Linear Estimation With an Incorrect Dispersion Matrix in Linear Models With a Common Linear Part

THOMAS MATHEW*

The linear model $(Y, X\beta, V)$, where X is any matrix satisfying the conditions $\mathcal{R}(U) \subset \mathcal{R}(X)$ and $\mathcal{R}(A') \subset \mathcal{R}(X')$, is considered, U and A being known matrices and $\mathcal{R}(\cdot)$ denoting range space. Nonnegative definite matrices V, for which every linear representation or some linear representation of the BLUE of $A\beta$ under $(Y, X\beta, V_1)$ continues to be its BLUE under $(Y, X\beta, V)$ for every such matrix X, are characterized. Conditions under which estimable linear parametric functions admit a BLUE for the variance components model are also given.

KEY WORDS: Ordinary least squares estimator; Best linear unbiased estimator; Variance components model.

1. INTRODUCTION AND SUMMARY

Consider the general linear model $Y = X\beta + e$ where $Y \in R^n$ is a random vector, X is the $n \times m$ design matrix, $\beta \in R^m$ is a vector of unknown parameters, and e is an $n \times 1$ random vector with $E(e) = \mathbf{0}$ and $\text{cov}(e) = V$ (possibly a singular matrix). Such a model will henceforth be denoted by the triplet $(Y, X\beta, V)$. An estimable linear paramatric function, ordinary least squares (OLS) estimator, and best linear unbiased estimator (BLUE) under the model $(Y, X\beta, V)$ are well-known terms.

The following notations are used in this article. For any matrix A, $\mathcal{R}(A)$ and $\mathcal{N}(A)$ denote respectively the vector space spanned by the columns (or the range space) and the null space of A, $r(A)$ denotes the rank of A, A^- denotes a generalized inverse of A, A_l^- denotes a least squares g-inverse of A, A^+ denotes the Moore-Penrose inverse of A (see Rao and Mitra 1971, Ch. 3) and $A^\perp$ denotes a matrix of maximum rank that satisfies $A'A^\perp = 0$.

Conditions under which the OLS estimators are also BLUE's have been investigated by Rao (1967), Zyskind (1967), Watson (1967), Kruskal (1968), and Haberman (1975). McElroy (1967) has given a necessary and sufficient condition so that for all models $(Y, X\beta, V)$ in which every element of the first column of X is equal to one, the OLS estimators are also BLUE. This problem has been considered in a more general framework by Zyskind (1969). He considers linear models, with $r(X) = p$, such that $\mathcal{R}(X)$ contains a particular vector subspace $\mathcal{R}(U)$ of dimension less than p, and gives general conditions on the form of V so that for all such models every OLS estimator is also BLUE. One important consequence of Zyskind's result is that if V satisfies appropriate conditions, then all parametric augmentations of the linear model $Y = U\beta + e$ will give rise to models for which the OLS estimators are also BLUE's.

We start with the set up described by Zyskind and consider all linear models $(Y, X\beta, V_1)$ with a common linear part $\mathcal{R}(U)$, that is, $r(X) = p < n$ such that $\mathcal{R}(U) \subset \mathcal{R}(X)$, U being a known matrix with $r(U) < p$. The class of such matrices is denoted by $C^p(U)$. As pointed out by Mitra and Moore (1973), the BLUE of an estimable parametric function may not have a unique linear representation under $(Y, X\beta, V_1)$ when V_1 is singular. In Section 2 we consider design matrices $X \in C^p(U)$ further satisfying the condition $\mathcal{R}(A') \subset \mathcal{R}(X')$, where A is a given matrix with $1 \le r(A) \le p$. The class of such matrices X is denoted by $C_A{}^p(U)$. We characterize matrices V such that every linear representation or some linear representation of the BLUE of $A\beta$ under $(Y, X\beta, V_1)$ continues to be its BLUE under $(Y, X\beta, V)$ also. It turns out, surprisingly, that if, for some nonnull matrix A, every linear representation of the BLUE $A\beta$ under $(Y, X\beta, V_1)$ continues to be its BLUE under $(Y, X\beta, V)$ for every $X \in C_A{}^p(U)$, then every linear representation of the BLUE of every estimable parametric function under $(Y, X\beta, V_1)$ continues to be its BLUE under $(Y, X\beta, V)$ for every $X \in C^p(U)$.

Section 3 deals with the variance components and covariance components models. We obtain conditions under which $A\beta$ admits a BLUE for every $X \in C_A{}^p(U)$ when the variance components and covariance components are unknown.

2. BLUE ESTIMATION WITH AN INCORRECT DISPERSION MATRIX

In this section, we characterize V such that every linear representation or some linear representation of the

* Thomas Mathew is senior research fellow, Applied Statistics, Surveys and Computing Division, Indian Statistical Institute, 203 B.T. Road, Calcutta-700035, India. The author wishes to express his sincere appreciation of the valuable comments and guidance of Dr. P. Bhimasankaram. He is deeply indebted to a referee for suggesting Lemmas A.2 and A.3 which simplified the proof of Theorem 2.1 considerably. Thanks are also due to an associate editor whose helpful suggestions have improved the presentation of the article.

© Journal of the American Statistical Association
June 1983, Volume 78, Number 382
Theory and Methods Section

BLUE of $A\beta$ under $(Y, X\beta, V_1)$ continues to be its BLUE under $(Y, X\beta, V)$ for every $X \in C_A{}^p(U)$.

Theorem 2.1. Every linear representation of the BLUE of $A\beta$ under $(Y, X\beta, V_1)$ is its BLUE under $(Y, X\beta, V)$ for every $X \in C_A{}^p(U)$ if and only if any one of the following equivalent conditions holds:

(i) $VU^\perp = \lambda V_1 U^\perp$ for some $\lambda \geq 0$
(ii) Every vector in $\mathcal{R}(U^\perp)$ is an eigenvector of V with respect to V_1
(iii) $V = \lambda V_1 + UBU'$, for some matrix B and scalar $\lambda \geq 0$ such that V is nonnegative definite.

Proof. It is well known that LY is the BLUE of $A\beta$ under $(Y, X\beta, V_1)$ if and only if $X'L = A'$ and $X^{\perp\prime} V_1 L = 0$. If (i) holds, then $X^{\perp\prime} V_1 L = 0 \Rightarrow X^{\perp\prime} VL = 0$ for every $X \in C_A{}^p(U)$, thus proving the "if" part. We shall prove the "only if" part by considering two cases.

Case 1: $\mathcal{R}(V_1) \subset \mathcal{R}(U)$. In this situation $X^{\perp\prime} V_1 = 0$ for every $X \in C_A{}^p(U)$ and hence we want to obtain sufficient conditions of V such that $X'L = A' \Rightarrow X^{\perp\prime} VL = 0$ for every $X \in C_A{}^p(U)$. Solving $X'L = A'$ and substituting in $X^{\perp\prime} VL = 0$, we get the result that the condition $X^{\perp\prime}VX'^{+}A' + X^{\perp\prime} VX^\perp Z = 0$ should hold for all Z and for every $X \in C_A{}^p(U) \Leftrightarrow VX^\perp = 0$ for every $X \in C_A{}^p(U) \Leftrightarrow VU^\perp = 0$, which is condition (i) with $\lambda = 0$.

Case 2: $\mathcal{R}(V_1) \not\subset \mathcal{R}(U)$. $X'L = A'$ and $X^{\perp\prime} V_1 L = 0 \Rightarrow X^{\perp\prime} VL = 0$ only if the equations $X'L = A'$, $X^{\perp\prime}V_1L = 0$ and $X^{\perp\prime}VL = 0$ are jointly consistent. A necessary and sufficient condition for this is

$$\mathcal{R}(A: 0: 0)' \subset \mathcal{R}(X: V_1X^\perp: VX^\perp)'$$
$$\Leftrightarrow \mathcal{N}(X: V_1X^\perp: VX^\perp) \subset \mathcal{N}(A: 0: 0) \quad (2.1)$$

We now show that (2.1) is equivalent to

$$\mathcal{R}(X) \cap \mathcal{R}(V_1X^\perp: VX^\perp) = \{0\}. \quad (2.2)$$

That (2.2) $\Rightarrow$ (2.1) is trivial since $\mathcal{N}(X: 0: 0) \subset \mathcal{N}(A: 0: 0)$. Now suppose (2.1) holds and there exists a nonnull vector $x \in \mathcal{R}(X) \cap \mathcal{R}(V_1X^\perp: VX^\perp)$. One can then construct a matrix X_o such that $\mathcal{R}(X_o) = \mathcal{R}(X)$, $\mathcal{R}(A') \subset \mathcal{R}(X_o')$, and if $x = X_o a$, then $Aa \neq 0$ (see Lemma A.2 in the Appendix). This contradicts (2.1); hence (2.1) $\Leftrightarrow$ (2.2) and, applying Lemma A.3, we get $\mathcal{R}(VX^\perp) \subset \mathcal{R}((V_1 + V)X^\perp)$, which must hold for every $X \in C_A{}^p(U)$. In view of Lemma A.1, we see that a necessary and sufficient condition for this is $VU^\perp = \alpha(V_1 + V)U^\perp$ for some $\alpha \geq 0 \Leftrightarrow (1 - \alpha) VU^\perp = \alpha V_1 U^\perp$. Since $\mathcal{R}(V_1) \not\subset \mathcal{R}(U)$, we get $0 \leq \alpha < 1$ and hence $VU^\perp = \lambda V_1 U^\perp$, where $\lambda = \alpha/(1 - \alpha)$. This completes the proof of part (i) of the theorem. Part (ii) of the theorem is a restatement of part (i) of the theorem and the equivalence of (i) and (iii) is easily established.

Corollary 2.1. Every linear representation of the BLUE of $A\beta$ under $(Y, X\beta, V_1)$ is its BLUE under $(Y, X\beta, V)$ for every matrix X of rank p satisfying $\mathcal{R}(A') \subset \mathcal{R}(X')$ if and only if $V = \lambda V_1$.

It is interesting to note that the condition on V stated in Theorem 2.1 does not involve the matrix A. Thus we have another corollary.

Corollary 2.2. Every linear representation of the BLUE of $X\beta$ under $(Y, X\beta, V_1)$ continues to be its BLUE under $(Y, X\beta, V)$ for every $X \in C^p(U)$ if and only if V satisfies any one of the equivalent conditions given in Theorem 2.1.

Corollary 2.3. (Zyskind 1969). Consider the linear model $(Y, X\beta, V)$ where $X \in C^p(U)$. In each of these models all OLS estimators are also corresponding BLUE's if and only if anyone of the following equivalent conditions holds:

1. $VU^\perp = \lambda U^\perp$ for some $\lambda \geq 0$.
2. Every vector in $\mathcal{R}(U^\perp)$ is an eigenvector of V.
3. $V = \lambda I + UBU'$ for some matrix B and scalar $\lambda \geq 0$ such that V is nonnegative definite.

Remark. Putting $U = 1_n = (1, 1, \ldots, 1)'$, the representation given in condition 3 of Corollary 2.3 becomes $V = \lambda I + b\, 1_n 1_n'$, where $\lambda \geq 0$ and b is a scalar such that V is nonnegative definite. This representation for V has been obtained by McElroy (1967) under the assumption of full rank of the design matrix and nonsingularity of the dispersion matrix.

Theorem 2.2.

(i) If $\mathcal{R}(V_1) \subset \mathcal{R}(U)$, then at least one linear representation of the BLUE of $A\beta$ under $(Y, X\beta, V_1)$ is its BLUE under $(Y, X\beta, V)$ for every $X \in C_A{}^p(U)$ and for arbitrary V.
(ii) If $\mathcal{R}(V_1) \not\subset \mathcal{R}(U)$, then at least one linear representation of the BLUE of $A\beta$ under $(Y, X\beta, V_1)$ is its BLUE under $(Y, X\beta, V)$ for every $X \in C_A{}^p(U)$ if and only if every linear representation is so.

Proof. We want conditions on V under which there exists L which satisfies $X'L = A'$, $X^{\perp\prime}V_1L = 0$, and $X^{\perp\prime}VL = 0$ for every $X \in C_A{}^p(U)$. A necessary and sufficient condition for this to hold is

$$\mathcal{R}(A: 0: 0)' \subset \mathcal{R}(X: V_1X^\perp: VX^\perp)' \quad (2.3)$$

for every $X \in C_A{}^p(U)$. If $\mathcal{R}(V_1) \subset \mathcal{R}(U)$, then $X^{\perp\prime}V_1 = 0$ for every $X \in C_A{}^p(U)$. In this case (2.3) simplifies to $\mathcal{R}(A: 0)' \subset \mathcal{R}(X: VX^\perp)'$ and it is easy to see that this condition holds for any nonnegative definite matrix V and for every $X \in C_A{}^p(U)$. This proves part (i) of the theorem, since (2.3) is equivalent to (2.1). The proof of part (ii) follows easily.

Corollary 2.4.

1. If $\mathcal{R}(V_1) \subset \mathcal{R}(U)$, then at least one linear representation of the BLUE of every estimable parametric function under $(Y, X\beta, V_1)$ is its BLUE under $(Y, X\beta, V)$ for every $X \in C^p(U)$ and for arbitrary V.
2. If $\mathcal{R}(V_1) \not\subset \mathcal{R}(U)$, then at least one linear representation of the BLUE of $X\beta$ under $(Y, X\beta, V_1)$ is its BLUE under $(Y, X\beta, V)$ for every $X \in C^p(U)$ if and only if every linear representation is so.

For every $X \in C^p(U)$, let G_X be a least squares g-inverse of $V_1 + XX'$ and consider the linear representation $X(X' G_X X)^- X' G_X Y$ of the BLUE of $X\beta$ under $(Y, X\beta, V_1)$. If $\mathcal{R}(V_1) \not\subset \mathcal{R}(U)$, then it follows from Corollary 2.4 that if $X(X' G_X X)^- X' G_X Y$ is the BLUE of $X\beta$ under $(Y, X\beta, V)$ for every $X \in C^p(U)$, then every linear representation of the BLUE of $X\beta$ under $(Y, X\beta, V_1)$ will also be its BLUE under $(Y, X\beta, V)$, and $VU^\perp = \lambda V_1 U^\perp$ for some $\lambda \geq 0$. Now we assume that $\mathcal{R}(V_1) \subset \mathcal{R}(U)$ and we shall obtain conditions on V so that $X(X' G_X X)^- X' G_X Y$ is the BLUE of $X\beta$ under $(Y, X\beta, V)$ for every $X \in C^p(U)$.

Theorem 2.3. Suppose $\mathcal{R}(V_1) \subset \mathcal{R}(U)$ and let G_X be a least squares g-inverse of $V_1 + XX'$. Then $X(X' G_X X)^- X' G_X Y$ is the BLUE of $X\beta$ under $(Y, X\beta, V)$ for every $X \in C^p(U)$ if and only if $VU^\perp = \lambda U^\perp$ for some $\lambda \geq 0$.

Proof. We want conditions on V so that

$X' G_X V X^\perp = 0$ for every $X \in C^p(U)$

$\Leftrightarrow (V_1 + XX') G_X V X^\perp = 0$ for every $X \in C^p(U)$

(using the fact that $\mathcal{R}(V_1 + XX') = \mathcal{R}(X)$)

$\Leftrightarrow G_X'(V_1 + XX') V X^\perp = 0$ for every $X \in C^p(U)$

(using the fact that $(V_1 + XX')G_X$ is symmetric since G_X is $(V_1 + XX')_l^-$)

$\Leftrightarrow X' V X^\perp = 0$ for every $X \in C^p(U)$

(premultiplying by X')

$\Leftrightarrow \mathcal{R}(VX^\perp) \subset \mathcal{R}(X^\perp)$ for every $X \in C^p(U)$

$\Leftrightarrow VU^\perp = \lambda U^\perp$

for some $\lambda \geq 0$, using Lemma A.1. Theorem 2.3 is thus established.

3. THE VARIANCE COMPONENTS MODEL AND THE COVARIANCE COMPONENTS MODEL

The variance components model is a linear model with $E(Y) = X\beta$ and $V(Y) = \sum_{i=1}^k \sigma_i^2 V_i$. Here the $\sigma_i^2 (i = 1, 2, \ldots, k)$ are unknown parameters and the V_i $(i = 1, 2, \ldots, k)$ are known nonnegative definite matrices. If $V(Y) = \sum_{i=1}^k U_i' W U_i$, where U_i $(i = 1, 2, \ldots, k)$ are known $s \times n$ matrices and $W = ((w_{ij}))$ is an unknown nonnegative definite matrix of order $s \times s$, then the linear model is known as the covariance components model. Mitra and Moore (1973) have shown that for $i \leq j = 1, 2, \ldots, s$, there exist $\tilde{w}_{ij}$ (depending on w_{ij}) satisfying $V(Y) = \sum_{i=1}^k U_i' W U_i = \sum_{i \leq j} \tilde{w}_{ij} V_{ij}$, where the V_{ij}'s are known nonnegative definite matrices for $i \leq j = 1, 2, \ldots, s$.

In this section, we obtain the conditions under which $A\beta$ admits a BLUE under the variance components model or the covariance components model for every $X \in C_A{}^p(U)$. For a given linear model, specifically for a fixed X, this problem has been considered for the variance components model by Seely and Zyskind (1971) and Mitra and Moore (1973, 1976) and for the covariance components model by Mitra and Moore (1973). We now prove the following theorem.

Theorem 3.1.

(i) $A\beta$ admits a BLUE under the model $(Y, X\beta, \sum_{i=1}^k \sigma_i^2 V_i)$ for every $X \in C_A{}^p(U)$ if and only if for $i = 1, 2, \ldots, k$, $V_i U^\perp = \lambda_i V_o U^\perp$ for some $\lambda_i \geq 0$, where $V_o = \sum_{i=1}^k V_i$.

(ii) For $i \leq j = 1, 2, \ldots, s$, let V_{ij} be as given in the beginning of Section 3. Then $A\beta$ admits a BLUE under the model $(Y, X\beta, \sum_{i=1}^k U_i' W U_i)$ for every $X \in C_A{}^p(U)$ if and only if for each $i \leq j = 1, 2, \ldots, s$ $V_{ij} U^\perp = \lambda_{ij} V_o U^\perp$ for some $\lambda_{ij} \geq 0$, where $V_o = \sum_{i \leq j} V_{ij}$.

(iii) $X\beta$ admits a BLUE under the model $(Y, X\beta, \sum_{i=1}^k \sigma_i^2 V_i)$ (or under $(Y, X\beta, \sum_{i=1}^k U_i' W U_i)$) for every $X \in C^p(U)$ if and only if the condition stated in (i) (or (ii), respectively) holds.

Proof. $A\beta$ admits a BLUE under $(Y, X\beta, \sum_{i=1}^k \sigma_i^2 V_i)$ for every $X \in C_A{}^p(U)$ if and only if for every $X \in C_A{}^p(U)$ there exists a matrix L satisfying $X' L = A'$ and $X^{\perp\prime} V_i L = 0$ for $i = 1, 2, \ldots, k$ or, equivalently,

$$\mathcal{R}(A : 0 : 0, \ldots, 0)' \subset \mathcal{R}(X : V_1 X^\perp : V_2 X^\perp : \cdots : V_k X^\perp)' \quad (3.1)$$

for every $X \in C_A{}^p(U)$. As in the proof of Theorem 2.1, it can be shown that (3.1) is equivalent to the condition $\mathcal{R}(V_i X^\perp) \subset \mathcal{R}(V_o X^\perp)$, $i = 1, 2, \ldots, k$ for every $X \in C_A{}^p(U)$, where $V_o = \sum_{i=1}^k V_i$. Part (i) of Theorem 3.1 now follows from Lemma A.1. Part (ii) is proved similarly. The proof of part (iii) is clear in view of the results in Section 2.

Remark 3.1. When $A\beta$ admits a BLUE under the variance components model $(Y, X\beta, \sum_{i=1}^k \sigma_i^2 V_i)$, its BLUE could be computed as $A(X' G_o X)^- X' G_o Y$, G_o being a g-inverse of $V_o + XX'$, where $V_o = \sum_{i=1}^k V_i$.

Remark 3.2. Suppose V is any nonnegative definite matrix in the k-dimensional linear space $\mathcal{L}(V_1, V_2, \ldots, V_k)$ spanned by the nonnegative definite matrices $V_1, V_2, \ldots, V_k$. Then V can be written as $V = \sum_{i=1}^k a_i V_i$, where all the a_i's may not be nonnegative. Using arguments similar to those given in the proof of part (i) of Theorem 3.1 (i), it can be shown that whenever $X \in C_A{}^p(U)$, $A\beta$ admits a BLUE under $(Y, X\beta, V)$ for every nonnegative definite matrix $V \in \mathcal{L}(V_1, V_2, \ldots, V_k)$ if and only if $V_i U^\perp = \lambda_i V_o U^\perp$ for some $\lambda_i \geq 0$ and for $i = 1, 2, \ldots, k$, where $V_o = \sum_{i=1}^k V_i$. When this condition is satisfied, the estimator $A(X' G_o X)^- X' G_o Y$ is a BLUE of $A\beta$, where G_o is a g inverse of $V_o + XX'$. Mitra and Moore (1976) have pointed out that even if a linear parametric function $l' X\beta$ does not admit a BLUE under $(Y, X\beta, V)$, V being any nonnegative definite matrix in $\mathcal{L}(V_1, V_2, \ldots, V_k)$, the estimator $l' X (X' G_o X)^- X' G_o Y$ is an admissible estimator of $l' X \beta$; that is, no linear unbiased estimator of $l' X\beta$ can be better than $l' X (X' G_o X)^- X' G_o Y$.

As an example consider the matrix $V_{a,b} = aI + b\, 1_n 1_n'$, where 1_n is the column vector with each element unity and a and b any two real numbers that make $V_{a,b}$ a nonnegative definite matrix. If $U = 1_n$, $V_1 = I$, and $V_2 = 1_n 1_n'$, then it is clear that $V_1 U^{\perp} = (V_1 + V_2) U^{\perp}$ and $V_2 U^{\perp} = 0$. Hence in view of Remark 3.2, we see that $X\beta$ admits a BLUE under the model $(Y, X\beta, V_{a,b})$ for every $X \in C^p(1_n)$ and the BLUE can be computed as $X(X'(I + 1_n 1_n')^{-1}X)^{-}X'(I + 1_n 1_n')^{-1}Y$.

Remark 3.3. If $U = 0$, then from Theorem 3.1 it is clear that $A\beta$ $(A \neq 0)$ admits a BLUE under the variance components model or the covariance components model for every X of a particular rank $p < n$ satisfying $\mathcal{R}(A') \subset \mathcal{R}(X')$ if and only if the dispersion of Y is known either completely or up to a positive scalar multiplier.

APPENDIX: SOME ALGEBRAIC RESULTS

Lemma A.1. If V_1 and V are nonnegative definite matrices, then $\mathcal{R}(V X^{\perp}) \subset \mathcal{R}(V_1 X^{\perp})$ for all $X \in C_A{}^p(U)$ if and only if $VU^{\perp} = \lambda V_1 U^{\perp}$ for some $\lambda \geq 0$.

Proof. Theorem 1 in Zyskind (1969) states that $\mathcal{R}(VX^{\perp}) \subset \mathcal{R}(V_1X^{\perp})$ if and only if $w' V_1X^{\perp} = 0 \Rightarrow w' VX^{\perp} = 0$. This is equivalent to demanding that $V_1 w \in \mathcal{R}(X) \Rightarrow V w \in \mathcal{R}(X)$; we therefore want conditions on V such that $V_1 w \in \mathcal{R}(X) \Rightarrow V w \in \mathcal{R}(X)$ for every $X \in C_A{}^p(U)$. Since $p < n$, if $V_1 w \in \mathcal{R}(U)$ and $V w \notin \mathcal{R}(U)$, we can choose $X \in C_A{}^p(U)$ such that $V_1 w \in \mathcal{R}(X)$ and $V w \notin \mathcal{R}(X)$. Thus we necessarily have

$$\mathcal{R}(VU^{\perp}) \subset \mathcal{R}(V_1 U^{\perp}) \tag{A.1}$$

Now let $w_1 = U^{\perp}z$. Let the matrix C be such that $\mathcal{R}(C) = \mathcal{R}(U : V_1w_1 : V w_1)^{\perp}$, and consider the matrix X satisfying $\mathcal{R}(A') \subset \mathcal{R}(X')$ and $\mathcal{R}(X) = \mathcal{R}(U : V_1w_1 : C_1)$, where C_1 is chosen in such a way that $\mathcal{R}(C_1) \subset \mathcal{R}(C)$ and $r(X) = p$. For such an X, $V_1w_1 \in \mathcal{R}(X)$, and hence we should have $V w_1 \in \mathcal{R}(X)$. Thus $V w_1 = U\alpha_1 + \lambda V_1 w_1 + C_1\alpha_2$ for some vectors α_1 and α_2 and scalar λ. From the choice of C_1 it then follows that $V w_1 = U\alpha_1 + \lambda V_1 w_1$. However, from (A.1), $V w_1 = VU^{\perp}z \in \mathcal{R}(V_1U^{\perp})$ and $\mathcal{R}(V_1U^{\perp}) \cap \mathcal{R}(U) = \{0\}$; therefore $V w_1 = \lambda V_1 w_1$. Thus $V U^{\perp}z = \lambda V_1 U^{\perp}z$ for every z and hence $VU^{\perp} = \lambda V_1 U^{\perp}$. Since premultiplication by $U^{\perp\prime}$ leaves nonnegative definite matrices on both sides, we should have $\lambda \geq 0$. This completes the proof of Lemma A.1.

Remark A.1. If V_1 and V are nonnegative definite matrices, then $\mathcal{R}(V X^{\perp}) \subset \mathcal{R}(V_1 X^{\perp})$ for all $X \in C^p(U)$ if and only if $VU^{\perp} = \lambda V_1 U^{\perp}$, for some $\lambda \geq 0$.

Lemma A.2. Let A and X be given matrices with $r(A) \leq r(X)$ and let x be a non-null vector in $\mathcal{R}(X)$. Then there exists a matrix X_o satisfying $\mathcal{R}(X_o) = \mathcal{R}(X)$, $\mathcal{R}(A') \subset \mathcal{R}(X_o')$, and if $x = X_o a$, then $A a \neq 0$.

Proof. Let the columns of $(x : S)$ be a basis of $\mathcal{R}(X)$, the rows of T_1 be a basis of the row space of A, and $T = (T_1' : T_2')'$ be a matrix of linearly independent rows. For some j, let the jth column of T be $(1, 0, 0, \ldots, 0)'$, which is possible since the rows of T are non-null. Put $X_o = (x : S)T$. Then the jth column of X_o is x, or, in other words if a is the jth column of the identity matrix, then $X_o a = x$. However, $T_1 a = (1, 0, \ldots, 0)' \neq 0$. Hence $A a \neq 0$.

Lemma A.3. Let V_1 and V be $n \times n$ nonnegative definite matrices and X be an $n \times m$ matrix. Then the following conditions are equivalent.

(i) $\mathcal{R}(X : 0 : 0)' \subset \mathcal{R}(X : V_1X^{\perp} : V X^{\perp})'$.
(ii) $\mathcal{N}(X : V_1X^{\perp} : V X^{\perp}) \subset \mathcal{N}(X : 0 : 0)$.
(iii) $\mathcal{R}(X) \cap \mathcal{R}(V_1 X^{\perp} : V X^{\perp}) = \{0\}$.
(iv) $\mathcal{R}(V_1 X^{\perp} : V X^{\perp}) = \mathcal{R}((V_1 + V) X^{\perp})$.
(v) $\mathcal{R}(V X^{\perp}) \subset \mathcal{R}((V_1 + V) X^{\perp})$.

Proof. Equivalence of (i), (ii) with (iii), and (iv) with (v) is fairly straightforward. To show that (iv) $\Rightarrow$ (iii) we observe that on account of Theorems 2.1 and 2.3 of Mitra and Puri (1979), $(V + V_1)X^{\perp} = \{(V + V_1) - \mathcal{S}(V + V_1)\} X^{\perp}$, where $\mathcal{S}(V + V_1)$ is the shorted nonnegative definite matrix $(V + V_1)$ and $\mathcal{S} = \mathcal{R}(X)$. Furthermore,

$$\mathcal{R}\{(V + V_1) - \mathcal{S}(V + V_1)\} \cap \mathcal{R}(X) = \{0\}.$$

Hence (iv) $\Rightarrow$ (iii) since $\mathcal{R}(VX^{\perp} : V_1 X^{\perp}) \subset \mathcal{R}(V + V_1)$ and $\mathcal{R}((V + V_1) X^{\perp}) \subset \mathcal{R}(V X^{\perp} : V_1 X^{\perp})$; if (iv) is not true, there exists a non-null vector in $\mathcal{R}(V X^{\perp} : V_1 X^{\perp})$ that is also in $\mathcal{R}(\mathcal{S}(V + V_1))$. This contradicts (iii) since $\mathcal{R}(\mathcal{S}(V + V_1)) \subset \mathcal{S} = \mathcal{R}(X)$. Hence (iii) and (iv) are equivalent.

[*Received October 1981. Revised August 1982.*]

REFERENCES

HABERMAN, S.J. (1975), "How Much Do Gauss-Markov and Least Squares Estimators Differ? A Coordinate Free Approach," *The Annals of Statistics*, 3, 982–990.

KRUSKAL, W. (1968), "When Are Gauss-Markov and Least Squares Estimators Identical? A Coordinate Free Approach," *The Annals of Mathematical Statistics*, 39, 70–75.

McELROY, F.W. (1967), "A Necessary and Sufficient Condition That Ordinary Least Squares Estimators Are Best Linear Unbiased," *Journal of the American Statistical Association*, 62, 1302–1304.

MITRA, S.K., and MOORE, B.J. (1973), "Gauss-Markov Estimation With an Incorrect Dispersion Matrix," *Sankhya*, A, 35, 139–152.

——— (1976), "Linear Estimation in General Linear Models," in *Essays in Probability and Statistics*, eds. S. Ikeda et al., Tokyo: Shinko Tsuho Co., 189–201.

MITRA, S.K., and PURI, M.L. (1979), "Shorted Operators and Generalized Inverses of Matrices," *Linear Algebra and Its Applications*, 25, 45–56.

RAO, C.R. (1967), "Least Squares Theory Using an Estimated Dispersion Matrix and Its Application to Measurement of Signals," *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, 1, 355–372.

RAO, C.R., and MITRA, S.K. (1971), *Generalized Inverse of Matrices and Its Applications*, New York: John Wiley.

SEELEY, J., and ZYSKIND, G. (1971), "Linear Spaces and Minimum Variance Unbiased Estimation," *The Annals of Mathematical Statistics*, 42, 691–703.

WATSON, G.S. (1967), "Linear Least Square Regression," *The Annals of Mathematical Statistics*, 38, 1679–1699.

ZYSKIND, G. (1967), "On Canonical Forms, Nonnegative Covariance Matrices, and Best and Simple Least Squares Estimators in Linear Models," *The Annals of Mathematical Statistics*, 38, 1092–1110.

——— (1969), "Parametric Augmentations and Error Structures Under Which Certain Least Squares and Analysis of Variance Procedures Are Also Best," *Journal of the American Statistical Association*, 64, 1353–1368.