A Note on an Iterative Method for Generalized Inversion of Matrices*

By Adi Ben-Israel

The iterative method of Schulz [4], [3] for matrix inversion was generalized in [1] as follows:

THEOREM 1. *The sequence of matrices defined by*

$$X_{k+1} = X_k(2P_{R(A)} - AX_k) \qquad (k = 0, 1, \cdots) \tag{1}$$

where X_0 is an $n \times m$ complex matrix satisfying

$$X_0 = A^*B_0\,, \quad B_0 \text{ some nonsingular } m \times m \text{ matrix}, \tag{2}$$

$$X_0 = C_0A^*, \quad C_0 \text{ some nonsingular } n \times n \text{ matrix}, \tag{3}$$

$$\| AX_0 - P_{R(A)} \| < 1, \quad (\| \ \| \text{ any matrix norm } [3]), \tag{4}$$

$$\| X_0A - P_{R(A^*)} \| < 1, \tag{5}$$

converges to the generalized inverse A^+ of A.

As pointed out in [1], the computational significance of the method (1) is limited by the need for knowledge of $P_{R(A)}$ (and of $P_{R(A^*)}$ if condition (5) is to be checked). This difficulty is evaded in the following theorem.

THEOREM 2. *Let A be an arbitrary (nonzero) complex $m \times n$ matrix of rank r and let*

$$\lambda_1(AA^*) \geqq \lambda_2(AA^*) \geqq \cdots \geqq \lambda_r(AA^*)$$

denote the nonzero eigenvalues of AA^*. If the real scalar α satisfies

$$0 < \alpha < \frac{2}{\lambda_1(AA^*)} \tag{6}$$

then the sequence defined by:

$$X_0 = \alpha A^* \tag{7}$$

$$X_{k+1} = X_k(2I - AX_k) \qquad (k = 0, 1, \cdots). \tag{8}$$

converges to A^+ as $k \to \infty$.

Proof. X_0 defined by (7), (6) satisfies (2), (3), (4) and (5). To prove that X_0 of (7), (6) satisfies (4) we note that AA^+ $(=P_{R(A)})$ and AA^* are commuting Hermitian matrices with the same range space. The eigenvalues of the $m \times m$ matrix: $AX_0 - P_{R(A)} = \alpha AA^* - AA^+$ are therefore

$$\begin{cases} \alpha\lambda_i(AA^*) - 1 & (i = 1, \cdots, r) \\ 0 & (i = r + 1, \cdots, m) \end{cases} \tag{9}$$

Received August 5, 1965. Revised August 26, 1965.

* Research sponsored by the Office of Naval Research, Contract Nonr-1228(10), Project NR 047-021 and by the U.S. Army Research Office—Durham, Contract No. DA-31-124-ARO-D-322.

and, by (6), are all: <1 in absolute value:

$$(10) \qquad | \lambda_i(\alpha AA^* - AA^+) | < 1 \qquad (i = 1, \cdots, m)$$

similarly

$$(11) \qquad | \lambda_i(\alpha A^*A - A^+A) | < 1 \qquad (i = 1, \cdots, n).$$

(Indeed the nonzero eigenvalues of $(\alpha AA^* - AA^+)$, $(\alpha A^*A - A^+A)$ are identical.) With the lub$_S$-norm [3, p. 44] in (4) and (5), both hold because of (10) and (11). (Actually (10) and (11) suffice for the convergence of (8).)

Now the process (1) initiated with: $X_0 = \alpha A^*$ retains the form [1, Eq. (12)]:

$$(12) \qquad X_k = C_kA^* \qquad (k = 1, 2, \cdots)$$

and since

$$(13) \qquad A^*P_{R(A)} = A^*$$

it follows that:

$$(14) \qquad X_k(2P_{R(A)} - AX_k) = X_k(2I - AX_k) \qquad (k = 0, 1, \cdots)$$

and the convergence of (8) follows from that of (1). Q.E.D.

Remarks.

a) Similarly, the sequence defined by

$$(15) \qquad X_{k+1} = (2I - X_kA)X_k \qquad (k = 0, 1, \cdots)$$

with $X_0 = \alpha A^*$, converges to A^+.

b) In using the method (8) it is not necessary to compute $\lambda_1(AA^*)$: Writing

$$AA^* = (b_{ij}) \qquad (i, j = 1, \cdots, m)$$

we conclude from the Gershgorin theorem, [3] that:

$$\lambda_1(AA^*) \leqq \max_{i=1,\cdots,m} \left\{ \sum_{j=1}^{m} | b_{ij} | \right\}.$$

Condition (6) can therefore be replaced, e.g. by

$$(16) \qquad 0 < \alpha < \frac{2}{\max_{i=1,\cdots,m} \left\{ \sum_{j=1}^{m} | b_{ij} | \right\}}.$$

c) Examples and applications will be given in [2].

Technicon, Israel Institute of Technology
Haifa, Israel

Northwestern University
Evanston, Illinois

1. A. Ben-Israel, "An iterative method for computing the generalized inverse of an arbitrary matrix," *Math. of Comp.*, v. 19, 1965, pp. 452–455.

2. A. Ben-Israel & D. Cohen, "On iterative computation of generalized inverses and associated projections," *J. Soc. Indust. Appl. Math. Ser. Numer. Anal.* (To appear.)

3. A. S. Householder, *The Theory of Matrices in Numerical Analysis*, Blaisdell, New York, 1964.

4. G. Schulz, "Iterative Berechnung der reziproken Matrix," *Z. Angew. Math. Mech.*, v. 13, 1933, pp. 57–59.