A GENERALIZED INVERSE FOR MATRICES

By R. PENROSE

Communicated by J. A. Todd

Received 26 July 1954

This paper describes a generalization of the inverse of a non-singular matrix, as the unique solution of a certain set of equations. This generalized inverse exists for any (possibly rectangular) matrix whatsoever with complex elements‡. It is used here for solving linear matrix equations, and among other applications for finding an expression for the principal idempotent elements of a matrix. Also a new type of spectral decomposition is given.

In another paper its application to substitutional equations and the value of hermitian idempotents will be discussed.

Notation. Capital letters always denote matrices (not necessarily square) with complex elements. The conjugate transpose of the matrix A is written A^*. Small letters are used for column vectors (with an asterisk for row vectors) and small Greek letters for complex numbers.

The following properties of the conjugate transpose will be used:

$$A^{**} = A,$$

$$(A + B)^* = A^* + B^*,$$

$$(\lambda A)^* = \bar{\lambda} A^*,$$

$$(BA)^* = A^* B^*,$$

$$AA^* = 0 \ implies \ A = 0.$$

The last of these follows from the fact that the trace of AA^* is the sum of the squares of the moduli of the elements of A. From the last two we obtain the rule

$$BAA^* = CAA^* \ implies \ BA = CA, \tag{1}$$

since $\qquad (BAA^* - CAA^*)(B - C)^* = (BA - CA)(BA - CA)^*.$

Similarly $\qquad BA^*A = CA^*A \ implies \ BA^* = CA^*. \tag{2}$

THEOREM 1. *The four equations*

$$AXA = A, \tag{3}$$

$$XAX = X, \tag{4}$$

$$(AX)^* = AX, \tag{5}$$

$$(XA)^* = XA, \tag{6}$$

have a unique solution for any A.

‡ Matrices over more general rings will be considered in a later paper.

Proof. I first show that equations (4) and (5) are equivalent to the single equation

$$XX^*A^* = X. \tag{7}$$

Equation (7) follows from (4) and (5), since it is merely (5) substituted in (4). Conversely, (7) implies $AXX^*A^* = AX$, the left-hand side of which is hermitian. Thus (5) follows, and substituting (5) in (7) we get (4). Similarly, (3) and (6) can be replaced by the equation

$$XAA^* = A^*. \tag{8}$$

Thus it is sufficient to find an X satisfying (7) and (8). Such an X will exist if a B can be found satisfying

$$BA^*AA^* = A^*.$$

For then $X = BA^*$ satisfies (8). Also, we have seen that (8) implies $A^*X^*A^* = A^*$ and therefore $BA^*X^*A^* = BA^*$. Thus X also satisfies (7).

Now the expressions A^*A, $(A^*A)^2$, $(A^*A)^3$, ... cannot all be linearly independent, i.e. there exists a relation

$$\lambda_1 A^*A + \lambda_2 (A^*A)^2 + \ldots + \lambda_k (A^*A)^k = 0, \tag{9}$$

where $\lambda_1, \ldots, \lambda_k$ are not all zero. Let λ_r be the first non-zero λ and put

$$B = -\lambda_r^{-1}\{\lambda_{r+1}I + \lambda_{r+2}A^*A + \ldots + \lambda_k (A^*A)^{k-r-1}\}.$$

Thus (9) gives $B(A^*A)^{r+1} = (A^*A)^r$, and applying (1) and (2) repeatedly we obtain $BA^*AA^* = A^*$, as required.

To show that X is unique, we suppose that X satisfies (7) and (8) and that Y satisfies $Y = A^*Y^*Y$ and $A^* = A^*AY$. These last relations are obtained by respectively substituting (6) in (4) and (5) in (3). (They are (7) and (8) with Y in place of X and the reverse order of multiplication and must, by symmetry, also be equivalent to (3), (4), (5) and (6).) Now

$$X = XX^*A^* = XX^*A^*AY = XAY = XAA^*Y^*Y = A^*Y^*Y = Y.$$

The unique solution of (3), (4), (5) and (6) will be called the generalized inverse of A (abbreviated g.i.) and written $X = A^\dagger$. (Note that A need not be a square matrix and may even be zero.) I shall also use the notation $\lambda^\dagger$ for scalars, where $\lambda^\dagger$ means λ^{-1} if $\lambda \neq 0$ and 0 if $\lambda = 0$.

In the calculation of $A^\dagger$ it is only necessary to solve the two unilateral linear equations $XAA^* = A^*$ and $A^*AY = A^*$. By putting $A^\dagger = XAY$ and using the fact that XA and AY are hermitian and satisfy $AXA = A = AYA$ we observe that the four relations $AA^\dagger A = A$, $A^\dagger AA^\dagger = A^\dagger$, $(AA^\dagger)^* = AA^\dagger$ and $(A^\dagger A)^* = A^\dagger A$ are satisfied. Relations satisfied by $A^\dagger$ include

$$A^\dagger A^{\dagger *}A^* = A^\dagger = A^*A^{\dagger *}A^\dagger \left.\right\}$$

and

$$A^\dagger AA^* = A^* = A^*AA^\dagger, \left.\right\} \tag{10}$$

these being (7), (8) and their reverses.

Note that the properties $(A+B)^* = A^* + B^*$ and $(\lambda A)^* = \bar{\lambda} A^*$ were not used in the proof of Theorem 1.

LEMMA. 1·1 $A^{\dagger\dagger} = A$;

1·2 $A^{*\dagger} = A^{\dagger *}$;

1·3 if A is non-singular $A^{\dagger} = A^{-1}$;

1·4 $(\lambda A)^{\dagger} = \lambda^{\dagger} A^{\dagger}$;

1·5 $(A^*A)^{\dagger} = A^{\dagger} A^{\dagger *}$;

1·6 if U and V are unitary $(UAV)^{\dagger} = V^* A^{\dagger} U^*$;

1·7 if $A = \Sigma A_i$, where $A_i A_j^* = 0$ and $A_i^* A_j = 0$, whenever $i \neq j$ then
   $A^{\dagger} = \Sigma A_i^{\dagger}$;

1·8 if A is normal $A^{\dagger} A = A A^{\dagger}$ and $(A^n)^{\dagger} = (A^{\dagger})^n$;

1·9 A, A^*A, $A^{\dagger}$ and $A^{\dagger} A$ all have rank equal to trace $A^{\dagger} A$.

Proof. Since the g.i. of a matrix is always unique, relations $1 \cdot 1, \ldots, 1 \cdot 7$ can be verified by substituting the right-hand side of each in the defining relations for the required g.i. in each case. For 1·5 we require (10) and for 1·7 we require the fact that $A_i A_j^{\dagger} = 0$ and $A_i^{\dagger} A_j = 0$ if $i \neq j$. This follows from $A_j^{\dagger} = A_j^* A_j^{\dagger *} A_j^{\dagger}$ and $A_i^{\dagger} = A_i^{\dagger} A_i^{\dagger *} A_i^*$. The first part of 1·8, of which the second part is a consequence, follows from 1·5 and (10), since $A^{\dagger} A = (A^*A)^{\dagger} A^*A$ and $A A^{\dagger} = (AA^*)^{\dagger} AA^*$. (The condition for A to be normal is $A^*A = AA^*$.)‡ To prove 1·9 we observe that each of A, A^*A, $A^{\dagger}$ and $A^{\dagger} A$ can be obtained from the others by pre- and post-multiplication by suitable matrices. Thus their ranks must all be equal and $A^{\dagger} A$ being idempotent has rank equal to its trace.

Relation 1·5 provides a method of calculating the g.i. of a matrix A from the g.i. of A^*A, since $A^{\dagger} = (A^*A)^{\dagger} A^*$. Now A^*A, being hermitian, can be reduced to diagonal form by a unitary transformation. Hence $A^*A = UDU^*$, where U is unitary and $D = \operatorname{diag}(\alpha_1, \ldots, \alpha_n)$. Thus $D^{\dagger} = \operatorname{diag}(\alpha_1^{\dagger}, \ldots, \alpha_n^{\dagger})$ and $(A^*A)^{\dagger} = UD^{\dagger}U^*$ by 1·6.

An alternative proof of the existence of the g.i. of a square matrix can be obtained from Autonne's (1) result that any square matrix A can be written in the form VBW, where V and W are unitary and B is diagonal§. Now $B^{\dagger}$ exists since it can be defined similarly to $D^{\dagger}$ above. It is thus sufficient to verify that $A^{\dagger} = W^* B^{\dagger} V^*$ as in 1·6. A rectangular matrix can be treated by bordering it with zeros to make it square.

Notice that $A^{\dagger}$ is a continuous function of A if the rank of A is kept fixed, since in the singular case the polynomial in (9) could be taken to be the characteristic function of A^*A. The value of r, being the nullity of A^*A (as A^*A is hermitian), is thus determined by the rank of A (by 1·9). Also $A^{\dagger}$ varies discontinuously when the rank of A changes since rank $A = $ trace $A^{\dagger} A$ by 1·9.

‡ This result, like several of the earlier ones, is a simple consequence of the construction of $A^{\dagger}$ given in Theorem 1. It seems desirable, however, to show that it is also a direct consequence of (3), (4), (5) and (6).

§ A simple proof of this result is as follows: since AA^* and A^*A are both hermitian and have the same eigenvalues there exists a unitary matrix T such that $TAA^*T^* = A^*A$. Hence TA is normal and therefore diagonable by unitary transformation.

THEOREM 2. *A necessary and sufficient condition for the equation* $AXB = C$ *to have a solution is*

$$AA^{\dagger}CB^{\dagger}B = C,$$

in which case the general solution is

$$X = A^{\dagger}CB^{\dagger} + Y - A^{\dagger}AYBB^{\dagger},$$

where Y *is arbitrary.*

Proof. Suppose X satisfies $AXB = C$. Then

$$C = AXB = AA^{\dagger}AXBB^{\dagger}B = AA^{\dagger}CB^{\dagger}B.$$

Conversely, if $C = AA^{\dagger}CB^{\dagger}B$ then $A^{\dagger}CB^{\dagger}$ is a particular solution of $AXB = C$ (i.e. we may treat $A^{\dagger}$ formally as an inverse of A in order to find a particular solution).

For the general solution we must solve $AXB = 0$. Now any expression of the form $X = Y - A^{\dagger}AYBB^{\dagger}$ satisfies $AXB = 0$ and conversely if $AXB = 0$ then $X = X - A^{\dagger}AXBB^{\dagger}$. Theorem 2 follows. Note that the only property required of $A^{\dagger}$ for Theorem 2 is $AA^{\dagger}A = A$.

COROLLARY 1. *The general solution of the vector equation* $Px = c$ *is*

$$x = P^{\dagger}c + (I - P^{\dagger}P)y,$$

where y *is arbitrary, provided that the equation has a solution.*‡

The general linear matrix equation $A_{(1)}XB_{(1)} + \dots + A_{(m)}XB_{(m)} = C$ can also be solved by using a g.i., since this equation can be written in the form

$$\sum_{j,k} \left\{ \sum_{r=1}^{m} \alpha_{(r)ij}\beta_{(r)kl} \right\} \xi_{jk} = \gamma_{il},$$

where $A_{(r)} = (\alpha_{(r)ij})$, etc. If X and C are regarded as vectors, this is an equation of the type considered in Corollary 1.

COROLLARY 2. *A necessary and sufficient condition for the equations* $AX = C$, $XB = D$ *to have a common solution is that each equation should individually have a solution and that* $AD = CB$.

Proof. The condition is obviously necessary. To show that it is sufficient, put

$$X = A^{\dagger}C + DB^{\dagger} - A^{\dagger}ADB^{\dagger},$$

which is a solution if the required conditions $AA^{\dagger}C = C$, $DB^{\dagger}B = D$, $AD = CB$ are satisfied.

This result is due to Cecioni (3).

Hermitian idempotents. A *hermitian idempotent* matrix (abbreviated h.i.) is one satisfying $EE^* = E$, that is, $E = E^*$ and $E^2 = E$.

LEMMA. 2·1 $A^{\dagger}A, AA^{\dagger}, I - A^{\dagger}A, I - AA^{\dagger}$ *are all h.i.;*

2·2 *if E is h.i.* $E^{\dagger} = E$;

2·3 *K is idempotent if and only if there exist h.i.'s E and F such that* $K = (FE)^{\dagger}$ *in which case* $K = EKF$.

‡ See also Bjerhammar (2).

Proof. The proofs of 2·1 and 2·2 are evident. To prove 2·3 suppose $K^2 = K$, then, by 1·1, $K = \{(K^\dagger K)(KK^\dagger)\}^\dagger$. Thus $E = KK^\dagger$ and $F = K^\dagger K$ will do. Conversely, if $K = (FE)^\dagger$ then K is of the form $EFPEF$ (since $Q^\dagger = Q^*(Q^{\dagger*}Q^\dagger Q^{\dagger*})Q^*$ by (10)). Hence $K = EKF$, so that $K^2 = E(FE)^\dagger FE(FE)^\dagger F = E(FE)^\dagger F = K$.

Principal idempotent elements of a matrix. For any square matrix A there exists a unique set of matrices K_λ defined for each complex number λ such that

$$K_\lambda K_\mu = \delta_{\lambda\mu} K_\lambda, \tag{11}$$

$$\Sigma K_\lambda = I, \tag{12}$$

$$A K_\lambda = K_\lambda A, \tag{13}$$

$$(A - \lambda I) K_\lambda \text{ is nilpotent,} \tag{14}$$

the non-zero K_λ's being the *principal idempotent elements* of A (see Wedderburn (7)).‡ Unless λ is an eigenvalue of A, K_λ is zero so that the sum in (12) is finite, though summed over all complex λ. The following theorem gives an 'explicit' formula for K_λ in terms of g.i.'s.

THEOREM 3. *If*

$$E_\lambda = I - \{(A - \lambda I)^n\}^\dagger (A - \lambda I)^n \quad and \quad F_\lambda = I - (A - \lambda I)^n \{(A - \lambda I)^n\}^\dagger,$$

where n is sufficiently large (e.g. the order of A), then the principal idempotent elements of A are given by $K_\lambda = (F_\lambda E_\lambda)^\dagger$, and n can be taken as unity if and only if A is diagonable.

Proof. First, suppose A is diagonable and put

$$E_\lambda = I - (A - \lambda I)^\dagger (A - \lambda I)$$

and

$$F_\lambda = I - (A - \lambda I)(A - \lambda I)^\dagger.$$

Then E_λ and F_λ are h.i. by 2·1 and are zero unless λ is an eigenvalue of A by 1·3.

Now

$$(A - \mu I) E_\mu = 0 \quad and \quad F_\lambda(A - \lambda I) = 0, \tag{15}$$

so that $\mu F_\lambda E_\mu = F_\lambda A E_\mu = \lambda F_\lambda E_\mu$. Thus

$$F_\lambda E_\mu = 0 \quad if \quad \lambda \neq \mu. \tag{16}$$

Putting $K_\lambda = (F_\lambda E_\lambda)^\dagger$ we have, by 2·3,

$$K_\lambda = E_\lambda (F_\lambda E_\lambda)^\dagger F_\lambda. \tag{17}$$

Hence $K_\lambda K_\mu = \delta_{\lambda\mu} K_\lambda$ by (16). Also (17) gives

$$F_\lambda K_\mu E_\nu = \delta_{\lambda\mu} \delta_{\mu\nu} F_\lambda E_\lambda. \tag{18}$$

Any eigenvector z_α of A corresponding to the eigenvalue α satisfies $E_\alpha z_\alpha = z_\alpha$. Since A is diagonable, any column vector x conformable with A is expressible as a sum of eigenvectors; i.e. it is expressible in the form (a finite sum over all complex λ)

$$x = \Sigma E_\lambda x_\lambda.$$

‡ The existence of such K_λ's can be established as follows: let $\phi(\xi) \equiv (\xi - \lambda_1)^{n_1} \dots (\xi - \lambda_r)^{n_r}$ be, say, the minimum function of A where the factors $(\xi - \lambda_i)^{n_i}$ are mutually coprime. Then if $\phi(\xi) \equiv (\xi - \lambda_i)^{n_i} \psi_i(\xi)$ there exist polynomials $\chi_i(\xi)$ such that $\Sigma \chi_i(\xi) \psi_i(\xi) \equiv 1$. Set $K_{\lambda_i} = \chi_i(A)\psi_i(A)$ with the other K_λ's zero.

Similarly, if y^* is conformable with A, it is expressible as

$$y^* = \Sigma y_\lambda^* F_\lambda.$$

Now
$$y^*(\Sigma K_\mu) x = (\Sigma y_\lambda^* F_\lambda)(\Sigma K_\mu)(\Sigma E_\nu x_\nu)$$
$$= \Sigma y_\lambda^* F_\lambda E_\lambda x_\lambda = (\Sigma y_\lambda^* F_\lambda)(\Sigma E_\nu x_\nu)$$
$$= y^* x$$

by (18) and (16). Hence $\Sigma K_\mu = I$. Also, from (15) and (17) we obtain

$$AK_\lambda = \lambda K_\lambda = K_\lambda A. \tag{19}$$

Thus conditions (13) and (14) are satisfied. Also

$$A = \Sigma \lambda K_\lambda. \tag{20}$$

Conversely, if it is not assumed that A is diagonable but that $\Sigma K_\lambda = I$, K_λ being defined as above (i.e. with $n = 1$), then $x = \Sigma K_\lambda x$ gives x as a sum of eigenvectors of A, since (19) was deduced without assuming the diagonability of A.

If A is not diagonable it seems more convenient simply to prove that for any set of K_λ's satisfying (11), (12), (13) and (14) each $K_\lambda = (F_\lambda E_\lambda)^\dagger$, where F_λ and E_λ are defined as in the theorem.

Now, if the K_λ's satisfy (11), (12), (13) and (14) they must certainly satisfy

$$\Sigma K_\lambda = I$$

and
$$(A - \lambda I)^n K_\lambda = 0 = K_\lambda (A - \lambda I)^n,$$

where n is sufficiently large (if n is the order of A, this will do).

From this and the definitions of E_λ and F_λ given in the theorem we obtain

$$E_\lambda K_\lambda = K_\lambda = K_\lambda F_\lambda. \tag{21}$$

Using Euclid's algorithm we can find P and Q (polynomials in A) such that

$$I = (A - \lambda I)^n P + Q(A - \mu I)^n,$$

provided that $\lambda \neq \mu$. Now $F_\lambda (A - \lambda I)^n = 0 = (A - \mu I)^n E_\mu$. Hence $F_\lambda E_\mu = 0$ if $\lambda \neq \mu$, which with (21) gives $F_\lambda K_\mu = 0 = K_\lambda E_\mu$ whenever $\lambda \neq \mu$. But $\Sigma K_\lambda = I$. Thus

and
$$\left. \begin{array}{c} F_\lambda K_\lambda = F_\lambda \\ K_\lambda E_\lambda = E_\lambda. \end{array} \right\} \tag{22}$$

Using (21) and (22) it is a simple matter to verify that $(F_\lambda E_\lambda)^\dagger = K_\lambda$. (Note that $K_\lambda = I - \{(I - E_\lambda)(I - F_\lambda)\}^\dagger$ also.)

Incidentally, this provides a proof of the uniqueness of K_λ.‡

COROLLARY. *If A is normal, it is diagonable and its principal idempotent elements are hermitian.*

Proof. Since $A - \lambda I$ is normal we can apply 1·8 to the definitions of E_λ and F_λ in Theorem 3. This gives

$$E_\lambda = I - (A - \lambda I)^\dagger (A - \lambda I) \quad \text{and} \quad F_\lambda = I - (A - \lambda I)(A - \lambda I)^\dagger,$$

whence A is diagonable. Also $E_\lambda = F_\lambda$, so that $K_\lambda = E_\lambda$ is hermitian.

‡ In fact, $\Sigma K_\lambda = I$ and $(A - \lambda I)^n K_\lambda = 0$ are sufficient for uniqueness.

For this result see, for example, Drazin (4).

If A is normal (20) gives $A = \Sigma \lambda E_\lambda$, and thus using 1·7, 1·4 and 2·2 we obtain $A^\dagger = \Sigma \lambda^\dagger E_\lambda$.

A new type of spectral decomposition. Unless A is normal, there does not appear to be any simple expression for $A^\dagger$ in terms of its principal idempotent elements. However, this difficulty is overcome with the following decomposition, which applies equally to rectangular matrices.

THEOREM 4. *Any matrix A is uniquely expressible in the form*

$$A = \sum_{\alpha > 0} \alpha U_\alpha,$$

this being a finite sum over real values of α, where $U_\alpha^\dagger = U_\alpha^*$, $U_\alpha^* U_\beta = 0$ and $U_\alpha U_\beta^* = 0$ if $\alpha \neq \beta$.‡

Proof. The condition on the U's can be written comprehensively as

$$U_\alpha U_\beta^* U_\gamma = \delta_{\alpha\beta} \delta_{\beta\gamma} U_\alpha.$$

For $U_\alpha U_\alpha^* U_\alpha = U_\alpha$ implies $U_\alpha^* U_\alpha U_\alpha^* = U_\alpha^*$, whence by uniqueness, $U_\alpha^\dagger = U_\alpha^*$ (as $U_\alpha U_\alpha^*$ and $U_\alpha^* U_\alpha$ are both hermitian). Also $U_\alpha U_\alpha^* U_\beta = 0$ and $U_\alpha U_\beta^* U_\beta = 0$ respectively imply $U_\alpha^* U_\beta = 0$ and $U_\alpha U_\beta^* = 0$ by (2). Define

$$E_\lambda = I - (A^*A - \lambda I)^\dagger (A^*A - \lambda I).$$

The matrix A^*A is normal, being hermitian, and is non-negative definite. Hence the non-zero E_λ's are its principal idempotent elements and $E_\lambda = 0$ unless $\lambda \geq 0$.

Thus
$$A^*A = \Sigma \lambda E_\lambda \quad \text{and} \quad (A^*A)^\dagger = \Sigma \lambda^\dagger E_\lambda.$$

Hence
$$A^\dagger A = (A^*A)^\dagger A^*A = \Sigma \lambda^\dagger \lambda E_\lambda = \sum_{\lambda > 0} E_\lambda.$$

Put $U_\alpha = \alpha^{-1} A E_{\alpha^2}$ if $\alpha > 0$ and $U_\alpha = 0$ otherwise. Then

$$\Sigma \alpha U_\alpha = \sum_{\alpha > 0} \alpha \alpha^{-1} A E_{\alpha^2} = A \sum_{\lambda > 0} E_\lambda = A A^\dagger A = A.$$

Also if $\alpha, \beta, \gamma > 0$
$$\begin{aligned} U_\alpha U_\beta^* U_\gamma &= \alpha^{-1} \beta^{-1} \gamma^{-1} A E_{\alpha^2} E_{\beta^2} A^* A E_{\gamma^2} \\ &= \alpha^{-1} \beta^{-1} \gamma^{-1} \gamma^2 A E_{\alpha^2} E_{\beta^2} E_{\gamma^2} \\ &= \delta_{\alpha\beta} \delta_{\beta\gamma} \alpha^{-1} A E_{\alpha^2} = \delta_{\alpha\beta} \delta_{\beta\gamma} U_\alpha. \end{aligned}$$

To prove the uniqueness suppose

$$A = \sum_{\alpha > 0} \alpha V_\alpha, \quad \text{where} \quad V_\alpha V_\beta^* V_\gamma = \delta_{\alpha\beta} \delta_{\beta\gamma} V_\alpha.$$

Then it is easily verified that the non-zero expressions $V_\alpha^* V_\alpha$, for $\alpha > 0$, and $I - \sum_{\beta > 0} V_\beta^* V_\beta$ are the principal idempotent elements of A^*A corresponding respectively to the eigenvalues α^2 and 0. Hence $V_\alpha^* V_\alpha = E_{\alpha^2}$, where $\alpha > 0$, giving

$$U_\alpha = \alpha^{-1} (\sum_{\beta > 0} \beta V_\beta) V_\alpha^* V_\alpha = V_\alpha.$$

Polar representation of a matrix. It is a well-known result (closely related to Autonne's) that any square matrix is the product of a hermitian with a unitary matrix (see, for example, Halmos (5)).

‡ Thus $A^\dagger = \Sigma \alpha^\dagger U_\alpha^*$.

Put $H = \Sigma \alpha U_\alpha U_\alpha^*$. Then H is non-negative definite hermitian, since $U_\alpha = 0$ unless $\alpha > 0$, and $H^2 = AA^*$. (Such an H must be unique.) We have $H^\dagger H = HH^\dagger = AA^\dagger$.

Now $AA^\dagger$ and $A^\dagger A$ are equivalent under a unitary transformation, both being hermitian and having the same eigenvalues; i.e. there is a unitary matrix W satisfying $WA^\dagger A = AA^\dagger W$. Putting $V = H^\dagger A + W - WA^\dagger A$ we see that $VV^* = I$ and $A = HV$. This is the *polar representation* of A.

The polar representation is unique if A is non-singular but not if A is singular. However, if we require $A = HU$, where H is (as before) non-negative definite hermitian, $U^\dagger = U^*$ and $UU^* = H^\dagger H$, the representation is always unique‡ and also exists for rectangular matrices. The uniqueness of H follows from

$$AA^* = HUU^*H = H(H^\dagger H)H = H^2,$$

and that of U from $\qquad H^\dagger A = H^\dagger HU = UU^\dagger U = U.$

The existence of H and U is established by $H = \Sigma \alpha U_\alpha U_\alpha^*$ and $U = \Sigma U_\alpha$.

If we put $G = \Sigma \alpha U_\alpha^* U_\alpha$ we obtain the alternative representation $A = UG$.

I wish to thank Dr M. P. Drazin for useful suggestions and for advice in the preparation of this paper.

REFERENCES

(1) AUTONNE, L. Sur les matrices hypohermitiennes et sur les matrices unitaires. *Ann. Univ. Lyon*, (2), 38 (1915), 1–77.
(2) BJERHAMMAR, A. Rectangular reciprocal matrices, with special reference to geodetic calculations. *Bull. géod. int.* (1951), pp. 188–220.
(3) CECIONI, F. Sopra operazioni algebriche. *Ann. Scu. norm. sup. Pisa*, 11 (1910), 17–20.
(4) DRAZIN, M. P. On diagonable and normal matrices. *Quart. J. Math.* (2), 2 (1951), 189–98.
(5) HALMOS, P. R. *Finite dimensional vector spaces* (Princeton, 1942).
(6) MURRAY, F. J. and VON NEUMANN, J. On rings of operators. *Ann. Math., Princeton*, (2), 37 (1936), 141–3.
(7) WEDDERBURN, J. H. M. *Lectures on matrices* (Colloq. Publ. Amer. math. Soc. no. 17, 1934).

ST JOHN'S COLLEGE
  CAMBRIDGE

‡ See also Murray and von Neumann (6). In fact $U^\dagger = U^*$ follows from $UU^* = H^\dagger H$.