Sankhyā : The Indian Journal of Statistics
1975, Volume **37**, Series A, Pt. 3, pp. 428–438

GENERALIZED INVERSES OF THE FUNDAMENTAL BORDERED MATRIX USED IN LINEAR ESTIMATION

By FRANK J. HALL and CARL D. MEYER, JR.
North Carolina State University, Raleigh, North Carolina, U.S.A.

SUMMARY. All blocks of g-inverses of a bordered matrix of the form $B = \begin{bmatrix} V & X \\ X' & 0 \end{bmatrix}$ are completely characterized, and it is shown that these blocks are completely independent of each other. A new form for a B^- is presented and some new results concerning invariance and uniqueness of some important expressions involving blocks in a B^- are proven.

1. Introduction

For a general linear model $(Y, X\beta, \sigma^2 V)$ where X may be deficient in rank and V may be singular, C. R. Rao (1971, Theorem 3.1) has shown that the problem of inference from the linear model can be completely solved once one has obtained a g-inverse for the matrix

$$B = \begin{bmatrix} V & X \\ X' & 0 \end{bmatrix}. \qquad \dots \quad (1.1)$$

Although there has been a substantial amount of research conducted that concerns the matrix B and g-inverses, B^-, of B, we feel that there are still many facts about the structure of g-inverses for B which have not yet been discovered. For example, a great deal of attention has been given to the problem of obtaining a particular form for an entire B^- but very little attention has been devoted to the structure of the various submatrices that might appear in a B^-. One of the purposes of this paper is to show that the submatrices of B^- are entirely independent of each other and then to completely characterize the various classes of matrices which are blocks in a g-inverse for B, and thus characterize all g-inverses for B. As pointed out by Rao (1971), there is a need for efficient algorithms for computing g-inverses of B. As a consequence of our work, it will follow that the submatrices of any g-inverse for B may be computed separately and independently so that the sizes of the matrices involved in a computational scheme for a B^- may be greatly reduced, and thus opening the door to faster and more efficient algorithms. The equations on which such computations must be based are given in Theorem 4.2. It is hoped that our work may be useful in the future development of computational methods for g-inverses of B.

In addition, we present a new form for a B^- along with some results concerning the invariance and uniqueness of some terms which appear in a g-inverse for B.

2. Notation

The notation used by Rao and Mitra (1971) will be the notation adopted in this paper. In addition, B will always denote the matrix (1.1) where V is nonnegative definite. If V is $n\times n$ and X is $n\times r$ then every $n\times n$ matrix which appears as an upper left hand block in some g-inverse for B will be called a *C_{11}-matrix.* Likewise, those $n\times r$ matrices which appear as an upper right hand block in some B^- are called *C_{12}-matrices*; those $r\times n$ matrices which appear as lower left hand blocks in a B^- are called *C_{21}-matrices*; and those $r\times r$ matrices which are lower right hand blocks in a B^- are called *C_{22}-matrices.* For example, if

$$B^- = \begin{bmatrix} Q_{n\times n} & U_{n\times r} \\ L_{r\times n} & T_{r\times r} \end{bmatrix}$$

is a g-inverse for B then Q is a C_{11}-matrix, U is a C_{12}-matrix, L is C_{21}-matrix, and T is C_{22}-matrix.

For a given A^-, E_A will always denote the matrix $I-AA^-$ and F_A will denote $I-A^-A$.

3. Results on C_{11}-matrices

In the first theorem we exhibit a new form for a particular g-inverse of B. This form is simpler in structure than many of the previous forms which have been given (see Pringle and Rayner (1970) and Rao (1972)). Many of the techniques of this paper arose from a consideration of this new form.

Theorem 3.1 : *Let X^-, X'^- and $(E_XVF_{X'})^-$ be any choices of g-inverses of X, X', and $E_XVF_{X'}$, respectively. Then, a g-inverse of B is given by*

$$B^- = \left[\begin{array}{c|c} 0 & X'^- \\ \hline X^- & -X^-VX'^- \end{array}\right] + \begin{bmatrix} I \\ -X^-V \end{bmatrix} Q[I : -VX'^-] \qquad \ldots \quad (3.0)$$

where $Q = F_{X'}(E_XVF_{X'})^-E_X$.

Proof : The proof is by direct computation and is omitted.

Remark : Note that for each g-inverse, X^-, of X, the matrix $X^{-\prime}$ is a g-inverse for X'. In computational considerations, it is desirable to use this choice of g-inverse for X'. If this is done then $F_{X'} = E'_X$ so that Q in (3.0) assumes the more symmetrical form

$$Q = E'_X(E_XVE'_X)^-E_X.$$

Throughout this paper, one may employ the above mentioned option and take $F_{X'} = E'_X$, thus simplifying some of our statements and notation. However, for the sake of generality, we have stated all of our results using both $F_{X'}$ and E_X.

It should be pointed out that Theorem 3.1 is a special case of more general results obtained by Meyer (1973). Also, (3.0) may be obtained from Theorem 6 in Meyer's 1970 paper by a simple permutation.

In the next theorem we characterize the class of C_{11}-matrices.

Theorem 3.2 : *Let X^- and X'^- be any choices of g-inverses for X and X', respectively. Then, Q is a C_{11}-matrix iff Q satisfies*

$$E_X(V-VQV) = 0 \qquad \text{... (3.1)}$$

$$X'QV = 0 \qquad \text{... (3.2)}$$

$$VQX = 0 \qquad \text{... (3.3)}$$

and

$$X'QX = 0. \qquad \text{... (3.4)}$$

Furthermore, (3.1) *can be replaced by*

$$(V-VQV)F_{X'} = 0. \qquad \text{... (3.5)}$$

Proof : Let

$$W = \left[\begin{array}{c|c} 0 & X'^- \\ \hline X^- & -X^-VX'^- \end{array}\right]$$

and

$$Y = \left[\begin{array}{c|c} Q & -QVX'^- \\ \hline -X^-VQ & X^-VQVX'^- \end{array}\right].$$

By direct multiplication

$$BWB = \left[\begin{array}{c|c} V-E_XVF_{X'} & X \\ \hline X' & 0 \end{array}\right]$$

and

$$BYB = \left[\begin{array}{c|c} E_XVQVF_{X'} & E_XVQX \\ \hline X'QVF_{X'} & X'QX \end{array}\right].$$

Hence, if (3.1-3.4) hold (or if (3.2—3.4) and (3.5) hold) then $B(W+Y)B = B$ and Q is a C_{11}-matrix. The converse can easily be inferred from Rao's (1972) results.

Remark : It is clear from the proof of Theorem 3.1 that (3.1-3.4) can be replaced by

$$E_X VQVF_{X'} = E_X VF_{X'} \quad \dots \quad (3.6)$$

$$X'QVF_{X'} = 0 \quad \dots \quad (3.7)$$

$$E_X VQX = 0 \quad \dots \quad (3.8)$$

and

$$X'QX = 0. \quad \dots \quad (3.9)$$

If the restriction that Q be a *nonnegative definite* C_{11}-matrix is imposed in Theorem 3.2 then both (3.2) and (3.3) may be eliminated and (3.4) may be replaced by the condition $QX = 0$.

We are now in a position to establish a general form for C_{11}-matrices.

Theorem 3.3 : *Let X^-, X'^-, and K^- be any choices of g-inverses, where $K = E_X VF_{X'}$. If Q is a C_{11}-matrix then there exists a matrix Z and a g-inverse $\hat{K}$ of K so that*

$$Q = Z(I-KK^-)E_X + F_{X'}(I-K^-K)Z + F_{X'}\hat{K}E_X. \quad \dots \quad (3.10)$$

Conversely, for every matrix Z and every g-inverse $\hat{K}$ of K, the matrix Q in (3.10) *is a C_{11}-matrix.*

Proof : It is straightforward to show that if Q is of the form (3.10), then (3.6-3.9) hold, so that Q is a C_{11}—matrix.

Conversely, suppose Q is a C_{11}-matrix. Then, (3.1-3.4) are satisfied. Now,

$$VF_{X'}QV = VQV - VX'^-X'QV = VQV \quad \dots \quad (3.11)$$

since $X'QV = 0$ from (3.2). Similarly,

$$VQE_X V = VQV \quad \dots \quad (3.12)$$

from (3.3). Thus, we have

$$KQK = K \quad \dots \quad (3.13)$$

from (3.1) and (3.11-3.12).

We now show that if $Z = Q$

and

$$\hat{K} = QKK^- + K^-KQ - Q,$$

then (3.10) is true. Observe that $K\hat{K}K = K$ from (3.13). Now,

$$\begin{aligned} X'QK &= X'QE_X VF_{X'} \quad \dots \quad (3.14) \\ &= X'QVF_{X'} - X'QXX^-VF_{X'} \\ &= 0-0 = 0 \end{aligned}$$

from (3.2) and (3.4). Similarly,

$$KQX = 0. \qquad \dots \quad (3.15)$$

Using (3.14-3.15), it is now easy to show that

$$Q(I-KK^-)E_X+F_{X'}(I-K^-K)Q+F_{X'}(QKK^-+K^-KQ-Q)E_X = Q.$$

Thus, (3.10) is true and the proof is completed.

We can now answer the interesting question of when does there exist a *unique* C_{11}-matrix ?

Theorem 3.4 : *Let* X^- *and* X'^- *be any choices of g-inverses and let* $K = E_X V F_{X'}$. *The following statements are equivalent.*

There exists a unique C_{11}*-matrix.* … (3.16)

$$R(E_X V) = R(E_X). \qquad \dots \quad (3.17)$$

$$\mathcal{M}(E_X V) = \mathcal{M}(E_X). \qquad \dots \quad (3.18)$$

$$R(VF_{X'}) = R(F_{X'}). \qquad \dots \quad (3.19)$$

$$\mathcal{M}(F'_{X'}V) = \mathcal{M}(F'_{X'}). \qquad \dots \quad (3.20)$$

$F_{X'}K^-E_X$ *is invariant among all choices of* K^-. … (3.21)

Proof : Before we begin to prove the theorem, we make the following observations.

$$\mathcal{M}(F_{X'}) = \mathcal{N}(X') = \mathcal{M}(E'_X). \qquad \dots \quad (3.22)$$

$$F_{X'}E'_X = E'_X \text{ and } E'_X F_{X'} = F_{X'}. \qquad \dots \quad (3.23)$$

$$E_X = 0 \text{ if and only if } F_{X'} = 0. \qquad \dots \quad (3.24)$$

$$\mathcal{M}(K) = \mathcal{M}(E_X V) \text{ and } \mathcal{M}(K') = \mathcal{M}(F'_{X'} V). \qquad \dots \quad (3.25)$$

For conformable nonzero matrices A, B, and C such that BA^-C is invariant among all choices of A^-, it must be the case that

$$\mathcal{M}(B') \subseteq \mathcal{M}(A') \text{ and } \mathcal{M}(C) \subseteq \mathcal{M}(A). \qquad \dots \quad (3.26)$$

It is clear that (3.22) is true and (3.23) and (3.24) are immediate consequences of (3.22). To see that (3.25) is true, note that (3.22) implies that

$$\mathcal{M}(K) = \mathcal{M}(E_X V E'_X) \text{ and } \mathcal{M}(K') = \mathcal{M}(F'_{X'} V F_{X'})$$

so that by writing $V = A'A$,

we obtain $\mathcal{M}(K) = \mathcal{M}(E_X A'AE'_X) = \mathcal{M}((AE'_X)'AE'_X) = \mathcal{M}((AE'_X)')$

$$= \mathcal{M}(E_X A') = \mathcal{M}(E_X V)$$

and $\mathcal{M}(K') = \mathcal{M}(F'_{X'}A'AF_{X'}) = \mathcal{M}(F'_{X'}A') = \mathcal{M}(F'_{X'} V)$.

Statement (3.26) is known to be true and is given as Ex. 14, p. 43, in Rao and Mitra (1971). Although it is not explicitly stated in Rao and Mitra, it is clear that (3.26) may not hold when either A, B, or C is a zero matrix.

We now proceed with the proof of theorem. It is clear that (3.17) $\Longleftrightarrow$ (3.18) and (3.19) $\Longleftrightarrow$ (3.20). Assume now that (3.17) holds and use (3.22) and (3.25) to obtain

$$R(VF_{X'}) = R(F'_{X'}\, V) = R(K) = R(E_X V) = R(E_X) = R(F_{X'}).$$

Hence (3.17) $\Longrightarrow$ (3.19). By a similar argument, it is easy to show that (3.19) $\Longrightarrow$ (3.17). We now have shown (3.17) $\Longleftrightarrow$ (3.18) $\Longleftrightarrow$ (3.19) $\Longleftrightarrow$ (3.20). Assume now that (3.21) is true. If K, E_X, $F_{X'}$ are all nonzero matrices, then (3.26) yields

$$\mathcal{M}(E_X) \subseteq \mathcal{M}(K) = \mathcal{M}(E_X V) \subseteq \mathcal{M}(E_X)$$

and hence (3.18) follows. If either $E_X = 0$ or $F_{X'} = 0$ then (3.18) trivially follows. If $K = 0$ then (3.21) implies $F_{X'}HE_X = 0$ for all $n \times n$ matrices H, so that in particular, $F_{X'}E'_XE_X = 0$. By virtue of (3.23), it follows that $E_X = 0$ and thus (3.18) again is true. Therefore, in all cases, we have shown that (3.21) $\Longrightarrow$ (3.18). Conversely, assume (3.18). Thus, (3.20) also holds. By virtue of (3.25) and Lemma 2.2.4, p. 21, of Rao and Mitra it easily follows that (3.18) $\Longrightarrow$ (3.21), and hence we have shown (3.18) $\Longleftrightarrow$ (3.21). To complete the proof, note that by using (3.10), one obtains that (3.16) $\Longrightarrow$ (3.21). Conversely, (3.21) implies (3.18) and (3.20) so that by using (3.25) it follows that

$$KK^-E_X = E_X \text{ and } F_{X'}K^-K = F_{X'}.$$

Hence, by (3.10), we have that (3.21) $\Longrightarrow$ (3.16), and the proof is now complete.

Theorem 3.5 : *The following statements are equivalent* :

(1) $\mathcal{M}(V) \subseteq \mathcal{M}(X)$.

(2) $\mathcal{N}(X') \subseteq \mathcal{N}(V)$.

(3) 0 *is a* C_{11}-*matrix.*

Proof : The proof is straightforward and is omitted.

4. C_{12}, C_{21}, C_{22}-MATRICES AND THE INVARIANCE OF VQV

Our first theorem of this section shows that the product VQV is invariant among all C_{11}-matrices Q. Using this invariance, we will be able to fully characterize the other blocks of g-inverses of B.

Theorem 4.1 : *The product VQV is invariant for all C_{11}-matrices Q. In fact, if $V = A'A$ and $\mathcal{S} = \mathcal{M}(A/\mathcal{N}(X'))$ (i.e., the range of A restricted to $\mathcal{N}(X')$) then*

$$VQV = A'P_{\mathcal{S}}A = A'P_{AE_X'}A \quad \ldots \quad (4.1)$$

for every C_{11}-matrix Q.

Proof: Let X^- be a g-inverse for X and use $X^{-\prime}$ as a g-inverse for X' so that the term K given in Theorem 3.3 can be written as $K = EVE'$ where $E = I - XX^-$. Let K^- become g-inverse for K and let Q be a C_{11}-matrix. By Theorem 3.3, we know that there exists another g-inverse, $\hat{K}$, of K and a matrix Z such that

$$Q = Z(I-KK^-)E + E'(I-K^-K)Z + E'\hat{K}E.$$

Since $\mathcal{M}(K) = \mathcal{M}(EV)$, it follows that

$$(I-KK^-)EV = 0 \quad \ldots \quad (4.2)$$

and

$$VE'(I-K^-K) = 0. \quad \ldots \quad (4.3)$$

By Lemma 2.2.6 of Rao and Mitra (1971), $VE'\hat{K}EV$ is invariant under the choice of g-inverse $\hat{K}$ of K. Thus, choose $\hat{K} = K^\dagger$ so that

$$\begin{aligned} VE'\hat{K}EV &= A'AE'(EA'AE')^\dagger EA'A \\ &= A'(EA')^\dagger(EA')A = A'P_{AE'}A. \end{aligned} \quad \ldots \quad (4.4)$$

The result (4.1) now easily follows from (4.2—4.4) and the fact that $\mathcal{M}(AE') = \mathcal{S}$

Remark: An alternate proof of (4.1) will be included in the doctoral dissertation of Hall.

The next result gives necessary and sufficient conditions for a matrix to be a C_{12}, C_{21}, or C_{22}-matrix. But, more important, it establishes the amazing fact that the blocks of g-inverses of B are completely independent of each other.

Theorem 4.2 : *Let*

$$D = V - VQV$$

where Q is any C_{11}-matrix. (Note that in view of Theorem 4.1, D is invariant among all choices for Q.) Each of the following statements is true.

$$U \text{ is a } C_{12}\text{-matrix iff} \quad \ldots \quad (4.5)$$

$$X'UX' = X' \text{ and } VUX' = D.$$

$$L \text{ is a } C_{21}\text{-matrix iff } XLX = X \text{ and } XLV = D. \quad \ldots \quad (4.6)$$

$$T \text{ is a } C_{22}\text{-matrix iff } XTX' = -D. \quad \ldots \quad (4.7)$$

Moreover, if Q is any C_{11}-matrix, U is any C_{12}-matrix, L is any C_{21}-matrix and T is any C_{22}-matrix, then the composite matrix

$$M = \begin{bmatrix} Q & U \\ L & T \end{bmatrix}$$

is always a g-inverse for B.

Proof: We will prove (4.5). Suppose first that U satisfies the two equations in (4.5) and let

$$M_U = \begin{bmatrix} Q & U \\ X^-(I-VQ) & -X^-DX'^- \end{bmatrix}$$

where $Q = F_{X'}(E_X V F_{X'})^- E_X$. (Note that M_U is just (3.0) with the (1, 2)-block replaced by U.) From Theorem 3.1 we know Q is a C_{11}-matrix so that Theorem 3.2 guarantees that

$$VQVF_{X'} = VF_{X'}$$

and therefore

$$D = DX^-X'.$$

Using this, it is a simple matter to show that M_U is a g-inverse for B and hence U is a C_{12}-matrix. Conversely, suppose U is a C_{12}-matrix. Then there exist matrices M_{11}, M_{21}, and M_{22} such that

$$\begin{bmatrix} M_{11} & U \\ M_{21} & M_{22} \end{bmatrix}$$

is a g-inverse for B. In Theorem 2.3 Rao (1972) proves that it must be the case that

$$X'UX' = X'$$

and

$$VM_{11}V + VUX' = V.$$

The desired result now follows by virtue of Theorem 4.1. To prove (4.6) and (4.7), one constructs the matrices

$$M_L = \begin{bmatrix} Q & (I-QV)X'^- \\ L & -X^-DX'^- \end{bmatrix}$$

and

$$M_T = \begin{bmatrix} Q & (I-QV)X'^- \\ X^-(I-VQ) & T \end{bmatrix}$$

and proceeds in a manner similar to that used in the proof of (4.5). To prove the second part of the theorem, suppose that Q is any C_{11}-matrix, U is any C_{12}-matrix, L is any C_{21}-matrix, and T is any C_{22}-matrix. That the composite matrix

$$M = \begin{bmatrix} Q & U \\ L & T \end{bmatrix}$$

is always a g-inverse for B follows by direct calculation using (3.2-3.4) along with (4.5-4.7).

Theorem 4.2 must be considered to be important from a computational standpoint since it guarantees that each block of a g-inverse for B may be calculated as a separate entity without regard to any other block which might appear in a g-inverse for B. Furthermore, Theorem 3.2 and Theorem 4.2 provide the equations on which these calculations must be based. Throughout our development, the common thread is the invariant term D. If a C_{11}-matrix has been previously calculated then D is readily available. However, Theorem 4.1 shows that it is not necessary to always first obtain a C_{11}-matrix in order to calculate D. The computational aspects of obtaining D have not yet been studied but this is a topic which warrants further investigation.

The special case in which $D = 0$ may be of some interest. In the next theorem, we characterize those matrices B for which this situation occurs.

Theorem 4.3 : *Let D denote the invariant term $D = V - VQV$ where Q is any C_{11}-matrix. The following statements are equivalent.*

$$D = 0 \text{ (i.e., } Q \textit{ is a g-inverse for } V). \qquad \ldots \quad (4.8)$$

$$0 \textit{ is a } C_{22}\textit{-matrix}, \qquad \ldots \quad (4.9)$$

$$R(V) = R(VF_{X'}). \qquad \ldots \quad (4.10)$$

$$R(V) = R(E_X V). \qquad \ldots \quad (4.11)$$

$$R(V) = Tr\,(VQ). \qquad \ldots \quad (4.12)$$

Proof : The fact that (4.8) $\Longleftrightarrow$ (4.9) is a consequence of (4.7). The other equivalences are consequences of (4.1) and the fact that

$$R(E_X V) = R(VF_{X'}) = \mathrm{Tr}(VQ).$$

The details are left to the reader.

Corollary 4.1 : *For the linear model* $(Y,\ X\beta,\ \sigma^2 V)$ *where* $D = 0$, *the following statements are true.*

(i) *The BLUE of an estimable parametric function* $p'\beta$ *is* $p'\hat{\beta}$ *where*

$$\hat{\beta} = (X^- - X^- V + X^- V P_X)Y.$$

(ii) *If* $p'\beta$ *is estimable then* $p'\hat{\beta} = p'\beta$ *with probability one.*

(iii) *An unbiased estimator of* σ^2 *is*

$$\hat{\sigma}^2 = \frac{Y'(I-P_X)Y}{R(V)}.$$

Proof : Use Rao's (1971) Theorem 3.1 together with the following facts. If $D = 0$ Theorem 3.2 implies $I-P_X$ is a C_{11}-matrix. But then if Q is any C_{11}-matrix, $X^- - X^- VQ$ satisfies the two equations in (4.6) and hence is a C_{21}-matrix. Putting $Q = I-P_X$ yields that $X^- - X^- V + X^- VP_X$ is a C_{21}-matrix, and hence (i) follows. Because 0 is a C_{22}-matrix, $V(p'\hat{\beta}) = 0$ and hence (ii) follows. From (4.12), we obtain that $R(V) = \mathrm{Tr}(VQ)$ so that (iii) follows, because $R(V : X) - R(X) = \mathrm{Tr}(VQ)$.

Remark : Note that in the case $D = 0$, the matrix $Q = F_{X'} E_X$ is a C_{11}-matrix and may be used in place of $I-P_X$ in Corollary 4.1.

5. Final remarks

1. The results of this paper hold for complex matrices (with the appropriate changes).

2. Using (4.5-4.7), one can obtain general forms for C_{12}, C_{21}, and C_{22}-matrices similar to (3.10). Various forms will be included in the dissertation of Hall. Using such forms, one has at his disposal all of the g-inverses of B.

3. In a future publication the results of this paper and other results will be extended to a Hilbert space setting.

4. We express thanks to the referee for some valuable suggestions which helped to provide a more elegant proof of Theorem 4.1.

References

Meyer, C. D. (1970) : Generalized inverses of block triangular matrices. *Siam J. Appl. Math.*, **19**, 741-750.

——— (1971) : Representations for (1)- and (1, 2)-inverses for partitioned matrices. *Linear Algebra and Its Applications*, **4**, 221-232.

——— (1973) : Generalized inverses and ranks of block matrices. *Siam J. Appl. Math.*, **25**, 597-602.

Pringle, P. M. and Rayner, A. A. (1970) : Inverses of a bordered matrix with applications to the theory of constrained linear models. *Siam Review*, **12**, 107-115.

Rao, C. R. (1971) : Unified theory of linear estimation. *Sankhyā*, Series A, **33**, Part 4, 371-394.

——— A note on the IPM method in the unified theory of linear estimation. *Sankhyā* Series A, **34**, Part 3, 285-288.

Rao, C. R. and Mitra, S. K. (1971) : *Generalized Inverse of Matrices and Its Applications*, John Wiley and Sons, Inc., New York, N. Y.

Paper received : April, 1973.
Revised : April, 1974.