A New Proof of the Hansen-Bliek-Rohn Optimality Result

Rohn, Jiří
2014

Institute of Computer Science
Academy of Sciences of the Czech Republic

A New Proof of the Hansen-Bliek-Rohn Optimality Result

Jiří Rohn
http://uivtx.cs.cas.cz/~rohn

Technical report No. V-1212

20.03.2014

Pod Vodárenskou věží 2, 182 07 Prague 8, phone: +420 266 051 111, fax: +420 286 585 789,
e-mail:rohn@cs.cas.cz

Institute of Computer Science
Academy of Sciences of the Czech Republic

A New Proof of the Hansen-Bliek-Rohn Optimality Result

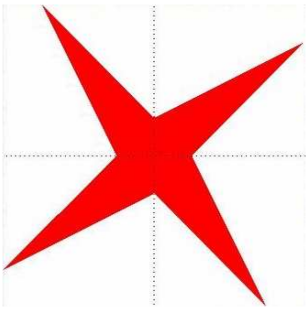
Jiří Rohn[1]
http://uivtx.cs.cas.cz/~rohn

Technical report No. V-1212

20.03.2014

Abstract:

We present a new proof of the Hansen-Bliek-Rohn optimality result for interval linear equations with unit midpoint.[2]



Keywords:
Interval linear equations, unit midpoint, Hansen-Bliek-Rohn optimality result.

[1] This work was supported with institutional support RVO:67985807.

[2] Above: logo of interval computations and related areas (depiction of the solution set of the system $[2,4]x_1 + [-2,1]x_2 = [-2,2]$, $[-1,2]x_1 + [2,4]x_2 = [-2,2]$ (Barth and Nuding [1])).

1 Introduction

For a system of interval linear equations $\mathbf{A}x = \mathbf{b}$, where $\mathbf{A}$ is an $n \times n$ interval matrix and $\mathbf{b}$ is an interval n-vector, the interval hull is defined as

$$\mathbf{x}(\mathbf{A}, \mathbf{b}) = \bigcap_{\mathbf{X}(\mathbf{A},\mathbf{b}) \subseteq [x,y]} [x, y],$$

where

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

i.e., as the narrowest interval vector containing the solution set $\mathbf{X}(\mathbf{A}, \mathbf{b})$. Computing the interval hull is NP-hard [11]. Yet it was shown by Hansen [3], Bliek [2] and Rohn [6] that the hull can be expressed by relatively simple closed-form formulae in case that the system matrix has unit midpoint, i.e., is of the form $\mathbf{A} = [I - \Delta, I + \Delta]$, where I is the unit matrix. However, the proof of this result is by no means straightforward. The formulae not using interval arithmetic were proved in [6], [8] and those formulated in terms of interval arithmetic by Ning and Kearfott [5] (using the result from [6]) and by Neumaier [4].

In this report we present another proof of the optimality result, based on a new characterization of the interval hull (Theorem 1). We give an interval-arithmetic-free version (Theorem 3) and an interval arithmetic version (Theorem 4), both in new formulations aimed at minimizing the number of auxiliary variables.

Notation used: $\operatorname{diag}(M)$ denotes the diagonal of a matrix M, $M_{k\bullet}$ its kth row, T_z is the diagonal matrix with diagonal vector z, $a \circ b$ stands for the Hadamard (entrywise) product of vectors a, b and a/b for their Hadamard division, minimum/maximum of a finite number of vectors is taken entrywise, I is the identity matrix and e is the vector of all ones.

2 Interval hull

We shall later make use of the following characterization of the interval hull.

Theorem 1. *Let $\mathbf{A} = [A_c - \Delta, A_c + \Delta]$ be regular. Then for each $z \in \{-1, 1\}^n$ the matrix equation*

$$QA_c - |Q|\Delta T_z = I$$

has a unique solution Q_z and for each right-hand side $\mathbf{b} = [b_c - \delta, b_c + \delta]$ there holds

$$\mathbf{x}(\mathbf{A}, \mathbf{b}) = \big[\min_{z\in\{-1,1\}^n} (Q_z b_c - |Q_z|\delta), \max_{z\in\{-1,1\}^n} (Q_z b_c + |Q_z|\delta) \big]. \tag{2.1}$$

Proof. The first part of the theorem is the assertion of [10, Thm. 1], the second one follows from [7, Thm. 2] if we take $Z = \{-1, 1\}^n$ there. □

3 Matrices Q_z

In this section we show that the matrices Q_z can be expressed explicitly in case of an interval matrix of the form $\mathbf{A} = [I - \Delta, I + \Delta]$. The result, as well as the subsequent ones, is formulated in terms of the matrix

$$M = (I - \Delta)^{-1}.$$

The assumption $M \geq I$ is equivalent to regularity of $[I-\Delta, I+\Delta]$, see [10].

Theorem 2. *Let $M \geq I$. Then for each $z \in \{-1,1\}^n$ the matrix Q_z is given rowwise by*

$$(Q_z)_{k\bullet} = \begin{cases} M_{k\bullet}T_z & \text{if } z_k = \ \ 1, \\ \nu_k(M_{k1}, \ldots, -M_{kk}, \ldots, M_{kn})T_z & \text{if } z_k = -1, \end{cases} \tag{3.1}$$

where

$$\nu_k = \tfrac{1}{2M_{kk}-1} \qquad (k = 1, \ldots, n).$$

Proof. The expression for $z_k = 1$ is contained in [10, Thm. 2]. The formula for $z_k = -1$ was given in the same theorem as

$$(Q_z)_{k\bullet} = ((\mu_k - 1)M_{k\bullet} - \mu_k e_k^T)T_z,$$

where

$$\mu_k = \tfrac{2M_{kk}}{2M_{kk}-1} \qquad (k = 1, \ldots, n).$$

Considering the fact that

$$\begin{aligned} (\mu_k - 1)M_{k\bullet} - \mu_k e_k^T &= (\mu_k - 1)(M_{k1}, \ldots, M_{kk} - \tfrac{\mu_k}{\mu_k - 1}, \ldots, M_{kn}) \\ &= (\mu_k - 1)(M_{k1}, \ldots, -M_{kk}, \ldots, M_{kn}) \\ &= \nu_k(M_{k1}, \ldots, -M_{kk}, \ldots, M_{kn}), \end{aligned}$$

we arrive at the desired result. □

4 Optimality result

The Hansen-Bliek-Rohn optimality result gives an explicit formula for the interval hull of an interval linear system of the form

$$\mathbf{I}x = \mathbf{b},$$

where $\mathbf{I} = [I - \Delta, I + \Delta]$.

Theorem 3. *Let $M \geq I$. Then for each right-hand side $\mathbf{b} = [b_c - \delta, b_c + \delta]$, denoting $d = \operatorname{diag}(M)$, $x_* = d \circ b_c$ and $x^* = M(|b_c| + \delta)$, we have*

$$\mathbf{x}(\mathbf{I}, \mathbf{b}) = [\min\{\underset{\sim}{x},\, \underset{\sim}{x}/(2d - e)\},\, \max\{\tilde{x},\, \tilde{x}/(2d - e)\}], \tag{4.1}$$

where

$$\begin{aligned} \underset{\sim}{x} &= x_* - (x^* - |x_*|), \\ \tilde{x} &= x_* + (x^* - |x_*|). \end{aligned}$$

Comment. In (4.1) we use (twice) the Hadamard division of vectors.

Proof. Denote $[\underline{x}, \overline{x}] = \mathbf{x}(\mathbf{A}, \mathbf{b})$. Let $k \in \{1, \ldots, n\}$. We shall first derive a formula for $\overline{x}_k$. From (2.1) we have

$$\overline{x}_k = \max_{z \in \{-1,1\}^n} (Q_z b_c + |Q_z|\delta)_k = \max_{z \in \{-1,1\}^n} ((Q_z)_{k\bullet} b_c + |Q_z|_{k\bullet}\delta),$$

so that according to (3.1) for each $z \in \{-1,1\}^n$ we must consider two cases: $z_k = 1$ and $z_k = -1$.

If $z_k = 1$, then by Theorem 2

$$\begin{aligned}(Q_z)_{k\bullet} b_c + |Q_z|_{k\bullet}\delta &= M_{k\bullet} T_z b_c + M_{k\bullet}\delta \\ &= \sum_{j\neq k} M_{kj} z_j (b_c)_j + M_{kk}(b_c)_k + M_{k\bullet}\delta \\ &\le \sum_{j\neq k} M_{kj} |(b_c)_j| + M_{kk}(b_c)_k + M_{k\bullet}\delta.\end{aligned}$$

Introducing the vector $\overline{z}(k) \in \{-1,1\}^n$ by

$$\overline{z}(k)_j = \begin{cases} 1 & \text{if } j = k, \\ 1 & \text{if } j \neq k \text{ and } (b_c)_j \ge 0, \\ -1 & \text{if } j \neq k \text{ and } (b_c)_j < 0 \end{cases} \qquad (j = 1, \dots, n),$$

we can write

$$\sum_{j\neq k} M_{kj} |(b_c)_j| + M_{kk}(b_c)_k + M_{k\bullet}\delta = M_{k\bullet} T_{\overline{z}(k)} b_c + M_{k\bullet}\delta = (Q_{\overline{z}(k)})_{k\bullet} b_c + |Q_{\overline{z}(k)}|_{k\bullet}\delta,$$

hence for each $z \in \{-1,1\}^n$ with $z_k = 1$ we have

$$(Q_z)_{k\bullet} b_c + |Q_z|_{k\bullet}\delta \le (Q_{\overline{z}(k)})_{k\bullet} b_c + |Q_{\overline{z}(k)}|_{k\bullet}\delta,$$

and the upper bound is obviously attained.

If $z_k = -1$, then, again by Theorem 2,

$$\begin{aligned}(Q_z)_{k\bullet} b_c + |Q_z|_{k\bullet}\delta &= \nu_k (M_{k1}, \dots, -M_{kk}, \dots, M_{kn}) T_z b_c + \nu_k M_{k\bullet}\delta \\ &= \nu_k \sum_{j\neq k} M_{kj} z_j (b_c)_j + \nu_k M_{kk}(b_c)_k + \nu_k M_{k\bullet}\delta \\ &\le \nu_k \sum_{j\neq k} M_{kj} |(b_c)_j| + \nu_k M_{kk}(b_c)_k + \nu_k M_{k\bullet}\delta \\ &= \nu_k (M_{k1}, \dots, -M_{kk}, \dots, M_{kn}) T_{\underline{z}(k)} b_c + \nu_k M_{k\bullet}\delta \\ &= (Q_{\underline{z}(k)})_{k\bullet} b_c + |Q_{\underline{z}(k)}|_{k\bullet}\delta\end{aligned}$$

where we have employed the vector $\underline{z}(k)$ given by

$$\underline{z}(k)_j = \begin{cases} -1 & \text{if } j = k, \\ 1 & \text{if } j \neq k \text{ and } (b_c)_j \ge 0, \\ -1 & \text{if } j \neq k \text{ and } (b_c)_j < 0 \end{cases} \qquad (j = 1, \dots, n),$$

hence for each $z \in \{-1,1\}^n$ with $z_k = -1$ we have

$$(Q_z)_{k\bullet} b_c + |Q_z|_{k\bullet}\delta \le (Q_{\underline{z}(k)})_{k\bullet} b_c + |Q_{\underline{z}(k)}|_{k\bullet}\delta,$$

and the upper bound is again obviously attained. In this way we have proved the formula

$$\overline{x}_k = \max\{Q_{\overline{z}(k)})_{k\bullet} b_c + |Q_{\overline{z}(k)}|_{k\bullet}\delta,\ (Q_{\underline{z}(k)})_{k\bullet} b_c + |Q_{\underline{z}(k)}|_{k\bullet}\delta\}.$$

Now,

$$\begin{aligned}(Q_{\overline{z}(k)})_{k\bullet}b_c + |Q_{\overline{z}(k)}|_{k\bullet}\delta &= \sum_{j\neq k} M_{kj}|(b_c)_j| + M_{kk}(b_c)_k + M_{k\bullet}\delta \\ &= M_{k\bullet}(|b_c|+\delta) + M_{kk}((b_c)_k - |b_c|_k) \\ &= (x_* + x^* - |x_*|)_k \\ &= \tilde{x}_k\end{aligned}$$

and similarly

$$\begin{aligned}(Q_{\underline{z}(k)})_{k\bullet}b_c + |Q_{\underline{z}(k)}|_{k\bullet}\delta &= \nu_k \sum_{j\neq k} M_{kj}|(b_c)_j| + \nu_k M_{kk}(b_c)_k + \nu_k M_{k\bullet}\delta \\ &= \nu_k(M_{k\bullet}(|b_c|+\delta) + M_{kk}((b_c)_k - |b_c|_k)) \\ &= \nu_k(x_* + x^* - |x_*|)_k \\ &= \nu_k\tilde{x}_k\end{aligned}$$

which together gives

$$\overline{x}_k = \max\{\tilde{x}_k, \nu_k\tilde{x}_k\}.$$

Since

$$\nu_k\tilde{x}_k = \tilde{x}_k/(2M_{kk}-1),$$

we finally obtain

$$\overline{x} = \max\{\tilde{x}, \tilde{x}/(2d-e)\},$$

where we have used the Hadamard (entrywise) division of vectors.

To prove the formula for $\underline{x}$, consider the system $\mathbf{I}x = -\mathbf{b}$, where $\mathbf{I} = [I-\Delta, I+\Delta]$ as before and $-\mathbf{b} = \{-b \mid b \in \mathbf{b}\} = [-b_c-\delta, -b_c+\delta]$. Then $\mathbf{X}(\mathbf{I}, -\mathbf{b}) = -\mathbf{X}(\mathbf{I}, \mathbf{b})$, hence $\mathbf{x}(\mathbf{I}, -\mathbf{b}) = [-\overline{x}, -\underline{x}]$. Now we can apply the previously derived formula for the upper bound of the interval hull:

$$-\underline{x} = \max\{-d\circ b_c + M(|b_c|+\delta) - |d\circ b_c|, (-d\circ b_c + M(|b_c|+\delta) - |d\circ b_c|)/(2d-e)\},$$

hence

$$\begin{aligned}\underline{x} &= \min\{d\circ b_c - M(|b_c|+\delta) + |d\circ b_c|,\ (d\circ b_c - M(|b_c|+\delta) + |d\circ b_c|)/(2d-e)\} \\ &= \min\{x_* - x^* + |x_*|,\ (x_* - x^* + |x_*|)/(2d-e)\} \\ &= \min\{\underset{\sim}{x},\ \underset{\sim}{x}/(2d-e)\}.\end{aligned}$$

This proves that

$$\mathbf{x}(\mathbf{I}, \mathbf{b}) = [\min\{\underset{\sim}{x},\ \underset{\sim}{x}/(2d-e)\},\ \max\{\tilde{x},\ \tilde{x}/(2d-e)\}].$$

□

Using the interval arithmetic, we can bring the result to yet simpler form.

Theorem 4. *Let $M \geq I$. Denoting $d = \operatorname{diag}(M)$, $x_* = d\circ b_c$ and $x^* = M(|b_c|+\delta)$, we have*

$$\mathbf{x}(\mathbf{I}, \mathbf{b}) = \frac{\langle x_*,\ x^* - |x_*|\rangle}{\langle d,\ d-e\rangle}. \tag{4.2}$$

Comment. In (4.2) we use the Hadamard (entrywise) division of interval vectors and their midpoint-radius representation, i.e., $\langle a, b\rangle = [a-b, a+b]$.

Proof. Because $\underset{\sim}{x} \le \tilde{x}$ and $\nu > 0$, we can write (4.1) as

$$\mathbf{x}(\mathbf{I}, \mathbf{b}) = [\min\{\underset{\sim}{x}/e,\ \underset{\sim}{x}/(2d-e), \tilde{x}/e,\ \tilde{x}/(2d-e)\},\ \max\{\underset{\sim}{x}/e,\ \underset{\sim}{x}/(2d-e), \tilde{x}/e,\ \tilde{x}/(2d-e)\}],$$

which is the Hadamard division performed in interval arithmetic:

$$\mathbf{x}(\mathbf{I}, \mathbf{b}) = \frac{[\underset{\sim}{x},\ \tilde{x}]}{[e,\ 2d-e]}. \tag{4.3}$$

Since

$$[\underset{\sim}{x},\ \tilde{x}] = [x_* - (x^* - |x_*|),\ x_* + (x^* - |x_*|)] = \langle x_*,\ x^* - |x_*|\rangle$$

and

$$[e,\ 2d-e] = \langle d,\ d-e\rangle,$$

(4.3) implies (4.2). □

The Hansen-Bliek-Rohn *optimality result* should not be misunderstood for the Hansen-Bliek-Rohn *enclosure*, see [9].

Bibliography

[1] W. Barth and E. Nuding, *Optimale Lösung von Intervallgleichungssystemen*, Computing, 12 (1974), pp. 117–125. 1

[2] C. Bliek, *Computer Methods for Design Automation*, PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, July 1992. 1

[3] E. R. Hansen, *Bounding the solution of interval linear equations*, SIAM Journal on Numerical Analysis, 29 (1992), pp. 1493–1503. 1

[4] A. Neumaier, *A simple derivation of the Hansen-Bliek-Rohn-Ning-Kearfott enclosure for linear interval equations*, Reliable Computing, 5 (1999), pp. 131–136. 1

[5] S. Ning and R. B. Kearfott, *A comparison of some methods for solving linear interval equations*, SIAM Journal on Numerical Analysis, 34 (1997), pp. 1289–1305. 1

[6] J. Rohn, *Cheap and tight bounds: The recent result by E. Hansen can be made more efficient*, Interval Computations, 4 (1993), pp. 13–21. 1

[7] J. Rohn, *An algorithm for computing the hull of the solution set of interval linear equations*, Linear Algebra and Its Applications, 435 (2011), pp. 193–201. 1

[8] J. Rohn, *The Hansen-Bliek optimality result as a consequence of the general theory*, Technical Report 1104, Institute of Computer Science, Academy of Sciences of the Czech Republic, Prague, February 2011. `http://uivtx.cs.cas.cz/~rohn/publist/hansen.pdf`. 1

[9] J. Rohn, *Compact form of the Hansen-Bliek-Rohn enclosure*, Technical Report 1157, Institute of Computer Science, Academy of Sciences of the Czech Republic, Prague, March 2012. `http://uivtx.cs.cas.cz/~rohn/publist/compactforms.pdf`. 5

[10] J. Rohn, *Explicit form of matrices Q_z for an interval matrix with unit midpoint*, Technical Report 1206, Institute of Computer Science, Academy of Sciences of the Czech Republic, Prague, February 2014. `http://uivtx.cs.cas.cz/~rohn/publist/report1206.pdf`. 1, 2

[11] J. Rohn and V. Kreinovich, *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. 1