EXTENSIONS OF THE
FRISCH–WAUGH–LOVELL THEOREM*

Jürgen Groß

Department of Statistics, University of Dortmund
Vogelpothsweg 87, D–44221 Dortmund, Germany

e-mail gross@amadeus.statistik.uni-dortmund.de

AND

Simo Puntanen

Department of Mathematics, Statistics and Philosophy
FI–33014 University of Tampere, Finland

e-mail sjp@uta.fi

Abstract

In this paper we introduce extensions of the so-called Frisch–Waugh–Lovell Theorem. This is done by employing the close relationship between the concept of linear sufficiency and the appropriate reduction of linear models. Some specific reduced models which demonstrate alternatives to the Frisch–Waugh–Lovell procedure are discussed.

Keywords: best linear unbiased estimation; Frisch–Waugh–Lovell Theorem; linear sufficiency; orthogonal projector; partitioned linear model; reduced linear model.

2000 Mathematics Subject Classification: 62J05, 62H12.

---

1. Introduction

Let $\mathbb{R}_{m,n}$ denote the set of $m \times n$ real matrices. The symbols $\mathbf{A}'$, $\mathbf{A}^+$, $\mathcal{C}(\mathbf{A})$, $\mathcal{C}(\mathbf{A})^\perp$ $\mathcal{N}(\mathbf{A})$, and $\mathrm{r}(\mathbf{A})$ will stand for the transpose, 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 $\mathcal{C}(\mathbf{A}^\perp) = \mathcal{N}(\mathbf{A}') = \mathcal{C}(\mathbf{A})^\perp$. Further, $\mathbf{P_A} = \mathbf{A}\mathbf{A}^+$ denotes the orthogonal projector (with respect to the standard inner product) onto $\mathcal{C}(\mathbf{A})$, and $\mathbf{M_A} = \mathbf{I} - \mathbf{P_A}$. In particular, we denote $\mathbf{P}_i = \mathbf{P}_{\mathbf{X}_i}$, $\mathbf{M}_i = \mathbf{I} - \mathbf{P}_i$, $i = 1, 2$. For matrices $\mathbf{A}$ and $\mathbf{B}$ with the same number of rows, $(\mathbf{A} : \mathbf{B})$ denotes the partitioned matrix with $\mathbf{A}$ and $\mathbf{B}$ as submatrices.

Consider a general Gauss-Markov model denoted by

$$(1.1) \qquad \mathcal{M} = \{\mathbf{y}, \mathbf{X}\boldsymbol{\beta}, \sigma^2 \mathbf{V}\}, \quad \mathsf{E}(\mathbf{y}) = \mathbf{X}\boldsymbol{\beta}, \quad \mathrm{cov}(\mathbf{y}) = \sigma^2 \mathbf{V},$$

where $\mathbf{X}$ is a known $n \times p$ matrix, $\boldsymbol{\beta}$ is a $p \times 1$ vector of unknown parameters, $\mathbf{V}$ is a known $n \times n$ nonnegative definite matrix, and $\sigma^2 > 0$ is an unknown scalar. $\mathsf{E}(\cdot)$ and $\mathrm{cov}(\cdot)$ denote expectation and dispersion of a random vector argument. It is assumed that the model is *consistent*, that is,

$$(1.2) \qquad \qquad \mathbf{y} \in \mathcal{C}(\mathbf{X} : \mathbf{V}),$$

see Rao (1971) and Feuerverger and Fraser (1980).

A vector of parametric functions $\mathbf{K}\boldsymbol{\beta}$, where $\mathbf{K} \in \mathbb{R}_{k,p}$, is estimable under the model $\mathcal{M}$ if and only if $\mathcal{C}(\mathbf{K}') \subseteq \mathcal{C}(\mathbf{X}')$. The best linear unbiased estimator (BLUE) for an estimable vector of parametric functions $\mathbf{K}\boldsymbol{\beta}$ is given by $\mathbf{G}\mathbf{y}$, where $\mathbf{G} \in \mathbb{R}_{k,n}$ is any solution to the set of equations

$$(1.3) \qquad \qquad \mathbf{G}(\mathbf{X} : \mathbf{V}\mathbf{X}^\perp) = (\mathbf{K} : \mathbf{0}),$$

see, e.g., Rao (1973, p. 282).

By partitioning $\mathbf{X} = (\mathbf{X}_1 : \mathbf{X}_2)$ so that $\mathbf{X}_1$ has p_1 columns and $\mathbf{X}_2$ has p_2 columns with $p = p_1 + p_2$, and by accordingly writing $\boldsymbol{\beta} = (\boldsymbol{\beta}_1', \boldsymbol{\beta}_2')'$, we can express $\mathcal{M}$ in a partitioned form

$$(1.4) \qquad \qquad \mathcal{M} = \{\mathbf{y}, \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2, \sigma^2 \mathbf{V}\}.$$

Regarding $\boldsymbol{\beta}_1$ as a nuisance parameter, our interest focuses on estimation of a vector of estimable parametric functions $\mathbf{K}_2\boldsymbol{\beta}_2 = \mathbf{K}\boldsymbol{\beta} = (\mathbf{K}_1 : \mathbf{K}_2)\boldsymbol{\beta}$, where $\mathbf{K}_1 = \mathbf{0}$. As noted by Groß and Puntanen (2000, Lemma 1), $\mathbf{K}_2\boldsymbol{\beta}_2$ is estimable under $\mathcal{M}$ if and only if $\mathcal{C}(\mathbf{K}_2') \subseteq \mathcal{C}(\mathbf{X}_2'\mathbf{M}_1)$, where $\mathbf{M}_1 = \mathbf{I} - \mathbf{P}_1$ is the orthogonal projector onto $\mathcal{N}(\mathbf{X}_1')$. Hence, the BLUE for an arbitrary estimable vector $\mathbf{K}_2\boldsymbol{\beta}_2$ may easily be computed from the BLUE of $\mathbf{M}_1\mathbf{X}_2\boldsymbol{\beta}_2$. Let $\mathbf{X}$ have full rank and consider the model equation

$$(1.5) \qquad \mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}.$$

Premultiplying this equation by $\mathbf{M}_1$ yields the reduced model

$$(1.6) \qquad \{\mathbf{M}_1\mathbf{y}, \mathbf{M}_1\mathbf{X}_2\boldsymbol{\beta}_2, \operatorname{cov}(\mathbf{M}_1\boldsymbol{\varepsilon})\}.$$

The so-called Frisch–Waugh–Lovell Theorem states that the ordinary least squares estimator (OLSE) of $\boldsymbol{\beta}_2$ under this reduced model equals the OLSE of $\boldsymbol{\beta}_2$ under the original partitioned model. We may cite Davidson and MacKinnon (1993, p. 19) who use the name Frisch–Waugh–Lovell Theorem, after Frisch and Waugh (1933) and Lovell (1963): '…since those papers seem to have introduced, and then reintroduced, it to econometricians.'

A generalized version of the Frisch–Waugh–Lovell Theorem to the case of possibly singular $\mathbf{V}$ and possibly non-estimable $\boldsymbol{\beta}_2$ claims that every BLUE of $\mathbf{M}_1\mathbf{X}_2\boldsymbol{\beta}_2$ under the reduced model (1.6) remains BLUE of $\mathbf{M}_1\mathbf{X}_2\boldsymbol{\beta}_2$ under the partitioned model $\mathcal{M}$, see, e.g., Groß and Puntanen (2000, Theorem 4) and Bhimasankaram and Sengupta (1996, Theorem 6.1). In the following section a more general reduction procedure is carried out.

2. Reduced models

Since we are not interested in the parameter vector $\boldsymbol{\beta}_1$, we may consider a reduction of the model $\mathcal{M}$ by a transformation of $\mathbf{y}$ into $\mathbf{F}\mathbf{y}$, where $\mathbf{F}$ is any matrix such that $\mathbf{F}\mathbf{X}_1\boldsymbol{\beta}_1 = \mathbf{0}$ for all $\boldsymbol{\beta}_1 \in \mathbb{R}_{p_1 \times 1}$. Hence, we obtain the reduced model

$$(2.1) \qquad \mathcal{M}_r(\mathbf{F}) = \{\mathbf{F}\mathbf{y}, \mathbf{F}\mathbf{X}_2\boldsymbol{\beta}_2, \sigma^2\mathbf{F}\mathbf{V}\mathbf{F}'\} \quad \text{subject to } \mathbf{F}\mathbf{X}_1 = \mathbf{0}.$$

We are interested in the following two questions: (1) For which transformation matrices $\mathbf{F}$ (where $\mathbf{F}\mathbf{X}_1 = \mathbf{0}$) is every representation of the BLUE of $\mathbf{M}_1\mathbf{X}_2\boldsymbol{\beta}_2$ in the reduced model $\mathcal{M}_r(\mathbf{F})$ also BLUE of $\mathbf{M}_1\mathbf{X}_2\boldsymbol{\beta}_2$ in the partitioned model $\mathcal{M}$? (2) Are there representations of the BLUE of $\mathbf{M}_1\mathbf{X}_2\boldsymbol{\beta}_2$ in the partitioned model $\mathcal{M}$ which are also BLUE of $\mathbf{M}_1\mathbf{X}_2\boldsymbol{\beta}_2$ in the reduced model $\mathcal{M}_r(\mathbf{F})$?

Question (1) can be answered by employing the close relationship between the concept of linear sufficiency and the appropriate reduction of linear models. Recall that a linear transformation $\mathbf{F}\mathbf{y}$ (where not necessarily $\mathbf{F}\mathbf{X}_1 = \mathbf{0}$) is called linearly sufficient for an estimable vector of parametric functions $\mathbf{K}\boldsymbol{\beta}$ in model $\mathcal{M}$ if and only if there exists a matrix $\mathbf{A}$ such that $\mathbf{A}\mathbf{F}\mathbf{y}$ is BLUE of $\mathbf{K}\boldsymbol{\beta}$. In that case, every representation of the BLUE of $\mathbf{K}\boldsymbol{\beta}$ in the induced model

$$\mathcal{M}(\mathbf{F}) = \{\mathbf{F}\mathbf{y}, \mathbf{F}\mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{F}\mathbf{X}_2\boldsymbol{\beta}_2, \sigma^2\mathbf{F}\mathbf{V}\mathbf{F}'\},$$

is also BLUE in the partitioned model $\mathcal{M}$, cf. Baksalary and Kala (1986), and Drygas (1983).

As mentioned in the introduction, the choice $\mathbf{F} = \mathbf{M}_1$ is valid to satisfy the required property that the BLUE of $\mathbf{M}_1\mathbf{X}_2\boldsymbol{\beta}_2$ in the reduced model $\mathcal{M}_r(\mathbf{M}_1)$ is also BLUE in the partitioned model $\mathcal{M}$. The model $\mathcal{M}_r(\mathbf{M}_1)$ has been studied for example in Nurhonen and Puntanen (1992), Puntanen (1996, 1997), Bhimasankaram *et al.* (1996, 1997, 1998) and Groß and Puntanen (2000). The obtained results can be seen as generalizations of the Frisch–Waugh–Lovell Theorem to the case of possibly singular $\mathbf{V}$ and possibly non-estimable $\boldsymbol{\beta}_2$.

As noted above, it is our intention here to demonstrate that further choices for $\mathbf{F}$ are possible. An algebraic characterization of such matrices is provided in the following theorem.

Theorem 1. *Let $\mathbf{F}$ be any matrix satisfying the conditions*

(2.2) $\mathbf{F}\mathbf{X}_1 = \mathbf{0},$

(2.3) $\mathcal{N}(\mathbf{F}) \cap \mathcal{C}(\mathbf{M}_1\mathbf{X}_2) = \{\mathbf{0}\},$

(2.4) $\mathcal{C}(\mathbf{F}\mathbf{X}_2) \cap \mathcal{C}[\mathbf{F}\mathbf{V}(\mathbf{X}_1 : \mathbf{X}_2)^{\perp}] = \{\mathbf{0}\}.$

Then the following four statements hold:

(i) $\mathbf{M}_1\mathbf{X}_2\boldsymbol{\beta}_2$ *is estimable in the reduced model* $\mathcal{M}_r(\mathbf{F})$;

(ii) $\mathbf{Fy}$ *is linearly sufficient for* $\mathbf{M}_1\mathbf{X}_2\boldsymbol{\beta}_2$ *in the partitioned model* $\mathcal{M}$;

(iii) *every representation of the BLUE of* $\mathbf{M}_1\mathbf{X}_2\boldsymbol{\beta}_2$ *in the reduced model* $\mathcal{M}_r(\mathbf{F})$ *is also the BLUE of* $\mathbf{M}_1\mathbf{X}_2\boldsymbol{\beta}_2$ *in the partitioned model* $\mathcal{M}$;

(iv) *there exists at least one representation of the BLUE of* $\mathbf{M}_1\mathbf{X}_2\boldsymbol{\beta}_2$ *in the partitioned model* $\mathcal{M}$ *which is also BLUE of* $\mathbf{M}_1\mathbf{X}_2\boldsymbol{\beta}_2$ *in the reduced model* $\mathcal{M}_r(\mathbf{F})$.

Proof. By using the fact that

$$\mathrm{r}(\mathbf{M}_1\mathbf{X}_2) - \mathrm{r}(\mathbf{FM}_1\mathbf{X}_2) = \dim[\mathcal{N}(\mathbf{F}) \cap \mathcal{C}(\mathbf{M}_1\mathbf{X}_2)],$$

see, e.g., Corollary 6.2 in Marsaglia and Styan (1974), it follows that (2.3) is equivalent to $\mathcal{C}(\mathbf{X}_2'\mathbf{M}_1) = \mathcal{C}(\mathbf{X}_2'\mathbf{M}_1\mathbf{F}')$. Since (2.2) may be expressed as $\mathbf{FM}_1 = \mathbf{F}$ we obtain $\mathcal{C}(\mathbf{X}_2'\mathbf{M}_1) \subseteq \mathcal{C}(\mathbf{X}_2'\mathbf{F}')$, which is the condition for estimability of $\mathbf{M}_1\mathbf{X}_2\boldsymbol{\beta}_2$ in $\mathcal{M}_r(\mathbf{F})$; thus statement (i) is proved. Moreover, in view of (2.2), the condition $\mathcal{C}(\mathbf{X}_2'\mathbf{M}_1) = \mathcal{C}(\mathbf{X}_2'\mathbf{M}_1\mathbf{F}')$ may equivalently be written as

$$\mathcal{C}(\mathbf{K}') = \mathcal{C}(\mathbf{X}'\mathbf{F}'), \quad \mathbf{K} = (\mathbf{0} : \mathbf{M}_1\mathbf{X}_2), \quad \mathbf{X} = (\mathbf{X}_1 : \mathbf{X}_2).$$

Using condition (2.4) and $\mathcal{C}(\mathbf{K}') = \mathcal{C}(\mathbf{X}'\mathbf{F}')$, Corollary 1 in Baksalary and Kala (1986) implies statement (ii). As a consequence, statement (iii) follows immediately from Baksalary and Kala (1986, Theorem 1). For statement (iv) we note that in view of (ii) there exists a matrix $\mathbf{A}$ such that $\mathbf{AFy}$ is BLUE of $\mathbf{M}_1\mathbf{X}_2\boldsymbol{\beta}_2$ in model $\mathcal{M}$. Hence,

$$(2.5) \qquad \mathbf{AF}(\mathbf{X}_1 : \mathbf{X}_2) = (\mathbf{0} : \mathbf{M}_1\mathbf{X}_2), \quad \mathbf{AFV}(\mathbf{X}_1 : \mathbf{X}_2)^{\perp} = \mathbf{0}.$$

From the former condition in (2.5) we have

$$(2.6) \qquad \mathbf{AFX}_2 = \mathbf{M}_1\mathbf{X}_2.$$

Moreover, the latter condition in (2.5) can be expressed as

$$(2.7) \qquad \mathcal{C}[\mathbf{FV}(\mathbf{X}_1 : \mathbf{X}_2)^\perp] \subseteq \mathcal{N}(\mathbf{A}).$$

Since (2.2) may alternatively be expressed as $\mathcal{C}(\mathbf{F}') \subseteq \mathcal{N}(\mathbf{X}_1')$, we obtain

$$\mathcal{C}[\mathbf{F}'(\mathbf{FX}_2)^\perp] = \mathcal{C}(\mathbf{F}') \cap \mathcal{N}(\mathbf{X}_2') \subseteq \mathcal{N}(\mathbf{X}_1') \cap \mathcal{N}(\mathbf{X}_2') = \mathcal{C}[(\mathbf{X}_1 : \mathbf{X}_2)^\perp].$$

Hence, (2.7) implies

$$(2.8) \qquad \mathcal{C}[\mathbf{FVF}'(\mathbf{FX}_2)^\perp] \subseteq \mathcal{N}(\mathbf{A}),$$

or, equivalently,

$$(2.9) \qquad \mathbf{AFVF}'(\mathbf{FX}_2)^\perp = \mathbf{0}.$$

Conditions (2.6) and (2.9) show that $\mathbf{AFy}$ is BLUE for $\mathbf{M}_1\mathbf{X}_2\boldsymbol{\beta}_2$ under the reduced model $\mathcal{M}_r(\mathbf{F})$, thus confirming statement (iv). $\blacksquare$

Note that a possible choice for the matrix $(\mathbf{X}_1 : \mathbf{X}_2)^\perp$ in condition (2.4) is given by

$$(2.10) \qquad (\mathbf{X}_1 : \mathbf{X}_2)^\perp = \mathbf{I} - \mathbf{P}_{(\mathbf{X}_1:\mathbf{X}_2)} = \mathbf{M}_{(\mathbf{X}_1:\mathbf{X}_2)} = \mathbf{M}_1\mathbf{Z},$$

where, for convenience,

$$(2.11) \qquad \mathbf{Z} = \mathbf{M}_{\mathbf{M}_1\mathbf{X}_2} = \mathbf{I} - \mathbf{M}_1\mathbf{X}_2(\mathbf{M}_1\mathbf{X}_2)^+.$$

3. Some specific reduced models

We have already emphasized the fact that the BLUE of $\mathbf{M}_1\mathbf{X}_2\boldsymbol{\beta}_2$ in the reduced model $\mathcal{M}_r(\mathbf{M}_1)$ is also BLUE in the partitioned model $\mathcal{M}$. Indeed, it is clear that the conditions (2.2) and (2.3) from Theorem 1 are satisfied for $\mathbf{F} = \mathbf{M}_1$. Moreover, condition (2.4) with the choice $(\mathbf{X}_1 : \mathbf{X}_2)^{\perp} = \mathbf{M}_1\mathbf{Z}$ becomes

$$(3.1) \qquad \mathcal{C}(\mathbf{M}_1\mathbf{X}_2) \cap \mathcal{C}[\mathbf{M}_1\mathbf{V}\mathbf{M}_1(\mathbf{I} - \mathbf{P}_{\mathbf{M}_1\mathbf{X}_2})] = \{\mathbf{0}\}.$$

Condition (3.1) is a special case of the general relation

$$(3.2) \qquad \mathcal{C}(\mathbf{B}) \cap \mathcal{C}(\mathbf{A}\mathbf{B}^{\perp}) = \{\mathbf{0}\},$$

here $\mathbf{A}$ is symmetric nonnegative definite and $\mathbf{B}$ is arbitrary with the same number of rows as $\mathbf{A}$, see e.g., Rao (1974, Lemma 2.1). Hence Theorem 1 covers the Frisch–Waugh–Lovell Theorem.

As a more involved procedure one may, after orthogonally projecting $\mathbf{y}$ onto $\mathcal{C}(\mathbf{X}_1)^{\perp}$, orthogonally project $\mathbf{M}_1\mathbf{y}$ onto $\mathcal{C}(\mathbf{V}_*\mathbf{Z})^{\perp}$, where $\mathbf{V}_* = \mathbf{M}_1\mathbf{V}\mathbf{M}_1$ and $\mathbf{Z} = \mathbf{M}_{\mathbf{M}_1\mathbf{X}_2}$. It is easily seen that

$$(3.3) \qquad \mathbf{M}_{\mathbf{V}_*\mathbf{Z}}\mathbf{M}_1 = \mathbf{M}_1\mathbf{M}_{\mathbf{V}_*\mathbf{Z}},$$

so that the matrix

$$(3.4) \qquad \mathbf{F} = \mathbf{M}_{\mathbf{V}_*\mathbf{Z}}\mathbf{M}_1$$

is the orthogonal projector onto $\mathcal{C}(\mathbf{X}_1)^{\perp} \cap \mathcal{C}(\mathbf{V}_*\mathbf{Z})^{\perp}$. The matrix $\mathbf{F}$ in (3.4) satisfies the conditions of Theorem 1. In view of (3.3), clearly, $\mathbf{F}\mathbf{M}_1 = \mathbf{F}$, which is equivalent to (2.2). Moreover, $\mathbf{F}\mathbf{V}\mathbf{M}_1\mathbf{Z} = \mathbf{F}\mathbf{V}_*\mathbf{Z} = \mathbf{0}$, showing that (2.4) is satisfied. To see (2.3) let ℓ be any vector in $\mathcal{N}(\mathbf{F}) \cap \mathcal{C}(\mathbf{M}_1\mathbf{X}_2)$. Then $\mathbf{M}_1\mathbf{M}_{\mathbf{V}_*\mathbf{Z}}\ell = \mathbf{0}$, where $\ell = \mathbf{M}_1\mathbf{X}_2\mathbf{u}$ for some vector $\mathbf{u}$. Combining these identities yields

$$\mathbf{M}_1\mathbf{M}_{\mathbf{V}_*\mathbf{Z}}\mathbf{M}_1\mathbf{X}_2\mathbf{u} = \mathbf{0},$$

which is equivalent to $\mathbf{M}_{\mathbf{V}_*\mathbf{Z}}\mathbf{M}_1\mathbf{X}_2\mathbf{u} = \mathbf{0}$, i.e.,

$$\mathbf{M}_1\mathbf{X}_2\mathbf{u} \in \mathcal{C}(\mathbf{V}_*\mathbf{Z}).$$

But since in view of (3.2), $\mathcal{C}(\mathbf{M}_1\mathbf{X}_2) \cap \mathcal{C}(\mathbf{V}_*\mathbf{Z}) = \{\mathbf{0}\}$, it follows $\mathbf{M}_1\mathbf{X}_2\mathbf{u} = \boldsymbol{\ell} = \mathbf{0}$, showing $\mathcal{N}(\mathbf{F}) \cap \mathcal{C}(\mathbf{M}_1\mathbf{X}_2) = \{\mathbf{0}\}$. Hence (2.3) is satisfied.

Let us now turn our attention to a different procedure. Consider the linear model

$$(3.5) \qquad\qquad \mathcal{M}_1 = \{\mathbf{y}, \mathbf{X}_1\boldsymbol{\beta}_1, \sigma^2\mathbf{V}\},$$

which is a 'reduction' of the partitioned linear model $\mathcal{M}$ by ignoring $\mathbf{X}_2\boldsymbol{\beta}_2$. If we assume model $\mathcal{M}_1$ to be consistent, then $\mathbf{y} \in \mathcal{C}(\mathbf{X}_1 : \mathbf{V})$. If in addition we wish to avoid any conflict with the original model $\mathcal{M}$, then we have to assume $\mathcal{C}(\mathbf{X}_1 : \mathbf{V}) = \mathcal{C}(\mathbf{X}_1 : \mathbf{X}_2 : \mathbf{V})$, or, equivalently,

$$(3.6) \qquad\qquad \mathcal{C}(\mathbf{X}_2) \subseteq \mathcal{C}(\mathbf{X}_1 : \mathbf{V}).$$

One representation of the BLUE of $\mathbf{X}_1\boldsymbol{\beta}_1$ in model $\mathcal{M}_1$ is given by

$$(3.7) \qquad \mathbf{X}_1(\mathbf{X}_1'\mathbf{T}_1^+\mathbf{X}_1)^+\mathbf{X}_1'\mathbf{T}_1^+\mathbf{y}, \quad \mathbf{T}_1 = \mathbf{V} + \mathbf{X}_1\mathbf{X}_1'.$$

If $\mathbf{y} \in \mathcal{C}(\mathbf{X}_1 : \mathbf{V})$, then the 'raw' residual of an arbitrary representation of the BLUE of $\mathbf{X}_1\boldsymbol{\beta}_1$ in model $\mathcal{M}_1$ can be written as

$$(3.8) \qquad [\mathbf{I} - \mathbf{X}_1(\mathbf{X}_1'\mathbf{T}_1^+\mathbf{X}_1)^+\mathbf{X}_1'\mathbf{T}_1^+]\mathbf{y} = \mathbf{V}\mathbf{M}_1(\mathbf{M}_1\mathbf{V}\mathbf{M}_1)^+\mathbf{M}_1\mathbf{y}.$$

As we will demonstrate below, the choice

$$(3.9) \qquad\qquad \mathbf{F} = \mathbf{V}\mathbf{M}_1(\mathbf{M}_1\mathbf{V}\mathbf{M}_1)^+\mathbf{M}_1$$

satisfies the conditions of Theorem 1 when (3.6) is satisfied. Note that the choice of $\mathbf{F}$ in (3.9) and the corresponding reduced model $\mathcal{M}_r(\mathbf{F})$ have also been discussed by Bhimasankaram, Shah and Saha Ray (1998) without explicitly mentioning properties (i) to (iv) of Theorem 1. It is clear that for $\mathbf{V}_* = \mathbf{M}_1\mathbf{V}\mathbf{M}_1$ we have $\mathbf{V}_*^+ = \mathbf{M}_1\mathbf{V}_*^+ = \mathbf{V}_*^+\mathbf{M}_1 = \mathbf{M}_1\mathbf{V}_*^+\mathbf{M}_1$, so that $\mathbf{F}$ from (3.9) is equal to $\mathbf{F} = \mathbf{V}\mathbf{V}_*^+$.

Now, let $\mathbf{F}$ be the matrix from (3.9). Then clearly $\mathbf{F}\mathbf{X}_1 = \mathbf{0}$, which is condition (2.2) in Theorem 1. Since $\mathcal{N}(\mathbf{F})$ is easily seen to coincide with $\mathcal{N}(\mathbf{V}\mathbf{M}_1)$, and since in view of (3.6) we have $\mathcal{C}(\mathbf{M}_1\mathbf{X}_2) \subseteq \mathcal{C}(\mathbf{M}_1\mathbf{V}) = \mathcal{N}(\mathbf{V}\mathbf{M}_1)^\perp$, it follows that $\mathcal{N}(\mathbf{F}) \cap \mathcal{C}(\mathbf{M}_1\mathbf{X}_2) = \{\mathbf{0}\}$, which is condition (2.3) in Theorem 1. It remains to show that (2.4) from Theorem 1 is satisfied. For this, let ℓ be a vector in $\mathcal{C}(\mathbf{F}\mathbf{X}_2) \cap \mathcal{C}(\mathbf{F}\mathbf{V}\mathbf{M}_1\mathbf{Z})$. Since $\mathbf{F}\mathbf{V}\mathbf{M}_1 = \mathbf{V}\mathbf{M}_1$, there exist vectors $\mathbf{u}$ and $\mathbf{v}$ such that

$$(3.10) \qquad \ell = \mathbf{V}\mathbf{M}_1(\mathbf{V}_*)^+\mathbf{M}_1\mathbf{X}_2\mathbf{u} = \mathbf{V}\mathbf{M}_1\mathbf{Z}\mathbf{v}.$$

Premultiplying (3.10) by $\mathbf{M}_1$ and using $\mathcal{C}(\mathbf{M}_1\mathbf{X}_2) \subseteq \mathcal{C}(\mathbf{V}_*)$ yields

$$(3.11) \qquad \mathbf{M}_1\ell = \mathbf{M}_1\mathbf{X}_2\mathbf{u} = \mathbf{V}_*\mathbf{Z}\mathbf{v}.$$

In view of (3.2), it follows that (3.11) can hold only if $\mathbf{M}_1\ell = \mathbf{0}$. The latter means that $\ell = \mathbf{X}_1\mathbf{w}$ for some vector $\mathbf{w}$. Hence,

$$(3.12) \qquad \ell = \mathbf{X}_1\mathbf{w} = \mathbf{V}\mathbf{M}_1\mathbf{z}, \quad \mathbf{z} = \mathbf{M}_{\mathbf{M}_1\mathbf{X}_2}\mathbf{v}.$$

Now (3.2) implies that (3.12) can hold only if $\ell = \mathbf{0}$. This confirms condition (2.4) in Theorem 1.

As a slightly different alternative to the matrix $\mathbf{F}$ from (3.9) one may choose

$$(3.13) \qquad \mathbf{G} = \mathbf{V}^{1/2}\mathbf{M}_1(\mathbf{M}_1\mathbf{V}\mathbf{M}_1)^+\mathbf{M}_1 = \mathbf{V}^{1/2}\mathbf{V}_*^+,$$

where $\mathbf{V}^{1/2}$ denotes the uniquely determined symmetric nonnegative definite square root of $\mathbf{V}$. It is of some interest to observe that $\mathbf{G}\mathbf{V}\mathbf{G}' = \mathbf{P}_{\mathbf{V}^{1/2}\mathbf{M}_1}$ and

$$(3.14) \qquad \mathbf{G}\mathbf{V}\mathbf{G}'\mathbf{G}\mathbf{X}_2 = \mathbf{P}_{\mathbf{V}^{1/2}\mathbf{M}_1}\mathbf{G}\mathbf{X}_2 = \mathbf{G}\mathbf{X}_2,$$

which means that under the model $\{\mathbf{G}\mathbf{y}, \mathbf{G}\mathbf{X}_2\boldsymbol{\beta}_2, \sigma^2\mathbf{G}\mathbf{V}\mathbf{G}'\}$ we have the equality between $\mathrm{OLSE}(\mathbf{G}\mathbf{X}_2\boldsymbol{\beta}_2)$ and $\mathrm{BLUE}(\mathbf{G}\mathbf{X}_2\boldsymbol{\beta}_2)$; cf. Puntanen and Styan (1989).

References

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

[2] P. Bhimasankaram and R. Saha Ray, *On a partitioned linear model and some associated reduced models*, Linear Algebra and Its Applications **264** (1997), 329–339.

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

[4] P. Bhimasankaram, K.R. Shah and R. Saha Ray, *On a singular partitioned linear model and some associated reduced models*, Journal of Combinatorics, Information & System Sciences **23** (1998), 415–421.

[5] R. Davidson and J.G. MacKinnon, *Estimation and Inference in Econometrics*, Oxford University Press 1993.

[6] H. Drygas, *Linear sufficiency and some applications in multilinear estimation*, Journal of Multivariate Analysis **16** (1985), 71–84.

[7] A. Feuerverger and D.A.S. Fraser, *Categorial information and the singular linear model*, Canadian Journal of Statistics **8** (1980), 41–45.

[8] R. Frisch and F.V. Waugh, *Partial time regressions as compared with individual trends*, Econometrica **1** (1933), 387–401.

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

[10] M.C. Lovell, *Seasonal adjustment of economic time series and multiple regression analysis*, Journal of the American Statistical Association **58** (1963), 993–1010.

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

[12] M. Nurhonen and S. Puntanen, *A property of partitioned generalized regression*, Communications in Statistics, Theory and Methods **21** (1992), 1579–1583.

[13] S. Puntanen, *Some matrix results related to a partitioned singular linear model*, Communications in Statistics, Theory and Methods **25** (1996), 269–279.

[14] S. Puntanen, *Some further results related to reduced singular linear models*, Communications in Statistics, Theory and Methods **26** (1996), 375–385.

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

[16] C.R. Rao, *Unified theory of linear estimation*, Sankhyā, Ser. A, 33 (1971), 371–394; Corrigenda, Sankhyā, Ser. A, 34, 194,477.

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

[18] C.R. Rao, *Projectors, generalized inverses and the BLUE's*, Journal of the Royal Statistical Society, Ser. B **36** (1974), 442–448.