SIAM J. Appl. Math.
Vol. 15, No. 2, March, 1967
Printed in U.S.A.

ON CONSTRUCTION AND PROPERTIES OF THE GENERALIZED INVERSE*

ARMEN G. FISHER†

1. Introduction. Suppose Z is any $n \times p$ matrix of rank r, and Z^* is its conjugate transpose. A unique $p \times n$ generalized inverse of Z, denoted $Z^\dagger$, has the four properties:

$$(1) \qquad \begin{array}{ll} \text{(a) } ZZ^\dagger Z = Z, & \text{(b) } Z^\dagger ZZ^\dagger = Z^\dagger, \\ \text{(c) } (ZZ^\dagger)^* = ZZ^\dagger, & \text{(d) } (Z^\dagger Z)^* = Z^\dagger Z. \end{array}$$

The properties and applications of $Z^\dagger$ were extensively studied in [1], [3], [6], [7], [10] and [11]. A nonunique "weak" form of $Z^\dagger$, without property (c), was defined and used in [5]. A "normalized" form of $Z^\dagger$, without property (d), was defined in [14]. However, in solving linear equations, the essential property of $Z^\dagger$ is (a). A nonunique generalized inverse Z^-, of rank at least r, is then any $p \times n$ matrix with property (a). It was shown in [10] that, if $Q = PZ$ has any solution P (that is, Q lies in the row space of Z) then QZ^- is a solution, and if $ZP = Q$ has a solution P then Z^-Q is a solution. We know the row spaces of Z and Z^*Z are the same, and the column spaces of Z^* and Z^*Z are the same. Then $Z = Z(Z^*Z)^-Z^*Z$ and thus $(Z^*Z)^-Z^* = Z^-$ (cf. [14]). Similarly, from the conjugate transpose of the equation $Z^*Z(Z^*Z)^-Z^* = Z^*$, we see that $((Z^*Z)^-)^*Z^* = Z^-$. Any matrix $(Z^*Z)^-Z^*$ has rank r, since

$$r = \operatorname{rank} Z(Z^*Z)^-Z^*Z \leqq \operatorname{rank} (Z^*Z)^-Z^* \leqq \operatorname{rank} Z^* = r.$$

In §2 we construct a matrix Z^- of arbitrary rank. In §3 we prove that $(Z^*Z)^-Z^*$ is a "normalized" generalized inverse of Z. In §4 we offer simplified proofs of some matrix results which are used in the construction and application of a form of $(Z^*Z)^-$ familiar in least squares theory.

2. Rao [12] constructed a generalized inverse Z_m^- of maximum rank $m = \min(n, p)$. We construct Z_{r+q}^- with arbitrary rank $r + q$, where $0 \leqq q \leqq \min(p - r, n - r)$. Suppose, after a suitable rearrangement of rows and columns, Z is partitioned

$$\left(\begin{array}{c|c} Z_{11} & Z_{12} \\ \hline Z_{21} & Z_{22} \end{array}\right),$$

* Received by the editors June 6, 1966, and in revised form September 9, 1966.

† Industrial Administration Program, Union College, Poughkeepsie, New York. This work was offered in partial fulfillment of the requirements for the Ph.D. in Statistics at Rutgers, The State University, New Brunswick, New Jersey.

where Z_{11} is an $r \times r$ nonsingular matrix. Suppose also that

$$V = \left(\begin{array}{c|c} V_{11} & V_{12} \\ \hline V_{21} & V_{22} \end{array}\right)$$

is any $(p - r) \times p$ matrix whose rows form a basis for the orthogonal complement of the row space of Z, partitioned so that V_{11} is $q \times r$. Then $Z_{11}V_{11}^* + Z_{12}V_{12}^* = 0$ and therefore $V_{11}^* = -Z_{11}^{-1}Z_{12}V_{12}^*$, and thus V_{12}^* has rank q. It follows that the $p \times n$ matrix

$$Z_{r+q}^- = \left(\begin{array}{c|c|c} Z_{11}^{-1} & V_{11}^* & 0 \\ \hline 0 & V_{12}^* & 0 \end{array}\right),$$

with $n - r - q$ null columns, is a generalized inverse of Z, with rank $r + q$.

3. The matrix $Z(Z^*Z)^-Z^*$ is used in the analysis of linear statistical models (cf. [4]). It is easily shown to be idempotent. Rao [13] has shown it to be unique. It is in fact the unique orthogonal projector onto the column space of Z. We show that $(Z^*Z)^-Z^*$ is a normalized generalized inverse of Z by proving property (c) in (1), that $Z(Z^*Z)^-Z^*$ is Hermitian. (Thus only the upper triangular portion of $Z(Z^*Z)^-Z^*$ need be computed and stored.) We first prove a characterization of Hermitian matrices. We let Re (c) denote the real component of a complex number c.

LEMMA 1. *If S is a square matrix, then*

(i) $\mathrm{Re}\,(\mathrm{Tr}\,(S^2)) \leqq \mathrm{Tr}\,(SS^*)$,

(ii) $\mathrm{Re}\,(\mathrm{Tr}\,(S^2)) = \mathrm{Tr}\,(SS^*)$ *if and only if* $S = S^*$, *and*

(iii) $S^2 = SS^*$ *if and only if* $S = S^*$.

Proof. For any skew-Hermitian matrix Q, $Q = -Q^*$, and $\mathrm{Tr}\,(Q^2) = -\mathrm{Tr}\,(QQ^*)$ is real $\leqq 0$. Therefore $\mathrm{Tr}\,(Q^2) = 0$ if and only if Q is null. But $(S - S^*)$ is skew-Hermitian, so $\mathrm{Tr}\,(S - S^*)^2 = 0$ if and only if $S = S^*$. Then $0 \geqq \mathrm{Tr}\,(S - S^*)^2 = \mathrm{Tr}\,(S^2) + \mathrm{Tr}\,(S^{*2}) - 2\mathrm{Tr}\,(SS^*)$. Here $\mathrm{Tr}\,(S^{*2}) = \overline{\mathrm{Tr}\,(S^2)}$, thus $\mathrm{Tr}\,(S^2) + \mathrm{Tr}\,(S^{*2}) = 2\,\mathrm{Re}\,(\mathrm{Tr}\,(S^2))$, and $\mathrm{Re}\,(\mathrm{Tr}\,(S^2)) \leqq \mathrm{Tr}\,(SS^*)$, with equality if and only if $S = S^*$. Part (iii) follows directly from (ii).

Remark. Result (i) yields the Schwarz inequality for real vectors X and Y. That is,

$$\begin{aligned}(X'Y)^2 &= \mathrm{Tr}\,X'YX'Y = \mathrm{Tr}\,YX'YX' \\ &\leqq \mathrm{Tr}\,YX'(YX')' = \mathrm{Tr}\,YX'XY' = \mathrm{Tr}\,Y'YX'X = (Y'Y)(X'X).\end{aligned}$$

LEMMA 2. *$Z(Z^*Z)^-Z^*$ is Hermitian idempotent.*

Proof. Let $Z(Z^*Z)^-Z^* = S$. It is easily verified that $S^*S = S = S^2$. Thus the result follows from Lemma 1.

4. Suppose F is any $p \times p$ matrix such that $FZ^*Z = (m_{ij})$ is in Hermite row-echelon form (cf. [9]) but with the rows suitably interchanged (cf. [12]) so that:

$$m_{ij} = 0 \text{ if } i > j; \quad m_{ii} = 0 \text{ or } 1;$$

$$\text{if } m_{ii} = 0, \quad \text{then} \quad m_{ij} = 0, \qquad j = 1, \cdots, p;$$

$$\text{if } m_{jj} = 1, \quad \text{then} \quad m_{ij} = 0, \qquad i = 1, \cdots, p, \quad i \neq j.$$

Then FZ^*Z is idempotent and $F = (Z^*Z)^-$ has rank p. Furthermore, after suitable rearrangement of the columns of Z, FZ^*Z may be written in the form $\left(\begin{array}{c|c} I_r & L \\ \hline 0 & 0 \end{array}\right)$, where L is $r \times (p - r)$. Then $V = (-L^* \mid I_{p-r})$ is a $(p - r) \times p$ basis for the orthogonal complement of the identical row spaces of Z and Z^*Z. Suppose a "constraints matrix" U is any $(p - r) \times p$ matrix such that VU^* is nonsingular. It is easily verified, by contradiction, that $(Z^* \mid U^*)$ has full rank p. If we then solve the equation $V(Z^* \mid U^*) = (0 \mid VU^*)$ for V, and multiply through by $(VU^*)^{-1}$, we get

$$(VU^*)^{-1}V = U(Z^*Z + U^*U)^{-1}. \tag{2}$$

The following results, some previously shown by more involved arguments, follow directly from (2): $U(Z^*Z + U^*U)^{-1}U^* = I$, $Z(Z^*Z + U^*U)^{-1}U^* = 0$ (Chipman [2]), $Z(Z^*Z + U^*U)^{-1}Z^*Z = Z$, and thus $(Z^*Z + U^*U)^{-1} = (Z^*Z)^-$ (John [8] and Rao [12]).

Remark. In statistical work, the real matrix $(Z'Z)^-Z'$ may be computed by reducing $Z'Z$ to Hermite canonical form. That is,

$$F(Z'Z \mid Z') = ((Z'Z)^-Z'Z \mid (Z'Z)^-Z'). \tag{3}$$

However, the program for (3) often overflows a modern computer with statistical problems of even modest size. Suppose Z is arbitrarily partitioned $(Z_1 \mid Z_2)$, where Z_1 is $n \times p_1$, Z_2 is $n \times p_2$. It was shown independently by Rohde [14] and Fisher [4] that $(Z'Z)^-$ may be stated in partitioned form. Then (cf. [4]) we may show that a partitioned form of $(Z'Z)^-Z'$ is

$$Z^- = \left(\frac{Z_1^-(I_n - Z_2(DZ_2)^-D)}{(DZ_2)^-D}\right), \tag{4}$$

where $Z_1^- = (Z_1'Z_1)^-Z_1'$ is $p_1 \times n$ and $D = Z_2'(I - Z_1Z_1^-)$ is $p_2 \times n$. To confirm (4), we premultiply and postmultiply (4) by Z, noting that $DZ_1 = 0$ and that, since the row spaces of Z_2 and DZ_2 are identical, then $Z_2(DZ_2)^-DZ_2 = Z_2$. The component matrices Z_1^- and $(DZ_2)^-D$ may then be computed by means of the program for (3). That is, using (3), we reduce $(Z_1'Z_1 \mid Z_1')$ to $((Z_1'Z_1)^-Z_1'Z_1 \mid Z_1^-)$, and reduce $(DZ_2 \mid D)$ to $((DZ_2)^-DZ_2 \mid (DZ_2)^-D)$.

REFERENCES

[1] A. Ben-Israel and A. Charnes, *Contributions to the theory of generalized inverses*, this Journal, 11 (1963), pp. 667–699.

[2] J. S. Chipman, *Least squares with insufficient observations*, J. Amer. Statist. Assoc., 59 (1964), pp. 1078–1111.

[3] R. E. Cline, *Representations for the generalized inverse of a partitioned matrix*, this Journal, 12 (1964), pp. 588–600.

[4] A. G. Fisher, *Some techniques for analysis and interpretation of estimable functions in incomplete fixed-factor designs*, Doctoral thesis, Rutgers University, 1966.

[5] A. J. Goldman and M. Zelen, *Weak generalized inverses and minimum variance linear unbiased estimation*, J. Res. Nat. Bur. Standards Sect. B, 68 (1964), pp. 151–172.

[6] T. N. E. Greville, *Some applications of the pseudoinverse of a matrix*, SIAM Rev., 2 (1960), pp. 15–22.

[7] ———, *Fitting functions of several independent variables*, this Journal, 9 (1961), pp. 109–115.

[8] P. W. John, *Pseudo-inverses in the analysis of variance*, Ann. Math. Statist., 35 (1965), pp. 895–896.

[9] M. Marcus and H. Minc, *A Survey of Matrix Theory and Matrix Inequalities*, Allyn and Bacon, Boston, 1964, p. 34.

[10] R. Penrose, *A generalized inverse for matrices*, Proc. Cambridge Philos. Soc., 51 (1955), pp. 406–413.

[11] C. M. Price, *The matrix pseudoinverse and minimal variance estimates*, SIAM Rev., 6 (1964), pp. 115–120.

[12] C. R. Rao, *Linear Statistical Inference and Its Applications*, John Wiley, New York, 1965, pp. 24–30, 183, 189.

[13] ———, *On the generalized inverse of a matrix with applications*, J. Roy. Statist. Soc. Ser. B, 24 (1962), pp. 152–158.

[14] C. A. Rohde, *Generalized inverses of partitioned matrices*, this Journal, 13 (1965), pp. 1033–1035.