ACTA ET COMMENTATIONES UNIVERSITATIS TARTUENSIS DE MATHEMATICA
Volume 10, 2006

Linear sufficiency and completeness in the partitioned linear model

Jarkko Isotalo and Simo Puntanen

Abstract. In this paper we consider the estimation of $\mathbf{X}_1\boldsymbol{\beta}_1$ under the partitioned linear model $\{\mathbf{y},\, \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2,\, \sigma^2\mathbf{V}\}$. In particular, we consider linear sufficiency and linear completeness of $\mathbf{X}_1\boldsymbol{\beta}_1$. We give new characterizations for linear sufficiency of $\mathbf{X}_1\boldsymbol{\beta}_1$, and define and characterize linear completeness in a case of estimation of $\mathbf{X}_1\boldsymbol{\beta}_1$. We also introduce a predictive approach for obtaining the best linear unbiased estimator of $\mathbf{X}_1\boldsymbol{\beta}_1$, and subsequently, we give the linear analogues of the Rao–Blackwell and Lehmann–Scheffé Theorems in the context of estimating $\mathbf{X}_1\boldsymbol{\beta}_1$.

1. Introduction

In this paper we consider the partitioned linear model

$$\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon},$$

or shortly,

$$\mathcal{M}_{12} = \{\mathbf{y},\, \mathbf{X}\boldsymbol{\beta},\, \sigma^2\mathbf{V}\} = \{\mathbf{y},\, \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2,\, \sigma^2\mathbf{V}\},$$

where $\mathrm{E}(\mathbf{y}) = \mathbf{X}\boldsymbol{\beta}$, $\mathrm{E}(\boldsymbol{\varepsilon}) = \mathbf{0}$, $\mathrm{cov}(\mathbf{y}) = \mathrm{cov}(\boldsymbol{\varepsilon}) = \sigma^2\mathbf{V}$. We denote the expectation vector and covariance matrix, respectively, by $\mathrm{E}(\cdot)$ and $\mathrm{cov}(\cdot)$.

In the model $\mathcal{M}_{12}$ the vector $\mathbf{y}$ is an $n \times 1$ observable random vector, $\boldsymbol{\varepsilon}$ is an $n \times 1$ random error vector, $\mathbf{X}$ is a known $n \times p$ matrix, partitioned columnwise as $\mathbf{X} = (\mathbf{X}_1 : \mathbf{X}_2)$ with $\mathbf{X}_1$ $(n \times p_1)$ and $\mathbf{X}_2$ $(n \times p_2)$, $\boldsymbol{\beta} = (\boldsymbol{\beta}_1', \boldsymbol{\beta}_2')'$ is a $p \times 1$ vector of unknown parameters, correspondingly with $\boldsymbol{\beta}_1$ $(p_1 \times 1)$ and $\boldsymbol{\beta}_2$ $(p_2 \times 1)$, $\sigma^2 > 0$ is an unknown scalar, and $\mathbf{V}$ is a known $n \times n$ nonnegative definite matrix.

Furthermore, let $\mathbb{R}_{m,n}$ denote the set of $m \times n$ real matrices and $\mathbb{R}_m = \mathbb{R}_{m,1}$. The symbols $\mathbf{A}'$, $\mathbf{A}^-$, $\mathbf{A}^+$, $\mathscr{C}(\mathbf{A})$, $\mathscr{C}(\mathbf{A})^\perp$, $\mathscr{N}(\mathbf{A})$, and $\mathrm{r}(\mathbf{A})$ will

Received October 15, 2006.
2000 *Mathematics Subject Classification.* 62J05, 62B05.
Key words and phrases. Best linear unbiased estimation, Gauss–Markov model, linear sufficiency, linear completeness, linear estimation.

stand for the transpose, a generalized inverse, the Moore–Penrose inverse, the column space, the orthogonal complement of the column space, the null space, and the rank, respectively, of $\mathbf{A} \in \mathbb{R}_{m,n}$. By $\mathbf{A}^{\perp}$ we denote any matrix satisfying $\mathscr{C}(\mathbf{A}^{\perp}) = \mathscr{N}(\mathbf{A}') = \mathscr{C}(\mathbf{A})^{\perp}$. Further we will write $\mathbf{P}_{\mathbf{A}} = \mathbf{A}\mathbf{A}^{+} = \mathbf{A}(\mathbf{A}'\mathbf{A})^{-}\mathbf{A}'$ to denote the orthogonal projector (with respect to the standard inner product) onto $\mathscr{C}(\mathbf{A})$, and $\mathbf{M}_{\mathbf{A}} = \mathbf{I} - \mathbf{P}_{\mathbf{A}}$ to denote the orthogonal projector onto $\mathscr{C}(\mathbf{A})^{\perp}$, where $\mathbf{I}$ denotes the identity matrix. In particular,

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

Moreover, by $(\mathbf{A} : \mathbf{B})$ we denote the partitioned matrix with $\mathbf{A} \in \mathbb{R}_{m,n}$ and $\mathbf{B} \in \mathbb{R}_{m,k}$ as submatrices.

Let us consider the estimation of the linear parametric function $\mathbf{X}_1\boldsymbol{\beta}_1$ under the partitioned model $\mathcal{M}_{12}$. The parametric function $\mathbf{X}_1\boldsymbol{\beta}_1$ is said to be estimable under the model $\mathcal{M}_{12}$ if it has a linear unbiased estimator, i.e., there exists a matrix $\mathbf{G}$ such that

$$\mathrm{E}(\mathbf{G}\mathbf{y}) = \mathbf{G}\mathbf{X}\boldsymbol{\beta} = \mathbf{G}\mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{G}\mathbf{X}_2\boldsymbol{\beta}_2 = \mathbf{X}_1\boldsymbol{\beta}_1 \quad \text{for all } \boldsymbol{\beta} \in \mathbb{R}_p,$$

or equivalently, if $\mathbf{G}(\mathbf{X}_1 : \mathbf{X}_2) = (\mathbf{X}_1 : \mathbf{0})$. Therefore, we can conclude the following:

$$\mathbf{X}_1\boldsymbol{\beta}_1 \text{ is estimable} \iff \mathscr{C}\begin{pmatrix} \mathbf{X}_1' \\ \mathbf{0} \end{pmatrix} \subset \mathscr{C}\begin{pmatrix} \mathbf{X}_1' \\ \mathbf{X}_2' \end{pmatrix},$$

which is equivalent to $\mathscr{C}(\mathbf{X}_1) \cap \mathscr{C}(\mathbf{X}_2) = \{\mathbf{0}\}$.

Moreover, a linear statistic $\mathbf{G}\mathbf{y}$ is the Best Linear Unbiased Estimator, BLUE, for the estimable parametric function $\mathbf{X}_1\boldsymbol{\beta}_1$ if, for any other unbiased linear estimator $\mathbf{F}\mathbf{y}$, the difference $\mathrm{cov}(\mathbf{F}\mathbf{y}) - \mathrm{cov}(\mathbf{G}\mathbf{y})$ is a nonnegative definite matrix, i.e.,

$$\mathrm{cov}(\mathbf{F}\mathbf{y}) - \mathrm{cov}(\mathbf{G}\mathbf{y}) \geq \mathbf{0} \qquad \text{for all } \mathbf{F}\mathbf{y} \text{ such that } \mathrm{E}(\mathbf{F}\mathbf{y}) = \mathbf{X}_1\boldsymbol{\beta}_1.$$

It is well known that $\mathbf{G}\mathbf{y}$ is the BLUE of $\mathbf{X}_1\boldsymbol{\beta}_1$ if and only if $\mathbf{G}$ satisfies the fundamental equation of the BLUE:

$$\mathbf{G}(\mathbf{X}_1 : \mathbf{X}_2 : \mathbf{V}\mathbf{X}^{\perp}) = (\mathbf{X}_1 : \mathbf{0} : \mathbf{0}). \tag{1.1}$$

Proof of (1.1) is given by, e.g., Drygas (1970, p. 50), Rao (1973, p. 282), and recently Baksalary (2004). From the equation (1.1), explicit representations for the BLUE of $\mathbf{X}_1\boldsymbol{\beta}_1$ are obtainable. For example, the BLUE of $\mathbf{X}_1\boldsymbol{\beta}_1$ can be written as

$$\mathrm{BLUE}(\mathbf{X}_1\boldsymbol{\beta}_1 \mid \mathcal{M}_{12}) = (\mathbf{X}_1 : \mathbf{0})(\mathbf{X}'\mathbf{W}^{-}\mathbf{X})^{-}\mathbf{X}'\mathbf{W}^{-}\mathbf{y},$$

where $\mathbf{W} = \mathbf{V} + \mathbf{X}\mathbf{X}'$. Note also that we assume the model $\mathcal{M}_{12}$ being consistent in a sense that

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

almost surely [see Groß (2004, Section 2)].

In this paper, we consider properties of those linear statistics $\mathbf{Ty}$, which are linearly sufficient for $\mathbf{X}_1\boldsymbol{\beta}_1$ under the partitioned model $\mathcal{M}_{12}$. A linear statistic $\mathbf{Ty}$ is said to be linearly sufficient for $\mathbf{X}_1\boldsymbol{\beta}_1$ if there exists a matrix $\mathbf{A}$ such that $\mathbf{ATy}$ is the BLUE of $\mathbf{X}_1\boldsymbol{\beta}_1$ under the model $\mathcal{M}_{12}$, see, e.g., Baksalary and Kala (1986, Definition 1). The concept of linear sufficiency was introduced by Baksalary and Kala (1981) and Drygas (1983) while considering those linear statistics, which are "sufficient" for the expected value $\mathbf{X}\boldsymbol{\beta}$ under the model $\mathcal{M}_{12}$. Baksalary and Kala (1981) and Drygas (1983) showed that a linear statistic $\mathbf{Ty}$ is linearly sufficient for $\mathbf{X}\boldsymbol{\beta}$ under the model $\mathcal{M}_{12}$ if and only if the column space inclusion

$$\mathscr{C}(\mathbf{X}) \subset \mathscr{C}(\mathbf{WT}') \tag{1.2}$$

holds. Müller (1987) later proved that the condition (1.2) is equivalent to the inclusion

$$\mathscr{N}(\mathbf{T}) \cap \mathscr{C}(\mathbf{W}) \subset \mathscr{C}(\mathbf{VX}^{\perp}).$$

Linear sufficiency of the given estimable parametric function $\mathbf{K}'\boldsymbol{\beta}$ was considered by Baksalary and Kala (1986). They proved that $\mathbf{Ty}$ is linearly sufficient for $\mathbf{K}'\boldsymbol{\beta}$ if and only if the null space inclusion

$$\mathscr{N}(\mathbf{TX} : \mathbf{TVX}^{\perp}) \subset \mathscr{N}(\mathbf{K}' : \mathbf{0}) \tag{1.3}$$

holds. In particular, when considering linear sufficiency of $\mathbf{X}_1\boldsymbol{\beta}_1$, Baksalary and Kala's condition (1.3) becomes

$$\mathscr{N}(\mathbf{TX}_1 : \mathbf{TX}_2 : \mathbf{TVX}^{\perp}) \subset \mathscr{N}(\mathbf{X}_1 : \mathbf{0} : \mathbf{0}). \tag{1.4}$$

In this paper, we give further characterizations of linear sufficiency of the parametric function $\mathbf{X}_1\boldsymbol{\beta}_1$. We also consider linear completeness in a case of estimation of $\mathbf{X}_1\boldsymbol{\beta}_1$, and prove that a statistic $\mathbf{Ty}$ is simultaneously linearly sufficient and linearly complete for $\mathbf{X}_1\boldsymbol{\beta}_1$ if and only if it is a linearly minimal sufficient statistic for $\mathbf{X}_1\boldsymbol{\beta}_1$.

In Section 4, we consider a predictive approach for obtaining the best linear unbiased estimator of $\mathbf{X}_1\boldsymbol{\beta}_1$. Sengupta and Jammalamadaka (2003, Chapter 11) gave an interesting study on the linear version of the general estimation theory, including the linear analogues to the Rao–Blackwell and Lehmann–Scheffé Theorems when considering the estimation of the expected value $\mathbf{X}\boldsymbol{\beta}$ under the model $\mathcal{M}_{12}$. In this paper, we give the corresponding linear analogues of the Rao–Blackwell and Lehmann–Scheffé Theorems in the context of estimating $\mathbf{X}_1\boldsymbol{\beta}_1$.

This paper is closely connected to the paper Isotalo and Puntanen (2006) concerning linear sufficiency and completeness of the given estimable parametric function $\mathbf{K}'\boldsymbol{\beta}$. However, the results given in this paper are not just applied results from more general results given in Isotalo and Puntanen (2006). In fact, the results given in Isotalo and Puntanen (2006) have been obtained by first reparametrizing of the original linear model $\mathcal{M}_{12}$ into particularly partitioned linear model and then by using the results given in this paper. Hence we feel that considering linear sufficiency and completeness in the context of estimating $\mathbf{X}_1\boldsymbol{\beta}_1$ has its own merits apart from the results concerning linear sufficiency and completeness of $\mathbf{K}'\boldsymbol{\beta}$.

2. Linear sufficiency

Let us consider the estimation of the parametric function $\mathbf{X}_1\boldsymbol{\beta}_1$ under the model $\mathcal{M}_{12}$. Then the part $\mathbf{X}_2\boldsymbol{\beta}_2$ can be seen as a nuisance factor in the partitioned model

$$\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}. \tag{2.1}$$

Now by premultiplying the full model equation (2.1) by the orthogonal projector $\mathbf{M}_2$, the vector of nuisance parameters $\boldsymbol{\beta}_2$ can explicitly be removed from the partitioned model $\mathcal{M}_{12}$. This premultiplication yields the so-called reduced model:

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

Moreover, a generalized version of the so-called Frisch–Waugh–Lovell Theorem now states that the BLUE of $\mathbf{X}_1\boldsymbol{\beta}_1$ under the reduced model $\mathcal{M}_{12\cdot 2}$ equals almost surely the BLUE of $\mathbf{X}_1\boldsymbol{\beta}_1$ under the partitioned model $\mathcal{M}_{12}$, see, e.g., Groß and Puntanen (2000, Theorem 4) and Bhimasankaram and Sengupta (1996, Theorem 6.1). Thus the linear statistic $\mathbf{M}_2\mathbf{y}$ is linearly sufficient for $\mathbf{X}_1\boldsymbol{\beta}_1$ under the model $\mathcal{M}_{12}$.

In addition to the orthogonal projector $\mathbf{M}_2$, the matrix

$$\dot{\mathbf{M}}_2 = \mathbf{M}_2(\mathbf{M}_2\mathbf{W}_1\mathbf{M}_2)^{-}\mathbf{M}_2, \tag{2.2}$$

where $\mathbf{W}_1 = \mathbf{V} + \mathbf{X}_1\mathbf{X}_1$, plays also an important role in subsequent considerations. Note that for matrices $\mathbf{A}$ and $\mathbf{B}$ such that $\mathscr{C}(\mathbf{A}') \subset \mathscr{C}(\mathbf{W}_1)$ and $\mathscr{C}(\mathbf{B}) \subset \mathscr{C}(\mathbf{W}_1)$, the product $\mathbf{A}\dot{\mathbf{M}}_2\mathbf{B}$ is invariant with respect to the choice of $(\mathbf{M}_2\mathbf{W}_1\mathbf{M}_2)^{-}$, and hence in such cases the generalized inverse can be chosen to be symmetric without loss of generality.

The following lemma now gives a useful column space equality related to the matrix $\dot{\mathbf{M}}_2$. A corresponding equality in a case of estimation of $\mathbf{K}'\boldsymbol{\beta}$ is given by Isotalo and Puntanen (2006, Lemma 2).

Lemma 1. *Let $\mathbf{X}_1\boldsymbol{\beta}_1$ be an estimable parametric function under the model $\mathcal{M}_{12}$, and let $\dot{\mathbf{M}}_2$ be the matrix given in* (2.2)*. Then the column space equality*

$$\mathscr{C}(\mathbf{X}_2 : \mathbf{V}\mathbf{X}^{\perp})^{\perp} = \mathscr{C}(\dot{\mathbf{M}}_2\mathbf{X}_1 : \mathbf{M}_2 - \dot{\mathbf{M}}_2\mathbf{W}_1\mathbf{M}_2)$$

holds.

Proof. First, we use the fact that

$$\mathscr{C}(\mathbf{X}_2 : \mathbf{V}\mathbf{X}^{\perp})^{\perp} = \mathscr{C}(\mathbf{X}_2 : \mathbf{W}\mathbf{M}_{\mathbf{X}})^{\perp} = \mathscr{C}(\mathbf{M}_2) \cap \mathscr{C}(\mathbf{W}\mathbf{M}_{\mathbf{X}})^{\perp} = \mathscr{C}[\mathbf{M}_2(\mathbf{M}_2\mathbf{W}_1\mathbf{M}_{\mathbf{X}})^{\perp}],$$

see Rao and Mitra (1971, Complement 7, p. 118).

Consider a vector $\mathbf{u}$ such that $\mathbf{u}'\mathbf{M}_2(\mathbf{M}_2\mathbf{W}_1\mathbf{M}_{\mathbf{X}})^{\perp} = \mathbf{0}$. Then there exists a vector $\mathbf{a}$ such that $\mathbf{u}'\mathbf{M}_2 = \mathbf{a}'\mathbf{M}_{\mathbf{X}}\mathbf{W}_1\mathbf{M}_2$, and hence

$$\begin{aligned}\mathbf{u}'\dot{\mathbf{M}}_2\mathbf{X}_1 &= \mathbf{a}'\mathbf{M}_{\mathbf{X}}\mathbf{W}_1\mathbf{M}_2(\mathbf{M}_2\mathbf{W}_1\mathbf{M}_2)^{-}\mathbf{M}_2\mathbf{X}_1\\ &= \mathbf{a}'\mathbf{M}_{\mathbf{X}}\mathbf{M}_2\mathbf{W}_1\mathbf{M}_2(\mathbf{M}_2\mathbf{W}_1\mathbf{M}_2)^{-}\mathbf{M}_2\mathbf{X}_1\\ &= \mathbf{a}'\mathbf{M}_{\mathbf{X}}\mathbf{M}_2\mathbf{X}_1 = \mathbf{0},\end{aligned}$$

where we have used the identity $\mathbf{M}_2\mathbf{W}_1\mathbf{M}_2(\mathbf{M}_2\mathbf{W}_1\mathbf{M}_2)^{-}\mathbf{M}_2\mathbf{X}_1 = \mathbf{M}_2\mathbf{X}_1$. Since also

$$\begin{aligned}\mathbf{u}'(\mathbf{M}_2 - \dot{\mathbf{M}}_2\mathbf{W}_1\mathbf{M}_2) &= \mathbf{a}'\mathbf{M}_{\mathbf{X}}\mathbf{W}_1\mathbf{M}_2(\mathbf{M}_2 - \dot{\mathbf{M}}_2\mathbf{W}_1\mathbf{M}_2)\\ &= \mathbf{a}'\mathbf{M}_{\mathbf{X}}\mathbf{W}_1\mathbf{M}_2 - \mathbf{a}'\mathbf{M}_{\mathbf{X}}\mathbf{W}_1\mathbf{M}_2 = \mathbf{0},\end{aligned}$$

the inclusion $\mathscr{C}(\mathbf{X}_2 : \mathbf{V}\mathbf{X}^{\perp})^{\perp} \supset \mathscr{C}(\dot{\mathbf{M}}_2\mathbf{X}_1 : \mathbf{M}_2 - \dot{\mathbf{M}}_2\mathbf{W}_1\mathbf{M}_2)$ holds.

To prove the reverse inclusion, let $\mathbf{u}$ be such that

$$\mathbf{u}'\dot{\mathbf{M}}_2\mathbf{X}_1 = \mathbf{0} \quad \text{and} \quad \mathbf{u}'(\mathbf{M}_2 - \dot{\mathbf{M}}_2\mathbf{W}_1\mathbf{M}_2) = \mathbf{0}. \tag{2.3}$$

Then the former condition in (2.3) implies that there exists a vector $\mathbf{b}$ such that

$$\mathbf{u}'\dot{\mathbf{M}}_2 = \mathbf{u}'\dot{\mathbf{M}}_2\mathbf{M}_2 = \mathbf{b}'\mathbf{M}_{\mathbf{M}_2\mathbf{X}_1} = \mathbf{b}'\mathbf{M}_{\mathbf{M}_2\mathbf{X}_1}\mathbf{M}_2 = \mathbf{b}'\mathbf{M}_{\mathbf{X}},$$

since $\mathbf{M}_{\mathbf{M}_2\mathbf{X}_1}\mathbf{M}_2 = \mathbf{M}_{\mathbf{X}}$. Therefore, based on the latter condition in (2.3), $\mathbf{u}'\mathbf{M}_2 = \mathbf{b}'\mathbf{M}_{\mathbf{X}}\mathbf{W}_1\mathbf{M}_2$, and thus

$$\mathbf{u}'\mathbf{M}_2(\mathbf{M}_2\mathbf{W}_1\mathbf{M}_{\mathbf{X}})^{\perp} = \mathbf{b}'\mathbf{M}_{\mathbf{X}}\mathbf{W}_1\mathbf{M}_2(\mathbf{M}_2\mathbf{W}_1\mathbf{M}_{\mathbf{X}})^{\perp} = \mathbf{0},$$

which shows that the reverse inclusion $\mathscr{C}(\mathbf{X}_2 : \mathbf{V}\mathbf{X}^{\perp})^{\perp} \subset \mathscr{C}(\dot{\mathbf{M}}_2\mathbf{X}_1 : \mathbf{M}_2 - \dot{\mathbf{M}}_2\mathbf{W}_1\mathbf{M}_2)$ also holds. □

Using Lemma 1 we can obtain the following representation of the BLUE of $\mathbf{X}_1\boldsymbol{\beta}_1$:

$$\mathrm{BLUE}(\mathbf{X}_1\boldsymbol{\beta}_1 \mid \mathcal{M}_{12}) = \mathbf{X}_1(\mathbf{X}_1'\dot{\mathbf{M}}_2\mathbf{X}_1)^{-}\mathbf{X}_1'\dot{\mathbf{M}}_2\mathbf{y}. \tag{2.4}$$

See Isotalo and Puntanen (2006, Corollary 4) for a more general proof of (2.4).

Next, in the following theorem, we give our further characterizations of linear sufficiency of the estimable parametric function $\mathbf{X}_1\boldsymbol{\beta}_1$.

Theorem 1. *Let $\mathbf{X}_1\boldsymbol{\beta}_1$ be an estimable parametric function under the model $\mathcal{M}_{12}$, and let $\dot{\mathbf{M}}_2$ be the matrix given in* (2.2)*. Then a linear statistic* $\mathbf{Ty}$ *is linearly sufficient for $\mathbf{X}_1\boldsymbol{\beta}_1$ if and only if any of the following equivalent statements holds:*

(a) $\mathscr{N}(\mathbf{TX}_1 : \mathbf{TX}_2 : \mathbf{TVX}^{\perp}) \subset \mathscr{N}(\mathbf{X}_1 : \mathbf{0} : \mathbf{0})$,
(b) $\mathscr{C}(\mathbf{W}\dot{\mathbf{M}}_2\mathbf{X}_1) \subset \mathscr{C}(\mathbf{WT}')$,
(c) $\mathscr{N}(\mathbf{T}) \cap \mathscr{C}(\mathbf{W}) \subset \mathscr{N}(\mathbf{X}_1'\dot{\mathbf{M}}_2) \cap \mathscr{C}(\mathbf{W})$,
(d) *there exists a matrix* $\mathbf{L}$ *such that* $\mathbf{LTy} = \mathbf{X}_1'\dot{\mathbf{M}}_2\mathbf{y}$ *almost surely.*

Proof. Part (a) is Baksalary and Kala's (1986, Theorem 1) condition (1.4).

Part (a), together with Lemma 1, implies that there exists a matrix $\mathbf{A}$ such that

$$\mathscr{C}(\mathbf{T}'\mathbf{A}') \subset \mathscr{C}(\mathbf{X}_2 : \mathbf{VX}^{\perp})^{\perp} = \mathscr{C}(\dot{\mathbf{M}}_2\mathbf{X}_1 : \mathbf{M}_2 - \dot{\mathbf{M}}_2\mathbf{W}_1\mathbf{M}_2),$$

and hence

$$\mathscr{C}(\mathbf{WT}'\mathbf{A}') \subset \mathscr{C}(\mathbf{W}\dot{\mathbf{M}}_2\mathbf{X}_1). \tag{2.5}$$

Since also

$$\begin{aligned} \mathrm{r}(\mathbf{W}\dot{\mathbf{M}}_2\mathbf{X}_1) &= \mathrm{r}(\mathbf{M}_2\mathbf{W}\dot{\mathbf{M}}_2\mathbf{X}_1) = \mathrm{r}(\mathbf{M}_2\mathbf{X}_1) \\ &= \mathrm{r}(\mathbf{X}_1) = \mathrm{r}(\mathbf{ATX}_1) \leq \mathrm{r}(\mathbf{ATW}), \end{aligned} \tag{2.6}$$

conditions (2.5) and (2.6) together imply part (b).

Part (b), on the other hand, implies that $\mathscr{C}(\mathbf{WT}')^{\perp} \subset \mathscr{C}(\mathbf{W}\dot{\mathbf{M}}_2\mathbf{X}_1)^{\perp}$, and furthermore, $\mathscr{C}[\mathbf{W}(\mathbf{WT}')^{\perp}] \subset \mathscr{C}[\mathbf{W}(\mathbf{W}\dot{\mathbf{M}}_2\mathbf{X}_1)^{\perp}]$, which, based on Rao and Mitra (1971, Complement 7, p. 118), is equivalent to part (c).

From part (c), it follows that $\mathscr{N}(\mathbf{TW}) \subset \mathscr{N}(\mathbf{X}_1'\dot{\mathbf{M}}_2\mathbf{W})$ holds. Thus there exists a matrix $\mathbf{B}$ such that

$$\mathbf{BTW} = \mathbf{X}_1'\dot{\mathbf{M}}_2\mathbf{W}, \tag{2.7}$$

and since $\mathbf{y} \in \mathscr{C}(\mathbf{W})$ almost surely, the equation (2.7) is equivalent to part (d).

Lastly, if part (d) holds, then, based on the equation (2.4), the equality

$$\begin{aligned} \mathbf{X}_1(\mathbf{X}_1'\dot{\mathbf{M}}_2\mathbf{X}_1)^{-}\mathbf{LTy} &= \mathbf{X}_1(\mathbf{X}_1'\dot{\mathbf{M}}_2\mathbf{X}_1)^{-}\mathbf{X}_1'\dot{\mathbf{M}}_2\mathbf{y} \\ &= \mathrm{BLUE}(\mathbf{X}_1\boldsymbol{\beta}_1 \mid \mathscr{M}_{12}) \end{aligned}$$

holds almost surely, showing that $\mathbf{Ty}$ is linearly sufficient for $\mathbf{X}_1\boldsymbol{\beta}_1$. □

The statistic $\mathbf{M}_2\mathbf{y}$ is not only linearly sufficient for $\mathbf{X}_1\boldsymbol{\beta}_1$, but also has an interesting invariance property in a sense that for every linear transformation $\mathbf{T}$ such that $\mathbf{Ty}$ is linearly sufficient for $\mathbf{X}_1\boldsymbol{\beta}_1$, the statistic $\mathbf{TM}_2\mathbf{y}$ is also linearly sufficient for $\mathbf{X}_1\boldsymbol{\beta}_1$. The following corollary proves this invariance property of $\mathbf{M}_2\mathbf{y}$.

Corollary 1. *Let* $\mathbf{X}_1\boldsymbol{\beta}_1$ *be an estimable parametric function under the model* $\mathscr{M}_{12}$*. If a linear statistic* $\mathbf{Ty}$ *is linearly sufficient for* $\mathbf{X}_1\boldsymbol{\beta}_1$*, then* $\mathbf{TM}_2\mathbf{y}$ *is linearly sufficient for* $\mathbf{X}_1\boldsymbol{\beta}_1$*.*

Proof. A statistic $\mathbf{Ty}$ is linearly sufficient for $\mathbf{X}_1\boldsymbol{\beta}_1$ if and only if there exists a matrix $\mathbf{A}$ such that $\mathbf{AT}(\mathbf{X}_1 : \mathbf{X}_2 : \mathbf{VX}^{\perp}) = (\mathbf{X}_1 : \mathbf{0} : \mathbf{0})$. Thus there exists a matrix $\mathbf{B}$ such that $\mathbf{AT} = \mathbf{BM}_2 = \mathbf{ATM}_2$, which proves our claim. □

We may also consider linear minimal sufficiency of the parametric function $\mathbf{X}_1\boldsymbol{\beta}_1$. In view of Baksalary and Kala (1986), a linearly sufficient statistic $\mathbf{Ty}$ is called linearly minimal sufficient for $\mathbf{X}_1\boldsymbol{\beta}_1$ under the model $\mathcal{M}_{12}$, if for any other linearly sufficient statistic $\mathbf{Sy}$, there exists a matrix $\mathbf{A}$ such that $\mathbf{Ty} = \mathbf{ASy}$ almost surely. Applying Baksalary and Kala's (1986, Theorem 2) condition for the estimation of $\mathbf{X}_1\boldsymbol{\beta}_1$, a linear statistic $\mathbf{Ty}$ is linearly minimal sufficient for $\mathbf{X}_1\boldsymbol{\beta}_1$ if and only if the null space equality

$$\mathscr{N}(\mathbf{TX}_1 : \mathbf{TX}_2 : \mathbf{TVX}^{\perp}) = \mathscr{N}(\mathbf{X}_1 : \mathbf{0} : \mathbf{0}) \tag{2.8}$$

holds. The following theorem gives our characterizations for linearly minimal sufficiency of $\mathbf{X}_1\boldsymbol{\beta}_1$.

Theorem 2. *Let $\mathbf{X}_1\boldsymbol{\beta}_1$ be an estimable parametric function under the model $\mathcal{M}_{12}$, and let $\dot{\mathbf{M}}_2$ be the matrix given in* (2.2). *Then a linear statistic $\mathbf{Ty}$ is linearly minimal sufficient for $\mathbf{X}_1\boldsymbol{\beta}_1$ if and only if any of the following equivalent statements holds:*

(a) $\mathscr{N}(\mathbf{TX}_1 : \mathbf{TX}_2 : \mathbf{TVX}^{\perp}) = \mathscr{N}(\mathbf{X}_1 : \mathbf{0} : \mathbf{0})$,
(b) $\mathscr{C}(\mathbf{W}\dot{\mathbf{M}}_2\mathbf{X}_1) = \mathscr{C}(\mathbf{WT}')$,
(c) $\mathscr{N}(\mathbf{T}) \cap \mathscr{C}(\mathbf{W}) = \mathscr{N}(\mathbf{X}_1'\dot{\mathbf{M}}_2) \cap \mathscr{C}(\mathbf{W})$,
(d) $\mathbf{Ty}$ *is linearly sufficient for $\mathbf{X}_1\boldsymbol{\beta}_1$, and there exists a matrix $\mathbf{L}$ such that $\mathbf{Ty} = \mathbf{LX}_1'\dot{\mathbf{M}}_2\mathbf{y}$ almost surely.*

Proof. Part (a) is Baksalary and Kala's (1986, Theorem 2) condition (2.8), and part (b) can be proved by using the same reasoning as Drygas (1983, Theorem 3.4).

Let $\mathbf{Ty}$ be linearly minimal sufficient for $\mathbf{X}_1\boldsymbol{\beta}_1$. Since $\mathbf{X}_1'\dot{\mathbf{M}}_2\mathbf{y}$ is linearly sufficient for $\mathbf{X}_1\boldsymbol{\beta}_1$, there exists a matrix $\mathbf{A}$ such that the equality $\mathbf{Ty} = \mathbf{AX}_1'\dot{\mathbf{M}}_2\mathbf{y}$ holds almost surely, i.e., $\mathbf{TW} = \mathbf{AX}_1'\dot{\mathbf{M}}_2\mathbf{W}$. Thus $\mathscr{C}(\mathbf{W}\dot{\mathbf{M}}_2\mathbf{X}_1)^{\perp} \subset \mathscr{C}(\mathbf{WT}')^{\perp}$, implying the inclusion

$$\mathscr{N}(\mathbf{X}_1'\dot{\mathbf{M}}_2) \cap \mathscr{C}(\mathbf{W}) \subset \mathscr{N}(\mathbf{T}) \cap \mathscr{C}(\mathbf{W}).$$

Hence, together with part (c) at Theorem 1, part (c) follows.

If part (c) holds, then $\mathbf{TW} = \mathbf{BX}_1'\dot{\mathbf{M}}_2\mathbf{W}$ for some matrix $\mathbf{B}$, and furthermore, based on part (d) at Theorem 1, for any other linearly sufficient statistic $\mathbf{Sy}$ it holds that

$$\mathbf{Ty} = \mathbf{BLSy} \tag{2.9}$$

almost surely, for some matrix $\mathbf{L}$. Since part (c) clearly implies linear sufficiency of $\mathbf{Ty}$, the equality (2.9) proves linear minimal sufficiency of $\mathbf{Ty}$.

The equivalence between linear minimal sufficiency and part (d) is clear, and thus our claim is proved.

□

In view of part (a) at Theorem 2, it is easy to see that a statistic $\mathbf{Ty}$ is linearly minimal sufficient for $\mathbf{X}_1\boldsymbol{\beta}_1$ if and only if

$$(\mathbf{TX}_2 : \mathbf{TVX}^{\perp}) = (\mathbf{0} : \mathbf{0}) \quad \text{and} \quad \mathscr{C}(\mathbf{X}_1'\mathbf{T}') = \mathscr{C}(\mathbf{X}_1') \tag{2.10}$$

hold. Also, it is clear from Theorem 2 that the statistic $\mathbf{X}_1'\dot{\mathbf{M}}_2\mathbf{y}$ is linearly minimal sufficient for $\mathbf{X}_1\boldsymbol{\beta}_1$. Note that the statistic $\mathbf{X}'\mathbf{W}^-\mathbf{y}$ is linearly minimal sufficient for the expected value $\mathbf{X}\boldsymbol{\beta}$, and hence $\mathbf{X}_1'\dot{\mathbf{M}}_2\mathbf{y}$ can be considered as a corresponding representation for linearly minimal sufficient statistic in a case of estimation of the parametric function $\mathbf{X}_1\boldsymbol{\beta}_1$.

3. Linear completeness

Besides sufficiency, another important concept of mathematical statistics is the concept of completeness. Under the linear model $\mathcal{M}_{12}$, we may consider the concept of linear completeness. While considering the estimation of the expected value $\mathbf{X}\boldsymbol{\beta}$, Drygas (1983) introduced the notion of linear completeness, and called a linear statistic $\mathbf{Ty}$ linearly complete if for every $\mathbf{LTy}$, such that $\mathrm{E}(\mathbf{LTy}) = \mathbf{0}$, it follows that $\mathbf{LTy} = \mathbf{0}$ almost surely.

Drygas (1983) showed that a statistic $\mathbf{Ty}$ is linearly complete if and only if the column space inclusion

$$\mathscr{C}(\mathbf{TV}) \subset \mathscr{C}(\mathbf{TX})$$

holds. Also, Drygas (1983) proved that a linear statistic $\mathbf{Ty}$ is linearly minimal sufficient for $\mathbf{X}\boldsymbol{\beta}$ if and only if it is simultaneously linearly sufficient and linearly complete for $\mathbf{X}\boldsymbol{\beta}$.

In this paper, we consider a generalization of the concept of linear completeness that would be more applicable for the case of estimating the parametric function $\mathbf{X}_1\boldsymbol{\beta}_1$. We want to generalize the concept of linear completeness in such way that a statistic $\mathbf{Ty}$ can be linearly minimal sufficient for the parametric function $\mathbf{X}_1\boldsymbol{\beta}_1$ if and only if it is linearly sufficient and linearly complete for $\mathbf{X}_1\boldsymbol{\beta}_1$.

Note also that the property $\mathrm{E}(\mathbf{LTy}) = \mathbf{0}$ in Drygas' (1983) definition can equivalently be characterized as the expected value $\mathrm{E}(\mathbf{LTy})$ being independent of $\boldsymbol{\beta}$.

Definition 1. *Let* $\mathbf{X}_1\boldsymbol{\beta}_1$ *be an estimable parametric function under the model* $\mathcal{M}_{12}$. *Then a linear statistic* $\mathbf{Ty}$ *is called linearly complete for* $\mathbf{X}_1\boldsymbol{\beta}_1$ *if for every linear transformation of it,* $\mathbf{LTy}$, *such that the expected value* $\mathrm{E}(\mathbf{LTy})$ *does not dependent on* $\boldsymbol{\beta}_1$ *under the model* $\mathcal{M}_{12}$, *it follows that* $\mathbf{LTy} = \mathbf{0}$ *almost surely.*

The following theorem gives our characterization of linear completeness for $\mathbf{X}_1\boldsymbol{\beta}_1$.

Theorem 3. *Let* $\mathbf{X}_1\boldsymbol{\beta}_1$ *be an estimable parametric function under the model* $\mathcal{M}_{12}$. *Then a linear statistic* $\mathbf{Ty}$ *is linearly complete for* $\mathbf{X}_1\boldsymbol{\beta}_1$ *if and only if*

$$\mathscr{C}(\mathbf{TX}_2 : \mathbf{TV}) \subset \mathscr{C}(\mathbf{TX}_1)$$

holds.

Proof. In Definition 1, the expected value $\mathrm{E}(\mathbf{LTy})$ does not depend on $\boldsymbol{\beta}_1$ under the model $\mathcal{M}_{12}$ if and only if $\mathbf{LTX}_1 = \mathbf{0}$, and $\mathbf{LTy} = \mathbf{0}$ almost surely if and only if $\mathbf{LT}(\mathbf{X}_1 : \mathbf{X}_1 : \mathbf{V}) = (\mathbf{0} : \mathbf{0} : \mathbf{0})$. Thus, according to Definition 1, $\mathbf{Ty}$ is linearly complete for $\mathbf{X}_1\boldsymbol{\beta}_1$ if and only if for every $\mathbf{L}$ such that $\mathbf{LTX}_1 = \mathbf{0}$, it follows that $\mathbf{LT}(\mathbf{X}_2 : \mathbf{V}) = (\mathbf{0} : \mathbf{0})$, or in other words, if and only if for every $\mathbf{L}$ such that $\mathscr{C}(\mathbf{L}') \subset \mathscr{C}(\mathbf{TX}_1)^{\perp}$, it follows that $\mathscr{C}(\mathbf{L}') \subset \mathscr{C}(\mathbf{TX}_2 : \mathbf{TV})^{\perp}$. However, this holds if and only if $\mathscr{C}(\mathbf{TX}_1)^{\perp} \subset \mathscr{C}(\mathbf{TX}_2 : \mathbf{TV})^{\perp}$, i.e., if and only if $\mathscr{C}(\mathbf{TX}_2 : \mathbf{TV}) \subset \mathscr{C}(\mathbf{TX}_1)$. □

Based on Theorem 3, it is easy to verify that if a statistic $\mathbf{Ty}$ is linearly complete for $\mathbf{X}_1\boldsymbol{\beta}_1$, then also the statistic $\mathbf{TM}_2\mathbf{y}$ is linearly complete for $\mathbf{X}_1\boldsymbol{\beta}_1$. Hence the statistic $\mathbf{M}_2\mathbf{y}$ has this sort of invariance property concerning also linear completeness.

Next, we give our main theorem concerning sufficiency, completeness and minimal sufficiency.

Theorem 4. *Let* $\mathbf{X}_1\boldsymbol{\beta}_1$ *be an estimable parametric function under the model* $\mathcal{M}_{12}$. *Then a linear statistic* $\mathbf{Ty}$ *is linearly sufficient and linearly complete for* $\mathbf{X}_1\boldsymbol{\beta}_1$ *if and only if it is linearly minimal sufficient for* $\mathbf{X}_1\boldsymbol{\beta}_1$.

Proof. Let $\mathbf{Ty}$ be linearly sufficient and linearly complete for $\mathbf{X}_1\boldsymbol{\beta}_1$. Then, based on sufficiency, there exists a matrix $\mathbf{A}$ such that

$$(\mathbf{ATX}_1 : \mathbf{ATX}_2 : \mathbf{ATVX}^{\perp}) = (\mathbf{X}_1 : \mathbf{0} : \mathbf{0}). \tag{3.1}$$

On the other hand, completeness implies that,

$$\mathscr{C}(\mathbf{TX}_2 : \mathbf{TVX}^{\perp}) \subset \mathscr{C}(\mathbf{TX}_2 : \mathbf{TV}) \subset \mathscr{C}(\mathbf{TX}_1),$$

i.e.,

$$(\mathbf{TX}_2 : \mathbf{TVX}^{\perp}) = (\mathbf{TX}_1\mathbf{B}_1 : \mathbf{TX}_2\mathbf{B}_2) \tag{3.2}$$

for some $\mathbf{B} = (\mathbf{B}_1 : \mathbf{B}_2)$. Combining (3.1) and (3.2) gives that $\mathbf{ATX}_1(\mathbf{B}_1 : \mathbf{B}_2) = \mathbf{X}_1(\mathbf{B}_1 : \mathbf{B}_2) = (\mathbf{0} : \mathbf{0})$. Thus also the equality

$$(\mathbf{TX}_1\mathbf{B}_1 : \mathbf{TX}_1\mathbf{B}_2) = (\mathbf{TX}_2 : \mathbf{TVX}^{\perp}) = (\mathbf{0} : \mathbf{0})$$

holds, which together with $\mathscr{C}(\mathbf{X}_1'\mathbf{T}') = \mathscr{C}(\mathbf{X}_1')$ implies linear minimal sufficiency of $\mathbf{X}_1\boldsymbol{\beta}_1$.

If $\mathbf{Ty}$ is linearly minimal sufficient for $\mathbf{X}_1\boldsymbol{\beta}_1$, then

$$\begin{aligned}\mathscr{C}(\mathbf{TX}_2 : \mathbf{TV}) \subset \mathscr{C}(\mathbf{TX}_1 : \mathbf{TX}_2 : \mathbf{TV}) &= \mathscr{C}(\mathbf{TX}_1 : \mathbf{TX}_2 : \mathbf{TVX}^{\perp}) \\ &= \mathscr{C}(\mathbf{TX}_1 : \mathbf{0} : \mathbf{0}) = \mathscr{C}(\mathbf{TX}_1),\end{aligned}$$

proving linear completeness. □

4. Predictive approach

It was shown by Drygas (1983) and Müller (1987) that a linear statistic $\mathbf{Ty}$ is linearly sufficient for the expected value $\mathbf{X}\boldsymbol{\beta}$ under the model $\mathcal{M}_{12}$ if and only if the Best Linear Predictor, BLP, of $\mathbf{y}$ based on $\mathbf{Ty}$ does not depend on the vector of parameters $\boldsymbol{\beta}$. That is, $\mathbf{Ty}$ is linearly sufficient for $\mathbf{X}\boldsymbol{\beta}$ if and only if

$$\mathrm{BLP}(\mathbf{y} \mid \mathbf{Ty}) = \mathbf{X}\boldsymbol{\beta} + \mathbf{FT}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) = \mathbf{FTy}, \tag{4.1}$$

almost surely, where $\mathbf{F}$ is any solution of the equation $\mathbf{FTVT}' = \mathbf{VT}'$.

Sengupta and Jammalamadaka (2003, Chapter 11) actually used the property (4.1) as a definition for linear sufficiency, and then gave an interesting study on the linear version of the general estimation theory, including the linear analogues to the Rao–Blackwell and Lehmann–Scheffé Theorems.

However, if a statistic $\mathbf{Ty}$ is linearly sufficient for $\mathbf{X}_1\boldsymbol{\beta}_1$, then the BLP of $\mathbf{y}$ based on $\mathbf{Ty}$ can depend on the vector of parameters $\boldsymbol{\beta}$. This can happen, because linearly sufficient statistic $\mathbf{Ty}$ for $\mathbf{X}_1\boldsymbol{\beta}_1$ may contain only information about subvector $\boldsymbol{\beta}_1$ and not about the whole parameter vector $\boldsymbol{\beta}$. But if we only consider linear sufficiency of $\mathbf{X}_1\boldsymbol{\beta}_1$ under the reduced model $\mathcal{M}_{12\cdot 2} = \{\mathbf{M}_2\mathbf{y},\, \mathbf{M}_2\mathbf{X}_1\boldsymbol{\beta}_1,\, \sigma^2\mathbf{M}_2\mathbf{VM}_2\}$, i.e., linear sufficient statistics of the form $\mathbf{TM}_2\mathbf{y}$, a similar kind of predictive approach for obtaining the best linear unbiased estimator of $\mathbf{X}_1\boldsymbol{\beta}_1$ can be used as was used by Sengupta and Jammalamadaka (2003, Chapter 11) in a case of estimation of the expected value $\mathbf{X}\boldsymbol{\beta}$.

Since we now consider linear sufficiency under the reduced model $\mathcal{M}_{12\cdot 2}$, we can immediately use Drygas' (1983) condition (1.2) to characterize linear sufficiency of $\mathbf{X}_1\boldsymbol{\beta}_1$. That is, a linear statistic $\mathbf{TM}_2\mathbf{y}$ is linearly sufficient for $\mathbf{X}_1\boldsymbol{\beta}_1$ if and only if

$$\mathscr{C}(\mathbf{M}_2\mathbf{X}_1) \subset \mathscr{C}(\mathbf{M}_2\mathbf{W}_1\mathbf{M}_2\mathbf{T}') \tag{4.2}$$

holds. The following theorem gives now a predictive characterization for linear sufficient statistics of the form $\mathbf{TM}_2\mathbf{y}$.

Theorem 5. *Let $\mathbf{X}_1\boldsymbol{\beta}_1$ be an estimable parametric function under the model $\mathcal{M}_{12}$. Then a linear statistic $\mathbf{TM}_2\mathbf{y}$ is linearly sufficient for $\mathbf{X}_1\boldsymbol{\beta}_1$ if and only if the best linear predictor of $\mathbf{M}_2\mathbf{y}$ based on $\mathbf{TM}_2\mathbf{y}$ does not depend* *on $\boldsymbol{\beta}$. That is, a linear statistic $\mathbf{TM}_2\mathbf{y}$ is linearly sufficient for $\mathbf{X}_1\boldsymbol{\beta}_1$ if and only if*

$$\mathrm{BLP}(\mathbf{M}_2\mathbf{y} \mid \mathbf{TM}_2\mathbf{y}) = \mathbf{FTM}_2\mathbf{y}$$

almost surely, where $\mathbf{F}$ is any solution of the equation

$$\mathbf{FTM}_2\mathbf{VM}_2\mathbf{T}' = \mathbf{M}_2\mathbf{VM}_2\mathbf{T}'. \tag{4.3}$$

Proof. Let $\mathbf{TM}_2\mathbf{y}$ be linearly sufficient for $\mathbf{X}_1\boldsymbol{\beta}_1$. Then, based on the inclusion (4.2), there exists a matrix $\mathbf{A}$ such that $\mathbf{M}_2\mathbf{X}_1 = \mathbf{M}_2\mathbf{W}_1\mathbf{M}_2\mathbf{T}'\mathbf{A}$.

Let us consider the matrix

$$\mathbf{F} = \mathbf{M}_2\mathbf{W}_1\mathbf{M}_2\mathbf{T}'(\mathbf{TM}_2\mathbf{W}_1\mathbf{M}_2\mathbf{T}')^-. \tag{4.4}$$

Then clearly

$$\mathbf{FTM}_2\mathbf{W}_1\mathbf{M}_2\mathbf{T}' = \mathbf{M}_2\mathbf{W}_1\mathbf{M}_2\mathbf{T}'. \tag{4.5}$$

Since now

$$\begin{aligned}\mathbf{FTM}_2\mathbf{X}_1 &= \mathbf{FTM}_2\mathbf{W}_1\mathbf{M}_2\mathbf{T}'\mathbf{A}\\ &= \mathbf{M}_2\mathbf{W}_1\mathbf{M}_2\mathbf{T}'\mathbf{A} = \mathbf{M}_2\mathbf{X}_1,\end{aligned} \tag{4.6}$$

the equation (4.5) becomes

$$\begin{aligned}\mathbf{FTM}_2\mathbf{W}_1\mathbf{M}_2\mathbf{T}' &= \mathbf{FTM}_2\mathbf{VM}_2\mathbf{T}' + \mathbf{M}_2\mathbf{X}_1\mathbf{X}_1'\mathbf{M}_2\mathbf{T}'\\ &= \mathbf{M}_2\mathbf{VM}_2\mathbf{T}' + \mathbf{M}_2\mathbf{X}_1\mathbf{X}_1'\mathbf{M}_2\mathbf{T}',\end{aligned} \tag{4.7}$$

implying that $\mathbf{FTM}_2\mathbf{VM}_2\mathbf{T}' = \mathbf{M}_2\mathbf{VM}_2\mathbf{T}'$, i.e., $\mathbf{F}$ is a solution of the equation (4.3). Thus in view of equations (4.6) and (4.7), the $\mathrm{BLP}(\mathbf{M}_2\mathbf{y} \mid \mathbf{TM}_2\mathbf{y})$ has the following representation:

$$\begin{aligned}\mathrm{BLP}(\mathbf{M}_2\mathbf{y} \mid \mathbf{TM}_2\mathbf{y}) &= \mathbf{M}_2\mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{FT}(\mathbf{M}_2\mathbf{y} - \mathbf{M}_2\mathbf{X}_1\boldsymbol{\beta}_1)\\ &= \mathbf{FTM}_2\mathbf{y},\end{aligned}$$

i.e., the $\mathrm{BLP}(\mathbf{M}_2\mathbf{y} \mid \mathbf{TM}_2\mathbf{y})$ does not depend on $\boldsymbol{\beta}$.

Conversely, assume that $\mathrm{BLP}(\mathbf{M}_2\mathbf{y} \mid \mathbf{TM}_2\mathbf{y})$ does not depend on $\boldsymbol{\beta}$, (or actually on $\boldsymbol{\beta}_1$), i.e., $\mathbf{FTM}_2\mathbf{X}_1 = \mathbf{M}_2\mathbf{X}_1$ for any matrix $\mathbf{F}$ such that the equation (4.3) holds.

Since

$$\begin{aligned}\mathscr{C}(\mathbf{M}_2\mathbf{X}_1) \subset \mathscr{C}(\mathbf{M}_2\mathbf{W}_1\mathbf{M}_2) \subset \mathscr{C}&[\mathbf{M}_2\mathbf{W}_1\mathbf{M}_2 : (\mathbf{T}')^\perp]\\ &= \mathscr{C}[\mathbf{M}_2\mathbf{W}_1\mathbf{M}_2\mathbf{T}' : (\mathbf{T}')^\perp],\end{aligned}$$

there exists a matrix $\mathbf{B} = (\mathbf{B}_1 : \mathbf{B}_2)$ such that

$$\mathbf{M}_2\mathbf{X}_1 = \mathbf{M}_2\mathbf{W}_1\mathbf{M}_2\mathbf{T}'\mathbf{B}_1 + (\mathbf{T}')^\perp\mathbf{B}_2,$$

and then

$$\begin{aligned}\mathbf{FTM}_2\mathbf{X}_1 &= \mathbf{FT}(\mathbf{M}_2\mathbf{W}_1\mathbf{M}_2\mathbf{T}'\mathbf{B}_1 + (\mathbf{T}')^{\perp}\mathbf{B}_2)\\ &= \mathbf{FT}(\mathbf{M}_2\mathbf{V}\mathbf{M}_2\mathbf{T}'\mathbf{B}_1 + \mathbf{M}_2\mathbf{X}_1\mathbf{X}_1'\mathbf{M}_2\mathbf{T}'\mathbf{B}_1)\\ &= \mathbf{M}_2\mathbf{V}\mathbf{M}_2\mathbf{T}'\mathbf{B}_1 + \mathbf{M}_2\mathbf{X}_1\mathbf{X}_1'\mathbf{M}_2\mathbf{T}'\mathbf{B}_1\\ &= \mathbf{M}_2\mathbf{W}_1\mathbf{M}_2\mathbf{T}'\mathbf{B}_1.\end{aligned}$$

This shows that $\mathbf{M}_2\mathbf{W}_1\mathbf{M}_2\mathbf{T}'\mathbf{B}_1 = \mathbf{M}_2\mathbf{X}_1$, which is equivalent to the column space inclusion (4.2), and hence proves linear sufficiency of $\mathbf{TM}_2\mathbf{y}$. □

In the general setup, the uniformly minimum variance unbiased estimators are based on the Rao–Blackwell and Lehmann–Scheffé Theorems. Correspondingly, linear versions of above concepts can be defined and then used as an alternative route to obtain the BLUEs. In the following theorems, we give linear analogues to the Rao–Blackwell and Lehmann–Scheffé Theorems in a case of estimation of $\mathbf{X}_1\boldsymbol{\beta}_1$.

Theorem 6. *Let $\mathbf{X}_1\boldsymbol{\beta}_1$ be an estimable parametric function under the model $\mathcal{M}_{12}$. Furthermore, let $\mathbf{TM}_2\mathbf{y}$ be linearly sufficient for $\mathbf{X}_1\boldsymbol{\beta}_1$, and $\mathbf{Ly}$ be an unbiased estimator for $\mathbf{X}_1\boldsymbol{\beta}_1$. Then the best linear predictor of $\mathbf{Ly}$ based on $\mathbf{TM}_2\mathbf{y}$, $\mathrm{BLP}(\mathbf{Ly} \mid \mathbf{TM}_2\mathbf{y})$, is a linear unbiased estimator for $\mathbf{X}_1\boldsymbol{\beta}_1$ with*

$$\mathrm{cov}(\mathbf{Ly}) - \mathrm{cov}[\mathrm{BLP}(\mathbf{Ly} \mid \mathbf{TM}_2\mathbf{y})] \geq \mathbf{0}.$$

Proof. Since $\mathbf{Ly}$ is an unbiased estimator, $\mathbf{L} = \mathbf{AM}_2$ for some matrix $\mathbf{A}$. Therefore, because of linear sufficiency of $\mathbf{TM}_2\mathbf{y}$,

$$\begin{aligned}\mathrm{BLP}(\mathbf{Ly} \mid \mathbf{TM}_2\mathbf{y}) &= \mathrm{BLP}(\mathbf{AM}_2\mathbf{y} \mid \mathbf{TM}_2\mathbf{y})\\ &= \mathbf{A}\cdot\mathrm{BLP}(\mathbf{M}_2\mathbf{y} \mid \mathbf{TM}_2\mathbf{y}) = \mathbf{AFTM}_2\mathbf{y},\end{aligned}$$

where $\mathbf{F}$ is any solution of the equation (4.3). This shows that $\mathrm{BLP}(\mathbf{Ly} \mid \mathbf{TM}_2\mathbf{y})$ is a linear estimator. Clearly, $\mathrm{BLP}(\mathbf{Ly} \mid \mathbf{TM}_2\mathbf{y})$ is unbiased for $\mathbf{X}_1\boldsymbol{\beta}_1$, and $\mathbf{Ly} - \mathrm{BLP}(\mathbf{Ly} \mid \mathbf{TM}_2\mathbf{y})$ is uncorrelated with $\mathbf{TM}_2\mathbf{y}$, see Christensen(2001, Section 3.1). Thus

$$\begin{aligned}\mathrm{cov}(\mathbf{Ly}) &= \mathrm{cov}[\mathbf{Ly} - \mathrm{BLP}(\mathbf{Ly} \mid \mathbf{TM}_2\mathbf{y}) + \mathrm{BLP}(\mathbf{Ly} \mid \mathbf{TM}_2\mathbf{y})]\\ &= \mathrm{cov}[\mathbf{Ly} - \mathrm{BLP}(\mathbf{Ly} \mid \mathbf{TM}_2\mathbf{y})] + \mathrm{cov}[\mathrm{BLP}(\mathbf{Ly} \mid \mathbf{TM}_2\mathbf{y})]\\ &\geq \mathrm{cov}[\mathrm{BLP}(\mathbf{Ly} \mid \mathbf{TM}_2\mathbf{y})].\end{aligned}$$

□

Theorem 7. *Let $\mathbf{X}_1\boldsymbol{\beta}_1$ be an estimable parametric function under the model $\mathcal{M}_{12}$. Furthermore, let $\mathbf{TM}_2\mathbf{y}$ be linearly sufficient and linearly complete for $\mathbf{X}_1\boldsymbol{\beta}_1$, and $\mathbf{Ly}$ be an unbiased estimator for $\mathbf{X}_1\boldsymbol{\beta}_1$. Then*

$$\mathrm{BLUE}(\mathbf{X}_1\boldsymbol{\beta}_1 \mid \mathcal{M}_{12}) = \mathrm{BLP}(\mathbf{Ly} \mid \mathbf{TM}_2\mathbf{y})$$

almost surely.

Proof. Denote $\mathbf{ATM_2y} = \mathrm{BLP}(\mathbf{Ly} \mid \mathbf{TM_2y})$, and let $\mathbf{Sy}$ be an unbiased estimator for $\mathbf{X}_1\boldsymbol{\beta}_1$ with $\mathrm{cov}(\mathbf{ATM_2y}) - \mathrm{cov}(\mathbf{Sy}) \geq \mathbf{0}$. Let $\mathbf{BTM_2y} = \mathrm{BLP}(\mathbf{Sy} \mid \mathbf{TM_2y})$, and therefore $\mathbf{BTM_2y}$ is also an unbiased estimator for $\mathbf{X}_1\boldsymbol{\beta}_1$ with $\mathrm{cov}(\mathbf{ATM_2y}) - \mathrm{cov}(\mathbf{BTM_2y}) \geq \mathbf{0}$. However, since $\mathrm{E}[(\mathbf{A} - \mathbf{B})\mathbf{TM_2y}] = \mathbf{0}$, linear completeness of $\mathbf{TM_2y}$ implies that $\mathbf{ATM_2y} = \mathbf{BTM_2y}$ almost surely, and thus also $\mathrm{cov}(\mathbf{ATM_2y}) = \mathrm{cov}(\mathbf{BTM_2y})$. □

5. Example

Let us consider the following seemingly unrelated regression model $\mathcal{M}_{12} = \{\mathbf{y}, \mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{V}\}$, where

$$\mathbf{y} = \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix}, \quad \mathbf{X} = (\mathbf{X}_1 : \mathbf{X}_2) = \begin{pmatrix} \mathbf{X}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_{22} \end{pmatrix},$$
$$\boldsymbol{\beta} = \begin{pmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \end{pmatrix}, \quad \mathbf{V} = \begin{pmatrix} \mathbf{V}_{11} & \mathbf{V}_{12} \\ \mathbf{V}_{21} & \mathbf{V}_{22} \end{pmatrix}.$$

Then $\mathbf{X}_{11}\boldsymbol{\beta}_1$ is estimable under the model $\mathcal{M}_{12}$. Note that the orthogonal projection, i.e., the ordinary least squares estimator

$$\mathbf{P}_{\mathbf{X}_{11}}\mathbf{y}_1 = \mathbf{P}_{11}\mathbf{y}_1 = \mathbf{X}_{11}(\mathbf{X}_{11}'\mathbf{X}_{11})^-\mathbf{X}_{11}'\mathbf{y}_1$$

under the small model $\{\mathbf{y}_1, \mathbf{X}_{11}\boldsymbol{\beta}_1, \sigma^2\mathbf{V}_{11}\}$ is also an unbiased estimator of $\mathbf{X}_{11}\boldsymbol{\beta}_1$ under the full model $\mathcal{M}_{12}$.

Let us now start considering linear minimal sufficiency of $\mathbf{X}_{11}\boldsymbol{\beta}_1$. Denote $\mathbf{M}_{22} = \mathbf{M}_{\mathbf{X}_{22}}$, $\dot{\mathbf{M}}_{22} = \mathbf{M}_{22}(\mathbf{M}_{22}\mathbf{V}_{22}\mathbf{M}_{22})^-\mathbf{M}_{22}$, and $\mathbf{W}_{11} = \mathbf{V}_{11} + \mathbf{X}_{11}\mathbf{X}_{11}'$, and let a linear statistic $\mathbf{TM_2y}$ have the following form:

$$\begin{aligned} \mathbf{TM_2y} &= (\mathbf{T}_1 : \mathbf{T}_2)\begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_{22} \end{pmatrix}\begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix} = \mathbf{T}_1\mathbf{y}_1 + \mathbf{T}_2\mathbf{M}_{22}\mathbf{y}_2 \\ &= \mathbf{X}_{11}'(\mathbf{W}_{11} - \mathbf{V}_{12}\dot{\mathbf{M}}_{22}\mathbf{V}_{21})^-[\mathbf{y}_1 - \mathbf{V}_{12}\dot{\mathbf{M}}_{22}\mathbf{y}_2], \end{aligned}$$

where

$$\begin{aligned} \mathbf{T}_1\mathbf{y}_1 &= \mathbf{X}_{11}'(\mathbf{W}_{11} - \mathbf{V}_{12}\dot{\mathbf{M}}_{22}\mathbf{V}_{21})^-\mathbf{y}_1, \\ \mathbf{T}_2\mathbf{M}_{22}\mathbf{y}_2 &= -\mathbf{X}_{11}'(\mathbf{W}_{11} - \mathbf{V}_{12}\dot{\mathbf{M}}_{22}\mathbf{V}_{21})^-\mathbf{V}_{12}\dot{\mathbf{M}}_{22}\mathbf{y}_2. \end{aligned}$$

Then $\mathbf{TM_2y}$ is linearly minimal sufficient for $\mathbf{X}_{11}\boldsymbol{\beta}_1$ (and also for $\mathbf{X}_1\boldsymbol{\beta}_1$) under the seemingly unrelated regression model $\mathcal{M}_{12}$. To see this, note that $\mathscr{C}(\mathbf{X}_{11}) \subset \mathscr{C}(\mathbf{W}_{11} - \mathbf{V}_{12}\dot{\mathbf{M}}_{22}\mathbf{V}_{21})$ since

$$\begin{aligned} &\mathscr{C}\left[(\mathbf{I} : -\mathbf{V}_{12}\dot{\mathbf{M}}_{22})\begin{pmatrix} \mathbf{X}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_{22} \end{pmatrix}\right] \\ &\quad \subset \mathscr{C}\left[(\mathbf{I} : -\mathbf{V}_{12}\dot{\mathbf{M}}_{22})\begin{pmatrix} \mathbf{W}_{11} & \mathbf{V}_{12} \\ \mathbf{V}_{21} & \mathbf{V}_{22} \end{pmatrix}(\mathbf{I} : -\mathbf{V}_{12}\dot{\mathbf{M}}_{22})'\right], \end{aligned}$$

and hence

$$\mathbf{X}_{11}'(\mathbf{W}_{11} - \mathbf{V}_{12}\dot{\mathbf{M}}_{22}\mathbf{V}_{21})^{-}(\mathbf{I} : -\mathbf{V}_{12}\dot{\mathbf{M}}_{22})\begin{pmatrix}\mathbf{V}_{11}\mathbf{M}_{11} & \mathbf{V}_{12}\mathbf{M}_{22}\\ \mathbf{V}_{21}\mathbf{M}_{11} & \mathbf{V}_{22}\mathbf{M}_{22}\end{pmatrix} = (\mathbf{0} : \mathbf{0}), \tag{5.1}$$

$$\mathscr{C}\left[(\mathbf{X}_{11}' : \mathbf{0})\begin{pmatrix}\mathbf{T}_1' & \mathbf{0}\\ \mathbf{M}_{22}\mathbf{T}_2' & \mathbf{0}\end{pmatrix}\right] = \mathscr{C}(\mathbf{X}_{11}' : \mathbf{0}), \tag{5.2}$$

where $\mathbf{M}_{11} = \mathbf{M}_{\mathbf{X}_{11}}$. The equations (5.1) and (5.2) together show, in view of (2.10), linear minimal sufficiency of $\mathbf{TM}_2\mathbf{y}$.

We can now use a predictive approach, i.e., Theorem 7, to convert an unbiased estimator $\mathbf{P}_{11}\mathbf{y}_1$ into the BLUE of $\mathbf{X}_{11}\boldsymbol{\beta}_1$. That is,

$$\mathrm{BLUE}(\mathbf{X}_{11}\boldsymbol{\beta}_1 \mid \mathscr{M}_{12}) = \mathrm{BLP}(\mathbf{P}_{11}\mathbf{y}_1 \mid \mathbf{TM}_2\mathbf{y}) = \mathbf{FTM}_2\mathbf{y},$$

where $\mathbf{F}$ is any solution of the equation

$$\mathbf{FTM}_2\mathbf{VM}_2\mathbf{T}' = \mathbf{P}_{11}(\mathbf{V}_{11} : \mathbf{V}_{12})\mathbf{M}_2\mathbf{T}'. \tag{5.3}$$

Correspondingly to the representation (4.4), one solution for the equation (5.3) with respect to $\mathbf{F}$ is

$$\begin{aligned}\mathbf{F} &= \mathbf{P}_{11}(\mathbf{W}_{11} : \mathbf{V}_{12})\mathbf{M}_2\mathbf{T}'\left[\mathbf{TM}_2\begin{pmatrix}\mathbf{W}_{11} & \mathbf{V}_{12}\\ \mathbf{V}_{21} & \mathbf{V}_{22}\end{pmatrix}\mathbf{M}_2\mathbf{T}'\right]^{-}\\ &= \mathbf{X}_{11}[\mathbf{X}_{11}'(\mathbf{W}_{11} - \mathbf{V}_{12}\dot{\mathbf{M}}_{22}\mathbf{V}_{21})^{-}\mathbf{X}_{11}]^{-} = \mathbf{X}_{11}(\mathbf{X}_{11}'\boldsymbol{\Sigma}_{11\cdot 22}^{-}\mathbf{X}_{11})^{-},\end{aligned}$$

where $\boldsymbol{\Sigma}_{11\cdot 22} = \mathbf{W}_{11} - \mathbf{V}_{12}\dot{\mathbf{M}}_{22}\mathbf{V}_{21}$. Thus the BLUE of $\mathbf{X}_{11}\boldsymbol{\beta}_1$ under the seemingly unrelated regression model $\mathscr{M}_{12}$ has the following representation:

$$\begin{aligned}\mathrm{BLUE}(\mathbf{X}_{11}\boldsymbol{\beta}_1 \mid \mathscr{M}_{12}) &= \mathbf{X}_{11}(\mathbf{X}_{11}'\boldsymbol{\Sigma}_{11\cdot 22}^{-}\mathbf{X}_{11})^{-}\mathbf{X}_{11}'\boldsymbol{\Sigma}_{11\cdot 22}^{-}\mathbf{y}_1\\ &\quad - \mathbf{X}_{11}(\mathbf{X}_{11}'\boldsymbol{\Sigma}_{11\cdot 22}^{-}\mathbf{X}_{11})^{-}\mathbf{X}_{11}'\boldsymbol{\Sigma}_{11\cdot 22}^{-}\mathbf{V}_{12}\dot{\mathbf{M}}_{22}\mathbf{y}_2.\end{aligned}$$

References

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

Baksalary, J. K., and Kala, R. (1981), *Linear transformations preserving best linear unbiased estimators in a general Gauss–Markoff model*, Ann. Statist. **4**, 913–916.

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

Bhimasankaram, P., and Sengupta, D. (1996), *The linear zero functions approach to linear models*, Sankhyā Ser. B **58**, 338–351.

Christensen, R. (2001), *Advanced Linear Modeling*, Second Edition, Springer-Verlag, New York.

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

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

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

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

Isotalo, J., and Puntanen, S. (2006), *Linear sufficiency and completeness of the estimable parametric function in the general Gauss–Markov model*, University of Tampere, Dept. of Mathematics, Statistics & Philosophy, Report A 370.

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

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

Rao, C. R., and Mitra, S. K. (1971), *Generalized Inverse of Matrices and Its Applications*, John Wiley & Sons, Inc., New York – London – Sydney.

Sengupta, D., and Jammalamadaka, S. R. (2003), *Linear Models. An Integrated Approach*, World Scientific Publishing Co., River Edge, NJ.

Department of Mathematics, Statistics & Philosophy, FI-33014 University of Tampere, Finland

E-mail address: jarkko.isotalo@uta.fi

E-mail address: simo.puntanen@uta.fi