# HIGHER ORDER ITERATIONS FOR MOORE-PENROSE INVERSES ${ }^{\dagger}$ 

SHWETABH SRIVASTAVA* AND D.K. GUPTA


#### Abstract

A higher order iterative method to compute the Moore-Penrose inverses of arbitrary matrices using only the Penrose equation (ii) is developed by extending the iterative method described in [1]. Convergence properties as well as the error estimates of the method are studied. The efficacy of the method is demonstrated by working out four numerical examples, two involving a full rank matrix and an ill-conditioned Hilbert matrix, whereas, the other two involving randomly generated full rank and rank deficient matrices. The performance measures are the number of iterations and CPU time in seconds used by the method. It is observed that the number of iterations always decreases as expected and the CPU time first decreases gradually and then increases with the increase of the order of the method for all examples considered.


AMS Mathematics Subject Classification : 65H05, 65F10.
Key words and phrases : Moore-Penrose inverse, Random matrices, Convergence analysis, Residual, Null space, Matrix rank.

## 1. Introduction

The theory of generalized inverses has seen a substantial growth over the past few decades. Many applications of statistics, prediction theory, control analysis and numerical analysis often require computation of generalized inverses. The Moore-Penrose inverse is one of the most important generalized inverses of arbitrary singular square or rectangular (real or complex) matrix. It has been extensively studied by many researchers $[4,2,6,9,5,3]$ and many methods are proposed in the literature. Accordingly, it is important both practically and theoretically to find good higher order algorithms for computing a Moore-Penrose inverse of a given arbitrary matrix. Let $\mathbb{C}^{m \times n}$ and $\mathbb{C}_{r}^{m \times n}$ denote the set of

[^0]all complex $(m \times n)$ matrices and all complex $(m \times n)$ matrices with rank $r$, respectively. For $A \in \mathbb{C}^{m \times n}$, it is denoted by $A^{\dagger}$ and defined as
\[

$$
\begin{aligned}
A A^{\dagger} & =P_{R(A)} \\
A^{\dagger} A & =P_{R\left(A^{\dagger}\right)}
\end{aligned}
$$
\]

where, $R(A)$ and $P_{R(A)}$ denote the range space of $A$ and orthogonal projection on $R(A)$ respectively. The unique matrix $A^{\dagger}$ satisfies the following four equations
(i) $A X A=A,\left(\right.$ ii) $X A X=X,($ iii $)(A X)^{*}=A X,(i v)(X A)^{*}=X A$

Both direct and iterative methods (cf. [11, 10, 2, 3, 9]) can be used to compute $A^{\dagger}$. One of the most commonly used direct methods is the Singular Value Decomposition (SVD) method. For $A \in \mathbb{C}^{m \times n}$, the SVD method is a factorization of the form

$$
A=U \Sigma V^{*}
$$

where, $U$ is an $m \times m$ complex unitary matrix, $\Sigma$ is an $m \times n$ rectangular diagonal matrix with nonnegative real numbers on the diagonal and $V$ is an $n \times n$ complex unitary matrix. The $A^{\dagger}$ can then be written as

$$
A^{\dagger}=V \Sigma^{\dagger} U^{*}
$$

where, $\Sigma^{\dagger}$ is the generalized inverse of $\Sigma$ obtained by replacing every nonzero diagonal entry in $\Sigma$ by its reciprocal and then transposing the resulting matrix. This method is very accurate but time-intensive since it requires a large amount of computational resources, especially in the case of large matrices. The most frequently used iterative method for approximating $A^{-1}$ is the method studied by Householder [14] by analyzing successive improvements of a matrix $X$ to solve $A X=M, A$ nonsingular, using the equation

$$
\begin{equation*}
X_{k+1}=X_{k}+C_{k}\left(M-A X_{k}\right), k \geq 0 \tag{2}
\end{equation*}
$$

One version of the above has the matrices $C_{k}$ project on the $i$ th row, where $i$ cycles through all the rows of $A$. By taking $M=I$ and $C_{k}=V_{k}$, the famous Newton's method originated in [21] is

$$
\begin{equation*}
V_{k+1}=V_{k}\left(2 I-A V_{k}\right), k \geq 0 \tag{3}
\end{equation*}
$$

As usual, $I$ denotes the identity matrix of an appropriate order. If $L$ is the desired limit matrix and $V_{k}$ is the $k$-th estimate of $L$, then the convergence properties of examined iterative method can be studied with the aid of error $\operatorname{matrix} E_{k}=V_{k}-L$. If an iterative method is expressible as a simple matrix formula, $E_{k+1}$ is sum of several terms

- zero order term consisting of a matrix which does not depend on $E_{k}$,
- one or more first order matrix terms in which $E_{k}$ or its conjugate transpose $E_{k}^{*}$ appears only once,
- higher-order terms in which $E_{k}$ or $E_{k}^{*}$ appears at least twice

All suitable algorithms have a zero-order term equal to 0 . Hence, the firstorder term determine the terminal convergence properties. The eigenvalues of $A \in \mathbb{C}_{r}^{n \times n}$ are given by

$$
\begin{equation*}
\lambda_{1}(A) \geq \cdots \geq \lambda_{r}(A)>\lambda_{r+1}(A)=\cdots \lambda_{n}(A)=0 \tag{4}
\end{equation*}
$$

It is further established that the eigenvalues of $I-A V_{0}$ must have magnitude less than 1 to ensure the convergence of (3). Since the residuals $R_{k}=I-A V_{k}$ in (3) satisfies $\left\|R_{k+1}\right\| \leq\|A\|\left\|R_{k}\right\|^{2}$, Newton's method is a second order iterative method [7]. Ben-Israel [11] used (3) and the starting value $X_{0}=\alpha A^{*}$, where $\alpha$ satisfies

$$
\begin{equation*}
0<\alpha<\frac{2}{\lambda_{1}\left(A A^{*}\right)} . \tag{5}
\end{equation*}
$$

The iterative scheme (3) is generalized by the iteration $U_{k+1}=U_{k}\left(2 P_{R}(A)-\right.$ $A U_{k}$ ), which converges to $A^{\dagger}$. Ben-Israel and Charnes [13] proved that the sequence

$$
\begin{equation*}
Y_{k}=\alpha \Sigma_{i=0}^{k} A^{*}\left(I-\alpha A A^{*}\right)^{i}, k=0,1, \ldots \tag{6}
\end{equation*}
$$

converges to $A^{\dagger}$ under the assumption (5). Li et al.[6] established a family of iterative methods to compute the approximate inverse of a square matrix and inner inverse of a non-square matrix given by

$$
\begin{equation*}
V_{q+1}=V_{q}\left(k I-\frac{k(k-1)}{2} A V_{q}+\ldots+(-1)^{k-1}\left(A V_{q}\right)^{k-1}\right), k=2,3, \ldots \tag{7}
\end{equation*}
$$

Chen and Wang [4] showed that this can be extended to compute $A^{\dagger}$ with higher orders. Katskis et.al.[8] developed a much faster computational method to calculate the $A^{\dagger}$ of singular square matrices and of rectangular matrices $A$. This method has significantly better accuracy than the already proposed methods and works for full and sparse matrices. Weiguo et al.[5] improves on the generalized properties of a family of iterative methods to compute the approximate inverses of square matrices and fixed inner inverses of rectangular matrices proposed in [6]. They have also established that this fixed inverse is the Moore-Penrose inverse of the considered matrix. Vasilios et.al.[3] also presented a very fast and reliable method to compute Moore-Penrose inverse. By using a general framework where analytic functions of scalers are first developed and then matrices substituted for the scalers, Katsaggelos and Efstratiadis [12] produced a convergence faster than quadratic, for restricted initial estimates.

In this paper, a higher order iterative method to compute the Moore-Penrose inverses of arbitrary matrices using only the Penrose equation (ii) is developed by extending the iterative method described in [1]. Convergence properties as well as the error estimates of the method are studied. The efficacy of the method is demonstrated by working out four numerical examples, two involving a full rank simple and an ill-conditioned Hilbert matrix, whereas, the other two involving full rank and rank deficient randomly generated matrices. The performance measures are the number of iterations and CPU time in seconds used by the
method. It is observed that the number of iterations always decreases as expected and the CPU time first decreases gradually and then increases with the increase of the order of the method for all examples considered.

The paper is organized as follows. Section 1 is the introduction. In Section 2, the higher order iterative method for computing the Moore-Penrose generalized inverse of an arbitrary rectangular matrix is described. Some Lemmas and the convergence analysis of the method are also established. In Section 3, the efficacy of the method is demonstrated by working out four numerical examples, two involving a full rank simple and an ill-conditioned Hilbert matrix, whereas, the other two involving full rank and rank deficient randomly generated matrices. Finally, conclusions are included in Section 4.

## 2. Higher order iteration with convergence Analysis

In this section, higher order iterative methods and their convergence analysis to compute the Moore-Penrose inverse $A^{\dagger}$ of an arbitrary complex matrix $A \in \mathbb{C}^{m \times n}$ is described. Let $A^{t}, A^{*}, \mu(A), \nu(A)$ and $\operatorname{rank}(A)$, represent the transpose, the conjugate transpose, the range space, the null space and the rank of the matrix $A \in \mathbb{C}^{m \times n}$, respectively.

Definition 2.1 ([19]). Let $A \in \mathbb{C}^{m \times n}$, with column vectors $v_{1}, v_{2}, \ldots, v_{n}$, the set

$$
\mu(A)=\left\{c_{1} v_{1}+c_{2} v_{2}+\ldots+c_{n} v_{n} \mid c_{1}, c_{2}, \ldots, c_{n} \text { are scalers }\right\}
$$

are called the Range (the Image, or the Column Space) of $A$. The set

$$
v(A)=\left\{x \in C^{n} \mid A x=0\right\}
$$

is called the Null Space (or the Kernel) of $A$ and for any $A \in \mathbb{C}^{m \times n}, B \in \mathbb{C}^{s \times t}$, we call

$$
\begin{gathered}
\mu(A, B)=\left\{X \in \mathbb{C}^{m \times t} \mid X=A Y B, \forall Y \in \mathbb{C}^{n \times s}\right\} \\
\nu(A, B)=\left\{Y \in \mathbb{C}^{n \times s} \mid A Y B=0\right\}
\end{gathered}
$$

the Range of $(A, B)$ and the Null Space of $(A, B)$ respectively.
We shall now describe the higher order iterative method extending the method described in [1] for computing the Moore-Penrose generalized inverses. Petković et al. [1] proposed a quadratically convergent iterative method for computing $A^{\dagger}$ based on Penrose equations (ii) and (iv) as follows. Let

$$
X^{*}=(X A X)^{*}=X^{*}(X A)^{*}=X^{*} X A
$$

Hence, for arbitrary $\beta \in \mathbb{R}$, we get

$$
X^{*}=X^{*}-\beta\left(X^{*} X A-X^{*}\right)=X^{*}(I-\beta X A)+\beta X^{*}
$$

or equivalently,

$$
X=(I-\beta X A)^{*} X+\beta X
$$

This leads to the following iterative method

$$
\begin{equation*}
X_{k+1}=(1+\beta) X_{k}-\beta X_{k} A X_{k}, \quad k=0,1, \ldots \tag{8}
\end{equation*}
$$

where, $X_{0}=\beta A^{*}$ for an appropriate real number $\beta$. For $\beta<1$, the method (8) has a linear convergence while for $\beta=1$ its convergence is quadratic. The first-order and the second-order error terms of (8) are

$$
\begin{equation*}
\text { error }_{1}=(1-\beta) E_{k}, \text { error }_{2}=-\beta E_{k} A E_{k} \tag{9}
\end{equation*}
$$

where, $E_{k}=X_{k}-A^{\dagger}$ is the error matrix. Now, using only the Penrose equations (ii) given by $X=X A X$ and for arbitrary $\beta \in \mathbb{R}$, we get

$$
X=X+\beta(2 X-3 X A X+X A X A X)
$$

or, equivalently

$$
X=X+\beta X\left(2 I-3 A X+(A X)^{2}\right)
$$

This leads to the following third order extension of method (8)

$$
\begin{equation*}
X_{k+1}=X_{k}+\beta X_{k}\left(2 I-3 A X_{k}+\left(A X_{k}\right)^{2}\right), k=0,1, \ldots \tag{10}
\end{equation*}
$$

for $X_{0}=\beta A^{*}$. Continuing in a similar manner, this can further be extended to the $p$ th order for $p \geq 2$, given by
$X_{k+1}=X_{k}+\beta X_{k}\left[\left(I-A X_{k}\right)+\left(I-A X_{k}\right)^{2}+\ldots+\left(I-A X_{k}\right)^{p-1}\right], k=0,1,2, \ldots$
for $X_{0}=\beta A^{*}$, where $\beta$ is an appropriate real number. The following Lemmas will be used for establishing the convergence of these iterative methods.

Lemma 2.2. For all $k \geq 0$, the sequence $\left\{X_{k}\right\}$ generated by (10) and (11) satisfies
(1) $X_{k} A=\left(X_{k} A\right)^{*}$
(2) $X_{k} A A^{\dagger}=X_{k}$
(3) $A^{\dagger} A X_{k}=X_{k}$

Proof. This Lemma can be proved by induction. Clearly, (1) holds for $k=0$ as $X_{0} A=\beta A^{*} A=\left(X_{0} A\right)^{*}$. Assume that it holds for some $k$. To show that it also holds for $k+1$, we consider

$$
\begin{aligned}
\left(X_{k+1} A\right)^{*} & =\left(X_{k} A\right)^{*}+\beta\left\{2\left(X_{k} A\right)^{*}-3\left(\left(X_{k} A\right)^{*}\right)^{2}+\left(\left(X_{k} A\right)^{*}\right)^{3}\right\} \\
& =X_{k} A+\beta\left(2 X_{k} A-3\left(X_{k} A\right)^{2}+\left(X_{k} A\right)^{3}\right) \\
& =X_{k} A+\beta\left(2 X_{k}-3 X_{k} A X_{k}+X_{k}\left(A X_{k}\right)^{2}\right) A \\
& =X_{k+1} A
\end{aligned}
$$

Hence it holds for all $k \geq 0$. Likewise for $p$ th order method (11), we get

$$
\begin{aligned}
\left(X_{k+1} A\right)^{*} & =\left\{X_{k} A+\beta X_{k}\left[\left(I-A X_{k}\right)+\ldots\left(I-A X_{k}\right)^{p-1}\right] A\right\}^{*} \\
& =\left(X_{k} A\right)^{*}+\beta\left[\left(X_{k}\left(I-A X_{k}\right) A\right)^{*}+\ldots+\left(X_{k}\left(I-A X_{k}\right)^{p-1} A\right)^{*}\right] \\
& =X_{k} A+\beta\left[X_{k}\left(I-A X_{k}\right) A+\ldots+X_{k}\left(I-A X_{k}\right)^{p-1} A\right] \\
& =X_{k} A+\beta X_{k}\left[\left(I-A X_{k}\right)+\ldots\left(I-A X_{k}\right)^{p-1}\right] A \\
& =X_{k+1} A
\end{aligned}
$$

Clearly, (2) trivially holds for $k=0$. Let it also holds for some $k$. To show that it also holds for $k+1$, we get

$$
\begin{aligned}
X_{k+1} A A^{\dagger} & =X_{k} A A^{\dagger}+\beta X_{k}\left(2 I-3 A X_{k}+\left(A X_{k}\right)^{2}\right) A A^{\dagger} \\
& =X_{k}+\beta X_{k}\left(2 I-3 A X_{k}+\left(A X_{k}\right)^{2}\right) \\
& =X_{k+1}
\end{aligned}
$$

Likewise for $p$ th order method (11), we get

$$
\begin{aligned}
X_{k+1} A A^{\dagger} & =X_{k} A A^{\dagger}+\beta X_{k}\left[\left(I-A X_{k}\right)+\ldots+\left(I-A X_{k}\right)^{p-1}\right] A A^{\dagger} \\
& =X_{k}+\beta X_{k}\left[\left(I-A X_{k}\right)+\ldots+\left(I-A X_{k}\right)^{p-1}\right] \\
& =X_{k+1}
\end{aligned}
$$

Proceeding in a similar manner, (3) can easily be proved for (10) and (11). Hence, the Lemma 2.1 is proved for all $k \geq 0$.
Theorem 2.3. Let $0 \neq A \in \mathbb{C}^{m \times n}, X=A^{\dagger}$, the initial approximation $X_{0}=$ $\beta A^{*}, \beta \in(0,1]$ and its residual $R_{0}=\left(X_{0}-X\right) A$ satisfy $\left\|R_{0}\right\|<1$. The sequence $\left\{X_{k}\right\}$ generated by (10) starting with $X_{0}=\beta A^{*}$ converges to the Moore-Penrose inverse $A^{\dagger}$. It has linear convergence for $\beta \neq 1$ and third order convergence for $\beta=1$. Its first, second and the third order error terms are given by

$$
\begin{equation*}
\text { error }_{1}=(1-\beta) E_{k}, \text { error }_{2}=0 \text { and } \text { error }_{3}=\beta E_{k}\left(A E_{k}\right)^{2} \tag{12}
\end{equation*}
$$

where, $E_{k}=X_{k}-A^{\dagger}$ denotes the error matrix.
Proof. Using Lemma 2.2 and substituting for $X_{k+1}$, we get

$$
\left\|X_{k+1}-X\right\|=\left\|X_{k+1} A X-X A X\right\| \leq\left\|X_{k+1} A-X A\right\|\|X\|
$$

Using Lemma 2.2 and (10), we get

$$
\begin{aligned}
X_{k+1} A-X A & =\left(X_{k}+\beta X_{k}\left(2 I-3 A X_{k}+\left(A X_{k}\right)^{2}\right) A-X A\right. \\
& =X_{k} A-X A+\beta\left(2 X_{k} A-3\left(X_{k} A\right)^{2}+\left(X_{k} A\right)^{3}\right) \\
& =X_{k} A-X A+\beta\left(\left(X_{k} A-X A\right)^{3}-\left(X_{k} A-X A\right)\right)
\end{aligned}
$$

Thus, the sequence of residual matrices defined by $R_{k}=X_{k} A-X A$ satisfies the following recurrence relation

$$
\begin{equation*}
R_{k+1}=(1-\beta) R_{k}+\beta R_{k}^{3} \tag{13}
\end{equation*}
$$

Let $s_{k}=\left\|R_{k}\right\|$. Now, for the convergence of the sequence $\left\{X_{k}\right\}$, we require that $s_{k} \rightarrow 0$ as $k \rightarrow \infty$. This can be shown by mathematical induction. It trivially holds for $\mathrm{k}=0$, since, $s_{0}=\left\|R_{0}\right\|=\left\|X_{0} A-X A\right\|<1$. Let it holds for some $k$, i.e., $s_{k}<1$. To show that it also holds for $k+1$, we take norm on (13) and get

$$
\begin{equation*}
s_{k+1} \leq \beta s_{k}^{3}+(1-\beta) s_{k}<\beta s_{k}+(1-\beta) s_{k}=s_{k}<1 \tag{14}
\end{equation*}
$$

Thus, $s_{k}$ is a monotonically decreasing bounded sequence converging to $s$ as $k \rightarrow \infty$ and $0 \leq s<1$. From (14), we get

$$
s \leq \beta s^{3}+(1-\beta) s
$$

This gives either $s=0$ or $s \geq 1$. Thus, $s=0$. This completes the proof that $s_{k} \rightarrow 0$ when $k \rightarrow \infty$. Thus, $X_{k} \rightarrow X$ as $k \rightarrow \infty$. Substituting $X_{k}=X+E_{k}$ in (10) and rearranging the terms, we get $E_{k+1}$ given by

$$
\begin{array}{r}
E_{k+1}=(1+2 \beta) E_{k}-2 \beta E_{k} A X-3 \beta X A E_{k}+2 \beta X A E_{k} A X-3 \beta E_{k} A E_{k}+ \\
2 \beta E_{k} A E_{k} A X+\beta X\left(A E_{k}\right)^{2}+\beta E_{k}\left(A E_{k}\right)^{2}
\end{array}
$$

This leads to

$$
\begin{aligned}
& \text { error }_{1}=(1-\beta) E_{k} \\
& \text { error }_{2}=0
\end{aligned}
$$

and

$$
\text { error }_{3}=\beta E_{k}\left(A E_{k}\right)^{2}
$$

One can easily see that the method is linear convergent for $\beta \neq 1$ and of third order convergent for $\beta=1$ as error $_{1}$ and error $_{2}$ are equal to 0 .

Theorem 2.4. Let $0 \neq A \in \mathbb{C}^{m \times n}, X=A^{\dagger}$, the initial approximation $X_{0}=$ $\beta A^{*}, \beta \in(0,1]$ and its residual $R_{0}=\left(X_{0}-X\right) A$ satisfy $\left\|R_{0}\right\|<1$. The sequence $\left\{X_{k}\right\}$ generated by (11) starting with $X_{0}=\beta A^{*}$ converges to the Moore-Penrose inverse $A^{\dagger}$. It has linear convergence for $\beta \neq 1$ and pth order convergence for $\beta=1$, where, $p \geq 2$ is a positive integer. All error terms of the method vanishes except the first and the pth order error terms given by

$$
\begin{equation*}
\text { error }_{1}=(1-\beta) E_{k}, \text { error }_{p}=(-1)^{p-1} \beta E_{k}\left(A E_{k}\right)^{p-1} \tag{15}
\end{equation*}
$$

Proof. Using Lemma 2.2 and substituting for $X_{k+1}$, we get

$$
\left\|X_{k+1}-X\right\|=\left\|X_{k+1} A X-X A X\right\| \leq\left\|X_{k+1} A-X A\right\|\|X\|
$$

Again using Lemma 2.2 and (11), this gives

$$
X_{k+1} A-X A=\left(X_{k}+\beta X_{k}\left[\left(I-A X_{k}\right)+\left(I-A X_{k}\right)^{2}+\ldots+\left(I-A X_{k}\right)^{p-1}\right]\right) A-X A
$$

Thus, the sequence of residual matrices defined by $R_{k}=X_{k} A-X A$ satisfies the following recurrence relation

$$
\begin{equation*}
R_{k+1}=(1-\beta) R_{k}+(-1)^{p-1} \beta R_{k}^{p} \tag{16}
\end{equation*}
$$

Proceeding similar to theorem 2.3, it can be proved that $X_{k} \rightarrow X$ as $k \rightarrow \infty$. Substituting $X_{k}=X+E_{k}$ in (11) and rearranging the terms, we get $E_{k+1}$ given by

$$
E_{k+1}=(1+\beta) E_{k}-\beta X A E_{k}-\beta E_{k} A X+(-1)^{p-1} \beta E_{k}\left(A E_{k}\right)^{p-1}
$$

On simplification, we get all error terms equal to zero except error $_{1}$ and error $_{p}$ given by

$$
\begin{aligned}
& \text { error }_{1}=(1-\beta) E_{k} \\
& \text { error }_{p}=(-1)^{p-1} \beta E_{k}\left(A E_{k}\right)^{p-1}
\end{aligned}
$$

One can easily see that the method is linearly convergent for $\beta \neq 1$ and of $p$ th order convergent for $\beta=1$ as all other error terms vanishes. Hence, the theorem is proved.

We must note here that the convergence of the above methods (10) and (11) require the condition $\left\|\left(\beta A^{*}-X\right) A\right\|<1$. To verify this, we need an equivalent condition which does not contain the Moore-Penrose inverse $X$ or $A^{\dagger}$. We follow exactly the same way as done in [1] by using the following Lemma.

Lemma 2.5 ([1]). Let the eigenvalues of matrix $A^{*} A$ satisfies (4). Condition $\rho\left(\left(\beta A^{*}-X\right) A\right)<1$ is satisfied under the assumption

$$
\begin{equation*}
\max _{1 \leq i \leq r}\left|1-\beta \lambda_{i}\left(A^{*} A\right)\right|<1 \tag{17}
\end{equation*}
$$

To establish this Lemma, we need the following Lemmas.
Lemma 2.6 ([15]). Let $M \in \mathbb{C}^{n \times n}$ and $\epsilon>0$ be given. There is at least one matrix norm $\|$. || such that

$$
\rho(M) \leq\|M\| \leq \rho(M)+\epsilon
$$

where $\rho(M)=\max \left\{\left|\lambda_{1}(M)\right|, \ldots,\left|\lambda_{n}(M)\right|\right\}$ denotes the spectral radius of $M$.
Lemma 2.7 ([18]). If $P \in \mathbb{C}^{n \times n}$ and $S \in \mathbb{C}^{n \times n}$ are such that $P=P^{2}$ and $P S=S P$ then

$$
\rho(P S) \leq \rho(S)
$$

Remark 2.1. For any $A \in \mathbb{C}^{m \times n}$, the sequence generated by our higher order iterative methods starting with $X_{0}=\beta A^{*}$ are convergent for any $\beta$ satisfying $0<\beta<2 / \sigma^{2}(A)$, where $\sigma(A)=\|A\|_{2}$.

## 3. Numerical examples

In this section, four numerical examples are worked out to demonstrate the efficacy of our $p$ th order method for various values of $p$. The mean CPU time in second and the number of iterations are measured as performance parameters. For full rank matrices the number of iterations used are compared by plotting figures with $x$-axis representing the values of $p$ and $y$-axis representing the number of iterations. For randomly generated matrices, we have tested 50 times matrices and the mean CPU time in second are measured and tabulated for values of $2 \leq p \leq 7$ and the number of iterations are compared by plotting figures with $x$-axis representing the times of choice of random matrices and $y$ axis representing the number of iterations only for $p=2$ and $p=3$. The value of $\epsilon=10^{-7}$ and the maximum number of iterations equal to 100 are taken as the termination criteria. All the examples are run on an Intel core 2 Duo processor running at 2.80 GHz and using MATLAB 7.4 (R2009b).
Example 3.1. Consider the matrix $A$ of order $(5 \times 4)$ given by

$$
A=\left(\begin{array}{llll}
0.2794 & 0.1676 & 0.0645 & 0.2326 \\
0.0065 & 0.2365 & 0.2274 & 0.1261 \\
0.2271 & 0.1430 & 0.1009 & 0.2867 \\
0.1265 & 0.1015 & 0.1806 & 0.2846 \\
0.2773 & 0.0632 & 0.0503 & 0.1979
\end{array}\right)
$$

The choice $\beta=0.60$ satisfies the convergence criteria given by

$$
\max _{1 \leq i \leq 4}\left|1-\beta \lambda_{i}\left(A^{*} A\right)\right|=0.9988<1
$$

since the eigenvalues of the matrix $A^{*} A$ are

$$
\left(\lambda_{1}, \lambda_{2}, \lambda_{3}, \lambda_{4}\right)=(0.0020,0.0146,0.0832,0.6152)
$$

The iterative method (11) generates a sequence of iterates $\left\{X_{k}\right\}$ converging to the Moore-Penrose generalized inverse $A^{\dagger}$ given by

$$
A^{\dagger}=\left(\begin{array}{ccccc}
-0.2165 & 1.4802 & -4.9702 & -1.3732 & 8.4865 \\
5.0277 & 1.8673 & 4.1653 & -4.6975 & -6.3778 \\
-5.3215 & 4.5524 & -8.4278 & 3.4688 & 10.5748 \\
0.8566 & -4.0180 & 6.9330 & 3.0649 & -7.8449
\end{array}\right)
$$

The comparison of number of iterations are plotted in Figure 1. It can be observed from Figure 1 that the iterative method (11) converges to the MoorePenrose generalized inverse $A^{\dagger}$ in 36 iterations for $p=2$ and as expected as the order $p$ increases, it reduces to 25 for $p=8$.


Figure 1. Number of iterations versus the value of $p$ for example 3.1

Example 3.2. Consider a $(30 \times 30)$ matrix $A$ whose elements are generated randomly from $[-0.2,0.2]$. Taking the termination criteria as
$\max \left\{\left\|A X_{k} A-A\right\|_{F},\left\|X_{k} A X_{k}-X_{k}\right\|_{F},\left\|\left(A X_{k}\right)^{*}-A X_{k}\right\|_{F},\left\|\left(X_{k} A\right)^{*}-X_{k} A\right\|_{F}\right\} \leq \epsilon$
where, $\|\cdot\|_{F}$ stands for the Frobenius-norm of a matrix, Table 1 shows the comparison of mean CPU time in seconds for for $2 \leq p \leq 7$. One can easily see that with the increase of $p$, the CPU time decreases initially and then starts increases after $p>4$. The comparison of number of iterations are plotted in Figure 2 for $p=2$ and $p=3$. Similar results are also observed for higher order random matrices, for example, $(40 \times 40)$ matrix $A$ whose elements are also randomly generated from $[-0.2,0.2]$.

Table 1. Comparison of CPU time

| p | $\mathrm{p}=2$ | $\mathrm{p}=3$ | $\mathrm{p}=4$ | $\mathrm{p}=5$ | $\mathrm{p}=6$ | $\mathrm{p}=7$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| CPU time | $2.1387 e-5$ | $1.6948 e-5$ | $1.6245 e-5$ | $1.6554 e-5$ | $1.7205 e-5$ | $1.8946 e-5$ |



Figure 2. Comparison of number of iterations for example 3.2

Example 3.3. Consider a $(100 \times 50)$ rank deficient matrix $A$ whose elements are generated randomly from $[-0.2,0.2]$. Taking the termination criteria as

$$
\max \left\{\left\|A X_{k} A-A\right\|_{F},\left\|X_{k} A X_{k}-X_{k}\right\|_{F},\left\|\left(A X_{k}\right)^{*}-A X_{k}\right\|_{F},\left\|\left(X_{k} A\right)^{*}-X_{k} A\right\|_{F}\right\} \leq \epsilon
$$

where, $\|\cdot\|_{F}$ stands for the Frobenius-norm of a matrix, Table 2 shows the comparison of mean CPU time in seconds for for $2 \leq p \leq 7$. One can easily see that with the increase of $p$, the CPU time decreases initially and then starts increases after $p \geq 4$. The comparison of number of iterations are plotted in Figure 3 for $p=2$ and $p=3$.


Figure 3. Comparison of number of iterations for example 3.3

Table 2. Comparison of CPU time

| p | $\mathrm{p}=2$ | $\mathrm{p}=3$ | $\mathrm{p}=4$ | $\mathrm{p}=5$ | $\mathrm{p}=6$ | $\mathrm{p}=7$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| CPU time | $2.5287 e-5$ | $1.7793 e-5$ | $1.8346 e-5$ | $1.9451 e-5$ | $2.0135 e-5$ | $2.1137 e-5$ |

Example 3.4. Consider the ill-conditioned Hilbert matrix $A$ of order $(5 \times 5)$ given by

$$
A=\left(\begin{array}{lllll}
1.0000 & 0.5000 & 0.3333 & 0.2500 & 0.2000 \\
0.5000 & 0.3333 & 0.2500 & 0.2000 & 0.1667 \\
0.3333 & 0.2500 & 0.2000 & 0.1667 & 0.1429 \\
0.2500 & 0.2000 & 0.1667 & 0.1429 & 0.1250 \\
0.2000 & 0.1667 & 0.1429 & 0.1250 & 0.1111
\end{array}\right)
$$

The choice $\beta=0.80$ satisfies the convergence criteria given by

$$
\max _{1 \leq i \leq 5}\left|1-\beta \lambda_{i}\left(A^{*} A\right)\right|=0.9999<1
$$

since the eigenvalues of the matrix $A^{*} A$ are

$$
\left(\lambda_{1}, \lambda_{2}, \lambda_{3}, \lambda_{4}, \lambda_{5}\right)=(2.4556,0.0435,0.0001,9.357 e-8,1.081 e-11)
$$

The iterative method (11) generates a sequence of iterates $\left\{X_{k}\right\}$ converging to the Moore-Penrose generalized inverse $A^{\dagger}$ given by

$$
A^{\dagger}=\left(\begin{array}{ccccc}
20 & -300 & 1050 & -1400 & 630 \\
-300 & 4800 & -18900 & 26880 & -12600 \\
1050 & -18900 & 79380 & -117600 & 56700 \\
-1400 & 26880 & -117600 & 179200 & -88200 \\
630 & -12600 & 56700 & -88200 & 44100
\end{array}\right)
$$

The comparison of number of iterations are plotted in Figure 4. It can be observed from Figure 4 that the iterative method (11) converges to the MoorePenrose generalized inverse $A^{\dagger}$ in 51 iterations for $p=2$ and as expected as the order $p$ increases, it reduces to 19 for $p=10$.


Figure 4. Number of iterations versus the value of $p$ for example 3.4

## 4. Conclusions

A quadratically convergent iterative method proposed by [1] is extended to a family of higher order iterative methods to compute the Moore-Penrose generalized inverses of arbitrary singular or rectangular (real or complex) matrices. This extension is carried out by using only the Penrose equation (ii) in place of (ii) and (iv) as used in [1]. Convergence properties as well as the error estimations are studied. The efficacy of the method is demonstrated by working out four numerical examples involving full rank and rank deficient randomly generated matrices and ill-conditioned Hilbert matrix. The performance in terms of computational time and the number of iterations are evaluated with respect to the order of the iterative methods. It is observed that the number of iterations
always decreases as expected and the CPU time first decreases gradually and then increases with respect to the order of the iterative methods for all examples.

## References

1. Marko D. Petković and Predrag S.Stanimirović, Iterative method for computing MoorePenrose inverse based on Penrose equations, Journal of Computational and Applied Mathematics 235 (2011), 1604-1613.
2. W. Guo and T. Huang, Method of elementary transformation to compute Moore-Penrose inverse, Appl. Math. Comput. 216 (2010), 1614-1617.
3. Vasilios N. Katsikis, Dimitrios Pappas and Athanassios Petralias, An improved method for the computation of the Moore-Penrose inverse matrix, Appl. Math. Comput. 217 (2011), 9828-9834.
4. Haibin Chen and Yiju Wang, A Family of higher-order convergent iterative methods for computing the Moore-Penrose inverse, Appl. Math. Comput. 218 (2011), 4012-4016.
5. Li Weiguo, Li Juan and Qiao Tiantian, A family of iterative methods for computing MoorePenrose inverse of a matrix, Linear Algebra and its Applications 438 (2013), 47-56.
6. Weiguo Li and Zhi Li, A family of iterative methods for computing the approximate inverse of a square matrix and inner inverse of a non-square matrix, Appl. Math. Comput. 215 (2010), 3433-3442
7. V. Pan and R. Schreiber, An improved Newton iteration for the generalized inverse of a matrix with applications, SIAM J. Sci. Statist. Comput. 12 (1991), 1109-1130.
8. V.N. Katsikis, D. Pappas and T. Huang, Fast computing of the Moore-Penrose inverse matrix, Electronic Journal of Linear Algebra 17 (2008), 637-650.
9. F. Toutounian and A. Ataei, A new method for computing Moore-Penrose inverse matrices, Journal of Computational and applied Mathematics 228 (2009), 412-417.
10. P. Courrieu, Fast Computation of Moore-Penrose Inverse matrices, Neural Information Processing-Letters and Reviews 8 (2005), 25-29.
11. A. Ben-Israel and T. N. E. Greville, Generalized Inverses: Theory and Applications, Springer-Verlag, NewYork, 2003.
12. A.K.Katsaggelos and S.N. Efstratiadis, A class of iterative signal restoration algorithms, IEE Trans. Acoust. Speech Signal Process, 38 (1990), 778-793.
13. A. Ben-Israel and A.Charnes, Contribution to the theory of generalized inverses, J.Soc.Indust.Appl.Math. 11 (1963), 667-669.
14. A.S.Householder, The Theory of Matrices in Numerical Analysis, Dover Publications, Newyork, 1964.
15. R.A. Horn and C.R. Johnson, Matrix Analysis Cambridge University Press, Cambridge, UK, 1986.
16. F.J. Hall and C.D. Meryer Jr., Generalized inverses of the fundamental bordered matrix used in linear estimation, Sankhya Ser. A 37 (1975), 428-438.
17. B. Nacevska, Iterative methods for computing generalized inverses and splittings of operators, Appl. Math. Comput., 208 (2009), 186-188.
18. P.S.Stanimirovic and D.S. Cvetkovic-illic, Succesive matrix squaring algorithm for computing outer inverses, Appl. Math. Comput., 203 (2008), 19-29.
19. W. Sun and Y. Yuan, Optimization Theory and Methods- Nonlinear Programming, Springer, New York, 2006.
20. Stephen L. Campbell, and Carl D. Meyer, Generalized Inverses of Linear Transformations, SIAM, 2009.
21. G.Schultz, iterative Berechmmg der reziproken matrix, Angew. Math. Mech., 13 (1933), 57-59.

Shwetabh Srivastava is a Ph.D. research scholar in the Department of Mathematics, Indian Institute of Technology Kharagpur, India. His research interests include Theory of Generalized inverses and their applications.
Department of Mathematics, IIT Kharagpur, India.
e-mail: shwetabh20059@gmail.com
D.K. Gupta is a senior professor in the Department of Mathematics, Indian Institute of Technology Kharagpur, India. His research interests are Numerical Analysis and Computer Science.
Department of Mathematics, IIT Kharagpur, India.
e-mail: dkg@maths.iitkgp.ernet.in


[^0]:    Received March 10, 2013. Revised April 22, 2013. Accepted April 26, 2013. ${ }^{*}$ Corresponding author. ${ }^{\dagger}$ This work was supported by the research grant of the Indian Institute of Technology Kharagpur, India.
    (c) 2014 Korean SIGCAM and KSCAM.

