SOLVING SYSTEMS OF LINEAR INTERVAL EQUATIONS

by

J. Rohn, Prague

Abstract. This paper is a short survey of methods for computing bounds on solutions of a system of linear equations with square matrix whose coefficients as well as the right-hand side components are given by real intervals.

0. Introduction

In this paper we give a short survey of results on computing the exact bounds for components of solutions of a system of n linear equations in n variables whose coefficients and right-hand sides are prescribed by real intervals (obtained as a result of rounding off, truncating or data errors). We are primarily interested in methods for computing the exact bounds on solutions. There are many good methods for computing sufficiently close outer estimations, which we do not survey here; an interested reader is referred to monograhs by Alefeld and Herzberger [1] and Deif [7] and to the survey paper by Neumaier [12]. We omit here the proofs which can be found in [21] and [22], or in preprints [17], [18], [19].

In section 1 we sum up the basic theoretical results and show that the minimum (maximum) component values can be computed when taking into account only a finite number of vectors. Methods for computing these vectors are surveyed in section 2, while a special case in which the bounds can be expressed explicitly is handled in section 3. An application of these results to the problem of evaluating the exact bounds for coefficients of the inverse interval matrix is given in section 4.

Basic notation: coefficients of a matrix A are denoted by A_{ij}. If $A = (A_{ij})$, then the absolute value of A is defined by $|A| = (|A_{ij}|)$. The inequalities $A \geq 0$ $(A > 0)$ are to be understood componentwise, A^T denotes the transpose matrix. The same notations also apply to vectors.

1. Bounding the solutions

Let $A^I = \{A;\ A_c - \Delta \leq A \leq A_c + \Delta\}$ be an n x n interval matrix and $b^I = \{b;\ b_c - \delta \leq b \leq b_c + \delta\}$ an interval n-vector $(\Delta \leq 0,\ \delta \leq 0)$. For the system of linear interval equations $A^I x = b^I$, the solution set X is defined by

$$X = \{x;\ Ax = b,\ A \in A^I,\ b \in b^I\}.$$

Throughout the paper, we shall be interested in methods for computing the vectors $\underline{x} = (\underline{x}_i)$, $\overline{x} = (\overline{x}_i)$ defined by

$$\underline{x}_i = \min\{x_i;\ x \in X\}$$
$$\overline{x}_i = \max\{x_i;\ x \in X\} \qquad (i = 1, \ldots, n), \quad (1.1)$$

ISBN 0-12-505630-3

giving the exact bounds for the components of solution vectors. If A^I is regular (which means that each $A \in A^I$ is regular), then X is compact (Beeck [3]), so that $\underline{x}_i$, x_i (i = 1, ..., n) are well defined. Testing regularity of A^I is generally a difficult problem (see [21], section 5). Fortunately, the sufficient regularity condition (Beeck [5])

$$\rho(D) < 1, \qquad (1.2)$$

where ρ denotes the spectral radius and $D = |A_c^{-1}|\Delta$, usually works in practical examples. Interval matrices satisfying (1.2) are called strongly regular (Neumaier [12]). We shall first set out to describe the solution set X. The following basic result is due to Oettli and Prager [16]:

Theorem 1.1. We have

$$X = \{x;\ |A_c x - b_c| \le \Delta|x| + \delta\}.$$

The solution set is generally nonconvex; for example, see [2], [8], [14]. However, the intersection of X with each orthant is a convex polyhedron. To see this, define for each $x \in X$ its signature vector $\operatorname{sgn} x \in R^n$ by $(\operatorname{sgn} x)_i = 1$ if $x_i \ge 0$ and $(\operatorname{sgn} x)_i = -1$ otherwise, and let T_z denote the diagonal matrix with diagonal vector z. Then for $z = \operatorname{sgn} x$, we have $|x| = T_z x$, hence the intersection of X with the orthant $R_z^n = \{x \in R^n;\ T_z x \ge 0\}$ is according to theorem 1.1 given by

$$(A_c - \Delta T_z)x \le b_c + \delta$$

$$(A_c + \Delta T_z)x \ge b_c - \delta$$

$$T_z x \ge 0.$$

Oettli [15] therefore proposed using a linear programming procedure in each orthant to compute $\underline{x}_i$, $\overline{x}_i$, a method later investigated also by Cope and Rust [6]. The necessity of solving a number of linear programming problems ($n2^{n+1}$ in the worst case) makes this approach generally disadvantageous.

Another method, proposed in [17], [18], [21], is based on this theorem (notation : $Y = \{y \in R^n;\ |y_j| = 1 \text{ for each } j\}$) :

Theorem 1.2. Let A^I be regular. Then for each $y \in Y$, the nonlinear equation

$$A_c x - b_c = T_y(\Delta|x| + \delta) \qquad (1.3)$$

has a unique solution $x_y \in X$ and there holds

$$\text{Conv } X = \text{Conv}\{x_y;\ y \in Y\}.$$

The proof of this theorem reveals an unsuspected conneciton of our problem with the linear complementarity theory; an interested reader is referred to [211]. Since $|T_y| = E$ = a unit matrix for each $y \in Y$, the x_y's are just all solutions of the nonlinear equation $|A_c x - b_c| = \Delta|x| + \delta$. Methods for computing the x_y's will be described in seciton 2. Since Conv X is a convex polyhedron, each minimum (maximum) in (1.1) must be achieved at some vertex of Conv X, i.e., in view of theorem 1.2, at some x_y. In this way we obtain formulae

$$\underline{x}_i = \min\{(x_y)_i\ ;\ y \in Y\}$$
$$\overline{x}_i = \max\{(x_y)_i\ ;\ y \in Y\} \qquad (i = 1, ..., n)$$

involving only a finite number of vectors. In the most unfavorable case, computation of all 2^n vectors x_y may be inescapable. However, there exists a class of regular interval matrices for which the number of x_y's to be computed can be reduced down to at most 2n. A regular interval matrix A^I is called inverse stable if for each $i, j \in \{1, ..., n\}$, either $A_{ij}^{-1} \geq 0$ for each $A \in A^I$, or $A_{ij}^{-1} \leq 0$ for each $A \in A^I$ holds. For such an A^I, we may define vectors $y(i) \in Y$ $(i = 1, ..., n)$ by

$$(y(i))_j = \begin{cases} 1 & \text{if } A_{ij}^{-1} \geq 0 \text{ for each } A \in A^I \\ -1 & \text{otherwise} \qquad (j=1, ..., n) \end{cases}$$

Then we have

Theorem 1.3)[21]). Let A^I be inverse stable. Then there holds

$$\underline{x}_i = (x_{-y(i)})_i$$
$$\overline{x}_i = (x_{y(i)})_i \qquad (i = 1, ..., n).$$

Hence at most 2n vectors x_y are to be computed. The inequality

$$C\,|A_c^{-1}| \leq |A_c^{-1}|\Delta, \qquad (1.4)$$

where $C = D(E-D)^{-1}$ (with $D = |A_c^{-1}|\Delta$ as before) is a sufficient inverse stability condition ([21]), recommended for use when solving practical examples, where Δ is usually of small norm and inverse stability often occurs.

In the special case of interval matrices satisfying

$$T_z A^{-1} T_y \geq 0 \text{ for each } A \in A^I \tag{1.5}$$

(i.e., $A_{ij}^{-1} \geq 0$ if $z_i y_j = 1$ and $A_{ij}^{-1} \leq 0$ if $z_i y_j = -1$) for some fixed $z, y \in Y$, we have $y(i) = y$ if $z_i = 1$ and $y(i) = -y$ if $z_i = -1$, hence

$$\begin{aligned} \underline{x}_i &= \min\{(x_y)_i, (x_{-y})_i\} \\ \overline{x}_i &= \max\{(x_y)_i, (x_{-y})_i\} \end{aligned} \qquad (i = 1, \ldots, n).$$

If A^I is inverse nonnegative (by definition, $A^{-1} \geq 0$ for each $A \in A^I$; holds iff $(A_c - \Delta)^{-1} \geq 0$ and $(A_c + \Delta)^{-1} \geq 0$, see [10], [20]), then (1.5) is satisfied with $z = y = e$, where $e = (1, 1, \ldots, 1)^T$, and we have $\underline{x} = x_{-e}$, $\overline{x} = x_e$, a result obtained by Beeck [5]. If, moreover, $(A_c + \Delta)^{-1} (b_c - \delta) \geq 0$ holds, then the bounds may be expressed explicitly by

$$\underline{x} = (A_c + \Delta)^{-1} (b_c - \delta)$$

$$\overline{x} = (A_c - \Delta)^{-1} (b_c + \delta)$$

(see [20]; for special cases, Barth and Nuding [2], Beeck [4]).

2. Computing the x_y's

As stated in theorem 1.2, for each $y \in Y$, x_y is the unique solution of the equation

$$A_c x - b_c = T_y(\Delta|x| + \delta) . \tag{2.1}$$

We shall first describe a general method for computing x_y. Set $z = \operatorname{sgn} x$ and denote

$$A_{yz} = A_c - T_y \Delta T_z$$

$$b_y = b_c + T_y \delta ,$$

then (2.1) can be equivalently written as

$$\begin{aligned} A_{yz} x &= b_y \\ T_z x &\geq 0 . \end{aligned} \tag{2.2}$$

The following algorithm for solving (2.2) is a modification of Murty's algorithm [11] for solving the associated linear complementarity problem:

Algorithm ([21]).

0. Select a $z \in Y$ (recommended : $z = \mathrm{sgn}(A_c^{-1}b_y)$).
1. Solve $A_{yz}x = b_y$.
2. If $T_z x \geq 0$, set $x_y := x$ and terminate.
3. Otherwise find
   $k = \min\{j;\ z_j x_j < 0\}$.
4. Set $z_k := -z_k$ and go to step 1.

Theorem 2.1 ([21]). Let A^I be regular. Then the algorithm is finite for each $y \in Y$ and for an arbitrary starting $z \in Y$ in step 0.

If all the coefficients of A_c^{-1} are nonzero, if Δ and δ are sufficiently small in norm and if the algorithm is started in step 0 as recommended, then termination occurs when passing for the first time through step 2. Otherwise, especially if started improperly in step 0, the algorithm may solve up to 2^n linear systems to find x_y (for an example, see [21]). Therefore the algorithm, although general, may be found inappropriate in practical computations.

An iterative method for computing x_y may be constructed when observing that (2.1) can be rearranged to an equivalent fixed-point equation

$$x = D_y|x| + d_y \qquad (2.3)$$

where $D_y = A_c^{-1}T_y\Delta$, $d_y = A_c^{-1}b_y$. To solve (2.3), we may employ either Jacobi iterations

$$x_y^o = d_y$$

$$x_y^{k+1} = D_y|x_y^k| + d_y \quad (k = 0, 1, \dots)$$

or Gauss-Seidel iterations

$$\tilde{x}_y^o = d_y$$

$$\tilde{x}_y^{k+1} = L_y|\tilde{x}_y^{k+1}| + Q_y|\tilde{x}_y^k| + d_y \qquad (k = 0, 1, \dots),$$

where $D_y = L_y + Q_y$ is a triangular decomposition of D_y, with L_y having zero diagonal entries. If A^I is strongly regular, then $x_y^k \to x_y$, $\tilde{x}_y^k \to x_y$ [21]. Since

$$|x_y - x_y^k| \leq C|x_y^k - x_y^{k-1}| \leq CD^k|d_y|$$

for each $k \geq 1$ (similarly for x_y^k), iterative methods are suitable for problems with small values of $\rho(D)$.

As pointed out to the author by Neumaier [13], one may avoid computing the exact inverse A_c^{-1} (required in (2.3)) when using an approximate inverse B and employing Krawczyk [9] iterations

$$\hat{x}_y^{o} = d_y$$

$$\hat{x}_y^{k+1} = (E - BA_c)\hat{x}_y^{k} + BT_y\Delta|\hat{x}_y^{k}| + Bb_y \qquad (k = 0, 1, \dots)$$

which converge to x_y provided

$$\rho(|B|\Delta + |E - BA_c|) < 1$$

holds, a condition which is satisfied if A^I is strongly regular and B is a sufficiently close approximation of A_c^{-1}. Obviously, also a Gauss-Seidel version of Krawczyk iterations may be given.

Consider now an important special class of regular interval matrices satisfying

$$\Delta = qp^T \qquad (2.4)$$

for some nonnegative (column) vectors q, p (i.e., if $q \neq 0$ and $p \neq 0$, then Δ is of rank one). Assume, moreover, that q and p are so small that the whole solution set X lies in a single orthant; as proved in [22], this is the case if the inequality

$$p^T(|x_c| + \bar{\delta})\bar{q} + (1 - p^T\bar{q})\bar{\delta} + (p^T\bar{q})|x_c| < |x_c| \qquad (2.5)$$

holds, where we have denoted

$$x_c = A_c^{-1}b_c$$
$$\bar{q} = |A_c^{-1}|\, q$$
$$\bar{\delta} = |A_c^{-1}|\delta.$$

Now, using $z = \operatorname{sgn} x_c$, we have $|x_y| = T_z x_y$ for each $y \in Y$ and from (2.3) we obtain

$$x_y = x_c + A_c^{-1}T_y\delta + \alpha A_c^{-1}T_y q$$

where $\alpha = p^T T_z x_y$. Premultiplying the above equation by $p^T T_z$, computing α and substituting back, we get

$$x_y = x_c + A_c^{-1}T_y\delta + \frac{p^T|x_c| + p_T T_z A_c^{-1}T_y\delta}{1 - p_T T_z A_c^{-1}T_y q}(A_c^{-1}T_y q) \qquad (2.6)$$

In the special case of $\delta = \gamma q$ for some real $\gamma > 0$, (2.6) simplifies to

$$x_y = x_c + \frac{p^T|x_c| + \gamma}{1 - p^T T_z A_c^{-1} T_y q} (A_c^{-1} T_y q) \qquad (2.7)$$

Some applications of (2.6), (2.7) are given in [22].

3. Explicit formulae for $\underline{x}$, $\bar{x}$

For inverse stable interval matrices with radius Δ of the form qp^T, we may use formulae (2.6) for x_y derived at the end of the preceding section in conjunction with theorem 1.3 to obtain explicit formulae for $\underline{x}_i$, $\bar{x}_i$.

Denote $\bar{p}^T = p^T|A_c^{-1}|$. Then the sufficient inverse stability condition (1.4) has the form

$$\bar{q}\bar{p}^T + (p^T\bar{q})\,|A_c^{-1}| < |A_c^{-1}|. \qquad (3.1)$$

Further, for each $i \in \{1, ..., n\}$ denote

$$\lambda_i = p^T T_z A_c^{-1} T_{y(i)} q$$

$$\mu_i = p^T T_z A_c^{-1} T_{y(i)} \delta$$

where, as before, $z = \text{sgn}\, x_c$ and $y(i)$ is the signature vector of the i-th row of A_c^{-1}. Then there holds

Theorem 3.1 ([22]). Let A^I, b^I satisfy (2.4), (2.5), (3.1). Then for each $i \in \{1, ..., n\}$ we have

$$\underline{x}_i = (x_c)_i - \bar{\delta}_i - \frac{(p^T|x_c| - \mu_i)\bar{q}_i}{1 + \lambda_i}$$

$$\bar{x}_i = (x_c)_i + \bar{\delta}_i + \frac{(p^T|x_c| + \mu_i)\bar{q}_i}{1 - \lambda_i}$$

As a special case, consider linear interval systems $A^I x = b^I$ satisfying

$$\Delta_{ij} = \beta = \text{const}$$
$$\delta_i = \gamma = \text{const}$$

for each i, j; this corresponds to the above situation with $q = e$, $p = \beta e$, $\delta = \gamma e$. Introducing

$$r = |A_c^{-1}|e$$
$$s^T = e^T|A_c^{-1}|$$
$$v_i = z^T A_c^{-1} y(i),$$

we may reformulate (2.5), (3.1) as

$$\beta[\|x_c\| \, r + \|r\| \, |x_c|] + \gamma_r < |x_c| \qquad (3.21)$$
$$\beta[rs^T + \|r\| \, |A_c^{-1}|\,] < |A_c^{-1}|$$

where we used the norm $\|x\| = \|x\|_1 = \Sigma_i |x_i|$. Then the formulae for $\underline{x}_i$, $\bar{x}_i$ take on this simple form ([22]):

$$\underline{x}_i = (x_c)_i - \frac{(\beta \|x_c\| + \gamma) r_i}{1 + \beta v_i}$$

$$\bar{x}_i = (x_c)_i + \frac{(\beta \|x_c\| + \gamma) r_i}{1 - \beta v_i} \qquad (3.3)$$

($i = 1, \ldots, n$). These formulae have a number of consequences. We shall mention here only one of them : loss of significant decimals due to data rounding.

Assume that both left- and right-hand side coefficients of a system of linear equations $Ax = b$ have been rounded off down to π decimals, giving a system $A_c x_c = b_c$. For each i, we will be looking for the maximal integer ε_i satisfying

$$|x_i - (x_c)_i| \leq 5 \times 10^{-(\varepsilon_i + 1)}$$

if nothing more is known of A, b but that A_c, b_c are their rounded-off values up to π decimals. Applying formulae (3.3) with $\beta = \gamma = 5 \times 10^{-(\pi+1)}$, we get

$$\varepsilon_i = \pi - \lfloor \log_{10} \frac{(\|x_c\| + 1)\, r_i}{1 - \beta\, |v_i|} \rfloor_o$$

where we denoted $[a]_o = \min\{k;\ a \le k,\ k \text{ integer}\}$. Hence if $\pi \to \infty$, then the value of $\pi - \varepsilon_i$, representing the loss of significant decimals, tends to a finite value σ_i.

In none of the numbers $\log_{10}((\|x\| + 1)\bar{r}_i)$ is integer, $\bar{r} = |A^{-1}|e$ which is probably the case in practical computations, then for

$$\sigma = \max_i \sigma_i$$

we obtain this expression in terms of x and A^{-1} :

$$\sigma = \lceil \log_{10}((\|x\| + 1)\, \|A^{-1}\|_\infty) \rceil .$$

Another consequence of (3.3) may be found in [22] .

4. Inverse interval matrix

For a regular interval matrix A^I, the inverse interval matrix $B^I = [\underline{B}, \overline{B}]$ is defined by

$$\underline{B}_{ij} = \min\{A_{ij}^{-1}\ ;\ A \in A^I\}$$

$$\overline{B}_{ij} = \max\{A_{ij}^{-1}\ ;\ A \in A^I\}$$

(i, j, = 1, ..., n). The following theorem shows a general method for computing B^I, together with a necessary and sufficient regularity condition:

Theorem 4.1 ([21]). Let A_c be regular. Then A^I is regular if and only if for each $y \in Y$, the matrix equation

$$B = D_y |B| + A_c^{-1} \tag{4.1}$$

has a unique solution B_y. In this case, there holds

$$\underline{B}_{ij} = \min\{(B_y)_{ij};\ y \in Y\}$$

$$\overline{B}_{ij} = \max\{(B_y)_{ij};\ y \in Y\}$$

(i, j, = 1, ... , n). If A^I is inverse stable, then in view of theorem 1.3 we have

$$\underline{B}_{ij} = (B_{-y(i)})_{ij}$$

$$\overline{B}_{ij} = (B_{y(i)})_{ij}$$

for each i, j, hence only at most 2n matrices B_y must be computed If A^I is strongly regular, then the equation (4.1) can be solved by Jacobi, Gauss-Seidel or Kraczyk iterations, as described in section 2.

If the coefficients are given with a uniform absolute error, i.e. $\Delta_{ij} = \beta = \text{const}$ for each i, j, then from (3.3) we obtain formulae [22]

$$\underline{B}_{ij} = (A_c^{-1})_{ij} - \frac{\beta r_i s_j}{1 + \beta v_{ij}}$$

$$\overline{B}_{ij} = (A_c^{-1})_{ij} + \frac{\beta r_i s_j}{1 - \beta v_{ij}}$$

where

$$v_{ij} = \tilde{y}(j)^T A_c^{-1} y(i),$$

$\tilde{y}(j)$ being the signature vector of the j-th column of A_c^{-1}. These formulae are valid if β satisfies (3.22).

Arguing as in section 3, we may show that the maximal loss of significant decimals in components of the inverse matrix due to data rounding is given by

$$\sigma_A = [\log_{10}(\|A^{-1}\|_\infty \|A^{-1}\|_1)]_0$$

provided none of the numbers $\log_{10}(r_i s_j)$ is integer [22].

References

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

[2] W. Barth and E. Nuding, Optimale Losung von Intervall-gleichungssystemen, Computing 12: 117-125, (1974).

[3] H. Beeck, Charakterisierung der Lösungsmenge von Intervallgleichungssystemen, Z. Angew. Math. Mech 53: T181-T182 (1973).

[4] H. Beeck, Zur scharfen Aussenabschätzung der Lösungsmenge bei linearen Intervallgleichungssystemen, Z. Angew. Math. Mech. 54: T208-T209 (1974).

[5] H. Beeck, Zur Problematik der Hüllenbestimmung von Intervallgleichungssystemen, in Interval Mathematics, Lecture Notes in Computer Science 29 (K. Nickel, Ed.), Springer-Verlag (1975).

[6] J. Cope and B. Rust, Bounds on solutions of linear systems with inaccurate data, SIAM J. Numer. Anal. 16: 950-963 (1979).

[7] A. Deif, Sensitivity Analysis in Linear Systems, Springer Verlag, Berlin-Heidelberg-New York (1986).

[8] E. Hansen, On linear algebraic equations with interval coefficients, in Topics in Interval Analysis (E. Hansen, Ed.) Oxford University Press, Oxford, (1969).

[9] R. Krawczyk, Newton-Algorithmen zur Bestimmung von Nullstellen mit Fehlerschranken, Computing 4: 187-201 (1969).

[10] J. Kuttler, A fourth-order finite-difference approximation for the fixed membrane eigenproblem, Math. of Computation 25: 237-256 (1971).

[11] K. Murty, Note on a Bard-type scheme for solving the complementarity problem, Opsearch 11: 123-130 (1974).

[12] A. Neumaier, Linear Interval Equations, in Interval Mathematics 1985, Lecture Notes in Computer Science (K. Nickel, Ed.), Springer Verlag (1985).

[13] A. Neumaier, private communication, 1986.

[14] K. Nickel, Die Überschätzung des Wertebereichs einer Funktion in der Intervallrechnung mit Anwendungen auf lineare Gleichungssysteme, Computing 18: 15-36, (1977).

[15] W. Oettli, On the solution set of a linear system with inaccurate coefficients, SIAM J. Numer. Anal. 2: 115-118, (1965).

[16] W. Oettli and W. Prager, Compatibility of approximate solution of linear equations with given error bounds for coefficients and right-hand sides, Numer. Math. 6: 405-409 (1964).

[17] J. Rohn, Solving interval linear systems, Freiburger Intervall-Berichte 84/7 : 1-14, (1984).

[18] J. Rohn, Proofs to "Solving interval linear systems", Freiburger Intervall-Berichte 84/7: 17-30, (1984).

[19] J. Rohn, Some results on interval linear systems, Freiburger Intervall-Berichte 85/4: 93-116, (1985).

[20] J. Rohn, Inverse-positive interval matrices, Z. Angew. Math. Mech. 67 : T492-T493 (1987).

[21] J. Rohn, Systems of linear interval equations, to appear.

[22] J. Rohn, Formulae for exact bounds on solutions of linear systems with rank one perturbations, Freiburger Intervall-Berichte 87/6 : 1-20 (1987).