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 60th birthday

1. INTRODUCTION

Let

\[(0) \quad A^I x = b^I\]

be an interval linear system with an n x n 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

\[\underline{x}_i = \min \{x_i \mid x \in X\}\]

\[\overline{x}_i = \max \{x_i \mid x \in X\} \quad (i = 1, \ldots, n),\]

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 \(2^R(n+1)\) linear n x n 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 $2 \times 2$ 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 $A^I = \{ \Delta \mid \Delta \leq \Lambda \leq \bar{\Lambda} \}$, where $\Delta = (\Delta_{ij})$, $\Lambda = (\Lambda_{ij})$ are $n \times n$ matrices and let $b^I = \{ b \mid b \leq \bar{b} \leq \overline{b} \}$, $\bar{b} = (\bar{b}_1)$, $\overline{b} = (\overline{b}_1)$ being $n$-vectors. Further, let

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

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

$$\begin{align*}
(A_{yz})_{ij} &= \Delta_{ij} \text{ if } y_i z_j = 1 \\
&= \Lambda_{ij} \text{ if } y_i z_j = -1 \\
&= \bar{\Lambda}_{ij} \text{ if } y_i = 1 \\
&= \overline{\Lambda}_{ij} \text{ if } y_i = -1 \\
(b_y)_i &= \bar{b}_i \text{ if } y_i = 1 \\
&= \overline{b}_i \text{ if } y_i = -1
\end{align*}$$

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

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

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

Our basic result is then formulated as follows:
Theorem 1. Let $A^T$ be nonsingular and let for each $y \in Y$ there exist a $z \in Y$ such that the solution $x_y$ of the system

$$A_y z = b$$

satisfies

$$x_y^z \geq 0.$$  

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

$$x_i = \min \{ x_{y_1} \mid y \in Y \}$$

$$\bar{x}_i = \max \{ x_{y_1} \mid y \in Y \} \quad (i = 1, \ldots, n).$$

The proof 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 $x_0 \in X$, so that $A x_0 = b$ for some $A \in A^T$, $b \in b^T$. For each $r \in \{0, 1, \ldots, n\}$ and $y \in Y$, the nx2n system

$$(A(x_1 - x_2))^r = b$$

$$(A_{y} x_1 - A_{y} x_2)^r = b_y$$

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 $A_y x_1 - A_y x_2 = b_y$, hence for the vectors $x_1, x_2$ given by $x_{11} = \max \{ x_{y_1}, 0 \}$, $x_{21} = \max \{ -x_{y_1}, 0 \}$ ($i = 1, \ldots, n$) we have $x_1 \geq 0, x_2 \geq 0, x_1 - x_2 \in W$ and (1), (2) provide for $A_{y} x_1 - A_{y} x_2 = b_y$. Thus let $1 \leq r \leq n$ and $y \in Y$; define $y', y'' \in Y$ by $y'_r = -1, y''_r = y'_r$ ($j \neq r$) and $y'_r = 1, y''_r = y'_r$ ($j \neq r$). 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 $f(0) = (\Lambda(x_1^0 - x_2^0))_r \preceq (\Lambda x_1^0 - \Lambda x_2^0)_r = (\Lambda x_1' - \Lambda x_2')_r = b_r$ and $f(1) = (\Lambda(x_1^1 - x_2^1))_r \succeq b_r$. Hence, there is a $t_0 \in [0, 1]$ with $f(t_0) = b_r$. Put
\[x_1 = t_0 x_1' + (1 - t_0) x_1^0,\]
\[x_2 = t_0 x_2' + (1 - t_0) x_2^0,\]
so that $x_1$ and $x_2$ are nonnegative and
\[(4) \quad x_1 - x_2 = t_0 (x_1' - x_2') + (1 - t_0) (x_1^0 - x_2^0),\]
which immediately gives $x_1 - x_2 \in \mathbb{W}$. From the definition of $t_0$, we have $\Lambda(x_1 - x_2)_r = b_r$. If $1 \leq i \leq r$, then $(4)$ gives $\Lambda(x_1 - x_2)_i = t_0 b_i + (1 - t_0) b_i = b_i$. If $r + 1 \leq i \leq n$, then $y_1 = y_1' = y_1^0$, hence
\[(\Lambda y_1 x_1 - \Lambda y_2 x_2)_i = t_0 (\Lambda y_1 x_1' - \Lambda y_2 x_2')_i + (1 - t_0) (\Lambda y_1 x_1^0 - \Lambda y_2 x_2^0)_i = b_1.\] Hence $x_1, x_2$ is a nonnegative solution to the $(r, r)$-system satisfying $x_1 - x_2 \in \mathbb{W}$, which completes the inductive proof. Taking now $r = n$, we get that there are $x_1, x_2$ satisfying $\Lambda(x_1 - x_2) = b$ and $x_1 - x_2 \in \mathbb{W}$. Then the nonsingularity of $\Lambda$ implies $x_0 = x_1 - x_2$, hence $x_0 \in \mathbb{W}$. This proves $X \subset \mathbb{W}$; 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 $x_0 \in X$ 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 $z \in Z$ such that the vector $x_0 = A_y^{-1} b_y$ satisfies $x_0^z \succeq 0$. This may be a difficult task in the general case; the heuristic algorithm for computing $x_0$ described below performed well on small size examples, although it is probably generally not prevented from cycling.
Algorithm (for computing $x_y$ for a given $y \in Y$):

Step 0: Set $z_i = e$.

Step 1: Solve $A_y x = b_y$.

Step 2: If $x^z \geq 0$, 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, and (ii) in all three examples, if $x_1 = x_{y1}$ for some $y$ and $i$, then $x_i = (x_{y1})_i$.

Example 1 (Barth and Nuding [3]).

\[
\begin{align*}
(2, 4)x_1 + [-2, 1]x_2 &= [-2, 2] \\
[-1, 2]x_1 + [2, 4]x_2 &= [-2, 2]
\end{align*}
\]

First, we set $y_1 = (1, 1)$ and $z_i = (1, 1)$. Then $A_y x = b_y$ has the form

\[
\begin{align*}
4x_1 + x_2 &= -2 \\
2x_1 + 4x_2 &= -2
\end{align*}
\]

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

\[
\begin{align*}
2x_1 - 2x_2 &= -2 \\
-x_1 + 2x_2 &= -2
\end{align*}
\]

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

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:

\[ \bar{x} = (-4,-4) \]
\[ \overline{\bar{x}} = (4,4). \]

**Example 2 (Nickel [11]).**

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

Here, we have:

\[ x_{(1,1)} = \left( \frac{23}{13}, -\frac{10}{13} \right) \]
\[ x_{(1,-1)} = (4,8) \]
\[ x_{(-1,1)} = \left( \frac{45}{13}, -\frac{40}{13} \right) \]
\[ x_{(-1,-1)} = (10,5) \]

thus:

\[ \bar{x} = \left( \frac{23}{13}, -\frac{40}{13} \right) \]
\[ \overline{\bar{x}} = (10,8). \]

**Example 3 (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:

\[ x_{(1,1)} = (-12,24), \quad x_{(1,-1)} = (-120,240), \quad x_{(-1,1)} = (90,-60), \quad x_{(-1,-1)} = (60,90) \]

which gives:

\[ \bar{x} = (-120,-60) \]
\[ \overline{\bar{x}} = (90,240). \]
4. EDGE POINTS

A system of the form

\[ A_{yz}x = b_y \]
\[ x^2 \geq 0, \]

appearing in Theorem 1, may seem strange at 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 does not 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 re-formulation of Theorem 2 in [4]:

**Lemma.** \( x \in X \) if and only if there is a \( z \in Y \) such that \( x \) satisfies

\[ A_{ez}x \leq b \]
\[ A_{ez}x \geq b \]
\[ x^2 \geq 0. \]

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

**Theorem 2.** Let \( x \in \mathbb{R}^n \) and let \( x_i \neq 0 \) \( (i = 1, \ldots, n) \). Then, \( x \) is an edge point of \( X \) if and only if it satisfies (5) for some \( y, z \in Y \).

**Proof.** 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 \( x_1 \geq 0, x_2 \geq 0 \). Take an \( i \) with \( y_i = -1 \); then Lemma gives \( (A_{yz}x_i)_1 = (A_{ez}x_1)_1 \leq b_i \) and similarly \( (A_{yz}x_2)_1 \leq b_i \). Assume at least one of these inequalities holds sharply; then we have \( (A_{yz}x)_i < b_i = b_{y_i} \), a contradiction. Hence

\[ (A_{yz}x_1)_1 = (A_{yz}x_2)_1 = (A_{yz}x)_1. \]
If \( y_1 = 1 \), 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 \( x^\infty \geq 0 \), so that \( A_{yz}x \leq b \), \( A_{yz}x \geq b \). Put
\[ J_1 = \{ i | (A_{xyz})_1 = b_1 \} \]
\[ J_2 = \{ i | (A_{xyz})_1 < b_1, (A_{xyz})_1 = b_1 \} \],
then \( J_1 \cap J_2 = \emptyset \). We prove \( J_1 \cup J_2 = \{1, \ldots, n\} \). Assume it is not so and consider the system (obviously, \( J_1 \cup J_2 \neq \emptyset \))
\[ (A_{xyz})_i = 0 \quad (i \in J_1) \]
\[ (A_{xyz})_i = 0 \quad (i \in J_2). \]
Since its number of equations is less than \( n \), it possesses a non-trivial solution \( x_0 \). Now choose a \( d_0 > 0 \) such that \( (x + d_0x_0)^\infty \geq 0 \),
\[ d_0 |(A_{xyz})_i| < b_1 - (A_{xyz})_i \]
for each \( i \) with \( (A_{xyz})_i < b_1 \) and
\[ d_0 |(A_{xyz})_i| < (A_{xyz})_i - b_1 \]
for each \( i \) with \( (A_{xyz})_i > b_1 \). Then the whole segment connecting the points \( x_1 = x - d_0x_0 \), \( x_2 = x + d_0x_0 \) 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 \( J_1 \cup J_2 = \{1, \ldots, n\} \). Now define \( y \in Y \) as follows:
\[ y_1 = -1 \text{ if } i \in J_1, \]
\[ y_1 = 1 \text{ if } i \in J_2. \]
Then we have \( A_{yz}x = b^\prime_y \), 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 $A^T$ 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 $A^T$. 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 $\hat{x}$ and $\hat{x}$, the current $x_y$ may be dropped out.

REFERENCES


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