Institute of Computer Science
Academy of Sciences of the Czech Republic

# Compact Form of the Hansen-Bliek-Rohn Enclosure

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

Technical report No. V-1157

24.03.2012

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

# Compact Form of the Hansen-Bliek-Rohn Enclosure

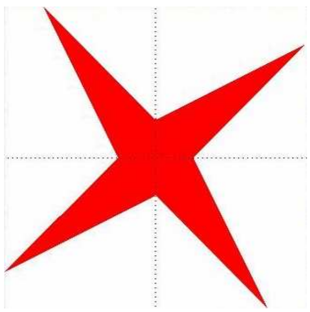
Jiří Rohn
http://uivtx.cs.cas.cz/~rohn

Technical report No. V-1157

24.03.2012

Abstract:

We show that using Hadamard product/division and interval arithmetic operations, the Hansen-Bliek-Rohn enclosure for both square and rectangular cases can be formulated in a surprisingly simple compact one-line form.[1]



Keywords:
Interval linear equations, Hansen-Bliek-Rohn enclosure, solution set, enclosure, interval hull, strong regularity.

[1]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 and notation

Hansen [4] and Bliek [2] published in 1992 almost simultaneously a closed-form explicit enclosure of the solution set of a system of interval linear equations. Both their proofs were not quite rigorous; a rigorous proof was supplied a year later in [8]. Ning and Kearfott [6] and Neumaier [5] later formulated the result in shorter form using interval arithmetic.

In this report we make one step further in three directions. First, we show that the result by Ning, Kearfott, and Neumaier, formulated in entrywise form, can be further simplified and brought to a vectorized form when using the Hadamard product/division. Second, we do not use the concept of $H$-matrices, but simply use equivalent formulation of strong regularity in the form of nonnegativity of certain inverse. And third, we generalize the result to rectangular matrices.

The main result for rectangular systems is given here in two versions. In Theorem 5 we formulate the assumptions in terms of general matrices $R$ and $M$ required to satisfy certain inequality only; the formulation of Theorem 6 is more specific, using instead of vague $R$ and $M$ exact inverses of certain matrices. In particular, the formulation contains $A_c^\dagger$, the pseudoinverse of $A_c$; replacing it simply by $A_c^{-1}$, we get the formerly-known result for square systems (Theorem 8). A reformulation of the Hansen-Bliek-Rohn optimality result is given in Theorem 9.

Theorem 6 is recommended for practical use for rectangular as well as square interval linear systems.

We use the following notation. $I$ is the unit matrix and $e$ is the vector of all ones. For a matrix $M = (m_{ij}) \in \mathbb{R}^{n\times n}$, $\operatorname{diag}(M) = (m_{11}, \ldots, m_{nn})^T$ stands for the diagonal of $M$. For two vectors $a = (a_i)_{i=1}^n$, $b = (b_i)_{i=1}^n$, $a \circ b = (a_i b_i)_{i=1}^n$ denotes the Hadamard product and $a/b = (a_i/b_i)_{i=1}^n$ the "Hadamard division". If $\underline{a}, \overline{a} \in \mathbb{R}^n$, $\underline{a} \le \overline{a}$, then the set

$$\mathbf{a} = \{\, x \mid \underline{a} \le x \le \overline{a} \,\} = [\underline{a}, \overline{a}]$$

is called an interval $n$-vector. We also use a so-called midpoint-radius notation

$$\mathbf{c} = \langle a, b\rangle = [a - b, a + b]$$

assuming that $b \ge 0$. If necessary (mainly when using interval arithmetic operations), we identify the interval vector $\mathbf{c} = \langle a, b\rangle$ with the vector of intervals $\mathbf{c} = (\langle a_i, b_i\rangle)_{i=1}^n$. Clearly, these are different mathematical objects, but there exists a trivial isomorphism between them. In this sense we define a Hadamard division of two interval vectors as

$$\frac{\langle a, b\rangle}{\langle c, d\rangle} = \left(\frac{\langle a_i, b_i\rangle}{\langle c_i, d_i\rangle}\right)_{i=1}^n \tag{1.1}$$

where the right-hand side division is performed in interval arithmetic. In fact, the componentwise division (1.1) will be our main tool used for expressing the Hansen-Bliek-Rohn enclosure in a compact one-line form (formula (6.2) and its consequent variants).

Next we recollect several basic terms and reformulate them for the rectangular case.

## 2 Solution set

Given an $m \times n$ interval matrix $\mathbf{A} = [A_c - \Delta, A_c + \Delta]$ and an interval $m$-vector $\mathbf{b} = [b_c - \delta, b_c + \delta]$, the solution set of the system of interval linear equations

$$\mathbf{A}x = \mathbf{b} \tag{2.1}$$

is defined as

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

The Oettli-Prager theorem [7] in today's formulation [3] asserts that

$$\mathbf{X}(\langle A_c, \Delta\rangle, \langle b_c, \delta\rangle) = \{ x \mid |A_c x - b_c| \le \Delta|x| + \delta \}. \tag{2.2}$$

We can immediately note the following property.

**Theorem 1.** *The solution set $\mathbf{X}(\mathbf{A}, \mathbf{b})$ is closed.*

Indeed, let $x_n \in \mathbf{X}(\mathbf{A}, \mathbf{b})$ for $n = 1, 2, \ldots$ and let $x_n \to x^*$. Due to the Oettli-Prager theorem we have

$$|A_c x_n - b_c| \le \Delta|x_n| + \delta$$

for each $n$, and taking the limit for $n \to \infty$ we obtain

$$|A_c x^* - b_c| \le \Delta|x^*| + \delta,$$

which shows that $x^* \in \mathbf{X}(\mathbf{A}, \mathbf{b})$. This proves that $\mathbf{X}(\mathbf{A}, \mathbf{b})$ is closed.

To achieve boundedness of $\mathbf{X}(\mathbf{A}, \mathbf{b})$, we need an additional assumption.

## 3 Compactness

An $m \times n$ matrix $A$ is said to have full column rank if its columns are linearly independent; this already implies that $m \ge n$. We shall later essentially use the fact that if $A$ has full column rank, then $A^T A$ is nonsingular. An interval matrix $\mathbf{A}$ is said to have full column rank if each $A \in \mathbf{A}$ has full column rank. We have this sufficient condition.

**Theorem 2.** *If $A_c$ has full column rank and*

$$(I - |A_c^\dagger|\Delta)^{-1} \ge 0$$

*holds, then $\mathbf{A}$ has full column rank.*

Let us note that here $A_c^\dagger = (A_c^T A_c)^{-1} A_c^T$ because $A_c$ has full column rank so that the above condition could also be written in the form

$$(I - |(A_c^T A_c)^{-1} A_c^T|\Delta)^{-1} \ge 0,$$

but we have preferred the shorter version. Theorem 2 could be proved independently but we shall see later that it is a direct consequence of Theorem 6 below. What we need for the sequel is the following assertion.

**Theorem 3.** *If* $\mathbf{A}$ *has full column rank, then the solution set* $\mathbf{X}(\mathbf{A}, \mathbf{b})$ *is compact for each right-hand side* $\mathbf{b}$*.*

Consider first the set

$$\mathbf{X}_{\mathbf{lsq}}(\mathbf{A}, \mathbf{b}) = \{ x \mid x = (A^T A)^{-1} A^T b \text{ for some } A \in \mathbf{A},\ b \in \mathbf{b} \}$$

(in fact, the set of least squares solutions). Because the mapping $(A, b) \mapsto (A^T A)^{-1} A^T b$ is a continuous mapping of the compact Cartesian product $\mathbf{A} \times \mathbf{b}$ onto $\mathbf{X}_{\mathbf{lsq}}(\mathbf{A}, \mathbf{b})$, its image $\mathbf{X}_{\mathbf{lsq}}(\mathbf{A}, \mathbf{b})$ is compact.

Now consider the set $\mathbf{X}(\mathbf{A}, \mathbf{b})$. If $x \in \mathbf{X}(\mathbf{A}, \mathbf{b})$, then $Ax = b$ for some $A \in \mathbf{A}$ and $b \in \mathbf{b}$, hence $A^T Ax = A^T b$ and $x = (A^T A)^{-1} A^T b \in \mathbf{X}_{\mathbf{lsq}}(\mathbf{A}, \mathbf{b})$. Thus $\mathbf{X}(\mathbf{A}, \mathbf{b})$, as a closed subset of a compact set $\mathbf{X}_{\mathbf{lsq}}(\mathbf{A}, \mathbf{b})$, is compact.

Next we formulate a "preconditioning theorem".

**Theorem 4.** *For each*[2] $R \in \mathbb{R}^{n \times m}$ *we have*

$$\mathbf{X}(\langle A_c, \Delta \rangle, \langle b_c, \delta \rangle) \subseteq \mathbf{X}(\langle RA_c, |R|\Delta \rangle,\ \langle Rb_c, |R|\delta \rangle). \tag{3.1}$$

Indeed, if $x \in \mathbf{X}(\langle A_c, \Delta \rangle, \langle b_c, \delta \rangle)$, then by the Oettli-Prager theorem (2.2) it satisfies

$$|A_c x - b_c| \leq \Delta |x| + \delta,$$

hence

$$|RA_c x - Rb_c| \leq |R||A_c x - b_c| \leq |R|(\Delta |x| + \delta) = |R|\Delta |x| + |R|\delta$$

which, again by the Oettli-Prager theorem, means that $x \in \mathbf{X}(\langle RA_c, |R|\Delta \rangle,\ \langle Rb_c, |R|\delta \rangle)$.

The transformation of the system

$$\langle A, \Delta \rangle x = \langle b, \delta \rangle \tag{3.2}$$

to the system

$$\langle RA_c, |R|\Delta \rangle x = \langle Rb_c, |R|\delta \rangle \tag{3.3}$$

is called preconditioning. If $R = A_c^\dagger$, we speak of preconditioning by the midpoint pseudoinverse. It is important to notice that whereas the system matrix of (3.2) is rectangular $m \times n$, that one of (3.3) is square $n \times n$.

# 4 Enclosures

Any interval vector $[x, y]$ satisfying

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

is called an enclosure of the solution set $\mathbf{X}(\mathbf{A}, \mathbf{b})$. If $\mathbf{A}$ has full column rank, then the solution set $\mathbf{X}(\mathbf{A}, \mathbf{b})$ is compact, thus bounded, and infinitely many enclosures exist. Any enclosure of the solution set of (3.3) is also an enclosure of the solution set of (3.2). This is the main idea behind preconditioning because the preconditioned system (3.3) may have a specific structure (as e.g. unit midpoint) which can be employed in construction of an efficient enclosure.

[2] "Preconditioner".

## 5 Interval hull

If $\mathbf{A}$ has full column rank, then the intersection of all enclosures

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

is called the interval hull of the solution set; obviously it is uniquely determined and it constitutes the minimal enclosure with respect to inclusion:

$$\mathbf{x}(\mathbf{A},\mathbf{b}) = \min_{\text{w.r.t. “}\subseteq\text{”}} \{\, [x,y] \mid \mathbf{X}(\mathbf{A},\mathbf{b}) \subseteq [x,y] \,\}.$$

Computing the interval hull is NP-hard [3], therefore in practical computations we usually resort to computing enclosures. Two obvious requirements (that, however, may contradict each other) are sufficient tightness of the resulting enclosure (i.e., closeness to the interval hull) and computability in polynomial time. This forms the subject matter of the next three sections.

## 6 Enclosure: rectangular case

First we give an explicit description of an enclosure for the case of a rectangular interval matrix $\mathbf{A}$ in (2.1).

**Theorem 5.** *Let $R \in \mathbb{R}^{n\times m}$ and $0 \le M \in \mathbb{R}^{n\times n}$ be arbitrary matrices satisfying*

$$M(I-G) \ge I, \tag{6.1}$$

*where $G = |I - RA_c| + |R|\Delta$. Then $\mathbf{A}$ has full column rank and we have*

$$\mathbf{X}(\mathbf{A},\mathbf{b}) \subseteq \frac{\langle d\circ x_c,\, M(|x_c| + |R|\delta) - |d\circ x_c|\rangle}{\langle d - d\circ r,\, d - |d\circ r| - h - e\rangle}, \tag{6.2}$$

*where*

$$\begin{aligned} d &= \operatorname{diag}(M), \\ r &= \operatorname{diag}(I - RA_c), \\ h &= \operatorname{diag}(M(I-G) - I), \\ x_c &= Rb_c. \end{aligned}$$

The right-hand side of (6.2) is the Hadamard division of two interval vectors written in the midpoint-radius notation and as such is again an interval vector which forms an enclosure of the solution set $\mathbf{X}(\mathbf{A},\mathbf{b})$. This is a vectorized version of the main result in [12] where it was formulated entrywise, and also for a square $\mathbf{A}$ only. But since both $G$ and $M$ are square, the proof carries over to rectangular interval matrices as well[3].

[3] I cannot resist disclosing here that the two-page paper [12] has been rejected by three journals as “too short”, “a note only”, “nothing new” etc., after which I resigned at its journal publication. Yet an attentive reader will probably notice that the proof given there is by no means trivial because it uses several tricks aimed at keeping the enclosure as tight as possible; this also explains the appearance of $r$ and $h$.

In the next theorem we show how to specify the choice of $R$ and $M$. This is the most useful formulation of all those presented here because it serves both rectangular and square cases.

**Theorem 6** *Let $A_c$ have full column rank and let $M = (I - |A_c^\dagger|\Delta)^{-1} \geq 0$. Then $\mathbf{A}$ has full column rank and denoting $d = \mathrm{diag}(M)$ and $x_c = A_c^\dagger b_c$, we have*

$$\mathbf{X}(\mathbf{A}, \mathbf{b}) \subseteq \frac{\langle d \circ x_c,\, M(|x_c| + |A_c^\dagger|\delta) - |d \circ x_c|\rangle}{\langle d,\, d - e\rangle}. \tag{6.3}$$

Indeed, since $A_c$ has full column rank, we have $A_c^\dagger = (A_c^T A_c)^{-1} A_c^T$, hence $R = A_c^\dagger$ and $M = (I - |A_c^\dagger|\Delta)^{-1}$ satisfy the assumptions of Theorem 5 and $r = 0$, $h = 0$, whereby (6.2) turns into (6.3).

The right-hand side expression in (6.3) still involves interval arithmetic operations. Getting rid of them, we obtain the following formulation.

**Theorem 7** *Let $A_c$ have linearly independent columns and let $M = (I - |A_c^\dagger|\Delta)^{-1} \geq 0$. Denoting $d = \mathrm{diag}(M)$ and $x_c = A_c^\dagger b_c$, we have*

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

*where*

$$\begin{aligned} \underset{\sim}{x} &= d \circ x_c - (M(|x_c| + |A_c^\dagger|\delta) - |d \circ x_c|), \\ \tilde{x} &= d \circ x_c + (M(|x_c| + |A_c^\dagger|\delta) - |d \circ x_c|). \end{aligned}$$

The rearrangement has been made possible due to the definition of interval arithmetic multiplication and the fact that $\underset{\sim}{x} \leq \tilde{x}$.

## 7 Enclosure: square case

From this section on we shall consider the case of a square interval matrix $\mathbf{A}$ in (2.1). Replacing $A_c^\dagger$ by $A_c^{-1}$ in Theorem 6, we obtain the following theorem.

**Theorem 8** *Let $A_c$ be nonsingular and let $M = (I - |A_c^{-1}|\Delta)^{-1} \geq 0$. Then each $A \in \mathbf{A}$ is nonsingular and denoting $d = \mathrm{diag}(M)$ and $x_c = A_c^{-1} b_c$, we have*

$$\mathbf{X}(\mathbf{A}, \mathbf{b}) \subseteq \frac{\langle d \circ x_c,\, M(|x_c| + |A_c^{-1}|\delta) - |d \circ x_c|\rangle}{\langle d,\, d - e\rangle}. \tag{7.1}$$

Inner bounds on the endpoints $\underline{x}$, $\overline{x}$ of the above enclosure were given in [9], p. 41, but the proof was never published.

## 8 Interval hull

Next we show that the enclosure (7.1) becomes the interval hull in the special case of the system matrix of the form

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

i.e., having unit midpoint.

**Theorem 9** *Let* $M = (I - \Delta)^{-1} \geq 0$. *Then each* $A \in \mathbf{I}$ *is nonsingular and denoting* $d = \mathrm{diag}(M)$, *we have*

$$\mathbf{x}(\mathbf{I}, \mathbf{b}) = \frac{\langle d \circ b_c,\ M(|b_c| + \delta) - |d \circ b_c| \rangle}{\langle d,\ d - e \rangle}. \tag{8.1}$$

It is instructive to compare the compact form of (8.1) with its original form in [8]. This comparison clearly shows what reduction in formulae and notation has been achieved due to use of interval arithmetic.

**Theorem 10** *Let* $\varrho(\Delta) < 1$. *Then for each* $i \in \{1, \ldots, n\}$ *we have*

$$\underline{x}_i = \min\{\underset{\sim}{x}_i, \nu_i \underset{\sim}{x}_i\},$$

$$\overline{x}_i = \max\{\tilde{x}_i, \nu_i \tilde{x}_i\},$$

*where*

$$\begin{aligned}
\underset{\sim}{x}_i &= -x_i^* + m_{ii}(b_c + |b_c|)_i \\
\tilde{x}_i &= \quad x_i^* + m_{ii}(b_c - |b_c|)_i \\
x_i^* &= (M(|b_c| + \delta))_i \\
\nu_i &= \frac{1}{2m_{ii} - 1} \in (0, 1]
\end{aligned}$$

*and*

$$M = (I - \Delta)^{-1} = (m_{ij}) \geq 0.$$

Another proof of Theorem 10 was given in [11].

## 9 Backward step: from hull to enclosure

To get a full picture, we show how Theorem 9 can be proved from Theorem 10. If $\mathbf{A}$ is square with nonsingular $A_c$, taking $R = A_c^{-1}$, from (3.1) we get

$$\mathbf{X}(\mathbf{A}, \mathbf{b}) \subseteq \mathbf{X}(\langle I, |A_c^{-1}|\Delta \rangle,\ \langle A_c^{-1} b_c, |A_c^{-1}|\delta \rangle)$$

and substituting $\Delta \mapsto |A_c^{-1}|\Delta$, $b_c \mapsto A_c^{-1} b_c = x_c$ and $\delta \mapsto |A_c^{-1}|\delta$ in Theorem 9, we obtain the assertion of Theorem 8.

## 10 The condition (6.1)

Last, we shall consider the condition (6.1) under which all the results work. Avoiding $G$, the condition can be written as

$$M(I - |I - RA_c| - |R|\Delta) \geq I. \tag{10.1}$$

For the square case it has been proved in [10] that *the inequality (10.1) has a solution $R$ and $M \geq 0$ if and only if $A_c$ is nonsingular and*

$$\varrho(|A_c^{-1}|\Delta) < 1 \tag{10.2}$$

*holds; if this is the case, then it also has the particular solution $R_0 = A_c^{-1}$, $M = (I - |A_c^{-1}|\Delta)^{-1} \geq 0$.* This shows that the condition (10.1), unfortunately, does not reach beyond the scope of strongly regular interval matrices. It remains a big challenge to cross this border.

# 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] M. Fiedler, J. Nedoma, J. Ramík, J. Rohn, and K. Zimmermann, *Linear Optimization Problems with Inexact Data*, Springer-Verlag, New York, 2006. 2, 4

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

[5] 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

[6] 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

[7] W. Oettli and W. Prager, *Compatibility of approximate solution of linear equations with given error bounds for coefficients and right-hand sides*, Numerische Mathematik, 6 (1964), pp. 405–409. 2

[8] 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, 6

[9] J. Rohn, *A handbook of results on interval linear problems*, 2005. Internet text available at `http://uivtx.cs.cas.cz/~rohn/publist/!handbook.pdf`. 5

[10] J. Rohn, *A characterization of strong regularity of interval matrices*, Electronic Journal of Linear Algebra, 20 (2010), pp. 717–722. `http://www.math.technion.ac.il/iic/ela/ela-articles/articles/vol20_pp717-722.pdf`. 7

[11] 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`. 6

[12] J. Rohn, *A perturbation theorem for linear equations*, Technical Report 1103, Institute of Computer Science, Academy of Sciences of the Czech Republic, Prague, January 2011. `http://uivtx.cs.cas.cz/~rohn/publist/perturb.pdf`. 4