ON THE INTERVAL HULL OF THE SOLUTION SET OF AN INTERVAL LINEAR SYSTEM

Jiří Rohn
Charles University
Prague, Czechoslovakia

Dedicated to Prof. Dr. Rudolf Krawczyk on his 60^{th} birthday

1. INTRODUCTION

Let

(0) $$A^I x = b^I$$

be an interval linear system with an nxn interval matrix A^I. The set

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

is usually called the solution set of (0). If A^I is nonsingular (which means that each $A \in A^I$ is nonsingular), then X is closed, bounded and connected [4], but generally not convex and not an interval [3]. The narrowest interval containing X, i.e. the interval $[\underline{x}, \overline{x}]$ given by

$$\underline{x}_i = \min\{x_i \mid x \in X\}$$
$$\overline{x}_i = \max\{x_i \mid x \in X\} \qquad (i = 1,\dots,n),$$

is called the interval hull of X. There is a number of results concerning the problem of computing the interval hull under special assumptions (see [1] - [14]). Less is known of the general case. Nickel [13] pointed out that the interval hull of X can be computed by solving $2^{n(n+1)}$ linear nxn systems (in ordinary, not interval, arithmetic). In this paper, we propose a method which reduces the number of linear systems to be solved to a number between 2^p and 2^{p+q} where p is the number of equations in (0) containing at least one nondegenerate interval coefficient and q is the number of columns of A^I having the same property. As shown in section 3, the method performed well on examples with 2x2 matrices. The present lack of a broader computational experience does not allow to judge of the efficiency of the method in general case.

2. BASIC RESULT

We begin with some notations. Let $A^I = \{A \mid \underline{A} \leq A \leq \overline{A}\}$, where $\underline{A} = (\underline{a}_{ij})$, $\overline{A} = (\overline{a}_{ij})$ are nxn matrices and let $b^I = \{b \mid \underline{b} \leq b \leq \overline{b}\}$, $\underline{b} = (\underline{b}_i)$, $\overline{b} = (\overline{b}_i)$ being n-vectors. Further, let

$$Y = \{y \in R^n \mid |y_j| = 1,\ j = 1,\dots,n\}\ ,$$

so that Y contains 2^n elements. For each $y \in Y$, $z \in Y$ define an nxn matrix A_{yz} and an n-vector b_y by

$$\begin{aligned}(A_{yz})_{ij} &= \overline{a}_{ij} \text{ if } y_i z_j = 1\\ &= \underline{a}_{ij} \text{ if } y_i z_j = -1 \qquad (i,j = 1,\dots,n),\\ (b_y)_i &= \underline{b}_i \text{ if } y_i = 1\\ &= \overline{b}_i \text{ if } y_i = -1 \qquad (i = 1,\dots,n).\end{aligned}$$

For $x \in R^n$ and $z \in Y$ we define an n-vector x^z by

$$(x^z)_j = z_j x_j \qquad (j = 1,\dots,n).$$

Finally, we denote by e the n-vector $(1, \dots, 1)$ and $f = -e$, so that $e \in Y$ and $f \in Y$.

Our basic result is then formulated as follows:

<u>Theorem 1.</u> Let A^I be nonsingular and let for each $y \in Y$ there exist a $z \in Y$ such that the solution x_y of the system

(1) $$A_{yz}x = b_y$$

satisfies

(2) $$x_y^z \geqslant 0.$$

Then the interval hull $[\underline{x}, \overline{x}]$ of the solution set X is given by

(3) $$\underline{x}_i = \min \left\{ x_{yi} \mid y \in Y \right\}$$
$$\overline{x}_i = \max \left\{ x_{yi} \mid y \in Y \right\} \qquad (i = 1,\dots,n).$$

The <u>proof</u> employs the idea of the constructive part of the proof of Theorem 1 in [15]. Let W be the convex hull of the points x_y, $y \in Y$. First we prove that $X \subset W$. To this end, take an $x_o \in X$, so that $Ax_o = b$ for some $A \in A^I$, $b \in b^I$. For each $r \in \{0, 1, \dots, n\}$ and $y \in Y$, the nx2n system

$$(A(x_1 - x_2))_i = b_i \qquad (i = 1,\dots,r)$$
$$(A_{ye}x_1 - A_{yf}x_2)_i = b_{yi} \qquad (i = r+1,\dots,n)$$

will be called an (r,y)-system. We shall prove by induction on r that each (r,y)-system has a nonnegative solution x_1, x_2 satisfying $x_1 - x_2 \in W$. If $r = 0$, then a (0,y)-system has the form $A_{ye}x_1 - A_{yf}x_2 = b_y$, hence for the vectors x_1, x_2 given by $x_{1i} = \max\left\{x_{yi}, 0\right\}$, $x_{2i} = \max\left\{-x_{yi}, 0\right\}$ $(i = 1,\dots,n)$ we have $x_1 \geqslant 0$, $x_2 \geqslant 0$, $x_1 - x_2 \in W$ and (1), (2) provide for $A_{ye}x_1 - A_{yf}x_2 = b_y$. Thus let $1 \leqslant r \leqslant n$ and $y \in Y$; define y', $y'' \in Y$ by $y'_r = -1$, $y'_j = y_j$ $(j \neq r)$ and $y''_r = 1$, $y''_j = y_j$ $(j \neq r)$. Due to the inductive assumption, the $(r-1, y')$-system has a nonnegative solution x'_1, x'_2 satisfying $x'_1 - x'_2 \in W$ and similarly the $(r-1, y'')$-system has a nonnegative solution x''_1, x''_2 with $x''_1 - x''_2 \in W$. Define a real function f of one real variable by

$$f(t) = (A(t(x'_1 - x'_2) + (1-t)(x''_1 - x''_2)))_r.$$

Then, we have $f(0) = (A(x_1'' - x_2''))_r \leq (\overline{A}x_1'' - \underline{A}x_2'')_r = (A_{y''e}x_1'' - A_{y''f}x_2'')_r = \underline{b}_r \leq b_r$ and $f(1) = (A(x_1' - x_2'))_r \geq (\underline{A}x_1' - \overline{A}x_2')_r = (A_{y'e}x_1' - A_{y'f}x_2')_r = \overline{b}_r \geq b_r$, hence there is a $t_o \in [0, 1]$ with $f(t_o) = b_r$. Put

$$x_1 = t_o x_1' + (1 - t_o)x_1''$$
$$x_2 = t_o x_2' + (1 - t_o)x_2'',$$

so that x_1 and x_2 are nonnegative and

(4) $$x_1 - x_2 = t_o(x_1' - x_2') + (1 - t_o)(x_1'' - x_2''),$$

which immediately gives $x_1 - x_2 \in W$. From the definition of t_o we have $(A(x_1 - x_2))_r = b_r$. If $1 \leq i < r$, then (4) gives $(A(x_1 - x_2))_i = t_o b_i + (1 - t_o)b_i = b_i$; if $r+1 \leq i \leq n$, then $y_i = y_i' = y_i''$, hence $(A_{ye}x_1 - A_{yf}x_2)_i = t_o(A_{y'e}x_1' - A_{y'f}x_2')_i + (1 - t_o)(A_{y''e}x_1'' - A_{y''f}x_2'')_i = b_{yi}$. Hence x_1, x_2 is a nonnegative solution to the (r,y)-system satisfying $x_1 - x_2 \in W$, which completes the inductive proof. Taking now $r = n$, we get that there are x_1, x_2 satisfying $A(x_1 - x_2) = b$ and $x_1 - x_2 \in W$. Then the nonsingularity of A implies $x_o = x_1 - x_2$, hence $x_o \in W$. This proves $X \subset W$; since the interval $[\underline{x}, \overline{x}]$ given by (3) satisfies $W \subset [\underline{x}, \overline{x}]$, we have $X \subset [\underline{x}, \overline{x}]$. On the other hand, since $x_y \in X$ for each $y \in Y$, $[\underline{x}, \overline{x}]$ must be the narrowest interval containing X, hence $[\underline{x}, \overline{x}]$ is the interval hull of X, Q. E. D. ■

Theorem 1 shows a way how to compute the (exact) interval hull. However, it requires for each $y \in Y$ to find a $z \in Y$ such that the vector $x_y = A_{yz}^{-1}b_y$ satisfies $x_y^z \geq 0$. This may be a difficult task in the general case; the heuristic algorithm for computing x_y described below performed well on small size examples, although it is probably generally not prevented from cycling:

<u>Algorithm</u> (for computing x_y for a given $y \in Y$):

Step 0: Set $z:=e$.

Step 1: Solve $A_{yz}x = b_y$.

Step 2: If $x^z \geqslant 0$, set $x_y:=x$. Stop!

Step 3: Set $z_k:=-z_k$ for each k with $z_k x_k < 0$ and return to Step 1.

This algorithm combined with Theorem 1 gives a method for computing the interval hull. Several examples are shown in the next section.

3. EXAMPLES

Three examples with 2x2 matrices are computed here. Two observations were made: (i) the algorithm always stopped after solving at most two systems, (ii) in all three examples, if $\underline{x}_i = x_{yi}$ for some y and i, then $\overline{x}_i = (x_{-y})_i$.

<u>Example 1</u> (Barth and Nuding [3]).

$$[2,4]\,x_1 + [-2,1]\,x_2 = [-2,2]$$
$$[-1,2]\,x_1 + [2,4]\,x_2 = [-2,2]$$

First, we set $y:=(1,1)$ and $z:=(1,1)$. Then $A_{yz}x = b_y$ has the form

$$4x_1 + x_2 = -2$$
$$2x_1 + 4x_2 = -2$$

and its solution $x_1 = -\frac{3}{7}$, $x_2 = -\frac{2}{7}$ does not satisfy $x^z \geqslant 0$. Hence we set $z:=(-1,-1)$ (Step 3 of the algorithm) and solve

$$2x_1 - 2x_2 = -2$$
$$-x_1 + 2x_2 = -2$$

which gives the solution $x_1 = -4$, $x_2 = -3$ satisfying $x^z \geqslant 0$.

Thus we get

$$x_{(1,1)} = (-4,-3).$$

In a similar way we obtain

$$x_{(1,-1)} = (-3,4)$$
$$x_{(-1,1)} = (3,-4)$$
$$x_{(-1,-1)} = (4,3)$$

and Theorem 1 gives

$$\underline{x} = (-4,-4)$$
$$\overline{x} = (4,4).$$

<u>Example 2</u> (Nickel [13]).

$$[2,4]x_1 + [-2,-1]x_2 = [8,10]$$
$$[2,5]x_1 + [4,5]x_2 = [5,40]$$

Here, we have

$$x_{(1,1)} = (\tfrac{21}{13}, -\tfrac{10}{13})$$
$$x_{(1,-1)} = (4,8)$$
$$x_{(-1,1)} = (\tfrac{45}{13}, -\tfrac{40}{13})$$
$$x_{(-1,-1)} = (10,5),$$

thus

$$\underline{x} = (\tfrac{21}{13}, -\tfrac{40}{13})$$
$$\overline{x} = (10,8).$$

<u>Example 3</u> (Hansen [11]).

$$[2,3]x_1 + [0,1]x_2 = [0,120]$$
$$[1,2]x_1 + [2,3]x_2 = [60,240].$$

Here we obtain $x_{(1,1)} = (-12,24)$, $x_{(1,-1)} = (-120,240)$, $x_{(-1,1)} = (90,-60)$, $x_{(-1,-1)} = (60,90)$, which gives $\underline{x} = (-120,-60)$ and $\overline{x} = (90,240)$.

4. EDGE POINTS

A system of the form

(5) $$A_{yz}x = b_y \qquad x^z \geqslant 0,$$

appearing in Theorem 1, may seem strange at the first glance. In this section, we shall give some geometric interpretation to the points satisfying (5). We introduce this notion: a point $x \in X$ is said to be an edge point of X if there <u>does not</u> exist a pair of different points x_1, x_2 such that the segment connecting x_1 and x_2 lies in X and $x = \frac{1}{2}(x_1 + x_2)$. For a characterization of the edge points we need the following lemma, which is a mere reformulation of Theorem 2 in [4]:

<u>Lemma.</u> $x \in X$ if and only if there is a $z \in Y$ such that x satisfies

$$A_{fz}x \leqslant \overline{b} \qquad A_{ez}x \geqslant \underline{b} \qquad x^z \geqslant 0.$$

Now, we have (assuming again A^I is nonsingular):

<u>Theorem 2.</u> Let $x \in R^n$ and let $x_i \neq 0$ $(i = 1,\dots,n)$. Then, x is an edge point of X if and only if it satisfies (5) for some $y,z \in Y$.

<u>Proof.</u> The "if" part: Let x satisfy (5) and assume x is not an edge point of X so that there are x_1, $x_2 \in X$, $x_1 \neq x_2$, such that $x = \frac{1}{2}(x_1 + x_2)$; moreover, they can be chosen so closely to x so that $x_1^z \geqslant 0$, $x_2^z \geqslant 0$. Take an i with $y_i = -1$; then Lemma gives $(A_{yz}x_1)_i = (A_{fz}x_1)_i \leqslant \overline{b}_i$ and similarly $(A_{yz}x_2)_i \leqslant \overline{b}_i$. Assume at least one of these inequalities holds sharply; then we have $(A_{yz}x)_i < \overline{b}_i = b_{yi}$, a contradiction. Hence

(6) $$(A_{yz}x_1)_i = (A_{yz}x_2)_i = (A_{yz}x)_i.$$

If $y_i = 1$, then a similar reasoning again gives (6). Hence $A_{yz}x_1 = A_{yz}x_2 = A_{yz}x$, which implies $x_1 = x_2 = x$, a contradiction.

The "only if" part: Assume x is an edge point. Then there is a unique $z \in Y$ with $x^z \geqslant 0$, so that $A_{fz}x \leqslant \overline{b}$, $A_{ez}x \geqslant \underline{b}$. Put

$$J_1 = \{\, i \mid (A_{fz}x)_i = \overline{b}_i \}$$
$$J_2 = \{\, i \mid (A_{fz}x)_i < \overline{b}_i,\ (A_{ez}x)_i = \underline{b}_i \} ,$$

then $J_1 \cap J_2 = \emptyset$. We prove $J_1 \cup J_2 = \{1,\dots,n\}$. Assume it is not so and consider the system (obviously, $J_1 \cup J_2 \neq \emptyset$)

$$(A_{fz}x_*)_i = 0 \qquad (i \in J_1)$$
$$(A_{ez}x_*)_i = 0 \qquad (i \in J_2).$$

Since its number of equations is less than n, it possesses a non-trivial solution x_o. Now choose a $d_o > 0$ such that $(x \pm d_o x_o)^z \geqslant 0$,

$$d_o \left|(A_{fz}x_o)_i\right| < \overline{b}_i - (A_{fz}x)_i$$

for each i with $(A_{fz}x)_i < \overline{b}_i$ and

$$d_o \left|(A_{ez}x_o)_i\right| < (A_{ez}x)_i - \underline{b}_i$$

for each i with $(A_{ez}x)_i > \underline{b}_i$. Then the whole segment connecting the points $x_1 = x - d_o x_o$, $x_2 = x + d_o x_o$ lies in X, $x_1 \neq x_2$ and $x = \frac{1}{2}(x_1 + x_2)$, hence x is not an edge point. This contradiction shows that $J_1 \cup J_2 = \{1,\dots,n\}$. Now define $y \in Y$ as follows:

$$y_i = -1 \text{ if } i \in J_1,$$
$$y_i = 1 \text{ if } i \in J_2.$$

Then we have $A_{yz}x = b_y$, which completes the proof. ■

Theorems 1 and 2, if combined, show that the edge points of the solution set X play a similar role as the vertices of convex polytopes. Notice that all the x_y's in the above examples 1 - 3 are edge points of the respective solution sets.

5. DISCUSSION

A closer look into the form of the systems (1) shows that the number of such systems to be examined lies between 2^p and 2^{p+q}, where p is the number of equations in (0) containing at least one nondegenerate interval coefficient and q is the number of columns of A^I with the same property. In fact, if the i-th equation does not contain a nondegenerate interval coefficient, then all its coefficients are real numbers and the change of the sign of y_i does not affect the form of (1); similarly for the j-th column of A^I. This shows that the number of mutually different b_y's is 2^p and the number of mutually different systems (1) is at most 2^{p+q}. Under special assumptions, the number of systems (1) to be solved can be essentially less, cf. Garloff [7].

Further, it is not necessary to store all the x_y's during the computation: after updating $\underline{x}$ and $\overline{x}$, the current x_y may be dropped out.

REFERENCES

[1] J. Albrecht, Monotone Iterationsfolgen und ihre Verwendung zur Lösung linearer Gleichungssysteme, Numer. Math. 3(1961), 345-358

[2] G. Alefeld, J. Herzberger, Über die Verbesserung von Schranken für die Lösung bei linearen Gleichungssystemen, Angew. Informatik 3(1971), 107-112

[3] W. Barth, E. Nuding, Optimale Lösung von Intervallgleichungssystemen, Computing 12(1974), 117-125

[4] H. Beeck, Charakterisierung der Lösungsmenge von Intervallgleichungssystemen, ZAMM 53(1973), T181-T182

[5] H. Beeck, Zur scharfen Aussenabschätzung der Lösungsmenge bei linearen Intervallgleichungssystemen, ZAMM 54(1974), T208-T209

[6] H. Beeck, Zur Problematik der Hüllenbestimmung von Intervallgleichungssystemen, in: Interval Mathematics (K. Nickel, Ed.), Springer-Verlag, Berlin-Heidelberg-New York 1975

[7] J. Garloff, Totally Nonnegative Interval Matrices, in: Interval Mathematics 1980 (K. Nickel, Ed.), Academic Press, New York 1980

[8] E. Hansen, Interval Arithmetic in Matrix Computations, Part I, J. SIAM Numer. Anal., Ser. B, Vol. 2, No. 2, 1965, 308-320

[9] E. Hansen, R. Smith, Interval Arithmetic in Matrix Computations, Part II, SIAM J. Numer. Anal., Vol. 4, No. 1, 1967, 1-9

[10] E. Hansen, On the Solution of Linear Algebraic Equations with Interval Coefficients, Lin. Alg. and Its Apps. 2(1969), 153-165

[11] E. Hansen, On Linear Algebraic Equations with Interval Coefficients, in: Topics in Interval Analysis (E. Hansen, Ed.), Clarendon Press, Oxford 1969

[12] K.-U. Jahn, Eine Theorie der Gleichungssysteme mit Intervall-Koeffizienten, ZAMM 54(1974), 405-412

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

[14] W. Oettli, On the Solution Set of a Linear System with Inaccurate Coefficients, J. SIAM Numer. Anal., Ser. B, Vol. 2, No. 1, 1965, 115-118

[15] J. Rohn, Strong Solvability of Interval Linear Programming Problems, Computing 26(1981), 79-82

Author's address: Jiří Rohn, Faculty of Mathematics and Physics, Charles University, Malostranské nám. 25, 11800 Prague 1, Czechoslovakia