The Annals of Statistics
1981, Vol. 9, No. 3, 633–637

UNBIASED AND MINIMUM-VARIANCE UNBIASED ESTIMATION OF ESTIMABLE FUNCTIONS FOR FIXED LINEAR MODELS WITH ARBITRARY COVARIANCE STRUCTURE[1]

BY DAVID A. HARVILLE

Iowa State University

Consider a general linear model for a column vector y of data having $E(y) = X\alpha$ and $\text{Var}(y) = \sigma^2 H$, where α is a vector of unknown parameters and X and H are given matrices that are possibly deficient in rank. Let $b = Ty$, where T is any matrix of maximum rank such that $TH = \phi$. The estimation of a linear function of α by functions of the form $c + a'y$, where c and a are permitted to depend on b, is investigated. Allowing c and a to depend on b expands the class of unbiased estimators in a nontrivial way; however, it does not add to the class of linear functions of α that are estimable. Any minimum-variance unbiased estimator is identically [for y in the column space of (X, H)] equal to the estimator that has minimum variance among strictly linear unbiased estimators.

1. Introduction. Take y to be an $n \times 1$ data vector having $E(y) = X\alpha$ and $\text{Var}(y) = \sigma^2 H$, where α is a $p \times 1$ parameter vector, σ^2 is a known or unknown scalar, and X and H are known matrices. The parameter space is $\{\alpha : \alpha \in R^p\}$ or $\{(\alpha, \sigma^2) : \alpha \in R^p, \sigma^2 > 0\}$, depending on whether σ^2 is known or unknown. No assumptions are made about rank (X) or rank (H). This setup is referred to as the *general Gauss-Markov model.*

Take T to be any matrix satisfying $TH = \phi$ whose rank is a maximum, equal to $n - \text{rank}(H)$, and let $b = Ty$ and $A = TX$. It can be shown that $\text{rank}(A) = \text{rank}(X, H) - \text{rank}(H)$. Further, let $L = I - A^- A$, or take L to be any other matrix satisfying $AL = \phi$ whose rank equals $p + \text{rank}(H) - \text{rank}(X, H)$. (For any matrix B, B^- will denote an arbitrary generalized inverse of B, i.e., any solution to $BB^-B = B$, and $\mathscr{C}(B)$ will denote the column space of B.)

We have that $\text{Var}(b) = \phi$, implying that

$$A\alpha = b \quad \text{with probability 1.}$$

Once y is observed, we know the right side, as well as the coefficient matrix, of the equation $A\alpha = b$, which "must" be satisfied by α. Based on this fact, several writers, e.g., Zyskind and Martin (1969), Rao (1972), and Kempthorne (1976), have asserted that, to ensure that a function of the form $c + a'y$ estimate unbiasedly a linear parametric function $\lambda'\alpha$, we need only require that $E(c + a'y) = \lambda'\alpha$ for those α for which $A\alpha = b$ rather than for all α. This is the same as requiring that

$$(1.1) \qquad c = (\lambda' - a'X)A^- b \qquad \text{and} \qquad (\lambda' - a'X)L = \phi \qquad \text{whenever} \quad b \in \mathscr{C}(A).$$

[Actually, these writers confine themselves to "homogeneous" functions $a'y$, which is equivalent to superimposing the condition $c = 0$ on conditions (1.1).]

Conditions (1.1) differ from the seemingly more restrictive conditions

Received May, 1978; revised August, 1979.

[1] This research was supported by the Air Force Office of Scientific Research, Air Force Systems Command, USAF, under Grant No. AFOSR 76-3037.

AMS 1970 *subject classification.* 62J05.

Key words and phrases. Linear models, best linear unbiased estimation, singular covariance matrices, Gauss-Markov theorem.

(1.2) $$c = 0 \quad \text{and} \quad \lambda' - a'X = \phi$$

claimed by various other writers, e.g., Albert (1972), to be necessary and sufficient for $c + a'y$ to estimate unbiasedly $\lambda'\alpha$. This apparent contradiction can be resolved by observing that it is implicit in the approach that leads to conditions (1.1) that c and a be allowed to be functionally dependent on y (through b), while conditions (1.2) are obtained by specifying that c and a be functionally independent of y (and of α).

The purpose of the present article is to derive necessary and sufficient conditions for the unbiasedness and minimum-variance unbiasedness of estimators of $\lambda'\alpha$ having the form $c(b) + [a(b)]'y$, where $c(\cdot)$ is a function and $a(\cdot)$ is a vector-valued function.[2] The form $c(b) + [a(b)]'y$ is (in general) nonlinear in y, however

(1.3) $$c(b) + [a(b)]'y = c(A\alpha) + [a(A\alpha)]'y \quad \text{with probability 1.}$$

2. Unbiased representations. Using result (1.3), we find that

$$E\{c(b) + [a(b)]'y\} = E\{c(A\alpha) + [a(A\alpha)]'y\} = c(A\alpha) + [a(A\alpha)]'X\alpha$$

and thus that the condition

$$E\{c(b) + [a(b)]'y\} = \lambda'\alpha \quad \text{for all} \quad \alpha \in R^p$$

is equivalent to the condition

$$c(\gamma) + [a(\gamma)]'X\alpha = \lambda'\alpha \quad \text{for all } \alpha \text{ such that} \quad A\alpha = \gamma \quad \text{and all} \quad \gamma \in \mathscr{C}(A).$$

Now, applying standard results on constrained linear estimation, see, e.g., Rao 1973, Section 4a.9, we obtain the following theorem.

THEOREM 1. *Under the general Gauss-Markov model, $c(b) + [a(b)]'y$ is an unbiased estimator of $\lambda'\alpha$ if and only if*

(2.1) $$c(b) = [\lambda - X'a(b)]'A^-b \quad \textit{and} \quad [\lambda - X'a(b)]'L = \phi \quad \textit{for all} \quad b \in \mathscr{C}(A)$$

or, equivalently, if and only if, for some vector-valued function $k(\cdot)$,

$$c(b) = [k(b)]'b \quad \textit{and} \quad \lambda - X'a(b) = A'k(b) \quad \textit{for all} \quad b \in \mathscr{C}(A).$$

Note that conditions (2.1) are essentially the same as the unbiasedness conditions (1.1) obtained by acting as though b were a vector of constants and α were subject to constraints $A\alpha = b$. Note also that, if $c(\cdot)$ and $a(\cdot)$ satisfy conditions (2.1), then, defining

$$c_1(b) = c(b) - [\lambda - X'a(b)]'A^-b \quad \text{and} \quad a_1(b) = a(b) + T'(A^-)'[\lambda - X'a(b)],$$

we have that

$$c_1(b) + [a_1(b)]'y = c(b) + [a(b)]'y \quad \text{for all} \quad y \in R^n$$

and that

$$c_1(b) = 0 \quad \text{and} \quad \lambda' - [a_1(b)]'X = \phi \quad \text{for all} \quad b \in \mathscr{C}(A).$$

Thus, if an estimator has a representation of the form $c(b) + [a(b)]'y$ which satisfies conditions (2.1), it has a second representation of the same form which satisfies the conditions

(2.2) $$c(b) = 0 \quad \text{and} \quad \lambda' - [a(b)]'X = \phi \quad \text{for all} \quad b \in \mathscr{C}(A).$$

Conditions (2.2) resemble the unbiasedness conditions (1.2) obtained when attention is

[2] This class of estimators was considered also by Rao (1979), in a paper submitted for publication subsequent to the present paper.

restricted to functions $c + a'y$ where c and a do not depend on y (or α). If we were to restrict attention to forms $c(b) + [a(b)]'y$ that satisfy conditions (2.2) rather than the more general conditions (2.1), no estimators would be lost, however some representations of estimators having multiple representations would be sacrificed.

We say that $\lambda'\alpha$ is estimable if there exists a function having the form $c(b) + [a(b)]'y$ that is an unbiased estimator of $\lambda'\alpha$. Clearly, $\lambda'\alpha$ is estimable if and only if $\lambda \in \mathscr{C}(X')$. Thus, allowing c and a in the form $c + a'y$ to depend on b does not expand the class of estimable parametric functions.

3. Minimum-variance unbiased representations. To derive conditions that are necessary and sufficient for the form $c(b) + [a(b)]'y$ to comprise a minimum-variance unbiased estimator of an estimable function $\lambda'\alpha$, we apply the following result.

THEOREM 2. *Suppose that V is a symmetric nonnegative definite matrix, that K is any matrix, and that μ is a column vector such that $\mu \in \mathscr{C}(K')$.*

(i) *The linear system*

$$\begin{pmatrix} V & K \\ K' & \phi \end{pmatrix} \begin{pmatrix} a \\ \rho \end{pmatrix} = \begin{pmatrix} \phi \\ \mu \end{pmatrix} \tag{3.1}$$

is consistent, i.e., *has a solution for a and ρ.*

(ii) *A necessary and sufficient condition for a to minimize the quadratic form $a'Va$, subject to the linear constraints $K'a = \mu$, is that a comprise the first component of some solution to system* (3.1).

(iii) *Every vector a that comprises the first component of some solution to system* (3.1) *is generated from a particular solution a^* by putting $a = a^* - d$ and letting d range over all solutions to $[V, K]'d = \phi$.*

The results that make up Theorem 2 were reported in a series of papers by C. R. Rao. Simple proofs of parts (i) and (ii) can be found in Kempthorne's (1976) article. Part (iii) can be established by noting that, if d and τ satisfy

$$\begin{pmatrix} V & K \\ K' & \phi \end{pmatrix} \begin{pmatrix} d \\ \tau \end{pmatrix} = \begin{pmatrix} \phi \\ \phi \end{pmatrix},$$

then $Vd = -K\tau$, implying that $d'Vd = -d'K\tau = \phi$ and thus that $Vd = \phi$.

We find that

$$\operatorname{Var}\{c(b) + [a(b)]'y\} = \sigma^2[a(A\alpha)]'Ha(A\alpha).$$

Thus, among representations of the form $c(b) + [a(b)]'y$, a representation is a uniformly minimum-variance unbiased estimator of an estimable function $\lambda'\alpha$ if and only if $c(\cdot)$ and $a(\cdot)$ are such that, for each $b \in \mathscr{C}(A)$, $c(b) = [\lambda - X'a(b)]'A^-b$ and $a(b)$ minimizes $[a(b)]'Ha(b)$ subject to the constraint $[\lambda - X'a(b)]'L = \phi$. Applying parts (i) and (ii) of Theorem 2, we obtain the following result.

THEOREM 3. *Assume the general Gauss-Markov model, and suppose that $\lambda'\alpha$ is estimable. The linear system*

$$\begin{pmatrix} H & XL \\ L'X' & \phi \end{pmatrix} \begin{pmatrix} a \\ \rho \end{pmatrix} = \begin{pmatrix} \phi \\ L'\lambda \end{pmatrix} \tag{3.2}$$

is consistent, i.e., *has one or more solutions for a and ρ; and, among representations of the form $c(b) + [a(b)]'y$, a representation is a uniformly minimum-variance unbiased estimator of $\lambda'\alpha$ if and only if, for each $b \in \mathscr{C}(A)$, $c(b) = [\lambda - X'a(b)]'A^-b$ and $a(b)$ comprises the first component of some solution to the system* (3.2).

In deriving estimators having uniformly minimum variance, among estimators with representations of the form $c(b) + [a(b)]'y$ that are unbiased for $\lambda'\alpha$ [i.e., representations satisfying conditions (2.1)], we could have restricted our search to representations satisfying conditions (2.2). It follows from the results of Section 2 that we would have obtained exactly the same minimum-variance unbiased estimators as before, though certain representations of those estimators having multiple representations would have been lost. Applying Theorem 2, we have that, among representations of the form $c(b) + [a(b)]'y$ that satisfy conditions (2.2), the representations that have uniformly minimum variance are those such that, for each $b \in \mathscr{C}(A)$, $c(b) = 0$ and $a(b)$ comprises the first component of some solution to the necessarily consistent linear system

$$\begin{pmatrix} H & X \\ X' & \phi \end{pmatrix} \begin{pmatrix} a \\ \rho \end{pmatrix} = \begin{pmatrix} \phi \\ \lambda \end{pmatrix}. \tag{3.3}$$

Theorem 3 characterizes representations having minimum variance among representations of the form $c(b) + [a(b)]'y$ that are unbiased for $\lambda'\alpha$. Part (iii) of Theorem 2 can be used to obtain an alternative characterization. Clearly, a vector d satisfies the condition $(H, XL)'d = \phi$ if and only if $d = T's$ for some vector s. Thus, taking a^* to be the first component of any solution to system (3.3) [in which case a^* is also the first component of some solution to system (3.2)], we find that, among representations of the form $c(b) + [a(b)]'y$ that are unbiased for $\lambda'\alpha$, the representations having minimum variance are those for which there exists a vector-valued function $s(\cdot)$ such that

$$c(b) = [s(b)]'b \quad \text{and} \quad a(b) = a^* - T's(b) \quad \text{for all} \quad b \in \mathscr{C}(A). \tag{3.4}$$

It follows from the characterization (3.4) that, if $t_1(y)$ and $t_2(y)$ are any two of the minimum-variance unbiased estimators of $\lambda'\alpha$ of the form $c(b) + [a(b)]'y$, then $t_1(y) = t_2(y)$ for all y for which $b \in \mathscr{C}(A)$ or equivalently [since, for any vector r, $Ar = b$ if and only if $y - Xr \in \mathscr{C}(H)$] for all $y \in \mathscr{C}(X, H)$.

It can be shown that the rank of the coefficient matrix of system (3.2) equals rank(H) + rank(XL) and that rank (XL) = rank(X) + rank(H) − rank(X, H).

4. Example. Let $y = (y_1, y_2, y_3, y_4)'$ and $\alpha = (\alpha_1, \alpha_2)'$, and put

$$X = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

We can take

$$T = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \text{in which case} \quad A = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad b = \begin{pmatrix} y_2 \\ y_3 \\ y_4 \end{pmatrix}.$$

A generalized inverse of A is

$$A^- = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

We can choose $L = I - A^-A = \phi$.

Consider the estimation of $\lambda'\alpha$ with $\lambda' = (1, 0)$, i.e., of α_1. The function $t(y) = y_1 - y_1y_3 - y_1y_4 + y_3^2 + y_3y_4$ can be written as $c(b) + [a(b)]'y$ with

$$c(b) = 0 \quad \text{and} \quad [a(b)]' = (1 - y_3 - y_4, 0, y_3, y_3). \tag{4.1}$$

It is easy to verify that this $c(\cdot)$ and $a(\cdot)$ satisfy conditions (2.1) and thus that $t(y)$ is an unbiased estimator of α_1.

The functions $c(\cdot)$ and $a(\cdot)$ as given by (4.1) do not satisfy conditions (2.2); however the estimator $t(y)$ can also be written as $c_1(b) + [a_1(b)]'y$ with

$$c_1(b) = y_4(y_3 - y_2) \qquad \text{and} \qquad [a_1(b)]' = (1 - y_3 - y_4, y_4, y_3, 0),$$

and $c_1(\cdot)$ and $a_1(\cdot)$ do satisfy conditions (2.2).

The general form for first components of solutions to system (3.2) is $a = (0, k_1, k_2, k_3)'$, where k_1, k_2 and k_3 are arbitrary. Thus, among representations of the form $c(b) + [a(b)]'y$, the representations that are uniformly minimum-variance unbiased estimators of $\lambda'\alpha$ are those such that, for all b having $y_2 = y_3$,

$$[a(b)]' = [0, k_1(b), k_2(b), k_3(b)] \qquad \text{and} \qquad c(b) = [1 - k_1(b) - k_2(b)]y_2 - k_3(b)y_4, \tag{4.2}$$

where $k_1(\cdot)$, $k_2(\cdot)$, and $k_3(\cdot)$ are arbitrary functions. Making the substitution (4.2), we get

$$c(b) + [a(b)]'y = y_2 + (y_3 - y_2)k_2(b),$$

which reduces to y_2 for $y \in \mathscr{C}(X, H)$.

Acknowledgment. The author is indebted to the referees and the Associate Editor for several valuable suggestions.

REFERENCES

ALBERT, A. (1972). *Regression and the Moore-Penrose Pseudoinverse.* Academic, New York.

KEMPTHORNE, O. (1976). Best linear unbiased estimation with arbitrary variance matrix. In *Essays in Probability and Statistics.* (S. Ikeda et al., eds.). Shinko Tsusho, Tokyo

RAO, C. R. (1972). Some recent results in linear estimation. *Sankhyā Ser. B* **34** 369–378.

RAO, C. R. (1973). *Linear Statistical Inference and Its Applications*, 2nd ed. Wiley, New York.

RAO, C. R. (1979). Estimation of parameters in the singular Gauss-Markoff model. *Commun. Statist.* A**8** 1353–1358.

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.

DEPARTMENT OF STATISTICS
IOWA STATE UNIVERSITY
SNEDECOR HALL
AMES, IOWA 50011