On the Solution of Linear Algebraic Equations Involving Interval Coefficients *

S. Markov, E. Popova and Ch. Ullrich†
Institute of Biophysics, Bulgarian Academy of Sciences
e-mail: smarkov@bgearn.bitnet
† Institute of Informatics, University of Basel
e-mail: ullrich@ifi.unibas.ch

Abstract

We discuss the solution to the interval algebraic system $\mathbf{A}\times\mathbf{x} = \mathbf{b}$ involving interval $n\times n$ matrix $\mathbf{A}$ and interval vector $\mathbf{b}$ in directed interval arithmetic involving improper intervals. We give some new relations for directed intervals, which form the basis for a directed interval matrix algebra. Using such relations we prove convergence of an iterative method, formulated by L. Kupriyanova, under simple explicit conditions on the interval matrix $\mathbf{A}$. We propose an iterative numerical algorithm for the solution to a class of interval algebraic systems $\mathbf{A}\times\mathbf{x} = \mathbf{b}$. Cramer-type formula for a special case of real matrices and interval right-hand side are used for the computation of an initial approximation for the iteration method. A *Mathematica* function performing the proposed algorithm is described.

1 Introduction

A linear algebraic system $\mathbf{A}x = \mathbf{b}$ involving intervals in the $n \times n$-matrix $\mathbf{A}$ and/or in the right-hand side n-vector $\mathbf{b}$, relates to four different problems, resp. solution sets [14]–[16], [20]:

i) the *united solution set* is the set of solutions of all real (degenerate, thin) systems $Ax = b$ with $A \in \mathbf{A}$ and $b \in \mathbf{b}$, i. e.,

$$\begin{aligned}\Sigma_{\exists\exists}(\mathbf{A},\mathbf{b}) &= \{x \in R^n \mid (\exists A \in \mathbf{A})(\exists\, b \in \mathbf{b})(Ax = b)\} \\ &= \{\, x \in R^n \mid \mathbf{A}x \cap \mathbf{b} \neq \emptyset\,\};\end{aligned} \tag{1}$$

ii) the *tolerable solution set* is the set of all real vectors x such that for every real $A \in \mathbf{A}$ the real vector Ax is contained in the interval vector $\mathbf{b}$, that is

$$\begin{aligned}\Sigma_{\forall\exists}(\mathbf{A},\mathbf{b}) &= \{\, x \in R^n \mid (\forall A \in \mathbf{A})(\exists\, b \in \mathbf{b})(Ax = b)\,\} \\ &= \{\, x \in R^n \mid (\forall A \in \mathbf{A})(Ax \in \mathbf{b})\} = \{\, x \in R^n \mid \mathbf{A}x \subseteq \mathbf{b}\};\end{aligned} \tag{2}$$

*This paper has been partly supported by the Swiss National Science Foundation and published in: *Iterative Methods in Linear Algebra, II*, S. Margenov, P. Vassilevski (Eds.), IMACS Series in Computational and Applied Mathematics, **3** 1996, pp. 216–225.

iii) the *controlled solution set* is the set of all real vectors $x \in R^n$, such that for any $b \in \mathbf{b}$ we can find the corresponding $A \in \mathbf{A}$ satisfying $Ax = b$ (see [14]);

$$\begin{aligned} \Sigma_{\exists\forall}(\mathbf{A}, \mathbf{b}) &= \{ x \in R^n \mid (\exists A \in \mathbf{A})(\forall b \in \mathbf{b})(Ax = b) \} \\ &= \{ x \in R^n \mid (\forall b \in \mathbf{b})(\mathbf{A}x \ni b) = \{x \in R^n \mid \mathbf{A}x \supseteq \mathbf{b}\}; \end{aligned} \qquad (3)$$

iv) the *interval algebraic solution* is an interval vector $\mathbf{x}$ which substituted in the expression $\mathbf{A} \times \mathbf{x}$, using interval arithmetic, results in $\mathbf{b}$, that is

$$\mathbf{A} \times \mathbf{x} = \mathbf{b}. \qquad (4)$$

When talking about interval algebraic solution we have to specify the interval arithmetic used. The interval algebraic solutions do not exist in general in the ordinary interval space [13]. An appropriate interval arithmetic for solving algebraic equations is the generalized interval arithmetic involving improper intervals. We shall call this arithmetic directed interval arithmetic, since the support set of this arithmetic is the set of both proper *and* improper intervals, which we call *directed* intervals [9]–[10]. Directed interval arithmetic is the natural arithmetic for the solution of algebraic equations, since it is obtained from the arithmetic for normal intervals via algebraic completion. Solving interval algebraic equations in ordinary interval arithmetic can be compared to solving real linear algebraic (systems of) equations using only positive real numbers.

It has been noticed by S. Shary and by some other authors [3], [6], [14]–[16], [20] that the (directed) interval algebraic solution (4) is closely related to (some of) the solutions of problems i)–iii). Also, several convenient formulas for symbolic manipulation with directed intervals have been found [2], [8]–[10]. On the other side S. Shary [15], A. Zakharov [18]–[20], V. Zyuzin [21], [22] and L. Kupriyanova [6] formulated some computational tools for the solution to (4). Based on these investigations, the purpose of this work is:

i) to present several new useful rules for algebraic manipulations in directed interval (matrix) arithmetic;

ii) using these rules to give a new proof for convergence of an iteration method, proposed by L. Kupriyanova, under explicit conditions on the input data $\mathbf{A}$;

iii) to formulate Cramer-type formulas for the solution of problem (4) for a special case of real matrix;

iv) using the obtained results to formulate a numerical algorithm and implement it in the computer algebra system *Mathematica*.

2 Directed interval arithmetic

Directed interval arithmetic [3]–[5], [8]–[10] presents an algebraic completion of the interval arithmetic for compact intervals on the real line, providing thus a convenient tool for solving interval algebraic equations. We give below some basic concepts of directed interval arithmetic.

Denote by $D = \{[a, b] \mid a, b \in R\}$ the set of ordered couples of real numbers. We call the elements of D directed (or generalized) intervals and the real numbers a, b are called their components. Directed intervals will be denoted by boldface letters. The first component of

$\mathbf{a} \in D$ is further denoted by a^-, and the second by a^+, so that $\mathbf{a} = [a^-, a^+]$. Thus $a^\lambda \in R$ with $\lambda \in \Lambda = \{+,-\}$ is the first or second component of $\mathbf{a} \in D$ depending on the value of λ. In what follows the binary variable λ will be sometimes expressed as a "product" of two binary variables, $\lambda = \mu\nu$, $\mu, \nu \in \Lambda$, assuming that $++ = -- = +$, $+- = -+ = -$. The directed interval $\mathbf{a} = [a^-, a^+]$ is called proper (normal, ordinary), if $a^- \leq a^+$, degenerated if $a^- = a^+$ (in this case we write $\mathbf{a} = a = [a,a]$), and improper if $a^- > a^+$. The set of all proper intervals is denoted by $I(R)$, the set of degenerated intervals is equivalent to R and is denoted again by R, and the set of improper intervals is denoted by $\overline{I(R)}$. We have $D = I(R) \cup \overline{I(R)}$.

The directed interval $\mathbf{a} = [a^-, a^+] \in D$ defines a binary variable "direction" by $\tau(\mathbf{a}) = \{+,$ if $a^- \leq a^+; -,$ if $a^- > a^+\}$. To every directed interval $\mathbf{a} = [a^-, a^+] \in D$ we assign the proper interval $a = \mathrm{pro}(\mathbf{a}) = \{[a^-, a^+]$, if $\tau(\mathbf{a}) = +$; $[a^+, a^-]$, if $\tau(\mathbf{a}) = -\}$; we have $\mathrm{pro}(\mathbf{a}) = [a^{-\tau(\mathbf{a})}, a^{\tau(\mathbf{a})}]$. Following S. Shary we shall call the proper interval $\mathrm{pro}(\mathbf{a}) = a$ the *proper projection* of $\mathbf{a}$. Obviously, $a^{-\tau(\mathbf{a})} < a^{\tau(\mathbf{a})}$.

Let $Z = \{\mathbf{a} \in I(R) \mid a^- \leq 0 \leq a^+\}$, $Z^* = \{\mathbf{a} \in I(R) \mid a^- < 0 < a^+\}$, $\overline{Z} = \{\mathbf{a} \in \overline{I(R)} \mid a^+ \leq 0 \leq a^-\}$, $\overline{Z}^* = \{\mathbf{a} \in \overline{I(R)} \mid a^+ < 0 < a^-\}$, $\mathcal{T} = Z \cup \overline{Z}$, $\mathcal{T}^* = Z^* \cup \overline{Z}^*$. In $D^* = D \setminus \mathcal{T}^*$ we define the functional "sign" of a directed interval $\sigma : D^* \to \Lambda$, for $\mathbf{a} \in D^* \setminus \{0\}$ by $\sigma(\mathbf{a}) = \{+,$ if $a^- \geq 0$ and $a^+ \geq 0$; $-,$ if $a^- \leq 0$ and $a^+ \leq 0\}$, and for zero argument by $\sigma([0,0]) = \sigma(0) = +$. In particular, σ is well defined over R. The sign σ is not defined for intervals from $\mathcal{T}^*$. Obviously, $\sigma(\mathbf{a}) = \sigma(p(\mathrm{pro}(\mathbf{a})) = \sigma(a)$.

The operations $+, \times$ are defined in D by:

$$\mathbf{a} + \mathbf{b} = [a^- + b^-, a^+ + b^+], \quad \mathbf{a}, \mathbf{b} \in D, \tag{5}$$

$$\mathbf{a} \times \mathbf{b} = \begin{cases} [a^{-\sigma(b)} b^{-\sigma(a)},\ a^{\sigma(b)} b^{\sigma(a)}], & \mathbf{a}, \mathbf{b} \in D^*, \\ [a^{\delta} b^{-\delta},\ a^{\delta} b^{\delta}], \quad \delta = \sigma(a), & \mathbf{a} \in D^*,\ \mathbf{b} \in \mathcal{T}^*, \\ [a^{-\delta} b^{\delta},\ a^{\delta} b^{\delta}], \quad \delta = \sigma(b), & \mathbf{a} \in \mathcal{T}^*,\ \mathbf{b} \in D^*. \end{cases} \tag{6}$$

$$\mathbf{a} \times \mathbf{b} = \begin{cases} [\min\{a^- b^+, a^+ b^-\},\ \max\{a^- b^-, a^+ b^+\}], & \mathbf{a}, \mathbf{b} \in Z^*, \\ [\max\{a^- b^-, a^+ b^+\},\ \min\{a^- b^+, a^+ b^-\}], & \mathbf{a}, \mathbf{b} \in \overline{Z}^*, \\ 0, \quad (\mathbf{a} \in Z^*,\ \mathbf{b} \in \overline{Z}^*) \vee (\mathbf{a} \in \overline{Z}^*,\ \mathbf{b} \in Z^*). \end{cases} \tag{7}$$

Formulae (6), (7), presenting component-wise the product in D are equivalent to the definitions given in [4], [5]. Recall that the arithmetic operations for normal intervals $[a,b], [c,d] \in I(R)$ are defined by $[a,b] \star [c,d] = \{x \star y \mid a \leq x \leq b, c \leq y \leq d\}$, $\star \in \{+,-,\times,/\}$ [1]. The restrictions of the expressions (5), (6)–(7) in $I(R)$ produce the familiar addition, resp. multiplication, for normal intervals.

We shall omit the symbol $\times$ if one of the multipliers in a product $\mathbf{a} \times \mathbf{b}$ is degenerate. From (6) for $\mathbf{a} = [a,a] = a \in R$, $\mathbf{b} \in D$ we have $a \times \mathbf{b} = a\mathbf{b} = [ab^{-\sigma(a)}, ab^{\sigma(a)}]$. The operator *negation* (negative element) is defined by $-\mathbf{b} = (-1)\mathbf{b} = [-b^+, -b^-]$. The restriction of the composite operation $\mathbf{a} + (-1)\mathbf{b} = \mathbf{a} + (-\mathbf{b}) = [a^- - b^+, a^+ - b^-]$, for $\mathbf{a}, \mathbf{b} \in I(R)$ is the familiar subtraction of intervals.

The algebraic systems $(D, +)$ and $(D \setminus \mathcal{T}, \times)$ are groups [4], [5]. Denote by $-_h\mathbf{a}$ the inverse element of $\mathbf{a} \in D$ with respect to addition, called *opposite element*, and by $1/_h\mathbf{a}$ the inverse element of $\mathbf{a} \in D \setminus \mathcal{T}$ with respect to "$\times$". For the inverse elements we have the component-wise presentations $-_h\mathbf{a} = [-a^-, -a^+]$, for $\mathbf{a} \in D$, and $1/_h\mathbf{a} = [1/a^-, 1/a^+]$, for $\mathbf{a} \in D \setminus \mathcal{T}$.

The group properties of $(D,+)$, $(D \setminus \mathcal{T}, \times)$ allow solving of algebraic equations. For $\mathbf{a}, \mathbf{b} \in D$ the unique solution to the equation $\mathbf{a} + \mathbf{x} = \mathbf{b}$ is $\mathbf{x} = \mathbf{b} + (-_h \mathbf{a})$. Similarly, for $\mathbf{a} \in D \setminus \mathcal{T}, \mathbf{b} \in D$ the unique solution to the equation $\mathbf{a} \times \mathbf{y} = \mathbf{b}$ is $\mathbf{y} = \mathbf{b} \times (1/_h \mathbf{a})$.

In addition to the operators negation $-\mathbf{a} = [-a^+, -a^-]$ and opposite element $-_h \mathbf{a} = [-a^-, -a^+]$ we define the operator *dual element* by $\overline{\mathbf{a}} = \overline{[a^-, a^+]} = [a^+, a^-]$. The negative, the opposite and the dual elements are related in the following way:

$$\overline{\mathbf{a}} = -_h(-\mathbf{a}) = -(-_h \mathbf{a}). \tag{8}$$

The operator inversion defined in $D \setminus \mathcal{T}$ by analogy to the operator $-\mathbf{a}$ in (8) as $1/\mathbf{a} = \overline{1/_h \mathbf{a}} = 1/_h \overline{\mathbf{a}} = [1/a^+, 1/a^-]$, satisfies

$$1/_h(1/\mathbf{a}) = 1/(1/_h \mathbf{a}) = \overline{\mathbf{a}}. \tag{9}$$

The operation division $\mathbf{a} \times (1/\mathbf{b})$ for $\mathbf{a} \in D$, $\mathbf{b} \in D \setminus \mathcal{T}$, is denoted by $\mathbf{a}/\mathbf{b}$. From (8) and (9) we obtain for the inverse operators w. r. t. addition and multiplication $-_h \mathbf{a} = -\overline{\mathbf{a}}$, resp., $1/_h \mathbf{a} = 1/\overline{\mathbf{a}}$. Using these presentations, we can present the unique solutions of the equations $\mathbf{a} + \mathbf{x} = \mathbf{b}$, $\mathbf{a} \times \mathbf{y} = \mathbf{b}$, $\mathbf{a} + \mathbf{c} \times \mathbf{z} = \mathbf{b}$ in the form $\mathbf{x} = \mathbf{b} + (-\overline{\mathbf{a}}) = \mathbf{b} - \overline{\mathbf{a}}$, resp. $\mathbf{y} = \mathbf{b} \times (-\overline{\mathbf{a}}) = \mathbf{b}/\overline{\mathbf{a}}$, $\mathbf{z} = (\mathbf{b} - \overline{\mathbf{a}})/\overline{\mathbf{c}}$. The inverse elements $-_h \mathbf{a}$, $1/_h \mathbf{a}$ generate the operations $\mathbf{a} -_h \mathbf{b} = \mathbf{a} + (-_h \mathbf{b}) = \mathbf{a} + (-\overline{\mathbf{b}}) = \mathbf{a} - \overline{\mathbf{b}}$, $\mathbf{a}/_h \mathbf{b} = \mathbf{a} \times (1/_h \mathbf{b}) = \mathbf{a} \times (1/\overline{\mathbf{b}}) = \mathbf{a}/\overline{\mathbf{b}}$.

To summarize, the algebra $\mathcal{K} = (D, +, \times)$ involves the operations subtraction $\mathbf{a} - \mathbf{b}$, division $\mathbf{a}/\mathbf{b}$, the operator dual element $\overline{\mathbf{a}}$, and the operations $\mathbf{a} - \overline{\mathbf{b}}$, $\mathbf{a}/\overline{\mathbf{b}}$, $\overline{\mathbf{a}} - \mathbf{b}$, $\overline{\mathbf{a}}/\mathbf{b}$. Similarly, we can compose $\mathbf{a} + \overline{\mathbf{b}}$, $\mathbf{a} \times \overline{\mathbf{b}}$, $\overline{\mathbf{a}} + \mathbf{b}$, $\overline{\mathbf{a}} \times \mathbf{b}$ etc.

In addition to the notation $\overline{\mathbf{a}}$, for the operator dual element we shall also use the notation $\mathbf{a}_- = \overline{\mathbf{a}}$. For the sake of uniformity we shall write $\mathbf{a}_+ = \mathbf{a}$. Using the functional notation $\mathbf{a}_\lambda$ with $\lambda = \{+, -\}$, we can formulate a simple distributive-like relation in D^*, which is more convenient than the one formulated in table form by E. Kaucher [5].

Proposition 2.1 *(Conditionally Distributive Law) For* $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{a} + \mathbf{b} \in D^*$ *we have*

$$(\mathbf{a} + \mathbf{b}) \times \mathbf{c}_{\sigma(a+b)} = (\mathbf{a} \times \mathbf{c}_{\sigma(a)}) + (\mathbf{b} \times \mathbf{c}_{\sigma(b)}). \tag{10}$$

Relation (10) can be written in the following equivalent forms [2], [8]

$$\begin{aligned}
(\mathbf{a} + \mathbf{b}) \times \mathbf{c} &= (\mathbf{a} \times \mathbf{c}_{\sigma(a)\sigma(a+b)}) + (\mathbf{b} \times \mathbf{c}_{\sigma(b)\sigma(a+b)}) \\
&= \begin{cases} (\mathbf{a} \times \mathbf{c}) + (\mathbf{b} \times \mathbf{c}), & \sigma(\mathbf{a}) = \sigma(\mathbf{b}) \ \ (= \sigma(\mathbf{a} + \mathbf{b})), \\ (\mathbf{a} \times \mathbf{c}) + (\mathbf{b} \times \overline{\mathbf{c}}), & \sigma(\mathbf{a}) = -\sigma(\mathbf{b}),\ \sigma(\mathbf{a} + \mathbf{b}) = \sigma(\mathbf{a}), \\ (\mathbf{a} \times \overline{\mathbf{c}}) + (\mathbf{b} \times \mathbf{c}), & \sigma(\mathbf{a}) = -\sigma(\mathbf{b}),\ \sigma(\mathbf{a} + \mathbf{b}) = \sigma(\mathbf{b}). \end{cases}
\end{aligned}$$

Note that, $\sigma(\mathbf{a}) = \sigma(a) = \sigma(\mathrm{pro}(\mathbf{a}))$. Moreover, we have $\sigma(\mathbf{a} + \mathbf{b}) = \sigma(a + b)$.

Corrolary 2.1 *In the special case when* $\mathbf{c}$ *is degenerate,* $\mathbf{c} = c \in R$*, (10) becomes*

$$c(\mathbf{a} + \mathbf{b}) = c\mathbf{a} + c\mathbf{b}. \tag{11}$$

In the special case when in (10) the intervals $\mathbf{a}$ *and* $\mathbf{b}$ *are degenerate,* $\mathbf{a} = a \in R$*,* $\mathbf{b} = b \in R$*, we have*

$$(a + b)\mathbf{c}_{\sigma(a+b)} = a\mathbf{c}_{\sigma(a)} + b\mathbf{c}_{\sigma(b)}. \tag{12}$$

From the conditionally-distributive law we see that if $\sigma(\mathbf{a}) = \sigma(\mathbf{b})$, we have $(\mathbf{a}\times\mathbf{c})+(\mathbf{b}\times\mathbf{c}) = (\mathbf{a}+\mathbf{b})\times\mathbf{c}$, that is we can extract the (common) multiplier $\mathbf{c}$ out of brackets in this case. In the case $\sigma(\mathbf{a}) = -\sigma(\mathbf{b})$ we cannot extract $\mathbf{c}$ out of the brackets. Denote by $\nu : D \longrightarrow R$ the functional defined by $\nu(\mathbf{a}) = \mathbf{a} + \overline{\mathbf{a}} = a^- + a^+$. We have $\nu(\mathbf{a}\times\mathbf{b}) = (\mathbf{a}\times\mathbf{b}) + \overline{(\mathbf{a}\times\mathbf{b})} = (\mathbf{a}\times\mathbf{b})+(\overline{\mathbf{a}}\times\overline{\mathbf{b}})$. In particular, for $a \in R$, $\mathbf{b} \in D$, $\nu(a\mathbf{b}) = a\mathbf{b}+a\overline{\mathbf{b}} = a(\mathbf{b}+\overline{\mathbf{b}}) = a\nu(\mathbf{b}) \in R$.

Proposition 2.2 *For* $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{a}+\mathbf{b}, \mathbf{a}+(-\mathbf{b}) \in D^*$ *we have*
i) if $\sigma(\mathbf{a}) = \sigma(\mathbf{b})$ *then*

$$(\mathbf{a}\times\mathbf{c}) + (\mathbf{b}\times\mathbf{c}) = (\mathbf{a}+\mathbf{b})\times\mathbf{c}; \tag{13}$$

ii) if $\sigma(\mathbf{a}) = -\sigma(\mathbf{b})$ *then* $(\mathbf{a}\times\mathbf{c})+(\mathbf{b}\times\mathbf{c})$ *is equivalent to one of the following expressions*

$$\begin{array}{ll} (\mathbf{a}+(-\mathbf{b}))\times\mathbf{c}+\nu(\mathbf{b}\times\mathbf{c}), & (\mathbf{a}+(-\overline{\mathbf{b}}))\times\mathbf{c}+\nu(\mathbf{c})\mathbf{b}; \\ ((-\mathbf{a})+\mathbf{b})\times\mathbf{c}+\nu(\mathbf{a}\times\mathbf{c}), & ((-\overline{\mathbf{a}})+\mathbf{b})\times\mathbf{c}+\nu(\mathbf{c})\mathbf{a}. \end{array} \tag{14}$$

We shall also need to be able to extract $\mathbf{c}$ or $\overline{\mathbf{c}}$ out of brackets in expressions of the form $(\mathbf{a}\times\mathbf{c}) + (\mathbf{b}\times\overline{\mathbf{c}})$. The following proposition holds true

Proposition 2.3 *For* $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{a}+\mathbf{b}, \mathbf{a}+(-\mathbf{b}) \in D^*$ *we have*
i) if $\sigma(\mathbf{a}) = -\sigma(\mathbf{b})$ *then*

$$(\mathbf{a}\times\mathbf{c}) + (\mathbf{b}\times\overline{\mathbf{c}}) = \begin{cases} (\mathbf{a}+\mathbf{b})\times\mathbf{c}, \ \sigma(\mathbf{a}+\mathbf{b}) = \sigma(\mathbf{a}), \\ (\mathbf{a}+\mathbf{b})\times\overline{\mathbf{c}}, \ \sigma(\mathbf{a}+\mathbf{b}) = \sigma(\mathbf{b}); \end{cases} \tag{15}$$

ii) if $\sigma(\mathbf{a}) = \sigma(\mathbf{b}) = \sigma^*$ *then*

$$(\mathbf{a}\times\mathbf{c}) + (\mathbf{b}\times\overline{\mathbf{c}}) = \begin{cases} (\mathbf{a}+(-\mathbf{b}))\times\mathbf{c}+\nu(\mathbf{b}\times\overline{\mathbf{c}}), & \sigma(\mathbf{a}+(-\mathbf{b})) = \sigma^* \\ ((-\mathbf{a})+\mathbf{b})\times\mathbf{c}+\nu(\mathbf{a}\times\mathbf{c}), & \sigma(\mathbf{a}+(-\mathbf{b})) = -\sigma^*. \end{cases}$$

The special case $\mathbf{a}, \mathbf{b} \in R$ in Propositions 2.2 and 2.3 can be summarized in the following corrolary

Corrolary 2.2 *For* $a, b \in R$, $\mathbf{c} \in D^*$ *we have*

$$a\mathbf{c} + b\mathbf{c} = \begin{cases} (a+b)\mathbf{c}, & \sigma(a) = \sigma(b), \\ (b-a)\mathbf{c} + a\nu(\mathbf{c}), & \sigma(a) = -\sigma(b); \end{cases}$$

$$a\mathbf{c} + b\overline{\mathbf{c}} = \begin{cases} (a+b)\mathbf{c}, & \sigma(a) = -\sigma(b) = \sigma(a+b), \\ (a+b)\overline{\mathbf{c}}, & \sigma(a) = -\sigma(b) = -\sigma(a+b), \\ 0, & a = -b, \\ (b-a)\mathbf{c} + a\nu(\mathbf{c}), & \sigma(a) = \sigma(b) = \sigma(a-b). \end{cases}$$

Remark. The expression $(b-a)\mathbf{c}+a\nu(\mathbf{c})$ to be met twice in the right hand sides of the above equalities can be replaced by the expression $(a-b)\mathbf{c} + b\nu(\mathbf{c})$.

Operations between matrices of directed intervals are defined similarly to matrix operations involving normal intervals. Sum (difference) of two interval matrices of identical size is an interval matrix of the same size formed by component-wise sums (differences). If

$\mathbf{A} = (\mathbf{a}_{ij}) \in D^{m \times l}$ and $\mathbf{B} = (\mathbf{b}_{ij}) \in D^{l \times n}$, then the product of the directed interval matrices $\mathbf{A}$ and $\mathbf{B}$ is the matrix $\mathbf{C} = (\mathbf{c}_{ij}) \in R^{m \times n}$ with $\mathbf{c}_{ij} = \sum_{k=1}^{l} \mathbf{a}_{ik}\mathbf{b}_{kj}$. This defines, in particular, problem (4) where the expression $\mathbf{A} \times \mathbf{x}$ is a product of two interval matrices: namely, $\mathbf{A}$, $\mathbf{x}$ are directed interval matrices of order $n \times n$, $n \times 1$ resp. and the result $\mathbf{y} = \mathbf{A} \times \mathbf{x}$ is a $n \times 1$ directed interval matrix.

Although D is not a linear space it can be normed in the usual way [5]. As a norm we may take $\| \mathbf{x} \| = \max\{| x^- |, | x^+ |\}$. The norm is defined for vectors and matrices in the usual way. For instance, for $\mathbf{A} = (\mathbf{a}_{ik}) \in D^{n \times n}$, $\| \mathbf{A} \| = \max_i \{\sum_{k=1}^{n} \| \mathbf{a}_{ik} \|\}$. For the product of two interval matrices we have $\| \mathbf{A} \times \mathbf{B} \| \leq \| \mathbf{A} \| \| \mathbf{B} \|$. A metric in D^n is defined by $\| \mathbf{x} -_h \mathbf{y} \| = \| \mathbf{x} - \overline{\mathbf{y}} \|$ for $\mathbf{x}, \mathbf{y} \in D^n$. The following proposition holds true

Proposition 2.4 *For* $\mathbf{a}, \mathbf{b}, \mathbf{c} \in D^n$, $\| (\mathbf{c} \times \mathbf{a}) -_h (\mathbf{c} \times \mathbf{b}) \| \leq \| \mathbf{c} \| \| \mathbf{a} -_h \mathbf{b} \|$.

In the sequel we shall make use of the following fixed-point theorem, which is a generalization of the fixed-point theorem given in [7]

Proposition 2.5 *Let* $\mathbf{U} : D_1 \to D_1$, $D_1 \subseteq D^n$, *be a contraction mapping in the sense that there exists* $q \in R$, $0 < q < 1$, *such that* $\| \mathbf{U}(\mathbf{x}) -_h \mathbf{U}(\mathbf{y}) \| < q \| \mathbf{x} -_h \mathbf{y} \|$, *for every* $\mathbf{x}, \mathbf{y} \in D_1$. *Then* $\mathbf{U}$ *possesses a fixed-point* $\mathbf{x}^* \in D_1$ *which is the limit of the sequence* $\mathbf{x}^{(l+1)} = \mathbf{U}(\mathbf{x}^{(l)})$, $l = 0, 1, \ldots$, *with any* $\mathbf{x}^{(0)} \in D_1$.

The proof follows the classical proof using properties of "$-_h$" such as $(\mathbf{a} -_h \mathbf{b}) + (\mathbf{b} -_h \mathbf{c}) = \mathbf{a} -_h \mathbf{c}$.

3 Applications to systems of interval algebraic equations

For $\mathbf{A} \in D^{n \times n}$ we denote $\mathbf{D}(\mathbf{A}) = (\mathbf{d}_{ij})$ with $\mathbf{d}_{ii} = \mathbf{a}_{ii}$, $\mathbf{d}_{ij} = 0, i \neq j$. For $\mathbf{D} = \mathbf{D}(\mathbf{A})$ denote $\mathbf{D}^{-1} = (\mathbf{d}^*_{ij})$, $\mathbf{d}^*_{ij} = 1/\overline{\mathbf{a}_{ii}}$, $\mathbf{d}^*_{ij} = 0, i \neq j$. Clearly $\mathbf{D}^{-1} \times \mathbf{D} = 1$. In [6] the following iteration method has been proposed

$$\mathbf{x}_i \ := \ (\mathbf{b}_i -_h \sum_{j=1, j \neq i}^{n} \mathbf{a}_{ij} \times \mathbf{x}_j)/\overline{\mathbf{a}_{ii}}, \quad i = 1, \ldots, n. \tag{16}$$

It has been proved [6] that the iterative process (16) converges to the solution to (4) under special restrictions on the input data $\mathbf{A}$, $\mathbf{b}$ given in implicit form and a rather restrictive choice of the initial approximation. We formulate explicit conditions on the matrix $\mathbf{A}$ and prove that under these conditions (16) converges to the solution to (4) with an arbitrary initial approximation and arbitrary right-hand side $\mathbf{b}$. We first rewrite (16) in matrix form:

$$\mathbf{x} \ := \ \mathbf{D}^{-1} \times (\mathbf{b} -_h (\mathbf{A} -_h \mathbf{D}) \times \mathbf{x}), \qquad \mathbf{D} = \mathbf{D}(\mathbf{A}). \tag{17}$$

Proposition 3.1 *If* $\| \mathbf{D}(\mathbf{A})^{-1} \| < q < 1$, $\| \mathbf{A} -_h \mathbf{D}(\mathbf{A}) \| < q < 1$, *then (4) has a solution* $\mathbf{x}^* \in D^n$ *and method (17) converges to* $\mathbf{x}^*$ *for any* $\mathbf{b} \in D^n$ *and any initial approximation* $\mathbf{x}^{(0)} \in D^n$.

Proof. For $\mathbf{x} \in D^n$ denote $\mathbf{B}(\mathbf{x}) = \mathbf{D}^{-1} \times (\mathbf{b} -_h (\mathbf{A} -_h \mathbf{D}) \times \mathbf{x})$. For $\mathbf{x}, \mathbf{y} \in D^n$ we have

$$
\begin{aligned}
\| \mathbf{B}(\mathbf{x}) -_h \mathbf{B}(\mathbf{y}) \| &= \| \mathbf{D}^{-1} \times (\mathbf{b} -_h (\mathbf{A} -_h \mathbf{D}) \times \mathbf{x}) -_h \mathbf{D}^{-1} \times (\mathbf{b} -_h (\mathbf{A} -_h \mathbf{D}) \times \mathbf{y}) \| \\
&\leq \| \mathbf{D}^{-1} \| \| (\mathbf{b} -_h (\mathbf{A} -_h \mathbf{D}) \times \mathbf{x}) -_h (\mathbf{b} -_h (\mathbf{A} -_h \mathbf{D}) \times \mathbf{y}) \| \\
&= \| \mathbf{D}^{-1} \| \| (\mathbf{A} -_h \mathbf{D}) \times \mathbf{y} -_h (\mathbf{A} -_h \mathbf{D}) \times \mathbf{x} \| \\
&\leq \| \mathbf{D}^{-1} \| \| \mathbf{A} -_h \mathbf{D} \| \| \mathbf{y} -_h \mathbf{x} \| < q^2 \| \mathbf{y} -_h \mathbf{x} \|,
\end{aligned}
$$

using Proposition 2.4. The inequality $\| \mathbf{B}(\mathbf{x}) -_h \mathbf{B}(\mathbf{y}) \| < q^2 \| \mathbf{y} -_h \mathbf{x} \|$ shows that $\mathbf{B}$ is a contraction mapping. This combined with Proposition 2.5 proves the theorem. □

Proposition 3.2 *Let* $\mathbf{A} = A = (a_{i,k}) \in R^{n \times n}$ *be a real matrix and let the numbers* $a_{i,k}\Delta_{i,k}$, *where* $\Delta_{i,k}$ *is the subdeterminant of* $a_{i,k}$, *have constant signs for all* $i, k = 1, 2, ..., n$. *Then for the solution to* $A\mathbf{x} = \mathbf{b}$ *the following Cramer-type formula holds:*

$$
(\mathbf{x}_i)_{\sigma(\Delta)} = \frac{1}{\Delta} \sum_{i=1}^{n} (-1)^{i+k} \Delta_{ik} (\mathbf{b}_i)_{\lambda_{i,k}} \stackrel{Def}{=} \frac{1}{\Delta} \begin{vmatrix} a_{11} & \dots & \mathbf{b}_1 & \dots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{1n} & \dots & \mathbf{b}_n & \dots & a_{nn} \end{vmatrix}, \quad (18)
$$

where $\lambda_{i,k} = (-)^{i+k} = \{+,\ i+k\ \text{even};\ -,\ i+k\ \text{odd}\}$.

The proof is obtained using the propositions in section 2. A class of matrices satisfying the conditions of the theorem is the class of Wandermond matrices, appearing in interpolation theory. This makes the above formula suitable for the exact solution of identification problems in an interval interpolation setting [9, 11]. If the conditions of the theorem do not hold, formula (18) can still be used for the computation of an initial approximation to the solution as proposed in the next section.

4 Numerical algorithm

The results of the previous section allow us to formulate the following

Numerical Algorithm

1. Check the conditions $\| \mathbf{D}^{-1} \| < 1$, $\| \mathbf{A} -_h \mathbf{D} \| < 1$.

2. Compute an initial approximation $\mathbf{x}^{(0)}$ applying formulae (18) for the interval algebraic problem mid$(\mathbf{A})\mathbf{x} = \mathbf{b}$.

3. Using $\mathbf{x}^{(0)}$ iterate according to

$$
\mathbf{x}^{(k+1)} := \mathbf{D}^{-1} \times \left(\mathbf{b} -_h (\mathbf{A} -_h \mathbf{D}) \times \mathbf{x}^{(k)} \right), \quad k = 0, 1, \dots
$$

The above iterative method was implemented in the computer algebra system *Mathematica* [17] using an experimental package for generalized interval arithmetic [12]. A Mathematica function `AlgebraicIntervalSolve[m_, b_, opts___]` returns a list with the computed approximations for the algebraic solution to a given interval system with matrix $\mathbf{A}$ and right-hand-side vector $\mathbf{b}$ of directed intervals.

As initial approximation, `AlgebraicIntervalSolve` uses an appropriate initial approximation of the algebraic solution. `AlgebraicIntervalSolve` automatically starts iterations from the solution (18) to the system $\mathrm{mid}(\mathbf{A})\mathbf{x} = \mathbf{b}$ where $\mathrm{mid}(\mathbf{A}) = \left((a_{ij}^- + a_{ij}^+)/2\right)$. One can give the `AlgebraicIntervalSolve` an other initial approximation of the algebraic solution by using the option `InitialApproximation`.

Several options can be used to control the number of iterations performed by the function `AlgebraicIntervalSolve`. First, one can set `MaxIterations` to specify the maximum number of iterations that `AlgebraicIntervalSolve` should use. If `AlgebraicIntervalSolve` does not find a good solution in the number of steps that have been specified, it returns the last values that have been computed. These values can be used as `InitialApproximation` if one needs to continue the iterations. To check if an acceptable solution has been found, `AlgebraicIntervalSolve` iterates and sees whether the differences in the end-points of the interval components between two successive approximations of the algebraic solution are within the accuracy specified by the option `AccuracyGoal`.

option name	*default value*	*function*
InitialApproximation	Automatic	approximate solution to the system using mid-points of the directed input intervals of matrix **A**
WorkingPrecision	$MachinePrecision	number of digits of precision to be kept during the computations
AccuracyGoal	10^-6	accuracy to which two successive approximations differ
MaxIterations	$IterationLimit	maximum number of iterations to be performed in finding approximations to the algebraic solution
IterationList	False	gives the list of approximations obtained at each step of the iterative process

Table 1. Options for `AlgebraicIntervalSolve`.

`AlgebraicIntervalSolve` uses the precision specified by the option `WorkingPrecision` in the computational process. For the purpose of debugging one can ask `AlgebraicInterval Solve` to give the whole list of approximations obtained at each step of the iterative process by using the option `IterativeList`. Table 1 gives the options for `AlgebraicIntervalSolve` and their default values.

5 Acknowledgements

We wish to thank S. Shary and L. Kupriyanova for kindly sending us their latest papers [6], [16] before the latter have been published.

References

[1] Alefeld, G., Herzberger, J.: *Introduction to Interval Computations.* Academic Press, New York, 1983.

[2] Dimitrova, N.; Markov, S. M.; Popova, E.: *Extended Interval Arithmetics: New Results and Applications.* In: Computer arithmetic and enclosure methods (Eds. L. Atanassova, J. Herzberger), North-Holland, Amsterdam, 1992, 225–232.

[3] Gardeñes, E.; Trepat, A.: *Fundamentals of SIGLA, an Interval Computing System over the Completed Set of Intervals.* Computing, **24**, 1980, 161–179.

[4] Kaucher, E.: *Über Eigenschaften und Anwendungsmöglichkeiten der erweiterten Intervallrechnung und des hyperbolischen Fastkörpers über R.* Computing Suppl. *1*, 1977, 81–94.

[5] Kaucher, E.: *Interval Analysis in the Extended Interval Space IR.* Computing Suppl. *2*, 1980, 33–49.

[6] Kupriyanova, L. V.: *Inner estimation of the united solution set to interval linear algebraic system*, to appear in Reliable Computing.

[7] Markov, S. M.: *Some Applications of the Extended Interval Arithmetic to Interval Iterations.* Computing Suppl. *2*, 1980, 69–84.

[8] Markov, S. M.: *Extended Interval Arithmetic Involving Infinite Intervals.* Mathematica Balkanica, New Series, *6*, 3, 1992, 269–304.

[9] Markov, S. M.: *Some Interpolation Problems Involving Interval Data.* Interval Computations, 3, 1993, 164–182.

[10] Markov, S. M.: *On directed interval arithmetic and its applications.* Proc. Intern. Conf. "Real Numbers and Computers", Saint-Etienne, U. of Mines, France, April 4–6, 1995 (Eds. J.-C. Bajard, D. Michelucci, J.-M. Moreau and J.-M. Muller), 249–260.

[11] Markov, S. M. and Popova, E. D.: *Linear Interpolation and Estimation using Interval Analysis.* In: Bounding Approaches to System Identification (Eds. M. Milanese, J. P. Norton, H. P.-Lahanier, E. Water), Plenum Press, London, N. Y. (to appear)

[12] Popova, E. D.: *Generalized Interval Arithmetic in Mathematica.* preprint 1995.

[13] Ratschek, H.; Sauer, W.: *Linear Interval Equations.* Computing, 28, 1982, 105–115.

[14] Shary, S. P.: *On controlled solution set of interval algebraic systems*, Interval Computations, 6, 1992, 66–75.

[15] Shary, S. P.: *Solving the tolerance problem for interval linear equations*, Interval Computations, 1, 1994, 4–22.

[16] Shary, S. P.: *Algebraic Approach to the Interval Linear Static Identification, Tolerance and Control Problems, or One More Application of Kaucher Arithmetic.* (to appear in Interval Computations).

[17] Wolfram, S.: *Mathematica: A System for Doing Mathematics by Computer.* Second Edition, Addison-Wesley, 1991.

[18] Zakharov A. V.: *Construction of Interval Algebraic Solution in Extended Interval Arithmetic.* Proceedings All Union Conference "Actual Problems of Applied Mathematics", Vol. 3, Saratov State University, Saratov, 1991, 311–317.

[19] Zakharov A. V. and Shokin, Yu. I.: *Algebraic Interval Solution of System of Linear Interval Equations* $Ax = b$ *and* $Ax + d = b$, Preprint Computing Centre Siberian Branch of Acad. of Sci. of USSR, No. 5, Krasnoyarsk, 1987, 17p. (in Russian).

[20] Zakharov A. V. and Shokin, Yu. I.: *Synthesis of the control systems under conditions of interval uncertainity in parameters of their mathematical models*, Doklady Akademii Nauk, 299, 1988, No. 2, 292–295 (in Russian).

[21] Zyuzin, V. S.: *On a way of finding two-sided approximation to the solution of system of linear interval equations*, in Differential Equations and the Theory of Functions, Saratov State University, Saratov, 1987, pp. 28–32 (in Russian).

[22] Zyuzin, V.: *An iterative method for solving system of algebraic segment equations of the first order*, in Differential Equations and the Theory of Functions, Saratov State University, Saratov, 1989, 72–82 (in Russian).