—NOTES—

ON THE COMPUTATION OF A GENERALIZED INVERSE OF A MATRIX*

By C. DONALD LA BUDDE (*New York University, University Heights*)
AND
G. R. VERMA (*University of Rhode Island*)

1. Introduction. The notion of the generalized inverse was introduced by Moore [4]. Unaware of the work of Moore, Bjerhammer [2] and Penrose [5] reintroduced the notion independently. This notion of generalized inverse is useful in finding solutions of linear equations in the case where the number of unknowns is different from the number of equations or the coefficient matrix is singular. Generalized inverses also provide a uniform approach to least squares theory and bivariate interpolation including the cases where the related linear equations are of the kind mentioned above. An extensive account of generalized inverses may be found in Ben-Israel [1] and La Budde [3]. In this note we describe a simple, noniterative algorithm for computing the generalized inverse of an arbitrary rectangular matrix. The method is essentially a modification of the Givens method for computing the inverse of a matrix. Although the algorithm is easily generalized to the complex case, we will restrict ourselves to the case of real matrices. The method discussed in this paper is essentially a modification of Givens' method for finding inverses of matrices. It is noniterative and requires only a fixed, predetermined number of operations, depending only on the numbers of rows and columns of the matrix.

2. Definitions and notation. Capital English letters A, B, C, etc. will represent matrices, with the usual notation A^t, B^t, C^t, etc. for the transposes. A^+, B^+, C^+ will represent the generalized inverses of the matrices A, B, C. I represents the identity matrix. Small English letters a, b, c, etc. will represent column vectors and a^t, b^t, c^t will represent row vectors. Real scalars will be represented by small Greek letters α, β, δ, etc., and positive integers will be represented by the letters i, j, k, l, m, n, p.

3. The basic result.

THEOREM 1. *Let A be any real $m \times n$ matrix with $m \geq n$. Then there are two orthogonal matrices U, $V(U^tU = UU^t = I, V^tV = VV^t = I)$ and a triangular (upper or lower) invertible $l \times l$ matrix B such that:*

$$UAV = B; \qquad m = n = l \tag{3.1}$$

$$UAV = \begin{bmatrix} B \\ - \\ 0 \end{bmatrix} \qquad m > n = l \tag{3.2}$$

*Received April 17, 1968.

$$(3.3) \qquad UAV = \begin{bmatrix} B & C \\ \hline 0 & 0 \end{bmatrix} \qquad m > n > l$$

and C is an $m \times (n - l)$ matrix. We first note that we need only consider the case where $m \geq n$: if $m < n$ we can apply Theorem 1 to A^t.

Proof. If $m = n$ and A in invertible then we can take $V = I$ and U to be a product $U = U^{(n-1)} U^{(n-2)} \cdots U^{(i)} \cdots U^{(2)} U^{(1)}$ where $U^{(i)} = I - 2w^{(i)}w^{(i)t}$. Let

$$(3.4) \qquad A^{(0)} = A, \qquad A^{(i)} = U^{(i)}A^{(i-1)}, \qquad B = A^{(n-1)}$$

and let the elements of $w^{(i)}$, $A^{(i)}$ be $(w_j^{(i)})$, $(\alpha_{jk}^{(i)})$ respectively. It suffices to consider the ith step as indicated by the second of Eqs. (3.4). If we define

$$\sigma^{(i-1)} = \left(\sum_{j=i}^{m} (\alpha_{ji}^{(i-1)})^2 \right)^{1/2},$$

$$(3.5) \qquad w_j^{(i)} = 0, \qquad j < i,$$

$$w_i^{(i)} = ((1/2)(1 + (|\alpha_{ii}^{(i-1)}|/\sigma^{(i-1)})))^{1/2},$$

$$w_j^{(i)} = (\operatorname{sgn}(\alpha_{ii}^{(i-1)})\alpha_{ji}^{(i-1)})/(2w_i^{(i)}\sigma^{(i-1)}) \qquad i < j \leq m,$$

then

$$(3.6) \qquad \alpha_{ji}^{(i)} = 0 \quad \text{for} \quad i < j \leq n.$$

Hence $B = A^{(n-1)}$ will be upper triangular and invertible.

If $m > n$ and A has rank n, then we may also take $V = I$ and apply the construction of the previous case to obtain the form (3.2) with B upper triangular and invertible with $l = m$.

If A has rank $l < m$, then we take V to be a product of permutation matrices $V = P^{(1)}P^{(2)} \cdots P^{(i)} \cdots P^{(n-2)}P^{(n-1)}$. We modify the algorithm given by equations (3.4) as follows. Let $C^{(i)}$ have elements $(\gamma_{jk}^{(i)})$ and let

$$(3.7) \qquad A^{(0)} = A, \qquad A^{(i-1)}P^{(i)} = C^{(i)}, \qquad U^{(i)}C^{(i)} = A^{(i)}.$$

The $U^{(i)}$ are defined as before with the elements $(\gamma_{jk}^{(i)})$ replacing the elements $(\alpha_{jk}^{(i-1)})$ in the definitions. It remains only to define the permutations $P^{(i)}$. Let $k^{(i)}$ be an integer such that

$$(3.8) \qquad \sum_{p=i}^{m} (\alpha_{pk^{(i)}}^{(i-1)})^2 \geq \sum_{p=i}^{m} (\alpha_{pk}^{(i-1)})^2$$

for all k, $i \leq k \leq n$. Then $P^{(i)}$ will be defined as that permutation matrix which, when multiplied on the left by a matrix A, has the affect of interchanging the ith and $k^{(i)}$th columns of A. With this definition of $P^{(i)}$, it can easily be seen that the end result of the algorithm defined by Eqs. (3.7) must be of the form (3.3).

THEOREM 2. *Let D be an $m \times n$ matrix, $m \leq n$ of the form*

$$(3.9) \qquad D = \begin{bmatrix} B \\ \hline 0 \end{bmatrix}$$

where B is n × n and invertible. Then

$$D^{+} = (B^{t} \mid 0). \tag{3.10}$$

Let D be an m × n matrix, m ≥ n, of the form

$$D = \left[\begin{array}{c|c} B & C \\ \hline 0 & 0 \end{array}\right] \tag{3.11}$$

where B is l × l, l < n and invertible. Then

$$D^{+} = \left[\begin{array}{c|c} B^{t}(BB^{t} + CC^{t})^{-1} & 0 \\ \hline C^{t}(BB^{t} + CC^{t})^{-1} & 0 \end{array}\right]. \tag{3.12}$$

(Note that since B is invertible, BB^{t}, and hence $BB^{t} + CC^{t}$ is positive definite and invertible).

Proof. We can verify that D^{+} is the generalized inverse for D by direct substitution in the two Penrose [2] equations which uniquely define D^{+}, namely

$$D^{+}D^{+t}D^{t} = D^{+} \qquad D^{+}DD^{t} = D^{t}. \tag{3.13}$$

Finally we note the Penrose result: if U, V are two orthogonal matrices and A is any matrix, then $(UAV)^{+} = V^{t}A^{+}U^{t}$.

It can now easily be seen how the three results of this section can be combined together to construct a computational algorithm for computing the generalized inverse of an arbitrary rectangular matrix A. If the number of columns of A is greater than the number of rows, apply the algorithm to A^{t}, and take the transpose of A^{t+} noting that $A^{+t} = A^{t+}$.

4. Numerical examples. In this section we give two numerical examples. We measure the accuracy of the computation by the deviation of our computed generalized inverse from the four Penrose conditions (the four error matrices). The numbers in the parentheses should be read as exponents, i.e. + .1000(+02) is $.1 \times 10^{2}$ and −.2000(−01) is $-.2 \times 10^{-1}$.

Input Matrix $A(6 \times 6)$. Rank = 4.

	1	2	3	4	5	6
1	+.1000(+01)	+.2000(+01)	+.3000(+01)	+.4000(+01)	+.5000(+01)	+.6000(+01)
2	+.1000(+01)	+.2000(+01)	+.3000(+01)	+.4000(+01)	+.5000(+01)	+.7000(+01)
3	+.1000(+01)	+.2000(+01)	+.3000(+01)	+.4000(+01)	+.6000(+01)	+.6000(+01)
4	+.1000(+01)	+.2000(+01)	+.3000(+01)	+.5000(+01)	+.5000(+01)	+.6000(+01)
5	+.5000(+01)	+.1000(+02)	+.1500(+02)	+.2000(+02)	+.2500(+02)	+.3000(+02)
6	+.6000(+01)	+.1200(+02)	+.1800(+02)	+.2400(+02)	+.3000(+02)	+.3600(+02)

Generalized Inverse A^{+}.

	1	2	3	4	5	6
1	+.1843(−01)	−.4286(+00)	−.3571(+00)	−.2857(+00)	+.9217(−01)	+.1106(+00)
2	+.3687(−01)	−.8571(+00)	−.7143(+00)	−.5714(+00)	+.1843(+00)	+.2212(+00)
3	+.5530(−01)	−.1286(+01)	−.1071(+01)	−.8571(+00)	+.2765(+00)	+.3318(+00)
4	−.1613(−01)	+.3636(−14)	+.5551(−16)	+.1000(+01)	−.8065(−01)	−.9677(−01)
5	−.1613(−01)	+.5954(−14)	+.1000(+01)	−.2776(−16)	−.8065(−01)	−.9677(−01)
6	−.1613(−01)	+.1000(+01)	−.3678(−14)	−.5274(−15)	−.8065(−01)	−.9677(−01)

Error Matrices (zero if the computation is exact).

I. $AA^{+}A - A$

	1	2	3	4	5	6
1	−.1422(−11)	−.2845(−11)	−.4268(−11)	−.5715(−11)	−.7145(−11)	−.8505(−11)
2	−.1423(−11)	−.2843(−11)	−.4263(−11)	−.5706(−11)	−.7136(−11)	−.8491(−11)
3	−.1533(−11)	−.3066(−11)	−.4599(−11)	−.6154(−11)	−.7694(−11)	−.9161(−11)
4	−.1620(−11)	−.3240(−11)	−.4861(−11)	−.6508(−11)	−.8130(−11)	−.9686(−11)
5	−.7123(−11)	−.1425(−10)	−.2137(−10)	−.2861(−10)	−.3577(−10)	−.4263(−10
6	−.8457(−11)	−.1691(−10)	−.2538(−10	−.3397(−10)	−.4242(−10)	−.5059(−10)

II. $A^{+}AA^{+} - A^{+}$

	1	2	3	4	5	6
1	+.1301(−16)	−.3622(−14)	−.6384(−15)	−.2415(−14)	+.2776(−16)	+.1388(−16)
2	−.1180(−15)	−.2068(−14)	−.3345(−14)	+.9992(−15)	−.6523(−15)	−.7772(−15)
3	+.1769(−15)	−.1221(−13)	−.8438(−14)	−.6398(−14)	+.8743(−15)	+.1069(−14)
4	−.3335(−14)	+.6558(−14)	+.3763(−14)	−.8188(−15)	−.1669(−13)	−.2001(−13)
5	−.1972(−14)	+.9681(−14)	+.4441(−15)	+.1273(−14)	−.9853(−14)	−.1182(−13)
6	−.1197(−15)	+.3331(−14)	+.7088(−15)	+.3676(−14)	−.5690(−15)	−.5967(−15)

III. $A^{+}A - (A^{+}A)^{*}$

	1	2	3	4	5	6
1	+.0000(+00)	−.4441(−15)	−.2220(−14)	+.1695(−12)	+.7906(−13)	−.4020(−13)
2	+.4441(−15)	+.0000(+00)	+.3775(−14)	+.3468(−12)	+.1588(−12)	−.7239(−13)
3	+.2220(−14)	−.3775(−14)	+.0000(+00)	+.5194(−12)	+.2403(−12)	−.1186(−12)
4	−.1695(−12)	−.3468(−12)	−.5194(−12)	+.0000(+00)	−.5400(−12)	−.1185(−11)
5	−.7906(−13)	−.1588(−12)	−.2403(−12)	+.5400(−12)	+.0000(+00)	−.6619(−12)
6	+.4020(−13)	+.7239(−13)	+.1186(−12)	+.1185(−11)	+.6619(−12)	+.0000(+00)

IV. $AA^{+} - (AA^{+})^{*}$

	1	2	3	4	5	6
1	+.0000(+00)	+.5172(−13)	−.6897(−14)	+.8132(−14)	−.4441(−15)	−.4441(−15)
2	−.5172(−13)	+.0000(+00)	−.6971(−13)	−.5757(−13)	−.2176(−12)	−.2938(−12)
3	+.6997(−14)	+.6971(−13)	+.0000(+00)	+.6745(−14)	+.2950(−13)	−.1416(−13)
4	−.8132(−14)	+.5757(−13)	−.6745(−14)	+.0000(+00)	−.1502(−13)	−.2170(−13)
5	+.4441(−15)	+.2176(−12)	−.2950(−13)	+.1502(−13)	+.0000(+00)	−.2887(−14)
6	+.4441(−15)	+.2938(−12)	+.1416(−13)	+.2170(−13)	+.2887(−14)	+.0000(+00)

Input Matrix $A(6 \times 3)$. Rank = 3.

	1	2	3
1	+.1000(+01)	+.2000(+01)	+.3000(+01)
2	+.3000(+01)	+.1000(+01)	+.2000(+01)
3	+.2000(+01)	+.3000(+01)	+.1000(+01)
4	+.3000(+01)	+.2000(+01)	+.1000(+01)
5	+.1000(+01)	+.3000(+01)	+.2000(+01)
6	+.2000(+01)	+.1000(+01)	+.3000(+01)

Generalized Inverse A^{+}.

	1	2	3	4	5	6
1	−.1389(+00)	+.1944(+00)	+.2778(−01)	+.1944(+00)	−.1389(+00)	+.2778(−01)
2	+.2778(−01)	−.1389(+00)	+.1944(+00)	+.2778(−01)	+.1944(+00)	−.1389(+00)
3	+.1944(+00)	+.2778(−01)	−.1389(+00)	−.1389(+00)	+.2778(−01)	+.1944(+00)

Error Matrices (zero if the computation is exact).

I. $AA^{+}A - A$

	1	2	3
1	+.1016(−11)	+.1000(−11)	+.1004(−11)
2	+.6743(−12)	+.6672(−12)	+.6768(−12)
3	+.2882(−12)	+.2853(−12)	+.2951(−12)
4	+.3026(−12)	+.3020(−12)	+.3131(−12)
5	+.6453(−12)	+.6353(−12)	+.6415(−12)
6	+.1030(−11)	+.1015(−11)	+.1021(−11)

II. $A^{+}AA^{+} - A^{+}$

	1	2	3	4	5	6
1	+.5967(−15)	−.4718(−15)	−.6939(−15)	−.9576(−15)	+.8327(−16)	+.3305(−15)
2	−.1224(−14)	−.1499(−14)	−.1762(−14)	−.1774(−14)	−.1485(−14)	−.1235(−14)
3	+.2777(−13)	+.3004(−13)	+.2932(−13)	+.3021(−13)	+.2802(−13)	+.2887(−13)

III. $A^{+}A - (A^{+}A)^{*}$

	1	2	3
1	+.0000(+00)	+.1590(−13)	−.3522(−12)
2	−.1590(−13)	+.0000(+00)	−.3636(−12)
3	+.3522(−12)	+.3636(−12)	+.0000(+00)

IV. $AA^{+} - (AA^{+})^{*}$

	1	2	3	4	5	6
1	+.0000(+00)	+.3041(−13)	+.5795(−13)	+.5852(−13)	+.2859(−13)	+.8743(−15)
2	−.3041(−13)	+.0000(+00)	+.3013(−13)	+.3007(−13)	+.1943(−15)	−.3045(−13)
3	−.5795(−13)	−.3013(−13)	+.0000(+00)	−.5967(−15)	−.2948(−13)	−.5798(−13)
4	−.5852(−13)	−.3007(−13)	+.5967(−15)	+.0000(+00)	−.2866(−13)	−.5905(−13)
5	−.2859(−13)	−.1953(−15)	+.2948(−13)	+.2866(−13)	+.0000(+00)	−.2838(−13)
6	−.8743(−15)	+.3045(−13)	+.5798(−13)	+.5905(−13)	+.2838(−13)	+.0000(+00)

Acknowledgments. The authors are indebted to Dr. R. C. Sahni for many helpful discussions. This work was supported by the National Aeronautics and Space Administration, Washington, D.C. 20546, and the Air Pollution Division of the Bureau of State Services, Public Health Service, Bethesda, Maryland, U. S. A.

References

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

[2] A. Bjerhammer, *Application of calculus of matrices to method of least squares with special reference to geodetic calculations*, Trans. Roy. Inst. Tech., Stockholm, **49**, 1–86 (1951)

[3] C. D. La Budde, *Notes on generalized inverses of operators*. Part I, Dept. of Chemistry, New York University, New York, 1968

[4] E. H. Moore, Bull. Amer. Math. Soc. **26**, 394–395 (1920)

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