UNIVERSAL ITERATIVE METHODS FOR COMPUTING GENERALIZED INVERSES

Predrag S. Stanimirović and Dragan S. Djordjević

Abstract. In this paper we construct a few iterative processes for computing $\{1,2\}$ inverses of a linear bounded operator, based on the hyper-power iterative method or the Neumann-type expansion. Under the suitable conditions these methods converge to the $\{1,2,3\}$ or $\{1,2,4\}$ inverses. Also, we specify conditions when the iterative processes converge to the Moore-Penrose inverse, the weighted Moore-Penrose inverse or to the group inverse. A few error estimates are derived. The advantages of the introduced methods over the Tanabe's method [16] for computing the reflexive generalized inverses are also investigated.

Key words: generalized inverses, Moore–Penrose inverse, hyper–power method, Neumann–type expansion

1. Introduction

Let X and Y be two finite dimensional complex Hilbert spaces and let $A : X \to Y$ be a linear operator. There are well–known properties of generalized inverses of A:

$$(1)\quad AXA = A, \quad (2)\quad XAX = X, \quad (3)\quad (AX)^* = AX, \quad (4)\quad (XA)^* = XA.$$

For a subset $\mathcal{S}$ of the set $\{1,2,3,4\}$, the set of operators obeying the conditions contained in $\mathcal{S}$ is denoted by $A\{\mathcal{S}\}$. An operator in $A\{\mathcal{S}\}$ is called an $\mathcal{S}$-inverse of A and is denoted by $A^{(\mathcal{S})}$. In particular, for any A, the set $A\{1,2,3,4\}$ consists of a single element, the Moore-Penrose inverse of A, denoted by $A^\dagger$ [9]. The group inverse $A^\#$ is the unique operator which satisfies (1), (2) and

$$AA^\# = A^\# A. \tag{5}$$

Any element from the class $A\{1,2\}$ is also called the reflexive generalized inverse of A.

1991 *Mathematics Subject Classification.* 65F20, 65F30, 15A24.

Let $A = PQ$ be a full–rank decomposition of A, where P and Q are two full rank linear operators. Then all $\{1,2\}$ inverses of A can be represented in the form

$$X = W_1(QW_1)^{-1}(W_2P)^{-1}W_2 = W_1(W_2AW_1)^{-1}W_2, \tag{6}$$

where W_1 and W_2 are suitable choosen operators, such that QW_1 and W_2P are invertible [13].

The weighted Moore-Penrose inverse is investigated in [3], [12]. For the sake of completeness we restate here the main results of these papers. Let there be given positive-definite (and hermitian) operators M and N. For any operator A there exists the unique solution $X = A^{\dagger}_{M,N} \in A\{1,2\}$ satisfying the following equations in X [3], [12]:

$$(3M) \qquad (MAX)^* = MAX \qquad\qquad (4N) \qquad (XAN)^* = XAN.$$

If $A = PQ$ is a full rank factorization of A, then [12]:

$$A^{\dagger}_{M,N} = (QN)^*(Q(QN)^*)^{-1}((MP)^*P)^{-1}(MP)^*. \tag{7}$$

The hyper-power method of the order 2 dates back to the well-known paper of Schulz [14]. Altman devised the hyper-power method of any arbitrary order $q \geq 2$, for inverting a nonsingular bounded operator on a Banach space [1]. In [10] the convergence of the same method is proved under the condition which is weaker than the one assumed in [1], and some better error estimates are derived.

Zlobec [19] and Petryshyn [11] showed that the qth order hyper-power iterative method with $q \geq 2$ for the determination of inverses of nonsingular matrices and linear operators, can be generalized to the determination of the Moore-Penrose inverse of an arbitrary matrix, or a bounded linear operator with closed range.

Zlobec in [19] defined the following two hyper-power iterative methods of an arbitrary high order $q \geq 2$:

$$\begin{aligned} T_k &= I_X - Y_kA, \\ Y_{k+1} &= (I_X + T_k + \cdots + T_k^{q-1})Y_k, \quad k = 0,1,\ldots \end{aligned}$$

$$\begin{aligned} T'_k &= I_Y - AY'_k, \\ Y'_{k+1} &= Y'_k(I_Y + T'_k + \cdots + {T'_k}^{q-1}), \quad k = 0,1,\ldots \end{aligned}$$

It is well–known [19], that if we take

$$Y_0 = Y'_0 = \alpha A^*, \qquad 0 < \alpha \leq \frac{2}{\operatorname{tr}(A^*A)},$$

then $\lim\limits_{k\to\infty} Y_k = \lim\limits_{k\to\infty} Y'_k = A^{\dagger}$.

The process which generates the sequence Y_k is more superior then the process which generates the sequence Y'_k in the case $m > n$ [6]. The hyper-power iterative method of the order 2 is studied in [15] in view of the singular value decomposition of a matrix. In [6] the hyper-power iterative method is adapted for computing $A^\dagger B$, where A and B are arbitrary complex matrices with equal number of rows.

The paper is organized as follows. In Section 2 we construct iterative methods for computing the reflexive generalized inverses of a linear operator. These methods are based on the hyper-power iterative methods. We select two arbitrary matrices and adequate initial values for these methods to generate different generalized inverses for the concerned operator. In Section 3 we give a few error estimates, look for the optimal value of the parameter α and show that the method is self-correcting. In Section 4 we develop analogous iterative methods which arise from the Neumann-type expansion and compare our method with the Tanabe's method [16]. Finally, we give several examples which illustrate our theory.

2. Iterative methods

In the following lemma we introduce an improvement of the hyper-power iterative method, and construct iterative method which generates all of the reflexive generalized inverses.

Lemma 2.1. *Let* $\operatorname{rank}(A) = r \geq 2$ *and* W_1, W_2 *are two arbitrary operators, such that* W_2AW_1 *is invertible operator. If* $q \geq 2$ *is an integer, then the following two iterative processes:*

$$Y_0 = Y'_0 = \alpha(W_2AW_1)^*, \qquad 0 < \alpha \leq \frac{2}{\operatorname{tr}((W_2AW_1)^*W_2AW_1)},$$

$$\begin{aligned}
T_k &= I_X - Y_kW_2AW_1,\\
Y_{k+1} &= (I_X + T_k + \cdots + T_k^{q-1})Y_k,\\
X_{k+1} &= W_1Y_{k+1}W_2 \quad k = 0, 1, \dots
\end{aligned}$$

$$\begin{aligned}
T'_k &= I_Y - W_2AW_1Y'_k,\\
Y'_{k+1} &= Y'_k(I_Y + T'_k + \cdots + {T'_k}^{q-1}),\\
X'_{k+1} &= W_1Y'_{k+1}W_2 \quad k = 0, 1, \dots
\end{aligned}$$

generate the class of $\{1, 2\}$ *inverses of* A.

Proof. Using the results form [19], we conclude

$$\lim_{k\to\infty} Y_k = \lim_{k\to\infty} Y'_k = (W_2AW_1)^\dagger = (W_2AW_1)^{-1}.$$

According to (6), it is obvious that

$$\lim_{k\to\infty} X_k = \lim_{k\to\infty} X'_k = X = W_1(W_2AW_1)^{-1}W_2 \in A\{1, 2\}. \qquad \square$$

Therefore, we just formed two iterative processes for computing all of the $\{1,2\}$ inverses of A. However, under the suitable conditions, we can get some iterative methods for computing $\{1,2,3\}$ or $\{1,2,4\}$ inverses, the Moore-Penrose inverse, the weighted Moore-Penrose inverse or the group inverse of A. For the sake of simplicity we use the following notation: $B = W_2AW_1$, $C = AW_1$, $D = W_2A$.

Theorem 2.1. *Let* $\operatorname{rank}(A) = r \geq 2$ *and* QW_1, W_2P *be invertible operators.*

(a) *If* W_2 *is an unitary operator with respect to the considered scalar product and*
$$0 < \alpha \leq \min\left\{\frac{2}{\operatorname{tr}(B^*B)}, \frac{2}{\operatorname{tr}(C^*C)}\right\},$$
then $X_k \to X = W_1(AW_1)^\dagger \in A\{1,2,3\}$ *as* $k \to \infty$.

(b) *If* W_1 *is an unitary operator with respect to the considered scalar product and*
$$0 < \alpha \leq \min\left\{\frac{2}{\operatorname{tr}(B^*B)}, \frac{2}{\operatorname{tr}(D^*D)}\right\},$$
then $X'_k \to X = (W_2A)^\dagger W_2 \in A\{1,2,4\}$ *as* $k \to \infty$.

(c) *If (a) and (b) are valid, then* $X_k \to A^\dagger$.

(d) *If (b) is valid and* $W_2 = P^*$, *then* $X'_k \to X = A^\dagger$.

(e) *If (a) is valid and* $W_1 = Q^*$, *then* $X_k \to X = A^\dagger$.

(f) *If* $W_1 = Q^*$ *and* $0 < \alpha \leq \dfrac{2}{\operatorname{tr}(B^*B)}$,
then $X_k \to X = Q^*(W_2AQ^*)^{-1}W_2 \in A\{1,2,4\}$.

(g) *If* $W_2 = P^*$ *and* $0 < \alpha \leq \dfrac{2}{\operatorname{tr}(B^*B)}$,
then $X_k \to X = W_1(P^*AW_1)^{-1}P^* \in A\{1,2,3\}$.

(h) *If* $W_1 = Q^*$, $W_2 = P^*$ *and* $0 < \alpha \leq \dfrac{2}{\operatorname{tr}(B^*B)}$,
then $X_k \to A^\dagger$.

(i) *If* $W_1 = (QN)^*$, $W_2 = (MP)^*$ *and* $0 < \alpha \leq \dfrac{2}{\operatorname{tr}(B^*B)}$,
then $X_k \to A^\dagger_{M,N}$.

(j) *In the case* $m = n$, $W_1 = P$, $W_2 = Q$ *and* $0 < \alpha \leq \dfrac{2}{\operatorname{tr}(B^*B)}$,
we get $X_k \to A^\#$.

Proof. (a) Obviously,
$$X_{k+1} = W_1Y_{k+1}W_2 = W_1(I_X + T_k + \cdots + T_k^{q-1})Y_kW_2.$$

Let $Z_k = Y_k W_2$. Then

$$Z_{k+1} = Y_{k+1} W_2 = (I_X + T_k + \cdots + T_k^{q-1}) Y_k W_2 = (I_X + T_k + \cdots + T_k^{q-1}) Z_k,$$

where $T_k = I_X - Y_k W_2 A W_1 = I_X - Z_k(AW_1)$. Since

$$Z_0 = Y_0 W_2 = \alpha W_1^* A^* W_2^* W_2 = \alpha W_1^* A^* = \alpha (AW_1)^*,$$

we have [19] $Z_k \to (AW_1)^\dagger$, as $k \to \infty$. This implies

$$T_k^l = (I_X - Z_k A W_1)^l \underset{k\to\infty}{\longrightarrow} (I_X - (AW_1)^\dagger A W_1)^l, \quad \text{for} \quad l = 1, 2, \ldots$$

Since $((AW_1)^\dagger A W_1)^2 = (AW_1)^\dagger A W_1$, we get

$$\begin{aligned} T_k^l &\underset{k\to\infty}{\longrightarrow} I_X - \binom{l}{1}(AW_1)^\dagger AW_1 + \binom{l}{2}(AW_1)^\dagger AW_1 + \cdots + (-1)^l \binom{l}{l}(AW_1)^\dagger AW_1 \\ &= I_X - (AW_1)^\dagger AW_1. \end{aligned}$$

Hence

$$\begin{aligned} Z_{k+1} &= (I_X + T_k + \cdots + T_k^{q-1}) Z_k \underset{k\to\infty}{\longrightarrow} \\ &\quad \left[I_X + (I_X - (AW_1)^\dagger AW_1) + \cdots + (I_X - (AW_1)^\dagger AW_1) \right] (AW_1)^\dagger = \\ &\quad = (AW_1)^\dagger. \end{aligned}$$

Now, it follows that $X_k \to W_1(AW_1)^\dagger = X$, as $k \to \infty$. Since $(AW_1(AW_1)^\dagger)^* = AW_1(AW_1)^\dagger$, we get $(AX)^* = AX$ and X is an $\{1,2,3\}$ inverse for A.

(b) Let W_1 be unitary. For the sequences

$$X'_{k+1} = W_1 Y'_{k+1} W_2 = W_1 Y'_k (I_Y + T'_k + \cdots + {T'_k}^{q-1}) W_2$$

and $Z'_k = W_1 Y'_k$ we have

$$Z'_{k+1} = W_1 Y'_{k+1} = W_1 Y'_k (I_Y + T'_k + \cdots + {T'_k}^{q-1}).$$

Since $T'_k = I_Y - W_2 A Z'_k$ and $Z'_0 = W_1 Y'_0 = \alpha W_1 W_1^* A^* W_2^* = \alpha (W_2 A)^*$, we use the method from (a) to conclude that $Z'_k \to (W_2 A)^\dagger$ and $X'_{k+1} \to (W_2 A)^\dagger W_2 = X$. Since $((W_2 A)^\dagger W_2 A)^* = (W_2 A)^\dagger W_2 A$, we get $(XA)^* = XA$, and consequently, X is $\{1,2,4\}$ inverse of A.

(c) Follows from (a) and (b).

(d) In (b) we obtain $X_k \to (W_2A)^\dagger W_2 \in A\{1,2,4\}$. We mention one important property of the Moore–Penrose inverse [2], [4]:

$$(UV)^\dagger = V^\dagger U^\dagger \quad \Longleftrightarrow \quad U^\dagger UVV^*U^* = VV^*U^* \text{ and } VV^\dagger U^*UV = U^*UV. \tag{8}$$

We take $U = W_2P$ and $V = Q$ in the expression $AX = PQ(W_2PQ)^\dagger W_2$. Operator W_2P is invertible, and $Q^\dagger$ is the right inverse of the full rank operator Q. So the right side of (8) is valid in this case. Now, we get

$$AX = PQQ^\dagger(W_2P)^{-1}W_2 = P(W_2P)^{-1}W_2.$$

Also $(AX)^* = W_2^*(P^*W_2^*)^{-1}P^*$. Now, if $W_2 = P^*$, we get $(AX)^* = AX$. On the other hand, if (b) is valid, then $X = (W_2A)^\dagger W_2$ is an $\{1,2,4\}$ inverse of A, and we immediately conclude $X = A^\dagger$.

(e) If W_2 is an unitary operator, we have that $\lim\limits_{k\to\infty} X_k = X = W_1(AW_1)^\dagger \in A\{1,2,3\}$. Using $U = P$ and $V = QW_1$, we conclude that (8) is valid. Consequently $XA = W_1(PQW_1)^\dagger PQ = W_1(QW_1)^{-1}P^\dagger PQ = W_1(QW_1)^{-1}Q$. Also $(XA)^* = Q^*(W_1^*Q^*)^{-1}W_1^*$. Obviously, if $W_1 = Q^*$, we get $X = A^\dagger$.

(f), (g) Follows from (6) and the well–known results [13]:

the general solution of the equations (1), (2), (4) is given by

$$X = Q^*(QQ^*)^{-1}(W_2P)^{-1}W_2 = Q^*(W_2AQ^*)^{-1}W_2;$$

the general solution of the equations (1), (2), (3) is given by

$$X = W_1(QW_1)^{-1}(P^*P)^{-1}P^* = W_1(P^*AW_1)^{-1}P^*.$$

(h) Follows from $A^\dagger = Q^*(QQ^*)^{-1}(P^*P)^{-1}P^*$ [3], [13].

(i) Comparing (6) and (7), we conclude that $A^\dagger_{M,N}$ can be selected from the class $A\{1,2\}$ using $W_1 = (QN)^*$ and $W_2 = (MP)^*$.

(j) Follows from $A^\# = P(QP)^{-2}Q$ [5]. $\square$

In the case $\operatorname{rank}(A) = 1$ we can use the next known proposition [19]:

Proposition 2.1. *Let A be of the rank $r = 1$. Then*

$$A^\dagger = \frac{1}{\operatorname{tr}(A^*A)}A^*.$$

Theorem 2.2. *Let* $\operatorname{rank}(A) = 1$ *and* QW_1, W_2P *be invertible operators.*

(a) $X = W_1(AW_1)^\dagger \in A\{1,2,3\}$ *is given by*

$$X = \frac{1}{\operatorname{tr}((AW_1)^* AW_1)} W_1(AW_1)^*.$$

(b) $Y = (W_2A)^\dagger W_2 \in A\{1,2,4\}$ *is given by*

$$Y = \frac{1}{\operatorname{tr}((W_2A)^* W_2A)} (W_2A)^* W_2.$$

(c) $Z = W_1(W_2AW_1)^{-1}W_2 \in A\{1,2\}$ *is presented by*

$$Z = \frac{1}{\operatorname{tr}((W_2AW_1)^* W_2AW_1)} W_1(W_2AW_1)^* W_2.$$

Proof. (a) Since QW_1 is invertible and P is left invertible, we have that $AW_1 = PQW_1 \neq 0$, which implies $\operatorname{rank}(AW_1) \geq 1$. Using $\operatorname{rank}(AW_1) \leq \operatorname{rank}(A) = 1$, we get $\operatorname{rank}(AW_1) = 1$. Thus, from Proposition 2.1. we get

$$(AW_1)^\dagger = \frac{1}{\operatorname{tr}((AW_1)^* AW_1)} (AW_1)^*.$$

Now we prove that $X = W_1(AW_1)^\dagger$ is an $\{1,2,3\}$ inverse of A. It is easy to see that $W_1(AW_1)^\dagger$ is an $\{2,3\}$ inverse of A. In order to prove that $AXA = A$ holds, we use the property (8) with $U = P$ and $V = QW_1$. Thus

$$AXA = PQW_1(PQW_1)^\dagger PQ = PQW_1(QW_1)^{-1}P^\dagger PQ = PQ = A.$$

(b) In a similar way can be proved that $\operatorname{rank}(W_2A) = 1$. We just need to prove that $AYA = A$. We also use (8) with $U = W_2P$ and $V = Q$. Then

$$AYA = PQ(W_2PQ)^\dagger W_2PQ = PQQ^\dagger(W_2P)^{-1}W_2PQ = PQ = A. \qquad \square$$

3. Error bounds

Since $X = W_1(W_2AW_1)^{-1}W_2$ and $X_k = W_1Y_kW_2$, we have

$$\|X - X_k\| \leq \|W_1\| \cdot \|(W_2AW_1)^{-1} - Y_k\| \cdot \|W_2\|.$$

So we just have to make bounds for $\|(W_2AW_1)^{-1} - Y_k\|$. Operator $B = W_2AW_1$ is invertible, so B^*B is invertible and positive–definite. The spectrum of B^*B is $\sigma(B^*B) = \{\lambda_1, \lambda_2, \dots \lambda_r\}$ and we can take that

$$0 < \lambda_1 < \lambda_2 < \cdots < \lambda_r.$$

We need some matrix norm $\|\cdot\|$ such that $\|T_0\| < 1$. Let $\|\cdot\|_{sp}$ be the spectral norm, i.e.

$$\|C\|_{sp} = \sqrt{\max \lambda(C^*C)},$$

where $\max \lambda(C^*C)$ denotes the greatest eigenvalue of C^*C.

Lemma 3.1. *Let* $T_0 = I - \alpha B^*B$ *and*

$$0 < \alpha \leq \frac{2}{\operatorname{tr}(B^*B)}.$$

If $\operatorname{rank}(A) \geq 2$ *then* $\|T_0\|_{sp} < 1$.

Proof. Since $T_0^*T_0 = I - 2\alpha B^*B + \alpha^2(B^*B)^2$, we get

$$\lambda \in \sigma(B^*B) \Longleftrightarrow 1 - 2\alpha\lambda + \alpha^2\lambda^2 = (\alpha\lambda - 1)^2 \in \sigma(T_0^*T_0).$$

So $(\alpha\lambda - 1)^2 < 1$ if and only if $\alpha\lambda < 2$ for all $\lambda \in \sigma(B^*B)$. We see that $\|T_0\|_{sp} < 1$ is valid for $0 < \alpha < \dfrac{2}{\lambda_r}$. Since $\operatorname{tr}(B^*B) = \sum\limits_{j=1}^{r} \lambda_j > \lambda_r$, we get

$$\frac{2}{\operatorname{tr}(B^*B)} < \frac{2}{\lambda_r}. \qquad \square$$

We use Lemma 3.1 and the error estimates for the norm $\|(W_2AW_1)^{-1} - Y_k\|$, given in [10], [8], [7], to verify the next theorem:

Theorem 3.1. *Providing that assumptions of Theorem 2.1. are valid, we get:*

(a) $\|X - X_k\|_{sp} \leq \dfrac{\|Y_kT_k\|_{sp}}{1 - \|T_k\|_{sp}} \|W_1\|_{sp}\|W_2\|_{sp}$,

(b) $\|X - X_k\|_{sp} \leq \|T_{k-1}\|_{sp}^{q-1} \dfrac{\|Y_{k-1}T_{k-1}\|_{sp}}{1 - \|T_{k-1}\|_{sp}} \|W_1\|_{sp}\|W_2\|_{sp}$,

(c) $\|X - X_k\|_{sp} \leq \|T_0\|_{sp}^{q^k} \dfrac{\|Y_0\|_{sp}}{1 - \|T_0\|_{sp}} \|W_1\|_{sp}\|W_2\|_{sp}$,

(d) $\|X - X_k\|_{sp} \leq \dfrac{\|Y_k'\|_{sp}\|T_k'\|_{sp}}{1 - \|T_k'\|_{sp}} \|W_1\|_{sp}\|W_2\|_{sp}$,

(e) $\|X - X_k\|_{sp} \leq \dfrac{\|Y_k'\|_{sp}\|T_{k-1}'\|_{sp}^q}{1 - \|T_{k-1}'\|_{sp}^q} \|W_1\|_{sp}\|W_2\|_{sp}$,

(f) $\|X - X_k\|_{sp} \leq \dfrac{\|Y_k'\|_{sp}\|T_{k-1}'\|_{sp}\|(T_{k-1}')^{q-1}\|_{sp}}{1 - \|T_{k-1}'\|_{sp}} \|W_1\|_{sp}\|W_2\|_{sp}$,

(f) *The order of convergence of the defined processes is* q*, i.e.*

$$\|X - X_{k+1}\|_{sp} = O\left(\|X - X_k\|_{sp}^q\right), \quad k \to \infty.$$

If W_1 and W_2 are unitary operators, then $X = A^\dagger$, by Theorem 2.1. So we immediately get the error bounds for $\|A^\dagger - X_k\|_{sp}$, since $\|W_1\|_{sp} = \|W_2\|_{sp} = 1$.

Now we look for the optimal value of α.

Proposition 3.1. *Let* $\operatorname{rank}(A) \geq 2$. *In each case of Theorem 2.1., the optimal value of* α *is the upper bound for the interval given in that case.*

Proof. Since I and αBB^* are commuting selfadjoint operators, the proof is identical to the corresponding in [19, Proposition 1]. □

Remark 3.1. It is well-known that the hyper–power method for computing the inverse of an invertible operator is self–correcting, but it is not self–correcting for computing generalized inverses [17], [18]. There is the well-known Zielke's iterative refinement process which solves the self–correcting problem. Namely, the iterative refinement for computing the Moore–Penrose inverse of A has the form:

$$\begin{aligned}\tilde{X}_k &= A^* X_k^* X_k X_k^* A^* \\ \tilde{T}_k &= I - \tilde{X}_k A \\ X_{k+1} &= (I + \tilde{T}_k + \cdots + \tilde{T}_k^{q-1})\tilde{X}_k.\end{aligned}$$

This modification is not necessary in each step.

Our iterative method for computing the sequence Y_k is self–correcting, so this method is self–correcting for computing the Z_k and X_k. We do not need any iterative refinements. So we solve, on an elementary way, the self–correcting problem for the iterative computation of generalized inverses.

4. Using the Neumann-type expansion

It is well–known [16], [19] that the q-th order hyper–power method generates the partial sums of the infinite series

$$\sum_{i=0}^{\infty} \left[(I - X_0 A)^i X_0 \right] \quad \text{or} \quad \sum_{i=0}^{\infty} \left[X_0 (I - A X_0)^i \right],$$

i.e.

$$X_k = \sum_{i=0}^{q^k - 1} \left[(I - X_0 A)^i X_0 \right] \quad \text{or} \quad X_k = \sum_{i=0}^{q^k - 1} \left[X_0 (I - A X_0)^i \right].$$

In the case $\rho(I - X_0 A) < 1$ the inverse A^{-1} of a nonsingular matrix admits the Neumann-type expansion [16]

$$A^{-1} = \sum_{i=0}^{\infty} \left[(I - X_0 A)^i X_0 \right].$$

Similarly, in the case $\rho(I - A X_0) < 1$

$$A^{-1} = \sum_{i=0}^{\infty} \left[X_0 (I - A X_0)^i \right].$$

Zlobec [19] shown that $A^{\dagger}$ can be computed by means of the infinite series under the assumption $0 < \alpha \leq \frac{2}{\mathrm{tr}(A^*A)}$.

Our strategy is to adapt the infinite series in order to compute $(W_2AW_1)^{-1}$. In this way, we develop corresponding iterative method for computing the reflexive generalized inverses $W_1(W_2AW_1)^{-1}W_2$. We determine conditions for the defined method to generate the $\{1,2,3\}$, $\{1,2,4\}$ inverses, the Moore-Penrose or the group inverse.

If $q \geq 2$, $\mathrm{rank}(A) = r \geq 2$, $B = W_2AW_1$, $C = AW_1$, $D = W_2A$, the iterative processes, based on the Neumann-type expansion, are defined as follows:

$$
\begin{aligned}
X_0 &= X_0' = \alpha B^*, \qquad 0 < \alpha \leq \frac{2}{\mathrm{tr}(B^*B)},\\
X_k &= W_1 \cdot \sum_{i=0}^{q^k-1} \left[(I - X_0B)^i \, X_0 \right] \cdot W_2, \qquad \text{or}\\
X_k' &= W_1 \cdot \sum_{i=0}^{q^k-1} \left[X_0 \, (I - BX_0)^i \right] \cdot W_2, \quad k = 0, 1, \ldots
\end{aligned}
$$

The following results, analogous to the results from Theorem 2.1. can be easily verified.

Theorem 4.1. *Let* $\mathrm{rank}(A) = r \geq 2$ *and* QW_1, W_2P *be invertible operators.*

(a) *If* W_2 *is an unitary operator and*

$$0 < \alpha \leq \min\left\{ \frac{2}{\mathrm{tr}(B^*B)}, \frac{2}{\mathrm{tr}(C^*C)} \right\},$$

then $X_k \to X = W_1(AW_1)^{\dagger}$ *as* $k \to \infty$ *and* X *is an* $\{1,2,3\}$ *inverse of* A.

(b) *If* W_1 *is an unitary operator and*

$$0 < \alpha \leq \min\left\{ \frac{2}{\mathrm{tr}(B^*B)}, \frac{2}{\mathrm{tr}(D^*D)} \right\},$$

then $X_k' \to X = (W_2A)^{\dagger}W_2 \in A\{1,2,4\}$ *as* $k \to \infty$.

(c) *If (a) and (b) are valid, then* $X_k \to A^{\dagger}$.

(d) *If (b) is valid and* $W_2 = P^*$, *then* $X_k' \to X = A^{\dagger}$.

(e) *If (a) is valid and* $W_1 = Q^*$, *then* $X_k \to X = A^{\dagger}$.

(f) *If* $W_1 = Q^*$ *and* $0 < \alpha \leq \dfrac{2}{\mathrm{tr}(B^*B)}$,

then $X_k \to X = Q^*(W_2AQ^*)^{-1}W_2 \in A\{1,2,4\}$.

(g) *If* $W_2 = P^*$ *and* $0 < \alpha \leq \dfrac{2}{\operatorname{tr}(B^*B)}$,
then $X_k \to X = W_1(P^*AW_1)^{-1}P^* \in A\{1,2,3\}$.

(h) *If* $W_1 = Q^*$, $W_2 = P^*$ *and* $0 < \alpha \leq \dfrac{2}{\operatorname{tr}(B^*B)}$,
then $X_k \to A^\dagger$.

(i) *If* $W_1 = (QN)^*$, $W_2 = (MP)^*$ *and* $0 < \alpha \leq \dfrac{2}{\operatorname{tr}(B^*B)}$,
then $X_k \to A^\dagger_{M,N}$.

(j) *In the case* $m = n$, $W_1 = P$, $W_2 = Q$ *and* $0 < \alpha \leq \dfrac{2}{\operatorname{tr}(B^*B)}$,
we get $X_k \to A^\#$.

Proof. Follows from

$$\begin{aligned} X_k &= W_1 \cdot \sum_{i=0}^{q^k-1} \left[(I - X_0B)^i \, X_0 \right] \cdot W_2 = \\ &= W_1 \cdot \sum_{i=0}^{q^k-1} \left[(I - X_0W_2AW_1)^i \, X_0W_2 \right], \quad k = 0, 1, \dots \end{aligned}$$

$$\begin{aligned} X_k' &= W_1 \cdot \sum_{i=0}^{q^k-1} \left[X_0 \, (I - BX_0))^i \, X_0 \right] \cdot W_2 = \\ &= \sum_{i=0}^{q^k-1} \left[W_1X_0 \, (I - W_2AW_1X_0)^i \right] \cdot W_2, \quad k = 0, 1, \dots \end{aligned}$$ □

Remark 4.1. Now we give some comparisons with the paper of Tanabe [16]. In [16] there are given necessary and sufficient conditions for the starting approximation X_0 of the hyper-power iterative method or the Neumann-type series, ensuring convergence of these methods to an arbitrary reflexive generalized inverse. Advantages of this paper related to [16] are:

(a) It is more convenient to use the initial conditions in Theorem 2.1. than the conditions from [16, Theorem 2.1.].
(b) We give a few error estimates.
(c) We know exact conditions ensuring convergence of defined processes to the $\{1,2,3\}$ inverses, the $\{1,2,4\}$ inverses, the Moore-Penrose inverse, the group inverse or the weighted Moore–Penrose inverse of A.

5. Examples

Example 5.1. Consider the matrix $A = \begin{pmatrix} 1 & 0 \\ -1 & 0 \\ 0 & 1 \end{pmatrix}$. If we select unitary matrix $W_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, and $W_2 = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$, we get

$$\begin{aligned} B &= W_2AW_1 = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}, \qquad \operatorname{tr}(BB^*) = 3, \\ D &= W_2A = \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}, \qquad \operatorname{tr}(DD^*) = 3, \\ \alpha &= \min\left\{\frac{2}{3}, \frac{2}{3}\right\}. \end{aligned}$$

Using the package MATHEMATICA we construct the following iterative process of the order 4:

$$\begin{aligned} Y_0 &= \alpha B^* = \begin{pmatrix} \frac{2}{3} & 0 \\ \frac{2}{3} & -\frac{2}{3} \end{pmatrix}, \\ T_0 &= I - BY_0 = \begin{pmatrix} -\frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} \end{pmatrix}, \\ Y_1 &= Y_0(I + T_0 + T_0^2 + T_0^3) = \begin{pmatrix} \frac{56}{81} & \frac{56}{81} \\ 0 & -\frac{56}{81} \end{pmatrix}. \end{aligned}$$

$$\begin{aligned} T_1 &= I - BY_1 = \begin{pmatrix} \frac{25}{81} & 0 \\ 0 & \frac{25}{81} \end{pmatrix}, \\ Y_2 &= Y_1(I + T_1 + T_1^2 + T_1^3) = \begin{pmatrix} \frac{42656096}{43046721} & \frac{42656096}{43046721} \\ 0 & -\frac{42656096}{43046721} \end{pmatrix}. \end{aligned}$$

In a similar way can be obtained

$$Y_3 = \begin{pmatrix} \frac{3433683797009448119270886198656}{3433683820292512484657849089821} & \frac{3433683797009448119270886198656}{3433683820292512484657849089821} \\ 0 & -\frac{3433683797009448119270886198656}{3433683820292512484657849089821} \end{pmatrix}.$$

Now we get the following sequence $X_k = W_1Y_kW_2$:

$$X_1 = \begin{pmatrix} 0 & -\frac{56}{81} & 0 \\ \frac{56}{81} & \frac{56}{81} & \frac{56}{81} \end{pmatrix},$$

$$X_2 = \begin{pmatrix} 0 & -\frac{42656096}{43046721} & 0 \\ \frac{42656096}{43046721} & \frac{42656096}{43046721} & \frac{42656096}{43046721} \end{pmatrix},$$

$$X_3 = \begin{pmatrix} 0 & -\frac{3433683797009448119270886198656}{3433683820292512484657849089821} & 0 \\ \frac{3433683797009448119270886198656}{3433683820292512484657849089821} & \frac{3433683797009448119270886198656}{3433683820292512484657849089821} & \frac{3433683797009448119270886198656}{3433683820292512484657849089821} \end{pmatrix}.$$

We have obtained a sequence converging to $X = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \in A\{1,2,4\}$.

Example 5.2. For the same matrices A, W_1 and $W_2 = P^* = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, we get

$$\begin{aligned} B &= W_2AW_1 = \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix}, \qquad \operatorname{tr}(BB^*) = 5, \\ D &= W_2A = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \operatorname{tr}(DD^*) = 5, \\ \alpha &= \min\left\{\frac{2}{5}, \frac{2}{5}\right\}. \end{aligned}$$

The following sequence $X_k = W_1Y_kW_2$ can be obtained:

$$X_1 = \begin{pmatrix} \frac{272}{625} & -\frac{272}{625} & 0 \\ 0 & 0 & \frac{544}{625} \end{pmatrix},$$

$$X_2 = \begin{pmatrix} \frac{76272421952}{152587890625} & -\frac{76272421952}{152587890625} & 0 \\ 0 & 0 & \frac{152544843904}{152587890625} \end{pmatrix},$$

The first row of X_3 is

$$\frac{27105054312137439165995305396124309893190067 2}{54210108624275221700372640043497085571289062 5} \qquad -\frac{271050543121374391659953053961243098931900672}{542101086242752217003726400434970855712890625} \qquad 0$$

and the second row of X_3 is

$$0 \quad 0 \quad \frac{542101086242748783319906107922486197863801344}{542101086242752217003726400434970855712890625}.$$

The limit of this sequence is $X = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} = A^\dagger$.

Example 5.3. For the same matrix A and $W_1 = Q^* = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $W_2 = P^* = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, we can generate the same sequence X_k as in Example 5.2.

Example 5.4. Now we expand Example 5.1. to include illustrations for the values of α less than the optimal, or for values which do not produce the convergence.

Modifying only $\alpha = 1/2 < 2/3$ in Example 5.1., we obtain the following sequence:

$$X_1 = \begin{pmatrix} \frac{3}{16} & -\frac{11}{16} & \frac{3}{16} \\ \frac{11}{16} & \frac{1}{2} & \frac{11}{16} \end{pmatrix},$$

$$X_2 = \begin{pmatrix} \frac{987}{65536} & -\frac{63939}{65536} & \frac{987}{65536} \\ \frac{63939}{65536} & \frac{7869}{8192} & \frac{63939}{65536} \end{pmatrix},$$

$$X_3 = \begin{pmatrix} \frac{10610209857723}{18446744073709551616} & -\frac{18446726906029374051}{18446744073709551616} & \frac{10610209857723}{18446744073709551616} \\ \frac{18446726906029374051}{18446744073709551616} & \frac{2305839536977439541}{2305843009213693952} & \frac{18446726906029374051}{18446744073709551616} \end{pmatrix}.$$

The obtained sequence converges to the same matrix as in Example 5.1., but the convergence is slower.

Similarly, using $\alpha = 1/3 < 1/2 < 2/3$ we obtain the following:

$$X_1 = \begin{pmatrix} \frac{7}{27} & -\frac{47}{81} & \frac{7}{27} \\ \frac{47}{81} & \frac{26}{81} & \frac{47}{81} \end{pmatrix},$$

$$X_2 = \begin{pmatrix} \frac{726103}{14348907} & -\frac{39522143}{43046721} & \frac{726103}{14348907} \\ \frac{39522143}{43046721} & \frac{37343834}{43046721} & \frac{39522143}{43046721} \end{pmatrix},$$

$$X_3 = \begin{pmatrix} \frac{83909608561183162716808087}{1144561273430837494885949696427} & -\frac{3433276514496608404104017015327}{3433683820292512484657849089281} & \frac{83909608561183162716808087}{1144561273430837494885949696427} \\ \frac{3433276514496608404104017015327}{3433683820292512484657849089281} & \frac{3433024785670924854615866591066}{3433683820292512484657849089281} & \frac{3433276514496608404104017015327}{3433683820292512484657849089281} \end{pmatrix}.$$

The values of α, greather than the optimal one leads to the divergence of the methods. For example, $\alpha = 1 > 2/3$ imply the following divergent sequence:

$$X_1 = \begin{pmatrix} -3 & 1 & -3 \\ -1 & 2 & -1 \end{pmatrix},$$

$$X_2 = \begin{pmatrix} -987 & 609 & -987 \\ -609 & 378 & -609 \end{pmatrix},$$

$$X_3 = \begin{pmatrix} -10610209857723 & 6557470319841 & -10610209857723 \\ -6557470319841 & 4052739537882 & -6557470319841 \end{pmatrix}.$$

Finally, a small decreasing of the parameter α near the optimal value $\frac{2}{\mathrm{tr}(B^*B)}$ imply slightly slowing down the speed of the convergence. For example, $\alpha = 2/3 - 0.001$ generates the following sequence:

$$X_1 = \begin{pmatrix} 0.00221225 & -0.692092 & 0.00221225 \\ 0.692092 & 0.68988 & 0.692092 \end{pmatrix},$$

$$X_2 = \begin{pmatrix} 0.000255561 & -0.991009 & 0.000255561 \\ 0.991009 & 0.990753 & 0.991009 \end{pmatrix},$$

$$X_3 = \begin{pmatrix} 7.1253910^{-10} & -1. & 7.1253910^{-10} \\ 1. & 1. & 1. \end{pmatrix}.$$

REFERENCES

[1] Altman, M., *An optimum cubically convergent iterative method of inverting a linear bounded operator in Hilbert space*, Pacific J. Math. **10** (1960), 1107–113.

[2] Arghiriade, E., *Remarques sur l'inverse generalise d'un produit de matrices*, Lincei-Rend. Sc. fis. mat. e nat. **XLII** (1967), 621–625.

[3] Ben-Israel, A. ; Grevile, T.N.E., *Generalized inverses: Theory and applications*, Wiley-Interscience, New York, 1974.

[4] Bouldin, R.H., *The pseudoinverse of a product*, SIAM J. Appl. Math. **24 No 4** (1973), 489–495.

[5] Cline, R.E., *Inverses of rank invariant powers of a matrix*, SIAM J. Numer. Anal. **5 No 1** (1968), 182–197.

[6] Garnett, J., Ben-Israel, A. and Yau, S.S., *A hyperpower iterative method for computing matrix products involving the generalized inverse*, SIAM J. Numer. Anal. **8** (1971), 104–109.

[7] Herzberger, J., *Using error-bounds hyperpower methods to calculate inclusions for the inverse of a matrix*, BIT **30** (1990), 508–515.

[8] Milovanović, G.V., *Numerical analysis, part I*, Naučna knjiga, Beograd, 1985 (in Serbian).

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

[10] Petryshyn, W.V., *On the inversion of matrices and linear operators*, Proc. Amer. Math. Soc. **16** (1965), 893–901.

[11] Petryshyn, W.V., *On generalized inverses and on the uniform convergence of* $(I - \beta K)^n$ *with application to iterative methods*, J. Math. Anal. Appl. **18** (1967), 417–439.

[12] Pyle, L.D., *The weighted generalized inverse in nonlinear programming-active set selection using a variable-metric generalization of the simplex algorithm*, International simposium on extremal methods and systems analysis, Lecture Notes in Economics and Mathematical Systems: Vol. 174, 1977, pp. 197–231.

[13] Radić, M, *Some contributions to the inversions of rectangular matrices*, Glasnik Matematički **1 (21) -No. 1** (1966), 23–37.

[14] Schulz, G., *Iterative Berechnung der reziproken Matrix*, Zeitsch. Angew. Math. Mech. **13** (1933), 57–59.

[15] Söderström, T. and Stewart, G.W., *On the numerical properties of an iterative method for computing the Moore-Penrose generalized inverse*, SIAM J. Numer. Anal. **11** (1974), 61–74.

[16] Tanabe, K., *Neumann-type expansion of reflexive generalized inverses of a matrix and the hyperpower iterative method*, Linear Algebra Appl. **10** (1975), 163–175.

[17] Zielke, G., *Iterative refinement of generalized matrix inverses now practicable*, SIGNUM Newsletter **13.4** (1978), 9–10.

[18] Zielke, G., *A survey of generalized matrix inverses*, Computational Mathematics, Banach center Publications **13** (1984), 499–526.

[19] Zlobec, S., *On computing the generalized inverse of a linear operator*, Glasnik Matematički **2(22) No 2** (1967), 265–271.

University of Niš
Faculty of Philosophy
Department of Mathematics
Ćirila i Metodija 2
18000 Niš
YUGOSLAVIA

Acta Math. Hungarica **79 (3)** (1998), 253–268.