Stud. Univ. Babeş-Bolyai Math. 59(2014), No. 4, 533–542

Enclosing the solution set of overdetermined systems of interval linear equations

Szilvia Huszárszky and Lajos Gergó

Abstract. We describe two methods to bound the solution set of full rank interval linear equation systems $\mathbf{A}x = \mathbf{b}$ where $\mathbf{A} \in \mathbb{IR}^{m\times n}$, $m \geq n$ is a full rank interval matrix and $\mathbf{b} \in \mathbb{IR}^m$ is an interval vector. The methods are based on the concept of generalized solution of overdetermined systems of linear equations. We use two type of preconditioning the $m \times n$ system: multiplying the system with the generalized inverse of the midpoint matrix or with the transpose of the midpoint matrix. It results an $n \times n$ system which we solve using Gaussian elimination or the method provided by J. Rohn in [8]. We give some examples in which we compare the efficiency of our methods and compare the results with the interval Householder method [11].

Mathematics Subject Classification (2010): 65G06.

Keywords: Interval linear equation, overdetermined, preconditioning.

1. Introduction

An interval matrix, $\mathbf{A}$, is a matrix whose elements are intervals, an interval vector, $\mathbf{b}$, is a vector whose components are intervals. Let $\mathbf{A} = [\underline{A}, \overline{A}]$ be an $m \times n$ interval matrix and $\mathbf{b} = [\underline{b}, \overline{b}]$ an m-dimensional interval vector. We suppose that $m \geq n$ and the interval matrix $\mathbf{A}$ has full rank, i.e., all real matrices $A \in \mathbf{A}$ have full rank. Consider the set of linear equations

$$\mathbf{A}x = \mathbf{b}. \tag{1.1}$$

The set of solutions of such problem is given by

$$\sum(\mathbf{A}, \mathbf{b}) = \left\{ \widetilde{x} \in \mathbb{R}^n \mid \exists A \in \mathbf{A}, \exists b \in \mathbf{b} : \; \|A\widetilde{x} - b\| = \min_{x\in\mathbb{R}^n} \|Ax - b\| \right\},$$

i.e., the minimalization of $\|Ax - b\|$ for any $A \in \mathbf{A}$ and any $b \in \mathbf{b}$.

This paper was presented at the 10th Joint Conference on Mathematics and Computer Science (MaCS 2014), May 21-25, 2014, Cluj-Napoca, Romania.

In recent years, much attention has been paid to systems of interval linear equations (1.1) with square interval matrices (see, for example, [1], [3], [4] [5]). Lot of works were performed to compute an enclosure interval vector of the set $\sum(\mathbf{A},\mathbf{b})$ which becomes as follows,

$$\sum(\mathbf{A},\mathbf{b}) = \{x \in \mathbb{R}^n \mid \exists A \in \mathbf{A}, \exists b \in \mathbf{b} : \ Ax = b\}.$$

The solution set is generally of a complicated non-convex structure. In practical computations, therefore, we look for an enclosure of it, i.e., for an interval vector $\mathbf{x}$ satisfying

$$\sum(\mathbf{A},\mathbf{b}) \subseteq \mathbf{x}.$$

If $\mathbf{A}$ is regular, then the intersection of all enclosures of $\sum(\mathbf{A},\mathbf{b})$ forms an interval vector which is called the interval hull of $\sum(\mathbf{A},\mathbf{b})$. If $\mathbf{A}$ is singular, then $\sum(\mathbf{A},\mathbf{b})$ is either empty, or unbounded and the interval hull is not defined in this case.

We note that it is especially interest if the matrix of the interval linear equation system is non-squared. Several papers have been published in this topic. Further details can be found for example in [4], [11]. In the present paper, we propose additional methods for the full rank case. The purpose of this work is to give an enclosure interval vector of the solution set of the overdetermined systems of interval linear equations. These methods are based on Hansen's preconditioning and the concept of generalized solution of full rank overdetermined systems of linear equations.

In the next section, we introduce some notation. In subsection 3.1 we describe variations of the preconditioning. This preconditioning results an $n \times n$ interval linear equation system. In subsection 3.2 we describe which methods have been used to solve the square interval linear equation system. In section 4 we give some examples in which we compare the efficiency of our methods and compare the results with the preconditioning interval Householder method [11] and with the interval Cholesky method [12] applied to the symmetric $n \times n$ interval linear system $\mathbf{A}^T\mathbf{A}x = \mathbf{A}^T\mathbf{b}$.

2. Notations and operations

We denote the set of real compact intervals by $\mathbb{IR}$ whose elements are $[a] = [\underline{a}, \overline{a}] = \{x \in \mathbb{R} \mid \underline{a} \le x \le \overline{a}\}$, for $\underline{a} \le \overline{a}$ and $\underline{a}, \overline{a} \in \mathbb{R}$. The set of $m \times n$ matrices over the real compact intervals is denoted by $\mathbb{IR}^{m\times n}$.

Let $[a] = [\underline{a}, \overline{a}]$ and $[b] = [\underline{b}, \overline{b}]$ are real compact intervals and let $* \in \{+, -, \cdot, :\}$. Then arithmetic operations on intervals are defined by [1]

$$[a] * [b] = \{x * y \mid \underline{a} \le x \le \overline{a}, \ \underline{b} \le y \le \overline{b}\}.$$

It is assumed that $0 \notin [b]$ in the case of division. We note that $[a]*[b]$ is a real compact interval and

$$[a] * [b] = [\min\{\underline{a} * \underline{b}, \ \underline{a} * \overline{b}, \ \overline{a} * \underline{b}, \ \overline{a} * \overline{b}\}, \ \max\{\underline{a} * \underline{b}, \ \underline{a} * \overline{b}, \ \overline{a} * \underline{b}, \ \overline{a} * \overline{b}\}].$$

For $\mathbf{A}, \mathbf{B} \in \mathbb{IR}^{m\times n}$, $\mathbf{A} \pm \mathbf{B}$ is the $m \times n$ interval matrix whose elements are $\mathbf{A}_{ij} \pm \mathbf{B}_{ij}$. If $\mathbf{A} \in \mathbb{IR}^{m\times n}$ and $\mathbf{B} \in \mathbb{IR}^{n\times r}$ than $\mathbf{A} \cdot \mathbf{B}$ is the $m \times r$ interval matrix whose elements are given by

$$(\mathbf{A}\cdot\mathbf{B})_{ij}=\sum_{k=1}^{n}\mathbf{A}_{ik}\cdot\mathbf{B}_{kj}.$$

For $\mathbf{A}\in\mathbb{IR}^{m\times n}$ and $\mathbf{b}\in\mathbb{IR}^{n}$, $\mathbf{A}\cdot\mathbf{b}$ is an m-dimensional interval vector whose components are defined by

$$(\mathbf{A}\cdot\mathbf{b})_{i}=\sum_{k=1}^{n}\mathbf{A}_{ik}\cdot\mathbf{b}_{k}.$$

If $[a]\in\mathbb{IR}$ and $\mathbf{b}\in\mathbb{IR}^{n}$, $[a]\cdot\mathbf{b}$ is an interval vector whose components are given by

$$([a]\cdot\mathbf{b})_{i}=[a]\cdot\mathbf{b}_{i}.$$

For an interval matrix $\mathbf{A}=[\underline{A},\overline{A}]$ we define

$$A_c:=\frac{1}{2}(\underline{A}+\overline{A})$$

the midpoint matrix whose each element $(A_c)_{ij}$ corresponds to the midpoint of the element $\mathbf{A}_{ij}$ of $\mathbf{A}$. Let

$$\Delta:=\frac{1}{2}(\overline{A}-\underline{A})$$

denote the radius matrix whose each element Δ_{ij} corresponds to the radius of the element $\mathbf{A}_{ij}$ of $\mathbf{A}$. Then $\underline{A}=A_c-\Delta$ and $\overline{A}=A_c+\Delta$, so that we also can write $\mathbf{A}=[A_c-\Delta,A_c+\Delta]$. Similarly, for an interval vector $\mathbf{b}=[\underline{b},\overline{b}]$

$$b_c:=\frac{1}{2}(\underline{b}+\overline{b})$$

the midpoint vector and

$$\delta:=\frac{1}{2}(\overline{b}-\underline{b})$$

the radius vector thus $\mathbf{b}=[b_c-\delta,b_c+\delta]$.

3. Solving overdetermined linear interval systems

3.1. Preconditioning

We now describe two ways to obtain the preconditioning matrix. As we have seen in the square case, when solving the system $\mathbf{A}x=\mathbf{b}$ of linear equations using interval version of methods such as Gaussian elimination, it is generally advisable to precondition the system. The most commonly used method of preconditioning is to multiply by an approximate inverse of A_c. The products $A_c^{-1}\mathbf{A}$ and $A_c^{-1}\mathbf{b}$ are computed using interval arithmetic. The solution set of the preconditioned equation

$$A_c^{-1}\mathbf{A}x=A_c^{-1}\mathbf{b}$$

contains the solution set of the original equation (see [1]).

When the interval elements of $\mathbf{A}$ and $\mathbf{b}$ are narrow, preconditioning increases the size of the solution set only slightly. When the intervals are wide, preconditioning can substantially increase the size of the solution set. If preconditioning is not used, interval widths generally grow so fast during the solution process that the final results are of little use. This preconditioning was introduced by E.R. Hansen in [2].

In our case, when the matrix of the interval linear equation is not square, $m > n$, two way due to the preconditioning. Our first method of preconditioning is to multiply by the generalized inverse of the midpoint matrix of $\mathbf{A}$. Since the system we consider is overdetermined and all $A \in \mathbf{A}$ has full rank, the generalized inverse of A_c is given by

$$A_c^+ = (A_c^T A_c)^{-1} A_c^T.$$

The products $\widetilde{\mathbf{A}}_\mathbf{1} := A_c^+ \mathbf{A}$ and $\widetilde{\mathbf{b}}_\mathbf{1} := A_c^+ \mathbf{b}$ are computed using interval arithmetic. After the preconditioning, we get the following $n \times n$ interval linear equation system

$$\widetilde{\mathbf{A}}_\mathbf{1} x = \widetilde{\mathbf{b}}_\mathbf{1} \tag{3.1}$$

which we can solve by one of the several existing method.

The second way of preconditioning is comes from the idea of the point (non-interval) case. Our second method of preconditioning is to multiply by the transpose of A_c. Just like in the previous case, the products $\widetilde{\mathbf{A}}_\mathbf{2} := A_c^T \mathbf{A}$ and $\widetilde{\mathbf{b}}_\mathbf{2} := A_c^T \mathbf{b}$ are computed using interval arithmetic. After the preconditioning we get the following $n \times n$ interval linear equation system

$$\widetilde{\mathbf{A}}_\mathbf{2} x = \widetilde{\mathbf{b}}_\mathbf{2} \tag{3.2}$$

which we can solve by one of the several existing method.

As we will see in section 4, the bounds on the solution set of (3.1) is usually narrower then the bounds on the solution set of (3.2). On the other hand, the calculation of the transpose of the midpoint matrix requires fewer arithmetic operations than the calculation of the generalized inverse of A_c.

3.2. Bounding the solution of interval linear equations

Preconditioning described above results an $n \times n$ interval linear equation system. Several methods were developed to compute an enclosure interval vector of the set $\sum(\widetilde{\mathbf{A}}, \widetilde{\mathbf{b}})$. For detailed we refer to [6], [7], [8].

First algorithm was used to bound the solution set of (3.1) and (3.2) was the method provided by J. Rohn in [8]. This algorithm either computes the interval hull of the solution set of the system of interval linear equations $\widetilde{\mathbf{A}} x = \widetilde{\mathbf{b}}$, or finds a singular matrix $S \in \widetilde{\mathbf{A}}$. It has been proved the algorithm terminates in finite number of steps for each $n \times n$ interval matrix $\widetilde{\mathbf{A}}$ and for each n-dimensional interval vector $\widetilde{\mathbf{b}}$. In this algorithm we have to solve an equation of the form

$$Ax + B|x| = b \tag{3.3}$$

where $A, B \in \mathbb{R}^{n \times n}$ and $b \in \mathbb{R}^n$, which is called an absolute value equation. A very efficient algorithm for the solution of equation (3.3) was described by J. Rohn in [9], [10].

Since the interval matrix $\widetilde{\mathbf{A}}_\mathbf{i}$ ($i \in \{1, 2\}$) is non-singular, Gaussian elimination also can be used to bound the solution set of (3.1) and (3.2). The direct generalization of the Gaussian algorithm was described in [1]. As we will see in example 4.1 sometimes

Gaussian elimination gives the same result as Rohn's algorithm. Generally Rohn's method provide better enclosure of the solution set.

4. Numerical examples

Let $\mathbf{A} \in \mathbb{IR}^{m\times n}$ be a full rank interval matrix where $m > n$ and $\mathbf{b} \in \mathbb{IR}^m$. Let $A_c = Q \cdot R$ be the QR factorization of A_c. The $n \times n$ triangular real matrix R_1 is obtained by dropping from R the last $m - n$ rows. Let $\mathbf{B} := (Q^T\mathbf{A}) \cdot R_1^{-1}$ and $\mathbf{c} := Q^T\mathbf{b}$. The subvector of c_c obtained by dropping the last $m - n$ components is denoted by x_0. Let the interval vector $\mathbf{d}$ is given by d_c is the subvector of c_c obtained by replacing the first n components by zeros and $\delta_d := \Delta_B \cdot |x_0| + \delta_c$. Let the interval vector $\mathbf{h}$ is given by $h_c := d_c$ and $\delta_h := \Delta_B \cdot x_0$. The following results was showed by A.H. Bentbib in [11].

$$\sum(\mathbf{A},\mathbf{b}) \subseteq R_1^{-1} \cdot \sum(\mathbf{B},\mathbf{c}) \subseteq R_1^{-1} \cdot \left(x_0 + \sum(\mathbf{B},\mathbf{d})\right)$$

and

$$\sum(\mathbf{A},\mathbf{b}) \subseteq R_1^{-1} \cdot \left(\sum(\mathbf{B},c_c) + \sum(\mathbf{B},\widetilde{\mathbf{c}})\right) \subseteq$$
$$\subseteq R_1^{-1} \cdot \left(x_0 + \sum(\mathbf{B},\mathbf{h}) + \sum(\mathbf{B},\widetilde{\mathbf{c}})\right)$$

where $\widetilde{\mathbf{c}}$ denote the interval vector whose component $\widetilde{\mathbf{c}}_i$ corresponds to the centered interval $[-(\delta_c)_i, (\delta_c)_i]$.

Let the interval vector given by the Householder method applied to the overdetermined full rank interval linear equation system $\mathbf{A}x = \mathbf{b}$ is denoted by $\mathbf{x}_H(\mathbf{A},\mathbf{b})$. We denote by $\mathbf{x}_{Ch}(\mathbf{A},\mathbf{b})$ the interval vector given by the interval Cholesky method [12] applied to symmetric $n \times n$ interval linear equation system $\mathbf{A}^T\mathbf{A}x = \mathbf{A}^T\mathbf{b}$. A.H. Bentbib compared the interval vectors $\mathbf{v}_1 = \mathbf{x}_H(\mathbf{A},\mathbf{b})$, $\mathbf{v}_2 = R_1^{-1}\mathbf{x}_H(\mathbf{B},\mathbf{c})$, $\mathbf{v}_3 = R_1^{-1}(x_0 + \mathbf{x}_H(\mathbf{B},\mathbf{d}))$, $\mathbf{v}_4 = R_1^{-1}(\mathbf{x}_H(\mathbf{B},\mathbf{c_c}) + \mathbf{x}_H(\mathbf{B},\widetilde{\mathbf{c}}))$ and $\mathbf{v}_5 = R_1^{-1}(x_0 + \mathbf{x}_H(\mathbf{B},\mathbf{h}) + \mathbf{x}_H(\mathbf{B},\widetilde{\mathbf{c}}))$ which all contains $\sum(\mathbf{A},\mathbf{b})$.

Let us denote by $\mathbf{x}_1$ the interval vector given by J. Rohn's method [8] applied to the square interval linear equation system $A_c^+\mathbf{A}x = A_c^+\mathbf{b}$, by $\mathbf{x}_2$ the interval vector given by J. Rohn's method applied to the square interval linear equation system $A_c^T\mathbf{A}x = A_c^T\mathbf{b}$. Let $\mathbf{x}_3$ denote the enclosure interval vector of the solution set of the square interval linear equation system $A_c^+\mathbf{A}x = A_c^+\mathbf{b}$ and $\mathbf{x}_4$ denote the enclosure interval vector of the solution set of the square interval linear equation system $A_c^T\mathbf{A}x = A_c^T\mathbf{b}$ by using Gaussian elimination.

Example 4.1. Let us consider the following interval linear system

$$\begin{pmatrix} [0.1,\ 0.3] & [0.9,\ 1.1] \\ [8.9,\ 9.1] & [0.4,\ 0.6] \\ [0.9,\ 1.1] & [6.9,\ 7.1] \end{pmatrix} \cdot x = \begin{pmatrix} [0.8,\ 1.2] \\ [-0.2,\ 0.2] \\ [1.8,\ 2.2] \end{pmatrix}.$$

The following results was published by A.H. Bentbib:

$$\mathbf{x}_{Ch} = \begin{pmatrix} [-0.0642,\ 0.0285] \\ [0.2408,\ 0.3692] \end{pmatrix}, \quad \mathbf{v}_1 = \begin{pmatrix} [-0.0761,\ 0.0362] \\ [0.2199,\ 0.4024] \end{pmatrix},$$

$$\mathbf{v}_2 = \begin{pmatrix} [-0.0616,\ 0.0294] \\ [0.2579,\ 0.3485] \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} [-0.0558,\ 0.0232] \\ [0.2560,\ 0.3486] \end{pmatrix},$$

$$\mathbf{v}_4 = \begin{pmatrix} [-0.0620,\ 0.0296] \\ [0.2564,\ 0.3485] \end{pmatrix}, \quad \mathbf{v}_5 = \begin{pmatrix} [-0.0558,\ 0.0232] \\ [0.2560,\ 0.3486] \end{pmatrix}.$$

We have the following results:

$$\mathbf{x}_1 = \begin{pmatrix} [-0.0451,\ 0.0121] \\ [0.2614,\ 0.3447] \end{pmatrix}, \quad \mathbf{x}_2 = \begin{pmatrix} [-0.0532,\ 0.0186] \\ [0.2557,\ 0.3516] \end{pmatrix},$$

$$\mathbf{x}_3 = \begin{pmatrix} [-0.0451,\ 0.0121] \\ [0.2614,\ 0.3447] \end{pmatrix}, \quad \mathbf{x}_4 = \begin{pmatrix} [-0.0532,\ 0.0188] \\ [0.2544,\ 0.3517] \end{pmatrix}.$$

Let e denote the real vector whose all components are equal to 1 and by E we denote the real $m \times n$ matrix whose elements are all equal to 1. We illustrate the real vectors $\underline{x}$ and $\overline{x}$ according to the index of component i which is varying from 1 to n.

Example 4.2. (See Figures 1-4.) Let $\mathbf{A} \in \mathbb{IR}^{m\times n}$ is given by

$$A_c = rand(m,n) + 4E - 2I, \qquad \Delta = \varepsilon_1 \cdot E.$$

The interval vector $\mathbf{b} \in \mathbb{IR}^m$ is given by

$$b_c = A_c \cdot e, \qquad \delta = \varepsilon_2 \cdot e.$$

Example 4.3. (See Figures 5-7.) Let $\mathbf{A} \in \mathbb{IR}^{m\times n}$ is given by

$$A_c = rand(m,n) + 3I, \qquad \Delta = \varepsilon_1 \cdot E.$$

The interval vector $\mathbf{b} \in \mathbb{IR}^m$ is given by

$$b_c = A_c \cdot e, \qquad \delta = \varepsilon_2 \cdot e.$$

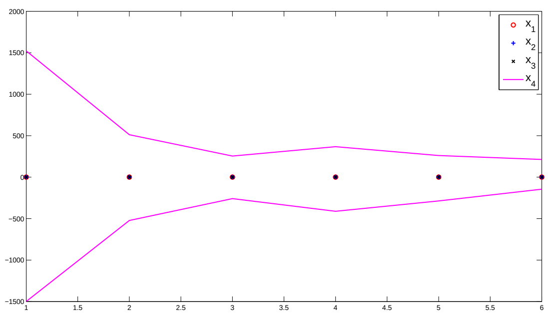

Figure 1. For $m = 10$, $n = 6$, $\varepsilon_1 = 10^{-3}$, $\varepsilon_2 = 10^{-3}$.

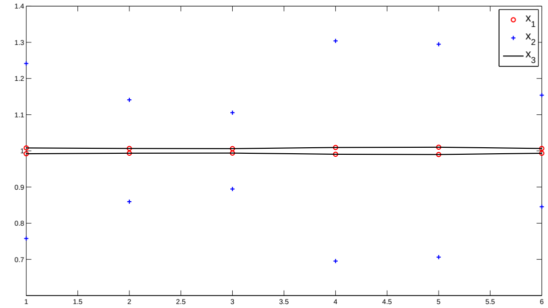

Figure 2. For $m = 10$, $n = 6$, $\varepsilon_1 = 10^{-3}$, $\varepsilon_2 = 10^{-3}$.

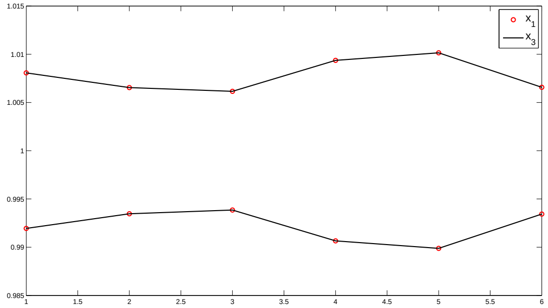

Figure 3. For $m = 10$, $n = 6$, $\varepsilon_1 = 10^{-3}$, $\varepsilon_2 = 10^{-3}$.

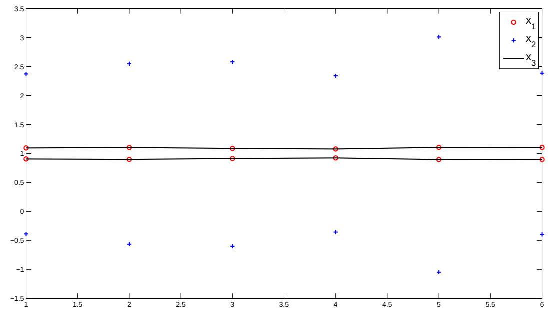

Figure 4. For $m = 10$, $n = 6$, $\varepsilon_1 = 10^{-3}$, $\varepsilon_2 = 10^{-1}$.

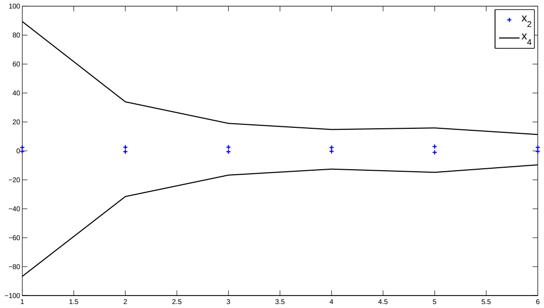

Figure 5. For $m = 10$, $n = 6$, $\varepsilon_1 = 10^{-3}$, $\varepsilon_2 = 10^{-1}$.

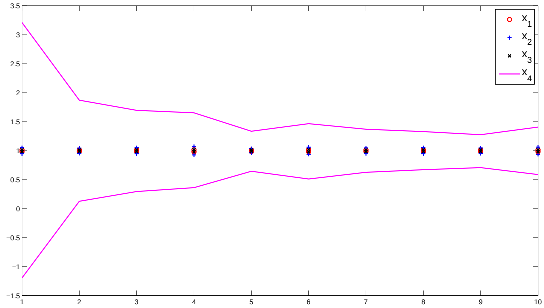

Figure 6. For $m = 20$, $n = 10$, $\varepsilon_1 = 10^{-3}$, $\varepsilon_2 = 10^{-2}$.

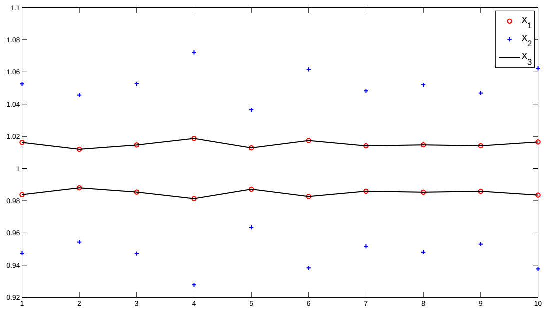

Figure 7. For $m = 20$, $n = 10$, $\varepsilon_1 = 10^{-3}$, $\varepsilon_2 = 10^{-2}$.

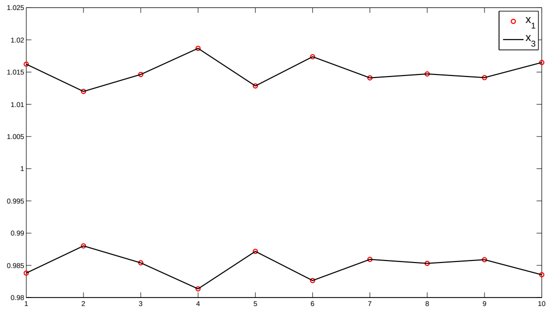

Figure 8. For $m = 20$, $n = 10$, $\varepsilon_1 = 10^{-3}$, $\varepsilon_2 = 10^{-2}$.

We note that in our experiments (see Example 4.2, 4.3) we got the same results for $\mathbf{x}_1$ and $\mathbf{x}_3$ using two different methods to solve equation (3.1). Namely, we applied Rohn's method and Gaussian elimination. It would be a very interesting question how we can characterize those systems where the equality holds.

References

[1] Alefeld, G., Herzberger, J., *Introduction to Interval Computations*, Academic Press, New York, 1983.

[2] Hansen, E.R., Smith, R.R., *Interval arithmetic in matrix computations, Part 2*, SIAM J. Numerical Analysis, **4**(1967), 1-9.

[3] Hansen, E.R., *Bounding the solution of interval linear equations*, SIAM J. Numerical Analysis, **29**(1992), 1493-1503.

[4] Ning, S., Kearfott, R.B., *A comparison of some methods for solving linear interval equations*, SIAM J. Numerical Analysis, **34**(1997), 1289-1305.

[5] Rohn, J., *Cheap and tight bounds: The recent result by E. Hansen can be made more efficient*, Interval Comput., **4**(1993), 13-21.

[6] Rohn, J., *Systems of linear interval equations*, Linear Algebra and Its Applications, **126**(1989), 39-78.

[7] Rohn, J., *Solvability of systems of linear interval equations*, SIAM J. Matrix Analysis Appl., **25**(2003), 237-245.

[8] Rohn, J., *An algorithm for computing the hull of the solution set of interval linear equations*, Linear Algebra and Its Applications, **435**(2011), 193-201.

[9] Rohn, J., *An algorithm for solving the absolute value equation*, Electronic Journal of Linear Algebra, **18**(2009), 589-599.

[10] Rohn, J., *An algorithm for solving the absolute value equation: An improvement*, Technical Report 1063, Institute of Computer Science, Academy of Sciences of the Czech Republic, Prague, January 2010.

[11] Bentbib, A.H., *Solving the full rank interval least squares problem*, Applied Numerical Mathematics, **41**(2002), 283-294.

[12] Alefeld, G., Mayer G., *The Cholesky method for interval data*, Linear Algebra and Its Applications, **194**(1993), 161-182.

Szilvia Huszárszky
Eötvös Loránd University
Faculty of Informatics, Department of Numerical Analysis
1/C, Pázmány Péter sétány, 1117 Budapest, Hungary
e-mail: szilvia.huszarszky@gmail.com

Lajos Gergó
Eötvös Loránd University
Faculty of Informatics, Department of Numerical Analysis
1/C, Pázmány Péter sétány, 1117 Budapest, Hungary
e-mail: gergo@inf.elte.hu