Characterizations of the Best Linear Unbiased Estimator in the General Gauss-Markov Model with the Use of Matrix Partial Orderings

Jerzy K. Baksalary*
Department of Mathematical and Statistical Methods
Academy of Agriculture in Poznań
Wojska Polskiego 28
PL-60-637 Poznań, Poland

and

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

Submitted by F. Pukelsheim

ABSTRACT

Under the general Gauss-Markov model $\{\mathbf{Y}, \mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{V}\}$, two new characterizations of BLUE($\mathbf{X}\boldsymbol{\beta}$) are derived involving the Löwner and rank-subtractivity partial orderings between the dispersion matrix of BLUE($\mathbf{X}\boldsymbol{\beta}$) and the dispersion matrix of $\mathbf{Y}$. As particular cases of these characterizations, three new criteria for the equality between OLSE($\mathbf{X}\boldsymbol{\beta}$) and BLUE($\mathbf{X}\boldsymbol{\beta}$) are given.

1. INTRODUCTION AND PRELIMINARIES

Consider the general Gauss-Markov model, denoted by

$$\mathrm{M} = \{\mathbf{Y}, \mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{V}\}, \tag{1.1}$$

in which $\mathbf{Y}$ is an $n \times 1$ observable random vector with expectation $E(\mathbf{Y}) = \mathbf{X}\boldsymbol{\beta}$

*Research partially supported by Grant No. CPBP 01.02.IV.10.
†This research was completed while the author was an acting senior researcher of the Academy of Finland.

655 Avenue of the Americas, New York, NY 10010 0024-3795/90/$3.50

and with dispersion matrix $D(\mathbf{Y}) = \sigma^2\mathbf{V}$, where $\mathbf{X}$ is an $n \times p$ nonnull known matrix, $\boldsymbol{\beta}$ is a $p \times 1$ vector of unknown parameters, $\mathbf{V}$ is an $n \times n$ known symmetric nonnegative definite matrix, and σ^2 is an unknown positive scalar. The matrices $\mathbf{X}$ and $\mathbf{V}$ are both allowed to be of arbitrary rank. Moreover, it is assumed throughout the note that the model (1.1) is consistent [cf. Rao (1971, p. 378; 1973a, p. 297)], or, in other words, that the inference base is not self-contradictory [cf. Feuerverger and Fraser (1980, p. 44)], i.e.,

$$\mathbf{Y} \in \mathscr{R}(\mathbf{X}:\mathbf{V}), \tag{1.2}$$

where $\mathscr{R}(\mathbf{X}:\mathbf{V})$ denotes the range of the partitioned matrix $(\mathbf{X}:\mathbf{V})$.

As is well known, a statistic $\mathbf{FY}$ is said to be the best linear unbiased estimator (BLUE) of $\mathbf{X}\boldsymbol{\beta}$ if $E(\mathbf{FY}) = \mathbf{X}\boldsymbol{\beta}$ and $D(\mathbf{FY}) \leqslant_{\mathsf{L}} D(\mathbf{GY})$ for every $\mathbf{GY}$ such that $E(\mathbf{GY}) = \mathbf{X}\boldsymbol{\beta}$. Here $\mathbf{A} \leqslant_{\mathsf{L}} \mathbf{B}$ means that $\mathbf{A}$ is below $\mathbf{B}$ with respect to the Löwner partial ordering [cf. Marshall and Olkin (1979, p. 462)], i.e., that the difference $\mathbf{B} - \mathbf{A}$ is a symmetric nonnegative definite matrix. Since $E(\mathbf{Y}) = \mathbf{X}\boldsymbol{\beta}$, an obvious consequence of the definition is that a necessary condition for a statistic $\mathbf{FY}$ to represent BLUE($\mathbf{X}\boldsymbol{\beta}$) is the Löwner ordering

$$D(\mathbf{FY}) \leqslant_{\mathsf{L}} D(\mathbf{Y}). \tag{1.3}$$

The second matrix partial ordering utilized in this paper is rank subtractivity, defined as

$$\mathbf{A} \leqslant_{\mathsf{rs}} \mathbf{B} \quad \text{whenever} \quad \operatorname{rank}(\mathbf{B} - \mathbf{A}) = \operatorname{rank}(\mathbf{B}) - \operatorname{rank}(\mathbf{A}).$$

It was originally introduced by Hartwig (1980) and now is referred to in the literature also as the minus partial ordering; cf. Baksalary (1986), Baksalary, Pukelsheim, and Styan (1989), Carlson (1987), Hartwig and Styan (1986), Mitra (1986). From Marsaglia and Styan (1974, p. 288) and Cline and Funderlic (1979, p. 195) it follows that

$$\mathbf{A} \leqslant_{\mathsf{rs}} \mathbf{B} \quad \Leftrightarrow \quad \mathscr{R}(\mathbf{A}) \subseteq \mathscr{R}(\mathbf{B}),\, \mathscr{R}(\mathbf{A}') \subseteq \mathscr{R}(\mathbf{B}'),\ \text{and}\ \mathbf{A}\mathbf{B}^{-}\mathbf{A} = \mathbf{A}, \tag{1.4}$$

where $\mathbf{B}'$ and $\mathbf{B}^{-}$ denote the transpose and a generalized inverse of $\mathbf{B}$, respectively. It is known [cf. Baksalary and Hauke (1984, p. 35)], that if $\mathbf{A}, \mathbf{B}$ are both symmetric nonnegative definite matrices, then

$$\mathbf{A} \leqslant_{\mathsf{rs}} \mathbf{B} \quad \Rightarrow \quad \mathbf{A} \leqslant_{\mathsf{L}} \mathbf{B}. \tag{1.5}$$

See also a strengthened version of this result in Hartwig and Styan (1987, Theorem 2.1).

Perhaps the most widely known general characterization of $\text{BLUE}(\mathbf{X}\boldsymbol{\beta})$ is that given in Drygas (1970, p. 55) and Rao (1973b, p. 282):

$$\mathbf{FY} = \text{BLUE}(\mathbf{X}\boldsymbol{\beta}) \text{ under M} \quad \Leftrightarrow \quad \mathbf{F}(\mathbf{X}:\mathbf{VZ}) = (\mathbf{X}:\mathbf{0}), \tag{1.6}$$

where $\mathbf{Z}$ is any matrix such that $\mathscr{R}(\mathbf{Z}) = \mathscr{N}(\mathbf{X}')$, the null space of $\mathbf{X}'$. According to Rao (1974), the equation on the right-hand side of (1.6) means that $\mathbf{F}$ is a projector onto $\mathscr{R}(\mathbf{X})$ along $\mathscr{R}(\mathbf{VZ})$, such a projector being unique only when $\mathscr{R}(\mathbf{X}:\mathbf{VZ}) = \mathbb{R}^n$. See also Rao (1978, Theorem 1) for representations of the general solution to the matrix equation involved.

Notice that the condition for unbiasedness of $\mathbf{FY}$ is stated in (1.6) in the form $\mathbf{FX} = \mathbf{X}$, which actually corresponds to considering $\boldsymbol{\beta}$ as free to vary over $\mathbb{R}^p$ or at least over $\mathscr{R}(\mathbf{X}')$, or, equivalently, to neglecting restrictions on $\boldsymbol{\beta}$ that arise from singularity of $\mathbf{V}$ when $\mathscr{R}(\mathbf{X}) \not\subseteq \mathscr{R}(\mathbf{V})$ but are not completely specified prior to an observation of $\mathbf{Y}$; cf. comments in Rao (1979, p. 1354) and Seely and Zyskind (1971, p. 693). In this context it should be emphasized that, as originally pointed out by Rao (1973a, pp. 297–298; 1976, p. 1033; 1985, p. 20), there is no loss in generality in adopting the stronger unbiasedness condition, for if $\mathbf{FY}$ is an unbiased estimator of $\mathbf{X}\boldsymbol{\beta}$ not satisfying $\mathbf{FX} = \mathbf{X}$, then there exists $\mathbf{GY}$ such that $\mathbf{GX} = \mathbf{X}$ and $\mathbf{GY} = \mathbf{FY}$ with probability one.

The purpose of this paper is to derive characterizations of $\text{BLUE}(\mathbf{X}\boldsymbol{\beta})$ with the use of the Löwner ordering (1.3) and the rank-subtractivity ordering

$$D(\mathbf{FY}) \leqslant_{\text{rs}} D(\mathbf{Y}). \tag{1.7}$$

These characterizations appear to be quite interesting. It turns out that in certain cases it is just sufficient to combine (1.3) or (1.7) with unbiasedness of $\mathbf{FY}$ to force $\mathbf{FY}$ to be a representation of $\text{BLUE}(\mathbf{X}\boldsymbol{\beta})$. In particular, this leads to new criteria for the equality between $\text{BLUE}(\mathbf{X}\boldsymbol{\beta})$ and the ordinary least-squares estimator of $\mathbf{X}\boldsymbol{\beta}$, the latter being defined as

$$\text{OLSE}(\mathbf{X}\boldsymbol{\beta}) = \mathbf{P}_{\mathbf{X}}\mathbf{Y}, \tag{1.8}$$

where $\mathbf{P}_{\mathbf{X}}$ denotes the orthogonal projector onto $\mathscr{R}(\mathbf{X})$.

2. RESULTS

In the sequel, we will utilize Lemma 2.1 of Rao (1974) indicating the possibility of decomposing $\mathscr{R}(\mathbf{X}:\mathbf{V})$ as

$$\mathscr{R}(\mathbf{X}:\mathbf{V}) = \mathscr{R}(\mathbf{X}) \oplus \mathscr{R}(\mathbf{VZ}), \tag{2.1}$$

where $\mathscr{R}(\mathbf{Z}) = \mathscr{N}(\mathbf{X}')$. See also Nordström (1985) for a survey of various decompositions of linear subspaces related to the model (1.1).

THEOREM. *Under the general Gauss-Markov model* $\mathrm{M} = \{\mathbf{Y}, \mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{V}\}$,

$$\mathbf{FY} = \mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta}) \textit{ for all } \mathbf{Y} \in \mathscr{R}(\mathbf{X}:\mathbf{V})$$

if and only if the conditions

(a) $\mathbf{FY}$ *is unbiased, i.e.*, $\mathbf{FX} = \mathbf{X}$,

(b) $\mathscr{R}(\mathbf{FV}) \subseteq \mathscr{R}(\mathbf{X})$, *i.e.*, $\mathbf{FV} = \mathbf{XK}$ *for some* $\mathbf{K}$,

hold along with any one of the following three conditions:

(c_1) $\mathbf{FV} = \mathbf{VF}'$,

(c_2) $D(\mathbf{FY}) \leqslant_{\mathsf{L}} D(\mathbf{Y})$, *i.e.*, $\mathbf{V} - \mathbf{FVF}'$ *is nonnegative definite*,

(c_3) $D(\mathbf{FY}) \leqslant_{\mathsf{rs}} D(\mathbf{Y})$, *i.e.*, $\mathrm{rank}(\mathbf{V} - \mathbf{FVF}') = \mathrm{rank}(\mathbf{V}) - \mathrm{rank}(\mathbf{FVF}')$.

Proof. Combining (c_1) with (b) yields $\mathscr{R}(\mathbf{VF}') \subseteq \mathscr{R}(\mathbf{X})$, which is clearly equivalent to $\mathbf{FVZ} = \mathbf{0}$. In view of (1.6), this establishes the sufficiency of the triplet (a), (b), (c_1). Conversely, the necessity of (a) and (c_2) is obvious. Moreover, the equation on the right-hand side of (1.6) implies that

$$\mathbf{Z}'\mathbf{F}(\mathbf{X}:\mathbf{VZ}) = \mathbf{Z}'(\mathbf{X}:\mathbf{0}) = (\mathbf{0}:\mathbf{0}).$$

Hence, in view of (2.1),

$$\mathbf{Z}'\mathbf{F}(\mathbf{X}:\mathbf{V}) = (\mathbf{0}:\mathbf{0}),$$

thus leading to (b). It is easily seen that (a) and (b) entail $\mathbf{F}^2\mathbf{V} = \mathbf{FV}$. Corollary 1 of Baksalary, Kala, and Kłaczyński (1983), which generalizes

Proposition 2 of Taylor (1976), asserts that if $\mathbf{F}^2\mathbf{V}=\mathbf{FV}$, then the Löwner ordering (c_2) is equivalent to condition (c_1). Further, (c_1) implies that

$$\mathscr{R}(\mathbf{FVF}') \subseteq \mathscr{R}(\mathbf{V}), \tag{2.2}$$

and (c_1) with $\mathbf{F}^2\mathbf{V}=\mathbf{FV}$ implies that

$$\mathbf{FVF}'\mathbf{V}^-\mathbf{FVF}' = \mathbf{F}^2\mathbf{V}(\mathbf{F}^2)' = \mathbf{FVF}'. \tag{2.3}$$

In view of (1.4), the conditions (2.2) and (2.3) are equivalent to (c_3). Consequently, since always (c_3) $\Rightarrow$ (c_2) [cf. (1.5)], it follows that if $\mathbf{F}^2\mathbf{V}=\mathbf{FV}$, then conditions ($c_1$), ($c_2$), and ($c_3$) are all equivalent, which concludes the proof. ∎

The characterization consisting of conditions (a), (b), and (c_1) is a minor generalization of Theorem 5.1 in Rao (1971) which asserts that, under the assumption $\operatorname{rank}(\mathbf{F}) \leqslant \operatorname{rank}(\mathbf{X})$, the equality $\mathbf{FY} = \operatorname{BLUE}(\mathbf{X}\boldsymbol{\beta})$ holds if and only if $\mathbf{FX}=\mathbf{X}$, $\mathbf{F}^2=\mathbf{F}$, $\mathbf{FV}=\mathbf{VF}'$, and $\mathbf{F}=\mathbf{XD}$, where $\mathbf{D}'$ is a minimum $\mathbf{V}$-seminorm generalized inverse of $\mathbf{X}'$. The remaining two characterizations, consisting of conditions (a), (b), (c_2) and (a), (b), (c_3), reveal new appearances of matrix partial orderings in problems of mathematical statistics. This seems to be particularly noteworthy with reference to the rank-subtractivity ordering, whose statistical applications have hitherto been discussed only in the context of the distribution theory of quadratic forms in normal variables; cf. Baksalary and Hauke (1984, Section 3) and Hartwig and Styan (1986, Section 3F).

The theorem above will now be applied to derive criteria for the equality between $\operatorname{OLSE}(\mathbf{X}\boldsymbol{\beta})$ and $\operatorname{BLUE}(\mathbf{X}\boldsymbol{\beta})$, the former being defined in (1.8).

COROLLARY. *Under the general Gauss-Markov model* $\mathrm{M} = \{\mathbf{Y}, \mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{V}\}$, *the following statements are equivalent*:

(a) $\operatorname{OLSE}(\mathbf{X}\boldsymbol{\beta}) = \operatorname{BLUE}(\mathbf{X}\boldsymbol{\beta})$ *for all* $\mathbf{Y} \in \mathscr{R}(\mathbf{X}:\mathbf{V})$,

(b) $\mathbf{P_X V} = \mathbf{V P_X}$,

(c) $\mathbf{P_X V P_X} \leqslant_{\mathsf{L}} \mathbf{V}$, *i.e.*, $\mathbf{V} - \mathbf{P_X V P_X}$ *is nonnegative definite*,

(d) $\mathbf{P_X V P_X} \leqslant_{\mathsf{rs}} \mathbf{V}$, *i.e.*, $\operatorname{rank}(\mathbf{V} - \mathbf{P_X V P_X}) = \operatorname{rank}(\mathbf{V}) - \operatorname{rank}(\mathbf{VX})$,

(e) $\operatorname{rank}(\mathbf{V} - \mathbf{P_X V P_X}) = \dim \mathscr{N}(\mathbf{X}') \cap \mathscr{R}(\mathbf{V})$.

Proof. Since $\mathbf{F} = \mathbf{P_X}$ obviously satisfies conditions (a) and (b) of the Theorem, its immediate consequence is that (a) $\Leftrightarrow$ (b) $\Leftrightarrow$ (c) $\Leftrightarrow$ (d). The equivalence of (d) to (e) follows from the equalities

$$\operatorname{rank}(\mathbf{Z}:\mathbf{V}) = \operatorname{rank}(\mathbf{Z}) + \operatorname{rank}(\mathbf{P_X V}) = \operatorname{rank}(\mathbf{Z}) + \operatorname{rank}(\mathbf{VX})$$

and

$$\operatorname{rank}(\mathbf{Z}:\mathbf{V}) = \operatorname{rank}(\mathbf{Z}) + \operatorname{rank}(\mathbf{V}) - \dim \mathcal{N}(\mathbf{X}') \cap \mathcal{R}(\mathbf{V});$$

cf., e.g., Marsaglia and Styan (1974, p. 274). ■

Exhaustive discussions of the problem of the equality between OLSE($\mathbf{X}\boldsymbol{\beta}$) and BLUE($\mathbf{X}\boldsymbol{\beta}$) have recently been given by Alalouf and Styan (1984), Puntanen (1987), and Puntanen and Styan (1989). This problem originated from the paper of Anderson (1948), and the first complete solution in the general case when both $\mathbf{X}$ and $\mathbf{V}$ may be deficient in rank was derived by Rao (1967, p. 364) in the form $\mathbf{X}'\mathbf{VZ} = \mathbf{0}$. The equivalence of this condition to (b) in the Corollary, as well as to seven other conditions, was established by Zyskind (1967, Theorem 2).

Criteria (c), (d), and (e) are new. The first of them asserts that OLSE($\mathbf{X}\boldsymbol{\beta}$) coincides with BLUE($\mathbf{X}\boldsymbol{\beta}$) if and only if it is not worse than $\mathbf{Y}$, the rough estimator of $\mathbf{X}\boldsymbol{\beta}$, with respect to the dispersion-matrix criterion or, since both $\mathbf{P_X Y}$ and $\mathbf{Y}$ are unbiased, with respect to the mean-square-error-matrix criterion. The necessity of this condition is obvious, but its sufficiency we find somewhat unexpected. Criteria (d) and (e) seem to be the first conditions expressed in terms of the ranks of the matrices $\mathbf{X}$ and $\mathbf{V}$ without referring to the spectral decomposition of $\mathbf{V}$, as does the rank criterion originally given by Anderson (1971, 1972) and then extended to the general Gauss-Markov model by Styan (1973) and Baksalary and Kala (1977).

REFERENCES

Alalouf, I. S. and Styan, G. P. H. 1984. Characterizations of the conditions for the ordinary least squares estimator to be best linear unbiased, in *Topics in Applied Statistics* (Chaubey, Y. P. and Dwivedi, T. D., Eds.), Dept. of Mathematics, Concordia Univ., Montreal, pp. 331–344.

Anderson, T. W. 1948. On the theory of testing serial correlation, *Skand. Aktuarietidskr.* 31:88–116.

Anderson, T. W. 1971. *The Statistical Analysis of Time Series*, Wiley, New York.

Anderson, T. W. 1972. Efficient estimation of regression coefficients in time series, in *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 1 (Le Cam, L. M., Neyman, J., and Scott, E. L., Eds.), Univ. of California Press, Berkeley, pp. 471–482.

Baksalary, J. K. 1986. A relationship between the star and minus orderings, *Linear Algebra Appl.* 82:163–167.

Baksalary, J. K. and Hauke, J. 1984. Inheriting independence and chi-squaredness under certain matrix orderings, *Statist. Probab. Lett.* 2:35–38.

Baksalary, J. K. and Kala, R. 1977. An extension of a rank criterion for the least squares estimator to be the best linear unbiased estimator, *J. Statist. Plann. Inference* 1:309–312.

Baksalary, J. K., Kala, R., and Kłaczyński, K. 1983. The matrix inequality $\mathbf{M} \geqslant \mathbf{B}^*\mathbf{MB}$, *Linear Algebra Appl.* 54:77–86.

Baksalary, J. K., Pukelsheim, F., and Styan, G. P. H. 1989. Some properties of matrix partial orderings, *Linear Algebra Appl.*, 119, 57–85.

Carlson, D. 1987. Generalized inverse invariance, partial orders, and rank-minimization problems for matrices, in *Current Trends in Matrix Theory* (Uhlig, F. and Grone, R., Eds.), Elsevier, New York, pp. 81–87.

Cline, R. E. and Funderlic, R. E. 1979. The rank of a difference of matrices and associated generalized inverses, *Linear Algebra Appl.* 24:185–215.

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

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

Hartwig, R. E. 1980. How to partially order regular elements, *Math. Japon.* 25:1–13.

Hartwig, R. E. and Styan, G. P. H. 1986. On some characterizations of the "star" partial ordering for matrices and rank subtractivity, *Linear Algebra Appl.* 82:145–161.

Hartwig, R. E. and Styan, G. P. H. 1987. Partially ordered idempotent matrices, in *Proceedings of the Second International Tampere Conference in Statistics* (Pukkila, T. and Puntanen, S., Eds.), Dept. of Mathematical Sciences, Univ. of Tampere, Tampere, pp. 361–383.

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

Marshall, A. W. and Olkin, I. 1979. *Inequalities: Theory of Majorization and Its Applications*, Academic, New York.

Mitra, S. K. 1986. The minus partial order and the shorted matrix, *Linear Algebra Appl.* 83:1–27.

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

Puntanen, S. 1987. On the relative goodness of ordinary least squares estimation in the general linear model, *Acta Univ. Tamper. Ser. A*, Vol. 216.

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.

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 (Le Cam, L. M. and Neyman, J., Eds.), Univ. of California Press, Berkeley, pp. 355–372.

Rao, C. R. 1971. Unified theory of linear estimation, *Sankhyā Ser. A* 33:371–394; Corrigenda, 34:194, 477 (1972).

Rao, C. R. 1973a. *Linear Statistical Inference and Its Applications*, 2nd ed., Wiley, New York (1st ed., 1965).

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. 1974. Projectors, generalized inverses and the BLUE's, *J. Roy. Statist. Soc. Ser. B* 36:442–448.

Rao, C. R. 1976. Estimation of parameters in a linear model, *Ann. Statist.* 4:1023–1037; Corrigendum, 7:696 (1979).

Rao, C. R. 1978. Choice of best linear estimators in the Gauss-Markoff model with a singular dispersion matrix, *Comm. Statist. A—Theory Methods* 7:1199–1208.

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

Rao, C. R. 1985. A unified approach to inference from linear models, in *Proceedings of the First International Tampere Seminar on Linear Statistical Models and Their Applications* (Pukkila, T. and Puntanen, S., Eds.), Dept. of Mathematical Sciences, Univ. of Tampere, Tampere, pp. 9–36.

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

Styan, G. P. H. 1973. When does least squares give the best linear unbiased estimate? in *Multivariate Statistical Inference* (Kabe, D. G. and Gupta, R. P., Eds.), North-Holland, Amsterdam, pp. 241–246.

Taylor, W. E. 1976. Prior information on the coefficients when the disturbance covariance matrix is unknown, *Econometrica* 44:725–739.

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.

Received 20 April 1988; final manuscript accepted 6 October 1988