On Overestimations Produced by the Interval Gaussian Algorithm

Dedicated to Prof. Dr. Gerhard Heindl on the occasion of his 60th birthday

JIŘÍ ROHN
Faculty of Mathematics and Physics, Charles University, Prague, Czech Republic, and Institute of Computer Science, Academy of Sciences, Prague, Czech Republic, e-mail: rohn@uivt.cas.cz

Abstract. Interval Gaussian algorithm is a popular method for enclosing solutions of linear interval equations. In this note we show that both versions of the method (with or without preconditioning) may yield large overestimations for arbitrarily small data widths even in case $n = 4$.

1. Introduction

As is well known, interval Gaussian algorithm for enclosing solutions of a system of linear interval equations

$$A^I x = b^I \tag{1}$$

(A^I square) consists in solving the system (1) by Gaussian algorithm performed in interval arithmetic (i.e., with real arithmetic operations being replaced by interval ones; see Alefeld and Herzberger [1] or Neumaier [3]). As in the real case, it is recommendable to use a partial pivoting strategy, i.e., among all interval coefficients eligible for pivot to choose that one having the greatest absolute value. If both the forward and backward step of the algorithm can be performed, then the algorithm (called feasible in this case) yields an enclosure of the solution set of (1). However, the algorithm may break down at some step if all the interval coefficients eligible for pivot contain zeros; such a situation may occur even if the interval matrix A^I is regular (Reichmann [4]).

In an attempt to enhance feasibility, Hansen and Smith [2] proposed to apply interval Gaussian algorithm to the preconditioned system

$$(A_c^{-1} \odot A^I)x = A_c^{-1} \odot b^I, \tag{2}$$

where A_c is the midpoint matrix of A^I and "$\odot$" denotes the usual matrix multiplication carried out in interval arithmetic. The idea behind preconditioning is quite simple: since all diagonal coefficients of $A_c^{-1} \odot A^I$ contain 1 whereas all nondiagonal ones contain 0, this structure, provided $A_c^{-1} \odot A^I$ is sufficiently narrow, may enforce feasibility of the interval Gaussian algorithm when applied to (2). Even more, the enclosure computed in this way is often tighter than that one obtained by direct application of the algorithm to the original system (1). Therefore, the preconditioned form of the interval Gaussian algorithm is usually preferred and has become a popular method for solving linear interval equations.

However, it turns out that the method, even if feasible, may be far from being optimal. In this note we show that already in the case $n = 4$ there exist examples with arbitrarily small data widths, arbitrarily large absolute overestimations, and with relative overestimations arbitrarily close to $\frac{1}{2}$ in the preconditioned case and arbitrarily close to 2 if preconditioning is not applied. This shows that interval Gaussian algorithm, in both its forms, may yield quite unsatisfactory results.

2. The example

For $\varepsilon > 0, \alpha > 0$ and $\beta > 1$, consider a linear interval system of the form

$$\begin{pmatrix} \frac{\varepsilon^2}{\alpha} & [-\varepsilon,\varepsilon] & [-\varepsilon,\varepsilon] & [-\varepsilon,\varepsilon] \\ 0 & \beta & 1 & 1 \\ 0 & 1 & \beta & 1 \\ 0 & 1 & 1 & \beta \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ [-\varepsilon,\varepsilon] \\ [-\varepsilon,\varepsilon] \\ [-\varepsilon,\varepsilon] \end{pmatrix}. \tag{3}$$

It comprises three parameters: ε is the radius of interval coefficients, whereas α and β will have to do with absolute and relative overestimations, respectively. Preconditioning (3) by the midpoint inverse

$$A_c^{-1} = \frac{1}{(\beta-1)(\beta+2)} \begin{pmatrix} \frac{(\beta-1)(\beta+2)\alpha}{\varepsilon^2} & 0 & 0 & 0 \\ 0 & \beta+1 & -1 & -1 \\ 0 & -1 & \beta+1 & -1 \\ 0 & -1 & -1 & \beta+1 \end{pmatrix}$$

yields the system

$$\begin{pmatrix} 1 & [-\frac{\alpha}{\varepsilon},\frac{\alpha}{\varepsilon}] & [-\frac{\alpha}{\varepsilon},\frac{\alpha}{\varepsilon}] & [-\frac{\alpha}{\varepsilon},\frac{\alpha}{\varepsilon}] \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \frac{(\beta+3)\varepsilon}{(\beta-1)(\beta+2)} \begin{pmatrix} 0 \\ [-1,1] \\ [-1,1] \\ [-1,1] \end{pmatrix} \tag{4}$$

in the form (2). The interval matrix A^I of (3), if written as $A^I = [A_c - \Delta, A_c + \Delta]$, satisfies

$$|A_c^{-1}|\Delta = \frac{\alpha}{\varepsilon}\begin{pmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

hence

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

(where for $A = (a_{ij})$ the absolute value is defined by $|A| = (|a_{ij}|)$, and $\varrho(A)$ denotes the spectral radius of A), so that A^I is strongly regular [3], with lowest possible value of spectral radius. Yet it turns out that neither the system (3), nor its preconditioned form (4) is well suited for solving by the interval Gaussian algorithm.

3. The result

First we derive some explicit formulae.

THEOREM 1. *For each $\varepsilon > 0$, $\alpha > 0$ and $\beta > 1$ the system (3) satisfies:*

(i) the exact upper bound on x_1 is

$$\overline{x}_1 = \frac{(3\beta+5)\alpha}{(\beta-1)(\beta+2)}, \tag{5}$$

(ii) the upper bound on x_1 computed by preconditioned interval Gaussian algorithm (with or without partial pivoting) is

$$\overline{\overline{x}}_1 = \frac{(3\beta+9)\alpha}{(\beta-1)(\beta+2)} \tag{6}$$

(and is equal to the exact upper bound on x_1 for the preconditioned system (4)),

(iii) the upper bound on x_1 computed by interval Gaussian algorithm without preconditioning (with or without partial pivoting) is

$$\overline{\overline{\overline{x}}}_1 = \frac{(3\beta^2+3\beta+2)\alpha}{\beta^2(\beta-1)}. \tag{7}$$

Proof. (i) Denote

$$A = \begin{pmatrix} \beta & 1 & 1 \\ 1 & \beta & 1 \\ 1 & 1 & \beta \end{pmatrix},$$

then A is nonsingular for $\beta > 1$ and satisfies

$$A^{-1} = \frac{1}{(\beta-1)(\beta+2)} \begin{pmatrix} \beta+1 & -1 & -1 \\ -1 & \beta+1 & -1 \\ -1 & -1 & \beta+1 \end{pmatrix}. \tag{8}$$

From the form of (3) it can be easily seen that for the exact bound on x_1 we have (with $e = (1,1,1)^T$ and $\|x\|_1 = e^T|x| = \sum_i |x_i|$)

$$\begin{aligned} \overline{x}_1 &= \frac{\alpha}{\varepsilon^2} \max\{\varepsilon e^T|x|;\ -\varepsilon e \le Ax \le \varepsilon e\} \\ &= \alpha \max\{\|A^{-1}x'\|_1;\ -e \le x' \le e\}, \end{aligned}$$

which in view of convexity of the norm implies

$$\overline{x}_1 = \alpha \max\{\|A^{-1}x'\|_1;\ |x'| = e\}. \tag{9}$$

Since A^{-1} is diagonally dominant for $\beta > 1$, for each x' satisfying $|x'| = e$ (i.e., a ± 1-vector) there holds $(A^{-1}x')_i > 0$ if $x'_i = 1$ and $(A^{-1}x')_i < 0$ if $x'_i = -1$, hence

$$\begin{aligned} \|A^{-1}x'\|_1 &= \sum_i |A^{-1}x'|_i = \sum_i x'_i (A^{-1}x')_i = (x')^T A^{-1} x' \\ &= \frac{3(\beta+2) - (e^T x')^2}{(\beta-1)(\beta+2)} \le \frac{3\beta+5}{(\beta-1)(\beta+2)} \end{aligned} \tag{10}$$

since $(e^T x')^2 \ge 1$ (x' is a ± 1-vector), and $(e^T x')^2 = 1$ e.g. for $x' = (1,1,-1)^T$, hence the bound is achieved and (9), (10) imply (5).

(ii) Interval Gaussian algorithm applied to the preconditioned system (4) (which consists of the backward step only) gives the upper bound (6), which is also the exact upper bound on x_1 for (4) since it is achieved at the system

$$\begin{pmatrix} 1 & -\frac{\alpha}{\varepsilon} & -\frac{\alpha}{\varepsilon} & -\frac{\alpha}{\varepsilon} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \frac{(\beta+3)\varepsilon}{(\beta-1)(\beta+2)} \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

(iii) The forward step of the interval Gaussian algorithm without preconditioning (which is the same with or without partial pivoting due to $\beta > 1$) results in the system

$$\begin{pmatrix} \frac{\varepsilon^2}{\alpha} & [-\varepsilon,\varepsilon] & [-\varepsilon,\varepsilon] & [-\varepsilon,\varepsilon] \\ 0 & \beta & 1 & 1 \\ 0 & 0 & \frac{\beta^2-1}{\beta} & \frac{\beta-1}{\beta} \\ 0 & 0 & 0 & \frac{(\beta-1)(\beta+2)}{\beta+1} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ [-\varepsilon,\varepsilon] \\ [-\frac{(\beta+1)\varepsilon}{\beta}, \frac{(\beta+1)\varepsilon}{\beta}] \\ [-\frac{(\beta+2)\varepsilon}{\beta}, \frac{(\beta+2)\varepsilon}{\beta}] \end{pmatrix},$$

and the backward step gives (7). □

Notice that the values of $\overline{x}_1$, $\overline{\overline{x}}_1$, $\overline{\overline{\overline{x}}}_1$ do not depend on ε. As a result, we obtain:

THEOREM 2. *For arbitrary data width $\varepsilon > 0$, for the system (3) we have:*

(i) for each $r \in (0, \frac{1}{2})$ and $K > 0$ there exist $\alpha > 0$ and $\beta > 1$ such that

$$\frac{\overline{\overline{x}}_1 - \overline{x}_1}{\overline{x}_1} > r \tag{11}$$

and

$$\overline{\overline{x}}_1 - \overline{x}_1 > K, \tag{12}$$

(ii) for each $r \in (0, 2)$ and $K > 0$ there exist $\alpha > 0$ and $\beta > 1$ such that

$$\frac{\overline{\overline{\overline{x}}}_1 - \overline{x}_1}{\overline{x}_1} > r$$

and

$$\overline{\overline{\overline{x}}}_1 - \overline{x}_1 > K.$$

Proof. (i) According to (5) and (6),

$$\lim_{\beta \to 1_+} \frac{\overline{\overline{x}}_1 - \overline{x}_1}{\overline{x}_1} = \lim_{\beta \to 1_+} \frac{4}{3\beta + 5} = \frac{1}{2},$$

hence for each $r \in (0, \frac{1}{2})$ there exists a $\beta > 1$ such that (11) holds, and a choice of

$$\alpha > \frac{1}{4}(\beta - 1)(\beta + 2)K$$

assures (12) to hold.

The proof of (ii) is quite analogous since

$$\lim_{\beta \to 1_+} \frac{\overline{\overline{\overline{x}}}_1 - \overline{x}_1}{\overline{x}_1} = \lim_{\beta \to 1_+} \frac{4\beta^2 + 8\beta + 4}{3\beta^3 + 5\beta^2} = 2$$

and the derivative of the rational function is negative at $\beta = 1$. □

4. Concluding remarks

We have shown that overestimations (11) and (12) may occur for systems of the form (3). But since $\overline{\overline{x}}_1$ is the *exact* upper bound on the solution of the preconditioned system (4) (Theorem 1, (ii)), the result of Theorem 2, (i) also holds true for *any* method based on solving a preconditioned linear interval system (2), as e.g. Rump's method in [5]. Hence, preconditioning may yield quite unsatisfactory results for systems with arbitrarily small data widths even in case $n = 4$. It is not known to the author whether a similar result may be proved for lower dimensions.

Acknowledgements

This work was supported by the Czech Republic Grant Agency under grant GAČR 201/95/1484.

References

1. G. Alefeld and J. Herzberger. *Introduction to Interval Computations.* Academic Press, New York, 1983.
2. E. Hansen and R. Smith. Interval arithmetic in matrix computations, Part II. *SIAM Journal on Numerical Analysis*, 4:1–9, 1967.
3. A. Neumaier. *Interval Methods for Systems of Equations.* Cambridge University Press, Cambridge, 1990.
4. K. Reichmann. Abbruch beim Intervall-Gauss-Algorithmus. *Computing*, 22:355–361, 1979.
5. S. M. Rump. Solving algebraic problems with high accuracy. In U. Kulisch and W. Miranker, editors, *A New Approach to Scientific Computation*, pages 51–120, New York, 1983. Academic Press.