : \mathbf{W}\mathbf{F}'\mathbf{Q}_{\mathbf{FX}_2}) &= \mathscr{C}(\mathbf{X}_2 : \mathbf{M}_2\mathbf{W}\mathbf{F}'\mathbf{Q}_{\mathbf{FX}_2}) \\ &= \mathscr{C}(\mathbf{X}_2) \oplus [\mathscr{C}(\mathbf{W}\mathbf{F}') \cap \mathscr{C}(\mathbf{W}\mathbf{M}_2)]\,.\end{aligned} \tag{1.89}$$

Our next step is to prove the equivalence of (1.89) and the linear sufficiency condition

$$\mathscr{C}(\mathbf{W}\dot{\mathbf{M}}_{2W}\mathbf{X}_1) \subset \mathscr{C}(\mathbf{W}\mathbf{F}')\,. \tag{1.90}$$

The equality (1.85), which is equivalent to (1.89), immediately implies (1.90). To go the other way, we observe that (1.90) implies

$$\mathscr{C}[\mathbf{W}\mathbf{M}_2(\mathbf{M}_2\mathbf{W}\mathbf{M}_2)^-\mathbf{M}_2\mathbf{X}_1] \subset \mathscr{C}(\mathbf{W}\mathbf{F}') \cap \mathscr{C}(\mathbf{W}\mathbf{M}_2) = \mathscr{C}(\mathbf{W}\mathbf{F}'\mathbf{Q}_{\mathbf{F}\mathbf{X}_2})\,, \tag{1.91}$$

where we have used Lemma 5. Premultiplying (1.91) by $\mathbf{M}_2$ and noting that

$$\mathbf{M}_2\mathbf{W}\mathbf{M}_2(\mathbf{M}_2\mathbf{W}\mathbf{M}_2)^+ = \mathbf{P}_{\mathbf{M}_2\mathbf{W}} \tag{1.92}$$

yields

$$\mathscr{C}(\mathbf{P}_{\mathbf{M}_2\mathbf{W}}\mathbf{M}_2\mathbf{X}_1) = \mathscr{C}(\mathbf{M}_2\mathbf{X}_1) \subset \mathscr{C}(\mathbf{M}_2\mathbf{W}\mathbf{F}'\mathbf{Q}_{\mathbf{F}\mathbf{X}_2})\,. \tag{1.93}$$

Using (1.93) we get

$$\mathscr{C}(\mathbf{X}_1) \subset \mathscr{C}(\mathbf{X}_1 : \mathbf{X}_2) = \mathscr{C}(\mathbf{X}_2 : \mathbf{M}_2\mathbf{X}_1) \subset \mathscr{C}(\mathbf{X}_2 : \mathbf{M}_2\mathbf{W}\mathbf{F}'\mathbf{Q}_{\mathbf{F}\mathbf{X}_2})\,, \tag{1.94}$$

and thus we have shown that (1.90) implies (1.89).

Now we can summarise our findings for further equivalent conditions for $\mathbf{Fy}$ being linearly sufficient for $\mathbf{X}_1\beta_1$:

Theorem 3. *Let* $\mu_1 = \mathbf{X}_1\beta_1$ *be estimable under* $\mathscr{M}_{12}$ *and* $\mathscr{M}_t$ *and let* $\mathbf{W} \in \mathscr{W}$. *Then*

$$\operatorname{cov}(\tilde{\mu}_1 \mid \mathscr{M}_{12}) \leq_{\mathrm{L}} \operatorname{cov}(\tilde{\mu}_1 \mid \mathscr{M}_t)\,. \tag{1.95}$$

Moreover, the following statements are equivalent:

(a) $\operatorname{cov}(\tilde{\mu}_1 \mid \mathscr{M}_{12}) = \operatorname{cov}(\tilde{\mu}_1 \mid \mathscr{M}_t)$.
(b) $\mathscr{C}(\mathbf{W}\dot{\mathbf{M}}_{2W}\mathbf{X}_1) \subset \mathscr{C}(\mathbf{W}\mathbf{F}')$.
(c) $\mathscr{C}(\mathbf{X}_1) \subset \mathscr{C}(\mathbf{X}_2 : \mathbf{W}\mathbf{F}'\mathbf{Q}_{\mathbf{F}\mathbf{X}_2}) = \mathscr{C}(\mathbf{X}_2 : \mathbf{M}_2\mathbf{W}\mathbf{F}'\mathbf{Q}_{\mathbf{F}\mathbf{X}_2})$.
(d) $\mathscr{C}(\mathbf{X}_1) \subset \mathscr{C}(\mathbf{X}_2) \oplus [\mathscr{C}(\mathbf{W}\mathbf{F}') \cap \mathscr{C}(\mathbf{W}\mathbf{M}_2)]$.
(e) $\mathbf{W}\mathbf{M}_2(\mathbf{M}_2\mathbf{W}\mathbf{M}_2)^-\mathbf{M}_2\mathbf{X}_1 = \mathbf{W}\mathbf{N}(\mathbf{N}'\mathbf{W}\mathbf{N})^-\mathbf{N}\mathbf{X}_1$, *where* $\mathbf{N} = \mathbf{F}'\mathbf{Q}_{\mathbf{F}\mathbf{X}_2}$.
(f) *The statistic* $\mathbf{Fy}$ *is linearly sufficient for* $\mathbf{X}_1\beta_1$ *under* $\mathscr{M}_{12}$.

If, in the situation of Theorem 3, we request $\mathbf{Fy}$ to be linearly sufficient for $\mathbf{X}_1\beta_1$ for *any* $\mathbf{X}_1$ (expecting though $\mathbf{X}_1\beta_1$ to be estimable), we get the following corollary.

Corollary 1. *Let* $\mu_1 = \mathbf{X}_1\beta_1$ *be estimable under* $\mathscr{M}_{12}$ *and* $\mathscr{M}_t$ *and let* $\mathbf{W} \in \mathscr{W}$. *Then the following statements are equivalent:*

(a) *The statistic* $\mathbf{Fy}$ *is linearly sufficient for* $\mathbf{X}_1\beta_1$ *under* $\mathscr{M}_{12}$ *for any* $\mathbf{X}_1$.
(b) $\mathscr{C}(\mathbf{W}) \subset \mathscr{C}(\mathbf{X}_2 : \mathbf{W}\mathbf{F}'\mathbf{Q}_{\mathbf{F}\mathbf{X}_2}) = \mathscr{C}(\mathbf{X}_2) \oplus \mathscr{C}(\mathbf{W}\mathbf{F}') \cap \mathscr{C}(\mathbf{W}\mathbf{M}_2)$.

Proof. The statistic $\mathbf{Fy}$ is linearly sufficient for $\mathbf{X}_1\beta_1$ under $\mathscr{M}_{12}$ for any $\mathbf{X}_1$ if and only if

$$\mathbf{W}^{+1/2}\mathbf{P}_{\mathbf{Z}}\mathbf{W}^{+1/2} = \mathbf{0}\,. \tag{1.96}$$

Now (1.96) holds if and only if

$$\mathscr{C}(\mathbf{W}^{+1/2}) \subset \mathscr{C}(\mathbf{Z})^{\perp} = \mathscr{C}(\mathbf{W}^{+1/2}\mathbf{X}_2 : \mathbf{Q}_{\mathbf{W}} : \mathbf{W}^{1/2}\mathbf{F}'\mathbf{Q}_{\mathbf{FX}_2}) . \tag{1.97}$$

Premultipying (1.97) by $\mathbf{W}^{1/2}$ yields an equivalent form

$$\mathscr{C}(\mathbf{W}) \subset \mathscr{C}(\mathbf{X}_2 : \mathbf{WF}'\mathbf{Q}_{\mathbf{FX}_2}) = \mathscr{C}(\mathbf{X}_2) \oplus \mathscr{C}(\mathbf{WF}') \cap \mathscr{C}(\mathbf{WM}_2) . \quad \square \tag{1.98}$$

1.5 Linear sufficiency under $\mathscr{M}_1$ vs. $\mathscr{M}_{12}$

Consider the small model $\mathscr{M}_1 = \{\mathbf{y}, \mathbf{X}_1\beta_1, \mathbf{V}\}$ and full model $\mathscr{M}_{12} = \{\mathbf{y}, \mathbf{X}_1\beta_1 + \mathbf{X}_2\beta_2, \mathbf{V}\}$. Here is a reasonable question: what about comparing conditions for

$$\mathbf{Fy} \in \mathrm{S}(\mu_1 \mid \mathscr{M}_1) \quad \text{versus} \quad \mathbf{Fy} \in \mathrm{S}(\mu_1 \mid \mathscr{M}_{12}) . \tag{1.99}$$

For example, under which condition

$$\mathbf{Fy} \in \mathrm{S}(\mu_1 \mid \mathscr{M}_1) \implies \mathbf{Fy} \in \mathrm{S}(\mu_1 \mid \mathscr{M}_{12}) . \tag{1.100}$$

There is one crucial matter requiring our attention. Namely in the small model $\mathscr{M}_1$ the response $\mathbf{y}$ is lying in $\mathscr{C}(\mathbf{W}_1)$ but in $\mathscr{M}_{12}$ the response $\mathbf{y}$ can be in a wider subspace $\mathscr{C}(\mathbf{W})$. How to take this into account? What about assuming that

$$\mathscr{C}(\mathbf{X}_2) \subset \mathscr{C}(\mathbf{X}_1 : \mathbf{V}) \ ? \tag{1.101}$$

This assumption means that adding the $\mathbf{X}_2$-part into the model does not carry $\mathbf{y}$ out of $\mathscr{C}(\mathbf{W}_1)$ which seems to be a logical requirement. In such a situation we should find conditions under which

$$\mathscr{C}(\mathbf{X}_1) \subset \mathscr{C}(\mathbf{W}_1\mathbf{F}') \tag{1.102}$$

implies

$$\mathscr{C}(\mathbf{W}_1\dot{\mathbf{M}}_{2W}\mathbf{X}_1) \subset \mathscr{C}(\mathbf{W}_1\mathbf{F}') . \tag{1.103}$$

We know that under certain conditions the BLUE of $\mathbf{X}_1\beta_1$ does not change when the predictors in $\mathbf{X}_2$ are added into the model. It seems obvious that in such a situation (1.102) and (1.103) are equivalent. Supposing that (1.101) holds, then, e.g., according to Haslett & Puntanen (2010, Th. 3.1),

$$\tilde{\mu}_1(\mathscr{M}_{12}) = \tilde{\mu}_1(\mathscr{M}_1) - \mathbf{X}_1(\mathbf{X}_1'\mathbf{W}_1^{+}\mathbf{X}_1)^{-}\mathbf{X}_1'\mathbf{W}_1^{+}\tilde{\mu}_2(\mathscr{M}_{12}) , \tag{1.104}$$

and hence

$$\tilde{\mu}_1(\mathscr{M}_{12}) = \tilde{\mu}_1(\mathscr{M}_1) \tag{1.105}$$

if and only if $\mathbf{X}_1'\mathbf{W}_1^{+}\tilde{\mu}_2(\mathscr{M}_{12}) = \mathbf{0}$, i.e.,

$$\mathbf{X}_1'\mathbf{W}_1^{+}\mathbf{X}_2(\mathbf{X}_2'\dot{\mathbf{M}}_1\mathbf{X}_2)^{-}\mathbf{X}_2'\dot{\mathbf{M}}_1\mathbf{y} = \mathbf{0} . \tag{1.106}$$

Requesting (1.106) to hold for all $\mathbf{y} \in \mathscr{C}(\mathbf{X}_1 : \mathbf{V})$ and using the assumption $\mathscr{C}(\mathbf{X}_2) \subset \mathscr{C}(\mathbf{X}_1 : \mathbf{V})$, we obtain

$$\mathbf{X}_1'\mathbf{W}_1^+\mathbf{X}_2(\mathbf{X}_2'\dot{\mathbf{M}}_1\mathbf{X}_2)^-\mathbf{X}_2'\dot{\mathbf{M}}_1\mathbf{X}_2 = \mathbf{0}\,, \tag{1.107}$$

i.e.,

$$\mathbf{X}_1'\mathbf{W}_1^+\mathbf{X}_2\mathbf{P}_{\mathbf{X}_2'} = \mathbf{X}_1'\mathbf{W}_1^+\mathbf{X}_2 = \mathbf{0}\,, \tag{1.108}$$

where we have used the fact $\mathscr{C}(\mathbf{X}_2'\dot{\mathbf{M}}_1\mathbf{X}_2) = \mathscr{C}(\mathbf{X}_2')$. Thus we have shown the equivalence of (1.105) and (1.108).

On the other hand, (1.102) implies (1.103) if and only if $\mathscr{C}(\mathbf{W}_1\dot{\mathbf{M}}_{2W}\mathbf{X}_1) \subset \mathscr{C}(\mathbf{X}_1)$, which is equivalent to

$$\mathscr{C}(\mathbf{W}_1\dot{\mathbf{M}}_{2W}\mathbf{X}_1) = \mathscr{C}(\mathbf{X}_1)\,, \tag{1.109}$$

because we know that $\mathrm{r}(\mathbf{W}_1\dot{\mathbf{M}}_{2W}\mathbf{X}_1) = \mathrm{r}(\mathbf{X}_1)$. Hence neither column spaces in (1.109) can be a proper subspace of the other. Therefore, as stated by Baksalary (1984, p. 23) in the case of $\mathbf{V} = \mathbf{I}_n$, either the classes of statistics which are linearly sufficient for μ_1 are in the models $\mathscr{M}_1$ and $\mathscr{M}_{12}$ exactly the same, or, if not, there exists at least one statistic $\mathbf{Fy}$ such that $\mathbf{Fy} \in \mathrm{S}(\mu_1 \mid \mathscr{M}_1)$ but $\mathbf{Fy} \notin \mathrm{S}(\mu_1 \mid \mathscr{M}_{12})$ and at least one statistic $\mathbf{Fy}$ such that $\mathbf{Fy} \in \mathrm{S}(\mu_1 \mid \mathscr{M}_{12})$ but $\mathbf{Fy} \notin \mathrm{S}(\mu_1 \mid \mathscr{M}_1)$.

Now (1.102) and (1.103) are equivalent if and only if (1.109) holds, i.e.,

$$\mathbf{M}_1\mathbf{W}_1\dot{\mathbf{M}}_{2W}\mathbf{X}_1 = \mathbf{0}\,. \tag{1.110}$$

Using (1.57), (1.110) becomes

$$\mathbf{M}_1\mathbf{X}_2(\mathbf{X}_2'\mathbf{W}_1^+\mathbf{X}_2)^-\mathbf{X}_2'\mathbf{W}_1^+\mathbf{X}_1 = \mathbf{0}\,. \tag{1.111}$$

Because $\mathrm{r}(\mathbf{M}_1\mathbf{X}_2) = \mathrm{r}(\mathbf{X}_2)$, we can cancel, on account of Marsaglia & Styan (1974, Th. 2), the matrix $\mathbf{M}_1$ in (1.111) and thus obtain

$$\mathbf{X}_2(\mathbf{X}_2'\mathbf{W}_1^+\mathbf{X}_2)^-\mathbf{X}_2'\mathbf{W}_1^+\mathbf{X}_1 = \mathbf{0}\,. \tag{1.112}$$

Premultiplying (1.111) by $\mathbf{X}_2'\mathbf{W}_1^+$ shows that (1.109) is equivalent to

$$\mathbf{X}_2'\mathbf{W}_1^+\mathbf{X}_1 = \mathbf{0}\,. \tag{1.113}$$

In (1.113) of course $\mathbf{W}_1^+$ can be replaced with any $\mathbf{W}_1^-$. Thus we have proved the following:

Theorem 4. *Consider the models $\mathscr{M}_{12}$ and $\mathscr{M}_1$ and suppose that $\mu_1 = \mathbf{X}_1\beta_1$ is estimable under $\mathscr{M}_{12}$ and $\mathscr{C}(\mathbf{X}_2) \subset \mathscr{C}(\mathbf{X}_1 : \mathbf{V})$. Then the following statements are equivalent:*

(a) $\mathbf{X}_1'\mathbf{W}_1^+\mathbf{X}_2 = \mathbf{0}$,
(b) $\mathrm{BLUE}(\mu_1 \mid \mathscr{M}_1) = \mathrm{BLUE}(\mu_1 \mid \mathscr{M}_{12})$ *with probability* 1,
(c) $\mathbf{Fy} \in \mathrm{S}(\mu_1 \mid \mathscr{M}_1) \iff \mathbf{Fy} \in \mathrm{S}(\mu_1 \mid \mathscr{M}_{12})$.

Overlooking the problem for $\mathbf{y}$ belonging to $\mathscr{C}(\mathbf{X} : \mathbf{V})$ or to $\mathscr{C}(\mathbf{X}_1 : \mathbf{V})$, we can start our considerations by assuming that (1.100) holds, i.e.,

$$\mathscr{C}(\mathbf{X}_1) \subset \mathscr{C}(\mathbf{W}_1\mathbf{F}') \implies \mathscr{C}(\mathbf{W}\dot{\mathbf{M}}_{2W}\mathbf{X}_1) \subset \mathscr{C}(\mathbf{W}\mathbf{F}') . \tag{1.114}$$

Choosing $\mathbf{F}' = \mathbf{W}_1^{-}\mathbf{X}_1$ we observe that $\mathbf{Fy} \in \mathrm{S}(\mu_1 \mid \mathscr{M}_1)$ for any choice of $\mathbf{W}_1^{-}$. Thus (1.114) implies that we must also have

$$\mathscr{C}(\mathbf{W}\dot{\mathbf{M}}_{2W}\mathbf{X}_1) \subset \mathscr{C}(\mathbf{W}\mathbf{W}_1^{-}\mathbf{X}_1) . \tag{1.115}$$

According to Lemma 3 of Baksalary & Mathew (1986), (for nonnull $\mathbf{W}\dot{\mathbf{M}}_{2W}\mathbf{X}_1$ and $\mathbf{X}_1$) the inclusion (1.115) holds for any $\mathbf{W}_1^{-}$ if and only if

$$\mathscr{C}(\mathbf{W}) \subset \mathscr{C}(\mathbf{W}_1) \tag{1.116}$$

holds along with

$$\mathscr{C}(\mathbf{W}\dot{\mathbf{M}}_{2W}\mathbf{X}_1) \subset \mathscr{C}(\mathbf{W}\mathbf{W}_1^{+}\mathbf{X}_1) . \tag{1.117}$$

Inclusion (1.116) means that $\mathscr{C}(\mathbf{X}_2) \subset \mathscr{C}(\mathbf{X}_1 : \mathbf{V})$, i.e., $\mathscr{C}(\mathbf{W}) = \mathscr{C}(\mathbf{W}_1)$, which is our assumption in Theorem 4. Thus we can also conclude the following.

Corollary 2. *Consider the models $\mathscr{M}_{12}$ and $\mathscr{M}_1$ and suppose that $\mu_1 = \mathbf{X}_1\beta_1$ is estimable under $\mathscr{M}_{12}$. Then the following statements are equivalent:*

(a) $\mathbf{X}_1'\mathbf{W}_1^{+}\mathbf{X}_2 = \mathbf{0}$ *and* $\mathscr{C}(\mathbf{X}_2) \subset \mathscr{C}(\mathbf{X}_1 : \mathbf{V})$.
(b) $\mathbf{Fy} \in \mathrm{S}(\mu_1 \mid \mathscr{M}_1) \iff \mathbf{Fy} \in \mathrm{S}(\mu_1 \mid \mathscr{M}_{12})$.

We complete this section by considering the linear sufficiency of $\mathbf{Fy}$ versus that of $\mathbf{FM}_2\mathbf{y}$.

Theorem 5. *Consider the models $\mathscr{M}_{12}$ and $\mathscr{M}_{12\cdot 2}$ and suppose that $\mu_1 = \mathbf{X}_1\beta_1$ is estimable under $\mathscr{M}_{12}$. Then*

(a) $\mathbf{Fy} \in \mathrm{S}(\mu_1 \mid \mathscr{M}_{12}) \implies \mathbf{FM}_2\mathbf{y} \in \mathrm{S}(\mu_1 \mid \mathscr{M}_{12})$.
(b) *The reverse relation in* (a) *holds* $\iff \mathscr{C}(\mathbf{FM}_2\mathbf{W}) \cap \mathscr{C}(\mathbf{FX}_2) = \{\mathbf{0}\}$.

Moreover, the following statements are equivalent:

(c) $\mathbf{FM}_2\mathbf{y} \in \mathrm{S}(\mathbf{X}_1\beta_1 \mid \mathscr{M}_{12})$,
(d) $\mathbf{FM}_2\mathbf{y} \in \mathrm{S}(\mathbf{X}_1\beta_1 \mid \mathscr{M}_{12\cdot 2})$,
(e) $\mathbf{FM}_2\mathbf{y} \in \mathrm{S}(\mathbf{M}_2\mathbf{X}_1\beta_1 \mid \mathscr{M}_{12\cdot 2})$.

Proof. To prove (a), we observe that $\mathbf{Fy} \in \mathrm{S}(\mu_1 \mid \mathscr{M}_{12})$ implies (1.91), i.e.,

$$\mathscr{C}(\mathbf{W}\dot{\mathbf{M}}_{2W}\mathbf{X}_1) \subset \mathscr{C}(\mathbf{W}\mathbf{F}'\mathbf{Q}_{\mathbf{FX}_2}) . \tag{1.118}$$

Now on account Lemma 5d, we have $\mathbf{M}_2\mathbf{F}'\mathbf{Q}_{\mathbf{FX}_2} = \mathbf{F}'\mathbf{Q}_{\mathbf{FX}_2}$. Substituting this into (1.118) gives $\mathscr{C}(\mathbf{W}\dot{\mathbf{M}}_{2W}\mathbf{X}_1) \subset \mathscr{C}(\mathbf{W}\mathbf{M}_2\mathbf{F}'\mathbf{Q}_{\mathbf{FX}_2})$, and so

$$\mathscr{C}(\mathbf{W}\dot{\mathbf{M}}_{2W}\mathbf{X}_1) \subset \mathscr{C}(\mathbf{W}\mathbf{M}_2\mathbf{F}') , \tag{1.119}$$

which is the condition for $\mathbf{FM}_2\mathbf{y} \in \mathrm{S}(\mu_1 \mid \mathscr{M}_{12})$, thus confirming our claim (a). For an alternative proof of (a), see Isotalo & Puntanen (2006, Cor. 1).

The reverse relation in (a) holds if and only if (1.119) implies (1.118), i.e.,

$$\mathscr{C}(\mathbf{WM}_2\mathbf{F}') \subset \mathscr{C}(\mathbf{WM}_2\mathbf{F}'\mathbf{Q}_{\mathbf{FX}_2}), \tag{1.120}$$

where of course only the equality is possible. Now

$$\mathrm{r}(\mathbf{WM}_2\mathbf{F}'\mathbf{Q}_{\mathbf{FX}_2}) = \mathrm{r}(\mathbf{WM}_2\mathbf{F}') - \dim \mathscr{C}(\mathbf{FM}_2\mathbf{W}) \cap \mathscr{C}(\mathbf{FX}_2), \tag{1.121}$$

and hence (1.120) holds if and only if $\mathscr{C}(\mathbf{FM}_2\mathbf{W}) \cap \mathscr{C}(\mathbf{FX}_2) = \{\mathbf{0}\}$ which proves our claim (b).

The condition (c), $\mathbf{FM}_2\mathbf{y} \in \mathrm{S}(\mu_1 \mid \mathscr{M}_{12})$, holds if and only if

$$\mathscr{C}[\mathbf{WM}_2(\mathbf{M}_2\mathbf{WM}_2)^-\mathbf{M}_2\mathbf{X}_1] \subset \mathscr{C}(\mathbf{WM}_2\mathbf{F}'). \tag{1.122}$$

Premultiplying (1.122) by $\mathbf{M}_2$ yields

$$\mathscr{C}(\mathbf{M}_2\mathbf{X}_1) \subset \mathscr{C}(\mathbf{M}_2\mathbf{WM}_2\mathbf{F}'), \tag{1.123}$$

which means that (e) holds. Premultiplying (1.123) by $\mathbf{WM}_2(\mathbf{M}_2\mathbf{WM}_2)^-$ yields (1.122) thus confirming the equivalence of (c) and (e).

Applying Lemma 2b we observe that (d) holds if and only if

$$\mathscr{C}\{\mathbf{M}_2\mathbf{X}_1[\mathbf{X}_1'\mathbf{M}_2(\mathbf{M}_2\mathbf{WM}_2)^-\mathbf{M}_2\mathbf{X}_1]^-\mathbf{X}_1'\mathbf{M}_2\} \subset \mathscr{C}(\mathbf{M}_2\mathbf{WM}_2\mathbf{F}'). \tag{1.124}$$

The fact that the left-hand side of (1.124) is $\mathscr{C}(\mathbf{M}_2\mathbf{X}_1)$ shows the equivalence of (d) and (c), and thus the proof is completed. □

1.6 Conclusions

If our interest is in all estimable parametric functions of β_1, as in Groß & Puntanen (2000), Markiewicz & Puntanen (2009), and Kala & Pordzik (2009), then we could concentrate on estimating $\mathbf{M}_2\mathbf{X}_1\beta_1$. This is due to the fact that $\mathbf{K}_1\beta_1$ is estimable if and only if $\mathscr{C}(\mathbf{K}_1') \subset \mathscr{C}(\mathbf{X}_1'\mathbf{M}_2)$; see, e.g., Groß & Puntanen (2000, Lemma 1). Hence, the BLUE for an arbitrary estimable vector $\mathbf{K}_1\beta_1$ may easily be computed from the BLUE of $\mathbf{M}_2\mathbf{X}_1\beta_1$. We may also mention that in the corresponding way, Baksalary (1984, 1987, §3.3, §5) considered the linearly sufficienct statistics for $\mathbf{X}_1'\mathbf{M}_2\mathbf{X}_1\beta_1$ under $\mathscr{M}_1$ and $\mathscr{M}_{12}$ when $\mathbf{V} = \mathbf{I}_n$.

Our interest in this paper has been focused on the estimation of $\mu = \mathbf{X}\beta$ and $\mu_1 = \mathbf{X}_1\beta_1$ under the linear model $\mathscr{M}_{12} = \{\mathbf{y}, \mathbf{X}_1\beta_1 + \mathbf{X}_2\beta_2, \mathbf{V}\}$. Here we need to assume that $\mathbf{X}_1\beta_1$ is estimable (and thereby also $\mathbf{X}_2\beta_2$ is estimable). We have characterized the linearly sufficient statistic $\mathbf{Fy}$ by using the covariance matrices of the BLUEs under $\mathscr{M}_{12}$ and under its transformed version $\mathscr{M}_t = \{\mathbf{Fy}, \mathbf{FX}_1\beta_1 + \mathbf{FX}_2\beta_2, \mathbf{FVF}'\}$. In particular, certain orthogonal projectors appear useful in our considerations. We have obtained new interesting proofs for some known results, like Lemma 2, and presented some new properties related to linear sufficiency. Particular attention has been paid to the condition under which adding new regressors (in $\mathbf{X}_2$) does not affect the linear sufficiency of $\mathbf{Fy}$. Similarly we have characterized linear sufficiency of $\mathbf{Fy}$ versus that of $\mathbf{FM}_2\mathbf{y}$ under the models $\mathscr{M}_{12}$ and $\mathscr{M}_{12\cdot 2}$.

As one of the referees of this paper stated, our considerations are based on effective matrix and column space properties and hence its relevance for applied data analysts may be a bit limited. However, we believe that in the long run the given column space properties may provide some new insights into the linear estimation theory.

Acknowledgements

Thanks go to the referees for construcive comments. Part of this research was done during the meeting of an International Research Group on Multivariate Models in the Mathematical Research and Conference Center, Będlewo, Poland, March 2016, supported by the Stefan Banach International Mathematical Center.

References

Baksalary, J. K. (1984): A study of the equivalence between a Gauss–Markoff model and its augmentation by nuisance parameters. Math. Operationsforsch. Stat. Ser. Stat. **15**, 2–35.

Baksalary, J. K. (1987): Algebraic characterizations and statistical implications of the commutativity of orthogonal projectors. Proceedings of the Second International Tampere Conference in Statistics. T. Pukkila & S. Puntanen, Eds., Univ. of Tampere, vol. A 184, pp. 113–142.

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

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

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

Dong, B., Guo, W., Tian, Y. (2014): On relations between BLUEs under two transformed linear models. J. Multivariate Anal. **13**, 279–292.

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

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

Haslett, S. J., Puntanen, S. (2010): Effect of adding regressors on the equality of the BLUEs under two linear models. J. Stat. Plan. Inf. **140**, 104–110.

Isotalo, J., Puntanen, S. (2006): Linear sufficiency and completeness in the partitioned linear model. Acta Comment Univ. Tartu Math. **10**, 53–67.

Isotalo, J., Puntanen, S. (2009): Linear sufficiency and completeness in the context of estimating the parametric function in the general Gauss–Markov model. J. Stat. Plan. Inf. **139**, 722–733.

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

Kala, R., Pordzik, P. R. (2009): Estimation in singular partitioned, reduced or transformed linear models. Stat. Pap. **50**, 633–638.

Kala, R., Puntanen, S., Tian, Y. (2017): Some notes on linear sufficiency. Stat. Pap., available online. http://dx.doi.org/10.1007/s00362-015-0682-2

Kala, R., Markiewicz, A., Puntanen, S. (2017): Some further remarks on the linear sufficiency in the linear model. Applied and Computational Matrix Analysis. Natália Bebiano, editor, Springer, to appear. http://dx.doi.org/10.1007/978-3-319-49984-0

Markiewicz, A., Puntanen, S. (2009): Admissibility and linear sufficiency in linear model with nuisance parameters. Stat. Pap. **50**, 847–854.

Markiewicz, A., Puntanen, S. (2015): All about the $\perp$ with its applications in the linear statistical models. Open Mathematics **13**, 33–50.

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

Müller, J. (1987): Sufficiency and completeness in the linear model. J. Multivar. Anal. **21**, 312–323.

Milliken, G. A., Akdeniz, F. (1977): A theorem on the difference of the generalized inverses of two nonnegative matrices. Comm. Statist. Theory Methods **6**, 1747–1751.

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

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

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

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