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

\[ X_{k+1} = X_k(2P_{R(A)} - AX_k) \quad (k = 0, 1, \ldots) \]

where \( X_0 \) is an \( n \times m \) complex matrix satisfying

\[ X_0 = A^*B_0, \quad B_0 \text{ some nonsingular } m \times m \text{ matrix,} \]

\[ X_0 = C_0A^*, \quad C_0 \text{ some nonsingular } n \times n \text{ matrix,} \]

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

\[ \| X_0A - P_{R(A^*)} \| < 1, \]

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

\[ \lambda_1(AA^*) \geq \lambda_2(AA^*) \geq \cdots \geq \lambda_r(AA^*) \]

denote the nonzero eigenvalues of \( AA^* \). If the real scalar \( \alpha \) satisfies

\[ 0 < \alpha < \frac{2}{\lambda_1(AA^*)} \]

then the sequence defined by:

\[ X_0 = \alpha A^* \]

\[ X_{k+1} = X_k(2I - AX_k) \quad (k = 0, 1, \ldots) . \]

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 \( AA^* \) 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

\[ \left\{ \begin{array}{ll}
\alpha \lambda_i(AA^*) - 1 & (i = 1, \ldots, r) \\
0 & (i = r + 1, \ldots, m)
\end{array} \right. \]

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. Reproduction of this paper in whole or in part is permitted for any purpose of the United States Government.
and, by (6), are all: $< 1$ in absolute value:

\[(10) \quad |\lambda_i(\alpha AA^* - AA^+)| < 1 \quad (i = 1, \ldots, m)\]

similarly

\[(11) \quad |\lambda_i(\alpha A^*A - A^+A)| < 1 \quad (i = 1, \ldots, n).\]

(Indeed the nonzero eigenvalues of $(\alpha AA^* - AA^+), (\alpha A^*A - A^+A)$ are identical.)

With the lub$_k$-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)]:

\[(12) \quad X_k = C_kA^* \quad (k = 1, 2, \ldots)\]

and since

\[(13) \quad A^*P_{R(A)} = A^*\]

it follows that:

\[(14) \quad X_k(2P_{R(A)} - AX_k) = X_k(2I - AX_k) \quad (k = 0, 1, \ldots)\]

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

**Remarks.**

a) Similarly, the sequence defined by

\[(15) \quad X_{k+1} = (2I - X_kA)X_k \quad (k = 0, 1, \ldots)\]

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

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

\[AA^* = (b_{ij}) \quad (i, j = 1, \ldots, m)\]

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

\[\lambda_1(\alpha A^*) \leq \max_{i=1, \ldots, m} \left\{ \sum_{j=1}^{m} |b_{ij}| \right\}.\]

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

\[(16) \quad 0 < \alpha < \frac{2}{\max_{i=1, \ldots, m} \left\{ \sum_{j=1}^{m} |b_{ij}| \right\}}.\]

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

Technion, Israel Institute of Technology
Haifa, Israel

Northwestern University
Evanston, Illinois