

# The method of interval entropy function for ill-conditioned system of linear equations

H. WANG

*Science of College, China University of Mining and Technology, Xuzhou, China*  
e-mail: hjwcumt@163.com

In this paper, a new algorithm for solving the ill-conditioned system of linear equations is given. The algorithm is based on the maximum entropy principle and the interval analysis theory. Moreover, it is proved that this algorithm is global convergent under mild conditions. The theoretic analysis and numerical results indicate that the algorithm is effective and reliable.

*Keywords:* numerical algebra, ill-conditioned system of linear equations, entropy function, interval algorithm.

## Introduction

As is known, it is necessary to seek the solution to the ill-conditioned system of linear equations in many fields, in such as geophysical indirect problem, graphic information processing, de-convolution, parameter estimation of a model, etc. Although some iterative methods can yield many good results theoretically, the validity of iteration and its condition of convergence are closely related. In many practical problems, when we seek the solution of ill-conditioned system of linear equations by iteration, the divergence and slow rate of convergence usually appear in actual calculation. So the usage of iterative method is subjected to certain limit in fact. In this paper, we introduce a new tool and construct a new algorithm, expecting to obtain new results.

We consider the following system of linear equations

$$Ax = b, \quad (0.1)$$

where  $A = \{a_{in}\}_{n \times n} \in \mathbb{R}^n$  is nonsingular and  $Cond(A) \gg 1$ ,  $x = (x_1, x_2, \dots, x_n)^T \in \mathbf{x}^{(0)} \subset \mathbb{R}^n$ ,  $b = (b_1, b_2, \dots, b_n)^T \in \mathbb{R}^n$ . Suppose, there exists  $x^* \in \mathbf{x}^{(0)}$  and  $x^*$  is an isolated solution of  $Ax = b$ .

To solve system (0.1) with the idea of unconstrained optimization, usually we converted problem of solving (0.1) into the following optimization problem:

$$\min_{x \in \mathbf{x}^{(0)}} f(x) = \|Ax - b\|_p, \quad (0.2)$$

where  $\|\cdot\|_p$  denotes the  $p$  norm of a vector space. Let  $p = \infty$ , then problem (0.2) is simplified as follows:

$$\min_{x \in \mathbf{x}^{(0)}} \max_{1 \leq i \leq n} \{|f_i(x)|\}, \quad (0.3)$$

where  $f_i(x) = \sum_{j=1}^n a_{ij}x_j - b_i$  ( $i = 1, 2, \dots, n$ ). Obviously, problem (0.3) is a nondifferentiable optimization problem.

Several interval methods are known that provide guaranteed bounds on the solutions of the unconstrained optimization problem (see, e. g., [1, 2]). Interval methods for global optimization are based on the branch-and-bound principle. These methods subdivide the search interval in subintervals (branches) and use bounds for the objective function to eliminate subintervals which cannot contain a global minimizer of given objective function. These interval methods do not necessarily provide the exact range of the function over the subintervals. Moreover, interval arithmetics is able to provide rigorous bounds on the range of the function over the subintervals. For a more thorough discussion (see [3–7]). In this paper, we propose a new idea of solving ill-conditioned system of linear equations that requires the following two steps.

First, we use the entropy function to reduce the minimax problem (0.3) to an approximate unconstrained optimization problem with a smooth objective function.

Second, we apply known interval techniques to solve the resulting unconstrained optimization problem.

As a particular case of this idea, we propose a new interval algorithm for solving this problem. The structure of the paper is as follows: In Section 1, we will describe entropy functions and their relationship with interval methods. In Section 2, the proposed algorithm is described, and the sufficient condition which guarantees that a given multidimensional interval contains exact solution of system (0.1) is given. In conclusions, we will show the numerical analysis results. Finally, the last section contains the conclusions and an outline of the directions for future research.

**Denotations.** The set of all interval vectors on  $\mathbb{R}^n$  is denoted by  $\mathbb{IR}^n$ , the set of all interval vectors on  $\mathbf{x}^{(0)}$  is denoted by  $\mathbb{I}(\mathbf{x}^{(0)})$ ,  $\forall \mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_n)^T \in \mathbb{IR}^n$ , let  $\mathbf{x}_i = [\underline{\mathbf{x}}_i, \overline{\mathbf{x}}_i]$ . Denote the midpoint, width, radius, and absolute value of  $\mathbf{x}$  by  $\text{mid}(\mathbf{x})$ ,  $\text{wid}(\mathbf{x})$ ,  $\text{rad}(\mathbf{x})$ , and  $|\mathbf{x}|$  respectively.

**Definition 0.1.** *Interval function  $\mathbf{f} : \mathbb{IR}^n \rightarrow \mathbb{IR}$  is called the interval extension of a real function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ , if*

- 1)  $\mathbf{f}(x) = f(x)$  for any  $x \in \mathbb{R}^n$  from the domain of  $f$ ;
- 2)  $\mathbf{f}(\mathbf{x})$  is monotonic by inclusion, i. e.,  $\mathbf{x}, \mathbf{y} \in \mathbb{IR}^n, \mathbf{x} \subseteq \mathbf{y} \implies \mathbf{f}(\mathbf{x}) \subseteq \mathbf{f}(\mathbf{y})$ .

Therefore, if  $\mathbf{f}(\mathbf{x})$  is an interval extension of the function  $f(x)$ , then always

$$\{f(x)|x \in \mathbf{x}\} \subseteq \mathbf{f}(\mathbf{x})$$

and we get an outer (by superset) estimate of the range of  $f$  over the bound  $\mathbf{x} \in \mathbb{IR}^n$ .

**Definition 0.2.** *An inclusion function  $\mathbf{f}$  of  $f : \mathbf{x} \rightarrow \mathbb{R}$  is called of convergence order  $\alpha > 0$  if*

$$\text{wid}(\mathbf{f}(\mathbf{x})) - \text{wid}(f(\mathbf{x})) = O(\text{wid}(\mathbf{x})^\alpha),$$

that is, if there exists a constant  $c \geq 0$  such that

$$\text{wid}(\mathbf{f}(\mathbf{x})) - \text{wid}(f(\mathbf{x})) \leq c(\text{wid}(\mathbf{x}))^\alpha,$$

where  $f(\mathbf{x}) = \{f(x)|x \in \mathbf{x}\}$ .

The further information about interval mathematics can be found in [1, 3].

## 1. Main Theory

If a problem is nondifferentiable, it is hard to solve it by traditional methods including derivative. Usually we handle it with a differentiable function that approximates the objective function. A maximum entropy function of uniform approximation to objective function is investigated in [4, 8–10].

### 1.1. Maximum Entropy Function

The maximum entropy function originated in the work of L. Boltzmann [11] who studied the relation between entropy and probability in physical systems. Subsequently, entropy concepts and ideas have been introduced in physics, chemistry, meteorology, information science, etc. The maximum entropy function, which was introduced by Jaynes [12] has proved to be useful for obtaining convergent upper bounds on nonsmooth objective functions.

**Definition 1.1.** Let  $p$  be a positive integer and let  $\varphi_i: D \rightarrow R^1$ ,  $i = 1, \dots, k$ , be given continuously differentiable functions, and let  $\varphi(x) = \max_{1 \leq i \leq k} \varphi_i(x)$ . The maximum entropy function  $\varphi_p: D \rightarrow R^1$  of  $\varphi_1, \dots, \varphi_k$  is defined by

$$\varphi_p(x) = \frac{1}{p} \ln \left\{ \sum_{i=1}^k \exp(p\varphi_i(x)) \right\}. \quad (1.1)$$

Obviously,  $\varphi_p(x)$  is also continuously differentiable on  $D$ .

**Lemma 1.1** [12]. Given  $\epsilon > 0$ ,  $\exists \bar{p} > 0$ , such that  $\forall p > \bar{p}$  and  $\forall x \in D$

$$|\varphi_p(x) - \varphi(x)| < \epsilon. \quad (1.2)$$

Also, if  $1 \leq p \leq q$ , then

$$\varphi_q(x) \leq \varphi_p(x). \quad (1.3)$$

**Proof.** Let  $v_i(x) = \exp(\varphi_i(x))$  ( $i = 1, \dots, k$ ) and let  $v(x) = (v_i(x))_{k \times 1}$ . For each  $x \in D$  ( $\forall p \geq 1$ )

$$\|v(x)\|_p = \left( \sum_{i=1}^k v_i(x)^p \right)^{1/p} = \left( \sum_{i=1}^k \exp(p\varphi_i(x)) \right)^{1/p}.$$

So

$$\varphi_p(x) = \ln(\|v(x)\|_p).$$

Now

$$\|v(x)\|_\infty = \lim_{p \rightarrow \infty} \|v(x)\|_p = \max_{1 \leq i \leq k} \{\exp(\varphi_i(x))\}.$$

So

$$\begin{aligned} \ln(\|v(x)\|_\infty) &= \max_{1 \leq i \leq k} \varphi_i(x), \\ \lim_{p \rightarrow \infty} \varphi_p(x) &= \max_{1 \leq i \leq k} \varphi_i(x). \end{aligned}$$

This proves (1.2). Finally, according to Jensen's inequality

$$(1 \leq p \leq q) \Rightarrow (\|v(x)\|_q \leq \|v(x)\|_p),$$

we obtain  $\ln(\|v(x)\|_q) \leq \ln\|v(x)\|_p$ , whence  $\varphi_q(x) \leq \varphi_p(x)$ . So, (1.3) is proved.

**Lemma 1.2.** *For all  $x \in D$ , we have*

$$\varphi(x) \leq \varphi_p(x) \leq \varphi(x) + \frac{\ln k}{p}. \quad (1.4)$$

**Proof.** It follows

$$\varphi_p(x) - \varphi(x) = \frac{1}{p} \ln \left\{ \sum_{i=1}^k \exp(p\varphi_i(x)) \right\} - \varphi(x) = \frac{1}{p} \ln \left\{ \sum_{i=1}^k \exp(p(\varphi_i(x) - \varphi(x))) \right\}.$$

By definition of  $\varphi(x)$ ,  $\varphi_i(x) - \varphi(x) \leq 0$  ( $i = 1, \dots, k$ ) and there is at least one  $j \in \{1, \dots, k\}$  such that  $\varphi_j(x) - \varphi(x) = 0$ . Therefore,  $1 \leq \sum_{i=1}^k \exp(p(\varphi_i(x) - \varphi(x))) \leq k$ , and we have

$$0 \leq \frac{1}{p} \ln \left\{ \sum_{i=1}^k \exp(p(\varphi_i(x) - \varphi(x))) \right\} \leq \frac{1}{p} \ln(k).$$

So, we complete the proof.

From Lemma 1.1 and Lemma 1.2, we conclude that

$$\lim_{p \rightarrow +\infty} \varphi_p(x) = \varphi(x).$$

Namely,  $\varphi_p(x)$  is uniform approximation of  $\max_{1 \leq i \leq k} \varphi_i(x)$ . Obviously, it is reasonable and feasible that we substitute a differentiable function  $\varphi_p(x)$  for  $\varphi(x)$ . Let  $g(x) = \max_{1 \leq i \leq n} \{|f_i(x)|\} = \max_{1 \leq i \leq n} \{f_i(x), -f_i(x)\}$ , with the idea of maximum entropy function, we can transform (0.3) into the approximative unconstrained differentiable optimization problem

$$\min_{x \in \mathbf{X}^{(0)}} f_p(x), \quad (1.5)$$

where  $f_p(x) = \frac{1}{p} \ln \left\{ \sum_{i=1}^n (\exp(pf_i(x)) + \exp(-pf_i(x))) \right\}$ ,  $p \geq 1$  is entropy factor.

**Theorem 1.1.** *Given  $\epsilon > 0$ ,  $\exists p_k > 0$ , such that  $p = p_k > \frac{2n}{\epsilon}$ ,*

$$0 \leq f_p(\hat{x}) - g(x^*) \leq \epsilon, \quad (1.6)$$

where  $\hat{x}$  is a solution of problem (1.5), and  $x^*$  is a solution of problem (0.3).

**Proof.** Because  $x^*$  is a solution of problem (0.3), we have

$$f_p(\hat{x}) - g(x^*) \geq f_p(x^*) - g(x^*) \geq 0$$

and  $\hat{x}$  is a solution of problem (1.5), so

$$g(x^*) \leq g(\hat{x}) \leq f_p(\hat{x}) \leq f_p(x^*).$$

From (1.4), we can prove

$$f_p(\hat{x}) - g(x^*) \leq f_p(x^*) - g(x^*) \leq \frac{\ln 2n}{p}.$$

Let  $p = p_k > \frac{\ln 2n}{\epsilon}$ , then

$$0 \leq f_p(\hat{x}) - g(x^*) \leq \epsilon.$$

So, we complete the proof.

Since we are interested in the minimum of the function  $g(x)$ , we will use the following corollary of Theorem 1.1.

**Corollary 1.1.** *We have*

$$\inf_{x \in \mathbf{x}^{(0)}} f_p(x) - \frac{\ln 2n}{p} \leq \inf_{x \in \mathbf{x}^{(0)}} g(x) \leq \inf_{x \in \mathbf{x}^{(0)}} f_p(x) \quad (1.7)$$

and, therefore,

$$\lim_{p \rightarrow \infty} \inf_{x \in \mathbf{x}^{(0)}} f_p(x) = \inf_{x \in \mathbf{x}^{(0)}} g(x). \quad (1.8)$$

Therefore, if  $p$  is large enough, the optimal solution of problem (1.5) sufficiently approximates that one of problem (0.3) in theory.

**Our main idea** is to solve the problem (0.3) by solving the corresponding problem (1.5) with interval analysis method and by using the corresponding minimum and Corollary 1.1 to estimate the minimum of problem (0.3), and at last we get the approximate solution of system (0.1).

## 1.2. Interval Extension

Now, we consider the interval extension of  $f_p(x)$ . Above all, we define interval extension of  $f_i(x)$  ( $i = 1, \dots, n$ ) as follows, respectively:

$$\mathbf{f}_i(\mathbf{x}) = \sum_{j=1}^n a_{ij} \mathbf{x}_j - b_i \quad (i = 1, \dots, n) \quad (1.9)$$

and we have  $\mathbf{f}_i(\mathbf{x}) = \{f_i(x) | x \in \mathbf{x}\}$ . Let

$$\underline{u}_i = \inf \mathbf{f}_i(\mathbf{x}), \quad \bar{u}_i = \sup \mathbf{f}_i(\mathbf{x}) \quad (i = 1, \dots, n).$$

Thus, we can define

$$\begin{aligned} \mathbf{f}_p(\mathbf{x}) &= \frac{1}{p} \ln \left\{ \sum_{i=1}^n (\exp(p\mathbf{f}_i(\mathbf{x})) + \exp(-p\mathbf{f}_i(\mathbf{x}))) \right\} = \\ &= \frac{1}{p} \left[ \ln \left\{ \sum_{i=1}^n (\exp(p\underline{u}_i) + \exp(-p\bar{u}_i)) \right\}, \ln \left\{ \sum_{i=1}^n (\exp(p\bar{u}_i) + \exp(-p\underline{u}_i)) \right\} \right], \quad (1.10) \\ \mathbf{f}'_p(\mathbf{x}) &= \frac{\sum_{i=1}^n a_i (\exp(p\mathbf{f}_i(\mathbf{x})) - \exp(-p\mathbf{f}_i(\mathbf{x})))}{\sum_{i=1}^n (\exp(p\mathbf{f}_i(\mathbf{x})) + \exp(-p\mathbf{f}_i(\mathbf{x})))}, \end{aligned}$$

where  $a_i = (a_{i1}, \dots, a_{in})$ ,  $i = 1, 2, \dots, n$ .

Obviously, the interval function  $\mathbf{f}_p(\mathbf{x})$  and  $\mathbf{f}'_p(\mathbf{x})$  are the interval extension of the function  $f_p(x)$  and the derivative  $\frac{\partial f_p(x)}{\partial x}$  on box  $\mathbf{x}$ .

Suppose that  $\mathbf{x}^{(0)}$  is subdivided into  $N$  segments  $\mathbf{x}^{(j)}$  ( $j = 1, \dots, N$ ) and  $\mathbf{x}^{(0)} = \cup_{j=1}^N \mathbf{x}^{(j)}$ , such that the width of  $\mathbf{x}^{(j)}$  satisfies  $\omega_N = \max_{1 \leq j \leq N} \|\text{wid}(\mathbf{x}^{(j)})\|_\infty \rightarrow 0 (N \rightarrow \infty)$ . Let  $\{p_N\}$  be monotonically increasing sequences of real numbers with  $p_N \geq 1$ ,  $N = 1, 2, \dots$ , such that  $\lim_{N \rightarrow \infty} p_N = +\infty$ , and let

$$a_N = \min_{1 \leq l \leq N} \{\text{inf}_{p_N}(\mathbf{x}^{(l)})\}. \quad (1.11)$$

**Theorem 1.2.** *Given  $\epsilon > 0$ ,  $\exists \bar{N} > 0$ , such that  $p_N \rightarrow \infty$  ( $\forall N > \bar{N}$ )*

$$\left| \min_{x \in \mathbf{x}^{(0)}} \max_{1 \leq i \leq n} \{|f_i(x)|\} - a_N \right| < \epsilon. \quad (1.12)$$

**Proof.** From (1.4), we know that there exists  $N_1 > 0$  such that ( $\forall N > N_1$ )

$$\left| \min_{x \in \mathbf{x}^{(0)}} \max_{1 \leq i \leq n} \{|f_i(x)|\} - \min_{x \in \mathbf{x}^{(0)}} f_{p_N}(x) \right| = \left| \min_{x \in \mathbf{x}^{(0)}} g(x) - \min_{x \in \mathbf{x}^{(0)}} f_{p_N}(x) \right| < \epsilon. \quad (1.13)$$

On the other hand, there exists  $\mathbf{x}^{(l^*)}$  such that

$$\text{inf}_{p_N}(\mathbf{x}^{(l^*)}) = \min_{1 \leq l \leq N} \{\text{inf}_{p_N}(\mathbf{x}^{(l)})\} = a_N.$$

Let

$$\underline{v} = \min_{x \in \mathbf{x}^{(l^*)}} f_i(x), \quad \bar{v} = \min_{x \in \mathbf{x}^{(l^*)}} (-f_i(x)),$$

then by (1.9), we have

$$\begin{aligned} \left| \min_{x \in \mathbf{x}^{(0)}} f_{p_N}(x) - a_N \right| &= \left| \min_{x \in \mathbf{x}^{(0)}} f_{p_N}(x) - \text{inf}_{p_N}(\mathbf{x}^{(l^*)}) \right| \leq \left| \min_{x \in \mathbf{x}^{(l^*)}} f_{p_N}(x) - \text{inf}_{p_N}(\mathbf{x}^{(l^*)}) \right| = \\ &= \left| \frac{1}{p_N} \ln \frac{\sum_{i=1}^n (\exp(p_N \cdot \underline{v}) + \exp(p_N \cdot \bar{v}))}{\sum_{i=1}^n (\exp(p_N \underline{\mathbf{f}}_i(\mathbf{x}^{(l^*)})) + \exp(-p_N \bar{\mathbf{f}}_i(\mathbf{x}^{(l^*)})))} \right| = \frac{1}{p_N} \ln 1 = 0. \end{aligned} \quad (1.14)$$

Therefore, for  $\bar{N} = N_1$ , from (1.13) and (1.14), we deduce that ( $N > \bar{N}$ )

$$\begin{aligned} \left| \min_{x \in \mathbf{x}^{(0)}} \max_{1 \leq i \leq n} \{|f_i(x)|\} - a_N \right| &\leq \left| \min_{x \in \mathbf{x}^{(0)}} \max_{1 \leq i \leq n} \{|f_i(x)|\} - \min_{x \in \mathbf{x}^{(0)}} f_{p_N}(x) \right| + \\ &+ \left| \min_{x \in \mathbf{x}^{(0)}} f_{p_N}(x) - a_N \right| \leq \epsilon + 0 \leq \epsilon. \end{aligned}$$

So, we complete this proof.

## 2. Interval Algorithm

In actual computations, it is important for our algorithm how to choose the initial interval (box) that includes the solution of system  $Ax = b$ . Here, we give the following result.

**Theorem 2.1.** *Let the matrix  $A$  be nonsingular and  $x^*$  be some isolated solution of the system  $Ax = b$ , then*

$$x^* \in E \cdot \frac{(\|b\|_\infty + \epsilon)n}{\lambda^{1/2}},$$

where  $E = (E_0, \dots, E_0)^\top$ ,  $E_0 = [-1, 1]$ ,  $\epsilon > 0$ , and  $\lambda$  is minimal eigenvalue of  $A^\top A$ .

**Proof.** We have

$$n \cdot \|Ax\|_\infty \geq \|Ax\|_2 = (x^\top A^\top Ax)^{1/2} \geq (\lambda x^\top x)^{1/2}.$$

Since  $A$  is nonsingular, then  $\lambda > 0$

$$\|Ax\|_\infty \geq \frac{(\lambda x^\top x)^{1/2}}{n}$$

and

$$\|Ax - b\|_\infty + \|b\|_\infty \geq \|Ax\|_\infty \geq \frac{\|x\|_2}{n} (\lambda)^{1/2}.$$

For all  $\epsilon > 0$ , if  $\|x\|_2 > \frac{(\|b\|_\infty + \epsilon)n}{\lambda^{1/2}}$ , then  $\|Ax - b\|_\infty > 0$ . So, we can deduce that

$$\|x\|_2 \leq \frac{(\|b\|_\infty + \epsilon)n}{\lambda^{1/2}},$$

furthermore, we have

$$x^* \in E \cdot \frac{(\|b\|_\infty + \epsilon)n}{\lambda^{1/2}}.$$

### 2.1. Test Operator

For  $Ax = b$  we define interval test operator as follows:

$$\mathcal{K}(\mathbf{x}) = Db + (I - DA)\mathbf{x}, \quad (2.1)$$

where  $D$  is an  $n \times n$  nonsingular real matrix,  $I$  is the  $n \times n$  unit matrix. It is easy to prove the following theorem.

**Theorem 2.2.** *For each  $\mathbf{x} \in \mathbf{x}^{(0)}$ :*

1) *every solution  $x^*$  of the system  $Ax = b$  within the box  $\mathbf{x}$  is also contained in  $\mathcal{K}(\mathbf{x})$ , so that  $x^* \in \mathbf{x} \cup \mathcal{K}(\mathbf{x})$ ;*

2) *if  $\mathbf{x} \cap \mathcal{K}(\mathbf{x}) = \phi$ , then the box  $\mathbf{x}$  contains no solutions of the system  $Ax = b$ ;*

3) *if  $\mathcal{K}(\mathbf{x}) \subset \mathbf{x}$ , then the box  $\mathbf{x}$  contains, with certainty, at least one solution of the system  $Ax = b$ ;*

4) *if  $\mathcal{K}(\mathbf{x}) \subset \text{int}(\mathbf{x})$  and  $A$  is nonsingular, then the box  $\mathcal{K}(\mathbf{x})$  contains exactly one solution of the system  $Ax = b$ .*

In actual application, we can also improve operator  $\mathcal{K}$  by applying the Gauss-Seidel technique

$$\mathcal{K}_i(\mathbf{x}) = p_i + \sum_{j=1}^n V_{ij} X_j \quad (i = 1, \dots, n), \quad (2.2)$$

where  $p = Db$ ,  $V = I - DA$ . So, we have

$$\tilde{\mathcal{K}}_i(\mathbf{x}) = p_i + \sum_{j=1}^{i-1} \mathcal{K}_j^*(\mathbf{x}) + \sum_{j=i+1}^n V_{ij} \mathbf{x}_j \quad (i = 1, \dots, n), \quad (2.3)$$

where  $\mathcal{K}_j^*(\mathbf{x}) = \tilde{\mathcal{K}}_j(\mathbf{x}) \cap \mathbf{x}_j$  ( $j = 1, \dots, n-1$ ). The expression of the box  $\tilde{\mathcal{K}}(\mathbf{x})$  can be rewritten in the following matrix form:

$$\tilde{\mathcal{K}}(\mathbf{x}) = p - L\mathcal{K}^*(\mathbf{x}) - U\mathbf{x}, \quad (2.4)$$

where  $\mathcal{K}^*(\mathbf{x}) = \tilde{\mathcal{K}}(\mathbf{x}) \cap \mathbf{x}$  and

$$L = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ a_{21} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ a_{n1} & \cdots & a_{n,n-1} & 0 \end{pmatrix}, \quad U = \begin{pmatrix} a_{11} - 1 & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} - 1 & \cdots & a_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nn} - 1 \end{pmatrix}.$$

$\tilde{\mathcal{K}}$  is the Gauss-Seidel interval test operator, and we can also prove that  $\tilde{\mathcal{K}}(\mathbf{x})$  has the same property as Theorem 2.2, and it has smaller excess-width under certain condition, Therefore, replacing  $\mathcal{K}(\mathbf{x})$  with  $\tilde{\mathcal{K}}(\mathbf{x})$  in the described algorithm, we usually can obtain better results.

In the process of actual computing, the matrix  $D$  can be chosen in different ways. For example, let  $D = (e_1, e_2, \dots, e_m, e_1, e_2, \dots, e_{n-m})^\top$ , where  $e_i \in R^n$ , or let  $D = \text{diag}(a_{11}, a_{22}, \dots, a_{nn})$  and so on. The choice of  $D$  affects the result of iteration.

**Example 1.** For

$$\begin{cases} 2x_1 - x_2 = 2, \\ -4x_1 + 7x_2 = -5, \end{cases}$$

where  $\mathbf{x} = ([0.8, 1.2], [-0.2, 0.3])^\top$ ,  $x^* = (1.1, 0.2)^\top$ .

Let

$$D = \begin{bmatrix} 1/2 & 0 \\ 0 & 1/7 \end{bmatrix},$$

then

$$\mathcal{K}(\mathbf{x}) = Db + (I - DA)\mathbf{x} = \begin{pmatrix} [0.9, 1.15] \\ \left[ \frac{-3.1}{7}, \frac{1}{4} \right] \end{pmatrix},$$

so, we have

$$\mathcal{K}(\mathbf{x}) \cap \mathbf{x} = ([0.9, 1.15], [0.2, 0.25])^\top$$

and

$$x^* \in \mathcal{K}(\mathbf{x}) \cap \mathbf{x}, \quad \text{wid}(\mathcal{K}(\mathbf{x}) \cap \mathbf{x}) < \text{wid}(\mathbf{x}).$$

If let  $\mathbf{x} = ([0.8, 1.2], [-0.2, 0])^\top$ , then

$$\mathcal{K}(\mathbf{x}) = Db + (I - DA)\mathbf{x} = \begin{pmatrix} [0.9, 1] \\ \left[ \frac{-4.4}{7}, \frac{-0.2}{7} \right] \end{pmatrix}$$

and

$$\mathcal{K}(\mathbf{x}) \cap \mathbf{x} = \left( [0.9, 1], \left[ 0.2, \frac{-0.2}{7} \right] \right)^\top.$$

Further, let  $\mathbf{x} = \mathcal{K}(\mathbf{x}) \cap \mathbf{x}$ , then

$$\mathcal{K}(\mathbf{x}) \cap \mathbf{x} = \phi,$$

so, we have result  $x^* \notin ([0.8, 1.2], [-0.2, 0])^\top$  according to Theorem 2.2.

## 2.2. Region Deletion Rules

Region deletion rules can reduce the computational cost and accelerate the convergence rate of algorithm.

**Deletion rule 1** (feasible region test):

If  $\exists j^* \in \{1, \dots, n\}$  such that  $0 \notin \mathbf{f}_{j^*}(\mathbf{x})$ , then discard  $\mathbf{x}$ , which is an infeasible region.

**Deletion rule 2** (midpoint test):

Let  $\bar{f}$  be an upper bound of  $\min_{x \in \mathbf{x}^{(0)}} f_p(x)$ . For  $\mathbf{x} \in \mathbb{I}(\mathbf{x}^{(0)})$ , if  $\underline{\mathbf{f}}_p(\mathbf{x}) > \bar{f}$ , then there is no optimal point in  $\mathbf{x}$ , and  $\mathbf{x}$  can be discarded.

In the process of actual computing,  $\bar{f}$  be selected and adjusted according to algorithm and interval extension of  $f_p(x)$ . In this paper, at first let  $\bar{f} = f_p(\text{mid}(\mathbf{x}))$  (suppose that feasible solution exists in  $\mathbf{x}^{(0)}$ . For  $\mathbf{x} \subseteq \mathbf{x}^{(0)}$ , we have

$$\bar{f} = \min\{\bar{f}, f_p(\text{mid}(\mathbf{x}))\}. \quad (2.5)$$

**Deletion rule 3** (monotonicity test):

1) if  $\bar{\mathbf{f}}_{pk}(\mathbf{x}) < 0$ , then let  $\mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_{k-1}, \bar{\mathbf{x}}_k, \mathbf{x}_{k+1}, \dots, \mathbf{x}_n)^\top$ ;

2) if  $\underline{\mathbf{f}}_{pk}(\mathbf{x}, \beta) > 0$ , then let  $\mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_{k-1}, \underline{\mathbf{x}}_k, \mathbf{x}_{k+1}, \dots, \mathbf{x}_n)^\top$ .

**Deletion rule 4:**

1) if  $\mathbf{x} \cap \mathcal{K}(\mathbf{x}) = \phi$ , then the box  $\mathbf{x}$  can be discarded;

2) if  $\mathbf{x} \cap \mathcal{K}(\mathbf{x}) \neq \phi$ , then the box  $\mathbf{x} = \mathbf{x} \cap \mathcal{K}(\mathbf{x})$ .

## 2.3. Interval Algorithm

With the preceding ideas in mind, we suggest the following interval algorithm for solving system (0.1). The algorithm starts from  $\mathbf{x}^{(0)}$  and generates a list L. The elements in the list L are of the form  $(\mathbf{x}, \underline{\mathbf{f}}_p, \bar{f})$  and are arranged in an increasing order of  $\underline{\mathbf{f}}_p$ , in which  $\bar{f}$  is selected and adjusted according to (2.5), and  $\underline{\mathbf{f}}_p = \underline{\mathbf{f}}_p(\mathbf{x})$ . Each  $\mathbf{x}$  can be obtained from the region of the first element in the list L in the following ways. The region is bisected normally to the coordinate direction, in which  $\mathbf{x}$  has the maximal width.

The detailed steps of the algorithm are the following.

**Step 1.** Set  $\mathbf{x} = \mathbf{x}^{(0)}$ ,  $f = f_p(\text{mid}(\mathbf{x}))$ ,  $\underline{\mathbf{f}}_p = \underline{\mathbf{f}}_p(\mathbf{x})$ ,  $\bar{f} = f$ ,  $B_0 = +\infty$ . Enter the item  $(\mathbf{x}, \underline{\mathbf{f}}_p, \bar{f})$  into the list L.

**Step 2.** If the list L is empty, then go to Step 10.

**Step 3.** Extract the first element  $(\mathbf{x}, \underline{\mathbf{f}}_p, \bar{f})$  from the list L and remove it from the list L. Set  $B = \min\{B_0, \underline{\mathbf{f}}_p\}$ .

**Step 4.** If  $\bar{f} - B < \epsilon$  for given tolerance  $\epsilon > 0$ , then proceed with Step 8.

**Step 5.** Bisect  $\mathbf{x}$  normal to the coordinate direction, in which  $\mathbf{x}$  has the maximal width, to obtain  $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}$  such that  $\mathbf{x} = \mathbf{x}^{(1)} \cup \mathbf{x}^{(2)}$ .

**Step 6:**

- 1) set  $i = 1$ ;
- 2) if there exists  $j \in \{1, \dots, n\}$  such that  $0 \notin \mathbf{f}_j(\mathbf{x}^{(i)})$ , then go to point (7);
- 3) compute  $\mathcal{K}(\mathbf{x}^{(i)})$ , let  $\mathbf{x}^{(i)} = \mathcal{K}(\mathbf{x}^{(i)}) \cap \mathbf{x}^{(i)}$ . If  $\mathbf{x}^{(i)} = \phi$ , then go to point (7);
- 4) if  $\bar{\mathbf{f}}'_{pk}(\mathbf{x}^{(i)}) < 0$ , then set  $\mathbf{x}^{(i)} = (\mathbf{x}_1^{(i)}, \dots, \mathbf{x}_{k-1}^{(i)}, \bar{\mathbf{x}}_k^{(i)}, \mathbf{x}_{k+1}^{(i)}, \dots, \mathbf{x}_n^{(i)})^\top$ ; if  $\underline{\mathbf{f}}'_{pk}(\mathbf{x}^{(i)}) > 0$ , then set  $\mathbf{x}^{(i)} = (\mathbf{x}_1^{(i)}, \dots, \mathbf{x}_{k-1}^{(i)}, \underline{\mathbf{x}}_k^{(i)}, \mathbf{x}_{k+1}^{(i)}, \dots, \mathbf{x}_n^{(i)})^\top$ ,  $k = 1, \dots, n$ ;
- 5) compute  $f = f_p(\text{mid}(\mathbf{x}^{(i)}))$ ;  $\underline{\mathbf{f}}_p = \underline{\mathbf{f}}_p(\mathbf{x}^{(i)})$ ,  $\bar{f} = \min\{\bar{f}, f\}$ ;
- 6) enter the item  $(\mathbf{x}^{(i)}, \underline{\mathbf{f}}_p, \bar{f})$  in proper order into the list L;
- 7) if  $i < 2$ , then let  $i = i + 1$ , then go to point (2).

**Step 7.** Discard all the elements  $(\mathbf{x}, \underline{\mathbf{f}}_p, \bar{f})$  from the list L that satisfy  $\underline{\mathbf{f}}_p > \bar{f}$ , return to Step 2.

**Step 8.** Print  $[\underline{\mathbf{f}}_p, \bar{f}]$  and  $\mathbf{x}$ .

**Step 9.** Set  $B_0 = B$ , and return to Step 2.

**Step 10.** End.

When working in actual computing process, it is necessary to keep in the mind the following items:

- 1) according to Corollary 1.1, let  $p > \frac{\ln 2n}{\epsilon}$ ;
- 2) in numerical computation, in order to avoid deletion of optimization solution that is on the boundary of interval, we use rounded interval arithmetics;
- 3) in this paper, we select

$$D = \begin{cases} a_{ii}, & a_{ii} \neq 0, \\ 1, & a_{ii} = 0. \end{cases}$$

## Conclusions

In order to test the efficiency of our algorithm, we consider the linear system  $Hx = b$  with Hilbert  $n \times n$  matrix. Recall that  $H$  is called the Hilbert matrix of order  $n$  if its entries  $h_{ij} = \left\{ \frac{1}{i+j-1} \right\}$ ,  $i, j = 1, \dots, n$ :

$$H = \begin{pmatrix} 1 & 1/2 & 1/3 & 1/4 & \cdots \\ 1/2 & 1/3 & 1/4 & 1/5 & \cdots \\ 1/3 & 1/4 & 1/5 & 1/6 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

The matrix is poor-conditioned even for small dimension. For this reason, serious difficulties arise when computing the solutions of linear equations system with this matrix because of the large sensitivity of the solution to errors in actual computing. Thus, the Hilbert matrix has been taken as a good test example for the algorithm of this paper. Let  $x = (x_1, \dots, x_n)^\top$ ,  $b = A(1, \dots, 1)^\top$  and  $\mathbf{x}_i^{(0)} = [-1, 2]$ ,  $i = 1, \dots, n$ . It is easy to check that the solution is  $x^* = (1, \dots, 1)^\top$ . To solve the above problem with traditional algorithms, we can draw the results as follows.

By the matrix factorization method, we can not obtain satisfactory solution because of the condition number  $Cond(A) \gg 1(n \geq 10)$ . Jacobi iterative method is divergent, Gauss-Seidel iterative and SOR iterative methods are convergent, but the rate of convergence of both methods is very slow, and it is hard to obtain the solution with precision less than  $10^{-4}$

because the spectral radius of iterative matrix is less than 1 and very close to 1. With our interval algorithm, let entropy factor be  $p = 10^6$  (or  $p$  be large enough), we can obtain the solution with any precision  $\epsilon > 0$ .

But it is necessary to note also the following problems:

1. Since of complexity of interval operations, we can spend more time with our algorithm to obtain a satisfying result (if given precision is  $\epsilon \leq 10^{-10}$  or smaller), but the interval method has the global convergence and obtains both the solution and its error simultaneously. So, the direction for future research is in obtaining additional theoretical results on convergence with new rules of region deletion and in finding the way to produce the box sequence  $\{x^{(k)}\}$  converging to the solution  $x^*$  without any a priori information.

2. To obtain the initial box  $\mathbf{x}^{(0)}$  by Theorem 2.1, we need to compute eigenvalues of  $A^T A$ . So, we expect to gain good result in future research, which is a direction for future research.

In this paper, a new algorithm based on the introduction of maximum entropy function and interval computation has been presented to generate the solution of ill-conditioned linear equations system. The algorithm can be applied to a number of problems such as solving systems of linear inequalities or nonlinear equations and nonlinear inequalities, global optimization, etc.

**Acknowledgements:** *The author thanks the referee for his/her valuable comments and suggestions which improved the original manuscript of this paper.*

## References

- [1] RATSCHKE H., ROKNE J. New computer methods for global optimization. Chichester: Ellis Horwood Limited, 1988.
- [2] WOLFE M.A. Interval methods for global optimization // Appl. Math. and Comput. 1996. Vol. 75. P. 179–206.
- [3] MOORE R.E. Methods and application of interval analysis. Philadelphia: SIAM, 1979.
- [4] SHEN Z., HUANG Z., WOLFE M.A. An interval maximum entropy method for a discrete minimax problem // Appl. Math. and Comput. 1997. Vol. 87. P. 49–68.
- [5] HANSEN E. Global optimization using interval analysis the multi-dimensional case // Appl. Math. and Comput. 1980. Vol. 35. P. 505–512.
- [6] HANSEN E. Global optimization using interval analysis. N. Y.: Marcel Dekker, Inc. 1992.
- [7] HANSEN E., WILLIAM G. Global optimization using interval analysis. Second Edition, Revised and Expanded. N. Y.: Marcel Dekker, Inc. 2004.
- [8] LI X. The coherent function method solving nonlinear program // Sciences in China (A). 1991. Vol. 12. P. 1283–1288.
- [9] LI X. The efficient methods for a sort of nondifferentiable optimization problem // Sciences in China(A). 1994. Vol. 24. P. 371–377.
- [10] WANG H., CAO D. Smoothing iterative method for a class of nonsmooth equations // Comput. Technol. 2006. Vol. 11, N 6. P. 1–9.
- [11] BOLTZMANN L. Weitere Studien über das Warmegleichgewicht unter Gasmolekulen // Wien. Akad. Sitz. 1972. Vol. 66. P. 275–370.
- [12] JAYNES E.T. Information theory and statistical mechanics // Phys. Rev. 1957. Vol. 106. P. 620–630.