Reliable Computing (2005) 11: 129–135

Linear Interval Equations: Midpoint Preconditioning May Produce a 100% Overestimation for Arbitrarily Narrow Data Even in Case $n = 4$

JIŘÍ ROHN
Institute of Computer Science, Czech Academy of Sciences, Pod vodárenskou věží 2, 182 07 Prague, Czech Republic, e-mail: rohn@cs.cas.cz

(Received: 22 March 2004; accepted: 30 April 2004)

Abstract. We construct a linear interval system $\mathbf{A}x = \mathbf{b}$ with a 4×4 interval matrix whose all proper interval coefficients (there are also some noninterval ones) are of the form $[-\varepsilon, \varepsilon]$. It is proved that for each $\varepsilon > 0$, the interval hull $[\underline{x}, \overline{x}]$ and interval hull of the midpoint preconditioned system $[\underline{\underline{x}}, \overline{\overline{x}}]$ satisfy $\overline{x}_1 = 0.6$ and $\overline{\overline{x}}_1 = 1.2$, hence midpoint preconditioning produces a 100% overestimation of $\overline{x}_1$ independently of ε in this case. The example was obtained as a result of an extensive MATLAB search.

1. Introduction

Let $\mathbf{A} = [A_c - \Delta, A_c + \Delta]$ be an $n \times n$ interval matrix and $\mathbf{b} = [b_c - \delta, b_c + \delta]$ an interval n-vector. Solving the linear interval system

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

usually means computing the *interval hull* of the solution set

$$X = \{x;\ Ax = b \text{ for some } A \in \mathbf{A},\ b \in \mathbf{b}\},$$

i.e., the narrowest interval vector $[\underline{x}, \overline{x}]$ containing X. Such an interval vector exists (i.e., X is bounded) if and only if the interval matrix $\mathbf{A}$ is regular, which, by definition, means that each $A \in \mathbf{A}$ is nonsingular. The basic theoretical result concerning computation of the interval hull in the general case is contained in the following theorem proved in [6]:

THEOREM 1.1. *Let* $\mathbf{A}$ *be regular. Then for each* $y \in Y$ *the nonlinear equation*

$$A_c x - b_c = T_y(\Delta|x| + \delta) \tag{1.2}$$

has exactly one solution x_y *and for the interval hull* $[\underline{x}, \overline{x}]$ *of the solution set of (1.1) there holds*

$$\underline{x} = \min_{y \in Y} x_y,$$

$$\overline{x} = \max_{y \in Y} x_y$$

(componentwise).

Here, the following notation is used: for $x = (x_i)$, the absolute value is defined by $|x| = (|x_i|)$, $Y = \{y \in \mathbb{R}^n;\ y_j \in \{-1, 1\} \text{ for each } j\}$ is the set of all ± 1-vectors in $\mathbb{R}^n$, and

$$T_y = \operatorname{diag}(y_1, y_2, \dots, y_n) = \begin{pmatrix} y_1 & 0 & \dots & 0 \\ 0 & y_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & y_n \end{pmatrix}$$

for each $y \in Y$. Unfortunately, computing the interval hull $[\underline{x}, \overline{x}]$ is NP-hard [9]; we shall also obtain this result as a by-product of Theorem 2.1 here. Therefore, for practical needs we must set the goal differently: instead of computing the exact interval hull $[\underline{x}, \overline{x}]$, we strive for obtaining an *enclosure* of the solution set X, i.e., an interval vector $[\underset{\sim}{x}, \tilde{x}]$ satisfying $X \subseteq [\underset{\sim}{x}, \tilde{x}]$ and as narrow as (computationally) possible.

One of the basic techniques developed for this purpose as early as in the 1960's is the so-called midpoint preconditioning. It is based on a simple observation that if $x \in X$, i.e. if $Ax = b$ for some $A \in \mathbf{A}$ and $b \in \mathbf{b}$, then also $A_c^{-1}Ax = A_c^{-1}b$. Since the Oettli-Prager theorem [5] implies that

$$\begin{aligned} \{A_c^{-1}A;\ A \in \mathbf{A}\} &= [I - |A_c^{-1}|\Delta,\ I + |A_c^{-1}|\Delta], \\ \{A_c^{-1}b;\ b \in \mathbf{b}\} &= [A_c^{-1}b_c - |A_c^{-1}|\delta,\ A_c^{-1}b_c + |A_c^{-1}|\delta], \end{aligned}$$

we can see that the above x belongs to the solution set X_p of the preconditioned system

$$[I - |A_c^{-1}|\Delta,\ I + |A_c^{-1}|\Delta]x = [A_c^{-1}b_c - |A_c^{-1}|\delta,\ A_c^{-1}b_c + |A_c^{-1}|\delta], \tag{1.3}$$

in other words, $X \subseteq X_p$ holds. The interval matrix in (1.3) is regular if and only if

$$\varrho(|A_c^{-1}|\Delta) < 1 \tag{1.4}$$

holds (i.e., if $\mathbf{A}$ is strongly regular). Under this condition, X_p is bounded and its interval hull $[\underline{\underline{x}}, \overline{\overline{x}}]$, which we call the *preconditioned interval hull*, satisfies $X \subseteq X_p \subseteq [\underline{\underline{x}}, \overline{\overline{x}}]$, so that it forms an enclosure of X. The usefulness of this technique was demonstrated by Hansen [2], Bliek [1], and Rohn [7] who showed that the bounds $\underline{\underline{x}}$ and $\overline{\overline{x}}$ can be expressed by closed-form formulae involving only computations of two inverses and of several matrix-vector products (for other derivations of this result, see also Ning and Kearfott [4] and Neumaier [3]).

As a rule, the enclosure $[\underline{\underline{x}}, \overline{\overline{x}}]$ computed in this way gives fairly good results if the spectral radius in (1.4) is small. It is the main goal of this paper to demonstrate that it might be not always so, which implies that preconditioning cannot be considered a universal remedy. In Section 3 we shall show that for each $\varepsilon > 0$ we can construct a linear interval system (1.1) with a 4×4 interval matrix $\mathbf{A}$ and a right-hand side interval vector $\mathbf{b}$ such that the radii of all the coefficients of $\mathbf{A}$ and $\mathbf{b}$ are less or equal than ε and $\overline{x}_1 = 0.6$, $\overline{\overline{x}}_1 = 1.2$ independently of ε. This means that $\overline{\overline{x}}_1$ overestimates $\overline{x}_1$ by 100% for arbitrarily narrow data, so that midpoint preconditioning gives an utterly false result in this case.

In Section 2 we precede this example by a theoretical result in which formulae for the quantities $\overline{x}$ and $\overline{\overline{x}}$ are derived, thereby giving a closer insight into the structure of the example in question. The norms

$$\begin{aligned} \|x\|_1 &= e^T|x| = \sum_i |x_i|, \\ \|A\|_{\infty,1} &= \max_{\|x\|_\infty = 1} \|Ax\|_1 = \max_{y \in Y} \|Ay\|_1 \end{aligned} \tag{1.5}$$

are used there (see [8] for the last formula in (1.5)).

2. A Theoretical Example

Given a nonsingular matrix $A \in \mathbb{R}^{(n-1)\times(n-1)}$ and a real number $\varepsilon > 0$, consider a linear interval system

$$\mathbf{A}x = \mathbf{b}, \tag{2.1}$$

where

$$\mathbf{A} = \begin{pmatrix} \varepsilon^2 & [-\varepsilon e^T, \varepsilon e^T] \\ 0 & A \end{pmatrix} \tag{2.2}$$

and

$$\mathbf{b} = \begin{pmatrix} 0 \\ [-\varepsilon e, \varepsilon e] \end{pmatrix} \tag{2.3}$$

with $e = (1, 1, \ldots, 1)^T \in \mathbb{R}^{n-1}$. This means that the centers and radii are given by

$$\begin{aligned} A_c &= \begin{pmatrix} \varepsilon^2 & 0^T \\ 0 & A \end{pmatrix}, \qquad & \Delta &= \begin{pmatrix} 0 & \varepsilon e^T \\ 0 & 0 \end{pmatrix}, \\ b_c &= \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \qquad & \delta &= \begin{pmatrix} 0 \\ \varepsilon e \end{pmatrix}, \end{aligned}$$

which implies that

$$|A_c^{-1}|\Delta = \begin{pmatrix} \frac{1}{\varepsilon^2} & 0^T \\ 0 & |A^{-1}| \end{pmatrix} \begin{pmatrix} 0 & \varepsilon e^T \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{\varepsilon} e^T \\ 0 & 0 \end{pmatrix},$$

so that

$$\varrho(|A_c^{-1}|\Delta) = 0,$$

hence not only the interval matrix **A** is strongly regular, but also the spectral radius of $|A_c^{-1}|\Delta$ attains the lowest possible value independently of ε. Next, $A_c^{-1}b_c = 0$ and

$$|A_c^{-1}|\delta = \begin{pmatrix} 0 \\ \varepsilon|A^{-1}|e \end{pmatrix}.$$

Now we can state the basic result concerning the system (2.1)–(2.3):

THEOREM 2.1. *Let A be nonsingular and let* $\varepsilon > 0$. *Then for the interval hull* $[\underline{x}, \overline{x}]$ *and for the preconditioned interval hull* $[\underline{\underline{x}}, \overline{\overline{x}}]$ *of the system (2.1)–(2.3) we have*

$$\overline{x} = -\underline{x} = \begin{pmatrix} \| A^{-1} \|_{\infty,1} \\ \varepsilon d \end{pmatrix}, \tag{2.4}$$

$$\overline{\overline{x}} = -\underline{\underline{x}} = \begin{pmatrix} \| d \|_1 \\ \varepsilon d \end{pmatrix}, \tag{2.5}$$

where $d = |A^{-1}|e$.

Proof. (*a*) Since **b** is symmetric about 0, the same holds for the solution set X of (2.1)–(2.3) (because if $x \in X$, then $A'x = b'$ for some $A' \in \mathbf{A}$ and $b' \in \mathbf{b}$, hence $A'(-x) = -b' \in \mathbf{b}$ and $-x \in X$), which implies that $\underline{x} = -\overline{x}$. Thus we are confined to evaluate $\overline{x}$ only. According to Theorem 1.1, we have

$$\overline{x} = \max_{y \in Y} x_y,$$

where for each $y \in Y$, x_y is the unique solution of the equation (1.2). Let us write $y = (y_1, y'^T)^T$, where $y' = (y_2, \ldots, y_n)^T$, and let us decompose x_y accordingly as $x_y = (x_1, x'^T)^T$. Then the equation (1.2) for the system (2.1)–(2.3) has the form

$$\begin{pmatrix} \varepsilon^2 & 0^T \\ 0 & A \end{pmatrix} \begin{pmatrix} x_1 \\ x' \end{pmatrix} = T_{\begin{pmatrix} y_1 \\ y' \end{pmatrix}} \left(\begin{pmatrix} 0 & \varepsilon e^T \\ 0 & 0 \end{pmatrix} \left| \begin{matrix} x_1 \\ x' \end{matrix} \right| + \begin{pmatrix} 0 \\ \varepsilon e \end{pmatrix} \right)$$

or equivalently

$$\begin{aligned} \varepsilon^2 x_1 &= y_1 \varepsilon e^T |x'|, \\ Ax' &= T_{y'} \varepsilon e = \varepsilon y', \end{aligned}$$

which gives

$$\begin{aligned} x' &= \varepsilon A^{-1} y', \\ x_1 &= y_1 e^T |A^{-1} y'| = y_1 \| A^{-1} y' \|_1, \end{aligned}$$

hence

$$x_y = \begin{pmatrix} y_1\|A^{-1}y'\|_1 \\ \varepsilon A^{-1}y' \end{pmatrix},$$

and from Theorem 1.1 in view of (1.5) we obtain

$$\overline{x} = \max_{y\in Y} x_y = \begin{pmatrix} \|A^{-1}\|_{\infty,1} \\ \varepsilon|A^{-1}|e \end{pmatrix} = \begin{pmatrix} \|A^{-1}\|_{\infty,1} \\ \varepsilon d \end{pmatrix}.$$

(*b*) Since the right-hand side of the preconditioned system (1.3) is again symmetric about 0, we again have $\underline{\underline{x}} = -\overline{\overline{x}}$. The equation (1.2) for the preconditioned system (1.3), (2.2), (2.3) has the form

$$\begin{pmatrix} x_1 \\ x' \end{pmatrix} = T_{\begin{pmatrix} y_1 \\ y' \end{pmatrix}} \left(\begin{pmatrix} 0 & \frac{1}{\varepsilon}e^T \\ 0 & 0 \end{pmatrix} \left| \begin{matrix} x_1 \\ x' \end{matrix} \right| + \begin{pmatrix} 0 \\ \varepsilon|A^{-1}|e \end{pmatrix} \right),$$

which gives

$$x' = T_{y'}\varepsilon|A^{-1}|e = \varepsilon T_{y'}d,$$

$$x_1 = y_1\frac{1}{\varepsilon}e^T|x'| = y_1 e^T|A^{-1}|e = y_1\|d\|_1,$$

hence

$$x_y = \begin{pmatrix} y_1\|d\|_1 \\ \varepsilon T_{y'}d \end{pmatrix}$$

and

$$\overline{\overline{x}} = \max_{y\in Y} x_y = \begin{pmatrix} \|d\|_1 \\ \varepsilon d \end{pmatrix},$$

which concludes the proof. □

Now we can see the main point: the values of $\overline{x}_1 = \|A^{-1}\|_{\infty,1}$ and $\overline{\overline{x}}_1 = e^T|A^{-1}|e$ are *independent of* ε. To achieve the result wanted, it remains to choose an appropriate matrix A. But before doing so, we note that the formula (2.4) yields another proof of the NP-hardness of computing the interval hull (proved originally in [9]): since computing the norm $\|\cdot\|_{\infty,1}$ is NP-hard (as proved in [8]), by (2.4) computing $\overline{x}_1$, and thus also $[\underline{x}, \overline{x}]$, is NP-hard as well.

3. The Example

Consider the example (2.1)–(2.3) with

$$A = \begin{pmatrix} 1 & -3 & -3 \\ -3 & 1 & -3 \\ -3 & -3 & 1 \end{pmatrix}, \tag{3.1}$$

or, explicitly written,

$$\begin{pmatrix} \varepsilon^2 & [-\varepsilon,\varepsilon] & [-\varepsilon,\varepsilon] & [-\varepsilon,\varepsilon] \\ 0 & 1 & -3 & -3 \\ 0 & -3 & 1 & -3 \\ 0 & -3 & -3 & 1 \end{pmatrix} x = \begin{pmatrix} 0 \\ [-\varepsilon,\varepsilon] \\ [-\varepsilon,\varepsilon] \\ [-\varepsilon,\varepsilon] \end{pmatrix}. \tag{3.2}$$

As a direct application of Theorem 2.1 we obtain:

THEOREM 3.1. *For each* $\varepsilon > 0$*, for the interval hull* $[\underline{x}, \overline{x}]$ *and for the preconditioned interval hull* $[\underline{\underline{x}}, \overline{\overline{x}}]$ *of the linear interval system (3.2) we have*

$$\overline{x} = -\underline{x} = \begin{pmatrix} 0.6 \\ 0.4\varepsilon \\ 0.4\varepsilon \\ 0.4\varepsilon \end{pmatrix}, \tag{3.3}$$

$$\overline{\overline{x}} = -\underline{\underline{x}} = \begin{pmatrix} 1.2 \\ 0.4\varepsilon \\ 0.4\varepsilon \\ 0.4\varepsilon \end{pmatrix}. \tag{3.4}$$

Proof. We are left with substituting

$$A^{-1} = \begin{pmatrix} 0.10 & -0.15 & -0.15 \\ -0.15 & 0.10 & -0.15 \\ -0.15 & -0.15 & 0.10 \end{pmatrix}$$

into (2.4) and (2.5), using (1.5) for evaluation of $\|A^{-1}\|_{\infty,1}$, which yields (3.3) and (3.4). □

4. Concluding Remarks

We have proved that for the system (3.2) there holds $\overline{\overline{x}}_1 = 2\overline{x}_1 = 1.2$ independently of ε, thereby justifying the statement made in the title of the paper. The matrix A in (3.1), although being of quite regular structure at a glance, was in fact found through extensive experiencing in MATLAB involving computation of several tens of thousands of randomly generated examples (of size 3×3) aimed at maximizing the value of

$$\frac{e^T|A^{-1}|e}{\|A^{-1}\|_{\infty,1}}. \tag{4.1}$$

For the best result found the ratio was slightly less than 2 and the coefficients of A were close to integers; then rounding to nearest integers produced the matrix (3.1) for which the value of (4.1) is 2. However, notice from (3.3), (3.4) that $\underline{\underline{x}}_i = \underline{x}_i$,

$\overline{\overline{x}}_i = \overline{x}_i$ for $i \geq 2$. Theorem 2.1 may yield another related results, but we have not pursued the matter any further.

Finally we would like to emphasize that the negative result presented here should not shed bad light on the idea of midpoint preconditioning as a whole. It only warns us that one should take the result with some care. A possible remedy could consist in computation of nonnegative vectors $\underline{d}, \overline{d}$ satisfying

$$\begin{aligned} \overline{\overline{x}} - \overline{d} &\leq \overline{x} \leq \overline{\overline{x}}, \\ \underline{\underline{x}} &\leq \underline{x} \leq \underline{\underline{x}} + \underline{d}, \end{aligned}$$

wherefrom one could recognize the amount of overestimation—an idea that seems to be worth studying.

References

1. Bliek, C.: *Computer Methods for Design Automation*, Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, 1992.
2. Hansen, E. R.: Bounding the Solution of Interval Linear Equations, *SIAM Journal on Numerical Analysis* **29** (1992), pp. 1493–1503.
3. Neumaier, A.: A Simple Derivation of the Hansen-Bliek-Rohn-Ning-Kearfott Enclosure for Linear Interval Equations, *Reliable Computing* **5** (2) (1999), pp. 131–136.
4. Ning, S. and Kearfott, R. B.: A Comparison of Some Methods for Solving Linear Interval Equations, *SIAM Journal on Numerical Analysis* **34** (1997), pp. 1289–1305.
5. Oettli, W. and Prager, W.: Compatibility of Approximate Solution of Linear Equations with Given Error Bounds for Coefficients and Right-Hand Sides, *Numerische Mathematik* **6** (1964), pp. 405–409.
6. Rohn, J.: Systems of Linear Interval Equations, *Linear Algebra and Its Applications* **126** (1989), pp. 39–78.
7. Rohn, J.: Cheap and Tight Bounds: The Recent Result by E. Hansen Can Be Made More Efficient, *Interval Computations* 4 (1993), pp. 13–21.
8. Rohn, J.: Computing the Norm $\|A\|_{\infty,1}$ Is NP-Hard, *Linear and Multilinear Algebra* **47** (2000), pp. 195–204.
9. Rohn, J. and Kreinovich, V.: Computing Exact Componentwise Bounds on Solutions of Linear Systems with Interval Data Is NP-Hard, *SIAM Journal on Matrix Analysis and Applications* **16** (1995), pp. 415–420.