Generalized Inverses and Solutions of Linear Systems *

John Z. Hearon**

(September 28, 1968)

For an arbitrary complex matrix A we consider (1) the set of all matrices B such that ABA = Aand AB is Hermitian and (2) the set of all matrices B such that ABA = A and BA is Hermitian. It is shown that if B is in (1) then x = By is a least-squares solution of Ax = y and that if B is in (2) then x = By is the solution of minimum Euclidian norm of the consistent systems Ax = y. The connection is exposed between the properties of the generalized inverses in (a) and (b) and the fact that among all matrices X satisfying AXA = A, that with minimum Euclidian norm is the Moore-Penrose inverse of A.

Key Words: Generalized inverse; linear systems; matrix.

1. Introduction

In this paper we discuss some properties of certain classes of generalized inverses, give some new results and some enlightening connections between some known results.

It is shown how to construct (a) the set of all matrices B such that ABA = A and AB is Hermitian and (b) the set of all matrices B such that ABA = A and BA is Hermitian. The generalized inverses in (a) include the normalized generalized inverse of Rohde [15]¹ and have been employed by Decell and Odell [3] in constructing the fixed point probability vector of regular or ergodic transition matrices. The generalized inverses in (b) include the Goldman-Zelen weak generalized inverse [4] and these authors have utilized a property of their inverse which we mention in the next paragraph.

It is well known [12] that given Ax = y, with A and y a given matrix and vector respectively, $x_0 = A^+y$, where A^+ is the Moore-Penrose generalized inverse, is the least-squares solution of minimum Euclidian norm; in particular, the solution of minimum norm when the system is consistent. In what follows, we show that if B is in (b), then $x_0 = By$ is the solution of minimum Euclidian norm when the system is consistent. It is further shown that if B is in (a), then x = By is a leastsquares solution of the system. We then show that the only matrix B such that $x_0 = By$ is the leastsquares solution of minimum norm is $B = A^+$. This is known [12]. The proof here shows the connection between this property and the weaker classes of generalized inverses mentioned above. All of the proofs assign a definitive role to a well-known property of perpendicular projections. The connection is shown between the properties of the generalized inverses in (a) and (b) and the fact that among all matrices B such that ABA = A, that with minimum Euclidian norm is A^+ .

2. Definitions and Preliminaries

We consider matrices with complex entries. For any matrix M we denote by $\rho(M)$, R(M), N(M) and M^* the rank, range, nullspace and conjugate transpose respectively of M. By the norm, ||M||, of any matrix M, we mean the Euclidian norm: $||M||^2 = \text{trace } (M^*M) = \text{trace } (MM^*)$. For any column vector x, $||x||^2 = x^*x$, and we observe that $||M||^2$ is the sum of the squared norms of the column vectors of M.

^{*}An invited paper.

^{**}Mathematical Research Branch, NIAMD, National Institutes of Health, Bethesda, Md. 20014.

¹ Figures in brackets indicate the literature references at the end of this paper.

We use the terminology for generalized inverses previously employed [7, 8, 9]. For a given matrix A, $C_1(A)$ is the set of all matrices B such that ABA = A; $C_2(A)$ is the set of all matrices B such that $B\epsilon C_1(A)$ and BAB = B; $C_3(A)$ is the set of all matrices $B\epsilon C_2(A)$ such that AB is Hermitian; $C_{3'}(A)$ is the set of all matrices $B\epsilon C_2(A)$ such that BB = B; $C_3(A)$ is the set of all matrices $B\epsilon C_2(A)$ such that BAB = B; $C_3(A)$ is the set of all matrices $B\epsilon C_2(A)$ such that BA is Hermitian; $C_4(A)$ is the set of all matrices B such that $B\epsilon C_3(A)$ and $B\epsilon C_{3'}(A)$. We call a matrix $B\epsilon C_i(A)$ a C_i -inverse of A, i = 1, 2, 3, 3', 4. We note that $C_4(A)$, the intersection of $C_3(A)$ and $C_{3'}(A)$, is a single matrix, the unique Moore-Penrose generalized inverse [11]. We observe that if $B\epsilon C_i(A)$ then AB and BA are projections and that $\rho(AB) = \rho(BA) = \rho(A)$ [15]. In particular, if i is 3, 3' or 4, then AB, BA or both, respectively, are perpendicular projections. We reserve the notation A^+ for the C_4 -inverse of any matrix A.

We consider the system of equations

$$Ax = y, \tag{1}$$

where A and y are given. If $y \in R(A)$, i.e., the system is consistent, then it is known [6] that there exists a solution $x_0 \in R(A^*)$, which is uniquely determined and is the solution of minimum norm: $||x_0|| \leq ||x||$, where x is any solution, with equality only if $x = x_0$. The proof is rather simpler than that [6] cited. Any solution x can be written x = u + v, with $u \in R(A^*)$ and $v \in N(A)$. Then clearly u is a solution. Let $x_1 = u_1 + v_1$ and $x_2 = u_2 + v_2$ be any two solutions with $u_i \in R(A^*)$, $v_i \in N(A)$, i = 1, 2. Then $x_1 - x_2 = (u_1 - u_2) + (v_1 - v_2)$ is in N(A), which implies that $(x_1 - x_2) - (v_1 - v_2) = u_1 - u_2$ is in N(A). But $(u_1 - u_2) \in R(A^*)$ and hence $u_1 = u_2$. Thus any solution can be written $x = x_0 + v$, with $x_0 \in R(A^*)$ uniquely determined and $v \in N(A)$. Clearly $||x||^2 = ||x_0||^2 + ||v||^2 \ge ||x_0||^2$. In what follows, we therefore refer to the solution of minimum norm of a consistent system.

It is well known that if (1) is inconsistent, then any solution of

$$A^*Ax = A^*y \tag{2}$$

is a least-squares solution of (1). For, let x be any vector, x_1 satisfy (2), d = Ax - y and $d_1 = Ax_1 - y$. Then, by (2), $d_1^*A = 0$ and since $d_1 - d = A(x_1 - x)$ we have $d_1^*(d_1 - d) = 0$. Thus $d_1^*d_1 = d_1^*d = d^*d_1$. From the Cauchy inequality, $(d^*d)(d_1^*d_1) \ge (d_1^*d)^2$, we then have $d^*d \ge d_1^*d_1$, with equality if and only if $d = d_1$. If equality is the case, then $d^*A = 0$ and x is also a least-squares solution. Conversely, if x and x_1 are least-squares solutions then $d = d_1$, for then $(x - x_1) \in N(A^*A) = N(A)$ which implies $Ax = Ax_1$. Thus every solution of (2) gives the same vector of deviations d = Ax - y, and, of necessity, the same minimum sum of squares. By what has been said in the last paragraph above, there exists a unique least-squares solution of minimum norm. Penrose [12] calls this least-squares solution the "best approximate solution" of (1), but he applies this concept and definition to more general systems than (1).

If (1) is consistent, any solution is a least-squares solution (zero sum of squares) and the solution of minimum norm is the least-squares solution of minimum norm. Thus the Penrose definition of "best approximate solution" is meaningful, and always exists for the system (1), without reference to the consistency of (1).

Lemma 1 gives a well-known property of perpendicular projections and Lemma 2 is an easy consequence of Lemma 1. We include both for ready reference and clarity.

LEMMA 1. Let E be a projection. Then $||x|| \ge ||Ex||$, for every x, if and only if $E = E^*$. In that case, ||x|| = ||Ex|| if and only if $x \in R(E)$.

PROOF. The first statement is a known theorem [5, p. 147]. To see the second statement, let x = u + v with $u \in R(E)$ and $v \in N(E)$. Then $||Ex||^2 = x^* Ex = u^* u$. But $||x||^2 = x^* x = u^* u$ if and only if $x \in R(E)$.

LEMMA 2. Let X be any matrix and E and F be Hermitian projections. If $X_0 = EXF$, we have $||X_0|| \leq ||X||$, with equality if and only if $X = X_0$.

PROOF. Define Y = EX. Then by Lemma 1, $||y_i|| \le ||x_i||$ where y_i and x_i are the *i*th column of Y and X respectively. Thus we have $||Y|| \le ||X||$ with equality if and only if Y = X. Now $X_0^* = FY^*$, and by the same argument $||X_0|| \le ||Y||$, with equality if and only if $X_0 = Y$. Hence, $||X_0|| \le ||Y|| \le ||X||$, with equality only if $X_0 = X$.

3. Generalized Inverses and Solutions of Linear Systems

In what follows we will be interested in and make use of matrices $B \in C_1(A)$ such that either *AB* or *BA* is Hermitian. The next lemma gives conditions which such C_1 -inverses must satisfy and in fact a method of constructiong them.

LEMMA 3. Let A be an arbitrary, fixed matrix. Then every solution, B, of

$$\mathbf{A}^* \mathbf{A} \mathbf{B} = \mathbf{A}^* \tag{3}$$

is in $C_1(A)$ and is such that AB is a uniquely determined Hermitian projection. Conversely, every $B \in C_1(A)$ such that AB is Hermitian satisfies (3). Similarly every solution, B, of

$$AA^*B^* = A \tag{4}$$

is in $C_1(A)$ and such that BA is Hermitian. Conversely every $B \in C_1(A)$ such that BA is Hermitian satisfies (4).

PROOF. Define E = AB. Then (3) reads $A^*E = A^*$ and left multiplication by B^* gives $E^*E = E$, which shows $E = E^* = E^2$. But then (3) can be written $A^*B^*A^* = A^*$ which implies that $B\epsilon C_1(A)$. Since $E = E^*$ and it is known that $\rho(A) = \rho(E)$, we have R(E) = R(A) and E is completely determined by A. Conversely, if ABA = A and AB is Hermitian we have $B^*A^*A = A$ which gives (3). The statements regarding the solutions of (4) are proved in the same way.

In the proof of his Theorem 1, Penrose [11] includes the second part of Lemma 3. Bose [2] has previously shown, for the real case, that if *B* is any solution of (3) then *AB* is a unique Hermitian projection. The proof is quite different (cf. Rohde [14]) and neither Bose nor Rohde drew the conclusion that $B \in C_1(A)$. As before ([9]), we call a matrix *B* a *strict* C_1 -inverse of *A* if $B \in C_1(A)$ and $B \notin C_2(A)$, or equivantly ([7, 15]), if $B \in C_1(A)$ and $\rho(B) > \rho(A)$. There exist strict C_1 -inverses of *A* which satisfy (3) or (4) or both. For, if we write

$$A = Q \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} R \tag{5}$$

with Q and R unitary, where D is square, nonsingular and $\rho(A) = \rho(D)$, then [9] any $B \epsilon C_1(A)$ must have the form

$$B = R^* \begin{bmatrix} D^{-1} & U \\ V & W \end{bmatrix} Q^*$$
(6)

where the entries of U, V, and W are arbitrary. It is readily verified that AB is Hermitian if and only if U=0; that BA is Hermitian if and only if V=0; and that when either or both of these conditions are satisfied, $\rho(B) > \rho(A)$ if $W \neq 0$. Thus there exist solutions of (3) which are not in $C_3(A)$, solutions of (4) not in $C_{3'}(A)$ and solutions of both (3) and (4) other than the C_4 -inverse of A.

The generalized inverses required by the theorem of Decell and Odell [3] are precisely those matrices satisfying (3), i.e., those matrices of the form (6) with U=0.

Penrose [11] has shown that KXM = P is consistent if and only if P = KLPNM, where L is any matrix in $C_1(K)$ and N is any matrix in $C_1(M)$. In that case, $X_0 = K^+PM^+$ is the solution of minimum norm [12]. The following theorem shows that, in the consistent case, the solution of minimum norm can be obtained by the use of generalized inverses in a weaker class than the C_4 -inverse.

THEOREM 1. Let K, M, and P be given matrices such that KXM = P is consistent. Then the solution of minimum norm is $X_0 = LPN$, where L is any matrix in $C_1(K)$ such that LK is Hermitian and N is any matrix in $C_1(M)$ such that MN is Hermitian.

PROOF. Let X be any solution. Then $X_0 = LPN = LKXMN$ is also a solution. Since LK and MN are Hermitian projections, we have by Lemma 2, that $||X_0|| \le ||X||$, with equality if and only if $X = X_0$.

Ben-Israel and Wersan [1] have shown that the solution of (3) with minimum norm is the C_4 -inverse of A. We show how this follows from Theorem 1.

COROLLARY 1. If A is a given matrix, the solution of $A^*AX = A^*$ of minimum norm is the C₄-inverse of A.

PROOF. Let $K \in C_1(A^*A)$ and KA^*A be Hermitian. Then, by Theorem 1, $X_0 = KA^*$ is the solution of minimum norm. It is known [9] that $X_0 \in C_3(A)$. Since $X_0A = KA^*A$ is Hermitian, we have $X_0 \in C_4(A)$.

In view of Lemma 3, Corollary 1 shows that among all matrices $B \in C_1(A)$ such that AB is Hermitian, the C_4 -inverse of A has minimum norm. In fact among *all* matrices in $C_1(A)$, the C_4 -inverse of A has minimum norm. A proof of this has been indicated by Kalman [10]. We give this result below as a corollary of Theorem 1.

COROLLARY 2. If A is a given matrix, the solution of AXA = A of minimum norm is the C₄-inverse of A.

PROOF. Clearly AXA = A is consistent. By Theorem 1, the solution of minimum norm is $X_0 = B_1AB_2$ where B_1 and B_2 are in $C_1(A)$ and are such that B_1A and AB_2 are Hermitian. By a known theorem [7], $X_0 \in C_2(A)$, and plainly $AX_0 = AB_2$ and $X_0A = B_1A$ are Hermitian. Hence $X_0 \in C_4(A)$.

REMARK. One construction of the C_4 -inverse of A given by Penrose [11] is in fact $B = B_1AB_2$, where B_1 and B_2 are any solutions of (4) and (3), respectively. This is a special instance of the theorem [7] that $B = B_1AB_2$ is in $C_2(A)$ when B_1 and B_2 are *arbitrary* elements of $C_1(A)$. In particular, if AB_2 is Hermitian, then $B \in C_3(A)$ and if B_1A is Hermitian, then $B \in C_{3'}(A)$ ([cf. 9]). We could also prove Corollary 2 by observing that among all matrices of the form (6),

$$B = R * \begin{bmatrix} D^{-1} & 0 \\ 0 & 0 \end{bmatrix} Q^*$$

has minimum norm and that, according to the discussion after (6), this matrix is A^+ .

The next theorem gives a subset of $C_1(A)$ such that elements of this subset select the solution of minimum norm when (1) is consistent. It also gives a subset of $C_1(A)$ which selects a least-squares solution of (1) and shows that the C_4 -inverse of A is the only matrix which enjoys both properties.

THEOREM 2. Consider the system

$$\mathbf{A}\mathbf{x} = \mathbf{y} \tag{s}$$

where A and y are a given matrix and vector respectively.

(i) Let $B \in C_1(A)$ and BA be Hermitian. If (s) is consistent, then $x_0 = By$ is the solution of minimum norm. Conversely, if $x_0 = By$ is the solution of minimum norm for every y such that (s) is consistent, then $B \in C_1(A)$ and BA is Hermitian.

(ii) Let $B \in C_1(A)$ and AB be Hermitian. Then x = By is a least-squares solution. Conversely, if x = By is a least-squares solution for every y, then $B \in C_1(A)$ and AB is Hermitian.

(iii) Let B be the C_4 -inverse of A. Then $x_0 = By$ is the least-squares solution of minimum norm. Conversely, if $x_0 = By$ is the least-squares solution of minimum norm for every y, then B is the C_4 -inverse of A.

PROOF. (i) It follows from Theorem 1 that $x_0 = By$ is the solution of minimum norm if $B \in C_1(A)$ and BA is Hermitian. Conversely, suppose $x_0 = By$ to be the solution of minimum norm for each $y \in R(A)$. Then $Ax_0 = ABy = y$, for all $y \in R(A)$ implies ABAz = Az, for arbitrary z, and so $B \in C_1(A)$. Thus far we have only used the fact that By is a solution. For any z, let y = Az, then we have By = BAzand ||By|| = ||BAz||. If By is the solution of minimum norm, $||By|| = ||BAz|| \le ||z||$, and that BA is Hermitian follows from Lemma 1. (ii) The least-squares solutions of (s) are the solutions of $A^*Ax = A^*y$. We show that x = By with *B* as in (ii) satisfies this equation. We have $A^*Ax = A^*ABy = A^*B^*A^*y = A^*y$, since $AB = B^*A^*$ and ABA = A. Conversely, if $A^*ABy = A^*y$ for every *y*, we have $A^*AB = A^*$ and, from Lemma 3, $B \in C_1(A)$ and *AB* is Hermitian.

(iii) First, let $B = A^+$. Then by (ii), x_0 is a least-squares solution. Now let x be any least-squares solution. Then $A^*Ax = A^*y$ from which $BB^*A^*Ax = BB^*A^*y$. But $B^*A^* = AB$ and we have $BAx = By = x_0$. Since BA is Hermitian, $||x_0|| \le ||x||$ follows from Lemma 1. Conversely, suppose that for every y, $x_0 = By$ is the minimum norm solution of $A^*Ax = A^*y$. By Theorem 1 we then have $x_0 = By = KA^*y$ for all y, where $K \in C_1(A^*A)$ and KA^*A is Hermitian. Thus $B = KA^*$, which as in the proof of Corollary 1 is the C_4 -inverse of A.

REMARK. In the proof of (i) we have essentially proved that if $B \epsilon C_1(A)$, then x = By is a solution of (s) whenever $y \epsilon R(A)$ and that, conversely, if x = By is a solution of (s) for every $y \epsilon R(A)$ then $B \epsilon C_1(A)$. This was proved by Bose [2] (cf. [14]) and in fact Rao [13] takes this as a definition of $C_1(A)$.

The content of (iii) of Theorem 2 is known. For, Penrose [12] shows that $X_0 = A^+C$ is the least-squares solution of minimum norm of AX = C. By observing that A^+ is the least-squares solution of minimum norm of AX = I he shows that A^+ is the only matrix with the required properties. We prove (iii) here to show how the result stems directly from the Hermitian character of A^+A and AA^+ and to show the relation of (iii) to (i) and (ii).

It is clear that any $B \epsilon C_{3'}(A)$ meets the conditions of (i) of Theorem 2. It is precisely that property of the $C_{3'}$ -inverse which was exploited by Goldman and Zelen [4]. Their proof however is essentially different in character.² With Theorem 2 in hand, we can give a rather brief version of their proof of the Gauss theorem (Theorem 1 of [4]) utilizing a weaker class of generalized inverse than the $C_{3'}$ -inverse. The term *estimable* is used here as it is used in [4].

THEOREM 3. Let y be a vector of random variables such that the variance-covariance matrix is $\sigma^2 I$ and the expected value of y is $A\zeta$. The minimum variance unbiased linear estimate $\hat{\theta}$ of any estimable function $\theta = q^*\zeta$ is given by $\hat{\theta} = q^*By = q^*x$, where B is any matrix in $C_1(A)$ such that AB is Hermitian and x is any solution of (2).

PROOF. A linear estimate $\hat{\theta} = p^* y$ of θ is unbiased if and only if the expected value $E(\hat{\theta})$ of $\hat{\theta}$ is $E(\hat{\theta}) = p^* A \zeta = \theta$ for all ζ , i.e., $A^* p = q$. Thus θ is estimable if and only if $q \epsilon R(A^*)$, and since the variance of $\hat{\theta}$ is $(p^* p) \sigma^2$, $\hat{\theta}$ gives the minimum variance unbiased estimate if and only if p is the minimum norm solution of the consistent system $A^* p = q$. According to (i) of Theorem 2, this is equivalent to $p = B^* q$ where $B^* \epsilon C_1(A^*)$ and $B^* A^*$ is Hermitian, i.e., $B \epsilon C_1(A)$ and AB is Hermitian. Thus $\hat{\theta} = p^* y = q^* B y$. But for any $q \epsilon R(A^*)$ we further have $q^* B y = q^* x$, where x is any solution of (2). For, by (ii) of Theorem 2, By is a solution of (2) and any solution of (2) can be written x = By + v with $v \epsilon N(A)$.

It is of interest to note that the general solution of (2) can be written as $x = KA^*y + (I - KA^*A)z$, where K is any C_1 -inverse of A^*A and z is arbitrary.³ This, in a different notation, is eq (3.3) of Goldman and Zelen [4], except that they require that $K \in C_2(A^*A)$. As noted previously in this paper, and shown in [9], $KA^* \in C_3(A)$. Thus Goldman and Zelen in fact employ both the C_3 -inverse and the C_3 -inverse and these are included in the set of all matrices which meet the conditions of (ii) and (i) of Theorem 2, respectively.

3. References

- Ben-Israel, A., Wersan, S. J., An elimination method for computing the generalized inverse of an arbitrary complex matrix, J. Assn. for computing Machinery 10, 532-537 (1963).
- [2] Bose, R. C., Lecture Notes on Analysis of Variance, University of North Carolina, Chapel Hill, 1959.
- [3] Decell, H. P., Jr., and Odell, P. L., The fixed point probability vector of regular or ergodic transition matrices, J. Am. Stat. Assoc. **62**, 600–602 (1967).

² Their argument is given immediately preceding their equation (3.6) and they make explicit use of the fact that $C_2(A)$ is contained in $C_{a'}(A)$.

³ This is a straightforward application of Penrose's Corollary 1 [11] and we observe with Penrose that the only property of the C_4 -inverse required for the proof of the corollary is $B \in C_4(A)$ implies $B \in C_1(A)$.

- [4] Goldman, A. J., and Zelen, M., Weak generalized inverses and minimum variance linear unbiased estimation, J. Res. NBS 68B (Math. and Math. Phys.), No. 4, 151-172 (1964).
- [5] Halmos, P., Finite-dimensional Vector Spaces (D. Van Nostrand, New York, 1958).
- [6] Hamburger, H. L., and Grimshaw, M. E., Linear Transformation in n-dimensional Vector Space (Cambridge University Press, 1951).
- [7] Hearon, J. Z., Construction of *EPr* generalized inverses by inversion of nonsingular matrices, J. Res. NBS 71B (Math. and Math. Phys.), Nos. 2 and 3, 57-60 (1967).
- [8] Hearon, J. Z., A generalized matrix version of Rennie's inequality, J. Res. NBS **71B** (Math. and Math. Phys.), Nos. 2 and 3, 61-64 (1967).
- [9] Hearon, J. Z., and Evans, J. W., On spaces and maps of generalized inverses, J. Res. NBS 72B (Math. Sci.), No. 2, 103-107 (1968).
- [10] Kalman, R. E., New Methods and Results in Linear Prediction and Filtering Theory, Technical Report 61-1, Research Institute for Advanced Studies, Baltimore, Md. (1960).
- [11] Penrose, R., A generalized inverse for matrices, Proc. Cambridge Philos. Soc. 51, 406-418 (1955).
- [12] Penrose, R., On best approximate solutions of linear matrix equations, Proc. Cambridge Philos. Soc. 52, 17-19 (1956).
- [13] Rao, C. R., A note on a generalized inverse of a matrix with applications to problems in mathematical statistics, J. Royal Stat. Soc. (b) 24, 152–158 (1962).
- [14] Rohde, C. A., Contributions to the Theory, Computation and Application of Generalized Inverses, Ph.D. Thesis, North Carolina State University, Raleigh (1964).
- [15] Rohde, C. A., Some results on generalized inverses, SIAM Rev. 8, 201-205 (1966).

(Paper 72B4-280)