MATHEMATICAL STATISTICS
BANACH CENTER PUBLICATIONS, VOLUME 6
PWN—POLISH SCIENTIFIC PUBLISHERS
WARSAW 1980

A NOTE ON AHLERS AND LEWIS' REPRESENTATION OF THE BEST LINEAR UNBIASED ESTIMATOR IN THE GENERAL GAUSS–MARKOFF MODEL*

JERZY K. BAKSALARY and RADOSŁAW KALA

Department of Mathematical and Statistical Methods, Academy of Agriculture, Poznań, Poland

1. Introduction

Let the triplet

$$(\mathbf{y}, \mathbf{X\beta}, \mathbf{V}) \tag{1}$$

denote the general Gauss–Markoff model in which the $n\times 1$ observable random vector $\mathbf{y}$ has $E(\mathbf{y}) = \mathbf{X\beta}$ as its expectation and $D(\mathbf{y}) = \mathbf{V}$ as its dispersion matrix; $\mathbf{X}$ is an $n\times p$ known matrix of arbitrary rank, $\mathbf{\beta}$ is a $p\times 1$ vector of unknown parameters, and $\mathbf{V}$ is an $n\times n$ nonnegative definite symmetric matrix, known or known except for a positive constant multiplier. Assume that the model is consistent (see Rao [5] for a condition). Further, following Watson [8], assume that the vector of parameters is decomposed as

$$\mathbf{\beta} = \mathbf{\beta}_1+\mathbf{\beta}_2, \tag{2}$$

where: $\mathbf{\beta}_1 \in \mathscr{C}(\mathbf{X}')$, the column space of the transpose of $\mathbf{X}$, and $\mathbf{\beta}_2 \in \mathscr{C}^{\perp}(\mathbf{X}')$, the orthogonal (under the standard inner product) complement of $\mathscr{C}(\mathbf{X}')$. It is known that $\mathbf{\beta}_1$ is unbiasedly estimable in model (1) and that $\mathbf{c}'\mathbf{\beta} = \mathbf{c}'\mathbf{\beta}_1$ for every estimable functional $\mathbf{c}'\mathbf{\beta}$. Also, it is clear that

$$\mathbf{\beta}_1 = \mathbf{X}^+\mathbf{X\beta},$$

where $\mathbf{X}^+$ denotes the Moore–Penrose inverse of $\mathbf{X}$.

In the case of a singular $\mathbf{V}$, there are many representations of the best linear unbiased estimator (BLUE) of $\mathbf{\beta}_1$ given in the literature. Besides referring to the representation proposed by Ahlers and Lewis [1], on which the main interest of this note is focused, we shall use the formula due to Albert [4],

$$\hat{\mathbf{\beta}}_1 = \mathbf{X}^+\mathbf{y}-\mathbf{X}^+\mathbf{V}^{1/2}(\mathbf{Q}\mathbf{V}^{1/2})^+\mathbf{y}, \tag{3}$$

* This research was accomplished during the authors' stay at the Stefan Banach International Mathematical Center.

where

(4) $$\mathbf{Q} = \mathbf{I} - \mathbf{X}\mathbf{X}^{+},$$

the formula due to Rao [6],

(5) $$\hat{\boldsymbol{\beta}}_1 = \mathbf{X}^{+}\mathbf{P}_{\mathbf{X}|\mathbf{VQ}}\mathbf{y},$$

where $\mathbf{P}_{\mathbf{X}|\mathbf{VQ}}$ is a projector on $\mathscr{C}(\mathbf{X})$ along $\mathscr{C}(\mathbf{VQ})$ (see p. 444 of Rao's paper for the definition of such a projector), and also the formula due to Zyskind and Martin [10],

(6) $$\hat{\boldsymbol{\beta}}_1 = (\mathbf{X}'\mathbf{V}^{\#}\mathbf{X})^{+}\mathbf{X}'\mathbf{V}^{\#}\mathbf{y},$$

where $\mathbf{V}^{\#}$ is some suitably chosen g-inverse of $\mathbf{V}$. Formula (6) simplifies to

(7) $$\hat{\boldsymbol{\beta}}_1 = (\mathbf{X}'\mathbf{V}^{+}\mathbf{X})^{+}\mathbf{X}'\mathbf{V}^{+}\mathbf{y}$$

if and only if

(8) $$\mathscr{C}(\mathbf{X}) \subset \mathscr{C}(\mathbf{V}).$$

Note that Albert's result (3) can be viewed as one in which the least squares estimator (LSE) of $\boldsymbol{\beta}_1$, equal to $\mathbf{X}^{+}\mathbf{y}$, is adjusted by another estimator, $-\mathbf{X}^{+}\mathbf{V}^{1/2}(\mathbf{Q}\mathbf{V}^{1/2})^{+}\mathbf{y}$, to provide the BLUE of $\boldsymbol{\beta}_1$. Zyskind and Martin's result, however, can be considered as concordant with Aitken's [2] idea of expressing the BLUE by a solution of some generalized normal equations.

Ahlers and Lewis [1] have stated that the BLUE of $\boldsymbol{\beta}_1$ admits the representation

(9) $$\tilde{\boldsymbol{\beta}}_1 = \mathbf{L}_1\mathbf{y} + \mathbf{L}_2\mathbf{y},$$

where

(10) $$\mathbf{L}_1 = (\mathbf{X}'\mathbf{V}^{+}\mathbf{X})^{+}\mathbf{X}'\mathbf{V}^{+}$$

and

(11) $$\mathbf{L}_2 = (\mathbf{X}'\mathbf{U}\mathbf{X})^{+}\mathbf{X}'\mathbf{U},$$

with

(12) $$\mathbf{U} = \mathbf{I} - \mathbf{V}\mathbf{V}^{+}.$$

Note that, similarly to (3), this result can be viewed as an adjustment of one estimator, $\mathbf{L}_1\mathbf{y}$, being the BLUE of $\boldsymbol{\beta}_1$ under condition (8), by another estimator, $\mathbf{L}_2\mathbf{y}$, the latter having the zero dispersion matrix.

The purpose of this note is two-fold: (i) to show that, in general, formula (9) does not provide the BLUE of $\boldsymbol{\beta}_1$, and (ii) to give a necessary and sufficient condition under which it works correctly.

2. Counterexample

That Ahlers and Lewis' formula (9) does not lead to the BLUE of $\boldsymbol{\beta}_1$ in general case of model (1), can be seen on a simple example, given by Albert [3], p. 92, in which

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \mathbf{V} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},$$

and the vector $\boldsymbol{\beta}$ of unknown parameters consists of one element only. Then, it is obvious that the component $\boldsymbol{\beta}_2$ in the decomposition (2) vanishes, and hence the entire $\boldsymbol{\beta}$ is estimable.

Using formulae (10) to (12) and the fact that, in the example, $\mathbf{V} = \mathbf{V}^{+}$, we obtain

$$\mathbf{L}_1 = (1)^{+}[1 \;\; 0] = [1 \;\; 0]$$

and

$$\mathbf{L}_2 = (1)^{+}[0 \;\; 1] = [0 \;\; 1].$$

Thus, formula (9) gives

$$\tilde{\boldsymbol{\beta}} = (y_1 + y_2),$$

which, in contradiction to Ahlers and Lewis' statement, is not the BLUE of $\boldsymbol{\beta}$, this indeed being equal to (y_2) (cf. Albert [3], p. 92). What more, it may be seen that $E(\tilde{\boldsymbol{\beta}}) = 2\boldsymbol{\beta}$, i.e., $\tilde{\boldsymbol{\beta}}$ does not constitute even a LUE of $\boldsymbol{\beta}$.

3. Necessary and sufficient condition

Although formula (9) is not true in general, nevertheless it works correctly in some special cases of model (1). For instance, if $\mathscr{C}(\mathbf{X}) \subset \mathscr{C}(\mathbf{V})$, then $\mathbf{UX} = \mathbf{0}$, and hence $\mathbf{L}_2 = \mathbf{0}$, giving $\tilde{\boldsymbol{\beta}}_1 = \mathbf{L}_1\mathbf{y}$, which coincides with (7), a particular case of Zyskind and Martin's representation (6). Moreover, if $\mathbf{V} = \mathbf{0}$, then also $\mathbf{V}^{+} = \mathbf{0}$, and hence $\mathbf{L}_1 = \mathbf{0}$ while $\mathbf{L}_2 = \mathbf{X}^{+}$. Thus $\tilde{\boldsymbol{\beta}}_1 = \mathbf{X}^{+}\mathbf{y}$, which is the LSE of $\boldsymbol{\beta}_1$, but in this case (see Zyskind [9], p. 1099) simultaneously is BLUE. Therefore, it is natural to inquire about a general condition under which formula (9) provides the BLUE of $\boldsymbol{\beta}_1$. The answer is given in the following

THEOREM. *Let* $\boldsymbol{\beta}_1$ *be the estimable part of the vector of parameters in the general Gauss–Markoff model* $(\mathbf{y}, \mathbf{X}\boldsymbol{\beta}, \mathbf{V})$. *A necessary and sufficient condition for*

(13) $$\tilde{\boldsymbol{\beta}}_1 = (\mathbf{L}_1 + \mathbf{L}_2)\mathbf{y},$$

where $\mathbf{L}_1$ *and* $\mathbf{L}_2$ *are specified as in* (10) *and* (11), *to be the* BLUE *of* $\boldsymbol{\beta}_1$ *is*

(14) $$\mathscr{C}(\mathbf{X}\mathbf{X}'\mathbf{V}) \subset \mathscr{C}(\mathbf{V}).$$

Proof. In view of (5) it is clear that (13) will represent the BLUE of $\boldsymbol{\beta}_1$ if and only if

$$(\mathbf{L}_1 + \mathbf{L}_2)\mathbf{y} = \mathbf{X}^{+}P_{\mathbf{X}|\mathbf{VQ}}\mathbf{y}$$

for every $\mathbf{y}$ for which the model is consistent, i.e. (see Rao [5], Lemma 2.1), for every $\mathbf{y} \in \mathscr{C}\{(\mathbf{X} \,\vdots\, \mathbf{V})\}$, or (see Rao [6], Lemma 2.1 (ii)) for every $\mathbf{y} \in \mathscr{C}\{(\mathbf{X} \,\vdots\, \mathbf{VQ})\}$. But this is equivalent to the requirement for the relations

(15) $$(\mathbf{L}_1 + \mathbf{L}_2)\mathbf{X}\mathbf{a} = \mathbf{X}^{+}\mathbf{P}_{\mathbf{X}|\mathbf{VQ}}\mathbf{X}\mathbf{a}$$

and

(16) $$(\mathbf{L}_1 + \mathbf{L}_2)\mathbf{V}\mathbf{Q}\mathbf{b} = \mathbf{X}^{+}\mathbf{P}_{\mathbf{X}|\mathbf{VQ}}\mathbf{V}\mathbf{Q}\mathbf{b}$$

to hold for every $\mathbf{a}$ and every $\mathbf{b}$, respectively. Then, since $\mathbf{P}_{\mathbf{X}|\mathbf{VQ}}\mathbf{X} = \mathbf{X}$ and $\mathbf{P}_{\mathbf{X}|\mathbf{VQ}}\mathbf{VQ} = \mathbf{0}$, (15) and (16) become

$$\mathbf{L}_1\mathbf{X}+\mathbf{L}_2\mathbf{X} = \mathbf{X}^+\mathbf{X} \tag{17}$$

and

$$\mathbf{L}_1\mathbf{VQ} = \mathbf{0}. \tag{18}$$

Thus, it remains to establish that (14) implies (17) and (18), and vice versa.

To show the first implication, apply (10) and (11) to check that both components on the left-hand side of (17) are idempotent matrices, and thus projectors. Furthermore, note that (14) implies $\mathbf{V}^+\mathbf{XX}' = \mathbf{WV}^+$ for some $\mathbf{W}$. Hence, in view of (10), (11), (12), and the symmetry of $\mathbf{L}_2\mathbf{X}$,

$$\mathbf{L}_1\mathbf{XL}_2\mathbf{X} = (\mathbf{X}'\mathbf{V}^+\mathbf{X})^+\mathbf{X}'\mathbf{V}^+\mathbf{XX}'\mathbf{L}_2' = (\mathbf{X}'\mathbf{V}^+\mathbf{X})^+\mathbf{X}'\mathbf{WV}^+\mathbf{UX}(\mathbf{X}'\mathbf{UX})^+ = \mathbf{0}.$$

Similarly, $\mathbf{L}_2\mathbf{XL}_1\mathbf{X} = \mathbf{0}$. Therefore, by Rao and Mitra's [7], Theorem 5.1.2, the left-hand side of (17) is the projector on $\mathscr{C}(\mathbf{L}_1\mathbf{X})\oplus\mathscr{C}(\mathbf{L}_2\mathbf{X})$ along $\mathscr{N}(\mathbf{L}_1\mathbf{X})\cap\mathscr{N}(\mathbf{L}_2\mathbf{X})$, where $\mathscr{N}(\cdot)$ stands for the null space of a matrix argument. But $\mathscr{C}(\mathbf{L}_1\mathbf{X}) = \mathscr{C}(\mathbf{X}'\mathbf{V}^+)$ and $\mathscr{C}(\mathbf{L}_2\mathbf{X}) = \mathscr{C}(\mathbf{X}'\mathbf{U})$, and then it is easy to verify that

$$\mathscr{C}(\mathbf{L}_1\mathbf{X})\oplus\mathscr{C}(\mathbf{L}_2\mathbf{X}) = \mathscr{C}(\mathbf{X}')$$

and

$$\mathscr{N}(\mathbf{L}_1\mathbf{X})\cap\mathscr{N}(\mathbf{L}_2\mathbf{X}) = \mathscr{C}^{\perp}(\mathbf{X}').$$

Therefore $\mathbf{L}_1\mathbf{X}+\mathbf{L}_2\mathbf{X}$ turns out to be the orthogonal projector on $\mathscr{C}(\mathbf{X}')$, thus being equal to $\mathbf{X}^+\mathbf{X}$. This shows that (14) implies (17). To prove that (14) also implies (18), note that, under (14), $\mathbf{XX}'\mathbf{V} = \mathbf{VZ}$ for some $\mathbf{Z}$. Then

$$\mathscr{C}(\mathbf{VV}^+\mathbf{X}) = \mathscr{C}(\mathbf{VV}^+\mathbf{XX}'\mathbf{VV}^+) = \mathscr{C}(\mathbf{VZV}^+) = \mathscr{C}(\mathbf{XX}'\mathbf{VV}^+) \subset \mathscr{C}(\mathbf{X}),$$

and hence $\mathbf{X}'\mathbf{V}^+\mathbf{V} = \mathbf{TX}'$ for some $\mathbf{T}$. Thus, by (10) and (4),

$$\mathbf{L}_1\mathbf{VQ} = (\mathbf{X}'\mathbf{V}^+\mathbf{X})^+\mathbf{X}'\mathbf{V}^+\mathbf{VQ} = (\mathbf{X}'\mathbf{V}^+\mathbf{X})^+\mathbf{TX}'\mathbf{Q} = \mathbf{0},$$

and so the proof of the sufficiency of (14) is completed.

Conversely, by the repeated use of Rao and Mitra's [7], Theorem 5.1.2, it follows from (17) that $\mathbf{L}_1\mathbf{XL}_2\mathbf{X} = \mathbf{0}$, or

$$(\mathbf{X}'\mathbf{V}^+\mathbf{X})^+\mathbf{X}'\mathbf{V}^+\mathbf{X}(\mathbf{X}'\mathbf{UX})^+\mathbf{X}'\mathbf{UX} = \mathbf{0}.$$

This, premultiplied by $\mathbf{V}^+\mathbf{X}$ and postmultiplied by $\mathbf{X}'\mathbf{U}$, implies that

$$\mathbf{V}^+\mathbf{XX}'\mathbf{U} = \mathbf{0},$$

i.e., in view of (12), that $\mathscr{C}(\mathbf{XX}'\mathbf{V}^+) \subset \mathscr{C}(\mathbf{V}^+)$. But this can easily be seen as being equivalent to (14), which concludes the proof.

References

[1] C. W. Ahlers and T. O. Lewis, *Linear estimation with a positive semidefinite covariance matrix*, Industrial Mathematics 21 (1971), pp. 23–27.

[2] A. C. Aitken, *On least squares and linear combination of observations*, Proc. Roy. Soc. Edinburgh A 55 (1935), pp. 42–48.

[3] A. Albert, *Regression and the Moore–Penrose pseudoinverse*, Academic Press, New York 1972.

[4] —, *The Gauss–Markov theorem for regression models with possibly singular covariances*, SIAM J. Appl. Math. 24 (1973), pp. 182–187.

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

[6] —, *Projectors, generalized inverses and the* BLUE's, J. Roy. Statist. Soc. B 36 (1974), pp. 442–448.

[7] C. R. Rao and S. K. Mitra, *Generalized inverse of matrices and its applications*, Wiley, New York 1971.

[8] G. S. Watson, *Prediction and the efficiency of least squares*, Biometrika 59 (1972), pp. 91–98.

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

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

Presented to the semester
MATHEMATICAL STATISTICS
September 15–December 18, 1976