NORTH-HOLLAND

Some Further Remarks on the Singular Linear Model

Simo Puntanen
Department of Mathematical Sciences
University of Tampere
P.O. Box 607
FIN-33101 Tampere, Finland

and

Alastair J. Scott
Department of Statistics
University of Auckland
Private Bag 92019
Auckland, New Zealand

Submitted by George P. H. Styan

ABSTRACT

Inspired by a recent article of Searle (1994), we consider some specific features of estimation of $\mathbf{X\beta}$ in the general linear model $\{\mathbf{y}, \mathbf{X\beta}, \mathbf{V}\}$, where the model matrix $\mathbf{X}$ need not have full column rank and the dispersion matrix $\mathbf{V}$ can be singular. Particular attention is paid to the problems related to the invariance of some matrix expressions with respect to the choice of generalized inverses. These invariance properties yield interesting matrix algebra while characterizing the best linear unbiased estimator of $\mathbf{X\beta}$.

1. INTRODUCTION

Let the triplet

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

denote the general linear model, in which $\mathbf{y}$ is an $n \times 1$ observable random vector with expectation vector $\mathbf{X\beta}$ and dispersion matrix $\mathscr{V}\!ar(\mathbf{y}) = \sigma^2\mathbf{V}$,

LINEAR ALGEBRA AND ITS APPLICATIONS 237/238:313–327 (1996)

0024-3795/96/$15.00
SSDI 0024-3795(95)00560-E

where an $n \times p$ matrix $\mathbf{X}$ and $n \times n$ nonnegative definite matrix $\mathbf{V}$ are known, while $\boldsymbol{\beta}$ is a $p \times 1$ vector of unknown parameters; the positive scalar σ^2 is also unknown, but it has no role in this paper and therefore we may set $\sigma^2 = 1$.

The matrices $\mathbf{X}$ and $\mathbf{V}$ are both allowed to be of arbitrary rank, but it is assumed, throughout the paper, that the model (1.1) is consistent [cf. Rao (1973a, p. 297)], or, in other words, that the inference base is not self-contradictory [cf. Feuerverger and Fraser (1980, p. 44)], i.e.,

$$\mathbf{y} \in \mathscr{C}(\mathbf{X}:\mathbf{V}) \qquad \text{with probability } 1, \tag{1.2}$$

where $\mathscr{C}(\mathbf{X}:\mathbf{V})$ denotes the column space of the partitioned matrix $(\mathbf{X}:\mathbf{V})$. Hence, in the sequel, whenever we have a statement concerning a linear estimator $\mathbf{Ay}$, this statement is to be understood to be valid for all $\mathbf{y} \in \mathscr{C}(\mathbf{X}:\mathbf{V})$.

When both $\mathbf{X}$ and $\mathbf{V}$ are of full rank, the best linear unbiased estimator (BLUE) of $\mathbf{X}\boldsymbol{\beta}$ is then expressible in the form [Aitken (1935, p. 45)]

$$\text{BLUE}(\mathbf{X}\boldsymbol{\beta}) = \mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y}, \tag{1.3}$$

where $\mathbf{X}'$ denotes the transpose of $\mathbf{X}$. Here $\text{BLUE}(\mathbf{X}\boldsymbol{\beta})$ is understood as an estimator $\mathbf{By}$ such that $\mathbf{BX} = \mathbf{X}$ and the difference $\mathscr{V}\!ar(\mathbf{Ay}) - \mathscr{V}\!ar(\mathbf{By})$ is nonnegative definite for every $\mathbf{A}$ satisfying $\mathbf{AX} = \mathbf{X}$. The ordinary least squares estimator (OLSE) of $\mathbf{X}\boldsymbol{\beta}$ is defined as

$$\text{OLSE}(\mathbf{X}\boldsymbol{\beta}) = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = \mathbf{P}_{\mathbf{X}}\mathbf{y} = \mathbf{Hy}; \tag{1.4}$$

here $\mathbf{P}_{\mathbf{X}} = \mathbf{H}$ is the orthogonal projector onto the column space of $\mathbf{X}$.

If matrices $\mathbf{X}$ and $\mathbf{V}$ are both deficient in rank, what changes we have to make in (1.3) and (1.4) in order to obtain the BLUE and OLSE? It is well known that in (1.4) we can replace $(\mathbf{X}'\mathbf{X})^{-1}$ by any generalized inverse $(\mathbf{X}'\mathbf{X})^{-} \in \{(\mathbf{X}'\mathbf{X})^{-}\}$, where $\{\mathbf{A}^{-}\}$ denotes the set of all generalized inverses of $\mathbf{A}$, i.e., $\{\mathbf{A}^{-}\} = \{\mathbf{G} : \mathbf{AGA} = \mathbf{A}\}$. In this case, the orthogonal projector onto the column space of $\mathbf{X}$ is

$$\mathbf{P}_{\mathbf{X}} = \mathbf{H} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'. \tag{1.5}$$

But to obtain the BLUE using generalized inverses in (1.3) is not that simple. Consider, for example, the following estimator of $\mathbf{X}\boldsymbol{\beta}$:

$$\hat{\boldsymbol{\mu}}(\mathbf{V}^{+}) = \mathbf{X}(\mathbf{X}'\mathbf{V}^{+}\mathbf{X})^{-}\mathbf{X}'\mathbf{V}^{+}\mathbf{y}, \tag{1.6}$$

where $\mathbf{V}^{+}$ denotes the (unique) Moore-Penrose inverse of $\mathbf{V}$ and $(\mathbf{X}'\mathbf{V}^{+}\mathbf{X})^{-}$ is a member of the class $\{(\mathbf{X}'\mathbf{V}^{+}\mathbf{X})^{-}\}$. One important problem in this context is whether (1.6) is invariant with respect to the choice of a generalized inverse $(\mathbf{X}'\mathbf{V}^+\mathbf{X})^-$. This is clearly not always the case. One further problem related to the estimator (1.6) is that it must be understood in such a way that in this expression we have chosen one *fixed* member from the class $\{(\mathbf{X}'\mathbf{V}^+\mathbf{X})^-\}$ and hence we have a properly defined estimator. If we choose another member from the class $\{(\mathbf{X}'\mathbf{V}^+\mathbf{X})^-\}$, then we might get another estimator.

There will be several situations where we are interested in the invariance of a product of the type $\mathbf{AB}^-\mathbf{C}$ with respect to the choice of generalized inverse $\mathbf{B}^-$. The following important result is proved by Rao and Mitra (1971, Lemma 2.2.4 and Complement 14, p. 43); see also Rao, Mitra, and Bhimasankaram (1972, Lemma 1), Mitra and Odell (1986, Lemma 1.1).

LEMMA 1.1. *Let* $\mathbf{A} \neq \mathbf{0}$ *and* $\mathbf{C} \neq \mathbf{0}$. *Then*

(a) $\mathbf{AB}^-\mathbf{C} = \mathbf{AB}^+\mathbf{C}$ *for every* $\mathbf{B}^-$ *if and only if* $\mathscr{C}(\mathbf{A}') \subset \mathscr{C}(\mathbf{B}')$ *and* $\mathscr{C}(\mathbf{C}) \subset \mathscr{C}(\mathbf{B})$, *and in particular,*

(b) $\mathbf{AA}^-\mathbf{C} = \mathbf{C}$ *for some* $\mathbf{A}^-$ *(and hence for every* $\mathbf{A}^-$*) if and only if* $\mathscr{C}(\mathbf{C}) \subset \mathscr{C}(\mathbf{A})$.

In the sequel we will also use the following rank formula [cf. Zyskind and Martin (1969, p. 1194), Marsaglia and Styan (1974, p. 276)]:

$$\operatorname{rank} \mathbf{AB} = \operatorname{rank} \mathbf{A} - \dim\left[\mathscr{C}(\mathbf{A}') \cap \mathscr{C}(\mathbf{B}^\perp)\right], \tag{1.7}$$

where $\mathbf{B}^\perp$ is any matrix such that $\mathscr{C}(\mathbf{B}^\perp)$ is the null space of the transpose of $\mathbf{B}$.

We note, in view of Lemma 1.1, that $\mathbf{H}$ in (1.5) is clearly invariant with respect to the choice of $(\mathbf{X}'\mathbf{X})^-$, and furthermore, $\hat{\boldsymbol{\mu}}(\mathbf{V}^+)$ in (1.6) is invariant with respect to the choice of a generalized inverse $(\mathbf{X}'\mathbf{V}^+\mathbf{X})^-$ if and only if

$$\mathscr{C}(\mathbf{X}') \subset \mathscr{C}(\mathbf{X}'\mathbf{V}^+\mathbf{X}). \tag{1.8}$$

Since $\mathscr{C}(\mathbf{X}'\mathbf{V}^+\mathbf{X}) = \mathscr{C}(\mathbf{X}'\mathbf{V})$ and $\mathscr{C}(\mathbf{X}'\mathbf{V}) \subset \mathscr{C}(\mathbf{X}')$, (1.8) becomes $\mathscr{C}(\mathbf{X}') = \mathscr{C}(\mathbf{X}'\mathbf{V})$, or equivalently

$$\operatorname{rank} \mathbf{X} = \operatorname{rank} \mathbf{X}'\mathbf{V}. \tag{1.9}$$

Therefore, by means of (1.7), the estimator $\hat{\boldsymbol{\mu}}(\mathbf{V}^+)$ is invariant with respect to the choice of a generalized inverse $(\mathbf{X}'\mathbf{V}^+\mathbf{X})^-$ if and only if

$$\mathscr{C}(\mathbf{X}) \cap \mathscr{C}(\mathbf{V}^\perp) = \{\mathbf{0}\}. \tag{1.10}$$

Recently Searle (1994) gave (as he states on p. 139) "some new forms and shortened proofs for results on the linear model with singular dispersion matrix." The purpose of our paper is to give some further simplifications and generalizations of Searle's results. Particular attention is paid to the problems related to the invariance of some specific matrix expression with respect to the choice of generalized inverses. Our considerations are strongly based on the use of the consistency condition (1.2) and Lemma 1.1. Their use offers an interesting approach to problems raised by Searle. In Section 3, we consider so-called natural restrictions, which are consequences of the consistency condition (1.2).

2. PROPERTIES OF THE BLUE

As mentioned in the introductory section, the estimator $\mathbf{By}$ is said to be the BLUE of $\mathbf{X}\boldsymbol{\beta}$ whenever the dispersion matrix of $\mathbf{By}$ is smallest, in the Löwner sense, among all linear unbiased estimators of $\mathbf{X}\boldsymbol{\beta}$. One characterization [cf. Drygas (1970, p. 55), Rao (1973b, p. 282)] of $\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta})$ is

$$\mathbf{By} = \mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta}) \qquad \text{under} \quad \mathscr{M} = \{\mathbf{y}, \mathbf{X}\boldsymbol{\beta}, \mathbf{V}\} \tag{2.1}$$

if and only if

$$\mathbf{B}(\mathbf{X} : \mathbf{VM}) = (\mathbf{X} : \mathbf{0}); \tag{2.2}$$

here $\mathbf{M} = \mathbf{I} - \mathbf{H}$, with $\mathbf{H}$ defined as in (1.5) and $\mathbf{I}$ denoting the $n \times n$ identity matrix. It is noteworthy that the matrix $\mathbf{B}$ satisfying (2.2) is unique if and only if the columns of $(\mathbf{X} : \mathbf{VM})$ span the whole Euclidean n-dimensional space $\mathfrak{R}^n$. On the other hand, since $\mathbf{y} \in \mathscr{C}(\mathbf{X} : \mathbf{V}) = \mathscr{C}(\mathbf{X} : \mathbf{VM})$, the numerical value of $\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta})$ is unique with probability 1, and furthermore, (1.2) can be replaced by

$$\mathbf{y} \in \mathscr{C}(\mathbf{X} : \mathbf{VM}) \qquad \text{with probability } 1. \tag{2.3}$$

One representation for $\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta})$, as noted by Searle (1994, p. 140) and Albert (1973, p. 182), is

$$\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta}) = \left[\mathbf{H} - \mathbf{HVM}(\mathbf{MVM})^{+}\mathbf{M}\right]\mathbf{y} \tag{2.4a}$$

$$= \mathrm{OLSE}(\mathbf{X}\boldsymbol{\beta}) - \mathbf{HVM}(\mathbf{MVM})^{+}\mathbf{My}. \tag{2.4b}$$

The well-known matrix identity

$$\mathbf{M}(\mathbf{MVM})^{+}\mathbf{M} = \mathbf{M}(\mathbf{MVM})^{+} = (\mathbf{MVM})^{+}\mathbf{M} = (\mathbf{MVM})^{+} \tag{2.5}$$

offers other equivalent ways to express (2.4). The result (2.5) comes at once from the identity $\mathbf{A}^{+} = (\mathbf{A}'\mathbf{A})^{+}\mathbf{A}' = \mathbf{A}'(\mathbf{AA}')^{+}$ for any matrix $\mathbf{A}$. Note further that from (2.4b) it is easy to conclude that the equality between BLUE($\mathbf{X\beta}$) and OLSE($\mathbf{X\beta}$) holds if and only if [cf. Rao (1967) and Zyskind (1967)]

$$\mathbf{HVM} = \mathbf{0}. \tag{2.6}$$

Searle (1994, p. 141) raises the question of whether

$$\left[\mathbf{H} - \mathbf{HVM}(\mathbf{MVM})^{-}\mathbf{M}\right]\mathbf{y} \tag{2.7}$$

is invariant with respect to the choice of generalized inverse $(\mathbf{MVM})^{-}$, and hence, in view of (2.4), is the BLUE of $\mathbf{X\beta}$. There are various ways to study this invariance. We use the consistency condition (2.3) and Lemma 1.1, and then the proof becomes fairly short and simple. It is crucial to note that when discussing the invariance under the choice of a generalized inverse of a representation like (2.7), invariance is to be interpreted as invariance for all $\mathbf{y}$ satisfying the consistency condition $\mathbf{y} \in \mathscr{C}(\mathbf{X}:\mathbf{V})$. Similarly, two linear estimators $\mathbf{Ay}$ and $\mathbf{By}$ are to be regarded equal if they have the same value for all $\mathbf{y} \in \mathscr{C}(\mathbf{X}:\mathbf{V})$.

THEOREM 2.1. *Under the model* $\mathscr{M} = \{\mathbf{y}, \mathbf{X\beta}, \mathbf{V}\}$, *we have*

(a) BLUE($\mathbf{X\beta}$) $= [\mathbf{H} - \mathbf{HVM}(\mathbf{MVM})^{-}\mathbf{M}]\mathbf{y}$ *for any choice of* $(\mathbf{MVM})^{-}$;
(b) BLUE($\mathbf{X\beta}$) $= [\mathbf{I} - \mathbf{VM}(\mathbf{MVM})^{-}\mathbf{M}]\mathbf{y}$ *for any choice of* $(\mathbf{MVM})^{-}$.

Proof. We first observe that (2.7) is invariant with respect to the choice of generalized inverse $(\mathbf{MVM})^{-}$ if and only if

$$\mathbf{My} \in \mathscr{C}(\mathbf{MV}). \tag{2.8}$$

By the consistency condition (2.3) we know that with probability 1, $\mathbf{y}$ is expressible as

$$\mathbf{y} = \mathbf{Xa} + \mathbf{VMb} \qquad \text{for some vectors } \mathbf{a} \text{ and } \mathbf{b}. \tag{2.9}$$

From (2.9) the requirement (2.8) follows at once, and (a) is proved. Replacing $\mathbf{H}$ by $\mathbf{I} - \mathbf{M}$ in (2.7) yields the following representation for the BLUE:

$$\text{BLUE}(\mathbf{X\beta}) = \left[\mathbf{I} - \mathbf{VM}(\mathbf{MVM})^{-}\mathbf{M}\right]\mathbf{y} - \mathbf{Cy}, \tag{2.10}$$

where $\mathbf{C} = \mathbf{M} - \mathbf{MVM}(\mathbf{MVM})^{-}\mathbf{M}$. Using again (2.9), we see that $\mathbf{Cy} = \mathbf{0}$ and hence one representation for the BLUE is

$$\text{BLUE}(\mathbf{X\beta}) = \left[\mathbf{I} - \mathbf{VM}(\mathbf{MVM})^{-}\mathbf{M}\right]\mathbf{y}, \tag{2.11}$$

which clearly is invariant for any choice of $(\mathbf{MVM})^{-}$. ■

From Theorem 2.1 we obtain the following formulas for the dispersion matrix of BLUE($\mathbf{X\beta}$):

$$\mathscr{V}\!ar[\text{BLUE}(\mathbf{X\beta})] = \mathbf{HVH} - \mathbf{HVM}(\mathbf{MVM})^{-}\mathbf{MVH} \tag{2.12a}$$

$$= \mathbf{V} - \mathbf{VM}(\mathbf{MVM})^{-}\mathbf{MV}, \tag{2.12b}$$

where both expressions are invariant with respect to the choice of $(\mathbf{MVM})^{-}$. If the rank of $\mathbf{V}$ is v, then we may write the spectral decomposition

$$\mathbf{V} = \mathbf{Q\Lambda Q}', \tag{2.13}$$

where $\mathbf{Q}$ is an $n \times v$ matrix such that $\mathbf{Q}'\mathbf{Q} = \mathbf{I}_v$ and $\mathbf{\Lambda}$ is a $v \times v$ diagonal matrix with the nonzero eigenvalues of $\mathbf{V}$ on its diagonal. Let us further denote the nonnegative definite symmetric square root of $\mathbf{V}$ as

$$\mathbf{V}^{1/2} = \mathbf{Q\Lambda}^{1/2}\mathbf{Q}' = \mathbf{T}, \qquad \text{say.} \tag{2.14}$$

Then (2.12b) can be written as

$$\mathscr{V}\!ar[\text{BLUE}(\mathbf{X\beta})] = \mathbf{T}(\mathbf{I} - \mathbf{P}_{\mathbf{TM}})\mathbf{T}. \tag{2.15}$$

Using (1.7), we obtain

$$\begin{aligned}\operatorname{rank}\{\mathscr{V}\!ar[\text{BLUE}(\mathbf{X\beta})]\} &= \operatorname{rank}[\mathbf{T}(\mathbf{I} - \mathbf{P}_{\mathbf{TM}})] \\ &= \operatorname{rank}\mathbf{T} - \dim[\mathscr{C}(\mathbf{T}) \cap \mathscr{C}(\mathbf{TM})] \\ &= \operatorname{rank}\mathbf{T} - \operatorname{rank}\mathbf{TM} \\ &= \dim[\mathscr{C}(\mathbf{V}) \cap \mathscr{C}(\mathbf{X})]. \end{aligned} \tag{2.16}$$

Hence it is clear that $\mathscr{V}\!ar[\text{BLUE}(\mathbf{X\beta})] = \mathbf{0}$ if and only if $\mathscr{C}(\mathbf{V}) \cap \mathscr{C}(\mathbf{X}) = \{\mathbf{0}\}$. In the following theorem we collect together some related results.

THEOREM 2.2. *Under the model* $\mathscr{M} = \{\mathbf{y}, \mathbf{X\beta}, \mathbf{V}\}$, *the following statements are equivalent*:

(a) $\mathbf{M}(\mathbf{MVM})^{-}\mathbf{M} \in \{\mathbf{V}^{-}\}$ *for any choice of* $(\mathbf{MVM})^{-}$;
(b) $\mathscr{C}(\mathbf{X}) \cap \mathscr{C}(\mathbf{V}) = \{\mathbf{0}\}$;
(c) $\mathscr{V}\!ar[\text{BLUE}(\mathbf{X\beta})] = \mathbf{0}$.

Furthermore,

(d) $\mathbf{M}(\mathbf{MVM})^{-}\mathbf{M} \in \{(\mathbf{MVM})^{-}\}$ *for any choice of* $(\mathbf{MVM})^{-}$;
(e) $\mathbf{M}(\mathbf{MVM})^{-}\mathbf{M} = (\mathbf{MVM})^{+}$ *for any choice of* $(\mathbf{MVM})^{-}$ *if and only if*

$$\mathscr{C}(\mathbf{X} : \mathbf{V}) = \mathfrak{R}^n; \tag{2.17}$$

(f) $[\mathbf{I} - \mathbf{V}(\mathbf{MVM})^{-}\mathbf{M}]\mathbf{y}$ *is invariant for any choice of* $(\mathbf{MVM})^{-}$ *if and only if*

$$\mathscr{C}(\mathbf{X}) \subset \mathscr{C}(\mathbf{V}^{\perp}); \tag{2.18}$$

(g) $[\mathbf{I} - \mathbf{V}(\mathbf{MVM})^{-}]\mathbf{y}$ *is invariant for any choice of* $(\mathbf{MVM})^{-}$ *if and only if* $\mathbf{X} = \mathbf{0}$.

Proof. The equivalence of (a), (b), and (c) follows from (2.12b) and (2.16); see also Werner (1987). Part (d) is easy to check. The matrix $\mathbf{M}(\mathbf{MVM})^{-}\mathbf{M}$ is invariant with respect to the choice of $(\mathbf{MVM})^{-}$ if and only if

$$\mathscr{C}(\mathbf{M}) \subset \mathscr{C}(\mathbf{MV}). \tag{2.19}$$

Now (2.19) is equivalent to $\operatorname{rank}\mathbf{M} = \operatorname{rank}\mathbf{MV}$, and hence, by (1.7), to

$$\mathscr{C}(\mathbf{M}) \cap \mathscr{C}(\mathbf{V}^{\perp}) = \{\mathbf{0}\}, \tag{2.20}$$

which further is equivalent to (2.17), in which case $\mathbf{M}(\mathbf{MVM})^{-}\mathbf{M} = (\mathbf{MVM})^{+}$ for any choice of $(\mathbf{MVM})^{-}$, and (e) is proved. To prove (f), we note that

$$\left[\mathbf{I} - \mathbf{V}(\mathbf{MVM})^{-}\mathbf{M}\right]\mathbf{y} \tag{2.21}$$

is invariant with respect to the choice of $(\mathbf{MVM})^{-}$ if and only if

$$\mathscr{C}(\mathbf{V}) \subset \mathscr{C}(\mathbf{MV}). \tag{2.22}$$

It is easy to show that (2.22) is equivalent to $\mathbf{V} = \mathbf{MV} = \mathbf{VM} = \mathbf{MVM}$, i.e., $\mathbf{HV} = \mathbf{0}$, which can also be expressed as (2.18). To prove (g), we note that the necessary and sufficient conditions for the invariance of $[\mathbf{I} - \mathbf{V}(\mathbf{MVM})^{-}]\mathbf{y}$ are the following:

$$\mathbf{HV} = \mathbf{0} \quad \text{and} \quad \mathbf{y} \in \mathscr{C}(\mathbf{MV}). \tag{2.23}$$

It is easy to see that (2.23) holds if and only if $\mathbf{X} = \mathbf{0}$, which is of no interest from a statistical point of view. ■

Let $(\mathbf{MVM})^{\sim}$ be a generalized inverse of $\mathbf{MVM}$. Then the matrix $\mathbf{M}(\mathbf{MVM})^{\sim}\mathbf{M} = (\mathbf{MVM})^{=}$, say, is a generalized inverse of $\mathbf{MVM}$, and

$$\left[\mathbf{I} - \mathbf{V}(\mathbf{MVM})^{=}\right]\mathbf{y} = \text{BLUE}(\mathbf{X}\boldsymbol{\beta}). \tag{2.24}$$

But as Searle (1994, p. 142) points out, the matrix $(\mathbf{MVM})^{=}$ in (2.24) cannot be replaced by an arbitrary generalized inverse of $\mathbf{MVM}$. Part (g) of Theorem 2.2 shows that such a replacement is possible if and only if $\mathbf{X} = \mathbf{0}$.

Note that if (2.18) holds, then $[\mathbf{I} - \mathbf{V}(\mathbf{MVM})^{-}\mathbf{M}]\mathbf{y} = \mathbf{Hy} = \text{BLUE}(\mathbf{X}\boldsymbol{\beta})$. We further note that it is easy to see that the vector $\mathbf{y}$ itself is the BLUE of $\mathbf{X}\boldsymbol{\beta}$ if and only if $\mathbf{VM} = \mathbf{0}$, or equivalently

$$\mathscr{C}(\mathbf{M}) \subset \mathscr{C}(\mathbf{V}^{\perp}). \tag{2.25}$$

The inclusion (2.25) is actually a necessary and sufficient condition for every unbiased estimator of $\mathbf{X}\boldsymbol{\beta}$ to be the BLUE.

Let us next consider the estimator of the type

$$\hat{\boldsymbol{\mu}}(\mathbf{W}) = \mathbf{X}(\mathbf{X}'\mathbf{W}\mathbf{X})^{-}\mathbf{X}'\mathbf{W}\mathbf{y}, \tag{2.26}$$

where $\mathbf{W}$ is now an arbitrary matrix such that $\mathbf{X}'\mathbf{W}\mathbf{X} \neq \mathbf{0}$. To satisfy the unbiasedness of $\hat{\boldsymbol{\mu}}(\mathbf{W})$, we must have

$$\mathbf{X}(\mathbf{X}'\mathbf{W}\mathbf{X})^{-}\mathbf{X}'\mathbf{W}\mathbf{X} = \mathbf{X}. \tag{2.27}$$

By Lemma 1.1, (2.27) holds if and only

$$\mathscr{C}(\mathbf{X}') \subset \mathscr{C}(\mathbf{X}'\mathbf{W}'\mathbf{X}). \tag{2.28}$$

Clearly (2.28) is equivalent to $\mathscr{C}(\mathbf{X}') = \mathscr{C}(\mathbf{X}'\mathbf{W}'\mathbf{X})$ and similarly to $\mathscr{C}(\mathbf{X}') = \mathscr{C}(\mathbf{X}'\mathbf{W}\mathbf{X})$. How about the invariance of $\hat{\boldsymbol{\mu}}(\mathbf{W})$ with respect to the choice of $(\mathbf{X}'\mathbf{W}\mathbf{X})^{-}$? Lemma 1.1 shows that the invariance holds if and only if both (2.28) and

$$\mathbf{X}'\mathbf{W}\mathbf{y} \in \mathscr{C}(\mathbf{X}'\mathbf{W}\mathbf{X}) \tag{2.29}$$

hold. Since (2.28) implies (2.29), we have proved the following theorem (cf. Baksalary and Puntanen, 1989, Lemma 1):

THEOREM 2.3. *Consider the model* $\mathscr{M} = \{\mathbf{y}, \mathbf{X}\boldsymbol{\beta}, \mathbf{V}\}$, *and assume that* $\mathbf{X}'\mathbf{W}\mathbf{X} \neq \mathbf{0}$. *Then the following three statements are equivalent*:

(a) $\hat{\boldsymbol{\mu}}(\mathbf{W}) = \mathbf{X}(\mathbf{X}'\mathbf{W}\mathbf{X})^{-}\mathbf{X}'\mathbf{W}\mathbf{y}$ *is an unbiased estimator for* $\mathbf{X}\boldsymbol{\beta}$;
(b) $\hat{\boldsymbol{\mu}}(\mathbf{W}) = \mathbf{X}(\mathbf{X}'\mathbf{W}\mathbf{X})^{-}\mathbf{X}'\mathbf{W}\mathbf{y}$ *is invariant for any choice of* $(\mathbf{X}'\mathbf{W}\mathbf{X})^{-}$;
(c) $\operatorname{rank}\mathbf{X} = \operatorname{rank}(\mathbf{X}'\mathbf{W}\mathbf{X})$.

Searle (1994, Theorem 1) provided a rather lengthy proof of Theorem 2.3 in the situation where $\mathbf{W}$ is nonnegative definite and symmetric.

The statistical analysis under the model $\mathscr{M} = \{\mathbf{y}, \mathbf{X}\boldsymbol{\beta}, \mathbf{V}\}$ simplifies when

$$\mathscr{C}(\mathbf{X}) \subset \mathscr{C}(\mathbf{V}) \tag{2.30}$$

holds. This was first noted by Goldman and Zelen (1964, p. 165), who asserted that the BLUE of $\mathbf{X}\boldsymbol{\beta}$ is then expressible in the form

$$\hat{\boldsymbol{\mu}}(\mathbf{V}^{+}) = \mathbf{X}(\mathbf{X}'\mathbf{V}^{+}\mathbf{X})^{-}\mathbf{X}'\mathbf{V}^{+}\mathbf{y}. \tag{2.31}$$

In view of Theorem 2.3, $\hat{\boldsymbol{\mu}}(\mathbf{V}^{+})$ is unbiased if and only if $\operatorname{rank}\mathbf{X} = \operatorname{rank}\mathbf{X}'\mathbf{V}$ or equivalently $\mathscr{C}(\mathbf{X}) \cap \mathscr{C}(\mathbf{V}^{\perp}) = \{\mathbf{0}\}$. This indeed holds if (2.30) holds. In order for $\hat{\boldsymbol{\mu}}(\mathbf{V}^{+})$ to be the BLUE, the following equality must hold [cf. (2.2)]:

$$\mathbf{X}(\mathbf{X}'\mathbf{V}^{+}\mathbf{X})^{-}\mathbf{X}'\mathbf{V}^{+}\mathbf{V}\mathbf{M} = \mathbf{0}. \tag{2.32}$$

Clearly (2.30) implies (2.32).

The observation of Goldman and Zelen (1964) was generalized by Mitra and Rao (1968, p. 286), who pointed out that representation (2.31) is true under (2.30) not only with $\mathbf{V}^{+}$, but with any generalized inverse $\mathbf{V}^{-}$. A complete solution to the problem of the validity of (2.31) was given by Zyskind and Martin (1969, p. 1196), who showed that (2.30) is not only sufficient, but also necessary. Actually, their Corollary 1.1 may be generalized as follows.

THEOREM 2.4. *For the general linear model* $\mathscr{M} = \{\mathbf{y}, \mathbf{X}\boldsymbol{\beta}, \mathbf{V}\}$, *consider the following two statements*:

(a) $\text{BLUE}(\mathbf{X}\boldsymbol{\beta}) = \mathbf{X}(\mathbf{X}'\mathbf{V}^{-}\mathbf{X})^{-}\mathbf{X}'\mathbf{V}^{-}\mathbf{y}$,
(b) $\mathscr{C}(\mathbf{X}) \subset \mathscr{C}(\mathbf{V})$.

Then (b) *implies* (a) *irrespective of the choices of* $\mathbf{V}^{-}$ *and* $(\mathbf{X}'\mathbf{V}^{-}\mathbf{X})^{-}$, *and conversely*, (b) *must hold whenever* (*a*) *is satisfied for some* $\mathbf{V}^{-}$ *such that* $\operatorname{rank}(\mathbf{X}'\mathbf{V}^{-}\mathbf{X}) = \operatorname{rank}(\mathbf{X}'\mathbf{V}^{-}\mathbf{V})$ *and some* $(\mathbf{X}'\mathbf{V}^{-}\mathbf{X})^{-}$.

Proof. Condition (b) guarantees the invariance of $\hat{\boldsymbol{\mu}}(\mathbf{V}^{-}) = \mathbf{X}(\mathbf{X}'\mathbf{V}^{-}\mathbf{X})^{-}\mathbf{X}'\mathbf{V}^{-}\mathbf{y}$ with respect to the choice of $\mathbf{V}^{-}$ and $(\mathbf{X}'\mathbf{V}^{-}\mathbf{X})^{-}$, and hence $\hat{\boldsymbol{\mu}}(\mathbf{V}^{-})$ equals $\hat{\boldsymbol{\mu}}(\mathbf{V}^{+})$, which is $\text{BLUE}(\mathbf{X}\boldsymbol{\beta})$ in this situation.

To prove the latter part, let $\mathbf{V}^{\sim}$ be a fixed generalized inverse of $\mathbf{V}$, such that

$$\operatorname{rank}(\mathbf{X}'\mathbf{V}^{\sim}\mathbf{X}) = \operatorname{rank}(\mathbf{X}'\mathbf{V}^{\sim}\mathbf{V}), \tag{2.33}$$

and let $\hat{\boldsymbol{\mu}}(\mathbf{V}^{\sim}) = \mathbf{X}(\mathbf{X}'\mathbf{V}^{\sim}\mathbf{X})^{-}\mathbf{X}'\mathbf{V}^{\sim}\mathbf{y}$ be the BLUE of $\mathbf{X}\boldsymbol{\beta}$. Then the unbiasedness of $\hat{\boldsymbol{\mu}}(\mathbf{V}^{\sim})$ means that

$$\operatorname{rank}(\mathbf{X}'\mathbf{V}^{\sim}\mathbf{X}) = \operatorname{rank}\mathbf{X}, \tag{2.34}$$

and hence (2.33) implies that

$$\operatorname{rank}(\mathbf{X}'\mathbf{V}^{\sim}\mathbf{V}) = \operatorname{rank}\mathbf{X}. \tag{2.35}$$

Furthermore, since $\hat{\boldsymbol{\mu}}(\mathbf{V}^{\sim})$ is the BLUE of $\mathbf{X}\boldsymbol{\beta}$, we must have

$$\mathbf{X}(\mathbf{X}'\mathbf{V}^{\sim}\mathbf{X})^{-}\mathbf{X}'\mathbf{V}^{\sim}\mathbf{V}\mathbf{M} = \mathbf{0}. \tag{2.36}$$

Premultiplying (2.36) by $\mathbf{X}'\mathbf{V}^{\sim}$ and using $\mathbf{X}'\mathbf{V}^{\sim}\mathbf{X}(\mathbf{X}'\mathbf{V}^{\sim}\mathbf{X})^{-}\mathbf{X}' = \mathbf{X}'$, we see that (2.36) is equivalent to

$$\mathbf{X}'\mathbf{V}^{\sim}\mathbf{V}\mathbf{M} = \mathbf{0}, \tag{2.37}$$

i.e.,

$$\mathscr{C}(\mathbf{V}\mathbf{V}^{\sim\prime}\mathbf{X}) \subset \mathscr{C}(\mathbf{X}). \tag{2.38}$$

Combining (2.35) with (2.38) yields

$$\mathscr{C}(\mathbf{V}\mathbf{V}^{\sim\prime}\mathbf{X}) = \mathscr{C}(\mathbf{X}), \tag{2.39}$$

which clearly implies condition (b) of the theorem. ■

As an example, let us consider the model $\{\mathbf{y}, (\mathbf{1} : \mathbf{X}_2), \mathbf{C}\}$, where $\mathbf{1} = (1, 1, \ldots, 1)'$ and $\mathbf{C} = \mathbf{I} - n^{-1}\mathbf{1}\mathbf{1}'$, that is, $\mathbf{C}$ is the centering matrix. In this situation, it is easy to observe that the ordinary least squares estimator of $\mathbf{X}\boldsymbol{\beta}$ equals BLUE($\mathbf{X}\boldsymbol{\beta}$), and hence (a) of Theorem 2.4 holds when we take $\mathbf{I}$ as a generalized inverse of $\mathbf{V}$. But now neither (b) nor the condition $\operatorname{rank}(\mathbf{X}'\mathbf{V}^{-}\mathbf{X}) = \operatorname{rank}(\mathbf{X}'\mathbf{V}^{-}\mathbf{V})$ holds. This explains why the rather artificial-seeming rank condition needs to be included with (a). We can extend this to the situation where the model matrix is $\mathbf{X} = (\mathbf{X}_1 : \mathbf{X}_2)$ and $\mathscr{V}\!ar(\mathbf{y}) = \mathbf{I} - \mathbf{X}_1(\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'$. This arises when $\mathbf{y}$ is the residual from fitting $\mathbf{X}_1$. If we ignore it and fit the full model, we still get the BLUE.

We note that choosing $\mathbf{V}^{\sim} = \mathbf{V}^{+}$ means that (2.33) holds, and so Theorem 2.4 could be formulated by replacing $\mathbf{V}^{-}$ by $\mathbf{V}^{+}$.

Searle (1994, Theorem 2) gives such a version of the above theorem, where he chooses $\mathbf{V}^{\sim} = \mathbf{V}_{\mathrm{rs}}^{-}$, a symmetric reflexive generalized inverse of $\mathbf{V}$. It can be shown that

$$\operatorname{rank}(\mathbf{X}'\mathbf{V}\mathbf{V}_{\mathrm{rs}}^{-}) = \operatorname{rank}(\mathbf{X}'\mathbf{V}_{\mathrm{rs}}^{-}\mathbf{X}) = \operatorname{rank}\mathbf{X}'\mathbf{V}, \tag{2.40}$$

and hence Searle's result comes as a corollary to our Theorem 2.4.

The condition (2.30) plays a similar role when considering a representation of the dispersion matrix of BLUE($\mathbf{X\beta}$). Mitra and Rao (1968, p. 286) pointed out that if (2.30) holds, then $\mathscr{V}\!ar$[BLUE($\mathbf{X\beta}$)] may be expressed in the same form as in the case of positive definite $\mathbf{V}$, with the only difference that $\mathbf{V}^{-1}$ is replaced by a generalized inverse $\mathbf{V}^{-}$. And again Zyskind and Martin (1969, p. 1196) showed that (2.30) is not only sufficient, but also necessary. We reformulate here (omitting the proof) their Corollary 1.3 as follows.

THEOREM 2.5. *For the general linear model* $\mathscr{M} = \{\mathbf{y}, \mathbf{X\beta}, \mathbf{V}\}$, *consider the following two statements*:

(a) $\mathscr{V}\!ar$[BLUE($\mathbf{X\beta}$)] $= \mathbf{X}(\mathbf{X}'\mathbf{V}^{-}\mathbf{X})^{-}\mathbf{X}'$,
(b) $\mathscr{C}(\mathbf{X}) \subset \mathscr{C}(\mathbf{V})$.

Then (b) *implies* (a) *irrespective of the choices of* $\mathbf{V}^{-}$ *and* $(\mathbf{X}'\mathbf{V}^{-}\mathbf{X})^{-}$, *and conversely*, (b) *must hold whenever* (a) *is satisfied for some* $\mathbf{V}^{-}$ *such that* $\operatorname{rank}(\mathbf{X}'\mathbf{V}^{-}\mathbf{X}) = \operatorname{rank}\mathbf{X}$ *and some* $(\mathbf{X}'\mathbf{V}^{-}\mathbf{X})^{-}$.

We will conclude this section with some considerations related to representations of the dispersion matrix of BLUE($\mathbf{X\beta}$). Specifically, it might be natural to ask for a "direct" proof of the equality between the expressions of (2.12) and of the type (a) of Theorem 2.5. We offer here a simple matrix-algebraic proof. Consider the decomposition $\mathbf{V} = \mathbf{Q\Lambda Q}'$ as in (2.13), and let us denote $\mathbf{V}^{1/2} = \mathbf{Q\Lambda}^{1/2}\mathbf{Q}'$ and $(\mathbf{V}^{+})^{1/2} = \mathbf{Q\Lambda}^{-1/2}\mathbf{Q}'$. Then $\mathbf{V}^{1/2}(\mathbf{V}^{+})^{1/2} = \mathbf{P}_{\mathbf{V}}$. It is well known that

$$\mathbf{P}_{\mathbf{A}} + \mathbf{P}_{\mathbf{B}} = \mathbf{P}_{(\mathbf{A}\,:\,\mathbf{B})} \quad \Leftrightarrow \quad \mathbf{A}'\mathbf{B} = \mathbf{0}. \tag{2.41}$$

Substituting $\mathbf{A} = (\mathbf{V}^{+})^{1/2}\mathbf{X}$ and $\mathbf{B} = \mathbf{V}^{1/2}\mathbf{M}$ into (2.41), we observe that the equality

$$\mathbf{V}^{1/2}\mathbf{M}(\mathbf{MVM})^{-}\mathbf{MV}^{1/2} + (\mathbf{V}^{+})^{1/2}\mathbf{X}(\mathbf{X}'\mathbf{V}^{+}\mathbf{X})^{-}\mathbf{X}'(\mathbf{V}^{+})^{1/2} = \mathbf{P}_{\mathbf{V}} \tag{2.42}$$

holds if and only if $\mathbf{X}'\mathbf{P}_{\mathbf{V}}\mathbf{M} = \mathbf{0}$. Pre- and postmultiplying (2.42) by $\mathbf{V}^{1/2}$, we conclude that (2.42) is equivalent to

$$\mathbf{VM}(\mathbf{MVM})^{-}\mathbf{MV} + \mathbf{P}_{\mathbf{V}}\mathbf{X}(\mathbf{X}'\mathbf{V}^{+}\mathbf{X})^{-}\mathbf{X}'\mathbf{P}_{\mathbf{V}} = \mathbf{V}. \tag{2.43}$$

If $\mathscr{C}(\mathbf{X}) \subset \mathscr{C}(\mathbf{V})$, then (2.43) holds, and we can delete $\mathbf{P}_\mathbf{V}$ from (2.43) and obtain the presentation (a) of Theorem 2.5. As a matter of fact, we have proved the following theorem:

THEOREM 2.6. *Under the model* $\mathscr{M} = \{\mathbf{y}, \mathbf{X}\boldsymbol{\beta}, \mathbf{V}\}$, *the following statements are equivalent*:

(a) $\mathbf{H}\mathbf{P}_\mathbf{V}\mathbf{M} = \mathbf{0}$;
(b) $\mathscr{V}\!ar[\text{BLUE}(\mathbf{X}\boldsymbol{\beta})] = \mathbf{P}_\mathbf{V}\mathbf{X}(\mathbf{X}'\mathbf{V}^+\mathbf{X})^-\mathbf{X}'\mathbf{P}_\mathbf{V}$.

For further discussion of the dispersion matrix of BLUE($\mathbf{X}\boldsymbol{\beta}$), the reader is referred to Baksalary, Puntanen, and Styan (1990).

3. NATURAL RESTRICTIONS

In their main result, Zyskind and Martin (1969) show that whatever the matrices $\mathbf{X}$ and $\mathbf{V}$ and relationships between them are, it is always possible to represent BLUE($\mathbf{X}\boldsymbol{\beta}$) and $\mathscr{V}\!ar[\text{BLUE}(\mathbf{X}\boldsymbol{\beta})]$ in the forms $\mathbf{X}(\mathbf{X}'\mathbf{V}^{\sim}\mathbf{X})^-\mathbf{X}'\mathbf{V}^{\sim}\mathbf{y}$ and $\mathbf{X}(\mathbf{X}'\mathbf{V}^{\sim}\mathbf{X})^-\mathbf{X}'$, respectively, with the use of a suitably chosen generalized inverse $\mathbf{V}^{\sim}$. From this point of view, therefore, the only difference between the models $\mathscr{M}$ which satisfy (2.30) and those which do not satisfy (2.30) is that the choice of $\mathbf{V}^{\sim}$ is irrelevant in the first situation.

The role of the condition (2.30) in statistical analysis under the model $\mathscr{M}$ seems to be more important from a different point of view. It has been pointed out by several authors that if $\mathbf{V}$ is singular, then $\boldsymbol{\beta}$ is subject to certain restrictions implied by the model structure, which will henceforth be called *natural restrictions*. They are expressed in the literature in various ways. We quote here only the form

$$\mathbf{F}'\mathbf{X}\boldsymbol{\beta} = \mathbf{F}'\mathbf{y}, \tag{3.1}$$

where $\mathbf{F}$ is any matrix such that $\mathscr{C}(\mathbf{F})$ coincides with the orthocomplement of $\mathscr{C}(\mathbf{V})$. The formulation (3.1) is used for instance by Kempthorne (1976, p. 207), Mitra and Rao (1968, p. 282), Rao (1973b, p. 279), and Seely and Zyskind (1971, p. 693). For two alternative formulations of the natural restrictions, we refer the reader to Alalouf (1978, p. 68) and Feuerverger and Fraser (1980, p. 43).

The interpretation of the equations (3.1) is somewhat murky, as they become completely specified restrictions on $\boldsymbol{\beta}$ only when $\mathbf{y}$ is replaced by its observed outcome. A consequence of the above is that there is no unique way of treating the equations (3.1) in statistical analysis under the model $\mathscr{M}$. In the problem of linear unbiased estimation, some papers take (3.1) into account, which leads in particular to defining $\mathbf{Ay}$ as an unbiased estimator of $\mathbf{X\beta}$ if $\mathbf{AX\beta} = \mathbf{X\beta}$ holds for every $\boldsymbol{\beta}$ satisfying (3.1); whereas some other papers ignore (3.1) and consider $\boldsymbol{\beta}$ to be free to vary over the entire space of $p \times 1$ vectors, in which case the unbiasedness condition retains its classical stronger form $\mathbf{AX} = \mathbf{X}$. The former approach further leads to the so-called "wider definition of BLUE($\mathbf{X\beta}$)"; cf., e.g., Harville (1981), Rao (1973b; 1979), and Schönfeld and Werner (1987). On the other hand, the crucial argument for the second approach is that the simplification involved in it does not lead to any loss of generality, because if there is an unbiased estimator $\mathbf{A}_*\mathbf{y}$ of $\mathbf{X\beta}$ not satisfying the equation $\mathbf{A}_*\mathbf{X} = \mathbf{X}$, then there exists $\mathbf{A}$ such that $\mathbf{AX} = \mathbf{X}$ and $\mathbf{Ay} = \mathbf{A}_*\mathbf{y}$ with probability one; cf., e.g., Kempthorne (1976) and Rao (1973a, pp. 297–298). In this context, the reader is also referred to Puntanen and Styan (1989, 1990), Christensen (1990), Harville (1990), and Baksalary, Rao, and Markiewicz (1992).

In view of the above, the most important role of the condition (2.30) is that it reconciles the two approaches mentioned above, as the equations (3.1) are trivially fulfilled for every model $\mathscr{M}$ which satisfies (2.30) in addition to the consistency condition $\mathbf{y} \in \mathscr{C}(\mathbf{X} : \mathbf{V})$. When the natural restrictions (3.1) vanish, i.e., when there is no deterministic component providing information on $\mathbf{X\beta}$, then the statistical analysis under the model $\mathscr{M}$ becomes so similar to that under the model with a positive definite $\mathbf{V}$ that it is justifiable to emphasize the fact that the condition (2.30) holds by calling the model "weakly singular." According to our knowledge, this term was proposed by Nordström (1985, p. 243) as a result of his considerations concerning a decomposition of $\mathscr{M}$; cf. Rao (1984, Section 6).

The authors wish to thank Jerzy K. Baksalary for very helpful discussions. This research was completed when the first author was a Senior Research Fellow of the Academy of Finland.

REFERENCES

Aitken, A. C. 1935. On least squares and linear combination of observations, *Proc. Roy. Soc. Edinburgh Sect. A* 55:42–49.

Alalouf, I. S. 1978. An explicit treatment of the general linear model with singular covariance matrix, *Sankhyā Ser. B* 40:65–73.

Albert, A. 1973. The Gauss-Markov theorem for regression models with possibly singular covariances, *SIAM J. Appl. Math.* 24:182–187.

Baksalary, J. K. and Puntanen, S. 1989. Weighted-least-squares estimation in the general Gauss-Markov model, in *Statistical Data Analysis and Inference* (Y. Dodge, Ed.), North-Holland, Amsterdam, pp. 355–368.

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

Baksalary, J. K., Rao, C. R., and Markiewicz, A. 1992. A study of the influence of the "natural restrictions" on estimation problems in the singular Gauss-Markov model, *J. Statist. Plann. Inference* 31:335–351.

Christensen, R. 1990. Comment on Puntanen and Styan (1989) (letter to the editor), *Amer. Statist.* 44:191–192.

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

Feuerverger, A. and Fraser, D. A. S. 1980. Categorical information and the singular linear model, *Canad. J. Statist.* 8:41–45.

Goldman, A. J. and Zelen, M. 1964. Weak generalized inverses and minimum variance linear unbiased estimation, *J. Res. Nat. Bur. Standards Sect. B* 68:151–172.

Harville, D. A. 1981. Unbiased and minimum-variance unbiased estimation of estimable functions for fixed linear models with arbitrary covariance structure, *Ann. Statist.* 9:633–637.

Harville, D. A. 1990. Comment on Puntanen and Styan (1989) (letter to the editor), *Amer. Statist.* 44:192.

Kempthorne, O. 1976. Best linear unbiased estimation with arbitrary variance matrix, in *Essays in Probability and Statistics: A Volume in Honor of Professor Junjiro Ogawa*, Shinko Tsusho, Tokyo, pp. 203–225.

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

Mitra, S. K. and Odell, P. L. 1986. On parallel summability of matrices, *Linear Algebra Appl.* 74:239–255.

Mitra, S. K. and Rao, C. R. 1968. Some results in estimation and tests of linear hypotheses under the Gauss-Markoff model, *Sankhyā Ser. A* 30:281–290.

Nordström, K. 1985. On a decomposition of the singular Gauss-Markov model, in *Linear Statistical Inference* (T. Caliński and W. Klonecki, Eds.), Springer-Verlag, Berlin, pp. 231–245.

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

Puntanen, S. and Styan, G. P. H. 1990. Reply to Christensen (1990), and Harville (1990) (letter to the editor), *Amer. Statist.* 44:192–193.

Rao, C. R. 1967. Least squares theory using an estimated dispersion matrix and its application to measurement of signals, in *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 1 (L. M. Le Cam and J. Neyman, Eds.), Univ. of California Press, Berkeley, pp. 355–372.

Rao, C. R. 1973a. *Linear Statistical Inference and Its Applications*, 2nd ed., Wiley, New York.

Rao, C. R. (1973b). 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. 1979. Estimation of parameters in the singular Gauss-Markoff model, *Comm. Statist. A—Theory Methods* 8:1353–1358.

Rao, C. R. 1984. Inference from linear models with fixed effects: Recent results and some problems, in *Statistics: An Appraisal* (H. A. David and H. T. David, Eds.), Iowa State U.P., Ames, pp. 345–369.

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

Rao, C. R., Mitra, S. K., and Bhimasankaram, P. 1972. Determination of a matrix by its subclasses of generalized inverses, *Sankhyā Ser. A* 34:5–8.

Schönfeld, P. and Werner, H. J. 1987. A note on C. R. Rao's wider definition of the BLUE in the general Gauss-Markov model, *Sankhyā Ser. B* 49:1–8.

Searle, S. R. 1994. Extending some results and proofs for the singular linear model, *Linear Algebra Appl.* 210:139–151.

Seely, J. and Zyskind, G. 1971. Linear spaces and minimum variance unbiased estimation, *Ann. Math. Statist.* 42:691–703.

Werner, H. J. 1987. C. R. Rao's IPM method: a geometric approach, in *New Perspectives in Theoretical and Applied Statistics* (M. L. Puri, J. P. Vilaplana, and W. Wertz, Eds.), Wiley, New York, pp. 367–382.

Zyskind, G. 1967. On canonical forms, non-negative covariance matrices and best and simple least squares linear estimators in linear models, *Ann. Math. Statist.* 38:1092–1109.

Zyskind, G. and Martin, F. B. 1969. On best linear estimation and a general Gauss-Markov theorem in linear models with arbitrary nonnegative covariance structure, *SIAM J. Appl. Math.* 17:1190–1202.

Received 5 March 1995; final manuscript accepted 4 August 1995