The Annals of Mathematical Statistics
1971, Vol. 42, No. 5, 1588–1594

NEW CRITERIA FOR ESTIMABILITY FOR LINEAR MODELS[1,2]

BY GEORGE A. MILLIKEN
Kansas State University

A new criterion for determining the estimability of linear combinations of the parameters of a linear model is established. The result consists of evaluating the trace of a matrix and thus only one number must be checked to determine estimability. The sums of squares necessary to test hypotheses about estimable linear combinations are also derived. Finally, a stepwise computational procedure to compute generalized inverses and matrix products involving generalized inverses is presented. Using the theory and computational techniques, a computer program can be developed to provide a complete analysis of the linear model using generalized inverses.

1. Introduction. Let

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta}+\mathbf{e} \tag{1.1}$$

be any linear model (LM) where $\mathbf{y}$ is a $n\times 1$ random observation vector, $\mathbf{X}$ is a $n\times p$ matrix of known constants (design matrix) of rank q ($q < n$), $\boldsymbol{\beta}$ is a $p\times 1$ vector of unknown parameters defined in E_p, and $\mathbf{e}$ is an unobserved random normal vector with mean $\mathbf{0}$ and covariance matrix $\sigma^2\mathbf{I}$ where σ^2 is positive and unknown. Linear combinations of the parameters, say $\mathbf{A}\boldsymbol{\beta}$, are defined to be estimable if the rows of the matrix $\mathbf{A}$ belong to the vector space spanned by the rows of the matrix $\mathbf{X}$ (Bose (1949)). The importance of the idea of estimability is that there exist unique best linear unbiased estimates (BLUE) of linear combinations of the parameters if the linear combinations are estimable. Generally the estimability condition is difficult to check; however, Searle (1966) states the condition as: the linear combinations $\mathbf{A}\boldsymbol{\beta}$ are estimable if and only if $\mathbf{A}(\mathbf{X}'\mathbf{X})^c\mathbf{X}'\mathbf{X} = \mathbf{A}$, where $(\mathbf{X}'\mathbf{X})^c$ is any matrix satisfying the matrix equation $\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^c\mathbf{X}'\mathbf{X} = \mathbf{X}'\mathbf{X}$ (the matrix $\mathbf{B}^c$ is called a conditional inverse of the matrix $\mathbf{B}$). To determine whether the rows of $\mathbf{A}$ belong to the row space of $\mathbf{X}$, one must check each element of the matrix product against the corresponding element of the matrix $\mathbf{A}$. There are several other references which have considered the problem of estimability of linear combinations of the parameters of a linear model. (Bose (1949); Chipman (1964); Graybill (1961); Rao (1962); Searle (1966)).

The results herein restate the estimability condition in terms of the trace of a matrix product, i.e., only one number needs to be computed and checked in order to establish the estimability of the linear combinations $\mathbf{A}\boldsymbol{\beta}$ as opposed to checking each element of the matrix product for Searle's procedure. The matrix product

Received March 25, 1970; revised February 11, 1971.

[1] This investigation was supported in part by PHS Training Grant No. 5 T01GM00725. A part of this paper is a portion of a dissertation presented in partial fulfillment of the requirements for the Ph.D. degree in Statistics at Colorado State University.

[2] Contribution No. 138 of the Department of Statistics and Computer Science, Kansas Agricultural Experiment Station.

considered involves generalized inverses of matrices (Penrose (1955)); $\mathbf{B}^-$ denotes the generalized inverse of the matrix $\mathbf{B}$ if $\mathbf{B}^-$ satisfies $\mathbf{BB}^-\mathbf{B} = \mathbf{B}$, $\mathbf{B}^-\mathbf{BB}^- = \mathbf{B}^-$, $(\mathbf{B}^-\mathbf{B})' = \mathbf{B}^-\mathbf{B}$ and $(\mathbf{BB}^-)' = \mathbf{BB}^-$.

The main results are presented in Section 2. Section 3 provides procedures for testing hypotheses about estimable linear combinations, and a stepwise procedure for computing the matrix products $\mathbf{B}^-\mathbf{B}$ and $\mathbf{BB}^-$ as well as $\mathbf{B}^-$ is developed in Section 4. The structure of these results enables the linear model to be easily analyzed via a computer.

2. Estimability.

The theorems to be proved concern the estimability of certain linear combinations of the parameters of the linear model (1.1). The normal equations are

$$\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{y} \tag{2.1}$$

and a general solution, $\hat{\boldsymbol{\beta}}$, to the normal equations is (Graybill (1969))

$$\hat{\boldsymbol{\beta}} = \mathbf{X}^-\mathbf{y} + (\mathbf{I} - \mathbf{X}^-\mathbf{X})\mathbf{h} \tag{2.2}$$

where $\mathbf{h}$ is any $p \times 1$ vector in E_p. If $\mathbf{A}\boldsymbol{\beta}$ are estimable linear combinations, it is known (Bose (1949)) that the BLUE's of the set are $\mathbf{A}\hat{\boldsymbol{\beta}}$ where $\hat{\boldsymbol{\beta}}$ is any solution as (2.2). The estimability condition can be stated in terms of the rank of the matrix product $\mathbf{X}(\mathbf{I} - \mathbf{A}^-\mathbf{A})$.

THEOREM 2.1. *For the* LM, *where the rank of* $\mathbf{X}$ *is* q, *the linear combinations* $\mathbf{A}\boldsymbol{\beta}$ *are estimable, where* $\mathbf{A}$ *is a* $k \times p$ *matrix of rank* k, *if and only if the rank of* $\mathbf{X}(\mathbf{I} - \mathbf{A}^-\mathbf{A})$ *is* $q - k$.

PROOF. Assume that the linear combinations $\mathbf{A}\boldsymbol{\beta}$ are estimable. By definition, there exists a $k \times n$ matrix $\mathbf{C}_0$ of rank k such that $\mathbf{A} = \mathbf{C}_0\mathbf{X}$. Let $\mathbf{X}_1$ be a $q \times p$ row basis of the matrix $\mathbf{X}$. Then there exists a $k \times q$ matrix $\mathbf{C}$ of rank k such that $\mathbf{A} = \mathbf{C}\mathbf{X}_1$. Let $\rho(\mathbf{B})$ denote the rank of the matrix $\mathbf{B}$. Then

$$\begin{aligned} \rho[\mathbf{X}(\mathbf{I} - \mathbf{A}^-\mathbf{A})] &= \rho[\mathbf{X}_1(\mathbf{I} - \mathbf{A}^-\mathbf{A})] \\ &= \rho[\mathbf{X}_1(\mathbf{I} - \mathbf{A}^-\mathbf{A})\mathbf{X}_1'] \\ &= \rho[\mathbf{X}_1(\mathbf{I} - \mathbf{X}_1'\mathbf{C}'(\mathbf{C}\mathbf{X}_1\mathbf{X}_1'\mathbf{C}')^{-1}\mathbf{C}\mathbf{X}_1)\mathbf{X}_1'] \end{aligned} \tag{2.3}$$

using the result $\mathbf{A}^-\mathbf{A} = \mathbf{A}'(\mathbf{AA}')^{-1}\mathbf{A}$. Since $\mathbf{X}_1\mathbf{X}_1'$ is a $q \times q$ positive definite matrix, it can be factored as $\mathbf{X}_1\mathbf{X}_1' = \mathbf{MM}'$ where $\mathbf{M}$ is a $q \times q$ nonsingular matrix. Using the matrix $\mathbf{M}$, (2.3) can be written as

$$\rho[\mathbf{MM}' - \mathbf{MM}'\mathbf{C}'(\mathbf{CMM}'\mathbf{C}')^{-1}\mathbf{CMM}'] = \rho(\mathbf{I}_q - (\mathbf{CM})^-(\mathbf{CM})]. \tag{2.4}$$

The matrix $\mathbf{I}_q - (\mathbf{CM})^-(\mathbf{CM})$ is idempotent. Hence its rank is equal to its trace, or

$$\rho[\mathbf{I}_q - (\mathbf{CM})^-(\mathbf{CM})] = \operatorname{tr}[\mathbf{I}_q - (\mathbf{CM})^-(\mathbf{CM})] = q - k.$$

Thus assuming the set $\mathbf{A}\boldsymbol{\beta}$ is estimable implies that

$$\rho[\mathbf{X}(\mathbf{I} - \mathbf{A}^-\mathbf{A})] = q - k. \tag{2.5}$$

Now assume (2.5) holds and use the above definitions of $\mathbf{X}_1$ and $\mathbf{C}$. Assuming (2.5) and applying (2.3) is equivalent to

$$\rho[\mathbf{X}_1(\mathbf{I}-\mathbf{A}^-\mathbf{A})\mathbf{X}_1'] = \rho[\mathbf{X}_1(\mathbf{I}-\mathbf{A}'(\mathbf{AA}')^{-1}\mathbf{A})\mathbf{X}_1'] = q-k. \tag{2.6}$$

In order for (2.5) to imply that the linear combinations $\mathbf{A}\boldsymbol{\beta}$ are estimable, it must be shown that (2.5) implies there exists a $k\times q$ matrix $\mathbf{C}$ of rank k such that $\mathbf{A} = \mathbf{CX}_1$. Assume that $\mathbf{A}$ can be expressed as

$$\mathbf{A} = \mathbf{CX}_1+\mathbf{B} \quad \text{where} \quad \mathbf{X}_1\mathbf{B} = \mathbf{0} \quad \text{and} \quad \mathbf{C} \quad \text{is a} \quad k\times q \quad \text{matrix.} \tag{2.7}$$

The matrix in (2.6) can be written, using (2.7), as

$$\begin{aligned} \mathbf{X}_1[\mathbf{I}-(\mathbf{X}_1'\mathbf{C}'+\mathbf{B}')(\mathbf{AA}')^{-1}(\mathbf{CX}_1+\mathbf{B})]\mathbf{X}_1' &= (\mathbf{X}_1\mathbf{X}_1')-(\mathbf{X}_1\mathbf{X}_1')\mathbf{C}'(\mathbf{AA}')^{-1}\mathbf{CX}_1\mathbf{X}_1' \\ &= \mathbf{MM}'-\mathbf{MM}'\mathbf{C}'(\mathbf{AA}')^{-1}\mathbf{CMM}' \\ &= \mathbf{M}[\mathbf{I}_q-\mathbf{M}'\mathbf{C}'(\mathbf{AA}')^{-1}\mathbf{CM}]\mathbf{M}'. \end{aligned} \tag{2.8}$$

The rank condition of (2.6) then becomes

$$\rho[\mathbf{I}_q-\mathbf{M}'\mathbf{C}'(\mathbf{AA}')^{-1}\mathbf{CM}] = q-k. \tag{2.9}$$

The matrix of (2.9) is nonnegative and hence there exists a $q\times q$ orthogonal matrix $\mathbf{P}$ such that

$$\mathbf{P}(\mathbf{I}_q-\mathbf{M}'\mathbf{C}'(\mathbf{AA}')^{-1}\mathbf{C}M)\mathbf{P}' = \mathbf{I}_q-\mathbf{D}_\lambda \tag{2.10}$$

where $\mathbf{D}_\lambda$ is a diagonal matrix of the characteristic roots of $\mathbf{M}'\mathbf{C}'(\mathbf{AA}')^{-1}\mathbf{CM}$. The rank condition of (2.9) implies that $\mathbf{D}_\lambda$ has k ones on the diagonal and zeros elsewhere. This means that $\mathbf{M}'\mathbf{C}'(\mathbf{AA}')^{-1}\mathbf{CM}$ is an idempotent matrix. Thus $\mathbf{M}'\mathbf{C}'(\mathbf{AA}')^{-1}\mathbf{CMM}'\mathbf{C}'(\mathbf{AA}')^{-1}\mathbf{CM} = \mathbf{M}'\mathbf{C}'(\mathbf{AA}')^{-1}\mathbf{CM}$ which implies $\mathbf{C}'(\mathbf{AA}')^{-1} \times \mathbf{CX}_1\mathbf{X}_1'\mathbf{C}'(\mathbf{AA}')^{-1}\mathbf{C} = \mathbf{C}'(\mathbf{AA}')^{-1}\mathbf{C}$. Pre- and post-multiplying by $(\mathbf{CC}')^{-1}\mathbf{C}$ and $\mathbf{C}'(\mathbf{CC}')^{-1}$ respectively, gives $(\mathbf{AA}')^{-1}\mathbf{CX}_1\mathbf{X}_1'\mathbf{C}'(\mathbf{AA}')^{-1} = (\mathbf{AA}')^{-1}$ or $\mathbf{AA}' = \mathbf{CX}_1\mathbf{X}_1'\mathbf{C}$. From (2.7) $\mathbf{A} = \mathbf{CX}_1+\mathbf{B}$; thus $\mathbf{AA}' = \mathbf{CX}_1\mathbf{X}_1'\mathbf{C}'+\mathbf{BB}'$ which implies $\mathbf{BB}' = \mathbf{0}$. But $\mathbf{BB}' = \mathbf{0}$ if and only if $\mathbf{B} = \mathbf{0}$, which implies $\mathbf{A} = \mathbf{CX}_1$. Since the rank of $\mathbf{A}$ is assumed to be k, the rank of the $k\times q$ matrix $\mathbf{C}$ must also be k. This completes the proof.

For unbalanced designs, it is generally difficult to determine whether a linear combination of $\boldsymbol{\beta}$ is estimable. The result in Theorem 2.1 is also difficult to check but the estimability condition can be reformulated in terms of the trace of a matrix.

THEOREM 2.2. *For the conditions of Theorem* 2.1, *the linear combinations* $\mathbf{A}\boldsymbol{\beta}$ *are estimable if and only if*

$$\operatorname{tr}[\mathbf{X}(\mathbf{I}-\mathbf{A}^-\mathbf{A})\{\mathbf{X}(\mathbf{I}-\mathbf{A}^-\mathbf{A})\}^-] = q-k. \tag{2.11}$$

PROOF. The matrix of (2.11) is idempotent, hence

$$\begin{aligned} \operatorname{tr}[\mathbf{X}(\mathbf{I}-\mathbf{A}^-\mathbf{A})\{\mathbf{X}(\mathbf{I}-\mathbf{A}^-\mathbf{A})\}^-] &= \rho[\mathbf{X}(\mathbf{I}-\mathbf{A}^-\mathbf{A})\{\mathbf{X}(\mathbf{I}-\mathbf{A}^-\mathbf{A})\}^-] \\ &= \rho[\mathbf{X}(\mathbf{I}-\mathbf{A}^-\mathbf{A})] = q-k. \end{aligned}$$

This is necessary and sufficient for $\mathbf{A}\boldsymbol{\beta}$ to be estimable linear combinations.

The condition for estimability in Theorem 2.2 is easily checked since the matrix manipulations can be performed by an electronic computer as is described in Section 4.

3. Testing hypotheses about estimable linear combinations. For the LM, consider testing

$$H_0 : \mathbf{A}\boldsymbol{\beta} = \mathbf{0} \quad \text{against} \quad H_a : \mathbf{A}\boldsymbol{\beta} \neq \mathbf{0} \tag{3.1}$$

where $\mathbf{A}$ is a $k \times p$ matrix of rank k and the linear combinations $\mathbf{A}\boldsymbol{\beta}$ are estimable. Thus by Theorem 2.1

$$\rho\,[\mathbf{X}(\mathbf{I}-\mathbf{A}^{-}\mathbf{A})] = q-k. \tag{3.2}$$

The principle of conditional error (Bose (1949)) is used to compute the sum of squares due to the hypothesis as follows. Obtain the sum of squares due to error for the LM. Next impose the hypothesis on the LM to obtain a restricted model. The sum of squares due to the hypothesis is the sum of squares due to error for the restricted model minus the sum of squares due to error for the LM.

THEOREM 3.1. *The restricted model used to obtain the sum of squares due to the null hypothesis of* (3.1), *where the* $\mathbf{A}\boldsymbol{\beta}$ *are estimable linear combinations, is*

$$\mathbf{y} = \mathbf{X}(\mathbf{I}-\mathbf{A}^{-}\mathbf{A})\boldsymbol{\beta}+\mathbf{e}. \tag{3.3}$$

PROOF. Let $\mathbf{F}$ be a $p \times \overline{p-k}$ full rank matrix such that $\mathbf{F}\mathbf{F}' = (\mathbf{I}-\mathbf{A}^{-}\mathbf{A})$. Note that $\mathbf{F}'\mathbf{F} = \mathbf{I}_{p-k}$. Let $\mathbf{L}'$ be the $p \times p$ matrix $[\mathbf{A}', \mathbf{F}]$; in which case, $\mathbf{L}^{-1} = [\mathbf{A}^{-}, \mathbf{F}]$. The LM can be written as

$$\mathbf{y} = \mathbf{X}\mathbf{L}^{-1}\mathbf{L}\boldsymbol{\beta}+\mathbf{e} = \mathbf{X}\mathbf{A}^{-}\mathbf{A}\boldsymbol{\beta}+\mathbf{X}\mathbf{F}\mathbf{F}'\boldsymbol{\beta}+\mathbf{e}. \tag{3.4}$$

When the null hypothesis is true, i.e., $\mathbf{A}\boldsymbol{\beta} = \mathbf{0}$, the model of (3.4) is

$$\mathbf{y} = \mathbf{X}\mathbf{F}\mathbf{F}'\boldsymbol{\beta}+\mathbf{e} = \mathbf{X}(\mathbf{I}-\mathbf{A}^{-}\mathbf{A})\boldsymbol{\beta}+\mathbf{e}. \tag{3.5}$$

Equation (3.5) is the desired model.

The sum of squares due to error for the restricted model of (3.3) is

$$SSE_R = \mathbf{y}'[\mathbf{I}-\mathbf{X}(\mathbf{I}-\mathbf{A}^{-}\mathbf{A})[\mathbf{X}(\mathbf{I}-\mathbf{A}^{-}\mathbf{A})]^{-}]\mathbf{y}. \tag{3.6}$$

The sum of squares due to error for the LM is

$$SSE = (\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}}) = \mathbf{y}'(\mathbf{I}-\mathbf{X}\mathbf{X}^{-})\mathbf{y}. \tag{3.7}$$

By the principle of conditional error, the sum of squares due to the hypothesis is

$$SSH_0 = SSE_R-SSE = \mathbf{y}'\{\mathbf{X}\mathbf{X}^{-}-\mathbf{X}(\mathbf{I}-\mathbf{A}^{-}\mathbf{A})[\mathbf{X}(\mathbf{I}-\mathbf{A}^{-}\mathbf{A})]^{-}\}\mathbf{y}. \tag{3.8}$$

THEOREM 3.2. *Let*

$$Q_1 = \sigma^{-2}\mathbf{y}'\{\mathbf{X}\mathbf{X}^{-}-\mathbf{X}(\mathbf{I}-\mathbf{A}^{-}\mathbf{A})[\mathbf{X}(\mathbf{I}-\mathbf{A}^{-}\mathbf{A})]^{-}\}\mathbf{y}$$

$$Q_2 = \sigma^{-2}\mathbf{y}'[\mathbf{I}-\mathbf{X}\mathbf{X}^{-}]\mathbf{y} \qquad \textit{and}$$

$$\mathbf{y} \sim \mathbf{N}_n(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I}_n).$$

Then Q_1 is distributed as a noncentral chi-square random variable with k degrees of freedom and noncentrality parameter $1/2\sigma^2\boldsymbol{\beta}'\mathbf{A}'\mathbf{A}'^{-}\mathbf{X}'(\mathbf{I}-\mathbf{X}(\mathbf{I}-\mathbf{A}^{-}\mathbf{A})[\mathbf{X}(\mathbf{I}-\mathbf{A}^{-}\mathbf{A})]^{-})\mathbf{X}\mathbf{A}^{-}\mathbf{A}\boldsymbol{\beta}$, *$Q_2$ is distributed as a central chi-square random variable with $n-k$ degrees of freedom, and Q_1 and Q_2 are independent.*

PROOF. For the quadratic form $\mathbf{y}'\mathbf{A}\mathbf{y}$ to be distributed as a chi-square random variable with r degrees of freedom, the matrix of the quadratic form, $\mathbf{A}$, must be idempotent of rank r. This condition is easily checked for Q_1 and Q_2. The noncentrality parameter of the quadratic form is $1/2\sigma^2 E(\mathbf{y}')\mathbf{A}E(\mathbf{y})$. The noncentrality parameters can be easily obtained for Q_1 and Q_2. The product of the two matrices of the quadratic forms is null, which is the necessary and sufficient condition for the two quadratic forms to be independent. This completes the proof.

4. A stepwise computing procedure. Matrix products of the form $\mathbf{B}\mathbf{B}^{-}$ and $\mathbf{B}^{-}\mathbf{B}$ need to be computed in order to check the estimability condition and to compute the sums of squares for testing hypotheses. This matrix product can be computed by an iterative technique, which only involves computing the generalized inverse of a vector.

THEOREM 4.1. *Let $\mathbf{X}$ be any $n\times p$ matrix of rank q where $q \leqq p \leqq n$ and let $[\mathbf{X}_1, \mathbf{X}_2]$ be any column-wise partition of $\mathbf{X}$. Then*

$$\mathbf{X}\mathbf{X}^{-} = \mathbf{X}_1\mathbf{X}_1^{-}+(\mathbf{I}-\mathbf{X}_1\mathbf{X}_1^{-})\mathbf{X}_2[(\mathbf{I}-\mathbf{X}_1\mathbf{X}_1^{-})\mathbf{X}_2]^{-}. \tag{4.1}$$

PROOF. It is known that for any conditional inverse of $(\mathbf{X}'\mathbf{X})$, $\mathbf{X}\mathbf{X}^{-} = \mathbf{X}(\mathbf{X}'\mathbf{X})^c\mathbf{X}'$. It can be shown that a conditional inverse of $(\mathbf{X}'\mathbf{X})$ is

$$(\mathbf{X}'\mathbf{X})^c = \begin{bmatrix} (\mathbf{X}_1'\mathbf{X}_1)^c+(\mathbf{X}_1'\mathbf{X}_1)^c\mathbf{X}_1'\mathbf{X}_2\mathbf{H}^c\mathbf{X}_2'\mathbf{X}_1(\mathbf{X}_1'\mathbf{X}_1)^c & -(\mathbf{X}_1'\mathbf{X}_1)^c\mathbf{X}_1'\mathbf{X}_2\mathbf{H}^c \\ & \\ -\mathbf{H}^c\mathbf{X}_2'\mathbf{X}_1(\mathbf{X}_1'\mathbf{X}_1)^c & \mathbf{H}^c \end{bmatrix} \tag{4.2}$$

where $\mathbf{H}^c$ is any conditional inverse of $\mathbf{H} = \mathbf{X}_2'(\mathbf{I}-\mathbf{X}_1\mathbf{X}_1^{-})\mathbf{X}_2$. Using this conditional inverse, $\mathbf{X}\mathbf{X}^{-}$ can be expressed as

$$\mathbf{X}\mathbf{X}^{-} = \mathbf{X}_1\mathbf{X}_1^{-}+(\mathbf{I}-\mathbf{X}_1\mathbf{X}_1^{-})\mathbf{X}_2[(\mathbf{I}-\mathbf{X}_1\mathbf{X}_1^{-})\mathbf{X}_2]^{-}.$$

The iterative procedure uses this result by partitioning $\mathbf{X}$ as $[\mathbf{x}_{10}, \mathbf{X}_{20}]$ where $\mathbf{x}_{10}$ is a $n\times 1$ vector and $\mathbf{X}_{20}$ is a $n\times \overline{p-1}$ matrix. Then by Theorem 4.1,

$$\begin{aligned}\mathbf{X}\mathbf{X}^{-} &= \mathbf{x}_{10}\mathbf{x}_{10}^{-}+[(\mathbf{I}-\mathbf{x}_{10}\mathbf{x}_{10}^{-})\mathbf{X}_{20}][(\mathbf{I}-\mathbf{x}_{10}\mathbf{x}_{10}^{-})\mathbf{X}_{20}]^{-} \\ &= \mathbf{x}_{10}\mathbf{x}_{10}^{-}+\mathbf{X}_1\mathbf{X}_1^{-},\end{aligned} \tag{4.3}$$

where $\mathbf{X}_1 = (\mathbf{I}-\mathbf{x}_{10}\mathbf{x}_{10}^{-})\mathbf{X}_{20}$. Next partition the $n\times\overline{p-1}$ matrix $\mathbf{X}_1$ as $\mathbf{X}_1 = [\mathbf{x}_{11}, \mathbf{X}_{21}]$ where $\mathbf{x}_{11}$ is a $n\times 1$ vector and $\mathbf{X}_{21}$ is a $n\times\overline{p-2}$ matrix. Again by Theorem 4.1,

$$\begin{aligned}\mathbf{X}\mathbf{X}^{-} &= \mathbf{x}_{10}\mathbf{x}_{10}^{-}+\mathbf{x}_{11}\mathbf{x}_{11}^{-}+[(\mathbf{I}-\mathbf{x}_{11}\mathbf{x}_{11}^{-})\mathbf{X}_{21}][(\mathbf{I}-\mathbf{x}_{11}\mathbf{x}_{11}^{-})\mathbf{X}_{21}]^{-} \\ &= \mathbf{x}_{10}\mathbf{x}_{10}^{-}+\mathbf{x}_{11}\mathbf{x}_{11}^{-}+\mathbf{X}_2\mathbf{X}_2^{-}, \text{ where } \mathbf{X}_2 = (\mathbf{I}-\mathbf{x}_{11}\mathbf{x}_{11}^{-})\mathbf{X}_{21}.\end{aligned} \tag{4.4}$$

Continue by partitioning the $n \times \overline{p-2}$ matrix $\mathbf{X}_2$ as $\mathbf{X}_2 = [\mathbf{x}_{12}, \mathbf{X}_{22}]$ and apply Theorem 4.1 to this partitioned matrix. This process must be carried out $p-1$ times where on the pth step the $\mathbf{X}_{p-1}$ matrix will be a $n \times 1$ vector. The generalized inverse of a vector $\mathbf{h}$ is $\mathbf{h}^- = \mathbf{h}'/\mathbf{h}'\mathbf{h}$ which is easily computed. The following theorem has just been proven.

THEOREM 4.2. *Let* $\mathbf{X}$ *be any* $n \times p$ *matrix of rank* q. *Then* $\mathbf{X}\mathbf{X}^- = \sum_{i=0}^{p-1} \mathbf{x}_{1i}\mathbf{x}_{1i}$ *where the* $\mathbf{x}_{1i}$ *vectors are defined above.*

THEOREM 4.3. *Let* $\mathbf{X}$ *be any* $n \times p$ *matrix of rank* q. *Partition* $\mathbf{X}$ *row-wise as*

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{bmatrix}. \text{ Then}$$

$$\mathbf{X}^-\mathbf{X} = \mathbf{X}_1{}'\mathbf{X}_1{}'^- + [(\mathbf{I} - \mathbf{X}_1{}'\mathbf{X}_1{}'^-)\mathbf{X}_2{}'][(\mathbf{I} - \mathbf{X}_1{}'\mathbf{X}_1{}'^-)\mathbf{X}_2{}']^-.$$

PROOF. By a property of the definition of the generalized inverse,

$$\mathbf{X}^-\mathbf{X} = (\mathbf{X}^-\mathbf{X})' = \mathbf{X}'\mathbf{X}'^- \quad \text{and} \quad \mathbf{X}' = [\mathbf{X}_1{}', \mathbf{X}_2{}'].$$

The result follows by applying Theorem 4.1 to the matrix $\mathbf{X}'$.

A technique for computing the generalized inverse of a matrix is obtained using the iterative procedure and Theorem 4.3.

THEOREM 4.4. *The* $p \times p$ *matrix* $\mathbf{X}'\mathbf{X} + \mathbf{I}_p - \mathbf{X}^-\mathbf{X}$ *is nonsingular and*

$$\mathbf{X}^- = (\mathbf{X}'\mathbf{X} + \mathbf{I}_p - \mathbf{X}^-\mathbf{X})^{-1}\mathbf{X}'.$$

PROOF. The $n \times p$ matrix $\mathbf{X}$ has q linearly independent rows and the $p \times p$ matrix $\mathbf{I}_p - \mathbf{X}^-\mathbf{X}$ has $p-q$ linearly independent rows. Since $\mathbf{X}[\mathbf{I}_p - \mathbf{X}^-\mathbf{X}] = \mathbf{0}$,

$$\mathbf{W} = \begin{bmatrix} \mathbf{X} \\ \mathbf{I}_p - \mathbf{X}^-\mathbf{X} \end{bmatrix}$$

has p linearly independent rows. But $\mathbf{W}'\mathbf{W} = \mathbf{X}'\mathbf{X} + \mathbf{I}_p - \mathbf{X}^-\mathbf{X}$ and is of rank p since $\mathbf{W}$ is of rank p. The inverse of $\mathbf{W}'\mathbf{W}$ can be expressed as $(\mathbf{W}'\mathbf{W})^{-1} = (\mathbf{X}'\mathbf{X})^- + \mathbf{I}_p - \mathbf{X}^-\mathbf{X}$, i.e.,

$$[(\mathbf{X}'\mathbf{X})^- + \mathbf{I}_p - \mathbf{X}^-\mathbf{X}][(\mathbf{X}'\mathbf{X}) + \mathbf{I}_p - \mathbf{X}^-\mathbf{X}] = \mathbf{I}_p.$$

Therefore $[\mathbf{X}'\mathbf{X} + \mathbf{I}_p - \mathbf{X}^-\mathbf{X}]^{-1}\mathbf{X}' = [(\mathbf{X}'\mathbf{X})^- + \mathbf{I}_p - \mathbf{X}^-\mathbf{X}]\mathbf{X}' = (\mathbf{X}'\mathbf{X})^-\mathbf{X}' = \mathbf{X}^-\mathbf{X}'^-\mathbf{X}' = \mathbf{X}^-\mathbf{X}\mathbf{X}^- = \mathbf{X}^-$.

This completes the proof.

The theoretical and computational results obtained herein enable one to easily analyze the LM as the iterative computational procedure can be programmed for computer application. The estimate of $\boldsymbol{\beta}$ can be computed as $\hat{\boldsymbol{\beta}} = \mathbf{X}^-\mathbf{Y}$ using Theorem 4.4, the estimability of linear combinations of the parameters can be determined and the sums of squares due to a hypothesis about estimable linear combinations can be computed by applying Theorems 4.2 and 4.3. In particular, the criterion of Theorem 2.2 implies that the trace of the matrix must be an integer. This condition is not significantly affected by machine rounding error for the above computational procedure and it has the advantage over Searle's condition in that only one number must be checked.

REFERENCES

BOSE, R. C. (1949). *Least squares aspects of analysis of variance*, Institute of Statistics Mimeo Series 9, Chapel Hill.

CHIPMAN, J. S. (1964). On least squares with insufficient observations. *J. Amer. Statist. Assoc.* **59** 1078–1111.

GRAYBILL, F. A. (1961). *An Introduction to Linear Statistical Models* **1**. McGraw-Hill, New York.

GRAYBILL, F. A. (1969). *Introduction to Matrices with Applications in Statistics*. Wadsworth, Belmont, California.

JOHN, P. W. M. (1964). Pseudo-inverses in the analysis of variance. *Ann. Math. Statist.* **35** 895–896.

LENTNER, M. M. (1969). Generalized least-squares estimation of a subvector of parameters in randomized fractional factorial experiments. *Ann. Math. Statist.* **40** 1344–1352.

PENROSE, R. (1955). A generalized inverse for matrices. *Proc. Cambridge Philos. Soc.* **51** 406–413.

PLACKETT, R. L. (1950). Some theorems in least squares. *Biometrika* **37** 149–157.

RAO, C. R. (1962). A note on a generalized inverse of a matrix with applications to problems in mathematical statistics. *J. Roy. Statist. Soc. Ser. B* **24** 152–158.

RAYNER, A. A. and PRINGLE, R. M. (1967). A note on generalized inverses in the linear hypothesis not of full rank. *Ann. Math. Statist.* **38** 271–273.

SEARLE, S. R. (1966). *Matrix Algebra for the Biological Sciences*. Wiley, New York.

ZYSKIND, GEORGE (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.