INTERVAL ITERATIVE METHODS FOR COMPUTING MOORE-PENROSE INVERSE *

Xian Zhang [1] Jianfeng Cai [2] Yimin Wei [3]

[1,3] School of Mathematical Sciences, Fudan University
Shanghai 200433, People's Republic of China and Key Laboratory of Mathematics for Nonlinear Sciences (Fudan University), Ministry of Education.
E-mail: 0018123@fudan.edu.cn, ymwei@fudan.edu.cn

[2] Department of Mathematics, The Chinese University of Hong Kong
Hong Kong, P.R. of China. E-mail: jfcai@math.cuhk.edu.hk

Abstract

In this paper, we import interval method to the iteration for computing Moore-Penrose inverse of the full row (or column) rank matrix. Through modifying the classical Newton iteration by interval method, we can get better numerical results. The convergence of the interval iteration is proven. We also give some numerical examples to compare interval iteration with classical Newton iteration.

Keywords: Interval method, Newton iteration, Moore-Penrose inverse.

AMS classifications: 15A18, 65F15, 65F10.

1 Introduction and preliminaries

The concept of interval analysis [3] is to compute with interval of real numbers in place of real numbers. While floating point arithmetic is affected by rounding errors, and can produce inaccuracy results, interval arithmetic has the advantage of giving rigorous bounds for the exact solution. One application is that some parameters are not known exactly but only known to lie within a certain interval; algorithms may be implemented using interval arithmetic with uncertain parameters as intervals to produce an interval which bounds all possible results. If the lower and upper bounds of the interval can be rounded up and down respectively, then the finite precision calculations can be performed intervals to give an enclosure of the exact solution. As for the interval iterative method for computing the Moore-Penrose inverse of a matrix A, if we can ensure that $A^\dagger$ contained by the interval solution at every iteration step, and the width of the intervals can be reduced, then we can get the enclosure with $A^\dagger$. The recent results on computing Moore-Penrose inverse and linear least squares problem can be found in [5-8, 13-22].

This paper gives some modified interval iterations which are the expansion of some work about interval computation completed by Alefeld and Herzberger ([1, 2]). Before the central discussion, some necessary knowledge about interval algorithm will be introduced.

A real interval $\mathbf{x}$ is a nonempty set of real numbers

$$\mathbf{x} = [\underline{x}, \overline{x}] = \{x \in \mathbb{R} : \underline{x} \leq x \leq \overline{x}\}$$

*Project 10471027 supported by the National Natural Science Foundation of China and Shanghai Education Committee.
Correspondence to: Yimin Wei (ymwei@fudan.edu.cn)

where $\underline{x}$ is called the infimum, and $\overline{x}$ is called the supremum.

The set of all interval over $\mathbb{R}$ is denoted by $\mathbb{IR}$

$$\mathbb{IR} = \{[\underline{x}, \overline{x}] : \underline{x}, \overline{x} \in \mathbb{R}, \underline{x} \leq \overline{x}\}.$$

The midpoint of $\mathbf{x}$ is given by

$$mid(\mathbf{x}) = \frac{1}{2}(\underline{x} + \overline{x})$$

and the radius of $\mathbf{x}$

$$rad(\mathbf{x}) = \frac{1}{2}(\overline{x} - \underline{x})$$

may also be used to define an interval $\mathbf{x}$. If an interval has zero radius it is called a point interval or thin interval, and contains a single point represented by

$$[x, x] \equiv x.$$

A thick interval has a radius greater than zero. If $\mathbf{x}$ is an interval

$$d(\mathbf{x}) = (\overline{x} - \underline{x}).$$

The absolute value of an interval $\mathbf{x}$ is defined as

$$\mid \mathbf{x} \mid = \max\{\mid x \mid, x \in \mathbf{x}\}.$$

An interval $\mathbf{x}$ is a subset of an interval $\mathbf{y}$, is denoted by $\mathbf{x} \subseteq \mathbf{y}$, if and only if $\underline{y} \leq \underline{x}$ and $\overline{y} \geq \overline{x}$.

An interval matrix is a matrix, whose elements are intervals. An interval matrix A is a subset of an interval matrix B, is denoted by $A \in B$, if and only if $A_{ij} \in B_{ij}, \quad for\ any\ i, j \in \mathbb{R}$, where A_{ij} is the elements of A, and B_{ij} is the elements of B. The midpoint matrix of an interval matrix A is

$$mid(A)_{ij} = mid(A_{ij})$$

where A_{ij} are the interval elements of A.

Given $\mathbf{x} = [\underline{x}, \overline{x}]$ and $\mathbf{y} = [\underline{y}, \overline{y}]$, the four elementary operations are defined by

$$\mathbf{x} + \mathbf{y} = [\underline{x} + \underline{y}, \overline{x} + \overline{y}]$$

$$\mathbf{x} - \mathbf{y} = [\underline{x} - \overline{y}, \overline{x} - \underline{y}]$$

$$\mathbf{x} \times \mathbf{y} = [\min\{\underline{x}\underline{y}, \underline{x}\overline{y}, \overline{x}\underline{y}, \overline{x}\overline{y}\}, \max\{\underline{x}\underline{y}, \underline{x}\overline{y}, \overline{x}\underline{y}, \overline{x}\overline{y}\}]$$

$$1/\mathbf{x} = [1/\overline{x}, 1/\underline{x}], \qquad if\ \ \underline{x} > 0\ \ or\ \ \overline{x} < 0$$

$$\mathbf{x} \div \mathbf{y} = \mathbf{x} \times 1/\mathbf{y}.$$

If $\mathbf{x} = [\underline{x}, \overline{x}]$ and $\mathbf{y} = [\underline{y}, \overline{y}]$ with $\underline{y} \leq 0 \leq \overline{y}$ and $\underline{y} < \overline{y}$, then the operation rules for division are as follows

$$\mathbf{x}/\mathbf{y} = \begin{cases} [\overline{x}/\underline{y}, \infty] & if\ \ \overline{x} \leq 0\ \ and\ \ \overline{y} = 0, \\ [-\infty, \overline{x}/\overline{y}] \cup [\overline{x}/\underline{y}, \infty] & if\ \ \overline{x} \leq 0\ \ and\ \ \underline{y} < 0 < \overline{y}, \\ [-\infty, \overline{x}/\overline{y}] & if\ \ \overline{x} \leq 0\ \ and\ \ \underline{y} = 0, \\ [-\infty, \infty] & if\ \ \underline{x} < 0 < \overline{x}, \\ [-\infty, \underline{x}/\underline{y}] & if\ \ \underline{x} \geq 0\ \ and\ \ \overline{y} = 0, \\ [-\infty, \underline{x}/\underline{y}] \cup [\underline{x}/\overline{y}, \infty] & if\ \ \underline{x} \geq 0\ \ and\ \ \underline{y} < 0 < \overline{y}, \\ [\underline{x}/\overline{y}, \infty] & if\ \ \underline{x} \geq 0\ \ and\ \ \underline{y} = 0. \end{cases}$$

The addition and subtraction of infinite or semi-infinite intervals are then defined by the follows:

$$\begin{aligned}
[\underline{x},\overline{x}]+[-\infty,\overline{y}] &= [-\infty,\overline{x}+\overline{y}],\\
[\underline{x},\overline{x}]+[\underline{y},\infty] &= [\underline{x}+\underline{y},\infty],\\
[\underline{x},\overline{x}]+[-\infty,\infty] &= [-\infty,\infty],\\
[\underline{x},\overline{x}]-[-\infty,\infty] &= [-\infty,\infty],\\
[\underline{x},\overline{x}]-[-\infty,\overline{y}] &= [\underline{x}-\overline{y},\infty],\\
[\underline{x},\overline{x}]-[\underline{y},\infty] &= [-\infty,\overline{x}-\underline{y}].
\end{aligned}$$

For addition and multiplication the associative and commutative laws hold. However

$$\mathbf{x}(\mathbf{y}+\mathbf{z}) \neq \mathbf{xy}+\mathbf{xz}$$

except for special cases, therefore the distributive law does not hold. Instead the sub-distributive law is

$$\mathbf{x}(\mathbf{y}+\mathbf{z}) \subseteq \mathbf{xy}+\mathbf{xz}.$$

Another example of one operation rule valid in real arithmetic that does not hold for interval computation. For example,

$$\mathbf{x}-\mathbf{x} \neq 0.$$

Let $\mathbf{x}=[2,3]$, which gives $\mathbf{x}-\mathbf{x}=[-1,1]\neq 0$. The reason is that an interval containing the difference between all possible results of two independent numbers lying within $\mathbf{x}$ is calculated, rather than the difference between two identical numbers.

An important result of interval analysis ia as follows:

Lemma 1.1. ([11]) If the function $f(z_1,z_2,\cdots,z_n)$ is an expression with a finite number $z_1,\cdots,z_n\in\mathbb{IR}$ and interval operations $(+,-,\times,\div)$, and if

$$x_1\subseteq z_1,\cdots,x_n\subseteq z_n$$

then

$$f(x_1,\cdots,x_n)\subseteq f(z_1,\cdots,z_n).$$

Definition 1.1. ([4, 12]) Let $X\in\mathbb{R}^{n\times m}$. If it satisfies the following conditions:

$$AXA=A,\quad XAX=X,\quad (AX)^T=AX,\quad (XA)^T=XA$$

where B^T is the transpose of B, then X is called the Moore-Penrose inverse of A and denoted by $A^\dagger$. We denote the orthogonal projector $P_{R(A)}=AA^\dagger$.

2 Classical iteration methods for computing $A^\dagger$

An iterative method for computing $A^\dagger$ is a set of instructions for generating a sequence $\{X_k : k=1,2,\cdots\}$ converges to $A^\dagger$. The instructions specify how to select the initial approximation X_0 and proceed from X_k to X_{k+1} for each k, and when to stop after having obtained a reasonable approximation of $A^\dagger$. The rate of convergence of such an iterative method is determined by the corresponding sequence of residuals

$$R_k=P_{R(A)}-AX_k,\qquad k=1,2,\cdots \tag{2.1}$$

which converges to 0 as $X_k\to A^\dagger$. An iteration method is said to be a $p\underline{th}-order$ method, for some $p>1$, if there is a positive constant c such that

$$\| R_{k+1} \| \leq c \| R_k \|,\qquad k=0,1,\cdots \tag{2.2}$$

for any multiplicative matrix norm.

In analogy with the nonsingular case, we consider iterative method of the type

$$X_{k+1} = X_k + C_k R_k, \qquad k = 0, 1, \cdots \tag{2.3}$$

where $\{C_k : k = 0, 1, \cdots\}$ is a suitable sequence, and X_0 is the initial approximation.

One objection to (2.3) as an iterative method for computing $A^\dagger$ is that (2.3) requires at each iteration the residual R_k, for which one needs the projection $P_{R(A)}$, whose computation is a task comparable to computing $A^\dagger$. This difficulty will be overcome here by choosing the sequence $\{C_k\}$ in (2.3) to satisfy

$$C_k = C_k P_{R(A)}, \qquad k = 0, 1, \cdots \tag{2.4}$$

For such a choice we have

$$\begin{aligned} C_k R_k &= C_k(P_{R(A)} - AX_k) \\ &= C_k(I - AX_k) \end{aligned} \tag{2.5}$$

and (2.3) can therefore be rewritten as

$$X_{k+1} = X_k + C_k T_k, \qquad k = 0, 1, \cdots \tag{2.6}$$

where

$$T_k = I - AX_k, \qquad k = 0, 1, \cdots \tag{2.7}$$

Lemma 2.1. ([4]) Let $0 \neq A \in \mathbb{R}^{m \times n}$ and let the initial approximation X_0 and its residual R_0 satisfy

$$X_0 \in R(A^T, A^T) = \{X \mid X = A^T B A^T, B \in \mathbb{R}^{m \times n}\} \tag{2.8}$$

and the spectral radius

$$\rho(R_0) < 1 \tag{2.9}$$

then for any integer $p \geq 2$, the sequence

$$\begin{aligned} X_{k+1} &= X_k(I + T_k + T_K^2 + \cdots + T_k^{p-1}) \\ &= X_k\left[I + (I - AX_k) + (I - AX_k)^2 + \cdots + (I - AX_k)^{p-1}\right] \end{aligned} \tag{2.10}$$

converge to $A^\dagger$ as $k \to \infty$, and the corresponding sequence of residual satisfies

$$\| R_{k+1} \| \leq \| R_k \|^p, \qquad k = 0, 1, \cdots \tag{2.11}$$

3 Interval iterative methods for computing $A^\dagger$

In this section, we will present the interval iterative methods for computing $A^\dagger$.

Theorem 3.1. Let $0 \neq A \in \mathbb{R}^{m \times n}$ $(m \leq n)$ and the rank of A is m. Let the initial approximation X_0 and its residual $R_0 = I - Am(X^{(0)})$ satisfy

$$m(X^{(0)}) \in R(A^T, A^T) = \{X \mid X = A^T B A^T, B \in \mathbb{R}^{m \times n}\} \tag{3.1}$$

and

$$\rho[I - Am(X^{(0)})] < 1 \tag{3.2}$$

where $m(X^{(0)})$ denotes the midpoint matrix of $X^{(0)}$. The sequence

$$X^{(k+1)} = m(X^{(k)}) \sum_{v=0}^{r-2} [I - Am(X^{(k)})]^v + X^{(k)}[I - Am(X^{(k)})]^{r-1}, k \geq 0 \tag{3.3}$$

converge to $A^\dagger$ as $k \to \infty$. I is the identity matrix and $X^{(k)} \in \mathbb{IR}^{n \times m}$, $(k = 0, 1, \cdots)$. $r > 1$ is a natural constant. It follows from Theorem 3.1 that

$$\|d(X^{(k+1)})\| \leq \gamma \|d(X^{(k)})\|^r, \gamma \geq 0 \tag{3.4}$$

where $\|.\|$ denotes the matrix norm, $d(.)$ denotes the distance matrix.

Proof: The midpoint matrix satisfies

$$m(X+Y) = m(X) + m(Y) \tag{3.5}$$

$$m(X-Y) = m(X) - m(Y) \tag{3.6}$$

where $X, Y \in \mathbb{IR}^{m\times n}$

$$m(BX) = Bm(X), \quad m(XC) = m(X)C \tag{3.7}$$

where $B \in \mathbb{R}^{m\times n}$, $X \in \mathbb{IR}^{n\times p}$ and $C \in \mathbb{R}^{p\times q}$.

If $B \in \mathbb{R}^{m\times n}$, it satisfies

$$m(B) = B. \tag{3.8}$$

From (3.3),(3.5),(3.6),(3.7) and (3.8), it is obvious that the sequence $\{X^{(k)}\}$ satisfies

$$m(X^{(k+1)}) = m(X^{(k)})\sum_{v=0}^{r-1}[I - Am(X^{(k)})]^v \tag{3.9}$$

where $m(X^{(k)}) \in \mathbb{R}^{n\times m}$, from (3.9) and Theorem 2.1, we can get

$$m(X^{(\infty)}) = A^\dagger. \tag{3.10}$$

Finally we will show that

$$X^{(k)} \to A^\dagger \iff m(X^{(k)}) \to A^\dagger. \tag{3.11}$$

It is obvious that the interval matrix sequence $\{X^{(k)}\}$ satisfies

$$d(X^{(k+1)}) \le d(X^{(k)}) \mid [I - Am(X^{(k)})]^{r-1} \mid \tag{3.12}$$

where $(d(X^{(k)}))_{ij} = d(X_{ij}^{(k)})$

$$\mid (I - Am(X^{(k)}))^{r-1} \mid_{ij} = \mid (I - Am(X^{(k)}))_{ij}^{r-1} \mid. \tag{3.13}$$

If $k \to \infty$, we can get

$$\lim_{k\to\infty} d(X^{(k)}) = 0. \tag{3.14}$$

Then from (3.10) and (3.14), it is obvious that

$$\lim_{k\to\infty} m(X^{(k)}) = A^\dagger \iff \lim_{k\to\infty} X^{(k)} = A^\dagger. \tag{3.15}$$

Since

$$\begin{aligned} d(X^{(k+1)}) &\le d(X^{(k)})|[I - Am(X^{(k)})]^{r-1}| \\ &= d(X^{(k)})|[AA^\dagger - Am(X^{(k)})]^{r-1}| \\ &\le d(X^{(k)})[|A||A^\dagger - m(X^{(k)})|]^{r-1} \\ &\le d(X^{(k)})2^{-(r-1)}[|A|d(X^{(k)})]^{r-1} \end{aligned} \tag{3.16}$$

we can get a matrix norm $\|.\|'$, and it satisfies

$$\|d(X^{(k+1)})\|' \le 2^{-(r-1)}\|A\|'^{r-1}\|d(X^{(k)})\|'^{r} \tag{3.17}$$

all the matrix norms satisfy

$$\gamma_1\|B\| \le \|B\|' \le \gamma_2\|B\|, \qquad \gamma_1 > 0, \gamma_2 > 0 \tag{3.18}$$

from (3.19),(3.20) and (3.21), we can get

$$\|d(X^{(k+1)})\|\gamma_1 \le 2^{-(r-1)}\gamma_2^{r-1}\|A\|^{r-1}\gamma_2^{r}\|d(X^{(k)})\|^{r} \tag{3.19}$$

so (3.4) is given. □

Similarly, we have another conclusion.

Theorem 3.2. Let $0 \neq A \in \mathbb{R}^{m\times n}$ $(m \geq n)$ and the rank is n. The initial approximation $X^{(0)}$ and its residual $R_0 = I - m(X^{(k)})A$ satisfy

$$m(X^{(0)}) \in R(A^T, A^T)$$

and

$$\rho[I - m(X^{(0)})A] < 1.$$

The sequence $\{X^{(k)}\}$

$$X^{(k+1)} = \left[\sum_{v=0}^{r-2}[I - m(X^{(k)})A]^v\right] m(X^{(k)}) + [I - m(X^{(k)})A]^{r-1}X^{(k)}$$

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

The proof of Theorem 3.2 is similar to that of Theorem 3.1.

As a matter of fact, from Theorem 3.1, we can know that

$$A^\dagger \in X^{(k)} \tag{3.20}$$

if

$$A^\dagger \in X^{(0)} \tag{3.21}$$

because

$$\begin{aligned} A^\dagger &= m(X^{(k-1)})\sum_{v=0}^{r-2}[I - Am(X^{(k-1)})]^v + A^\dagger[I - Am(X^{(k-1)})]^{r-1} \\ &\in m(X^{(k-1)})\sum_{v=0}^{r-2}[I - Am(X^{(k-1)})]^v + X^{(k-1)}[I - Am(X^{(k-1)})]^{r-1} \quad (3.22) \end{aligned}$$

from induction we can prove it.

All the proof is based on Theorem 3.1, Theorem 3.2 has similar conclusions.

From the proof we can see that the convergence statement is also valid even if $X^{(0)}$ is an arbitrary interval matrix not necessarily containing $A^\dagger$. In this case however, the iterates do not necessarily containing $A^\dagger$. We note that criterion (3.2) depending only on the midpoint matrix $m(X^{(0)})$ of the given inclusion matrix $X^{(0)}$. The width $d(X^{(0)})$ may on the other hand be arbitrary. This means that if one has a suitable approximation $m(X^{(0)})$ for $A^\dagger$ that satisfies $\rho[I - Am(X^{(0)})] < 1$, then using certain norm estimates one is always able to produce an interval matrix $X^{(0)}$ such that $A^\dagger \in X^{(0)}$. The sequence of iterates generated according to (3.3), then converges to $A^\dagger$ by Theorem 3.1.

Since the sequence of iterates from (3.3) always contains $A^\dagger$ according to (3.16), it seems natural to form the intersection of the new iterate with the previous iterate and then to continue the iteration with this new potentially improved iteration.

$$\begin{cases} Y^{(k+1)} = m(X^{(k)})\sum_{v=0}^{r-2}[I - Am(X^{(k)})]^v + X^{(k)}[I - Am(X^{(k)})]^{r-1}, k \geq 0, \\ X^{(k+1)} = Y^{(k+1)} \cap X^{(k)}. \end{cases} \tag{3.23}$$

Using this iteration procedure one obtains a monotonic sequence

$$X^{(0)} \supseteq X^{(1)} \supseteq X^{(2)} \cdots$$

of inclusion for $A^\dagger$. The following numerical example does show, however, that the convergence criterion (3.2) is not sufficient for convergence in general.

Theorem 3.3. Let $0 \neq A \in \mathbb{R}^{m\times n}$ $(m \leq n)$ and the rank of A is m. Let the initial approximation X_0 be an $n \times m$ interval matrix for which $A^{\dagger} \in X_0$ and its residual $R_0 = I - Am(X^{(0)})$ satisfy

$$m(X^{(0)}) \in R(A^T, A^T) = \{X \mid X = A^T B A^T, B \in \mathbb{R}^{m\times n}\} \tag{3.24}$$

and

$$\rho(|I - AX|) < 1 \ for \ all \ X \in X^{(0)} \tag{3.25}$$

where $m(X^{(0)})$ denotes the midpoint matrix of $X^{(0)}$. The sequence (3.23) converges to $A^{\dagger}$ as $k \to \infty$. I is the identity matrix and $X^{(k)} \in \mathbb{IR}^{n\times m}$ $(k = 0, 1, \cdots)$. $r > 1$ is a natural constant. The sequence $\{d(X^{(k)})\}$ may be bounded as follows, by using a matrix norm:

$$\|d(X^{(k+1)})\| \leq \gamma' \|d(X^{(k)})\|^r. \tag{3.26}$$

Proof: The sequence

$$X^{(0)} \supseteq X^{(1)} \supseteq X^{(2)} \cdots$$

of iterates always converges to an interval matrix X ([9]). We now show that under the assumptions of the theorem we must necessarily have $d(X) = 0$. We therefore define

$$Y = m(X) \sum_{v=0}^{r-2} [I - Am(X)]^v + X[I - Am(X)]^{r-1} \tag{3.27}$$

and obtain $X \subseteq Y$. Then we obtain $d(X) \leq d(Y)$. For $d(X)$ we get from (3.23) that

$$d(X)|I - Am(X)|^{r-1} \geq d(X)|[I - Am(X)]^{r-1}| = d(Y) \geq d(X) \tag{3.28}$$

which implies that

$$d(X)(I - |I - Am(X)|^{r-1}) \leq 0. \tag{3.29}$$

The assumption $\rho(|I - Am(X)|) < 1$ implies the existence of $[(I - |I - Am(X)|^{r-1})]^{-1}$. This inverse is also nonnegative. From (3.29), it follows that $d(X) \leq 0$.

As for the proof of (3.4) one first shows that the inequality $\|d(Y^{(k+1)})\|' \leq \gamma \|d(X^k)\|'^r$ holds for a monotonic and consistent matrix norm $\|.\|'$. From this it follows that the inequality

$$\|d(X^{(k+1)})\|' \leq \|d(Y^{(k+1)})\|' \leq \gamma \|d(X^k)\|'^r$$

is valid since $X^{(k+1)} \subseteq Y^{(k+1)}$ and the monotonic norm $\|.\|'$. Analogous to the proof of (3.18), we can use the equivalence theorem to prove the final statement.

In fact from the proof of (3.16), it follows that each iterate $X^{(k)}$ from (3.23) contains $A^{\dagger}$ if $X^{(0)}$ contains $A^{\dagger}$. The convergence criterion (3.25) depends on the width of the inclusion matrix $X^{(0)}$ for $A^{\dagger}$ as opposed to the criterion (3.2). Formulas may easily be given for this dependence. If, for example, an interval matrix $X^{(0)}$ satisfies the inequality $\|I - Am(X^{(0)})\| < 1$, for a monotonic and multiplicative norm $\|.\|$, then we have that

$$\|d(X^{(0)})\| < 2(1 - \|I - Am(X^{(0)})\|)/\|A\| \tag{3.30}$$

is a sufficient criterion for the statement that $\|I - AX\| < 1$ for all $X \in X^{(0)}$([2]). If (3.30) is now satisfied after a certain iteration step, then the iteration may be continued as the procedure (3.23). □

The computation of Moore-Penrose inverse of full column rank matrix also has the similar conclusions.

4 Numerical examples

From Theorem 3.1 we choose parameter $r = 2$. The classical iteration is the Newton method. We compute these examples by Intlab which is a toolbox of Matlab ([10]).

Example 4.1. We generate a random 50×60 matrix. We use classical Newton iteration and improve interval iteration to compute $A^\dagger$ to compare two algorithms. The initial approximation $X^{(0)} = \alpha * A^T$, and α is a constant interval. From (3.2) we know it will satisfy the convergence conditions if we specify a good α. In practice, we can use matrix norm to estimate α.

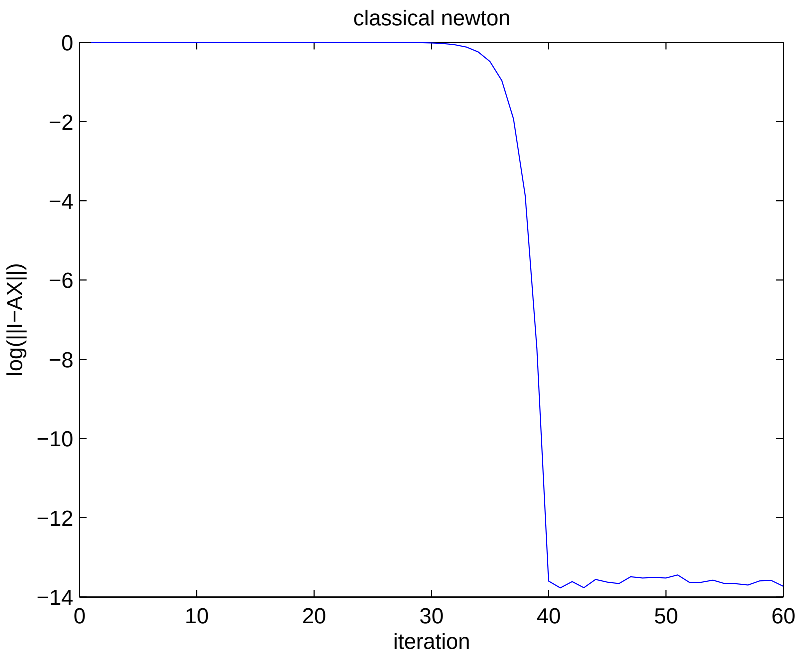

Figure 1: Classical Newton iteration

From these figures we can see the convergence of the upper and lower boundary of the interval iteration directly.

Figure 4 describes the difference of the accuracy between two methods, blue line represents the classical method result, red and green lines represent the upper boundary and lower boundary of the interval method. From the figure we can see that the lines of interval method is much smoother than that of classical method and the result of interval method has higher accuracy.

Example 4.2. We generate a random 500×600 matrix. We will compute its Moore-Penrose inverse by two methods to compare them. Other conditions are the same as Example 4.1.

Blue line represents the classical method result, red and green lines represent the upper boundary and lower boundary of the interval method.

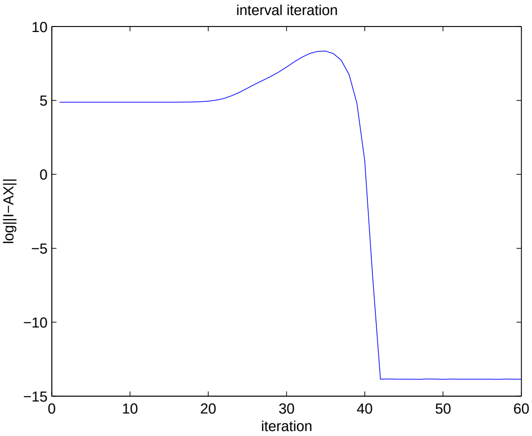

Figure 2: The upper boundary result of interval iteration

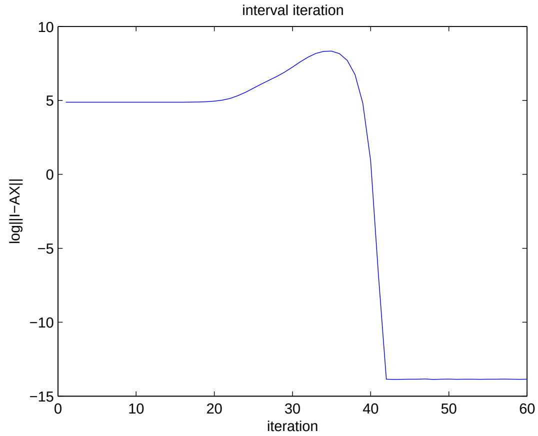

Figure 3: The lower boundary result of interval iteration

5 Concluding Remarks

In this paper we present the interval method for computing Moore-Penrose inverse of full row (or column) rank matrix. Compared with the classical iteration, the interval method can give higher accuracy results, but the computation time is longer. How to overcome these problems will be the future research topic.

References

[1] G. Alefeld, On the convergence of the higher order versions of D. J. Evans' implicit matrix inversion process, Z. Angew. Math. Mech., 64(1984) 413–418.

[2] G. Alefeld and J. Herzberger, *Introduction to Interval Computations*, Transl. from the German by Jon Rokne. (English) Computer Science and Applied Mathematics. New York-London etc.: Academic Press. XVIII, 333, 1983.

[3] G. Alefeld and G. Mayer, *Interval analysis: Theory and applications*, J. Comput. Appl. Math., 121 (2000) 421–464.

[4] A. Ben-Israel and T.N.E. Greville, *Generalized Inverses: Theory and Applications*, Wiley, New York, 1974; 2nd edition, Springer, New York, 2003.

[5] M. Gulliksson, X. Jin and Y. Wei, Perturbation bounds for constrained and weighted least squares problems, Linear Algebra Appl., 349 (2002) 221–232.

[6] M. Gulliksson, P. Wedin and Y. Wei, Perturbation identities for regularized Tikhonov inverses and weighted pseudoinverses, BIT, 40 (2000) 513–523.

[7] J. Ji and Y. Wei, A note on the sensitivity of the solution of the weighted linear least squares problem. Appl. Math. Comput., 145 (2003) 481–485.

[8] X. Jin and Y. Wei, *Numerical Linear Algebra and its Applications*, Science Press, Beijing, 2004.

[9] R.E. Moore, *Methods and Applications of Interval Analysis*, SIAM studies in Applied Mathematics, Philadelphia, 1979.

[10] S. Rump, INTLAB – INTerval LABoratory. (English) Csendes, Tibor (ed.), Developments in reliable computing. SCAN-98 conference, 8th international symposium on Scientific computing, computer arithmetic and validated numerics. Budapest, Hungary, September 22-25, 1998. Dordrecht: Kluwer Academic Publishers. 77–104 (1999).

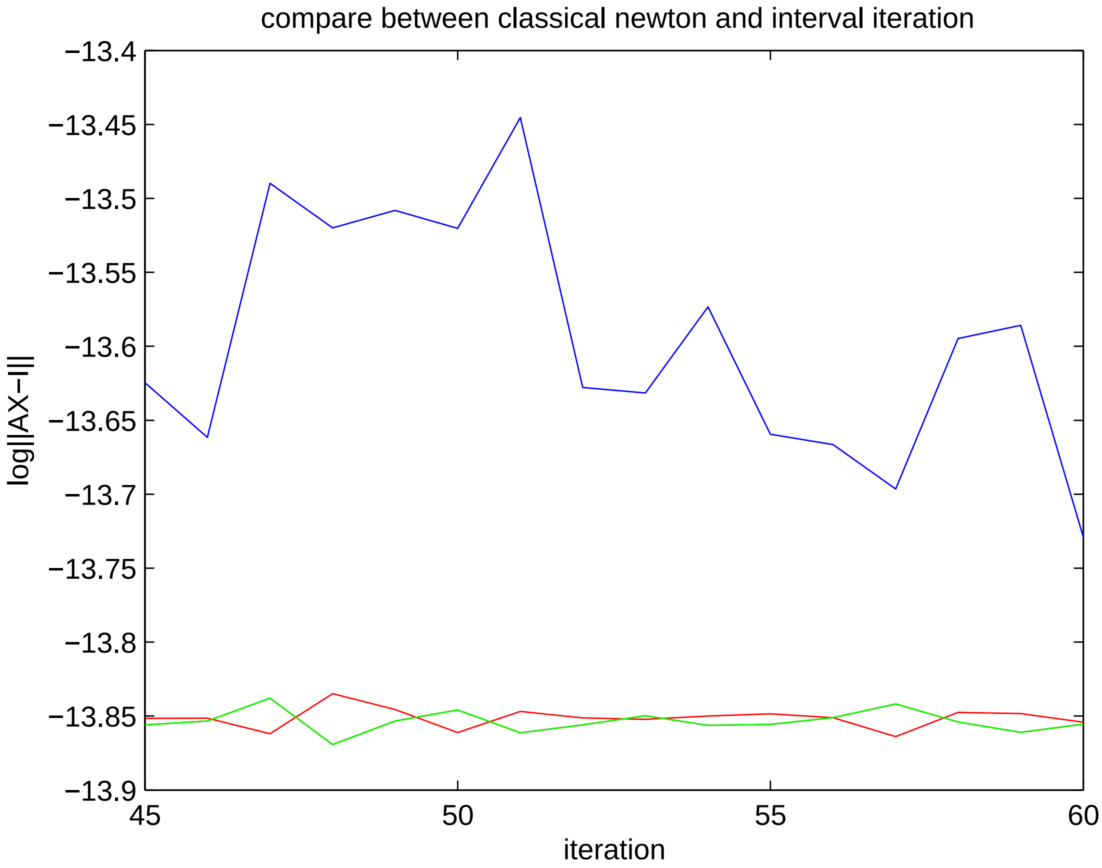

Figure 4: Result comparison

[11] S. Rump, Fast and parallel interval arithmetic, BIT, 39 (1999) 534–554.

[12] G. Wang, Y. Wei and S. Qiao, *Generalized Inverses: Theory and Computations*, Science Press, Beijing, 2004.

[13] Y. Wei, Recurrent neural networks for computing weighted Moore-Penrose inverse, Appl. Math. Comput., 116 (2000) 279–287.

[14] Y. Wei, The weighted Moore-Penrose inverse of modified matrices, Appl. Math. Comput., 122 (2001) 1–13.

[15] Y. Wei, J. Cai, and M.K. Ng, Computing Moore-Penrose inverses of Toeplitz matrices by Newton's iteration, Math. Comput. Modelling, 40 (2004) 181–191.

[16] Y. Wei and M.K. Ng, Weighted Tikhonov filter matrices for ill-posed problems, Appl. Math. Comput. 149 (2004) 411–422.

[17] Y. Wei and D. Wang, Condition numbers and perturbation of the weighted Moore-Penrose inverse and weighted linear least squares problem, Appl. Math. Comput., 145 (2003) 45–58.

[18] Y. Wei and G. Wang, PCR algorithm for parallel computing minimum-norm (T) least-squares (S) solution of inconsistent linear equations, Appl. Math. Comput., 133 (2002) 547–557.

[19] Y. Wei and H. Wu, Expression for the perturbation of the weighted Moore-Penrose inverse, Comput. Math. Appl., 39 (2000) 13–18.

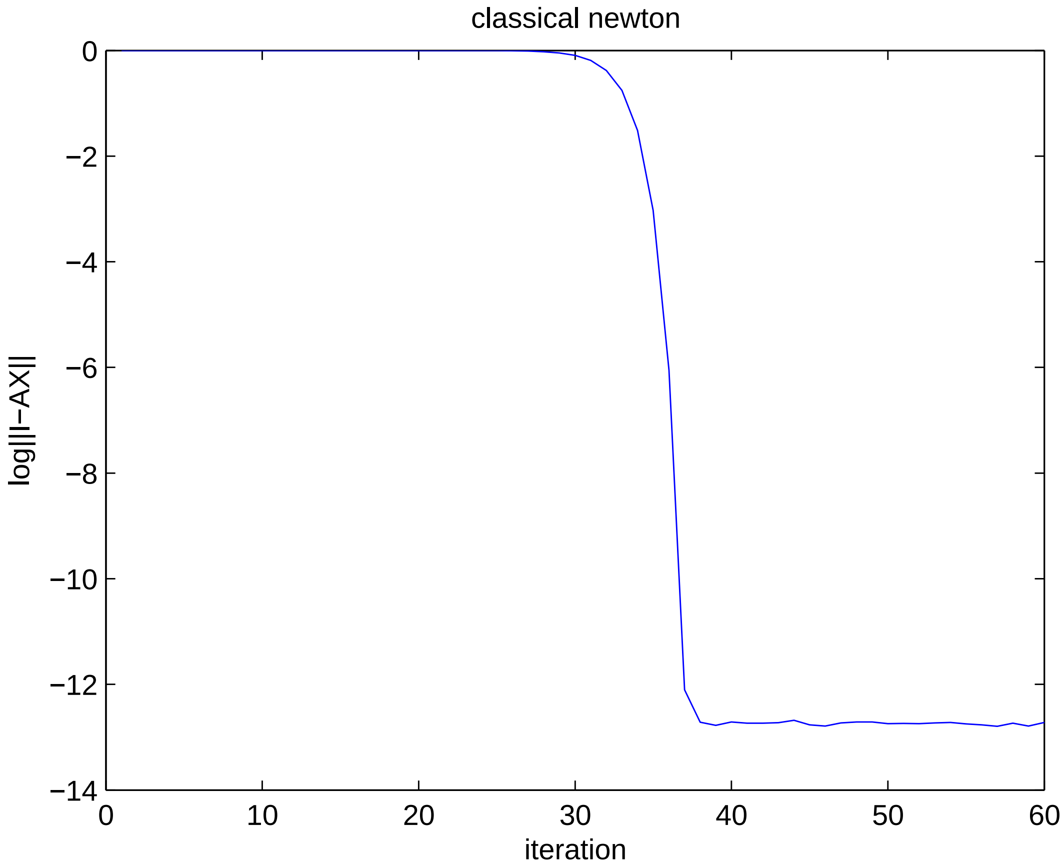

Figure 5: Classical Newton iteration

[20] Y. Wei, and H. Wu, The representation and approximation for the weighted Moore-Penrose inverse, Appl. Math. Comput., 121 (2001) 17–28.

[21] Y. Wei, H. Wu and J. Wei, Successive matrix squaring algorithm for parallel computing the weighted generalized inverse A^{+}_{MN}, Appl. Math. Comput., 116 (2000) 289–296.

[22] Y. Wei, X. Yu and R. Zhang, Preconditioned conjugate gradient method and generalized successive over relaxation method for the weighted least squares problems, Int. J. Comput. Math., 81 (2004), 203-214.

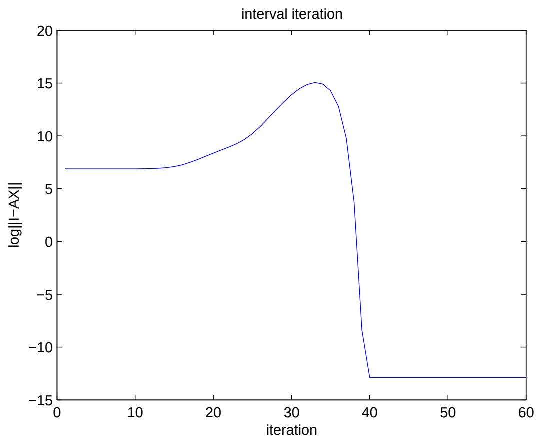

Figure 6: The upper boundary result of interval iteration

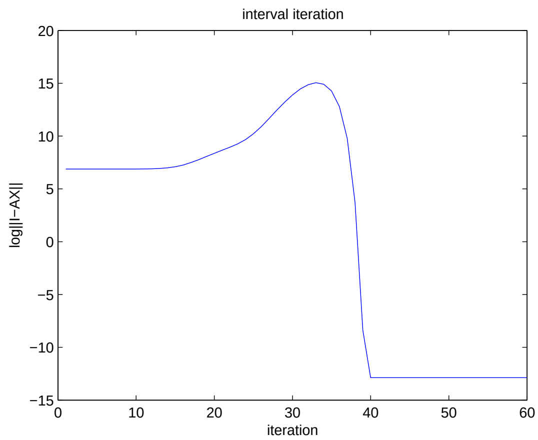

Figure 7: The lower boundary result of interval iteration

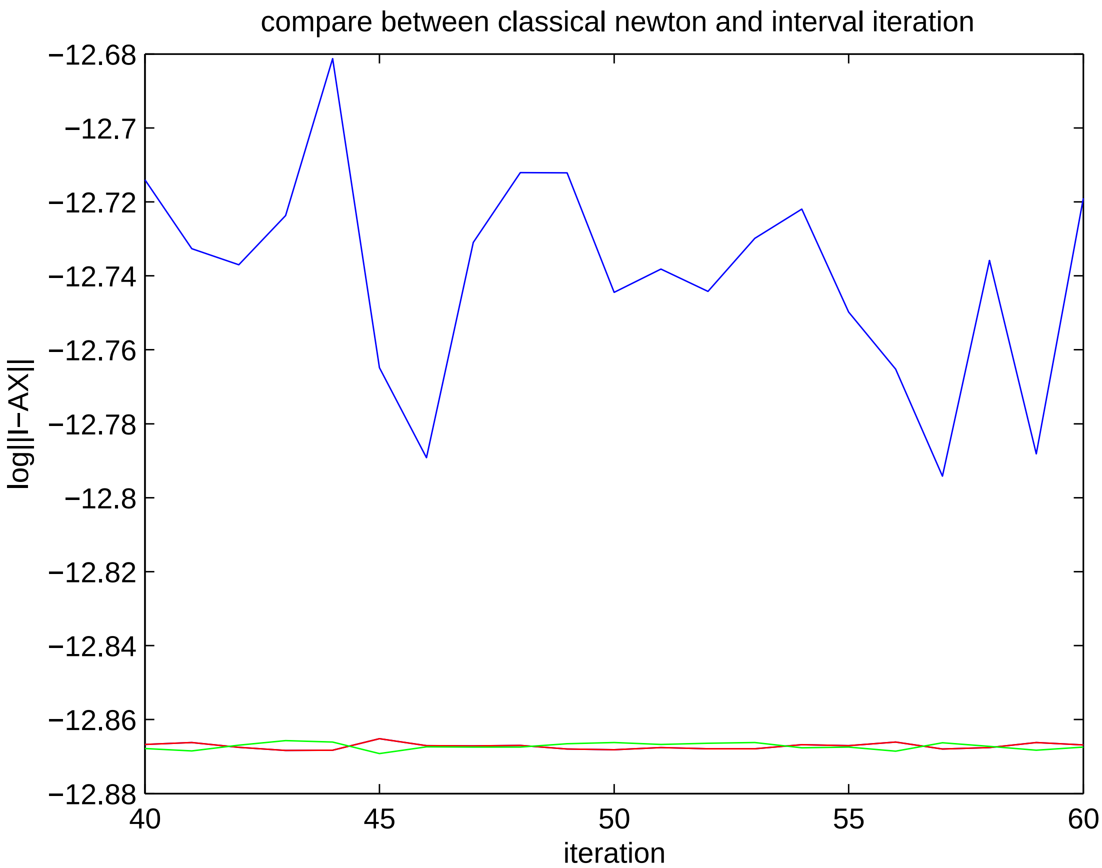

Figure 8: Result comparison