J. Soc. Indust. Appl. Math.
Vol. 12, No. 3, September, 1964
Printed in U.S.A.

REPRESENTATIONS FOR THE GENERALIZED INVERSE OF A PARTITIONED MATRIX*

RANDALL E. CLINE†

1. Introduction. The Moore-Penrose generalized inverse of a matrix was introduced by E. H. Moore [1] as the "general reciprocal" of a matrix, and by Penrose [2] as the "generalized inverse" of a matrix. This generalized inverse exists and is unique for every matrix (not necessarily square) with complex elements. The definitions employed by Moore and Penrose, although not identical, have been shown to be equivalent by Rado [3]. In addition, other authors have defined matrices which reduce to the generalized inverse. These include Bjerhammar [4], Lanzcos [5] and Greville [6].

Various expressions for the generalized inverse have been developed by a number of authors. These include Ben-Israel and Charnes [7], Ben-Israel and Wersan [8], Cline [9], [10], den Broeder and Charnes [11], den Broeder [12], Greville [13], [14], Penrose [15], and Pyle [16], [17]. In particular, Greville [13] has developed a representation for the generalized inverse of an arbitrary matrix, A_k, with k columns partitioned as $A_k = [A_{k-1}, a_k]$ where A_{k-1} designates the submatrix consisting of the first $k - 1$ columns and a_k is a single column. It is the purpose of this paper to show that Greville's representation can be extended to any matrix, A, partitioned as $A = [U, V]$ in which U and V are submatrices with l and $k - l$ columns, respectively, $0 \leqq l \leqq k$. Also, the converse problem of determining the generalized inverse of U, given the generalized inverse of A, is considered, and expressions in terms of partitions are developed.

2. Matrices of the form $A = [U, V]$. Adopting the definition employed by Penrose, the generalized inverse of an arbitrary matrix with complex elements is defined as follows: The four equations[1]

$$AXA = A, \tag{2.1}$$

$$XAX = X, \tag{2.2}$$

* Received by the editors October 17, 1963, and in revised form February 17, 1964.

† Optical and Radio Systems Laboratory, Institute of Science and Technology, The University of Michigan, Ann Arbor, Michigan. Presently at the U. S. Army Mathematics Center, The University of Wisconsin, Madison, Wisconsin. The material in this paper was submitted in 1963 to Purdue University as part of a thesis in partial fulfillment of the requirements for the degree of Doctor of Philosophy. Moreover a portion of the material was presented at the Symposium on Mathematical Programming held at the University of Chicago in June, 1962, and an abstract of the presentation is included in the proceedings of the symposium.

[1] The asterisk is used to designate the complex conjugate transpose of a matrix.

$$(AX)^* = AX, \tag{2.3}$$

$$(XA)^* = XA, \tag{2.4}$$

have a unique solution for every A. This unique solution is called the generalized inverse of A and is written $X = A^\dagger$.

Various useful properties of the generalized inverse are summarized in Lemma 1. Let I designate the n by n identity matrix and let θ designate the m by n null matrix, where m and n are determined by the context. Then we have, for arbitrary matrices,

LEMMA 1:

(a) $A = \theta(m \text{ by } n)$ *implies* $A^\dagger = \theta(n \text{ by } m)$.
(b) $A^{\dagger\dagger} = A$.
(c) $A^{\dagger *} = A^{*\dagger}$
(d) *If A has full column rank*, $A^\dagger = (A^*A)^{-1}A^*$.
(e) *If A is nonsingular*, $A^\dagger = A^{-1}$.
(f) *If U and V are unitary*, $(UAV)^\dagger = V^*A^\dagger U^*$.
(g) *If A is hermitian and idempotent*, $A^\dagger = A$.
(h) $A^\dagger A$, $AA^\dagger$, $I - A^\dagger A$ *and* $I - AA^\dagger$ *are all hermitian and idempotent.*
(i) $EF = \theta$ *implies* $F^\dagger E^\dagger = \theta$.
(j) *If P is hermitian and idempotent*, $(PQ)^\dagger = Q^\dagger P$ *whenever either* $PQ = Q$ *or P commutes with* Q, $Q^\dagger Q$ *and* $QQ^\dagger$.

(These relations are easily established by direct substitution in the defining equations for the generalized inverse, see, for example, [2], [6] and [10].)

Greville's form for the generalized inverse of a matrix $A_k = [A_{k-1}, a_k]$ can be stated as follows: Let $d_k = A_{k-1}^\dagger a_k$ and $c_k = a_k - A_{k-1}d_k$. Then

$$A_k^\dagger = \begin{bmatrix} A_{k-1}^\dagger - d_k b_k \\ b_k \end{bmatrix}, \tag{2.5}$$

where[2]

$$b_k = \begin{cases} c_k^\dagger, & \text{if} \quad c_k \neq \theta, \\ (1 + d_k^* d_k)^{-1} d_k^* A_{k-1}^\dagger, & \text{if} \quad c_k = \theta. \end{cases} \tag{2.6}$$

This form for $A_k^\dagger$ will now be extended to obtain representations for the generalized inverse of matrices $A = [U, V]$. We begin by combining (2.5) and (2.6) into a single expression.

Since c_k is a single column vector, $c_k \neq \theta$ implies $c_k^\dagger = (c_k^* c_k)^{-1} c_k^*$, by Lemma 1(d), and thus $c_k^\dagger c_k = 1$. Further, $c_k = \theta$ implies $c_k^\dagger = \theta$ and $c_k^\dagger c_k = 0$. Then we can rewrite b_k as

[2] Greville considered only matrices with real elements. It is an immediate consequence of the considerations in this paper, however, that the representation also holds for matrices with complex elements.

$$(2.7) \qquad b_k = c_k^\dagger + (1 - c_k^\dagger c_k)(1 + d_k^* d_k)^{-1} d_k^* A_{k-1}^\dagger$$

and obtain a single expression for the cases $c_k \neq \theta$ and $c_k = \theta$ in (2.6). Combining (2.5) and (2.7) then gives

$$(2.8) \qquad A_k^\dagger = \begin{bmatrix} A_{k-1}^\dagger - A_{k-1}^\dagger a_k c_k^\dagger - A_{k-1}^\dagger a_k(1 - c_k^\dagger c_k) k_1 a_k^* A_{k-1}^{\dagger *} A_{k-1}^\dagger \\ c_k^\dagger + (1 - c_k^\dagger c_k) k_1 a_k^* A_{k-1}^{\dagger *} A_{k-1}^\dagger \end{bmatrix},$$

where k_1 designates the quantity $(1 + d_k^* d_k)^{-1}$ and $a_k^* A_{k-1}^{\dagger *}$ is utilized in place of d_k^*. The expression in (2.8) exhibits the structure of the representations for the generalized inverse of matrices $A = [U, V]$.

Consider an arbitrary matrix $A = [U, V]$, where U and V have l and $k - l$ columns,[3] respectively, $0 \leqq l \leqq k$. Corresponding to c_k and k_1 in (2.8) let $C = (I - UU^\dagger)V$ and $K_1 = (I + V^* U^{\dagger *} U^\dagger V)^{-1}$, and let

$$(2.9) \qquad X_1 = \begin{bmatrix} U^\dagger - U^\dagger V C^\dagger - U^\dagger V(I - C^\dagger C) K_1 V^* U^{\dagger *} U^\dagger \\ C^\dagger + (I - C^\dagger C) K_1 V^* U^{\dagger *} U^\dagger \end{bmatrix}.$$

Then we have

THEOREM 1. *A necessary and sufficient condition that* $X_1 = A^\dagger$ *is that the matrices* $C^\dagger C$ *and* $V^* U^{\dagger *} U^\dagger V$ *commute.*

Proof. It will be shown that A and X_1 satisfy the defining equations (2.1), (2.2), (2.3), and (2.4), where the commutativity of $C^\dagger C$ and $V^* U^{\dagger *} U^\dagger V$ is utilized in order to conclude that $(X_1 A)^* = X_1 A$.

Using the definition of C and the relation $CC^\dagger C = C$ to simplify the resulting expression, block multiplication gives

$$(2.10) \qquad AX_1 = UU^\dagger + CC^\dagger.$$

Thus, since both $UU^\dagger$ and $CC^\dagger$ are hermitian $(AX_1)^* = AX_1$.

Now $U^\dagger C = (U^\dagger - U^\dagger U U^\dagger)V = \theta$ implies $C^\dagger U = \theta$, by Lemma 1(b) and (i). Whence

$$UU^\dagger V + CC^\dagger V = UU^\dagger V + CC^\dagger C = UU^\dagger V + C = V,$$

and the product $AX_1A = (AX_1)A$ becomes

$$AX_1A = [(UU^\dagger + CC^\dagger)U, (UU^\dagger + CC^\dagger)V] = [U, V] = A.$$

Similarly, $X_1AX_1 = X_1(AX_1)$ reduces to

$$X_1AX_1 = \begin{bmatrix} U^\dagger - U^\dagger V C^\dagger - U^\dagger V(I - C^\dagger C) K_1 V^* U^{\dagger *} U^\dagger \\ C^\dagger + (I - C^\dagger C) K_1 V^* U^{\dagger *} U^\dagger \end{bmatrix} = X_1,$$

since $U^\dagger(UU^\dagger + CC^\dagger) = U^\dagger$ and $C^\dagger(UU^\dagger + CC^\dagger) = C^\dagger$.

[3] In each representation the special cases $l = 0$ (that is, $A = V$) and $l = k$ (that is, $A = U$) will follow at once by inspection. It is tacitly assumed throughout the discussion that $0 < l < k$.

Finally, with $C^\dagger U = \theta$, $C^\dagger V = C^\dagger C$ and $U^{\dagger *}U^\dagger U = U^{\dagger *}(U^\dagger U)^* = (U^\dagger U U^\dagger)^* = U^{\dagger *}$, the product $X_1 A$ becomes

$$(2.11)\quad X_1 A = \begin{bmatrix} U^\dagger U - U^\dagger V(I - C^\dagger C)K_1 V^* U^{\dagger *} & U^\dagger V(I - C^\dagger C)K_1 \\ (I - C^\dagger C)K_1 V^* U^{\dagger *} & I - (I - C^\dagger C)K_1 \end{bmatrix},$$

where $K_1 V^* U^{\dagger *} U^\dagger V = I - K_1$ by definition of K_1. Suppose now that $C^\dagger C$ and $V^* U^{\dagger *} U^\dagger V$ commute. Then

$$(I - C^\dagger C)(I + V^* U^{\dagger *} U^\dagger V) = (I + V^* U^{\dagger *} U^\dagger V)(I - C^\dagger C),$$

and so

$$(2.12)\qquad K_1(I - C^\dagger C) = (I - C^\dagger C)K_1 .$$

Since both K_1 and $I - C^\dagger C$ are hermitian, this implies

$$(2.13)\qquad [K_1(I - C^\dagger C)]^* = K_1(I - C^\dagger C),$$

and it follows in (2.11) that $(X_1 A)^* = X_1 A$.

Thus we have shown that A and X_1 satisfy the relations $AX_1A = A$, $X_1 A X_1 = X_1$, $(AX_1)^* = AX_1$ and $(X_1 A)^* = X_1 A$, provided $C^\dagger C$ and $V^* U^{\dagger *} U^\dagger V$ commute, and so $X_1 = A^\dagger$.

Conversely, if $X_1 = A^\dagger$, then $(X_1 A)^* = X_1 A$, and (2.13) follows at once from (2.11). Hence (2.12) holds, and $C^\dagger C$ commutes with $V^* U^{\dagger *} U^\dagger V$.

The existence of matrices $A = [U, V]$ for which $C^\dagger C$ and $V^* U^{\dagger *} U^\dagger V$ do not commute can be shown by simple examples. Consequently, X_1 does not provide the most general form for $A^\dagger$. Before considering the general form, however, we will establish four corollaries to Theorem 1. Assume A has the form $A = [U, V]$, and again let $C = (I - UU^\dagger)V$ and $K_1 = (I + V^* U^{\dagger *} U^\dagger V)^{-1}$.

COROLLARY 1.1.

$$(2.14)\qquad A^\dagger = \begin{bmatrix} U^\dagger - U^\dagger V K_1 V^* U^{\dagger *} U^\dagger \\ C^\dagger + K_1 V^* U^{\dagger *} U^\dagger \end{bmatrix}$$

if and only if $C^\dagger C V^* U^{\dagger *} U^\dagger V = \theta$.

Proof. If $C^\dagger C V^* U^{\dagger *} U^\dagger V = \theta$, then $V^* U^{\dagger *} U^\dagger V$ and $C^\dagger C$ commute, and $X_1 = A^\dagger$. Also, $C^\dagger C V^* U^{\dagger *} U^\dagger V = \theta$ implies $C^{\dagger *} V^* U^{\dagger *} U^\dagger V C^\dagger = \theta$ and thus

$$(2.15)\qquad U^\dagger V C^\dagger = \theta.$$

Whereupon X_1 in (2.9) reduces at once to the right hand side of (2.14).

Conversely, if $A^\dagger$ has the form given in (2.14), then it follows from the equation $AA^\dagger A = A$ that

$$UU^\dagger V - UU^\dagger V K_1 V^* U^{\dagger *} U^\dagger V + V C^\dagger V + V K_1 V^* U^{\dagger *} U^\dagger V = V.$$

Using the relations $K_1 = I - K_1V^*U^{\dagger*}U^\dagger V$ and $C^\dagger V = C^\dagger C$, the definition of C now gives

$$C(I - K_1) + VC^\dagger C = C$$

and so

$$C^\dagger C(I - K_1) = \theta.$$

Hence

$$C^\dagger C(I + V^*U^{\dagger*}U^\dagger V) = C^\dagger C$$

and

$$C^\dagger CV^*U^{\dagger*}U^\dagger V = \theta.$$

Note in Corollary 1.1 that $C^\dagger CV^*U^{\dagger*}U^\dagger V = \theta$ is equivalent to the condition $VC^\dagger V = C$. If $C^\dagger CV^*U^{\dagger*}U^\dagger V = \theta$, then we have, using (2.15),

$$VC^\dagger V = VC^\dagger C = VC^\dagger C - UU^\dagger VC^\dagger C = CC^\dagger C = C.$$

Conversely, if $VC^\dagger V = C$, then

$$V^*U^{\dagger*}U^\dagger VC^\dagger C = V^*U^{\dagger*}U^\dagger VC^\dagger V = V^*U^{\dagger*}U^\dagger C = \theta,$$

and thus $C^\dagger CV^*U^{\dagger*}U^\dagger V = \theta$.

For the special case in which $C = \theta$, Corollary 1.1 reduces to

COROLLARY 1.2.

$$A^\dagger = \begin{bmatrix} U^\dagger - U^\dagger VK_1V^*U^{\dagger*}U^\dagger \\ K_1V^*U^{\dagger*}U^\dagger \end{bmatrix} \tag{2.16}$$

if and only if $C = \theta$.

Proof. That (2.14) reduces to (2.16) when $C = \theta$ is obvious. Conversely, if (2.16) holds, then $AA^\dagger A = A$ implies

$$UU^\dagger V - UU^\dagger VK_1V^*U^{\dagger*}U^\dagger V + VK_1V^*U^{\dagger*}U^{\dagger*}U^\dagger V = V,$$

which reduces to $C(I - K_1) = C$. Hence $CK_1 = \theta$ and so $C = \theta$.

Suppose now that $C^\dagger CV^*U^{\dagger*}U^\dagger V = V^*U^{\dagger*}U^\dagger V$. Then again $V^*U^{\dagger*}U^\dagger V$ and $C^\dagger C$ commute, and we obtain two more special cases of Theorem 1.

COROLLARY 1.3.

$$A^\dagger = \begin{bmatrix} U^\dagger - U^\dagger VC^\dagger \\ C^\dagger \end{bmatrix} \tag{2.17}$$

if and only if

$$C^\dagger CV^*U^{\dagger*}U^\dagger V = V^*U^{\dagger*}U^\dagger V. \tag{2.18}$$

Proof. From (2.1) and (2.3) it follows that $B^*BB^\dagger = B^*$ for every matrix B. Hence, taking $B = U^\dagger V$, the relation $C^\dagger CV^*U^{\dagger *}U^\dagger V = V^*U^{\dagger *}U^\dagger V$ implies that

$$C^\dagger CV^*U^{\dagger *} = V^*U^{\dagger *}, \tag{2.19}$$

and using (2.12) gives

$$(I - C^\dagger C)K_1V^*U^{\dagger *}U^\dagger = \theta.$$

Whereupon X_1 reduces to (2.17).

Conversely, if $A^\dagger$ has the form given in (2.17), then $A^\dagger A$ becomes

$$A^\dagger A = \begin{bmatrix} U^\dagger U & U^\dagger V(I - C^\dagger C) \\ \theta & C^\dagger C \end{bmatrix},$$

and $(A^\dagger A)^* = A^\dagger A$ implies $(I - C^\dagger C)V^*U^{\dagger *} = \theta$. This gives (2.19), from which the converse follows.

Analogous to the equivalence between the conditions $C^\dagger CV^*U^{\dagger *}U^\dagger V = \theta$ and $VC^\dagger V = C$ noted above, it is easily seen that $C^\dagger CV^*U^{\dagger *}U^\dagger V = V^*U^{\dagger *}U^\dagger V$ is equivalent to having $VC^\dagger V = V$. If $C^\dagger CV^*U^{\dagger *}U^\dagger V = V^*U^{\dagger *}U^\dagger V$, then it follows from (2.19) and the definition of C that

$$V^* - C^\dagger CV^*U^{\dagger *}U^* = V^* - V^*U^{\dagger *}U^*$$

or

$$V^* - C^\dagger C(V^* - C^*) = C^*.$$

Thus $V^* - C^\dagger CV^* = \theta$ and so $V = VC^\dagger C = VC^\dagger V$.

Conversely, if $VC^\dagger V = V$, then

$$V^*U^{\dagger *}U^\dagger V = V^*U^{\dagger *}U^\dagger VC^\dagger V = V^*U^{\dagger *}U^\dagger VC^\dagger C,$$

which implies

$$C^\dagger CV^*U^{\dagger *}U^\dagger V = V^*U^{\dagger *}U^\dagger V.$$

Note, in particular, that whenever C has full column rank we have $C^\dagger = (C^*C)^{-1}C^*$, by Lemma 1(d), and thus $C^\dagger C = I$. Hence $VC^\dagger V = VC^\dagger C = V$, and $A^\dagger$ has the form given in (2.17). Clearly, this is the case in Greville's form for $A_k^\dagger$, (2.5), when $c_k \neq \theta$ and $b_k = c_k^\dagger$ in (2.6). On the other hand, when $c_k = \theta$, Corollary 1.2 is applicable, and the form for $A_k^\dagger$ with $b_k = (1 + d_k^* d_k)^{-1} d_k^* A_{k-1}^\dagger$ follows directly from (2.16).

For the special case in which $C = V$, $VC^\dagger V = VV^\dagger V = V$, and Corollary 1.3 reduces to

COROLLARY 1.4.

$$(2.20) \qquad A^\dagger = \begin{bmatrix} U^\dagger \\ V^\dagger \end{bmatrix}$$

if and only if $C = V$.

Proof. If $C = V$, then $UU^\dagger V = \theta$ and

$$U^\dagger VC^\dagger = U^\dagger UU^\dagger VC^\dagger = \theta$$

in the form for $A^\dagger$, (2.17), which gives (2.20).

Conversely, if (2.20) holds, then it follows from the relation $AA^\dagger A = A$ that

$$UU^\dagger V + VV^\dagger V = V.$$

Hence $UU^\dagger V = \theta$ and $C = V$.

Let us now consider general forms for $A^\dagger$ in which it is not required that $C^\dagger C$ and $V^*U^{\dagger *}U^\dagger V$ commute. Let $\tilde{C}$ designate the expression

$$\tilde{C} = (I - VV^\dagger)U$$

obtained by interchanging the roles of U and V in $C = (I - UU^\dagger)V$. Also, let K and $\tilde{K}$ designate the dual expressions defined by

$$(2.21) \qquad K = [I + (I - C^\dagger C)V^*U^{\dagger *}U^\dagger V(I - C^\dagger C)]^{-1}$$

and

$$(2.22) \qquad \tilde{K} = [I + (I - \tilde{C}^\dagger \tilde{C})U^*V^{\dagger *}V^\dagger U(I - \tilde{C}^\dagger \tilde{C})]^{-1}.$$

(Note that both K and $\tilde{K}$, inverses of positive definite matrices, exist for every U and V.) Then we have

THEOREM 2. *The generalized inverse of any matrix,* $A = [U, V]$ *can be written in the following equivalent forms.*

$$\text{(a)} \quad A^\dagger = \begin{bmatrix} U^\dagger - U^\dagger VC^\dagger - U^\dagger V(I - C^\dagger C)KV^*U^{\dagger *}U^\dagger(I - VC^\dagger) \\ C^\dagger + (I - C^\dagger C)KV^*U^{\dagger *}U^\dagger(I - VC^\dagger) \end{bmatrix};$$

$$\text{(b)} \quad A^\dagger = \begin{bmatrix} U^\dagger - U^\dagger VC^\dagger - U^\dagger V(I - C^\dagger C)KV^*U^{\dagger *}U^\dagger(I - VC^\dagger) \\ V^\dagger - V^\dagger U\tilde{C}^\dagger - V^\dagger U(I - \tilde{C}^\dagger \tilde{C})\tilde{K}U^*V^{\dagger *}V^\dagger(I - U\tilde{C}^\dagger) \end{bmatrix};$$

$$\text{(c)} \quad A^\dagger = \begin{bmatrix} \tilde{C}^\dagger + (I - \tilde{C}^\dagger \tilde{C})\tilde{K}U^*V^{\dagger *}V^\dagger(I - U\tilde{C}^\dagger) \\ C^\dagger + (I - C^\dagger C)KV^*U^{\dagger *}U^\dagger(I - VC^\dagger) \end{bmatrix}.$$

Proof. Let X_0 designate the matrix

$$(2.23) \qquad X_0 = \begin{bmatrix} U^\dagger - U^\dagger VC^\dagger - U^\dagger V(I - C^\dagger C)L \\ C^\dagger + (I - C^\dagger C)L \end{bmatrix}$$

obtained from X_1 in (2.9) by replacing the quantity $K_1V^*U^{\dagger *}U^\dagger$ by an arbitrary matrix, L, of the same size. Then it follows immediately, using block multiplication, the definition of C, and the relation $C(I - C^\dagger C) = \theta$, that $AX_0 = UU^\dagger + CC^\dagger = AX_1$, and so we have $(AX_0)^* = AX_0$ and $AX_0A = A$ from the proof of Theorem 1.

Now forming $X_0AX_0 = X_0(AX_0)$, it is clear that $X_0AX_0 = X_0$ provided L satisfies

$$L(UU^\dagger + CC^\dagger) = L. \tag{2.24}$$

Similarly, forming X_0A gives

$$X_0A = \begin{bmatrix} U^\dagger U - U^\dagger V(I - C^\dagger C)LU & U^\dagger V(I - C^\dagger C)(I - LV) \\ (I - C^\dagger C)LU & C^\dagger C + (I - C^\dagger C)LV \end{bmatrix}$$

upon simplification of the submatrices, and it follows that $(X_0A)^* = X_0A$ provided L satisfies

$$[U^\dagger V(I - C^\dagger C)(I - LV)]^* = (I - C^\dagger C)LU \tag{2.25}$$

and also that both $U^\dagger V(I - C^\dagger C)LU$ and $(I - C^\dagger C)LV$ are hermitian. We will now show that the expression

$$L = KV^*U^{\dagger *}U^\dagger(I - VC^\dagger), \tag{2.26}$$

with K as defined in (2.21), satisfies these conditions.

Since $U^\dagger(I - VC^\dagger)UU^\dagger = U^\dagger$ and $U^\dagger(I - VC^\dagger)CC^\dagger = -U^\dagger VC^\dagger$, then L in (2.26) satisfies (2.24). Next observe that since $I - C^\dagger C$ is idempotent, it commutes with the matrix $I + (I - C^\dagger C)V^*U^{\dagger *}U^\dagger V(I - C^\dagger C)$, and thus with K. Whereupon, with both $I - C^\dagger C$ and K hermitian,

$$[(I - C^\dagger C)K]^* = (I - C^\dagger C)K,$$

and so

$$U^\dagger V(I - C^\dagger C)LU = U^\dagger V(I - C^\dagger C)KV^*U^{\dagger *}$$

is hermitian. Moreover, we have

$$(I - C^\dagger C)LV = (I - C^\dagger C)K(I - C^\dagger C)V^*U^{\dagger *}U^\dagger V(I - C^\dagger C),$$

or

$$(I - C^\dagger C)LV = (I - C^\dagger C)(I - K),$$

which implies that $(I - C^\dagger C)LV$ is hermitian. Finally, since

$$U^\dagger V(I - C^\dagger C)(I - LV) = U^\dagger V(I - C^\dagger C)K = [(I - C^\dagger C)KV^*U^{\dagger *}]^*$$

and

$$(I - C^\dagger C)LU = (I - C^\dagger C)KV^*U^{\dagger *},$$

then (2.25) holds for this choice of L.

Thus it has been shown that X_0 and A satisfy the relations $AX_0A = A$, $X_0AX_0 = X_0$, $(AX_0)^* = AX_0$, and $(X_0A)^* = X_0A$, provided L has the form given in (2.26), and so $X_0 = A^\dagger$. The form for $A^\dagger$ in (a) is obtained by replacing L in (2.23) by the expression in (2.26).

The forms for $A^\dagger$ in (b) and (c) are now easily established. Let $\tilde{A}$ designate the matrix $\tilde{A} = [V, U]$. Then it follows from (a) that $\tilde{A}^\dagger$ can be written as

$$(2.27)\qquad \tilde{A}^\dagger = \begin{bmatrix} V^\dagger - V^\dagger U\tilde{C}^\dagger - V^\dagger U(I - \tilde{C}^\dagger\tilde{C})\tilde{K}U^*V^{\dagger *}V^\dagger(I - U\tilde{C}^\dagger) \\ \tilde{C}^\dagger + (I - \tilde{C}^\dagger\tilde{C})\tilde{K}U^*V^{\dagger *}V^\dagger(I - U\tilde{C}^\dagger) \end{bmatrix},$$

where $\tilde{C}$ and $\tilde{K}$ are the dual expressions obtained from C and K by interchanging the roles of U and V. Since A and $\tilde{A}$ differ only by the order in which columns are written, there is a unitary permutation matrix P, say, such that $A = \tilde{A}P$. Then we have $A^\dagger = P^*\tilde{A}^\dagger$, by Lemma 1(f). Now P as a right multiplier permutes columns of $\tilde{A}$, and P^* as a left multiplier permutes rows of $\tilde{A}^\dagger$ in the same order, and it follows from (2.27) that $A^\dagger$ can be written in the form

$$(2.28)\qquad A^\dagger = \begin{bmatrix} \tilde{C}^\dagger + (I - \tilde{C}^\dagger\tilde{C})\tilde{K}U^*V^{\dagger *}V^\dagger(I - U\tilde{C}^\dagger) \\ V^\dagger - V^\dagger U\tilde{C}^\dagger - V^\dagger U(I - \tilde{C}^\dagger\tilde{C})\tilde{K}U^*V^{\dagger *}V^\dagger(I - U\tilde{C}^\dagger) \end{bmatrix}.$$

But $A^\dagger$ is unique. The forms for $A^\dagger$ in (b) and (c) are obtained now by equating the corresponding expressions for submatrices in (a) and (2.28).

It also follows from the symmetry exhibited by the expressions for $A^\dagger$ in Theorem 2(a) and (2.28) that Theorem 1 and each of its corollaries has a corresponding dual form in which the roles of U and V are interchanged.

Numerical examples are easily constructed to illustrate each of the forms which have been developed for $A^\dagger$. In fact, it was shown in [10] that such examples can be constructed using only matrices with elements $a_{ij} = 0$ and $a_{ij} = 1$.

3. Representations for $U^\dagger$ in terms of submatrices of $A^\dagger$. Consider an arbitrary matrix $A = [U, V]$, and assume $A^\dagger$ is known. Partition $A^\dagger$ as $A^\dagger = \begin{bmatrix} G \\ H \end{bmatrix}$ where G and H have the size of U^* and V^*, respectively. Then Theorem 3 provides an expression for $U^\dagger$ in terms of G, H, and related matrices.

THEOREM 3.

$$U^\dagger = G[I + V(I - HV)^\dagger H]$$
$$\cdot\{I - [H - (I - HV)(I - HV)^\dagger H]^\dagger[H - (I - HV)(I - HV)^\dagger H]\}.$$

Proof. We know from the expression in Theorem 2(a) that

$$G = U^\dagger - U^\dagger VC^\dagger - U^\dagger V(I - C^\dagger C)KV^*U^{\dagger *}U^\dagger(I - VC^\dagger)$$

and

$$H = C^\dagger + (I - C^\dagger C)KV^*U^{\dagger *}U^\dagger(I - VC^\dagger)$$

in the partition of $A^\dagger$ corresponding to the partition $A = [U, V]$. Then it follows using the relations employed in the proof of Theorem 2(a) that

$$GV = U^\dagger V(I - C^\dagger C)K \tag{3.1}$$

and

$$I - HV = (I - C^\dagger C)K. \tag{3.2}$$

Further, $I - C^\dagger C$ is idempotent and hermitian, and commutes with K. Therefore, since K is nonsingular, we have

$$(I - HV)^\dagger = K^{-1}(I - C^\dagger C),$$

by Lemma 1(j), which combined with (3.1) and (3.2) to give

$$GV(I - HV)^\dagger = U^\dagger V(I - C^\dagger C) \tag{3.3}$$

and

$$(I - HV)(I - HV)^\dagger = I - C^\dagger C. \tag{3.4}$$

Now since $C^\dagger CC^\dagger = C^\dagger$,

$$GV(I - HV)^\dagger H = U^\dagger V(I - C^\dagger C)KV^*U^{\dagger *}U^\dagger(I - VC^\dagger),$$

and so

$$G[I + V(I - HV)^\dagger H] = U^\dagger - U^\dagger VC^\dagger. \tag{3.5}$$

Moreover

$$(I - HV)(I - HV)^\dagger H = (I - C^\dagger C)KV^*U^{\dagger *}U^\dagger(I - VC^\dagger),$$

and thus

$$H - (I - HV)(I - HV)^\dagger H = C^\dagger. \tag{3.6}$$

Finally, since $U^\dagger C = \theta$, we have

$$U^\dagger = G[I + V(I - HV)^\dagger H](I - CC^\dagger)$$

from (3.5), which combines with (3.6) and the relation $C^{\dagger\dagger} = C$ to give the stated form for $U^\dagger$.

The following corollaries provide special forms for $U^\dagger$ corresponding to the forms for $A^\dagger$ in corollaries to Theorem 1. This correspondence is apparent by observing that the relations satisfied by V and C are simply the alternative statements of the conditions on $C^\dagger C$ and $V^*U^{\dagger*}U^\dagger V$ which were noted in §2.

COROLLARY 3.1. $U^\dagger = G[I + V(I - HV)^\dagger H]$ *if and only if* $VC^\dagger V = C$.

Proof. It follows from (3.5) that $U^\dagger = G[I + V(I - HV)^\dagger H]$ if and only if $U^\dagger VC^\dagger = \theta$. But this implies $VC^\dagger V = C$, and conversely.

COROLLARY 3.2. $U^\dagger = G(I - H^\dagger H)$ *if and only if* $VC^\dagger V = V$.

Proof. From Corollary 1.3 we have $G = U^\dagger - U^\dagger VC^\dagger$ and $H = C^\dagger$ if $VC^\dagger V = V$. Hence $G[I + V(I - HV)^\dagger H] = G$ in (3.5) and $H - (I - HV)(I - HV)^\dagger H = H$ in (3.6), and the general form for $U^\dagger$ in Theorem 3 reduces immediately to the above expression.

Conversely, since we can write the general form for G as $G = U^\dagger(I - VH)$, then

$$U^\dagger = U^\dagger(I - VH)(I - H^\dagger H) = U^\dagger(I - H^\dagger H)$$

if $U^\dagger = G(I - H^\dagger H)$. This gives $U^\dagger H^\dagger H = \theta$ and so

$$UU^\dagger H^* = UU^\dagger H^* H^{\dagger*} H^* = UU^\dagger H^\dagger HH^* = \theta,$$

by (2.1) and (2.3). Then $HUU^\dagger = \theta$, since $UU^\dagger$ is hermitian, and

$$\begin{aligned}\theta = HUU^\dagger = C^\dagger UU^\dagger &+ (I - C^\dagger C)KV^*U^{\dagger*}U^\dagger(I - VC^\dagger)UU^\dagger \\ &= (I - C^\dagger C)KV^*U^{\dagger*}U^\dagger,\end{aligned}$$

which implies $G = U^\dagger - U^\dagger VC^\dagger$ and $H = C^\dagger$. Whence $C^\dagger CV^*U^{\dagger*}U^\dagger V = V^*U^{\dagger*}U^\dagger V$, by Corollary 1.3, and thus $VC^\dagger V = V$.

COROLLARY 3.3. $U^\dagger = G$ *if and only if* $C = V$.

Proof. That $C = V$ implies $U^\dagger = G$ follows directly from Corollary 1.4.

Conversely, if $U^\dagger = G$, we have

$$U^\dagger VC^\dagger + U^\dagger V(I - C^\dagger C)KV^*U^{\dagger*}U^\dagger(I - VC^\dagger) = \theta, \tag{3.7}$$

by definition of G. Multiplying (3.7) on the right by $VC^\dagger$ then gives $U^\dagger VC^\dagger = \theta$, and thus

$$U^\dagger VK_1V^*U^{\dagger*}U^\dagger = \theta,$$

where $K_1 = (I + V^*U^{\dagger*}U^\dagger V)^{-1}$. Therefore $U^\dagger V(I - K_1) = \theta$ and $U^\dagger VV^*U^{\dagger*}U^\dagger V = \theta$, from which it follows that $U^\dagger V = \theta$ and $C = V$.

Observe in Corollary 3.1 that $VC^\dagger V = C$ if $C = \theta$. In this case $H = K_1V^*U^{\dagger*}U^\dagger$, and $I - HV = K_1$ is nonsingular, by definition of K_1. Conversely, if $I - HV$ is nonsingular, then $C^\dagger C = \theta$, by (3.4) which implies $C = \theta$. Thus it follows that $I - HV$ nonsingular is a necessary and sufficient condition that $A^\dagger$ has the form given in Corollary 1.2. Since we can have $VC^\dagger V = C$ but $C \neq \theta$, $I - HV$ nonsingular is only a sufficient condition that $U^\dagger$ has the form given in Corollary 3.1. In Corollaries 3.2

and 3.3, however, the necessary and sufficient conditions that $U^\dagger$ has the simplified forms can be restated in terms of V and H. This gives Corollaries 3.2(a) and 3.3(a).

COROLLARY 3.2(a). $U^\dagger = G(I - H^\dagger H)$ *if and only if* HV *is idempotent.*

Proof. From Corollary 3.2 it follows that we only need to show that HV idempotent implies $VC^\dagger V = V$, and conversely.

If $VC^\dagger V = V$, then $H = C^\dagger$, by Corollary 1.3 and

$$(HV)^2 = C^\dagger V C^\dagger V = C^\dagger V = HV.$$

Conversely, HV idempotent implies that $I - HV$ is idempotent. Therefore, since we also have $I - HV = (I - C^\dagger C)K$ hermitian, from the proof of Theorem 2(a), $(I - HV)^\dagger = I - HV$, by Lemma 1(g). Now $(I - HV)^\dagger = K^{-1}(I - C^\dagger C)$, and thus

$$K^{-1}(I - C^\dagger C) = (I - HV)(I - HV)^\dagger = I - C^\dagger C,$$

by (3.4). This gives

$$(I - C^\dagger C)V^* U^{\dagger *} U^\dagger V(I - C^\dagger C) = \theta,$$

by definition of K. Then $(I - C^\dagger C)V^* U^{\dagger *} = \theta$, and it follows from the proof of Corollary 1.3 and the remarks immediately thereafter that $VC^\dagger V = V$.

COROLLARY 3.3(a). $U^\dagger = G$ *if and only if* HV *is idempotent and* VH *is hermitian.*

Proof. The result follows from Corollary 3.3 by showing that HV idempotent and VH hermitian imply $C = V$, and conversely.

If $C = V$, then $H = V^\dagger$, by Corollary 1.4, and we have $HV = V^\dagger V$ idempotent and $VH = VV^\dagger$ hermitian from (2.1) and (2.3).

Conversely, suppose that HV is idempotent and VH is hermitian. Now we know from the proof of Corollary 3.2(a) that HV idempotent implies $VC^\dagger V = V$. Hence $H = C^\dagger$, by Corollary 1.3, and so $VHV = V$,

$$HVH = C^\dagger V C^\dagger = C^\dagger C C^\dagger = C^\dagger = H,$$

and

$$(HV)^* = (C^\dagger C)^* = C^\dagger C = HV.$$

Then with $(VH)^* = VH$, by hypothesis, H and V satisfy the defining equations for the generalized inverse. Therefore, $H = V^\dagger$. But $H = C^\dagger$. Hence $C = V$, by Lemma 1(b).

For the special case $A_k = [A_{k-1}, a_k]$ with $A_k^\dagger = \begin{bmatrix} G_{k-1} \\ h_k \end{bmatrix}$ in which G_{k-1} has $k - 1$ rows and h_k is a single row, Corollaries 3.1 and 3.2(a) combine to give a form for $A_{k-1}^\dagger$ corresponding to Greville's representation for $A_k^\dagger$ in

(2.5) and (2.6). Since $h_k a_k$ is a scalar, we have

$$A_{k-1}^{\dagger} = \begin{cases} G_{k-1}[I + (1 - h_k a_k)^{-1} a_k h_k], & \text{if } h_k a_k \neq 1, \\ G_{k-1}(I - h_k^{\dagger} h_k), & \text{if } h_k a_k = 1. \end{cases}$$

A form for $V^{\dagger}$ corresponding to each representation for $U^{\dagger}$ follows at once from the dual symmetry noted in §2. In each case we simply interchange U and V, G and H, and replace C by $\tilde{C}$.

REFERENCES

[1] E. H. Moore, *General analysis*, Mem. Amer. Philos. Soc., I, (1935).

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

[3] R. Rado, *Note on generalized inverses of matrices*, Ibid., 52 (1956), pp. 600–601.

[4] A. Bjerhammar, *Rectangular reciprocal matrices with special reference to geodetic calculations*, Bull. Géodésique, (1951), pp. 188–220.

[5] C. Lanczos, *Linear systems in self adjoint form*, Amer. Math. Monthly, 65 (1958), pp. 665–679.

[6] T. N. E. Greville, *The pseudoinverse of a rectangular or singular matrix and its application to the solution of systems of linear equations*, SIAM Rev., 1 (1959), pp. 38–43.

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

[8] A. Ben-Israel and S. J. Wersan, *A least-square method for computing the generalized-inverse of an arbitrary complex matrix*, O.N.R. Research Memo. No. 61, The Technological Institute, Northwestern University, (1962).

[9] R. E. Cline, *On the computation of the generalized inverse, $A^{\dagger}$, of an arbitrary matrix, A, and the use of certain associated eigenvectors in solving the allocation problem*, Preliminary Report, Statistical and Computing Lab., Purdue University, 1958.

[10] ———, *Representations for the generalized inverse of matrices with applications in linear programming*, Doctoral Thesis, Purdue University, 1963.

[11] G. G. den Broeder, Jr., and A. Charnes, *Contributions to the theory of a generalized inverse for matrices*, Res. Memo., ONR Project on Planning and Control of Industrial Operations, Purdue University, 1957.

[12] G. G. den Broeder, Jr., *Contributions to the theory of a generalized inverse for matrices*, unpublished, Purdue University.

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

[14] ———, *Note on fitting of functions of several independent variables*, this Journal, 9 (1961), pp. 109–115.

[15] R. Penrose, *On best approximate solutions of linear matrix equations*, Proc. Cambridge Philos. Soc., 52 (1956), pp. 17–19.

[16] L. D. Pyle, *A gradient projection method for solving programming problems using the generalized inverse, $A^{\dagger}$, and the eigenvectors $e^{(1)}, \cdots, e^{(n)}$ of $A^{\dagger}A$*, Preliminary Report, Statistical and Computing Lab., Purdue University, 1958.

[17] ———, *The generalized inverse in linear programming*, Doctoral Thesis, Purdue University, 1960.